PROCEEDINGS OF THE FIFTH WORKSHOP ON GENERAL RELATIVITY AND GRAVITATION

Symposion Hall, Nagoya University
January 22-25, 1996

edited by

Y. Nambu(secretary of this workshop),
Y. Eriguchi, T. Futamase,H. Ishihara, Y. Kojima,
K. Maeda, T. Nakamura, M. Shibata, A. Tomimatsu

PREFACE

The Fifth Workshop on General Relativity and Gravitation was held at Symposion Hall (Nagoya University) from January 22 to 25, 1996. The main purpose of this series of workshops is to review the latest observational and theoretical work in gravitational physics and relativistic astrophysics. A further purpose is to promote lively and stimulating interactions between researchers working in fields related to relativity and gravity -- including particle physics, astrophysics and cosmology.

110 researchers participated and 57 speakers gave 12 invited talks and 45 short contributions. Their research thema ranges over a wide field. We believe that there were mamy useful discussions and interactions between the participants.

This workshop was supported by a Grant-in-Aid for Creative Basic Research(No.07NP1101) from the Japanese Ministry of Education, Science and Culture. This proceedings volume was also partially supported by a Grant-in-Aid for Cooperative Research (No.06302021) from the Japanese Ministry of Education, Science and Culture.

Y. Nambu (*Secretary of this workshop*)	(Nagoya University)
Y. Eriguchi	(University of Tokyo)
T. Futamase	(Tohoku University)
H. Ishihara	(Tokyo Institute of Technology)
Y. Kojima	(Hiroshima University)
K. Maeda	(Waseda University)
T. Nakamura	(YITP)
M. Shibata	(Osaka University)
A. Tomimatsu	(Nagoya University)

CONTENTS

PREFACE..... iii

Monday, January 22

Afternoon Session

Chairman: H. Ishihara

Chaos in Stellar Systems
T. Tsuchiya..... 1

Structure Formation in the Universe and Non-linear Physics
J. Soda..... 18

Perturbative Lagragian Approach to Non-linear Clustering
M. Kasai..... *

Post-Zel'dovich Approximation in General Relativity
M. Morita..... 32

A Spinning Test Particle in a Relativistic Space Time
- Chaos and its Effect on Gravitational Wave Detection -
S. Suzuki and K. Maeda..... 40

(Coffee Break)

Chairman: J. Soda

Classical and Quantum Collapse of a Massless Scalar Field
A. Tomimatsu..... 50

Critical Behavior in the Brans-Dicke Theory of Gravitation
T. Chiba..... 64

Dynamics of Gravitating Magnetic Monopoles
N. Sakai..... 70

Inflation in the Inhomogeneous Universe
O. Iguchi and H. Ishihara..... 80

Density Perturbation Generated by Primordial Magnetic Fields
K. Tamai and A. Hosoya..... 89

Tuesday, January 23

Morning Session

Chairman: K. Maeda

Solar Neutrino and Atmosphere Neutrino Problem
A. Suzuki..... *

Present Status of Superstring and Theory of Gravitation
T. Eguchi..... *

(Coffee Break)

Anomalies in Gravitational Systems and Weak Expansion
N. Ikeda..... 97
Topologically Twisted 2-D Conformal Supergravity and BRST Gauge Fixing
N. Ano..... 107
Consistency of Matter Field Equations in Ashtekar Formalism
M. Tsuda and T. Shirafuji..... 113
Dynamics in the Ashtekar Gravity
H. Shinkai, G. Yoneda and A. Nakamichi..... 121

Afternoon Session
Chairman: T. Shiromizu
How Can We Understand a Small But Nonzero Cosmological Constant?
Y. Fujii..... 131
Compact Hyperbolic Universe and Singularities
A. Ishibashi, T. Koike and M. Siino..... 151
Dynamics of Compact Homogeneous Universe
M. Tanimoto..... 161
Quantum Stability of (2+1)-Spacetimes with Non-Trivial Topology:
Negative Curvature Cases
M. Siino..... 171
(Coffee Break)
Chairman: H. Shinkai
Causality Violation and Singularities
Kengo Maeda and A. Ishibashi..... 181
Construction of Gravitational Equation System Using a Non-coordinate Base
N. Iwanaga..... 190
Continuum Strong Coupling Expansion of Einstein Gravity coupled with a Scalar Field
Kayoko Maeda..... *
Dynamical Torsion and Torsion Potential
H. Xie and T. Shirafuji..... 202
Reflecting Condition and Cosmic Censorship
M. Nalita..... 211
Dilatonic Black Holes with Gauss-Bonnet Term
T. Torii, K. Maeda and H. Yajima..... 218
Non-Singular Superstring Cosmology
R. Easther..... 228

Wednesday, January 24
Morning Session

Chairman: T. Nakamura
Space-based Gravitational Wave Astronomy
N. Kawashima, R. Takahashi and S. Kawamura..... 238
MACHO Search by VLBI Observation
K. Oonishi..... *
Some Theoretical Issues in Gravitational Wave Physics and the Black-hole Perturbation Theory
M. Sasaki..... 246
(Coffee Break)
Chairman: Y. Kojima
Gravitational Wave by a Particle in Circular Orbits around a Schwarzshild Black Hole
T. Tanaka, H. Tagoshi and M. Sasaki..... 266
Post-Newtonian Hydrodynamic Equations Using the (3+1) Formalism
H. Asada, M. Shibata and T. Futamase..... 275
Evolution of Close Neutron Star Binaries - Tidal and Spin Effects -
W. Ogawaguchi and Y. Kojima..... 285
On the Stability of Synchronized Binary Neutron Stars in Post-Newtonian Approximation of General Relativity
M. Shibata..... 294

Afternoon Session
Chairman: M. Shibata
Black Hole Thermodynamics
O. Kaburaki..... 306
On the Ergoregion Instability - a New and General Method for Numerical Mode Analysis of Classical Massless Scalar Field -
S. Yoshida and Y. Eriguchi..... 326
Quasinormal Modes of Maximally Charged Black Holes
H. Onozawa, T. Mishima, T. Okamura and H. Ishihara..... 336
The Thermodynamic Second Law of 2D Dilaton Black Hole
K. Hirata , Y. Fujiwara and J. Soda..... 340
Field Theoretical Quantum Effects on the Kerr Geometry
Y. Satoh..... 350
(Coffee Break)
Chairman: T. Tanaka
Thermodynamic Properties of Two Dimensional R^2-Gravity
S. Ichinose..... 359
Analysis of the Back-Reaction on the Global Degrees of Freedom in (2+1)-Dimesional Spacetime

M. Seriu..... 370
Quantum Harmonic Oscillators in the Expanding Universe
A. Taruya and Y. Nambu..... 380
Low-Energy Interaction of a Cosmic String and an Extreme Dilatonic Black Hole
K. Shiraishi..... 393
Bremsstrahlung induced by the Cosmic Expansion
M. Hotta..... 400
Particle Production and Gravitino Abundance after Inflation
K.Kumekawa..... *
On Perturbation in a Self-Similar Spacetime
K. Tomita..... 402
Gravitational Collapse in Szekeres Universe
H. Mutoh, T. Hirai and K. Maeda..... 408

Thursday, January 25

Morning Session

Chairman: K. Yamamoto

ASCA Observations of Clusters of Galaxies
K. Yamashita..... 416
Supergravity Cosmology - Gravitino Problem, Polonyi Problem -
M. Kawasaki..... *
Thermal Inflation, the Moduli(Polonyi) Problem and Baryogenesis
E. D. Stewart..... 426
Mexican Hat Gravitational Lenses
P. Haines..... 430

(Coffee Break)

Chairman: M. Sasaki

Anisotropies of Cosmic Microwave Background and Observational Cosmology
N. Sugiyama..... 435
Evolution of Cosmological Perturbations during Reheating
T. Hamazaki and H. Kodama..... 444
Theory of Force-Free Black Hole Magnetosphere
T. Uchida..... 451
Collimation of Magnetosonic Waves near a Black Hole
K. Hirotani..... 458

* manuscript not avilable

Chaos in Stellar Systems

Toshio Tsuchiya
Department of Astronomy, Kyoto University, Kyoto 606

March 11, 1996

1 Introduction

Study of chaos has a long history. For example, Poincaré discussed the complex behaviour of the 3-body problem of celestial mechanics, in 1982. It is still very active, and applied to remarkably wide range of fields, not only physics, but also chemistry, biology, economics, and even for amusement park (Jurassic park). However, I have an impression that the outcome of the study is unclear. A part of interests in chaos are paid toward its mathematical properties. Thus before I started working on chaos, I had a simple question: how is chaos useful, or how important is it in studying Astrophysics. I guess that this question is somewhat common. Therefore, I will devote a half of this paper to a review of this issue. There I try to figure out substance of chaos in Stellar dynamics. And the result of our recent study about the evolution of gravitating systems using a one-dimensional model is given in the latter half.

2 Chaos in Hamiltonian Systems

2.1 Definitions of Chaos

The definition of chaos is not unique. I enumerate three main definitions of chaos beneath:

1. Unpredictability

   This is the most naive and intuitive concept of chaos. Even when one has sufficient information of the state at a instance, he/she cannot predict the state in a future. Of course, in some system, it is because random noise makes the evolution undeterministic. However, a more interesting example is the model of atmospheric motion suggested by E.N. Lorentz which has only three variables and the evolution of deterministic, but very sensitive dependence on the initial condition makes it actually unpredictable.

2. Positive Lyapunov Exponents

This means the existence of exponential instability. A system with N degrees of freedom has $2N$ Lyapunov exponents (λ_i, $i = 1, \ldots, N$) corresponding to the growth rates of the perturbations with respect to $2N$ independent direction in the phase space. The maximum Lyapunov exponent is determined as follows:

$$\lambda = \lim_{\substack{d(0)\to 0 \\ t\to\infty}} \frac{1}{t} \ln \frac{d(t)}{d(0)}, \tag{1}$$

where $d(t)$ and $d(0)$ are the separations in the phase space between two nearby orbits at times t and 0, respectively.

In more mathematical descriptions, this second definition of chaos is expressed that the Kolmogorov-Sinai entropy (KS entropy) is positive. I don't give the exact expression of KS entropy, h_k, but it is useful to give a simple relation which is valid in many usual case:

$$h_k = \sum_{\lambda_i > 0} \lambda_i. \tag{2}$$

This condition guarantees the exponential instability, and it induces the strong sensitivity to the initial conditions. That causes chaos.

3. Non-Integrability

This definition is only for Hamiltonian systems. Instead of explaining the non-integrability, I explain *integrability*.

A system with N degrees of freedom has $2N$ canonical variables. If one can find N independent integrals of motion I_i which satisfy the following equations,

$$\{I_i, I_k\}_{\mathrm{PB}} = 0, \qquad \{\,\}_{\mathrm{PB}} \text{ is the Poisson bracket,} \tag{3}$$

it is possible to transform the canonical variables into the action-angle variables, (θ_i, I_i), where I_i's are constant. In this case the system is called *integrable*. The orbits in the system make tori in the phase space, and show only regular motion. On the other hand, in the system which is not integrable orbits no longer show regular motion but stochastic. Figure 1 shows the basic concept of the regular and stochastic orbits. The concentric closed curves are the tori, the regular orbits, and the other regions filled with dots are the stochastic orbits.

This definition of chaos sounds different from the above two definition, but they are closely related. Motions on the tori do not posses any exponential instability. On the other hand the exponential instability causes stochastic diffusion of orbits in the phase space, to produce uniformly distributed *stochastic seas*.

2.2 Classification of System

2.2.1 Integrable systems

As described before, in this system all the orbits are regular. Thus the state of the system can be predicted however far in the future. Examples of such systems are the Kepler motion,

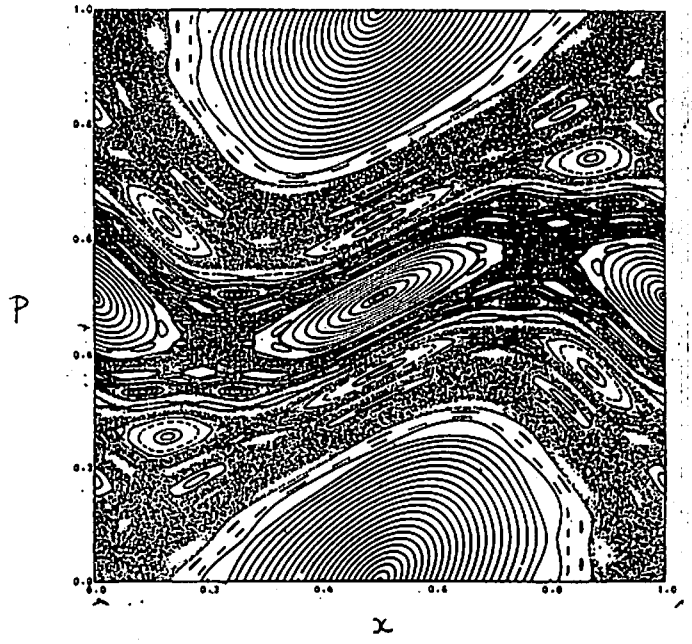


Figure 1: Orbital structure in phase space of the standard map system. The map from the old coordinates (x, p) to the new one (x', p') is defined by $p' = p + 1/2\pi \sin 2\pi x$, and $x' = x + p'$.

and harmonic oscillators. The motion in a static spherical potential is also regular, because there exists at least three independent integrals, the energy, and two components of angler momenta.

2.2.2 Ergodic systems

In a ergodic system, the time average and statistical average (space average) are equivalent. This property is one of the most important basics of statistical mechanics.

2.2.3 Mixing systems

There is a famous example given by Arnold and Avez[1] to explain the concept of mixing. Following it, we take a shaker that consists of 20% rum and 80% cola, representing the initial distribution of the "incompressible fluid" in phase space. If A is the region originally occupied by the rum, then for any part B of the shaker, the percentage of rum in B, after N repetitions of the act of stirring (see Fig. 2), is

$$\frac{\mu(\phi^n A \cap B)}{\mu(A)}, \tag{4}$$

where μ is a measure and ϕ is the mathematical representation of the act of stirring. In such a situation, physicists expect that, after the liquid has been stirred sufficiently often

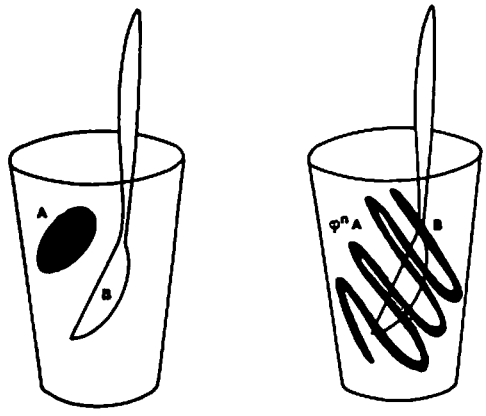


Figure 2: Mixing. This figure is taken from Arnold and Avez.

($n \to \infty$) every part of B, however small, will contain "approximately" 20% rum. When stated rigorously, this defines a *mixing system.*

The property of mixing is equivalent to *relaxation*, because the mixing leads a system to stationary, the maximum entropy state. Mixing implies ergodicity.

2.2.4 K-systems

A system is called a K-system if the KS entropy of the system is positive. In this case, as stated above, there exist exponential instabilities. Let us consider a subset of the phase space A. Due to the exponential instability, the measure of phase points which remain in the A decreases in time. In other words, the system loses the initial information. The physical meaning of KS entropy is a measure of rate to lose the information.

2.3 Intermediate zone between the order and chaotic systems

There is a inclusion relation between the systems stated above:

$$\text{Ergodicity} \supset \text{mixing systems} \supset \text{K-systems.} \tag{5}$$

Chaotic systems are always K-systems. Thus the chaos is important for statistical properties of the system. A strongly chaotic system is guaranteed applicability of statistical mechanics. However, there are systems which are intermediate between the order and chaos. In such system, sometimes peculiar phenomena are observed.

The historic problem was given by Fermi, Pasta and Ulam[2]. They studied an one-dimensional chain of weakly nonlinear harmonic oscillators. They had an expectation that the nonlinear term drives the system to ergodicity and then the system achieves the equipartition of energies. On the contrary, their numerical study revealed that the energy given in one mode initially was not transfered to all another mode but some, and after a

while the energy returned to the initial mode. It seemed that even though the system had many degrees of freedom, the ergodicity was broken.

After that many works are devoted to solve the problem. The main issue was that what happen in the system of intermediate between order and chaos.

2.3.1 KAM tori

In a near-integrable system, which can be described as a perturbation from a integrable system. If the Hamiltonian of the integrable (unperturbed) system is H_0, then the Hamiltonian of the perturbed system is $H = H_0 + \epsilon H_1$. In the unperturbed system, all orbits make tori. Since the perturbed system is no longer integrable, one might expect there is no torus in the phase space. On the contrary, what KAM (Kolmogorov, Arnold, and Moser) theory tells is the tori around stable periodic orbits in the unperturbed system remain tori in the perturbed system. Thus tori and chaotic orbits coexist in the phase space. Figure 1 was the example of the KAM tori.

2.3.2 Arnold diffusion

In a system with N degrees of freedom, tori form N-dimensional hyper-surface. With this hyper-surface, $N+1$-dimensional space can be divided into disconnected regions. Thus for $N > 2$, the tori cannot divide the equi-energy hyper-surface, which has $2N-1$ dimensions. In the system, all chaotic regions are connected. The *Arnold diffusion* is the motion of a phase point to escape from the inside of the shells of KAM tori.

2.3.3 Stagnant motions

There are remarkable regions where two different orbits with opposite characters, KAM tori and chaotic orbit, meet. In this region, orbits are chaotic, but behave like regular orbits in some duration due to the KAM tori nearby. The time scale for a orbit to stay near the KAM torus was estimated by Nekhoroshev[3], which is the longer for the nearer orbit to the torus. The KAM tori are, however, present in fractal manner. Therefore an orbit moves in a web of KAM tori like a Brownian motion. That is the Arnold diffusion. In the region intermediate KAM tori and chaotic orbits, which is called *stagnant layer*, the distribution of time for a orbit to escape from the layer is estimated as

$$P(T) \propto T^{-1}. \tag{6}$$

This means the expectation value of the time of stay in side the layer becomes infinite. That is why the motion is called *stagnant motion.* This phenomena cause the apparent break of ergodicity and usual application of statistical mechanics.

Now it is known that the stagnant motion plays an important roll in the FPU problem.

Another example of realisation of the stagnant motion is found in a motion of asteroids in the Solar system. Milani and Nobili[4] found by numerical studies that the eccentricity of 522 Helga shows no change for a duration 1000 times longer than the Lyapunov time. They call this phenomena *stable chaos.*

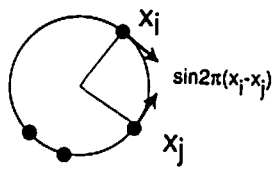


Figure 3: Konishi-Kaneko model. A two-particle interaction is schematically shown.

One more example is the work by which I was motivated to study chaos in Astrophysics. I will explain it in the next subsection.

2.4 Globally coupled standard map systems

Though ergodicity of a system is basic ground of statistical physics, we saw that near integrable system could break this property. In particular, statistical physics is so important in systems with many degrees of freedom. Thus in the next logical step, the study of chaos in systems with many degrees of freedom was noticed.

For systems with a short range interaction force, in other words when the particles couple only with the neighbours, some numerical work resulted that the volume of KAM tori is getting reduced as the number of degrees of freedom increases[5, 6].

On the other hand, though not so clear results are obtained, the systems with long range force show the different behaviours. Konishi and Kaneko[7] studied the roll of chaos in a globally coupled system by using a simple model. In their the map of coordinates of N particles, $(x_i, p_i) \mapsto (x_i', p_i')$ is given by

$$p_i' = p_i + \frac{K}{2\pi\sqrt{N-1}} \sum_{j=1} N \sin 2\pi(x_j - x_i), \qquad K > 0, \tag{7}$$

$$x_i' = x_i + p_i'. \tag{8}$$

This system is called the symplectic coupled map system, or sometimes simply Konishi-Kaneko system, among the researchers of astrophysics. This system is depicted as a set of particles distributed on a ring, moving under the attractive force ($K > 0$) which is proportional to the sine of the angle between two particles,(Fig. 3).

In the system, there are two quasi-equilibria. One is the uniform state, in which particles distribute uniformly on the ring, but the motions are random. The other is the clustered state, in which most particles make a clump and rotate on the ring as a whole.

From numerical simulations, they found that the system start with the uniform state transforms into the clustered state after a while, and then returns to the uniform state. The system repeated the transformation, which appeared similar to the transition of the excited modes in FPU problem.

Analysing microscopic dynamics of the system using the variation of the Lyapunov exponents revealed that the clustered state is less chaotic than the uniform state. They showed that the orbits of clustered states exhibited torus like structure. Figure 4 shows a

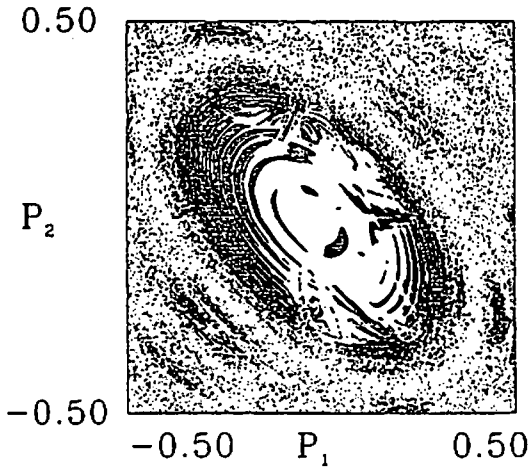


Figure 4: A 2D slice (p_1, p_2) of phase space. $N = 4$, $K = 0.3$. In the $(4 \times 2 =)$ 8-dimensional phase space, the slice is taken by setting the following six constraints; $x_1 = -0.075$, $x_2 = -0.025$, $x_3 = 0.025$, $p_3 = 0$, $\sum_1^4 x_i = \sum_1^4 p_i = 0$.

2-dimensional slice of the phase space of the system with $N = 4$. Two axes are values of p_1 and p_2, where the other coordinates each are given some fixed values. Time evolution of the orbits which start with the given (p_1, p_2) as the initial condition are calculated. Among the initial conditions, dots are plotted for the initial points such that the orbits show small growing rate of the instability are shown by dots. In other words, the dots indicate weakly chaotic orbits. It is clear that the dots make torus like structure. Thus the reason why the orbits are weakly chaotic is the orbits are located in the vicinity of KAM tori or at least the ruin of KAM tori.

This is a very simple model, but capable to study structure formation. In this case, structure formation (from the uniform to the clustered state) can be interpreted as the stagnant motion.

This system is not real, but have several common properties with gravitating systems. Both systems are

1. Hamiltonian systems,
2. many-body systems,
3. globally coupled by long range force.

Thus it is possible that the similar phenomena relating chaos is observed in gravitating systems. It is worthwhile to study what role chaos plays in gravitating many-body systems. Especially time scale is apt to be affected by existence of KAM structures.

Besides, analyzing tools of chaos reveals the *microscopic dynamics* of the system, which should be new information about states and evolution of the system.

Since my idea is fundamental, I employ the simplest model: the one-dimensional sheet model. In succeeding sections, I will explain the results of my work.

3 Evolution of one-dimensional self-gravitating many-body systems

Here I report a brief review of the series of works[8, 9, 10] with N. Gouda and T. Konishi.

3.1 Model

One-dimensional self-gravitating many-body systems consist of N identical parallel sheets which have uniform mass density m and infinite in extent in the (y, z) plane. We call the sheets *particles* in this paper. The particles are free to move along the x axis and accelerate as a result of their mutual gravitational attraction. The Hamiltonian of this system has the form:

$$H = \frac{m}{2}\sum_{i=1}^{N} v_i^2 + (2\pi G m^2)\sum_{i<j}|x_j - x_i|, \tag{9}$$

where m, v_i, and x_i are the mass (surface density), velocity, and position of ith particle, respectively.

This system is originally introduced as a model for the motion of stars in a direction normal to the disk of a highly flattened galaxy[11, 12]. In 1970s and 1980s this system was used to study mechanism of relaxation and evolution of galaxies. This system was not realistic model of galaxies but the reason why they used it was because the gravitational force is uniform in one-dimensional systems, thus the equations of motion are reduced to algebraic equations.

This model has some advantages over usual three-dimensional systems: no softening parameter is necessary, then we do not worry about the effect of softening, which may alter the evolution of the systems. Further, the force in one-dimensional systems are uniform, thus the equations of motion can be integrated analytically while no intersection of particles occurs. Therefore we can calculate the evolution without any truncation error, which is inevitable for three-dimensional systems. Since our main interest is chaos in the systems, these properties are important.

3.2 Numerical simulations

In our numerical simulations, we use the system size, the total mass, and crossing time t_c, as the units of length, mass, and time. In these units, N is the only parameter which discriminates the Hamiltonian.

Our initial condition, the water-bag distribution (Fig. 5a), has a homogeneous phase density, and is in a virial equilibrium. Putting particles in a rectangle region is just for convenience. In a few crossing times, the corners disappear and become round but still remains to be uniform. In such a short time scale, the system behaves like *collisionless*

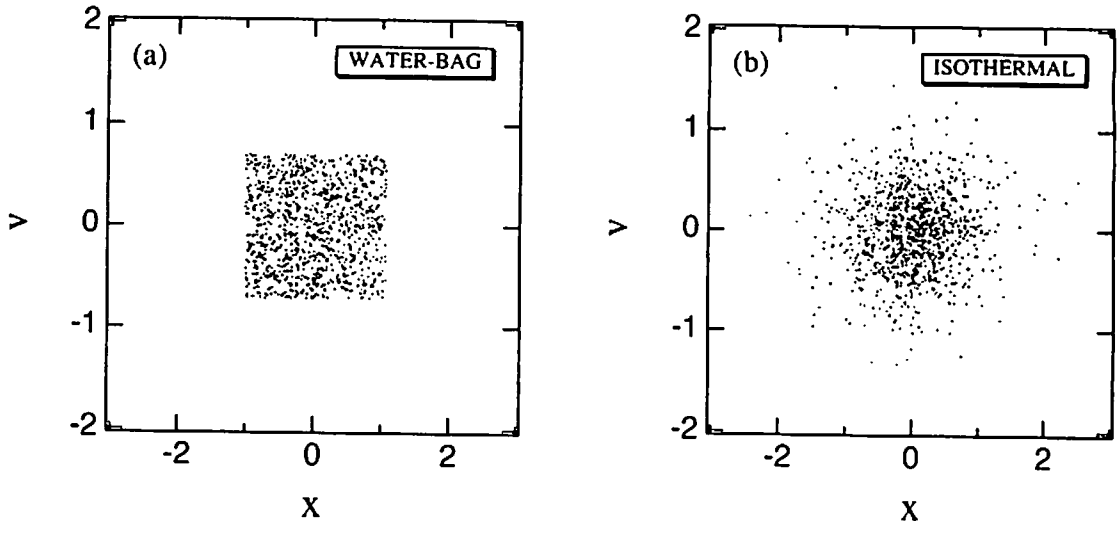


Figure 5: The water-bag(a) and the isothermal(b) distributions in one-body phase space.

system, and in that case the water-bag distribution is one of the stationary solutions. One of the other stationary solutions is the *isothermal* distribution (Fig. 5b). It is true thermal equilibrium, the maximum entropy state, and then all the other distributions should transform into the isothermal distribution. However, it is known that the water-bag has difficulty to relax to the thermal equilibrium[13].

We have studied the very long time evolution of the stationary water-bag distribution, and found it eventually transformed into the isothermal distribution[8]. At $t \sim 4 \times 10^6\, t_c$ the macroscopic distribution changed, but in a much shorter time scale we found another kind of relaxation. It can be seen in the microscopic dynamics. To see that, we concern with the individual (specific) energy fluctuation $\varepsilon_i(t)$, where i is index of the particles. If the system is ergodic, which is expected in usual thermal equilibrium, the infinite time average of $\varepsilon_i(t)$ gives unique value for the equipartition:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \varepsilon_i(t)dt = \overline{\varepsilon}_i \equiv \varepsilon_0. \tag{10}$$

Even in the thermal equilibrium, there exist the thermal fluctuations thus a finite time average,

$$\overline{\varepsilon}_i(t) = \frac{1}{t} \int_0^t \varepsilon_i(t')dt', \tag{11}$$

shows small deviation from the equipartition. Here we introduce a averaged deviation from equipartition until time t,

$$\Delta(t) \equiv \varepsilon_0^{-1} \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\overline{\varepsilon}_i(t) - \varepsilon_0)^2}. \tag{12}$$

The statistical theory tells us that if $\varepsilon_i(t)$ behaves like thermal noise, $\Delta(t) \propto t^{-1/2}$. Therefore we can make use of $\Delta(t)$ for test of thermalization.

Figure 6 shows the time variation of $\Delta(t)$. In the figure we can find two distinct time scales. The plateau at the beginning represents the collisionless phase, because in the collisionless phase the individual energies are conserved. After $t \sim 100$, $\Delta(t)$ begins to decrease as $t^{-1/2}$, which means that the fluctuation behaves as the same manner as the thermal noise.

This behaviour of $\Delta(t)$ is one of indirect evidences that the state is in the thermal equilibrium. We analyzed by two more methods. One is to examine convergence of the maximum Lyapunov exponent. Since the Lyapunov exponent is defined as a time average of local instability, its convergence means that the orbit shows ergodicity. Another method is the spectral analysis. The power spectrum density (PSD) of fluctuation reveals the motion of the phase point, which determines the state of the system, in the phase space. If PSD is proportional to f^{-2}, the motion is diffusion by the process of random walk, and once the phase point travels all region in the phase space, then the motion is ergodic and PSD becomes a white noise. These three analyses give quite the same time scale of relaxation. Therefore we show only result of the analysis of $\Delta(t)$.

The transition from constant of $\Delta(t)$ to the power law, $\Delta(t) \propto t^{-1/2}$ determines the *microscopic relaxation*. In the time scale where $\Delta(t) \propto t^{-1/2}$ the system has the same properties as the thermal equilibrium. However, what surprised us is the macroscopic distribution is the water-bag distribution and unchanged from the beginning.

If the water-bag distribution is the thermal-equilibrium, then no more change is expected and $\Delta(t)$ goes to zero as t increases. However, it is found that $\Delta(t)$ increases at some $10^6 t_c$. At that time the macroscopic distribution transforms from the water-bag into the isothermal distribution. It is the *macroscopic relaxation*. Since the water-bag distribution shows all properties of the thermal equilibrium, but it is different in the macroscopic distribution, it is called a *quasiequilibrium*.

3.3 Mechanism of two relaxations

Dependence of relaxation time on the number of particles gives us much insight about its mechanism. Figure 7 shows dependence of the microscopic relaxation time on the number of particles. The doted line stands for the linear dependence on N. Thus the microscopic relaxation time $\sim N t_c$. It can be explained that the microscopic relaxation is a *diffusion process* caused by random force created from thermal fluctuation of mean field distribution[14]. Even though the microscopic dynamics shows the same property as the thermal equilibrium, the macroscopic distribution still remains the water-bag one.

The mechanism of the macroscopic relaxation time is not so clear. The macroscopic relaxation has some different properties to the microscopic relaxation. In the case of the microscopic relaxation of the water-bag distributions, different microscopic distributions (created by different random seeds) yield the definite relaxation time scale. For the macroscopic relaxation, however, we found a distribution of relaxation time which has a range over an order of magnitude. Therefore the mechanism of the macroscopic relaxation is surmised to be different from that of the microscopic relaxation. In this case, the distribution

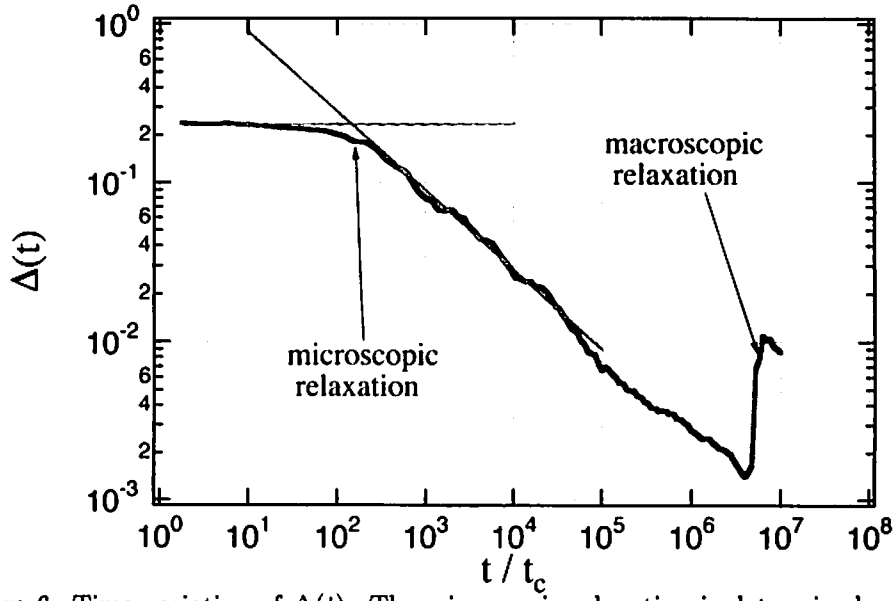


Figure 6: Time variation of $\Delta(t)$. The microscopic relaxation is determined as the time scale where two different slopes (constant and $t^{-1/2}$) meet in the figure. The macroscopic relaxation corresponds to the steep increase at about $t \sim 5 \times 10^6$.

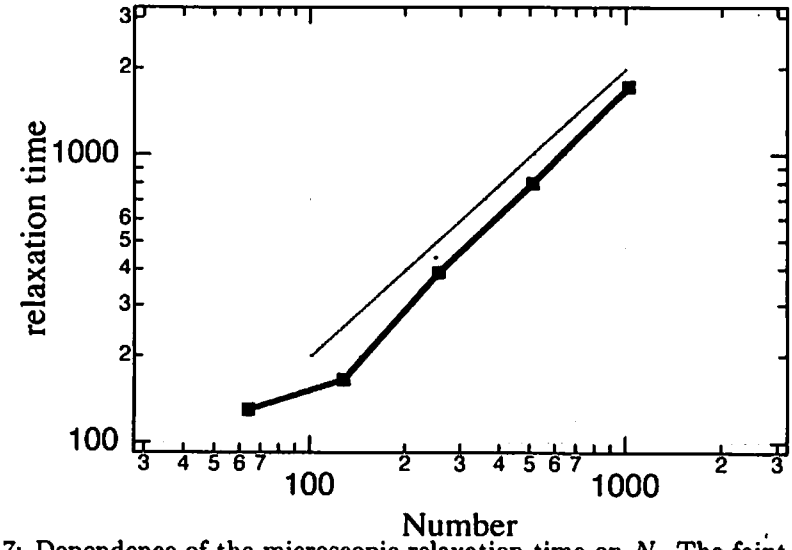


Figure 7: Dependence of the microscopic relaxation time on N. The faint line is $T \propto N$.

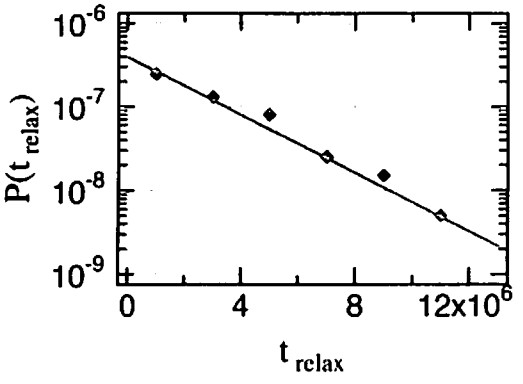


Figure 8: (left) Probability distribution of macroscopic relaxation time

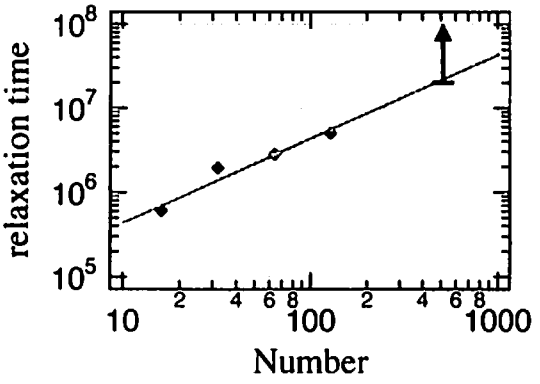


Figure 9: (right) Dependence of the macroscopic relaxation time on N.

of the relaxation time is a big clue to find the mechanism.

Figure 8 shows the probability distribution of the relaxation time of the system with $N = 64$. We took 100 different initial states, statistical ensemble, of the macroscopically same water-bag distribution, which are created by different seeds of the random number generator The relaxation time of each run, T_M, is determined by the figure of $\delta(t)$, but there are wide variations among the ensemble. The relaxation times are divided into bins with the interval of $2 \times 10^6 t_c$. The solid diamonds represent the results of the numerical simulations. The vertical axis is scaled such that $P(T_M)\,dT_M$ gives the fraction of the number of runs, which fall into the interval dT_M. It is clear from the figure, that the distribution has a exponential distribution,

$$P(T_M) = \frac{1}{\langle T_M \rangle} e^{-T_M/\langle T_M \rangle}, \quad \text{where,} \quad \langle T_M \rangle = 2.8 \times 10^6 t_c, \qquad \text{for } N = 64, \tag{13}$$

and this gives the expectation value of the relaxation time.

Before we proceed to speculate the mechanism it is useful to remind us the phase space dynamics of the system.

Any state of the system with N particles can be described by a certain point in the $2N$ dimensional phase space (Γ space). Each point yields some macroscopic distribution, such as the water-bag or the isothermal. In Γ space, there is a region where all phase points yield the water-bag distribution. At the beginning the phase point is located in the region and then it moves out of the region as the system evolves. The macroscopic distribution is the water-bag during the phase point stays in the region. When the point escapes from the region, the macroscopic distribution is transformed from the water-bag distribution into the isothermal distribution because the isothermal distribution is defined as the maximum entropy state, which means that it occupies the largest, and usually most region in the Γ space.

Now we should explain the following facts:

1. The water-bag exhibits several properties of thermal equilibrium, such as convergence of Lyapunov exponents, and equipartition of energy, which suggest ergodicity.

2. The probability distribution of the macroscopic relaxation time is a exponential function.

One possibility to explain these facts is that there exists a barrier which is an obstacle for a phase point to get away from the water-bag region. Due to this barrier a phase point is restricted in the region for a long time, then it travels all over the region as if the water-bag region is ergodic. However, the barrier is not perfect and there is a small gate from which a phase point in the water-bag region can escape. A phase point travels in the region in a very complicate way and almost ergodic, thus the point finds the gate and escape at random. Now we show a simple model which might explain the numerical results well as follows: as the simplest case, we assume that the escape probability is uniform in the region. Suppose that we have an ensemble of the phase points in the region. At the beginning the ensemble contains $n(0)$ points, but they escape from the region with constant rate $1/\langle T_M \rangle$, then the number of points which stay in the region at t decreases as $n(t)$. It is well known that $n(t)$ has the same form as eq.(13). Therefore this simplest model can explain the facts obtained by the simulations. This kind of stochastic escape is basically same as decay of unstable nuclear, which also results the exponential distribution of life time.

Next, in order to investigate the dependence of the time scale on the number of particle, the same procedure was applied to the system with different N; $N = 16$, 32, 128, and 512. Figure 9 shows the results. Especially for $N = 512$, we observed that 50 runs of the maximum integration until the time $T = 10^6 t_c$ did not relax. Thus we can not determine the time of the relaxation for $N = 512$, but by assuming the exponential probability distribution, we can restrict the region that the *true* relaxation time lies probably. The arrow indicated the region of 90% confident level. These data is approximated by a linear relaxation,

$$\langle T_M \rangle = 4 \times 10^4 \, N \, t_c, \tag{14}$$

which is shown by the dashed line.

3.4 Evolution in longer time scale

In our recent work[10], we investigated the evolution of longer time scale toward the thermal equilibrium. The initial conditions were taken in general, and found that there exist many quasiequilibria other than the water-bag. They experienced the same relaxation processes as the water-bag. In the time scale of the macroscopic relaxation the macroscopic distribution changed. However, the system did not became the isothermal distribution directly, but after experienced a transient state for a while the system went back to another quasiequilibrium, and repeated such transitions. From detailed analysis of the quasiequilibria and the transient states (I do not explain it here, please see Ref. [10]), we can speculate the

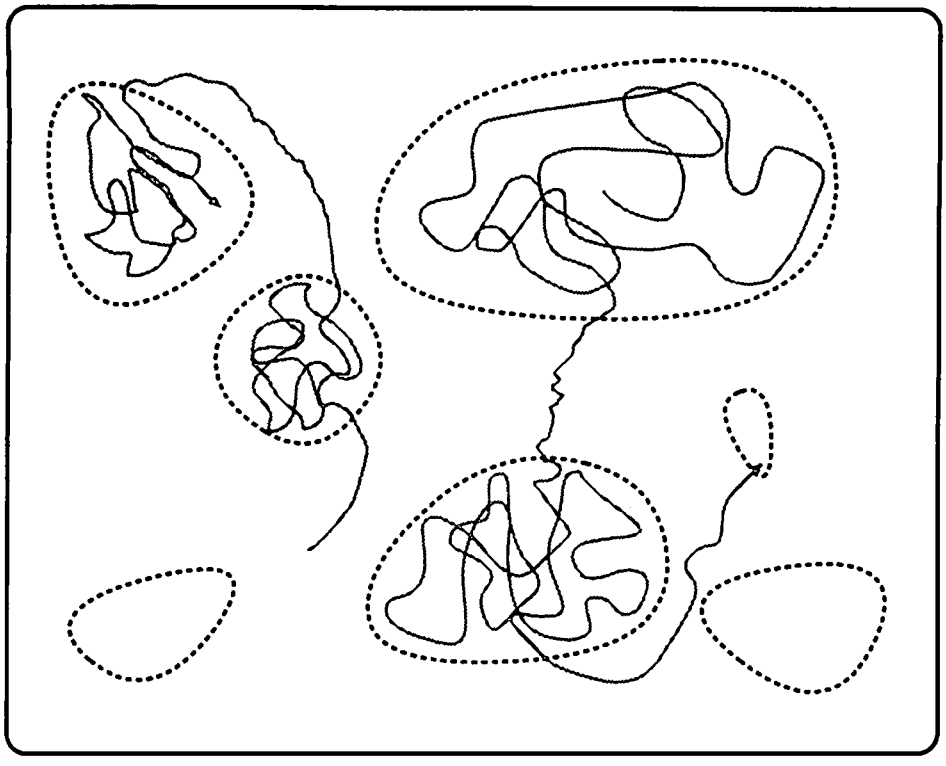


Figure 10: Basic idea of the structure of orbits in the phase space. The solid lines indicates the orbits, which describes the states of the system. The regions which enclosed by dashed curves corresponds to the quasiequilibria. The paths which connect the quasiequilibrium regions corresponds to the transition states. Actually these structures are fractal.

structure of the orbits in the phase space. Figure 10 shows the basic idea of the evolution in the phase space. There are many barriered region in the phase space, which are denoted by dashed circles in Fig. 10. The barriers which enclose the region obstruct the orbit to go out of the region, then the orbit travels all over the region ergodically. The barriers are, however, not complete, and the orbits escape trough small windows on the barriers. The escaped orbits enter the transition stage. In this stage the number of dimension of the orbits reduces as if the orbits are passing through a narrow corridor. And the orbits enter another barriered region and the system becomes another quasiequilibrium, again. The orbits repeat the same itenerancy again and again, and at last, the orbits travels all over the phase space. In this stage, the thermal equilibrium is established.

3.5 Conclusions for the study of one-dimensional systems

As conclusion we summarize the evolution of the one-dimensional self-gravitating many-body systems, based on the results which are obtained by the series of our works.

We suggest that there are roughly four stages of evolution in the system.

1. Virialization phase: The system which is far from the virial equilibrium experiences violent oscillation of the mean field. The violent relaxation leads the system to one of the virial (dynamical) equilibrium at the end of this stage. Many works were devoted to the study of this stage in 1970s and 1980s. In general, this violent oscillation ceases in several t_c.

2. Dynamical equilibria: This stage corresponds to the era with the time scale of several $t_c \lesssim t \lesssim N t_c$. The system is in a dynamical equilibrium such as the water-bag distribution, which we investigated in Paper I and II. In this stage the energies of the individual particles conserve. Small fluctuations of mean field are driving the system to the equipartition, but it is not yet effective in this time scale.

3. Quasiequilibria: After the change of energies driven by the fluctuation of the mean field becomes effective, the energy distribution achieves the equipartition. This process is called the microscopic relaxation. The time scale of the microscopic relaxation is about $N t_c$ irrespective with the initial conditions and after that the system is settled in a quasiequilibrium. This quasiequilibrium is realized by the orbit which is restricted in a part of the phase space. This region is enclosed by some kind of barrier which obstructs the orbit to go out of the region, thus the orbit eventually travels all over the barriered region and exhibits the nature of equilibria. That is why the mixing of energies occurs sufficiently but the global distribution does not change. The water-bag distribution is one of the quasiequilibria, but we found in this paper that there exist many similar quasiequilibria. The orbit which is initially located in one of the regions stays inside. It escapes from the region when it find a window to the outside. This happens stochastically, but the typical time scale is, in the case of the water-bag, $4 \times 10^4 N t_c$. As for other quasiequilibria, the time scale is about the same. When the orbit escapes from one of the barriered region, the global distribution transforms. This process is, we call, the macroscopic relaxation.

4. The thermal equilibrium: We found that the system in a quasiequilibrium transforms into another quasiequilibrium, experiencing a transition state in the midway between the quasiequilibria. The system wanders between the quasiequilibria and the transition states. Averaging this behavior over a time scale much greater than the macroscopic relaxation time, gives the isothermal distribution. In this time scale, the system becomes ergodic and then the thermal equilibrium.

4 Discussions

Our results are not sufficient to prove existence of the KAM tori. It is very difficult to figure out the precise structure of orbits in the phase space of such a system with many degrees of freedom. We found, however, that the phase space has rich structure, and is weakly separated into pieces. This structure persists and even is strengthen for larger N systems, because the relaxation times are proportional to N. The facts that the separation of the phase space is enhanced, and the Lyapunov exponent becomes smaller for the larger N system. Combined with the results of Reidl and Miller[15, 16] that dynamics is most chaotic for the system with about 30 particles, we conclude that the one-dimensional self-gravitating systems are the less chaotic for the larger N ($N \gtrsim 30$). This result is interesting in the view of the basic theory of statistical mechanics, because the strength of chaos and ergodicity of a system assures validity of the statistical mechanics. Since our system is one of the simplest models, it is very important to study more general systems such as three dimensional systems, which are realistic systems commonly existing in the universe.

References

[1] V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).

[2] E. Fermi, J. R. Pasta, and S. Ulam, *Collected Works of Enrico Fermi* (University of Chicago Press, Chicago, 1965), Vol. 2, p. 978.

[3] N. N. Nekhoroshev, Russian Math. Surveys **32**, 1 (1977).

[4] A. Milani and A. M. Nobili, Nature **357**, 569 (1992).

[5] G. Gallavotti, in *Scaling and self-similarity in physics*, edited by J. Fröhlich (Birkhäuser Boston Inc., Boston, 1983), p. 359.

[6] T. Konishi, Prog. Theor. Phys. Suppl. **98**, 19 (1989).

[7] T. Konishi and K. Kaneko, J.Phys.A: Math.Gen. **25**, 6283 (1992).

[8] T. Tsuchiya, T. Konishi, and N. Gouda, Phys. Rev. E **50**, 2607 (1994).

[9] T. Tsuchiya, N. Gouda, and T. Konishi, Phys. Rev. E **53**, 2210 (1996).

[10] T. Tsuchiya, N. Gouda, and T. Konishi, preprint (1996).

[11] J. H. Oort, Bull. Astr. Inst. Netherlands **6**, 289 (1932).

[12] G. Camm, Mon. Not. Roy. Astron. Soc **110**, 305 (1950).

[13] H. L. Wright, B. N. Miller, and W. E. Stein, Astrophys. Space Sci. **84**, 421 (1982).

[14] B. N. Miller, J. Stat. Phys. **63**, 291 (1991).

[15] C. J. Reidl, Jr. and B. N. Miller, Phys. Rev. E **48**, 4250 (1993).

[16] C. J. Reidl, Jr. and B. N. Miller, Phys. Rev. E **51**, 884 (1995).

Structure Formation in the Universe and Non-linear Physics

Jiro Soda

Department of Fundamental Sciences, FIHS
Kyoto University, Kyoto 606, Japan

Abstract

There are various cosmological situations where the nonlinear physics is important. In this paper, we would like to concentrate on the problem to explain the reason why the distribution of the galaxies shows the scaling behavior, $\xi(r) \sim r^{-1.8}$. First, the argument by Peebles is presented. To find the asymptotic form of the scaling function is the essence of his result. In the linear region, the free parameter in the self-similar solution is determined by assuming the initial power spectrum in the power law form $P(k) \sim k^n$. In the nonlinear region, the scaling function is determined by combining the self-similar solution with the pair conservation equation which is derived from BBGKY hierarchy. Consequently, the scaling exponent in the strong nonlinear regime is found as a function of the initial spectrum index n. Although this argument is attractive, the self-similar solution may not be stable under the small perturbations. Then, we will present the somewhat radical alternative attempt which uses the Burgers equation with noise term as a basic equation. Using the renormalization group method, the scaling exponents are calculated and related to the correlation functions. According to the properties of the noise, there are various possible exponents. We could regard these as the universal dynamics for the structure formation in the universe. We also discuss other connections between the structure formation in the universe and the nonlinear physics.

1 Introduction

Nowadays, cosmology has been considered as a subject of physics. One of the purpose of cosmology is to explain the present structure of the universe. Recent observations show various kind of structures, e.g., void, filament, etc.. It is usual to quantify these structures statistically. If one can know all of the n-point correlation functions, one can reconstruct the probability distribution functional of the density field. However, it is not possible to know all of them from the observational data. Hence, it is not necessary to calculate them theoretically, rather a good representative statistics is desired in practice. Usually, the 2-point correlation function or , equivalently, the power-spectrum is used for that purpose. It is well known that the two-point correlation function, $\xi(r)$, shows the scaling behavior in a certain range, i.e., $\xi(r) \sim r^{-1.8}$. However, it should be stressed that the two-point function is not enough to characterize the structure of the universe. Indeed, the CfA data shows the multi-fractality instead of the mono-fractality. Hence, we should consider various statistics to characterize the fertile structure of the universe. Many statistical quantities, such as count in cell, percolation and topology, etc. are invented. Different statistical descriptors measure different aspects of the clustering pattern revealed by a survey.

There are two aspects to understand the large scale structure of the universe, i.e., the quantification of the structure and the dynamical evolution of it. As for the quantification problem, we have mentioned already. Recent progress of the numerical apparatus have forced us to calculate the dynamics by computer. There is no doubt that it is tremendously strong tool to analyze the non-linear dynamics. However, it is desired to study the non-linear dynamics analytically, because it reveals the physics behind the numerical calculation. Of course, here, there are two possible attitude to explaining the present structure of the universe. One of them is to consider the present structure as a transient pattern, so it is meaningless to consider the scaling $\xi(r) \sim r^{-1.8}$ seriously. The other

possible attitude which we take in this paper is to consider the present structure as a consequence of the non-linear physics. As an example, we present two attempts to explain the scaling of the two-point correlation function. The argument by Peebles is well known. [1] As it is unique as an analytical result, we briefly review it. Peebles also discussed the renormalization group computation to investigate the large scale structure of the universe. [2] Recently, Berera and Fang [3] studied the structure formation problem by using the theory of the surface growth. They considered the reheating stage and concluded that the seed of density perturbation is generated by the stochastic fluctuations besides initial (inflationary) quantum fluctuations. Here, we consider the matter dominant stage. However, the spirit is the same, i.e., the structure formation in the universe can be classified as the universal dynamics. This phenomenological approach is our main concern.

The plan of this paper is as follows: In Sec.2, we review the BBGKY hierarchy and Peeble's argument of the scaling. This is the traditional method in a sense. In Sec.3, we would like to present the radical approach, i.e., KPZ phenomenology. Honestly speaking, we did not perform any new calculation. We only give the argument to connect known results. In the final section, other topics are discussed.

2 Peebles's argument

The theory of self-gravitating matter distributions is inherently nonlinear, and nonlinear problems have always been notoriously difficult to attack analytically. Given this, and the availability of modern first computing hardware and efficient algorithms for large scale N-body simulation, it is little surprise that for the last decade, analytic study of large scale cosmological structure in the full nonlinear regime has more or less dried up. In these circumstances, the argument based on BBGKY hierarchy [1] is valuable to consider seriously. Let us start with the fluid limit approximation. The basic

equation is the Vlasov equation

$$\frac{\partial f}{\partial t}+\frac{p^i}{ma^2}\frac{\partial f}{\partial x^i}-\frac{Gm^2}{a}\int d^3x'd^3p'\frac{x^i-x'^i}{|x-x'|^3}f(x',p')\frac{\partial f(x,p)}{\partial p^i}=0\ , \quad (1)$$

BBGKY hierarchy is obtained by taking the ensemble average of the above equation in the following way:

$$\frac{\partial f_N}{\partial t}+\frac{p^i}{ma^2}\frac{\partial f_N}{\partial x^i}-\frac{Gm^2}{a}\int d^3x'd^3p'\frac{x^i-x'^i}{|x-x'|^3}\frac{\partial f_{N+1}(x,p)}{\partial p^i}=0\ , \quad (2)$$

where $f_N=\underbrace{<f\cdots f>}_{N}$.

As the lowest equation, we obtain the pair conservation equation

$$\frac{\partial \xi}{\partial t}+\frac{1}{a}\frac{\partial}{\partial x^i}[(1+\xi)<v^i>]=0 \quad (3)$$

or , using $<v^i>=vx^i/x$,

$$\frac{\partial \xi}{\partial t}+\frac{1}{x^2a}\frac{\partial}{\partial x}[x^2(1+\xi)v]=0\ . \quad (4)$$

The self-similar solution of the BBGKY hierarchy takes the form

$$\xi(x,t)=\hat{\xi}(\frac{x}{t^\alpha})\ , \quad (5)$$

where $\hat{\xi}$ is the scaling function. To find the asymptotic form of this scaling function is the essence of the following discussion. In the linear region, it is determined by assuming the power spectrum, $P(k)\sim k^n$, as a initial condition. In the nonlinear region, it is determined by combining the self-similar solution with the pair conservation equation which is derived from BBGKY hierarchy. Consequently, the scaling exponent in the strong nonlinear regime is found as a function of the spectrum index n.

It is possible to connect the scaling exponent α with the initial condition, i.e., spectrum index n. Indeed,

$$<|\delta_k|^2> \sim \int d^3x\hat{\xi}(\frac{x}{t^\alpha})\exp ikx \quad (6)$$

$$\sim t^{3\alpha}\int d^3y\hat{\xi}(y)\exp it^\alpha ky \quad (7)$$

$$\sim t^{3\alpha}F(t^\alpha k) \quad (8)$$

$$\sim t^{\frac{4}{3}}k^n\ . \quad (9)$$

By comparison, it is easy to find

$$\alpha = \frac{4}{3n+9} \ . \tag{10}$$

Furthermore, if we consider the strong non-linear region,i.e., $\xi \gg 1$, it is expected that the peculiar velocity compensate the Hubble flow

$$v \sim -\dot{a}x \ . \tag{11}$$

Substituting the above equation into eq.(4), we can solve as

$$\xi(x,t) = a^3 g(ax) \ , \tag{12}$$

where g is the arbitrary function. By comparing eq.(12) with the self-similar solution (5), the scaling function is determined as

$$\xi(x,t) \sim (\frac{x}{t^\alpha})^{-\frac{6}{3\alpha+2}} \tag{13}$$

$$\sim r^{-\frac{9+3n}{5+n}} . \tag{14}$$

where we have used the relation (10).

The result mentioned above is very attractive in the sense that we can predict the scaling exponent in the non-linear region. However, the self-similarity is too restrictive to be plausible. Indeed, there is a work which shows that the self-similarity solution is not stable against the small perturbation. [4] So we will consider the dynamical scaling instead of the kinematical scaling in the next section.

3 KPZ phenomenology

In this section, we shall start with the basic equation for a self-gravitating dust fluid. As we are considering the density fluctuations within a Hubble scale, the Newtonian approximation is valid. Moreover, we will use the fluid approximation. The basic equations are

$$\frac{\partial \rho}{\partial t} + 3\frac{\dot{a}}{a}\rho + \frac{1}{a}\vec{\nabla}\cdot(\rho\vec{v}) = 0 \ , \tag{15}$$

$$\frac{\partial \vec{v}}{\partial t} + \frac{1}{a}\vec{v}\cdot\vec{\nabla}\vec{v} + \frac{\dot{a}}{a}\vec{v} = -\frac{1}{a}\vec{\nabla}\phi \ , \tag{16}$$

$$\nabla^2\phi = 4\pi G a^2 \rho_b \delta \ . \tag{17}$$

Let us define new variables:

$$\bar{\rho} = \rho a^3 \ , \vec{u} = \frac{\vec{v}}{a\dot{a}} \ . \tag{18}$$

Then eq.(15) becomes

$$\frac{\partial \bar{\rho}}{\partial a} + \vec{\nabla} \cdot (\rho \vec{u}) = 0 \ . \tag{19}$$

Also, eq.(16) yields

$$\frac{\partial \vec{u}}{\partial a} + \vec{u} \cdot \vec{\nabla} \vec{u} = -\frac{1}{a\dot{a}^2}(4\pi G a^2 \rho_b \vec{u} + \frac{1}{a}\vec{\nabla}\phi) \equiv \vec{P} \ . \tag{20}$$

Notice that

$$\vec{\nabla} \cdot \vec{P} = 4\pi G a[\rho - \rho_b(1 - a\vec{\nabla} \cdot \vec{u})] = 0 \ . \tag{21}$$

The right hand side vanishes if

$$\delta = -a\vec{\nabla} \cdot \vec{u} \ . \tag{22}$$

In this case, eq.(20) becomes

$$\frac{\partial \vec{u}}{\partial a} + \vec{u} \cdot \vec{\nabla} \vec{u} = 0 \ . \tag{23}$$

If we consider one-dimensional case, eq.(19) is

$$\frac{\partial}{\partial x}[\frac{\partial u}{\partial a} + u \nabla u] = 0 \ . \tag{24}$$

thus we obtained the Zeldovich exact solution. [5] In case of general dimensions, of course, it is merely approximation. In the dust fluid case, a caustics will arise soon. Hence, we artificially, include the viscosity term to circumvent the caustics.

$$\frac{\partial \vec{u}}{\partial a} + \vec{u} \cdot \vec{\nabla} \vec{u} = \nu \nabla^2 \vec{u} \ . \tag{25}$$

It seems legitimate to consider this procedure does not influence the large scale structure on which we are concentrating.

In the next step, we will add the noise term by hand. The physical origin of this noise is not discussed in this paper (It may be supernova explosion). Here, we shall consider the resulting equation as the phenomenological equation. When the dissipation is taken into account, it is natural to consider the noise. In this case, the velocity field becomes the stochastic variable. Remember that the cosmological fields should be the stochastic variables. Let us consider the noise effect in the following form:

$$\frac{\partial \vec{u}}{\partial a} + \vec{u} \cdot \vec{\nabla} \vec{u} = \nu \nabla^2 \vec{u} - \vec{\nabla} \eta \ , \tag{26}$$

where

$$< \eta(x,t)\eta(x',t') >= 2D\delta^d(x-x')\delta(t-t') \ . \tag{27}$$

If we perform the transformation of variable,

$$\vec{u} = -\vec{\nabla} h, \tag{28}$$

we obtain

$$\frac{\partial h}{\partial a} = \frac{1}{2}\lambda(\nabla h)^2 + \nu \nabla^2 h + \eta \tag{29}$$

This eq.(29) is well known as KPZ equation in the surface theory. [6, 7]

To begin with, we shall consider the linear theory

$$\frac{\partial h}{\partial t} = \nu \nabla^2 h + \eta \ . \tag{30}$$

If the function $h(x,t)$ is self-affine, then on rescaling it

$$x \to x' = bx, h \to h' = b^{\alpha} h, \tag{31}$$

we should obtain the function that is statistically indistinguishable from the original one. Since the function depends on time t as well, to compare two functions obtained at different moments we must also rescale the time,

$$t \to t' = b^z t. \tag{32}$$

Straightforward substitution yields

$$b^{\alpha - z}\frac{\partial h}{\partial t} = \nu b^{\alpha-2}\nabla^2 h + b^{-\frac{\alpha}{2}-\frac{z}{2}}\eta \ , \tag{33}$$

here we used the fact

$$< \eta(bx, b^z t)\eta(bx', b^z t') > = 2D\delta^d(bx - bx')\delta(b^z t - b^z t') \quad (34)$$
$$= 2Db^{-(d+z)}\delta^d(x - x')\delta(t - t') \ . \quad (35)$$

Multiplying both sides of eq.(35) with $b^{z-\alpha}$, we obtain

$$\frac{\partial h}{\partial t} = \nu b^{z-2}\nabla^2 h + b^{-\frac{\alpha}{2}+\frac{z}{2}-\alpha}\eta \ . \quad (36)$$

By imposing the scale invariance, we find

$$\alpha = \frac{2-d}{2}, \beta = \frac{2-d}{4}, z = 2. \quad (37)$$

To relate this results with the cosmological two-point function, notice that the relation (22) and (28). The result is

$$\xi \sim r^{-4+2\alpha}. \quad (38)$$

This expression gives the cosmological 2-point function, when we know the exponent α. In the linear theory, as is shown above, a simple scaling argument gives the value $\alpha = 1/2$. While, in the nonlinear theory, it does not work well, then we need more elaborate tool. Below, we will explain the renormalization group method as such a tool.

Now we shall consider the full nonlinear equation (29). It is convenient to work in the Fourier space. Using the Fourier transformation

$$h(\vec{k}, t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int^{k<\Lambda} \frac{d^d k}{(2\pi)^d} h(\vec{k}, \omega) \exp[i(\vec{k} \cdot \vec{x} - \omega t)] \ . \quad (39)$$

eq.(29) can be solved perturbatively. The lowest solution is given by

$$h_0(\vec{k}, \omega) = G_0(\vec{k}, \omega)\eta(\vec{k}, \omega) \ , \quad (40)$$

where

$$G_0(\vec{k}, \omega) = \frac{1}{\nu k^2 - i\omega} \ . \quad (41)$$

Formally, we have the integral equation

$$h(\vec{k},\omega) = G(\vec{k},\omega)\eta(\vec{k},\omega) \tag{42}$$

$$= G_0(\vec{k},\omega)\eta(\vec{k},\omega) \tag{43}$$

$$-\frac{\lambda}{2}G_0(\vec{k},\omega)\int\frac{d^dqd\Omega}{(2\pi)^{d+1}}\vec{q}\cdot(\vec{k}-\vec{q})h(\vec{q},\Omega)h(\vec{k}-\vec{q},\omega-\Omega)$$

To improve the perturbative solution, the renormalization group method is useful and, indeed, widely used in the quantum field theoretical problems. It is the renormalization group equation that is used to predict the scaling exponent of the cosmological 2-point function. Practically, we take the following scheme. Divide the momentum integral, $k \in [0,\Lambda]$, into two parts, one with high momentum, $k^> \in [\Lambda/b,\Lambda]$ and with low momenta, $k^< \in [0,\Lambda/b]$. The resulting integrals have a cut off Λ/b. Then, we shall rescale the system using $k \to bk$. At the same time, we must do rescaling, $h \to b^\alpha h$ and $t \to b^z t$. This procedure defines the renormalization group transformation. Now, let us move to concrete calculation. Straightforward perturbative calculation yields

$$G(\vec{k},0) = G_0(\vec{k},0) + \frac{\lambda^2 D}{\nu^2}G_0^2(\vec{k},0)\frac{d-2}{4d}k^2K_d\int dqq^{d-3} \tag{44}$$

Since $G_0(\vec{k},0) = 1/\nu k^2$, we define the effective viscosity as $G(\vec{k},0) = 1/\tilde{\nu}k^2$, then

$$\tilde{\nu} = \nu[1 - \frac{\lambda^2 D}{\nu^2}\frac{d-2}{4d}K_d\int dqq^{d-3}], \tag{45}$$

where $K_d = S_d/(2\pi)^2$ and S_d is the volume of d-dimensional sphere. After integrating out the high momentum, we obtain

$$\nu^< = \nu[1 - \delta l\frac{\lambda^2 D}{\nu^2}\frac{d-2}{4d}K_d], \tag{46}$$

where we have used $\delta l = \log b$ for convenience. The rescaling complete the renormalization group transformation

$$\tilde{\nu} = b^{z-2}\nu^< \tag{47}$$

$$= \nu^<[1 + \delta l(z-2)] \tag{48}$$

$$= \nu[1 + \delta l\{z - 2 + \frac{\lambda^2 D}{\nu^2}\frac{2-d}{4d}K_d] \tag{49}$$

Thus, we find the renormalization group equation

$$\frac{d\nu}{dl} = \nu[z - 2 + \frac{\lambda^2 D}{\nu^2}\frac{2-d}{4d}K_d]. \tag{50}$$

Due to Galilean invariance, there is no correction to the coupling constant λ. Hence, we obtain

$$\frac{d\lambda}{dl} = \lambda[\alpha + z - 2]. \tag{51}$$

Finally, the effective diffusion coefficient, $\tilde{D}$, is defined as

$$< h^*(\vec{k},\omega)h(\vec{k},\omega) >= 2\tilde{D}G_0(\vec{k},\omega)G_0(-\vec{k},-\omega). \tag{52}$$

Perturbative correction

$$D^< = D[1 + \delta l\frac{\lambda^2 D}{\nu^2}\frac{K_d}{4}] \tag{53}$$

and the rescaling

$$\tilde{D} = D^<[1 + \delta l(z - d - 2\alpha)] \tag{54}$$

gives the renormalization group equation for D as

$$\frac{dD}{dl} = D[z - d - 2\alpha + \frac{\lambda^2 D}{\nu^2}\frac{K_d}{4}]. \tag{55}$$

Thus, a set of the renormalization group equations is given by

$$\frac{d\nu}{dl} = \nu[z - 2 + K_d g^2\frac{2-d}{4d}] \tag{56}$$

$$\frac{dD}{dl} = D[z - d - 2\alpha + K_d\frac{g^2}{4}] \tag{57}$$

$$\frac{d\lambda}{dl} = \lambda[\alpha + z - 2]. \tag{58}$$

where $g^2 = \lambda^2 D/\nu^2$. The exponents can be obtained by searching for the fixed points of the flow equation, i.e., by considering

$$\frac{d\nu}{dl} = \frac{dD}{dl} = \frac{d\lambda}{dl} = 0\ . \tag{59}$$

Using eqs.(56) , (57) and (58), we can calculate the flow of the coupling constant

$$\frac{dg}{dl} = \frac{2-d}{2}g + K_d\frac{2d-3}{4d}g^3 \ . \tag{60}$$

The fixed point of this equation is obtained by taking $dg/dl = 0$.

For $d = 1$, there are two fixed points:

$$g_1^* = 0 \ , g_2^* = (\frac{2}{K_d})^{\frac{1}{2}} \ . \tag{61}$$

In this case, the relevant one is g_2^*. At the fixed point g_2^*, we obtain

$$z = \frac{3}{2}, \alpha = \frac{1}{2} \ . \tag{62}$$

Consequently, we find

$$\xi(r) \sim r^{-3} \ . \tag{63}$$

Of course, we would like to know the scaling exponent in the real world. In $d = 3$ case, unfortunately, there is no infrared fixed point in the weak region. The fixed point which we are interested in should lie on the strong coupling region. Therefore, the non-purturbative method is necessary to obtain the desired result. Thus, we encounter the difficulty again. If we consider the colored noise, however, the situation is very different [8] and it is easy to derive the exponent which coincides with the observed value. If we consider the power law distribution of the noise instead of the Gaussian distribution, we obtain the multi-fractal distribution. Therefore, it may be possible to explain the observed multi-fractal structure in the universe along this line of thought. As the result depends on the nature of the noise, the physical origin of the noise should be clarified. This task is the beyond of the scope of this paper. Further investigation to this direction must be interesting.

4 Discussion

We have discussed the possible physical explanation for the large scale structure. In particular, we have concentrated on the 2-point function which shows the scaling observationally.

First, we reviewed Peebles's argument which relies on the self-similar solution of the BBGKY hierarchy. If we assume the power law spectrum, $P(k) \sim k^n$, the correlation function becomes $\xi(r) \sim r^{-9+3n/5+n}$ in the strong non-linear region $v = -\dot{a}x$. However, the above self-similar solution might be unstable against the small perturbation.

As an alternative approach, we propose the phenomenological equation in which the noise term is artificially included. In the statistical physics, the concept of the universality class is useful to understand critical phenomena. There is also an attempt to study the quantum gravity as a universality class in the field theory. Our strategy is similar in the sense that we take effective theory to understand the cosmological dynamical scaling phenomena. As we have seen in this paper, there are many universality classes according to the nature of the noise. Observation determines which universality class our univer choose. We can also find another example in quantum chromodynamics. There, the effective theories are useful to understand the confinement phenomena.

As for other topics in non-linear physics, wavelet analysis is interesting as a data analysis method and an analytical tool for non-linear analysis. Many fundamental properties of physical systems have been described in terms of Fourier spectrum. However, since Fourier spectrum ignores the phase of each Fourier coefficients, it lacks information about positions of local events. Wavelet analysis is invented to overcome this problem. In Fourier analysis, the basis functions of expansion are the familiar sines and cosines. In wavelet analysis, [9, 10] the bases functions are somewhat more complicated localized functions, the so-called wavelet. By using this new method, various statistics can be defined. It should be noted that the wavelet analysis is especially powerful in the analysis of multi-fractal structure. [11] Hence, the nonlinear analysis on the basis of the wavelet analysis might be available as a tool for investigating the large scale structure of the universe.

The reductive perturbation method is also attractive as a non-linear method. We have previously studied the quasi-nonlinear evo-

lution of the density perturbation in Newtonian gravity by using the reductive perturbation method. [12] Especially, weak mode-mode coupling in a small range below Jeans wavelength is considered. We have shown that the basic equation for the acoustic wave reduces to a nonlinear Schrodinger equation. As a result, it turns out that there is a new kind of instability in the self-gravitating systems.

It should be stressed that the non-linear analytical method is important for future development of understanding the large scale structure of the universe.

Acknowledgments

The work is supported by MOnbusho Grant-in-Aid for Scientific Research No. 07740216.

References

[1] P.J.E.Peebles, The Large-Scale Structure of the Universe (Princeton University Press, 1980).

[2] P.J.E.Peebles, Ap. J. 297 (1985) 350.

[3] A. Berera and L.Z.Fang, Phys. Rev. Lett. 72 (1994) 458.

[4] L. Ruamsuwan and J.N. Fry, Ap.J. 396 (1992) 416.

[5] Ya.B.Zeldovich, Astrofizila 6 (1970) 119; J.Soda and Y.Suto, Ap. J. 396 (1992) 379; J.Soda, Prog. Theor. Phys. 90 (1993) 569.

[6] A.L.Barabasi and H.E.Stanley, Fractal concepts in Surface Growth (Cambridge University Press, 1995).

[7] M.Kardar, G.Parisi and Y.C.Zhang, Phys.Rev.Lett. 56 (1986) 889.

[8] E.Medina, T. Hwa, M.Kardar and Y.C. Zhang, Phys. Rev. A 39 (1989) 3053.

[9] C.K Chui, An introduction to Wavelet, (Academic Press, 1992).

[10] Y.Fujiwara and J. Soda, Wavelet analysis of one-dimensional cosmological fluctuations, KUCP-0082 (1995).

[11] J. F. Muzy, E. Bacry and A.Arneodo, International Journal of Bifurcation and Chaos, Vol.4, No.2 (1994) 245.

[12] Y. Fujiwara and J. Soda, Non-linear self-modulation in Newtonian Gravity, KUCP-0080 (1995).

Post-Zel'dovich approximation in general relativity

Masaaki Morita

Department of Physics, Hirosaki University, 3 Bunkyo-cho, Hirosaki, 036 Japan

Abstract

The gravitational instability of irrotational dust in the expanding universe is studied by the relativistic second-order approximation. Introducing the orthonormal tetrad, we obtain equations very similar to those given in the Lagrangian perturbation theory in Newtonian cosmology. The solutions of the nonlinear dynamics of cosmological perturbations are presented in the Einstein-de Sitter background. We propose a relativistic version of the Zel'dovich approximation and its higher-order corrections, say, "post-Zel'dovich" approximation.

1 Introduction

The investigation of the nonlinear clustering based on the gravitational instability scenario is quite important in the research of the large-scale structure formation in the universe. By using N-body codes it is possible to follow the general nonlinear evolution of initially small perturbations numerically, but an understanding of what has happened between input and output can often better be gained by analytical treatments. Perturbative approaches such as the linear perturbation theory in Newtonian cosmology [1] are well-known and often used as conventional analytic methods. These approaches, however, depend on the assumption that the absolute value of the density fluctuation $|\delta|$ is much smaller than unity, whereas in the real universe $|\delta| > 1$ up to the scale of superclusters. Zel'dovich [2] considered an approximation scheme which does not depend on the assumption. This scheme is known as the Zel'dovich approximation, which is now widely applied to the problems of the large-scale structure formation. Various analytical approaches have been compared with the numerical results statistically [3, 4, 5] and it has turned out that the Zel'dovich approximation gives the best fit to the numerical treatments. Buchert [6] developed the Lagrangian perturbative approximation and showed that the Zel'dovich approximation can be regarded as a subclass of its first-order

solutions. This work was extended to second-order [7] and even to third-order [8], giving some new useful information about self-gravitating systems.

But these and most other analytical treatments are Newtonian approaches, which are valid only for perturbations on scales much smaller than the horizon size. On super-horizon scales one needs a relativisic approach instead. The pioneering work was done by Lifshitz [9]. This work was extended to second-order by Tomita [10, 11, 12]. Yet in some of these works there remained a gauge problem that unphysical perturbations were included in the solutions. This problem was solved by the gauge-invariant formalism [13, 14]. The relativistic approaches mentioned above, however, again use the assumption $|\delta| \ll 1$. In order to develop a more realistic description for an inhomogeneous universe, Futamase [15, 16, 17] formulated the so-called cosmological post-Newtonian approximation to describe a clumpy ($|\delta| \gg 1$) universe. In his scheme the spatial averaging is introduced to get a background metric. But the relation of his approach to the linear perturbation theory is not very clear, partly because he adopted the harmonic gauge condition. Therefore, it is desirable to construct a scheme which does not use the assumption $|\delta| \ll 1$ and reproduce the linear perturbation theory in the limit $|\delta| \ll 1$.

Matarrese *et al.* [18] considered a "relativistic Lagrangian approach" in the case of vanishing the magnetic part of the Weyl tensor $H_{ab} = 0$ in the first-order, based on the fluid flow approach. It is extended to second-order [19, 20]. They claim some similarity to the Zel'dovich approximation, but their formulation is quite complicated even under the restriction $H_{ab} = 0$. Moreover, in the second-order H_{ab} inevitably appear. Thus, the second-order calculation by their formulation becomes complicated all the more.

Parry *et al.* [21] presented the nonlinear solution of the Hamilton-Jacobi equation for general relativity and reproduced the Zel'dovich approximation, using the gradient expansion technique. The "higher-order Zel'dovich approximation" is discussed in Refs. [22, 23]. In spite of its elegance, this method is classified as "long-wavelength approximation" and is applicable only when the characteristic scale of spatial fluctuations is much larger than the horizon size.

In this article, we develop a relativistic Zel'dovich-type approximation without any restrictions such as vanishing H_{ab} and the wavelength-specific assumption, based on the works by Kasai [24] and Russ *et al.* [25]. By using the tetrad formalism, we derive fully general relativistic equations very similar to those given in the Newtonian case, which

are solved in the Einstein-de Sitter background by an iteration method up to second-order. A general relativistic version of the Zel'dovich approximation is proposed and its higher-order corrections are also considered. Following Munshi *et al.* [26], we call this scheme "post-Zel'dovich" approximation.

2 Exposition of the Method

First, we summarize a general relativistic treatment to describe the nonlinear evolution of an inhomogeneous universe [24, 27, 28]. The models we consider contain irrotational dust with density ρ and four-velocity u^μ (and possibly a cosmological constant Λ). Neglecting the fluid pressure and the vorticity is a reasonable assumption in a cosmological context. In comoving synchronous coordinates, the line element can be written in the form

$$ds^2 = -dt^2 + g_{ij}dx^i dx^j \tag{1}$$

with $u^\mu = (1,0,0,0)$. Then the Einstein equations read

$$\frac{1}{2}\left[{}^3R^i{}_i + \left(K^i{}_i\right)^2 - K^i{}_j K^j{}_i \right] = 8\pi G\rho + \Lambda \,, \tag{2}$$

$$K^i{}_{j\|i} - K^i{}_{i\|j} = 0 \,, \tag{3}$$

$$\dot{K}^i{}_j + K^k{}_k K^i{}_j + {}^3R^i{}_j = (4\pi G\rho + \Lambda)\,\delta^i{}_j \,, \tag{4}$$

where ${}^3R^i{}_j$ is the three-dimensional Ricci tensor, $K^i{}_j = \frac{1}{2}g^{ik}\dot{g}_{jk}$ is the extrinsic curvature, $\|$ denotes the covariant derivative with respect to the three-metric g_{ij}, and an overdot ( ) denotes $\partial/\partial t$. The energy equation $u_\mu T^{\mu\nu}{}_{;\nu} = 0$ gives

$$\dot{\rho} + \rho K^i{}_i = 0 \tag{5}$$

with the solution

$$\rho = \rho(t_{in}, \boldsymbol{x}) \left(\frac{\det\left[g_{ij}(t_{in}, \boldsymbol{x})\right]}{\det\left[g_{ij}(t, \boldsymbol{x})\right]} \right)^{\frac{1}{2}} . \tag{6}$$

Let us introduce the scale factor function $a(t)$ which satisfies the Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\rho_b + \frac{\Lambda}{3} \,, \tag{7}$$

where the curvature constant k takes the value of $+1, 0, -1$ for closed, flat, and open spaces, respectively. If the spacetime is exactly Friedmann-Lemaître-Robertson-Walker

(FLRW), then we have

$$K^i{}_j = \frac{\dot{a}}{a}\delta^i{}_j, \qquad {}^3R^i{}_j = \frac{2k}{a^2}\delta^i{}_j \qquad \text{for FLRW}. \tag{8}$$

Therefore, the deviations from the FLRW models due to inhomogeneity are expressed by the peculiar part of the extrinsic curvature

$$V^i{}_j \equiv K^i{}_j - \frac{\dot{a}}{a}\delta^i{}_j\,, \tag{9}$$

which represents the deviation from the uniform Hubble expansion, and the deviation of the curvature tensor

$$\mathcal{R}^i{}_j \equiv a^2\,{}^3R^i{}_j - 2k\delta^i{}_j\,. \tag{10}$$

The procedure essential to develop the relativistic Zel'dovich-type approximation [24] is to introduce the following orthonormal tetrad

$$g_{\mu\nu} = \eta_{(a)(b)}\bar{e}^{(a)}_{\mu}\bar{e}^{(b)}_{\nu} \tag{11}$$

with

$$\bar{e}^{(0)}_{\mu} = u_\mu = (-1,0,0,0)\,, \qquad \bar{e}^{(\ell)}_{\mu} = \left(0,\bar{e}^{(\ell)}_{i}\right) \equiv \left(0,a(t)\,e^{(\ell)}_{i}\right) \quad \text{for } \ell = 1,2,3. \tag{12}$$

The spatial basis vectors are parallelly transported along each fluid line, i.e.,

$$\bar{e}^{(\ell)}_{\mu;\nu}\,u^\nu = 0\,. \tag{13}$$

In our choice of the tetrad components, it reads

$$\dot{e}^{(\ell)}_{i} = V^j_i\,e^{(\ell)}_{j} \qquad \text{or} \qquad V^i_j = e^i_{(\ell)}\,\dot{e}^{(\ell)}_{j}\,. \tag{14}$$

The Einstein equations can be rewritten in terms of these quantities as the following key equation

$$\frac{\partial}{\partial t}\left[a^3\left(\ddot{e}^{(\ell)}_{i} + 2\frac{\dot{a}}{a}\dot{e}^{(\ell)}_{i} - 4\pi G\rho_b e^{(\ell)}_{i}\right)\right] = a\left(P^{(\ell)}_{i} + Q^{(\ell)}_{i} + S^{(\ell)}_{i}\right)\,, \tag{15}$$

where

$$P^{(\ell)}_{i} = \frac{\partial}{\partial t}\left\{\frac{a^2}{4}\left[\left(V^k_k\right)^2 - V^k_n V^n_k\right]e^{(\ell)}_{i} - a^2\left(V^k_k V^j_i - V^j_k V^k_i\right)e^{(\ell)}_{j}\right\}\,, \tag{16}$$

$$Q^{(\ell)}_{i} = \left(V^j_k\mathcal{R}^k_i + \frac{1}{4}V^j_i\mathcal{R}^k_k - \frac{1}{2}\delta^j_i\,V^k_n\mathcal{R}^n_k\right)e^{(\ell)}_{j}\,, \tag{17}$$

and

$$S^{(\ell)}_i = \left[V^j_{\ i}{}^{|k}_{|k} + V^k_{\ k}{}^{|j}_{|i} - V^j_{k|i}{}^{|k} - V^k_{\ i}{}^{|j}_{|k} + k\left(3V^j_{\ i} - V^k_{\ k}\delta^j_{\ i}\right)\right] e^{(\ell)}_j \,. \tag{18}$$

Note that the left-hand-side of Eq. (15) is already linearized with respect to $e^{(\ell)}_i$, and all terms on the right-hand-side, except $S^{(\ell)}_i$, are manifestly nonlinear quantities. It has, therefore, a form suitable for solving it perturbatively by iteration. It should also be stressed that we have not used any approximation methods in deriving Eq. (15). Our treatment here is fully nonlinear and exact.

3 Perturbative approach

We solve perturbatively the key equation Eq. (15) by an iteration method. Hereafter, we restrict our consideration to the Einstein-de Sitter background, $k = \Lambda = 0$.

3.1 The first-order solutions

Linear perturbations are classified into scalar, vector, and tensor modes. In the first-order level, they do not couple with each other, and can be discussed separately. Here let us consider the scalar perturbations. The general form for the linearly perturbed triad in this case is

$$e^{(\ell)}_i = \delta^{(\ell)}_i + E^{(\ell)}_i = \delta^{(\ell)}_i + \delta^{(\ell)}_j \left(A\,\delta^j_{\ i} + B^{,j}_{\ ,i}\right) \,. \tag{19}$$

Let us write the first-order quantities with subscript (1). The first-order perturbed extrinsic curvature is

$$V^i_{(1)j} = \dot{B}^{,i}_{\ ,j} \,. \tag{20}$$

As was noted previously, it is apparent that the source terms $P^{(\ell)}_i$ and $Q^{(\ell)}_i$ are second-order quantities (and higher). Using Eq. (20), we also find that $S^{(\ell)}_i$ vanishes in linear order:

$$S^{(\ell)}_{(1)i} = \delta^{(\ell)}_j \left(V^j_{(1)i}{}^{,k}_{\ ,k} + V^k_{(1)k}{}^{,j}_{\ ,i} - V^j_{(1)k,i}{}^{,k} - V^k_{(1)i}{}^{,j}_{\ ,k}\right) = 0 \,. \tag{21}$$

Therefore, to first-order, the right-hand-side of the key equation (15) vanishes and it can be integrated to give

$$a^3 \left(\ddot{E}^{(\ell)}_i + 2\frac{\dot{a}}{a}\dot{E}^{(\ell)}_i - 4\pi G\rho_b E^{(\ell)}_i\right) = C^{(\ell)}_i(\boldsymbol{x}) \,. \tag{22}$$

Note that now it has the same form as the equation which governs the density contrast δ in conventional linear perturbation theory [1]. Using the growing mode $D^+(t) = a(t) = t^{2/3}$ and the decaying mode solutions $D^-(t) = t^{-1}$ respectively, we obtain the solutions in the form

$$B^i{}_j = t^{\frac{2}{3}}\,\Psi^i{}_{,j}(\boldsymbol{x}) + t^{-1}\,\Phi^i{}_{,j}(\boldsymbol{x})\,. \tag{23}$$

The scalar function A can be expressed in terms of $\Psi(\boldsymbol{x})$ as $A = \frac{10}{9}\Psi(\boldsymbol{x})$. Thus the first-order solution for the triad is

$$e^{(\ell)}_i = \left(1 + \frac{10}{9}\Psi(\boldsymbol{x})\right)\delta^{(\ell)}_i + \delta^{(\ell)}_j\left(t^{\frac{2}{3}}\,\Psi^{,j}_{,i}(\boldsymbol{x}) + t^{-1}\,\Phi^{,j}_{,i}(\boldsymbol{x})\right)\,. \tag{24}$$

Note that we have not assumed that the density contrast is small, in order to derive the solutions. Then, the Zel'dovich-type approximation can be composed from Eq. (6):

$$\rho = \rho_b\,\frac{\det[\,e^{(\ell)}_i(t_{in},\boldsymbol{x})\,]}{\det[\,e^{(\ell)}_i(t,\boldsymbol{x})\,]} \simeq \rho_b\,\frac{\det\left[\delta^{(\ell)}_i + E^{(\ell)}_i(t_{in},\boldsymbol{x})\right]}{\det\left[\delta^{(\ell)}_i + E^{(\ell)}_i(t,\boldsymbol{x})\right]}\,. \tag{25}$$

3.2 The second-order solutions

In order to avoid notational complexity, in this subsection we only deal with growing mode terms. Thus we begin with the following form

$$e^{(\ell)}_i = \left(1 + \frac{10}{9}\Psi(\boldsymbol{x})\right)\delta^{(\ell)}_i + t^{\frac{2}{3}}\,\delta^{(\ell)}_j\Psi^{,j}_{,i}(\boldsymbol{x}) + \varepsilon^{(\ell)}_i\,. \tag{26}$$

The second-order quantity $\varepsilon^{(\ell)}_i$ is decomposed into a transverse-traceless part and a remaining longitudinal part

$$\varepsilon^{(\ell)}_i = \delta^{(\ell)}_j\left(\beta^j{}_i + \chi^j{}_i\right)\;;\quad \chi^i{}_{j,i} = 0,\ \chi^i{}_i = 0\,. \tag{27}$$

Let us pay our attention to the key equation (15). To second-order it reads

$$\ddot{\varepsilon}^{(\ell)}_i + 2\frac{\dot{a}}{a}\dot{\varepsilon}^{(\ell)}_i - 4\pi G\rho_b\varepsilon^{(\ell)}_i = \frac{1}{a^3}c^{(\ell)}_i(\boldsymbol{x}) + \frac{1}{a^3}\int^t a\left(P^{(\ell)}_{(2)i} + Q^{(\ell)}_{(2)i} + S^{(\ell)}_{(2)i}\right)dt\,. \tag{28}$$

It is apparent that the source terms $P^{(\ell)}_{(2)i}$ and $Q^{(\ell)}_{(2)i}$ are quadratic with respect to the first-order quantities, hence contain neither $\beta^i{}_j$ nor $\chi^i{}_j$. Furthermore, we find that the longitudinal part of $S^{(\ell)}_{(2)i}$ does not contain $\beta^i{}_j$. Therefore, solutions for $\beta^i{}_j$ can be written as a linear combination of the homogeneous solution and the inhomogeneous solution in the presence of the given source terms:

$$\beta^i{}_j = t^{\frac{2}{3}}\,\psi^i{}_j(\boldsymbol{x}) + t^{\frac{4}{3}}\,\varphi^i{}_j(\boldsymbol{x})\,. \tag{29}$$

Once we obtain the temporal dependency of the solutions, their spatial dependency, i.e., $\psi^i{}_j(\boldsymbol{x})$ and $\varphi^i{}_j(\boldsymbol{x})$ are determined. To second-order

$$\psi^i{}_j = \frac{5}{9}\Psi^{,k}\Psi_{,k}\delta^i{}_j - \frac{10}{9}\left(\Psi^2\right)^{,i}{}_{,j} , \tag{30}$$

$$\varphi^i{}_j = \frac{3}{7}\left(\mu^k{}_k\,\delta^i{}_j - 4\mu^i{}_j\right) , \tag{31}$$

where

$$\mu^i{}_j \equiv \frac{1}{2}\left(\Psi^{,k}{}_{,k}\Psi^{,i}{}_{,j} - \Psi^{,i}{}_{,k}\Psi^{,k}{}_{,j}\right) . \tag{32}$$

The equation for $\chi^i{}_j$ can also be obtained from Eq. (28). To second-order, it gives for the transverse-traceless part

$$\ddot{\chi}^i{}_j + 3\frac{\dot{a}}{a}\dot{\chi}^i{}_j - \frac{1}{a^2}\nabla^2\chi^i{}_j = \mathcal{S}^i{}_j , \tag{33}$$

where

$$\mathcal{S}^i{}_j = \frac{3}{7}\mu^k{}_k{}^{,i}{}_{,j} + \frac{3}{7}\left(\mu^k{}_k\,\delta^i{}_j - 4\mu^i{}_j\right)^{,\ell}{}_{,\ell} \tag{34}$$

is a transverse and traceless tensor: $\mathcal{S}^i{}_i = 0$, $\mathcal{S}^i{}_{j,i} = 0$. This shows that gravitational waves are induced even if there are initially scalar perturbations only. Eq. (33) can be solved by use of the retarded Green function.

Finally, the post-Zel'dovich approximation can be composed as follows:

$$\rho = \rho_b \frac{\det[\,e^{(\ell)}_i(t_{in},\boldsymbol{x})\,]}{\det[\,e^{(\ell)}_i(t,\boldsymbol{x})\,]} \simeq \rho_b \frac{\det\left[\delta^{(\ell)}_i + E^{(\ell)}_i(t_{in},\boldsymbol{x}) + \varepsilon^{(\ell)}_i(t_{in},\boldsymbol{x})\right]}{\det\left[\delta^{(\ell)}_i + E^{(\ell)}_i(t,\boldsymbol{x}) + \varepsilon^{(\ell)}_i(t,\boldsymbol{x})\right]} . \tag{35}$$

4 Concluding remarks

In this article, we have developed the second-order perturbative approach to the non-linear evolution of irrotational dust universes in the framework of general relativity. We have presented the second-order solutions in a $k = 0$, $\Lambda = 0$ background, based on the tetrad formalism given by Kasai [24]. A relativistic version of the Zel'dovich-type approximation have been also proposed, which reproduces the conventional perturbation theories in the limit $|\delta| \ll 1$.

In our approach the extensions to $k \neq 0$, $\Lambda \neq 0$ cases and radiation universes ($p = \frac{1}{3}\rho$) are straightforward. These will be the subjects of future investigation.

References

[1] P. J. E. Peebles, *The Large Scale Structure of the Universe* (Princeton Univ. Press, 1980).

[2] Ya. B. Zel'dovich, Astron. Astrophys. **5**, 84 (1970).

[3] P. Coles, A. L. Melott, and S. F. Shandarin, Mon. Not. R. Astron. Soc. **260**, 765 (1993).

[4] A. L. Melott *et al.*, Mon. Not. R. Astron. Soc. **269**, 626 (1994).

[5] B. S. Sathyaprakash *et al.*, Mon. Not. R. Astron. Soc. **275**, 463 (1995).

[6] T. Buchert, Mon. Not. R. Astron. Soc. **254**, 729 (1992).

[7] T. Buchert and J. Ehlers, Mon. Not. R. Astron. Soc. **264**, 375 (1993).

[8] T. Buchert, Mon. Not. R. Astron. Soc. **267**, 811 (1994).

[9] E. M. Lifshitz, J. Phys. (Moscow) **10**, 116 (1946).

[10] K. Tomita, Prog. Theor. Phys. **37**, 831 (1967).

[11] K. Tomita, Prog. Theor. Phys. **45**, 1747 (1971).

[12] K. Tomita, Prog. Theor. Phys. **47**, 416 (1972).

[13] J. M. Bardeen, Phys. Rev. D **22**, 1882 (1980).

[14] H. Kodama and M. Sasaki, Prog. Theor. Phys. Suppl. **78**, 1 (1984).

[15] T. Futamase, Phys. Rev. Lett. **61**, 2175 (1988).

[16] T. Futamase, Mon. Not. R. Astron. Soc. **237**, 187 (1989).

[17] T. Futamase, Prog. Theor. Phys. **86**, 389 (1991).

[18] S. Matarrese, O. Pantano, and D. Saez, Phys. Rev. D **47**, 1311 (1993).

[19] S. Matarrese, O. Pantano, and D. Saez, Phys. Rev. Lett. **72**, 320 (1994).

[20] S. Matarrese, O. Pantano, and D. Saez, Mon. Not. R. Astron. Soc. **271**, 513 (1994).

[21] J. Parry, D. S. Salopek, and J. M. Stewart, Phys. Rev. D **49**, 2872 (1994).

[22] K. M. Croudace *et al.*, Astrophys. J. **423**, 22 (1994).

[23] D. S. Salopek *et al.*, Mon. Not. Astron. Soc. **271**, 1005 (1994).

[24] M. Kasai, Phys. Rev. D **52**, 5605 (1995).

[25] H. Russ, M. Morita, M. Kasai, and G. Börner, Phys. Rev. D, in press.

[26] D. Munshi, V. Sahni, and A. A. Starobinsky, Astrophys. J. **436**, 517 (1994).

[27] M. Kasai, Phys. Rev. Lett. **69**, 2330 (1992).

[28] M. Kasai, Phys. Rev. D **47**, 3214 (1993).

A Spinning Test Particle in A Relativistic Space Time

- Chaos and its Effect on Gravitational Wave Detection -

SHINGO SUZUKI[†] and KEI-ICHI MAEDA[††]
Department Physics, Waseda University, Shinjuku-ku, Tokyo 169, Japan

Abstract

We study the motion of a spinning test particle in Schwarzschild spacetime, and find chaotic behavior for a rapidly spinning particle. The critical value of spin is about $S = 0.6mM$ for $J = 4mM$, where m and M are the mass of a test particle and of a Schwarzschild black hole, respectively. This chaos is induced by spin-orbit coupling. The possibility that the chaotic behavior of the spinning particle may affect detection of gravitational waves from a coalescing binary is also discussed.

1 Introduction

Starting with the famous three body problem, many studies about chaos in celestial mechanics have been done and some important chaotic phenomena have been found in the Universe. Now chaos has become one of the most fundamental ideas to explain various non-linear phenomena in nature.

In the Universe there exist systems which are in a strong gravity field and whose dynamics have to be analyzed with Einstein's theory of gravitation. Because the equations of motion of the object in general relativity are different from those in Newtonian dynamics and there may exist some relativistic effects which make the motion more complicated, we may find new types of chaos in a strong gravitational field, which does not appear in Newtonian dynamics.

If we find chaotic behavior in a general relativistic system, it may affect some experimental tests of general relativity. One such test is the detection of the general relativistic frame-dragging effect[1]. In this experiment the spin-orbit and spin-spin interaction between the Earth and a gyroscope on the satellite hope to be detected. In [1], the authors have made the theoretical analysis about such a system and shown some numerical results, in which we find

[†]e-mail address : 64L070@cfi.waseda.ac.jp
[††]e-mail address : maeda@cfi.waseda.ac.jp

a very strange motion of the gyroscope. It is natural to think that such a strange motion is induced by spin effects.

Another test of general relativity is the detection of gravitational waves. Coalescing binary systems of neutron stars and/or black holes are promising sources of gravitational waves detected by large-scale laser interferometric gravitational observatories, such as the US LIGO and Italian-French VIRGO projects (see, e.g., [2] and references therein).

If we can detect the signal of gravitational waves emitted from such a system and compare it with theoretical templates, we may be able to determine a variety of astrophysical parameters of the sources such as their direction, distance, masses, spin, and so on[3][4]. In order to extract exact information about a source from an observed signal we need exact theoretical templates of the gravitational waveforms. In order to make such templates, it is very important to know the exact motion of the sources because the mass distribution is the most important factor in emission of gravitational waves. In post-Newtonian approximation, it is known that spin brings about some serious effects on the orbital motion of the sources. The motion is not necessarily confined to an orbital plane, due to the spin-orbit and spin-spin interaction force. The non-parallel spin motion will induce the precession of the orbital plane, resulting in modulation of a gravitational waveform[5][6]. In [5] it is shown that the orbital plane may behave in very strange way due to some spin effects. We do not verify whether or not any chaotic behavior occurs in their system. However from the analogy of the studies about spin effects in Newtonian dynamics, we may expect that a spin effect can make a motion chaotic. The gravitational waveform from a system with chaotic motion will be different from waveform from a regular system. If such a difference is enough large to be detected on the Earth, it will make it difficult to prepare complete theoretical templates.

Thus we believe that the study of the spin effect on the orbital evolution and the gravitational waves is very important from an observational viewpoint as well as the academic one. To clarify the pure effect of a spin, especially spin-orbit interaction, on the orbital evolution in general relativity we study the motion of a spinning test particle around a Schwarzschild black hole.

Throughout this paper we use units $c = G = 1$. We define the signature of the metric as $(-,+,+,+)$.

2 Chaotic Behavior of a spinning test particle

In this section we discuss the chaotic behavior of a spinning test particle in Schwarzschild spacetime.

2.1 Basic Equations

First of all, we briefly introduce the basic equations for a spinning test particle in relativistic spacetime.

The equations of motion of a spinning test particle are obtained by the "pole-dipole approximation" method[7].

$$\frac{dx^\mu}{d\tau} = v^\mu, \tag{1}$$

$$\frac{Dp^\mu}{D\tau} = -\frac{1}{2}R^\mu_{\nu\rho\sigma}v^\nu S^{\rho\sigma}, \tag{2}$$

$$\frac{DS^{\mu\nu}}{D\tau} = p^\mu v^\nu - p^\nu v^\mu, \tag{3}$$

where p^μ is the momentum of the particle, $S^{\mu\nu}$ is its spin tensor and τ is an affine parameter of the orbit, which is taken as proper time of the particle in this study. We also need a supplementary condition that gives a relation between v^μ and p^μ. We adopt the following condition[8],

$$p_\mu S^{\mu\nu} = 0, \tag{4}$$

which determines the center of mass of the particle consistently. Using (4) we can write down the relation between v^μ and p^μ explicitly, that is,

$$v^\mu = n\left[u^\mu + \frac{1}{2m^2\Delta}R_{\nu\lambda\rho\sigma}S^{\mu\nu}u^\lambda S^{\rho\sigma}\right], \tag{5}$$

$$\Delta = 1 + \frac{1}{4m^2}R_{\alpha\beta\gamma\delta}S^{\alpha\beta}S^{\gamma\delta}, \tag{6}$$

where n is a normalization constant determined by $v_\mu v^\mu = -1$, u^μ is a unit vector parallel to p^μ, and m, which is regarded as mass of the test particle, is defined by

$$m^2 = -p_\mu p^\mu. \tag{7}$$

This system has several conserved quantities. Regardless of the symmetry of a background spacetime, it is easy to show that m and $S^2 \equiv S^{\mu\nu}S_{\mu\nu}/2$ are constants of motion. If the spacetime allows an isometry, one can show that

$$C \equiv \xi^\mu p_\mu - \frac{1}{2}\xi_{\mu;\nu}S^{\mu\nu} \tag{8}$$

is also a conserved quantity, where ξ^μ is the Killing vector associated with the isometry.

2.2 Chaos in Schwarzschild Spacetime.

We shall analyze the motion of a spinning particle in Schwarzschild spacetime,

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1}dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2, \tag{9}$$

to study the effect of spin-orbit interaction. Because this spacetime is static and spherically symmetric, two Killing vectors $\xi^\mu_{(t)}$ and $\xi^\mu_{(\phi)}$ exist. Using these Killing vectors we define two conserved quantities (8) as

$$-E \equiv C_{(t)} = -p_t - \frac{M}{r^2}S^{tr}, \tag{10}$$

$$J_z \equiv C_{(\phi)} = p_\phi - r\sin^2\theta(S^{\phi r} - r\cot\theta S^{\theta\phi}). \tag{11}$$

E and J_z are interpreted as the energy and the z component of the total angular momentum of the particle, respectively. Because the spacetime is spherically symmetric, the x and y components of the total angular momentum of the particle should also be conserved. Without loss of generality, we can choose the z-direction to find the total angular momentum $(J_x, J_y, J_z) = (0, 0, J)$, where J is the total angular momentum. To study the motion of a test particle, if we can define the effective potential, the analysis becomes much easier. However, because the spin interaction is not described by a potential force, we may not be able to define the effective potential. Instead here we introduce a contour map in 2-dimensional space, on which the r-θ components of the momentum vanish. Imposing the condition $p^r = p^\theta = 0$ to the constraint equations(10), (11), we find the boundary of a region where the particle with energy E can move. This contour curve is defined by

$$E = V(r, \theta; J, S), \tag{12}$$

$$V(r, \theta; J, S) = mf^{\frac{1}{2}}(1+\Lambda^2)^{\frac{1}{2}} + \frac{M}{r}\Lambda\left(S^2 - \frac{J^2\cos^2\theta}{1+\Lambda^2}\right)^{\frac{1}{2}}, \tag{13}$$

where $f = 1 - 2M/r$ and

$$\Lambda = \frac{-mJr\sin\theta}{m^2r^2 - S^2f} + \left[\frac{m^2J^2r^2\sin^2\theta}{(m^2r^2 - S^2f)^2} - \frac{\left(J^2 - \frac{2MJ^2}{r}\cos^2\theta - S^2f\right)}{m^2r^2 - S^2f}\right]^{\frac{1}{2}}. \tag{14}$$

The particle with energy E can move in the area inside of the contour curve(12), i.e.,

$$E > V(r, \theta; J, S), \tag{15}$$

We shall call the function $V(r, \theta; J, S)$ the "effective potential" of a spinning test particle in Schwarzschild spacetime, although its gradient does not give a force. We find the "effective potential" V is classified into four types depending on J and S. The result is shown in Fig.1. Note that the contours of "effective potential" are depicted in the ρ-z plane, where $\rho = r\sin\theta$ and $z = r\cos\theta$.

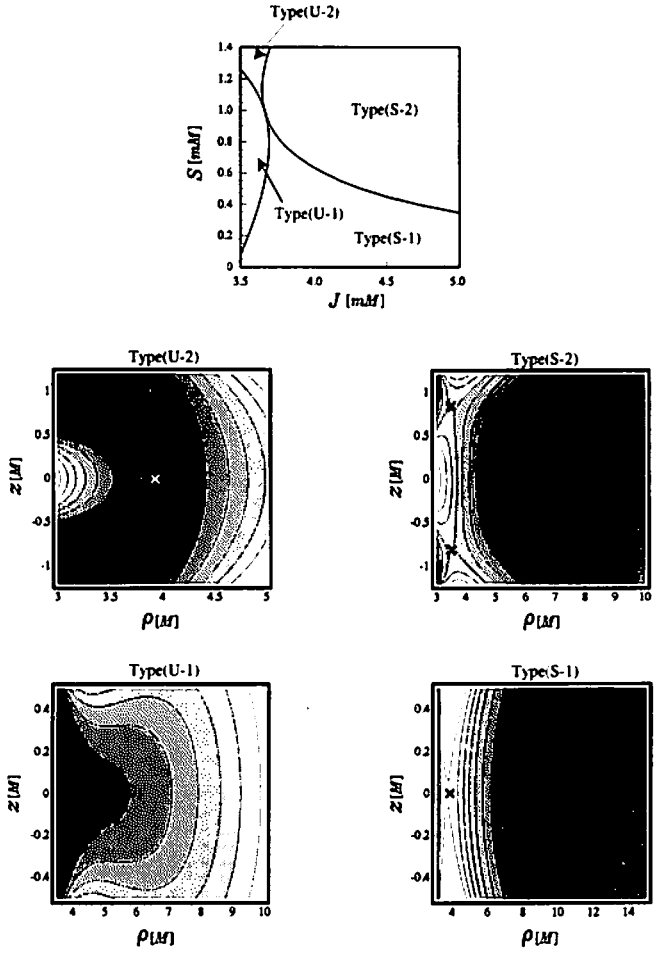


Fig. 1: Typical shapes of effective potential. The type(U-1) and type(U-2) have no bounded region, while the type(S-1) and type(S-2) have them. (U-2) has one saddle point on the equatorial plane. (S-2) and (S-1) have two saddle points off the equatorial plane and on the equatorial plane, respectively.

Because type(U-1) and type(U-2) potentials have no bounded region and any particle will fall into a black hole or escape to infinity, we are not interested in those potentials here. We analyze the type (S-1) and (S-2) potentials which have a bounded region. The type(S-1) potential has one saddle point on the equatorial plane, just as in the case of a spinless particle, which is the case that S/J is small. As S/J gets large and becomes larger than some critical value, the type(S-1) potential shifts to the type(S-2) potential, in which two saddle points appear off the equatorial plane.

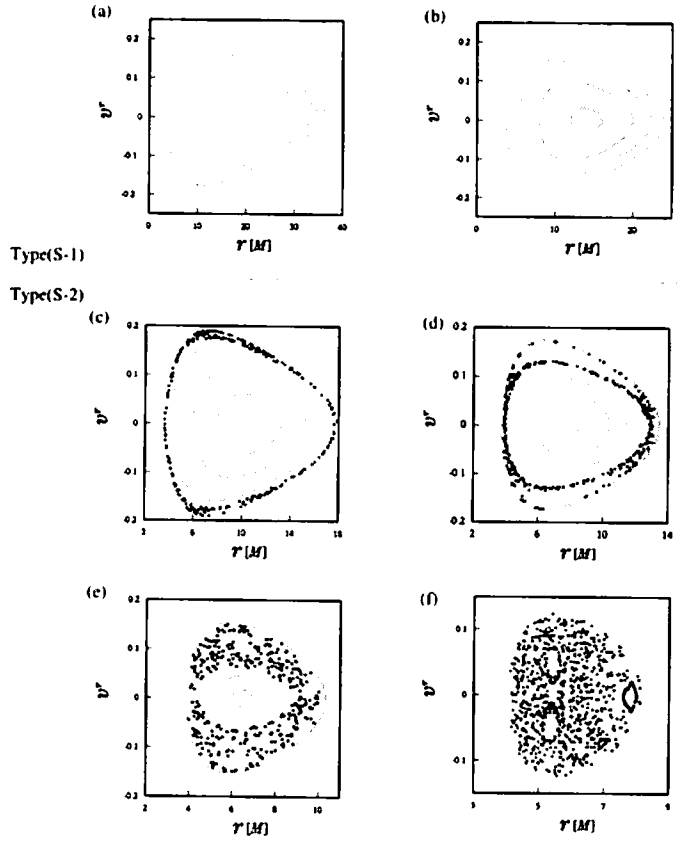


Fig. 2: Poincaré maps of the particle orbits for $J = 4mM$. (a) $S = 0.4mM$, $E = 0.97698396m$, (b) $S = 0.6mM$, $E = 0.96730999m$, (c) $S = 0.8mM$, $E = 0.95815568m$, (d) $S = 1mM$, $E = 0.94738162m$, (e) $S = 1.2mM$, $E = 0.93545565m$ and (f) $S = 1.4mM$, $E = 0.9229241m$.

In Fig.2 the Poincaré maps of the orbits in type(S-1) or (S-2) potentials are shown. These Poincaré maps are constructed by choosing the equatorial plane as a Poincaré section. The gray dots in these figures are on a torus in the phase space, which means such orbits are not chaotic. Black dots, which are not on a torus, denote a chaotic motion. Figs.2(a) and (b) are the Poincaré maps for the particle which moves in the type(S-1) potential. We find that all orbits are not chaotic. On the other hand, Figs.2(c)-(f) depict the Poincaré maps for the particle which moves in the type(S-2) potential. We find that some orbits become chaotic. From a numerical study, we obtain some conditions for the appearance of chaos. First, a particle should move in the type(S-2) potential, which means $S > 0.6mM$ for $J = 4mM$. Secondly, the particle energy must be high enough for it to be able to approach the saddle points. Thirdly, the initial condition should be chosen appropriately in order that the particle will approach the saddle points.

3 Effect on Observation of Gravitational Waves

3.1 Does Chaos Appear?

We find that a spin effect induces chaos. Then we discuss whether such chaotic behavior of a test particle affects observation of gravitational waves from a coalescing binary. When a particle moves in relativistic spacetime, it produces gravitational waves. Because the motion of the particle determines the gravitational wave from, it is expected that some influence will appear in the gravitational waves emitted from the system when the particle motion is chaotic.

As the system of a black hole and a spinning test particle emits the gravitational waves it loses energy and total angular momentum. If chaos occurs during the evolution of the system by energy and angular momentum losses, the next two conditions must be satisfied.

- The system passes through a region in a parameter space where it may become chaotic.
- The system stays in that region for a long enough time for its chaotic behavior to become conspicuous.

To see whether the those conditions are satisfied, the energy and angular momentum emission rates must be given. Here we use the quadrupole formula to estimate them. We assume that S is constant during the evolution and the system changes adiabatically. First we give E_0 and J_0 and integrate the equations of motion with appropriate initial condition from $r = r_0$. Then we calculate $\dot{E}$ and $\dot{J}$ by using the quadrupole formula. We set the state of the system after some time interval δt, which is sufficiently longer than the orbital period T, as

$$E_1 = E_0 + \dot{E}\delta t, \tag{16}$$

$$J_1 = J_0 + \dot{J}\delta t. \tag{17}$$

Repeating this estimation of the change of energy and angular momentum, we find an evolution sequence in the parameter space of E and J.

The result is shown in Fig.3. We set $S = 1mM$. The solid lines represent the energies of the saddle point($E_{\max}$) and of the minimum point($E_{\min}$) of the "effective potential". The bound orbit exists only for $E_{\min} < E < E_{\max}$. In Fig.3, we see that all evolutions end up with crossing the $E_{\max}$ line. It means that the particle has enough energy to go over the saddle point just before the end of adiabatic evolution. We then conclude that the first condition is satisfied.

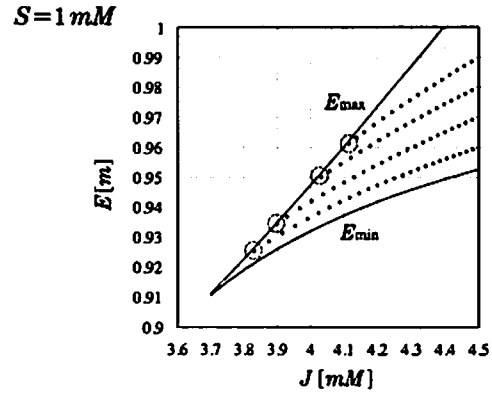


Fig. 3: The evolution of the system in the J-E plane. The orbits just before the end of the adiabatic evolution (encircled points) can become strongly chaotic .

To check the second condition, we have to calculate the Lyapunov exponent λ. The inverse of the Lyapunov exponent represents the time scale for the occurance of chaos. Hence we expect that chaotic behavior may appear during the evolution of the system if

$$\frac{1}{\lambda} < \Delta t. \tag{18}$$

Δt is the time interval when the system stays in the parameter space where the system can be chaotic. We calculate the Lyapunov exponent λ by the method developed by Sano and Sawada[9].

We find that $1/\lambda$ just before the end of adiabatic evolution is a few times longer than the average orbital period T, while the time scale of loss of energy or angular momentum is about 10~20 times longer than the average orbital period. Therefore we conclude there is enough time for the chaotic motion to appear just before the end of the adiabatic evolution of the system.

3.2 How Does Chaos Appear?

What can we say about the gravitational wave form from the system we consider? Due to a spin effect, the orbital plane is not fixed but precesses, even for a regular motion of the particle. This causes the modulation of the gravitational waves, especially in their amplitude[6]. How does chaotic behavior of the particle affect the wave forms?

Fig.4 shows the behavior of the orbital plane during the evolution of the system. Θ is the angle between the z axis and the direction of the orbital plane. In the early and middle stages of the evolution ((i) and (ii)) the particle motion is not chaotic. We find that the precession of the orbital plane is periodic. But just before the end of evolution, this periodicity is broken and the variation of Θ becomes chaotic. Such a clear difference in the behavior of the orbital plane between the regular and chaotic motions will change the gravitational wave form.

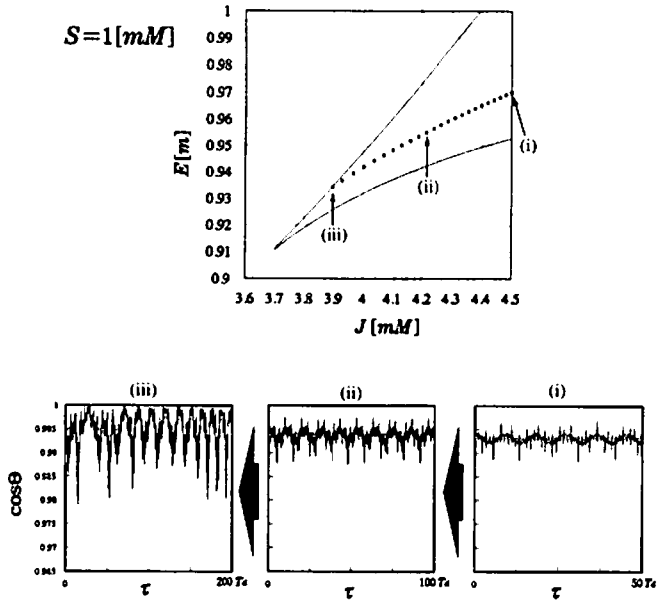


Fig. 4: The behavior of the orbital plane. In the early and middle stages of the evolution ((i) and (ii)), the precession is periodic. In the last stage of the evolution, however, the periodicity is broken and the precession becomes chaotic.

4 Summary and Discussion

We find that a spin-orbit interaction induces chaos. In realistic systems, we may find more complicated interactions, such as spin-spin or quadrupole interactions. We are interested in whether such additional terms enhance the chaotic behavior found here or not. To study the effect of spin-spin interaction, we have to analyze the motion of a spinning test particle in Kerr spacetime.

Secondly, in order to investigate the effect of chaos on the gravitational wave form precisely, it is necessary to estimate $\dot{E}$ and $\dot{J}$ more accurately with a relativistic method[10][11].

References

[1] S. Habib et al, Phys. Rev. **D50** 6068 (1994).

[2] K. S. Thorne, preprint gr-qc/9506084 and gr-qc/9506086.

[3] T. A. Apostlatos, Phy. Rev. **D52** 605 (1995).

[4] E. Poisson and C. M. Will, Phys. Rev. **D52** 848 (1995).

[5] T. A. Apostlatos, C. Cutler, G. J. Sussman and K. S. Thorne, Phys. Rev. **D49** 6274 (1994).

[6] L.E. Kidder, Phys. Rev. **D52** 821 (1995).

[7] A. Papapetrou, Proc. Roy. Soc. **A209** 248 (1951).

[8] W. G. Dixon, in "*Isolated Gravitating Systems in General Relativity*", edited by J. Ehlers, (North-Holland, Amsterdam, 1979), pp 156-219.

[9] M. Sano and Y. Sawada, Phys. Rev. Lett. **55** 1082 (1985).

[10] Y. Mino, M. Shibata and T. Tanaka, Phys. Rev. **D** (1996), to be published.

[11] T. Tanaka, Y. Mino, M. Sasaki and M. Shibata, preprint gr-qc/9602038.

Classical and Quantum Collapses of a Massless Scalar Field

Akira Tomimatsu[1]

Department of Physics, Nagoya University, Chikusa-ku, Nagoya 464-01, Japan

ABSTRACT

We develop classical and quantum descriptions of spherically symmetric gravitational collapse of a massless scalar field. Our purpose is to present some interesting examples of black hole dynamics. First, the four-metric and scalar field variables are restricted to a continuously self-similar functional form. The canonical quantization procedure is applied to the minisuperspace model, and the wave function of black hole evaporation is derived. We propose a clear mechanism of the quantum evaporation in terms of the time evolution of the wave function, and the evaporation time is semi-classically estimated. Next, we study the critical behavior near the threshold of black hole formation in the classical level. The mass-scaling relation with the critical exponent γ and the echoing with the universal period Δ have been found by numerical calculations of the near-critical collapse. Here we treat the periodic echoing (which represents a discretely self-similar field) as an oscillatory perturbation for the continuously self-similar field. This perturbation approach allows us to clarify a dynamical feature of the echoing near the singular center, and we can establish an analytical relation between γ and Δ.

[1]E-mail address: c42615a@nucc.cc.nagoya-u.ac.jp

I. INTRODUCTION

Gravitational collapse has attracted much attention in relation to various interesting problems in general relativity. For examples, the cosmic censorship hypothesis which requires the absence of naked singularities has been critically checked, by constructing realistic models of gravitational collapse. The emission of gravitational waves from collapsing stars also is an important astrophysical problem, which has inspired remarkable development of numerical relativity. Here we would like to emphasize another aspect of gravitational collapse, describing a dynamical stage of black hole formation. Various properties of stationary black holes have been extensively studied, one of which is a close analogy to the ordinary thermodynamic laws. However, a detailed understanding of evolutionary behaviors of dynamical black holes interacting with matter fields is still lacking. In particular, the validity and true implications of the black hole thermodynamics should be confirmed in the dynamical level. We may be able to establish important new features of black hole physics in terms of the dynamical models.

Spherically symmetric collapse of a massless scalar field (which is minimally coupled to gravity) is useful to consider dynamical black holes. Different end states can occur for one-parameter families of initial data of ingoing wave packets; while weak scalar fields bounce and disperse to infinity, strong fields form a black hole [1]. There exists a critical parameter value $p = p^*$ which separates the supercritical (black hole) solutions from the subcritical (dispersed wave) solutions. Choptuik [2] has numerically discovered a critical behavior near the threshold of black hole formation, which shows the universal mass-scaling relation

$$M_{BH} \propto (p - p^*)^\gamma , \quad \gamma \simeq 0.37 \tag{1}$$

and the echoing of the gravitational and scalar fields

$$Z(r,t) = Z(re^\Delta, te^\Delta) , \tag{2}$$

where Z stands for any field strength and $\Delta \simeq 3.44$ is the period in a logarithmic scale. (The relation (2) for a fixed Δ is called a discrete self-similarity. If it holds for arbitrary Δ, the field is continuously self-similar.) The change from the black hole (ordered) states to the dispersed wave (disordered) states may be regarded as a phase transition known in statistical physics. The origin of such critical phenomena appeared in near-critical black holes should be fully understood as a typical example of black hole dynamics.

The other example we consider here is a black hole evaporation as quantum dynamics. It is well-known that the Schwarzschild black hole is quantum-mechanically unstable owing

to the Hawking radiation. This thermal process leads to the information loss paradox in addition to the explosive evaporation

$$\dot{M}_{BH} \propto -1/M_{BH}^2 \tag{3}$$

in the small mass limit. However, this will be partly due to the classical Schwarzshild background geometry. To contain a consistent backreaction effect which makes the process fully dynamical, we must develop a quantization procedure without fixing the background geometry. If the above-mentioned scalar field collapse is quantized, we will obtain a good dynamical model to study the quantum evaporation of black holes..

The purpose of this talk is to discuss the classical and quantum dynamics of near-critical black holes and evaporating ones, by using a simple model of the collapsing scalar field with a critical parameter. The basic assumption of the model is a continuous self-similarity of the field variables [3,4,5]. In Sec.II, we construct the canonical formulation for the dynamical variables. Then, in Sec.III, we apply the canonical quantization procedure to this minisuperspace model of continuous self-similarity of metrics and scalar field configurations. Quantum dynamics of the spherical collapse with the supercritical initial data are described by a wave function which is defined on the minisuperspace and satisfies the Schrödinger equation. It is found that the supercritical initial state can evolve finally into an outgoing wave state without leaving any black hole state. We can propose a clear mechanism of quantum evaporation of black holes and estimate the evaporation time. Though this model is a good example of black hole evaporation, the proposed mechanism should be verified without the restriction of continuous self-similarity. Here we do not pursue the quantum dynamics in more extended models of scalar field collapse. The present step should be rather to develop the analysis of near-critical black holes in the classical level. In fact the continuously self-similar model fails to recover the critical behavior found numerically by Choptuik, although it has the critical parameter. In Sec.IV, we introduce the echoing with the period Δ as an oscillatory perturbation for the continuously self-similar configuration in the near-critical collapse. The important result is an analytical relation between the critical exponent γ of the mass-scaling law and the period Δ of the echoing, which is due to the inetraction between the oscillatory mode and the continuously self-similar field near the central singularity. We conclude that the essential features of near-critical dynamics can be well understood in terms of the perturbation approach.

II. CONTINUOUSLY SELF-SIMILAR MODEL

We consider a massless scalar field ϕ which is minimally coupled to the gravitational field. (In the following we use units such that $G = c = \hbar = 1$.) The Einstein-scalar action for the metric g_{ab} plus ϕ is given by

$$I = \frac{1}{16\pi}\int d^4x\sqrt{-g}(R - 2\partial^a\phi\partial_a\phi) . \tag{4}$$

In spherical symmetry, the line-element may be written in double-null form as

$$ds^2 = -2hdudv + r^2d\Omega^2 , \tag{5}$$

where $d\Omega^2$ refers to the unit sphere and h and r depend on the retarded and advanced null coordinates u and v. A time coordinate defined by $\xi = -v/u$ is used to assume the continuous self-similarity of the form

$$h = h(\xi) , \ r = -uf(\xi) , \ \phi = \phi(\xi) , \tag{6}$$

where the variable f is introduced as a self-similar field instead of the radius r of S_2. The allowed region of this metric form is assumed to be $u < 0$ and $v > 0$, where the hypersurface of a constant ξ becomes spacelike, and the other region $\xi < 0$ is replaced by the empty Minkowski spacetime [4,5]. This means that ingoing scalar waves from past null infinity start at $\xi = 0$, and the subsequent self-similar evolution lasts in a restricted region $0 < \xi < \infty$. The two boudaries $\xi = 0$ and $\xi = \infty$ are null surfaces. The initial boundary $\xi = 0$ covers both the path of the first ray $v = 0$ of ingoing scalar waves and the past null infinity $-u = \infty$. The boundary at the final time $\xi = \infty$ also covers both the path of the last ray $-u = 0$ of outgoing scalar waves and the future null infinity $v = \infty$.

The classical self-similar family of solutions is given by

$$h = 1/2 , \tag{7}$$

$$f^2 = \frac{1}{4}[(1-p^2)\xi^2 + 2\xi + 1] , \tag{8}$$

$$\phi = \frac{1}{2}\ln\frac{(1+p)\xi + 1}{(1-p)\xi + 1} , \tag{9}$$

where p is a critical parameter. To see the spacetime structure, it is convenient to calculate the quasi-local mass

$$M \equiv \frac{r}{2}(1 - \partial^a r\partial_a r) = -\frac{p^2}{8r}uv , \tag{10}$$

which is positive only in the region $\xi > 0$. We note that the apparent horizon $r = 2M$ exists at $\xi = \xi_A \equiv 1/(p^2 - 1)$ only when $p > 1$, and the supercritical collapse stops at the spacelike singularity $r = 0$ corresponding to $\xi = \xi_s \equiv 1/(p-1)$. In the subcritical collapse ($p < 1$), the scalar field completely disperses to infinity in the limit $\xi \to \infty$ without such a black hole formation.

We can construct the canonical formulation for the self-similar dynamics [6], if the classical Einstein-scalar equations are divided into the two equations of motion

$$\ddot{f} = -f\dot{\phi}^2 , \tag{11}$$

$$\frac{d}{d\xi}(f^2\dot{\phi}) = 0 , \tag{12}$$

and the one constraint

$$h = 2f\dot{f} - 2\xi(\dot{f}^2 - f^2\dot{\phi}^2) , \tag{13}$$

where dots mean derivatives with respect to ξ. It is easy to check that the conservation $\dot{h} = 0$ is automatically satisfied by virtue of the two equations of motion, which give the dynamical evolution of the independent variables f and ϕ. Note that for the self-similar fields the original action reduces to the form

$$I = \frac{1}{2}\int (f^2\dot{\phi}^2 - \dot{f}^2)d\xi , \tag{14}$$

which generates the canonical momenta conjugate to the variables f and ϕ

$$\Pi_f = -\dot{f} , \ \Pi_\phi = f^2\dot{\phi} , \tag{15}$$

and the Hamiltonian with the kinetic terms only

$$H = \frac{1}{2}(f^{-2}\Pi_\phi^2 - \Pi_f^2) . \tag{16}$$

By using the canonical variables, the constraint is rewritten into the form

$$h = 2(2\xi H - f\Pi_f) . \tag{17}$$

The calculation of the Poisson bracket of h and H can confirm that the metric component h is a constant of motion for the Hamiltonian. The physical importance of the Hamiltonian is clear, because it is related to the critical parameter p as follows,

$$H = \frac{1}{8}(p^2 - 1) . \tag{18}$$

The Hamiltonian H is a useful measure for separating the supercritical and subcritical solutions.

III. BLACK HOLE EVAPORATION

Now the canonical quantization of the scalar field collapse is straightforward [6], in which the canonical momenta are replaced by the operators

$$\hat{\Pi}_f = \frac{1}{i}\frac{\partial}{\partial f}, \quad \hat{\Pi}_\phi = \frac{1}{i}\frac{\partial}{\partial \phi}, \tag{19}$$

and the Hamiltonian operator has the form

$$\hat{H} = \frac{1}{2}(f^{-2}\hat{\Pi}_\phi^2 - f^{-1}\hat{\Pi}_f f \hat{\Pi}_f). \tag{20}$$

The metric operator $\hat{h}$ is also defined by

$$\hat{h} = 2(2\xi\hat{H} - \hat{\Pi}_f f), \tag{21}$$

which satisfies the equation

$$\frac{d\hat{h}}{d\xi} \equiv \frac{1}{i}[\hat{h}, \hat{H}] + \frac{\partial \hat{h}}{\partial \xi} = 0. \tag{22}$$

for the commutator bracket $[\hat{h}, \hat{H}]$. This means the consistency of the canonical quantization procedure on the minisuperspace covered by the coordinates f and ϕ.

The next task is to study the Hamiltonian eigenstates defined by

$$\hat{H}\Psi_E = E\Psi_E, \tag{23}$$

where E is the Hamiltonian eigenvalue. The eigenfunction can have the form

$$\Psi_E = \psi_{E,k}(f)\exp(ik\phi), \tag{24}$$

where k is the eigenvalue of the momentum operator $\hat{\Pi}_\phi$. Because the Hamiltonian has no potential term, we can easily obtain exact solutions of Ψ_E. By virtue of the form of a Klein-Gordon equation in two dimensions, the eigenvalue E is allowed in all the range $-\infty < E < \infty$. We consider separately the cases $E > 0$ and $E < 0$:

(1) For $E > 0$ we have a black hole state corresponding to the classical supercritical solution. This state which may be called a bound state shows a localized structure, because the eigenfunction $\psi_{E,k}$ given by the modified Bessel function decreases exponentially at large f as follows,

$$\psi_{E,k} = K_{ik}(f\sqrt{2E}) \sim \exp(-f\sqrt{2E}). \tag{25}$$

(2) For $E < 0$ there exist two linearly independent osillatory solutions. We denote them by the Hankel functions as follows,

$$\psi_{E,k} = H_{ik}^{(1)}(f\sqrt{-2E}) , \tag{26}$$

and

$$\psi_{E,k} = H_{ik}^{(2)}(f\sqrt{-2E}) . \tag{27}$$

If we consider the WKB approximate form $\exp(iS_E)$ of (27) in the limit $f \to \infty$, we can obtain the classical dynamics $\dot{f} = \sqrt{-2E}$ according to the equation

$$\Pi_f = \partial S_E/\partial f . \tag{28}$$

Recall that in the subcritical evolution initial ingoing waves bounce and are scattered away to infinity. The solution (27) corresponds to a quantum description of the subcritical state and the behavior in the asymptotic region $f \to \infty$ represents the evolution of f due to outgoing scalar waves at large ξ. Then the other solution (26) is interpreted to be the time reversal. These subcritical Hamiltonian eigenfunctions constitute a basic set of scattering states.

To discuss the time evolution of the wave function, it is important to remark a behavior of the metric eigenfunction defined by

$$\hat{h}\Psi_h = h\Psi_h , \tag{29}$$

where h is the metric eigenvalue. We can find the metric eigenfunction which shows no oscillatory behavior in the asymptotic region $f \to \infty$. This is a black hole state with the metric eigenvalue and is represented by a superposition of various black hole states corresponding to the Hamiltonian eigenstates. The remarkable point is that the metric eigenstate Ψ_h leads to a complete uncertainty of the Hamiltonian eigenvalue E, because $\hat{h}$ and $\hat{H}$ do not commute. Further we note that Ψ_h is dependent on the time coordinate ξ, and the time evolution occurs according to the Schrödinger equation

$$i\frac{\partial\Psi}{\partial\xi} = \hat{H}\Psi . \tag{30}$$

For the scattering states corresponding to the metric eigenstates the asymptotic WKB form in the limit $f \to \infty$ becomes oscillatory as follows,

$$\Psi_\eta \sim \exp(-if^2/2\xi) \times \exp(ik\phi) . \tag{31}$$

The factor $\exp(-if^2/2\xi)$ is identical with the asymptotic form of the time-dependent subcritical Hamiltonian eigenfunction

$$\exp(-i\xi E)H_{ik}^{(2)}(f\sqrt{-2E}) \sim \exp[-i(\xi E + f\sqrt{-2E})] , \tag{32}$$

if we eliminate the eigenvalue E by the help of the classical relation $\sqrt{-2E} = f/\xi$ which is valid in the limit $\xi \to \infty$. This implies that the wave function of such an asymptotic form describes a final stage of the quantum collapse toward an outgoing propagation of scalar waves, in which our knowledge of E is lost as a result of the time evolution according to the Schrödinger equation.

The final step to arrive at the black hole evaporation is to use the integral formula involving the Bessel function,

$$K_{ik}(f\sqrt{2E}) = \frac{1}{2}\int_{-\infty}^{0}(\frac{-E'}{E})^{ik/2}(E-E')^{-1}J_{ik}(f\sqrt{-2E'})dE' , \tag{33}$$

which is due to the breakdown of orthogonality of the Hamiltonian eigenfunctions. This expansion of the Hamiltonian eigenstate corresponding to $E > 0$ in terms of different Hamiltonian eigenstates corresponding to $E' < 0$ reveals an important property of quantum black holes. We can interpret the black hole state to be a special superposition of various scattering states. Then the asymptotic behavior of the eigenfunction K_{ik} at $f \to \infty$ will imply a well-localized wave packet rather than a bound state independent of scattering states. If so, the wave packet must spread with time by virtue of the quantum-mechanical uncertainty of Π_f. This is a quantum instability of black holes.

We can verify the decay process, by giving the initial condition $\Psi = K_{ik}(f\sqrt{2E})e^{ik\phi}$ of the wave function at $\xi = 0$. The solution of the time-dependent Schrödinger equation satisfying the initial condition has the form

$$\Psi = \frac{1}{2}\int_{-\infty}^{0}\exp(-i\xi E')(\frac{-E'}{E})^{ik/2}(E-E')^{-1}J_{ik}(f\sqrt{-2E'})dE'e^{ik\phi} . \tag{34}$$

To evaluate the integral with respect to E', we use the Wick rotaion $\tau \equiv -i\xi$ of the time coordinate. Then we obtain the wave function of black hole evaporation as follows,

$$\Psi = \psi(\tau, f)\exp(ik\phi) , \tag{35}$$

$$\psi = \frac{1}{2}(\frac{f^2}{2E})^{ik/2}\exp(E\tau)\int_{\tau}^{\infty}d\tilde{\tau}\tilde{\tau}^{-(1+ik)}\exp(-E\tilde{\tau} - \frac{f^2}{2\tilde{\tau}}) . \tag{36}$$

For a rough estimation of the integral it is useful to consider the minimum point $\tilde{\tau} = \tau_m \equiv f/\sqrt{2E}$ of the factor $E\tilde{\tau} + (f^2/2\tilde{\tau})$ in the exponential function:

(1) If $\tau < \tau_m$ and $f\sqrt{2E} \gg 1$, the main contribution to the integral with respect to $\tilde{\tau}$ is restricted to a narrow range near the minimum point, which corresponds to a classical trajectory characterized by the eigenvalue E. Then we obtain the approximate form of the black hole state written by

$$\psi \sim \exp(-i\xi E)\exp(-2E\tau_m) . \tag{37}$$

(2) The effective decay of the black hole state occurs only when the minimum point is outside the range of the integral ($\tau > \tau_m$). In particular, at the final stage $\tau \gg \tau_m$ of the time evolution the asymptotic behavior of ψ is approximately given by

$$\psi \sim \exp(-if^2/2\xi) . \tag{38}$$

Note that the wave function behaves like the WKB approximation of the metric eigenfunction (rather than the Hamiltonian eigenfunction) of the subcritical collapse. For this final state no black hole structure remains and the outgoing flux corresponding to the amplitude k becomes observable at future null infinity as a result of the black hole evaporation.

The above-mentioned behavior of the wave function motivates us to define semiclassically the decay (evaporation) time $\xi = \xi_d$ by the equation $\xi_d^2 = f^2(\xi_d)/2E$, which leads to

$$\xi_d = \frac{1}{2(p^2-1)}(1+\sqrt{2p^2-1}) . \tag{39}$$

We obtain the interesting inequality $\xi_s > \xi_d > \xi_A$, which means that the quantum effect becomes important after the apparent horizon is formed and before the singularity is formed. Further note that the ratio defined by

$$\frac{\xi_d - \xi_A}{\xi_A} = \frac{1}{2}(-1+\sqrt{2p^2-1}) \tag{40}$$

vanishes in the limit $p \to 1$ ($E \to 0$) and increases in proportion to p for large p. If the efficiency of black hole evaporation is measured in comparison with ξ_A and ξ_d, we can recover the usual result that the quantum process for larger black holes is less efficient.

We have constructed a quantum-mechanical model of the self-similar scalar field collapse and have derived the wave function of black hole evaporation as a solution of the time-dependent Schrödinger equation. Though the Hamiltonian eigenvalue E characterizes the geometrical structure of the initial black hole, the final state misses any information of the initial eigenvalue. This will be an information loss analogous to the prediction of the Hawking radiation. However, the evaporation process presented here shows no

explosive phenomenon at least in the time evolution of the wave function. The thermal picture of the Hawking radiation will break down in the quantum treatment of the dynamical black hole which is different from the Schwarzschild one. A more precise implication of the result for the Hawking radiation should be understood in quantum models beyond the framework of the continuously self-similar minisuperspace model.

VI. NEAR-CRITICAL BLACK HOLES

The purpose of this section is to study the critical behavior of dynamical black holes and to treat a deviation from the continuous self-similarity only in the classical level. We have mentioned that the continuously self-similar model is valid only in a restricted region of the spacetime. The boundary of the continuously self-similar spacetime, which is a path of the first ingoing null ray, must be matched with the flat spacetime to reqire the regularity of the center $r = 0$ before the beginning of spacelike singularity. This corresponds to an artificial truncation of scalar field distribution at the null boundary. However, Choptuik has numerically studied the near-critical collapses, by giving some smooth initial data of scalar field distributed over the regular center. Then the echoing with a special period Δ, which corresponds to a discrete self-similarity, has been found as a dynamical behavior different from the continuous self-similarity.

The regularity condition of the center during an early stage of the near-critical collapse is essential to determine the period Δ. For example, Gundlach [7] has investigated the scalar field collapse just at the critical parameter value $p = p^*$, by assuming the discrete self-similarity of the critical spacetime written by the metric

$$ds^2 = -\alpha^2 dt^2 + a^2 dr^2 + r^2 d\Omega^2 . \tag{41}$$

The beginning of singularity is adjusted to be $r = 0$, $t = 0$, and the self-similar and periodic variables are defined by

$$\xi \equiv \ln(r/-t) , \quad \tau \equiv \ln(-t) . \tag{42}$$

Then the Einstein-scalar equations allow the derivative $\partial Z/\partial \xi$ for any field $Z(\xi, \tau)$ to be expressed as functions of all the fields and their derivatives with respect to τ. This set of the field equations defines a Cauchy problem with the time coordinate ξ. It has been numerically shown that the boundary condition which requires the regularity both

at the center $\xi \to -\infty$ and some sonic line $\xi = \xi_0(\tau)$ can be satisfied only when the period is restricted to the eigenvalue $\Delta = 3.4439 \pm 0.0004$. Further Price and Pullin [8] have claimed that in each region near the two boundaries the critical solution can be well approximated by a flat spacetime scalar field solution, and a nonlinear matching across the transition region can yield the period Δ as the eigenvalue.

These approaches based on the eigenvalue problem have partly revealed the origin of discrete self-similarity. However, the structure of the critical spacetime, in particular, in relation to the singularity remains unclear. Here we would like to treat the singular center $r = 0$ in the spacetime written by the advanced null form of metric

$$ds^2 = -F^2(1 - \frac{2M}{r})dv^2 + 2Fdvdr + r^2 d\Omega^2 , \tag{43}$$

where M denotes the quasi-local mass. This is a useful form to obtain a simple picture of near-critical black holes with the apparent horizon $r = 2M$ very close to the singular center, and in the following the mass-scaling law will be shown to be due to the discretely self-similar behavior.

Our key idea is to assume the approximate validity of the continuously self-similar solution given in Sec.II with the critical parameter value $p = 1$, which for the metric (43) is rewritten into the form

$$\phi = \phi_0(\xi) \equiv \frac{1}{2} \ln \frac{\sqrt{1+\xi^2} - \xi}{\sqrt{1+\xi^2} + \xi} , \tag{44}$$

$$\frac{2M}{r} = m_0(\xi) \equiv \xi(\sqrt{1+\xi^2} - \xi) , \quad F = F_0(\xi) \equiv 1/\sqrt{\xi^2 + 1} , \tag{45}$$

where the self-similar variable is defined by $\xi = v/2r > 0$. The discrete self-similarity is introduced as an oscillatory perturbation for the continuously self-similar soltuion as follows,

$$\phi = \phi_0(\xi) + \phi_1 , \quad \phi_1 \equiv \Re(\varphi(\xi) e^{i\omega\tau}) , \tag{46}$$

where $\tau \equiv \ln v$ is the periodic variable. Of course, as previously mentioned, the discretely self-similar field must dominate at least in the range $v < 0$ for assuring the existence of a regular center. However, this does not necessarily mean that the perturbation treatment is meaningless also in the range $v > 0$. Recall that the absence of any apparent horizon around the singular center is a feature essential to the critical collapse. The background solution ϕ_0 is critical because of the behavior $2M/r \to 1/2$ at the singular center, which should be compared with the result that $2M > r$ for a black hole center and $M \propto r^3$ for a regular center. The perturbation treatment is partly verified, if the perturbed field ϕ_1 also can preserve such a critical feature.

The Klein-Gordon equation $\nabla^a\nabla_a\phi = 0$ for the perturbed field ϕ_1 is given by

$$(1-x)\frac{d}{dx}[(1-x^2)\frac{d\varphi}{dx} + i\omega(1+\frac{1}{x})\varphi] \;=\; (2x + \frac{i\omega}{x})\varphi - 2x\int_0^1 \varphi dx \;, \tag{47}$$

where $0 < x \equiv \xi/\sqrt{1+\xi^2} < 1$. The solution of the wave equation (47) can be written in the form

$$\varphi \;=\; \delta x \;+\; \sigma(x) \;, \tag{48}$$

where δ is a constant, and the function σ satisfying the constraint

$$\int_0^1 \sigma dx \;=\; (1+i\omega)\delta/2 \;. \tag{49}$$

is a solution of (47) in which the second term of the right-hand side is omitted. The important point of the solution is the asymptotic behavior in the limit $x \to 1$ (i.e., $r/v \ll 1$),

$$\sigma \;\sim (1-x)^n \;, \tag{50}$$

where for the exponent $n \equiv \alpha + i\beta$ we have

$$\alpha \;=\; \frac{1}{2\sqrt{2}}(4 - \omega^2 + \sqrt{\omega^4 - 4\omega^2 + 16})^{1/2} \;. \tag{51}$$

Note the inequality $1/2 < \alpha < 1$. This power law (50) of the oscillatory field ϕ_1 is crucial to determine the radius $r = r_A(v)$ of the apparent horizon, because the Einstein-scalar equations lead to the equation

$$\int_0^{r_A} F dr \;=\; \lim_{r\to 0} \frac{r(2r\partial_r\partial_v\phi + 2\partial_v\phi + F\partial_r\phi)}{r\partial_r^2\phi + \partial_r\phi} \;. \tag{52}$$

In this perturbation treatment (52) should be understood to be

$$r_A^2/v \;\simeq\; \lim_{r\to 0} \frac{r(2r\partial_r\partial_v\phi_1 + 2\partial_v\phi_1 + F_0\partial_r\phi_1 + F_1\partial_r\phi_0)}{r\partial_r^2\phi_0 + \partial_r\phi_0} \;, \tag{53}$$

where F_1 is the perturbed part of F. It is easy to see that the right-hand side of (53) vanishes by substitution of the solution σ with the power law. Then, we can conclude that no apparent horizon exists around the singular center.

We have found the essential feature (50) of the discretely self-similar field at the critical collapse. However, for the near-critical collapse, the apparent horizon with very small radius must form as a result of the violation of the power law in the region $r \le r_A$. Let us discuss such a change of the behavior in the near-critical collapse to derive the mass-scaling relation. Now the background field ϕ_0 should be modified into the near-critical

black hole solution of continuous self-similarity. This modification of the background field becomes important only in the region $(r/v)^2 \leq \epsilon \ll 1$, where $\epsilon \equiv p - p^*(p^* = 1)$. Therefore, it is sufficient to reconsider the equation for the perturbed part σ of discrete self-similarity under the approximation $y \equiv 1 - x \ll 1$, and we obtain

$$(y+\epsilon)\frac{d^2\sigma}{dy^2} + (1 - i\omega + \frac{\epsilon}{y})\frac{d\sigma}{dy} - \frac{2+i\omega}{2y}\sigma = 0 . \tag{54}$$

If the variable y is rescaled by $z \equiv -y/\epsilon$, (54) becomes the hypergeometric differential equation, and we obtain the solution which remains finite at $z = 0$ as follows,

$$\sigma = \kappa F(-n, n - i\omega, 1; z) , \tag{55}$$

where κ is a constant. The behavior of (55) in the region $-z \gg 1$ is given by

$$\sigma \simeq \kappa[\frac{\Gamma(2n - i\omega)}{\Gamma(1+n)\Gamma(n - i\omega)}(-z)^n + \frac{\Gamma(i\omega - 2n)}{\Gamma(1 - n + i\omega)\Gamma(-n)}(-z)^{-n+i\omega}] . \tag{56}$$

Note that the second term becomes negligible in the critical limit $\epsilon \to 0$. Then, the requirement that the solution recovers the asymptotic behavior (50) in the region $-z \gg 1$ leads to the choice $\kappa \sim \epsilon^n$. This implies that the deviation ϵ from the critical value generates the scalar field σ of discrete self-similarity with the amplitude of order of ϵ^α near the singular center, which is much larger than the amplitude ϵ of the continuously self-similar black hole field. Therefore, this part of the scalar field can work as a source of a more massive black hole. If so, the black hole solution of the perturbed field ϕ_1 which violates the discrete self-similarity should have the amplitude of the same order, and it is found to be

$$\phi_1 \sim \epsilon^\alpha / v^2 \tag{57}$$

in the limit $r \to 0$. By using this estimation of the scalar amplitude in (53), we obtain

$$r_A^2 \sim \epsilon^\alpha , \tag{58}$$

which leads to the relation between the critical exponent γ of the mass-scaling law and the frequency ω of the echoing as follows,

$$\gamma = \alpha/2 = \frac{1}{4\sqrt{2}}(4 - \omega^2 + \sqrt{\omega^4 - 4\omega^2 + 16})^{1/2} \tag{59}$$

In fact, for the value $\omega = 2\pi/\Delta = 1.824$ evaluated by Gundlach, we can derive the value $\gamma = 0.370$, which is in agreement with the result of numerical calculations.

In summary, we have understood that the near-critical black hole is formed as a result of the interaction between the continuously self-similar part and the oscillatory part with the period Δ. If the latter is absent, we have the mass-scaling relation $M_{BH} \propto (p-p^*)^{1/2}$ [4,5]. The latter contribution is important to increase the black hole mass to the value $M_{BH} \propto (p-p^*)^{\gamma}$. Though our analysis to treat the singular center is not useful to determine the period Δ, we have succeeded in finding the analytical $\gamma - \Delta$ relation and the new critical behavior $\sigma \sim (r^2/v^2)^n$ of the oscillatory part in the region $r/v \ll 1$ outside the apparent horizon, which should be confirmed by more detailed numerical calculations.

The quantization of the oscillatory part also is an interesting problem as a next step to study black hole evaporation. The perturbative treatment developed here in the classical level is a good approximation of the near-critical collapse. Then it will not be so difficult to clarify the quantum effect in terms of the canonical formulation presented in the previous sections. This will be done in future works.

ACKNOWLEDGMENTS

This work was partially supported by the Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan (No. 07640390).

REFERENCES

[1] D. Goldwirth and T. Piran, Phys. Rev. D **36**, 3575 (1987).

[2] M. W. Choptuik, Phys. Rev. Lett. **70**, 9 (1993).

[3] M. D. Roberts, Gen. Relativ. Grav. **21**, 907 (1989).

[4] P. R. Brady, Class. Quantum Grav. **11**, 1255 (1994).

[5] Y. Oshiro, K. Nakamura and A. Tomimatsu, Prog. Theor. Phys. **91**, 1265 (1994).

[6] A. Tomimatsu, Phys. Rev. D **52**, 4540 (1995).

[7] C. Gundlach, *The Choptuik spacetime as an eigenvalue problem*, gr-qc/9507054.

[8] R. H. Price and J. Pullin, *Analytic approximations to the spacetime of a critical gravitational collapse*, gr-qc/9601009.

Department of Physics
Kyoto University

March 1996

Critical Behavior in the Brans-Dicke Theory of Gravitation

Takeshi Chiba*

Department of Physics, Kyoto University,
Kyoto 606-01, Japan

The collapse of a massless scalar field in the Brans-Dicke theory of gravitation is studied numerically. By conformally transforming the Choptuik's solution into the Brans-Dicke frame, we find for $\omega > -3/2$ that at the critical solution shows discrete self-similarity, however, the critical exponent and the echoing parameter do depend on ω.

I. INTRODUCTION

Critical phenomena in black hole formation found by Choptuik [1] has renewed interest in the classical general relativistic black hole formation. Among the most interesting things in Choptuik's results are
(1)*Scaling*: Black hole mass exhibits a power law

$$M_{BH} \propto (p - p^*)^\beta, \tag{1.1}$$

where p is arbitrary parameter which characterize the strength of initial condition and p^* is the threshold value.
(2)*Echoing*: Configurations sufficiently close to critical show a discrete homotheticity (or scale invariance)

$$\phi(\rho - \Delta, \tau - \Delta) \simeq \phi(\rho, \tau), \tag{1.2}$$

where ρ and τ are logarithms of proper radius r and central proper time t. Echoing means that the features of critical solution are repeated on ever decreasing time length scales.
(3) *Universality*: The exponent $\beta \simeq 0.37$ and echoing parameter Δ are independent of any choice of initial data p.

It should be noted, however, that the universality mentioned there refers to the independence from the initial condition of the matter field considered. To investigate the dependence of the model of the matter, the collapse of gravitational wave [2] and radiation fluid collapse [3] were examined. For vacuum gravity, $\beta \simeq 0.37$ and $\Delta \simeq 0.6$, while for radiation fluid $\beta \simeq 0.36$ with Δ being arbitrary(i.e. continuous self-similarity). These results excite the expectation that the critical exponent β may be universal among matter fields having massless property, although quantum effects will destroy the phenomena against expectation [4].

In this paper, we will investigate another model dependence: *dependence on the theory of gravitation*. To see this we take the Brans-Dicke theory of gravitation [5] for its simplicity. The theory contains a parameter ω which controls the strength of gravity, and in the limit $\omega \to \infty$ the theory reduces to the Einstein theory. We consider the Brans-Dicke theory in vacuum. The scalar field examined here is thus the Brans-Dicke scalar field. By conformally transforming the Choptuik's solution into the Brans-Dicke frame, we find that a discrete self-similar solution appears, however, the critical exponent and the echoing parameter do depend on ω.

The collapse of massless scalar field in the Brans-Dicke theory is also interesting in the light of structural stability of Choptuik's solution, which is our next problem.

Our paper is organized as follows. In sec.2, basic equations and numerical procedures are given. We reproduce Choptuik solution. In sec.3, numerical results for Brans-Dicke theory are given. Sec.4 is devoted to summary.

---

*e-mail: chiba@tap.scphys.kyoto-u.ac.jp

II. COLLAPSE OF BRANS-DICKE SCALAR FIELD IN DOUBLE-NULL COORDINATES

We consider the collapse of the Brans-Dicke scalar field. The action is

$$S = \int d^4x\sqrt{-\tilde{g}}(\Phi\tilde{R} - \frac{\omega}{\Phi}(\tilde{\nabla}\Phi)^2). \tag{2.1}$$

By the conformal transformation $\tilde{g}_{ab} = g_{ab}/\Phi$, the action is reduced to that of a massless scalar field coupled to the Einstein gravity:

$$S = \int d^4x\sqrt{-g}(R - \frac{1}{2}(\nabla\phi)^2), \tag{2.2}$$

where $\phi = \sqrt{2\omega+3}\ln\Phi$. We treat the problem with the action Eq.(2.2) in terms of null initial value formulation. Since in the null initial value formulation the grid points are tied to ingoing light rays, overall size of the grid becomes smaller as system evolves and the resolution improves. Adaptive mesh refinement algorithm used by Choptuik is not always necessary.

We take the following line element

$$ds^2 = -a^2dudv + r^2d\Omega^2. \tag{2.3}$$

Double-null coordinates are not unique. There remain two degrees of gauge freedom which correspond to redefining u and v. We fix one of them by requiring $u = v$ on axis $r = 0$. The remaining gauge freedom will be fixed by the initial condition.

The equations of motion that derived from Eq.(2.2) are

$$rr_{uv} + r_ur_v + \frac{1}{4}a^2 = 0, \tag{2.4}$$

$$a^{-1}a_{uv} - a^{-2}a_ua_v - r^{-2}r_ur_v - \frac{1}{4}r^{-2}a^2 + \frac{1}{4}\phi_u\phi_v = 0, \tag{2.5}$$

$$r\phi_{uv} + r_u\phi_v + r_v\phi_u = 0. \tag{2.6}$$

And constraint equations are

$$r_{uu} - 2a^{-1}a_ur_u + \frac{1}{4}r\phi_u^2 = 0, \tag{2.7}$$

$$r_{vv} - 2a^{-1}a_vr_v + \frac{1}{4}r\phi_v^2 = 0, \tag{2.8}$$

where $r_u \equiv \partial r/\partial u$, for example. These equations are solved numerically. Since the usual iteration scheme for solving these equations does not work because of their nonlinearity, we adopt the first order scheme developed by Hamade and Stewart [6].

We introduce the following new variables

$$d = \frac{a_v}{a}, f = r_u, g = r_v, p = \phi_u, q = \phi_v. \tag{2.9}$$

The system of equations (2.3-2.7) can then be converted to a first order system:

$$f_v + r^{-1}(fg + \frac{1}{4}a^2) = 0, \tag{2.10}$$

$$d_u - r^{-2}(fg + \frac{1}{4}a^2) + \frac{1}{4}pq = 0, \tag{2.11}$$

$$p_v + r^{-1}(fq + gp) = 0, \tag{2.12}$$

$$q_u + r^{-1}(fq + gp) = 0, \tag{2.13}$$

$$g_v - 2dg + \frac{1}{4}rq^2 = 0, \tag{2.14}$$

and the equations defining variables

$$a_v - ad = 0, \tag{2.15}$$

$$r_v - g = 0, \tag{2.16}$$

$$\phi_v - q = 0. \tag{2.17}$$

We have to specify the boundary conditions on axis, $u = v$. Of course, $r = 0$ there and this implies $r_u + r_v = 0$, i.e., $f = -g$. Eq.(2.10) will be regular on axis if and only if $a = 2g$. Eq.(2.12) implies $p = q$ on axis. The boundary value for a and ϕ are obtained by requiring $a_r = \phi_r = 0$ there.

We need to give initial data d and q on the initial surface $u = 0$. The remaining gauge freedom mentioned in the second paragraph is the choice of d on $u = 0$. This is arbitrary, and we choose $d = 0$. As for the initial condition of q (or ϕ) on $u = 0$, we typically take

$$\phi(u = 0, v) = 1 + \phi_0 v^2 \exp(-(v - v_0)^2/v_1^2). \tag{2.18}$$

We then integrate outward from axis Eq.(2.15) to obtain a, Eq.(2.16) to obtain r, Eq.(2.14) to obtain g, Eq.(2.10) to obtain f and Eq.(2.12) to obtain p. Thus we set initial conditions.

The integration in the u-direction is done using an explicit difference algorithm. The integration in the v-direction is done using an implicit algorithm, which is necessary to ensure stability. However, the integration can be made explicit. Details of algorithm are given in [6].†

In Fig.1, we plot the scalar field at the center with marginally subcritical initial parameter as a function of logarithm of the proper time $\tau \equiv -\log(t^* - t)$, where t^* is the time when an infinitesimal black hole forms. We find that the scalar field oscillates with period $\Delta \simeq 3.43$ in τ-coordinate, in agreement with Choptuik.

Fig.2 shows the black hole mass for marginally supercritical evolution as a function of $\phi_0 - \phi_0^*$ with ϕ_0^* being the critical parameter. The least square fit shows the exponent $\beta \simeq 0.38$, in agreement with Choptuik.

These results confirm logical consistency and accuracy of our numerical code.

III. CRITICAL PHENOMENA IN THE BRANS-DICKE THEORY OF GRAVITATION

Next we study the behavior of scalar field for various values of ω, $125 \geq \omega > -3/2$ by conformally transforming the Choptuik's solution reproduced in the previous section. The results are shown in Figures and Table. Fig.3 shows the scalar filed at the center as a function of logarithm of the proper time $\tau \equiv -\log(t^* - t)$ for several values of ω. The scalar field oscillates greatly for smaller ω. Fig.4 shows the black hole mass for marginally supercritical evolution as a function of initial parameter for $2\omega + 3 = 2$. The least square fit gives the exponent $\beta \simeq 0.43$. Table summarizes the numerical results. As can be seen the exponent β depends strongly on ω, hence it is not universal quantity. This is the first numerical evidence of non-universality [7] for the collapse of single matter content. On the other hand the scaling parameter Δ depends weakly on ω.

Table Brans-Dicke parameter, critical exponents, and scaling parameters.

$2\omega+3$	β	Δ
0.50	0.50	3.41
1.0	0.46	3.43
2.0	0.43	3.45
5.0	0.41	3.46
10	0.38	3.46
50	0.38	3.47
200	0.38	3.47
1000	0.38	3.47

It is to be noted that critical phenomena in the Brans-Dicke theory does emerge in discrete self-similar manner. Thus stability analysis around a *continuous* self-similar solution gives no information for the stability of the actual dynamical solutions.

---

†However, slight modification is necessary to ensure that the system has second-order accuracy; Eq.(3.5) in their paper should be replaced with $z_n = \frac{1}{2}\left(\hat{z}_n + z_w + \frac{h}{2}(G(\hat{y}_n, \hat{z}_n) + G(y_w, z_w))\right)$. We thank T.Harada for pointing this out to us.

IV. SUMMARY

We have studied the collapse of the Brans-Dicke scalar field to examine the dependence of the critical phenomena on the theory of gravitation. By conformally transforming the Choptuik's solution into the Brans-Dicke frame, we find for $\omega > -3/2$ that at the critical solution shows discrete self-similarity. However, the critical exponent and the scaling parameter do depend on ω. Numerical dynamical solutions show discrete self-similarity. These are another examples of discrete self-similar critical solutions. Given these numerical examples, the critical phenomena are not *universal*. The apparent analogy to statistical system may be just analogy not fundamental.

There remain many unanswered questions in critical phenomena in black hole formation. For example, "Why does the critical solution exhibit discrete (or continuous) self-similarity? " or "Why is the critical solution independent of initial condition?" We have no idea at present, however, we hope to report on the resolution of these problems near future.

ACKNOWLEDGMENTS

We would like to thank Jiro Soda and Dieter Mason for enjoyable discussions and Misao Sasaki for useful comment on the conformal transformation. We are also grateful to H.Sato for continuous encouragement.

---

[1] M.W. Choptuik, *Phys. Rev. Lett.* **70** 9 (1992).
[2] A.M. Abrahams and C.R. Evans, *Phys. Rev. Lett.* **70** 2980 (1992).
[3] C.R. Evans and J.S. Coleman, *Phys. Rev. Lett.* **72** 1782 (1993).
[4] T. Chiba and M. Siino, *Disappearance of Critical Behavior in Semiclassical General Relativity*, submitted to CQG.
[5] C.B. Brans and R.H. Dicke, *Phys. Rev.* **124** 925 (1961); R.H.Dicke, *Phys. Rev.* **125** 2163 (1962).
[6] R.S. Hamade and J.M. Stewart, gr-qc/9506044.
[7] D. Mason, *Phys. Lett.* **B366** 82 (1996).

FIGURE CAPTIONS

Fig.1. The scalar field on the center with a marginally subcritical initial parameter as a function of logarithm of the proper time $\tau = -\log(t^* - t)$, where t^* is the time when an infinitesimal black hole forms. In this time scale the scalar field oscillates with a periodicity $\Delta \simeq 3.43$.

Fig.2. The mass of a formed black hole for supercritical solutions as a function of $p - p^*$ with p^* being the critical initial parameter. The least square fit shows the exponent $\beta \simeq 0.38$.

Fig.3. The scalar field on the center for $2\omega + 3 = 0.5, 1, 2$ with a marginally subcritical initial parameter as a function of logarithm of the proper time $\tau = -\log(t^* - t)$.

Fig.4. The mass of a formed black hole for supercritical solutions for $2\omega + 3 = 2$ as a function of $p - p^*$. The least square fit shows the exponent $\beta \simeq 0.43$.

Fig.1

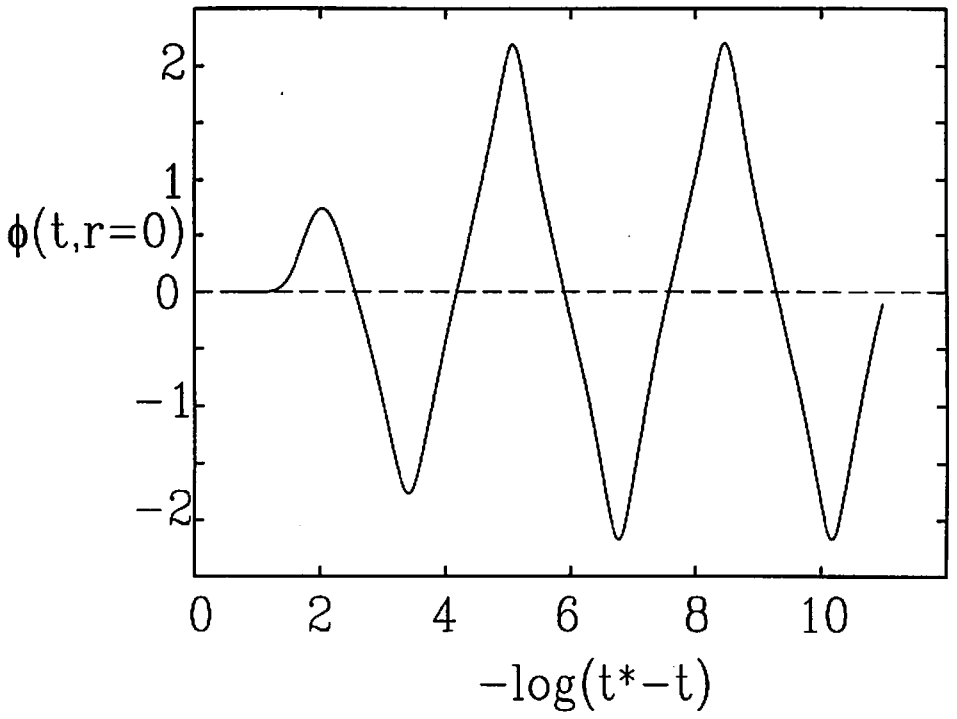


Fig.2

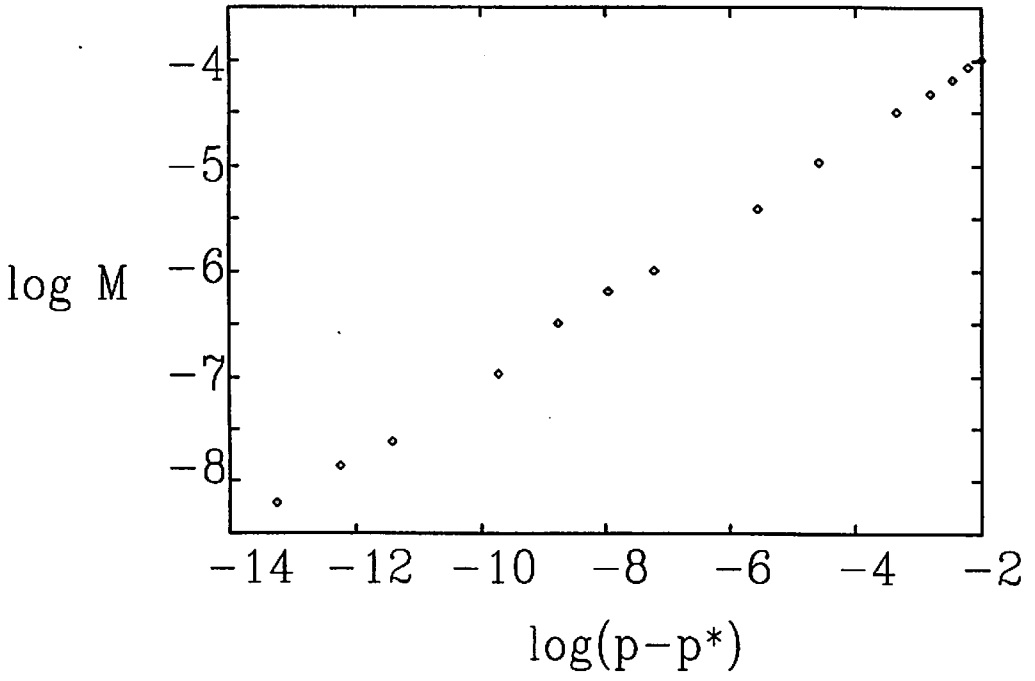


Fig.3

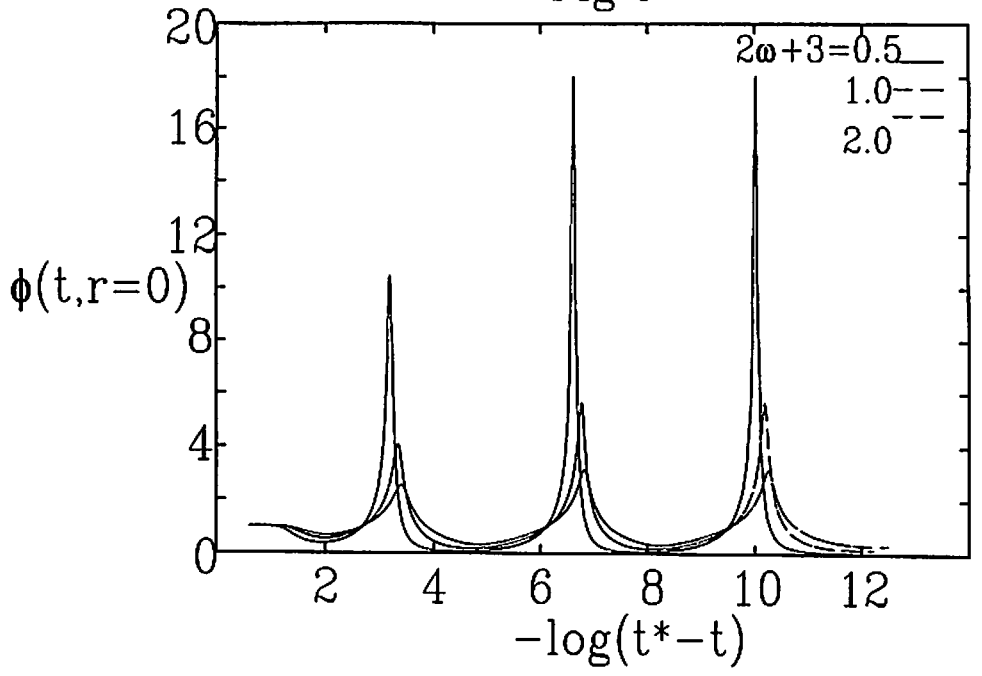
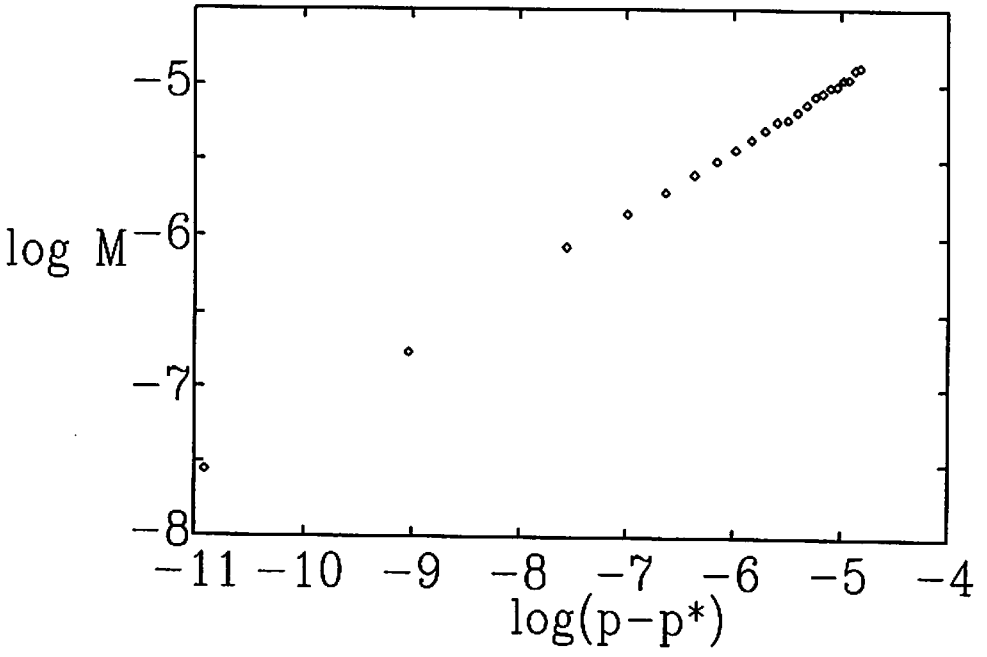

2ω+3=0.5
1.0
2.0
φ(t,r=0)
20
16
12
8
4
0
2
4
6
8
10
12
−log(t*−t)

Fig.4


log M
−5
−6
−7
−8
−11
−10
−9
−8
−7
−6
−5
−4
log(p−p*)

Dynamics of Gravitating Magnetic Monopoles*

NOBUYUKI SAKAI†

Department of Physics, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169

Abstract

According to previous work on magnetic monopoles, static regular solutions are nonexistent if the vacuum expectation value of the Higgs field η is larger than a critical value η_{cr}, which is of the order of the Planck mass. In order to understand the properties of monopoles for $\eta > \eta_{cr}$, we investigate their dynamics numerically. If η is large enough ($\gg \eta_{cr}$), a monopole expands exponentially and a wormhole structure appears around it, regardless of coupling constants and initial configuration. If η is around η_{cr}, there are three types of solutions, depending on coupling constants and initial configuration: a monopole either expands as stated above, collapses into a black hole, or comes to take a stable configuration.

1 Introduction

In recent years static and spherically symmetric solutions of the Einstein-Yang-Mills-Higgs system have been intensively studied in the literature [1,2,3]. One purpose of such investigation has been to understand the nature of black holes, especially in the context of the no-hair conjecture; it was shown that non-trivial black holes are stable and hence the monopole black hole could be one of the most plausible counterexamples. The other interest has been in the properties of particle-like solutions; it was shown that such regular monopoles exist only if the vacuum expectation value of the Higgs field η is less than a critical value η_{cr}, which is of the order of the Planck mass m_{Pl}. This result naturally gives rise to the next question: what is the fate of monopoles for $\eta > \eta_{cr}$?

Because the only static solution for $\eta > \eta_{cr}$ is the Reissner-Nordström black hole, we can expect that a monopole which is regular initially evolves into the Reissner-Nordström black hole. Even if this speculation is reasonable, we still do not know how the black hole formation occurs. One could imagine two alternatives: a monopole just shrinks, or its core continues to expand inside the black-hole horizon, just as a "child universe" [4].

Linde and Vilenkin independently pointed out the latter possibility in the context of the "topological inflation" model [5]. They claimed monopoles as well as other topological defects expand exponentially if $\eta > O(m_{\rm Pl})$. Their discussions for the Einstein-Higgs system were verified by our numerical simulation in [6]: we found that domain walls and global monopoles inflate if and only if $\eta \gtrsim 0.33 m_{\rm Pl}$. The next question on this monopole inflation is similarly what happens to magnetic monopoles in the Einstein-Yang-Mills-Higgs system. Because we cannot find an answer to the question only by analyzing static solutions, we investigate dynamic monopole solutions in this paper.

The plan of this paper is as follows. In Sec. II, we derive the basic equations and explain how we solve those dynamical equations numerically. In Sec. III, we offer analytic discussions and numerical results. Sec. IV is devoted to summary and discussions. In this paper we use the units $c = \hbar = 1$.

*This talk is based on N. Sakai, preprint, WU-AP/95/52, gr-qc/9512045, submitted to Phys. Rev. D.

†Electronic address: sakai@cfi.waseda.ac.jp

2 Basic Equations

The $SO(3)$ Einstein-Yang-Mills-Higgs system is described by

$$S = \int d^4x\sqrt{-g}\left[\frac{m_{\rm Pl}^2}{16\pi}\mathcal{R} - \frac{1}{4}(F^a_{\mu\nu})^2 - \frac{1}{2}(D_\mu\Phi^a)^2 - V(\Phi)\right], \tag{2.1}$$

$$V(\Phi) = \frac{1}{4}\lambda(\Phi^2 - \eta^2)^2, \quad \Phi \equiv \sqrt{\Phi^a\Phi^a}, \tag{2.2}$$

$$F^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + e\epsilon^{abc}A^b_\mu A^c_\nu, \quad D_\mu\Phi^a \equiv \nabla_\mu\Phi^a + e\epsilon^{abc}A^b_\mu\Phi^c, \tag{2.3}$$

where A^a_μ and $F^a_{\mu\nu}$ are the $SU(2)$ Yang-Mills field potential and its field strength, respectively. Φ^a is the real triplet Higgs field and $V(\Phi)$ is its potential. λ and e are the Higgs self coupling constant and the gauge coupling constant, respectively.

We assume a spherically symmetric spacetime and adopt the coordinate system,

$$ds^2 = -dt^2 + A^2(t,r)dr^2 + B^2(t,r)r^2(d\theta^2 + \sin^2\theta d\varphi^2). \tag{2.4}$$

For the matter fields, we adapt the 't Hooft-Polyakov ansatz in such a way that we can apply it to a time-dependent curved spacetime:

$$\Phi^a = \Phi(t,r)\hat{r}^a, \quad \hat{r}^a \equiv (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta). \tag{2.5}$$

$$A^a_\mu = \frac{\partial(\sqrt{g_{\theta\theta}}\hat{r}^b)}{\partial x^\mu}\epsilon^{abc}\hat{r}^c\frac{1-w(t,r)}{e\sqrt{g_{\theta\theta}}}. \tag{2.6}$$

With the metric (2.4) and the 't Hooft-Polyakov ansatz (2.5) and (2.6), we derive the basic equations:

$$\begin{aligned} -G^0_0 &\equiv K_2^2(2K - 3K_2^2) - \frac{2B''}{A^2B} - \frac{B'^2}{A^2B^2} + \frac{2A'B'}{A^3B} - \frac{6B'}{A^2Br} + \frac{2A'}{A^3r} - \frac{1}{A^2r^2} + \frac{1}{B^2r^2} \\ &= \frac{8\pi}{m_{\rm Pl}^2}\left[\frac{\dot{\Phi}^2}{2} + \frac{\Phi'^2}{2A^2} + \left(\frac{w\Phi}{Br}\right)^2 + V + \frac{1}{(eBr)^2}\left\{\dot{w}^2 + \frac{w'^2}{A^2} + \frac{1}{2}\left(\frac{w^2-1}{Br}\right)\right\}\right], \end{aligned} \tag{2.7}$$

$$\frac{1}{2}G_{01} \equiv K_2^{2\prime} + \left(\frac{B'}{B} + \frac{1}{r}\right)(3K_2^2 - K) = \frac{4\pi}{m_{\rm Pl}^2}\left[\dot{\Phi}\Phi' + \frac{2\dot{w}w'}{(eBr)^2}\right], \tag{2.8}$$

$$\begin{aligned} \frac{1}{2}(G^1_1 + G^2_2 + G^3_3 - G^0_0) &\equiv \dot{K} - (K^1_1)^2 - 2(K^2_2)^2 \\ &= \frac{8\pi}{m_{\rm Pl}^2}\left[\dot{\Phi}^2 - V + \frac{1}{(eBr)^2}\left\{\dot{w}^2 + \frac{w'^2}{A^2} + \frac{1}{2}\left(\frac{w^2-1}{Br}\right)\right\}\right], \end{aligned} \tag{2.9}$$

$$\ddot{\Phi} - K\dot{\Phi} - \frac{\Phi''}{A^2} - \left(-\frac{A'}{A} + \frac{2B'}{B} + \frac{2}{r}\right)\frac{\Phi'}{A^2} + \frac{2w^2\Phi}{B^2r^2} + \frac{dV}{d\Phi} = 0, \tag{2.10}$$

$$\ddot{w} - K^1_1\dot{w} - \frac{w''}{A^2} + \frac{A'w'}{A^3} - \frac{w(1-w^2)}{B^2r^2} + e^2\Phi^2 w = 0, \tag{2.11}$$

where the overdot and the prime denote the partial derivative with respect to t and r, respectively. We have introduced the extrinsic curvature tensor of a t = constant hypersurface, K_{ij}, whose components are given by

$$K^1_1 = -\frac{\dot{A}}{A}, \quad K^2_2\,(= K^3_3) = -\frac{\dot{B}}{B}, \tag{2.12}$$

and we have denoted its trace by $K \equiv K^i_i$.

As an initial configuration of the matter fields, we adopt the functional form of the static solution in a flat spacetime with $\lambda = 0$:

$$\Phi(t=0,r) = \Phi_{\rm flat}\Big(\frac{r}{c_\Phi}\Big) \equiv \eta\left[\frac{1}{\tanh(e\eta r/c_\Phi)} - \frac{1}{e\eta r/c_\Phi}\right],$$
$$w(t=0,r) = w_{\rm flat}\Big(\frac{r}{c_w}\Big) \equiv \frac{e\eta r/c_w}{\sinh(e\eta r/c_w)}, \tag{2.13}$$

where c_Φ and c_w are the initial size parameters. As to the time-derivative, we suppose $\dot\Phi(t=0,r) = \dot w(t=0,r) = 0$. In order to set up consistent initial data, we have to solve the constraint equations (2.7) and (2.8). At this point, there are four unknown variables, A, B, K and K_2^2, in the two constraint equations; two of the variables are arbitrarily chosen. One of the methods which is usually adopted is to assume $K = \text{const}$ and $A = B$. In this system, however, the condition of $K = \text{const} \neq 0$ is not appropriate because the far region is asymptotically flat. Further, in the range where there exists no static solution, we cannot fix $K = 0$ even momentarily. As an alternative, thereby, we suppose $A(t=0,r) = B(t=0,r) = 1$ and solve the constraint equations (2.7) and (2.8) to determine $K(t=0,r)$ and $K_2^2(t=0,r)$. This treatment is suitable for this system because we obtain

$$-\frac{K}{3} \approx -K_2^2 \approx \sqrt{\frac{8\pi}{3m_{\rm Pl}^2}\Big(\frac{\Phi'^2}{2} + \frac{\Phi^2}{r^2} + V\Big)}, \tag{2.14}$$

which approaches zero as r increases; we can construct an asymptotically flat spacetime without iterative integration. We have also assumed $K(t=0,r) < 0$: every point in the spacetime is locally expanding. The numerical boundary is fixed at $r = 30/(e\eta)^{-1}$ or $60/(e\eta)^{-1}$.

In order to solve the dynamical equations, we use a finite difference method with 2000 to 10,000 meshes. Now we have six dynamical variables: A, B, K, K_2^2, Φ and w. Equations (2.9)-(2.12) provide the next time-step of A, B, K, Φ and w, respectively. At each step, we integrate (2.8) in the r-direction to obtain K_2^2. In this way we have reduced spatial derivatives appearing in the equations, which may become seeds for numerical instability. The Hamiltonian constraint equation (2.7) remains unsolved during the evolution and is used for checking numerical accuracy. We stop numerical computation when some errors exceed a few percent.

In order to understand the spacetime structure from the numerical data, it is useful to observe the signs of the expansion of a null geodesic congruence. Nambu and Siino [7] also utilized this tool to study wormhole formation in a singlet scalar field system. For the metric (2.4), the expansion $\Theta_\pm$ is written as

$$\Theta_\pm = k^2_{\pm;2} + k^3_{\pm;3} = 2\Big[-K_2^2 \pm \frac{(Br)'}{ABr}\Big], \tag{2.15}$$

where $k^\mu_\pm = (-1, \pm A^{-1}, 0, 0)$ is an outgoing $(+)$ or ingoing $(-)$ null vector. We observe the signs of $\Theta_\pm$ at all points in the numerical spacetime. For later convenience, we define "RI" as the region where both Θ^+ and Θ^- are positive, and "RII" as the region where they are negative. We can interpret that the region around RI is de Sitter-like and the region around RII is Schwarzschild-like. The two-surface which bounds RI or RII is called an "apparent horizon". Later, we will use the term "black hole horizon" to refer to any boundary of RII.

And, we will use the term "cosmological horizon" in the sense that no information beyond it reaches the center of a monopole; only the innermost boundary of RI is called the cosmological horizon.

3 Numerical Results

Before we move on to numerical simulation, we offer a rough discussion on the effect of the gauge fields on the gravitational field. In our previous paper [6], we investigated the effect for static monopole solutions and found that the gauge fields generate an attractive force. This property is also described in the time-dependent coordinate system (2.4) as follows.

In a homogeneous and isotropic spacetime, the evolution of the scale factor $a(t)$ follows

$$\frac{3\ddot{a}}{a} = -\frac{4\pi}{m_{\rm Pl}^2}(\rho + 3p), \tag{3.1}$$

where ρ and p are the energy density and the pressure of a matter, respectively. Equation (3.1) indicates that the sign of $\rho + 3p$ determines whether the acceleration of the cosmic expansion is positive or negative. We can extend this discussion to general spacetimes: the sign of $\rho + \Sigma p_i \equiv -T_0^0 + T_i^i$ determines whether a local region expands with positive acceleration or not. A corresponding equation in the present system to (3.1) is (2.9). At the origin, (2.9) reduces to

$$\frac{3\ddot{A}}{A}\Big|_{r=0} = -\frac{4\pi}{m_{\rm Pl}^2}(\rho + \Sigma p_i)_{r=0} \equiv -\frac{4\pi}{m_{\rm Pl}^2}\Big(-2V + \frac{3w''^2}{e^2A^4}\Big)_{r=0} \tag{3.2}$$

In the case of global monopoles, the second term of the right-hand side disappears, and hence the central region is always locally de Sitter spacetime. If gauge fields exist, however, the local acceleration at the center also depends on the second term. Although the exact value of $\partial^2 w/(A\partial r)^2$ cannot be determined without solving the full dynamical equations, we can estimate its order by use of the static solution in a flat spacetime with $\lambda = 0$. Assuming $\partial^2 w/(A\partial r)^2 = b\, w''_{\rm flat}(r)$ with $b = O(1)$, we have

$$\rho + \Sigma p_i|_{r=0} = \frac{e^2\eta^4}{4}\Big(-\frac{\lambda}{e^2} + \frac{2}{3}b\Big), \tag{3.3}$$

which lets us understand how the local expansion of the spacetime in the center depends on λ/e^2. We see that, if $\lambda/e^2 \ll 1$, the monopole core is an attractive spacetime; while, if $\lambda/e^2 \gg 1$, it is repulsive like de Sitter spacetime. Of course, if the initial configuration is quite different from that of the static solutions, i.e., $b \neq O(1)$, the above discussion is not true. The dynamics may also depend on initial configuration.

In what follows, by use of the method in Sec. II, we will numerically integrate the field equations (2.7)-(2.11). To show the results, we define X as a proper distance in the radial direction: $X \equiv \int_0^r A dr$. We also define the boundaries of a monopole in two ways: $X_\Phi(t) = X$ at the position of $\Phi = \eta/2$ and $X_w(t) = X$ at the position of $w = 1/2$. We normalize time and length by the horizon scale defined as $H_0^{-1} \equiv (8\pi V(0)/3m_{\rm Pl}^2)^{-\frac{1}{2}}$.

First, we check our numerical code by solving the equations for the case of weak gravity. In Fig. 1 we set $\eta = 0.1 m_{\rm Pl}$ and $\lambda/e^2 = 0.1$, and give two initial configurations: $c_\Phi = c_w = 1$ and 0.5. We plot the trajectories of $X_\Phi(t)$. We find that the fields behaves stably; this reasonable result indicates that our numerical code works well.

From now on, we concentrate on the parameter range where no static solution exists. In Fig. 2 we set $\eta = 0.4 m_{\rm Pl}$ and $\lambda/e^2 = 0.1$, and give two initial configurations: $c_\Phi = c_w = 1$ in (a) and $c_\Phi = c_w = 10$ in (b). In Fig. 2(a)(b) we plot the trajectories of $X_\Phi(t)$ and of $X_w(t)$ as well as apparent horizons. The dynamics in these two cases contrast sharply: in (a) a monopole shrinks and black-hole horizons appear, while in (b) a cosmological horizon exists from the beginning and a monopole continues to expand. We also draw the distributions of $\rho + \Sigma p_i$ in (c) and in (d), which correspond to the results in (a) and in (b), respectively. In (c) the values around the center become negative at the beginning, but they bounce back to positive values, which confirms that the monopole core never inflates. On the other hand, in (d) the values of $\rho + \Sigma p_i$ around the center remain negative from the beginning. This behavior indicates that exponential expansion really occurs inside the monopole. These two results tell us that monopoles for $\eta > \eta_{\rm cr}$ tend to be dynamic, and their dynamics depends on initial configuration, contrary to the case of global monopoles.

As we will see soon, for larger η, monopoles are more likely to inflate rather than shrink. We show an example for larger η in Fig. 3; we set $\eta = 0.55 m_{\rm Pl}$, $\lambda/e^2 = 0.1$ and $c_\Phi = c_w = 1$. In Fig. 3(a) we plot the trajectories of $X_\Phi(t)$ and of $X_w(t)$ as well as apparent horizons. (Please also refer to Fig. 4, which is a schematic sketch of the spacetime structure.) From the beginning there are two apparent horizons, $S1$ and $S2$: $S1$ is the cosmological horizon. Later other two apparent horizons, $S3$ and $S4$, appear, and then $S2$ and $S4$ approach each other. These surfaces turn out to be black-hole horizons, $S2'$ and $S4'$, the moment they intersect. In Fig. 3(b) we draw the distributions of $\rho + \Sigma p_i$. Contrary to the case of the contracting monopole in Fig. 2(a), the values around the center are initially positive, but they become negative. This suggests that, if η is large enough, a monopole begins to expand exponentially even if its initial size is not so large. We also show in Fig. 3(c) the relation between the proper distance along the radial direction and the circumference radius, which indicates a wormhole structure really appears. Figure 4 lets us understand how the wormhole is created. Because the expanding core is causally disconnected from the outer region, such an isolated region is called a "child universe".

One may think that if $\eta > \eta_{\rm cr}$, a monopole either expands or collapses, as shown in Fig. 2 or 3. However, we find some cases where a monopole neither expands nor collapses. An example of such solutions is presented in Fig. 5 ($\eta = 0.3 m_{\rm Pl}$ and $\lambda/e^2 = 10$). Setting $c_\Phi = c_w = 1$, we show the evolution of Φ in (a) and that of w in (b), and the trajectories of $X_\Phi(t)$ and of $X_w(t)$ in (c). Although some oscillations remain outside the monopole, the core of the monopole approaches a stable configuration. We change the initial size in Fig. 5(d), finding monopoles with any initial size behave stably. These results indicate that there exist stationary solutions. The existence of such solutions looks surprising, because one may expect that all solutions in the parameter range where no static solution exists must be dynamical. The existence of "stationary" solutions, however, does not contradict the nonexistence of "static" solutions: the solutions in Fig. 5 cannot be described with a static coordinate system because the size of the monopole is greater than the cosmological horizon.

Finally we systematically survey the dynamics of monopoles for $0.05 \leq \eta/m_{\rm Pl} \leq 0.55$ and $0.1 \leq \lambda/e^2 \leq 10$ and summarize the solutions in the λ/e^2-$\eta/m_{\rm Pl}$ plane of Fig. 7. A square denotes a stable solution, as is the case in Fig. 1 or 5. A cross denotes the case where a monopole shrinks, as is the case in Fig. 2(a). A circle denotes the case where a monopole inflates and the wormhole structure appears, as is the case in Fig. 3. A dotted line indicates the maximum values of $\eta/m_{\rm Pl}$ versus λ/e^2, depicted approximately by use of Fig.6 in [1]. We

vary c_Φ and c_w from 1 to 10, and hence some parameter points are labeled as two symbols. We interpret these results as follows. In the case of $\lambda/e^2 > 1$, a monopole expands exponentially if $\eta \gtrsim 0.35 m_{Pl}$; this critical value has little dependence on λ/e^2 and initial configuration, and almost agrees with that for global monopoles [6]. This agreement is quite reasonable because the effect of the gauge fields is smaller as λ/e^2 is larger. Below the critical value, a monopole tends to take a stable configuration even in the theories where static solutions are nonexistent. In the case of $\lambda/e^2 < 1$, the dynamics also depend on λ/e^2 and initial configuration. In some cases the effect of the gauge fields becomes dominant and a monopole shrinks and becomes a black hole. These results are consistent with our analytic discussions at the beginning of this section.

4 Summary and Discussions

We have studied the dynamics of magnetic monopoles numerically. Our main purpose has been to understand the behavior of monopoles in the case where static solutions are nonexistent.

If η is large enough ($\gg \eta_{\rm cr}$), a monopole inflates and a wormhole structure appears around it. We have shown how the wormhole connected with a child universe is created. We should emphasize that a child universe can be generated without fine-tuned initial conditions in this model, contrary to the case of a trapped false vacuum bubble [4]. In the case of $\lambda/e^2 > 1$, the condition of inflation is $\eta \gtrsim 0.35 m_{Pl}$, which has little dependence on λ/e^2 and initial configuration. Below the critical value, a monopole tends to take a stable configuration even in the theories where static solutions are nonexistent. This is true for any initial configuration, which indicates the existence of stationary solutions. While, in the case of $\lambda/e^2 < 1$, the dynamics also depends on λ/e^2 and initial configuration. In some cases the effect of the gauge fields becomes dominant and a monopole collapses into a black hole.

We should notice that the condition of inflation was also estimated analytically in a simplified model by Tachizawa et al.[3] They discussed the global structure of a spacetime by regarding the inside the monopole core as de Sitter spacetime and the outside as Reissner-Nordström spacetime. They showed that the surface of the monopole core exceeds a cosmological horizon if $\eta > m_{\rm Pl}/\sqrt{3\pi} \approx 0.33 m_{\rm Pl}$. This condition almost agrees with the condition of inflation for most cases in our analysis; this agreement suggests that our numerical results are reasonable as well as that their simplified model is a good approximation in most cases. When the effect of the gauge fields is dominant to that of the Higgs field, however, the spacetime is not de Sitter-like, and then the validity of the simplified model is lost.

Our results as a whole support the discussions of Linde and Vilenkin [5], who pointed out the possibility of monopole inflation. Actually, we have found that inflation happens in most cases of $\eta > \eta_{\rm cr}$. Further, if the initial size of a monopole is large enough, the effect of the gauge fields is not important, as Linde mentioned. What we have clarified more about this subject is there are some cases where static solutions are nonexistent but monopoles do not continue to expand, as the results in Fig. 5. Although we did not state this fact in our previous paper [6], it is also true for global monopoles.

Acknowledgments

The author would like to thank J. Koga, A. Linde, K. Maeda, D. Maison, T. Tachizawa, T. Torii, and A. Vilenkin for useful discussions. Thanks are also due to P. Haines and W.

Rozycki for reading the manuscript. This work was supported partially by the Grant-in-Aid for Scientific Research Fund of the Ministry of Education, Science and Culture (No.07740226) and by a Waseda University Grant for Special Research Projects.

References

[1] P. Breitenlohner, P. Forgàcs, and D. Maison, Nucl. Phys. **B383**, 357 (1992).

[2] M.E. Ortiz, Phys. Rev. D **45**, R2586 (1992);
K. Lee, V.P. Nair, and E.J. Weinberg, *ibid.* **45**, 2751 (1992);
P. Breitenlohner, P. Forgàcs, and D. Maison, Nucl. Phys. **B442**, 126 (1995).

[3] T. Tachizawa, K. Maeda, and T. Torii, Phys. Rev. D **51**, 4054 (1995).

[4] K. Sato, M. Sasaki, H. Kodama, and K. Maeda, Prog. Theor. Phys. **65**, 1443 (1981);
K. Sato, H. Kodama, M. Sasaki, and K. Maeda, Phys. Lett. B **108**, 103 (1982).

[5] A. Linde, Phys. Lett. B **327**, 208 (1994); A. Linde and D. Linde, Phys. Rev. D **50**, 2456 (1994); A. Vilenkin, Phys. Rev. Lett. **72**, 3137 (1994).

[6] N. Sakai, H. Shinkai, T. Tachizawa, and K. Maeda, Phys. Rev. D **53** 655 (1996).

[7] Y. Nambu and M. Siino, Phys. Rev. D **46**, 5367 (1992).

Figure Captions

Fig. 1. Dynamics of a monopole for the case of weak gravity. We set $\eta = 0.1 m_{\rm Pl}$ and $\lambda/e^2 = 0.1$, and give two initial configurations: $c_\Phi = c_w = 1$ and 0.5. We plot the trajectories of $X_\Phi(t)$. The Higgs field behaves stably; these reasonable results indicates that our numerical code works well.

Fig. 2. Dynamics of a monopole for $\eta = 0.4 m_{\rm Pl}$ and $\lambda/e^2 = 0.1$. We assume two initial configurations: $c_\Phi = c_w = 1$ in (a) and $c_\Phi = c_w = 10$ in (b). In (a) and (b) we plot the trajectories of $X_\Phi(t)$ and of $X_w(t)$ as well as apparent horizons. In (a) a monopole shrinks and the black-hole horizons appear, while in (b) a cosmological horizon exists from the beginning and a monopole continues to expand. We also draw the distributions of $\rho + \Sigma p_i$ in (c) and in (d), which correspond the results in (a) and in (b), respectively. In (c) the values around the center get negative at the beginning, but they bounce back to positive values, which confirms that the monopole core never inflates. On the other hand, in (d) the values of $\rho + \Sigma p_i$ around the center remain negative from the beginning. This behavior indicates that exponential expansion really occurs inside the monopole.

Fig. 3. Dynamics of a monopole for $= 0.55 m_{\rm Pl}$, $\lambda/e^2 = 0.1$ and $c_\Phi = c_w = 1$. In (a) we plot the trajectories of $X_\Phi(t)$ and of $X_w(t)$ as well as apparent horizons. (Please also refer to Fig. 4, which is a schematic sketch of the spacetime structure.) From the beginning there are two apparent horizons, $S1$ and $S2$: $S1$ is the cosmological horizon. Later other two apparent horizons, $S3$ and $S4$, appear, and then $S2$ and $S4$ approach each other. These surfaces turn out to be black-hole horizons, $S2'$ and $S4'$, the moment they intersect. In (b) we draw the distributions of $\rho + \Sigma p_i$. The values around the center are initially positive, but they become negative. This suggests that, if η is large enough, a monopole begins to expand exponentially even if its initial size is not so large. We also show in (c) the relation between the proper

the proper distance along the radial direction and the circumference radius, which indicates a wormhole structure really appears.

Fig. 4. Schematic sketches of the spacetime structure. These figures are not generated from the numerical data, but they are based on the results presented in Fig. 3.

Fig. 5. Dynamics of a monopole for $\eta = 0.3m_{\rm Pl}$ and $\lambda/e^2 = 1$. Setting $c_\Phi = c_w = 1$, we show the evolution of Φ in (a) and that of w in (b), and the trajectories of $X_\Phi(t)$ and of $X_w(t)$ in (c). Although some oscillations remain outside the monopole, the core of the monopole approaches a stable configuration. We change the initial size in (d), finding monopoles with any initial size behave stably. These results indicate that there exist stationary solutions.

Fig. 6. Parameter plane of $\eta/m_{\rm Pl}$ and λ/e^2 in which we summarize our numerical results. A square (□) denotes a stable solution, as is the case in Fig. 1 or 6. A cross (×) denotes the case where a monopole shrinks, as is the case in Fig. 2(a). A circle (○) denotes the case where a monopole inflates and the wormhole structure appears, as is the case in Fig. 2(c) or 4. A dotted line indicates the maximum values of $\eta/m_{\rm Pl}$ versus λ/e^2, depicted approximately by use of Fig.6 in [1]. We vary c_Φ and c_w from 1 to 10, and hence some parameter points are labeled as two symbols.

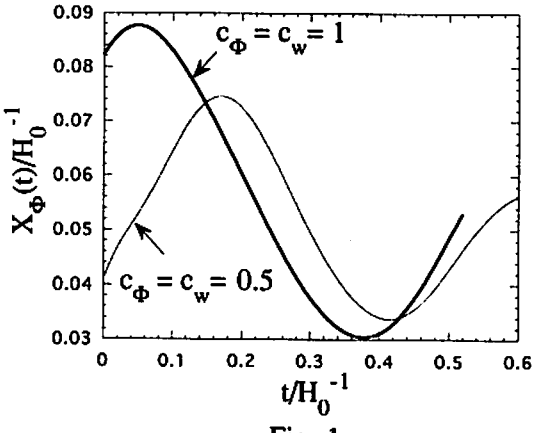


Fig. 1

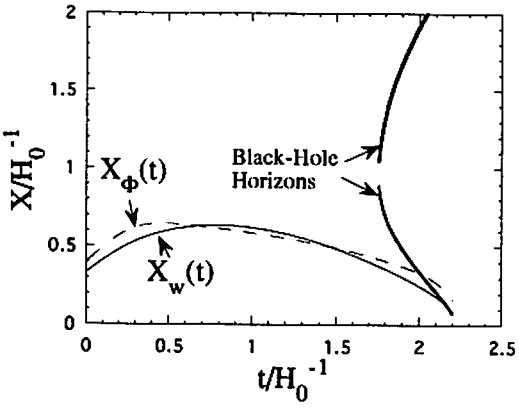


Fig.2(a)

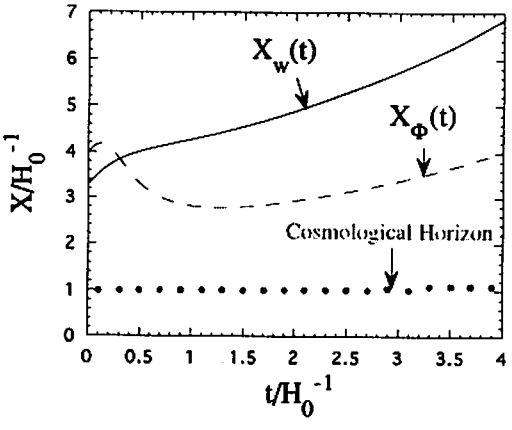


Fig.2(b)

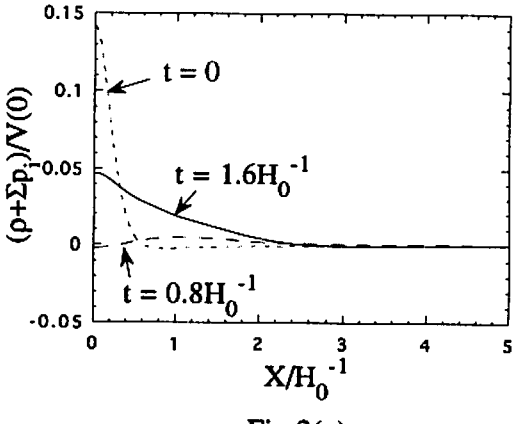


Fig.2(c)

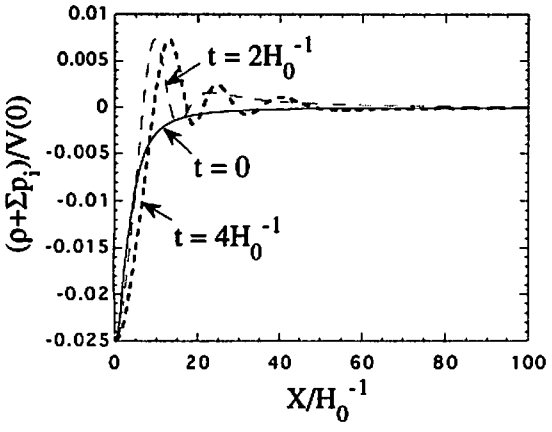


Fig.2(d)

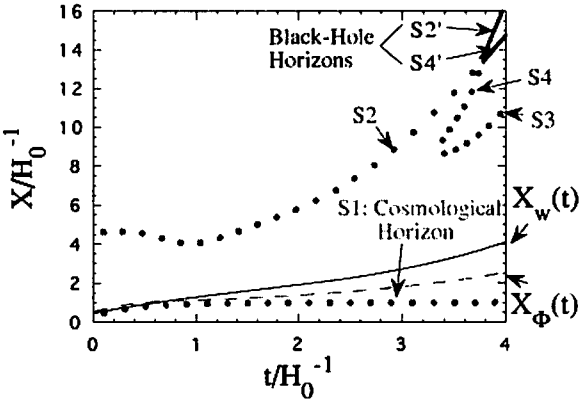


Fig.3(a)

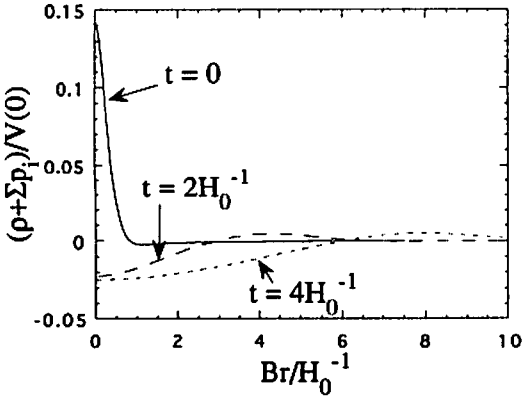


Fig.3(b)

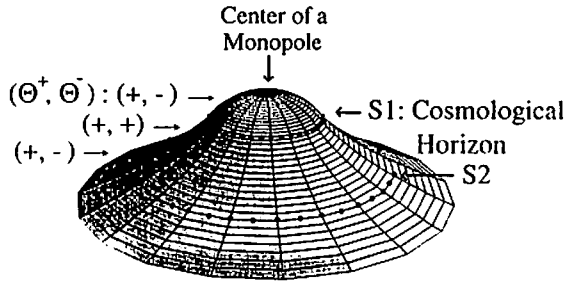


Fig.4(a)

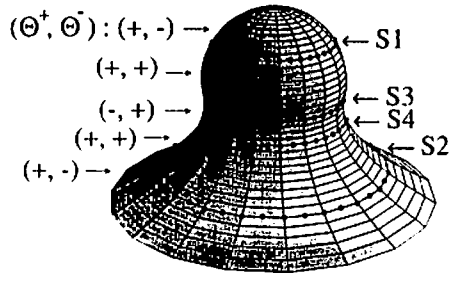


Fig.4(b)

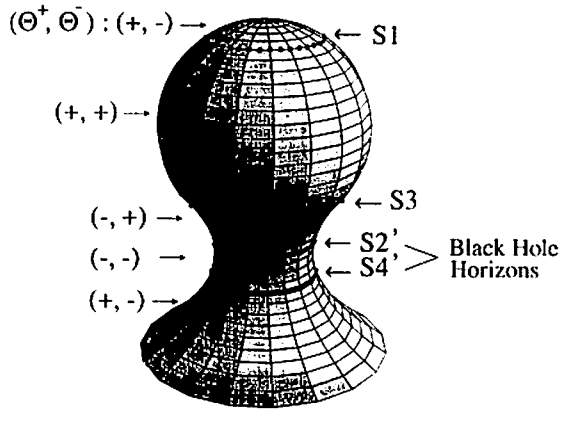


Fig.4(c)

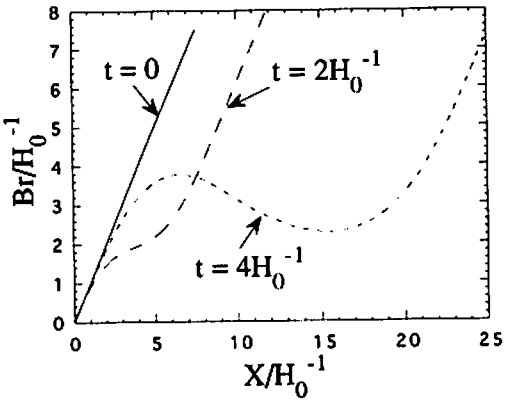


Fig.3(c)

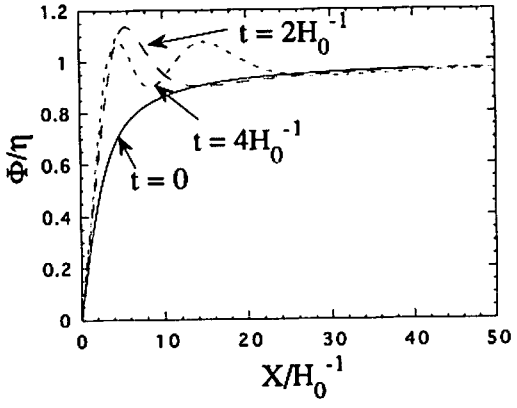


Fig.5(a)

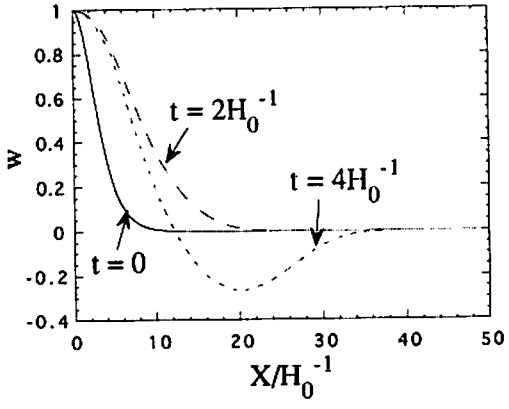


Fig.5(b)

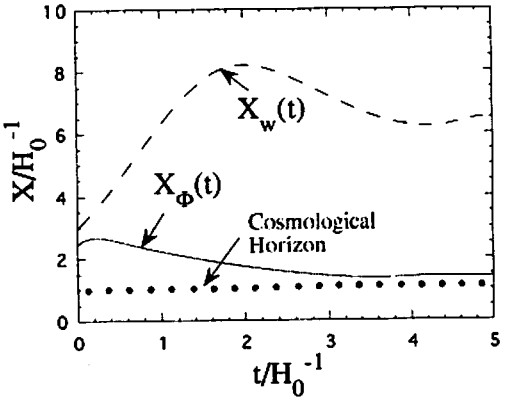


Fig.5(c)

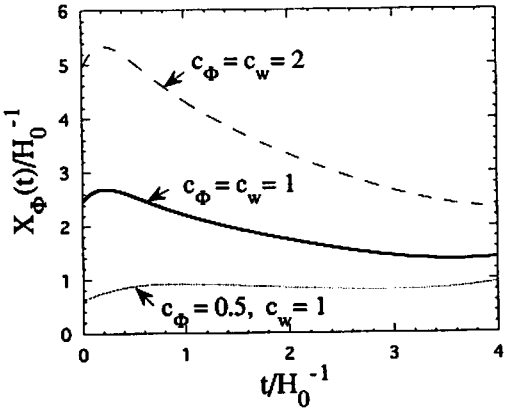


Fig.5(d)

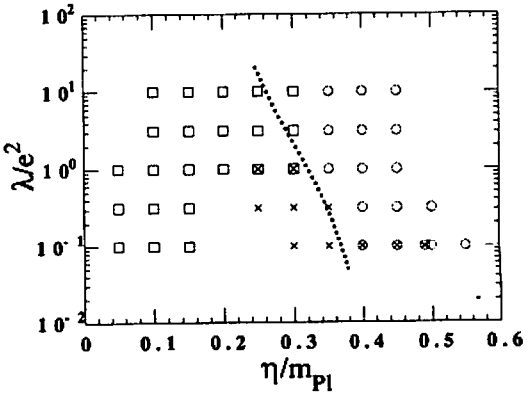


Fig.6

Inflation in the inhomogeneous universe

Osamu Iguchi and Hideki Ishihara

Department of Physics
Tokyo Institute of Technology, Meguroku, Tokyo 152, Japan

Abstract

We studied how the initial inhomogeneity of the spatial curvature affects the onset of inflation in the closed universe. We consider the model of a chaotic inflation which is driven by a massive scalar field. In order to construct an inhomogeneous universe model, we use the long wavelength approximation (the gradient expansion method). We show the condition of the inhomogeneity for the universe to enter the inflationary phase.

1 Introduction

Inflationary scenario explains the reason why the universe is homogeneous and isotropic on large scales. An important question concerning inflationary cosmology is whether inflation is generic or needs a fine tuning of initial conditions.

For the Friedmann-Robertoson-Walker (FRW) models with a massive scalar field, many authors [1] studied this problem. Belinsky *et al* [2] concluded that inflation is a general property of FRW models with a massive scalar field except for the closed one. In the homogeneous universe, whether the universe will enter the inflationary phase is almost determined by the competition between a positive spatial curvature and a effective positive cosmological constant which drives inflation.

The role of initial inhomogeneity on the occurrence of inflation was studied by a numerical simulation. This problem was investigated in the case of planar symmetry [3] and in the case of spherical symmetry [4]. In ref. [4], it is concluded that the crucial feature necessary for inflation is a sufficient high average field over a region of several horizon sizes.

We study how the initial inhomogeneity of the spatial curvature affects the onset of inflation in the closed universe by use of an alternative approach. We consider the model of chaotic inflation which is driven by a massive scalar field. In order to treat the inhomogeneity of the gravitational field coupled to a massive scalar field, we use the long wavelength approximation (the gradient expansion method).

The long wavelength iteration scheme is a method to construct an approximate solution of Einstein's equations which describes an inhomogeneous universe on large scales. This approximation is assumption that all spatial

gradients are small compared to time derivatives. In the iteration scheme, the spatial metric can be expanded as a sum of spatial tensors of increasing order in spatial gradients. Thus the approximation is called the gradient expansion. It was introduced by Lifschitz and Khalatnikov [5]. They investigated the behavior of the spacetime near the Big Bang. Using this method, Tomita [6] studied the evolution of irregularities on super-horizon scales. For the higher order solutions, Comer, Deruelle, Langlois *et al* [7] developed an iteration scheme of Einstein's equations and Salopek, Stewart *et al* [8] developed one of Hamilton-Jacobi equation for general relativity.

Recently the influence of initial inhomogeneity on the occurrence of inflation is studied by using the gradient expansion method [9, 10]. From investigations in the homogeneous universe model, it is important to study the inhomogeneity in the closed universe because a positive spatial curvature prevents universe from inflating. However the spatial curvature is treated as a small quantity in the gradient expansion scheme.

In this paper, we improve the approximation scheme in order to treat the non-small spatial curvature. Then we show the condition of the inhomogeneity of the spatial curvature for the universe to enter the inflationary phase.

2 The influence of inhomogeneity for inflation in the closed universe

In this section, we derive the first order equation of gradient expansion in the closed universe with a scalar field.

The Einstein equations for the gravitational field coupled to a scalar field, in a synchronous reference frame, read

$$\frac{1}{2}\dot{K} + \frac{1}{4}K^l_m K^m_l = \kappa\left[-\dot{\phi}^2 + V(\phi)\right], \tag{2.1}$$

$$R^j_i + \frac{1}{2\sqrt{\gamma}}\partial_t(\sqrt{\gamma}K^j_i) = \kappa\left[\partial_i\phi\partial^j\phi + V(\phi)\delta^j_i\right], \tag{2.2}$$

$$\frac{1}{2}(K^j_{i;j} - K_{;i}) = \kappa\dot{\phi}\partial_i\phi, \tag{2.3}$$

where R^j_i is the Ricci tensor associated with γ_{ij}, $K_{ij} \equiv \dot{\gamma}_{ij}$, $\gamma \equiv \det\gamma_{ij}$, a dot denotes the derivative with respect to t, a semicolon denotes the covariant derivative with respect to γ_{ij}, $V(\phi)$ is the potential of the scalar field and $\kappa \equiv 8\pi G$. Throughout this paper Latin letters will denote spatial indices. The equation of motion for the scalar field is

$$\ddot{\phi} + \frac{1}{2}K\dot{\phi} - \phi^{;l}_{;l} + \partial_\phi V(\phi) = 0. \tag{2.4}$$

At lowest order, we assume the space-dependence of the solution in the form

$$\overset{(0)}{\gamma}_{ij}(t,x) \equiv a^2(t,\Omega(x))h_{ij}, \tag{2.5}$$

$$\overset{(0)}{\phi}(t,x) \equiv \phi_0(t,\Omega(x)), \tag{2.6}$$

where h_{ij} is the metric of S^3. The spatial metric and the scalar field in the lowest order has the space-dependence by an arbitrary spatial function $\Omega(x)$.

Substituting Eq.(2.5) and Eq.(2.6) into Eq.(2.1), Eq.(2.2) and Eq.(2.4) and neglecting all spatial derivatives, we obtain the equations which a and ϕ_0 must satisfy:

$$\frac{\ddot{a}}{a} = \frac{\kappa}{3}\left[-\dot{\phi}^2 + V(\phi)\right], \tag{2.7}$$

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{1}{4a^2} = \frac{\kappa}{12}\left[\frac{1}{2}\dot{\phi}^2 + V(\phi)\right], \tag{2.8}$$

$$\ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} + \partial_\phi V(\phi) = 0. \tag{2.9}$$

These equations have the same form of the equation for the homogeneous model.

Here we show the behavior of the lowest order solution. We think the special model that $V(\phi) = \frac{m^2}{2}\phi^2$. For simplicity, we consider the case of $\frac{1}{2}\dot{\phi}^2 >> V(\phi)$ initially. At the early time, the asymptotic behavior of the scale factor and the scalar field is

$$a^2(t,x) = t^{\frac{2}{3}}\Omega^2(x), \tag{2.10}$$

$$\phi_0(t,x) = \sqrt{\frac{2}{3\kappa}}\left[\ln t + \frac{81}{54\Omega^2}t^{\frac{4}{3}}\right] + C, \tag{2.11}$$

where C is a constant of space and time variables. In the early stage, the scale factor is proportional to $\Omega(x)$ then the spatial dependence of the spatial curvature $R_{initial} \propto \Omega^{-2}$. The constant C corresponds to the freedom of initial value of the scalar field $\phi_{initial}$. Figure.1 shows the inflationary region and the recollapsing one by varying Ω and C. From Figure.1, we can see that $a(t,x)$ inflates in the region where Ω is big ($R_{initial}$ is small) and C is big ($\phi_{initial}$ is big). This result corresponds with the case of homogeneous universe because the lowest order equations are same as homogeneous ones.

At next order, we consider the parts of two spatial gradients of $\Omega(x)$. We will take corrections to the metric and the scalar field of the form

$$\overset{(1)}{\gamma}_{ij}(t,x) = a^2(t,\Omega(x))\left[\frac{1}{3}F(t,\Omega)\frac{\nabla_l\nabla^l\Omega}{\Omega}h_{ij} + \overline{F}(t,\Omega)\frac{\overline{\nabla_i\nabla_j\Omega}}{\Omega}\right.$$

$$+\frac{1}{3}G(t,\Omega)\frac{\nabla_l\Omega\nabla^l\Omega}{\Omega^2}h_{ij}+\overline{G}(t,\Omega)\frac{\overline{\nabla_i\Omega\nabla_j\Omega}}{\Omega^2}\Bigg], \tag{2.12}$$

$$\overset{(1)}{\phi}(t,x) = P(t,\Omega)\frac{\nabla_l\nabla^l\Omega}{\Omega}+Q(t,\Omega)\frac{\nabla_l\Omega\nabla^l\Omega}{\Omega^2}, \tag{2.13}$$

where ∇ denotes the covariant derivative with respect to h_{ij}, the over-line denotes its traceless part.

Substituting $\gamma_{ij}=\overset{(0)}{\gamma}_{ij}+\overset{(1)}{\gamma}_{ij}$ into Eq.(2.3) and comparing the coefficients of the first order derivative of Ω, we obtain

$$\dot{\overline{F}} = 2\left[\frac{\dot{a}'}{a}-\frac{\dot{a}a'}{a^2}\right]\Omega+\kappa\dot{\phi}_0\phi_0'\Omega, \tag{2.14}$$

where a prime denotes the derivative with respect to Ω. Note readers should pay attention to the commutation relation for the covariant derivative. Substituting γ_{ij} into Eq.(2.1) and comparing the coefficients of the second order derivative of Ω, we get

$$\ddot{F}+2H\dot{F} = 2\kappa\left[-2\dot{\phi}_0\dot{P}+m^2\phi_0P\right], \tag{2.15}$$

$$\ddot{\tilde{G}}+2H\dot{\tilde{G}} = 2\kappa\left[-2\dot{\phi}_0\dot{Q}+m^2\phi_0Q\right]+\dot{\overline{F}}^2. \tag{2.16}$$

where $\tilde{G}\equiv G+\overline{F}^2$. Similarly from Eq.(2.4), we have

$$\ddot{P}+3H\dot{P}+m^2P = -\frac{1}{2}\dot{\phi}_0\dot{F}+\frac{1}{a^2}\phi_0'\Omega, \tag{2.17}$$

$$\ddot{Q}+3H\dot{Q}+m^2Q = -\frac{1}{2}\dot{\phi}_0\dot{\tilde{G}}+\frac{1}{a^2}\left[\phi_0''+\frac{a'}{a}\phi_0'\right]\Omega^2. \tag{2.18}$$

From Eq.(2.2), we obtain

$$\ddot{F}+6H\dot{F} = 6\kappa m^2\phi_0P+\frac{4}{a^2}\left[F-\overline{F}+2\frac{a'}{a}\Omega\right], \tag{2.19}$$

$$\ddot{\overline{F}}+3H\dot{\overline{F}} = \frac{2a'}{a^3}\Omega, \tag{2.20}$$

$$\ddot{G}+3H\dot{G} = 6\kappa m^2\dot{\phi}_0Q+\frac{2\kappa}{a^2}\phi_0'^2\Omega^2 +\frac{4}{a^2}\left[\dot{G}+\overline{F}-\overline{F}'\Omega+\left(2\frac{a''}{a}-\frac{a'^2}{a^2}\right)\Omega^2\right], \tag{2.21}$$

$$\ddot{\overline{G}}+3H\dot{\overline{G}} = \frac{2\kappa}{a^2}\phi'^2\Omega^2 +\frac{2}{a^2}\left[\overline{G}+\overline{F}-\overline{F}'\Omega-\frac{1}{2}\overline{F}^2+\frac{a'\Omega}{2a}\overline{F}+\left(\frac{a''}{a}-\frac{2a'^2}{a^2}\right)\right] \tag{2.22}$$

We obtain the nine equations Eq.(2.14) $\sim$ Eq.(2.22) for the six variables $F, \overline{F}, \tilde{G}, \overline{G}, P$ and Q. These equations are consistent. In order to solve Eq.(2.17) $\sim$ Eq.(2.22), we need to know the time evolution of $a', a'', \dot{a}', \phi_0', \phi_0''$ and $\overline{F}'$. Differentiating Eq.(2.7) $\sim$ Eq.(2.9) and Eq.(2.14), we obtain the equations for them.

We choose initial conditions as follows:

$$F = \overline{F} = \overline{F}' = \dot{G} = \overline{G} = P = Q = 0,$$
$$\dot{F} = \frac{24}{7\Omega^2} t^{\frac{1}{3}}, \dot{\tilde{G}} = -\frac{12}{7\Omega^2} t^{\frac{1}{3}}, \dot{\overline{G}} = -\frac{3}{\Omega^2} t^{\frac{1}{3}},$$
$$\dot{P} = -\sqrt{\frac{3}{2\kappa}} \frac{6}{7\Omega^2} t^{\frac{1}{3}}, \dot{Q} = \sqrt{\frac{3}{2\kappa}} \frac{3}{7\Omega^2} t^{\frac{1}{3}}.$$

This condition means that we take the growing mode of the solution.

At each point, we define a local scale factor which includes up to the first order by

$$\begin{aligned} a_{local}^2 &\equiv \left[\det\left[\overset{(0)}{\gamma}_{ij} + \overset{(1)}{\gamma}_{ij}\right]\right]^{1/3} \\ &= a^2\left[1 + \frac{1}{3}F\frac{\nabla_l\nabla^l\Omega}{\Omega} + \frac{1}{3}\tilde{G}\frac{\nabla_l\Omega\nabla^l\Omega}{\Omega^2}\right]. \end{aligned} \tag{2.23}$$

We can follow the evolution of the local scale factor $a_{local}(t, x)$ by the spatial derivatives of Ω and the evolution of $a(t, \Omega)$, $F(t, \Omega)$ and $\tilde{G}(t, \Omega)$. The local expansion rate and acceleration rate are given by

$$\begin{aligned} \frac{\dot{a}_{local}}{a_{local}} &= \frac{\dot{a}}{a} + \frac{1}{6}\left[\dot{F}\frac{\nabla_l\nabla^l\Omega}{\Omega} + \dot{\tilde{G}}\frac{\nabla_l\Omega\nabla^l\Omega}{\Omega^2}\right], \\ \frac{\ddot{a}_{local}}{a_{local}} &= \frac{\ddot{a}}{a} + \frac{1}{6}\left[\ddot{F} + \frac{2\dot{a}}{a}\dot{F}\right]\frac{\nabla_l\nabla^l\Omega}{\Omega} + \frac{1}{6}\left[\ddot{\tilde{G}} + \frac{2\dot{a}}{a}\dot{\tilde{G}}\right]\frac{\nabla_l\Omega\nabla^l\Omega}{\Omega^2}. \end{aligned}$$

The time evolution of the above two quantities is determined by the value of $C, \Omega, \nabla_l\nabla^l\Omega/\Omega$ and $\nabla_l\Omega\nabla^l\Omega/\Omega$. We divide the four-dimensional parameter space into two regions, inflationary region and recollapsing region. The local scale factor with the parameters in the inflationary region enter the phase with accelerating expansion: $\dot{a}_{local}/a_{local} > 0$ and $\ddot{a}_{local}/a_{local} > 0$. On the other hand, the local scale factor with the parameters in the recollapsing region enter recollapsing phase: $\dot{a}_{local}/a_{local} < 0$ and $\ddot{a}_{local}/a_{local} < 0$.

While we compute the time evolution, we assume the approximation is valid when

$$\left|\left[\det\left(\overset{(0)}{\gamma}_{ij} + \overset{(1)}{\gamma}_{ij}\right) - \det\overset{(0)}{\gamma}_{ij}\right] / \det\overset{(0)}{\gamma}_{ij}\right| < 0.5,$$
$$\left|\overset{(1)}{R}(\gamma) / \overset{(0)}{R}(\gamma)\right| < 0.5.$$

3 Numerical results

Investigating the evolution of the local expansion rate and acceleration rate numerically, we can see the role of two spatial gradient, $\nabla_l\nabla^l\Omega/\Omega$ and $\nabla_l\Omega\nabla^l\Omega/\Omega^2$, on the occurrence of inflation. The result is shown in Figure.2. The positive $\nabla_l\nabla^l\Omega/\Omega$ helps the universe tends to enter the inflationary phase. The negative value of it helps the universe to recollapse. On the other hand, $\nabla_l\Omega\nabla^l\Omega/\Omega^2$ tends to prevent the onset of inflation.

From the result, we obtain the condition for the onset of inflation,

$$\frac{1}{\Omega^2}\left[1-50\frac{\nabla_l\nabla^l\Omega}{\Omega}+100\frac{\nabla_l\Omega\nabla^l\Omega}{\Omega^2}\right]<\psi(t), \tag{3.1}$$

where $\psi(t)$ is a constant of space and corresponds to the marginal value of the scalar field in the homogeneous universe model. In the early stage, the spatial curvature $R_{initial}$ is

$$R_{initial}=\frac{2}{t^{\frac{2}{3}}\Omega^2}\left[3-2\frac{\nabla_l\nabla^l\Omega}{\Omega}+\frac{\nabla_l\Omega\nabla^l\Omega}{\Omega^2}\right]. \tag{3.2}$$

This condition described by $R_{initial}$ is

$$R_{initial}\left[1+\frac{74}{3}\frac{\nabla_l\nabla^l R_{initial}}{R_{initial}}-\frac{145}{12}\frac{\nabla_l R_{initial}\nabla^l R_{initial}}{R^2_{initial}}\right]<\Psi(t), \tag{3.3}$$

where $\Psi(t)$ is a constant of space.

D.S.Goldwirth and T.Piran calculate how inhomogeneity influences the inflationary epoch numerically in the case of spherical symmetry by use of a special model. They conclude that the crucial feature necessary for inflation is a sufficiently high average value of the scalar field (suitable value for inflation in homogeneous universe) over a region of several horizon sizes. In order to compare their results with ours, we impose $\Omega^2(x)$ to spherically symmetry, $\Omega^2(\chi,\theta,\varphi)=\Omega^2(\chi)$, in the spherical coordinate where the metric h_{ij} is $h_{ij}=\mathrm{diag}[1,\sin^2\chi,\sin^2\chi\sin^2\theta]$. We consider the form of $\Omega^2(\chi)$ as follows

$$\Omega^2(\chi) = \Omega_0^2\left[1+\delta\Omega\left[\frac{1-\exp\left[-\frac{\cos^2\frac{\chi}{2}}{\Delta^2}\right]}{1-\exp\left[-\frac{1}{\Delta^2}\right]}\right]\right]. \tag{3.4}$$

The form of Ω^2 is shown in Figure.3. The distribution of Ω^2 depends on the three parameter Ω_0^2, $\delta\Omega$ and Δ which are constant of space and time. Δ describes the comoving width of the initial inhomogeneity of the spatial curvature. We study the conditions for entering the inflationary phase at

the origin when we vary Ω_0^2, $\delta\Omega$ and Δ keeping $\Omega_0^2[1+\delta\Omega/2] = \psi^{-1}(t)$ (for example C is 5.0 and Ω^2 is about 1.73015, see Figure.1) unchanged. From now on $m^2 = 0.1$ in $c = \kappa = 1$ unit.

At the origin universe can enter the inflationary phase whenever Δ is smaller than about 0.7. It means that the region of a suitable value of Ω^2 is over $\chi \leq 2.0$ (as a physical length 1.22 for $C = 5.0$, 1.6 for $C = 3.0$). The initial horizon size is $a(0,x)\chi = 0.3$ in our calculation. So we conclude the condition for inflation is a suitable value of Ω^2 over about a several times the initial horizon size at least. The above condition corresponds with Goldwirth and Piran's.

4 Conclusion

We studied how the initial inhomogeneity of the spatial curvature affects the onset of inflation in the closed universe by use of the gradient expansion from a locally closed Freidamnn model. We consider a chaotic inflation which is driven by a massive scalar field.

At lowest order, we assume the spatial dependence of the spatial metric and the scalar field is described by an arbitrary spatial function $\Omega(x)$ which describes the initial inhomogeneity of the spatial curvature. The equations for a and ϕ reduce to locally closed Friedmann equations. At next order, using the local scale factor a_{local}, we investigate the effect of inhomogeneous spatial curvature for inflation.

From numerical results, we obtain the condition for the onset of inflation.

In the case of spherically symmetry, we compare D.S.Goldwirth and T.Piran'results with ours. Our results recover them within the first order of the gradient expansion.

References

[1] See,e.g.,D.S.Goldwirth and T.Piran,Phys.Rep.**214**,223(1992)

[2] V.A.Belinski,H.Ishihara,I.M.Khalatnikov and H.Sato, Prog.theor.Phys.79,676(1988)

[3] H.Sinkai and K.Maeda,Phys.Rev.D49,6367(1994)

[4] D.S.Goldwirth and T.Piran,Phys.Rev.D40,3263(1989)

[5] E.M.Lifchitz and I.M.Khalatnikov,Adv.Phys.**12**,185(1963)

[6] K.Tomita,Prog.Theor.Phys.54,730(1975)

[7] G.L.Comer,D.Deruelle.D.Langlois and J.Parry,
Phys.Rev.D49,2759(1994)

[8] D.S.Salopek and J.M.Stewart,Class.Quant.Grav. 9,1943(1992)

[9] N.Deruuelle and D.S.Goldwirth,Phys.Rev.D51,1563(1995)

[10] Y.Nambu and A.Taruya,preprint.*Application of Gradient Expanion Universe*,(1994)

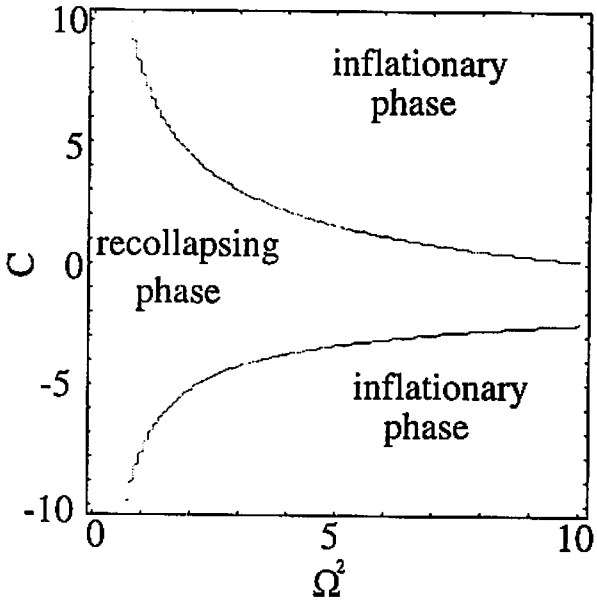


Figure 1: The behavior of the lowest order solution in (Ω^2, C) plane. We consider the case of $m^2 = 0.1$ in $c = \kappa = 1$ unit.

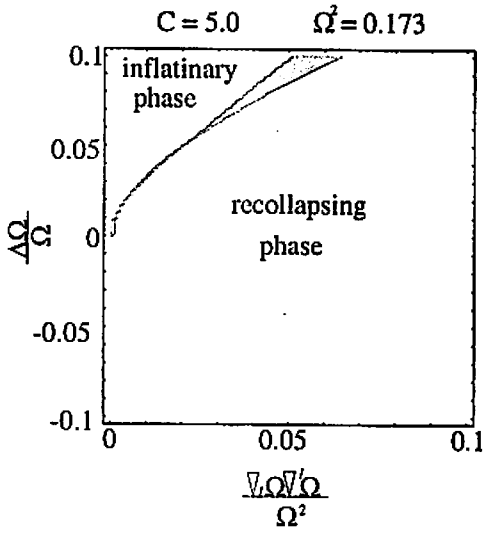


Figure 2: An inflationary region and a recollapsing region are shown in the parameter space ($\nabla_l\nabla^l\Omega/\Omega$, $\nabla_l\Omega\nabla^l\Omega/\Omega^2$). A meshed region is a region where the approximation break down.

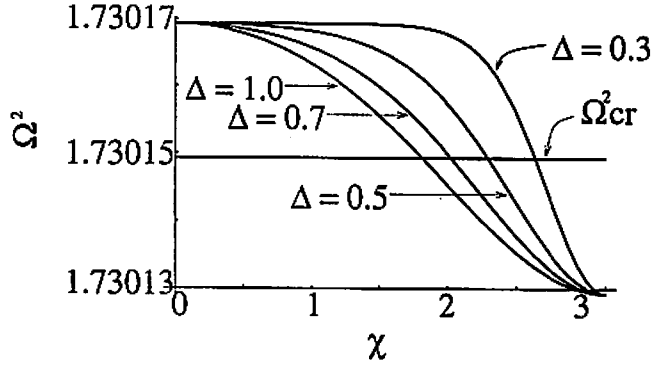


Figure 3: The distribution of Ω^2 when we vary Δ for the special case: $\Omega_0^2 = 1.73015$, $\delta\Omega = 0.000023$.

Density perturbation generated by primordial magnetic fields

Kazumi Tamai
E-mail address: kazumi@th.phys.titech.ac.jp,
Akio Hosoya
E-mail:ahosoya@th.phys.titech.ac.jp,
Department of Physics, Tokyo Institute of Technology, Oh-Okayama Meguro-ku, Tokyo 152, Japan

March 1, 1996

Abstract

Recently a mechanism generating the primordial magnetic field at the recombination time is considered. If such a magnetic field existed, protons entering into it would be trapped and as a result a density perturbation would be generated. It is assumed that the amplitude of the density perturbation is 10^{-3}. Then the scale of the density perturbation is that of the quasar if the present baryon density parameter is 0.01 and is that of the galaxy if it is 0.1.

1 Introduction

It is well known from the observation that the amount of the dark matter is 10 times as large as that of the luminous matter. However, it is less known what is the dark matter. The conventional theory of the structure formation based on the gravitational instability needs the cold dark matter [1]. But it has not yet been experimentally verified that the cold dark matter really exists. So it will be meaningful to consider any alternative theory of the structure formation without the cold dark matter.

Many years ago, Wasserman suggested that if the primordial magnetic field had existed, they might generate a density perturbation at the recombination time of the Universe since they became a gravitational source [2, 3, 4]. He assumed that the density perturbation was roughly 10^{-3} at the recombination time because it became 10^3 times larger and nonlinear until the present time. The correlation scale of the primordial magnetic field was assumed to be $10^{22}cm$ because the mass scale was that of the ordinary galaxy, 10^{11} solar mass. He concluded that the primordial magnetic field was 10^{-3} Gauss at the recombination time. But the difficulty is that the corresponding magnetic jeans length is $10^{23}cm$ and the magnetic field becomes unstable.

In this paper, we propose a simple mechanism which more efficiently generates the density perturbation of $\sim 10^{-3}$ by the primordial magnetic field. The basic idea is the following. Protons in the magnetic field do a synchrotron motion and are trapped inside the Larmor radius. So if there was the primordial magnetic field at the recombination time with a relatively large correlation scale in which the direction of the magnetic field was more or less aligned, protons entering the correlation region would do a synchrotron motion. Then they can move only along the magnetic field line. If the Larmor radius was smaller than the correlation scale,

protons would be trapped in the correlation region. There is a similar phenomenon to this mechanism. It is well known that charged particles coming from the Sun are trapped by the magnetic field of the Earth. The above mechanism would generate a density perturbation.

A mechanism generating the aligned magnetic field at the recombination time is considered by A. Hosoya and S. Kobayasi [5]. However, in the present paper we do not need to specify the creation mechanism of the magnetic field.

From the condition that the Larmor radius is smaller than the correlation scale and protons are trapped, the lower limit of the value of the magnetic field can be estimated by the correlation scale. If it is assumed that the density perturbation is 10^{-3} and the baryon density parameter is 0.1 or equivalently the number density of the protons at the recombination time is 10^3, the upper limit of the value of the correlation scale can be estimated to be $3 \times 10^{21} cm$ and the mass scale is 3×10^{11} solar mass which is that of the galaxy. Then the corresponding lower limit of the magnetic field is $3 \times 10^{-20} gauss$. If the baryon density parameter is 0.01 or the number density of the protons is $10^2 cm^{-3}$, the upper limit of the correlation scale is $3 \times 10^{20} cm$ and the mass scale is 3×10^8 solar mass which is that of the quasar(QSO) or the active galactic nuclei(AGN). Then the lower limit of the magnetic field is $3 \times 10^{-19} Gauss$. The correlation scale must be larger than the magnetic jeans length because the scale smaller than the magnetic jeans length is unstable. From that condition, the upper limit of the magnetic field is determined. It is concluded from the above results that our mechanism requires the magnetic field smaller than that of Wasserman's mechanism and whose scale is smaller than that observed by the COBE.

Let us examine a possible mechanism for protons to escape from the magnetic field by which the density perturbation decreases. If the Larmor radius was smaller than the correlation scale, but larger than the mean free path of a proton, the diffusion of protons caused by the Coulomb scattering had to be considered. The diffusion speed of a proton perpendicular to the magnetic field is proportional to the gradient of the pressure or the density of charged particles(protons and electrons) and inversely proportional to the correlation scale. So the effect of the diffusion is very small because the density perturbation considered here is 10^{-3} at most and the correlation scale is relatively large. Moreover the larger the density perturbation is, the larger the effect of gravity is. So the effect of the diffusion can be negligible. Neutral hydrogens are not trapped in the magnetic field. So the density perturbation may decrease as the increased number of neutral hydrogens escape from the magnetic field. The net number of hydrogens escaping from the correlation region per unit time is proportional to the density perturbation and the density of hydrogens. As the Universe becomes neutral and the density perturbation becomes large, that number may become large. But then the effect of the gravitation would become large and hydrogens could not escape from the correlation region. The effect of the escaping hydrogens turns out to be negligible as we show later.

Finally, we consider the evolution of the magnetic field. The magnetic field around the density perturbation would be enhanced after the turn-around epoch by a factor 10^{14} at a redshift $z \sim 2$ if the density parameter was 0.1 and the redshift of the galaxy formation was 4 [6]. Because it seems that the magnetic field is not so enhanced until the turn-around epoch, the magnetic field of $10^{-20} Gauss$ at the recombination would be about $10^{-6} Gauss$ at $z \sim 2$. From the observation, the inter galactic magnetic field is $10^{-6} Gauss$ at $z \sim 2$ [7]. This is consistent with the result expected from our mechanism.

2 Density perturbation generated by the primordial magnetic field

It is assumed that the primordial magnetic field, B had the correlation scale, ξ and was aligned at the recombination time. Protons would enter the magnetic field by the analogous way to the charged particles trapped by the magnetic field of the Earth.

particles from the Sun can enter into the closed magnetic field along the opened Protons coming along the magnetic field will make a helical motion with the Larmor radius r_s, of the synchrotron motion of protons (Figure 1). It is given by,

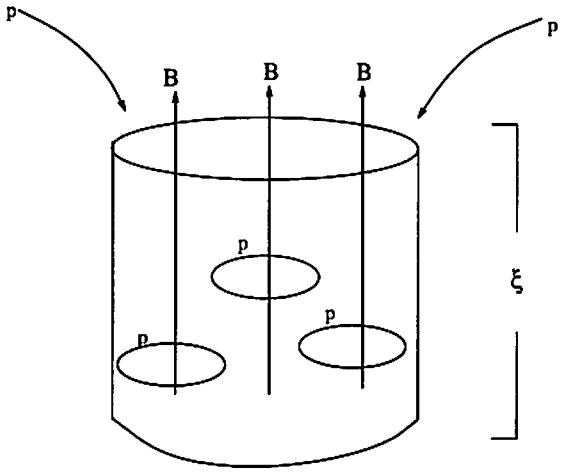


Figure 1: This figure represents protons trapped by the magnetic field.

$$\begin{aligned} r_s &= \frac{m_p v_p}{eB} \\ v_p &= \sqrt{3T_{rec}m_p} \sim \frac{c}{3} \times 10^{-4} \\ T_{rec} &= 4000K \end{aligned}$$

where e is the electric charge of a proton and m_p, v_p is the mass and the thermal velocity of a proton respectively. From the above equations, we obtain

$$r_s \sim 10^2 (1Gauss/B)cm \tag{1}$$

If $r_s < \xi$, protons would be trapped by the magnetic field inside the correlation region.

The number density of protons, n_p has not be determined yet because it is not known what is the dark matter. If the dark matter was all baryonic, the number density of protons at the recombination time, n_p^{rec} was $10^3 cm^{-3}$. If it was all nonbaryonic, n_p^{rec} was $10^2 cm^{-3}$.

The number of free protons decreases rapidly after the recombination time. When it is assumed from the Saha thermal ionization equilibrium equation that the fractional ionization, χ, decreases with the power of the redshift, z while it decreases rapidly and $\chi = 10^{-3}\left(\frac{n_p^{rec}}{10^2 cm^{-3}}\right)$ at $1+z = 1000$ [8], it may be given as below.

$$\chi = \left(\frac{1+z}{1400}\right)^{\alpha} \tag{2}$$

$$\alpha = \left(-3+\log\frac{n_p^{rec}}{10^2cm^{-3}}\right)/\log\frac{5}{7}$$
$$\sim 6.85\left(3-\log\frac{n_p^{rec}}{10^2cm^{-3}}\right)$$

The flux of protons is isotropic. So the number of protons per unit time which enter the correlation region is given by,

$$\frac{dN_p}{dt} = 4\pi\xi^2 n_p(t)v_p \quad (3)$$
$$n_p = \chi n_p^{rec}$$

The effect of the expansion of the Universe is negligible because the number of protons sufficiently decreases before that effect is considerable.

Using the relation, $t = t_0(1+z)^{-3/2}$ with $t_0 \sim 1.5\times10^{10}$ year, being the present age of the Universe, we can estimate the number of protons accumulated in the correlation region before χ becomes much smaller than 1 as,

$$N_p \sim \xi^2\left(\frac{n_p^{rec}}{10^2cm^{-3}}\right)\times10^{21} \quad (4)$$

The resultant density perturbation is then given by

$$\delta = \frac{N_p}{V}/10^3$$
$$\sim 3\xi^{-1}\left(\frac{n_p^{rec}}{10^2cm^{-3}}\right)\times10^{17} \quad (5)$$

where $V=\frac{4\pi}{3}\xi^3$, is the volume of the correlation region.

If δ is assumed to be more than 10^{-3} as we explained before, the upper limit of the scale of ξ is determined as,

$$\xi \le 3\left(\frac{n_p^{rec}}{10^2cm^{-3}}\right)\times10^{20}cm \quad (6)$$

Then the lower limit of the magnetic field is determined by the trapping condition for protons that r_s (1) is smaller than ξ as,

$$B \ge \frac{1}{3}\left(\frac{10^2cm^{-3}}{n^{rec}}\right)\times10^{-18}Gauss \quad (7)$$

The total mass density of the Universe at the recombination is about $10^3m_p(cm^{-3})$ whatever the dark matter is. So the mass scale of the density perturbation is,

$$M = 10^3m_p(cm^{-3})V$$
$$\sim 3\times10^8\left(\frac{\xi}{3\times10^{20}cm}\right)^3M_s$$
$$\le 3\times10^8\left(\frac{n_p^{rec}}{10^2cm^{-3}}\right)^3M_s \quad (8)$$

where M_s is the solar mass.

If the dark matter was all baryonic ($n_p^{rec}=10^3cm^{-3}$), the mass scale was $3\times10^{11}M_s$ and the density perturbation would become the size of the galaxy. If it was all nonbaryonic, the mass scale was 3×10^8M_s and it would become that of QSO or AGN.

The magnetic jeans length at the recombination, λ_B is given as,

$$\lambda_B = \frac{cB}{\rho_{tot}\sqrt{G}}$$

where ρ_{tot} is the total mean density of the Universe at the recombination time and about $10^{-21} gcm^{-3}$.

The correlation scale must be larger than the magnetic jeans scale. If it is not so, the density perturbation becomes unstable. So the upper limit of the magnetic field is determined as,

$$\begin{aligned} B &\leq \frac{\rho_{tot}\sqrt{G}\xi}{c} \\ &\sim 3\times 10^{-5}\Big(\frac{\xi}{3\times 10^{20}}\Big)\Big(\frac{10^2}{n_p^{rec}}\Big)\ (Gauss) \end{aligned} \tag{9}$$

3 Diffusion of protons

If r_s was smaller than ξ but larger than the mean free path of a proton, l_c, the diffusion of protons had to be considered because protons might escape from the region of the magnetic field and δ might vanish. The cross section of the Coulomb scattering for a proton with velocity v_p is,

$$\sigma_c \sim \pi \frac{e^4}{T_{rec}^2} \ln\lambda$$

where $\ln\lambda \sim 20$ is the Coulomb logarithm.

So the mean free path of a proton at the recombination time is,

$$\begin{aligned} l_c &= \frac{1}{n_p^{rec}\sigma_c} \\ &\sim \Big(\frac{10^2 cm^{-3}}{n_p^{rec}}\Big)\times 10^8 cm \end{aligned} \tag{10}$$

The condition that r_s (1)is smaller than l_c implies $B \geq (\frac{n_p^{rec}}{10^2 cm^{-3}})\times 10^{-6} Gauss$. This value is comparable to the upper limit determined by the magnetic jeans length. So one might worry if there exists an allowed region for the magnetic field. If B does not satisfy the above condition, the diffusion of protons must be considered. However, if we can show the diffusion length is much smaller than ξ, we can conclude the effect of the diffusion is negligible. So we will estimate the diffusion length of a proton and show that the diffusion is indeed negligible.

The collision frequency of an electron scattered by a proton is given by,

$$\nu_{ep} = \frac{v_p}{l_c}$$

When the plasma is quasi neutral and in equilibrium, the magnetic hydrodynamics equation of the plasma is,

$$\vec{\mathrm{j}} \times \vec{\mathrm{B}} - \vec{\nabla} P = 0 \tag{11}$$

$$P \equiv P_p + P_e$$

$$\vec{\mathrm{E}} + \vec{\mathrm{v}} \times \vec{\mathrm{B}} = \frac{m_e \nu_{ep}}{ne^2}\vec{\mathrm{j}} \tag{12}$$

where $\vec{\mathrm{E}}$ is the electric field and P_p, P_e is the pressure of protons and electrons respectively and n is the number density of protons or electrons and v is the velocity of the center of mass and $\vec{\mathrm{j}}$ is the current density. We can regard v as v_p because

proton mass is much larger than electron mass.

The equation of the state of a proton and an electron is,

$$\begin{aligned} P_p &= n_p T \\ P_e &= n_e T \end{aligned} \tag{13}$$

where T is the temperature.

If we use the cylindrical coordinate and direct $\vec{E}$ and $\vec{\nabla} P$ to the direction of r and $\vec{B}$ to that of z, from the above equations, we obtain the diffusion speed of a proton perpendicular to the magnetic field as,

$$\begin{aligned} u_\perp &= -\frac{m_e \nu_{ep} T_{rec}}{e^2 B^2 n_p} \frac{\partial (n_p + n_e)}{\partial r} \\ &\sim 3 \times 10^2 \Big(\frac{\delta}{10^{-3}}\Big)\Big(\frac{n_p^{rec}}{10^2 cm^{-3}}\Big)\Big(\frac{3 \times 10^{20} cm}{\xi}\Big)\Big(\frac{3 \times 10^{-19} Gauss}{B}\Big)^2 (cm/s) \\ &\sim 3 \times 10^2 \Big(\frac{n_p^{rec}}{10^2 cm^{-3}}\Big)^4 \Big(\frac{\delta}{10^{-3}}\Big)(cm/s) \end{aligned} \tag{14}$$

The diffusion may occur until the ionization fraction is not so small. The interval from $z = 1400$ to $z = 1000$ is,

$$\begin{aligned} c\delta t &= c t_0 (1000^{-1.5} - 1400^{-1.5}) \\ &\sim 10^{23} (cm) \end{aligned}$$

So the diffusion length of a proton until $z = 1000$ is,

$$\begin{aligned} l_d &= \delta t u_\perp \\ &\sim 10^{15} (\frac{\delta}{10^{-3}})(\frac{n_p^{rec}}{10^2 cm^{-3}})^4 (cm) \end{aligned}$$

In this paper we estimate the diffusion length under the condition that the ionization is 1 for simplicity until z = 1000. So the real value of the diffusion length may be smaller than the above value.

It is concluded that the diffusion of protons is negligible because $\xi \gg l_d$.

4 Escape of hydrogens

By the above discussions, protons are trapped inside the correlation region where the magnetic field exists and the diffusion of protons is negligible. But because neutral hydrogens are not trapped by the magnetic field, the density perturbation may decrease as the number of hydrogens increases. If the number density of hydrogens in the density perturbation is a little larger than the mean number density of hydrogens, the net number of the escaping hydrogens per unit time, $\frac{dN_H^{out}}{dt}$ is,

$$\begin{aligned} \frac{dN_H^{out}}{dt} &= 4\pi \xi^2 n_H \delta_H v_H \\ &= 4\pi \xi^2 n_p^{rec} \{1 - \chi\} \delta_H v_H \end{aligned}$$

where $v_H, n_H and \delta_H$ are the velocity of a hydrogen and the mean number density and the density perturbation of the hydrogens, respectively.

It is reasonable to assume that the velocity of a hydrogen is equal to that of a proton and the density perturbation of hydrogens is below 10^{-3} and the ionization

fraction is over 10^{-3}. Then the ratio of the net number of the escaping hydrogens per unit time to that of the incoming protons (3) is given by,

$$\frac{dN_H^{out}}{dt}/\frac{dN_p}{dt} = \frac{1-\chi}{\chi}\delta_H \leq 1$$

The real value of this ratio would be much smaller than the above value because as the density perturbation became large, the effect of the gravitational force was large and hydrogens could not escape from the correlation region. So the effect of escaping hydrogen is negligible.

5 Conclusion

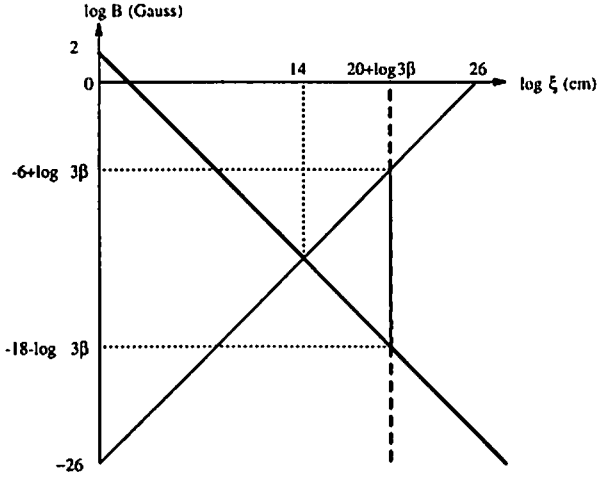


Figure 2: The bold solid line, the thin solid line and the dashed line represent $\xi = r_s, \xi = \lambda_B, \delta = 10^{-3}$ respectively and $\beta \equiv \frac{n_p^{rec}}{10^3 cm^{-3}}$. The structure may be formed by the primordial magnetic field with the value of B and ξ inside the dotted area.

Recently a theory generating the magnetic field at the recombination time is considered ([5]). We suggest that the primordial magnetic field may play a role in the structure formation. Protons entering into the magnetic field are trapped inside the Larmor radius. If the magnetic field existed at the recombination time, it would gather protons by the above mechanism and generate a density perturbation. The theory of the structure formation based on this mechanism does not need the cold dark matter as the one based on the gravitational instability theory and is more efficient than the scenario suggested by Wasserman.

We conclude that the primordial magnetic field indicated inside the dotted area in fig 2 would generate the density perturbation of more than 10^{-3} at the recombination time. Then the required magnetic field ranges from $3 \times 10^{-19}(\frac{10^3 cm^{-3}}{n_p^{rec}})$ to $3 \times 10^{-5}(\frac{n_p^{rec}}{10^3 cm^{-3}})$ Gauss in our scenario. The upper limit of the correlation scale of the magnetic field or the scale of the density perturbation is $3 \times 10^{20}(\frac{n_p^{rec}}{10^3 cm^{-3}})$ cm and then the mass scale is $3 \times 10^8(\frac{n_p^{rec}}{10^3 cm^{-3}})^3$ solar mass. So if the number density

of the protons at the recombination time was $10^2 cm^{-3}$, AGN or QSO would be generated. If it was $10^3 cm^{-3}$, the galaxy would be generated.

The effect of the diffusion of protons and the escaping neutral hydrogens is negligible.

The mechanism which generates the magnetic field at the recombination time suggests that the correlation scale would be over 10^{21} cm [5]. So it is consistent with our result.

References

[1] Y. Suto, N. Gouda and N. Sugiyama, Ap J Supplement. 74 (1990)

[2] I. Wasserman, Ap J. 224, 337 (1978)

[3] P.J.E. Peebles, The Large-Scale Structure of the Universe. Princeton University press (1980)

[4] E. Kim, A. Olinto and R. Rosner, astro-ph/9412070

[5] A. Hosoya and S. Kobayasi, gr-qc/9509001

[6] H. Lesch and M. Chiba, Astron. Astrophys.297, 305 (1995)

[7] E. Asseo and H. Sol, Phys Rep. 148, No6, 307 (1987)

[8] P.J.E. Peebles, Principles of Physical Cosmology. Princeton University press (1993)

Anomalies in Gravitational Systems and Weak Expansion

Noriaki IKEDA[1]
Research Institute for Mathematical Sciences
Kyoto University, Kyoto 606-01, Japan

1 Introduction

The perturbation calculation in curved spacetime is usually very complicated. Besides the theoretical difficulties like non-renormalizability of quantum gravity, it seems that technical difficulties also disturb further analysis of the quantum effects in curved spacetime.

In order to reconsider the quantum field theories in curved spacetime. we focus the anomalies, which are one of the direct quantum effects and are analyzed by many authors[1]. This talk is based on the paper[2].

Fujikawa's general standpoint that the anomalies come from the path-integral measure is taken. Then a general approach to anomaly in quantum field theory is newly formulated by use of the propagator theory in solving the heat-kernel equation. We regard the heat-kernel as a sort of the point-splitting regularization in the space(-time) manifold. We obtain some useful formulae which are valid for general anomaly calculation.

In the present talk, special emphasis is on the Weyl anomaly which is different from other anomalies in the following points. Weyl anomaly appears in general theories because it is directly related to the trace part of the energy-momentum tensor. It is essentially given by the β-function, which determines the scaling properties of a theory. It represents that Weyl anomaly has dynamical degree of freedom. Weyl anomaly, at present, does not seem to be understood only by the global geometrical (topological) analysis. This is related to the above fact. Therefore we can expect the Weyl anomaly contains richer dynamical information than other anomalies.

Motivated by this expectation, we newly formulate the general anomaly problem. It is based on (the coordinate version of) the *propagator approach* in the ordinary perturbative field theory[3]. All explanation is done by the familiar field theory language. We take the heat-kernel regularization for the ultraviolet divergences. General formulae of anomalies are obtained. Many applications are presented. Weyl anomalies, chiral anomalies(in the flat and curved space), local Lorentz anomaly and gravitational anomaly are explicitly derived by the paper[2]. Practical usefulness is stressed.

[1]This work is a collaboration with Shoichi ICHINOSE. address: Department of Physics, University of Shizuoka, Yada 52-1, Shizuoka 422, Japan

Generally anomaly terms (especially Weyl anomaly terms), including their coefficients, are fixed by calculating (1-loop) quantum fluctuation in the perturbative expansion. The explicit calculation becomes more and more complicated as we increase the space(-time) dimension. In n-dim gravitational theories, we must treat complicated higher-rank global SO(n) covariants and invariants such as $\partial_\mu\partial_\nu h_{\alpha\beta}$, $\partial_\mu\partial_\nu h_{\lambda\sigma}\cdot\partial_\mu\partial_\nu h_{\lambda\sigma}$, etc. in the weak-gravity expansion: $g_{\alpha\beta}=\delta_{\alpha\beta}+h_{\alpha\beta}$, $|h_{\alpha\beta}|\ll 1$, $(\alpha,\beta,\cdots=1,2,\cdots,n)$. In order to get rid of the obstacle, we introduce a graphical representation to treat those terms systematically. The detail of this representation appears in [2]. The representation makes it so easy to list up all independent terms. Although the present paper deals with $n=4$ cases only, in order to clearly show the usefulness, it is applicable to higher-dimensional cases. The application to six dimensional Weyl anomaly is discussed by a next paper[4].

The anomaly terms are often compared with the counter-terms. The former is independent of gauge whereas the latter is not. The former is related with the consistency of the theory whereas the latter is related with the renormalization of the theory. Although both come from the ultra-violet divergences, their most appropriate regularizations are different: the anomaly terms are commonly calculated by use of Pauli-Villars regularization (or its variants) whereas the counter-terms are usually calculated by use of the dimensional regularization. We will obtain some direct connection between the two quantities through the anomaly formula calculated by the heat-kernel regularization. The formula is very powerful as in the case of the counter-term formula by 'tHooft[5]. It is demonstrated that various anomalies of various theories are derived by the formulae.

2 New Formulation of Anomaly and Heat-Kernel Regularization

Let us explain the present formulation of anomalies taking a simple example : Weyl anomaly in n-dim Euclidean gravity-scalar coupled system.

$$\mathcal{L}[g_{\mu\nu},\phi]=\sqrt{g}(\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi+\frac{1}{2}qR\phi^2)\quad,$$
$$q=-\frac{n-2}{4(n-1)}\quad, \tag{1}$$

where $g_{\mu\nu}$ and ϕ are the metric field and the scalar field. This Lagrangian is invariant under the local Weyl transformation:

$$g^{\mu\nu}(x)'=e^{2\alpha(x)}g^{\mu\nu}(x)\quad,\quad\phi(x)'=e^{\frac{n-2}{2}\alpha(x)}\phi(x)\quad,\quad\tilde{\phi}(x)'=e^{-\alpha(x)}\tilde{\phi}(x)\quad, \tag{2}$$

where $\alpha(x)$ is the parameter of the local Weyl transformation, $g=\det g_{\mu\nu}$, and we introduce $\tilde{\phi}\equiv\sqrt[4]{g}\,\phi$ for the measure $\mathcal{D}\tilde{\phi}$ to be general coordinate invariant [6]. The partition function ,on the external gravitational field $g_{\mu\nu}$, is given by

$$Z[g_{\mu\nu}]=\int\mathcal{D}\tilde{\phi}\ \exp\{-S[g_{\mu\nu},\phi]\}\quad,\quad S[g_{\mu\nu},\phi]=\int d^nx\mathcal{L}[g_{\mu\nu},\phi]\quad. \tag{3}$$

We see the variation of $Z[g_{\mu\nu}]$ comes from the Jacobian factor in the Weyl transformation of integration measure [7]. The Jacobian is formally written as

$$\begin{aligned} J &\equiv \det \frac{\partial \tilde{\phi}'(y)}{\partial \tilde{\phi}(x)} = \det\ (e^{-\alpha(x)}\delta^n(x-y)\ ) \\ &= \exp(-\mathrm{Tr}\ [\alpha(x)\delta^n(x-y)] + O(\alpha^2)) \quad . \end{aligned} \tag{4}$$

In order to regularize the delta function $\delta^n(x-y)$, we introduce the following quantity.

$$\begin{aligned} &G(x,y;t) \equiv < x|e^{-t\vec{D}}|y > \quad , \quad t>0 \quad , \\ &\vec{D}_x \equiv \sqrt[4]{g}(-\nabla_x^2 + qR(x))\frac{1}{\sqrt[4]{g}} \quad , \end{aligned} \tag{5}$$

where t will be regarded as a regularization parameter and is called Schwinger's *proper time* [8]. The operator $\vec{D}_x$ is the hermitian differential (energy) operator which appears in the field equation for $\tilde{\phi}$. $G(x,y;t)$ satisfies the following differential equation.

$$(\frac{\partial}{\partial t} + \vec{D}_x)G(x,y;t) = 0 \quad , \quad t>0 \quad . \tag{6}$$

We solve this equation with the following initial condition.

$$\lim_{t\to+0}\ G(x,y;t) = \delta^n(x-y) \quad . \tag{7}$$

This equation expresses the regularization of the delta function $\delta^n(x-y)$ where the proper time t plays the role of the regularization parameter [9, 10]. For later general use, we write (6) and (7) in a more general form as,

$$\begin{aligned} &(\frac{\partial}{\partial t}\delta^{ij} + \vec{D}_x^{ij})G^{jk}(x,y;t) = 0 \quad , \quad t>0 \quad . \\ &\lim_{t\to+0}\ G^{ij}(x,y;t) = \delta^n(x-y)\delta^{ij} \quad . \quad i,j = 1,2,\cdots N \quad . \end{aligned} \tag{8}$$

where i and j are the field suffixes , such as a fermion suffix and a vector suffix. In the present example $N=1$.

For a general theory with the derivative couplings up to the second order, the operator $\vec{D}_x^{ij}$ can be always written as

$$\begin{aligned} &\vec{D}_x^{ij} = -\delta_{\mu\nu}\delta^{ij}\partial_\mu\partial_\nu - \vec{V}^{ij}(x) \quad , \\ &\vec{V}^{ij}(x) \equiv W_{\mu\nu}^{ij}(x)\partial_\mu\partial_\nu + N_\mu^{ij}\partial_\mu + M^{ij} \quad , \end{aligned} \tag{9}$$

where $W_{\mu\nu}^{ij}$, N_μ^{ij} and M^{ij} are external fields (background coefficient fields). In the present example, the above quantities are explicitly written as

$$\begin{aligned} &\vec{V}(x) = \sqrt[4]{g}(\nabla^\mu\nabla_\mu - qR)\frac{1}{\sqrt[4]{g}} - \delta_{\mu\nu}\partial_\mu\partial_\nu \quad , \\ &W_{\mu\nu} = g^{\mu\nu} - \delta_{\mu\nu} = -h_{\mu\nu} + h_{\mu\lambda}h_{\lambda\nu} + O(h^3) \quad , \\ &N_\lambda = -g^{\mu\nu}\Gamma_{\mu\nu}^\lambda - g^{\lambda\mu}\Gamma_{\mu\nu}^\nu = -\partial_\mu h_{\lambda\mu} + O(h^2) \quad , \\ &M = -qR + \frac{1}{4}g^{\mu\nu}\{\Gamma_{\mu\lambda}^\lambda\Gamma_{\nu\sigma}^\sigma + 2\Gamma_{\mu\nu}^\lambda\Gamma_{\lambda\sigma}^\sigma - 2\partial_\nu\Gamma_{\mu\lambda}^\lambda\} \\ &\quad = -q(\partial^2 h - \partial_\alpha\partial_\beta h_{\alpha\beta}) - \frac{1}{4}\partial^2 h + O(h^2) \quad , \end{aligned} \tag{10}$$

where $g_{\mu\nu} = \delta_{\mu\nu} + h_{\mu\nu}$, $h \equiv h_{\mu\mu}$. The usage of general coefficients $W^{ij}_{\mu\nu}, N^{ij}_{\mu}$ and M^{ij} ,instead of their concrete contents, makes it possible to obtain a general formula for anomalies.Let us solve the differential equation (8) perturbatively for the case of weak external fields $(W^{ij}_{\mu\nu}, N^{ij}_{\mu}, M^{ij})$. (In the present example, this corresponds to the perturbation around the *flat* space.) The differential equation (8) becomes

$$
\begin{aligned}
&(\frac{\partial}{\partial t} - \partial^2)G^{ij}(x,y;t) = \vec{V}^{ik}(x)G^{kj}(x,y;t) \quad , \\
&\partial^2 \equiv \delta_{\mu\nu}\partial_\mu\partial_\nu
\end{aligned}
\tag{11}
$$

In the following we suppress the field suffixes $i, j, \cdots$ and take the matrix notation. This equation (11) is the n-dim heat equation with the small perturbation $\vec{V}$. We prepare two quantities in order to obtain the solution.

i) Heat Equation

The heat equation:

$$
(\frac{\partial}{\partial t} - \partial^2)G_0(x,y;t) = 0 \quad , \quad t > 0 \tag{12}
$$

has the solution

$$
\begin{aligned}
G_0(x,y;t) = G_0(x-y;t) &= \int \frac{d^n k}{(2\pi)^n} \exp\{-k^2 t + ik^\mu (x-y)^\mu\}\ I_N \\
&= \frac{1}{(4\pi t)^{n/2}} e^{-\frac{(x-y)^2}{4t}}\ I_N \quad , \quad k^2 \equiv \sum_{\mu=1}^{n} (k^\mu)^2 \quad .
\end{aligned}
\tag{13}
$$

where I_N is the identity matrix of the size $N \times N$. G_0 satisfies the initial condition: $\lim_{t \to +0}\ G_0(x-y;t) = \delta^n(x-y)\ I_N$. We define

$$
G_0(x,y;t) = 0 \text{ for } t \leq 0 \quad . \tag{14}
$$

ii) Heat Propagator

The heat equation with the delta-function source defines the heat propagator.

$$
\begin{aligned}
&(\frac{\partial}{\partial t} - \partial^2)S(x,y;t-s) = \delta(t-s)\delta^n(x-y)\ I_N \quad , \\
&S(x,y;t) = S(x-y;t) = \theta(t)G_0(x-y;t) \quad , \\
&\qquad k^2 \equiv \sum_{\nu=1}^{\mu} k^\mu k^\mu \quad , \quad k \cdot x \equiv \sum_{\mu=1}^{n} k^\mu x^\mu \quad .
\end{aligned}
\tag{15}
$$

$\theta(t)$ is the *step function* defined by : $\theta(t) = 1$ for $t > 0$, $\theta(t) = 0$ for $t < 0$. $S(x-y;t)$ satisfies the initial condition: $\lim_{t \to +0} S(x-y;t) = \delta^n(x-y)\ I_N$ and $S(x,y;t) = 0$ for $t \leq 0$.

Now the formal solution of (11) with the initial condition (8) is given by

$$
G(x,y;t) = G_0(x-y;t) + \int d^n z \int_{-\infty}^{\infty} ds\ S(x-z;t-s)\vec{V}(z)G(z,y;s) \quad . \tag{16}
$$

$G(x,y;t)$ appears in both sides above. We can iteratively solve (16) as

$$\begin{aligned}
G(x,y;t) &= G_0(x-y;t) + \int S\vec{V}G_0 + \int S\vec{V}\int S\vec{V}G_0 + \cdots \quad , \\
G_1(x,y;t) &\equiv \int S\vec{V}G_0 = \int d^n z ds S(x-z;t-s)\vec{V}(z)G_0(z-y;s) \\
&= \int d^n z \int_0^t ds G_0(x-z;t-s)\vec{V}(z)G_0(z-y;s) \quad , \qquad (17) \\
G_2(x,y;t) &\equiv \int S\vec{V}\int S\vec{V}G_0 = \int d^n z' ds' S(x-z';t-s')\vec{V}(z') \\
&\quad \times \int d^n z ds S(z'-z;s'-s)\vec{V}(z)G_0(z-y;s) \\
&= \int d^n z' \int_0^t ds' G_0(x-z';t-s')\vec{V}(z') \\
&\quad \times \int d^n z \int_0^{s'} ds G_0(z'-z;s'-s)\vec{V}(z)G_0(z-y;s) \quad .
\end{aligned}$$

Higher-order terms are similarly obtained. Generally, in n-dim, the terms up to $G_{n/2}$ are practically sufficient for the anomaly calculation. The trace-part of (4) is given by putting $x = y$ in the above equations. where we introduce some scaled integration variables which are dimension-less: $r = \frac{s}{t}$, $w^\mu = (z-x)^\mu/\sqrt{t}$. Furthermore Further analysis will be done for each dimension.

3 Anomaly Formula in 4 Dimension

In this section we will obtain a formula for anomalies in 4 dim.

3.1 $G_0(0;t), G_1(x,x;t)$

From (13) with $n = 4$, we obtain

$$G_0(0;t) = \frac{1}{(4\pi t)^2} I_N \quad . \qquad (18)$$

From (17),

$$G_1(x,x;t) = \frac{1}{(4\pi)^4 t}\int d^4 w \int_0^1 dr \frac{1}{\{(1-r)r\}^2} e^{-\frac{w^2}{4(1-r)}} \vec{V}(x+\sqrt{t}w) e^{-\frac{w^2}{4r}} \quad . \qquad (19)$$

We notice t^0-terms only contribute to the anomalies. They correspond to log-divergent terms in the effective action: $\Gamma = \int_0^\infty \frac{dt}{t} G(x,x;t)$. (t^{-m}-terms ($m = 1,2,\cdots$) contribute to power-divergences, t^{+m}-terms vanish as $t \to +0$.) External fields, $W_{\mu\nu}, N_\mu$ and M in $\vec{V}(x+\sqrt{t}w)$ are expanded around $t = 0$ as,

$$\begin{aligned}
W_{\mu\nu}(x+\sqrt{t}w) &= W_{\mu\nu}(x) + \sqrt{t}w^\alpha \partial_\alpha W_{\mu\nu}(x) + \frac{t}{2} w^\alpha w^\beta \partial_\alpha \partial_\beta W_{\mu\nu}(x) + \cdots \\
N_\mu(x+\sqrt{t}w) &= N_\mu(x) + \sqrt{t}w^\alpha \partial_\alpha N_\mu(x) + \frac{t}{2} w^\alpha w^\beta \partial_\alpha \partial_\beta N_\mu(x) + \cdots \qquad (20) \\
M(x+\sqrt{t}w) &= M(x) + \sqrt{t}w^\alpha \partial_\alpha M(x) + \frac{t}{2} w^\alpha w^\beta \partial_\alpha \partial_\beta M(x) + \cdots \quad .
\end{aligned}$$

Then we can pick up t^0-part of (19) as follows.

$$
\begin{aligned}
G_1(x,x;t)|_{t^0} &= \frac{1}{(4\pi)^4}\int d^4w \int_0^1 dr \frac{1}{\{(1-r)r\}^2} e^{-\frac{w^2}{4r(1-r)}} \\
&\quad \times\{\ \frac{1}{4!}w^\lambda w^\sigma w^\tau w^\omega \partial_\lambda\partial_\sigma\partial_\tau\partial_\omega W_{\mu\nu}(x)(-\frac{\delta_{\mu\nu}}{2r}+\frac{w^\mu w^\nu}{4r^2}) \\
&\quad +\frac{1}{3!}w^\lambda w^\sigma w^\tau \partial_\lambda\partial_\sigma\partial_\tau N_\mu(x)(-\frac{w^\mu}{2r})+\frac{1}{2}w^\lambda w^\sigma\partial_\lambda\partial_\sigma M(x)\ \} \quad (21) \\
&= \frac{1}{(4\pi)^2 4!}\{\ -\frac{1}{5}\partial^2\partial^2 W_{\mu\mu}(x)+\frac{6}{5}\partial^2\partial_\mu\partial_\nu W_{\mu\nu}(x) \\
&\qquad -2\partial^2\partial_\mu N_\mu(x)+4\partial^2 M(x)\ \}\ .
\end{aligned}
$$

This is one part of the 4 dim anomaly formula. Let us derive the other part.

3.2 $G_2(x,x;t)$

From (17), we have

$$
\begin{aligned}
G_2(x,x;t) &= \frac{1}{(4\pi)^6}\int d^4v d^4u \int_0^1 dk \int_0^k dl \frac{1}{\{(1-k)(k-l)l\}^2} e^{-\frac{1}{4}(\frac{v^2}{1-k}+\frac{(v-u)^2}{k-l}+\frac{u^2}{l})} \\
&\quad \times\{\ \frac{1}{t}W_{\mu\nu}(x+\sqrt{t}v)(-\frac{\delta_{\mu\nu}}{2(k-l)}+\frac{(v-u)^\mu(v-u)^\nu}{4(k-l)^2}) \\
&\quad +\frac{1}{\sqrt{t}}N_\mu(x+\sqrt{t}v)(-\frac{(v-u)^\mu}{2(k-l)})+M(x+\sqrt{t}v)\ \} \qquad (22) \\
&\quad \times\{\ \frac{1}{t}W_{\lambda\sigma}(x+\sqrt{t}u)(-\frac{\delta_{\lambda\sigma}}{2l}+\frac{u^\lambda u^\sigma}{4l^2})+\frac{1}{\sqrt{t}}N_\lambda(x+\sqrt{t}u)(-\frac{u^\lambda}{2l})+M(x+\sqrt{t}u)\ \}\ .
\end{aligned}
$$

For the comparison with the (1-loop) counter-term formula, we first consider the case of 'flat' space: $W_{\mu\nu}=0$. In this case the t^0-part of (22) is given as

$$
\begin{aligned}
G_2^{flat}(x,x;t)|_{t^0} &= \frac{1}{(4\pi)^2}[\ -\frac{1}{24}\partial^2(N_\mu\cdot N_\mu) \\
&\quad +\frac{1}{48}(\partial_\mu N_\nu-\partial_\nu N_\mu)^2+\frac{1}{2}(M-\frac{1}{2}\partial_\mu N_\mu)^2\ ]\ . \qquad (23)
\end{aligned}
$$

Combining the above result and (21) for $W_{\mu\nu}=0$, $G^{flat}(x,x;t)|_{t^0}$ is expressed, up to the present approximation, as

$$
\begin{aligned}
G^{flat}(x,x;t)|_{t^0} &= \frac{1}{(4\pi)^2}[\ \frac{1}{6}\partial^2(M-\frac{1}{2}\partial_\mu N_\mu-\frac{1}{4}N_\mu\cdot N_\mu)\ ]+\Delta\mathcal{L}_{tH}+O((N,M)^3)\ , \\
\Delta\mathcal{L}_{tH} &\equiv \frac{1}{(4\pi)^2}[\ \frac{1}{48}(\partial_\mu N_\nu-\partial_\nu N_\mu)^2+\frac{1}{2}(M-\frac{1}{2}\partial_\mu N_\mu)^2\ ]\ , \qquad (24)
\end{aligned}
$$

where $\Delta\mathcal{L}_{tH}$ is 'tHooft's 1-loop counter-term formula [5] at the present approximation. This anomaly formula should be compared with the counter-term formula in the following points: 1) The total derivative terms have meaning in the anomaly formula; 2) Because the present approximation is weak-field expansion (of the 1-loop part) up to G_2, the cubic and quartic terms with respect to the external fields

(N_μ, M) do not appear and will appear in G_3 and G_4 ; 3) The linear terms appear as total derivatives; 4) Symmetries, with respect to interchange of suffixes, are not assumed for the external fields N_μ and M; 5) The counter-term formula is obtained using the dimensional regularization, whereas the present anomaly formula is obtained using the heat-kernel regularization. Although conformal anomalies were discussed in connection with the 1-loop counter-term formula[11], such a direct relation as above has not been known so far.

For the anomaly calculation of gravitational theories, we must consider the general case of $W_{\mu\nu}$. From the dimensional counting ($[W_{\mu\nu}]$=(Mass)^0 , $[N_\mu]$=(Mass)^1 , $[M]$=(Mass)^2 , $[\partial_\mu]$=(Mass)^1), we see the t^0-part of $G_2(x,x;t)$ has three types of terms: 1) $W \times (\partial\partial\partial W,\ \partial\partial N,\ \partial\partial M)$; 2) $(\partial W, N) \times (\partial\partial\partial W, \partial\partial N, \partial M)$; 3) $(\partial\partial W, \partial N, M) \times (\partial\partial W, \partial N, M)$. Among them, the most useful ones are type 3) terms, i.e. those terms which are composed only of (Mass)^2-dimensional quantities: $(\partial\partial W, \partial N, M)$,because they are sufficient to determine all anomaly terms[2]. Other types, 1) and 2), are also similarly evaluated, but practically they are not necessary for the anomaly calculation.

3.3 Graphical Representation of Anomaly Formula

Further evaluation of (22) is straightforward, but we need to treat many terms. Here we introduce a graphical method to express those terms. Recently it has been shown that the graphical representation is practically useful to treat invariants and covariants in general relativity[12]. Because the connectivity of suffixes is visually expressed, it is very easy to discriminate between independent terms and dependent ones. Here we apply the technique to the present case: to represent global SO(n) (n = 4 in the present case) covariants and invariants. We define the following graphical representation for $\partial_\alpha\partial_\beta W_{\mu\nu}$, $\partial_\alpha N_\mu$.

$\partial_\lambda\partial_\rho W_{\mu\nu}$ $\qquad\qquad$ $\partial_\lambda N_\mu$

Fig.1

All independent terms which could appear in (22) are graphically listed up in [2]. They are those terms which satisfy the following conditions: 1) Invariants with respect to the global $SO(n)$ (n = 4 in this section) transformation of the coordinate; 2) Dimension of (Mass)^4; 3) They are composed only of $\partial\partial W$, ∂N and M . Totally

26 terms appear. The final evaluation of (22) is given by Table 1 where the coefficients for all independent terms ,except the overall factor $1/(4\pi)^2$, are listed. The result (21) for G_1 and the result of Table 1 for G_2 constitute the anomaly formula in 4 dim.

Graph	Expression	Coeff.	Graph	Expression	Coeff.
$\bar{A}1$	$\partial_\sigma\partial_\lambda W_{\mu\nu}\cdot\partial_\sigma\partial_\nu W_{\mu\lambda}$	1/45	$\bar{E}1$	$\partial_\mu\partial_\lambda W_{\lambda\nu}\cdot\partial_\mu N_\nu$	1/12
$\bar{A}2$	$\partial_\sigma\partial_\lambda W_{\lambda\mu}\cdot\partial_\sigma\partial_\nu W_{\mu\nu}$	-2/45	$\bar{E}2$	$\partial_\mu\partial_\lambda W_{\lambda\nu}\cdot\partial_\nu N_\mu$	1/12
$\bar{A}3$	$\partial_\sigma\partial_\lambda W_{\lambda\mu}\cdot\partial_\mu\partial_\nu W_{\nu\sigma}$	-2/45	$\bar{E}3$	$\partial_\mu\partial_\nu W_{\lambda\lambda}\cdot\partial_\nu N_\mu$	0
$\bar{B}1$	$\partial_\nu\partial_\lambda W_{\sigma\sigma}\cdot\partial_\lambda\partial_\mu W_{\mu\nu}$	-1/90	$\bar{E}4$	$\partial^2 W_{\mu\nu}\cdot\partial_\nu N_\mu$	0
$\bar{B}2$	$\partial^2 W_{\lambda\nu}\cdot\partial_\lambda\partial_\mu W_{\mu\nu}$	1/180	$\bar{Q}R$	$\partial_\mu\partial_\nu W_{\mu\nu}\cdot\partial_\lambda N_\lambda$	-1/6
$\bar{B}3$	$\partial_\mu\partial_\nu W_{\lambda\sigma}\cdot\partial_\mu\partial_\nu W_{\lambda\sigma}$	1/180	$\bar{P}R$	$\partial^2 W_{\mu\mu}\cdot\partial_\nu N_\nu$	1/24
$\bar{B}4$	$\partial_\mu\partial_\nu W_{\lambda\sigma}\cdot\partial_\lambda\partial_\sigma W_{\mu\nu}$	1/180	$F1$	$\partial_\mu N_\nu\cdot\partial_\mu N_\nu$	-1/24
$\bar{Q}^2$	$(\partial_\mu\partial_\nu W_{\mu\nu})^2$	1/18	$F2$	$\partial_\mu N_\nu\cdot\partial_\nu N_\mu$	-1/24
$\bar{C}1$	$\partial_\mu\partial_\nu W_{\lambda\lambda}\cdot\partial_\mu\partial_\nu W_{\sigma\sigma}$	1/360	RR	$(\partial_\mu N_\mu)^2$	1/8
$\bar{C}2$	$\partial^2 W_{\mu\nu}\cdot\partial^2 W_{\mu\nu}$	1/144	$M\bar{P}$	$M\cdot\partial^2 W_{\mu\mu}$	-1/12
$\bar{C}3$	$\partial_\mu\partial_\nu W_{\lambda\lambda}\cdot\partial^2 W_{\mu\nu}$	-1/90	$M\bar{Q}$	$M\cdot\partial_\mu\partial_\nu W_{\mu\nu}$	1/3
$\bar{P}\bar{Q}$	$\partial^2 W_{\lambda\lambda}\cdot\partial_\mu\partial_\nu W_{\mu\nu}$	-1/36	MR	$M\cdot\partial_\mu N_\mu$	-1/2
$\bar{P}^2$	$(\partial^2 W_{\lambda\lambda})^2$	1/288	MM	$M\cdot M$	1/2

Table 1 Anomaly Formula for $(\partial\partial W,\partial N,M)^2$-part of $G_2(x,x;t)|_{t^0}$
The overall factor is $1/(4\pi)^2$. Graph names are defined in [2].

3.4 Weyl Anomaly in 4 Dimension

We apply the formula (21) and Table 1 to the Weyl anomaly calculation of the present example (1) with $n=4$. Here we introduce another graphical representation for the terms appearing in the weak-gravity expansion: $g_{\mu\nu}=\delta_{\mu\nu}+h_{\mu\nu}$. We represent $\partial_\mu\partial_\nu h_{\alpha\beta}$ as follows.

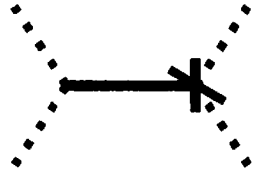


Fig.2

We focus on $(\partial\partial h)^2$-terms in the anomaly because this type terms come only from $G_1(x,x;t)|_{t^0}$ and $(\partial\partial W,\partial N,M)$-part of $G_2(x,x;t)|_{t^0}$. All possible terms that could appear in the Weyl anomaly, are graphically listed up in [2]. Totally 13 independent terms appear as listed in Table 2. Inserting the above expressions(10) into $(\partial\partial W,\partial N,M)^2$-part of $G_2(x,x;t)|_{t^0}$(Table 1) we obtain G_2-Coeff. of Table 2, where the overall factor $1/(4\pi)^2$ is omitted. Inserting h^2-part of $(W_{\mu\nu},N_\lambda,M)$,

defined in (10), into $G_1(x,x;t)|_t$(21) and picking up $(\partial\partial h)^2$-terms, we obtain G_1-Coeff. of Table 2.

Graph	Expression	G_2-Coeff.	G_1-Coeff.	G_1+G_2
$A1$	$\partial_\sigma\partial_\lambda h_{\mu\nu}\cdot\partial_\sigma\partial_\nu h_{\mu\lambda}$	$1/45$	$-7/180$	$-1/60$
$A2$	$\partial_\sigma\partial_\lambda h_{\lambda\mu}\cdot\partial_\sigma\partial_\nu h_{\mu\nu}$	$-1/45\cdot 8$	$-1/90$	$-1/72$
$A3$	$\partial_\sigma\partial_\lambda h_{\lambda\mu}\cdot\partial_\mu\partial_\nu h_{\nu\sigma}$	$-1/45\cdot 8$	0	$-1/360$
$B1$	$\partial_\nu\partial_\lambda h_{\sigma\sigma}\cdot\partial_\lambda\partial_\mu h_{\mu\nu}$	$-1/90$	$2/3\cdot 24$	$1/60$
$B2$	$\partial^2 h_{\lambda\nu}\cdot\partial_\lambda\partial_\mu h_{\mu\nu}$	$1/180$	$-4/15\cdot 24$	$-1/180$
$B3$	$\partial_\mu\partial_\nu h_{\lambda\sigma}\cdot\partial_\mu\partial_\nu h_{\lambda\sigma}$	$1/180$	$1/5\cdot 24$	$-1/72$
$B4$	$\partial_\mu\partial_\nu h_{\lambda\sigma}\cdot\partial_\lambda\partial_\sigma h_{\mu\nu}$	$1/180$	0	$1/180$
Q^2	$(\partial_\mu\partial_\nu h_{\mu\nu})^2$	0	0	0
$C1$	$\partial_\mu\partial_\nu h_{\lambda\lambda}\cdot\partial_\mu\partial_\nu h_{\sigma\sigma}$	$1/360$	$-1/144$	$-1/240$
$C2$	$\partial^2 h_{\mu\nu}\cdot\partial^2 h_{\mu\nu}$	$1/144$	$-1/15\cdot 24$	$1/240$
$C3$	$\partial_\mu\partial_\nu h_{\lambda\lambda}\cdot\partial^2 h_{\mu\nu}$	$-1/90$	$1/3\cdot 24$	$1/360$
PQ	$\partial^2 h_{\lambda\lambda}\cdot\partial_\mu\partial_\nu h_{\mu\nu}$	0	0	0
P^2	$(\partial^2 h_{\lambda\lambda})^2$	0	0	0

Table 2 Weyl Anomaly of 4 Dim Gravity-Scalar Theory: $(\partial\partial h)^2$-part
The overall factor is $1/(4\pi)^2$.

We consider the general properties of the solution $G(x,y;t)$. Its mass dimension is $[M]^n$ and is a general coordinate scalar density. We also consider Bianchi identities. This condition determines the anomaly terms as

$$\text{Weyl Anomaly} \quad = \frac{1}{(4\pi)^2}\sqrt{g}(\alpha_1\nabla^2 R+\beta_1 R^2+\beta_2 R_{\mu\nu}R^{\mu\nu}+\beta_3 R_{\mu\nu\lambda\sigma}R^{\mu\nu\lambda\sigma}). \quad (25)$$

in four dimension, where α_1, β_1, β_2, β_3 and γ_1 are constants to be determined.

Now we have evaluated the $(\partial\partial h)^2$-part of the Weyl anomaly completely. They are expressed by the $(\partial\partial h)^2$-part of the invariant quantities 26. The coefficients are obtained by the weak gravity expansion of the right-hand side and equating the result with that of Table 2. The result is

$$\alpha_1=-\frac{1}{180}\quad,\quad \beta_1=0\quad,\quad \beta_2=-\frac{1}{180}\quad,\quad \beta_3=\frac{1}{180}\quad. \quad (26)$$

Here we comment on the reason why the present formulae are sufficient to determine the anomaly terms completely. In $n=4$ dim case, the number of independent (general coordinate) invariants which could appear in the Weyl anomaly is 4 as shown in (26). Whereas the number of independent (global SO(4)) invariants of $(\partial\partial h)^2$ is 13. Since each term of $(\partial\partial h)^2$ gives an independent constraint, $(\partial\partial h)^2$-terms are sufficient to determine the Weyl anomaly. Generally, in n-dim, the number of independent terms which appear in the Weyl anomaly is far less than that of $(\partial\partial h)^{n/2}$-terms. Hence the formulae up to $G_{n/2}$ is sufficient. Furthermore we need not all terms contained in G_i : $i=1,2,\cdots,\frac{n}{2}$. In $n=4$ dim case, we have obtained only $(\partial\partial W,\partial N,M)^2$-type of G_2. Other types do not contribute to $(\partial\partial h)^2$-terms. This is the same situation in the general n-dim case. Although we explain for the case of the Weyl anomaly, it is valid for other anomalies.

4 Discussions

We have presented a systematic approach to general anomaly calculation. It is based on the propagator theory. Fujikawa's general standpoint about anomaly is taken. In the evaluation of the Jacobian due to the change of the path-integral measure, we take the heat-kernel regularization. The 4 dim general anomaly formulae are obtained. The known various anomalies and some relations between them are explicitly derived from the formulae. We have also presented the graphical technique for the representation of (global SO(n)) invariants and covariants appearing in the weak field expansion calculation. For the flat space(-time) case, the 4 dim anomaly formula reduces to the 'tHooft's 1-loop counter-term formula except total derivative terms.

Renormalization and anomaly are two outstandingly important aspects of quantum field theory. They have been giving us various rich information and helping us to understand the field theory. The direct relation between the 1-loop counter-term formula and the anomaly formula clearly shows the intimate relation between them.

References

[1] For review. K.Fujikawa,Lecture in Quantum Gravity and Cosmology, Kyoto Summer Inst.1985,ed.by Sato and Inami,World Scientific :106 (1985)

[2] S.Ichinose and N.Ikeda,preprint RIMS-1034, US-95-02

[3] J.D.Bjorken and S.D.Drell.*Relativistic Quantum Mechanics* (McGraw-Hill.New York,1964).

[4] S.Ichinose and N.Ikeda, in preparation,

[5] G.'tHooft,Nucl.Phys.**B62**,444(1973).

[6] K.Fujikawa,Nucl.Phys.**B226**,437(1983).

[7] K.Fujikawa,Phys.Rev.Lett.**42**,1195(1979);**44**,1733(1980);
Phys.Rev.**D21**.2848(1980);**D22**,1499(E)(1980).

[8] J.Schwinger,Phys.Rev.**82**,664(1951).

[9] E.Abdalla, M.C.B.Abdalla and K.D.Rothe,*Non-perturbative Methods in 2 Dimensional Quantum Field Theory*(World Scientific,Singapore, 1991).

[10] E.D'Hoker,*Tasi Lectures on Critical String Theory*, UCLA/92/TEP/30(1992).

[11] H.S.Tsao,Phys.Lett.**68B**,79(1977).

[12] S.Ichinose,Class.Quantum Grav.**12**,1021(1995).

TOPOLOGICALLY TWISTED 2-D CONFORMAL SUPERGRAVITY AND BRST GUAUGE FIXING †

Noriaki Ano *

Department of Physics, Rikkyo University,
Nishi-Ikebukuro, Tokyo171, Japan

ABSTRACT

It is mentioned that there exists a moduli space intrinsic to topologically twisted $osp(2|2)\oplus osp(2|2)$algebra in two dimensions through Noether current approach recently proposed by the auther. This means that N=2 'pure' conformal supergravity theory in two dimensions can be under twisting procedure.

1 Introduction

We know two typical stand points for constructing the cohomological field theory [1][2], i.e. topological twisting and BRST gauge fixing. The theory which we refer to as TFT deals with topological invariants, that is, a kind of the topological field theories [3]. Both approaches, topological twisting and BRST gauge fixing, result in the so-called moduli problem [2][4]. In the BRST approach, the relation between the moduli problem and TFT may be comparatively clear owing to the intrinsic constructing procedure where some moduli problem can be settled as the gauge fixing condition. In the topological twisting formalism, the above relation is not much clear, on the contrary. It seems that there has not been a common recognition on what the topological twist is really doing.

In the recent publication[5], it is shown that the topological twist is related to moduli problem intrinsic to the topologically twisted superalgebra, i.e. the so-called topological algebra [6]. To obtain the algebra, we perform topological twist of $osp(2|2)\oplus osp(2|2)$which can be identified to the finite N=2 conformal superalgebra in two dimensions [7]. The derivation of the moduli space inherent in the topological algebra is elaborated through a

† Talk given at workshop on General Relativity and Gravitation'95, Jan. 1996, Nagoya(Japan)

* E-mail: nano@rikkyo.ac.jp

gauge system of $osp(2|2) \oplus osp(2|2)$and vanishing Noether current with aid of induction of Lagrangians of the topological field theory found in the twisted algebra. In case of considering matter systems, it is known that there exist the different models with different moduli problems from the same supersymmetric system, i.e. the so-called A- and B-twisted models (for detailed explanation of A- and B-twist, see Ref.[8]). It is necessary to emphasize that we are indeed interested in the geometrical object of the topological algebra, i.e. gauge connections configuration alone, not the configuration associated with non geometrical objects. The vanishing Noether current is a resultant under the weak coupling limit on the path-integration with the TFT Lagrangians found in the topological algebra. It is claimed that the obtained relation between the topological twist and the moduli proble inherent in the twisted algebra seems a new characteristic of TFT, through considering the pure gauge system of interest as geometrical object.

We can see that the above mentioned characteristic enables us to be under twisting procedure for N=2 'pure' conformal supergravity theory in two dimensions [9], moreover. The above 'pure' means that the theory is composed of the gauge fields alone, i.e., without non geometrical objects. It is well-known that such a gravity theory is trivial, and there seems no way to twist the theory. In Ref.[9], notwithstanding, it is argued that a twisting procedure of $osp(2|2) \oplus osp(2|2)$theory results in twisting of N=2 'pure' conformal supergravity in two dimensions associated with $osp(2|2) \oplus osp(2|2)$gauge symmetry via coupling of two well-known constructing approaches for TFT, i.e., both the twisting and the BRST gauge fixing. Hence we will be confirmed in the above argument through the following contexts reviewing the twisted $osp(2|2) \oplus osp(2|2)$theory in Ref.[5], and this is just the chief aim, here.

2 Specific Relation Between Twisting and Moduli Problem

Now let us review the twited $osp(2|2) \oplus osp(2|2)$theory presented in Ref.[5]. Here, the topological twist means mixing of the representation space of internal symmetry of supersymmetry group with that of the symmetry group with respect to the space-time, i.e. spinor space. Practically speaking, the twisting results in identification of the representation space of internal group of N=2 supersymmetry with that of the local Lorentz group. Therefore, it is easy to perform twisting of $osp(2|2) \oplus osp(2|2)$algebra in table(1) to get the topological algebra. Notations adapted here are as follows: The 2-d Lorentzian metric is $g_{ab} = g^{ab} = -(-)^a\delta_{ab}$. The gamma matrics and the ϵ simbol are $\gamma^0 = i\sigma^2$, $\gamma^1 = \sigma^1$, $\gamma^5 = \sigma^3$, $\epsilon^{ab} = \epsilon_{ab}$ and $\epsilon^{01} = 1$. The spinor metric reads $\eta^{+-} = \eta_{-+} = -1$, $\eta^{-+} = \eta_{+-} = 1$ and $\eta^{++} = \eta^{--} = \eta_{++} = \eta_{--} = 0$, from the charge conjugation matrix:$C = \gamma^0$. The algebra associated with $osp(2|2) \oplus osp(2|2)$after the twist is presented in table(2). To obtain the twisted algebra, what we have to do mainly is to replace complex super chrages Q, $\bar{Q}$, S and $\bar{S}$ with Q^+, Q^-, S^+ and S^-, respectively, and

to make the modified algebra closed.

The indices "+, −" are raised and lowerd with the metric of the spinor space: $\eta_{\alpha\beta}$. The complex Weyl spinors $\varphi_\alpha, \bar{\varphi}_\alpha$ are substituted for $\varphi_\alpha^+, \varphi_\alpha^-$, that is, $\varphi_\alpha = \frac{i}{\sqrt{2}}\varphi_\alpha{}^+, \bar{\varphi}_\alpha = \frac{i}{\sqrt{2}}\varphi_\alpha{}^-$. In relation to the above replacement of the super charges, remaining manipulations are the modification of the definitions of local Lorentz M and Weyl D generators so that the four (0,0)-form fermionic generators $Q_\pm{}^\pm$, $S_\pm{}^\pm$, have no charges with respect to M and D. The representation space accompanied with the internal symmetry group have been put upon the spinor space. Therefore, the modified M, D generators must be direct sums with $so(2) \oplus so(2)$ genrators V and A. The solution to this constraint then resolves uniquely into $\tilde{M} = M + 2iV$ and $\tilde{D} = D + 2iA$. As can be seen from table(2), the step for closure of the algebra have been taken to obtain either left- or right-chiral part. The left-chiral part is the case in table(2). Here note that Q, S denote $Q = Q_+{}^+ + Q_-{}^-$, $S = S_+{}^+ + S_-{}^-$, respectively. The internal $SO(2) \otimes SO(2)$ symmetry still remains as global internal symmetry whose charge is the so-called ghost-number, the generators of which are defined by $G \equiv 2i(A - V)$, $\tilde{G} \equiv 2i(A + V)$. We can see that Q and $S_+{}^-$ increase the ghost number by one unit, while $Q_+{}^-$ and S decrease it by the same quantity.

Let us next pay attention to the connection fields associated with the topological algebra: $h_\mu = e_\mu^z P_z + f_\mu^z K_z + \omega_\mu \tilde{M} + b_\mu \tilde{D} - \frac{1}{2}(Q_+{}^-\varphi_{\mu-}{}^+ + S_+{}^-\psi_{\mu-}{}^+ + \psi_\mu Q + \phi_\mu S)$, where $\phi_{\mu+}{}^- = 0 = \psi_{\mu+}{}^-$, $e_\mu^z = 0 = f_\mu^z$, $\psi_{\mu+}{}^+ = -\psi_{\mu-}{}^- = -\psi_\mu$ and $\phi_{\mu+}{}^+ = -\phi_{\mu-}{}^- = -\phi_\mu$. The parameters are as follows: $h_\mu = \xi_P^z P_z + \xi_K^z K_z + \lambda_l \tilde{M} + \lambda_d \tilde{D} - \frac{1}{2}(Q_+{}^-\varepsilon_-{}^+ + S_+{}^-\varepsilon_-{}^+ + \varepsilon Q + \kappa S)$. The connections a_μ, v_μ which are intrinsic to the generators A, V, respectively, of the internal symmetry $so(2) \oplus so(2)$ have been also under modification as follows: $a_\mu = 2i\omega_\mu$, $v_\mu = 2ib_\mu$. The above connections obey the transformation laws (table(3)) where $\mathcal{D}_\mu \varepsilon_-{}^+ = (\partial_\mu - \omega_\mu + b_\mu)\varepsilon_-{}^+$, $\mathcal{D}_\mu \kappa_-{}^+ = (\partial_\mu - \omega_\mu - b_\mu)\kappa_-{}^+$, $\mathcal{D}_\mu \xi_P^z = (\partial_\mu - \omega_\mu + b_\mu)\xi_P^z$ and $\mathcal{D}_\mu \xi_K^z = (\partial_\mu - \omega_\mu - b_\mu)\xi_K^z$. Using the above gauge connections, a moduli space associated with the topological algebra can be derived just through discussion about the Lagrangian which will be found in the twisted algebra. We will be able to see that a moduli space of flat connections: $\mathcal{M}_{flat} = \{R^\star = 0\}/\mathcal{G}^\star$ is indeed derived from the algebra presented in table(2). Here, $\star$ denotes the label of the generators in table(2). Moreover, we can see that $\mathcal{M}_{flat} \cong \mathcal{M}_0$:

$$\mathcal{M}_0 = \{R^{\star\star} = 0\}/\mathcal{G}^{\star\star}, \tag{1}$$

where the double star $\star\star$ means the label of the generators $\tilde{M}$, $\tilde{D}$, Q, S. These four generators are the intersection of the left- and right-chiral part of the twisted algebra.

Let us then introduce the three (quasi-) Lagrangians of TFT in explanation of the derivation of $\mathcal{M}_0$:

$$\mathcal{L}_Q = \{Q, Q_+{}^-\}, \quad \mathcal{L}_S = \{S, S_+{}^-\}, \quad \mathcal{L}_{QS} = \{Q, S\}. \tag{2}$$

We see that the relations are defined on 2-manifold with boundary owing to the Stokes' theorem and it is possible for the corresponding theory to be in the case of 2-manifold without boundary. We can make the formulation on manifold without boundary by means of the inner product of the path-integrals as "in" and "out" states [10][11]. At the quantum

level, the three (quasi-) Lagrangians are invariant under the symmetry generated by the topological algebra A^*, because, if the Lagrangian is the exact form of BRST-like operator $\mathcal{Q}$, the path-integral of the $\mathcal{Q}$-exact form are trivial. The path-integral with the $\mathcal{Q}$-exact Lagrangian is independent of the coupling factor because of the same reason. Therefore, the zero coupling condition materializes a leading contribution to the path-integral, that is, the reduction to the moduli space. Such a limit induces vanishing Lagrangian condition which also means vanishing Noether current in the present case with the Lagrangians composed of the generators of the topological algebra.

Let us adopt $\mathcal{L}_{QS}$ as Lagrangian. We then see that $J^0_M = 0$ after the limitation on the path-integration with $\mathcal{L}_{QS}$, where J^0_M is obtained from the Noether current associated with the $osp(2|2) \oplus osp(2|2)$Yang-Mills action through the procedure of the topological twist. The point to be made is that the adequate informations to prescribe for the derivation of the moduli space of flat connections can be derived from the Noether current owing to the ambiguity of the current, so that we are free to add to J^0_M the following terms: $\partial_\alpha \theta^{A^*} R^{A^* \alpha\beta}$. We can define that $\theta^{A^*} = \mathcal{D}^\dagger \delta a^{A^*}$, where a^{A^*} is a connection 1-form associated with the topological algebra and $\mathcal{D}^\dagger = *\mathcal{D}*$ on 2-manifold. Therefore, the informations:

$$R^{A^*} = 0, \quad \mathcal{D}^\dagger \delta a^{A^*} = 0, \tag{3}$$

are obtained as conditions for the moduli space: $\mathcal{M}_{flat}$. As mentioned above, we can see that $\mathcal{M}_{flat} \cong \mathcal{M}_0$. After the collapse of the gauge orbits, the nilpotent fermionic operators: Q, $Q_+{}^-$, S, $S_+{}^-$ which are the ghost number carriers retain on the moduli space $\mathcal{M}_{flat}$. Let us next introduce the total Lagrangian:

$$\mathcal{L}_{total} = \mathcal{L}_Q + \mathcal{L}_S + \mathcal{L}_{QS} = \{\mathcal{QS}, \mathcal{QS}^\dagger\}, \tag{4}$$

where $\mathcal{QS} = Q + S_+{}^-$ and $\mathcal{QS}^\dagger = Q_+{}^- + S$. Under the weak coupling limit, the vanishing Lagrangian condition means that $\mathcal{L}_{total} = 0$. ¿From the standpoint of the operation of $\mathcal{L}_{total}$ on $\mathcal{M}_{flat}$, the components: e^z_μ, f^z_μ, $\psi_{\mu-}{}^+$, $\phi_{\mu-}{}^+$ are not in $\mathcal{M}_{flat}$. The fact is also supported by the relation: $\{Q, S\} \sim \tilde{M}$. That is, the following two relations:

$$\phi_{\mu-}{}^+ \underset{S}{\overset{Q}{\rightleftharpoons}} f^z_\mu, \; e^z_\mu \underset{S}{\overset{Q}{\rightleftharpoons}} \psi_{\mu-}{}^+, \tag{5}$$

are on the gauge orbits, and then this connections are collapsed together with the gauge orbits under the weak coupling limit. Therefore, $\mathcal{M}_{flat}$ is reduced to $\mathcal{M}_0$. We then claim that the moduli space intrinsic to the topologically twisted $osp(2|2) \oplus osp(2|2)$is really $\mathcal{M}_0$.

Clearly, our discussions, from which some geometrical informations associated with the topological algebra can be read, have been tacitly based on the path-integral recipe. The important question is then whether the path-integral under consideration is trivial or not. In TFT, the triviality depends on the details of the moduli space (or rigorously speaking, the classical configuration under the weak coupling limit on the path-integration). If the classical configuration corresponds to the full $\mathcal{M}_0$, the path-integral measure would be free

from the ghost number anomaly and contains the Grassman factors: $d\phi\, d\psi$. We can see that there exists a TFT observable $\mathcal{O}$ which assure that the path-integral is non-trivial:

$$\mathcal{O} = \int_{M^2} U \wedge *U = -2 \int_{M^2} d\mu\, \psi_0 \phi_1 \psi_1 \phi_0, \tag{6}$$

where $U = \psi \wedge \phi$ and $d\mu$ means 2-dimensional measure which is determined by a 2-dimensional metric: g_{ab}. It is easy to see that $\mathcal{O}$ is not dependent on g_{ab} (or rather, $d\mu$ dependence) and is gauge-invariant modulo Q- and S-exact, which also means that $\mathcal{O}$ is gauge invariant on the path-integration.

3 Conclusion

Let us make concluding remarks. We have reviewed the twisted $osp(2|2) \oplus osp(2|2)$theory in Ref.ano1. The twisted algebra includes the appropriate TFT's Lagrangians composed of the fermionic operators: Q, S, Q_+^- and S_+^-. They lead us to $\mathcal{M}_0$ under the weak coupling limit. The fact which have been clarified in the above discussions shows that the twisted algebra have a relation with a certain moduli problem, and that such a geometrical feature of the algebra is one of the interesting characteristics inherent in the topological twist. Hence we must learn that the above mentioned twisting approach upon $osp(2|2) \oplus osp(2|2)$theory is just regarded as the twisting of N=2 'pure' conformal supergravity theory in two dimensions.

Acknowledgment: The author would like to thank Professor H. Fujisaki for encouragement and enlightening discossions.

References

[1] E. Witten, Commun. Math. Phys. **117**, 353 (1988).

[2] E. Witten, Int. J. Mod. Phys. **A6**, 2775 (1991).

[3] See, for example, D. Birmingham, M. Blau, M. Rakowski and G. Thompson, Phys. Rep. **209**, 129 (1991), and references therein.
For recent developments, see, S. Cordes, G. Moore, and S. Ramgoolam, *Lectures on 2D Yang-Mills Theory, Equivariant Cohomology and Topological Field Theories*, YCTP-p11-94, hep-th/9411210, and references therein.

[4] J. Sonnenschein, Phys. Rev. **D42**, 2080 (1990).

[5] N. Ano, J. Math. Phys. **37**, 880 (1996).

[6] J. M. F. Labastida and P. M. Llatas, Nucl. Phys. **B379**, 220 (1992).

[7] J. MacCabe and B. Velikson, Phys. Rev. **D40**, 40 (1989).

[8] P. Fré and P. Soriani, "The N=2 Wonderland", World Scientific 1995, and references therein.

[9] N. Ano, "Two Dimensional N=2 Conformal Supergravity from TFT Viewpoint", to appear.

[10] See, for example, G. Thompson, *1992 Trieste Lectures on Topological Gauge Theory and Yang-Mills Theory*, in: Proc.High Energy Physics and Cosmology (Trieste, 1992), hep-th/9305120, and references therein.

[11] E. Witten, Nucl. Phys. **340**, 281 (1990).

Consistency of matter field equations in Ashtekar formulation

MOTOMU TSUDA AND TAKESHI SHIRAFUJI
Physics Department, Saitama University
Urawa, Saitama 338, Japan

Abstract

The chiral Lagrangian including more than two spin-3/2 fields becomes complex in general when the self-dual connection satisfies its equation of motion, and the imaginary part of the chiral Lagrangian gives the additional equation for spin-3/2 fields which gives rise to the inconsistency. This inconsistency can be removed by taking the tetrad to be complex and defining the total Lagrangian which is the sum of the chiral Lagrangian and its complex conjugate. It is possible to establish the right (left)-handed supersymmetry in its total Lagrangian as in the case of the chiral Lagrangian. We also comment on the canonical formulation of the total Lagrangian.

The Ashtekar refomulation of canonical gravity [1]-[4] can be derived from the chiral Lagrangian in which a tetrad and a complex self-dual connection are regarded as independent variables. Although the chiral Lagrangian is complex, its imaginary part vanishes in the source-free case when the tetrad is assumed to be real and the equation of self-dual connection is satisfied: The chiral Lagrangian then reduces to the Einstein-Hilbert Lagrangian. If the matter terms of integer spin fields exist, the chiral Lagrangian also becomes real as in the source-free case, because the Lagrangian of integer spin fields does not contain the self-dual connection. Moreover, it has been shown for spin-1/2 fields [4, 5] and simple supergravity [3, 6, 7] that the imaginary part of the chiral Lagrangian vanishes (up to a total divergence term). But in general, the chiral Lagrangian becomes complex for more than two spin-3/2 fields, and its imaginary part gives the additional equation for spin-3/2 fields which gives rise to the inconsistency [8]. We suggest in this letter one of the possible ways to evade the inconsistency of matter field equations.

We consider a spacetime manifold with a metric $g_{\mu\nu}$ constructed from a tetrad e^i_μ via $g_{\mu\nu} = e^i_\mu e^j_\nu \eta_{ij}$, and denote a self-dual connection by $A^{(+)}_{ij\mu} = A^{(+)}_{[ij]\mu}$ which satisfy

$$A^{(+)*}_{ij\mu} := \frac{1}{2}\epsilon_{ij}{}^{kl} A^{(+)}_{kl\mu} = iA^{(+)}_{ij\mu}, \tag{1}$$

where the $*$ operation means the duality operation in Lorentz indices. [1] In order to see the inconsistency of matter field equations, let us take N-(Majorana) Rarita-Schwinger fields ψ_μ ($:= \psi_\mu^I$ with I running from 1 to N) as example. The chiral Lagrangian density $\mathcal{L}^{(+)}$ including N-(Majorana) Rarita-Schwinger fields is

$$\mathcal{L}^{(+)} = \mathcal{L}_G^{(+)} + \mathcal{L}_{RS}^{(+)}. \tag{2}$$

Here $\mathcal{L}_G^{(+)}$ is the chiral gravitational Lagrangian density constructed from the tetrad and the self-dual connection:

$$\mathcal{L}_G^{(+)} = -\frac{i}{2} e\, \epsilon^{\mu\nu\rho\sigma} e_\mu^i e_\nu^j R_{ij\rho\sigma}^{(+)}, \tag{3}$$

where the unit with $8\pi G = c = 1$ is used, e denotes $\det(e_\mu^i)$ and the curvature of self-dual connection $R^{(+)ij}{}_{\mu\nu}$ is

$$R^{(+)ij}{}_{\mu\nu} := 2(\partial_{[\mu} A^{(+)ij}{}_{\nu]} + A^{(+)i}{}_{k[\mu} A^{(+)kj}{}_{\nu]}). \tag{4}$$

On the other hand, $\mathcal{L}_{RS}^{(+)}$ is the chiral Lagrangian density of N-(Majorana) Rarita-Schwinger fields, which is obtained by assuming that the fields ψ_μ are minimally coupled to gravity and that the Lagrangian is described by using only the self-dual connection:

$$\mathcal{L}_{RS}^{(+)} = -e\, \epsilon^{\mu\nu\rho\sigma} \overline{\psi}_{R\mu} \gamma_\rho D_\sigma^{(+)} \psi_{R\nu}, \tag{5}$$

where $\psi_R = (1/2)(1+\gamma_5)\psi$ and $D_\mu^{(+)}$ denotes the covariant derivative with respect to $A_{ij\mu}^{(+)}$:

$$D_\mu^{(+)} := \partial_\mu + \frac{i}{2} A_{ij\mu}^{(+)} S^{ij} \tag{6}$$

with S_{ij} standing for the Lorentz generator.

When the self-dual connection satisfies its equation of motion, the imaginary part of $\mathcal{L}^{(+)}$ vanishes for $N = 1$, while it does not vanish for $N \geq 2$. To see this, vary the $\mathcal{L}^{(+)}$ of (2) with respect to $A_{ij\mu}^{(+)}$ to obtain

$$\epsilon^{\mu\nu\rho\sigma} D_\nu^{(+)} H^{(+)ij}{}_{\rho\sigma} = -\frac{i}{2} e^{\mu\nu\rho\sigma} \left(e_\nu^{[i} \overline{\psi}_\rho \gamma^{j]} \psi_\sigma - \frac{i}{2} \epsilon^{ij}{}_{kl} e_\nu^k \overline{\psi}_\rho \gamma^l \psi_\sigma \right) \tag{7}$$

[1] Greek letters $\mu, \nu, \cdots$ are space-time indices, and Latin letters $i, j, \cdots$ are local Lorentz indices. We denote the Minkowski metric by $\eta_{ij} = \mathrm{diag}(-1,+1,+1,+1)$. The totally antisymmetric tensor ϵ_{ijkl} is normalized as $\epsilon_{0123} = +1$. The antisymmetrization of a tensor is denoted by $A_{[ij]} := (1/2)(A_{ij} - A_{ji})$.

with $H^{(+)ij}{}_{\mu\nu}$ being defined by $H^{(+)ij}{}_{\mu\nu} := e^i_{[\mu}e^j_{\nu]} - (i/2)\epsilon^{ij}{}_{kl}e^k_\mu e^l_\nu$, and $D^{(+)}_\mu e^i_\nu = \partial_\mu e^i_\nu + A^{(+)i}{}_{j\mu}e^j_\nu$. Equation (7) can be solved with respect to $A^{(+)}_{ij\mu}$ as

$$A^{(+)}_{ij\mu} = A^{(+)}_{ij\mu}(e) + K^{(+)}_{ij\mu}, \tag{8}$$

where $A^{(+)}_{ij\mu}(e)$ is the self-dual part of the Ricci rotation coefficients $A_{ij\mu}(e)$, while $K^{(+)}_{ij\mu}$ is that of $K_{ij\mu}$ given by

$$K_{ij\mu} := \frac{i}{4}(e^\rho_i e^\sigma_j e^k_\mu \,\overline{\psi}_\rho\gamma_k\psi_\sigma + e^\rho_i\,\overline{\psi}_\rho\gamma_j\psi_\mu - e^\rho_j\,\overline{\psi}_\rho\gamma_i\psi_\mu). \tag{9}$$

If the tetrad is real, substituing the solution of (8) back into the $\mathcal{L}^{(+)}$ of (2) gives the imaginary part of $\mathcal{L}^{(+)}$ as follows:

$$\mathrm{Im}\mathcal{L}^{(+)} = -\frac{1}{32}\sum_{I,J}\epsilon^{\mu\nu\rho\sigma}(\overline{\psi}^I_\mu\gamma_i\psi^I_\nu)(\overline{\psi}^J_\rho\gamma^i\psi^J_\sigma), \tag{10}$$

where we denote the indices I, J explicitly. Now the $\mathrm{Im}\mathcal{L}^{(+)}$ of (10) does not vanish except for $N = 1$. This non-vanishing $\mathrm{Im}\mathcal{L}^{(+)}$ yields the additional equation $\partial\mathrm{Im}\mathcal{L}^{(+)}/\partial\psi^I_\mu = 0$ in general, besides the Einstein equation and the Rarita-Schwinger equation obtained by taking variation of $\mathrm{Re}\mathcal{L}^{(+)}$ with respect to the tetrad and the (Majorana) Rarita-Schwinger fields. Therefore the difficulty of overdetermination for ψ^I_μ arises and the matter field equations generally become inconsistent. [2]

In order to remove this inconsistency of matter field equations, one could take $\mathcal{L}^{(+)}$ to be analytic for all fields; namely, the complex tetrad, the self-dual connection, the independent complex spinor fields ψ_R and $\overline{\psi}_R$. If we follow this procedure and take the tetrad to be complex in the Lagrangian of spin-1/2 fields, we have to take ψ and $\overline{\psi}$ to be independent in order to keep the matter field equations consistent. But ψ and $\overline{\psi}$ cannot be independent in QED or QCD, because if they are independent, the electric current, for example, will become complex.

So we adopt the following procedure as a simplest way to eliminate the term of (10): Namely, we add the complex conjugate of the chiral Lagrangian density, $\overline{\mathcal{L}^{(+)}}$, to the Lagrangian density $\mathcal{L}^{(+)}$, and define a total Lagrangian density $\mathcal{L}^{tot}$ by

$$\mathcal{L}^{tot} := \mathcal{L}^{(+)} + \overline{\mathcal{L}^{(+)}}. \tag{11}$$

[2] This inconsistency cannot be removed only by taking the tetrad to be complex in $\mathcal{L}^{(+)}$, because all terms in $\mathcal{L}^{(+)}$ (up to the term of (10)) merely becomes to be complex, and so the additional equations of ψ^I_μ also appear as in the case of the real tetrad.

Here we take the tetrad to be complex, because if the tetrad is real, $\mathcal{L}^{tot}$ reduces to the Lagrangian for ordinary Einstein gravity. The basic variables in $\mathcal{L}^{(+)}$ are the complex tetrad, the self-dual connection, the spinor fields ψ and its Dirac conjugate $\overline{\psi}$, while the basic variables in the $\mathcal{L}^{tot}$ of (11) are the set of $Q = (e, A^{(+)}, \psi)$ and their complex conjugate $\overline{Q}$.

From the definition of (11), the total gravitational Lagrangian density, $\mathcal{L}_G^{tot}$, is written by

$$\mathcal{L}_G^{tot} := -\frac{i}{2} e\, \epsilon^{\mu\nu\rho\sigma} e_\mu^i e_\nu^j R_{ij\rho\sigma}^{(+)} + \text{c.c.}, \tag{12}$$

where "c.c." means "the complex conjugate of the preceding term". It can be shown that $\mathcal{L}_G^{tot}$ describes the sum of complex general relativity and its complex conjugate when the self-dual connection (and its complex conjugate) satisfies its equations of motion. In the case of N-(Majorana) Rarita-Schwinger fields coupling, since the term of (10) does not contain the tetrad, it cancels with its complex conjugate in $\mathcal{L}^{tot}$. Accordingly, the inconsistency of matter field equations disappears. Furthermore we can show that $\mathcal{L}^{tot}$ describes the sum of complex Einstein gravity with four-fermion contact terms and its complex conjugate, after the equation of self-dual connection (and its complex conjugate) is solved.

The right (left)-handed supersymmetry in the chiral Lagrangian for $N = 1$ supergravity has been established [3, 6, 7]. In the same way, it is possible to establish right (left)-handed supersymmetry in $\mathcal{L}^{tot}$ for $N = 1$, if we now take $\psi_{R\mu}$ and $\overline{\psi}_{R\mu}$ in (5) to be independent and define the total Lagrangian density of (Majorana) Rarita-Schwinger fields by $\mathcal{L}_{RS}^{tot} := \mathcal{L}_{RS}^{(+)} + \overline{\mathcal{L}_{RS}^{(+)}}$; namely

$$\mathcal{L}_{RS}^{tot} = -e\, \epsilon^{\mu\nu\rho\sigma} \overline{\tilde{\psi}}_{R\mu} \gamma_\rho D_\sigma^{(+)} \psi_{R\nu} + \text{c.c.} \tag{13}$$

with $\overline{\tilde{\psi}}_{R\mu}$ and $\psi_{R\mu}$ being independent. Note that in special relativistic limit, the $\mathcal{L}_{RS}^{tot}$ of (13) becomes

$$L_{RS}^{tot} = \epsilon^{\mu\nu\rho\sigma} \overline{\tilde{\psi}}_\mu \gamma_5 \gamma_\rho \partial_\sigma \psi_\nu, \tag{14}$$

which can be diagonalized as

$$L_{RS}^{tot} = \epsilon^{\mu\nu\rho\sigma} (\overline{\psi}_\mu^1 \gamma_5 \gamma_\rho \partial_\sigma \psi_\nu^1 - \overline{\psi}_\mu^2 \gamma_5 \gamma_\rho \partial_\sigma \psi_\nu^2), \tag{15}$$

with $\psi_\mu^1 := (1/2)(\psi_\mu + \tilde{\psi}_\mu)$ and $\psi_\mu^2 := (1/2)(\psi_\mu - \tilde{\psi}_\mu)$. The minus sign in (15) means the appearance of negative energy states.

Since we take $\psi_{R\mu}$ and $\overline{\psi}_{R\mu}$ in (5) to be independent as stated above, we need two anticommuting Majorana spinor parameters α and $\tilde{\alpha}$ which generate the supersymmetry transformation. The total Lagrangian density $\mathcal{L}^{tot}$, which is the sum of (12) and (13), is invariant under the ψ-supersymmetry transformation generated by α,

$$\begin{aligned}
&\delta\psi_{R\mu} = 2D^{(+)}_{\mu}\alpha_R && \delta\overline{\tilde{\psi}}_{R\mu} = 0, \\
&\delta\psi_{L\mu} = 2D^{(+)}_{\mu}(\overline{A^{(+)}})\alpha_L && \delta\overline{\tilde{\psi}}_{L\mu} = 0, \\
&\delta e^i_{\mu} = -i\overline{\tilde{\psi}}_{R\mu}\gamma^i\alpha_R && \delta\overline{e^i_{\mu}} = -i\overline{\tilde{\psi}}_{L\mu}\gamma^i\alpha_L,
\end{aligned} \tag{16}$$

and the $\tilde{\psi}$-supersymmetry transformation generated by $\tilde{\alpha}$,

$$\begin{aligned}
&\tilde{\delta}\psi_{R\mu} = 0 && \tilde{\delta}\overline{\tilde{\psi}}_{R\mu} = 2\tilde{D}^{(-)}_{\mu}\overline{\tilde{\alpha}}_R, \\
&\tilde{\delta}\psi_{L\mu} = 0 && \tilde{\delta}\overline{\tilde{\psi}}_{L\mu} = 2\overline{\tilde{D}^{(-)}_{\mu}\tilde{\alpha}}_L, \\
&\tilde{\delta}e^i_{\mu} = -i\overline{\psi}_{L\mu}\gamma^i\tilde{\alpha}_L && \tilde{\delta}\overline{e^i_{\mu}} = -i\overline{\psi}_{R\mu}\gamma^i\tilde{\alpha}_R,
\end{aligned} \tag{17}$$

when we use the equation of self-dual connection (and its complex conjugate). [3] Here $\tilde{D}^{(-)}_{\mu}$ denotes the covariant derivative with respect to antiself-dual connection, $\tilde{A}^{(-)}_{ij\mu}$, and we assume that $\tilde{A}^{(-)}_{ij\mu}$ is the solution derived from the Lagrangian, $\tilde{\mathcal{L}}^{(-)} + \overline{\tilde{\mathcal{L}}^{(-)}}$ with $\tilde{\mathcal{L}}^{(-)}$ being

$$\tilde{\mathcal{L}}^{(-)} = \frac{i}{2}e\,\epsilon^{\mu\nu\rho\sigma}e^i_{\mu}e^j_{\nu}\tilde{R}^{(-)}_{ij\rho\sigma} + e\,\epsilon^{\mu\nu\rho\sigma}\overline{\psi}_{L\mu}\gamma_{\rho}\tilde{D}^{(-)}_{\sigma}\tilde{\psi}_{L\nu}, \tag{18}$$

where $\tilde{R}^{(-)ij}{}_{\mu\nu}$ is the curvature of antiself-dual connection. The commutator algebra of the above supersymmetry on the complex tetrad, for example, is easily calculated as

$$\begin{aligned}
&[\delta_1, \delta_2]e^i_{\mu} = 0 = [\tilde{\delta}_1, \tilde{\delta}_2]e^i_{\mu}, \\
&[\delta_1, \tilde{\delta}_2]e^i_{\mu} = 2iD_{\mu}(\overline{\tilde{\alpha}}_{2R}\gamma^i\alpha_{1R}).
\end{aligned} \tag{19}$$

Therefore, we have

$$[\delta^{tot}_1, \delta^{tot}_2]e^i_{\mu} = 2iD_{\mu}(\overline{\tilde{\alpha}}_{2R}\gamma^i\alpha_{1R} + \overline{\alpha}_{2L}\gamma^i\tilde{\alpha}_{1L}), \tag{20}$$

where $\delta^{tot} := \delta + \tilde{\delta}$. If the tetrad is real and $\psi_{R\mu} = \tilde{\psi}_{R\mu}$, the algebra of (20) coincides with that of $N = 1$ supergravity [9]. The commutator algebra on (Majorana) Rarita-Schwinger fields is now being calculated.

[3] We denote the covariant derivative with respect to $\overline{A^{(+)}_{ij\mu}}$ by $D^{(+)}_{\mu}(\overline{A^{(+)}}) := \partial_{\mu} + \frac{i}{2}\overline{A^{(+)}_{ij\mu}}S^{ij}$.

Finally we comment on the canonical formulation of $\mathcal{L}^{tot}$ in the source-free case. Although the Lagrangian $\mathcal{L}^{tot}$ involves the complex conjugate connection $\overline{A^{(+)}_{ij\mu}}$, the canonical formulation of $\mathcal{L}^{tot}$ in the source-free case can be expressed by using real canonical variables in a simple form. We assume spacetime manifold to be topologically $\Sigma \times R$, for some space-like submanifold Σ. Performing the Legendre transform, the $\mathcal{L}^{tot}_G$ of (12) can be written in canonical form:

$$\mathcal{L}^{tot}_G = -\mathrm{Tr}(\pi^{(+)a}\dot{A}^{(+)}_a) - \mathrm{Tr}(A^{(+)}_0\mathcal{G}^{(+)}) - N^a\mathcal{H}^{(+)}_a - N\mathcal{H}^{(+)} + \text{c.c.}, \tag{21}$$

where the first four terms of right hand side of (21) just come from $\mathcal{L}^{(+)}$. Here N^a is the complex shift vector, N is a density of weight -1 representing the complex lapse, and

$$\mathcal{G}^{(+)ij} := D^{(+)}_a\pi^{(+)ija} = \partial_a\pi^{(+)ija} + [A^{(+)}_a, \pi^{(+)a}]^{ij} \approx 0 \tag{22}$$

$$\mathcal{H}^{(+)}_a := -\mathrm{Tr}(\pi^{(+)b}R^{(+)}_{ab}) \approx 0 \tag{23}$$

$$\mathcal{H}^{(+)} := \mathrm{Tr}(\pi^{(+)a}\pi^{(+)b}R^{(+)}_{ab}) \approx 0 \tag{24}$$

are the Gauss, vector and Hamiltonian constraint, respectively, with $\pi^{(+)ija}$ the densitized momenta conjugate to $A^{(+)}_{ija}$. [4] The trace in (21) to (24) is defined as

$$\mathrm{Tr}(XY) := X_i{}^jY_j{}^k \qquad \mathrm{Tr}(XYZ) := X_i{}^jY_j{}^kZ_k{}^l. \tag{25}$$

If we define real canonical variables A_{ija} and π^{ija} from the $A^{(+)}_{ija}$ and its conjugate momenta $\pi^{(+)ija}$ as

$$A_{ija} := A^{(+)}_{ija} + \overline{A^{(+)}_{ija}}, \tag{26}$$

$$\pi^{ija} := \pi^{(+)ija} + \overline{\pi^{(+)ija}}, \tag{27}$$

and split the N^a and N into the real and imaginary part as $N^a_1 + iN^a_2$ and $N_1 + iN_2$, then the $\mathcal{L}^{tot}_G$ of (21) becomes

$$\mathcal{L}^{tot}_G = -\mathrm{Tr}(\pi^a\dot{A}_a) - \mathrm{Tr}(A_0\mathcal{G}) - N^a_1\mathcal{H}_a - N^a_2\mathcal{H}^*_a - N_1\mathcal{H} - N_2\mathcal{H}^*, \tag{28}$$

where

$$\mathcal{G}^{ij} := \mathcal{G}^{(+)ij} + \text{c.c.} = D_a\pi^{ija} = \partial_a\pi^{ija} + [A_a, \pi^a]^{ij} \approx 0 \tag{29}$$

[4] Latin letters $a, b, \cdots$ are spatial indices on Σ.

$$\mathcal{H}_a := \mathcal{H}_a^{(+)} + \text{c.c.} = -\text{Tr}(\pi^b R_{ab}) \approx 0 \tag{30}$$

$$\mathcal{H}_a^* := i(\mathcal{H}_a^{(+)} - \text{c.c.}) = -\text{Tr}(\pi^b R_{ab}^*) \approx 0 \tag{31}$$

$$\mathcal{H} := \mathcal{H}^{(+)} + \text{c.c.} = \text{Tr}(\pi^a \pi^b R_{ab}) \approx 0 \tag{32}$$

$$\mathcal{H}^* := i(\mathcal{H}^{(+)} - \text{c.c.}) = \text{Tr}(\pi^a \pi^b R_{ab}^*) \approx 0 \tag{33}$$

with the $*$ operation meaning the duality operation in Lorentz indices. These constraints of (29) to (33) form a first-class set, because the constraints of (22) to (24) are first-class. The canonical formulation of (28) is just the $SO(3,1;R)$ theory which is derived from a viewpoint of generalization of the $SO(3;C)$ Ashtekar formulation [10].

In the source-free case, the canonical formulation of $\mathcal{L}^{tot}$ is equivalent to that of $\mathcal{L}^{(+)}$. When (Majorana) Rarita-Schwinger fields are coupled, however, the canonical formulation of $\mathcal{L}^{tot}$ may possibly differ from that of $\mathcal{L}^{(+)}$, because the term of (10) has canceled with its complex conjugate in $\mathcal{L}^{tot}$. The canonical formulation of $\mathcal{L}^{tot}$ including (Majorana) Rarita-Schwinger fields is under study.

We would like to thank the members of Physics Department at Saitama University for discussions and encouragement.

References

[1] Ashtekar A 1986 *Phys. Rev. Lett.* **57** 2244; 1987 *Phys. Rev.* **D36** 1587

[2] Jacobson T and Smolin L 1987 *Phys. Lett.* **196B** 39; 1988 *Class. Quantum Grav.* **5** 583

[3] Ashtekar A 1988 *New Perspectives in Canonical Gravity* (Naples: Bibliopolis)

[4] Ashtekar A 1991 *Lectures on Non-Perturbative Canonical Gravity* (Singapore: World Scientific)

[5] Ashtekar A, Romano J D and Tate R S 1989 *Phys. Rev.* **D40** 2572

[6] Jacobson T 1988 *Class. Quantum Grav.* **5** 923

[7] Capovilla R, Dell J, Jacobson T and Mason L 1991 *Class. Quantum Grav.* **8** 41

[8] Tsuda M, Shirafuji T and Xie H 1995 *Class. Quantum Grav.* **12** 3067

[9] Freedman D Z, van Nieuwenhuizen P and Ferrara S 1976 *Phys. Rev.* **D13** 3214
Freedman D Z and van Nieuwenhuizen P 1976 *Phys. Rev.* **D14** 912

[10] Peldán P 1994 *Nucl. Phys.* **B430** 460

Dynamics in the Ashtekar Gravity

HISA-AKI SHINKAI* GEN YONEDA† and AKIKA NAKAMICHI‡

*,‡ *Dept. of Physics, Waseda University, Okubo 3-4-1, Shinjuku, Tokyo 169, Japan*

† *Dept. of Mathematics, Waseda University, Okubo 3-4-1, Shinjuku, Tokyo 169, Japan*

Abstract

We examine the advantages of the $SO(3)$-ADM (Ashtekar) formulation of general relativity, from the point of following the dynamics of the spacetime for the case of a vacuum with/without a cosmological constant. In the first part of this paper, we show how to treat the constraints and reality conditions, and propose new slicing conditions for evolving three-hypersurfaces. In the second part, we consider the mechanism for passing a degenerate point, which is one of the advantages in the Ashtekar formulation, and show that a 'deformed slice' approach enables us to 'pass' such point.

1 Introduction

A decade has passed since the proposal of the new formulation of general relativity by Ashtekar [1]. By using the special pair of variables, the framework has many advantages in the treatment of gravity. That is, the constraint equations which appear in the theory become low-order polynomials and do not contain the inverses of the variables, which enables us to treat the degenerate points. The theory also has the correct form for gauge theoretical features, and suggests possibilities for treating a quantum description of gravity nonperturbatively.

We examine these advantages of the $SO(3)$-ADM (Ashtekar) formulation of general relativity, from the point of following the dynamics of the spacetime for the case of a vacuum with/without a cosmological constant. We concentrate our attention on the dynamics of the classical (real Lorenzian) spacetime, which are often treated as the evolution of the three-hypersurfaces using the ADM formulation. In order to apply the Ashtekar formalism in classical general relativity, we need to solve the reality conditions for the metric and the extrinsic curvature additionally. We examined those in detail in [2] and proposed a new pair of variables which enables us to treat constraint equations differently.

In this paper, we discuss two features. The first one is the framework of the evolution schemes, especially keeping in mind the applications to numerical relativity. Our strategy is a "combined scheme" of ADM and Ashtekar (or modified Ashtekar) formulations, which are explained in detail in §3. We also discuss some expected advantages in applying this formulation to classical numerical relativity.

The second one is the possibility of passing degenerate points in the theory. Assuming that both all dynamical variables and the calculation time are finite, we seek the condition which enables us to 'pass' degenerate points. We take two approaches, named 'intersecting slice' and 'deformed slice', and those are described in §4.

*Electronic address: shinkai@cfi.waseda.ac.jp
†Electronic address: yoneda@cfi.waseda.ac.jp
‡Electronic address: akika@cfi.waseda.ac.jp

We use greek letters $(\mu, \nu, \rho, \cdots)$, which range over the four spacetime coordinates $0, \cdots, 3$, while uppercase Latin letters from the middle of the alphabet $(I, J, K, \cdots)$ range over the four internal SO(1,3) indices $(0), \cdots, (3)$. Lower case latin indices from the middle of the alphabet $(i, j, k, ...)$ range over the three spatial indices $1, \cdots, 3$, while lower case latin indices from the beginning of the alphabet $(a, b, c, ...)$ range over the three internal SO(3) indices $(1), \cdots, (3)$[1]. We use volume forms ϵ_{abc}; $\epsilon_{abc}\epsilon^{abc} = 3!$.

2 Brief review of the Ashtekar formulation

The key feature of Ashtekar's formulation of general relativity [1] is the introduction of a self-dual connection as one of the basic dynamical variables. Let us write the metric $g_{\mu\nu}$ using the tetrad, e^I_μ, and define its inverse, E^μ_I, by $g_{\mu\nu} = e^I_\mu e^J_\nu \eta_{IJ}$ and $E^\mu_I := e^J_\nu g^{\mu\nu}\eta_{IJ}$. We define a SO(3,C) self-dual connection

$$\mathcal{A}^a_\mu := \omega^{0a}_\mu - \frac{i}{2}\epsilon^a{}_{bc}\omega^{bc}_\mu, \tag{1}$$

where ω^{IJ}_μ is a spin connection 1-form (Ricci connection), $\omega^{IJ}_\mu := E^{I\nu}\nabla_\mu e^J_\nu$. Note that the extrinsic curvature, $K_{ij} = -(\delta_i{}^l + n_i n^l)\nabla_l n_j$ in the ADM formalism, where ∇ is a covariant derivative on Σ, satisfies the relation $-K_{ij}E^{ja} = \omega^{0a}_i$, when the gauge condition $E^0_a = 0$ is fixed. So $\mathcal{A}^a_i$ is also expressed by

$$\mathcal{A}^a_i = -K_{ij}E^{ja} - \frac{i}{2}\epsilon^a{}_{bc}\omega^{bc}_i. \tag{2}$$

The lapse function, N, and shift vector, N^i, are expressed as $E^\mu_0 = (\frac{1}{N}, -\frac{N^i}{N})$. Ashtekar treated the set $(\mathcal{A}^a_i, \tilde{E}^i_a)$ as basic dynamical variables, where $\tilde{E}^i_a$ is an inverse of the densitized triad defined by $\tilde{E}^i_a := eE^i_a$, and where $e := \det e^a_i$ is a density.

The Hilbert action takes the form

$$S = \int \mathrm{d}^4x[\dot{\mathcal{A}}^a_i\tilde{E}^i_a + \frac{i}{2}\underset{\sim}{N}\tilde{E}^i_a\tilde{E}^j_b F^c_{ij}\epsilon^{ab}{}_c - 2\Lambda\underset{\sim}{N}\det\tilde{E} - N^i F^a_{ij}\tilde{E}^j_a + \mathcal{A}^a_0\mathcal{D}_i\tilde{E}^i_a], \tag{3}$$

where $\underset{\sim}{N} := e^{-1}N$, Λ is the cosmological constant, $\mathcal{D}_i\tilde{E}^i_a := \partial_i\tilde{E}^i_a - i\epsilon_{ab}{}^c\mathcal{A}^b_i\tilde{E}^i_c$, and where $F^a_{\mu\nu}$ is the curvature 2-form, defined as $F^a_{\mu\nu} := \partial_\mu\mathcal{A}^a_\nu - \partial_\nu\mathcal{A}^a_\mu - \frac{i}{2}\epsilon^a{}_{bc}(\mathcal{A}^b \wedge \mathcal{A}^c)_{\mu\nu}$, and $\det\tilde{E}$ is defined to be $\det\tilde{E} = \frac{1}{6}\epsilon^{abc}\underset{\sim}{\epsilon}_{ijk}\tilde{E}^i_a\tilde{E}^j_b\tilde{E}^k_c$, where [2] $\epsilon_{ijk} := \epsilon_{abc}e^a_i e^b_j e^c_k$ and $\underset{\sim}{\epsilon}_{ijk} := e^{-1}\epsilon_{ijk}$.

Varying the action with respect to the non-dynamical variables $\underset{\sim}{N}$, N^i and $\mathcal{A}^a_0$ yields the constraint equations,

$$C_H = \frac{i}{2}\epsilon^{ab}{}_c\tilde{E}^i_a\tilde{E}^j_b F^c_{ij} - 2\Lambda\det\tilde{E} \approx 0, \tag{4a}$$

$$C_{Mi} = -F^a_{ij}\tilde{E}^j_a \approx 0, \tag{4b}$$

$$C_{Ga} = \mathcal{D}_i\tilde{E}^i_a \approx 0. \tag{4c}$$

The equations of motion for the dynamical variables ($\mathcal{A}^a_i$ and $\tilde{E}^i_a$) are

$$\dot{\mathcal{A}}^a_i = -i\epsilon^{ab}{}_c\underset{\sim}{N}\tilde{E}^j_b F^c_{ij} + N^j F^a_{ji} + \mathcal{D}_i\mathcal{A}^a_0 + 2\Lambda\underset{\sim}{N}\tilde{e}^a_i, \tag{5a}$$

$$\dot{\tilde{E}}^i_a = -i\mathcal{D}_j(\epsilon^{cb}{}_a\underset{\sim}{N}\tilde{E}^j_c\tilde{E}^i_b) + 2\mathcal{D}_j(N^{[j}\tilde{E}^{i]}_a) + i\mathcal{A}^b_0\epsilon_{ab}{}^c\tilde{E}^i_c, \tag{5b}$$

[1] We raise and lower μ, ν, ρ by $g^{\mu\nu}$ and $g_{\mu\nu}$(Lorenzian metric); I, J, K by $\eta^{IJ} = \mathrm{diag}(-1, 1, 1, 1)$ and η_{IJ}; i, j, k by γ^{ij} and γ_{ij}(3-metric).

[2] $\epsilon_{xyz} = e$, $\underset{\sim}{\epsilon}_{xyz} = 1$, $\epsilon^{xyz} = e^{-1}$, $\tilde{\epsilon}^{xyz} = 1$.

where $\mathcal{D}_j T^{ji}_a := \partial_j T^{ji}_a - i\epsilon^a{}_b{}^c \mathcal{A}^b_j T^{ji}_c$, for $T^{ij}_a + T^{ji}_a = 0$, and $\tilde{e}^a_i$ is also expressed as $\tilde{e}^a_i = \frac{1}{2}\epsilon^{abc}\epsilon_{ijk}\tilde{E}^j_b\tilde{E}^k_c$.

3 Procedure for evolving three-hypersurfaces

In this section, we discuss the framework of the evolution schemes, especially keeping in mind the applications to numerical relativity. §3.1 and §3.2 are digest version of [2].

3.1 Reality conditions

To ensure that the metric is real-valued, we need to impose two conditions; the first is that the doubly densitized contravariant metric $\tilde{\gamma}^{ij} := e^2\gamma^{ij}$ is real,

$$\Im(\tilde{E}^i_a\tilde{E}^{ja}) = 0, \qquad \textbf{metric reality condition} \tag{6a}$$

and the second condition is that the time derivative of $\tilde{\gamma}^{ij}$ is real,

$$\Im\{\partial_t(\tilde{E}^i_a\tilde{E}^{ja})\} = 0. \qquad \textbf{second metric reality condition} \tag{6b}$$

We denote these the "metric reality condition" and the "second metric reality condition" (extrinsic curvature reality condition), hereafter. Ashtekar *et al.* [3] discovered that, with the second metric reality condition (6b), the reality of the 3-metric and extrinsic curvature are automatically preserved under time evolution, as a consequence of the equations of motion. This means we need only solve both reality conditions (6a) and (6b) on the initial hypersurface.

Using the equations of motion for $\tilde{E}^i_a$ (5b), the gauge constraint (4c) and the first reality condition (6a), we can replace the second reality condition (6b) with a different constraint

$$W^{ij} := \Re(\epsilon^{abc}\tilde{E}^k_a\tilde{E}^{(i}_b\mathcal{D}_k\tilde{E}^{j)}_c) \approx 0, \tag{7}$$

which fixes six components of $\mathcal{A}^a_i$ and $\tilde{E}^i_a$. Moreover, in order to recover the original lapse function $N := \underset{\sim}{N}e$, we demand $\Im(N/e) = 0$, i.e. the density e be real and positive. This requires that e^2 be positive, i.e.

$$\det\tilde{E} > 0. \tag{8}$$

Note that this condition does not remove any degrees of freedom for the variables and is analogous to making the implicit assumption of $\det\gamma_{ij} > 0$ in the ADM formulation. In [2], we show the secondary condition of (8) is automatically satisfied.

Next, we show that rather stronger reality conditions are also useful in Ashtekar's formalism for recovering the real 3-metric and extrinsic curvature. These conditions are

$$\Im(\tilde{E}^i_a) = 0 \qquad \textbf{first triad reality condition} \tag{9a}$$

$$\text{and } \Im(\dot{\tilde{E}}^i_a) = 0 \qquad \textbf{second triad reality condition.} \tag{9b}$$

Using the equations of motion of $\tilde{E}^i_a$, the gauge constraint (4c), the metric reality conditions (6a), (6b) and the first condition (9a), we see that (9b) is equivalent to

$$\Re e(\mathcal{A}^a_0) = \partial_i(\underset{\sim}{N})\tilde{E}^{ia} + \frac{1}{2}e^{-1}e^b_i\underset{\sim}{N}\tilde{E}^{ja}\partial_j\tilde{E}^i_b + N^i\Re(\mathcal{A}^a_i). \tag{10}$$

From this expression we see that the second triad reality condition restricts the three components of the "triad lapse"[3] vector $\mathcal{A}_0^a$. Therefore (10) is not a restriction on the dynamical variables ($\mathcal{A}_i^a$ and $\tilde{E}_a^i$) but on a part of the slicing, which we should impose on each hypersurface. Thus the second triad reality condition does not restrict the dynamical variables any further than the second metric condition does.

3.2 Alternative treatment of constraint equations

The equations we need to solve for $\mathcal{A}_i^a$ and $\tilde{E}_a^i$ are the constraints (4a), (4b), (4c) and the reality conditions (7), (9a). Capovilla-Jacobson-Dell(CDJ)[4] solved C_H and C_M by introducing new variables. These reduced the 36 (real) independent components of $\mathcal{A}_i^a$ and $\tilde{E}_a^i$ to 28, or in CDJ's variables the 18 (real) independent components of ψ_{ab} are reduced to 10 (a symmetric and traceless tensor), which corresponds to the Weyl curvature Ψ_i. These are again restricted by C_G and the reality condition.

In contrast to CDJ's method, we make an alternative treatment of the gauge constraint (4c) and the second metric reality condition (7). For convenience, we assume that $\tilde{E}_a^i$ is real. This assumption (9a) restricts our choice of triad, but this constraint is not difficult to satisfy. We introduce the connection with double internal indices (note that here we do not use the densitized triad),

$$\mathcal{A}^{ab} := \mathcal{A}_i^a E^{ib}, \tag{11}$$

and express all the constraints with $(\mathcal{A}^{ab}, \tilde{E}_a^i)$ as the basic pair of variables [2]. The 6 real constraints of C_G are automatically satisfied if we impose the following conditions on the anti-symmetric part of $\mathcal{A}^{ab}$,

$$\Im m(\mathcal{A}^{[ab]}) = -\frac{1}{2e}\epsilon^{abc}\partial_i \tilde{E}_c^i. \tag{12}$$

$$\Re e(\mathcal{A}^{[ab]}) = 0 \tag{13}$$

Moreover, the second metric reality condition (7) is written as

$$\Im m(\mathcal{A}^{(ab)}) = \frac{1}{2}[E_d^j(\epsilon^{dac}\underset{\sim}{e}_i^b + \epsilon^{dbc}\underset{\sim}{e}_i^a)\partial_j \tilde{E}_c^i - \delta^{ab}E_d^j\epsilon^d{}_e{}^c\underset{\sim}{e}_i^e\partial_j\tilde{E}_c^i], \tag{14}$$

where $\underset{\sim}{e}_i^a$ is the inverse of $\tilde{E}_a^i$. From these expressions, we see that (4c) and (7) are satisfied if and only if $\mathcal{A}^{ab}$ satisfies (12), (13) and (14), after assuming (9a). We note that the imaginary part of $\mathcal{A}^{ab}$ consists of the triad and its spatial differential and that the real part of $\mathcal{A}^{ab}$ is symmetric.

These results become clearer if we compare $\Re e(\mathcal{A}^{ab})$ and the extrinsic curvature K_{ij} through the definition of $\mathcal{A}_i^a$, (2). From (2) we derive

$$\Re e(\mathcal{A}^{ab}) = -K_{ij}E^{ia}E^{jb}, \tag{15a}$$

$$\Im m(\mathcal{A}^{ab}) = -\frac{1}{2}\epsilon^a{}_{cd}\omega_i^{cd}E^{ib}. \tag{15b}$$

Since the extrinsic curvature is symmetric, we see $\Re e(\mathcal{A}^{ab})$ is also symmetric [(15a)]. After some calculation, we can see that (15b) is equivalent to (12) and (14). Moreover, we find from (2) that

$$K_{ij} = -\mathcal{A}_i^a e_{ja} - \frac{i}{2}\epsilon_{abc}\omega_i^{bc}e_j^a,$$

[3]This "triad lapse" was named by A.Ashtekar in private communication.

	ADM	Ashtekar		Ashtekar-CDJ[4]	Ashtekar-YS[2]
dynamical variables	γ_{ij}, K_{ij}	$\mathcal{A}^a_i, \tilde{E}^i_a$		$\mathcal{A}^a_i, \Psi_{ab}$	$\Re e[\mathcal{A}^{(ab)}], \tilde{E}^i_a$
gauge fixing variables	N, N^i	$N, N^i, \mathcal{A}^a{}_0$		$N, N^i, \mathcal{A}^a{}_0$	$N, N^i, \mathcal{A}^a{}_0$
constraints	$\mathcal{C}_H, \mathcal{C}_M$	$\mathcal{C}_H, \mathcal{C}_M, \mathcal{C}_G$		$\mathcal{C}_G$	$\mathcal{C}_H, \mathcal{C}_M$ $(\mathcal{C}_G \to \mathcal{A}^{[ab]})$
reality conditions	(none)	metric 1st 2nd	triad 1st 2nd	metric 1st 2nd	triad 1st metric 2nd $\to \Im m[\mathcal{A}^{(ab)}]$

Table 1: A list of alternative approaches for time evolution of the three-hypersurfaces.

so the reality of the extrinsic curvature, $\Im m(K_{ij}) = 0$, is equivalent to (15b). Consequently, $\Re e(\mathcal{C}_G) = 0$ [(12)] and $W^{ij} = 0$ [(14)] indicate that the extrinsic curvature is real and $\Im m(\mathcal{C}_G) = 0$ [(13)] indicates that the extrinsic curvature is symmetric.

When one has solved the 12 equations $\mathcal{C}_G = 0$ and $W^{ij} = 0$ for the 27 variables $\mathcal{A}^a_i$(complex) and $\tilde{E}^i_a$(real), 15 degrees of freedom remain. Introducing $\mathcal{A}^{ab}$ clarifies this remaining freedom; these are 6 degrees of freedom for $\Re e(\mathcal{A}^{(ab)})$ and 9 for $\tilde{E}^i_a$. Our task is now reduced to solving the other constraints (4a) and (4b) for the variables $\Re e(\mathcal{A}^{(ab)})$ and $\tilde{E}^i_a$, and the imaginary parts of (4a) and (4b) are trivially zero. The constraints $\mathcal{C}_H$ and $\mathcal{C}_{Mi}$ in terms of $\Re e(\mathcal{A}^{(ab)})$ and $\tilde{E}^i_a$ are shown in [2].

We expect that our variables are convenient for expressing the data on each hypersurface if we impose a reality condition. However, we remark that, like CDJ's variables, our variables ($\Re e(\mathcal{A}^{(ab)})$, $\tilde{E}^i_a$) are not canonical. Therefore, when we describe the equations of motion of our variables, we transform those of the canonical pair [e.g., (5a) and (5b)] into ours. Also note that the formulation is not polynomial. Consequently, our full set of equations consists of four constraint equations $\mathcal{C}_H$,$\mathcal{C}_{Mi}$ [together with definitions (12), (13) and (14)], and the equations of motion (5a), (5b).

3.3 Expected advantages in numerical relativity

In this subsection, we comment on the expected advantages in the applications for numerical relativity. Recently, Salisbury *et al.* [5] proposed the use of Ashtekar's and CDJ's formulations in the numerical treatment of general relativity. They showed a full set of equations using Ashtekar's connection and Newman-Penrose's Weyl scalar as dynamical variables. Their formulation is based on Ashtekar's variables and has the merit of showing the evolution of the Weyl scalar directly. However, it is not convenient for general numerical relativists, because the evolution of Ashtekar's connection does not give us a direct geometrical understanding.

We set our strategy such that we focus upon the ADM variables and use connection variables only when we have some advantages in developing the time slices. That is, we prepare the initial data within the ADM framework and use the Ashtekar or modified Ashtekar formulation (such

as CDJ, our variables discussed in the previous subsection, etc) for determining the slicing conditions and/or evolving slices. By this procedure, we do not need to solve reality conditions, because the reality is guaranteed in the initial ADM slice and is preserved during the evolution as we denoted in §3.1. In Table.1, we summarize the variables, constraints and reality conditions in the Ashtekar or modified Ashtekar formulation, together with the ADM formulation.

Next, we give some comments on the slicing conditions in the Ashtekar formulation. Since we have the additional gauge freedom of 'triad lapse' $\mathcal{A}_0^a$, there are wide varieties in choosing a slicing condition. Here, as an example, we take a constant-mean-curvature (CMC) or maximal slicing condition and its analogous slicing condition.

Let us define two scalars

$$\tilde{A} := \mathcal{A}_i^a \tilde{E}_a^i, \tag{16a}$$

$$A := \mathcal{A}_i^a E_a^i, \tag{16b}$$

and consider a condition of $\dot{\tilde{A}} = 0$ or $\dot{A} = 0$ as a slicing condition. When the triad reality condition is imposed, $\Re e(A^{ab})$ represents the extrinsic curvature. Therefore the latter condition is expected to have the properties of the CMC slicing condition.

The condition $\dot{A} = 0$ becomes

$$\begin{aligned} i\partial_j(\underset{\sim}{N})\epsilon^{cba}\tilde{E}_c^j\mathcal{A}^{ab} &= \underset{\sim}{N}[i(\partial_j\tilde{E}_b^i)\epsilon_a{}^{bc}E_c^j\mathcal{A}_i^a + \mathcal{A}^2 - \mathcal{A}^{ac}\mathcal{A}_{ac} + 6e\Lambda] \\ &\quad + N^i[(\partial_i\tilde{E}_a^j)\mathcal{A}_j^a - i\epsilon_a{}^{bc}\mathcal{A}_i^b\mathcal{A}^{ac}] - (\partial_j N^i)E_a^j\mathcal{A}_i^a + E_a^i\partial_i\mathcal{A}_0^a \end{aligned} \tag{17}$$

The real part of this (17) gives us the same CMC slicing condition in the ADM formulation (but in a rather complicated form for calculation). If we assume the triad reality condition (9a) and (9b) in treating the real part of (17), then we see the 'triad lapse' $\mathcal{A}_0^a$ contributes only in the $E_a^i\Re e(\partial_i\mathcal{A}_0^a)$ term, which is fixed by the condition (10). The fact of this disappearance of the $\mathcal{A}_0^a$ contribution is because the original CMC condition is free from the choice of $\mathcal{A}_0^a$. However, the imaginary part of (17) restricts a freedom of $\mathcal{A}_0^a$.

Similarly, the condition $\dot{\tilde{A}} = 0$ becomes

$$\begin{aligned} i(\partial_j\underset{\sim}{N})\epsilon_{ab}{}^c\tilde{E}_c^j\tilde{\mathcal{A}}^{ab} &= \underset{\sim}{N}[i(\partial_j\tilde{E}_b^i)\epsilon_a{}^{bc}\tilde{E}_c^j\mathcal{A}_i^a + \tilde{\mathcal{A}}^2 - \tilde{\mathcal{A}}^{ac}\tilde{\mathcal{A}}_{ac} + 6e^2\Lambda] + N^j[(\partial_j\tilde{E}_a^i)\mathcal{A}_i^a - i\epsilon_{abc}\mathcal{A}_j^b\tilde{\mathcal{A}}^{ac}] \\ &\quad + (\partial_j N^j)\tilde{\mathcal{A}} - (\partial_j N^i)\tilde{E}_a^j\mathcal{A}_i^a + (\partial_i\mathcal{A}_0^a)\tilde{E}_a^i. \end{aligned} \tag{18}$$

Since the definition of $\tilde{\mathcal{A}}$ is a contraction of the canonical pair, (18) is regarded as an analogy of the CMC slicing condition. The real part of (18) means $\sqrt{\det\gamma_{ij}}K$ =const. in terms of ADM variables, and the imaginary part restricts a component of the triad lapse.

The determination of which slicing condition works properly depends on the particular physical model. Further proposals and experiments of these slicing conditions in the Schwarzschild spacetime are now under preparation[6].

4 How can we pass a degenerate point?

In this section, we consider the possibilities of passing a degenerate point, which is one of the advantages in Ashtekar's formulation. After we describe our definitions of 'passing' degenerate point in §4.1, we consider two approaches; intersecting slice approach (§4.2) and deformed slice approach (§4.3).

4.1 Our definitions of 'passing' degenerate point

In the beginning, let us clarify some terminologies. A 'degenerate point' is the point in the spacetime where the density e of 3-space vanishes [4]. In the Ashtekar formulation, the density is defined as $e = \sqrt{\det \tilde{E}^i_a}$, and all the equations remain finite at points $e = 0$. From the point of dynamics, this fact suggests to us that we can 'pass' such a degenerate point. Here, we make clear the meaning of this possibility of 'passing'.

Note that, in the ADM formulation, this is impossible because the equations in the ADM include an inverse of the variables. We also remark that we can not transform the Ashtekar variables onto the ADM three-hypersurface Σ at the degenerate point. This is because the 3-metric γ_{ij} is given by $\gamma_{ij} := e^a_i e^a_j$, where $e^a_i := (E^i_a)^{-1}$ and $E^i_a := \tilde{E}^i_a/e$, and the last quantity diverges as $e \to 0$. Therefore the degenerate points in the Ashtekar formulation do not cover the real manifold.

In order to say 'pass' degenerate points, we demand the following four conditions: (a) Ashtekar variables $\tilde{E}^i_a, \mathcal{A}^a_i, \underset{\sim}{N}, N^i, \mathcal{A}^a_0$ must remain finite throughout the calculation, (b) the spatial derivatives of them must also be finite throughout the calculation (because they appear in the equations of motion) (c) the time differentiation of dynamical variables $\dot{\tilde{E}}^i_a$, $\dot{\mathcal{A}}^a_i$ must also be finite throughout the calculation, and (d) the calculation must be finished in finite coordinate time.

We take mainly two approaches. The first one, named 'intersecting slice approach', passes a degenerate point directly, and the second one, 'deformed slice approach' takes a complex path around a real degenerate point. The former is divided into two cases from the point whether we impose the reality conditions or not, while the latter assumes the complex values necessarily.

4.2 Intersecting slice approaches

For a while, let's consider first the case that we impose reality conditions. The problem is whether we can find a solution under our assumptions of (a)-(d) above. To say our conclusion first, we face at least two problems; *degenerate lapse* and *divergence of connection* ω^{0a}_i, and those may make this approach impossible, unfortunately.

First we discuss the *degenerate lapse problem.* The troublesome variable is the (inverse) densitized lapse $\underset{\sim}{N} := N/e_\Sigma$, where $e_\Sigma = \sqrt{\det \gamma_{ij}}$. Since $\underset{\sim}{N}$ is held finite [condition (a)], the ADM lapse $N = \underset{\sim}{N} e_\Sigma$ vanishes at the degenerate point $e_\Sigma = 0$. Notice that the ratio of the proper time to the coordinate time is not $\underset{\sim}{N}$ but N. Thus we are afraid that the calculation exhausts an infinite amount of time, i.e., the condition (d) is violated. Let us take a parameter of coordinate time t which becomes $t = 0$ at the degenerate point $e_\Sigma = 0$. Then the condition

[4] Here we assume that degenerate area is a point. Another cases are commented on in §5.

(d) is denoted by

$$\delta t = \int_{-t_0}^{t_1} dt = \int_{-\tau_0}^{\tau_1} \frac{d\tau}{N(\tau)} < \infty.^{5} \tag{19}$$

Here the range of the integral is arbitrary but includes the zero point $N(0) = 0$. Note that $N \geq 0$ since it is a lapse function. Compatibility of (19) and $N(0) = 0$ will be considerably restrictive. Let us call this condition the *degenerate lapse passing condition (DLPC).*

For a continuous function $N(t) : R \to [0, \infty)$ satisfying $N(0) = 0$, what are the conditions for N to satisfy $\int_0^1 \frac{dt}{N(t)} < \infty$? If $\dot{N}(t)$ exists for $0 < t$, we see $\limsup_{t \to 0+} \dot{N}(t) = \infty$ as follows. Put $k = \limsup_{t \to 0+} \dot{N}(t)$. [6] For any $\epsilon > 0$, there exists $\delta > 0$ such that $\dot{N}(t) \leq k + \epsilon$ for $0 \leq t \leq \delta$. We have $\dot{N}(t) \leq (k + \epsilon)$ for $0 \leq t \leq \delta$. Thus $N(t) \leq (k + \epsilon)t$ for $0 \leq t \leq \delta$. Then we see

$$\infty > \int_0^\delta \frac{\mathrm{d}t}{N(t)} \geq \int_0^\delta \frac{\mathrm{d}t}{(k+\epsilon)t} = \frac{1}{(k+\epsilon)} \int_0^\delta \frac{\mathrm{d}t}{t} = \frac{\infty}{(k+\epsilon)}$$

If $k < \infty$ then the last term goes to infinity, it causes contradiction. Thus $k = \infty$.

Since $N(t) \geq 0$, if we assume $N(t)$ is smooth, then $\dot{N}(0) = 0$. Thus smooth $N(t)$ does not satisfy the DLPC. If we assume $N(t) = |t|^s$ where s is constant, then the condition is $0 < s < 1$. For example, $N(t) = \sqrt{|t|}$ satisfies the DLPC.

Furthermore we will see that the density $|e|$ should also satisfy the DLPC as follows. We assume $e(0) = 0$ and $|\underset{\sim}{N}(t)| < M$ (bounded). Since $N(t) \geq 0$, we have $e(t)\underset{\sim}{N}(t) = |e(t)\underset{\sim}{N}(t)| = |e(t)|\ |\underset{\sim}{N}(t)|$. Then we see

$$\infty > \int_0^1 \frac{dt}{N(t)} = \int_0^1 \frac{dt}{e(t)\underset{\sim}{N}(t)} = \int_0^1 \frac{dt}{|e(t)|\ |\underset{\sim}{N}(t)|} > \frac{1}{M} \int_0^\delta \frac{dt}{|e(t)|}.$$

Note that the integral is bounded and monotonically increasing. Thus $|e(t)|$ also satisfies the DLPC [7]. Totally we see that N and $|e|$ satisfies DLPC which is considerably restrictive.

Secondly we discuss the problem *divergence of connection.* The connection variable $\mathcal{A}^a_i$ includes

$$\omega_i^{0a} = -K_{ij}E^{ja} = \frac{1}{e^2\underset{\sim}{N}}(\gamma_{ij} - D_iN_j - D_jN_i)\tilde{E}^j_a.$$

In order to examine this finiteness, we prepare

$$\omega_i^{0a}\tilde{E}^i_a = \frac{1}{\underset{\sim}{N}}\left(\frac{\partial_t(e^2)}{e^2} - D_i(N^i) - D_j(N^j)\right) = \frac{1}{\underset{\sim}{N}}\left(\frac{\partial_t(e^2)}{e^2} - \frac{\partial_j(e^2)}{e^2}N^j - 2\partial_i(N^i)\right).$$

Let us evaluate the last one. The term $1/\underset{\sim}{N}$ is bounded below since $\underset{\sim}{N}$ is finite, $1/\underset{\sim}{N} > 1/M$. Note that $f'(t)/f(t)$ diverges when $f(0) = 0$. The term $\frac{\partial_t(e^2)}{e^2} - \frac{\partial_j(e^2)}{e^2}N^j$ diverges. This is because, we take a parameter s whose direction is normal vector $\{1, -N^i\}$, then $\frac{\partial_t(e^2)}{e^2} - \frac{\partial_j(e^2)}{e^2}N^j$ equals

[5] Strictly, this is an improper Riemann integral.

[6] There are other possibilities than $\lim_{t \to 0+} \dot{N}(t) = \infty$. A counter-example is $N(t) = \sqrt{t}(2 + \sin(1/t))$.

[7] The region satisfying $e > 0$ and the region satisfying $e < 0$ are compatible when the manifold is divided by the region satisfying $e = 0$. Note also that e^2 is time differentiable, since $e^2 = \det \tilde{E}^i_a$. So if we assume $e(t) = t^s$, then $1/2 < s < 1$ must be satisfied.

$(de^2/ds)/e^2$, which diverges. The term $\partial_i(N^i)$ is finite since we assume (b). Thus we see that $\omega_i^{0a}\tilde{E}_a^i > \frac{1}{M}(divergent + finite)$, so $\omega_i^{0a}\tilde{E}_a^i$ diverges. If we assume ω_i^{0a} is finite, then $\omega_i^{0a}\tilde{E}_a^i$ must be finite. Therefore we see that ω_i^{0a} diverges.

This is a serious problem, since ω_i^{0a} is a part of $\mathcal{A}_i^a$. At least, if we assume triad reality, we have $\Re e(\mathcal{A}_i^a) = \omega_i^{0a}$, so divergence of ω_i^{0a} contradicts the finiteness of $\mathcal{A}_i^a$. Even if we relax the reality to metric reality condition, then $i\epsilon^a{}_{bc}\omega_i^{bc}$ must diverge in order to cancel out the divergence of ω_i^{0a}.

Next, let's consider the case without the reality conditions, i.e., apart from the Einstein frame and seek the possibilities of 'passing' within a complex manifold. In this situation, we naturally assume that the gauge variables ($\underset{\sim}{N}$, N^i and $\mathcal{A}_0^a$) which define the foliation of slices are complex. However, if we also assume that the slice goes across the degenerate point, then the above two problems arise and remain impractical as follows.

As for the degenerate lapse problem, if $N(\tau)$ is complex, the integral of (19) is also complex-valued. The real and imaginary parts of the integral must finitely converge (complex DLPC). We see that, as same as the above, the smooth complex N does not satisfy the complex DLPC. When we assume the form $N(t) = |t|^p + i|t|^q$ where p, q are constants, then the condition becomes $0 < \min(p, q) < 1$. As for the connection divergence, the fact that ω_i^{0a} diverges does not depend on the reality. Thus ω_i^{0a} diverges here. Since the two problems remain in this way, we conclude that our 'passing' conditions (a)-(d) are difficult to satisfy using the 'intersecting slice' approach.

4.3 Deformed slice approaches

The second approach, we call the 'deformed slicing approach', in which we make time slicing to avoid the degenerate point, by using complexibility of the Ashtekar variables [8]. We extend our slice to be in a complex manifold as in the second case in the previous approach, but here we do not intersect a degenerate point but avoid such points in the real section. In other words, the slice approaches the degenerate point, though it just goes around the point.

In general, when we take a detour in the complex plane, the uniqueness of the time evolution is not guaranteed. In order to discuss a physically meaningful path (recover the Einstein system in the far region), we concentrate on the path which starts from the real section, and evolves in the complex manifold only in the vicinity of a degenerate point. That is, imposing an 'asymptotic reality condition', such as $\Im m(N) \to 0, \Im m(N^i) \to 0$ and $\Im m(\tilde{E}_a^i \tilde{E}_a^j) \to 0$ both for spatial and time far limit, is expected to work and determine the dynamics properly. Here, we define an asymptotic region naively. When we apply this idea to a case of *one* degenerate point, we can easily define time in the far region as $t \to \pm\infty$.

When a coordinate position of a degenerate point is given, such as a caustic of a null fluid, this 'asymptotic reality condition' can be translated into 'time-reversal slicing condition'. This slicing condition is described by $K = 0$ in the ADM formulation. In the Ashtekar formalism, this condition is equivalent to $\dot{\tilde{E}}_a^i = 0$. It consists of two individual conditions; $\mathcal{R}e(\dot{\tilde{E}}_a^i) = 0$ and $\mathcal{I}m(\dot{\tilde{E}}_a^i) = 0$, where we remark that the second equation is the 'second triad reality condition'. If we take real-valued lapse $\underset{\sim}{N}$ and take a detour by using complex shift vector N^i then the residual gauge freedom A_0^a is determined by this additional 'time-reversal slicing' condition. Since the

[8] We assume that the degenerate points exist only in the real manifold.

path does not cross a degenerate point directly in this approach, we do not face the finite-time passing problem.

5 Discussions

We have studied the $SO(3)$-Ashtekar formulation from the point of pursuing the dynamics of spacetime, using evolutions of time-constant slices. We examined the difference between the reality conditions on the metric and on the triad, and demonstrated that the latter condition restricts a part of the gauge freedoms [$\Re e(\mathcal{A}_0^a)$]. When we apply this condition in time evolution problems (based on 3+1 decompositions), this restriction of gauge variables must be imposed at every time step. Having assumed the triad reality condition, we find a new variable, allowing us to solve $\mathcal{C}_G$ and the reality conditions simultaneously. Our variable clarifies the meanings of the additional constraint and the second reality condition, which express the symmetry and reality of the extrinsic curvature, respectively.

In the second part, we considered mechanisms for passing a degenerate point, which is one of the advantages in the Ashtekar formulation, and showed that a deformed slice approach enables us to 'pass' such points. Here we comment on a case in which a timelike degenerate area exists, like a black hole formed by a dyamical collapse of spacetime. To avoid such a degenerate line (or plane), we may make a deformed slice continuously for every step in time evolution assuming an asymptotic reality only in the spatially far region. However, we do not know yet whether we can avoid such a curvature singularity by deforming a slice into the complex plane. We can not reject the possibility that a degenerate area spreads over a deformed slice also. A detailed discussion will be reported elsewhere.

We thank Paul Haines and Richard Easther for their careful reading of our manuscript. This work was supported partially by the Grant-in-Aid for Scientific Research Fund of the Ministry of Education, Science, Sports and Culture No. 07854014, by the Grant-in-Aid for JSPS Fellow (No. 3327) and by a Waseda University Grant for Special Research Projects.

References

[1] A. Ashtekar, Phys. Rev. Lett. **57**, 2244 (1986); Phys. Rev. D **36**, 1587 (1987); *Lectures on Non-Perturbative Canonical Gravity* (Singapore, World Scientific, 1991).

[2] G. Yoneda and H. Shinkai, to be published in Class. Quantum Grav., WU-AP/47/95, gr-qc/9602026.

[3] A. Ashtekar, J.D. Romano, and R.S. Tate, Phys. Rev. **D40**, 2572 (1989).

[4] R. Capovilla, T. Jacobson, and J. Dell, Phys. Rev. Lett. **63**, 2325 (1989); Class. Quantum Grav. **8**, 59 (1991).

[5] D.C Salisbury, L.C. Shepley, A. Adams, D. Mann, L.Turvan and B. Turner, Class. Quantum Grav. **11**, 2789 (1994).

[6] H. Shinkai and G. Yoneda, in preparation.

How Can We Understand a Small But Nonzero Cosmological Constant?

Yasunori Fujii[1]
Nihon Fukushi University, Handa, Aichi, 475 Japan
and
ICRR, University of Tokyo, Tanashi, Tokyo, 188 Japan

The talk was given based on the two papers:

Y. Fujii, How natural is a small but nonzero cosmological constant, to be published in Astroparticle Physics, also available as gr-qc/9508029;

Y. Fujii, A Mechanism for a small but nonzero cosmological constant, to be published in the Proceedings of RESCEU Symposium, *The Cosmological Constant and the Evolution of the Universe*, November 7-10, 1995, University of Tokyo, also available as gr-qc/9602030,

which are herewith inlcuded in the following pages.

[1]fujii@handy.n-fukushi.ac.jp

How natural is a small but nonzero cosmological constant?

Yasunori Fujii*
Nihon Fukushi University, Handa, Aichi, 475 Japan
and
Institute for Cosmic Ray Research, University of Tokyo, Tanashi, Tokyo, 188 Japan

Abstract

Based on our previous attempt, we propose a better way to understand a small but nonzero cosmological constant, as indicated by a number of recent observational studies. We re-examine the assumptions of our model of two scalar fields, trying to explain the basic mechanism resulting in a series of mini-inflations occuring nearly periodically with respect to $\ln t$ with t the cosmic time. We also discuss how likely the solution of this type would be, depending on the choice of the parameters.

A growing number of different observations, notably the recent determination of the Hubble constant H_0 [1], seem to point to a suggestion that there is a possible *small but nonzero cosmological constant* Λ [2], with $\Omega_\Lambda \equiv \Lambda/\rho_{cr} \lesssim 1$, where $\rho_{cr} = (3/8\pi G)H_0^2$. This may not, however, be readily acceptable from a theoretical point of view, because introducing Λ has been considered to be highly *ad hoc*. Contrary to this long-held prejudice, on the other hand, it is widely recognized that a cosmological constant is an indispensable ingredient in many of the theoretical models of unification.† Unfortunately, they tend to predict Λ larger than the observed value, or its upper bound, by as much as 120 or so orders of magnitude. One of the possible ways out is to devise a theory in which the cosmological constant is not a true constant but decays like $\sim t^{-2}$, with t the cosmic time [4,5].

Notice that the theoretically natural size of Λ is of the order one in the Planckian unit system with $8\pi G = 1$,‡ while the present age of the Universe $t_0 \sim 10^{10}$y being of the order of 10^{60}. In this scenario of "a decaying cosmological constant," *today's cosmological "constant" is small* $\lesssim 10^{-120}$ *only because our Universe is old.* No unnaturally extreme fine-tuning of parameters is called for.

The scenario has been often formulated based on the models in which a scalar field plays a role; the time-dependent *effective* cosmological constant $\Lambda_{eff}(t)$ is in fact the energy density of this scalar field that couples to the ordinary matter only as weakly as gravity.

*E-mail address: ysfujii@tansei.cc.u-tokyo.ac.jp

†Some authors [3] assume $\Lambda = 0$ in the starting Lagrangian. This is protected, however, by no known symmetries against any likely perturbations.

‡By also choosing $\hbar = c = 1$, the units of length, time and the energy are 8.09×10^{-33}cm, 2.70×10^{-43}sec, and 2.44×10^{18}GeV, respectively.

This model implies, however, a complete *absence* of the cosmological *constant*. What is indicated by the observation is, on the contrary, the *presence* of a flat portion in the energy density as a function of t, a deviation from a smooth falling-off $\sim t^{-2}$. We came across, however, to a model which, by introducing another scalar field, would result in occasional flattenings of $\Lambda_{\rm eff}(t)$ thanks to a nonlinear nature of the cosmological equations [6]. The purpose of this note is to provide a simpler understanding of the mechanism.

It seems appropriate here to state our attitude. In view of the lack of the complete theoretical framework to derive all the details of the final results, we follow a heuristic approach, trying to see what the effective theory in 4 dimensions should be like, if it is to fit to what appears to be the effect of a small but nonzero cosmological constant. As it turns out, this is highly nontrivial, if the model is somehow related to modern unification theories. We list some of the main assumptions in Ref. [6].

First we assume the presence of a scalar field ϕ of the *dilaton*-type, having a non-minimal coupling, which is chosen, for the sake of simplicity, to be a Brans-Dicke type; $\phi^2 R$.[§] We then apply a conformal transformation (Weyl rescaling) to remove the non-minimal coupling. We do this for the technical convenience, at the moment, though the correct conformal frame (CF) should be selected according to what clock we use to describe the evolution of the Universe. In this connection we should notice that none of the realistic theories of gravity is conformally invariant, and that our conclusion remains true also in the original CF in which the non-minimal coupling is present.

As an important consequence of this transformation the Λ term in the original CF is converted to a potential of ϕ of the type $\Lambda e^{-\sigma/\kappa}$ where κ is a constant[¶] while σ is a transformed scalar field appropriate in the new CF. The σ field rolls down the slope toward infinity, ensuring the $\Lambda_{\rm eff}$, essentially the energy of σ, to fall off like $\sim t^{-2}$ after the inflationary era.

As a next step we introduce another scalar field Φ which has a specific interaction with σ but couples to conventional matter fields also as weakly as the gravitational interaction. We discovered an example of the interaction such that $\Lambda_{\rm eff}$, which is now the total energy density of the σ-Φ system, shows a repeated occurrence of leveling-off superimposed on the overall smooth fall-off $\sim t^{-2}$, coming barely to exceed the normal matter density, hence acting as a cosmological constant. In accordance with this the scale factor $a(t)$ exhibits an extra acceleration deviating from the overall smooth behavior $\sim t^{1/2}$ or $\sim t^{2/3}$. A typical solution of the cosmological equations was shown in Fig. 3 of Ref. [6].

The basic cosmological equations are ($k = 0$) [6]

$$3H^2 = \rho_{\rm s} + \rho_{\rm m}, \tag{1}$$

[§]Our ϕ is, following the standard notation in the conventional relativistic field theory, related to the original notation φ in Ref. [7] by $\varphi = \phi^2/8\omega$.

[¶]See Ref. [5] on how κ is constrained in order for the results to be consistent with realistic cosmology.

$$\ddot{\sigma} + 3H\dot{\sigma} - \exp(-\sigma/\kappa)\left[\frac{\Lambda}{\kappa} + \frac{1}{2}m^2\Phi^2\left(\frac{U}{\kappa} - \frac{dU}{d\sigma}\right)\right] = 0, \tag{2}$$

$$\ddot{\Phi} + 3H\dot{\Phi} + \exp(-\sigma/\kappa)m^2U\Phi = 0, \tag{3}$$

where

$$\rho_s = \frac{1}{2}\dot{\sigma}^2 + \frac{1}{2}\dot{\Phi}^2 + V, \tag{4}$$

and

$$V(\Phi,\sigma) = \exp(-\sigma/\kappa)\left[\Lambda + \frac{1}{2}m^2\Phi^2U(\sigma)\right], \tag{5}$$

with

$$U(\sigma) = 1 + B\sin(\omega\sigma). \tag{6}$$

Here κ, m, B and ω are constants. The exponential factor $\exp(-\sigma/\kappa)$ ($\sim \phi^{-4}$) comes typically from the Weyl rescaling, thus transforming the Λ-term into the potential, as mentioned before. The sinusoidal dependence in (6), on the other hand, has been introduced on the try-and-error basis. We later discuss how this specific form is favored. It is crucial to assume that the conformal transformation property of Φ is such that the same factor $\exp(-\sigma/\kappa)$ appears in front of the $\Phi^2U(\sigma)$ term.

For the matter density ρ_m we assume the mixture of relativistic and non-relativistic matters:

$$\rho_m = \rho_r a^{-4} + \rho_{nr}a^{-3}. \tag{7}$$

We admit that the result depends heavily on these assumptions. In view of the huge discrepancy of 120 orders of magnitude between conventional theory and the observation, however, variety of models seem to deserve consideration as working hypotheses. Notice that all of our parameters are essentially of the order one in Planckian units, appealing to *theoretical naturalness*.

One might still argue that we are introducing as much as what we want to come up with. We emphasize, however, that it is far from trivial to make a right choice on what to be introduced; otherwise it will not make sense no matter how much we bring in.

On using the new time variable τ defined by

$$\tau \equiv \ln t,$$

and also defining $b(\tau)$ by

$$a = e^b,$$

eqs. (1)-(3) are put into

$$3b'^2 = \frac{1}{2}\sigma'^2 + \frac{1}{2}\Phi'^2 + t^2(V + \rho_m), \tag{8}$$

$$\sigma'' + (3b' - 1)\sigma' - t^2\exp(-\sigma/\kappa)\left[\frac{\Lambda}{\kappa} + \frac{1}{2}m^2\Phi^2\left(\frac{U}{\kappa} - \frac{dU}{d\sigma}\right)\right] = 0, \tag{9}$$

$$\Phi'' + (3b' - 1)\Phi' + t^2\exp(-\sigma/\kappa)m^2U\Phi = 0, \tag{10}$$

where $'$ means differentiation with respect to τ. Notice the explicit occurrence of the time variable t^2 on the right-hand sides.

It is also worth noticing that (8)-(10) allow the asymptotic solution for $t \to \infty$ if U is chosen to be constant, namely $B = 0$:

$$a(t) = t^{2/3}, \quad \text{or} \quad b(\tau) = \frac{2}{3}\tau, \tag{11}$$

$$\sigma(t) = 2\kappa \ln\left(\sqrt{\frac{\Lambda}{2}}\frac{t}{\kappa}\right), \tag{12}$$

$$\Phi(\tau) = Ae^{-\tau/2}\sin(\tilde{m}\tau), \tag{13}$$

where A is an integration constant while

$$\tilde{m} = \sqrt{\frac{2\kappa^2 m^2}{\Lambda} - \frac{1}{4}}.$$

We learn that the Φ field, if decoupled from σ, would play no role in the asymptotic era and that τ might be a useful variable. From the solution (12) also follows that the combination

$$F(t,\sigma) \equiv t^2 \exp(-\sigma/\kappa)$$

tends to a constant ($= 2\kappa^2/\Lambda$) if $U = 1$. This implies that with non-constant $U(\sigma)$ this combination might be nontrivial. We show that this is indeed the case.

A typical solution obtained numerically is shown in Fig. 1, another example with parameters somewhat different from those used in Fig. 3 in Ref. [6]. Notice that we chose the "initial time" $t_1 = 10^{10}$, because, though the real classical cosmology had begun much earlier, we can conveniently avoid more details on the inflation era and the ensuing reheating process.

We first find in the plot (a), as already alluded, the scale factor $a(t)$ shows a series of "mini-inflations," each implying a rapid increase lasting during an opoch which is "short" in terms of τ, but could be quite "long" if it is measured in the ordinary time t; nearly as comparable as t itself. One of such epochs has been chosen to include the present time with $\lambda \equiv \log t \approx 60$, corresponding to $t \approx 10^{10}$y. (The scale factor resumes a usual expansion immediately beyond the frame.)

In the plot (b) we notice that ρ_s, the total energy of the σ-Φ system, and ρ_m, the ordinary matter energy density, fall off like $\sim t^{-2}$ as an over-all behavior, but interwinding each other. A closer look reveals that a mini-inflation occurs whenever ρ_s exceeds ρ_m.

Fig. 2 is the same plot as Fig. 1 (b) presented in a magnified scale around the present time; showing that ρ_s surpasses ρ_m, remaining nearly constant for a while, hence *imitating a cosmological constant with the size basically of the same order of magnitude as* $t^{-2} \sim \rho_{cr}$.

We also notice in Fig. 1(a) that each of these anomalous behaviors takes place toward the end of the "dormant period," during which both of the scalar fields come almost to standstill. To understand this behavior we first point out that the dormant period and its repetition are primarily due to the dynamics of the σ-Φ system; the "back-reaction" from the cosmological expansion has a rather minor effect.

Fig. 3 shows an example in which the cosmological part is cut off with the same parameters but with $3b' - 1$ replaced by 0.5. In spite of some differences, which, representing how much the cosmological effect could be, will be discussed shortly, comparing Fig. 1(a) and Fig. 3 is sufficient to convince ourselves that the "recycled dormant periods" could take place even without cosmological effect. From this point of view, we now focus upon more detailed analysis of the solution in the *isolated* σ -Φ system as a simplified mathematical model.

The initial value $\sigma_1 = 6.75442$ implies that $\omega\sigma_1 = 2\pi \times 10.75$; the σ field starts at one of the potential minima as given by $\sin(\omega\sigma)$ if viewed in the σ direction (see Fig. 2 of Ref. [6]), but on the slope in the direction of Φ. Also the "initial time" chosen to be $t_1 = 10^{10}$ implies $2\tau_1 = 46.052$, and hence $F(t_1, \sigma_1) = t_1^2 \exp(-\sigma_1/\kappa) = \exp(2\tau_1 - \sigma_1/\kappa) = \exp(3.303) = 27.19$, which is quite large. As a result, σ is pushed forward strongly. In this sense the system started from a "catapulting stage." The rapid increase of σ, however, makes $F(t, \sigma)$ small, as will be found by comparing the curve of σ and the straight line $2\kappa\tau$, also shown in Fig. 3. Soon σ is nearly free, going further until it is decelerated by the frictional term $3b' - 1 \approx 0.5$, finally to be trapped to another minimum of $\sin(\omega\sigma)$.

On the other hand, the Φ field, having pearched on the middle of the potential slope, is also catapulted downward, passing the central valley $\Phi = 0$ past, until the force dwindling again due to the decrease of $F(t, \sigma)$ and the cosmological friction stop it to an almost complete halt, hence the dormant period. The energy density ρ_s still continues to decrease according to $\sim t^{-3}$ before it stays constant.[‖]

Now with a virtually unchanging σ, the increasing τ makes $F(t, \sigma)$ non-negligible again, bringing the system back to the catapulting stage from which we started before. In this way the dormant period may repeat itself nearly periodically with respect to τ (instead of t), if the field configurations match sufficiently close to the previous values.

The "recycling," however, may fail if the matching turns out incomplete. Fig. 3 is in fact one of the patterns of such "short" recycling encountered most commonly. Toward the end of a dormant period, σ is "released" off the track before it is kicked hard by the force which has been building up. The system enters into an asymptotic behavior in which σ grows linearly while Φ decreases slowly toward $\Phi = 0$. If the same behavior occurs in the cosmological setting, the scale factor increases smoothly, resulting in no effective cosmological constant today.

On the other hand, much "longer" recycling is also rather common, as shown in Fig. 4. Suppose a recycling process lasts sufficiently long in the cosmological system. Then, as in the isolated σ-Φ system, ρ_s would stay nearly constant

[‖]See also Ref. [8] for a similar behavior.

toward the end of the dormant period, surpassing ρ_m, hence playing the role of a cosmological constant. When the scalar fields start moving as the factor $F(t,\sigma)$ increases, however, ρ_s begins to decrease, eventually nosediving beneath ρ_m. The Universe resumes an ordinary expansion again. This explains the behaviors shown in Figs. 1-2.

A question then arises how likely the solutions of sufficiently long recycling could be. An idea on the answer may be obtained again by studying the isolated σ-Φ model. We surveyed solutions of the isolated σ-Φ system by changing one of the initial values, Φ_1 for example, keeping other parameters and initial values fixed.** Solutions with shorter recycling exhibit basically the same patterns as in Fig. 3, while Fig. 4 is an example of sufficiently long recycling. Combining the solutions, we plotted in Fig. 5, the time of the end of recycling, t_e, against Φ_1 varied discretely. In spite of apparently rampant variation, we obtain solutions of long recycling ($\lambda_e > 62$) for 8 out of 23 choices of Φ_1. This, together with other limited but similar examples, seems to be an encouraging sign that the occurrence of the continued recycling is reasonably likely. The same "optimistic" view applies also if the cosmological effects are fully included.

We emphasize that the presence of minima of the potential as a funciton of σ is crucial. The leveling-off behavior is triggered by trapping σ. The form $\sin(\omega\sigma)$ is favored because it is ready to trap σ virtually at any time. Any other potential will probably be acceptable, from a "phenomenological" point of view, if it shares this property.

We have shown that a small but nonzero cosmological constant as required by the observations could result rather naturally due to a nonlinear nature of the scalar field equations. As a favored coupling, we tentativley suggested a Sine-Gordon-like interaction, though its real origin is yet to be discussed [6]. Furthermore having introduced two scalar fields, we have too many parameters, including the initial conditions, to allow unique predictions, or even a systematic survey of the solutions. In this sense our conclusion is still preliminary. The example in Figs. 1-2 merely illustrates how the results *can* be realistic, leaving many details yet to be worked out. We nevertheless have promising indications that the desired result comes about quite likely.

We certainly have to adjust some of the parameters in order to bring a mini-infaltion to the epoch including the present time, for example. The extent to which we are supposed to fine-tune them is rather *mild*, however.†† In other words, detailed analyses of various cosmological parameters at present, like H_0, t_0, Ω_Λ and q_0, will serve to constrain the parameters of the theory, as will be discussed elsewhere.

As one of the generic consequences of the present mechanism, we anticipate some past epochs to have emerged with significant amount of Ω_Λ. In the solution

**Presentation of the result has been made simpler by choosing parameters different from those used in Figs. 1-3, but the same as in Fig. 4 except for Φ_1.

††Changing Φ_1 from 0.212 to 0.210 in the solution of Figs. 1-2 would shift $\Omega_\Lambda = 0.67$ at $\lambda = 60.15$ ($t = 1.21 \times 10^{10}$y) to 0.71. However, $\Phi_1 = 0.200$ gives $\Omega_\Lambda \sim 1 - 10^{-20}$. On the other hand, $\Phi_1 = 0.2115$ yields $\Omega_\Lambda = 0.73$, H_0 =79 km/sec/Mpc at $t = 1.5 \times 10^{10}$y.

of Figs. 1-2, for example, we find $\Omega_\Lambda \gg 1$ for $27 \lesssim \lambda \lesssim 39$, while we have avoided the same at the epoch of primordial nucleo-synthesis ($\lambda \sim 45$). This may illustrate how the parameters and the initial conditions can be determined, in principle, also by looking into details of the past cosmological histories, which should deserve future studies.

We confined ourselves to the "primordial" cosmological constant prepared in the starting Lagrangian. It is yet to be seen if the same mechanism applies successfully to the vacuum energies associated with cosmological phase transitions at later times.

References

1. W. Freedman *et al.*, Nature **371**, 757 (1994). See also, M. Fukugita, C.J. Hogan and P.J.E. Peebles, Nature, **366**, 309 (1993).

2. See, for example, J.P. Ostriker and P.J. Steinhardt, Cosmic Concordance, and papers cited therein.

3. See, for exmaple, M. Fukugita and T. Yanagida, Model for the Cosmological Constant, YITP/K-1098 (1994).

4. A.D. Dolgov, in *The Very Early Universe*, Proceedings of Nuffield Workshop, England, 1982, edited by G.W. Gibbons and S.T. Siklos (Cambridge University Press, Cambridge, England, 1982): L.H. Ford, Phys. Rev. **D35**, 2339(1987).

5. Y. Fujii and T. Nishioka, Phys. Rev. **D42**, 361(1990).

6. Y. Fujii and T. Nishioka, Phys. Lett. **B254**, 347(1991).

7. C. Brans and R.H. Dicke, Phys. Rev. **124**, 925(1961).

8. Y. Fujii, M. Omote and T. Nishioka, Prog. Theor. Phys. **92**, 521(1994).

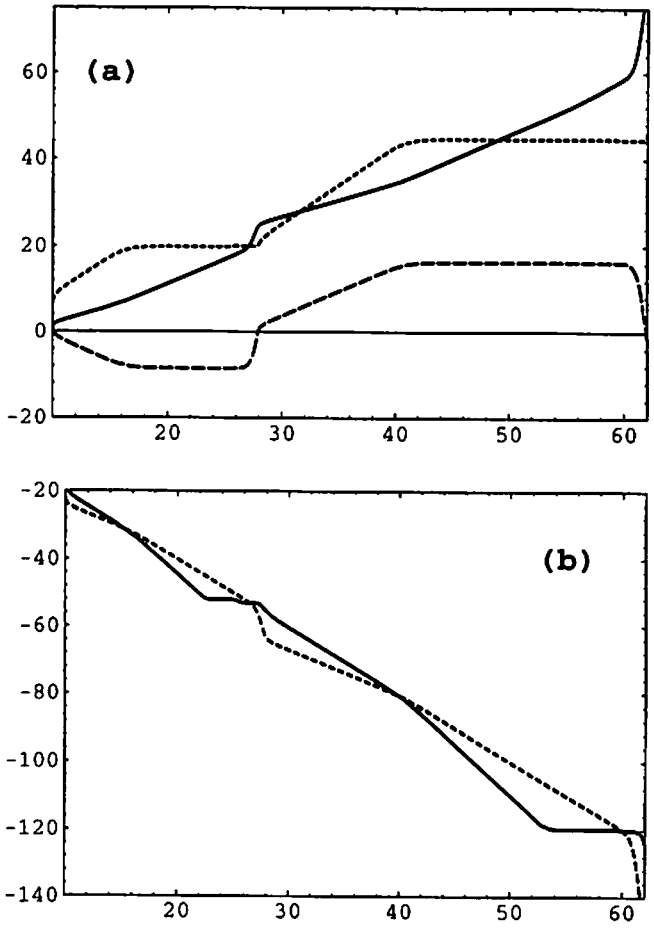


Figure 1: An example of the solution of (8)-(10). (a) Upper plot: $b = \ln a$ (solid), σ (dotted) and 2Φ (broken) are plotted against $\lambda = \log t = 0.434\tau$. The present age of the Universe supposed to be $(1.0\text{-}1.5)\times 10^{10}$y corresponds to 60.0-60.2 of λ in units of the Planck time. The parameters were chosen to be $\Lambda = 1, \kappa = 0.158, m = 4.75, B = 0.8, \omega = 10$ in the Planckian units. The initial values chosen conveniently at $t_1 = 10^{10}$ are $a = 1, \sigma_1 = 6.75442, \dot{\sigma}_1 = 0, \Phi_1 = 0.212, \dot{\Phi}_1 = 0, \rho_{r1} = 2.04 \times 10^{-21}, \rho_{nr1} = 4.46 \times 10^{-44}$; the last two being adjusted to give the "equal time" $\lambda_{eq} \sim 55$. The value of σ_1 corresponds to starting at a minimum of $\sin(\omega\sigma)$. (b) Lower plot: ρ_s (solid), the total energy density of σ and Φ, and ρ_m (dotted), the matter energy density, against $\lambda = \log t$.

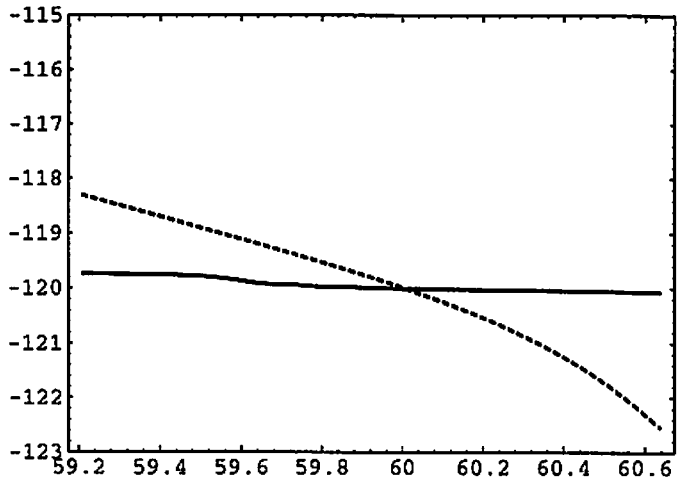


Figure 2: The same plot as in Fig. 1(b) but in a magnified scale of λ around the present time. We find $\Omega_\Lambda = 0.67$ and $H_0 = 81\text{km/sec/Mpc}$ at $\lambda = 60.15$ ($t = 1.21 \times 10^{10}\text{y}$).

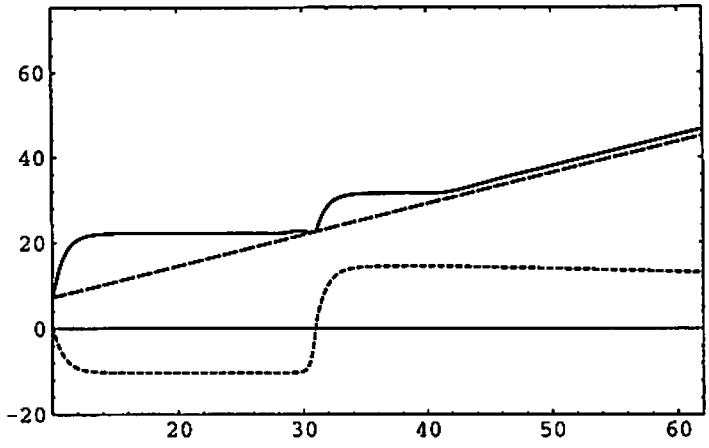


Figure 3: An example of the solution in the isolated σ-Φ system, in which recycling of the dormant periods ends prematurely at $\lambda \approx 41$. σ (solid) and 2Φ (dotted) are shown against $\lambda = \log t$. Also shown is $2\kappa\tau$ (broken) to be compared with σ. All the parameters remain the same as in Figs. 1-2, except for $3b' - 1$ replaced by 0.5.

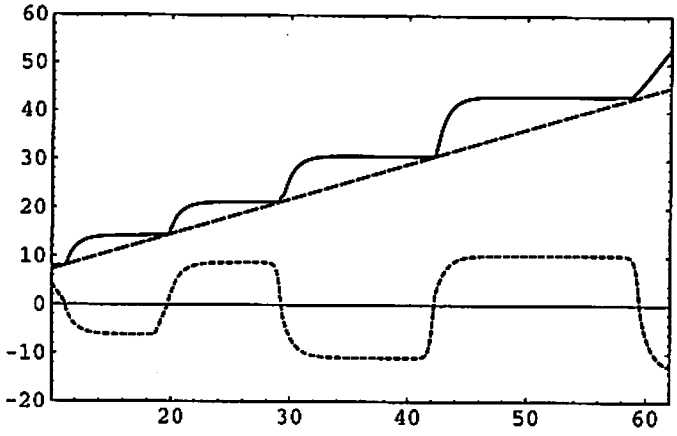


Figure 4: An example of the solution in the isolated σ-Φ system, showing long recycling. We choose $m = 5.0$ with other parameters as well as the symbols the same as in Fig. 3. The initial values at $\lambda_1 = 10$ are $\sigma_1 = 8.0, \dot{\sigma}_1 = 0, \Phi_1 = 2.1, \dot{\Phi}_1 = 0.19$, somewhat different from those in Fig. 3.

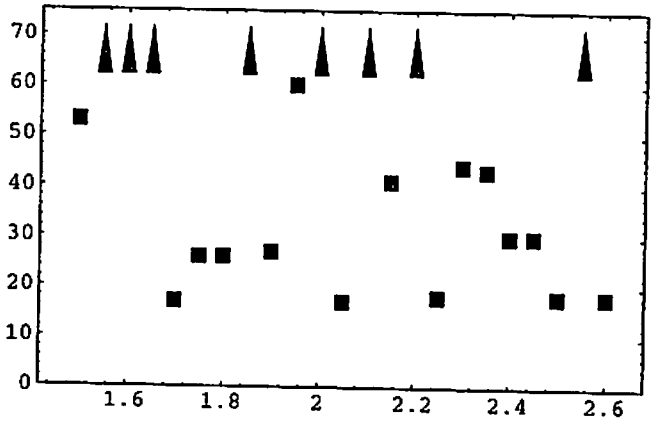


Figure 5: The time λ_e for the end of recycling in the isolated σ-Φ system is plotted against one of the initial values Φ_1, varied with spacing 0.05. The other initial values and the parameters are the same as in Fig. 4. The arrows indicate $\lambda_e > 62$.

A MECHANISM FOR A SMALL BUT NONZERO COSMOLOGICAL CONSTANT

Yasunori FUJII
Nihon Fukushi University, Handa, 470-32 Japan, and
Institute for Cosmic Ray Research, University of Tokyo, Tanashi, Tokyo, 188 Japan

I start with assuming that we do have a cosmological constant which is nearly as large as the critical density [1];

$$\Omega_\Lambda \equiv \frac{\Lambda}{3H_0^2} \sim 1.$$

How can it be so in the context of unified theories? This is the cosmological constant problem, which is in fact a 2-fold riddle:

• In almost any of the unification theories, we have a nonzero cosmological constant. Unfortunately the observed value of Λ is about 120 orders smaller than what we expect naturally from a theoretical point of view. This is the first part of the 2-fold riddle. If the above result on the size of Λ were only an upper bound, it might be sufficient to invent a theory in such a way that the cosmological constant goes away entirely at least today. An example is the simplest version of the decaying cosmological constant scenario [2].

• Suppose, however, Λ is in fact nonzero finite. Then we would face a tougher problem. We have so small a number that we find it almost impossible to keep it undisturbed by any perturbation no matter how small. This is the second part of the riddle.

The cosmological constant seems to be a very tiny but potentially extremely dangerous stumbling block on the road toward unification. I will try to outline how I can evade it by proposing a theoretical model which I believe is not entirely unnatural. I support the view that there is a final theory that unifies everything at the Planck scale. But I do not like to base my argument on the exact details of unified theories which are still yet to be fully worked out. Instead I start with assuming an effective theory in 4 dimensions, expecting in particular that some scalar fields would play an important role.

In many of the models of unified theories we find scalar fields which couple to the ordinary matter as weakly as gravitation; they are different from Higgs fields having stronger interactions. An example is the so-called dilaton field. Another example is the scalar field representing a size of internal compactified space, a remnant of higher-dimensional spacetime. These scalar fields are characterized by the "nonminimal" coupling. For the sake of illustration I give a simplified Lagrangian in 4-dimensions, as a reasonably good starting theory:

$$\mathcal{L} = \sqrt{-g}\left(\frac{1}{2}F(\phi)R - \frac{1}{2}(\partial\phi)^2 - \Lambda + L_{\text{matter}}\right), \tag{1}$$

where ϕ is the scalar field. $F(\phi)$ multiplied with R is the nonminimal coupling, and $F(\phi) = \xi\phi^2$ is the simplest choice due originally to Brans and Dicke, with ξ a constant related to BD's ω by $\xi\omega = 1/4$ (also $\varphi_{BD} = \xi\phi^2/2$). I introduce Λ at this level of the theory. This imples that I focus upon the *primordial* cosmological constant, rather than the vacuum energies due to the cosmological phase transitions in the later epochs.

Also I use a unit system with $c = \hbar = 8\pi G = 1$. The present age of the Universe considered to be somewhere around 10 Gy is nearly 10^{60} in units of the Planck time $\sim 2.7 \times 10^{-43}$sec.

Now we have to go through some of the complications concerning conformal transformations and conformal frames. But I will skip all of them. At the moment I simply apply a particular

conformal transformation to remove the nonminimal coupling, so that the original Lagrangian is put into the new form:

$$\mathcal{L} = \sqrt{-g}\left(\frac{1}{2}R - \frac{1}{2}(\partial\sigma)^2 - \Lambda e^{-\sigma/\kappa} + L_{\text{matter}}\right). \tag{2}$$

There are a number of remarks relevant here.

Notice first the absence of the nonminimal coupling just as was designed. Secondly, as a consequence, the *canonical* scalar field in this new conformal frame is σ which is different from but is related to the original ϕ by

$$\phi = e^{\sigma/4\kappa},$$

with κ a constant given by ξ. Thirdly, the Λ-term is multiplied by $\phi^{-4} = e^{-\sigma/\kappa}$. This implies that the Λ-term now *acts as a potential of* σ. We may expect that σ would fall off the exponential slope toward infinity. This would signal that any effect of Λ will decrease with time.

That this is indeed the case can be shown by integrating the cosmological equations:

$$3H^2 = \rho_s + \rho_m, \tag{3}$$

$$\ddot{\sigma} + 3H\dot{\sigma} + \frac{dV}{d\sigma} = 0, \tag{4}$$

$$\dot{\rho}_m + 4H\rho_m = 0, \tag{5}$$

where ρ_m is the matter density, while ρ_s is the density of the scalar field;

$$\rho_s = \frac{1}{2}\dot{\sigma}^2 + V,$$

with

$$V(\sigma) = \Lambda F(\phi)^{-2}.$$

I here choose the nonminimal coupling $F(\phi) = 1 + \xi\phi^2$ which is a little more complicated than BD's original suggestion; V is now slightly flatter near the origin (see Fig. 1 of Ref. [2]). This has an effect to bring about sufficient amount of the inflationary expansion of the primordial Universe.

The numerical integration gives in fact solutions in which the scale factor $a(t)$ starts with an exponential growth but followed by the power-law behavior (see Fig. 2 of Ref. [2]). This implies that there is no trace of the truly constant Λ in the asymptotic era.

Eq. (3) suggests that $\rho_s(t)$ may be interpreted as the effective cosmological constant Λ_{eff}. We also find that this Λ_{eff} behaves asymptotically like $\sim t^{-2}$, the same behavior as ρ_m, hence giving $\Lambda_{\text{eff}} \sim 10^{-120}$ today ($t \sim 10^{60}$). This may provide an answer to the first part of the riddle, but certainly short of replying the second part. The observations seem to tell us that ρ_s must decrease like t^{-2} as an overall behavior to ensure the size $\sim 10^{-120}$, but should deviate from the smooth fall-off, hopefully as a *leveling-off* behavior, which would *imitate* the constant Λ at least locally.

But what kind of theory can do the job? Without any clear clue, we decided to introduce another scalar field, called Φ [3]. But I soon realized that merely introducing Φ is not enough; nothing spectacular will happen unless Φ is coupled to σ in a nontrivial way. Again without any useful guide, we tried several candidates based on a try-and-error basis. Then we came across to an interction of the form

$$V(\sigma, \Phi) = e^{-\sigma/\kappa}\left[\Lambda + \frac{1}{2}m^2\Phi^2 U(\sigma)\right], \tag{6}$$

with

$$U(\sigma) = 1 + B\sin(\omega\sigma), \tag{7}$$

where B, ω and m are constants. This may look rather awkward. If B were zero, the second term of (6) would be simply a mass term of the Φ field multiplied by $e^{-\sigma/\kappa}$ expected to come from the conformal transformation.

With nonzero B, we find a parabolic slope in the Φ direction, whereas V falls off exponentially in the σ direction with the oscillation superimposed (see Fig. 4 of Ref. [3]). We may expect to roll down the potential slope toward $\sigma \to \infty$ probably with some smooth meandering behavior. However, quite surprisingly, the solution shows something striking. An example is shown in Fig. 1 [4].

The horizontal axis is $\log_{10} t$ in units of Planck time, so that we are now around 60. I also gave initial conditions of the classical cosmology at t_1 which I tentativly chose $= 10^{10}$. Also as an important rule of the game, I assume that all the constants in the theory as well as the initial values of the scalar fields are essentially of the order 1 in Planckian units.

Then as we find, ρ_s behaves as we wanted to see; an overall behavior $\sim t^{-2}$ and step-like leveling-offs. Corresponding to each of them, we have a mini-inflation; a rapid but temporary rise of the scale factor. I have two of them in this example.

I have adjusted parameters such that one of the mini-inflations shows up around the present epoch, as shwon in the zoomed-up view in Fig. 2. We have $\Lambda_{\rm eff}$ which is *small today because our Universe is old, not because of a fine-tuning.* I obtained the values t_0=12.1 Gy, $h = 0.81$ and $\Omega_\Lambda = 0.67$, just as an illustration.

I will discuss more about the characteristics of the solution. We have two mini-inflations before the present epoch in this solution. This number depends on the choice of parameters, however. For somewhat different values of parameters, I have five of them, for example. We also find that the whole behavior is a *nearly cyclic repetition* of the same pattern. It is nearly periodical if we plot them against *ln t rather than t* itself. This is a highly nonlinear effect, but I will try to give intuitive explanations.

Each pattern consists basically of two phases. First a "catapulting phase," in which both of the scalar fields are driven by the forces coming from the potential V. σ is pushed forward toward infinity while Φ toward the central valley at $\Phi = 0$. The increase of σ, however, makes the potential dwindle very quickly because of the factor $\exp(-\sigma/\kappa)$. For this reason and also due to the "cosmological frictional forces" provided by $-3H\dot{\sigma}$ and $-3H\dot{\Phi}$, the scalar fields are soon decelerated.

Now the σ slowed down sufficiently is trapped by the sinusoidal potential given by $\sin(\omega\sigma)$. In this way, the system enters eventually the "dormant phase," during which both of the scalar fields come to almost complete stop, and ρ_s stays constant, hence it is as if we had a truly constant Λ. This is the key of our mechanism for a small but nonzero cosmological constant.

Interestingly, this leveling-off does not last forever. The forces which once dwindled begin to build up again, bringing the system, rather likely, back again to the catapulting phase. In this way, a set of the two phases would repeat itself, a process called "recycling."

In this connection I point out that the real origin of the recycling behavior lies in the dynamics of the system of two scalar fields coupled to eahc other; the cosmological environment plays only minor roles.

In fact in the isolated σ-Φ system, I find typical solutions as in Fig. 3, showing many repetitions indeed. If the same behavior takes place in the cosmological setting, then we would have mini-inflations, and consequently the effect of a small but nonzero Λ.

For some other parameters, however, I find solutions as in Fig. 4. Recycling ends prematually. This is because the values of the scalar fields and their time-derivatives do not match sufficiently close to the previous values. If this premature "derailing" occurs in the cosmological environment, the Universe would evolve just smoothly; no anomalous effect. These two behaviors are the typical ones though there are some other variants as well.

The question is then how likely we get sufficiently long recycling. Frankly speaking this is a hard question because I have too many parameters; in addition to the constants of the theory itself, like B and ω, I have at least four initial values of the two scalar fields. Nevertheless I

have an impression that the chances of having a long recycling are rather high.

With the same model of the isolated σ-Φ model, I varied one of the initial values Φ_1 with other values and constants held fixed. In one such attempt, I changed Φ_1 from 1.5 to 2.6 with equal spacing 0.05, obtaining 8 solutions (out of 23) showing sufficiently long recycling (see Fig. 5 of Ref. [4]). This is a fair reflection of the general trend as far as I have tried.

To conclude we add a few more comments. My story may have sounded somewhat complicated and messy. But I point out that this mechanism has something in common with what is widely known as "relaxation oscillation," which is happening in everyday life. In playing violin, for example, one moves the bow rather slowly, still producing sounds of much higher frequencies. Friction is obviously crucial. There is no reason why something should not happen in Nature simply because it is complicated. The same should be true also in the Universe.

As a generic feature of our solutions, I have several mini-inflations. On this basis, I predict backward that the Universe may have experienced several Λ-dominated epochs. This can be dangerous. Suppose a calculation gives ρ_s which is non-negligible compared with ρ_m at the time of nucleosynthesis, for example. It may jeopardize the success of the standard theory. This has to be avoided. This is in fact what I did when I selcted out the examples shown before. In Fig. 1, for example, ρ_s is kept below ρ_m at $\log_{10} t \sim 45$, though ρ_s may be dominant at other epochs. This illustrates how I can constrain the theory by studying the past history of the Universe. The same argument can be applied to the values of H_0 and Ω_Λ at the present epoch. In any case my result at this moment is still away from the goal in terms of numerical fits.

I also admit that I have made several assumptions which I myself am not sure how to derive from more fundamental theories. In this respect I emphasize that my approach is phenomenological. I am asking what the fundamental theory should be like if the cosmological constant ceases to be a problem. I hope with Weinberg [5] that challenging the cosmological constant will open up a new breakthrough in our effort toward unification.

[1] See, for example, J.P. Ostriker and P.J. Steinhardt, *Cosmic Concordance*, and papers cited therein. These authors suggest a set of representative values; $H_0 = 65$ km/sec/Mpc and $\Omega_\Lambda = 0.65$.

[2] Y. Fujii and T. Nishioka, Phys. Rev. **D42**, 361(1990), and papers cited therein.

[3] Y. Fujii and T. Nishioka, Phys. Lett. **B254**, 347(1991).

[4] Y. Fujii, How natural is a small but nonzero cosmological constant? preprint, gr-qc/9508029; to be published in Particle Astrophysics.

[5] S. Weinberg, Rev. Mod. Phys. **61**, 1(1989).

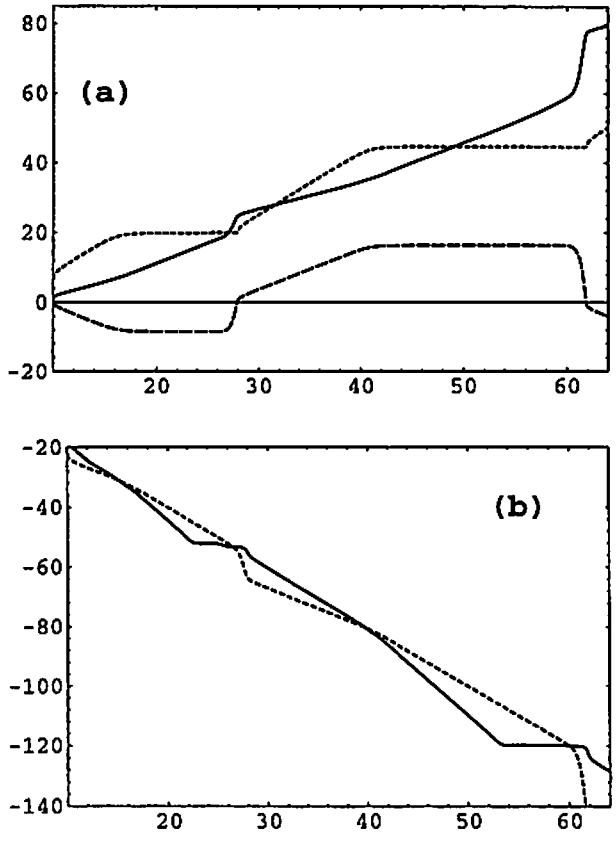


Figure 1: An example of the solution of (3)-(5). (a) Upper plot: $b = \ln a$ (solid), σ (dotted) and 2Φ (broken) are plotted against $\lambda \equiv \log_{10} t$. The present age of the Universe supposed to be $(1.0\text{-}1.5)\times 10^{10}$y corresponds to 60.0 - 60.2 of λ in units of the Planck time. The parameters were chosen to be $\Lambda = 1, \kappa = 0.158, m = 4.75, B = 0.8, \omega = 10$ in Planckian units. The initial values chosen conveniently at $t_1 = 10^{10}$ are $a = 1, \sigma_1 = 6.75442, \dot{\sigma}_1 = 0, \Phi_1 = 0.212, \dot{\Phi}_1 = 0, \rho_{r1} = 2.04 \times 10^{-21}, \rho_{nr1} = 4.46 \times 10^{-44}$; the last two being adjusted to give the "equal time" $\lambda_{eq} \sim 55$. The value of σ_1 corresponds to starting at a minimum of $\sin(\omega\sigma)$. (b) Lower plot: ρ_s (solid), the total energy density of σ and Φ, and ρ_m (dotted), the matter energy density, against λ.

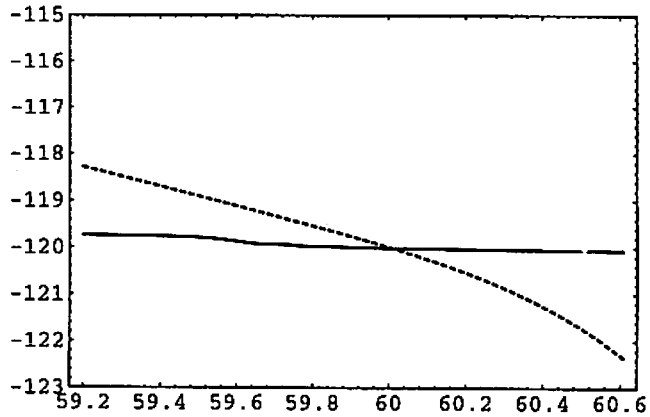


Figure 2: The same plot as in Fig. 1(b) but in a magnified scale of λ around the present time. We find $\Omega_\Lambda = 0.67$ and $H_0 = 81$km/sec/Mpc at $\lambda = 60.15$ ($t = 1.21 \times 10^{10}$y).

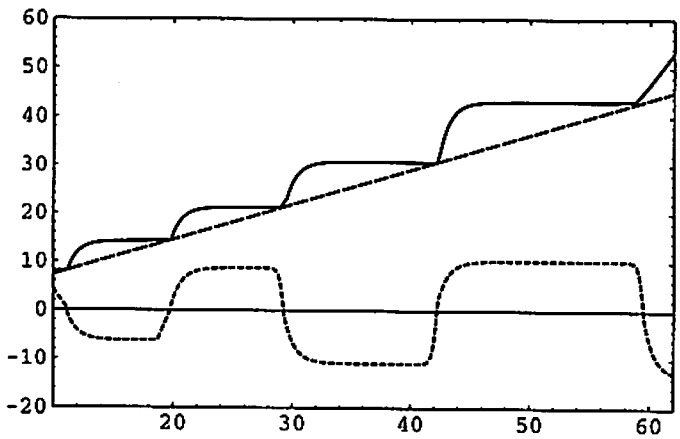


Figure 3: An example of the solutions in the isolated σ-Φ system, showing long recycling. We choose $m = 5.0$ with other constants as well as the symbols the same as in Fig. 1(a), except for $3b' - 1$ replaced by 0.5, and the added broken line for $2\kappa\tau$ to be compared with σ. The initial values at $\lambda_1 = 10$ are $\sigma_1 = 8.0, \dot{\sigma}_1 = 0, \Phi_1 = 2.1, \dot{\Phi}_1 = 0.19$.

of points of space-time. We investigate compact hyperbolic Friedmann-Robertson-Walker universe in an inflationary phase. Such models can be presented by *time* × *compact hyperbolic homogeneous hypersurfaces.*

A compact hyperbolic manifold is given as the quotient space of H^3 by the discrete subgroup Γ of its isometry group $SO(3,1)$ which acts on H^3 properly discontinuously. In general, the action of a group Γ on a manifold M must be properly discontinuous so that the quotient space M/Γ be a Hausdorff manifold. We call an action of Γ on M *properly discontinuous* if it satisfies the following conditions [3, 4];
(1) each point $q \in M$ has a neighbourhood U such that $\gamma(U) \cap U = \emptyset$ for each $\gamma \in \Gamma$ which is not the identity element, and
(2) if $q, r \in M$ are such that there is no $\gamma \in \Gamma$ with $\gamma(q) = r$, then there are neighbourhoods U and U' of q and r respectively such that there is no $\tilde{\gamma} \in \Gamma$ with $\tilde{\gamma}(U) \cap U' \neq \emptyset$.

Condition (1) implies that the quotient M/Γ is a manifold, and condition (2) implies that it is Hausdorff.

The universal covering manifold of the concerning compact universe is a hyperbolic inflationary universe, of which the maximal extension has a Cauchy horizon. When we assume that the actions of the covering transformation group analytically continue to the extended regions, we expect that the topology change takes place and closed timelike curves appear due to rearrangements of homogeneous slices by the extension and compactification. This is suggested by the work on the Lorentzian topology change of (2+1)-dimensional compact black hole geometry by Siino [2].

We naturally ask whether or not we can construct compact hypersurfaces as the quotient submanifold in the extended region because the homogeneous hypersurface becomes a Lorentzian submanifold dS^3. More precisely, does the Γ act on 3-de Sitter hypersurfaces properly discontinuously?

The Misner universe is a simple example of a compact universe and its extension, which is given as a quotient manifold in (1+1)-dimensions [3]. The maximally extended covering manifold is given by a part of 2-dimensional Minkowski space-time $(\widetilde{M} \setminus \{0\}, \tilde{\eta})$. We must subtract the point $\{o\}$ from whole the Minkowski space-time $(\widetilde{M}, \tilde{\eta})$ so that the quotient space become a manifold. Because of subtracting the point $\{o\}$ from $(\widetilde{M}, \tilde{\eta})$, all causal curves which direct to $\{o\}$ are incomplete. We define singularity as non-spacelike incomplete curve according to the definition in [3] and therefore, the maximally extended Misner universe has a singularity. In this paper, we call such type of singularity *a topological singularity.*

In case of the compact hyperbolic inflationary universe model, however, we will observe in the next section that singularities densely appear on the hypersurfaces dS^3 and the extended regions are no longer space-time manifolds. This concludes that one cannot analytically extend the space-time by extending both the universal cover and the action of $\Gamma \subset \mathrm{Isom}(M)$ on

it. If one regards this anomalous occurrence of singularities serious, one can say that the spatially compact hyperbolic inflationary universe cannot be extended beyond the Cauchy horizon. Then we have an interpretation for this anomalous phenomena such that this is an example which holds up *the strong cosmic censorship conjecture*, which, roughly speaking, states that a physical space-time is globally hyperbolic.

As the scenario about a birth of hyperbolic universe, one-bubble inflationary universe model [1] is an appealing one. However, it cannot be a covering manifold of a compact hyperbolic inflationary universe, since such a model has a past Cauchy horizon inside the bubble. Thus we conclude one-bubble inflationary hyperbolic universe scenario cannot be compatible with a compact hyperbolic inflationary universe model.

One also can attribute this anomaly to the high degrees of a given space-time symmetries. Then one can expect that, in this situation, topology change takes place owing to space-time extensions and compactifications because there is no isometry in realistic universe. In this case, physical implications of the extended regions are farther problems.

2 Anomalous occurrence of singularities

In this section, we show that an anomalous occurrence of topological singularities in the extended space-time.

To investigate what happen by the action of Γ in the extended region, for simplicity, we consider the $(2+1)$-dimensional case. Then the homogeneous isotropic hypersurfaces in the concernig universe model correspond to 2-hyperboilc space H^2 inside the Cauchy horizon and 2-de Sitter space-time dS^2 in the extended region. H^2 is imbedded in 3-Minkowski space-time $(E^{2,1}, \eta)$ as: $-(X^0)^2+(X^1)^2+(X^2)^2=-1$, by an imbedding $f : H^2 \longrightarrow E^{2,1}$ with the induced metric $g_H := f^*\eta$. dS^2 is described as; $-(X^0)^2+(X^1)^2+(X^2)^2=1$. Here the curvature radii are normalized to unity.

We assume that the action of a discrete subgroup Γ of $\mathrm{Isom}(H^2)$ on the H^2 is analytically continued to the extended regions and acts on the Lorentzian manifolds dS^2. When there exist at most finite number of accumulation points of $\{\gamma^n(x)\}$, we can subtract these points from dS^2 and obtain a quotient manifold, $\{dS^2 - \{\text{accumulation points}\}\}/\Gamma$ as seeing in the construction of the Misner universe. Therefore, we want to know how densely the accumulation points exist.

The $\mathrm{Isom}(H^2)$ consists of boosts and space-rotations of $\mathrm{Isom}(dS^3)$. Each element of the subgroup Γ must contain one of the boosts because the element consisted of only space-rotations has fixed points. First, we consider the actions of the simplest element $\gamma_o \in \Gamma$ on dS^2, where γ_o is represented in the

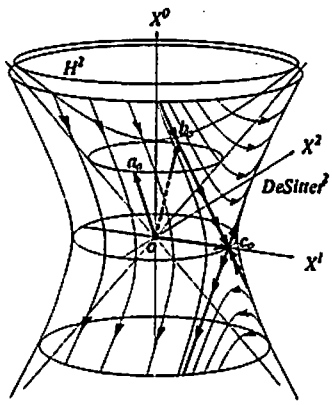


Figure 1: Geodesic flows of $\gamma_o{}^n$ are schematically depicted.

coordinate system $\{X^0, X^1, X^2\}$ as the following matrix;

$$\gamma_o(\zeta) := \begin{pmatrix} \cosh\zeta & 0 & \sinh\zeta \\ 0 & 1 & 0 \\ \sinh\zeta & 0 & \cosh\zeta \end{pmatrix}, \quad \zeta = constant. \tag{1}$$

This γ_o has three eigenvectors,

$$a_o := \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad b_o := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad c_o := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{2}$$

Two eigenvectors a_o and b_o are lying on the light cone; $-(X^0)^2 + (X^1)^2 + (X^2)^2 = 0$, and c_o points the point c_o on dS^2 described by eq.(2). Thus we observe that any points s on the null line through the point c_o converge to c_o by the action of $\gamma_o{}^n$ illustrated in Figure(1), that is, the infinite sequence $\{\gamma_o{}^n(s)\}$ has an accumulation point at the point c_o. For $^{\forall}\gamma_i(\neq \gamma_o) \in \Gamma$, $\gamma_i(c_o)$ is one of the eigenvectors of $\gamma_i \circ \gamma_o \circ \gamma_i{}^{-1} \in \Gamma$ and so the point on the dS^2 pointed by the vector $\gamma_i(c_o)$ is also an accumulation point of the sequence $\{(\gamma_i \circ \gamma_o \circ \gamma_i{}^{-1})^n(\gamma_i(s))\}$. Therefore, there exist many, countable infinite, accumulation points of $\{\gamma^n(s)\}$ on dS^2.

Now we show that the accumulation points, that are topological singularities, densely occur on dS^2. More precisely, we will show that;

If we assume that the actions of the covering transformation group Γ on H^2 are analytically continued on dS^2 in the extended region, then, for every neighbourhood O_c of an arbitrary point c on the dS^2, there exist $\gamma \in \Gamma$ and point $s \in dS^2$ such that an accumulation point c' of the infinite sequence $\{\gamma^n(s)\}$ is contained in O_c.

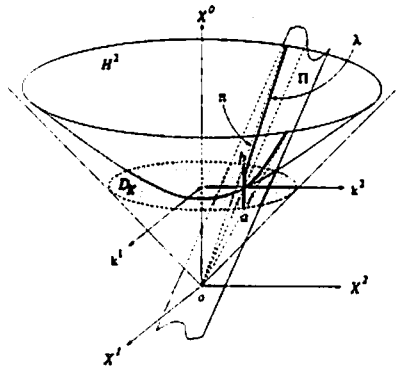


Figure 2:

We will prove this statement according to the following steps.

First, for every point c on dS^2, there exist an element $\gamma \in SO(2,1)$ such that c is one of the eigenvectors of γ.

If $\gamma \in \Gamma$, there exists $s \in dS^2$ such that c is the accumulation point of $\{\gamma^n(s)\}$. So, we only have to consider the case that $\gamma \notin \Gamma$.

The other two eigenvectors, a and b, are lying on the null cone. These two vectors uniquely determine a geodesic on the H^2. In fact, these two eigenvectors span a plane Π and the intersection of $\Pi \cap H^2$ is an geodesic curve λ in the H^2. The geodesic λ is projected to a straight line $\Pi \cap D_K$ on *Klein Disk* : $D_K := \{X^0 = 1, (X^1)^2 + (X^2)^2 < 1\}$. The projection $\pi : E^{2,1} \ni (X^0, X^1, X^2) \mapsto (k_1, k_2) \in D_K$ is defined as $k_i := \frac{X^i}{X^0}$, $k := ({k_1}^2 + {k_2}^2)^{1/2}$ with the induced metric: $ds^2 = [(1-{k_2}^2)d{k_1}^2 + (1-{k_1}^2)d{k_2}^2 + 2k_1k_2dk_1dk_2]/(1-k^2)^2$. We can identify H^2 and D_K by the diffeomorphism $\pi \circ f$. This projection π maps a geodesic curve of H^2 which is a hyperboloid in the $E^{2,1}$ to a straight line in the $E^{2,1}$ and doesn't preserve inner product.

The eigenvectors, a, b, direct to the points a, b, respectively, on the boundary $\partial D_K := \{X^0 = 1, (X^1)^2 + (X^2)^2 = 1\}$. We can see that the straight line from b to a on $\overline{D}_K := D_K \cup \partial D_K$ is the geodesic $\pi(\lambda)$.

As long as considering the case that $\gamma \notin \Gamma$, the geodesic λ does not close on H^2/Γ. Moreover, every line terminated at a and b on $\overline{D}_K$ does not close on H^2/Γ because if there exist a closed line which terminates at a and b on $\overline{D}_K$, there must exist an element of Γ corresponding to the closed curve and it is just γ itself, but now $\gamma \notin \Gamma$.

It can be observed that $X := \partial D_K \times \partial D_K \setminus \Delta_{\partial D_K}$ is one-to-one correspondence with the set of all geodesics on D_K, where $\Delta_{\partial D_K} := \{(p,p) | p \in \partial D_K\}$. Namely, a pair $(a, b(\neq a)) \in X$ can be identified to the geodesic on D_K terminated at $a, b \in \partial D_K$.

It is also confirmed that the positions of each eigenvector are closely

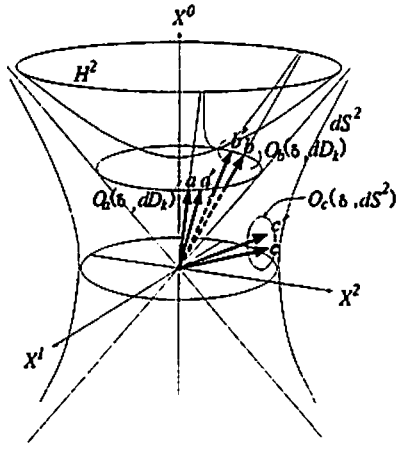


Figure 3:

related to each other. [4] Therefore, one can correspond an unique geodesic λ on H^2 to the point c on the dS^2. [5]

Second, choose a Riemannian metric $\mathrm{e} := (dX^0)^2 + (dX^1)^2 + (dX^2)^2$ on $E^{2,1}$. Take an open ball $B(c, \delta, E^{2,1}) := \{x \in E^{2,1} | \|x - c\|_{\mathrm{e}} < \delta\}$ with radius δ respected to e and define a neighbourhood of c on the dS^2 as $O_c(\delta, dS^2) := B(c, \delta, E^{2,1}) \cap dS^2$. Similarly, define neighbourhoods of a and b respectively on the ∂D_K as $O_a(\delta, \partial D_K) := B(a, \delta, E^{2,1}) \cap \partial D_K$ and $O_b(\delta, \partial D_K) := B(b, \delta, E^{2,1}) \cap \partial D_K$ corresponding to $O_c(\delta, dS^2)$ (see Figure (3)).

Now, what we want to show is reduced to the following lemma;

Lemma
$^\forall(a, b) \in X, ^\forall\delta(> 0)$,

[4]For example, choosing an appropriate coordinate system and a constant α, the eigenvector c can be represented as

$$c = \begin{pmatrix} \sinh\alpha \\ \cosh\alpha \\ 0 \end{pmatrix}. \tag{3}$$

Then, the other two are uniquely determined as

$$a = \begin{pmatrix} \cosh\alpha \\ \sinh\alpha \\ -1 \end{pmatrix}, \quad b = \begin{pmatrix} \cosh\alpha \\ \sinh\alpha \\ 1 \end{pmatrix}. \tag{4}$$

[5]If we constrain the point c on the part $(\{X^0 > 0\} \cup \{X^0 = 0, (X^1)^2 + (X^2)^2 = 1, X^1 > 0\} \cup \{X^0 = X^1 = 0, X^2 = 1\})$ of dS^2, the point c has one-to-one correspondence with a pair of points (a, b) on D_K.

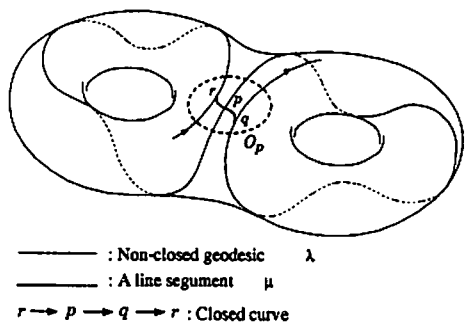


Figure 4: A compact Riemann surface with genus $p = 2$ is illustrated. λ is a non-closed geodesic and α is a line segment which is connecting two points q and r on the λ. Then, one can see a closed line $\{r \to p \to q \to r\}$ composed of the line segment $\{r \to p \to q\}$ of the geodesic λ and the line segment μ : $\{q \to r\}$.

$^{\exists}a' \in O_a(\delta, \partial D_K), ^{\exists} b' \in O_b(\delta, \partial D_K), ^{\exists}\gamma \in \Gamma, ^{\exists} p \in D_K$
s.t. $\lim_{n\to\infty} \gamma^n(p) =: a' \in O_a(\delta, \partial D_K) \lim_{n\to-\infty} \gamma^n(p) =: b' \in O_b(\delta, \partial D_K)$

We show that there exists a closed line on the compact Riemann surface H^2/Γ which directs to points $a' \in O_a(\delta, \partial D_K)$, $b' \in O_b(\delta, \partial D_K)$ on ∂D_K (see Figure (5) and Figure (4)). We can do make such a closed line by using a non-closed geodesic $\lambda(v) := \Pi \cap D_K$ affinely parameterized by v and ergodicity of geodesic flow on a compact Riemann surface(see Appendix). In Figure (4), a non-closed geodesic curve λ on a H^2/Γ is illustrated. Let p be a point on λ at $v = 0$ and take an arbitrary small open desk O_p whose radius $\epsilon \ll 1$ as a neighbourhood of p. From the ergodicity of geodesic λ, we can always find points $q = \lambda(l), l \gg 1$ and $r = \lambda(k), k < 0, |k| \gg 1$ on λ such that each of them is contained in O_p and the tangent vectors of λ at these points r, p, q, are sufficiently parallel to each other as depicted in Figure (4). Connecting the points r and q by a suitable line segment μ, we obtain a closed line $\{r \to p \to q \to r\}$ on the H^2/Γ. Corresponding to this closed line, there exists an element $\gamma \in \Gamma$.

It is shown in Figure (5) that the images of O_p, p, q, r, $\lambda \cap O_p$ and μ by γ on the Klein desk D_K which is the universal cover of H^2/Γ. On D_K, there are as many images of O_p as the number of the elements of Γ. Each of the images of O_p which contains the portion of the geodsic $\pi(\lambda)$ can be transformed each other by γ^n. Since ϵ is determined by the metric g_H on H^2, so the radii of its images on the Klein desk which are determined by the Euclid metric $\mathbf{e}$ on R^2 become smaller and smaller by the projection π, far and far the image aparts from the origin $\{k_1 = k_2 = 0\}$.

The ergodic theory of geodesic flow on H^2/Γ, we can taking the affine dis-

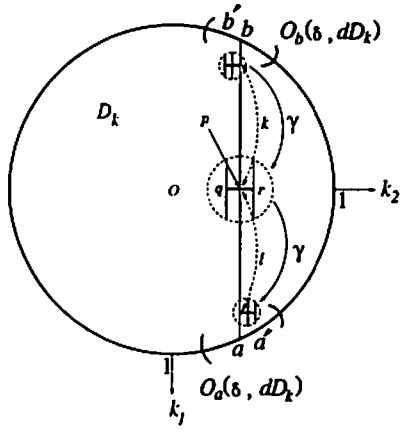


Figure 5:

tances l and k be sufficiently large so that $\lim_{n\to\infty}\gamma^n(p) =: a' \in O_a(\delta, \partial D_K)$ and $\lim_{n\to-\infty}\gamma^n(p) =: b' \in O_b(\delta, \partial D_K)$. Then, we have the two eigenvectors of γ, a' and b'. We can obtain the third eigenvector c' of γ directing to a point c' in the $\delta-$neighbourhood $O_c(\delta, dS^2)$. That is, the point c' is an accumulation point of $\{\gamma^n(s)\}$.

In $(3+1)$-dimensional case, the discrete subgroup of Isom(dS^4) spacetime has four eigenvectors. Two of them are null vectors corresponding to a and b in the $(2+1)$ case. The other two are spacelike corresponding to c and point to points on the hypersurfaces dS^3. In addition, we can also observe the ergodicity of geodesic flows on three compact hyperbolic Riemannian manifold H^3/Γ [5] (see Appendix A). Therefore, we obtain the same results as the $(2+1)$-dimensional case; topological singularities densely occur on the dS^3.

Acknowledgments

We are grateful to Professor Sadayoshi Kojima for useful advice and discussions.

Appendix: Ergodic theory [5]

Abstract Dynamical system:(M, μ, ϕ_t)
Let (M, μ) be a measure space with measure μ and $\phi_t : M \longrightarrow M$ be an one-parameter transformation group which is measure-preserving. Then a set (M, μ, ϕ_t) is called *dynamical system.*
Definition
Let f be a complex valued function on M.

The time-average: f^* is defined as

$$f^*(x) := \lim_{T\to+\infty} \frac{1}{T}\int_0^T f(\phi_t(x))dt, \quad x \in M, \quad t \in R.$$

The space-average: $\bar{f}$ is defined as

$$\bar{f} := \int_M f(x)d\mu.$$

Definition:
A dynamical system (M, μ, ϕ_t) is *ergodic* if almost everywhere $f^*(x) = \bar{f}$, for any f which is integrable with μ(i.e. $f \in L_1(M, \mu)$).

In our case, the dynamical system $(H^2/\Gamma, g_H, \Gamma)$ is ergodic. One can obtain the following statement from ergodicity of $(H^2/\Gamma, g_H, \Gamma)$.
Let $\alpha_x : [0, \infty) \ni\longmapsto \alpha_x(\tau) \in H^2/\Gamma$ be a geodesic determined by the starting point $\alpha_x(0) = x \in H^2/\Gamma$ and the tangent vector $\dot{\alpha}_x(0)$. Define the length of α_x by its affine parameter τ. Let A be *a measurable subset of* H^2/Γ and I be *a total length of the intersections of* $\alpha_x(\tau)$ *and* A, i.e. $I := \bigcup_{\tau\in R}\{\alpha_x(\tau) \cap A\}$. Take $f(x), x \in H^2/\Gamma$, in the definition above, as;

$$f_\tau(x) := \frac{I}{\tau : \text{the length of } \alpha_x \text{ of interval } [0, \tau]}.$$

Then, there exists a time-average:

$$f^*(x) = \lim_{\tau\to\infty} \frac{I}{\tau},$$

and a space-average:

$$\bar{f} = \frac{\text{measure of } A}{\text{measure of } H^2/\Gamma},$$

so that $f^*(x) = \bar{f}$ for any A and almost every $(x, \dot{\alpha}_x(0))$.

Theorem(Labatchevsky-Hadamard theorem [5, p62 + p77])
Let M be a compact, connected, negative-curvature Riemannan surface with genus $p(\geq 2)$ and T_1M be a unit tangent bundle of M. Then, the geodesic flows on T_1M are ergodic.

From this theorem, if one want, one can always find two points on a non-closed geodesic on compact Riemannian manifold H^2/Γ such that each point is arbitrary close to the other and the tangents at these points are arbitrary parallel to each other.

These ergodic characters of geodesic flows on compact Riemann surface are also observed in a three-dimensional compact Riemannian manifold as

summarized following theorem;

Theorem (Auslander, Green and Hahn [6])
Let V be an n-dimensional vector space over $\mathbf{C}$ and let G be a real Lie subgroup of $GL(n, \mathbf{C})$.
Suppose that either
a) $G \supset SL(n, \mathbf{C})$, $n > 1$, or
b) $G \supset Sp(n, \mathbf{C})$, n even.
Let Γ be a discrete subgroup of G with compact factor space $\Gamma \setminus G$. Then if $v \in V$, $v \neq 0$, the orbit $\Gamma(v)$ is dense in V.

References

[1] S.Coleman and F.De Luccia, (Phys.Rev.D21, (1980), 3305); J.R.Gott III, (Nature, 295, 28,(1982), 304); J.R.Gott III, (Phys.Lett.136 B, (1984), 157); M.Sasaki, T.Tanaka, K.Yamamoto and J.Yokoyama, (Phys.Lett.B317, (1993), 510); T.Tanaka, K.Yamamoto, M. Sasaki, (Proc.4th Workshop on G.R.G (1994)); T.Tanaka and M.Sasaki, (Phys.Rev.D, 50, (1994), 6444); M.Bucher and A.S.Goldhaber, (Phys.Rev.D, 52, (1995), 3314);

[2] M.Siino, : Class. Quantum Grav., 11, 1995 (1994)

[3] S.W.Hawking and G.F.R.Ellis,: (Cambridge University Press, Cambridge, 1973)

[4] S.Kobayashi and N.Nomizu, : (Interscience Pub., (1963))

[5] V.I.Arnold and A.Avez : (Yoshioka-syoten, (1972), in Japanese)

[6] L.Auslander, L.Green, and F.Hahn : (Ann. Math. Stud., Princeton Univ. Press, (1963))

Dynamics of compact homogeneous universes

MASAYUKI TANIMOTO[†], TATSUHIKO KOIKE[♭], AND AKIO HOSOYA[♮]

† *Department of Physics, Kyoto University, Kyoto 606-01, Japan. e-mail: prince@tap.scphys.kyoto-u.ac.jp*

♭ *Department of Physics, Keio University, Kanagawa 230, Japan. e-mail: koike@rk.phys.keio.ac.jp*

♮ *Department of Physics, Tokyo Institute of Technology, Tokyo 152, Japan. e-mail: ahosoya@th.phys.titech.ac.jp*

ABSTRACT

A general method of construction are given for compact locally homogeneous universes. A concrete example is presented with careful counting of dynamical degrees of freedom and explicit calculation of the Teichmüller parameters as time-dependent functions. We finally show that the dynamical system for the example, which includes the Teichmüller parameters as dynamical variables, possesses a Hamiltonian structure.

1 Introduction

Compact homogeneous universe $({}^{(4)}M, g_{ab})$ is a four-dimensional Lorentzian manifold which admits a foliation by leaves of compact locally homogeneous Riemannian three-manifold. This article is to give a detailed analysis of dynamics of such compact homogeneous universes in the context of general relativity. It is less known but of great importance that compact (locally) homogeneous Riemannian manifolds can be different even if they have same topology and same local geometry. Such "global" differences of geometry are known as the *Teichmüller deformations*, parametrized by the Teichmüller parameters, which are also dynamical variables of the compact homogeneous universe.

In Ref.[1] we presented a treatment of three-dimensional compact homogeneous Riemannian manifolds, where we (1) gave the possible eight types (a~h) of homogeneous universal covers, which are closely related to Thurston's eight geometries[2, 3], (2) classified compact quotients (a1/1,b/1,etc.), and (3) gave Teichmüller spaces, spanned by the Teichmüller parameters, by explicitly finding embeddings of covering groups in the isometry groups of the universal covers, which enable us to perform explicit calculations. Hence, we here first need to show how to "adapt" such knowledge of compact homogeneous three-manifolds in the context of relativity in four-dimensions. To this, we begin with considering a four-dimensional universal cover $({}^{(4)}\tilde{M}, \tilde{g}_{ab})$, and then make identifications in $({}^{(4)}\tilde{M}, \tilde{g}_{ab})$ so as to make each three-surface $(\tilde{M}_t, \tilde{h}_{ab})$

compact. In order to utilize our knowledge about compact homogeneous three-manifolds, this operation is translated into the condition that we make $(\tilde{M}_t, \tilde{h}_{ab})$ compact by the action of a discrete subgroup of the *extendible isometries* $\mathrm{Esom}\tilde{M}_t$, which are defined with respect to having natural extension as a subgroup of $\mathrm{Isom}({}^{(4)}\tilde{M})$, the isometries of $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$. The four-dimensional universal cover $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$ is chosen so that it satisfies Einstein's equation.

According to the method that we shall provide, we then construct an example of compact Bianchi II (the b/1 model), and show calculations of giving the time-development of the Teichmüller parameters. We then further develop our theory through the example of the b/1 model. We show that all the dynamical variables including the Teichmüller parameters can be transformed into a metric of the form of Bianchi II. We therefore find that the system possesses the natural Hamiltonian structure induced from the Einstein-Hilbert action.

We employ the abstract index notation [4] throughout this article.

2 Universal covers and identifications

In this section, we explore possible form of metric of a compact homogeneous universe, establish a method of construction of compact homogeneous universes, and discuss generic feature of the global deformations of compact homogeneous spatial sections of the universe.

We denote the universal cover of a compact homogeneous universe $({}^{(4)}M, \mathrm{g}_{ab})$ as $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$. Note that it comes from a property of covering map that a compact homogeneous universe inherits a natural metric from the universal cover $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$ (and the converse is also true). We thus first consider the universal cover metric $\tilde{\mathrm{g}}_{ab}$.

In the case of a Bianchi minimal geometry [1], we can take the 4-metric as

$$\mathrm{d}s^2 = -N^2(t,\mathrm{x})\mathrm{d}t^2 + h_{\alpha\beta}(t)(N^\alpha(t,\mathrm{x})\mathrm{d}t + \sigma^\alpha)(N^\beta(t,\mathrm{x})\mathrm{d}t + \sigma^\beta), \tag{1}$$

where (t,x) are local coordinates, σ^α the invariant 1-forms, and $\alpha, \beta, \ldots$ run from 1 to 3. Indeed, the spatial metric $h_{\alpha\beta}(t)\sigma^\alpha\sigma^\beta$ is homogeneous on each surface t =const.. But, of course, the three killing vectors k_α of the homogeneous 3-metric do, in general, not act on the 4-metric, i.e.

$$\mathcal{L}_{k_\alpha}(\mathrm{d}s^2) \neq 0. \tag{2}$$

In (2), k_α is regarded as the vector induced by the embeddings of the homogeneous 3-manifolds in ${}^{(4)}\tilde{M}$. Although we cannot exclude this general class of metrics by the method of "compactification" which we shall explain later, we in this article consider only the case where k_α acts four-dimensionally, i.e. the case $\mathcal{L}_{k_\alpha}(\mathrm{d}s^2) = 0$. In fact, such a subclass of metrics can express the most variety of compact universes. Hence, our metric is of the form

$$\mathrm{d}s^2 = -N^2(t)\mathrm{d}t^2 + h_{\alpha\beta}(t)(N^\alpha(t)\mathrm{d}t + \sigma^\alpha)(N^\beta(t)\mathrm{d}t + \sigma^\beta). \tag{3}$$

This metric becomes without loss of generality the following form

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + h_{\alpha\beta}(t)\sigma^\alpha\sigma^\beta. \tag{4}$$

For Bianchi class A and type V, we can diagonalize the metric components $h_{\alpha\beta}$'s in Eq.(4) in vacuum with respect to an invariant basis σ^α which is defined in appropriate local coordinates. In some cases, the components can become simpler than being diagonal. Note that even in spatially compact cases, we can find such local coordinates simplifying the metric in a neibourhood of any point in ${}^{(4)}M$, and can use these coordinates to find solutions of local equations such as the Einstein equation $R_{\mu\nu} = 0$. Also, for the universal cover, this is the case. We can therefore say

that the dynamical degrees of freedom carried by the universal cover coincides with those of the open model. That is, as for the open case, we need to exclude the degrees of freedom of HPDs [6, 1] from the universal cover to exclude the "gauge" freedom.

We express compact ${}^{(4)}M$ by taking identifications in ${}^{(4)}\tilde{M}$. The identifications act on each homogeneous 3-surface $(\tilde{M}_t, \tilde{h}_{ab})$ of t =const.. Let $\mathrm{Isom}\tilde{M}_t$ be the isometries of $(\tilde{M}_t, \tilde{h}_{ab})$. Then, we may make the homogeneous 3-manifold $(\tilde{M}_t, \tilde{h}_{ab})$ compact by the action of a discrete subgroup $\Gamma$of $\mathrm{Isom}\tilde{M}_t$, where possible Γ's of various universal covers are already given in Ref.[1]. Although this seems to be enough to specify the initial identifications, we should remark some subtle points to avoid possible confusions. Before elucidating them, we define the extendible isometries of $(\tilde{M}_t, \tilde{h}_{ab})$, $\mathrm{Esom}\tilde{M}_t \subset \mathrm{Isom}\tilde{M}_t$, as follows.

Definition 1 (Extendible isometry group) *Let $(\tilde{M}_t, \tilde{h}_{ab})$ be a spatial section of $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$. An extendible isometry is the restriction on $\tilde{M}_t$ of an isometry of $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$ which preserves $\tilde{M}_t$. They form a subgroup of* $\mathrm{Isom}\tilde{M}_t$. *We call it the extendible isometry group, and denote it as* $\mathrm{Esom}(\tilde{M}_t, {}^{(4)}\tilde{M})$, *or simply* $\mathrm{Esom}\tilde{M}_t$. *Obviously, an extendible isometry $a \in \mathrm{Esom}\tilde{M}_t$ has the natural extension on ${}^{(4)}\tilde{M}$ which is an element of* $\mathrm{Isom}{}^{(4)}\tilde{M}$ *and preserves $\tilde{M}_t$. We call such the natural extension on ${}^{(4)}\tilde{M}$ the extended isometry of a, or simply the extension of a.*

The first point to remark is that our compact universes are obtained by taking identifications in *four* dimensional universal covers, and so that, for any two points which are identified, there should exist an isometry of $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$ (*not* of $(\tilde{M}_t, \tilde{h}_{ab})$) which maps one to the other. For on $(\tilde{M}_t, \tilde{h}_{ab})$, this fact is interpreted as saying simply that the identifications must be implemented in the extendible isometries $\mathrm{Esom}\tilde{M}_t$. Moreover, when we are concerned with the dynamical degrees of freedom of the system, we must take conjugations for the identifications to exclude "gauge" freedom again, and it is now clear that the conjugations must also be taken by $\mathrm{Esom}\tilde{M}_t$. In view of initial value formulation, we may be able to summarize the above facts as follows; [1]

Proposition 1 *The identifications on an initial surface $(\tilde{M}_t, \tilde{h}_{ab})$ must be implemented in* $\mathrm{Esom}\tilde{M}_t$,

$$\Gamma \subset \mathrm{Esom}\tilde{M}_t, \tag{5}$$

to get a compact homogeneous universe out of the four-dimensional universal cover $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$, and the conjugations also must be taken by $\mathrm{Esom}\tilde{M}_t$. *Moreover, the identifications acting on whole $({}^{(4)}\tilde{M}, \tilde{\mathrm{g}}_{ab})$ are determined by the action of the extended isometry of Γ on ${}^{(4)}\tilde{M}$.*

Note that we can take conjugations for $(\tilde{M}_t, \tilde{h}_{ab})$ only by $\mathrm{Esom}\tilde{M}_t$, whereas the Teichmüller parameters are defined with respect to conjugations by full $\mathrm{Isom}\tilde{M}_t$. Roughly speaking, the difference between freedom of $\mathrm{Esom}\tilde{M}_t$ and $\mathrm{Isom}\tilde{M}_t$ appears as the freedom of giving initial 'velocities' of Teichmüller parameters. We will see this effect more explicitly in the example of compact Bianchi II in the next section.

(The reader may find in Ref.[9] more a detailed account to the subject discussed in this section.)

[1] Although proposition 1 is sufficient to prescribe how to make a compact homogeneous universe, one another view of how to "evolve" the identifications on an initial surface $(\tilde{M}_t, \tilde{h}_{ab})$ is worth noting. To this, we first note that the normal geodesics emerging from $\tilde{M}_t$ for a fixed value of t are unique, provided that they are parametrized by the proper time τ. We refer to the exponential map $\exp \tau n^a(t)$ which is defined with respect to the normal vector field $n^a(t)$ on $\tilde{M}_t$ as the *normal map*. It is easy to see that, to maintain consistency, *any two points mapped by the normal maps from identified two points on an initial surface should continue to be identified.* An advantage of this view is that once given a four-dimensional universal cover metric $\tilde{g}_{ab}$, we can determine how to evolve the identifications even if we do not know the isometries of the universal cover. Specifically with the metric (4), since the hypersurface-orthogonal geodesics are along the t-axes, we can say that *In terms of the globally defined coordinates $(t, \mathbf{x})$ of metric (4), if at the initial surface $t = t_0$ an identification is specified as $(t_0, \mathbf{x}) \simeq (t_0, a\mathbf{x})$, where a is a free action on the coordinate space, then at any time t we must have $(t, \mathbf{x}) \simeq (t, a\mathbf{x})$.* This may be useful as an alternative of the latter half of proposition 1.

3 Time-development of Teichmüller parameters

To get the Teichmüller parameters of a compact section (M_t, h_{ab}), we need to compare two mathematical representations, i.e., $(\tilde{M}_t, \tilde{h}_{ab})$ with the covering group Γ, and the standard universal cover $(\tilde{M}, \tilde{h}^{\rm std}_{ab})$ [2] with the covering group, A, parametrized by the Teichmüller parameters. Γ and A are generated by the same number, n, of generators, $\{g_i\}$ and $\{a_i\}$ $(i = 1, \cdots, n)$, respectively. $\{g_i\}$ and $\{a_i\}$ satisfy the same multiplication rule of an extendible isometry group. We can get the Teichmüller parameters by finding the automorphism of Esom$\tilde{M}$ which relates the two sets of generators.

We shall do this for compact Bianchi II (the b/1 model), where we can see the most typical calculation to get Teichmüller parameters. Although similar calculations are applicable straightforwardly to other types, e.g. compact Bianchi I, $\rm VI_0$, $\rm VII_0$, we will concentrate on the b/1 model to present a concrete example in a way as complete as possible.

Our universal cover metrics are synchronous (Eq.(4)) and diagonal.

The b/1 model: a compact model on Bianchi II geometry

We start with the multiplication rule of Nil (=Bianchi II group);

$$\begin{pmatrix} g^1 \\ g^2 \\ g^3 \end{pmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = \begin{pmatrix} g^1 + h^1 \\ g^2 + h^2 \\ g^3 + h^3 + g^1 h^2 \end{pmatrix}, \tag{6}$$

where $g, h \in$ Nil, and we shall use superscripts to denote the components of a group element. We use the same components $(x^1, x^2, x^3) \equiv (x, y, z)$ as coordinates of $\tilde{M}_t$. The action of Nil on $\tilde{M}_t$ is defined by the left action on $(x, y, z) \in$Nil. A Nil-invariant (diagonal) metric is given by

$$\mathrm{d}l^2 = h_{11}\mathrm{d}x^2 + h_{22}\mathrm{d}y^2 + h_{33}(\mathrm{d}z - x\mathrm{d}y)^2, \tag{7}$$

where $h_{\alpha\alpha}$ $(\alpha = 1 \sim 3)$ are constants, i.e. independent of (x, y, z). The four dimensional universal cover metric of our concern is of the form

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}l^2 \tag{8}$$

with $h_{\alpha\alpha}$ being functions of t. The vacuum solution is, of course, known, but we proceed with calculation, leaving $h_{\alpha\alpha}$ free, since they are complicated functions in the synchronous gauge and, moreover, it enable us to apply the result also to models other than the vacuum model.

We consider manifold "b/1 $(n = 1)$", classified in Ref.[1], which is probably the most stereotypical compact manifold modeled on Bianchi II geometry. The fundamental group π_1 is given by (See Eq.(118) in Ref.[1])

$$\pi_1 = \left\langle g_1, g_2, g_3; [g_1, g_2]g_3^{-1}, [g_1, g_3], [g_2, g_3] \right\rangle. \tag{9}$$

To represent the generators of π_1, g_i's, in Esom$\tilde{M}_t$ = Nil, we put them as

$$g_i = \begin{pmatrix} {g_i}^1 \\ {g_i}^2 \\ {g_i}^3 \end{pmatrix}, (i = 1 \sim 3), \tag{10}$$

[2] In Ref.[1], they were called the representative metrics.

and substitute these in the relations of π_1 (Eq.(9)). We then get the following;

$$g_1 = \begin{pmatrix} {g_1}^1 \\ {g_1}^2 \\ {g_1}^3 \end{pmatrix}, g_2 = \begin{pmatrix} {g_2}^1 \\ {g_2}^2 \\ {g_2}^3 \end{pmatrix}, g_3 = \begin{pmatrix} 0 \\ 0 \\ \bar{g}_3{}^3 \end{pmatrix}, \tag{11}$$

where $\bar{g}_3{}^3 \equiv {g_1}^1{g_2}^2 - {g_1}^2{g_2}^1 \neq 0$.

We then consider the possible conjugations by $\mathrm{Esom}\tilde{M}_t = \mathrm{Nil}$. For the conjugation of g_i's by $h = (h^1, h^2, h^3) \in \mathrm{Nil}$ [3], we have

$$hg_1h^{-1} = \begin{pmatrix} {g_1}^1 \\ {g_1}^2 \\ {g_1}^3 + h^1{g_1}^2 - h^2{g_1}^1 \end{pmatrix}, hg_2h^{-1} = \begin{pmatrix} {g_2}^1 \\ {g_2}^2 \\ {g_2}^3 + h^1{g_2}^2 - h^2{g_2}^1 \end{pmatrix}, hg_3h^{-1} = g_3. \tag{12}$$

We can make the third components of g_1 and g_2 zero if we take h as $h^1 = ({g_1}^3{g_2}^1 - {g_1}^1{g_2}^3)/\bar{g}_3{}^3$, $h^2 = ({g_1}^3{g_2}^2 - {g_1}^2{g_2}^3)/\bar{g}_3{}^3$. After all, our representation of π_1 in Nil reduces to

$$g_1 = \begin{pmatrix} {g_1}^1 \\ {g_1}^2 \\ 0 \end{pmatrix}, g_2 = \begin{pmatrix} {g_2}^1 \\ {g_2}^2 \\ 0 \end{pmatrix}, g_3 = \begin{pmatrix} 0 \\ 0 \\ \bar{g}_3{}^3 \end{pmatrix}. \tag{13}$$

The nonvanishing four independent components in these g_i's determine the initial identifications in the universal cover with metric (8).

To proceed further calculations, we here cite the definition given in Ref.[1] of the Teichmüller parameters for b/1 and some related properties. We denote the standard universal cover as $(\mathbf{R}^3, \tilde{h}^{\mathrm{std}}_{ab})$, where the standard metric $\tilde{h}^{\mathrm{std}}_{ab}$ is given by (Eq.(75) in Ref.[1])

$$\mathrm{d}l^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + (\mathrm{d}z - x\mathrm{d}y)^2. \tag{14}$$

Any compact homogeneous 3-manifold classified in b/1 is globally conformally isometric to manifold $(\mathbf{R}^3, \tilde{h}^{\mathrm{std}}_{ab})/A$, where A is a covering group whose generators are given by

$$a_1 = \begin{pmatrix} {a_1}^1 \\ 0 \\ 0 \end{pmatrix}, a_2 = \begin{pmatrix} {a_2}^1 \\ {a_2}^2 \\ 0 \end{pmatrix}, a_3 = \begin{pmatrix} 0 \\ 0 \\ {a_1}^1{a_2}^2 \end{pmatrix}. \tag{15}$$

Then, the Teichmüller parameters are $\tau = ({a_1}^1, {a_2}^1, {a_2}^2)$ (Eq.(129) in Ref.[1]). We can see that the map

$$s_\theta : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} R_\theta \begin{pmatrix} x \\ y \end{pmatrix} \\ z + \zeta_\theta(x, y) \end{pmatrix} \tag{16}$$

is a 1-parameter isometry for $(\mathbf{R}^3, \tilde{h}^{\mathrm{std}}_{ab})$, where R_θ is the rotation matrix by angle θ, and ζ_θ is defined by

$$\zeta_\theta(x, y) \equiv \frac{1}{2}((x^2 - y^2)\cos\theta - 2xy\sin\theta)\sin\theta. \tag{17}$$

We here remark that s_θ is not an element of $\mathrm{Esom}\tilde{M}_t$ but of $\mathrm{Isom}\tilde{M}_t$, and therefore $\mathrm{Esom}\tilde{M}_t \neq \mathrm{Isom}\tilde{M}_t$ in the b/1 model. For an element $h \in$Nil, conjugation by s_θ is given by

$$s_\theta \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} s_\theta^{-1} = \begin{pmatrix} R_\theta \begin{pmatrix} h^1 \\ h^2 \end{pmatrix} \\ h^3 + \zeta_\theta(h^1, h^2) \end{pmatrix}. \tag{18}$$

[3] For typographical convenience, we sometimes write components of group horizontally.

Note that metric (7) is rewritten as

$$dl^2 = \frac{h_{11}h_{22}}{h_{33}}(dx'^2 + dy'^2 + (dz' - x'dy')^2) \tag{19}$$

with

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{h_{33}}{h_{22}}}x \\ \sqrt{\frac{h_{33}}{h_{11}}}y \\ \frac{h_{33}}{\sqrt{h_{11}h_{22}}}z \end{pmatrix}, \tag{20}$$

where $h_{\alpha\alpha}$ are regarded as constants. If we view this coordinate transformation as a diffeomorphism and drop the constant conformal factor of metric (19), the resulting metric coincides with the standard metric (14). This diffeomorphism is obviously an element of the HPDs [6, 1], from the form of metric (19), so that the transformation $(x, y, z) \to (x', y', z')$ is an (outer-) automorphism of Nil. The image of g_i's, which acts on metric (19) (or metric (14)), is

$$g_1' = \begin{pmatrix} {g_1}^{1'} \\ {g_1}^{2'} \\ 0 \end{pmatrix}, g_2' = \begin{pmatrix} {g_2}^{1'} \\ {g_2}^{2'} \\ 0 \end{pmatrix}. \tag{21}$$

Here,

$${g_1}^{1'} = \sqrt{\frac{h_{33}}{h_{22}}}{g_1}^1, {g_1}^{2'} = \sqrt{\frac{h_{33}}{h_{11}}}{g_1}^2, {g_2}^{1'} = \sqrt{\frac{h_{33}}{h_{22}}}{g_2}^1, {g_2}^{2'} = \sqrt{\frac{h_{33}}{h_{11}}}{g_2}^2. \tag{22}$$

Generator g_3 is automatically determined by g_1 and g_2 (see Eq.(13)), so we will concentrate on g_1, g_2 and the images of them by automorphisms. Since Eq.(21) is not of the form of Eq.(15), it does not yet give the Teichmüller parameters. To get them, we take a conjugation of Eq.(21) by the (full) isometry of Nil, which is given by Nil itself with s_θ. We can "rotate" the two-dimensional vectors $({g_i}^{1'}, {g_i}^{2'})$ $(i = 1, 2)$ by conjugations by s_θ (Eq.(18)), leaving the third components zero by a conjugation by Nil like the way we obtained Eq.(13). So, we arrive at

$$a_1 = h s_{\theta_1} g_1' s_{\theta_1}^{-1} h^{-1} = \begin{pmatrix} \sqrt{({g_1}^{1'})^2 + ({g_1}^{2'})^2} \\ 0 \\ 0 \end{pmatrix} \tag{23}$$

and

$$a_2 = \begin{pmatrix} {g_2}^{1'}\cos\theta_1 - {g_2}^{2'}\sin\theta_1 \\ {g_2}^{1'}\sin\theta_1 + {g_2}^{2'}\cos\theta_1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{({g_1}^{1'})^2 + ({g_1}^{2'})^2}} \begin{pmatrix} {g_1}^{1'}{g_2}^{1'} + {g_1}^{2'}{g_2}^{2'} \\ {g_1}^{1'}{g_2}^{2'} - {g_1}^{2'}{g_2}^{1'} \\ 0 \end{pmatrix}, \tag{24}$$

where

$$\cos\theta_1 = \frac{{g_1}^{1'}}{\sqrt{({g_1}^{1'})^2 + ({g_1}^{2'})^2}}, \sin\theta_1 = \frac{-{g_1}^{2'}}{\sqrt{({g_1}^{1'})^2 + ({g_1}^{2'})^2}}, \tag{25}$$

and h is an element of Nil. Using Eq.(22), we obtain the final form of the Teichmüller parameters;

$$\begin{aligned} {a_1}^1 &= \sqrt{\frac{h_{33}}{h_{22}}({g_1}^1)^2 + \frac{h_{33}}{h_{11}}({g_1}^2)^2}, \\ {a_2}^1 &= \frac{1}{{a_1}^1}\left(\frac{h_{33}}{h_{22}}{g_1}^1{g_2}^1 + \frac{h_{33}}{h_{11}}{g_1}^2{g_2}^2\right), \\ {a_2}^2 &= \frac{\bar{g}_3{}^3}{{a_1}^1}\frac{h_{33}}{\sqrt{h_{11}h_{22}}}. \end{aligned} \tag{26}$$

In Eq.(26), parameters ${g_1}^1, {g_1}^2, {g_2}^1, {g_2}^2$ and hence ${\bar{g}_3}^3$ are constants, and $h_{\alpha\alpha}$'s are functions of t. The metric components h_{11}, h_{22}, and h_{33} are determined by Einstein's equation, and we must exclude the degrees of freedom of HPDs from them. Hence, the number of free parameters that the metric components can have coincides with the known number of degrees of freedom of the conventional (open) Bianchi models [7]. For the vacuum Bianchi II, the number of free parameters in the metric components is two. With the four parameters specifying the initial identifications, the total number of dynamical degrees of freedom of the present vacuum b/1 model is six. The dynamical variables are the Teichmüller parameters ${a_1}^1, {a_2}^1$, and ${a_2}^2$, *and* the total volume

$$v = ({\bar{g}_3}^3)^2 \sqrt{h_{11} h_{22} h_{33}}. \tag{27}$$

Remember that the Teichmüller parameters are defined with respect to the standard universal cover which is isometric to the universal cover $(\tilde{M}_t, \tilde{h}_{ab})$ up to a global conformal factor. In fact, it is clear that, if we know the values of them, we can completely construct the original compact 3-manifold.

4 Metric components as dynamical variables

In the first subsection, we present a heuristic argument to show that the dynamical variables of the b/1 model can be embedded into the solution of Einstein's equation, and thus possess the Hamiltonian structure. This argument will provide a firm ground for the geometrical interpretation in the second subsection.

4.1 The case of the b/1 model

Consider an HPD, $\phi : {}^{(4)}\tilde{M} \to {}^{(4)}\tilde{M}$, of Bianchi II. Then, the pullback of the invariant basis, σ^α $(\alpha = 1 \sim 3)$, of Bianchi II group is given by

$$\phi_* : \begin{pmatrix} \sigma^1 \\ \sigma^2 \\ \sigma^3 \end{pmatrix} \to \begin{pmatrix} {f^1}_1 & {f^1}_2 & 0 \\ {f^2}_1 & {f^2}_2 & 0 \\ {f^3}_1 & {f^3}_2 & {f^1}_1 {f^2}_2 - {f^1}_2 {f^2}_1 \end{pmatrix} \begin{pmatrix} \sigma^1 \\ \sigma^2 \\ \sigma^3 \end{pmatrix}, \tag{28}$$

where ${f^\alpha}_\beta$'s are constants, and

$$\sigma^1 \equiv \mathrm{d}x, \ \sigma^2 \equiv \mathrm{d}y, \ \sigma^3 \equiv \mathrm{d}z - x\mathrm{d}y. \tag{29}$$

In fact, the Maurer-Cartan relation is invariant under Eq.(28).

If we have a solution of the universal cover which satisfies Einstein's equation, then the pullback of it by Eq.(28) is also a solution of Einstein's equation. Such a solution would have the maximal number of arbitrary parameters. In the parameters, we can check that ${f^3}_1$ and ${f^3}_2$ span the gauge orbits generated by the two momentum constraints which would come out if all six metric components $h_{\alpha\beta}$ were supposed. However, the remaining four parameters in Eq.(28) are independent of the momentum constraints. What we would like to ask in the following is whether we can regard the four parameters as the four parameters, ${g_i}^j$'s, specifying the initial identifications.

Suppose that

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + h_{11}(\sigma^1)^2 + h_{22}(\sigma^2)^2 + h_{33}(\sigma^3)^2 \tag{30}$$

is the solution of the universal cover which does not contain the freedom of HPDs. Thus, the metric (30) possesses two arbitrary parameters. (See the last paragraph of the last section.) We

write the metric possessing "extra" four parameters due to Eq.(28) with $f^3{}_1 = f^3{}_2 = 0$ as

$$ds^2 = -dt^2 + H_{\alpha\beta}\sigma^\alpha\sigma^\beta. \tag{31}$$

Then, by direct substitution, we see that

$$\begin{aligned} H_{11} &= h_{11}(f^1{}_1)^2 + h_{22}(f^2{}_1)^2, \\ H_{22} &= h_{11}(f^1{}_2)^2 + h_{22}(f^2{}_2)^2, \\ H_{12} &= h_{11}f^1{}_1f^1{}_2 + h_{22}f^2{}_1f^2{}_2, \\ H_{33} &= (f^1{}_1f^2{}_2 - f^1{}_2f^2{}_1)^2h_{33}, \\ H &\equiv \det(H_{\alpha\beta}) = (f^1{}_1f^2{}_2 - f^1{}_2f^2{}_1)^4h, \ (h \equiv h_{11}h_{22}h_{33}). \end{aligned} \tag{32}$$

Note that, while this solution of Einstein's equation possesses six parameters, the solution of dynamical variables Eqs.(26) and (27) also has six parameters. Do these two "solutions" have explicit correspondence? Or more restrictively, are the four parameters $f^1{}_1, f^1{}_2, f^2{}_1$ and $f^2{}_2$ functions of the four parameters $g_1{}^1, g_1{}^2, g_2{}^1$ and $g_2{}^2$? The answer is yes. In fact, if we suppose

$$f^1{}_1 = g_1{}^1, f^1{}_2 = g_2{}^1, f^2{}_1 = g_1{}^2, f^2{}_2 = g_2{}^2, \tag{33}$$

then metric components (32) and the dynamical variables (26),(27) have the following explicit correspondence;

$$\begin{aligned} H_{11} &= \left(\frac{v}{(a_1{}^1)^2(a_2{}^2)^2}\right)^{2/3}(a_1{}^1)^2, \\ H_{22} &= \left(\frac{v}{(a_1{}^1)^2(a_2{}^2)^2}\right)^{2/3}((a_2{}^1)^2 + (a_2{}^2)^2), \\ H_{12} &= \left(\frac{v}{(a_1{}^1)^2(a_2{}^2)^2}\right)^{2/3}a_1{}^1a_2{}^1, \\ H_{33} &= \left(\frac{v}{(a_1{}^1)^2(a_2{}^2)^2}\right)^{2/3}(a_1{}^1)^2(a_2{}^2)^2. \end{aligned} \tag{34}$$

Or the inverse is;

$$a_1{}^1 = H_{33}\sqrt{\frac{H_{11}}{H}}, \ a_2{}^1 = \frac{H_{33}H_{12}}{\sqrt{HH_{11}}}, \ a_2{}^2 = \sqrt{\frac{H_{33}}{H_{11}}}, \ v = \sqrt{H}. \tag{35}$$

Note that $h_{\alpha\beta}$, $f^\alpha{}_\beta$ and $g_i{}^j$, which are "parts" of the two solutions (26), (27), and (32), do not appear in these relations explicitly.

Since metric components $H_{\alpha\beta}$ are solutions of Einstein's equation, we have found the embeddings of the dynamical variables into the Einstein's equation. Since the Bianchi II model belongs to class A that Ellis and MacCallum classified [5], the present dynamical variables do have the Hamiltonian structure induced from the reduced Einstein-Hilbert action. The Hamiltonian in terms of the Teichmüller parameters and the volume are given by explicit calculations as

$$\begin{aligned} \mathcal{H} &= \frac{1}{2v}\Big\{(a_1{}^1a_2{}^2v)^{4/3} + (a_1{}^1)^2(p^1{}_1)^2 + ((a_2{}^1)^2 + (a_2{}^2)^2)(p^2{}_1)^2 + (a_2{}^2)^2(p^2{}_2)^2 \\ &\quad +2a_1{}^1a_2{}^1p^1{}_1p^2{}_1 + a_1{}^1a_2{}^2p^1{}_1p^2{}_2 + a_2{}^1a_2{}^2p^2{}_1p^2{}_2\Big\} - \frac{3}{8}v(p_v)^2. \end{aligned} \tag{36}$$

Here $p^i{}_j$ and p_v are the conjugate momenta of $a_i{}^j$ and v, respectively. The Hamiltonian constraint $\mathcal{H} \approx 0$ reduces the number of degrees of freedom of the system by two, so that we again get $4 \times 2 - 2 = 6$ as the total number of degrees of freedom in the phase space.

4.2 The geometrical interpretation of the metric components

As we have seen, we can now view the metric components $H_{\alpha\beta}$ as dynamical variables, or in other words, the Teichmüller parameters have been "absorbed" in the universal cover metric. In this subsection, we give how we can geometrically interpret the absorption.

First, we make general discussion. Let $\tau = (\tau_1, \cdots, \tau_n)$ and Γ_τ be the Teichmüller parameters, and the covering group acting on the standard universal cover $(\tilde{M}, \tilde{h}^{\rm std}_{ab})$, respectively. For convenience, we choose the origin in the Teichmüller space arbitrarily, and denote it by τ_0 or simply 0. The projection map $\pi_\tau : \tilde{M} \to M$ is defined with respect to the action of Γ_τ on $\tilde{M}$. We denote the image of $(\tilde{M}, \tilde{h}^{\rm std}_{ab})$ by π_τ by $(M, h_{ab}|_\tau)$. Information of the Teichmüller deformations is contained in the projection map π_τ in this view.

On the other hand, for a Γ_τ we can define a diffeomorphism $\phi_\tau : \tilde{M} \to \tilde{M}$ such that

$$\forall a_\tau \in \Gamma_\tau,\ a_\tau = \phi_\tau \circ a_0 \circ \phi_\tau^{-1}. \tag{37}$$

We refer to ϕ_τ as a *Teichmüller diffeomorphism* (TD). Obviously, ϕ_τ maps the fundamental region of π_0 to that of π_τ. Hence, if we consider a universal cover $(\tilde{M}, \phi_{\tau *}\tilde{h}^{\rm std}_{ab})$, then the image of the new universal cover by the projection π_0 coincides with $(M, h_{ab}|_\tau)$. One thus find

$$h_{ab}|_\tau = \pi_{0*}^{-1}\phi_{\tau *}\tilde{h}^{\rm std}_{ab}. \tag{38}$$

With this view, the projection map is fixed, while the universal cover varies with τ. Note that this meets with that, in the b/1 model, the metric components $H_{\alpha\beta}$ of universal cover varies with the Teichmüller parameters as Eq.(34). Similar treatment has also been taken in (2+1)-gravity [8].

In general, there are many possibilities of TDs for a universal cover and a Γ_τ. Moreover, if we choose TDs arbitrarily and make spatial part of the four dimensional universal cover metric with Teichmüller parameters being functions of t, then the extrinsic curvature fails to have the necessary homogeneity that would extract a consistent slice out of the full phase space. One way to extract a consistent slice is to implement the TDs in a subgroup of HPDs.[4] (HPDs themselves comprise a group.) We call such TDs the *homogeneity preserving Teichmüller diffeomorphisms* (HPTDs).

Now we are in a position to apply our argument to the b/1 model. The standard universal cover metric is given by

$$dl^2 = (\sigma^1)^2 + (\sigma^2)^2 + (\sigma^3)^2, \tag{39}$$

where σ^α are defined in Eq.(29) with coordinates (x, y, z). Let $\tau = ({a_1}^1, {a_2}^1, {a_2}^2)$ and $\tau_0 = (1, 0, 1)$. Then, appropriate HPTDs are

$$\phi_\tau : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} {a_1}^1 x + {a_2}^1 y \\ {a_2}^2 y \\ \frac{1}{2}{a_2}^1{a_2}^2 y^2 + {a_1}^1{a_2}^2 z \end{pmatrix}, \tag{40}$$

which induce the map

$$\phi_{\tau *} : \begin{pmatrix} \sigma^1 \\ \sigma^2 \\ \sigma^3 \end{pmatrix} \to \begin{pmatrix} {a_1}^1 & {a_2}^1 & 0 \\ 0 & {a_2}^2 & 0 \\ 0 & 0 & {a_1}^1{a_2}^2 \end{pmatrix} \begin{pmatrix} \sigma^1 \\ \sigma^2 \\ \sigma^3 \end{pmatrix}. \tag{41}$$

[4] Another ways might be possible if non-trivial lapse and shifts are allowed. They may correspond to the way to employ metric (1) as a universal cover metric.

It is easy to observe that the diffeomorphisms of Eq.(41) certainly comprise a subgroup of HPDs. Also, we can check that the unit cube ($0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1$), which is a fundamental region of π_0, is mapped to a fundamental region of π_τ, showing ϕ_τ certainly comprises HPTDs.

The induced metric of Eq.(39) by ϕ_τ is given by the direct substitution of Eq.(41). Normalizing the induced metric to give v^2 as determinant, we again obtain the metric (31) with Eq.(34).

5 Conclusions

We have given two methods for constructing compact homogeneous universes. In the first one, we began with considering the four dimensional universal covers, and then took identifications and conjugations in it. Following this method, we also gave explicit calculations giving the time-development of the Teichmüller parameters for the b/1 model on the Bianchi II geometry. (Some other examples are presented in Ref.[9].)

In the second one, Teichmüller deformations were implemented into variations of the universal covers, and the four dimensional universal cover metric was written in terms of the Teichmüller parameters. As Ashtekar and Samuel first pointed out [6], the lack of Teichmüller parameters in open Bianchi models gives rise to discrepancies in counting the numbers of dynamical degrees of freedom, however the compact model treated in this article has no discrepancies, as we have seen. Moreover, while they did not mention how the metrics relate to the Teichmüller parameters, now that is manifest.

References

[1] T. Koike, M. Tanimoto and A. Hosoya, J. Math. Phys. **35**, 4855 (1994).

[2] W. P. Thurston, The Geometry and Topology of 3-manifolds (To be published by Princeton University Press).

[3] P. Scott, Bull. London Math. Soc. **15**, 401 (1983).

[4] R. M. Wald, General Relativity (University of Chicago, Chicago, 1984).

[5] G. F. R. Ellis and M. A. H. MacCallum, Commun. Math. Phys. **12**, 108 (1969)

[6] A. Ashtekar and J. Samuel, Class. Quantum Gravit. **8**, 2191 (1991).

[7] See, e.g., D. Kramer, H. Stephani, M. MacCallum, and E. Herlt, Exact Solutions of Einstein's Field Equations. (Cambridge Univ. Press, Cambridge, 1980)

[8] A. Hosoya and K. Nakao, Class. Quantum Gravit. **7**, 163 (1990); For recent status of (2+1)-gravity and more references, see e.g. S. Carlip, Lectures on (2+1)-Dimensional Gravity, gr-qc/9503024, (1995).

[9] M. Tanimoto, Thesis, unpublished (1996); M. Tanimoto, T. Koike and A. Hosoya, preprint, TIT/HEP-322/COSMO-68, gr-qc/9604056, (1996).

Quantum Stability of (2+1)-Spacetimes with Non-Trivial Topology: Negative Curvature Cases

Masaru Siino*
Department of Physics, Kyoto University
Kitashirakawa, Sakyoku, Kyoto 606-01, Japan

Quantum fields are investigated in the (2+1)-open-universes with non-trivial topologies by the method of images. The universes are locally de Sitter spacetime and anti-de Sitter spacetime. In the present article we study spacetimes whose spatial topologies are a torus with a cusp and a sphere with three cusps as a step toward the more general case. A quantum energy momentum tensor is obtained by the point stripping method. Though the cusps are no singularities, the latter cusps cause the divergence of the quantum field. This suggests that only the latter cusps are quantum mechanically unstable. Of course at the singularity of the background spacetime the quantum field diverges. Also the possibility of the divergence of topological effect by a negative spatial curvature is discussed. Since the volume of the negatively curved space is larger than that of the flat space, one see so many images of a single source by the non-trivial topology. It is confirmed that this divergence does not appear in our models of topologies. The results will be applicable to the case of three dimensional multi black hole [9].

I. INTRODUCTION

When we consider matter fields in a spacetime with a non-trivial topology, the boundary effects of quantum fields appear. This is one of main targets in quantum field theory in curved spacetimes. Such effects have been studied well in spatially flat spacetimes [1], but not so well in spatially curved spacetimes. This is because of the complexity of the topology which is allowed in such curved spatial section. In other words, there will be new topological effects of quantum field in the curved spatial section with a little complex topology.

To construct a space with a non-trivial topology, we identify the points of covering space by the discrete subgroup of the isometry of the space. Then we want to consider the covering space with an appropriately simple isometry group such that the topology has some interesting characteristics. As the space with simple isometry, there are S^n, R^n and H^n, so called closed-, flat- and open-universe. Furthermore, we decide to treat H^n since this hyperbolic space allows various topologies possessing interesting characteristics.

To treat this open-universe as a background spacetime, we determine its time evolution. For simplicity we consider maximally symmetric spacetimes of de Sitter spacetime with a hyperbolic chart or anti de Sitter spacetime in a Robertson-Walker coordinate. Their spatial sections are H^n. The de Sitter spacetime with the hyperbolic chart may be important in a cosmological sense. It is believed that the global feature of our universe is homogeneous and isotropic. If our observation suggests that the spatial curvature of our universe is negative, the background spacetime is locally open-universe. In the inflation, the de Sitter spacetime with a hyperbolic chart is a good model for cosmology [2]. If we prefer a universe with a finite volume, the de Sitter spacetime with a hyperbolic chart with non-trivial topology will become important.

The topology of the open-universe (in the present article, the 'open-universe' means not an open topology but only a negatively curved space) is well known in a two dimensional space. Then we construct the simple example of a two dimensional open-universe with interesting topology in (2+1)-de Sitter spacetimes and (2+1)-anti de Sitter spacetimes. Quantum scalar field is studied in these spacetimes using point stripping manner. The divergences of the quantum fields will be discussed.

In the section 2, we prepare simple model of the universes with interesting topologies in the (2+1)-de Sitter spacetime and the (2+1)-anti de Sitter spacetime. The quantum field is investigated in the section 3. The last section is devoted to a summary and discussions.

---

*e-mail: msiino@tap.scphys.kyoto-u.ac.jp, JSPS fellow

II. (2+1) OPEN-UNIVERSE WITH NON-TRIVIAL TOPOLOGY

A. Two Dimensional Universe

First of all, we develop topologies of two-dimensional spatial sections. For the simplicity of topologies we treat (2+1)-spacetime in the present article. For a cosmological reason and a further simplicity, spatial two-dimensional sections are assumed to be S^2, R^2 or H^2 corresponding to closed-, flat- or open-universe, respectively. In a two-dimensional space, the topologies of complete manifolds are classified by Euler numbers. This is calculated from the number of handles and cusps (see Fig.1). The cusp is a point at infinity with a needle-like structure. Here it should be emphasized that the cusp is no unnatural or artificial. These points at infinity are no singularity. There is no reason to avoid them in classical physics [3]. Furthermore, the role of the cusp might be rather essential as discussed in [4] [3].

From the Gauss-Bonnet theorem for a complete 2-manifold, the Euler number χ is given by

$$\frac{1}{4\pi}\int dv^{(2)}R = \chi = 2 - 2N(handle) - N(cusp), \tag{1}$$

where ${}^{(2)}R$ is a two-dimensional scalar curvature and $N(*)$ is the number of $*$. The signature of ${}^{(2)}R$ restricts the variety of the topology. Since the Euler number is less than 2 ($\chi = 2$ is for a sphere), the case of negative curvature allows various topologies (various numbers of handles and cusps). In such negative curvature space, we expect new topological effects of a quantum field. Since the cusp is an infinitely small structure, it may cause the divergence of the quantum field. The negative curvature space has a crucial characteristic that the volume of the space is larger than that of the flat space at a distant region. Then one will see so many images of sources because of a non-trivial topology. The method of images may suffer the difficulty of a divergence. To discuss these speculations, we must treat general topologies of the negative curvature space H^2. In the present article, however, only two simple cases of them can be investigated since these cases possess the cusps and the above mentioned characteristic.

To construct a non-trivial topology of the negative curvature space, we draw a polygon surrounded by geodesics on a hyperbolic space H^2 as a fundamental region and identify the geodesics. Poincaré model is one of the models of the hyperbolic space, which is conformally flat and a compact chart. The metric of the Poincaré model is

$$ds^2 = \frac{4(dr^2 + r^2 d\phi^2)}{(1-r^2)^2}, \tag{2}$$

whose spatial curvature is -1. In this model, geodesics are circles crossing a circle with $r = 1$ at right angles. This circle with $r = 1$ is a sphere at infinity corresponding to the infinity of H^2.

Now we give simple topologies including cusps. The simplest polygon producing non-trivial topology is a tetragon. Drawing the tetragon $\Box ABCD$ of a fundamental region as in Fig.2, the tessellation of H^2 by this tetragon becomes simplest. In Fig.2, a half of the tetragon, $\triangle ABC$ and $\triangle ACD$, is a regular triangle and its vertices A, B, C, D are on the sphere at infinity. H^2 are tessellated by parallel transformations of these triangles. The tessellation turns out to be a series of regular polygons whose centers are the origin of the Poincaré model. The triangles form a self-homothetic structure. Requiring orientability, there are two pairs of identifications for geodesics contained by the triangles, which provide complete manifold. One is

$$\triangle ABC \longrightarrow \triangle DCG \quad \text{by } \Lambda^{(1)}_{DCG} \tag{3}$$

$$\triangle ABC \longrightarrow \triangle LAD \quad \text{by } \Lambda^{(1)}_{LAD}, \quad \Lambda^{(1)} \in SO(2,1). \tag{4}$$

They generate a discrete subgroup $\Gamma^{(1)}$ of $SO(2,1)$. By these identifications, the topology of $H^2/\Gamma^{(1)}$ becomes a torus having a point at infinity, which is a cusp (see Fig.(3)). Of course, the Gauss-Bonnet theorem gives its Euler number as

$$\chi = 2 - 2N(handle) - N(cusp) = 2 - 2 - 1 = -1 = \frac{1}{4\pi}\int_{H^2/\Gamma^{(1)}} dv^{(2)}R. \tag{5}$$

The other is

$$\triangle ABC \longrightarrow \triangle GDC \quad \text{by } \Lambda^{(2)}_{GDC} \tag{6}$$

$$\triangle ABC \longrightarrow \triangle ADL \quad \text{by } \Lambda^{(2)}_{ADL}, \quad \Lambda^{(2)} \in SO(2,1). \tag{7}$$

They generate $\Gamma^{(2)} \subset SO(2,1)$. A resultant topology $H^2/\Gamma^{(2)}$ is a sphere with three cusps (Fig.3). The Euler number of it is

$$\chi = 2 - 2N(handle) - N(cusp) = 2 - 0 - 3 = -1 = \frac{1}{4\pi}\int_{H^2/\Gamma^{(2)}} dv^{(2)}R. \tag{8}$$

It is noted that these manifolds $H^2/\Gamma^{(1)}$ and $H^2/\Gamma^{(2)}$ are easier to handle than the double-torus which is a well known example of hyperbolic manifolds. For instance, though the fundamental group of our manifolds is (a,b), that of the double torus is $(a,b,c,d : aba^{-1}b^{-1}cdc^{-1}d^{-1} = 1)$ [5].

Here we express all elements of $\Gamma^{(i)}$ in an appropriate form for the rest of the present article. Using $\{\Lambda_1^{(i)}, \Lambda_2^{(i)}, \Lambda_3^{(i)}\}$ acting as

$$\triangle ABC \longrightarrow \triangle DCG \quad \text{by } \Lambda_1^{(1)} \tag{9}$$
$$\longrightarrow \triangle BIE \quad \text{by } \Lambda_2^{(1)} \tag{10}$$
$$\longrightarrow \triangle KFA \quad \text{by } \Lambda_3^{(1)} \tag{11}$$
$$\longrightarrow \triangle GDC \quad \text{by } \Lambda_1^{(2)} \tag{12}$$
$$\longrightarrow \triangle EBI \quad \text{by } \Lambda_2^{(2)} \tag{13}$$
$$\longrightarrow \triangle AKF \quad \text{by } \Lambda_3^{(2)}, \tag{14}$$

$^{\forall}T^{(i)} \in \Gamma^{(i)}$ is given by

$$T^{(i)} = R(\theta_j)(\Lambda_k^{(i)})^n. \tag{15}$$

For example, $T^{(i)}$ transform $\triangle ABC$ to $\triangle A'B'C'$ in Fig.2. $\triangle ABC$ is transformed to a $2n$-th outward position of triangles by $(\Lambda_k^{(i)})^n$ and rotated around the origin by $R(\theta_j)$ with an appropriate angle θ_j. k is selected from $1 \sim 3$ so that the orders of the vertices match between $\triangle ABC$ and $\triangle A'B'C'$.

For the rest of the article, we give $\Lambda_1^{(1)}$ in terms of the Poincaré model. Using $z = re^{i\theta}$, this representation of the isometry of H^2 becomes a subgroup of $SL(2,C)$ in the coordinate of the Poincaré model and is given by

$$z \to f(z) = \frac{az+b}{\bar{b}z+\bar{a}}, \quad \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \in SL(2,C). \tag{16}$$

$\Lambda_1^{(1)}$ is given by

$$f(z) = \frac{a_0 z + b_0}{\bar{b}_0 z + \bar{a}_0} \tag{17}$$

$$a_0 = \frac{1}{\sqrt{1-r_0^2}}\sqrt{i\frac{r_0 e^{i\pi/8}-1}{r_0 e^{-i\pi/8}-1}}, \quad b_0 = \frac{1}{\sqrt{1-r_0^2}}\frac{r_0 e^{i\pi/8}}{a_0}, \tag{18}$$

$$r_0 = \frac{1-\sin\frac{\pi}{8}}{\cos\frac{\pi}{8}} \tag{19}$$

Here we should note that $\Lambda^{(1)}$'s and $\Lambda^{(2)}$'s are in the different category of the Lorentz group $SO(2,1)$. It is known that all elements of $SO(2,1)$ are $SO(2,1)$-conjugate to an element of the following forms. We call them as standard forms. In the $SL(2,C)$ representation (16),

- 1) An elliptic element is conjugate to

$$T_e : \quad f(z) = e^{i\theta}z = \frac{e^{i\theta/2}z}{e^{-i\theta/2}} \tag{20}$$

  with one fixed point on H^2.

- 2) A parabolic element is conjugate to

$$T_p : \quad f(z) = \frac{(1+i)z+i}{-iz+(1-i)} \tag{21}$$

  with one fixed point on the sphere at infinity.

• 3) A hyperbolic element is conjugate to

$$T_h: \quad f(z) = \frac{z\cosh\beta + \sinh\beta}{z\sinh\beta + \cosh\beta}, \tag{22}$$

with two fixed point on the sphere at infinity.

These angle parameters θ, β are real numbers. Though $\Lambda^{(1)}$'s are conjugate the category 3), $\Lambda^{(2)}$'s are parabolic. We note that the fixed point of a parabolic element corresponds to a cusp produced by the parabolic element. It is revealed in the next section that these facts affect the quantum field.

B. de Sitter and anti de Sitter Spacetime with Non-trivial Topology

Now we consider (2+1)-spacetime whose certain spatial sections have above mentioned topologies. For simplicity, Teichmüller deformation [6] is not considered and every identifications are supposed to be done on the spatial 2-sections of the synchronous gauge. Then the local geometry of the spacetime is that of the open FRW-universe. For further simplicity, we assume the spacetime is maximally symmetric. Such (2+1)-spacetimes allowing spatial H^2-sections are Minkowski spacetime, de Sitter spacetime dS^3 and anti de Sitter spacetime AdS^3. Though we treat only dS^3 and AdS^3 in the present article, the investigation of the present article is easily applicable to the case of the Minkowski spacetime.

The isometry groups of dS^3 and AdS^3 are $SO(3,1)$ and $SO(2,2)$, respectively. The open chart of dS^3 and the RW(Robertson-Walker)-coordinate of AdS^3 determine the natural extension of $SO(2,1)$ which is the isometry of H^2 to $SO(3,1)$ or $SO(2,2)$ so that the extended action of $SO(2,1)$ preserves their time-slicing, respectively. The identifications $\Gamma \subset SO(2,1)$ on H^2 also extended to $\gamma \subset SO(3,1)$, $SO(2,2)$ on dS^3 and AdS^3, which preserves the time-slicing. γ provides non-trivial topology dS^3/γ or AdS^3/γ to the spacetimes.

For the next section, we imbed dS^3 and AdS^3 as a covering spacetime of the concerning spacetime with non-trivial topology into four dimensional flat spacetimes with signatures $(-+++)$ and $(--++)$, respectively. Using the coordinate of the Poincaré model for spatial sections,

$$ds^2 = dX^2 + dY^2 - dZ^2\begin{Bmatrix}+\\-\end{Bmatrix}dW^2 \tag{23}$$

$$X = \begin{Bmatrix}\sinh t\\ \cos t\end{Bmatrix}\frac{2r}{1-r^2}\cos\theta \tag{24}$$

$$Y = \begin{Bmatrix}\sinh t\\ \cos t\end{Bmatrix}\frac{2r}{1-r^2}\sin\theta \tag{25}$$

$$Z = \begin{Bmatrix}\sinh t\\ \cos t\end{Bmatrix}\frac{1+r^2}{1-r^2} \tag{26}$$

$$W = \begin{Bmatrix}\cosh t\\ \sin t\end{Bmatrix} \tag{27}$$

$$\Longrightarrow ds^2 = -dt^2 + \begin{Bmatrix}\sinh^2 t\\ \cos^2 t\end{Bmatrix}\frac{4(dr^2 + r^2 d\theta^2)}{(1-r^2)^2}, \tag{28}$$

where the upper case is dS^3 and the lower case is AdS^3. We treat only the spacetime with a unit curvature radius for simplicity because the absolute value of the curvature is not essential for the following investigation.

III. QUANTUM FIELD

In this section Quantum field is investigated in the spacetime with non-trivial topology whose covering spacetime is dS^3 or AdS^3. We introduce conformally coupled massless scalar field ϕ, with the action

$$S = -\int dx^3\sqrt{g}\left(\frac{1}{2}\partial^\mu\phi\partial_\mu\phi + \frac{1}{16}R\phi^2\right), \tag{29}$$

where R is the scalar curvature. The field equation in dS^3 or AdS^3 with $R_{\mu\nu} = \pm 2g_{\mu\nu}$ is

$$\left(\nabla^\mu\nabla^\nu \mp \frac{3}{4}\right)\phi = 0 \tag{30}$$

with $R = \pm 6$ for our dS^3 or AdS^3 with a unit curvature radius.

Now we consider the Hadamard Green functions in the covering spacetime dS^3 or AdS^3. According to Steif [7], they are given by

$$\bar{G}(x,y) = \frac{1}{4\pi}\frac{1}{|x-y|}, \tag{31}$$

where $|x-y|$ is a chordal distance between x and y in the four dimensional imbedding spacetime (24)~(27) and not a proper distance in dS^3 or AdS^3.

The Hadamard functions for spacetimes with non-trivial topologies $dS^3/\gamma^{(i)}$, $AdS^3/\gamma^{(i)}$ can be obtained from the Hadamard function for their covering spacetime (31) by the method of images. Since the images of y are generated by elements of $\gamma^{(i)}$, the Green function is

$$G^{(i)}(x,y) = \sum_{T^{(i)}\in\gamma^{(i)}, T^{(i)}\neq id} \bar{G}(x, T^{(i)}\circ y). \tag{32}$$

where the summation is over all elements of $\gamma^{(i)}$ except for identity. The identity is excluded to subtract all local contributions of the quantum field. This procedure ought to regularize the energy-momentum tensor of the quantum field.

When γ is Abelian group, the summation can be easily evaluated like the three dimensional black hole case [7] (for example, $AdS^3/\gamma_B H : \gamma_B H = (\Lambda_1^{(1)})^n$ is equivalent to the three dimensional black hole). On the other hand, our non-Abelian γ makes rigorous evaluations impossible. The simple universe shown in the previous section, however, allows us to evaluate some divergences. The abstract summation of (32) is decomposed into

$$G^{(i)}(x,y) = \sum_{n=1}^{\infty}\sum_{j=1}^{3\cdot 4^{n-1}}\sum_{k}^{3} \bar{G}(x, R(\theta_j)(\Lambda_k^{(i)})^n \circ y) \tag{33}$$

by eq.(15) A quantum energy-momentum tensor is given by

$$< T_{\mu\nu} >= \lim_{y\to x} \mathcal{D}_{\mu\nu} G(x,y), \tag{34}$$

in the point stripping method, where $\mathcal{D}_{\mu\nu}$ is a certain differential operator (see [7], for example). Hence, investigating the zero of the distance $|x - T(x)|$ and the summations of (33), we can discuss the divergences of $< T_{\mu\nu} >$ since the divergences of $< T_{\mu\nu} >$ come out from the divergences of the Green function.

First of all we discuss the characteristics of $\gamma^{(1)}$. As stated in the previous section, $T^{(1)} \in \gamma^{(1)}$ can be written as

$$T^{(1)} = A^{-1} T_h A$$

by an appropriate $A \in SO(2,1)$, where A corresponds to an isometric coordinate transformation of $\theta \to \theta'$ and $r \to r'$ in eq.(28). T_h in (22) corresponds to a Lorentz boost about X-Z direction in the imbedding spacetime. From imbedding equation (24)~(27), a chordal distance $|x - T^{(1)}(x)|$ is given by

$$|x - A^{-1}T_h A \circ x| = |x' - T_h \circ x'| \tag{35}$$

$$= \sqrt{(2\cosh 2\chi - 1)(X'^2 - Z'^2)} \tag{36}$$

$$= \begin{Bmatrix} \sinh t \\ \cos t \end{Bmatrix} \sqrt{(2\cosh 2\chi - 1)\frac{4r^2\cos^2\theta' - (1+r'^2)^2}{(1-r'^2)^2}}, \tag{37}$$

where $\{X', Y', Z', W'\}$ is an imbedding coordinate for $x' = A(x)$. T_h in the imbedding spacetime $\{X, Y, Z, W\}$ is given by the matrix,

$$\begin{pmatrix} \cosh\chi & 0 & \sinh\chi & 0 \\ 0 & 1 & 0 & 0 \\ \sinh\chi & 0 & \cosh\chi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{38}$$

From (24)~(27), χ is given by the parameter β of (22) as

$$\chi = 2\beta. \tag{39}$$

In the $SL(2,C)$ representation (16), for a hyperbolic element $T^{(1)}$

$$\chi = 2\cosh^{-1}\frac{Tr_{SL(2,C)}(T^{(1)})}{2} = 2\cosh^{-1}\frac{a+\bar{a}}{2}, \tag{40}$$

where $Tr_{SL(2,C)}$ is a trace about $SL(2,C)$ matrix. Especially, by the rotational symmetry of $\Lambda^{(1)}_{1\sim 3}$'s, all χ's of $\Lambda^{(1)}_{1\sim 3}$'s are the same. We write it as $\chi_0 = 2\cosh^{-1}((a_0+\bar{a_0})/2)$.

On the contrary, $T^{(2)}$ has a different standard form (21). With appropriate $B \in SO(2,1)$ and T_p in (21), $T^{(2)}$ is written as $B^{(-1)}T_pB$. We dare express the distance in the coordinate of the Poincaré model (24)~(27) to evaluate it at infinity. $|x - T^{(2)}(x)|$ becomes

$$|x - B^{-1}T_pB \circ x| = |x'' - T_p \circ x''| \tag{41}$$

$$= (Z'' + Y'') \tag{42}$$

$$= \left\{\begin{matrix}\sinh t \\ \cos t\end{matrix}\right\}\frac{2r''\sin\theta'' + 1 + r''^2}{1 - r''^2}, \tag{43}$$

where $\{X'', Y'', Z'', W''\}$ and r'', θ'' are coordinates for $x'' = B(x)$. When θ'' is $3\pi/2$, the distance vanishes at infinity ($r'' \to 1$). This point ($r'' = 1, \theta'' = 3\pi/2$) is a fixed point of T_p corresponding to a cusp. $T^{(2)}$ forms a cusp at x_c with $B(x_c) = (r'' = 1, \theta'' = 3\pi/2)$. Then $|x - T^{(2)}(x)|$ should vanish at the cusp on each timeslice, while $|x - T^{(1)}(x)|$ never approaches to zero except for the initial or final singularity.

There are three possibilities of divergences for the quantum field. The quantum field will diverge at the singularity of the background spacetime as in the case of three dimensional black hole [7]. Also the infinitely small structure of cusps will cause the divergence of the quantum field though the cusp is not a singularity. Furthermore, the summation of the images in the method of images is expected to diverge since the volume of the hyperbolic space is larger than that of the flat space at a distant region. One will see so many images of a source.

First two possibilities are investigated in the following. For $T^{(1)} \in \gamma^{(1)}$, eq.(37) tells that $|x - T^{(1)}(x)|$ vanishes at the planes $X'^2 - Z'^2 = 0$. Since $A \in SO(2,1)$ is a proper-Lorentz group in a part of the imbedding spacetime $\{X, Y, Z, W = \pm 1\}$, the planes $X'^2 - Z'^2 = 0$ cannot invade into the inside of a light cone at $\{0, 0, 0, \pm 1\}$ which is initial or final singularity of dS^3 and AdS^3. Moreover, one or two pairs of null geodesic, $\{X'^2 - Z'^2 = 0, Y' = 0, W' = \pm 1\}$, lie on the singularities. Since the quantum field diverges on these planes $< T_{\mu\nu} >$ diverge when one approaches to the singularity of the background spacetimes by the topological effect.

On the other hand, $T^{(2)} \in \gamma^{(2)}$ has a different characteristic about cusps. From eq.(43), $|x - T^{(2)}(x)|$ vanishes at the cusps on each time-slice, while $|x - T^{(1)}(x)|$ never vanishes on each time-slice even at their cusps. Therefore $< T_{\mu\nu} >$ is singular at the cusps of $dS^3/\gamma^{(2)}$ and $AdS^3/\gamma^{(2)}$ and regular at the cusps of $dS^3/\gamma^{(1)}$ and $AdS^3/\gamma^{(1)}$.

Finally we discuss the third possibility of the divergences estimating the summation over all transformations in (33). From a rotational symmetry around the origin of each time-slice ($X = Y = 0, Z = \left\{\begin{matrix}\sinh t \\ \cos t\end{matrix}\right\}, W = \left\{\begin{matrix}\cosh t \\ \sin t\end{matrix}\right\}$),

$$|x - R(\theta)(\Lambda^{(1)}_k)^n \circ x|_{x=origin} = |x - (\Lambda^{(1)}_k)^n \circ x|_{x=origin} \tag{44}$$

$$= \left\{\begin{matrix}\sinh t \\ \cos t\end{matrix}\right\}\sqrt{2\cosh(2n\chi_0) - 1}, \tag{45}$$

where we use $A^{-1}\Lambda^n A = (A^{-1}\Lambda A)^n$. If we consider the point $x = x_0$,

$$\left\{\begin{matrix}\sinh t \\ \cos t\end{matrix}\right\}\sqrt{2\cosh(2n\chi_0) - 1} - 2|x_0| \le |x - R(\theta)(\Lambda^{(1)}_k)^n|_{x=x_0} \tag{46}$$

$$\le \left\{\begin{matrix}\sinh t \\ \cos t\end{matrix}\right\}\sqrt{2\cosh(2n\chi_0) - 1} + 2|x_0| \tag{47}$$

For a sufficiently large n, $\sqrt{2\cosh(2n\chi_0) - 1}$ and $|x - R(\theta)(\Lambda^{(1)}_k)^n(x)|$ behaves as $e^{n|\chi_0|}$. Furthermore, $|x - R(\theta)(\Lambda^{(2)}_k)^n|$ also behaves as $e^{n|\chi_0|}$ for a large n, since the tessellation of $\gamma^{(2)}$ is the same tessellation as $\gamma^{(1)}$. From (17)(18)(19)(40) the summation about j and k in (33) is estimated as

$$\lim_{y\to x}\sum_{j=1}^{3\cdot 4^{n-1}}\sum_{k}^{3}\bar{G}(x,R(\theta_j)(T_k)^i\circ y)\sim\frac{4^n}{e^{n|\chi_0|}}=\left(\frac{4}{4.74}\right)^n=0.8438^n. \tag{48}$$

Though a rigorous estimation may be possible, it is too complicated and will give us no essential information. $< T_{\mu\nu} > \sim O(\sum_n 0.843^n)$ barely converge because of the exact value of χ_0. If the investigation could be done in other topology with negative curvature, a different value of χ might cause the divergences of the summation.

IV. SUMMARY AND DISCUSSION

In the present article, we have investigated new topological effects of a quantum field in a torus universe with a cusp and a sphere universe with three cusps. Their covering spacetime is de Sitter spacetime or anti de Sitter spacetime. The cusp is a point at infinity with regular local structure and needle-like global structure. The role of the cusp is discussed in [4] [3]. Three possibilities of divergences of the energy-momentum tensor have been studied.

First, a divergence appears on the coordinate singularity of the classical background spacetime, which is initial or final singularity in cosmological sense. This is similar to the case of the three dimensional black hole [7].

Next possibility is a divergence at the cusps. In the present article, we show there are two types of the cusps. One is a cusp made by hyperbolic transformations of $SO(2,1)$ and the other is made by a parabolic transformation. We observed that $< T_{\mu\nu} >$ diverge at the latter cusps, which are included in the sphere with three cusps. This aspect means that the latter cusp induce a quantum instability. Only the latter cusps will require a treatment in quantum gravity.

The last possibility is the divergence of the summation of images. This corresponds to the effect that we see more images of a source in a negatively curved universe than in a flat universe at a distant region. The summation, however, converges in the spacetimes given in the present article. The convergence of the image summation strongly depends on the values of a boost angle $|\chi|$ and the shape of a tessellation in the covering space. Though $e^{|\chi_0|}$ is 4.74, if $e^{|\chi_0|}$ were less than 4 with the same tessellation, $< T_{\mu\nu} >$ would diverge everywhere and the divergence is hard to remove. In the case of other topology, other $e^{|\chi|}$ and other tessellation may make the summation diverge. If so, it will turn out that there are topologies accepting quantum field and not accepting. When we consider a compact (without a cusp) topology with a negative curvature, such situation may occur, though the compact topology is very difficult to treat.

Recently Brill [9] shows that a three dimensional multi-black hole solution can be constructed in the three dimensional anti-de Sitter spacetime. It is easily found that $AdS^3/\gamma^{(2)}$ with a larger boost angle $|\chi'|$ than $|\chi_0|$ is regarded as a two-black hole solution. Of course, the summation of images converges in this solution.

By a regularization performed in the present article, we perfectly subtract local divergences. We, however, have observed the topological divergences. They cannot be regularized and will have physical meanings.

Can we carry out a similar investigation in other topology with a negative curvature. A compact topology (without a cusp) seems impossible since the tessellation is so complicated. On the other hand, it may be possible to treat other topologies with cusps. At least, we can decide whether each cusp causes a divergence of quantum field or not, knowing whether the identification providing the cusp is parabolic or hyperbolic. The divergence of summation of images sensitively depends on the shape of the tessellation in the covering space and will be difficult to treat without a sufficient symmetry. In (3+1)-dimension the similar investigation may be possible. There will be convenient models of non-trivial topology.

ACKNOWLEDGMENTS

We would like to thank Professor H. Sato and Dr. T. Tanaka for helpful discussions. The author thanks the Japan Society for the Promotion of Science for financial support. This work was supported in part by the Japanese Grant-in-Aid for Scientific Research Fund of the Ministry of Education, Science, Culture and Sports.

---

[1] R. Banach and J. S. Dowker, *J. Phys.* **A12** (1979)2545, B. S. Dewitt, C. F. Hart and C. J. Isham, *Physica* **96A** (1979)197, W. A. Hiscock and D. A. Konkowski, *Phys. Rev.* **D26** (1982)1225.

[2] J. R. Gott, III,*Nature (London)* **295** (1982)304, M. Sasaki, T. Tanaka, K. Yamamoto and J. Yokoyama, *Phys. Lett.* **B317** (1993)510.
[3] S. Ding, Y. Maeda and M. Siino, preprint TITHEP-282/COSMO-52, gr-qc/9503026.
[4] G. W. Gibbons, hep-th/9601075.
[5] C. Nash and S. Sen, *Topology and geometry for physics*, London, Academic Press, 1983.
[6] Y. Fujiwara and J. Soda, *Prog. Theor. Phys.* **83** (1990)733.
[7] A. R. Steif, *Phys. Rev.* **D49** (1994)585.
[8] M. Bañados, C. Teitelboim and J. Zanelli, *Phys. Rev. Lett.* **69** (1992) 1849, M. Bañados, M. Henneaux, C. Teitelboim and J. Zanelli, *Phys. Rev.* **D48** (1993) 1506.
[9] D. R. Brill, preprint UMDGR 96-47, gr-qc/9511022.

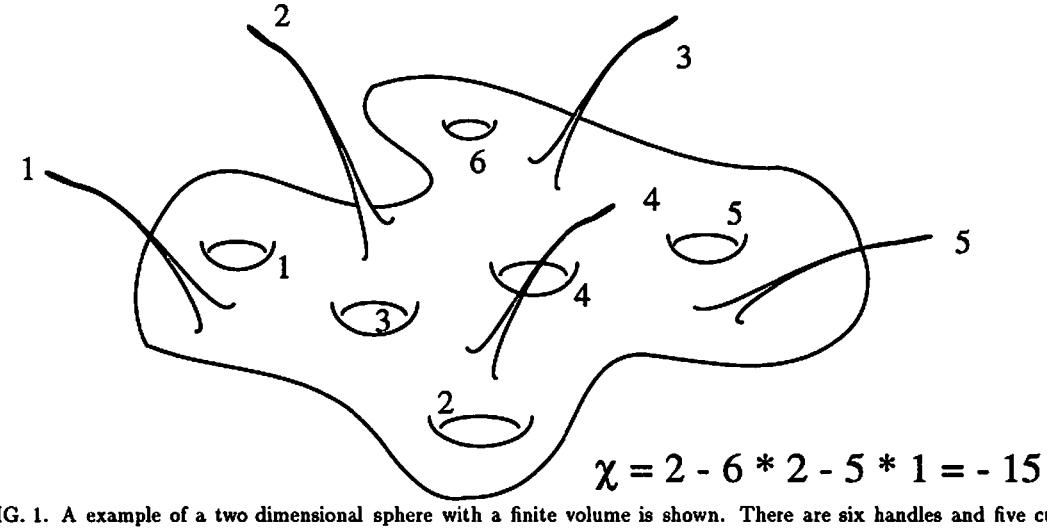


FIG. 1. A example of a two dimensional sphere with a finite volume is shown. There are six handles and five cusps. Its Euler number is $\chi = 2 - 2 * 6 - 1 * 5 = 15$.

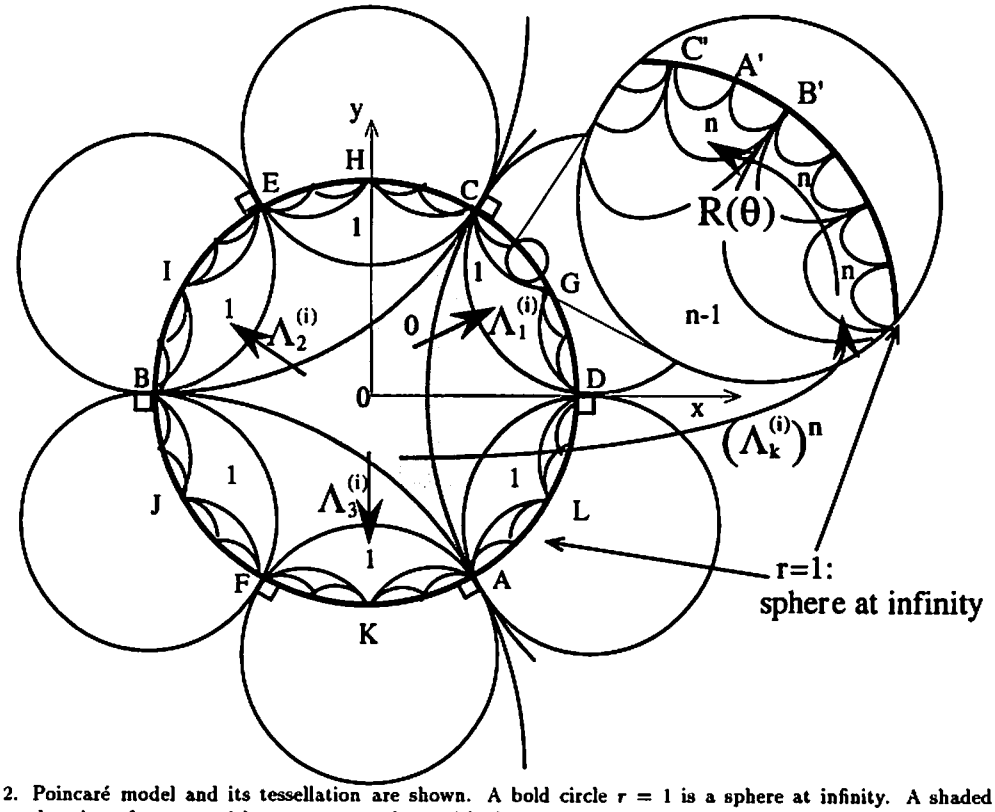


FIG. 2. Poincaré model and its tessellation are shown. A bold circle $r = 1$ is a sphere at infinity. A shaded region is a fundamental region of a torus with a cusp or a sphere with three cusps. $\triangle ABC$ is transformed to a n-th outward position by $(\Lambda_k^{(i)})^n$. Furthermore, $R(\theta)$ transform it to $\triangle A'B'C'$.

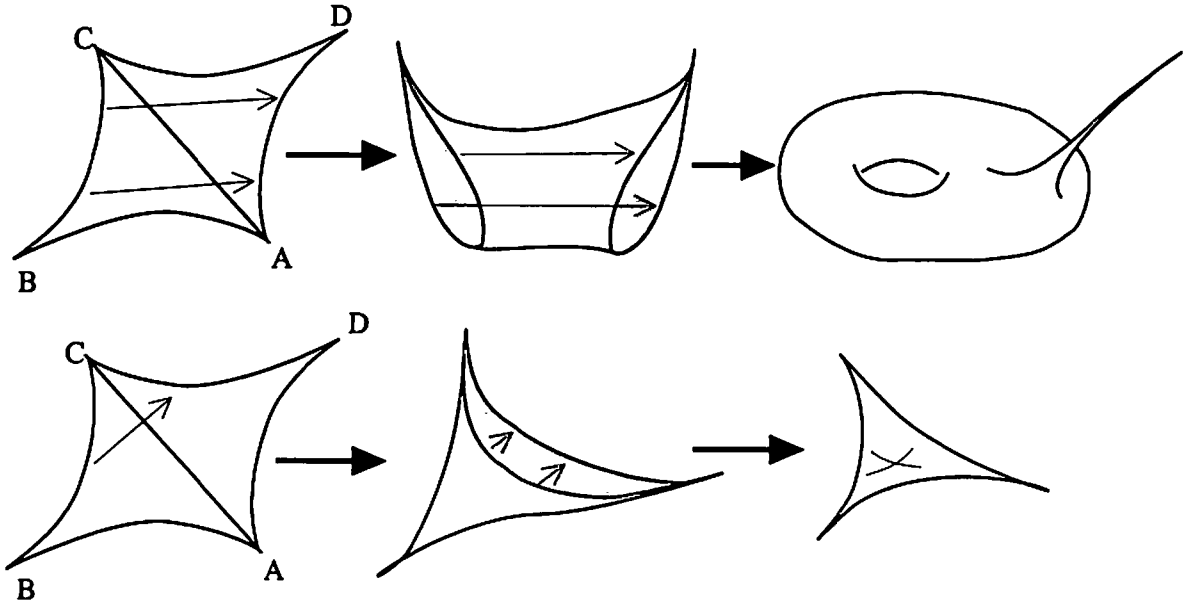


FIG. 3. The upper shows that a torus with a cusp is constructed from a tetragon. The lower shows a tetragon becomes a sphere with three cusps.

Causality violation and singularities

KENGO MAEDA [1] AND AKIHIRO ISHIBASHI [2]

Department of Physics, Tokyo Institute of Technology, Oh-Okayama Meguro-ku, Tokyo 152, Japan

ABSTRACT

We show that singularities necessarily occur when a boundary of causality violating set exists in a space-time under the physically suitable assumptions except the global causality condition in the Hawking-Penrose singularity theorems. Instead of the global causality condition, we impose some restrictions on the causality violating sets to show the occurrence of singularities.

1 Introduction

Space-time singularities have been discussed for a long time in general relativity. In 1970, Hawking and Penrose[1] showed that singularities, which mean causal geodesic incompleteness, could occur in a space-time under seemingly reasonable conditions in classical gravity. Their singularity theorem has an important implication that our universe has an initial singularity if we do not consider quantum effects. However, this theorem is physically unsatisfactory in the sense that the causality requirement everywhere in a space-time seems too restrictive. We can only experience local events and there is no guarantee that the causality holds in the entire universe. As is well known, Kerr type black holes have causality violating sets if the space-time is maximally extended. Therefore, it will be important to investigate occurrence of singularities in a space-time in which the global causality condition is violated.

There are some works on a causality violation concerned with the occurrence of singularity. Tipler[2,3] showed that any attempt to evolve closed timelike curves from an initial regular Cauchy data would cause singularities to form in a space-time. He presented a singularity theorem in which the global causality condition in the Hawking-Penrose theorem is replaced by the weaker one and adding the stronger energy condition. In his theorem his stronger energy condition is essential to the occurrence of singularities.

Kriele presented his singularity theorems in which causality violating sets are restricted but with usual energy condition in the Hawking-Penrose theorem instead of Tipler's energy condition. He showed that the causality violating set has incomplete null geodesics if its boundary is compact [4]. Kriele[5] also presented a generalization of the Hawking-Penrose singularity theorem. In his paper he showed that singularities would occur provided that causality holds at least in the future endpoints of the trapped set.

[1] e-mail:maeda@th.phys.titech.ac.jp
[2] e-mail:akihiro@th.phys.titech.ac.jp

Newman[6] found a black hole solution which had no singularities. This black hole solution is obtained by a suitable conformal transformation of the Gödel universe: One might consider his conclusion would suggest that causality violating set can prevent singularities from occurring. However, his case seems too special and even unphysical, because causality is violated in the entire space-time. It is physically more acceptable to assume that at least there must be causality preserving regions in a space-time. One can pick up the Taub-NUT universe as an example which contains both causality violating and preserving sets. In this universe, there exist singularities on the boundary of causality violating sets. This suggests that the boundary generates a geodesic incompleteness.

In this paper we shall show that the boundary of causality violating sets are essential to occurrence of singularities. We also discuss relation between our theorems and the Hawking-Penrose theorem.

In the next section, we briefly review Tipler's and Kriele's singularity theorems. In section 3, the definitions and the lemmas for discussing causal structure and singularities are listed up. We present our singularity theorems for partially causality violating space-times in section 4. Section 5 is devoted to summary.

2 Tipler's and Kriele's theorems

We review Tipler's and Kriele's theorems in this section. In addition, we discuss how causality violation is related to singularities in these theorems.

First, we quote Tipler's theorem.

Tipler's theorem(1977)
A space-time (M, g) cannot be null geodesically complete if
(1) $R_{ab}K^aK^b \geq 0$ for all null vectors K^a;
(2) there is a closed trapped surface in M;
(3) the space-time is asymptotically deterministic, and the Einstein equations hold;
(4) the partial Cauchy surface defined by (3) is non-compact.

Here the asymptotically deterministic condition in the condition (3) is defined as follows.

Definition
A space-time (M, g) is said to be *asymptotically deterministic* if
(i) (M, g) contains a partial Cauchy surface S such that
(ii) either $H(S) = H^+(S) \cup H^-(S)$ is empty, or if not, then

$$\lim_{s \to a} [\inf T_{ab}K^aK^b] > 0$$

on at least one of the null geodesic generators $\gamma(s)$ of $H(S)$, where a is the past limit of the affine parameter along γ if $\gamma \in H^+(S)$, and the future limit if $\gamma \in H^-(S)$. (K^a is the tangent vector to γ.)

This condition has been introduced by following reasons. In the case that the formation of a Cauchy horizon $H^+(S)$ is due to causality violation, one would expect that the region where $H^+(S)$ begins would contain enough matter (the condition (ii)) which causes gravitational field sufficiently strong so as to tip over the light cones and eventually leads to causality violation. We can regard this condition as a special type of energy condition which dispenses with the causality condition in the Hawking-Penrose singularity theorem.

In the following sections, we shall impose some conditions on causality violating sets to replace global causality condition in the Hawking-Penrose theorem instead of imposing this energy condition.

Next, we quote some definitions and Kriele's theorems [4, 5].

Definition

•*focal point*

Let S be a locally spacelike surface (not necessary achronal surface) and let us consider a future directed null geodesic, $\beta(t)$, from S parameterized by t. If for any point $\beta(t)$ such that $t \geq t_1$, there exists an arbitrarily close timelike curve from S to the point $\beta(t)$, then β is called a *focal point* to S.

•*Generalized future horismos of S*

Generalized future horismos of S, denoted by $e^+(S, M)$, is a closure of all future null geodesics β from S which have no focal points. (The future end points of $e^+(S, M)$ correspond to the focal points.)

•*cut locus*: $cl(S, M, +)$

The set of future end points of $e^+(S, M)$.

•*almost closed causal curve*

Choose an arbitrary Riemannian metric h of M. Let α be a curve and β be a reparametrization of α with $h(\beta', \beta') = 1$. Then α is called *almost closed* if there exists an $X \in \beta'(t)$ such that for every neighbourhood U of X in the tangent bundle, TM, there exists a deformation γ of β in $\pi_{TM}(U)$ which yields a closed curve and satisfies $\gamma(t) \in \pi(U) \Rightarrow \gamma'(t) \in U$.

Kriele's theorem

Theorem 1(1990)

(M, g) is causal geodesically incomplete if:

(1) $R_{ab}K^aK^b \geq 0$ for every causal vector K^a and the generic condition is satisfied.

(2) (a) there exists a closed locally spacelike but not necessarily achronal trapped surface S or (b) there exists a point r such that on every past (or every future) null geodesic from r the divergence θ of the null geodesics from r becomes negative or (c) there exists a compact achronal set S without edge.

(3) neither $cl(S, M, +)$ (respectively $cl(r, M, +)$) nor any $cl(D, M, -)$, where D is a compact topological submanifold (possibly with boundary) with $D \cap S \neq \emptyset$ (respectively $r \in D$), contains any almost closed causal curve that is a limit curve of a sequence of closed timelike curves.

This theorem is the maximum generalization of the Hawking-Penrose theorem in the sense that causality may be violated in the almost all regions except the cut locus. In this theorem causality violation seems to play a role of keeping the space-time under consideration from having singularities.

In theorem 2 below it is shown that there exist singularities when the causality violating set is compact even if there is no trapped surface. However, one cannot see which causes singularities, the compactness of the causality violating set or causality violation itself.

Theorem 2(1989)

Let (M, g) be a space-time with chronology violating set V that satisfies

(1) $R_{ab}K^aK^b \geq 0$ for every null vector K^a and the generic condition is satisfied.

(2) V has a compact closure but $M - V \neq \emptyset$.

Then V is empty or $\dot{V}$ is generated by almost closed but incomplete null geodesics.

3 Preliminaries

We consider a space-time (M, g), where M is a four-dimensional connected differentiable manifold and g is a Lorentzian and suitably differentiable metric. In this section, we quote some definitions and useful lemmas from (HE)[1] for the discussion of causal structure and space-time singularities.

Definition (HE)
A point p is said to be a *limit point* of an infinite sequence of non-spacelike curves l_n if every neighbourfood of p intersects an infinite number of the l_n's.
A non-spacelike curve l is said to be a *limit curve* of the sequence l_n if there is a subsequence l'_n of the l_n such that for every $p \in l$, l'_n converges to p.

Proposition 1 (HE 6.4.1)
The chronology violating set V of M is the disjoint union of the form $I^+(q) \cap I^-(q), q \in M$.

Lemma 1 (HE 6.2.1)
Let O be an open set and let l_n be an infinite sequence of non-spacelike curves in O which are future-inextendible in O. If $p \in O$ is a limit point of l_n, then through p there is a non-spacelike curve l which is future-inextendible in O and which is a limit curve of the l_n.

Proposition 2 (HE 4.5.10)
If p and q are joined by a non-spacelike curve $l(v)$ which is not a null geodesic they can also be joined by a timelike curve.

Proposition 3 (HE 4.4.5)
If $R_{ab}K^aK^b \geq 0$ everywhere and if at $p = \gamma(v_1)$, $K^cK^dK_{[a}R_{b]cd[e}K_{f]}$ is non-zero, there will be v_0 and v_2 such that $q = \gamma(v_0)$ and $r = \gamma(v_2)$ will be conjugate along $\gamma(v)$ provided $\gamma(v)$ can be extended to these values.

Proposition 4 (HE 4.5.12)
If there is a point r in (q, p) conjugate to q along $\gamma(t)$ then there will be a variation of $\gamma(t)$ which will give a timelike curve from q to p.

Proposition 5 (HE 6.4.6)
If M is null geodesically complete, every inextendible null geodesic curve has a pair of conjugate points, and chronology condition holds on M, then the strong causality condition holds on M.

Proposition 6 (HE 6.4.7)
If the strong causality condition holds on a compact set φ, there can be no past-inextendible non-spacelike curve totally or partially past imprisoned in φ.

Prop.5 physically means that the chronology condition is equivalent to the strong causality condition if energy conditions are satisfied.

4 The theorem

Generally, one can consider either of the following two cases in which causality violating sets and their boundaries exist. One is that there are *closed* null geodesic curves lying on the boundary or closed non-spacelike curves which pass through at least one point on the boundary. [3] The other is that there is no *closed* non-spacelike curve which passes through a point on the boundary. Here we have used the word *closed curve* in a specific sense that the curve is closed and moreover one lap length of the curve does not diverge.

We will show in each case that such a space-time has singularities in what follows.

Theorem 1
If a space-time (M, g) is null geodesically complete, then the following three conditions cannot be all satisfied together:
(a) There exists a chronology violating region V which does not coincide with the whole space-time, i.e. $M - V \neq \emptyset$,
(b) every inextendible non-spacelike geodesic in (M, g) contains a pair of conjugate points,
(c) there exists at least one point p on the boundary of V such that each closed timelike curve through a point in the $V \cap \varepsilon$ can be entirely contained in some compact set K. (ε is an arbitrary small neighbourhood of p.)

As mentioned above, if the condition (c) is satisfied, roughly speaking, one can always pick out an infinite sequence such that the one lap length of each closed timelike curve does not diverge and their shape does not change abruptly when a point on each closed curve approaches to the boundary of V.

This condition (c) is satisfied, for example, on the causality violating sets which cause *compactly generated Cauchy horizons* [7]. Causality violating sets of the Taub-NUT universe also satisfy the condition (c) because whose boundaries contain closed null geodesics. Therefore we can apply *Theorem 1* to the Taub-NUT universe, which indeed has singularities.

This condition (c) does not require that the boundary $\dot{V}$ is compact. Thus *Theorem 1* is essentially different from the Kriele's theorem 2.

Proof.
The chronology violating set V is an open set by *Prop 1*. If $V \neq \emptyset$, from the condition (a), we can find a boundary set $\dot{V}$ in $M - V$. Let us consider a sequence of points q_n in $V \cap \varepsilon$ which converges to p ($\lim_{n\to\infty} q_n = p$). By the definition of V there is a closed timelike curve l_n through q_n. From the condition (c), there exists a compact set K such that each l_n is entirely contained in $K \cap V$. Let l be a limit curve of the sequence l_n which passes through the limit point p. Choosing a suitable parameter of each l_n so that l_n is inextendible, the limit curve l is also non-spacelike inextendible curve in $K \cap \bar{V}$ by *Lemma 1*.

Let us consider the case that the limit point $p \in \dot{J}^+(q)$, $q \in V$ without loss of generality. This limit curve must also be contained in $K \cap \bar{V}$ because of the condition (c). Therefore l is totally past and future imprisoned in $K \cap \bar{V}$. If

[3] The case that whole null generators of the boundary are closed or imprisoned is similar to the situation which Hawking considered [7]. When he discussed the chronology violating sets appearing in a bounded region of general space-time without curvature singularities, he introduced the notion of *the compactly generated Cauchy horizon* defined as a Cauchy horizon such that all the past directed null geodesic generators enter and remain within a compact set. This is analogous to the existence of closed null curves on the boundary of V. He asserted that one cannot make such a Cauchy horizon while the weak energy condition is satisfied. This also supports our claims.

some point p' of l which is in the past of p is contained in V, there exists a closed non-spacelike curve but not null geodesic through p. Because one can connect the limit point p to some point c in V in the future of the p with some non-spacelike curve λ, one can always find a closed non-spacelike curve but not a null geodesic one such that $p \to c \to q \to p' \to p$ as depicted in Figure 1. This curve can be varied to a closed timelike curve through p by *Prop.2*. This contradicts with the achronality of the boundary $\dot{V}$ in which p is contained. Therefore any point of l in the past of p is not contained in V, but in the compact set $\dot{J}^+(q) \cap K$. If the null geodesic generator l of $\dot{J}^+(q)$ through p is closed, this generator has no future and past end points. Then l has pair conjugate points from *Prop.3* if l is complete. This contradicts with the achronality of the boundary $\dot{V}$. Therefore this null geodesic generator l is not closed but past imprisoned in the compact set $\dot{J}^+(q) \cap K$. Let $p_n \in \{K \cap (M - \bar{V})\}$ be an infinite sequence which converges to $p \in l$ and $r_n \in \{K \cap (M - \bar{V})\}$ be another infinite sequence such that $r_n \in \dot{J}^-(p_n)$ and converges to the point $r(\neq p)$ on l. Then one can take an infinite sequence of curves λ_n such that each of which is an inextendible null geodesic through p_n and r_n. If M is null geodesically complete, each λ_n can be extended into the open region $\{\dot{J}^-(p_n) \cap I^-(K)\}$ because each λ_n is entirely contained in $M - \bar{V}$ where the strong causality condition holds by using *Prop. 5*. Therefore, the limit curve λ of λ_n, which is an inextendible null geodesic curve through p and r from *Lemma 1*, is not imprisoned in the compact set $\{K \cap \dot{J}^+(q)\} \subset \{K \cap (M - V)\}$. However, this contradicts the fact that λ coincides with l by reparametrization of affine parameter since both of them are null geodesics through the two points p and r. Otherwise, there exists a null curve broken at p and r which is lying on $\dot{V}$, and it can be deformed to a timelike curve. This contradicts the achronality of $\dot{V}$. □

Combining *Theorem 1* and the Hawking-Penrose theorem [1], we immediately get the following corollary.

Corollary
If a space-time (M, g) is causally complete, then the following conditions cannot all hold:
(1) every inextendible non-spacelike geodesic contains a pair of conjugate points,
(2) the chronology condition holds everywhere on (M, g) or even if chronology condition is violated somewhere, such a region satisfies the condition (c),
(3) there exists a future-(or past-)trapped set S.

We have considered the case that a chronology violating set satisfies the condition (c). However, the causality violating sets in the Kerr black hole do not satisfy the condition (c). So we cannot apply our *Theorem 1* to the Kerr solution. However, we could still prove the existence of singularities if a given space-time satisfies the condition below.

Condition (c')
Let each chronology violating set be V_i. Any V_i is causally separated from $V_{j \neq i}$, i.e. $(\dot{J}^+(q) \cup \dot{J}^-(q)) \cap V_{j \neq i} = \emptyset$ for all $q \in V_i$.

For a space-time (M, g) which satisfies this condition (c') but not the condition (c), we can apply Kriele's theorem 1, taking a set S in his theorem 2 as $\dot{J}^+(q) \cap \dot{J}^-(q)$. In usual, we expect that the set $\dot{J}^+(q) \cap \dot{J}^-(q)$ is compact, which may have the topology S^2. However, there is a case that $\dot{J}^+(q) \cap \dot{J}^-(q)$ has non-compact topology. For example, in the case that $\dot{J}^+(q) \cap \dot{J}^-(q)$ has

topology $S^1 \times R$, we can regard the quotient space $e^+(\dot{J}^+(q) \cap \dot{J}^-(q))/R$ as the $e^+(S)$ in the Kriele's theorem 1, because the relevant thing in his theorem is that $e^+(S)$ is compact.

We obtain the following theorem.

Theorem 2
A space-time (M, g) which satisfies the conditions (a), (b), and either (c) or (c') is null or timelike geodesically incomplete.

As easily verified from the Penrose diagram of the Kerr solution, the condition (c') is satisfied for the Kerr solution. This theorem is applicable to the Kerr solution which indeed has singularities.

proof.
We suppose that (M, g) is null or timelike geodesically complete. We only have to prove the case that the condition (c') hold but the condition (c) does not. In such a space-time (M, g), every null geodesic generator on $\dot{V}$ is not closed.

Now we consider a non-closed null geodesic on $\dot{V}$. This null geodesic belongs to $\dot{J}^+(q)$ or $\dot{J}^-(q)$, as V is $I^+(q) \cap I^-(q)$ $(q \in V)$. Let this null geodesic belong to $\dot{J}^+(q)$ without losing generality. If this null geodesic has a past end point, it must be q. Let us take a point $p(\neq q)$ on this null geodesic and also let it be on the $\dot{V}$. Because $q \in V$, there is a closed timelike curve through q. This means that a timelike curve from q to p exists by *Prop.2*. Therefore, p belongs to $I^+(q)$. This contradicts $p \in \dot{V}$. If this null geodesic has no past end point, it is inextendible in the past. If the boundary of V contains this null geodesic entirely, from the condition (b), this boundary can be connected by timelike curves by *Prop.4*. This is also contradiction to the achronality of $\dot{V}$. Hence, let us consider the case that the boundary of V does not contain the whole segment of this null geodesic, that is, the null geodesic has an end point on the compact surface $S := \dot{J}^+(q) \cap \dot{J}^-(q)$. Extending this null geodesic beyond the future end point, we obtain an inextendible null geodesic lying on $\dot{J}^+(q)$ and call it outgoing. We also obtain an inextendible null geodesic belongs to $\dot{J}^-(q)$ and call it ingoing. The outgoing null geodesic has a pair of conjugate points from the condition (b). One of the conjugate points is on the segment lying on the $\dot{V}$. The other is on the segment lying on the $\dot{J}^+(q) - \dot{V}$. The ingoing null geodesic on the $\dot{J}^-(q)$ also has a pair of conjugate points in the same way as the outgoing case. Thus, S plays the same role as the trapped surface in the Kriele's theorem 1. From the condition (c'), the condition (3) of Kriele's theorem 1 is satisfied in the cut locuses of S, intersections of outgoing and ingoing null geodesics, because the condition (c) is not satisfied (if the condition (c) is satisfied, there exists an almost closed causal curve.). Therefore we can show the existence of singularities from Kriele's theorem 1. □

5 Conclusions and discussion

We have shown that *the boundaries of causality preserving and violating regions cause singularities in a physical space-time.*

We would like to emphasize that *Theorem 1* supplements the Hawking-Penrose theorem in the sense that the global causality condition is relaxed to some degree and instead, the condition (c) or the condition (c') is imposed on chronology violating sets. Roughly speaking, it is possible for observers to talk about the existence of singularities assuming that *our space-time has a causality*

preserving region, which conforms to our experience.

Whether the quasi-global condition (c') is removable or not is still an open question.

As well as the Hawking-Penrose theorem, our theorem cannot predict where singularities exist and how strong they are, which are left for future investigations.

Acknowledgements

We would like to thank Prof. A.Hosoya and S.Ding for useful discussions and helpful suggestions. We are also grateful to Prof. H.Kodama for kind advice and critical comments.

References

[1] S.W.Hawking and G.F.R.Ellis, *The large scale structure of space-time* (1973)

[2] F.J.Tipler, *Annals of physics* 108,1-36 (1977)

[3] F.J.Tipler, *Phys. Rev. Lett. Vol.67,No.14* 879-882 (1976)

[4] M.Kriele, *Class.Quantum Grav.6* 1607-1611 (1989)

[5] M.Kriele, *Proc.R.Soc.Lond* 431,451-464 (1990)

[6] R.P.A.C. Newman *General Relativity and Gravitation, Vol. 21, No.10* 981-995 (1989)

[7] S.W.Hawking, *Phys.Rev.D Vol.46,No.2* 603-611 (1992)

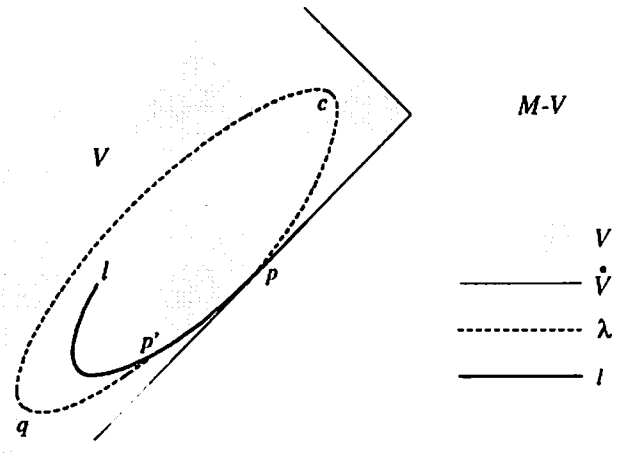


Figure 1: In the case that the past points of the limit curve l go into the V, we can find a closed non-spacelike non-geodesic curve like a $p \to c \to q \to p' \to p$, which is the union of λ and a segment $p' \to p$.

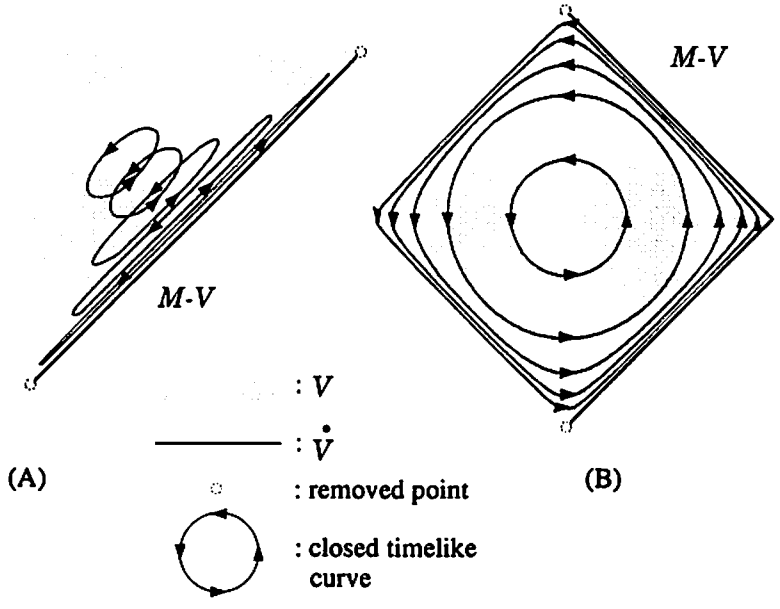


Figure 2: Examples of the space-time in which the limit curve of infinite sequence of closed timelike curves do not close.

Construction of gravitional equation system using a non-coordinate base.

National Space Development Agency of Japan,
Office of Research and Development, Noriki Iwanaga[注1]

Abstract

By examining the Jacobi identity using a non-coordinate base, we derive a support equation which replaces Bianchi identities. It makes possible to show that gravitational field has an effect similar to Faraday's electromagnetic inducting phenomena. And, it also makes possible to constitute a Lagrangian density and a source equation according to the line of gauge theoretical discussion. From this fact, it is understood that gravitational equation system of general relativity originally has a quasi-gauge theoretical structure.

1. Introduction

Generally the system of equation of the gauge fields, e.g. electromagnetic field, Yan-Mills field and gravitational field, are composed of source equation and support equation. Reviewing the derivation processes and the physical contents of the two kinds of field equations themselves, we make clear the problems on the formulation of the gravitional equation system.

To begin with, we focus on the support equations of each field. They are all derived by the same procedure viewing from the framework of gauge theory. That is, we have only to substitute commutation relation of the covariant differentiation into the Jacobi identity [1]. The support equations thus obtained have the same forms except for the difference in the differential order of the field strengths.

Next, we focus on the source equations of each field. They are derived by applying the variational principle to the action integral used the quadratic Lagrangian density on the field strengths [2]. Neither the source equation of the electromagnetic field nor that of the Yan-Mills field are the exceptions. However, the source equation of the general relativistic gravitational field, i.e. Einstein equation, is not subject to this rule. The reason why we have said so is as follows: In general relativity Riemann curvature tensor is regarded as the intensity of the gravitational field. The Lagrangian density which leads to Einstein equation is called scalar curvature density and is the first quantity on the curvature tensor. It follows from this reason that the system of equation of the general relativistic gravitational field is composed of gauge theoretical support equation and non-gauge theoretical source equation. This shows that there exists the lack of unity between two kinds of field equations in their theoretical background.

To clear this lack of unity, we may choose a quadratic Lagrangian density on the curvature tensor [3]. If we use an action integral with such a Lagrangian density, we can derive a gauge theoretical source equation through the variation principle. The source equation thus obtained becomes the differential equation of the fourth order with respect to field variable, i.e. metric tensor, while Einstein equation is of the second order. In addition to the difference of tne order, the solutions of the former source equation are considerably different from those of the latter in property. While Einstein equation explains many experimental facts, the gauge theoretical source equation of the fourth order doesn't. Therefore, the gauge theoretical source equation must not be considered as physically valid. This shows that the conventional way of constructing the gravitational equation system unified from the gauge theoretical point of view. That is, it is not a right direction to regard the curvature tensor as the field strength, to constitute the quadratic Lagrangian density on it and to deduce the gauge theoretical source equation of the gravitional field instead of Einstein equation. This conclusion is justified by a difficulty in the quantization of the R^2 theory [4]. The difficulty is that, in the theory using the quadratic Lagrangian density on the curvature tensor, the ghost of Planck mass appears and the unitarity

[注1] Sengen 2-1-1, tukuba city, ibaraki 305 e-mail; iwanaga@rd.tksc.nasda.go.jp

breaks perturbationally.

In addition to the problem on the formulation, the system of equation of the general relativistic gravitational field has another difficulty. Another difficulty is the hardness to grasp of the physical content of the equations themselves.

It roughly says that the field equations give two kinds of information. One is information about the behaviour of the field in giving the distribution of the conservative quantities (charge and mass, etc.) as a source. Another is information read directly from the forms of the field equations. This information is proper to the fields themselves, not relevant for the way of the distribution of the conservative quantity. The contents are, for example, the way of the interaction with other objects and/or the dynamical behaviour of the fields themselves. The system of equation of the electromagnetic field has a sufficient form to give us both information. However, though the system of equation of the general relativistic gravitational field can tell us the first information, it can not tell us the second. For example, Faraday's electromagnetic inducting phenomena can be read directly from Maxwell's magnetic equation. On the other hand, similar inducting phenomena can not be read directly from the gravitational support equation, i.e. Bianchi identities.

Here again, we mention clearly the problems on the conventional formulation on the gravitational equation system :
(1) General relativistic source equation, i.e. Einstein equation, is not derived from the Lagrangian density constructed along the gauge theory. The source equation, which is made from the quadratic Lagrangian density on the field strength in order to overcome above mentioned defect, does not have the physical validity.
(2) Neither the general relativistic equation system nor the equation system constituted within the framework of the gauge theory can foretells the physical property of the gravitational field directly from their forms.

To begin with, we consider the method of solving the problem (1). We know that the graviational field is invariant under translation and local Lorentz transformation. Since tetrad vector is covariant under both transformations, we regard it as a basic variable of the field [5]. And, we shall maintain the gauge theoretical principle that the Lagrangian density has quadratic part of the field strength and the part gives the source equations. Then the freedom left us is the way of selecting the field strength and divergence term which should be added to the Lagrangian density. The way of selecting the field strength is limited to the quantity of the first order lower than the curvature tensor such as non-holonomic object and Ricci's coefficient. It is because the appropriate source equation is of second order from the fact that though Einstein equation of the second order has physical validity, the fourth order differential equation mentioned above doesn't have it. Then we make a quadratic expression using such a quantity selected as the field strength and add some divergence terms to the expression in order to recover the broken invariance. Further if we request to an action integral using such a Lagrangian density on the invariance, we will obtain the Lagrangian density that leads to the source equation with physical validity. Therefore, the key point is a possibility to discuss the innevitable derivation of the field strength desired.

Next, we consider the method of solving the problem (2). It is not reasonable that while the Maxwell's magnetic equation can tell us the behaviour of the electromagnetic field, only the system of equation of the gravitational field equation systen can not sufficiently do. If we aquire the appropriate forms of every field equation system, we should be able to read the dynamical behaviour of every field directly from it. Therefore, we should rewrite the conventional form of the gravitational field equation system into a desired form. The reasons for thinking above is as follows ; First, the electromagnetic field is considered as one physical aspect belonged to space-time itself which appears when a conservative quantity of charge is put in space-time. Similarily the gravitational field is considered as another physical aspect belonged to space-time itself which appears when a conservative quantity of energy is put in space-time. It is possible, therefore, to regard each field as an embodiment of the different physical properties belonged to one and the only space-time respevtively. From this, we intuitively imagine that the field equation systems should be able to formulated in

such forms as the common property in both fields remains. To put it concretely, if we adopt an adequate quantities of the gravitational field and reconstitute the field equation system in a form which is similar to that of the electromagnetic equation system, then directly from it we should be able to read the dynamical behaviour of the gravitational field. This expectation is backed up from the viewpoint of the quantum-theorety on the gauge field. In a series of paper [6] on the covariant canonical quantization of the gravitational field, Nakanishi shows that the gravitational field equation system has a mathematical structure analogous not to that of the Yan-Mills field but to that of the electromagnetic field. This fact can be considered to suggest that the possibility of constituting the gravitational field equation system in the same way of constituting the electromagnetic one before the quantization.

The guiding principles of solving the problem (1) and (2) are as follows: First, we extract from the theoretical structure the quantity which is regarded as the field strength and only first order lower than the curvature tensor. Second, we constitute a Lagrangian density which has quadratic part of this quantity as well as divergence part and is invariant under the translation and the local Lorentz transformation. Third, we complete the formulation by deriving the source equation with a form similar to that of the electromagnetic equation system.

To begin with, we priorly look for the quantity which is considered as the field strength and has a Faraday tensor-like form. It will be good to discuss using a base vector in which the desired geometric quantities covariant under two transformations, e.g. tetrad vector, appear naturally. (The tetrad vector is a basic quantity of the field). Such a base is called non-coordinate base and reduces to a usunal coordinate base in the limit [7]. Then in the starting point in deriving the field equation system, we shall introduce the non-coordinate base. Namely, we formulate the gravitational field equation system using the non-coordinate base from the begining.

In chapter 2, the derivation process of the Bianchi identities as a support equation is analized. Since the Bianchi identities are originated from the Jacobi identity, we examine the latter identity using the non-coordinate base. As a result, a new support equation is expected to deduce whose form is similar to that of Maxwell's magnetic equation. By rewriting it in the three-dimensional vector form, it becomes possible to direcly foretell the inducting behavior of the gravitational field. And, that expression tells us that the non-holonomic object plays an important role as a quantity which reflects some property of the gravitational field.

In chapter 3, we reset the way of deciding the field strength in reference to the result of chapter 2 and adopt the non-holonomic object as the field strength, and then start the formulation. By applying invariant variational theory to a Lagrangian density constituted by the non-holonomic object, we decide the form of the Lagrangian density with the invariance under translation and local Lorentz transformation. Further we derive the source equation from the variational principle and make a comparison with Einstein equation. As a result, it is found that our formulation with appropriate parameters agrees with that of the general relativity. That is to say, our formulation is found to be physically valid.

2-1. Derivation of support equation

Viewing the derivation process of Bianchi identities from the gauge theoretical point of view, we can recognize that they are derived from the Jacobi identity

$$0 \equiv [D,[D,D]]\psi + [D,[D,D]]\psi + [D,[D,D]]\psi \qquad (2\cdot1).$$

Here the ψ is supposed to be any vector field and the $D\psi$ be its covariant differentiation in convenience. The First and second Bianchi identities described with the component of a coordinate base,

$$R_{[\lambda\mu]}{}^{\alpha} \equiv 0 \qquad (2\cdot2)$$

$$D_{[\lambda}R_{\mu]\alpha}{}^{\beta} \equiv 0 \qquad (2\cdot3),$$

are deduced from the Jacobi identity shown with the component of the coordinate base

$$0 \equiv [D_\lambda,[D_\mu,D_\nu]]\psi^\alpha + [D_\mu,[D_\nu,D_\lambda]]\psi^\alpha + [D_\nu,[D_\lambda,D_\mu]]\psi^\alpha$$

where

$$D_\mu\psi^\alpha = (\delta_\beta{}^\alpha\partial_\mu + \left\{{}^{\alpha}_{\mu\beta}\right\})\psi^\beta$$

The first important point is that (2·1) holds on a special base called local Lorentz base. This base constitutes some local Lorentz systems with Minkowski metric tensor through the relation : $e_i \cdot e_j = \eta_{ij}$, and that the base is in a cathegory of non-coordinate basis.

Even more important is that (2·1) holds on *any* non-coordinate basis $e_j = (\delta_j{}^m\partial_m)$ (i, j, m=0,1,2,3) as well as on the local Lorentz basis. Therefore, we examine the derivation process of the Bianchi identities from the Jacobi identity on a non-coordinate base.

Before turning to a closer examination, we look over the properties of the non-coordinate basis briefly. Since a non-coordinate base reduces to a coordinate base in the limit, the former has more menerality than tha latter. And the former gives the geometrical objects called metric tensor and non-holonomic object :

$$g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j, \qquad g^{ij} = \mathbf{e}^i \cdot \mathbf{e}^j$$

$$\Omega_{\mu\nu}{}^i = \partial_{[\mu}e_{\nu]}{}^i$$

The non-coordinate base under consideration is assumed to the following orthogormal relation in convenience.

$$e_i{}^\lambda e^j{}_\lambda = \delta_i^j, \qquad e_m{}^\alpha e^m{}_\beta = \delta_\beta^\alpha$$

Let us now return to the subject. Using the non-coordinate base, (2·1) gives

$$0 \equiv [D_k,[D_i,D_j]]\psi^h + [D_i,[D_j,D_k]]\psi^h + [D_j,[D_k,D_i]]\psi^h$$

where

$$D_i\psi^k = (\delta_i{}^k\partial_i + \Gamma_{ai}{}^k)\psi^l$$

If we substitute Ricci's relation

$$[D_i,D_j]\psi^k = R_{ijl}{}^k\psi^l$$

into above equation, we obtain

$$0 \equiv R_{[kji]}{}^m D_m\psi^h + D_{[k}R_{ij]m}{}^h\psi^m \qquad (2\cdot4).$$

Usually we make use of independency of $\psi^m$and $D_m\psi^h$ here and extract from (2·4) the Bianchi identities

$$R_{[kji]}{}^m \equiv 0 \qquad (2\cdot5)$$

$$D_{[k}R_{ij]m}{}^h \equiv 0 \qquad (2\cdot6)$$

written with the component of the non-coordinate base. However, we do not take such a usual procedure but proceed to calculate (2·4). First, we substitute

$$R_{kji}{}^h = 2\partial_{[k}\Gamma_{j]i}{}^h + 2\Gamma_{[kl|m|}{}^h\Gamma_{j]i}{}^l + 2\Omega_{kj}{}^m\Gamma_{mi}{}^h$$

into (2·4) and make an expression with respect to a connection coefficient $\Gamma_{ji}{}^k$. Secondly, substituting

$$\Gamma_{ji}{}^h = \left\{{}^{h}_{ji}\right\} - \Omega_{ji}{}^h + g_{ji}g^{hk}\Omega_{ik}{}^l + g_{il}g^{hk}\Omega_{jk}{}^l$$

into the result derived above, we obtain

$$0 \equiv -2(\partial_{[l}\Omega_{kj]}{}^m - 2\Omega_{[sn|}{}^m\Omega_{kj]}{}^n)\partial_m\psi^h$$

From above expression, we obtain

$$\partial_{[l}\Omega_{kj]}{}^{h} - 2\Omega_{[lm]}{}^{h}\Omega_{kj]}{}^{n} = 0$$

or equivalently,

$$\partial_{[\omega}\Omega_{\nu\mu]}{}^{i} = 0 \qquad (2\text{-}7).$$

These are the identities deduced from the Jacobi identity. They also become a necessary and sufficient condition for (2-8) and (2-9) mathematically. Therefore, they can play a role of support equation instead of the Bianchi identities.

It is very difficult to know the behaviour of the gravitational field directly from the Bianchi identity. However, since the form of the support equation (2-7) is similar to that of Maxwell's magnetic equation

$$\partial_{[\lambda}F_{\mu\nu]} = 0 \qquad (2\text{-}8)$$

which is the support equation of the electromagnetic field, it is easy to investigate the physical content of the gravitational field. If we substitute the definition of the non-holonomic object

$$(\Omega_{\mu\nu}{}^{h}) = \begin{bmatrix} \Omega_{00}{}^{h} & \Omega_{0\alpha}{}^{h} \\ \Omega_{\alpha 0}{}^{h} & \Omega_{\alpha\beta}{}^{h} \end{bmatrix} = \frac{1}{2c^2}\begin{bmatrix} 0 & -a^{(h)}{}_{\beta} \\ a^{(h)}{}_{\beta} & b^{(h)}{}_{\alpha\beta} \end{bmatrix} \qquad (2\text{-}9)$$

into (2-7) and act a quantity

$$\varepsilon^{\gamma\alpha\beta} = \frac{1}{\sqrt{\gamma}}\delta_{\gamma\alpha\beta}$$

$$\gamma = \det(\gamma_{\alpha\beta}) \quad \gamma_{\alpha\beta} = g_{\alpha\beta} - \frac{g_{0\alpha}g_{0\beta}}{g_{00}}$$

we obtain

$$div\,\mathbf{b}^{(h)} = 0 \qquad (2\text{-}10)$$

$$\frac{1}{\sqrt{\gamma}}\frac{\partial}{c\partial t}(\sqrt{\gamma}\,\mathbf{b}^{(h)}) + rot\,\mathbf{a}^{(h)} = 0 \qquad (2\text{-}11).$$

Here

$$\mathbf{b}^{(h)} = (b^{(h)\alpha}) = (\frac{1}{2}\varepsilon^{\alpha\beta\gamma}b^{(h)}{}_{\beta\gamma}) = (\frac{1}{2\sqrt{\gamma}}\delta_{\alpha\beta\gamma}b^{(h)}{}_{\beta\gamma})$$

$$rot\,\mathbf{a}^{(h)} = (rot^{\gamma}\mathbf{a}^{(h)}) = \left(\frac{1}{2}\varepsilon^{\gamma\alpha\beta}(\partial_{\alpha}a^{(h)}{}_{\beta} - \partial_{\beta}a^{(h)}{}_{\alpha})\right) = \left(\frac{1}{\sqrt{\gamma}}\delta_{\gamma\alpha\beta}\partial_{\alpha}a^{(h)}{}_{\beta}\right)$$

$$rot\,\mathbf{a}^{(h)} = (rot^{\gamma}\mathbf{a}^{(h)}) = \left(\frac{1}{2}\varepsilon^{\gamma\alpha\beta}(\partial_{\alpha}a^{(h)}{}_{\beta} - \partial_{\beta}a^{(h)}{}_{\alpha})\right) = \left(\frac{1}{\sqrt{\gamma}}\delta_{\gamma\alpha\beta}\partial_{\alpha}a^{(h)}{}_{\beta}\right)$$

Since both (2-10) and (2-11) are have considerably similar to the three-dimensional vector equations of (2-8)

$$div\,\mathbf{B} = 0$$

$$\frac{\partial\mathbf{B}}{c\partial t} + rot\,\mathbf{E} = 0$$

in their forms, the same physical content between the gravitational field and the electromagnetic field is expected.

The $\mathbf{b}^{(h)}$ field in (2-10) corresponds to magnetic flux density field B. (2-10) means that there exists no material source which gives the divergence of the $\mathbf{b}^{(h)}$ field. However, as is seen from the source equation (3-20), material energy-momentum which corresponds to the flow state of real material source gives $\mathbf{b}^{(h)}$ as well as $\mathbf{a}^{(h)}$. From this fact, (2-10) shows that gravitional monopole does not exist in the sense that the "static"

state of real material source which gives the $b^{(h)}$ field by itself does not exist .

(2·11) means that the spatial change of the $a^{(h)}$ field is induced by the time variation of the $b^{(h)}$ field. This makes us expect that in the matter system moving in the gravitational field there appears the gravitational effect due to the $b^{(h)}$ and $a^{(h)}$ fields. This is similar kind of phenomena to Faraday's electromagnetic induction appered in the matter system moving in electromagnetic field .

(2·7) or (2·10) and (2·11) give us the direct physical meanings above mentioned thanks to the style analogous to that of Maxwell's magnetic equation. In addition to it, however, they indirectly indicate the possibility to exist local energy-momentum of the gravitational field. We can derive the Maxwell's magnetic equation as a necessary condition to hold the energy conservation law when the energy distribution of the magnetic field changes in a space region under consideration and some mount of energy outgoes[8]. From this fact, (2·7) may be considered to become the necessary condition to hold the conservation law when we examine the conservation law under the premise of the local energy-momentum tensor of the gravitational field[12] being defined. That is to say, (2·7) indicates that the flow of energy requested by the conservation law exists when the energy of the gravitational field changes with time. To put it the other way round, the existence of (2·7) shows the possibility of giving the definition of the local energy-momentum tensor which satisfies the energy-momentum conservation law of the gravitational field.

If there supposed to exist a resemblance of physical meanings between (2·7) and (2·8), there should exist a resemblance of physical meanings between the $\Omega_{\mu\nu}{}^{i}$ (or $a^{(h)}$ and $b^{(h)}$) field in (2·7) and the $F_{\mu\nu}$ (or electric field E and magnetic flux density field B) field in (2·8). In order to confirm this, we compare the equation of motion of a charged particle moving in the electromagnetic field

$$\frac{dE}{dt} = e(\mathbf{E}\cdot\mathbf{v}) \tag{2·12}$$

$$\frac{d\mathbf{p}}{dt} = e(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}) \tag{2·13}$$

with that of a particle moving in the gravitational field

$$\frac{dE}{dt} = -m_0\mathbf{a}^{(0)}\cdot\mathbf{v} \tag{2·14}$$

$$\frac{d\mathbf{p}}{dt} = m_0(-\mathbf{a}^{(0)} - \frac{\mathbf{v}}{c}\times\mathbf{b}^{(0)}) + m_0(-\frac{\mathbf{v}}{c}\times\mathbf{a} + \mathbf{a}_v) \tag{2·15}$$

where

$$\mathbf{a} = (a^{\alpha}) = (\frac{1}{2}\varepsilon^{\alpha\beta\gamma}a_{\beta,\gamma}) \qquad \mathbf{a}_v = (\mathbf{a}_{v\alpha}) = (a_{(\alpha,\beta)}\frac{v^{\beta}}{c})$$

The particles above introduced have 4-momentum P_i=(·E/c, p) and 4-velosity U^i=(c, v). We can derive (2·14) and (2·15) by three steps: The first step is to substitute (2·9) into the following expression

$$\frac{dP_i}{d\tau} = 2\Omega_{i(j,k)}P^kU^j$$

This is an equation of motion described on a Lorentz base stated before. The second step is to carry out a weak approximation such as

$$\left|v^{\alpha}/c\right| \le 1$$

$$e^{i}{}_{\mu} = \delta^{i}{}_{\mu} + h^{i}{}_{\mu} \qquad \textit{where} \qquad \left|h^{i}{}_{\mu}\right| \le 1$$

[12] As we see in chapter 3, (2·7) becomes a necessary condition to derive the Lagrangian density and the source equation when we reconstitute the general relativity according to the step of gauge theory.

The third step is to obtain the following 3-D vector version using a symmetrization:

$$\frac{dE}{dt} = -m_0 a^{(0)}{}_\alpha v^\alpha$$

$$\frac{dp_\alpha}{dt} = m_0(-a^{(0)}{}_\alpha - b^{(0)}{}_{\alpha\beta}\frac{v^\beta}{c}) + m_0(-a_{[\alpha,\beta]}\frac{v^\beta}{c} + a_{(\alpha,\beta)}\frac{v^\beta}{c})$$

(2·14) and (2·15) are arised from these equations.

The E and v/c×B terms in (2·12) and (2·13) indicate electric force and Lorentz force respectively acting on the charged particle which is moving in the electromagnetic field. On the other hand, the $a^{(0)}$ and $v/c \times b^{(0)}$ terms indicate two kinds of meanings respectively. One is Newtonian force and gravitional magnetic force (i.e. Lense·Thiring effect) when the particle is moving in the gravitational field. Another is centrifugal force and Corioli's force when the particle exists in a noninertial frame. The $a_{(\alpha,\beta)}v^\beta/c$ term indicates the shearing and expanding effects on the particle. This can be understood if $a_{(\alpha,\beta)}$ is resolved into a traceless part and a trace part

$$a_{(\alpha,\beta)} = \hat{a}_{\alpha,\beta} + \frac{1}{3}a g_{\alpha\beta}$$

$$\hat{a}^\alpha{}_\alpha = 0 \qquad a = a^\alpha{}_\alpha$$

The gravitational effect due to the $a_{(\alpha,\beta)}$ term can not be extracted explicitly when the equation of motion is handled with metric tensor as a field variable.

Now we find that $\Omega_{\mu\nu}{}^0$ ($a^{(0)}$ and $b^{(0)}$) corresponds to $F_{\mu\nu}$(E and B), but $\Omega_{\mu\nu}{}^c$ ($a^{(c)}$ and $b^{(c)}$) does not correspond to any quantity with respect to the electromagnetic field. And that, while the relations

$$div\, b^{(0)} = 0$$

$$\frac{1}{\sqrt{\gamma}}\frac{\partial}{c\partial t}(\sqrt{\gamma}\, b^{(0)}) + rot\, a^{(0)} = 0$$

correspond to Maxwell's magnetic equation (2·8), the relatoins

$$div\, b^{(c)} = 0$$

$$\frac{1}{\sqrt{\gamma}}\frac{\partial}{c\partial t}(\sqrt{\gamma}\, b^{(c)}) + rot\, a^{(c)} = 0$$

don't correspond to any relations about electromagnetic field. The $\Omega_{\mu\nu}{}^c$ ($a^{(c)}$ and $b^{(c)}$) show that the dynamic effect of the gravitational field is more various than that of the electromagnetic field. However, it is interesting that the behaviour of $\Omega_{\mu\nu}{}^c$ is regulated by the same equation as that of $\Omega_{\mu\nu}{}^0$ ($a^{(0)}$ and $b^{(0)}$) in their forms.

Finally we mention the usability of the newly obtained support equation. We have discussed that the non·holonomic object describes the electromagnetic·like effect of the gravitational field. However, the discussion on the electromagnetic·like effect of the other quantities has been done in the past. For example, the content of a paper [9] is recollected which is written by Christian and Sachs et al. The following is pointed out there.
(1)If there is the Wyle curvature tensor in a cosmological model, the shadow of the light source which globularly spreads should traject elliptically not circularlly.
(2) There exists an observational evidence to support this in our space.
And then, electric field and magnetic field part of the Wyle curvature tensor are introduced as the quantity which explains this strain effect.

It is physically interesting to investigate the possibility to describe the strain effect without the electric field and magnetic field part of the Wyle curvature tensor by using the quantity $\Omega_{\mu\nu}{}^{0}$ appeared in expression

$$\partial_{[\lambda}\Omega_{\mu\nu]}{}^{0} \equiv 0$$

3. Derivation of source equation

As we mentioned in chapter 1, Einstein equation

$$G^{\mu\nu} = kT^{\mu\nu} \tag{3·1}$$

has a physical validity but doesn't have a status as a source equation followed to the gauge theory, if we regard the curvature tensor $R_{\nu\mu\lambda}{}^{\kappa}$ as a field strength through the conventional support equation, i.e. Bianchi identities. It is because that the Einstein equation is derived from Scalar curvature density

$$\mathbf{R} = \delta_{\beta}{}^{\mu}\sqrt{-g}\,g^{\alpha\nu}R_{\mu\nu\alpha}{}^{\beta}$$

which is the linear Lagrangian density on the curvatur tensor.

On the other hand, the source equation derived from the quadratic Lagrangian density on the curvatur tensor such as

$$\mathbf{L} = \sqrt{-g}(aR^{2} + bR^{\mu\nu}R_{\mu\nu} + cR^{\mu\nu\alpha\beta}R_{\mu\nu\alpha\beta} + \cdots\cdots)$$

has a status as a source equation which is constituted according to the gauge theory. However, it doesn't have physical validity oppsite to the Einstein equation. The aim of this chapter is thus to clear this contradiction. Namely, we derive a source equation which is constructed according to the step of gauge theory and is also physically valid.

As we showed in chapter 2, we derived the support equation (2·13) which replaces the Bianchi identities as a result of consideration on the Jacobi identity using the non-coordinate base. This equation represents us a viewpoint to regard the tetrad vector $e^{i}{}_{\lambda}$ and non-holonomic object $\Omega_{\mu\nu}{}^{k}$ as a basic variable and a field strength respectively. Thus we change the conditions that (1) the field strength satisfies the support equation derived from the Jacobi identity and (2) it is defined by commutation relation between the covariant differential operators and (3) it is covariant under the general coordinate transformations, to those that (1) the field strength satisfies the support equation derived from the Jacobi identity and (2)' it is defined by commutation relation between the "non-holonomic" differential operators and (3)' it is invariant under the "holonomic" coordinate transformations. After that, we formulate the Lgrangian density and the source equation on the standpoint that the field strength is the non-holonomic object $\Omega_{\mu\nu}{}^{k}$.

To begin with, we constitute a Lagrangian density such as

$$\begin{aligned} \mathbf{L}_{R} &= \mathbf{L}_{G} + \partial_{\mu}\tilde{\mathbf{d}}^{\mu} \\ \mathbf{L}_{G} &= E(\alpha\Omega^{ijk}\Omega_{ijk} + \beta\Omega^{ijk}\Omega_{jik} + \gamma\Omega^{i}\Omega_{i}) \\ \tilde{\mathbf{d}}^{\mu} &= Ee_{i}{}^{\mu}(\varepsilon\Omega^{ji}{}_{j} + \rho\delta^{ijl}\Omega_{jil}) \end{aligned} \tag{3·2}$$

according to the gauge theoretical procedure in the sense mentioned in chapter 1. Here $\alpha, \beta, \gamma, \varepsilon, \rho$ are constant parameters. (3·2) includes quadratic part $\mathbf{L}_{G}$ and divergence part of the field strength, i.e. non-holonomic object $\Omega_{\mu\nu}{}^{k}$. Since we adopt the non-holonomic object as the field strength, the $\mathbf{L}_{G}$ term itself is not invariant under the local Lorentz transformation. The divergence part should be added to $\mathbf{L}_{G}$ in order to recover this

destroyed invariance and constitute a desired Lagrangian density L_R. Why is then the added term limited to the form of divergence? It is because we intend to recover the invariance according to the gauge theoretical rule that the source equation should be derived from the quadratic part of the Lagrangian density.

Next, we request to the Lagrangian density (3·2) the invariance under the parity transformation of the gravitational field. Then (3·2) may be replaced by

$$L_R = L_G + \partial_\mu d^\mu \qquad d^\mu = \varepsilon E \Omega^\mu \tag{3-3}$$

Furthermore, we deduce the condition between constant parameter $\alpha, \beta, \gamma, \varepsilon$ in order to make (3·3) a more suitable Lagrangian density. For this purpose, we use groups of identities which are derived from the invariance under translation and local Lorentz transformation of the gravitational field. The groups of identities are obtained from the request the invariance of action integral

$$S = \int L_R d^4x$$

under the Poincare gauge transformation[10][13]

$$\delta x^\mu = x'^\mu - x^\mu = \varepsilon^\mu(x)$$

$$\delta e^k{}_\lambda = -\varepsilon^\mu{}_{,\lambda} e^k{}_\mu + \omega^{mn} i \frac{1}{2} (S^l{}_{mn})^k{}_l e^l{}_\lambda$$

As a result of calculation, we obtain the following identities:

$$\varepsilon^\nu : \quad D_\mu {}^G B_k{}^\mu = 0 \tag{3-4}$$

$$\omega^{mn} : \quad {}^G B_{[mn]} = 0 \tag{3-5}$$

$$\varepsilon^\nu : \quad \partial_\mu B^\mu{}_\nu = 0 \tag{3-6}$$

$$\varepsilon^\nu{}_{,\mu} : \quad B^\mu{}_\nu + \partial_\lambda C^{[\lambda,\mu]}{}_\nu = 0 \tag{3-7}$$

$$\omega^{mn} : \quad \partial_\mu P^\mu{}_{mn} = 0 \tag{3-8}$$

$$\omega^{mn}{}_{,\mu} : \quad P^\mu{}_{mn} + \partial_\nu Q^{[\nu,\mu]}{}_{mn} = 0 \tag{3-9}$$

where

$${}^G B_k{}^\mu = [L_G]_k{}^\mu = \frac{\delta L_G}{\delta e^k{}_\mu} \tag{3-10}$$

$$D_\mu = \partial_\mu - i \frac{1}{2} S^l{}_{mn} \Gamma_\mu{}^{mn}$$

$$B^\mu{}_\nu = {}^G B_\nu{}^\mu + {}^G \tilde{t}_\nu{}^\mu + \partial_\lambda (2\delta^{[\lambda}{}_\nu d^{\mu]})$$

$${}^G \tilde{t}_\nu{}^\mu = \frac{\partial L_G}{\partial \partial_\mu e^k{}_\lambda} \partial_\nu e^k{}_\lambda - \delta^\mu{}_\nu L_G$$

$$C^{\mu,\lambda}{}_\nu = \frac{\partial L_G}{\partial e^k{}_{\lambda,\mu}} e^k{}_\nu + \frac{\partial d^\mu}{\partial e^k{}_\lambda} e^k{}_\nu + \frac{\partial d^\mu}{\partial e^k{}_{\alpha,\nu}} e^k{}_{\alpha,\nu} + \frac{\partial d^\mu}{\partial e^k{}_{\lambda,\alpha}} e^k{}_{\nu,\alpha}$$

---

[13] $\varepsilon^\mu(x)$ and $\omega^{mn}(x) = -\omega^{nm}(x)$ are the infinitesimal functions related to translation and rotation respectively. The quantity $(S^l{}_{mn})^k{}_l$ is defined by

$$(S^l{}_{mn})^k{}_l = -i 2 \delta^k{}_{[m} \eta_{n]l}$$

$$\mathbf{P}^{\mu}{}_{mn} = \frac{\partial \mathbf{L}_G}{\partial e^{k}{}_{\lambda,\mu}} e^{l}{}_{\lambda} i\frac{1}{2}(S^{1}{}_{mn})^{k}{}_{l} + \frac{\partial \mathbf{d}^{\mu}}{\partial e^{k}{}_{\lambda}} e^{l}{}_{\lambda} i\frac{1}{2}(S^{1}{}_{mn})^{k}{}_{l} + \frac{\partial \mathbf{d}^{\mu}}{\partial e^{k}{}_{\lambda,\nu}} e^{l}{}_{\lambda,\nu} i\frac{1}{2}(S^{1}{}_{mn})^{k}{}_{l}$$

$$\mathbf{Q}^{\mu,\nu}{}_{mn} = \frac{\partial \mathbf{d}^{\mu}}{\partial e^{k}{}_{\lambda,\nu}} e^{l}{}_{\lambda} i\frac{1}{2}(S^{1}{}_{mn})^{k}{}_{l}$$

(3·4), (3·6) and (3·7) are the identities originated from the translational invariance of the gravitational field, while (3·5), (3·8) and (3·9) are originated from the local Lorentz invariance. If we use field equation

$$^{G}\mathbf{B}_{\nu}{}^{\mu} = \mathbf{T}_{\nu}{}^{\mu} \tag{3·11}$$

we obtain conservation laws of energy·momentum and angular·momentum of matter from (3·4) (or (3·6)) and (3·5) (or (3·8)) respectively.

In order to fix the parameters α–ε, we use the relation

$$-(2\alpha-\beta)\Omega^{i}{}_{[mn]} + (\beta - \frac{1}{2}\varepsilon)\Omega_{mn}{}^{i} - (\gamma+\varepsilon)\delta^{i}{}_{[m}\Omega_{n]} \equiv 0 \tag{3·12}$$

calculated from (3·9), and the relation

$$E\{-(2\alpha-\beta)(\partial_{i}+2\Omega_{i})\Omega^{i}{}_{[mn]} - (2\beta+\gamma)\partial_{[m}\Omega_{n]}\} \equiv 0 \tag{3·13}$$

calculated from (3·5) (or (3·8)). Since (3·12) gives

$$2\alpha+\beta+3\gamma+2\varepsilon \equiv 0$$

and (3·13) gives

$$2\alpha-\beta \equiv 0 \qquad 2\beta+\gamma \equiv 0$$

we can fix the following ratio between the parameters:

$$\beta = 2\alpha \qquad \gamma = -4\alpha \qquad \varepsilon = +4\alpha \tag{3·14}$$

From (3·3) and (3·14), we obtain a poincare·gauge invariant Lagrangian density

$$\mathbf{L}_R = \alpha E(\Omega^{ijk}\Omega_{ijk} + 2\Omega^{ij}{}_{j}\Omega_{ik} - 4\Omega^{i}\Omega_{i}) + 4\alpha\partial_{\mu}(E\Omega^{\mu}) \tag{3·15}$$

If we fix the parameter α, we must compare the source equation to be derived below with Newtonian approximated equation or experimental result.

Substituting (3·15) into (3·10) and (3·11), we obtain a source equation

$$\partial_{\mu}\mathbf{H}_{l}{}^{\lambda\mu} + 2\Omega_{\mu l}{}^{\nu}\mathbf{H}_{\nu}{}^{\lambda\mu} - e_{l}{}^{\lambda}\mathbf{L} = \mathbf{T}_{l}{}^{\lambda} \tag{3·16}$$

from variational principle. This is the same source equation as the Riemannian limit of that which has been derived in the conventional Poincare·gauge theory[3]. (3·16) is gained from the variational equation

$$\frac{\delta \mathbf{L}}{\delta e^{i}{}_{\mu}} = \frac{\partial \mathbf{L}}{\partial e^{i}{}_{\mu}} - \partial_{\nu}\frac{\partial \mathbf{L}}{\partial \partial_{\nu} e^{i}{}_{\mu}} = 0$$

with tetrad vector as a basic variable. On the other hand, if we use the variational equation

$$\frac{\delta \mathbf{L}}{\delta e_{a}{}^{\mu}} = \frac{\partial \mathbf{L}}{\partial e_{a}{}^{\mu}} - (\partial_{b}+2\Omega_{b})\frac{\partial \mathbf{L}}{\partial \partial_{b} e_{a}{}^{\mu}} = 0$$

with the reciprocal vector as a basic variable, we obtain

$$\partial_{j}H_{l}^{mj} + H_{l}^{ij}\Omega_{ij}{}^{m} + 2H_{j}^{mi}(\Omega_{i}{}^{j} + \Omega_{i}\delta_{l}{}^{j}) - \delta_{l}{}^{m}L = T_{l}^{m} \tag{3·17}$$

where

$$H_{l}^{ij} = 2\alpha(-\Omega^{ij}{}_{l} + 2\Omega_{l}{}^{[ij]} - 4\delta_{l}{}^{[i}\Omega^{j]m}{}_{m}) \tag{3·18}$$

It is natural that (3·16) and (3·17) are equivalent.

Because of the tetrad being the basic variable, (3·17) can be resolved into antisymmetrized equation

$$0 = T_{[mn]} \tag{3·19}$$

and symmetrized equation

$$4\alpha\{(\partial_i + 2\Omega_i)\Omega^i{}_{(m.n)} - \eta_{mn}(\partial_i + \Omega_i)\Omega^i + \partial_{(m}\Omega_{n)}$$

$$-\frac{1}{2}\Omega^{pq}{}_m\Omega_{pq.n} + 2\Omega_{mp.q}\Omega_n{}^{(p.q)}\} - \eta_{mn}L = T_{(mn)} \tag{3·20}$$

It is easy to understand from (3·5) that the left handside of (3·19) vanish. This means that angular-momentum of matter doesn't reflect the geometry of space-time.

Both (3·20) and Einstein equation are the differential equations with the second order. Therefore, as a first step to examine the physical validity of (3·20), it is important to compare it with Einstein equation. Tetrad formed Einstein equation is

$$G_{(ij)} = -2\partial_k\Omega^k{}_{(ij)} - 2\partial_{(i}\Omega_{j)} - 4\Omega_i{}^{kl}\Omega_{j(kl)} + \Omega^{kl}{}_i\Omega_{kj} - 4\Omega^k{}_{(i}\Omega_{j)}$$

$$+\eta_{ij}(2\partial_k\Omega^k + 2\Omega^k\Omega_k + \frac{1}{2}\Omega^{klm}\Omega_{klm} + \Omega_{klm}\Omega^{mlk}) = \kappa T_{(mn)}$$

Therefore, if we adopt

$$\alpha = -\frac{1}{2\kappa}, \qquad \beta = -\frac{1}{\kappa}, \qquad \gamma = \frac{2}{\kappa} \tag{3·21}$$

both equations agree with in their forms and have the same physical contents. This shows that the souce equation (3·20) with (3·21) is constructed according to the gauge theoretical procedure in the sense mentioned in Chapter 1, and has physical validity. To put it the other way round, this means that the general relativistic equation system itself, i.e. the support equation (2·13) plus Einstein equation (3·1), is originally constructed according to the quasi-gauge syructure in the same sense.

4. Concluding remarks

As a result of examining the Jacobi identity by introducing the non-coordinate base, we obtained the support equation (2·13) which replaced the Bianchi identities. With respect to (2·13), we investigated two kinds of things stated below.

First, we intended to examine the gravitational effect which were deduced from (2·13). For this purpose, we gave the 3-D vector forms of (2·13) and equation of motion of a particle to compare with the corresponding electromagnetic equations. It followed from this comparison that as there existed an inducting effect in the electromagnetic field, so did in the gravitational field. This is important in that we found a gravitational behaviour from the support equation itself without examining the solution of the source equation.

Secondly, we took the standpoint to regard the tetrad vector and the non-holonomic object as a basic variable and a field strength respectively on the grounds of the support equation (2·13) and we formulated the Laglangian density and the source equation according to the step of gauge theory. To put it concretely, after deriving the invariant Lgrangian density under Parity, translational and local Lorentz transformations of space-time from the candidates which were constructed by quadratic and divergence part of the non-holonomic object, we obtained the source equation (3·20) from variational principle. Since symmetrical part of this equation agrees with Einstein equation, our gravitational equation system, i.e. the support equation (2·13) plus the source equation (3·20), has a physical validity and a gauge thoretical structure in the sence mentioned in this paper. This means that the field equation system of general relativity also has a quasi-gauge theoretical structure in the same sence.

References

[1] R.Utiyama, Itroduction to General Gauge Field Theory, Iwanami, (1987) 106
R.Utiyama, Prog.Theor.Phys., 64(1980)2220
[2] R.Utiyama, The Theory of General Relativity, Shoukabou, (1980)125
[3] F.W.Hehl et al, In General Relativity and Gravitation, Plenum Press, (1980) 329
R.Utiyama, Prog.Theor.Phys., 72 (1984) 83
S.Kichenassamy, Ann. Phys., 168 (1986) 404
[4] N. Nakanishi, Study on Elementary Particle Theory, 1(1990)
[5] R.Utiyama, Phys.Rev., 101 (1956)1597
[6] N.Nakanishi & I.Ojima, Covariant Operator Formalism of Gauge Theories and Quantum Gravity, World Scientific, (1990)
[7] J.A.Schouten, Ricci-Calculus, Springer, Berlin, (1954) 99
J.A.Schouten, Tensor Analysis for Physicists, 2nd ed. Clarendon Press, (1954) 81
[8] I.Imai, A consideration on elecromagnetics, Science Publication, (1990)82
[9] J.Kristien & R.K.Sachs, Ap. J., 143 (1966) 379
J.A.Tyson, F.Valdes & R.A.Wenk, Ap. J., 349 (1990)1
L.D.Gunnarsenn.L.D, Proceedings of the workshop on general relativity and gravitation, Tokyo, (1991)88
[10] T.W.B.Kibble, J.Math.Phys., 2 (1961) 212
K.Hayasi & T.Nakano, Prog. Theor. Phys., 38(2) (196 7) 491
K.Hayasi, Prog. Theor. Phys., 39(2) (1968)494
K.Hayasi & A.Bregmann, Ann. Phys., 75 (1973) 562
T.Kawai, Phy. Rev. D, 48(12) (1993) 5668

Dynamical torsion and torsion potential

HONG-JUN XIE and TAKESHI SHIRAFUJI

Physics Department, Saitama University
Urawa, Saitama 338, Japan

Abstract

We find a generalized tetrad which plays the role of a potential for torsion and makes torsion dynamic. Starting from the Einstein-Cartan action with torsion, we get two field equations, Einstein equation and torsion field equation by using metric tensor and torsion potential as independent variables; in the former equation torsion plays the role of a matter field, and the latter shows that torsion is a field which can propagate in vacuum. We also discuss properties of local transformations of torsion potential and give a simple example in which torsion potential is described by a scalar field.

1. Introduction

The gauge theory of gravitational field was first proposed by Utiyama [1] and later developed by Kibble [2] and Sciama [3]. The Einstein-Cartan-Sciama-Kibble (ECSK) theory of gravity uses Einstein-Cartan Lagrangian $\sqrt{-g}R$ made of the spin connection, but it does not require torsion to vanish. Rather, the torsion is treated as an independent variable along with the metric. The torsion is not dynamic in the ECSK theory, however, being algebraically determined by the local spin density: Namely, the torsion is always a pointwise function of the source field, and it is usually called *frozen torsion*.

Two approaches have been proposed to construct the theory of dynamical torsion. One is the quadratic Lagrangian approach to Poincaré gauge field theory initiated by Hayashi [4] in 1968. Further Hayashi and Shirafuji [5,6], and Hehl and Von der Heyde [7] pursued this approach. In this theory torsion satisfies second-order differential equations and can propagate in vacuum. The other is the theory of new general relativity studied from geometrical and observational point of view by Hayashi and Shirafuji [8]. In this theory the concept called absolute palallelism plays the fundemental role: It requires vanishing local connection and hence vanishing curvature, and attributes gravity to torsion instead of to curvature.

In this paper we propose a new approach to dynamical torsion. Introducing the sixteen variables that we call "torsion potential", we require symmetric condition for the local connection. Then we find that torsion can be expressed in terms of first-order derivatives of the torsion potential. The curvature is not vanishing in the present theory in contrast with new general relativity based on absolute palallelism. Starting from the Einstein-Cartan action with torsion, but regarding metric tensor and the torsion potential as independent variables, we obtain the Einstein equation and the torsion field equation.

In Sec.2 we show how to introduce torsion potential and how to represent torsion and connection by it. In Sec.3 we give field equations for metric tensor and torsion potential. In Sec.4 we discuss properties of local transformations of torsion potential, which differ from local Lorentz transformations. In the final section we give a simple example in which torsion potential is constructed from a scalar field.

2. Torsion potential

We consider connection $\Gamma^{\mu}{}_{\rho\sigma}$ with torsion

$$\Gamma^{\mu}{}_{\rho\sigma} = \mathring{\Gamma}^{\mu}{}_{\rho\sigma} + S^{\mu}{}_{\rho\sigma} , \tag{2.1}$$

where $\mathring{\Gamma}^{\mu}{}_{\rho\sigma}$ is torsionless connection coefficients and $S^{\mu}{}_{\rho\sigma}$ is a contorsion tensor defined by torsion $T^{\mu}{}_{\rho\sigma}$ $(=\Gamma^{\mu}{}_{\rho\sigma}-\Gamma^{\mu}{}_{\sigma\rho})$ in the form

$$S_{\mu\rho\sigma} := \frac{1}{2}\left(T_{\mu\rho\sigma} + T_{\rho\sigma\mu} + T_{\sigma\rho\mu}\right) . \tag{2.2}$$

When one takes the variation of the Einstein-Cartan action with torsion, one usually regards metric tensor $g_{\mu\nu}$ and contorsion $S^{\mu\nu}{}_{\rho}$ as independent variables. This is unnatural, however, because these two tensors are different in nature from each other: In fact, the metric tensor plays the role of a potential, while contorsion is a quantity like a force.

In order to find the potential of torsion, let us define a linear transformation

$$\overline{e}_{A} = t^{\mu}{}_{A}\,\overline{\partial}_{\mu} , \qquad \theta^{A} = t^{A}{}_{\mu}\,dx^{\mu} , \tag{2.3}$$

where $\overline{\partial}_\mu$ is the natural basis and $\overline{e}_A$ an arbitrary one with θ^A being the 1-form dual to $\overline{e}_A$ ($A = 0 \sim 3$). Here $t^A{}_\mu$ and its inverse $t^\mu{}_A$ are some arbitrary functions of coordinates satisfying the condition that the determinant $t = \det(t^A{}_\mu)$ is non zero. Obviously they satisfy

$$t^\mu{}_A\, t^A{}_\nu = \delta^\mu_\nu\,, \qquad\qquad t^A{}_\mu\, t^\mu{}_B = \delta^A_B\,. \tag{2.4}$$

On this new basis the metric tensor becomes

$$g_{AB} = \overline{e}_A \cdot \overline{e}_B = t^\mu{}_A\, t^\nu{}_B\, g_{\mu\nu}\,. \tag{2.5}$$

Here we do *not assume* that g_{AB} coincides with the Minkowski metric η_{AB}. Accordingly, the field $t^A{}_\mu$ is not a tetrad in the ordinary sense.

Now we define connection coefficients with respect to the bases $\overline{\partial}_\mu$ and $\overline{e}_A$ as follows:

$$\nabla\,\overline{\partial}_\nu = \Gamma^\mu{}_\nu\,\overline{\partial}_\mu\,, \qquad\qquad \nabla\,\overline{e}_B = \theta^A{}_B\,\overline{e}_A \tag{2.6}$$

with the connection 1-forms $\Gamma^\mu{}_\nu$ and $\theta^A{}_B$ being

$$\Gamma^\mu{}_\nu := \Gamma^\mu{}_{\nu\lambda}\,dx^\lambda = \Gamma^\mu{}_{\nu A}\,\theta^A\,, \tag{2.7}$$

$$\theta^A{}_B := \theta^A{}_{BC}\,\theta^C = \theta^A{}_{B\lambda}\,dx^\lambda\,, \tag{2.8}$$

respectively. The two connection coefficients are related to each other by

$$\theta^A{}_{BC} = t^A{}_\mu\, t^\nu{}_B\, t^\lambda{}_C\, \Gamma^\mu{}_{\nu\lambda} - t^\nu{}_B\, t^\lambda{}_C\, \partial_\lambda\, t^A{}_\nu\,. \tag{2.9}$$

In conformity with this, the total covariant derivative of $t^A{}_\mu$ with respect to indices A and μ is vanishing:

$$\begin{cases} \mathcal{D}_\lambda\, t^A{}_\nu := \partial_\lambda\, t^A{}_\nu - \Gamma^\mu{}_{\nu\lambda}\, t^A{}_\mu + \theta^A{}_{B\lambda}\, t^B{}_\nu = 0\,, \\ \mathcal{D}_\lambda\, t^\mu{}_B := \partial_\lambda\, t^\mu{}_B + \Gamma^\mu{}_{\nu\lambda}\, t^\nu{}_B - \theta^A{}_{B\lambda}\, t^\mu{}_A = 0\,. \end{cases} \tag{2.10}$$

The commutator of base vectors $\overline{e}_A$ and the exterior derivative of 1-form θ^A are respectively given by

$$[\overline{e}_B\,, \overline{e}_C] = f^A{}_{BC}\,\overline{e}_A\,, \qquad\qquad d\theta^A = -\frac{1}{2}\, f^A{}_{BC}\,\theta^B \wedge \theta^C\,, \tag{2.11}$$

where $f^A{}_{BC}$, which are called Ricci rotation coefficients in the uaual tetrad case, are

$$f^A{}_{BC} = (t^\mu{}_B\, t^\nu{}_C - t^\mu{}_C\, t^\nu{}_B)\, \partial_\nu\, t^A{}_\mu\,. \tag{2.12}$$

On the basis $\vec{e}_A$ the torsion 2-form T^A and the curvature 2-form $R^A{}_B$ are defined by

$$T^A := d\theta^A + \theta^A{}_B \wedge \theta^B = -\frac{1}{2}\, T^A{}_{BC}\, \theta^B \wedge \theta^C\,, \tag{2.13}$$

$$R^A{}_B := d\theta^A{}_B + \theta^A{}_C \wedge \theta^C{}_B = \frac{1}{2}\, R^A{}_{BCD}\, \theta^C \wedge \theta^D\,, \tag{2.14}$$

where the torsion tensor and the curvature tensor are respectively given by

$$T^A{}_{BC} = t^A{}_\mu\, t^\nu{}_B\, t^\lambda{}_C\, T^\mu{}_{\nu\lambda} = \theta^A{}_{BC} - \theta^A{}_{CB} + f^A{}_{BC}\,, \tag{2.15}$$

$$R^A{}_{BCD} = t^A{}_\mu\, t^\nu{}_B\, t^\rho{}_C\, t^\sigma{}_D\, R^\mu{}_{\nu\rho\sigma}\,. \tag{2.16}$$

We note that $g_{AB} \neq \eta_{AB}$ by assumption, therefore $\theta_{ABC} \neq -\theta_{BAC}$, indicating that θ_{ABC} is not a spin connection. In order to compare with the torsionless case and for convenience of later use, let us separate curvature tensor $R^\mu{}_{\nu\rho\sigma}$ into torsionless part and torsion part

$$R^\mu{}_{\nu\rho\sigma} = \overset{\circ}{R}{}^\mu{}_{\nu\rho\sigma} + \tilde{R}^\mu{}_{\nu\rho\sigma}\,, \tag{2.17}$$

where $\overset{\circ}{R}{}^\mu{}_{\nu\rho\sigma}$ is the Riemannian curvature tensor, and $\tilde{R}^\mu{}_{\nu\rho\sigma}$ is the torsion part of curvature tensor which is defined by

$$\tilde{R}^\mu{}_{\nu\rho\sigma} := \overset{\circ}{\nabla}_\rho S^\mu{}_{\nu\sigma} - \overset{\circ}{\nabla}_\sigma S^\mu{}_{\nu\rho} + S^\mu{}_{\lambda\rho} S^\lambda{}_{\nu\sigma} - S^\mu{}_{\lambda\sigma} S^\lambda{}_{\nu\rho}\,. \tag{2.18}$$

Here $\overset{\circ}{\nabla}_\rho$ denotes the covariant derivative defined by using $\overset{\circ}{\Gamma}{}^\mu{}_{\nu\rho}$.

Until now we have not yet imposed any restriction on $t^A{}_\mu$ and $\theta^A{}_{BC}$. From now on, however, we suppose that the following two conditions are satisfied:

(i) The metric condition

$$\mathcal{D}_\lambda g_{\mu\nu} = 0\,, \tag{2.19}$$

(ii) and the symmetric condition

$$\theta^A{}_{BC} = \theta^A{}_{CB} \qquad \text{or equivalently} \qquad \theta^A{}_B \wedge \theta^B = 0\,. \tag{2.20}$$

The condition (i) implies that $\overset{\circ}{\Gamma}{}^\mu{}_{\rho\sigma}$ is nothing but the Christoffel symbol. The condition (ii) is a generalization of the absolute palallelism: In fact, in the latter

4. Local transformation properties of the torsion potential

In this section we consider properties of the torsion potential under local transformations other than coordinate transformations. Let $L^A{}_B$ be some transformation functions,

$$\hat{t}^A{}_\mu = L^A{}_B\, t^B{}_\mu\,, \qquad \hat{t}^\mu{}_B = L^{-1\,A}{}_B\, t^\mu{}_A\,. \tag{4.1}$$

The 1-form θ^A and connection $\theta^A{}_B$ are then transformed like

$$\hat{\theta}^A = L^A{}_B\,\theta^B\,, \tag{4.2}$$

$$\hat{\theta}^A{}_B = L^A{}_C\, L^{-1\,D}{}_B\,\theta^C{}_D + L^A{}_C\, dL^{-1\,C}{}_B\,. \tag{4.3}$$

Here $\hat{\theta}^A{}_B\,(=\hat{\theta}^A{}_{BC}\hat{\theta}^C)$ is a new connection, which we require to satisfy the condition (ii); namely,

$$\hat{\theta}^A{}_{BC} = \hat{\theta}^A{}_{CB} \qquad \text{or equivalently} \qquad \hat{\theta}^A{}_B \wedge \hat{\theta}^B = 0\,. \tag{4.4}$$

Then we see that transformation functions $L^A{}_B$ must satisfy the constraint

$$d_B L^A{}_C = d_C L^A{}_B \qquad \text{or} \qquad L^B{}_C d_D L^{-1\,A}{}_B = L^B{}_D d_C L^{-1\,A}{}_B \tag{4.5}$$

with $d_B = t^\mu{}_B\,\partial_\mu$. Under the transformation (4.1) satisfying (4.5), the torsion and curvature 2-forms transform as follows:

$$\hat{T}^A = L^A{}_B\, T^B\,, \tag{4.6}$$

$$\hat{R}^A{}_B = L^A{}_C\, L^{-1\,D}{}_B\, R^C{}_D\,. \tag{4.7}$$

However, the torsion equation (3.8) is not invariant under local transformations with the constraint (4.5). The invariance of (3.8) requires a new condition

$$L^A{}_B d_D L^{-1\,M}{}_C = L^A{}_D d_B L^{-1\,M}{}_C\,. \tag{4.8}$$

This constraint is stronger than (4.5). To see this, we consider an infinitesimal transformation

$$L^A{}_B = \delta^A_B + \varepsilon^A{}_B \tag{4.9}$$

with $|\varepsilon^A{}_B| \ll 1$. The metric tensor g_{AB} transforms as

$$\hat{g}_{AB} = g_{AB} - \varepsilon_{AB} - \varepsilon_{BA}\,. \tag{4.10}$$

Here we lower the index of $\varepsilon^A{}_B$ using g_{AB}. If the transformation (4.9) satisfies the constraint (4.8), we have that

$$d_B\,\varepsilon^A{}_C = 0\,, \tag{4.11}$$

namely, $\varepsilon^A{}_B$ must be some constant parameters. Therefore we see that transformations $\{L^A{}_B\}$ make a globe general linear group $GL(4,R)$ which keep the field equation (3.8) being invariant. This situation is similar to the theory of new general relativity in which only the globe Lorentz transformations have been allowed.

5. Scalar field as the torsion potential

Now we consider as a simple example the torsion potential made of a scalar field. Namely, let us assume that the torsion potential takes the form

$$t^A{}_\mu = \delta^A_\mu \varphi\,, \qquad\qquad t^\mu{}_A = \delta^\mu_A \varphi^{-1} \tag{5.1}$$

with φ being a nonvanishing scalar field. Due to Kronecker's δ^A_μ in (5.1), the index A is assumed to undergo the transformation $L^A{}_B = \delta^A_\nu \delta^\mu_B\,(\partial x^\nu/\partial x'^\mu)$ under coordinate transformations, and accordingly we need not distinguish indices A from μ. The metric tensor is given by

$$g_{AB} = \varphi^{-2}\,\delta^\mu_A\,\delta^\nu_B\,g_{\mu\nu}\,, \tag{5.2}$$

which can be identified with a conformal rescaling. The torsion and contorsion tensors are expressed by

$$T_{\mu\nu\lambda} = g_{\mu\nu}\varphi^{-1}\partial_\lambda\varphi - g_{\mu\lambda}\varphi^{-1}\partial_\nu\varphi\,, \tag{5.3}$$

$$S^{\mu\nu}{}_\lambda = (g^{\mu\rho}\delta^\nu_\lambda - g^{\nu\rho}\delta^\mu_\lambda)\,\varphi^{-1}\partial_\rho\varphi\,. \tag{5.4}$$

If we substitute $S^{\mu\nu}{}_\lambda$ of (5.4) to the full equation (3.5), we will encounter a difficulty, because 16 equations cannot be satisfied by only one unknow function φ. Thus, we will choose a different way: Namely, we substitute $S^{\mu\nu}{}_\lambda$ to the Lagrangian (3.1). Then we have

$$\mathcal{L}_G = \sqrt{-g}\left(\overset{\circ}{R} - 6g^{\mu\nu}\varphi^{-2}\partial_\mu\varphi\partial_\nu\varphi\right). \tag{5.5}$$

Taking variation of (5.5) with respect to $g_{\mu\nu}$ and φ, we get

$$\overset{\circ}{G}{}^{\mu\nu} = \frac{6}{\varphi^2}\left(g^{\mu\rho}g^{\nu\sigma} - \frac{1}{2}\,g^{\mu\nu}g^{\rho\sigma}\right)\partial_\rho\varphi\,\partial_\sigma\varphi\,, \tag{5.6}$$

$$\overset{\circ}{\Box}\varphi - \frac{1}{\varphi}\,g^{\mu\nu}\partial_\mu\varphi\,\partial_\nu\varphi = 0\,. \tag{5.7}$$

But we expect to violate this conjecture when black holes *completely* evaporate. Since black hole evaporation is *semi-classical* theory, this has not been a object of cosmic censorship conjecture till now.

We want to believe that this conjecture is true in semi-classical level. Our purpose is to prove, under reasonable conditions, the non-existence of the naked s The essence of the proof is that space time which contains the completely evaporating black hole does not satisfy *the reflecting condition*

Our notation and fundamental definitions will be as those of Hawking and Ellis[2].

2 Preliminaries

By space time we mean a pair (M, g) where M is a connected orientable four dimensional Hausdorff C^∞ manifold and g is a C^∞ Lorentz met By a partial Cauchy surface S we mean a spacelike hypersurface in (M, g) which no nonspacelike curve intersects more than once. In the following several conditions are defined.

(a) **causality condition**

Definition 1 *We say that strong causality condition holds at a point $p \in M$ if every neighborhood of p contains a neighborhood of p which no nonspacelike curve inters*

Definition 2 *We say that strong causality condition holds in space time (M, g) if it holds for every point of M.*

Roughly speaking, the strong causality condition implies that there exist no almost closed causal curve.

(b) **convergence condition**

The usual convergence conditions are required for all region of space time. In semi-classical theory these conditions does not hold. Here we define the new convergence condition.

Definition 3 *We say that the partially null convergence condition holds if* $R_{ab}k^a k^b \geq 0$ *for every null vector whose direction is parallel to the null geodesic generators of the event horizon (This is defined below).*

Above condition together with the Raychaudhuri equation[2] that relates Ricci curvature and the expansion of a null geodesic congruence implies

(c) **event horizon and evaporation condition**

For defining the event horizon and its complete evaporation we must suppose a few conditions for space time. To do this we recall Wald's theorem which implies that space time is not globally hyperbolic if black holes completely evaporate.

Theorem 1 *[6] Let* (M, g_{ab}) *be a time-oriented space time, and let* S_1 *and* S_2 *be partial Cauchy surfaces. Suppose that*

(1) there is a point $p \in D^+(S_1)$ *such that* $p \notin (J^-(S_2) \cup J^+(S_2))$,

(2) $J^+(K) \cap S_2$ *has compact closure, where* $K \equiv S - (D^-(S_2) \cap S_1)$.

Then $S_2 \not\subset D^+(S_1)$.

¿From above theorem we can define the black hole, the event horizon and complete evaporation of the black hole.

Definition 4 *We say a black hole evaporation completely between S_1 and S_2 that, under the above theorem's assumption, $B = M - \{(J^+(S_1) \cup J^-(S_1)) \cap$ $(where J^-(S_1) \cap J^+(S_2) = \emptyset$ and we call a black hole, B, and a event horizon, $\dot{B}$. Further a expansion θ of every null geodesic generator of the event horizon is negative and $\lim_{t \to t_{eva}} \theta(t) = -\infty$.*

We call above condition *the completely evaporating condition.*

Remarks 1 *We know that the order of differentiability becomes a problem for the evaporating black hole. But we assume that the metric is C^∞, i.e. the Raychaudhuri equation and the expansion are useful since it seems that differentiability is not importa*

For example, Hawking and Ellis say:[2]

"..., the order of differentiability of the metric is probably not physically significant. Since one can never measure the metric exactly, but only with some margin of error, one could never determine that there was an actual discontinuity in its der

3 Theorem

In the proof of the our main theorem it is important that space time which contains the complete evaporating black hole is not only non-globally hyperbolic but also non-reflecting. Therefore we must recall the reflecting condition[3] before the theorem.

Definition 5 *Call space time (M,g) reflecting if the following condition holds: for all events x and y in (M,g), $I^+(y) \supseteq I^+(x)$ if and only if $I^-(x) \supseteq I^-(y)$.*

Clarke and Joshi said the above definition in another way that is requirement for achronal boundaries.

Proposition 1 *[1] The following are equivalent:*
(a) (M,g) is reflecting,
(b) $y \in \dot{I}^+(x)$ iff $x \in \dot{I}^-(y)$.

Hereafter we use the above **proposition 1** (b) to require the reflecting condition.

Theorem 2 *Space time that black holes evaporate completely does not hold the reflecting condition.*

Sketch of proof: By definition there exist points that are in $\dot{I}^-(S_2)$ and that are not connected by causal curves with S_2. This implies that the reflecting condition violates. □

¿From **theorem 2** the following corollary holds.

Corollary 1 *Every null geodesic generator of the event horizon is future incomplete.*

By the above mentioned facts we prove the main theorem.

Theorem 3 *Let S_1 and S_2 be partial Cauchy surfaces in time-orientable space time (M, g). the following three conditions cannot hold synchronously.*

(1) the partially null convergence condition,

(2) the strong causality condition,

(3) the completely evaporation condition.

Sketch of proof : Since space time we consider violates the reflecting condition, by **corollary 1**, on the same achronal boundary there exist a past incomplete null geodesic and a future incomplete null geodesic that is a generator of the black hole event horizon. We can construct a null geodesic sequence that has limit curves which are the above null geodesics. By the partially null convergence condition the expansion of each null geodesics of this sequence is strictly decreasing function. This fact and the completely evaporation condition are contradict by continuity. See [4] for details. □

Finally we conclude the following corollary.

Corollary 2 *Black holes cannot completely evaporate.*

References

[1] C.J.S.Clarke and P.S.Joshi, Class.Quantum Grav.**5**(1988)19.

[2] S.W.Hawking and G.F.R.Ellis, The large scale structure of space time (Cambridge University Press, Cambridge, 1973).

[3] S.W.Hawking and R.K.Sachs, Commun.Math.Phys.**35**(1974)287.

[4] M.Nalita, in preparation.

[5] R.Penrose, "Singularities and time-asymmetry" in General relativity: An Einstein centenary survey, edited by S.W.Hawking and W.Israel (Cambr

[6] R.M.Wald, "Black Holes, Singularities and Predictability" in Quantum Theory of Gravity, edited by S.M.Christensen (Adam Hilger Press, Bristol,

Dilatonic Black Holes with Gauss-Bonnet Term

TAKASHI TORII [a], KEI-ICHI MAEDA[b] and HIROKI YAJIMA[c],

Department of Physics, Waseda University, Shinjuku-ku, Tokyo 169, Japan

Abstract

We discuss black holes in an effective theory of a superstring model, which includes a dilaton field, a gauge field and the Gauss-Bonnet(GB) term. Assuming U(1) or SU(2) symmetry as a gauge field, we find four types of spherically symmetric solutions, i.e., a neutral, an electrically charged, a magnetically charged and a "colored" black hole. For each black hole there is a critical mass below which no solution exists. The curvature at the horizon diverges in the limit of the critical mass and a naked singularity appears there. We also study the thermodynamical properties of those solutions. Since the temperature is finite even in the limit of the critical mass, the black hole may eventually evolve into a naked singularity via the Hawking evaporation process.

(a) electronic mail: 64L514@cfi.waseda.ac.jp, (b) electronic mail: maeda@cfi.waseda.ac.jp
(c) electronic mail: 695L1079@cfi.waseda.ac.jp

1 Introduction

One of the most fascinating dreams for all physicists is the unification of all fundamental forces, i.e., electromagnetic, weak, strong and gravitational interactions. The electromagnetic and weak interactions are successfully unified as the Weinberg-Salam theory. The strong interaction is described by quantum chromodynamics (QCD) and is likely to be unified with the Weinberg-Salam theory into a grand unified theory in the context of gauge theory. The gravitational interaction, however, is not yet included, in spite of a great deal of effort. The most promising candidate for a unified theory of all interactions is a superstring theory, which may unify everything without any divergences.

String theory should be studied in extreme situations, since the difference from the conventional field theory appears notably at very high energy scales. Most of the work on such applications, so far, has been performed by using an effective field theory derived from a string theory which contains the leading or next leading order of the inverse string tension α'. One such application is string cosmology. It is shown that some puzzles in Einstein cosmology might be solved with a string theory. For example the singularity theorem demands the initial singularity of the universe in the Einstein gravity, while a string inspired model can remove it and provide a non-singular cosmology[1, 2].

Another application is a black hole physics. The first study was made by Gibbons and one of the present authors in the Einstein-Maxwell-Dilaton (EMD) system[3] and the same solution was discussed in the different coordinate system[4]. They found a static spherically symmetric black hole solution (GM-GHS solution) with a dilaton hair. Since the dilaton hair cannot appear without electromagnetic hair, it is classified as secondary hair. After this work, many solutions were discussed in various models and the following question may arose. How are the black hole solutions affected if the Lagrangian contains the higher-curvature term, which comes from the next leading order of α'? Callan et al.[5]

discussed black hole solutions in the theory with both the higher-curvature term $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$ and the dilaton field, and Mignemi and Stewart[6] took both the GB (Gauss-Bonnet) term and the dilaton field in account. In their work field variables were expanded by the inverse string tension α' and the first order term of α' are taken into account and by using this perturbative method they found analytic solutions. Campbell et al., Mignemi and Stewart tried to clarify the effect of the axion field[7, 8, 9]. In all these models containing the higher-curvature term they used perturbative methods. In this paper we present the black hole solutions obtained without any perturbative method, assuming the GB higher curvature term, and we clarify the effect of the higher curvature term.

A new black hole solution called a colored black hole, which has a non-Abelian hair, was discovered in Einstein-Yang-Mills system[10] soon after the existence of a particle-like Bartnik-McKinnon(BM) solution had been pointed out in the same system[11]. These solutions are governed by a balance between the attractive force of gravity and the repulsive force of the Yang-Mills(YM) field. Hence if gravity is absent these objects cannot exist. In this sense those structures are of a new type. Although both the BM particle and the colored black hole were found to be unstable against radial linear perturbation[12], they showed us a new aspect of black hole physics and forced us to reconsider the black hole no-hair conjecture.

After these solutions, a variety of self-gravitating structures and black hole solutions with a non-Abelian field were found in static spherically symmetric spacetime[13, 14, 15]. One of them is a dilatonic BM particle and a dilatonic colored black hole solution in the Einstein-Yang-Mills-Dilaton (EYMD) system[16, 17]. They are direct extensions of the GM-GHS solution. Non-Abelian black holes are expected to be very small, so we have to discuss other contributions, for example the GB term and/or the moduli field, if the fundamental theory is described by a string model. Donets and Gal'tsov showed that a particle-like solution does not exist in the EYMD system with the GB term[19]. Then, they put a numerical factor β in front of the GB term, where $\beta = 1$ correspond to the effective sting theory. They showed that there is a critical value $\beta_{cr} = 0.37$ beyond which no particle-like solution exists. However, we expect that the black hole solution may exist in the same system, so we also discuss the case of the SU(2) Yang-Mills field.

This paper is organized as follows. We show a model and field equations in section 2, and present various types of new solutions (neutral, electrically charged, magnetically charged and "colored" black holes) in section 3. In section 4 we study the thermodynamical properties of those black holes. Section 5 includes discussions and some remarks.

2 A Model and Field Equations

We shall consider the static spherically symmetric black holes is the model

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2\kappa^2}R - \frac{1}{2\kappa^2}(\nabla\phi)^2 - \frac{1}{6}e^{-2\gamma\phi}H^2 + \frac{\alpha'}{16\kappa^2}e^{-\gamma\phi}\left(\beta\hat{R}^2 - \mathrm{Tr}F^2\right)\right], \qquad (1)$$

where $\kappa^2 = 8\pi G$. This type of action comes from the low-energy limit of the heterotic string theory, where ϕ is a dilaton field, $\gamma = \sqrt{2}\kappa$ is its coupling constant and H is a three form which can be expressed by a pseudo-scalar field, or axion. When a solution does not have a dyon and the spacetime

is static and spherically symmetric, which is the case we will consider, the axion field H vanishes. F is the field strength of the gauge field expressed by its potential A. We adopt the following form

$$A = a(r)\tau_r dt - [1 + w(r)]\,\tau_\phi d\theta + [1 + w(r)]\,\tau_\theta \sin\theta d\phi, \tag{2}$$

which is the most generic form of a spherically symmetric and static one. Comparing (2) with the potential of the U(1) gauge field, we find that a and w play the roll of an electric and a magnetic field potential, respectively. For U(1) gauge field, we set $\tau_r = 1$ and $\tau_\theta = \tau_\phi = 0$ formally after calculating the field strength.

The GB term, $\hat{R}^2$, is

$$\hat{R}^2 = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} - 4R_{\mu\nu}R^{\mu\nu} + R^2. \tag{3}$$

This combination is introduced to cancel anomalies and has the advantage that the higher derivatives of metric functions do not appear in the field equations. Setting $\alpha' = 16\pi G/g^2$, g denotes a gauge coupling constant. β is a numerical factor adopted in Ref. [19], but we fix $\beta = 1$.

We assume a static spherically symmetric ansatz, and we use the Schwarzschild type metric written as

$$ds^2 = -\left(1 - \frac{2Gm(r)}{r}\right)\mathrm{e}^{-2\delta(r)}dt^2 + \left(1 - \frac{2Gm(r)}{r}\right)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{4}$$

Varying the action (1) and substituting ansätze (2), (4), we derive the field equations,

$$\delta' = h^{-1}\left[-\frac{\kappa^2}{2}\tilde{r}\phi'^2 - \frac{\mathrm{e}^{-\gamma\phi}}{4\tilde{r}}\left\{\mathrm{e}^{2\delta}B^{-2}\tilde{a}^2w^2 + w'^2 + \frac{2\tilde{m}}{\tilde{r}}(\gamma^2\phi'^2 - \gamma\phi'')\right\}\right], \tag{5}$$

$$\begin{aligned}\tilde{m}' = h^{-1}\Bigg[&\frac{\kappa^2}{4}B\tilde{r}^2\phi'^2 + \frac{\mathrm{e}^{-\gamma\phi}}{16}\left\{\mathrm{e}^{2\delta}\left(\tilde{r}^2\tilde{a}'^2 + 2B^{-1}\tilde{a}^2w^2\right) + 2Bw'^2 + \frac{(1-w^2)^2}{\tilde{r}^2}\right.\\ &\left.+4B\frac{2\tilde{m}}{\tilde{r}}(\gamma^2\phi'^2 - \gamma\phi'') + 8\gamma\phi'\frac{m}{\tilde{r}^2}\left(B - \frac{\tilde{m}}{\tilde{r}}\right)\right\}\Bigg],\end{aligned} \tag{6}$$

$$\begin{aligned}&\left[\mathrm{e}^{-\delta}\tilde{r}^2B\phi'\right]' - \frac{\mathrm{e}^{-\gamma\phi}}{8\kappa^2}\gamma\mathrm{e}^{-\delta}\tilde{r}^2\left[\mathrm{e}^{2\delta}\left\{\tilde{a}'^2 + \frac{2\tilde{a}^2w^2}{\tilde{r}}B\right\} - \left\{\frac{2\tilde{w}^2}{\tilde{r}^2}B + \frac{(1-w^2)^2}{\tilde{r}^4}\right\}\right.\\ &\left.+\frac{4}{\tilde{r}^2}\left\{2f^2 + B(\delta'f + f')\right\} + \frac{4}{\tilde{r}^2}(\delta'f - f')\right] = 0,\end{aligned} \tag{7}$$

$$\left[\mathrm{e}^{\delta}\tilde{r}^2\mathrm{e}^{-\gamma\phi}a'\right]' - 2\mathrm{e}^{\delta}\mathrm{e}^{-\gamma\phi}B^{-1}\tilde{a}w^2 = 0, \tag{8}$$

$$\left[\mathrm{e}^{-\delta}B\mathrm{e}^{-\gamma\phi}w'\right]' + \mathrm{e}^{\delta}\mathrm{e}^{-\gamma\phi}B^{-1}\tilde{a}'^2w + \mathrm{e}^{-\delta}\mathrm{e}^{-\gamma\phi}\frac{w(1-w^2)}{\tilde{r}^2} = 0, \tag{9}$$

Here we have used the dimensionless variables $\tilde{r} = r/\sqrt{\alpha'}$, $\tilde{m} = Gm/\sqrt{\alpha'}$, $\tilde{a} = a/\sqrt{\alpha'}$. A prime in the field equations denotes the derivative with respect to $\tilde{r}$, and

$$f = h^{-1}\left[\frac{m}{\tilde{r}^2} + \frac{\kappa^2}{4}B\tilde{r}\phi'^2 - \frac{\mathrm{e}^{-\gamma\phi}}{16\tilde{r}}\left\{\mathrm{e}^{2\delta}\left(\tilde{r}^2\tilde{a}'^2 - 2B^{-1}\tilde{a}^2w^2\right) - 2Bw'^2 + \frac{(1-w^2)^2}{\tilde{r}^2}\right\}\right], \tag{10}$$

$$h = 1 + \frac{\mathrm{e}^{-\gamma\phi}}{2\tilde{r}}\gamma\phi'\left(B - \frac{\tilde{m}}{\tilde{r}}\right), \tag{11}$$

$$B = 1 - \frac{2\tilde{m}}{\tilde{r}}. \tag{12}$$

As for boundary conditions on the event horizon and at spatial infinity, for the metric functions, we impose following three conditions: (i) Asymptotic flatness at spatial infinity. (ii) The existence of a regular horizon r_h, (iii) The nonexistence of singularities outside the event horizon, and for the field functions, we set

$$\phi \to 0, \tag{13}$$

$$a \to 0, \tag{14}$$

$$w \to \begin{cases} \pm 1 & \text{(globally magnetically charged solution).} \\ 0 & \text{(globally magnetically uncharged solution),} \end{cases} \tag{15}$$

as $r \to \infty$. These conditions guarantee the total energy of the present system to be finite.

3 Dilatonic Black Hole with Gauss-Bonnet Term

3.1 Neutral Black Hole

First we consider the neutral solution, which is the simplest case. The neutral black hole solution can be obtained by setting $a \equiv 0$ and $w \equiv \pm 1$, which satisfy the field equations (8) and (9). From Eq. (7) we find the following relation on the event horizon, i.e.,

$$\phi'^2 - \frac{\phi'}{A\gamma} + \frac{3}{\kappa^2} = 0, \tag{16}$$

where $A = e^{-\gamma\phi}/4\tilde{r_h}^2$. Hence if we choose $\phi(r_h)$ for the fixed radius of event horizon $\tilde{r_h}$, we obtain $\phi'_{\pm}(r_h)$. We integrate the field equations (5)$\sim$(9) from the horizon. We show the behavior of the field functions of neutral black holes in Fig. 1. We find the solutions with a regular horizon only when we choose $\phi'(r_h) = \phi'_+$ which is the larger root of equation (16).

The dilaton field of the smaller black holes varies more rapidly than that of the larger black holes. This means that the string effect becomes more important for the smaller black hole. Since we have normalized the radial coordinate by α', then, large $\tilde{r_h}$ has two possibilities, either r_h is large or α' is small. In the former case, a black hole becomes so large that we can treat it macroscopically and the next-leading term in α' does not affect its structure. In the latter case, the $O(\alpha')$ term is obviously negligible.

We see that the mass function decreases near the horizon, i.e., there is a region where the effective energy density becomes negative (see Fig.1(b)). On the event horizon the first derivative of $\tilde{m}$ is

$$\tilde{m}' = -\frac{\kappa^2}{24A^2\gamma^2}\left(1 \pm \sqrt{1 - 12A^2\frac{\gamma^2}{\kappa^2}}\right)^2. \tag{17}$$

The right hand side of this equation is negative definite, hence the function m always decreases in the vicinity of the horizon. It is an essential point for existence of the neutral black holes that the energy density becomes negative[18]. If the dominant energy condition is satisfied, no non-trivial solution can exist by the no-hair theorem for a scalar field. However in our situation this is not the case. When the radius of the event horizon gets small, the lapse function δ decreases rapidly around the horizon. This behavior will become important in the later discussion.

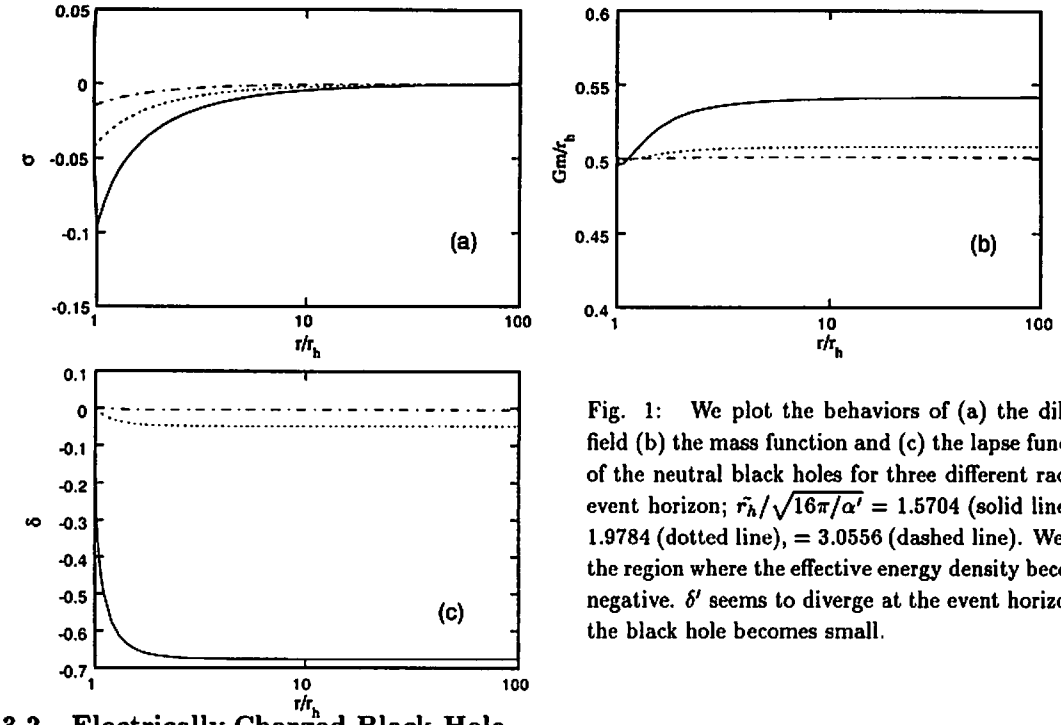


Fig. 1: We plot the behaviors of (a) the dilaton field (b) the mass function and (c) the lapse function of the neutral black holes for three different radii of event horizon; $\tilde{r_h}/\sqrt{16\pi/\alpha'} = 1.5704$ (solid line), = 1.9784 (dotted line), = 3.0556 (dashed line). We find the region where the effective energy density becomes negative. δ' seems to diverge at the event horizon as the black hole becomes small.

3.2 Electrically Charged Black Hole

Electrically charged black hole solutions are obtained by setting $w \equiv \pm 1$ and $\tilde{a}w \equiv 0$. The latter corresponds to the vanishing of the self-interaction terms which arise from the non-Abelian effect. In this case equation (8) can be integrated as

$$\tilde{a}' = \mathrm{e}^{-\delta}\mathrm{e}^{\gamma\phi}\frac{\tilde{Q}_e}{\tilde{r}^2}, \tag{18}$$

where $\tilde{Q}_e$ is a constant of integration and corresponds to an electric charge.

The behaviors of the field functions are almost same as those of the neutral case. We show $M - r_h$ relations of black holes with $\tilde{Q}_e = 0.0(neutral)$, 0.4, 1.0 in Fig. 2. Note that there is a critical mass M_{cr}, below which no solution exists, for each branch. The existence of the critical mass is also known in Reissner-Nordström black holes with a fixed charge. In that case the event and the inner horizon coincide and the black hole becomes extreme at the critical mass. However our black holes are found not to be extreme but to appear a naked singularity at the critical points[18]. We also find that $\phi'(r_h)$ is a multiple root of Eq. (16) at the critical mass. With this fact we can show that $\phi''(r_h)$ and $\delta'(r_h)$ diverge. We have also checked that $I = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$ diverges at the horizon. A naked singularity appears as a result.

3.3 Magnetically Charged Black Hole

Next we turn to the magnetically charged solutions. They can be obtained by putting $\tilde{a} \equiv 0$ and $w \equiv 0$. Note that they are not the 'tHooft-Polyakov type solution in Ref. [15] but only have a U(1)

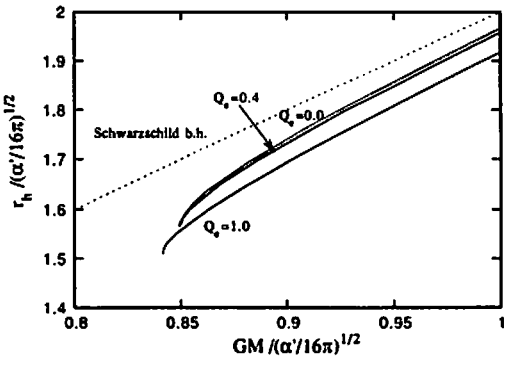


Fig. 2: The mass-horizon radius diagram for the neutral black holes ($Q_e = 0$) and the electrically charged black holes ($Q_e = 0.4$, 1.0). There are critical masses, below which no solution exists, for each curve. We can see that all curves become vertical at the critical mass.

charge like the dual solution of U(1) electrically charged one. In this case the value of the magnetic charge is quantized as $Q_m = 1.0$ from the theory. The behavior of the field equations is similar to that of the uncharged and electrically charged solutions. We plot the $M - r_h$ relation in Fig. 3. We discuss the properties of this black hole in the next subsection together with the "colored" case.

3.4 Dilatonic Colored Black Hole with Gauss-Bonnet Term

In this subsection we extend the gauge field from U(1) gauge group to assume SU(2) one. Here we set the 'tHooft-Polyakov ansatz $\tilde{a} \equiv 0$, i.e., purely magnetic YM field strength exists. Although the no-hair theorem for spherical monopole and dyons in SU(2) Einstein-Yang-Mills was proved[20], it is not clear whether this kind of no hair theorem holds in our model.

Since $w'(r_h)$ is determined by $\phi(r_h)$ and $w(r_h)$, we have only one shooting parameter, $w(r_h)$, which should be fixed by an iterative integration. Here we assume $w(r_h) > 0$ without loss of generality.

Under the above conditions we find a discrete family of regular black hole solutions which are characterized by the node number n of the YM potential, just as for the colored black hole. We call those solutions "colored" black holes as well. Since the YM field damps faster than $\sim 1/r^2$, those black holes have no global color charge related to the gauge field, like other non-Abelian black holes, with the exception of a monopole black hole. Hence we use double-quotation marks to imply that the new solutions are not globally charged but locally charged. The dilaton field and metric functions are similar to those of other solutions. For example there is a region where the effective energy density becomes negative.

The characteristic feature is that the YM potential w is almost scale invariant. In the limit of $\tilde{r_h} \to \infty$ the YM field equations decouple from the Einstein equations and Eq. (9) becomes

$$\tilde{r}(\tilde{r} - 2\tilde{M})w'' + 2\tilde{M}w' + w(1 - w^2) = 0, \tag{19}$$

Which is the field equation in a fixed Schwarzschild background spacetime. The YM field can have a nontrivial configuration although it makes no contribution to the black hole structure. The solution of this differential equation $w = w^*(\tilde{r}) = w^*(r/r_h)$, becomes scale invariant. This scale invariance can be seen approximately in our numerical solutions. The configuration of the YM potentials appears to be almost the same when they are drawn as the function of r/r_h. A similar behavior is found for

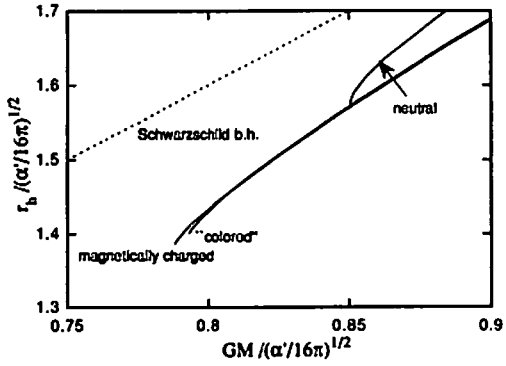


Fig. 3: The mass-horizon radius diagram for the neutral black holes, the magnetically charged black holes ($Q_m = 1.0$) and "colored" black holes ($n = 1$). There is a critical mass, where radius of event horizon keeps finite. We can see that each curves is not vertical at the critical mass. This is due to the non-trivial transformation between the electrically and the magnetically charged black holes

large black holes in the EYMD system[16]. It turns out that this approximation becomes valid at the almost same horizon radius.

We show the $M - r_h$ relation of neutral black holes, magnetically charged black holes ($\bar{Q_m} = 1.0$) and "colored" black holes with node number $n = 1$ in Fig. 3. There is also critical mass M_{cr} for each branches. We expect that naked singularities appear at the these critical points as neutral solutions.

Although the neutral and the electrically charged branches in Fig. 2 become vertical at critical points, we can see the magnetically charged and "colored" branches are not vertical but have finite inclination. The reason seems to be understood as follows. The magnetically charged black hole is obtained by setting $Q_m \equiv 1 - w^2 = 1$. If we set

$$\hat{Q}_e = e^{-\gamma\phi} Q_m, \tag{20}$$

in the field equation (5)~(9), we can recover the electrically charged solutions with the "electric charge" $\hat{Q}_e$. Note that the duality transformation of Ref. [3, 4]

$$\bar{F} = -e^{-2\gamma\phi} * F, \quad \bar{\phi} = -\phi, \tag{21}$$

does not applied in our model because of existence of the GB term. Suppose that there are several magnetically charged black holes in the same branch. Then we perform the transformation (20), we can obtain the black holes with the "electric charge" $\hat{Q}_e$. Since this transformation contains the dilaton field, which is the function of the radial coordinate r, the "electric charge" $\hat{Q}_e$ also depends on r. This dependence is different for each black hole, because the dilaton fields are non-trivial. Although each black hole has the same "electric charge" $\hat{Q}_e = Q_m$ at the infinity, each has different "electric charge" $\hat{Q}_e = e^{-\gamma\phi(r_h)} Q_m$ at the horizon. While the black hole mass is defined at infinity, the radius of event horizon is determined near the origin where $\hat{Q}_e$ has different value for each black hole. This is the reason why the branch of magnetically charged black holes is not vertical at the critical point. We expect that the same mechanism works although the structure of the "colored" black hole is different from that of magnetic case .

4 Thermodynamical properties

In this section we investigate the thermodynamical properties of the dilatonic black holes with the GB term. There are several reasons why we are interested in this topic. As we know, Kerr-Newman black hole is regarded as a thermodynamical object. Its thermodynamical properties are confirmed when we take into account quantum effects, which may become important for small black holes. Although black hole thermodynamics in the non-Einstein theories is not well understood, we can define the temperature of the black holes in our model and show that they obey the first low of thermodynamics. We can also discuss the evolution of our black holes via the evaporation processes. Then we have two possible scenarios. One possibility is that the temperature of the black hole vanishes at finite critical mass and the state of the black hole cannot reach this critical point by finite physical processes. This scenario is similar to that of the Reissner-Nordström black hole. Another possible scenario is that the temperature stays finite and the black hole's mass approaches the critical mass, and finally a naked singularity appears in the hypersurface where the event horizon was located. Which scenario descrives our new black holes?

The GM-GHS solutions and the non-Abelian black holes have a characteristic thermodynamical property. That is, a discontinuity of the heat capacity of the GM-GHS solution appears, depending on the coupling constant of the dilaton field γ. Its critical value is just what we use in this paper[3]. Similar properties were obtained in the EYMD system[16]. This is another reason why we investigate the thermodynamical properties of our new black holes.

The Hawking temperature is given as

$$T = \frac{1}{4\pi r_h} e^{-\delta(r_h)} \left[1 - 2\tilde{m}'(r_h)\right], \tag{22}$$

for the metric (4). The inverse temperature $\beta \equiv 1/T$ versus the gravitational mass is shown in Fig. 4. We show the branches of neutral, electrically charged ($\tilde{Q}_e = 1.0$) and "colored" black holes. All branches have finite temperatures for all mass range. Hence if a new black hole exists, its evaporation process does not stop and a naked singularity may be formed in any case.

This is physically undesirable, so we should investigate the evaporation process carefully. Recall that the extreme black hole with $\gamma > \sqrt{2}\kappa$ in the EMD system shows an infinite temperature. Holzhey and Wilczek showed that the potential, through which created particles travel away to infinity, grows infinitely high in the extreme limit. They expected that the emission rate may be suppressed to a finite value[21], but this was shown to be incorrect[22]. In our case if the potential barrier becomes infinitely large at the critical point, the emission rate might be suppressed to zero though the temperature of the black hole remains finite, as they expected in their case. Then evaporation stops and no naked singularity will be realized. We performed the same analysis and calculated the potential barrier in the background of the new solutions. However the potential remains finite against our expectation, even when the solutions approach to the critical point. Hence we conclude that the evaporating process does not stop and we must eventually be faced with a naked singularity.

The temperatures of all new black holes are always higher than that of a Schwarzschild black hole, even if the new solution is charged. This is because of the GB term, which has tendency to make the

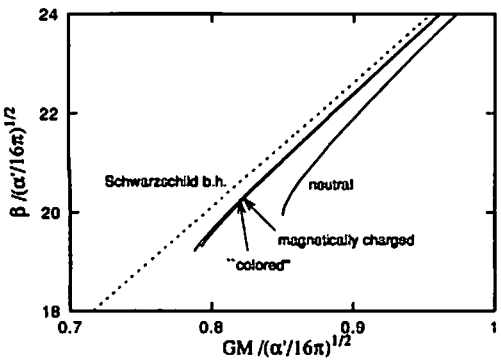


Fig. 4: The mass-temperature relations for neutral black holes, the electrically charged black holes ($Q_e = 1.0$) and the "colored" black hole ($n = 1$). The temperatures keep finite in all mass range for all branches, hence the evaporating process is not stopped and a naked singularity appears. We can see that heat capacities are always negative as in a Schwarzschild black hole case.

temperature of the black hole higher. The dominance of the GB term also appears when we discuss the heat capacity. It is always negative for all new black holes in spite of the existence of the gauge field. Hence we can conclude that most on the thermodynamical properties of the new black hole are dominated by the GB term.

5 Concluding Remarks

We investigate the dilatonic black holes with GB term and found four types of solutions, i.e., neutral, electrically charged, magnetically charged and "colored" black holes. The behaviors of field functions of those black holes are almost all the same. This may be because the gravitational structure is dominated by the GB term. Several previously unknown properties are found by putting the GB term in the action. For example, the effective energy density becomes negative near the event horizon. All types of dilatonic black holes with the GB term have critical masses below which no solution exists. At this critical point a naked singularity appears. From the study of their thermodynamical properties and the effective potential for massless scalar field, the evaporation process will not be stopped, and hence it seems inevitable that naked singularity will appeare. The heat capacity is always negative, like the Schwarzschild black hole, even if the black hole has a global charge.

However we have to note that these results are obtained by using the model (1) which includes only the leading terms of the expansion parameter α'. This expansion is not valid for solutions near the critical point. Hence there is a possibility that a naked singularity is removed by taking into account the higher or all orders in α'.

There is also a difference determined by the metric frame. In this paper we have worked only in the Einstein frame. However some results may change if we go into the string frame since the conformal transformation includes a nontrivial dilaton field. In particular if the conformal transformation becomes singular at the critical point, a naked singularity might be removed. Therefore we studied the system again in the string frame, but this gives no qualitative difference from Einstein frame.

Several problems are still left. One of them is the black hole entropy. In the model including the higher curvature term the entropy is not described by a quarter of the area of the event horizon. We are interested to know whether the entropy vanishes or not when the mass of the black hole approaches

the critical mass. This is under investigation. Another problem is the stability of the new black holes. We leave this as an open question.

– Acknowledgments –

This work was supported partially by the Grant-in-Aid for Scientific Research Fund of the Ministry of Education, Science and Culture (No. 06302021 and No. 06640412), by the Grant-in-Aid for JSPS Fellows (No. 074767), and by the Waseda University Grant for Special Research Projects.

References

[1] I. Antoniadis, J. Rizos and K Tamvakis, Nucl. Phys. B **415**, 497 (1994).

[2] R. Easther, K. Maeda and D. Wands, WU-AP/50/95, SUSSEX-AST-95/9-1, hep-th/9509074; R. Easther (these proceedings).

[3] G. W. Gibbons and K. Maeda, Nucl. Phys. B **298**, 741 (1988).

[4] D.Garfinkle, G.T.Horowitz and A. Strominger, Phys. Rev. D **43**, 3140 (1991).

[5] C.G.Callan, R.C.Myers and M.J.Perry, Nucl. Phys. B **311**, 673(1988/89).

[6] S.Mignemi and N.R.Stewart, Phys. Rev. D **47**, 5259 (1993).

[7] B.A.Campbell, N.Kaloper, R.Madden and K.A.Olive, Nucl. Phys. B **51**, 137 (1993).

[8] S.Mignemi, Phys. Rev. D **51**, 934 (1995).

[9] B.A.Campbell, N.Kaloper and K.A.Olive, Phys. Lett. B **285**, 199 (1992).

[10] M.S. Volkov and D.V. Galt'sov, Pis'ma Zh. Eksp. Teor. Fiz. **50**, 312 (1989); Sov. J. Nucl. Phys. **51**, 747 (1990); P. Bizon, Phys. Rev. Lett. **64**, 2844, (1990); H.P. Künzle and A.K. Masoud-ul-Alam, J. Math. Phys. **31**, 928 (1990).

[11] R. Bartnik and J. McKinnon, Phys. Rev. Lett. **61**, 141.

[12] D. V. Galt'sov and M. S. Volkov, Phys. Lett. A **162**, 14 (1992); N. Straumann and Z.-H. Zhou, Phys. Lett. B **234**, 33 (1990); Z.-H. Zhou and N. Straumann, Nucl. Phys. B **234**, 180 (1991); P. Bizon, Phys. Lett. B **259**, 53 (1991); P. Bizon and R.M. Wald, Phys. Lett. B **259**, 173 (1991).

[13] As a review paper see K. Maeda, Jounal of the Korean Phys. Soc. **28**, S468, (1995).

[14] K. Maeda, T. Tachizawa, T. Torii and T. Maki, Phys. Rev. Lett. **72**, 450, (1994); T. Torii, K. Maeda and T. Tachizawa, Phys. Rev. D**51**, 1510, (1995).

[15] T. Tachizawa, K. Maeda, and T. Torii, Phys. Rev. D **51**, 4054 (1995).

[16] T. Torii and K. Maeda, Phys. Rev. D **48**, 1643 (1993).

[17] G. Levrelashvili and D. Maison, Phys. Lett. B **295**, 67 (1992); Nucl. Phys. B **410**, 407 (1993); hep-th/9307159; E. E. Donets and D. V. Gal'tsov, Phys. Lett. B **302**, 411 (1993).

[18] P. Kanti, N. E. Mavromatos, J. Rizos, K. Tamvalis and E. Winstanley, hep-th/9511071.

[19] E. E. Donets and D. V. Gal'tsov, Phys. Lett. B **352**, 261 (1995).

[20] A. A. Ershov and D. V. Gal'tsov, Phys. Lett. A **150**, 159 (1993); P. Bizon and Q. T. Popp, Class. Quantum Grav. **9**, 193 (1992).

[21] C. F. E. Holzhey and F. Wilczek, Nucl. Phys. B. **380**, 447 (1992).

[22] J. Koga and K. Maeda, Phys. Rev. D **52**, 7066 (1995).

Non-singular Superstring Cosmology

Richard Easther

easther@cfi.waseda.ac.jp

Department of Physics, Waseda University,
3-4-1 Okubo, Shinjuku-ku, Tokyo, Japan.

Abstract

We study the cosmological implications of the one-loop terms in the string expansion. In particular, we present non-singular solutions which interpolate between a contracting universe and an expanding universe. We demonstrate the existence of a class of solutions which are marginally inflationary near the Planck scale, providing a mechanism for alleviating the initial conditions problems peculiar to a closed FRLW universe.

1 Introduction

The study of cosmology leads inexorably to an epoch in which the energy density of the universe approaches the Planck scale. While superstring theory is the leading candidate for a description of fundamental physics valid at these energies, it has not developed to the point where it can be used to construct detailed cosmological scenarios. Despite this, the generic properties of superstring models may be extracted, and their cosmological consequences investigated. In particular, since string theory has the ability to tame the infinities inherent in low-energy physics, it is reasonable to expect that the singularity associated with the big bang in general relativistic cosmology will be absent from a stringy cosmological model. Furthermore, the dualities associated with the full superstring action have been used by Gasperini, Veneziano and others to motivate the "pre-big-bang scenario" [1, 2]. This is a non-singular cosmology in which many of the problems associated with the "standard" big bang are alleviated by a period of accelerated contraction. The fundamental requirement

for such a theory is the existence of non-singular or "branch-changing" [3] solutions which smoothly interpolate between a contracting universe and an expanding one without passing through a singularity.

At the same time, many cosmologists have adopted an alternative strategy, which is to approximate the full string action with the first terms of the perturbative string expansion. When taking this route the non-perturbative features of string theory are implicitly discarded, so this approach will clearly break down at the highest energies. The lowest order, or "tree-level", action [4, 5, 6] has received the most attention.[1] In particular, Easther, Maeda and Wands [7] analyze the cosmological solutions derived from the tree-level action, including contributions from the dilaton, central charge and axion, with the assumption of a four dimensional maximally symmetric spacetime and conclude that all solutions of this system are singular at either or both early and late times. Furthermore, Kaloper, Madden and Olive have constructed no-go theorems for branch changing solutions in generic tree-level actions [8, 9]. Consequently, it is highly likely that solutions of the tree-level sting action are all singular, and that this approximation omits too much to adequately describe stringy cosmological models.

However, two features of the tree-level solutions offer hope that the picture may look very different at higher orders in the string expansion. Firstly, as solutions to the tree-level equations of motion approach a singularity the dilaton, ϕ, always becomes arbitrarily large. The contribution of the n-th loop terms to the action is proportional to $e^{n\phi}$ so as the tree-level solution becomes singular the higher order terms will dominate and these have the potential to remove the singularities present in the tree-level solutions. Secondly, even at tree-level we find solutions where the scale factor makes non-singular oscillations, implying that the strong energy condition is violated by the string matter. While this does require the existence of globally non-singular solutions, the violation of the strong energy condition is a necessary condition for the construction of branch changing solutions.

Clearly, the next step in the study of the low-energy string action is to add the one-loop order terms. These have received less attention than the tree-level case. However, Antoniadis, Rizos and Tamvakis [10] have demonstrated the existence of non-singular solutions with one loop string and moduli corrections with a spatially flat background in which the scale factor increases monotonically from a (non-zero) constant value. Note, these do not constitute branch-changing solutions as there is no transition between contraction and expansion.

In this paper we extend the work of Antoniadis, Rizos and Tamvakis to Robertson-Walker metrics with non-zero spatial curvature. With this extension we will be able to demonstrate the existence of branch changing solutions which connect a contracting universe to an expanding one. Furthermore, we consider the possibility of inflationary solutions to the one-loop equations of motion. Any non-singular solution is

[1] See ref. [7] for a full list of references.

necessarily inflationary, in the sense that at some time the growth of the scale factor will be accelerating at some time. However, we find a new class of marginally inflationary solutions in spacetimes with positive spatial curvature, which may provide a mechanism for solving the initial conditions problem inherent to spatially closed Robertson-Walker spacetimes. A longer account of this work is contained in [11].

2 Action and Equations of Motion

We take as our starting point the action [10, 12, 13]

$$S = \int d^4x\sqrt{-g}\left\{\frac{1}{2}R + \frac{1}{4}(D\phi)^2 + \frac{3}{4}(D\sigma)^2 + \frac{1}{16}\left[\lambda e^{\phi} - \delta\xi(\sigma)\right]R^2_{\rm GB}\right\}. \qquad (1)$$

where R is the Ricci scalar, ϕ is the dilaton and σ is a modulus field and $R^2_{\rm GB}$ is the Gauss-Bonnet combination,

$$R^2_{\rm GB} = R_{\mu\nu\kappa\lambda}R^{\mu\nu\kappa\lambda} - 4R_{\mu\nu}R^{\mu\nu} + R^2. \qquad (2)$$

Terms which do not contribute to the equations of motion for a Robertson-Walker metric have been dropped. The co-efficient λ is a positive quantity, related to the four dimensional string coupling, while δ is fixed by the trace anomaly of the $N = 2$ sectors of the theory. It will be important that δ can take both positive and negative values. The potential, $\xi(\sigma)$ is defined in terms of the Dedekind η function,

$$\xi(\sigma) = \ln\left[2e^{\sigma}\eta^4(ie^{\sigma})\right] \qquad (3)$$

where η is

$$\eta(\tau) = q^{1/12}\prod_{n=1}^{\infty}(1-q^{2n}), \qquad q = e^{i\pi\tau}. \qquad (4)$$

Anticipating that only first and second derivatives of ξ appear in the equations of motion, we note that

$$\xi'(\sigma) = 1 - \frac{\pi e^{\sigma}}{3} + 8\pi e^{\sigma}\sum_{n=1}^{\infty}\frac{ne^{-2n\pi e^{\sigma}}}{1-e^{-2n\pi e^{\sigma}}}. \qquad (5)$$

Despite its appearance, this is an odd function of σ and is well approximated by

$$\xi' \sim -\frac{2\pi}{3}\sinh(\sigma). \qquad (6)$$

where a prime denotes differentiation with respect to σ. We have used this approximate form of ξ' in our numerical calculations, as it preserves the symmetries and asymptotic form of the full potential but is much faster to compute.

Previously Antoniadis, Rizos and Tamvakis examined the cosmological solutions for this system that have a spatially flat Robertson Walker metric. We now extend their work to include the possibility that the spatial hypersurfaces have non-zero curvature, and so the appropriate ansatz for the line element is

$$ds^2 = dt^2 - e^{2\omega(t)}\left[\frac{1}{1-kr^2} + r^2\left(d\theta^2 + \sin^2(\theta)d\phi^2\right)\right]. \qquad (7)$$

The equations of motion are

$$3\dot{\omega}^2 + \tfrac{3k}{e^{2\omega}} - \tfrac{1}{4}\dot{\phi}^2 - \tfrac{3}{4}\dot{\sigma}^2 + 24\dot{f}\dot{\omega}^3 + \tfrac{24k\dot{f}\dot{\omega}}{e^{2\omega}} = 0, \tag{8}$$

$$2\ddot{\omega} + 3\dot{\omega}^2 + \tfrac{k}{e^{2\omega}} + \tfrac{1}{4}\dot{\phi}^2 + \tfrac{3}{4}\dot{\sigma}^2 + 16\dot{f}\dot{\omega}^3 + 8\ddot{f}\dot{\omega}^2 + 16\dot{f}\ddot{\omega}\dot{\omega} + \tfrac{8k\ddot{f}}{e^{2\omega}} = 0, \tag{9}$$

$$\ddot{\phi} + 3\dot{\omega}\dot{\phi} - 2\tfrac{df}{d\phi}R^2_{\rm GB} = 0, \tag{10}$$

$$\ddot{\sigma} + 3\dot{\omega}\dot{\sigma} - \tfrac{2}{3}\tfrac{df}{d\sigma}R^2_{\rm GB} = 0, \tag{11}$$

where a dot denotes differentiation with respect to t. The Gauss-Bonnet term is

$$R^2_{\rm GB} = 24\left(\ddot{\omega} + \dot{\omega}^2\right)\left(\dot{\omega}^2 + \frac{k}{e^{2\omega}}\right) \tag{12}$$

and f is defined to be

$$f = \frac{1}{16}\left[\lambda e^{\phi} - \delta\xi(\sigma)\right]. \tag{13}$$

By setting $k = 0$ we recover the system studied by Antoniadis, Rizos and Tamvakis. If f vanishes the system reduces to two free scalar fields, minimally coupled to gravity.

The equations above consist of three second order equations and a constraint, giving a total of five degrees of freedom. In order to facilitate the analysis we isolate each of the second derivative terms (remembering that $\ddot{f}$ implicitly contains $\ddot{\phi}$ and $\ddot{\sigma}$), on the left hand side, giving the following system

$$\ddot{\omega} = -\dot{\omega}^2 - \left(\dot{\omega}^2 + \frac{k}{e^{2\omega}}\right)\chi, \tag{14}$$

$$\ddot{\phi} = -3\dot{\omega}\dot{\phi} - 3\lambda e^{\phi}\left(\dot{\omega}^2 + \frac{k}{e^{2\omega}}\right)^2\chi, \tag{15}$$

$$\ddot{\sigma} = -3\dot{\omega}\dot{\sigma} + \delta\xi'(\sigma)\left(\dot{\omega}^2 + \frac{k}{e^{2\omega}}\right)^2\chi, \tag{16}$$

where

$$\chi = \frac{8 + \lambda\dot{\phi}^2 e^{\phi} - \delta\dot{\sigma}^2\xi''}{4 + 2(\lambda\dot{\phi}e^{\phi} - \delta\dot{\sigma}\xi')\dot{\omega} + (\dot{\omega}^2 + ke^{-2\omega})^2(3\lambda^2 e^{2\phi} + \delta^2\xi'^2)}. \tag{17}$$

If we wish, we can eliminate any one of the variables by inserting the constraint. However, it will be more convenient to work with the equations as they are given above, and retain the constraint for checking the consistency of our numerical calculations.

We now turn to the analysis of the possible types of solution which these equations possess. We begin with a discussion of the different asymptotic forms of the solutions, and in the next section use numerical techniques to investigate the general solution properties.

3 Asymptotic Solutions

In general this system of equations must be solved numerically. However, considerable insight into the late-time behavior of this system can be gained solely through analytic considerations.

We are particularly interested in "branch changing" solutions, where a large contracting universe is smoothly converted into an expanding one and so the scale factor

passes through a local minimum. Hence we begin our investigation by considering the possible extrema of ω. When the scale factor, $a = e^{\omega}$, passes through a local minimum, its second derivative, $\ddot{a}$, must be positive. Since $\ddot{a} > 0$ is the defining condition for inflation it follows that any branch changing solution is also inflationary. Note, this is the minimal requirement for inflation and the astrophysical constraints that must be satisfied by a successful inflationary model are much more stringent.

To show that it is possible for ω (and therefore a) to pass through a local minimum, consider the constraint in the form,

$$3\left(\dot{\omega}^2 + \frac{k}{e^{2\omega}}\right)\left(1 + 8f\dot{\omega}\right) - \frac{1}{4}\dot{\phi}^2 - \frac{3}{4}\dot{\sigma}^2 = 0. \tag{18}$$

For $k = 0$ with $\dot{\omega} = 0$ we must have $\dot{\sigma} = \dot{\phi} = 0$ as well. This is an exact, unstable, static solution to the equations of motion where the values of ω, σ and ϕ are all arbitrary constants. The non-singular solutions of Antoniadis, Rizos and Tamkvakis can be regarded as the consequence of making a small perturbation to this exact solution. If the spatial hypersurfaces have negative curvature, $k = -1$ and the constraint cannot be satisfied when $\dot{\omega} = 0$. Consequently, the scale factor is again monotonic, but the constant solution no longer exists.

Lastly, if $k = 1$ the constraint may be satisfied with $\dot{\omega} = 0$ and $\dot{\phi}, \dot{\sigma} \neq 0$, and so the scale factor can possess extremal values. The type of extrema follows from the sign of $\ddot{\omega}$ when $\dot{\omega} = 0$. When $\dot{\omega} = 0$ the denominator of χ must be positive, so

$$\mathrm{Sign}(\ddot{\omega}|_{\dot{\omega}=0}) = -\mathrm{Sign}(8 + \lambda\dot{\phi}^2 e^{\phi} - \delta\dot{\sigma}^2\xi''). \tag{19}$$

While λ is physically restricted to positive values, δ is not. Since ξ'' is negative for all values of σ, it follows that if $\delta \geq 0$, $\ddot{\omega}$ will be positive at an extremum of ω, but can take either sign if $\delta < 0$. Thus when $k = 1$ and $\delta < 0$ the scale factor can possess local minima, and therefore make one or more non-singular oscillations. However, this does not establish the existence of globally non-singular solutions. In the next section, however, we will display a specific non-singular solution and probe the region of "initial conditions space" for which the universe is non-singular.

Consider the possible asymptotic forms of the solutions to the equations of motion. For $k = -1$ we know that ω is monotonic. There is a trivial exact solution,

$$\omega = \ln t, \tag{20}$$

$$\phi = \phi_0, \tag{21}$$

$$\sigma = \sigma_0. \tag{22}$$

This describes an empty, curvature dominated universe. It is of limited physical interest, and since it applies when both λ and δ are zero it is not related to the one-loop terms we are considering here.

Now consider the evolution of a spacetime with positive spatial curvature, or $k = 1$. In this case the scale factor need not be monotonic, raising the possibility

of branch-changing solutions. In order to investigate the possible solutions that can grow to arbitrarily large sizes when $k = 1$, consider three possible types of late time behavior:

$$\begin{aligned} &\text{Type I} \quad \ddot{a}(t) > \epsilon, t > T, \\ &\text{Type II} \quad \ddot{a}(t) < -\epsilon, t > T, \\ &\text{Type III} \quad \ddot{a}(t) \to 0, t \to \infty, \end{aligned}$$

where ϵ is a positive real number. Type I describes a universe that inflates forever. In this case, the curvature terms will become negligible and the universe will asymptotically resemble the $k = 0$ case already analyzed by Antoniadis, Rizos and Tamvakis. However, their work shows that there are no solutions for $k = 0$ which inflate forever, and we therefore deduce that when $k = 1$ there are no solutions which have the asymptotic form of Type I, since their existence would lead to a contradiction with the results derived when $k = 0$. The second case describes non-inflationary expansion. However in this case the curvature term will dominate, and the universe will reach a maximum size and recontract. Such a solution does not necessarily approach a singularity, since the scale factor could evolve through a local minimum, but it cannot expand indefinitely in this manner.

Thus we are left with Type III. If a was exactly proportional to time, then $\ddot{a} = 0$ and the Gauss-Bonnet term would vanish identically. This linear solution is not a feature of the tree level equations of motion, and cannot exist at one-loop either. However, there is class of solutions which has the asymptotic behavior

$$\begin{aligned} \ddot{a}(t) &\to 0, \\ \phi(t) &\to \phi_0, \\ \sigma(t) &\to \sigma_0 \pm 2\log|t| \end{aligned} \tag{23}$$

where a_0, ϕ_0 and σ_0 are constants and we have assumed that ϕ (and hence e^{ϕ}) is not large. The properties of this solution can be investigated via numerical integration (which is done in the following section). An analytic discussion of the asymptotic form of $a(t)$ requires us to consider terms of the form $t(\ln t)^n$ whose second derivative tends to zero at late times, while allowing the Gauss-Bonnet term to remain non-trivial. We do not present this work here, but note that from a qualitative point of view the growth of $\xi'(\sigma)$ as $|\sigma|$ becomes large means the one-loop corrections have a major impact on the evolution of the modulus, even as the Gauss-Bonnet term becomes very small.

4 Numerical Results

The asymptotic solution, equation (23), was presented without a detailed derivation. However, in Fig. (1) we show a specific example of a solution which has this asymptotic form as $t \to \pm\infty$. This solution represents a specific example of a non-singular,

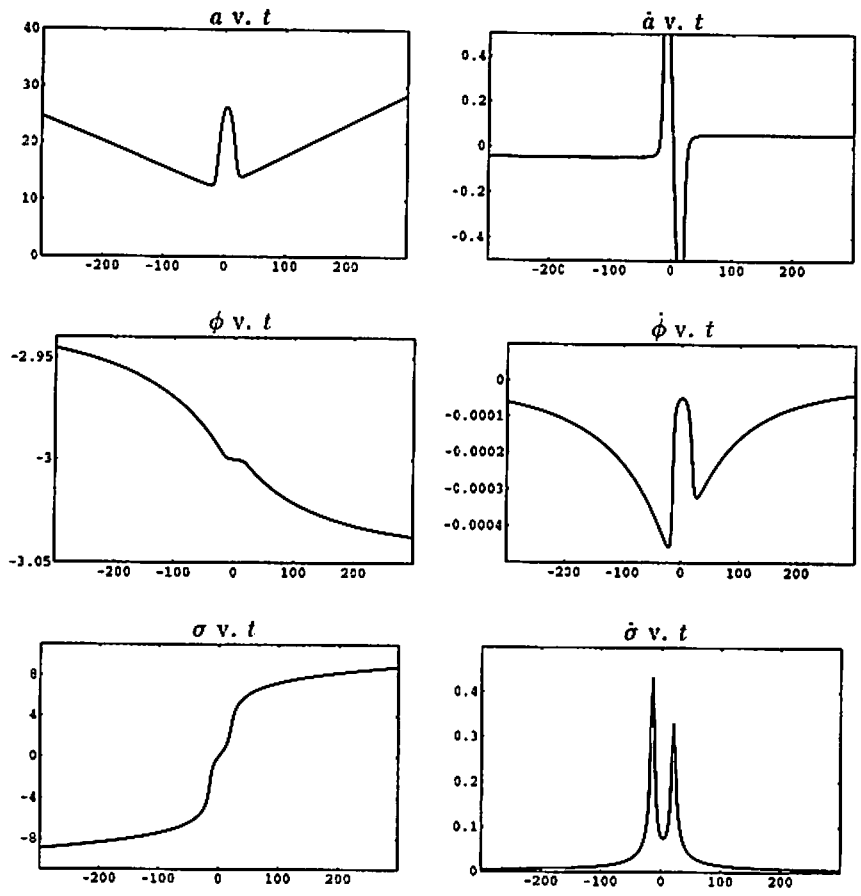


Figure 1: A particular non-singular solution is displayed, where $\delta = -64 \times 3/\pi$ and $\lambda = 1$. The chosen initial values are $\sigma = 0$, $\phi = -3$, $\dot{\phi} = -5 \times 10^{-4}$, $\dot{\sigma} = 0.08$ and $\dot{\omega} = 0.01$, with the constraint requiring $\omega_0 = 3.2115$. The scale factor has a minimum value of approximately 12.5, and approaches linear expansion as $t \to \pm\infty$.

branch-changing solution. Other solutions exist that interpolate between an initially singular solution and eventual linear expansion. At tree level, all solutions to the equations of motion with $k = 1$ are singular in the past and in the future.

Having demonstrated the existence of a specific non-singular solution, however, we wish to determine whether these solutions are generic, or correspond to a special (measure-zero) region of parameter space. For the $k = 0$ case Antoniadis, Rizos and Tamvakis gave a phase-space analysis. However, the additional complexity introduced by allowing non-zero spatial curvature makes this problem less tractable when $k = 1$ and we employ a different method.

Specifically, we want to study the basin of attraction for the non-singular solutions, which is the volume of "initial conditions space" that evolves towards the linearly expanding solution (in either the past or the future), rather than a singular collapse. This approach was recently used by Cornish and Levin [14] to study the

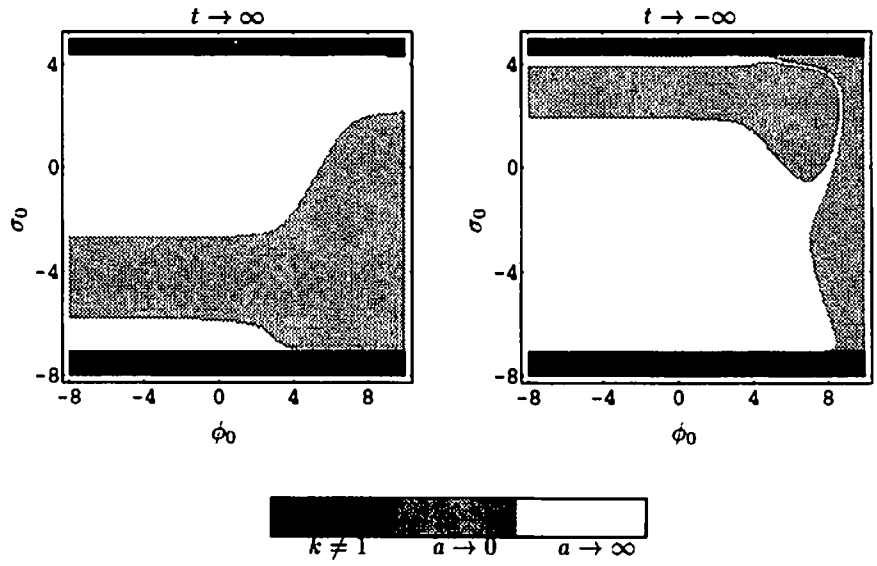


Figure 2: These plots show the asymptotic form of the solutions to the equations of motion when the dilaton and modulus have the initial values ϕ_0 and σ_0, and the other parameters are the same as for the solution depicted in Fig. 1. Points where $a \to \infty$ approach the linearly expanding solution (white), while $a \to 0$ imply the existence of a singularity (grey). Choices of initial conditions which would require $k \neq 1$ are not plotted (black). The structure shown in the top right of the $t \to -\infty$ panel is "real", but it corresponds to the strong coupling ($\phi > 0$) region where the one-loop action is unlikely to hold. The numerical integrations were carried out at 240 × 240 equally spaced points in the (ϕ_0, σ_0) plane.

basins of attraction for two field inflation.

Since the initial conditions space is five dimensional, it is most convenient to consider two dimensional slices through it. In practice we take a 240 × 240 grid of points in the (ϕ_0, σ_0) plane, with the other parameters fixed, solve the equations of motion numerically for each choice of initial data and deduce the asymptotic form in the past and future. In Fig. 2 a piece of the (ϕ_0, σ_0) plane is shown, and the choices of initial conditions which lead to linear expansion in both the past and future are apparent.

In Fig 3 the initial values of (ϕ_0, σ_0) for which a non-singular solution results are plotted. This is just the intersection of the values of (ϕ_0, σ_0) which lead to a linearly expanding solution in the past and in the future. In addition, the subset of initial conditions space in which the higher loop terms are expected to be smaller than the terms included in the action (on the basis of dimensional analysis) are superimposed. This region is defined heuristically, and the definition adopted here is that the one-loop approximation is trusted if the scale factor is always greater than unity and the kinetic terms ($\dot{\omega}$, $\dot{\phi}$) and $\dot{\sigma}$ are always less than unity (in Planckian units) for a given solution.

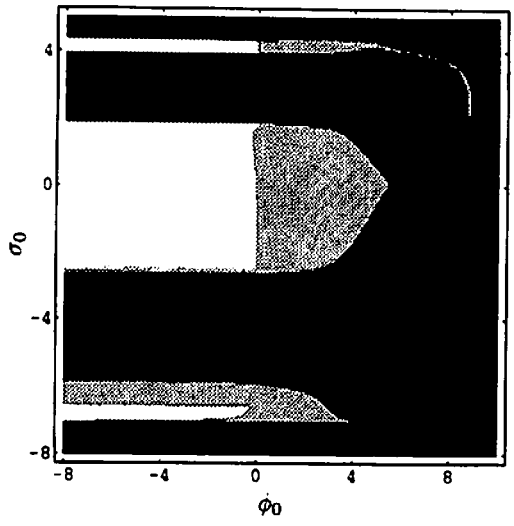


Figure 3: This plot shows the initial conditions which lead to a non-singular universe, with the parameter values used in Fig. 2. All of the non-singular solutions correspond to a large, contracting universe that smoothly evolves into an expanding solution. Initial conditions leading to a singular solution are given in black. The remaining set is divided into the grey region for solutions which are non-singular, but where the higher-loop terms may be significant, and the white region where the one-loop terms are probably dominant.

Interestingly, this region appears to possess a fractal structure (which is not displayed in detail here), indicating the presence of a strange attractor and chaos in the equations of motion derived for this one-loop superstring cosmology. This will be the subject of a future publication, and we will not consider it further here. For our present purposes, we merely wish to point out that the region of parameter space in which the one-loop terms are not dominated by higher order solutions is a non-trivial one, lending support to the belief that the non-singular solutions found here are unlikely to be removed by higher order corrections.

5 Discussion

In this paper we have demonstrated the existence of non-singular cosmological models derived from the one-loop approximation to the superstring action. In addition, we have estimated the likely impact of higher order terms by dimensional analysis and can have confidence that these solutions correspond to regions of parameter space where the one-loop action will not be dominated by higher order corrections.

While the inflationary scenario is typically invoked to provide the initial conditions required by the standard model of a hot big bang, one problem not addressed by inflation is that a spatially closed FLRW universe typically has a lifetime on the order of the Planck scale, which may not allow enough time for inflation to begin.

However, the linearly expanding solution displayed here for the $k = 1$ case alleviates this problem, as it allows a closed universe to expand indefinitely. At some point there must be a transition to "normal" inflation, which cannot be described by the model being studied here.

The model presented here provides interesting new examples of non-singular superstring cosmologies, and for the first time we have found cosmological solutions derived from the perturbative superstring action which smoothly interpolate between a contracting universe and an expanding one. Such solutions have applications to the pre-big-bang cosmological scenario and the suppression of singularities in string motivated theories of gravitation.

References

[1] M. Gasperini and G. Veneziano, Astropart. Phys. **1**, 317 (1993).

[2] M. Gasperini and G. Veneziano, Phys. Rev. D **50**, 2519 (1994).

[3] R. Brustein and G. Veneziano, Phys. Lett. B **329**, 429 (1994).

[4] E. S. Fradkin and A. A. Tseytlin, Phys. Lett. B **158**, 316 (1985).

[5] C. G. Callan, D. Friedan, E. J. Martinec, and M. J. Perry, Nucl. Phys. B **262**, 593 (1985).

[6] C. Lovelace, Nucl. Phys. B **273**, 413 (1986).

[7] R. Easther, K. Maeda, and D. Wands, Phys. Rev. D (in press) (1996).

[8] N. Kaloper, R. Madden, and K. A. Olive, Nucl. Phys. B **452**, 677 (1995).

[9] N. Kaloper, R. Madden, and K. A. Olive, hep-th/9510117 (1995).

[10] I. Antoniadis, J. Rizos, and K. Tamvakis, Nucl. Phys. B **415**, 497 (1994).

[11] R. Easther and K. Maeda, WU-AP/58-96 (1996).

[12] I. Antoniadis, E. Gava, and K. S. Narain, Nucl. Phys. B **383**, 93 (1992).

[13] I. Antoniadis, E. Gava, and K. S. Narain, Phys. Lett. B **283**, 209 (1992).

[14] N. J. Cornish and J. J. Levin, astro-ph/9510010 (1995).

SPACE-BASED GRAVITATIONAL WAVE ASTRONOMY

Nobuki Kawashima (ISAS), Ryutaro Takahashi (NAO) and Seiji Kawamura (Caltech)

ABSTRACT

Space-based gravitational wave astronomy is reviewed, particularly based on LISA (Laser Interferometer Space Antenna) which ESA (European Space Agnecy) has chosen as one of 3 candidates for the Cornerstone missions whith be launched in 2010-2020. Comparing with the ground based antenna, it is pointed out that the space-based antenna has an advantage of much more target objects inace and also better S/N and it can be a short cut to the gravitational wave astronomy. The comparison is also made between the space orbital antenna and the one on the lunar base.

1. Importance of Gravitational Wave Detection

The importance of gravitational wave detection is summarized as follows :

Establish Gravitational Wave Astronomy

-------Open a new window to the universe

1) World of 10^{8} tons / cm^{3} (Center of neutron stars)

2) Birth and evolution of the universe before it became clear to the electromagnetic wave astronomy

2. Development of Laser Interferometer Gravitational Wave Antenna on the Ground

Extensive efforts are going on the ground for the detection of the gravitational wave on the earth aiming for the first detection :

i) Proto-type Antenna

1) Fabry-Perot Type 40m : CALTECH(U.S.A.) 10m : Univ. of Glasgow(England)
20m : National Astron. Obs.(NAO)(Japan)

2) Delay-line Type 30m : Max Planck Inst.(Germany)
100m(TENKO-100) : Inst. Space Astronautical Sci. (ISAS) (Japan)

ii) Construction of km class GW Observatories has started:

They will start to operate from 2000.

1) LIGO (U.S.A.): 4 km x 2 (Hanford Washington, Livingston Louisiana)

2) VIRGO(French-Italian collaboration): 3 km(Pisa, Italy)

At the same time, GEO-600(German-English Collaboration) and TAMA-300(Japan) are aiming to be completed before LIGO and VIRGO.

3. Strong Support to the Gravitational Wave Detector Development

To establish the gravitational wave astronomy is supported not only by the science community of gravitational physics but also by communities of elementary particle physics, high energy physics, accelerators, cosmology etc. It is anticipated that the gravitational wave astronomy may do something which neither electromagetic wave astronomy nor high energy accelerators can do.

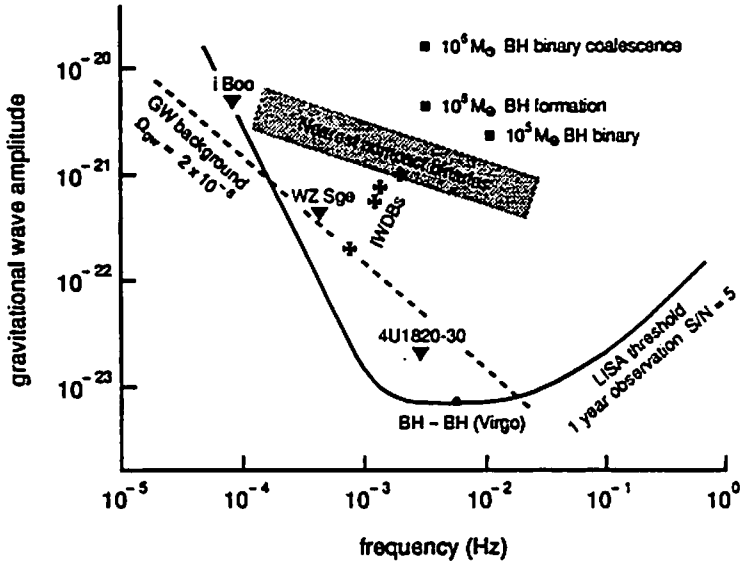


Fig.1. Target objects and its signal level for space-based gravitational wave observation [1]

One of the evidence is that in LIGO, the total budget recently inflated from M$250 to M$360 and it has been able to survive approved by NSF.

4. Development into Space--------21st Century

Objects which the space laser interferometer gravitational wave antenna aims :

	Objects	Frequency Range
Ground	Coalescence of binaru pulsars	50^{-1} kHz
	Supernovae	
Space (Orbital)		
	Binary pulsar in our galaxy	$10 - 10^{-4}$ Hz^{-1}
	Coalescence of giant black holes	$10\ Hz^{-4}$
	Background radiation	$10 - 10^{-4}$ Hz^{-1}

(The world from the Big Bang within 300,000 years, which cannot be seen by the electromagnetic waves)

5. Moon or Orbital ?

5-1 Advantages for the Lunar Laser Interferometer Gravitational Wave Antenna

i) Vacuum : No vacuum tanks and tubes and pumping system which occupy the major part of the construction cost on the ground is needed

ii) No scattering light effect

iii) Craters : Naturally provided most suitable structure for the interferometer free from the effect of the lunar curvature for km-class interferometers so that no heavy civil engineering work is required, which occupies another major part of the construction cost on the ground.

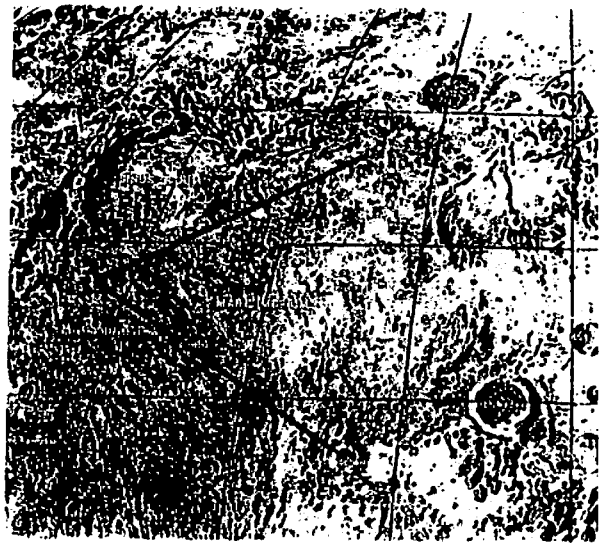


Fig.2. laser interferometer gravitational wave antenna on the lunar surface which utilizes the naturally provided advantageous structure

4-2 Disadvantage

The lunar surface is not so quiet in respect of seimic noises. Seismic noises arising from the internal activity of the moon and the thermal change of lunar surface material at he boundary of day and night due to its large temperature change [2].

From the discussion above, it is clear that the orbital antenna is much superior to the lunar one so long as it is discussed from only the scientific point of view.

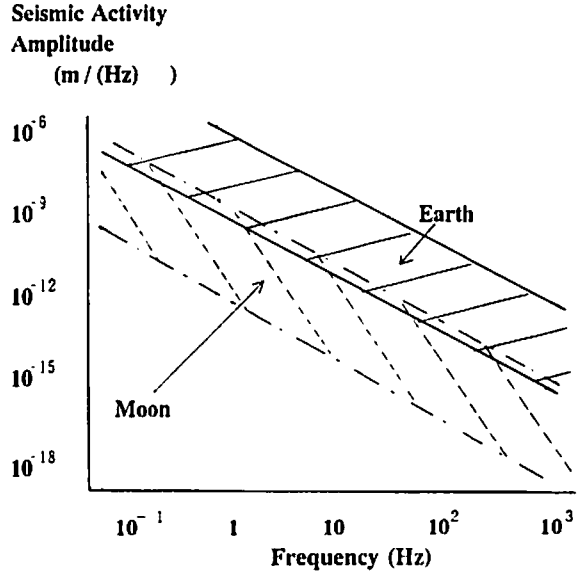


Fig.3. Seismic noise environment on the lunar surface

6. Background and Past Progress of Space-based Antenna

From the fall of 1993 till the fall of 1994, ESA(European Space Agency) investigated possible missions of M 3 / M 4(Medium scale mission to be launched in around 2003) and Cornerstone mission(Large scale mission to be done in 2010) [3].

A space orbital laser interfero-meter gravitational antenna LISA (Laser Interferometer Space An-tenna) has been selected as one of 3 candidates for the corner stone mission(Other 2 are Mercury Orbiter and Spectrometric Survey of the Sky) [4].

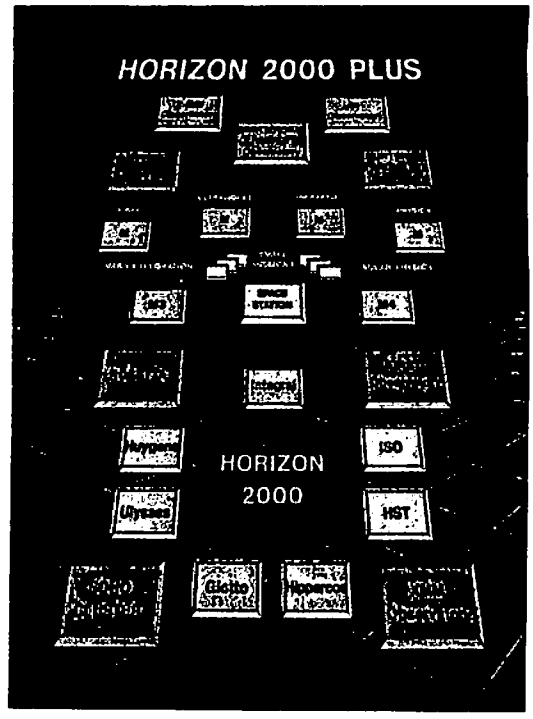


Fig.4. Space science missions scheduled/planned in early 21st century [3]

7. LISA (Laser Interferometer Space Antenna)

7-1 Configuration

LISA places 4-6 spacecrafts in the solar orbit with the separate ar difference of 5 10^6 km. In the center, 2 spacecrafts(c-1,c-2) are placed and 1 or 2 are at the ends. The angle between 2 arms is 60

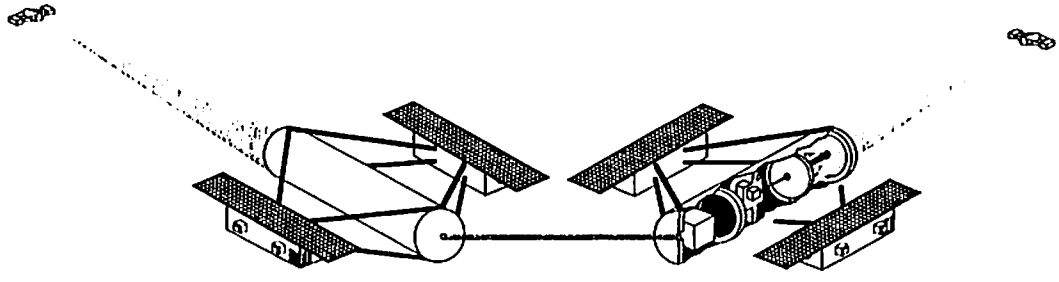


Fig.5. LISA (Laser Interferometer Space Antenna) one of candidates for the cornerstone mission around 2010-2020 [4]

degrees. In space mission, once launched, it is impossible to repair so that one spacecraft among 4 spacecrafts works as a redundancy and two spacecrafts for the 6 spacecraft configuration.

7-2 Orbit

The main orbital plane of LISA 4-6 spacecrafts lies on the earth's orbit around the sun(1 AU) and 20 degrees behind the earth. It is inclined 60 degrees to the ecliptic plane and there the interferometer is formed with an arm length of 5 million km.

This orbit has been selected that the time variation of the relative distance between spacecrafts are minimum.

Distance between the center and the end(arm length)		5×10^6	km
Distance between the center spacecrafts		10	km
Accuracy of the arm length	<	1	km
Accuracy of the difference of the arm length	<	20	m
Time variation of the difference of the arm length	<	7	mm/s

Table 1 Parameters of LISA[4]

7-3 Interferometer and Optical Components

The interferometer is an Michelson type and a laser oscillator is installed in each of two central space-crafts C-1 and C-2 for redundancy. One of the difference in the space antenna from the ground one is that the laser light emitted from the center comes back not reflected by a mirror but a transponder. Very weak light received at the end spacecraft is amplified

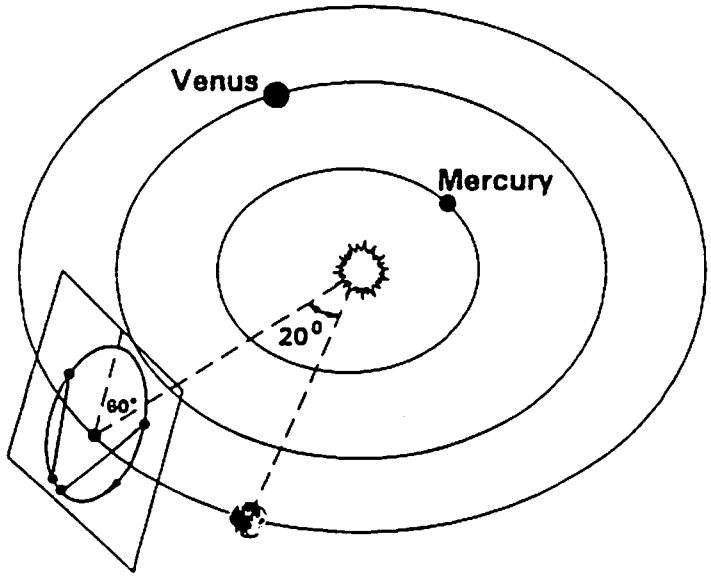


Fig. 6. Orbit of LISA system

in phased-locked in the laser light frequency and sent back to the center.

Mirror	: 38cm ϕ Cassegrain telescope	Surface roughness < λ /30
Laser	: Diode laser-pumped Nd-YAG 2 W	single frequency, single mode
	Frequency stability<10 Hz/(Hz)$^{1/2}$	(limited by the temperature change)

Tabel 2 Laser and mirror in LISA

7-4 Drag-free Configuration

Another important requirement to the space orbit laser interferometer is the drag-free configuration in which any drag force(non-gravitational force) to the spacecraft such as the solar radiation pressure should be compensated.

The requirement is $< 1.5 \cdot 10^{-15}$ m·s^{-2}/(Hz)$^{1/2}$

The drag-free control is realized by placing a reference sphere at the center gravity of the spacecraft free from any drag forces and by controlling the relative location of the spacecraft with respect to the reference sphere unchanged.

7-5 Nano-radian Control

The attitude control and the alignment control of the laser beam is so severe as nano-radian(10^{-9} rad).

Alignment control	:	nano－rad control
Pointing error(Abs.)		< 0.006 arcsec
Pointing stability		< 0.001 arcsec/(Hz)$^{1/2}$

Table 3 Requirement to the attitude in LISA

7-6 Micro-thruster

A micro-thruster is required to realize the control described above. In order for it, Cs electric field emission type micro-thruster called FEEP has been developed and partially verified in the lab. for nano-radian control [5].

7-7 Thermal Control

Furthermore, very severe thermal control is also one of key technologies in the space laser interferometer gravitational wave antenna. The thermal deformation of optical components and support structure should be kept as small as possible.

The thermal control level required is :

$$\delta T < 10^{-6} \cdot K/(Hz)^{1/2}$$

7-7 Sensitivity and Noises

The overall sensitivity and the noise figures are shoun in Table 4 :

Displacement : $\delta x < 25\ pm\ /(Hz)^{1/2}$ (@ 10^{-3} Hz - 10^{-1} Hz)

Acceleration : $\delta a < 10^{-15}$ m/sec $/(Hz)^{1/2}$ (@ 10^{-4} Hz - 10^{-3} Hz)

Shot noise	14	pm / $(Hz)^{1/2}$
Beam pointing fluctiations	4	
Laser phase measurement and offset lock	4	
Scattered light and other effects	11	

Table 4 Sensitivity and noise figures of LISA [4]

A schematic layout of the spacecraft configuration is shown in Fig. 7.
Characteristic parameters of other sub-systems are as follows :

Power supply : Ga-As solar panel 2.4 m^2
Battery : 250 Wh
Communication : X-band up/down link
RF output 5 W
30 cm high gain antenna

Table 5 Basic parameters of the spacecraft

Arian V will be used for a launch of the total spacecraft weight of about 1 ton to the earth orbit around the sun.

8.Realizability of the Space Laser Interferometer Gravitational Wave Antenna

As already mentioned

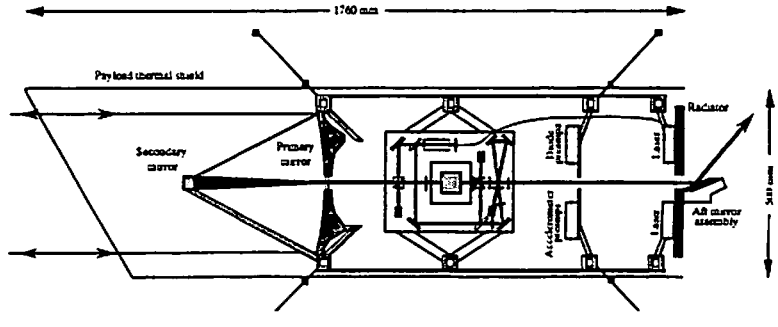


Fig.7. Schematic layout of the spacecraft configuration of LISA [4]

in Section 5, the space antenna has an advantage of much more objects in space and S/N is also much better than the ground based antenna even in its advanced stage. Therefore, from the view point of the establishment of gravitational wave astronomy, it can be said that the space antenna is much nearer to it.

In Sec.4, we compared the space orbital antenna and the one on the lunar surface. There, it was concluded that from only purely scientific point of view, the orbital antenna is much superior, however, which will be realized earlier is not self-evident. In both cases, it is a big project and is subject to other non-scientific factors. The orbital antenna can only be constructed for its own and purely scientific. It should compete with science missions of other disciplines. It is not easy in a severe environment surrounding space program. Some says it is more than the age of Winter and "the ice age."

While, the laser interferometer on the lunar surface is solely dependent how the human lunar activity is developed in future. Though the Apollo mission was so great that right now a further development is forgotten, the recent movement of "Return to the moon" in Japan [6] and Europe cannot be neglected. In Japan, the lunar exploration has been chosen as one of important future space missions in the recent 4-th amendment of Space Policy of Japan. This is in the scope of an unmanned construction of a manned lunar base in future. ESA has also started the lunar exploration program.

It will not be a dream to construction of a laser interferometer on the lunar surface in 2010s if these movement is accelerated further.

REFERENCES

[1] ESA LISA Working Group : Proposal for Laser Interferometer Gravitational Wave Detector in Space, MPQ 177, (1994)

[2] Y.Nakamura, G.V.Latham and H.J.Dorman, "Apollo Lunar Seismic Experiment-Final Summary", J. Geophys. Res., Vol.87, A117-A123, 1982

[3] ESA, "Horizon 2000 Plus" ESA Report, 1994

[4] ESA LISA Working Group "LISA for gravitational wave measurements" , ESA Assessment Study Report, SCI(94)6, 1994

[5] J.Gonzalez, G.Saccoccia and H.von Rohden, "Field Emission Propulsion", Proc. 23rd Intern'l Electr. Propulsion Conf. (Seattle 1993), 1993

[6] Lunar and Planetary Sciety, "Building the Manned Lunar Station : An Unmanned Formula" (Proposal to the Space Policy of Japan), The Lunar & Planetary Sciety, 1994

Some theoretical issues in gravitational wave physics and the black-hole perturbation approach

Misao Sasaki

Department of Earth and Space Science, Osaka University, Toyonaka, Osaka 560, Japan

(February 1996)

Abstract

There are now several on-going projects to detect gravitational waves from possible astrophysical sources and gravitational wave physics has become one of the important fields of research in cosmology and astrophysics. However, our understanding of gravitational waves is still very premature at present. In this talk, I discuss two different matters in gravitational wave physics. In the first part, I present my personal view on several important theoretical problems which prevent us from clear understanding of physics of gravitational waves. In the second part, I report on recent progress made in the black-hole perturbation approach, by which we may gain more insights into physics of gravitational waves in relativistic situations.

I. SOME THEORETICAL PROBLEMS

Gravitational waves are the propagating mode of the spacetime metric. The existence of them was pointed out soon after Einstein discovered general relativity. In the weak gravity limit, gravitational waves are described by the tranceverse-traceless (TT) components of the metric and propagate with the speed of light. They are generated whenever matter distribution changes dynamically except for perfectly spherically symmetric motion. This simple picture of gravitational waves is qualitatively valid in many realistic situations. However,

once we try to quantify phenomena associated with gravitational waves in good accuracy, there are many unsolved problems. In the following, I point out some of them.

A. Fundamental difficulties

The above picture of gravitational waves is valid only at the linear level. That is, if we write the metric as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}\,, \tag{1.1}$$

where $\eta_{\mu\nu}$ is the flat metric, the TT components of $h_{\mu\nu}$ $(= h^{TT}_{\mu\nu})$ describe gravitational waves only when the magnitude of $h_{\mu\nu}$ is much smaller than unity. The difficulty arises when the linear approximation breaks down. Because of the non-linear nature of the Einstein equations, it is not at all trivial to separate out the part of the metric that describes gravitational waves. Furthermore, the intrinsic difficulty is due to the covariance of the Einstein equations, which is the essential part of the theory. Because of covariance, it is known that the spacetime metric can be always put in the form of the flat metric locally at any spacetime point, which is a mathematical expression for the equivalence principle. This implies that even in the linear approximation limit, one cannot ascribe physical meaning to $h^{TT}_{\mu\nu}$ locally.

B. Generation problem

Since the notion of gravitational waves is not locally defined, there is no well-defined energy momentum tensor of gravitational waves $T^{\mu\nu}_{GW}$. Nevertheless, $T^{\mu\nu}_{GW}$ can be defined if averaged over a spacetime region with linear length scale several times the characteristic wavelength of the waves [1]. Thus a system will lose energy if gravitational waves are generated.

Consider a system with mass M, length scale R and characteristic matter velocity v. The gravity is weak if $GM/Rc^2 \ll 1$ and the system is said to be a slow motion source of gravitational waves if $v \ll c$. For a system approximately in virial equilibrium, we have

$GM/R \approx v^2$. In the weak gravity, slow motion source limit, the linear approximation is valid and the TT components of the metric at distance far away from the source ($r \gg R$) are known to be given by the so-called quadrupole formula (see e.g., [2]):

$$h_{ij}^{TT} = \frac{2G}{rc^4}\left[\ddot{I}_{ij}(t - r/c)\right]^{TT}, \tag{1.2}$$

where $I_{ij}(t)$ is the quadrupole moment of the system and the superscript TT on the square parenthesis denotes the TT projection of the quantity. Then the energy loss rate of the system is determined by the energy carried out by gravitational waves given by

$$\left\langle \left(\frac{dE}{dt}\right)_{GW} \right\rangle = \frac{G}{5c^5}\left\langle \overset{(3)}{I}_{ij}\overset{(3)}{I}^{ij} \right\rangle, \tag{1.3}$$

where $I\!\!I_{ij} = I_{ij} - \frac{1}{3}\delta_{ij}I^k{}_k$ and $\langle\cdots\rangle$ denotes time averaging over several periods of gravitational waves.

The above formulas for h_{ij}^{TT} and $\langle(dE/dt)_{GW}\rangle$ are valid in the leading order in v^2/c^2. It is possible to include higher order corrections to these formulas in the so-called post-Newtonian approach to gravitational waves [3]. However, the post-Newtonian approach is intrinsically incapable of dealing with fully relativistic sources of gravitational waves, i.e., for sources with strong gravity and/or fast motion. In such a situation, there are no known formula that describes the generation of gravitational waves.

C. Propagation problem

The picture that gravitational waves propagate with speed of light is only valid when the spacetime curvature can be neglected over linear scale of many wavelengths. Thus if we denote the characteristic wavelength by λ and the characteristic curvature radius by L, it is valid only when $\lambda \ll L$. If one considers an isolated source, this holds only in the wave zone where $r \gg \lambda$. However, it does not hold for the whole range of the wave zone to infinity, not to mention the region near the source [3].

To see this in a little more detail, let us consider the Einstein equations in vacuum. If

we chose the harmonic gauge $(\sqrt{-g}g^{\mu\nu})_{,\nu} = 0$, the vacuum Einstein equations are written in the form,

$$\Box_{flat}\,\bar{h}^{\mu\nu} = \Lambda^{\mu\nu}(h), \tag{1.4}$$

where $\Box_{flat}$ is the flat spacetime d'Alembertian, $\bar{h}^{\mu\nu} = h^{\mu\nu} - 1/2\eta^{\mu\nu}h^{\alpha}{}_{\alpha}$ and $\Lambda^{\mu\nu}$ is the non-linear term that contains quadratic and higher order terms in $h_{\mu\nu}$. Thus the picture that gravitational waves propagate with speed of light along the flat spacetime light cone is valid only when the non-linear corrections can be neglected. Such a situation holds in the wave zone. However, it will not hold if one considers the wave zone to infinity. This is because $h_{\mu\nu}$ contains not only the gravitational wave part but also the non-propagating longitudinal part like that describes the Newtonian potential of the source. Since the Newtonian potential decreases only as $1/r$, it gives rise to the famous Coulomb logarithmic correction to the waves. This is usually called the tail effect. In more explicit terms, it is the effect of curvature scattering. Of course, as one becomes closer to the source, there exist not only the effect of curvature scattering but also the effect of non-linear wave-wave interactions. Thus gravitational waves do not propagate only along the light cone. They also permeate into the region inside the light cone.

Again, these non-linear effects can be studied in the post-Newtonian approach if the source satisfies the weak gravity, slow motion source condition. However, not much is known or studied in the case of strong gravity source.

D. Radiation reaction

When a system emits gravitational waves, there will be radiation reaction to the system. In the weak gravity, slow motion limit, it is known that one can choose a gauge in which the metric takes the Newtonian form and the reaction is described by a single reaction potential ϕ_{react}. It is given by

$$\phi_{react} = \frac{G}{5c^5}\overset{(5)}{I}_{ij}x^i x^j, \tag{1.5}$$

where x^i are the locally Cartesian coordinates with $x^i = 0$ being the center-of-mass of the system [2].

However, the notion of locally defined reaction force as given in the above cannot exists in fully relativistic situations, simply because it is against the very nature of covariance or of the equivalence principle; the reaction must be necessarily non-local in nature. One case in which this non-locality of the reaction force can be made manifest is the test particle limit of particle motion in a given background spacetime. In this case, apart from a subtlety associated with the divergence arising from the self-gravity of the particle, it can be shown that the first order correction to the equation of motion takes the form [4,5],

$$m\frac{D}{d\tau}u^\alpha(\tau) = \frac{Gm^2}{c^6}\int_{-\infty}^{\tau} d\tau'\, v^\alpha{}_{\mu\nu}\left(z(\tau), z(\tau')\right) u^\mu(\tau')u^\nu(\tau), \tag{1.6}$$

where m is the mass of the particle, $z^\mu(\tau)$ is its world line with τ being the proper time, $u^\mu = dz^\mu/d\tau$ and $v^\alpha{}_{\mu\nu}(z,x)$ is a certain bi-tensor constructed from the tensor Green function of the background spacetime, which hence depends on the background curvature. Apparently, the deviation from the geodesic motion due to the radiation reaction depends on the entire history of the particle motion.

For general cases, it is completely unknown if there exists a covariant formula such as above which describes the radiation reaction even in a formal level.

II. BLACK-HOLE PERTURBATION APPROACH

Now let us turn to the black-hole perturbation approach. Black holes are most relativistic astronomical objects and we have exact solutions of the Einstein equations that describe black holes. Hence black holes are often used to examine various relativistic astrophysical phenomena as well as to test various theories. As for gravitational wave physics, studies of the perturbations of black-hole spacetimes will give us insight into the nature of gravitational waves in the case of strong gravity such as the effect of curvature scattering in fully non-linear regime. Furthermore, the existence of the so-called quasi-normal modes of black holes [6]

enables us to study the intrinsic physical properties of black holes in detail and perhaps has profound implications to the radiation reaction problem. Thus it is hoped that the black-hole perturbation approach can (at least partially) solve some of the unsolved problems discussed in the first part of this talk.

In the following, however, we do not consider these issues but confine our attention to a topic of gravitational radiation from coalescing compact binaries. Recently this has became a subject of great interest because one of the most promising targets of the planned future gravitational wave detectors such as LIGO [7,8]/VIRGO [9] is the gravitational radiation from compact binaries just before their coalescence, the so-called "last three minutes" [10].

A. Dawn of Gravitational Wave Physics

Recently, many gravitational wave experiment groups over the world are aiming to construct long arm-length interferometric gravitational wave detectors. Among such projects are LIGO [7,8] and VIRGO [9]. These detectors are expected to have sensitivity of $h \lesssim 10^{-21}$ over a broad frequency band ($10 \sim 1000\,\mathrm{Hz}$) and will be in operation by the beginning of the next century.

Among the possible sources of gravitational waves which can be detected by these detectors, coalescing compact binaries at their final stage of a few minutes are the most promising candidates. One reason is that we expect such events to occur $\sim 3/\mathrm{yr}$ within 200Mpc [11]. Another reason is that the amplitude of gravitational waves from an inspiraling binary at distance of r Mpc will be $h \sim 10^{-21}(100\mathrm{Mpc}/r)$, hence can be detected by LIGO and VIRGO if such events occur. Finally, since most of the binaries are expected to be settled down to quasi-circular orbits due to the radiation reaction by the time they come into the LIGO/VIRGO frequency band, it is relatively easy to make accurate theoretical predictions of the gravitational waveforms.

Gravitational radiation from coalescing compact binaries at its inspiral stage have a characteristic waveform, called a "chirp" signal, with both the frequency and amplitude

increasing rapidly until the final coalescing stage begins [8]. The last few milliseconds of coalescence are the stage which is not yet well-understood and is an issue of numerical relativity. But after that stage when the coalesced object is expected to become a black hole, the gravitational radiation will again have a characteristic waveform of a damped oscillation with frequency and decay rate determined by the mass and angular momentum of the black hole, the so-called quasi-normal mode oscillations [6].

Turning back to the inspiral state, since how the chirp signal develops in time depends on the rate of gravitational radiation, which further depends on orbital parameters of a binary, it brings us rich information about mass, spin and other physical quantities of the binary stars. In addition, together with the detected amplitude of gravitational waves, we will have an accurate measurement of the distance to the source, by which we may be able to determine cosmological parameters of the universe. Furthermore, provided we have an accurate theoretical prediction of the evolutionary behavior of inspiraling binaries based on general relativity, the observed data can be used to test the validity of general relativity or constrain alternative theories of gravity.

Also we may be able to obtain direct information about physics of high density matter such as the equation of state of neutron stars from the observed waveforms of the coalescing stage. The detection of quasi-normal mode oscillations of a coalesced object will be a direct evidence of the existence of a black hole.

All of these anticipations make us feel that the next century is an era of gravitational wave physics and astronomy. However, as mentioned in the above, it is necessary for us to have a definite and accurate theory of gravitational radiation from inspiraling compact binaries to realize these anticipations.

Observationally, physical information contained in the detected gravitational radiation is extracted out by the matched filtering technique, that is, by cross-correlating the incoming noisy signal with theoretical templates. If the signal and templates get out of phase with each other by one cycle as the waves sweep through the LIGO/VIRGO band, their cross correlation will be significantly reduced. This means that it is essential to construct theoretical

templates which are accurate to better than one cycle during entire frequency sweep through the LIGO/VIRGO band in order to make optimal use of the observed interferometric data. In a very inspiring paper by Cutler et al. [10], entitled "The last three minutes", this point was clearly emphasized. Thus, much effort has been recently made to construct accurate theoretical templates [12].

To construct theoretical templates, the post-Newtonian approximations are usually employed to solve the Einstein equations. Up to now, the gravitational wave luminosity from a binary in quasi-circular orbits has been calculated up through 2.5PN order, i.e, $O(v^5)$ beyond Newtonian quadrupole formula where v is the orbital velocity of the binary [13,14]. However, based on numerical calculations of the gravitational radiation from a particle in circular orbit around a non-rotating black hole, Cutler et al. [15] showed that evaluation of the gravitational wave luminosity to a post-Newtonian order much higher than presently achieved level might be required. Then in order to find out the necessary post-Newtonian order, the same problem was investigated by Tagoshi and Nakamura [16] with much higher accuracy, to 4PN order and concluded that the accuracy to at least 3PN order is required for the construction of effective theoretical templates. In addition, they found logarithmic terms in the luminosity at 3PN and 4PN orders.

These previous studies using the perturbation equation of a black hole show its powerfulness to the problem. Although restricted by the condition that $\mu \ll M$, where μ and M are the reduced mass and total mass, respectively, of the system, because it takes full account of relativistic effects by nature, we can spell out the relevant post-Newtonian effects from perturbation calculations in a rather straightforward manner. Hence it plays a complementary role to the standard post-Newtonian approach and provides a useful guideline for higher post-Newtonian calculations.

To strengthen the black hole perturbation approach further, it is then much desirable to develop an analytical method in which coefficients of the post-Newtonian expansion of the luminosity, for example, are evaluated exactly so that the results can be compared with those by the standard post-Newtonian calculations without any ambiguity. Poisson [17]

first developed such a method and calculated the luminosity to 2PN order from a particle in circular orbits around a non-rotating black hole. Then extending Poisson's method, a more systematic method was developed by Sasaki [18] and analytical waveforms and luminosity up through 4PN order were analytically derived by Tagoshi and Sasaki [19]. The results were in excellent agreement with those of Tagoshi and Nakamura [16]. Recently, the method has been extended to the case of a rotating black hole by Shibata et al. [20]. They have obtained the energy and angular momentum luminosities to 2.5PN order from a particle in circular orbits with small inclination angle, which hence clarified the next leading order effects of spin-orbit coupling. For slightly eccentric orbits around a Kerr black hole, Tagoshi has done the calculation to 2.5PN order [21]. For circular orbits around a Kerr black hole, the calculation to 4PN order has been done by Tagoshi et al. [22]. Furthermore, extention of the method to the case of a spinning particle has been done by Tanaka et al. [23] and the luminosity to 2.5PN order has been obtained for circular orbits, which includes the effect of spin-spin coupling.

In the next subsection, I review the black-hole perturbation approach based on the Teukolsky [24] and Regge-Wheeler equations [25]. (Due to limitation of space, however, I focus on the case of a non-rotating black hole in this talk.) Then in the proceeding subsections, I summarize these recent results and discuss future issues. For notational simplicity, we set $c = G = 1$ in the following.

B. Regge-Wheeler-Teukolsky formalism

Let us consider the case when a particle of mass μ is in a circular orbit around a Schwarzschild black hole of mass $M \gg \mu$. The gravitational waves radiated out to infinity from the system is then described by the fourth Newman-Penrose quantity ψ_4 [24], which is related to the two independent modes of gravitational waves h_+ and $h_\times$ at infinity as

$$\psi_4 = \frac{1}{2}(\ddot{h}_+ - i\ddot{h}_\times). \tag{2.1}$$

On the Schwarzschild background, it can be decomposed as

$$\psi_4 = \frac{1}{r^4}\sum_{\ell m\omega} R_{\ell m\omega}(r)\,{}_{-2}Y_{\ell m}(\theta,\varphi)e^{-i\omega t}, \tag{2.2}$$

where ${}_{-2}Y_{\ell m}$ are the $s=-2$ spin-weighted spherical harmonics. The radial function $R_{\ell m\omega}$ satisfies the inhomogeneous Teukolsky equation [24],

$$\left[\Delta^2\frac{d}{dr}\left(\frac{1}{\Delta}\frac{d}{dr}\right) - U(r)\right] R_{\ell m\omega}(r) = T_{\ell m\omega}(r), \tag{2.3}$$

where

$$U(r) = \frac{r^2}{\Delta}[\omega^2 r^2 - 4i\omega(r-3M)] - (\ell-1)(\ell+2), \quad \Delta = r(r-2M), \tag{2.4}$$

and $T_{\ell m\omega}$ is the source term determined by the energy momentum tensor of the particle.

To solve Eq.(2.3), we employ the Green function method. Then $R_{\ell m\omega}$ at $r\to\infty$ takes the form,

$$\begin{aligned} R_{\ell m\omega}(r\to\infty) &= \frac{r^3 e^{i\omega r^*}}{2i\omega B^{in}_{\ell\omega}}\int_{2M}^{\infty} dr R^{in}_{\ell\omega}(r)T_{\ell m\omega}(r)\Delta^{-2} \\ &\equiv r^3 e^{i\omega r^*}\tilde{Z}_{\ell m\omega}, \end{aligned} \tag{2.5}$$

where $R^{in}_{\ell\omega}(r)$ is a homogeneous solution which satisfies the boundary condition,

$$R^{in}_{\ell\omega}(r) = \begin{cases} D_{\ell\omega}\Delta^2 e^{-i\omega r^*} & \text{for } r^*\to-\infty\,, \\ r^3 B^{out}_{\ell\omega}e^{i\omega r^*} + r^{-1}B^{in}_{\ell\omega}e^{-i\omega r^*} & \text{for } r^*\to+\infty\,, \end{cases} \tag{2.6}$$

where $r^* = r + 2M\ln(r/2M-1)$. In the case of a circular orbit with radius $r=r_0$, $T_{\ell m\omega}(r)$ takes the form,

$$T_{\ell m\omega} \sim [a_0\delta(r-r_0) + a_1\delta'(r-r_0) + a_2\delta''(r-r_0)]\,\delta(\omega-m\Omega), \tag{2.7}$$

where Ω is the orbital angular frequency. Hence what we need to know are the behavior of $R_{\ell m\omega}$ around $r=r_0$ and its incident amplitude $B^{in}_{\ell\omega}$. Also because of Eq.(2.7), the amplitude $\tilde{Z}_{\ell m\omega}$ takes the form,

$$\tilde{Z}_{\ell m\omega} = Z_{\ell m}\delta(\omega-m\Omega). \tag{2.8}$$

In terms of $Z_{\ell m}$, the gravitational wave form at infinity is given by

$$h_+ - ih_\times = -\frac{2}{r}\sum_{\ell m}\frac{1}{\omega^2}Z_{\ell m}\,{}_{-2}Y_{\ell m}(\theta,\varphi)e^{-i\omega(t-r^*)}, \tag{2.9}$$

and the luminosity is given by

$$\frac{dE}{dt} = \sum_{\ell=2}^{\infty}\sum_{m=1}^{\ell} \mid Z_{\ell m} \mid^2 /2\pi\omega^2, \tag{2.10}$$

where $\omega = m\Omega$.

Now since the radius of the orbit and the angular velocity are related as

$$\frac{M}{r_0} = (r_0\Omega)^2 \equiv v^2, \tag{2.11}$$

where v is the orbital velocity, we have the small non-dimensional parameters of the problem, $r_0\omega = O(v)$ and $M\omega = O(v^3)$. Originally the parameter $r_0\omega$ represents the slowness of the particle motion, hence plays a role of the post-Newtonian expansion parameter. The parameter $M\omega$ represents the strength of the gravity, hence plays a role of the post-Minkowski parameter. In the present case, since the particle is in bound orbits, these parameters are related to each other as shown above through the orbital velocity of the particle.

Thus our task reduces to calculating the ingoing-wave Teukolsky function $R^{\rm in}_{\ell\omega}$ at $r\omega = O(v) \ll 1$ as well as to extracting out its incident amplitude $B^{\rm in}_{\ell m\omega}$ to a required order of $M\omega = O(v^3)$.

For technical reasons, however, it is much easier to deal with the Regge-Wheeler equation, rather than the Teukolsky equation, by the transformation [26,27],

$$R^{\rm in}_{\ell\omega} \;\rightarrow\; R^{\rm in}_{\ell\omega} = \Delta\left(\frac{d}{dr^*} + i\omega\right)\frac{r^2}{\Delta}\left(\frac{d}{dr^*} + i\omega\right) rX_{\ell\omega}. \tag{2.12}$$

Then $X_{\ell\omega}$ satisfies the homogeneous Regge-Wheeler equation [25],

$$\left[\frac{d^2}{dr^{*2}} + \omega^2 - V_\ell(r)\right] X_{\ell\omega}(r) = 0, \tag{2.13}$$

where

$$V_\ell(r) = \left(1 - \frac{2M}{r}\right)\left(\frac{\ell(\ell+1)}{r^2} - \frac{6M}{r^3}\right). \tag{2.14}$$

Corresponding to Eq.(2.6), we have the asymptotic forms of $X^{in}_{\ell\omega}$ as

$$X^{in}_{\ell\omega}(r) = \begin{cases} C_{\ell\omega} e^{-i\omega r^*}, & r^* \to -\infty \\ A^{out}_{\ell\omega} e^{i\omega r^*} + A^{in}_{\ell\omega} e^{-i\omega r^*}, & r^* \to +\infty. \end{cases} \tag{2.15}$$

where $A^{in}_{\ell\omega}$, $A^{out}_{\ell\omega}$ and $C_{\ell\omega}$ are respectively related to $B^{in}_{\ell\omega}$, $B^{out}_{\ell\omega}$ and $D_{\ell\omega}$ defined in Eq.(2.6) as

$$\begin{aligned} B^{in}_{\ell\omega} &= -\frac{c_0}{4\omega^2} A^{in}_{\ell\omega}, \\ B^{out}_{\ell\omega} &= -4\omega^2 A^{out}_{\ell\omega}, \\ D_{\ell\omega} &= \frac{c_0}{16(1-2iM\omega)(1-4iM\omega)M^3} C_{\ell\omega}, \end{aligned} \tag{2.16}$$

where $c_0 = (\ell-1)\ell(\ell+1)(\ell+2) - 12iM\omega$.

To make the structure of the equation more transparent, we rescale the independent variable r to $z = r\omega$ and rewrite the homogeneous Regge-Wheeler equation as

$$\left[\frac{d^2}{dz^{*2}} + 1 - \left(1 - \frac{\epsilon}{z}\right)\left(\frac{\ell(\ell+1)}{z^2} - \frac{3\epsilon}{z^3}\right)\right] X^{in}_{\ell} = 0, \tag{2.17}$$

where $z^* = z + \epsilon \ln(z-\epsilon) = r^*\omega + \epsilon \ln \epsilon$ with $\epsilon \equiv 2M\omega$, and we have suppressed the index ω since it is trivially absorbed in ϵ and z. As noted previously, the post-Newtonian expansion corresponds to expanding X^{in}_{ℓ} with respect to ϵ and evaluating X^{in}_{ℓ} at $z \ll 1$ as well as A^{in}_{ℓ} to required orders in ϵ. This procedure to the first order of ϵ was carried out by Poisson [17]. However, as can be seen from the structure of Eq.(2.17), the equation would become very complicated if we go beyond the first order in this procedure. Further, it becomes rather unclear how to impose the correct boundary condition that $X^{in}_{\ell} \propto e^{-iz^*}$ at horizon ($z \to \epsilon$) if we expand the equation naively in powers of ϵ, since z^* involves ϵ in itself.

These difficulties were then resolved by Sasaki [18] by the following choice of the dependent variable. Setting

$$X^{in}_{\ell} = e^{-i\epsilon \ln(z-\epsilon)} z \xi_{\ell}(z), \tag{2.18}$$

we find that Eq.(2.17) becomes

$$\left[\frac{d}{dz^2} + \frac{2}{z}\frac{d}{dz} + \left(1 - \frac{\ell(\ell+1)}{z^2}\right)\right] \xi_{\ell} = \epsilon e^{-iz} \frac{d}{dz}\left[\frac{1}{z^3}\frac{d}{dz}\left(e^{iz} z^2 \xi_{\ell}(z)\right)\right]. \tag{2.19}$$

Since ϵ appears only as an overall factor on the r.h.s., this form of the equation is most suited for an iterative treatment. Thus expanding ξ_ℓ with respect to ϵ as

$$\xi_\ell(z) = \sum_{n=0}^{\infty} \epsilon^n \xi_\ell^{(n)}(z), \tag{2.20}$$

we obtain the recursive equations,

$$\left[\frac{d}{dz^2} + \frac{2}{z}\frac{d}{dz} + \left(1 - \frac{\ell(\ell+1)}{z^2}\right)\right] \xi_\ell^{(n)}(z) = \epsilon e^{-iz} \frac{d}{dz}\left[\frac{1}{z^3}\frac{d}{dz}\left(e^{iz} z^2 \xi_\ell^{(n-1)}(z)\right)\right], \tag{2.21}$$

with the lowest order solution given by

$$\xi_\ell^{(0)} = \alpha^{(0)} j_\ell + \beta^{(0)} n_\ell, \tag{2.22}$$

where j_ℓ and n_ℓ are the usual spherical Bessel functions. The boundary condition is that $\xi_\ell^{(0)}$ be regular at $z = 0$ [17,18]. Hence $\beta^{(0)} = 0$ and for convenience we set $\alpha_\ell^{(0)} = 1$.

Putting the Regge-Wheeler equation in the above form helped us a lot to calculate the first and second order solutions by iteration [18]. Using a similar transformation of the Teukolsky equation [28], one obtains similar iterative equations for a Kerr black hole and the first and second order solutions are expressed also in closed analytical form [20], which are sufficient to calculate the gravitational wave luminosity to 4PN order.

However, after the present workshop, we were informed that Mano and Takasugi [29] have succeeded in developing a much better systematic method to obtain the solution by using the Coulomb wave functions and the hypergeometric functions. In their method, calculations to a very high order, including the effect of black hole absorption, are straightforward and surprisingly simple. Since this makes the original method absurd, we will not go into the details of it, but just present the results.

C. Gravitational Wave Luminosity up through 4PN Order

Here we only show the final result of the luminosity to 4PN order for circular orbits around a Kerr black hole by summarizing the results obtained in [19,20,22,23]. The orbits are assumed to be on the equatorial plane with the angular frequency Ω. The black hole

has mass M and the angular momentum J_{BH}. For the case of spinning particle, the z-component of the spin angular momentum is j_z, and the spin-dependent part is calculated only to 2.5PN order in the first order in the magnitude of spin [23]. Further in this case, we note the orbit precesses slightly with inclination proportional to the radial component of the spin, but it does not affect the luminosity. The result is

$$\begin{aligned}\left\langle\frac{dE}{dt}\right\rangle = \left(\frac{dE}{dt}\right)_N \Bigg\{ & 1 - \frac{1247}{336} v'^2 + \left(4\pi - \frac{11}{4} q - \frac{5}{4}\hat{s}\right) v'^3 \\ & + \left(-\frac{44711}{9072} + \frac{33}{16} q^2 + \frac{31}{8} q\hat{s}\right) v'^4 + \left(\frac{-8191\,\pi}{672} - \frac{59}{16} q - \frac{13}{16}\hat{s}\right) v'^5 \\ & + \Bigg(\frac{6643739519}{69854400} - \frac{1712\,\gamma}{105} + \frac{16\,\pi^2}{3} - \frac{3424\,\ln 2}{105} \\ & \quad - \frac{1712}{105}\ln v' - \frac{65\,\pi}{6} q + \frac{611}{504} q^2\Bigg) v'^6 \\ & + \left(\frac{-16285\,\pi}{504} + \frac{162035}{3888} q + \frac{65\,\pi}{8} q^2 - \frac{71}{24} q^3\right) v'^7 \\ & + \Bigg(-\frac{323105549467}{3178375200} + \frac{232597\,\gamma}{4410} - \frac{1369\,\pi^2}{126} + \frac{39931\,\ln 2}{294} - \frac{47385\,\ln 3}{1568} \\ & \quad + \frac{232597}{4410}\ln v' - \frac{359\,\pi}{14} q + \frac{22667}{4536} q^2 + \frac{17}{16} q^4\Bigg) v'^8 \Bigg\}, \end{aligned} \tag{2.23}$$

where

$$\left(\frac{dE}{dt}\right)_N := \frac{32}{5}\left(\frac{\mu}{M}\right)^2 v'^{10}, \tag{2.24}$$

and

$$v' := (M\Omega)^{1/3}, \qquad q := \frac{J_{BH}}{M^2}, \qquad \hat{s} := \frac{j_z}{\mu M}. \tag{2.25}$$

D. Future Issues

Here I discuss several issues of the Regge-Wheeler-Teukolsky approach to be solved in order to make this approach more realistic and physically more fruitful.

We have succeeded in obtaining the luminosity to 4PN order in the post-Newtonian expansion by the Regge-Wheeler-Teukolsky approach. In the first place, it is of interest to proceed the calculation to a much higher order, including the effect of black hole absorption.

In one respect, it will clarify the convergence property of the post-Newtonian approximation much better than what we know now and will give us more insight into the problem of matching the post-Newtonian spacetime with the fully relativistic spacetime, which is necessarily solved in order to give reasonable initial data of a coalescing compact binary to be used in numerical relativity. In another respect, it will give us detailed information of how the black hole's spacetime geometry affects the luminosity and waveforms of the gravitational radiation emitted by a small mass particle, hence will helps us to understand necessary ingredients to reconstruct the geometry out of the information contained in the waves [30].

However our approach is limited by the condition that one of the binary stars has a much smaller mass than the other, i.e., $\mu \ll M$. In order to make this approach more fruitful, it is then necessary to extend it so that it can take account of finite μ/M effects. Concerning this point, it is important to include the effect of radiation reaction in this approach. At the moment, our approach relies on the adiabatic approximation of the reaction. But even if we accept this approximation, there remains one important problem; that is the backreaction to the Carter constant. In the case of a non-rotating black hole, we may assume the orbit to be in the equatorial plane without loss of generality. Then the energy and the z-component of the angular momentum completely determines the orbit. Hence knowing the energy and angular momentum fluxes averaged over several orbital periods is enough to solve the adiabatic evolution of the orbit. But this will not be the case once we consider a rotating black hole, since the inclination of the orbit away from the equatorial plane becomes meaningful, which is represented by a non-zero value of the Carter constant. The problem is that the Carter constant is not associated with a Killing vector of the geometry, hence no field quantity representing the Carter constant has been known. Thus one cannot calculate the backreaction to the Carter constant by just evaluating the gravitational waves at infinity. In this sense, a simple adiabatic approximation scheme seems to break down anyway as soon as we take the rotation into account. Furthermore, if we take into account spin of the small mass particle, the situation gets worse; the equation of motion is not integrable in this case

from the beginning. In order to clarify the validity limit of the adiabatic approximation, as well as to find out the backreaction to the Carter constant, we need to formulate a method in which the non-adiabatic backreaction can be taken into account. In other words, we need to derive a backreaction force term in the equation of motion for the particle. This will be analogous to the post-Newtonian formulation which includes the back-reaction potential to a required order of accuracy.

Whether such a backreaction force term can be obtained for general orbits is very non-trivial. However, at least for a Kerr black hole, we know from the classic work by Chrzanowski [31] that the tensor Green function can be obtained in a separable form. Hence we may calculate the perturbed metric at any spacetime point. Then once we obtain a reliable method of regularizing the self-energy, we will be able to derive the backreaction force term. In this connection, for quasi-periodic orbits, it may be possible to justify the earlier work of Gal'tsov [32], in which the backreaction force term is assumed to be derivable from the radiative Green function, i.e., the retarded minus advanced Green functions. Recently, Ori has proposed another way of deriving the reaction force [33]. He has shown that the reaction force calculated solely in terms of the retarded Green function but averaged over many periods gives the same result as that of Gal'tsov as long as the energy and angular momentum are concerned. Finally, DeWitt-Brehme's covariant approach [4,5] suggests that one should use the radiative Green function but with the integral domain bounded to the past lightcone of the world point of the particle, see Eq.(1.6). However, It is not clear at all at present which approach gives the correct answer for the Carter constant, if ever.

Another important but difficult problems is to take into account the effect non-linear in μ/M in the equation of motion. In the post-Newtonian approach, we know this effect appears already in 2PN order in the luminosity formula [12], well before the radiation reaction comes into play. To my knowledge, however, no serious attempt has been made in this direction so far.

III. SUMMARY

We will have a new era of gravitational physics and astronomy in the next century. Gravitational waves will be detected, the data will be taken by detectors at various places in the world, and will be carefully analyzed through the international networks of the detectors. In order to optimize the use of such data, it is essential that we understand physics of gravitational waves as deep as possible. In the light of this present situation, I first reviewed some of unsolved conceptual as well as technical problems in gravitational wave physics. Then I reported on some recent progress made in the black-hole perturbation approach. It is hoped that this approach will fill some gaps in our knowledge and lead us to a better understanding of gravitational waves.

ACKNOWLEDGEMENTS

I would like to thank H. Asada, Y. Mino, T. Nakamura, E. Poisson, M. Shibata, T. Tagoshi and T. Tanaka for fruitful discussions and collaborations. This work was supported in part by the Monbu-sho Grant-in-Aid Scientific Research for Priority Area, No.04234104.

REFERENCES

[1] , R.A. Isaacson, Phys. Rev. **166**, 1263 (1968); **166**, 1272 (1968).

[2] C.W. Misner, K.S. Thorne and J.A. Wheeler, *Gravitation*, (Freeman, San Francisco, 1973) Chap. 35-36.

[3] K.S. Thorne, Rev. Mod. Phys. **52**, 299 (1980).

[4] Y. Mino, Master thesis (in Japanese) (1995).

[5] Y. Mino, M. Sasaki and T. Tanaka, in preparation (1996).

[6] S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Clarendon, Oxford, 1983).

[7] A. Abramovici et al., Science, **256**, 325 (1992).

[8] K.S. Thorne, in *Proceedings of the Eighth Nishinomiya-Yukawa Memorial Symposium: Relativistic Cosmology*, ed. M. Sasaki (Universal Academy Press, Tokyo, 1994), p.67, and references therein.

[9] C. Bradaschia et al., Nucl. Instrum. & Methods, **A289**, 518 (1990).

[10] C. Cutler et al. Phys. Rev. Lett. **70**, 2984 (1993).

[11] E.S. Phinney, Astrophys. J. **380**, L17 (1991).

[12] C.M. Will, in *Proceedings of the Eighth Nishinomiya-Yukawa Memorial Symposium: Relativistic Cosmology*, ed. M. Sasaki (Universal Academy Press, Tokyo, 1994), p.83, and references therein.

[13] L. Blanchet, T. Damour, B.R. Iyer, C.M. Will and A. G. Wiseman, Phys. Rev. Lett. **74**, 3515(1995).

[14] L. Blanchet, *Gravitational-radiation energy losses in coalescing compact binaries to five halves post-Newtonian order*, Observatoire de Meudon preprint (1995).

[15] C. Cutler, L.S. Finn, E. Poisson and G.J. Sussman, Phys. Rev. D**47**, 1511 (1993).

[16] H. Tagoshi and T. Nakamura, Phys. Rev. **D49**, 4016 (1994).

[17] E. Poisson, Phys. Rev. **D47**, 1497 (1993).

[18] M. Sasaki, Prog. Theor. Phys. **92**, 17 (1994).

[19] H. Tagoshi and M. Sasaki, Prog. Theor. Phys. **92**, 745(1994).

[20] M. Shibata, M. Sasaki, H. Tagoshi and T. Tanaka, Phys. Rev. **D51**, 1646(1994).

[21] H. Tagoshi, Prog. Theor. Phys. **93**, 307 (1995).

[22] H. Tagoshi, M. Shibata, T. Tanaka, and M. Sasaki, *Post-Newtonian expansion of gravitational waves from a particle in circular orbits around a rotating black hole: Up to $O(v^8)$ beyond the quadrupole formula*, Caltech preprint GRP-434/Osaka University preprint OU-TAP-28 (1996).

[23] T. Tanaka, Y. Mino, M. Sasaki and M. Shibata, *Gravitational waves from a spinning particle in circular orbits around a rotating black hole*, Osaka University preprint OU-TAP-27 (1996).

[24] S. A. Teukolsky, Astrophys. J. **185** (1973) 635.

[25] T. Regge and J. A. Wheeler, Phys. Rev. **108**, 1063 (1957).

[26] S. Chandrasekhar, Proc. R. Soc. London **A343**, 289 (1975).

[27] M. Sasaki and T. Nakamura, Phys. Lett. **87A**, 85 (1981).

[28] M. Sasaki and T. Nakamura, Prog. Theor. Phys. **67**, 1788 (1982);
M. Sasaki and T. Nakamura, Phys. Lett. **89A**, 68 (1982).

[29] S. Mano and E. Takasugi, in preparation (1996).

[30] F.D. Ryan, L.S. Finn and K.S. Thorne, *Testing the Black-Hole No-Hair Theorem by Gravitational-Wave Measurements*, Phys. Rev. Lett. in preparation (1996);
F. Ryan, *Gravitational waves from the inspiral of a compact object into a massive,*

axisymmetric body with arbitrary multipole moments, Caltech preprint (1995).

[31] P.L. Chrzanowski, Phys. Rev. **D11**, 2042 (1975).

[32] D. Gal'tsov, J. Phys. **A15**, 3737 (1982).

[33] A. Ori, Phys. Lett. **A202**, 347 (1995); *Radiative evolution of the Carter constant for generic orbits around a Kerr black hole*, preprint (1995).

Gravitational waves by a particle in circular orbits around a Schwarzshild black hole

Takahiro Tanaka, Hideyuki Tagoshi[1] and Misao Sasaki

Department of Earth and Space Science, Osaka University, Toyonaka 560, Japan

[1] *California Institute of Technology, Theoretical Astrophysics, Pasadena, CA 91125, and National Astronomical Observatory, Mitaka, Tokyo 181, Japan*

Using the post-Newtonian expansion technique of the gravitational wave perturbation around a Schwarzshild black hole, we calculate the energy flux of gravitational waves induced by a particle moving in circular orbits. We calculate the energy flux emitted as gravitational waves up to v^{11} order beyond Newtonian, where $v = (M/r_0)^{1/2}$

I. INTRODUCTION

As was reported in the contribution by Sasaki in this proceedings, gravitational waves from the coalescing compact binaries are the most promising candidates which will be able to be detected by the near-future, ground based laser interferometric gravitational wave detectors such as LIGO and VIRGO.

If a neutron star or a small black hole spirals into a massive black hole with mass $< 300M_\odot$, the inspiral wave form will be detected by above detectors. When a signal of gravitational waves is detected, we will try to extract parameters of binaries, such as masses and spins etc., from inspiral wave forms by matched filtering technique [1]. In this method, parameters of binaries are determined by cross-correlating the noisy signal from detectors with theoretical templates. If the signal and the templates lose phase with each other by one cycle over $\sim 10^3 - 10^4$ cycles as the waves sweep through the LIGO/VIRGO band, their cross correlation will be significantly reduced. This means that, in order to extract the information optimally, we need to make theoretical templates which are accurate to better than one cycle during entire sweep through the LIGO/VIRGO band [1].

To calculate inspiraling wave forms from coalescing binaries, the standard method is the post-Newtonian expansion of the Einstein equations, in which the orbital velocity v of binaries is assumed to be small compared to the speed of light. Although the post-Newtonian calculation technique will be developed to apply to the higher order calculation, it will become more and more difficult and complicated. Thus, it would be very helpful if we could have another reliable method to calculate the higher order post-Newtonian corrections.

As an alternative method, the post-Newtonian expansion of the black hole perturbation was developed as was reported in Sasaki's contribution. There one considers gravitational waves from a particle of mass μ orbiting a black hole of mass M assuming $\mu \ll M$. Although this method is restricted to the case of $\mu \ll M$, we can calculate very high order post-Newtonian corrections of gravitational waves by means of relatively simple analysis compared to the standard post-Newtonian analysis.

Since LIGO and VIRGO will be able to detect a signal of gravitational waves from binaries of mass less than $\sim 300 M_\odot$, it is important to make a template for such a binary. The frequency of gravitational waves for such a massive binary, however, comes into the frequency band for LIGO and VIRGO at $r/M \sim 16(100 M_\odot/M)^{2/3}$, i.e., highly relativistic region. We do not know whether the convergence property of the post-Newtonian approximation is good or not in such a highly relativistic motion. As was reported by Tagoshi and Sasaki [2], the post-Newtonian convergence of the total orbital phase during the detectable frequency band for LIGO and VIRGO can be very slow for these binaries. Hence, it is an urgent problem to clarify up to what point the convergence property of the post-Newtonian expansion is good. For this purpose, we study the energy loss rate of the binary up to v^{11} order in this paper.

The present small contribution is organized as follows. In section 2, we show the general formulas and conventions used in this paper very briefly. The full description of the formalism will be found in Tagoshi and Sasaki [2] and references therein. In section 3, we show the expressions of the energy flux extended to $O(v^{11})$. In section 4, we briefly discuss the implication of our results.

Throughout this paper we use the units of $c = G = 1$.

II. GENERAL FORMULATION

For the reader's convenience, we attach a minimum amount of the explanation about the black hole perturbation formalism in order to make the convention definite. For the reader who wants to know the details, see the reference [2]. (Some typographical errors in the formulas in this reference are corrected here.) We consider the case when a particle of small mass μ travels a circular orbit around a Schwarzschild black hole of mass $M \ll \mu$.

To calculate the gravitational luminosity, we consider the inhomogeneous Teukolsky equation,

$$\left[\Delta^2 \frac{d}{dr}\left(\frac{1}{\Delta}\frac{d}{dr}\right) - U(r)\right] R_{\ell m\omega}(r) = T_{\ell m\omega}(r), \tag{2.1}$$

where

$$U(r) = \frac{r^2}{\Delta}\left[\omega^2 r^2 - 4i\omega(r-3M)\right] - (\ell-1)(\ell+1), \quad \Delta = r(r-2M), \tag{2.2}$$

and $T_{\ell m\omega}$ is the source term which reflects the energy momentum tensor of the small particle. We omit the explicit form of $T_{\ell m\omega}(r)$ here.

We solve Eq. (2.1) by the Green function method. For this purpose, we need a homogeneous solution $R^{in}_{\ell\omega}$ of Eq. (2.1) which satisfies the following boundary condition,

$$R^{in}_{\ell\omega} = \begin{cases} D_{\ell\omega}\Delta^2 e^{-i\omega r^*} & \text{for } r^* \to -\infty, \\ r^3 B^{out}_{\ell\omega}\Delta^2 e^{i\omega r^*} + r^{-1}B^{in}_{\ell\omega}\Delta^2 e^{-i\omega r^*} & \text{for } r^* \to +\infty, \end{cases} \tag{2.3}$$

where $r^* = r + 2M\ln(r/2M - 1)$. Then the outgoing-wave solution of Eq. (2.1) at infinity with the appropriate boundary condition at horizon is given by

$$\begin{aligned} R_{\ell m\omega}(r \to \infty) &= \frac{r^3 e^{i\omega r^*}}{2i\omega B^{in}_{\ell\omega}} \int_{2M}^{\infty} dr R^{in}_{\ell\omega} T_{\ell m\omega}(r)\Delta^{-2} \\ &=: r^3 e^{i\omega r^*}\tilde{Z}_{\ell m\omega}. \end{aligned} \tag{2.4}$$

In the case of a circular orbit, the specific energy $\tilde{E}$ and angular momentum $\tilde{L}$ of the particle are given by

$$\tilde{E} = (r_0 - 2M)/\sqrt{r_0(r_0 - 3M)}, \tag{2.5}$$

and

$$\tilde{L} = \sqrt{Mr_0}/\sqrt{1 - 3M/r_0}, \tag{2.6}$$

where r_0 is the orbital radius. The angular frequency is given by $\Omega = (M/r_0^3)^{1/2}$. Defining ${}_sb_{\ell m}$ by

$$\begin{aligned} {}_0b_{\ell m} &= \frac{1}{2}\left[(\ell-1)\ell(\ell+1)(\ell+2)\right]^{1/2} {}_0Y_{\ell m}\left(\frac{\pi}{2}, 0\right)\tilde{E}r_0/(r_0 - 2M), \\ {}_{-1}b_{\ell m} &= \left[(\ell-1)(\ell+2)\right]^{1/2} {}_{-1}Y_{\ell m}\left(\frac{\pi}{2}, 0\right)\tilde{L}/r_0, \\ {}_{-2}b_{\ell m} &= {}_{-2}Y_{\ell m}\left(\frac{\pi}{2}, 0\right)\tilde{L}\Omega, \end{aligned} \tag{2.7}$$

where ${}_nY_{\ell m}(\theta, \varphi)$ are the spin-weighted spherical harmonics, $\tilde{Z}_{\ell m\omega}$ is found to take the form,

$$\tilde{Z}_{\ell m\omega} = Z_{\ell m}\delta(\omega - m\Omega), \tag{2.8}$$

where

$$\begin{aligned} Z_{\ell m} = \frac{\pi}{i\omega r_0^2 B^{in}_{\ell\omega}}\Bigg\{ &\Bigg[-{}_0b_{\ell m} - 2i\,{}_{-1}b_{\ell m}\left(1 + \frac{i}{2}\omega r_0^2/(r_0 - 2M)\right) \\ &+ i\,{}_{-2}b_{\ell m}\omega r_0(1 - 2M/r_0)^{(-2)}\left(1 - M/r_0 + \frac{1}{2}i\omega r_0\right)\Bigg] R^{in}_{\ell m} \\ &+ \left[i\,{}_{-1}b_{\ell m} - {}_{-2}b_{\ell m}\left(1 + i\omega r_0^2/(r_0 - 2M)\right)\right] r_0 R^{in\prime}_{\ell\omega}(r_0) + \frac{1}{2}\,{}_{-2}b_{\ell m} r_0^2 R^{in\prime\prime}_{\omega}(r_0)\Bigg\} \quad . \end{aligned} \tag{2.9}$$

In terms of the amplitudes $Z_{\ell m}$, the gravitational wave luminosity is given by

$$\frac{dE}{dt} = \sum_{\ell=2}^{\infty} \sum_{m=\ell}^{\ell} |Z_{\ell m}|^2 / 2\pi\omega^2, \tag{2.10}$$

where $\omega = m\Omega$. Thus the only remaining task is to calculate the series expansion of the ingoing-wave Teukolsky function $R^{in}_{\ell\omega}$ in terms of r and the Wronskian $B^{in}_{\ell m}$.

In the next section we give the explicit expressions of these quantities.

III. RESULTS

As noted in the Sasaki's contribution, a new method to calculate the Teukolsky function in post-Newtonian expansion was found by Suzuki, Mano and Takasugi [3]. We have derived the following results by using the old formalism developed by Sasaki [4] but the details of derivation is too complicated and seems less valuable to be noted here after a new powerful method was developed. Thus here will be the final results only.

The post-Newtonian expansion of the Teukolsky function in the near zone, where $z := \omega r$ is small, is given by

$$\begin{aligned}
R^{in}_{2\omega} = {} & \left(\frac{4z^4}{5} + \frac{8i}{15} z^5 - \frac{22z^6}{105} - \frac{2i}{35} z^7 + \frac{23z^8}{1890} + \frac{2i}{945} z^9 - \frac{13z^{10}}{41580} \right. \\
& \left. - \frac{i}{24948} z^{11} + \frac{59z^{12}}{12972960} + \frac{i}{2162160} z^{13} - \frac{83z^{14}}{1945944000} - \frac{i}{277992000} z^{15} \right) \\
& + \left(\frac{-8z^3}{5} - \frac{3i}{5} z^4 - \frac{8z^5}{63} - \frac{13i}{90} z^6 + \frac{109z^7}{1890} + \frac{341i}{22680} z^8 \right. \\
& \left. - \frac{9403z^9}{3118500} - \frac{293i}{594000} z^{10} + \frac{38963z^{11}}{567567000} + \frac{75529i}{9081072000} z^{12} \right) \epsilon \\
& + \left(\frac{4z^2}{5} + \frac{123317z^4}{36750} + \frac{231479i}{110250} z^5 - \frac{889954z^6}{1157625} - \frac{454499i}{2315250} z^7 \right. \\
& + \frac{215321483z^8}{5501034000} + \frac{35106811i}{5501034000} z^9 - \frac{214z^4 \log z}{525} \\
& \left. - \frac{428i}{1575} z^5 \log z + \frac{1177z^6 \log z}{11025} + \frac{107i}{3675} z^7 \log z - \frac{2461z^8 \log z}{396900} - \frac{107i}{99225} z^9 \log z \right) \epsilon^2 \\
& + \left(\frac{-66823z^3}{12250} - \frac{99851i}{55125} z^4 - \frac{504569z^5}{694575} - \frac{2488639i}{3969000} z^6 + \frac{428z^3 \log z}{525} \right. \\
& \left. + \frac{107i}{350} z^4 \log z + \frac{428z^5 \log z}{6615} + \frac{1391i}{18900} z^6 \log z \right) \epsilon^3 \\
& + \left(\frac{471487z^2}{220500} - \frac{263i}{1260} z^3 - \frac{214z^2 \log z}{525} \right) \epsilon^4, \\
R^{in}_{3\omega} = {} & \left(\frac{4z^5}{21} + \frac{2i}{21} z^6 - \frac{2z^7}{63} - \frac{i}{135} z^8 + \frac{29z^9}{20790} + \frac{i}{4620} z^{10} - \frac{47z^{11}}{1621620} \right. \\
& \left. - \frac{i}{294840} z^{12} + \frac{23z^{13}}{64864800} + \frac{i}{29937600} z^{14} \right)
\end{aligned}$$

$$
\begin{aligned}
&+\left(\frac{-10\,z^4}{21}-\frac{53\,i}{315}\,z^5+\frac{z^6}{210}-\frac{i}{90}\,z^7+\frac{751\,z^8}{155925}+\frac{1483\,i}{1247400}\,z^9\right.\\
&\left.-\frac{23\,z^{10}}{102960}-\frac{367\,i}{10810800}\,z^{11}\right)\epsilon\\
&+\left(\frac{8\,z^3}{21}+\frac{i}{14}\,z^4+\frac{40337\,z^5}{46305}+\frac{79099\,i}{185220}\,z^6-\frac{12562147\,z^7}{91683900}-\frac{4840537\,i}{157172400}\,z^8-\right.\\
&\left.\frac{26\,z^5\log z}{441}-\frac{13\,i}{441}\,z^6\log z+\frac{13\,z^7\log z}{1323}+\frac{13\,i}{5670}\,z^8\log z\right)\epsilon^2\\
&+\left(\frac{-2\,z^2}{21}-\frac{182981\,z^4}{92610}-\frac{3753697\,i}{5556600}\,z^5+\frac{65\,z^4\log z}{441}+\frac{689\,i}{13230}\,z^5\log z\right)\epsilon^3,
\end{aligned}
$$

$$
\begin{aligned}
R^{in}_{4\omega}=&\left(\frac{2\,z^6}{63}+\frac{4\,i}{315}\,z^7-\frac{13\,z^8}{3465}-\frac{8\,i}{10395}\,z^9\right.\\
&\left.+\frac{71\,z^{10}}{540540}+\frac{i}{54054}\,z^{11}-\frac{37\,z^{12}}{16216200}-\frac{i}{4054050}\,z^{13}\right)\\
&+\left(\frac{-2\,z^5}{21}-\frac{4\,i}{135}\,z^6+\frac{142\,z^7}{51975}-\frac{31\,i}{51975}\,z^8+\frac{929\,z^9}{2702700}+\frac{8\,i}{96525}\,z^{10}\right)\epsilon\\
&+\left(\frac{5\,z^4}{49}+\frac{97\,i}{4410}\,z^5+\frac{958223891\,z^6}{6051137400}+\frac{239560304\,i}{3781960875}\,z^7\right.\\
&\left.-\frac{1571\,z^6\log z}{218295}-\frac{3142\,i}{1091475}\,z^7\log z\right)\epsilon^2\\
&+\left(\frac{-20\,z^3}{441}-\frac{i}{196}\,z^4\right)\epsilon^3,\\
R^{in}_{5\omega}=&\left(\frac{2\,z^7}{495}+\frac{2\,i}{1485}\,z^8-\frac{7\,z^9}{19305}-\frac{i}{15015}\,z^{10}+\frac{17\,z^{11}}{1621620}+\frac{i}{737100}\,z^{12}\right)\\
&+\left(\frac{-7\,z^6}{495}-\frac{67\,i}{17325}\,z^7+\frac{1831\,z^8}{4054050}-\frac{43\,i}{4054050}\,z^9\right)\epsilon\\
&+\left(\frac{28\,z^5}{1485}+\frac{59\,i}{14850}\,z^6\right)\epsilon^2,\\
R^{in}_{6\omega}=&\left(\frac{8\,z^8}{19305}+\frac{16\,i}{135135}\,z^9-\frac{4\,z^{10}}{135135}-\frac{2\,i}{405405}\,z^{11}\right)\\
&+\left(\frac{-32\,z^7}{19305}-\frac{163\,i}{405405}\,z^8\right)\epsilon,\\
R^{in}_{7\omega}=&\frac{8\,z^9}{225225}+\frac{2\,i}{225225}\,z^{10},
\end{aligned}
\tag{3.1}
$$

and the Wronskian for $\ell=2$, and 3 is given by

$$
\begin{aligned}
B^{in}_{2\omega}=&\frac{i}{\omega^2}e^{-i\epsilon(\log 2\epsilon+\gamma)}\left(3-\frac{3}{4}i\epsilon\right)\left[1+\epsilon\left(\frac{5\,i}{3}-\frac{\pi}{2}-i\,(\gamma+\log 2)\right)\right.\\
&+\epsilon^2\left(\frac{-457\,i}{420}\,\pi+\frac{5\,\pi^2}{24}+\frac{457\,(\gamma+\log 2)}{210}+\frac{i}{2}\,\pi\,(\gamma+\log 2)-\frac{(\gamma+\log 2)^2}{2}\right)\\
&+\epsilon^3\left(\frac{343\,i}{216}+\frac{107\,\pi}{252}+\frac{491\,i}{1260}\,\pi^2-\frac{\pi^3}{16}+\frac{i}{3}\,\zeta(3)+\frac{107\,i}{126}\,(\gamma+\log 2)-\frac{47\,\pi\,(\gamma+\log 2)}{35}\right.
\end{aligned}
$$

$$-\frac{5\,i}{24}\pi^2\,(\gamma+\log 2)-\frac{47\,i}{35}(\gamma+\log 2)^2+\frac{\pi\,(\gamma+\log 2)^2}{4}+\frac{i}{6}(\gamma+\log 2)^3\Bigg)\Bigg],$$

$$B^{in}_{3\omega}=-\frac{1}{\omega^2}e^{-i\epsilon(\log 2\epsilon+\gamma)}\left(15-\frac{3}{4}i\epsilon\right)\Bigg[1+\epsilon\left(\frac{13\,i}{6}-\frac{\pi}{2}-i\,(\gamma+\log 2)\right)$$
$$+\epsilon^2\left(\frac{-26\,i}{21}\pi+\frac{5\,\pi^2}{24}+\frac{52\,(\gamma+\log 2)}{21}+\frac{i}{2}\pi\,(\gamma+\log 2)-\frac{(\gamma+\log 2)^2}{2}\right)$$
$$+\epsilon^3\Bigg(\frac{3623\,i}{1080}+\frac{169\,\pi}{504}+\frac{481\,i}{1008}\pi^2-\frac{\pi^3}{16}+\frac{i}{3}\zeta(3)+\frac{169\,i}{252}\,(\gamma+\log 2)-\frac{39\,\pi\,(\gamma+\log 2)}{28}$$
$$-\frac{5\,i}{24}\pi^2\,(\gamma+\log 2)-\frac{39\,i}{28}(\gamma+\log 2)^2+\frac{\pi\,(\gamma+\log 2)^2}{4}+\frac{i}{6}(\gamma+\log 2)^3\Bigg)\Bigg]. \qquad (3.2)$$

In the expression of $B^{in}_{\ell\omega}$, we have assumed that ω is positive. The expression corresponding to negative ω can be obtained by using the relation that $B^{in}_{\ell-\omega}=B^{in*}_{\ell\omega}$. For $\ell\geq 4$, we only need the forms of $B^{in}_{\ell\omega}$ valid up to $O(\epsilon)$, which can be found in Tagoshi and Sasaki [2].

Then it is straightforward to obtain the luminosity of the gravitational waves. We just give the total luminosity here:

$$\frac{dE}{dt}=\left(\frac{dE}{dt}\right)_N\Bigg[1-\frac{1247}{336}v^2+4\,\pi\,v^3-\frac{44711}{9072}v^4-\frac{8191\,\pi}{672}v^5$$
$$+\left(\frac{6643739519}{69854400}-\frac{1712\,\gamma}{105}-\frac{1712\log v}{105}+\frac{16\,\pi^2}{3}-\frac{3424\,\log 2}{105}\right)v^6-\frac{16285\,\pi}{504}v^7$$
$$+\Bigg(-\frac{323105549467}{3178375200}+\frac{232597\,\gamma}{4410}+\frac{232597\log v}{4410}$$
$$-\frac{1369\,\pi^2}{126}+\frac{39931\,\log 2}{294}-\frac{47385\,\log 3}{1568}\Bigg)v^8$$
$$+\left(\frac{265978667519\,\pi}{745113600}-\frac{6848\,\gamma\,\pi}{105}-\frac{6848\log v\,\pi}{105}-\frac{13696\,\pi\,\log 2}{105}\right)v^9$$
$$+\Bigg(-\frac{2500861660823683}{2831932303200}+\frac{916628467\,\gamma}{7858620}+\frac{916628467\log v}{7858620}-\frac{424223\,\pi^2}{6804}$$
$$-\frac{83217611\,\log 2}{1122660}+\frac{47385\,\log 3}{196}\Bigg)v^{10}$$
$$+\Bigg(\frac{8399309750401\,\pi}{101708006400}+\frac{177293\,\gamma\,\pi}{1176}+\frac{177293\log v\,\pi}{1176}$$
$$+\frac{8521283\,\pi\,\log 2}{17640}-\frac{142155\,\pi\,\log 3}{784}\Bigg)v^{11}$$

$$=\left(\frac{dE}{dt}\right)_N\Bigg[1-3.711309523809524\,v^2+12.56637061435917\,v^3$$
$$-4.928461199294533\,v^4-38.29283545469344\,v^5$$

$$+(115.7317166756113 - 16.3047619047619\log v)\ v^6$$

$$-101.5095959597416\,v^7+$$

$$(-117.5043907226773 + 52.74308390022676\log v)\ v^8$$

$$+(719.1283422334299 - 204.8916808741229\log v)\ v^9$$

$$+(-1216.906991317042 + 116.6398765941094\log v)\ v^{10}$$
$$+(958.934970119567 + 473.6244781742307\log v)\ v^{11} + \cdots\Bigg] \quad , \qquad (3.3)$$

where $(dE/dt)_N \equiv (32/5)(\mu/M)^2 v^{10}$ is the luminosity given by using the well-known quadrupole formula.

IV. DISCUSSION

We calculated the energy flux induced by a particle of small mass in a circular orbit around a Schwarzschild black hole up to $O(v^{11})$ beyond the quadrupole formula analytically.

We first compare our results with the numerical estimate given by Tagoshi and Nakamura [5]. The relative difference $\log_{10}|((dE/dt)^{(n)} - (dE/dt)_{num})/(dE/dt)_{num}|$, where $(dE/dt)^{(n)}$ is the energy flux calculated by the formula (3.3) truncated at v^n order and $(dE/dt)_{num}$ is that obtained by Tagoshi and Nakamura numerically, is plotted in figure 1. The axis of abscissas represents the radius of the orbit of the small mass particle. The numbers on the lines in figure 1 are the post-Newtonian order, n, in $(dE/dt)^{(n)}$. This comparison shows that the convergence is monotonic in the presented range ($r_0 > 100$) and that the numerical estimate given by Tagoshi and Nakamura is very accurate. For smaller radius we compared our present results with the data used in Tagoshi et. al. [6]. It was plotted in figure 2, which is the same plot as figure 1 but is plotted for smaller r_0. From these plots, we can see that the series with $n = 3m - 1$, where m is an integer, seems steadily converging. Several spikes observed in these figures are nothing but the points at which $(dE/dt)^{(n)}$ coincides with $(dE/dt)_{num}$.

For the innermost stable circular orbit ($r_0 = 6$), at which the post-Newtonian correction becomes most important. the convergence is slow. Even the formula valid up to $O(v^{11})$ still has relative error of 4%. As was pointed out by Tagoshi and Sasaki, this fact leads to a relatively large error in the estimate of N, where N is the total cycle of the gravitational waves from an inspiraling compact binary during sweep through the LIGO detection band,

$$N = \int_{r_f}^{r_i} dr \frac{\Omega}{\pi} \frac{dE/dr}{|dE/dt|}. \tag{4.1}$$

Here r_i and r_f is the initial and final orbital separations of binary, respectively. When $\mu \ll M$, dE/dr can be evaluated by the relation (2.5) and$|dE/dt|$ by the energy loss rate (3.3). To examine the effect on N due to the error in the estimate of the energy loss rate by using the present black hole perturbation analysis, we introduce $\Delta N^{(n)}$ which is defined by

$$\Delta N = \frac{5}{32\pi} \left| \int_{M/r_i}^{M/r_f} dv \frac{a_n v^n}{v^6} \right|, \tag{4.2}$$

where a_n is the coefficient of the $O(v^n)$ term in the gravitational wave luminosity, i.e., $\sum a_k v^k = (dE/dt)/(dE/dt)_N$.

For $(1.4M_\odot, 10M_\odot)$ binary, $r_i \sim 70M$ and $r_f = 6M$. Then we obtain

$$\begin{aligned} &\Delta N^{(6)} \sim 20, \quad \Delta N^{(7)} \sim 4, \quad \Delta N^{(8)} \sim 1, \\ &\Delta N^{(9)} \sim 2, \quad \Delta N^{(10)} \sim 1, \quad \Delta N^{(11)} \sim 0.4. \end{aligned} \tag{4.3}$$

Here we have confirmed again that the convergence in this case is not rapid. However, in this case, as μ/M is small, the standard post-Newtonian calculation, in which the both stars are supposed to have comparable masses, will be improved much when the leading order correction in μ/M which can be obtained by the black hole perturbation is taken into account. In the case of nearly equal mass binary, such as $(1.4M_\odot, 1.4M_\odot)$ neutron star binary, the next order correction in μ/M can be large. But, fortunately, the convergence with respect to the post-Newtonian expansion is not so bad in this case. Thus the standard post-Newtonian calculation with the correction obtained by the black hole perturbation will give a rather good template in both cases. As the standard post-Newtonian calculation has been already completed up to $O(v^5)$ by Blanchet, in order to improve the template, it will be important to examine the next order correction both in μ/M beyond the linear order and in v beyond $O(v^5)$.

---

[1] C. Cutler et al., Phys. Rev. Lett. **70**, 2984 (1993).

[2] H. Tagoshi and M. Sasaki, Prog. Theor. Phys. **92**, 745(1994).

[3] private communication.

[4] M. Sasaki, Prog. Theor. Phys. **92**, 17 (1994).

[5] H. Tagoshi and T. Nakamura, Phys. Rev.**D49**, 4016 (1994).

[6] H. Tagoshi, M. Shibata, T.Tanaka, and M. Sasaki, Osaka university preprint, OU-TAP-28 (1996).

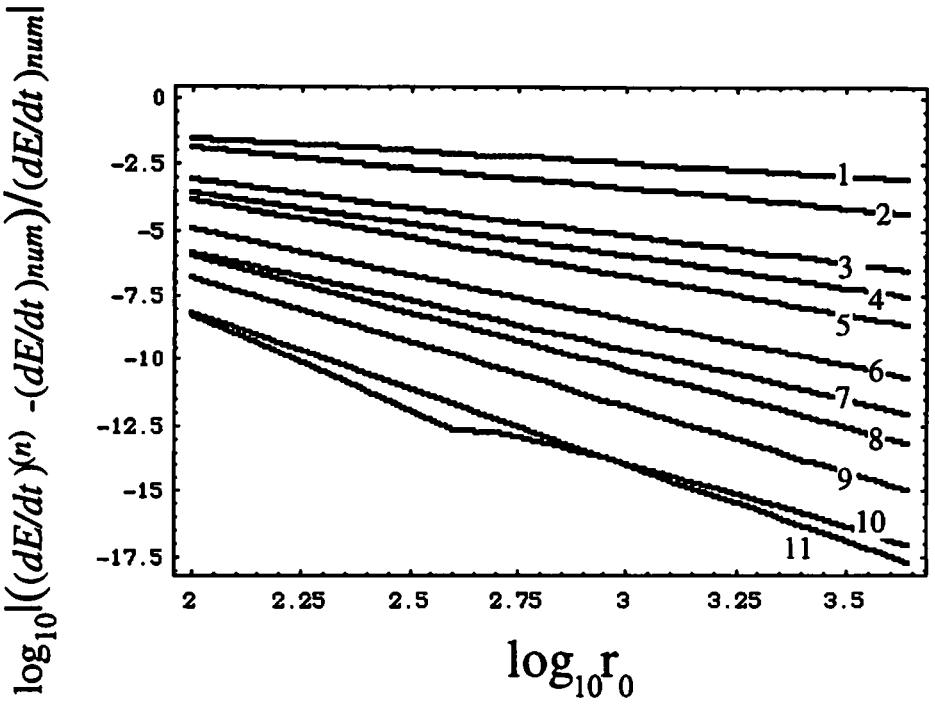


Fig.1 The relative error of the post-Newtonian formula of the energy flux versus the orbital radius. The numbers in the figure show the post-Newtonian order taken in the estimate of the energy flux.

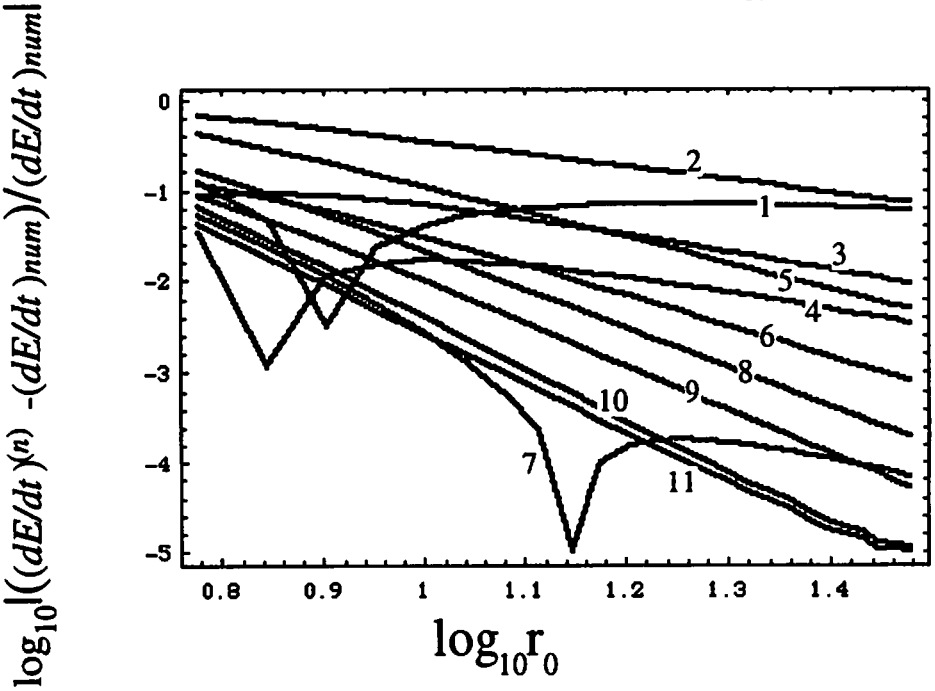


Fig.2 The same plots for smaller orbital radius.

February 21, 1996

Post-Newtonian Hydrodynamic Equations Using the (3+1) Formalism*

HIDEKI ASADA[1], MASARU SHIBATA[1] AND TOSHIFUMI FUTAMASE[2]

[1] *Department of Earth and Space Science, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan*

[2] *Astronomical Institute, Graduate School of Science, Tohoku University, Sendai 980-77, Japan*

ABSTRACT

The post-Newtonian(PN) hydrodynamic equations in the (3+1) formalism are obtained up through the 2.5PN order including the quadrupole gravitational radiation reaction. These equations are valid in the various slice conditions although we adopt a kind of transverse gauge condition to determine the shift vector. In particular, we describe methods to solve the 2PN tensor potential which arises from the spatial 3-metric. Our formulaton in the PN approximation using the (3+1) formalism will be useful for simulations providing an initial data for the final merging phase of coalescing binary neutron stars which can be treated only by fully general relativistic simulations.

---

* This is based on H. Asada, M. Shibata and T Futamase, *Post-Newtonian Hydrodynamic Equations Using the (3+1) Formalism in General Relativity*, submitted to Prog. Theor. Phys.

1. Introduction

Kilometer-size interferometric gravitational wave detectors, such as LIGO[1] and VIRGO[2], are now in construction aiming at direct detection of gravitational waves from relativistic astrophysical objects. Coalescing binary neutron stars are the most promising sources of gravitational waves for such detectors. After a long time emission of gravitational waves, the orbital separation of binary becomes comparable to the radius of the neutron star. Then, each of binary neutron stars begins to behave as a hydrodynamic object, not as a point particle, because they are tidally coupled each other. Recently, Lai, Rasio and Shapiro[4] have pointed out that such a tidal coupling of binary neutron stars is very important for their evolution in the final merging phase because the tidal effect causes the instability to the circular motion of them. Also important is the general relativistic gravity because in such a phase, the orbital separation is larger than $\sim 10\%$ of the Schwarzschild radius of the system. Thus, we need not only a hydrodynamic treatment, but also general relativistic one to study the final phase of binary neutron stars.

Fully general relativistic simulation is sure to be the best method, but it is also one of the most difficult ones. Although much effort has been focused and much progress can be expected there[3], it will take a long time until numerical relativistic calculations become reliable. One of the main reasons is that we do not know the behavior of the geometric variables in the strong field around coalescing binary neutron stars. Owing to this, we do not know what sort of gauge condition is useful and how to give an appropriate general relativistic initial condition for coalescing binary neutron stars.

The other reason is a technical one: In numerical relativistic simulations, gravitational waves are generated when we need to cover a region $L > \lambda \sim$ the wavelength in order to perform accurate simulations. This is in contrast with the case of Newtonian and/or PN simulations, in which we only need to cover a region $(\lambda >) L > R \sim$ the orbital separation. At present, we had better search other methods to prepare an initial condition for binary neutron stars.

The PN approximation in the fuid was pioneered by Chandrasekhar et. al.[6] who obtained the equation of motion up to the 2.5PN order. However, their expressions include potentials for non-compact sources which should be solved rather carefully in practical cases. On the other hand, using the ADM gauge, Blanchet et. al. obtained the (1+2.5)PN formula[7] which only consists of potentials for compact sources. Their formula actually works well in simulations of the coalescing binary neutron stars.[8] Our aim is to establish the formulation up to the 2.5PN hydrodynamic equation for a fairly wide class of gauge conditions and, in particular, to provide methods to solve the 2PN tensor potential which was not treated by Blanchet et. al.[7].

We use the units of $c = G = 1$ in this paper. Greek and Latin indices take $0, 1, 2, 3$ and $1, 2, 3$, respectively.

2. (3+1) Formalism for Post-Newtonian Approximation

2.1 (3+1) Formalism

In the (3+1) formalism, the metric is split as

$$g_{\mu\nu} = \gamma_{\mu\nu} - \hat{n}_\mu \hat{n}_\nu, \tag{2.1}$$

where $\hat{n}_\mu = (-\alpha,\ 0)$. Then the line element is written as

$$ds^2 = -(\alpha^2 - \beta_i\beta^i)dt^2 + 2\beta_i dt dx^i + \gamma_{ij}dx^i dx^j. \tag{2.2}$$

To distinguish the wave part from the non-wave part(for example, Newtonian potential) in the metric, we use $\tilde{\gamma}_{ij} = \psi^{-4}\gamma_{ij}$ instead of γ_{ij} where $\psi = \det(\gamma_{ij})^{1/12}$ so that $\det(\tilde{\gamma}_{ij}) = 1$ is satisfied. Using the extrinsic curvature, K_{ij}, we define $\tilde{A}_{ij}$ as

$$\tilde{A}_{ij} \equiv \psi^{-4}A_{ij} \equiv \psi^{-4}\Big(K_{ij} - \frac{1}{3}\gamma_{ij}K\Big). \tag{2.3}$$

The Einstein equation is split into the constraint equations and the evolution equations. The formers are the so-called Hamiltonian and momentum constraints which respectively become

$$\tilde{\Delta}\psi = \frac{1}{8}\mathrm{tr}\tilde{R}\psi - 2\pi\rho_H\psi^5 - \frac{\psi^5}{8}\Big(\tilde{A}_{ij}\tilde{A}^{ij} - \frac{2}{3}K^2\Big), \tag{2.4}$$

$$\tilde{D}_j(\psi^6\tilde{A}^j{}_i) - \frac{2}{3}\psi^6\tilde{D}_i K = 8\pi\psi^6 J_i, \tag{2.5}$$

where we introduce $\rho_H = T_{\mu\nu}n^\mu n^\nu$ and $J_j = -T_{\mu\nu}n^\mu\gamma^\nu_j$, and $\tilde{D}_i$, $\tilde{\Delta}$ and $\mathrm{tr}\tilde{R}$ are the covariant derivative, Laplacian and the scalar curvature with respect to $\tilde{\gamma}_{ij}$. Evolution equations for the spatial metric and extrinsic curvature are respectively

$$\frac{\partial}{\partial n}\tilde{\gamma}_{ij} = -2\alpha\tilde{A}_{ij} + \tilde{\gamma}_{il}\frac{\partial\beta^l}{\partial x^j} + \tilde{\gamma}_{jl}\frac{\partial\beta^l}{\partial x^i} - \frac{2}{3}\tilde{\gamma}_{ij}\frac{\partial\beta^l}{\partial x^l}, \tag{2.6}$$

$$\begin{aligned}\frac{\partial}{\partial n}\tilde{A}_{ij} =& \frac{1}{\psi^4}\Big[\alpha\Big(R_{ij} - \frac{1}{3}\gamma_{ij}\mathrm{tr}R\Big) - \Big(\tilde{D}_i\tilde{D}_j\alpha - \frac{1}{3}\tilde{\gamma}_{ij}\tilde{\Delta}\alpha\Big) - \frac{2}{\psi}\Big(\psi_{,i}\alpha_{,j} + \psi_{,j}\alpha_{,i} - \frac{2}{3}\tilde{\gamma}_{ij}\tilde{\gamma}^{kl}\psi_{,k}\alpha_{,l}\Big)\Big] \\ &+\alpha(K\tilde{A}_{ij} - 2\tilde{A}_{il}\tilde{A}^l{}_j) + \frac{\partial\beta^m}{\partial x^i}\tilde{A}_{mj} + \frac{\partial\beta^m}{\partial x^j}\tilde{A}_{mi} - \frac{2}{3}\frac{\partial\beta^m}{\partial x^m}\tilde{A}_{ij} - 8\pi\frac{\alpha}{\psi^4}\Big(S_{ij} - \frac{1}{3}\gamma_{ij}S^l{}_l\Big),\end{aligned} \tag{2.7}$$

$$\frac{\partial}{\partial n}\psi = \frac{\psi}{6}\Big(-\alpha K + \frac{\partial\beta^i}{\partial x^i}\Big), \tag{2.8}$$

$$\frac{\partial}{\partial n}K = \alpha\Big(\tilde{A}_{ij}\tilde{A}^{ij} + \frac{1}{3}K^2\Big) - \frac{1}{\psi^4}\tilde{\Delta}\alpha - \frac{2}{\psi^5}\tilde{\gamma}^{kl}\psi_{,k}\alpha_{,l} + 4\pi\alpha(S^i{}_i + \rho_H), \tag{2.9}$$

where R_{ij}, γ, S_{ij} and $\frac{\partial}{\partial n}$ denote, respectively, the Ricci tensor with respect of γ_{ij}, determinant of γ_{ij}, $T_{kl}\gamma^k{}_i\gamma^l{}_j$ and $\frac{\partial}{\partial t} - \beta^i\frac{\partial}{\partial x^i}$.

In order to clarify the wave property of $\tilde{\gamma}_{ij}$, we impose a kind of transverse gauge* to h_{ij} as

$$h_{ij,j} = 0. \tag{2.10}$$

Finally, we consider the matter as the perfect fluid, $T^{\mu\nu} = (\rho + \rho\varepsilon + P)u^\mu u^\nu + Pg^{\mu\nu}$, where u^μ, ρ, ε and P are the four velocity, the mass density, the specific internal energy and the pressure.

2.2 Post-Newtonian approximation

Each metric variable, which is relevant to the present paper, is expanded as

$$\begin{aligned}
\psi &= 1 + {}_{(2)}\psi + {}_{(4)}\psi + {}_{(6)}\psi + {}_{(7)}\psi + \dots, \\
\alpha &= 1 - U + \left(\frac{U^2}{2} + X\right) + {}_{(6)}\alpha + {}_{(7)}\alpha + \dots, \\
\beta^i &= {}_{(3)}\beta_i + {}_{(5)}\beta_i + {}_{(6)}\beta_i + {}_{(7)}\beta_i + {}_{(8)}\beta_i + \dots, \\
h_{ij} &= {}_{(4)}h_{ij} + {}_{(5)}h_{ij} + \dots, \\
\tilde{A}_{ij} &= {}_{(3)}\tilde{A}_{ij} + {}_{(5)}\tilde{A}_{ij} + {}_{(6)}\tilde{A}_{ij} + \dots, \\
K &= {}_{(3)}K + {}_{(5)}K + {}_{(6)}K + \dots,
\end{aligned} \tag{2.11}$$

where subscripts denote the PN order(c^{-n}) and U is the Newtonian potential satisfying

$$\Delta_{flat} U = -4\pi\rho. \tag{2.12}$$

From Eqs.(2.6) and (2.7), the wave equation for h_{ij} is written as

$$\begin{aligned}
\Box\, h_{ij} =& \left(1 - \frac{\alpha^2}{\psi^4}\right)\Delta_{flat} h_{ij} + \left(\frac{\partial^2}{\partial n^2} - \frac{\partial^2}{\partial t^2}\right) h_{ij} \\
&+ \frac{2\alpha}{\psi^4}\Big[-\frac{2\alpha}{\psi}\left(\tilde{D}_i\tilde{D}_j - \frac{1}{3}\tilde{\gamma}_{ij}\tilde{\Delta}\right)\psi + \frac{6\alpha}{\psi^2}\left(\tilde{D}_i\psi\tilde{D}_j\psi - \frac{1}{3}\tilde{\gamma}_{ij}\tilde{D}_k\psi\tilde{D}^k\psi\right) \\
&\quad - \left(\tilde{D}_i\tilde{D}_j - \frac{1}{3}\tilde{\gamma}_{ij}\tilde{\Delta}\right)\alpha - \frac{2}{\psi}\left(\tilde{D}_i\psi\tilde{D}_j\alpha + \tilde{D}_j\psi\tilde{D}_i\alpha - \frac{2}{3}\tilde{\gamma}_{ij}\tilde{D}^k\psi\tilde{D}_k\alpha\right)\Big] \\
&+ 2\alpha^2\left(K\tilde{A}_{ij} - 2\tilde{A}_{il}\tilde{A}^l{}_j\right) + 2\alpha\left(\beta^m_{,i}\tilde{A}_{mj} + \beta^m_{,j}\tilde{A}_{mi} - \frac{2}{3}\beta^m_{,m}\tilde{A}_{ij}\right) \\
&- 16\pi\frac{\alpha^2}{\psi^4}\left(S_{ij} - \frac{1}{3}\gamma_{ij}S^l{}_l\right) - \frac{\partial}{\partial n}\left(\beta^m_{,i}\tilde{\gamma}_{mj} + \beta^m_{,j}\tilde{\gamma}_{mi} - \frac{2}{3}\beta^m_{,m}\tilde{\gamma}_{ij}\right) + 2\frac{\partial\alpha}{\partial n}\tilde{A}_{ij} \\
\equiv& \tau_{ij},
\end{aligned} \tag{2.13}$$

where $\Box = -\frac{\partial^2}{\partial t^2} + \Delta_{flat}$.

* Hereafter, we call this condition merely the transverse gauge.

Thus, ${}_{(4)}h_{ij}$ is determined from

$$\Delta_{flat}{}_{(4)}h_{ij} = {}_{(4)}\tau_{ij}, \tag{2.14}$$

while ${}_{(5)}h_{ij}$ is derived from

$${}_{(5)}h_{ij}(t) = \frac{1}{4\pi}\frac{\partial}{\partial t}\int {}_{(4)}\tau_{ij}(t,\mathbf{y})d^3y. \tag{2.15}$$

Up to the 2.5PN order, the hydrodynamic equations become

$$\begin{aligned}
\frac{\partial S_i}{\partial t} + \frac{\partial (S_i v^j)}{\partial x^j} &= -\Big(1 + 2U + \frac{5}{4}U^2 + 6{}_{(4)}\psi + X\Big)P_{,i} \\
&+ \rho_*\Big[U_{,i}\Big\{1 + \varepsilon + \frac{P}{\rho} + \frac{3}{2}v^2 - U + \frac{5}{8}v^4 + 4v^2U + \Big(\frac{3}{2}v^2 - U\Big)\Big(\varepsilon + \frac{P}{\rho}\Big) + 3{}_{(3)}\beta_j v^j\Big\} \\
&- X_{,i}\Big(1 + \varepsilon + \frac{P}{\rho} + \frac{v^2}{2}\Big) + 2v^2{}_{(4)}\psi_{,i} - {}_{(6)}\alpha_{,i} - {}_{(7)}\alpha_{,i} \\
&+ v^j\Big\{{}_{(3)}\beta_{j,i}\Big(1 + \varepsilon + \frac{P}{\rho} + \frac{v^2}{2} + 3U\Big) + {}_{(5)}\beta_{j,i} + {}_{(6)}\beta_{j,i}\Big\} + {}_{(3)}\beta_{j}{}_{(3)}\beta_{j,i} \\
&+ \frac{1}{2}v^j v^k({}_{(4)}h_{jk,i} + {}_{(5)}h_{jk,i}) + O(c^{-8})\Big],
\end{aligned} \tag{2.16}$$

$$\frac{\partial H}{\partial t} + \frac{\partial (Hv^j)}{\partial x^j} = -P\Big[v^j{}_{,j} + \frac{\partial}{\partial t}\Big(\frac{1}{2}v^2 + 3U\Big) + \frac{\partial}{\partial x^j}\Big\{\Big(\frac{1}{2}v^2 + 3U\Big)v^j\Big\} + O(c^{-5})\Big], \tag{2.17}$$

where $S_i = \alpha\psi^6(\rho + \rho\varepsilon + P)u^0u_i$, $S^0 = \alpha\psi^6(\rho + \rho\varepsilon + P)(u^0)^2$ and $H = \alpha\psi^6\rho\varepsilon u^0$.

3. Strategy to obtain 2PN tensor potential

In this section, we describe methods to solve the equation for the 2PN tensor potential ${}_{(4)}h_{ij}$. Although Eq.(2.14) is formally solved as

$${}_{(4)}h_{ij}(t,\mathbf{x}) = -\frac{1}{4\pi}\int \frac{{}_{(4)}\tau_{ij}(t,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}d^3y, \tag{3.1}$$

it seems difficult to estimate this integral in practice since ${}_{(4)}\tau_{ij} \to O(r^{-3})$ for $r \to \infty$ and the integral is taken all over the space. Thus it is desirable to replace this equation by some tractable forms in numerical evaluation. In the following, we show two approaches: In section 3.1, we change Eq.(3.1) into the form in which the integration is performed only over the matter distribution like as in the Newtonian potential. In section 3.2, we propose a method to solve Eq.(2.14) as the boundary value problem.

3.1 DIRECT INTEGRATION METHOD

The explicit form of ${}_{(4)}\tau_{ij}$ is

$$
\begin{aligned}
{}_{(4)}\tau_{ij} =& -2\hat{\partial}_{ij}(X + 2{}_{(4)}\psi) + U\hat{\partial}_{ij}U - 3U_{,i}U_{,j} + \delta_{ij}U_{,k}U_{,k} - 16\pi\left(\rho v^i v^j - \frac{1}{3}\delta_{ij}\rho v^2\right) \\
& - \left({}_{(3)}\dot{\beta}_{i,j} + {}_{(3)}\dot{\beta}_{j,i} - \frac{2}{3}\delta_{ij}{}_{(3)}\dot{\beta}_{k,k}\right),
\end{aligned}
\tag{3.2}
$$

where $\hat{\partial}_{ij} \equiv \frac{\partial^2}{\partial x^i \partial x^j} - \frac{1}{3}\delta_{ij}\Delta_{flat}$. Since ${}_{(3)}\dot{\beta}_i$ is written as

$$
{}_{(3)}\dot{\beta}_i = \dot{p}_i - (X + 2{}_{(4)}\psi)_{,i} - \left\{\int \frac{\left(\rho v^2 + 3P - \rho U/2\right)}{|\mathbf{x}-\mathbf{y}|} d^3y\right\}_{,i}, \tag{3.3}
$$

the source term, ${}_{(4)}\tau_{ij}$, is split into five parts ${}_{(4)}\tau_{ij} = {}_{(4)}\tau_{ij}^{(S)} + {}_{(4)}\tau_{ij}^{(U)} + {}_{(4)}\tau_{ij}^{(C)} + {}_{(4)}\tau_{ij}^{(\rho)} + {}_{(4)}\tau_{ij}^{(V)}$, where we introduce

$$
\begin{aligned}
{}_{(4)}\tau_{ij}^{(S)} =& -16\pi\left(\rho v^i v^j - \frac{1}{3}\delta_{ij}\rho v^2\right), \\
{}_{(4)}\tau_{ij}^{(U)} =& UU_{,ij} - \frac{1}{3}\delta_{ij}U\Delta_{flat}U - 3U_{,i}U_{,j} + \delta_{ij}U_{,k}U_{,k}, \\
{}_{(4)}\tau_{ij}^{(C)} =& 4\frac{\partial}{\partial x^j}\int \frac{(\rho v^i)^{\cdot}}{|\mathbf{x}-\mathbf{y}|} d^3y + 4\frac{\partial}{\partial x^i}\int \frac{(\rho v^j)^{\cdot}}{|\mathbf{x}-\mathbf{y}|} d^3y - \frac{8}{3}\delta_{ij}\frac{\partial}{\partial x^k}\int \frac{(\rho v^k)^{\cdot}}{|\mathbf{x}-\mathbf{y}|} d^3y, \\
{}_{(4)}\tau_{ij}^{(\rho)} =& \hat{\partial}_{ij}\int \ddot{\rho}|\mathbf{x}-\mathbf{y}| d^3y, \\
{}_{(4)}\tau_{ij}^{(V)} =& 2\hat{\partial}_{ij}\int \frac{\left(\rho v^2 + 3P - \rho U/2\right)}{|\mathbf{x}-\mathbf{y}|} d^3y.
\end{aligned}
\tag{3.4}
$$

Thus it becomes clear that ${}_{(4)}h_{ij}$ and ${}_{(5)}h_{ij}$ as well as ${}_{(4)}\tau_{ij}$ are expressed in terms of matter variables only and thus do not depend on slicing conditions. Then, we define ${}_{(4)}h_{ij}^{(*)} = \Delta_{flat}^{-1}{}_{(4)}\tau_{ij}^{(*)}$, where the symbol '$*$' denotes S, U, C, ρ and V.

First, since ${}_{(4)}\tau_{ij}^{(S)}$ is a compact source, we immediately obtain

$$
{}_{(4)}h_{ij}^{(S)} = 4\int \frac{\left(\rho v^i v^j - \frac{1}{3}\delta_{ij}\rho v^2\right)}{|\mathbf{x}-\mathbf{y}|} d^3y.
$$

Second, we consider the following equation

$$
\Delta_{flat}G(\mathbf{x}, \mathbf{y_1}, \mathbf{y_2}) = \frac{1}{|\mathbf{x}-\mathbf{y_1}||\mathbf{x}-\mathbf{y_2}|}. \tag{3.5}
$$

It is possible to write ${}_{(4)}h_{ij}^{(U)}$ as integrals over the matter using this function, G. Eq.(3.5) has solutions[9], $\ln(r_1 + r_2 \pm r_{12})$, where $r_1 = |\mathbf{x}-\mathbf{y_1}|, r_2 = |\mathbf{x}-\mathbf{y_2}|, r_{12} = |\mathbf{y_1}-\mathbf{y_2}|$. Note that

$\ln(r_1+r_2-r_{12})$ is not regular on the interval between $\mathbf{y_1}$ and $\mathbf{y_2}$, while $\ln(r_1+r_2+r_{12})$ is regular on the matter. Thus we use $\ln(r_1+r_2+r_{12})$ as a Green function. Using this function, $UU_{,ij}$ and $U_{,i}U_{,j}$ are rewritten as

$$\begin{aligned} UU_{,ij} =& \Delta_{flat}\int d^3y_1 d^3y_2 \rho(\mathbf{y_1})\rho(\mathbf{y_2})\frac{\partial^2}{\partial y_1^i \partial y_1^j}\ln(r_1+r_2+r_{12}), \\ U_{,i}U_{,j} =& \Delta_{flat}\int d^3y_1 d^3y_2 \rho(\mathbf{y_1})\rho(\mathbf{y_2})\frac{\partial^2}{\partial y_1^i \partial y_2^j}\ln(r_1+r_2+r_{12}). \end{aligned} \tag{3.6}$$

Thus we can express ${}_{(4)}h_{ij}^{(U)}$ using the integral over the matter as

$$\begin{aligned} {}_{(4)}h_{ij}^{(U)} = & \int d^3y_1 d^3y_2 \rho(\mathbf{y_1})\rho(\mathbf{y_2}) \\ & \Big[\Big(\frac{\partial^2}{\partial y_1^i \partial y_1^j}-\frac{1}{3}\delta_{ij}\Delta_1\Big)-3\Big(\frac{\partial^2}{\partial y_1^i \partial y_2^j}-\frac{1}{3}\delta_{ij}\Delta_{12}\Big)\Big]\ln(r_1+r_2+r_{12}), \end{aligned} \tag{3.7}$$

where we introduce $\Delta_1 = \frac{\partial^2}{\partial y_1^k \partial y_1^k}$ and $\Delta_{12} = \frac{\partial^2}{\partial y_1^k \partial y_2^k}$. Using relations $\Delta_{flat}|\mathbf{x}-\mathbf{y}| = 2/|\mathbf{x}-\mathbf{y}|$ and $\Delta_{flat}|\mathbf{x}-\mathbf{y}|^3 = 12|\mathbf{x}-\mathbf{y}|$, ${}_{(4)}h_{ij}^{(C)}$, ${}_{(4)}h_{ij}^{(\rho)}$ and ${}_{(4)}h_{ij}^{(V)}$ are solved as

$${}_{(4)}h_{ij}^{(C)} = 2\frac{\partial}{\partial x^i}\int(\rho v^j)^{\cdot}|\mathbf{x}-\mathbf{y}|d^3y + 2\frac{\partial}{\partial x^j}\int(\rho v^i)^{\cdot}|\mathbf{x}-\mathbf{y}|d^3y + \frac{4}{3}\delta_{ij}\int\ddot{\rho}|\mathbf{x}-\mathbf{y}|d^3y, \tag{3.8}$$

$${}_{(4)}h_{ij}^{(\rho)} = \frac{1}{12}\frac{\partial^2}{\partial x^i \partial x^j}\int\ddot{\rho}|\mathbf{x}-\mathbf{y}|^3 d^3y - \frac{1}{3}\delta_{ij}\int\ddot{\rho}|\mathbf{x}-\mathbf{y}|d^3y, \tag{3.9}$$

$${}_{(4)}h_{ij}^{(V)} = \frac{\partial^2}{\partial x^i \partial x^j}\int\Big(\rho v^2+3P-\frac{\rho U}{2}\Big)|\mathbf{x}-\mathbf{y}|d^3y - \frac{2}{3}\delta_{ij}\int\frac{\Big(\rho v^2+3P-\rho U/2\Big)}{|\mathbf{x}-\mathbf{y}|}d^3y. \tag{3.10}$$

As a result, we obtain ${}_{(4)}h_{ij} = {}_{(4)}h_{ij}^{(S)} + {}_{(4)}h_{ij}^{(U)} + {}_{(4)}h_{ij}^{(C)} + {}_{(4)}h_{ij}^{(\rho)} + {}_{(4)}h_{ij}^{(V)}$.

3.2 Treatment as a boundary value problem

The above expression for ${}_{(4)}h_{ij}$ is quite interesting because it only consists of integrals over the matter. However, in actual numerical simulations, it will take a very long time to perform the direct integration. Therefore, we also propose other strategies where Eq.(2.14) is solved as the boundary value problem. Here, we would like to emphasize that the boundary condition should be imposed at $r(=|\mathbf{x}|) \gg |\mathbf{y_1}|, |\mathbf{y_2}|$, but r does not have to be greater than λ, where λ is a typical wave length of gravitational waves. We only need to impose $r > R$(a typical size of matter). This means that we do not need a large amount of grid numbers compared with the case of fully general relativistic simulations, in which we require $r > \lambda \gg R$.

First of all, we consider the equation

$$\Delta_{flat}\Big({}_{(4)}h_{ij}^{(S)} + {}_{(4)}h_{ij}^{(U)}\Big) = {}_{(4)}\tau_{ij}^{(S)} + {}_{(4)}\tau_{ij}^{(U)}. \tag{3.11}$$

Since its source term behaves as $O(r^{-6})$ at $r \to \infty$, this equation can be accurately solved under the boundary condition at $r > R$ as

$$\begin{aligned} {}_{(4)}h_{ij}^{(S)} + {}_{(4)}h_{ij}^{(U)} = {} & \frac{2}{r}\Big(\ddot{I}_{ij} - \frac{1}{3}\delta_{ij}\ddot{I}_{kk}\Big) \\ & + \frac{2}{3r^2}\Big(n^k\dddot{I}_{ijk} - \frac{1}{3}\delta_{ij}n^k\dddot{I}_{llk} + 2n^k(\dot{S}_{ikj} + \dot{S}_{jki}) - \frac{4}{3}\delta_{ij}n^k\dot{S}_{lkl}\Big) + O(r^{-3}), \end{aligned} \tag{3.12}$$

where $I_{ij} = \int \rho x^i x^j d^3x$, $I_{ijk} = \int \rho x^i x^j x^k d^3x$ and $S_{ijk} = \int \rho(v^i x^j - v^j x^i)x^k d^3x$.

${}_{(4)}h_{ij}^{(C)}$, ${}_{(4)}h_{ij}^{(\rho)}$ and ${}_{(4)}h_{ij}^{(V)}$ can be rewritten as

$${}_{(4)}h_{ij}^{(C)} = 2\int(\rho v^j)_{\cdot}\frac{x^i - y^i}{|\mathbf{x}-\mathbf{y}|}d^3y + 2\int(\rho v^i)_{\cdot}\frac{x^j - y^j}{|\mathbf{x}-\mathbf{y}|}d^3y - \frac{4}{3}\delta_{ij}\int(\rho v^k)_{\cdot}\frac{x^k - y^k}{|\mathbf{x}-\mathbf{y}|}d^3y, \tag{3.13}$$

$$\begin{aligned} {}_{(4)}h_{ij}^{(\rho)} = {} & \frac{1}{4}\frac{\partial^2}{\partial x^i \partial x^j}\int \rho v^k v^l \frac{(x^k - y^k)(x^l - y^l)}{|\mathbf{x}-\mathbf{y}|}d^3y + \frac{1}{3}\delta_{ij}\int(\rho v^k)_{\cdot}\frac{x^k - y^k}{|\mathbf{x}-\mathbf{y}|}d^3y \\ & + \frac{1}{2}\Big\{\frac{\partial}{\partial x^i}\int P'\frac{(x^j - y^j)}{|\mathbf{x}-\mathbf{y}|}d^3y + \frac{\partial}{\partial x^j}\int P'\frac{(x^i - y^i)}{|\mathbf{x}-\mathbf{y}|}d^3y\Big\} \\ & - \frac{1}{8}\Big\{2\int \rho\frac{U_{,j}(x^i - y^i) + U_{,i}(x^j - y^j)}{|\mathbf{x}-\mathbf{y}|}d^3y \\ & \quad + x^k\frac{\partial}{\partial x^i}\int \rho\frac{U_{,k}(x^j - y^j)}{|\mathbf{x}-\mathbf{y}|}d^3y + x^k\frac{\partial}{\partial x^j}\int \rho\frac{U_{,k}(x^i - y^i)}{|\mathbf{x}-\mathbf{y}|}d^3y\Big\}, \end{aligned} \tag{3.14}$$

$$\begin{aligned} {}_{(4)}h_{ij}^{(V)} = {} & \frac{1}{2}\Big[\frac{\partial}{\partial x^i}\int\Big(\rho v^2 + 3P - \frac{\rho U}{2}\Big)\frac{x^j - y^j}{|\mathbf{x}-\mathbf{y}|}d^3y + \frac{\partial}{\partial x^j}\int\Big(\rho v^2 + 3P - \frac{\rho U}{2}\Big)\frac{x^i - y^i}{|\mathbf{x}-\mathbf{y}|}d^3y\Big] \\ & - \frac{2}{3}\delta_{ij}\int\frac{\big(\rho v^2 + 3P - \rho U/2\big)}{|\mathbf{x}-\mathbf{y}|}d^3y, \end{aligned} \tag{3.15}$$

where $P' = P + \rho v^2/4 + \rho U_{,l}y^l/4$. From the above relations, ${}_{(4)}h_{ij}^{(C)}$, ${}_{(4)}h_{ij}^{(\rho)}$ and ${}_{(4)}h_{ij}^{(V)}$ become

$${}_{(4)}h_{ij}^{(C)} = 2({}_{(3)}\dot{P}^j x^i + {}_{(3)}\dot{P}^i x^j - Q_{ij}) + \frac{4}{3}\delta_{ij}\Big(\frac{Q_{kk}}{2} - {}_{(3)}\dot{P}^k x^k\Big), \tag{3.16}$$

$$\begin{aligned} {}_{(4)}h_{ij}^{(\rho)} = {} & \frac{1}{4}\frac{\partial^2}{\partial x^i \partial x^j}\Big(V_{kl}^{(\rho v)}x^k x^l - 2V_k^{(\rho v)}x^k + V^{(\rho v)}\Big) + \frac{1}{3}\delta_{ij}\Big({}_{(3)}\dot{P}_k x^k - \frac{Q_{kk}}{2}\Big) \\ & + \frac{1}{2}\Big\{\frac{\partial}{\partial x^i}\Big(V^{(P)}x^j - V_j^{(P)}\Big) + \frac{\partial}{\partial x^j}\Big(V^{(P)}x^i - V_i^{(P)}\Big)\Big\} \\ & - \frac{1}{8}\Big\{2\Big(x^i V_j^{(\rho U)} + x^i V_i^{(\rho U)} - V_{ij}^{(\rho U)} - V_{ji}^{(\rho U)}\Big) \\ & \quad + x^k\frac{\partial}{\partial x^i}\Big(x^j V_k^{(\rho U)} - V_{kj}^{(\rho U)}\Big) + x^k\frac{\partial}{\partial x^j}\Big(x^i V_l^{(\rho U)} - V_{li}^{(\rho U)}\Big)\Big\}, \end{aligned} \tag{3.17}$$

$$ {}_{(4)}h_{ij}^{(V)} = \frac{1}{2}\Big(Q_{,j}^{(I)}x^i + Q_{,i}^{(I)}x^j - Q_{i,j}^{(I)} - Q_{j,i}^{(I)}\Big) + \frac{1}{3}Q^{(I)}\delta_{ij}, \tag{3.18} $$

where

$$ \begin{aligned}
&\Delta_{flat(3)}P_i = -4\pi\rho v^i, \\
&\Delta_{flat}Q_{ij} = -4\pi\Big\{x^j(\rho v^i)^{\cdot} + x^i(\rho v^j)^{\cdot}\Big\}, \\
&\Delta_{flat}Q^{(I)} = -4\pi\Big(\rho v^2 + 3P - \frac{1}{2}\rho U\Big), \\
&\Delta_{flat}Q_i^{(I)} = -4\pi\Big(\rho v^2 + 3P - \frac{1}{2}\rho U\Big)x^i, \\
&\Delta_{flat}V_{ij}^{(\rho v)} = -4\pi\rho v^i v^j, \\
&\Delta_{flat}V_i^{(\rho v)} = -4\pi\rho v^i v^j x^j, \\
&\Delta_{flat}V^{(\rho v)} = -4\pi\rho(v^j x^j)^2, \\
&\Delta_{flat}V^{(P)} = -4\pi P', \\
&\Delta_{flat}V_i^{(P)} = -4\pi P' x^i, \\
&\Delta_{flat}V_i^{(\rho U)} = -4\pi\rho U_{,i}, \\
&\Delta_{flat}V_{ij}^{(\rho U)} = -4\pi\rho U_{,i}x^j.
\end{aligned} \tag{3.19} $$

Therefore, ${}_{(4)}h_{ij}^{(C)}$, ${}_{(4)}h_{ij}^{(\rho)}$ and ${}_{(4)}h_{ij}^{(V)}$ can be derived from the above potentials which satisfy the Poisson equations with compact sources.

We note that, instead of the above procedure, we may solve the Poisson equation for ${}_{(4)}h_{ij}$ carefully imposing the boundary condition for $r \gg R$ as

$$ \begin{aligned}
{}_{(4)}h_{ij} =& \frac{1}{r}\Big\{\frac{1}{4}I_{ij}^{(2)} + \frac{3}{4}n^k\Big(n^i I_{kj}^{(2)} + n^j I_{ki}^{(2)}\Big) \\
&- \frac{5}{8}n^i n^j I_{kk}^{(2)} + \frac{3}{8}n^i n^j n^k n^l I_{kl}^{(2)} + \frac{1}{8}\delta_{ij}I_{kk}^{(2)} - \frac{5}{8}\delta_{ij}n^k n^l I_{kl}^{(2)}\Big\} \\
&+ \frac{1}{r^2}\Big[\Big\{-\frac{5}{12}n^k I_{ijk}^{(2)} - \frac{1}{24}(n^i I_{jkk}^{(2)} + n^j I_{ikk}^{(2)}) + \frac{5}{8}n^k n^l(n^i I_{jkl}^{(2)} + n^j I_{ikl}^{(2)}) \\
&- \frac{7}{8}n^i n^j n^k I_{kll}^{(2)} + \frac{5}{8}n^i n^j n^k n^l n^m I_{klm}^{(2)} + \frac{11}{24}\delta_{ij}n^k I_{kll}^{(2)} - \frac{5}{8}\delta_{ij}n^k n^l n^m I_{klm}^{(2)}\Big\} \\
&+ \Big\{\frac{2}{3}n^k(\dot{S}_{ikj} + \dot{S}_{jki}) - \frac{4}{3}(n^i\dot{S}_{jkk} + n^j\dot{S}_{ikk}) \\
&+ 2n^k n^l(n^i\dot{S}_{jkl} + n^j\dot{S}_{ikl}) + 2n^i n^j n^k\dot{S}_{kll} + \frac{2}{3}\delta_{ij}n^k\dot{S}_{kll}\Big\}\Big] + O(r^{-3}).
\end{aligned} \tag{3.20} $$

It should be noticed that ${}_{(4)}h_{ij}$ obtained in this way becomes meaningless at the far zone because Eq.(2.14), from which ${}_{(4)}h_{ij}$ is derived, is valid only in the near zone.

4. Summary

We have developed a formalism for the hydrodynamic equation accurate up to 2.5PN order. For the sake of an actual numerical simulation, we carefully consider methods to solve the various metric quantities, especially, the 2PN tensor potential ${}_{(4)}h_{ij}$. We found it possible to solve them by using standard numerical methods. Thus, the formalism developed here will be useful also in numerical calculations.

Moreover, we would like to emphasize that, from the 2PN order, the tensor part of the 3-metric, $\tilde{\gamma}_{ij}$, cannot be neglected even if we ignore gravitational waves. Recently, Wilson and Mathews[10] presented numerical equilibrium configurations of binary neutron stars using a semi-relativistic approximation, in which they assume the spatially conformal flat metric as the spatial 3-metric, *i.e.*, $\tilde{\gamma}_{ij} = \delta_{ij}$. Thus, in their method, a 2PN term, h_{ij}, was completely neglected. This means that their results unavoidably have an error of the 2PN order which will become $\sim (M/R)^2 \sim 1 - 10\%$. If we hope to obtain a general relativistic eqiulibrium configuration of binary neutron stars with a better accuracy(say less than 1%), we should take into account the tensor part of the 3-metric.

Acknowledgment

We thank T. Nakamura for his suggestion to pursue this problem and useful discussions. We also thank M. Sasaki and T. Tanaka for helpful discussion.

REFERENCES

1. A. Abramovici et al. Science, **256**(1992), 325;
   K. S. Thorne, in proceedings of the eighth Nishinomiya-Yukawa memorial symposium on Relativistic Cosmology, edited by M. Sasaki(Universal Academy Press, Tokyo, 1994), pp.67.
2. C. Bradaschia et al., Nucl. Instrum. Method Phys. Res. Sect. **A289**(1990), 518.
3. T. Nakamura, in the proceedings of the eighth Nishinomiya-Yukawa memorial symposium on Relativistic Cosmology(ref.1), pp.155.
4. D. Lai, F. Rasio and S. L. Shapiro, Astrophys. J. **420** (1994), 811; Astrophys. J. supplement **88**(1993), 205.
5. L. E. Kidder, C. M. Will and A. G. Wiseman, Phys. Rev. **D47**(1993), 3281.
6. S. Chandrasekhar and Y. Nutku, Astrophys. J. **158**(1969), 55; see their series of papers cited therein.
7. L. Blanchet, T. Damour and G. Schäfer, Mon. Not. R. Astr. Soc. **242**(1990), 289.
8. K. Oohara and T. Nakamura, Prog. Theor. Phys. **83**(1990), 906; **86**(1991), 73; **88**(1992), 307;
   M. Shibata, T. Nakamura and K. Oohara, Prog. Theor. Phys. **88**(1992), 1079; **89**(1993), 809.
9. T. Ohta. H. Okamura, T. Kimura and K. Hiida, Prog. Thor. Phys. **51**(1974), 1598.
10. J. R. Wilson and G. J. Mathews, Phys. Rev. Lett. **75**(1995), 4161.

Evolution of Close Neutron Star Binaries – Tidal and Spin Effects –

Wataru Ogawaguchi[1] and Yasufumi Kojima[2]

[1] Department of Physics, Tokyo Metropolitan University
Hachioji, Tokyo, 192-03 JAPAN

[2] Department of Physics, Hiroshima University
Higashi-Hiroshima, Hiroshima, 739 JAPAN

February 29, 1996

Abstract

We have calculated evolution of neutron star binaries towards the coalescence driven by gravitational radiation. The hydrodynamical effects as well as the general relativistic effects are important in the final phase. All corrections up to post$^{2.5}$-Newtonian order as well as the tidal effect are included in the orbital motion. The star is approximated by a simple Newtonian stellar model called the affine star model. Stellar spins and angular momentum are assumed to be aligned, but the magnitude of initial spins are assumed to be variable. We have showed how the internal stellar structure affects variations of the spins, and the orbital motion of the binary just before the contact. The gravitational wave forms from the last a few revolutions significantly depend on the stellar structure.

1 Introduction

For direct detection of gravitational waves by the laser interferometers such as LIGO [1], the most promising sources are coalescing neutron star binaries. If we can observe up to a few hundred Mpc, we have a few events by a year. Theoretically, it is important to understand the final phase of the binary coalescence, because the preparation of the gravitational wave forms as the template is necessary for the matched filter analysis. When the separation L of the binary is much larger than the stellar radius R, each star can be regarded as a point mass. The study to get the matched filter is currently in

progress using the higher order post-Newtonian approximation [2, 3]. In particular, the test particle limit is extensively studied by the perturbation of the black hole space-time (See e.g., Ref. [4, 5]). When L is a few times R, hydrodynamical effects become important as well as higher post-Newtonian effects. The only method to examine the final stage may be solving the Einstein equations numerically [6, 7].

One of hydrodynamical effects before the contact of the binary is tidal force. Lai, Rasio & Shapiro (Ref.[8, 9, 10]) calculated the dynamical evolution of the binary of simplified stellar models under the Newtonian gravity. They showed that the orbit significantly deviates from that of two point masses. The reason is that there is additional contribution from the stellar quadrupole moment, which causes marginally stable circular orbit at $L = (2 \sim 3)R$. When the separation becomes below the critical distance, the radial infall velocity significantly accelerates as well as the radiation reaction force. However, since the tidal potential depends on L^{-6}, the effect is significant only for close binary evolved with higher post-Newtonian correction. Natural question is whether the contribution is still significant, when other post-Newtonian corrections are included. The effect of the spinning star, i.e., spin-orbit and spin-spin interaction, may be also important since the tidal torque changes the spin. They are respectively post$^{1.5}$-Newtonian and post^{2}-Newtonian order in the magnitude. We take all these corrections and the radiation reaction force into consideration, and study how the binary evolves at $L < 15R$. We adopt simplified stellar models called affine star models. In this approximation, the fluid displacement is limited to a certain class of the motion, i.e., uniform expansion, rotation and quadrupole oscillation. The dynamical degrees of freedom of a star are reduced to finite number, and the equations of motion are not partial differential equations, but ordinary differential equations. Using this approximation, we can easily simulate the final phase of the binary with various effects concerning spin and quadrupole moment.

2 Numerical Methods

2.1 Orbital motion

In order to solve orbital motion of the binary, we use the Hamiltonian formalism (Ref.[11]). Regarding the star as a point mass with $M_a (a = 1, 2)$, we will examine the relative orbital motion described by the total mass M_T and reduced mass μ. We choose the orbital plane as the equatorial plane of the polar coordinate (r, θ, ϕ). The Newtonian, the first and second post-Newtonian equations of motion are determined by the following Hamiltonians.

$$H_N = \frac{1}{2\mu}\left(p_r^2 + \frac{p_\phi^2}{r^2}\right) - \frac{G\mu M_T}{r}, \tag{1}$$

$$H_{PN} = \frac{3\nu - 1}{8c^2\mu^3}\left(p_r^2 + \frac{p_\phi^2}{r^2}\right)^2 - \frac{GM_T}{2c^2\mu r}\left\{(3+\nu)\left(p_r^2 + \frac{p_\phi^2}{r^2}\right) + \nu p_r^2\right\} + \frac{G^2\mu M_T^2}{2c^2 r^2}, \tag{2}$$

$$
\begin{aligned}
H_{P^2N} &= \frac{1-5\nu+5\nu^2}{16c^4\mu^5}\left(p_r^2+\frac{p_\phi^2}{r^2}\right)^3 \\
&+\frac{GM_T}{8c^4\mu^3 r}\left\{\left(5-20\nu-3\nu^2\right)\left(p_r^2+\frac{p_\phi^2}{r^2}\right)^2-2\nu^2p_r^2\left(p_r^2+\frac{p_\phi^2}{r^2}\right)-3\nu^2p_r^4\right\} \quad (3)\\
&+\frac{G^2M_T^2}{2c^4\mu r^2}\left\{(5+8\nu)\left(p_r^2+\frac{p_\phi^2}{r^2}\right)+3\nu p_r^2\right\}-\frac{G^3(1+3\nu)\mu M_T^3}{4c^4r^3},
\end{aligned}
$$

where $\nu = \mu/M_T$.

If the star has mass quadrupole moment, then the Newtonian gravitational potential has an additional contribution ($\sim 1/r^3$). The relative motion is affected by the monopole-quadrupole interaction,

$$
H_T = -\frac{G}{2r^3}\left(3n^in^j-\delta^{ij}\right)\left\{M_2\left(\mathrm{I}_1\right)_{ij}+M_1\left(\mathrm{I}_2\right)_{ij}\right\}, \quad (4)
$$

where $\mathbf{n} = \mathbf{r}/r$ and $(\mathbf{I}_a)_{ij}$ is mass quadrupole moment of star a (a=1,2).

The spin-orbit and the spin-spin interaction in appropriate coordinate system can be written as

$$
H_{SO} = \frac{G}{c^2r^3}\mathbf{L}_N\cdot\left\{\left(2+\frac{3M_2}{2M_1}\right)\mathbf{J}_1+\left(2+\frac{3M_1}{2M_2}\right)\mathbf{J}_2\right\}, \quad (5)
$$

$$
H_{SS} = \frac{G}{c^2r^3}\left\{3\left(\mathbf{J}_1\cdot\mathbf{n}\right)\left(\mathbf{J}_2\cdot\mathbf{n}\right)-\left(\mathbf{J}_1\cdot\mathbf{J}_2\right)\right\}, \quad (6)
$$

where $\mathbf{J}_a$ is spin angular momentum of star a and $\mathbf{L}_N$ is orbital angular momentum of the Newtonian order, i.e., $\mathbf{L}_N = \mathbf{r}\times\mathbf{p}$.

2.2 Affine star model

Because the spin and the quadrupole moment, that are determined by the stellar structure, appeared in the hamiltonians as seen in the previous section, we adopted so-called the affine star for the stellar model. (See Ref.[12, 13]) The binary consists of two affine stars (a =1,2). For simplicity, we omit the suffix a and use the Cartesian coordinate system, whose origin is the mass center of a single star.

The position of a fluid element inside the star, $x_i(t)$ is specified by the Lagrangian map from its initial position $\hat{x}_i$ as

$$
x_i(t) = q_{ia}(t)\hat{x}_a, \quad (7)
$$

where q_{ia} is 3×3 matrix and $q_{ia}(0) = \delta_{ia}$. The initial sphere of radius R is transformed into an ellipsoid as

$$
S_{ij}^{-1}x_ix_j = R^2, \quad (8)
$$

where S_{ij}^{-1} is the inverse of a symmetric matrix, $S_{ij} = q_{ia}q_{ja}$. The eigenvalues of S_{ij} determine three principal axes of the ellipsoid.

The dynamics of the fluid motion can be determined in the Hamiltonian formalism by the generalized coordinate q_{ia} and its conjugate momentum $p_{ia} = I_0\dot{q}_{ia}$, where the inertial moment I_0 is described $I_0 = \frac{1}{3}\int \hat{x}_a\hat{x}_a dM$. We assume a polytropic equation of state, $P = \kappa\rho^{1+(1/n)}$. Then, the initial state is a spherically symmetric star with mass M and radius R, so that we have $I_0 = \frac{1}{5}\kappa_n MR^2$, where κ_n is a constant relating with polytropic index n. The Hamiltonian concerning the internal fluid motion is,

$$H_I = T + U + \Phi. \tag{9}$$

The kinetic energy T is a function of p_{ia} given by

$$T = \frac{1}{2}\int v_i v_i dM = \frac{1}{2I_0}p_{ia}p_{ia}. \tag{10}$$

The internal energy U is a function of q_{ia} described by

$$U = \int \frac{nP}{\rho}dM = U_0|\mathrm{q}|^{(-1/n)}, \tag{11}$$

where $|\mathrm{q}|$ is determinant of q_{ia}, and U_0 is the initial internal energy. The self-gravitational energy Φ can be calculated by an elliptical integral (e.g., Ref.[14])

$$\Phi = \frac{1}{2}\Phi_0 S_{ik}\int_0^\infty \frac{(\mathrm{S}+u\delta)^{-1}_{ik}}{|\mathrm{S}+u\delta|^{\frac{1}{2}}}du, \tag{12}$$

where Φ_0 is the self-gravity of spherically symmetric star. Since the pressure force and self-gravity assumed to balance at the initial state, we have

$$U_0 = -\frac{\Phi_0}{3} = \frac{GM^2}{(5-n)\,R}. \tag{13}$$

The quadrupole moment and the spin of the affine star model are the function of q_{ia} and p_{ia} given by,

$$I_{ij} = I_0 q_{ia}q_{ja}, \tag{14}$$

$$J_i = \varepsilon_{ijk}q_{ja}p_{ka}. \tag{15}$$

2.3 Equations of motion and gravitational radiation

The dynamical degrees of freedom of the system are reduced to finite number in our approximation. The variables are fluid variables (q_{ij}, p_{ij}) and orbital variables (r, ϕ, p_r, p_ϕ). The system is determined by the following Hamiltonian,

$$H = H_{I1} + H_{I2} + H_N + H_{PN} + H_{P^2N} + H_T + H_{SO} + H_{SS}. \tag{16}$$

Then,the evolution of $\mathbf{J}_a$ (a=1,2) is governed by the following formula,

$$\frac{d\mathbf{J}_a}{dt} = \frac{G}{c^2 r^3}\left\{\left(2+\frac{3}{2}\frac{M_{a+1}}{M_a}\right)\mathbf{L}_N - \mathbf{J}_{a+1} + 3(\mathbf{n}\cdot\mathbf{J}_{a+1})\,\mathbf{n}\right\}\times\mathbf{J}_a - \frac{3GM_{a+1}}{r^3}(\mathbf{n}\times(\mathbf{I}_a:\mathbf{n})), \tag{17}$$

where the subscript $a+1=3$ means $a=1$. The first term of the right hand side of eq.(17) represents the precession due to the spin-orbit and spin-spin coupling. When the spins and angular momentum are aligned in the same direction, this term vanishes and the direction of the spins is fixed. The second term shows that the magnitude of the spin changes through the tidal torque. The system (16) is conserved one. We add the effect of the gravitational radiation by the lowest quadrupole formula, which is post$^{2.5}$-Newtonian order (e.g., Ref.[15] [16]).

We describe two polarization modes of gravitational waves by h_+ and $h_\times$. They are written by the distance to the source D and the directional angle Θ. Here, we show the case of $\Theta=0$ such as follows.

$$h_+ = \frac{G}{c^4 D}\frac{d^2}{dt^2}\left[\mu r^2\cos 2\phi + \sum_{a=1,2}\{(\mathrm{I}_a)_1-(\mathrm{I}_a)_2\}\cos 2(\phi+\theta_a)\right], \tag{18}$$

$$h_\times = \frac{G}{c^4 D}\frac{d^2}{dt^2}\left[\mu r^2\sin 2\phi + \sum_{a=1,2}\{(\mathrm{I}_a)_1-(\mathrm{I}_a)_2\}\sin 2(\phi+\theta_a)\right], \tag{19}$$

where $(\mathrm{I}_a)_i$ are the principal inertial momentum and θ_a is the lag angle of the ellipsoid a, which is defined by the angle between the major axis of the ellipsoid and the direction to the companion star. Tidally locked state corresponds to $\theta_a=0$.

3 Results

We have calculated the evolution of close binary from a circular orbit with the separation $r\simeq 15R$ to the contact, using two identical affine stars. The most natural orbit near $r\simeq 15R$ is circular one, because the eccentricity is diminished by the gravitational radiation as the binary orbit shrinks. As for the internal structure subject to the tidal force, we adopt the irrotational Roche-Riemann ellipsoids. The solution basically represents tidally locked, static figure in the inertial frame, and additionally includes tidally locked, but spinning figure such as the Maclaurin-Jacobi-Roche solutions (Ref.[17]). The stellar model is set up as $GM/(Rc^2)\simeq 0.2$ with polytropic index $n=0.5$ or $n=1$, corresponding to the typical neutron star of $M\sim 1.4M_\odot$ and $R\sim 10.4$Km. Both stellar spins are assumed to be parallel to the orbital angular momentum, but the initial magnitude of spin is variable. This magnitude are parametlized by $\chi=(|\mathbf{J}_a|/|(\mathrm{I}_a)_3|\,\Omega_o)_{r=15R}$, and we have numerically calculated the subsequent evolution with $\chi=-2,-1,-0.5,0,0.5,1,2$.

3.1 Tidal distortion

From our simulations, the minimum separation distance is $r \sim 2.4R$ for $n = 0.5$, and $r \sim 2.3R$ for $n = 1$. The axes a_1 and a_2 are located on the orbital plane, while a_3 is the direction of the orbital angular momentum. The a_1-axis is initially chosen as the separation direction n of the binary, but the direction does not exactly follow the orbital motion during the dynamical evolution. The orbital revolution at small r becomes so fast that the principal axis slightly deviates from the direction n, i.e., there is a time lag in the angle. However, the lag angle is small, so that the star is elongated almost in the separation direction and compressed in other two directions. The quadrupole deformation becomes significant for $r < 4R$ and grows as $|a_i - R| \propto r^{-3}$. The internal structure also affects the shape of the ellipsoids. The model with $n = 1$ is more centrally condensed than that with $n = 0.5$, so that the shape is less deformed. The difference is, however, a few percent in the magnitude for all χ values.

3.2 Orbital motions

The binary evolution driven by the gravitational radiation is calculated in terms of three approximations to estimate various forces. (1) All post^2-Newtonian, spin effects and tidal effects are included. (2) The post^2-Newtonian effects are included, but the stars are regarded as point masses. (3)Purely Newtonian gravity is used, but the tidal effects is included. Though the spin-spin effects for any χ and spin-orbit force with negative χ are repulsion, but post-Newtonian corrections and tidal force are dominative attraction forces. So that they significantly change the final approaching velocity. In particular, tidal force determines the minimum critical separation distance and therefore is important.

The absolute magnitude of the velocity shifts, when the post-Newtonian corrections are included. The ratio, i.e., the direction of colliding stars, may be more useful than the absolute values as the initial conditions of the hydrodynamical simulation after the coalescence. The approaching velocity increases up to 20 percent of the orbital angular velocity, i.e., the velocity ratio $-V_r/V_\phi > 0.2$ at contact. The ratio further increases with decrease of the polytropic index n, because the tidal force is larger.

When the binary separation is large enough, the binary evolves at much longer time-scale compared with that of the internal motions. The stellar figure is tidally locked at this stage. As the time-scale of the orbital motion decreases, the tidal lock becomes loose and the lag angle becomes non-zero. The lag angles amount to a few degrees just before the contact. This suggests that both stars at the coalescence touch each other not in the longest axis of the ellipsoid, but the direction slightly deviates.

3.3 Gravitational wave forms

The tidal correction is a small correction in the amplitude, but it also causes the phase lag. As the time proceeds, the wave form deviates from that of point masses. The deviation of the $n = 0.5$ model is larger than that of the $n = 1$ model, because of larger tidal deformation. The phase shift becomes significant only for the last a few revolutions, and the amount of the phase shift is less than π. The spin effects also cause the phase shifts, and the amount of the phase shift is more than π.

3.4 Evolution of spin

The stellar spin is modified through the tidal torque, eq.(17). We have calculated the evolution of the spin from various initial states, as seen in Figure 1. The variations of spins are quite small for at large separation, but monotonically grow only in the final stage, $r < 4R$. The stars significantly deviate from the spherical state there, and then tidal torque becomes quite large. The final spin depends on the the initial conditions, but some stars are spun up to almost maximumly rotating state. We show the total change of spins $\Delta J = J_{final} - J_{initial}$ with all χ, in Figure 2.

4 Discussion

The final state of the binary corresponds to $v_\phi \sim 0.3c$, so that the tidal correction seems to be $(0.3)^5 \sim 0.002$, but is much large in the actual problem. The discrepancy comes from the existence of the marginal circular radius, which is the hydrodynamical instability point (Ref.[8]). Two stars significantly accelerate in the radial direction below the critical radius. The tidal effects at the last a few cycles cause large radial infall velocity, lag angle and spin-up. The higher post-Newtonian corrections or fully relativistic hydrodynamical calculation will be required to study much more correct final phase. The topic treated here is preliminary, but will be useful as the guide for the future theoretical work and the observations by the advanced type of the laser interferometer. Our paper is in preparation. [18]

Acknowledgements

This work was supported in part by the Japanese Grant-in-Aid for Scientific Research of the Ministry of Education, Science and Culture (No.05218208, 06740225).

References

[1] A. Abramovici et al., *Science* **256**, 325 (1992).

[2] W. Lincoln & C. Will, *Phys.Rev.D* **42**, 1123 (1990).

[3] W. Junker & G. Schäfer, *MNRAS.* **254**, 146 (1992).

[4] H. Tagoshi & T. Nakamura, *Phys.Rev.D* **49**, 4016 (1994).

[5] H. Tagoshi & M. Sasaki, *Prog.Teor.Phys.* **92**, 745 (1994).

[6] T. Nakamura & K. Oohara, *Prog.Theor.Phys.* **86**, 73 (1991).

[7] T. Nakamura, *in: Relativistic Cosmology*,
eds M. Sasaki, Universal Academy Press,
Inc.–Tokyo,Japan (1994), p.155.

[8] D. Lai, F. A. Rasio & S. L. Shapiro, *ApJ.* **420**, S11 (1994).

[9] D. Lai, F. A. Rasio & S. L. Shapiro, *ApJ.* **423**, 344 (1994).

[10] D. Lai, F. A. Rasio & S. L. Shapiro, *ApJ.* , in press (1995).

[11] T. Damour & G. Schafer, *IL NUOVO CIMENTO* **101B**, 127 (1988).

[12] B. Carter & J. P. Luminet, *MNRAS.* **212**, 23 (1985).

[13] J. P. Luminet & B. Carter, *ApJS.* **61**, 219 (1986).

[14] S. Chandrasekhar, *Ellipsoidal Figures of Equilibrium*
Yale University Press, New Haven, in USA. (1969).

[15] G. Schafer, *ANNALS OF PHYSICS* **161**, 81 (1985).

[16] K. D. Kokkotas & G. Schafer, *MNRAS* **275**, 301 (1995).

[17] C. S. Kochanek, *ApJ.* **398**, 234 (1992).

[18] W. Ogawaguchi & Y. Kjima, in preparation (1996).

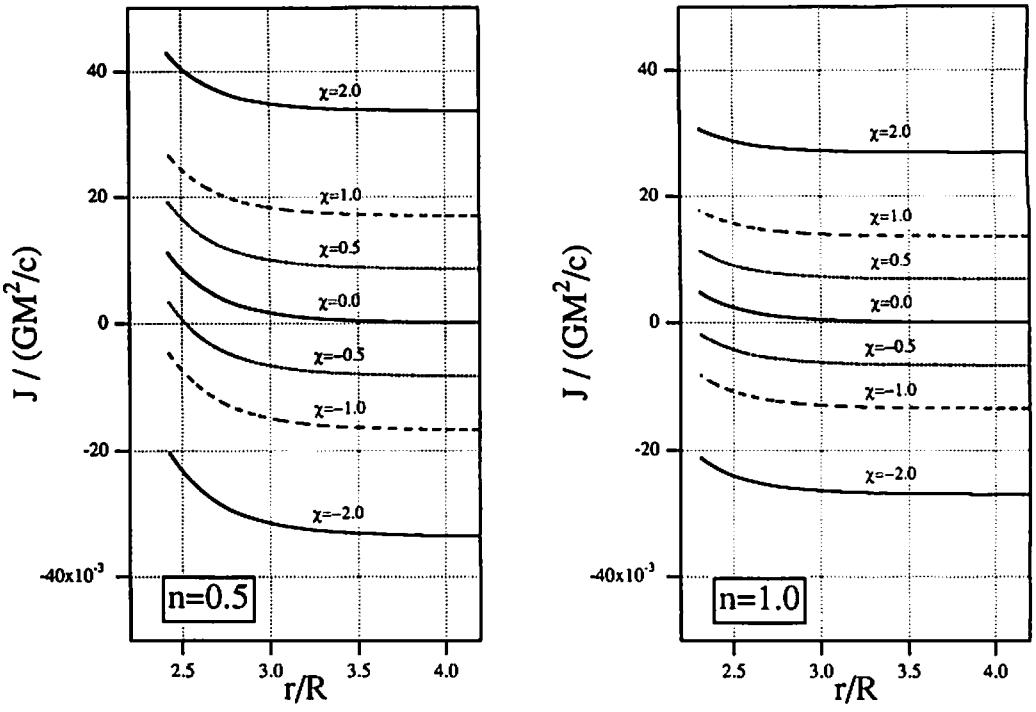


Figure 1: Variation of stellar spin for all χ values.

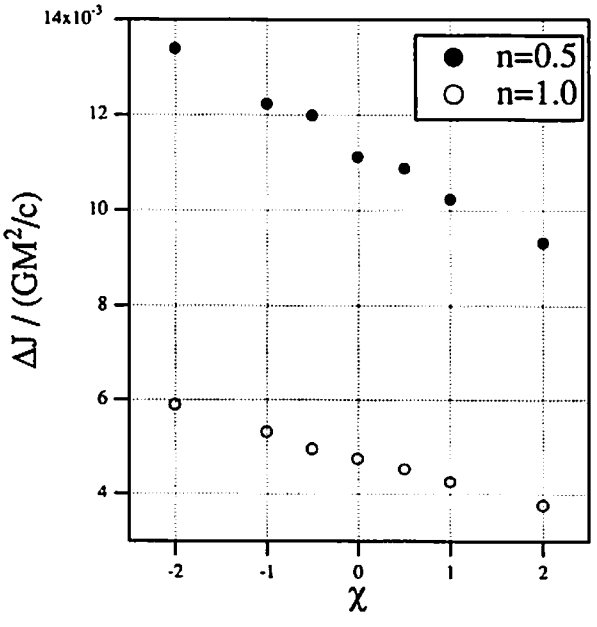


Figure 2: Total spin-up vs χ.

February 16, 1996

On the stability of synchronized binary neutron stars in post-Newtonian approximation of general relativity

MASARU SHIBATA

Department of Earth and Space Science, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan

ABSTRACT

We investigate stability property of binary neutron stars(BNSs) just before the merging in the post-Newtonian(PN) approximation. Stability analysis is performed making use of numerical solutions of the equilibrium configuration of synchronized BNSs which are obtained by means of the numerical scheme developed in a previous paper. NSs are modeled by means of the polytropic equation of state with the polytropic exponent $\Gamma = 2$ or 3. From numerical calculations, we find that like in the Newtonian case, in the PN approximation, the secular instability will occur for synchronized BNSs at a critical angular velocity Ω_{crit} before the surfaces of each star of BNSs come into contact. It is also found that Ω_{crit} in the PN approximation is $\sim 10 - 15\%$ larger than that in the Newtonian case for a realistic NS of mass $M \sim 1.4 M_\odot$ and radius $R \sim 10 - 20$km. Implication of this property to the orbital evolution of BNSs just before merging is discussed.

1. Introduction

The last stage of coalescing binary neutron stars(BNSs), which emits gravitational waves of frequency $10 - 1000$Hz, is one of the most promising sources of gravitational waves for kilo-meter size laser interferometric detectors such as LIGO[1]and VIRGO[2]. Evolutions of these compact BNSs are as follows: When the orbital separation of BNSs is sufficiently large compared with the NS radius, BNSs evolve in the adiabatic manner radiating gravitational waves in much longer time scale than the orbital period because the general relativistic(GR) gravity is not so strong. In such an inspiraling phase, they can also be regarded approximately as two point masses because the hydrodynamic effect is not important[3]. On the other hand, when the orbital separation of BNSs becomes comparable to the NS radius, the hydrodynamic effect becomes important for the evolution of BNSs, and also the GR gravity becomes very strong. In such a merging phase, each NS of binaries does not behave as a point mass, and also the binary evolves not in the adiabatic manner, but in the dynamical manner to merge.

These evolution scenarios mean that the nature of the signal of gravitational waves changes around a transition region between the inspiraling and merging phases. Gravitational waves emitted at this transition region have an important information about the NS radius[4,5] which can be used to determine the equation of state(EOS) of the NS matter[6]. Thus, investigations of gravitational waves in the merging phase, in particular, around the transition region from the inspiraling phase to the merging phase are very important. The purpose of this paper is to investigate the transition region incorporating the GR effect as well as the hydrodynamic one.

As a trigger of the transition, recently, the tidal effect was pointed out by Lai et al.[7,8]: They showed that when the orbital separation of BNSs becomes small, each star of binary is significantly deformed by the tidal force of the companion star. In such a case, the tidal field due to the deformation of each star is generated[9]and as a result, the circular orbit of BNSs becomes unstable.

In the case when we consider the tidal effect in the binary system, we need to know the structure of each star because the tidal effect to each star depends on its structure. Thus, when we consider the tidal effect on BNSs, we need to take into account the GR effects because the GR gravity plays an important role to determine the structure of NS. This means that in order to know whether the tidal effect is important or not, we must investigate the evolution of BNSs just before the merging taking into account not only the hydrodynamic effect, but also the effect of the GR gravity to each star of binary. For this purpose, we perform the PN calculation to obtain equilibrium solution for *synchronized* BNSs in circular orbits, and investigate the stability property of BNSs.

As pointed out by Kochanek or Bildsten and Cutler[10], the synchronized binaries are realized only when the viscosity of the NS matter is extremely large. Hence, to know the stability property for realistic BNSs, we need to investigate BNSs in non-synchronized orbits

taking into account the case where the circulation of the system conserves. However, we do not have a numerical method to obtain the equilibrium configuration of non-synchronized BNSs accurately, in contrast with the case of synchronized BNSs. Fortunately, the stability property of the non-synchronized binary seems to be similar to that of the synchronized one in the Newtonian case as shown by Lai et al[7,8]. Therefore, investigation of the stability property of synchronized BNSs would be a guideline of that for the other types of BNSs.

The paper is organized as follows. In section 2, we review the basic equations to obtain the PN configuration of uniformly rotating bodies, and define the various quantities in the first PN approximation. In section 3, we show numerical results for synchronized BNSs in circular orbits. Paying attention to the angular momentum and energy of the binary systems as a function of the orbital separation, we consider the stability of BNSs just before the merging, and the GR effect to it. Throughout of this paper, c and G, respectively, denote the speed of light and the gravitational constant.

2. Basic equations

As shown in a previous paper[12], in the first PN approximation, the equilibrium configuration of a uniformly rotating body of a polytropic equation of state(EOS),

$$P = (\Gamma - 1)\rho\varepsilon = K\rho^{\Gamma}, \tag{2.1}$$

is obtained from the following sets of equations[11,12];

$$\frac{K\Gamma}{\Gamma-1}\rho^{\Gamma-1} - \frac{1}{2c^2}\Big(\frac{K\Gamma}{\Gamma-1}\rho^{\Gamma-1}\Big)^2 = U - \frac{X_0}{c^2} + \Big\{\frac{R^2}{2} + \frac{1}{c^2}\Big(2R^2U - X_\Omega + \hat{\beta}_\varphi\Big)\Big\}\Omega^2 + \frac{R^4}{4c^2}\Omega^4 + \text{Const.}, \tag{2.2}$$

and

$$\Delta U = -4\pi G\rho, \tag{2.3}$$

$$\Delta\hat{P}_1 = -4\pi G\rho x\ , \tag{2.4}$$

$$\Delta\hat{P}_2 = -4\pi G\rho y\ , \tag{2.5}$$

$$\Delta X_0 = 4\pi G\rho\Big(\varepsilon + 2U + \frac{3P}{\rho}\Big), \tag{2.6}$$

$$\Delta X_\Omega = 8\pi G\rho R^2, \tag{2.7}$$

where ρ, ε, P, K, Γ and Ω denote the mass density, the specific internal energy, the pressure, the polytropic constant, the polytropic exponent and the angular velocity($d\varphi/dt$), respectively. R^2 denotes $x^2 + y^2$, and $\hat{\beta}_\varphi$ becomes

$$\hat{\beta}_\varphi = -\Big[\frac{7}{2}\Big(x\hat{P}_1 + y\hat{P}_2\Big) + \frac{1}{2}\Big(x^2\hat{P}_{2,y} + y^2\hat{P}_{1,x} - xy\hat{P}_{1,y} - xy\hat{P}_{2,x}\Big)\Big]. \tag{2.8}$$

In this paper, we will obtain the equilibrium configuration of BNSs of equal masses. So that, the numerical method to obtain them is the same as that in a previous paper[12].

In the PN approximation, we have the following conserved equantities;

(1)The conseved mass

$$M_* = \int \rho^* d^3x, \tag{2.9}$$

where ρ^* is the conserved mass density, which is defined as

$$\rho^* = \rho\Big\{1 + \frac{1}{c^2}\Big(\frac{v^2}{2} + 3U\Big)\Big\}. \tag{2.10}$$

(2)The ADM mass

$$M_{ADM} = \int \rho\Big\{1 + \frac{1}{c^2}\Big(v^2 + \frac{5}{2}U + \varepsilon\Big)\Big\}d^3x. \tag{2.11}$$

(3)The total energy[13]

$$E = \int \rho\Big\{\varepsilon + \frac{v^2}{2} - \frac{1}{2}U + \frac{1}{c^2}\Big(\frac{5}{8}v^4 + \frac{5}{2}v^2U + \frac{1}{2}\beta_i v^i + \varepsilon v^2 + \frac{P}{\rho}v^2 + 2\varepsilon U - \frac{5}{2}U^2\Big)\Big\}d^3x. \tag{2.12}$$

(4) The total angular momentum[13]

$$J = \int \rho\Big[v_\varphi\Big\{1 + \frac{1}{c^2}\Big(v^2 + 6U + \varepsilon + \frac{P}{\rho}\Big)\Big\} + \frac{\beta_\varphi}{c^2}\Big]d^3x. \tag{2.13}$$

M_* conserves throughout the whole evolution of the system even if there exists a dissipation process such as the emission of gravitational waves. Thus, in the case when we consider a sequence of a constant M_*, it may be regarded as an evolution sequence of the system. We note that E is derived from $M_{ADM} - M_*$ in the second PN calculation. In the following, we use these quantities to argue the stability property of BNSs. For the sake of convenience, we also define the position of the mass center for each star of binary as

$$x_g^i = \frac{1}{M_*}\int \rho^* x^i d^3x. \tag{2.14}$$

It should be noted that the above expressions of M_{ADM}, E and J are not unique expressions for the ADM mass, energy and angular momentum in the first PN approximation. For example, these may be written as

$$\begin{aligned}
M'_{ADM} &= \int \rho^*\Big\{1 + \frac{1}{c^2}\Big(\frac{v^2}{2} - \frac{1}{2}U + \varepsilon\Big)\Big\}d^3x = M_{ADM} + O(c^{-4}),\\
E' &= \int \rho^*\Big\{\varepsilon + \frac{v^2}{2} - \frac{1}{2}U + \frac{1}{c^2}\Big(\frac{3}{8}v^4 + \frac{5}{4}v^2U + \frac{1}{2}\beta_i v^i + \frac{1}{2}\varepsilon v^2 + \frac{P}{\rho}v^2 - \varepsilon U - U^2\Big)\Big\}d^3x\\
&= E + O(c^{-4}),\\
J' &= \int \rho^*\Big[v_\varphi\Big\{1 + \frac{1}{c^2}\Big(\frac{v^2}{2} + 3U + \varepsilon + \frac{P}{\rho}\Big)\Big\} + \frac{\beta_\varphi}{c^2}\Big]d^3x = J + O(c^{-4}).
\end{aligned} \tag{2.15}$$

In the following section, we use Eqs.(2.12) and (2.13) instead of Eqs.(2.15), and substitute

the numerical results of ρ, U, and so on, directly into Eqs.(2.12) and (2.13) to estimate E and J. In this case, numerical values of E and J depend on their definitions because the higher order PN terms are unavoidably involved in unexpected forms. Thus, E and J defined above are the conserved quantities only in the limit, $v^2/c^2, U/c^2, \varepsilon/c^2 \ll 1$, and otherwise they are not.

It is also appropriate to note that a solution obtained from Eq.(2.2) is the good approximation of a GR solution only in the case when the PN correction is small. This is because we solve Eq.(2.2) without using the PN expansion in contrast with the treatment by Chandrasekhar[11], and as a result, the solution involves not only correct first PN terms, but also extra higher PN terms, which are not correct terms in general. Thus, the solutions obtained from Eq.(2.2) by the above procedure are the correct approximate solutions of a GR solution only in the limiting case, $v^2/c^2, U/c^2, \varepsilon/c^2 \ll 1$. In other cases, it is not the approximate solution in the strict sense.

Finally, we define the physical units for the sake of convenience as

$$R_s \equiv \frac{GM_\odot}{c^2} = 1.477\text{km} \quad \text{and} \quad \rho_s \equiv \frac{M_\odot}{R_s^3} = 6.173 \times 10^{17}\text{g/cm}^3, \tag{2.16}$$

where $M_\odot = 1.989 \times 10^{33}$g is the solar mass.

3. Numerical Results

In this paper, we choose $\Gamma = 2$ or 3 as the polytropic exponent. According to models of the EOS for NSs, Γ becomes $2 \lesssim \Gamma \lesssim 3$[16], so that this assumption is reasonable. In each case, we define the polytropic constant as

$$K = \begin{cases} 2\pi^{-1} G r_2^2 & \text{for } \Gamma = 2, \\ 2.524 G M_\odot^{-1} r_3^5 & \text{for } \Gamma = 3, \end{cases} \tag{3.1}$$

where r_2 and r_3 are parameters. If we consider a spherical star in the Newtonian theory, r_2 and $a \equiv r_3 (M/M_\odot)^{1/5}$, where M is the mass of the spherical star, become radii of polytropic stars of the polytrope exponents $\Gamma = 2$ and 3, respectively.

Since in a previous paper[12], we showed various numerical results of BNSs of $\Gamma = 2$, we mainly present numerical results for $\Gamma = 3$ polytrope in this paper. First of all, in figs.1, we show the total energy and angular momentum as a function of the orbital separation r_g/a, where r_g denotes the distance between two center of masses, for the Newtonian configuration. In the figures, E and J are shown in units of $E/(GM^2/4a)$ and $J/(G(M/2)^{3/2}a^{1/2})$. Innermost circles(i.e., the circles of the minimum r_g/a) denote the case in which the surfaces of each star of BNSs come into contact. Like in the case of $\Gamma = 2$[14,12], it is found that the minimums

of E and J appear at a critical point r_{crit}. As pointed out by Lai et al.[7,8], this indicates that the secular instability occurs at r_{crit}. Since this is not the dynamical instability point, it does not mean that r_{crit} is the radius of the innermost stable orbit(ISCO), and it only means that BNSs cannot maintain uniformly rotating orbits for $r < r_{crit}$. However, as shown by Lai et al.[7,8], the dynamical instability point will always appear very near the secular instability one, so that we may expect that the ISCO locates at $\lesssim r_{crit}$.

Then, we discuss about the stability property of the PN configurations. First of all, in Table 1, we show the several quantities of spherical stars in the PN approximation: r_A is the areal radius which is defined as

$$r_A = r_{cood}\{1 + U(r_{cood})\}, \tag{3.2}$$

where r_{cood} is the coordinate radius. From Table 1, we can find that the PN effect makes the radius of spherical stars small, and this effect is stronger for smaller r_3(i.e., softer EOS). This property plays an important role in the following discussions. In fig.2, we show relations between J and r_g/a_* for BNSs of mass $M_* = 1.4M_\odot$ and $r_3/R_s = 8, 10, 15$ and 40(filled circles, open circles, filled squares and open squares) in the PN cases as well as the Newtonian sequence(dotted circles). Note that the Newtonian sequence may be regarded as the case $r_3/R_s \to \infty$ in the PN approximation. Here, $r_g = 2|x_g^i|$(x_g^i is defined in Eq.(2.14)) and $a_* \equiv r_3(M_*/M_\odot)^{1/5}$. E and J are shown in units of $E/(GM_*^2/4a_*)$ and $J/(G(M_*/2)^{3/2}a_*^{1/2})$. Fig.2 shows that the minimums exist at some critical separations, r_{crit}. Like in the case of $\Gamma = 2$[12], r_{crit} for the PN configurations is $\sim 10-20\%$ smaller than that in the Newtonian case for realistic models of NSs. The reason is mainly that in the PN approximation, the radius of each star of BNSs becomes small due to the PN correction compared with the Newtonian case. From Table 1, we find that r_{cood}/a is $\sim 85\%$ for $r_3 = 8R_s$ and $\sim 90\%$ for $r_3 = 15R_s$. This is consistent with the results shown in fig.2.

Let us investigate the influence of this property to Ω_{crit}. In fig.3(a), we show the relation between $f \equiv \Omega_{crit}/\pi$, which is the frequency of gravitaitonal waves at r_{crit}, and $r_A c^2/GM_{ADM}$, which denotes the compactness of spherical stars, for BNSs of $M_* = 2.8M_\odot$. Filled circles and squares denote f for $\Gamma = 3$ and 2 in the PN approximation, respectively, and dotted and dashed lines are those in the Newtonian case for $\Gamma = 3$ and 2, respectively. Note that in the Newtonian cases, Ω_{crit} scales as $r^{-3/2}$. In fig.3(b), we also show $f = \Omega_{crit}/\pi$ as a function of r_3/R_s for the case of $\Gamma = 3$. It is soon found that when we compare Ω_{crit} in the PN approximation with that in the Newtonian case fixing the compactness of NSs or the EOS, the former is always larger than the latter, and for realistic models of $5 \lesssim r_A c^2/GM_{ADM} \lesssim 10$, the deviation becomes about $\gtrsim$ 100Hz. Let us argue the reason for this property in the following.

In the first PN approximation, the angular frequency of two point masses in circular

orbits becomes[3]

$$\Omega = \sqrt{\frac{GM}{r^3}}\left(1 - \frac{3-\eta}{2}\frac{GM}{rc^2}\right), \tag{3.3}$$

where M and η are the total mass of system and the ratio of the reduced mass to the total mass, respectively. In this paper, $\eta = 1/4$. Thus, we may write Ω_{crit} as

$$\Omega_{crit} = \sqrt{\frac{GM}{r_{crit}^3}}\left(1 - \frac{11}{8}\frac{GM}{r_{crit}c^2}\right) + \delta\Omega, \tag{3.4}$$

where $\delta\Omega$ denotes corrections due to the tidal filed, the spin angular momentum of each star, and so on[7]. In the following, we regard it as a small quantity because in the Newtonian case, $\delta\Omega/\Omega_{crit}$ is at most a few %.

Since r_{crit} in the PN approximation is smaller than that in the Newtonian case, we write it as $r_{crit} = r_{crit,N}(1-\delta)$, where $r_{crit,N}$ denotes r_{crit} for a Newtonian sequence and δ is a constant $\sim 0.1 - 0.2$. Using this expression, Eq.(3.4) is rewritten as

$$\Omega_{crit} \simeq \Omega_{crit,N}\left(1 + \frac{3}{2}\delta - \frac{11}{8}\frac{GM}{r_{crit}c^2}\right) + \delta\Omega, \tag{3.5}$$

where $\Omega_{crit,N}$ is Ω_{crit} for a Newtonian sequence. Present numerical calculations show that the second term in the bracket is always larger than the third term. If we assume that $\delta\Omega$ in the PN approximation is small corrections like in the Newtonian case and it may be safely neglected, Ω_{crit}/π is always larger than $\Omega_{crit,N}/\pi$: For $5 \lesssim r_A c^2/GM_{ADM} \lesssim 10$, Ω_{crit} is $\sim 10 - 15\%$ larger than $\Omega_{crit,N}$. Although this argument is very rough, this seems to explain the quantitative difference $\Omega_{crit} - \Omega_{crit,N}$ essentially.

Finally, we briefly discuss the implication of this result to the location of the ISCO. As mentioned above, r_{crit} denotes the point where the secular instability occurs, so it does not denote the point of the ISCO. The ISCO will be determined by the dynamical instability point. However, the above simple argument for the frequency shift due to the PN correction seems to hold for the frequency at which the dynamical instability occurs, f_{dyn}. Thus, we may conjecture that the PN correction will increase f_{dyn}.

REFERENCES

1. A. Abramovici et al. Science, **256**(1992), 325;
   K. S. Thorne, in Proceedings of the 8-th Nishinomiya-Yukawa Memorial Symposium on Relativistic Cosmology, edited by M. Sasaki(Universal Academy Press, Tokyo, 1994), p67.
2. C. Bradaschia et al., Nucl. Instrum. Method Phys. Res. Sect. **A289**, 518(1990).
3. C. M. Will, in Proceedings of the 8-th Nishinomiya-Yukawa Memorial Symposium on Relativistic Cosmology(ref.1), p83.
4. C. Cutler et al., Phys. Rev. Lett. **70**, 2984(1993).
5. X. Zhuge, J. M. Centrella, and S. L. W. McMillan, Phys. Rev. **D50**, 6247(1994).
6. L. Lindblom, Astrophys. J. **398**, 569(1992).
7. D. Lai, F. A. Rasio and S. L. Shapiro, Astrophys. J. suppl. **88**, 205(1993).
8. D. Lai, F. A. Rasio and S. L. Shapiro, Astrophys. J. **420**, 811(1994); Astrophys. J. Lett. **406**, L63(1993).
9. S. Chandrasekhar, *Ellipsoidal Figures of Equilibrium*(Dover, 1969).
10. C. S. Kochanek, Astrophys. J. **398**, 234(1992):
    L. Bildsten and C. Cutler, Astrophys. J. **400**, 175(1992).
11. S. Chandrasekhar, Astrophys. J. **148**, 621(1967).
12. M. Shibata, Osaka University Preprint, OU-TAP 26, submitted to Phys. Rev. D.
13. S. Chandrasekhar, Astrophys. J. **142**, 1488(1965).
14. F. A. Rasio and S. L. Shapiro, Astrophys. J. **401**, 226(1992); **432**, 242(1994).
15. K. Murata, M. Natori and Y. Karaki, *Large-scale Numerical Simultion*, (Iwanami, 1990, in Japanese).
16. For example, S. L. Shapiro and S. A. Teukolsky, *Black Holes, White Dwarfs and Neutron Stars*, (NY, Wiley, 1983).

Table 1

Various quantities of spherical stars of mass $M_* = 1.4M_\odot$ for $\Gamma = 3$ polytrope with different polytropic exponents(i.e., different r_3)in the PN approximation. r_A and r_{cood} are the areal radius and the coordinate radius, respectively.

r_3/R_s	8	9	10	12	15	20
r_A/R_s	8.035	9.087	10.14	12.27	15.45	20.75
r_{cood}/R_s	7.115	8.135	9.170	11.23	14.37	19.61
$M_{ADM}/M_\odot$	1.343	1.346	1.348	1.353	1.358	1.365
$r_A c^2/GM_{ADM}$	5.982	6.752	7.523	9.068	11.37	15.20

Table 2

Critical angular velocities(Ω_{crit}/π) and separations(r_{crit}/a_*) for $\Gamma = 3$ polytrope with different polytropic exponents(i.e., different r_3)in the PN approximation. Ω_{crit}/π is shown in units of Hz. Note that each value involves an error of order 10^{-2}.

r_3/R_s	8	9	10	12	15	20	40
Ω_{crit}/π	970	798	675	504	350	222	77
r_{crit}/a_*	2.73	2.78	2.81	2.87	2.95	3.02	3.16

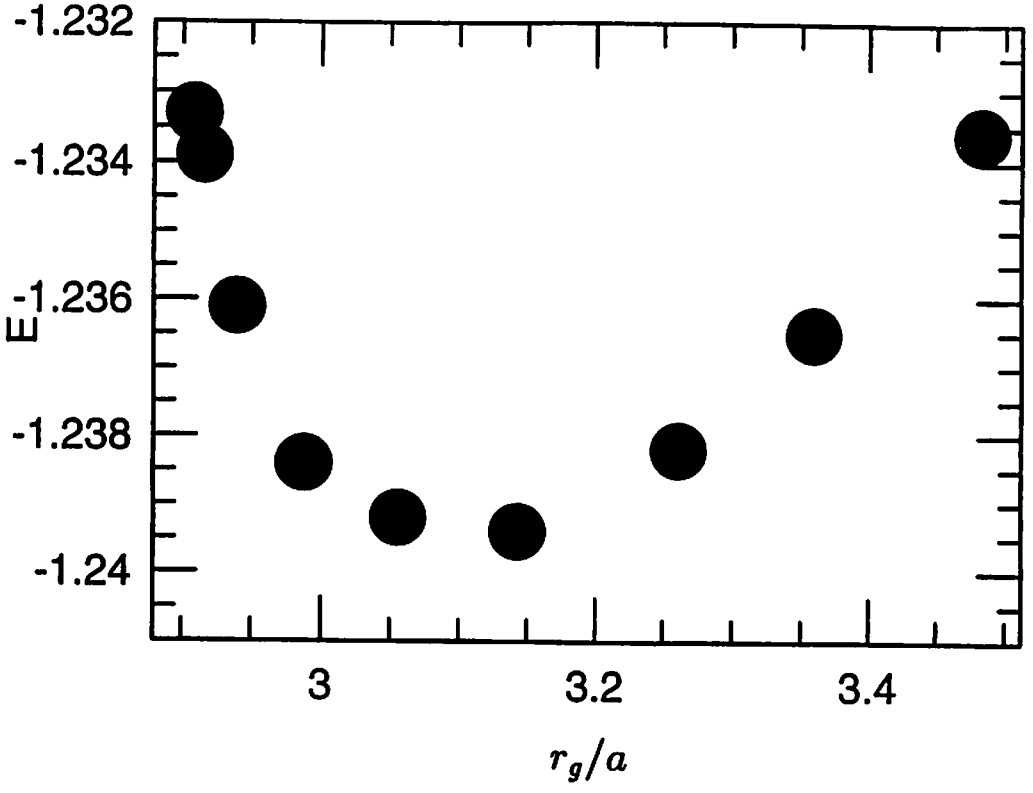
Figure captions

Figs.1 The energy(E) and the angular momentum(J) as a function of r_g/a for BNSs of $\Gamma = 3$ in the Newtonian case. The energy and the angular momentum are shown in units of $E/(GM^2/4a)$ and $J/(G(M/2)^{3/2}a^{1/2})$. Here, M and a are the total mass of the system and $r_3(M/2M_\odot)^{1/5}$, respectively.

Fig.2 The angular momentum as a function of r_g/a_* for BNSs of $\Gamma = 3$ and several EOSs in the PN approximation. The angular momentum is shown in units of $J/(G(M_*/2)^{3/2}a_*^{1/2})$, where $M_* = 2.8M_\odot$ and $a_* = r_3(M_*/2M_\odot)^{1/5}$, respectively. Filled circles, open circles, filled squares, and open squares denote the PN sequence of $r_3/R_s = 8$, 10, 15, and 40, respectively. Dotted circles denote the Newtonian sequence, which may be regarded as the case $r_3/R_s \to \infty$.

Figs.3(a) The frequency of gravitational waves at a critical separation r_{crit}, $f = \Omega_{crit}/\pi$, as a function of the compactness of the spherical NS, $r_A c^2/GM_{ADM}$. Filled circles and squares denote f of the PN sequences for $\Gamma = 3$ and 2, respectively. Dotted and dashed lines show f of the Newtonian sequences for $\Gamma = 3$ and 2, respectively.

Figs.3(b) $f = \Omega_{crit}/\pi$ as a function of r_3/R_s. Filled circles and dotted line denote f of the PN and Newtonian sequences, respectively.



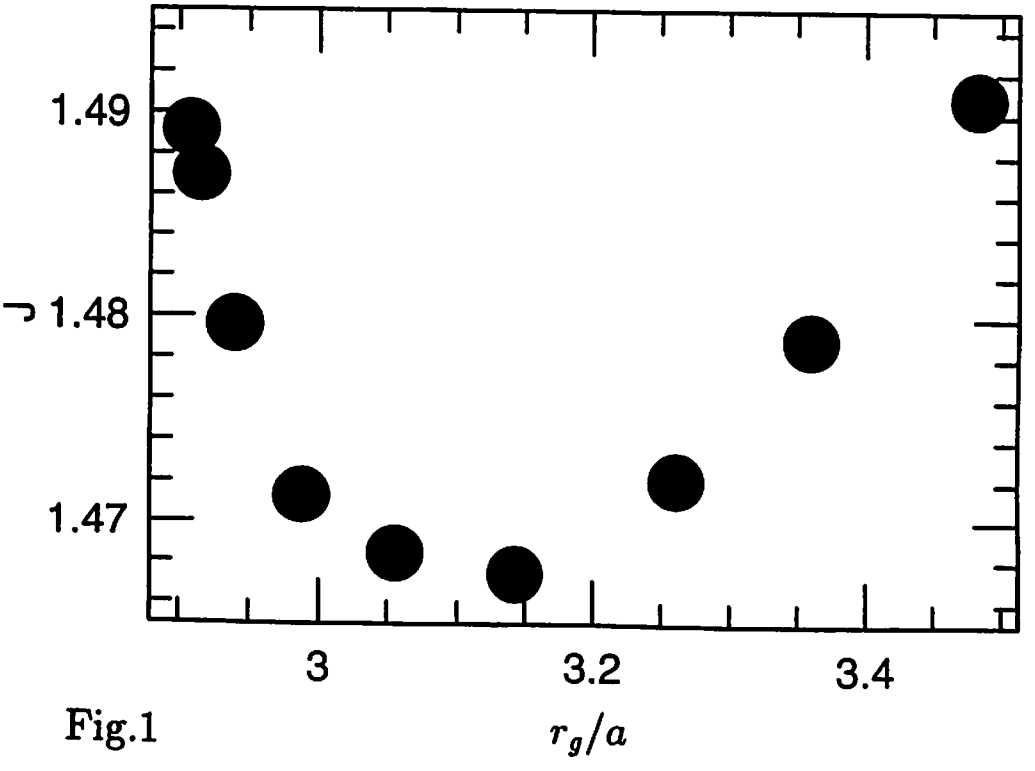


Fig.1

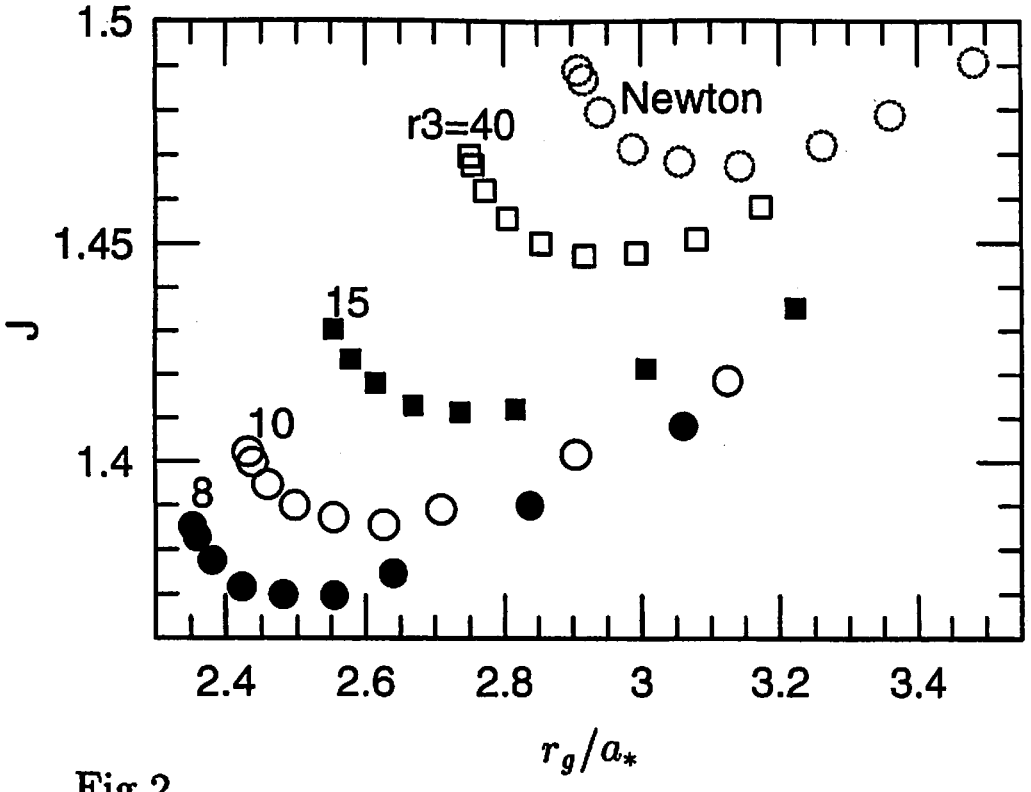


Fig.2

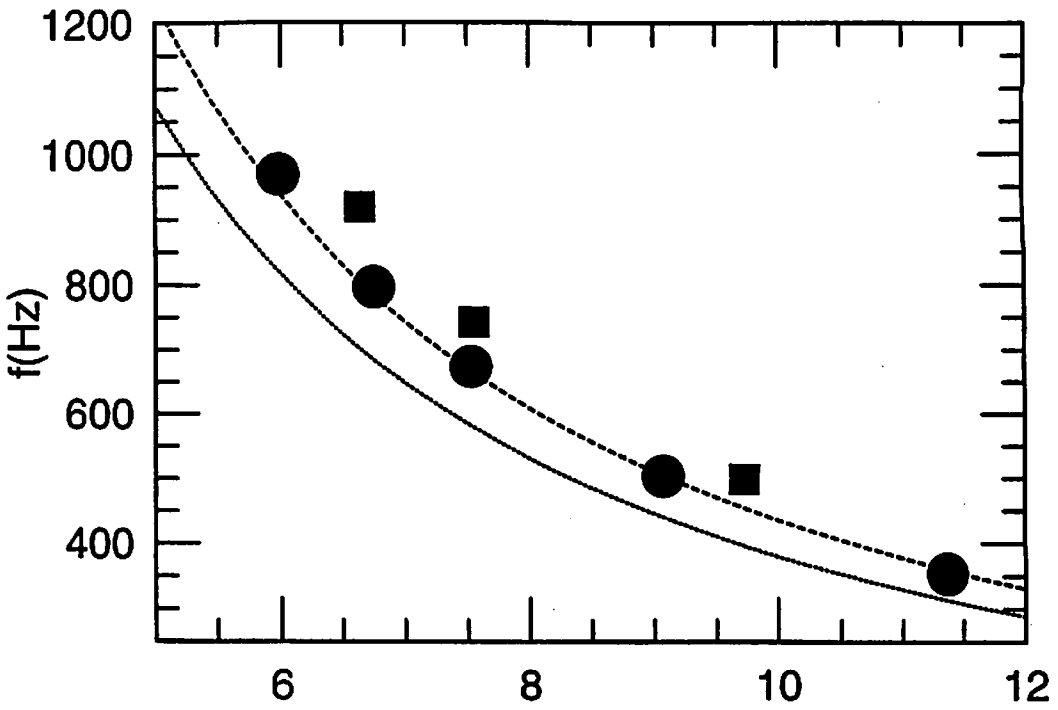

1200
1000
800
600
400
f(Hz)
6
8
10
12
$r_{AC}c^2/GM_{ADM}$

Fig.3(a)

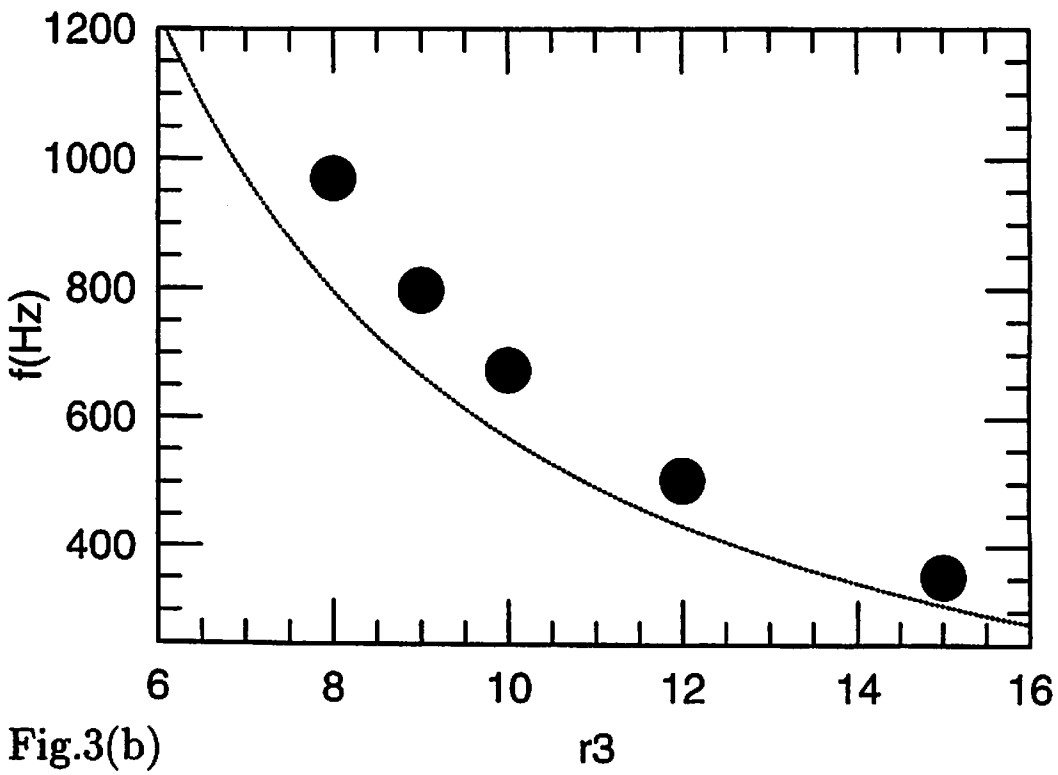

1200
1000
800
600
400
f(Hz)
6
8
10
12
14
16
r3

Fig.3(b)

Black Hole Thermodynamics

Osamu Kaburaki

Astronomical Institute, Graduate School of Science, Tohoku University,

Aoba, Sendai 980-77, Japan

The emergence and development of the black hole thermodynamics are briefly reviewed. As a selected topic of recent interest in this field, the critical behavior of Kerr-Newman holes is discussed in some detail. The scaling laws of the "first kind" are all satisfied at the extremal limit. The scaling laws of the "second kind" suggest a long-time correlation among the evaporating particles.

1. Emergence of Black Hole Thermodynamics (Introduction)

The analogy between the area of a black hole and entropy was first taken seriously by Bekenstein [1]. According to the area theorem by Hawking [2], the total area of black holes in a system under consideration does not decrease in time. This is reminiscent of the second law of thermodynamics: the entropy of an isolated system never decreases. Based on various considerations on the reasonability of regarding this analogy as physically meaningful, Bekenstein proposed a proportionality between the two quantities. He also derived the corresponding first law of black hole thermodynamics from the expression of the area A of a black hole in terms of its mass M, angular momentum J (the norm of the vector J and hence is positive or zero) and charge Q (this can be positive, zero or negative):

$$dM = \frac{\kappa}{8\pi} dA + \Omega dJ + \Phi dQ, \tag{1}$$

where κ, Ω and Φ are the surface gravity, angular velocity and electrostatic potential, respectively, of the hole. A similar equation was also derived through more sophisticated technics in the black hole mechanics [3].

From this equation, it turns out that if a black hole has entropy proportional to its area then it has a temperature proportional to its surface gravity. There seemed, however,

that a serious problem existed in this interpretation. If the hole is a thermal body whose temperature is specified by its surface gravity then it must emit thermal radiation, but this is impossible since nothing can come out from inside of the horizon. This dilemma was resolved when Hawking discovered [4] through his quantum field theoretical calculation that there are particles evaporating from the black hole. The distributions of such particles are characterized by a temperature

$$T = \frac{\hbar}{2\pi k_B c}\,\kappa, \tag{2}$$

now called the Bekenstein-Hawking temperature, where c is the light velocity, and $\hbar$ and k_B are Planck's and Boltzmann's constants, respectively. Then the relation between the area and entropy is determined as

$$S = \left(\frac{k_B c^3}{4\hbar G}\right) A = \frac{k_B}{4}\frac{A}{l_P^2}, \tag{3}$$

where G is Newton's gravitational constant and l_P is the Planck length.

The temperature given above becomes zero for extremal black holes, i.e. when $a^2 + Q^2 = M^2$ with $a = J/M$. There are many circumstantial evidence for the unattainability of this state by any finite series of quasi-stationary processes. Although this fact suggests the validity of the third law also in the black hole case, there is no rigorous proof so far. The third "law" may be regarded as a restricted version of the cosmic censorship hypothesis [5] which inhibits the production of naked singularity in our universe by any process. The readers are referred to the review papers [6] for these early stages of the development of black hole thermodynamics.

Black hole thermodynamics is essentially a standpoint of distant observers. The Bekenstein-Hawking temperature is proportional to the redshifted surface gravity measured by a distant observer but it is different from the locally measured temperature at an arbitrary distance from a black hole. The "membrane paradigm" [7] provides a powerful conceptual foundation in dealing with a black hole from the standpoint of such an observer. By slightly stretching the horizon of a hole one can ignore the complicated past history of infalling matter which is piling up just above the horizon owing to the time dilatation, as well as the interior of the horizon. Thus the stretched horizon forms a boundary of the spacetime relevant to the distant observer. Surprisingly enough, careful examinations of the boundary conditions on such a stretched horizon show that, in

the 3 + 1 decomposition of the spacetime by a distant observer, the boundary behaves like a 2-dimensional membrane with various physical properties such as viscosity, electric resistivity and surface pressure. All the thermodynamic quantities can be attributed consistently also to the membrane.

A serious objection to the above membrane viewpoint may be that for a freely falling observer there is no object like a membrane and the spacetime extends all the way to the central singularity. Therefore, the former viewpoint could be misleading or at least incomplete. However, a recently proposed concept of "black hole complimentarity" [8, 9] seems to overcome the above objection and provide a powerful support to the membrane paradigm. This asserts that the above viewpoints are both correct and are in fact complementary to each other in the sense used in quantum mechanics. Actually, the impossibility of communicating each other what they had observed has recently been demonstrated [10] confirming the validity of this idea.

2. Development of Black Hole Thermodynamics

Nowadays it is very difficult to look over the whole range of black hole thermodynamics and give a comprehensive summary or review, since the related topics spread over many branches of black hole physics. Another source of the difficulties is how far should we include the topics under the name of black hole thermodynamics. For these reasons, only a superficial introduction is given of the recent developments in this field. Interested readers are recommended to consult the original papers, though the references given here are also incomplete. Some review papers dealing with somewhat different aspects of the black hole thermodynamics can be found in the proceedings cited as ref. [11].

A. Equilibrium States and Their Stabilities

1) Formulation of equilibrium states

The formulation of the black hole thermodynamics based on the Euclidian actions was proposed by Gibbons and Hawking [12] in which the partition function was calculated from the saddle point contribution. This method was later used extensively by York and his collaborators [13] in discussing the formulation and stability of the black hole solutions as thermodynamic states. They specified the relations of the reduced action to

various thermodynamic potentials and showed the equivalence of thermodynamic stability to mechanical one.

Although the above works are restricted to the case of Kerr-Newman holes (almost to Schwarzschild holes), the black hole solutions of Einstein's equation coupled to various fields other than the electromagnetic field were also found. They are the dilaton [14], Yang-Mills [15], Skyrme [16], Proca [17] fields and so on. Except the delaton case, there appear the "exotic hairs" in addition to the Kerr-Newman hairs. The thermodynamic expressions of these solutions are possible as expected. It was pointed out in a series of solutions coupled with modified dilaton fields [18] that the extremal Kerr-Newman holes shows a peculiarity. Indeed, in the limit of extremal holes entropy tends to zero but temperature remains finite for non-Kerr-Newman cases (specified by a continuous parameter $a \neq 1$) while temperature tends to zero but entropy remains finite in the Kerr-Newman case ($a = 1$). As another kind of hairs, the possibility of the "quantum hair" on black holes due to quantum interference was discussed in ref. [19].

2) Stability

The stability of Kerr-Newman states were systematically examined under various thermodynamic environments [20], through the turning point method. A change of stability occurs at a turning point where the tangent of the curve of an equilibrium series plotted in a two-dimensional plane of a conjugate pair of thermodynamic variables changes sign through $\pm\infty$. The essential result of these stability analyses is that the stability of the same equilibrium series can be different under different thermodynamic environments. The independence of different modes associated with different hairs were demonstrated in the Kerr-Newman case [21]. The general expressions for the variances of fluctuating variables were given in the same paper. These results can be extended to the cases of arbitrary number of hairs as far as the equilibrium states are specified only by such hairs or by an equivalent set of the same number of parameters.

The stability analyses from a viewpoint of catastrophe theory were also developed. The correspondence of a turning point to a cusp-type catastrophe was pointed out in ref. [21]. The extensive applications of higher-dimensional catastrophes to the stability problems can be seen in the cases of exotic black holes where the equilibrium states need other parameters than their hairs to be specified [22].

B. Extremal Kerr-Newman Holes and Critical Phenomenon

1) Limit of thermodynamic description

The heat capacity of a Kerr-Newman black holes at constant angular momentum and charge, C_{JQ}, becomes zero in the limit of extremality. This means that the temperature rises (or decreases) infinitely associated with the absorption (or emission) of infinitesimal amount of heat and therefore the break down of thermodynamic description of the process [18]. Reflecting this fact, the variance of the temperature fluctuation diverges, i.e. $< (\Delta T)^2 > /T = k_B/C_{JQ} \to \infty$. The break down of thermodynamics in this limit is, in fact, due to the presence of a critical point there (see below). Therefore another scheme such as the scaling laws or renormalization groups should be employed. The quantum effects too becomes non-negligible in this limit since $\hbar\omega > k_B T$, where ω is for example the characteristic frequency of thermodynamic fluctuations.

2) Critical phenomenon

The divergences of thermodynamic fluctuations in some kind of quantities were interpreted repeatedly as the presence of a second-order phase transition in the series of Kerr-Newman black hole states. These divergences are classified into two classes in each of which the state where this divergence occur is different and further the origin of the divergence is also different. In the first class [23, 24], the fluctuation in entropy $< (\Delta S)^2 > = k_B C_{JQ}$, for example, diverges at the state where C_{JQ} diverges in the midway from the Schwarzschild state to the extremal state. In the second class [25, 26], the divergence is due to the vanishing of, e.g., C_{JQ} at the extremal limit as seen in the previous paragraph. Of the two classes of divergence, however, the first one corresponds in fact to that at a turning point under the canonical environment while the second, to a real critical point [26]. The appearance of divergent fluctuations at a turning point were discussed in ref. [21]. Some possible interpretations of this critical phenomenon in the latter case were discussed in ref. [27, 26].

Lousto [24] interpreted the divergences of the first class as the indication of a critical phenomenon of black holes and claimed that he had shown the validity of various scaling laws among various critical exponents. He also claimed that the effective dimension of a black hole as a thermodynamic system is 2 as determined from the scaling laws and that this confirms the hole's membrane aspect. The idea and results are interesting enough but

unfortunately the divergences here do not correspond to a critical phenomenon. Therefore, it is of great interest to apply the same idea to the real critical phenomenon appearing at the extremal limit. This will be done in section 5.

C. Black Hole Entropy

1) Entropy proportional to area

The proportionality of black hole entropy to the area was a great surprise in the common sense of physics and various efforts have been devoted to understand this fact. Wald and others [28] have interpreted black hole entropy as a Nöther charge corresponding to the Killing field on the horizon. A merit of this interpretation is the applicability of the same idea also to dynamical holes. Bañados et al. [29] have tried to understand it from the topological viewpoint of the Euler class of the Euclidian black holes. Another interesting work in this connection is that of Srendnicki [30] in which a virtual sphere is placed in a massless scalar field. He showed that averaging over the freedom of a subsystem which occupies either inside or outside the sphere yields the entropy proportional to its area.

2) One-loop quantum corrections

Quantum corrections to the classical entropy discussed above have been calculated at the one-loop level by many authors, for example, through semi-classical Einstein equation [31] or the Euclidian action [32]. Most of these works agree in that the quantum corrections at this level give the results of the form, $S = 4\pi M^2 - \sigma \ln M$ and $T = (8\pi M)^{-1}[1 + \sigma(8\pi M^2)^{-1}]$ for a Schwarzschild hole, where σ is a constant.

3) Statistical mechanical foundations

Most of the considerations of black hole entropy in the framework of statistical mechanics were given from the membrane viewpoint of black holes. Zurek and Thorne [33] have introduced the concept of thermal atmosphere [7] around the stretched horizon. They regarded the entropy as residing in this atmosphere. Recently the entropy has been interpreted as the logarithm of the number of degrees of freedom on the membrane which are to be quantized [34].

4) Entropy of extremal holes

As discussed in the subsection of black hole equilibrium states, the non-zero entropy of extremal Kerr-Newman holes seems somewhat peculiar. The assertions have appeared recently [35, 36] that the entropy of an extremal Kerr-Newman hole is actually zero. This

result is obtained in ref. [36] by introducing an inner boundary and thereby changing the topology of the spacetime.

D. Evaporation of Black Holes

1) Problem of information loss

A serious problem associated with the black hole evaporation has been pointed out by Hawking [37] soon after the discovery of his own radiation. If the particles evaporating from a Schwarzschild hole is completely thermal and the evaporation process continues until the whole black hole vanishes, the information carried by infalling matter to form a black hole is lost from the universe since the thermal state is the most random state coceivable. In order to make the calculations much simpler, this problem has been treated mainly in the lower-dimensional (i.e., 1+1 or 2+1) black hole models. However, a question always hangs around such calculations as for to what extent they reflect the case of 4-dimensional black holes. A compact review of lower-dimensional black holes can be found in ref. [38].

Although many papers have appeared in the literature, the present status of this problem is still controversial. The possible ways of resolution may be summarized as follows.

- The loss of information actually occurs. This is the most drastic case among others. The predictability of quantum mechanics is then imperfect.
- The complete evaporation of a black hole does not occur. In this case, the lost information can be considered as remaining in the remnant [39, 8].
- The complete evaporation does occur but evaporated particles are not perfectly thermal. Therefore a long-time correlation remains in the radiation carrying information. This case is supported by some considerations from the membrane paradigm [40].
- Other than the above.

2) Non-equilibrium thermodynamics

Black holes in a vacuum are not in equilibrium states when the presence of evaporating particles are taken into account, however slow the process may be. The first paper written

from this viewpoint is that of Candelas and Sciama [41]. They showed that the emission of thermal particles by a Schwarzschild hole in the gravitational-wave mode can be understood in terms of the fluctuation-dissipation theorem. After a long interval, this idea has been revived by Crescimanno [42] in the case of weakly charged Reissner-Nordström holes. The power spectrum of charge-charge correlation has been calculated there and the validity of the fluctuation-dissipation theorem has been pointed out also in this case.

The generalization of these results to the case of arbitrary Kerr-Newman holes and the construction of irreversible black-hole thermodynamics have been performed recently by the present author [43]. The evaporation process of a Kerr-Newman hole can be regarded as a relaxation process to its Schwarzschild counterpart with nearly the same mass, in which angular momentum and charge are mainly emitted. The Langevin equations describing the evolutions of angular momentum and charge are derived on the grounds of thermodynamic considerations. The dynamic response functions of a general black hole and the spectra of fluctuating quantities are both derived from these equations of evolution. Comparing the latter with the imaginary part of the former, on can confirm the validity of the fluctuation-dissipation theorem for the evaporation process of a general hole, in the classical limit.

3) Back reaction of evaporating particles

The back reaction of the evaporating particles to the Kerr-Newman geometry is usually calculated through the semi-classical approach in which the Einstein equation is solved by including the source term arising from the expectation values of the quantum fields of evaporating particles. The reactions has been calculated to one-loop level and the resulting correction to the thermodynamic quantities are mentioned in the previous subsection. However, the calculations in this approach are fairly complicated and hence they are almost restricted to Schwarzschild holes.

Recently, another method called the "black-hole seismology" has been proposed in ref. [42]. According to this paper, the random-force term due to the back reaction of evaporating particles is added on the left-hand side of the Einstein equation and, corresponding to this, the energy-momentum tensor of the evaporating fields itself appears on the right-hand side. Such equations will be discussed being reduced to the Langevin-type equations but, at present, no paper along this line has been published yet.

3. Basic Properties of Equilibrium Kerr-Newman States

Before going into the discussion of the critical phenomenon of black holes, we summarize the parameter expressions [20, 43] of the equilibrium thermodynamic quantities. They are indispensable to a perspective discussion of arbitrary Kerr-Newman states.

In Boyer-Lindquist coordinates (r, θ, φ), the metric of a general Kerr-Newman black hole is written as

$$ds^2 = -\frac{\Delta}{\Sigma}(dt - a\sin^2\theta\, d\varphi)^2 + \frac{\sin^2\theta}{\Sigma}[(r^2+a^2)d\varphi - a\, dt]^2 + \frac{\Sigma}{\Delta}dr^2 + \Sigma\, d\theta^2, \qquad (4)$$

where

$$\Delta = r^2 - 2Mr + a^2 + Q^2, \qquad \Sigma = r^2 + a^2\cos^2\theta, \qquad a = J/M \geq 0. \qquad (5)$$

We use Planck units in which $c = G = \hbar = 8\pi k_B = 1$ hereafter in this paper. However, k_B and $\hbar$ are sometimes written explicitly, instead of $1/8\pi$ and 1, when it is convenient to do so.

The location of the horizon, $r_H = M \pm \sqrt{M^2 - a^2 - Q^2} \equiv r_\pm$, is obtained from the equation $\Delta = 0$, with the larger and smaller roots, r_+ and r_-, being the outer and inner horizons, respectively. It is convenient for many purposes to introduce non-dimensional parameters, $j = a/r_H \geq 0$, $q = Q/r_H$ and $h = \sqrt{a^2+Q^2}/r_H$. Of the three parameters, only two are independent since they satisfy $j^2 + q^2 = h^2$. Reminding the definition of r_H, one can see that h describes a series of the Kerr-Newman states: starting from the extremal limit $h = 1$ (i.e., $a^2 + Q^2 = M^2$) where the outer and inner horizons coincide, h decreases on the outer horizons and increases on the inner horizons toward the Schwarzschild limit where $h = 0$ and $h = \infty$, respectively, (i.e., $a^2 + Q^2 = 0$). As a more convenient second parameter, we introduce α such that $j = h\cos\alpha$ and $q = h\sin\alpha$ $(-\pi/2 \leq \alpha \leq \pi/2)$. This parameter measures the ratio of the charge to angular momentum since $\tan\alpha = Q/a$.

In terms of these parameters, the relation $\Delta = 0$ is rewritten as

$$h^2 - 2\frac{M^2}{J}h\cos\alpha + 1 = h^2 - 2\frac{M}{Q}h\sin\alpha + 1 = 0. \qquad (6)$$

From these expressions, we have

$$J = \frac{2h\cos\alpha}{1+h^2}\, M^2, \qquad Q = \frac{2h\sin\alpha}{1+h^2}\, M. \qquad (7)$$

Further, we summarize below the definitions of other thermodynamic quantities and their parameter expressions:

$$S = \frac{A_\pm}{4} = \frac{1}{8}(2Mr_H - Q^2) = \frac{1 + h^2\cos^2\alpha}{2(1+h^2)^2}\, M^2, \tag{8}$$

$$T = \frac{\pm 16\pi\sqrt{M^2 - a^2 - Q^2}}{A_\pm} = \frac{1 - h^4}{1 + h^2\cos^2\alpha}\, M^{-1} \equiv \beta^{-1}, \tag{9}$$

$$\Omega = \frac{4\pi J}{MA_\pm} = \frac{h(1+h^2)\cos\alpha}{2(1+h^2\cos^2\alpha)}\, M^{-1}, \qquad \Phi = \frac{4\pi Q r_H}{A_\pm} = \frac{h\sin\alpha}{1+h^2\cos^2\alpha}, \tag{10}$$

$$\mu = \beta\Omega = \frac{h\cos\alpha}{2(1-h^2)}, \qquad \phi = \beta\Phi = \frac{h\sin\alpha}{1-h^4}\, M, \tag{11}$$

where $A_\pm$, S, T, Ω and Φ are the horizon areas, entropy, temperature, angular velocity and electrostatic potential, respectively, of the black hole. Thus all thermodynamic quantities are expressed in terms of three parameters, h, α and M. Note that these expressions are valid not only for the outer horizons ($h \leq 1$) but also for the inner horizons ($h \geq 1$). At present, however, the direct relevance of the inner horizon thermodynamics to a distant observer remains rather unclear.

The mass formula of a Kerr-Newman black hole is the equation of state from the thermodynamic viewpoint. There are two different expressions of this equation. One is Christodoulou's expression [44]

$$M = \left[2S + \frac{J^2}{8S} + \frac{Q^2}{2} + \frac{Q^4}{32S}\right]^{1/2}, \tag{12}$$

and the other is Smarr's expression [45]

$$M = 2TS + 2\Omega J + \Phi Q. \tag{13}$$

From the former expression, we can verify that the entropy $S = S(M,\ J,\ Q)$ is a generalized homogeneous function, i.e.,

$$S(\lambda^{1/2}M,\ \lambda J,\ \lambda^{1/2}Q) = \lambda S(M,\ J,\ Q) \tag{14}$$

where λ is an arbitrary positive number.

There are three kinds of susceptibilities for a Kerr-Newman black hole. Each kind corresponds to a mode of energy exchange, i.e., thermal, mechanical or electrostatic mode.

Each kind of susceptibility can take different values according to different boundary conditions. The boundary conditions are specified by the presence or absence of the contact to a reservoir of each mode. Therefore, for a selected kind of susceptibility whose value is to be determined, there are four possibilities as for the remaining two channels: (c, c), (o, c), (c, o) and (o, o) where c and o stand for "closed" and "open", respectively.

The susceptibilities in thermal, mechanical and electrostatic modes are heat capacity C_a, moment of inertia I_b and electric capacitance K_c, respectively, where a, b and c represent the set of the fixed quantities under different boundary conditions. Their explicit expressions are

$$\begin{aligned} C_a &\equiv -\beta^2 \left(\frac{\partial M}{\partial \beta}\right)_a = -\beta M \left(\frac{B_1}{A_1}, \frac{B_2}{A_2}, \frac{B_3}{A_3}, \frac{B_4}{A_4}\right), \\ I_b &\equiv \beta \left(\frac{\partial J}{\partial \mu}\right)_b = \frac{J}{\Omega} \left(\frac{B_1}{B_2}, \frac{A_1}{A_2}, \frac{B_3}{B_4}, \frac{A_3}{A_4}\right), \\ K_c &\equiv \beta \left(\frac{\partial Q}{\partial \phi}\right)_c = \frac{Q}{\Phi} \left(\frac{B_1}{B_3}, \frac{A_1}{A_3}, \frac{B_2}{B_4}, \frac{A_2}{A_4}\right), \end{aligned} \tag{15}$$

with

$$\begin{aligned} \{a\} &= \{(J, Q), (\mu, Q), (J, \phi), (\mu, \phi)\}, \\ \{b\} &= \{(M, Q), (\beta, Q), (M, \phi), (\beta, \phi)\}, \\ \{c\} &= \{(M, J), (\beta, J), (M, \mu), (\beta, \mu)\}, \end{aligned} \tag{16}$$

where

$$\begin{aligned} A_1 &= (1+j^2)[1-2(h^2+2j^2)-3h^4], \\ A_2 &= (1+h^2)[1-3h^2+(5+h^2)j^2], \\ A_3 &= 1-h^4-(5+10h^2+9h^4)j^2+2(1+3h^2)j^4, \\ A_4 &= (1+h^2)[1-h^2+3(1+h^2)j^2-2j^4], \\ B_1 &= (1+j^2)(1-h^2)^2, \\ B_2 &= (1+j^2)[(1-h^2)^2+4j^2], \\ B_3 &= (1+j^2)[1+3h^4-2(1+h^2)j^2], \\ B_4 &= (1+j^2)[1+3h^4+2(1-h^2)j^2]. \end{aligned} \tag{17}$$

4. Critical Behavior (Lousto's Work)

In a series of papers [24] Lousto examined the so called critical behavior of Kerr-Newman black holes at the point where C_{JQ}, I_{TQ}^{-1} and K_{TJ}^{-1} diverge. As is pointed out in section 2-B, however, this point is actually a turning point instead of a critical point associated with a second-order phase transition. Nevertheless, the outline of his work is described here because it contains some important ideas.

The above susceptibilities are written explicitly as,

$$C_{JQ} = T\left(\frac{\partial S}{\partial T}\right)_{JQ} = \frac{MTS^3}{\pi J^2 + (\pi/4)Q^4 - T^2S^3} \propto \frac{B_1}{A_1}, \tag{18}$$

$$I_{TQ}^{-1} = \left(\frac{\partial \Omega}{\partial J}\right)_{TQ} \propto \frac{\pi(2\Phi Q - M)(1 - 4\pi TM)}{S^2[1 - 2\pi TM + 4\pi^2T^2(6M^2 + Q^2)]} \propto \frac{A_2}{A_1}, \tag{19}$$

$$K_{TJ}^{-1} = \left(\frac{\partial \Phi}{\partial Q}\right)_{TJ} \propto \frac{A_3}{A_1}. \tag{20}$$

In the above equations, the last terms are expressions in terms of our A's and B's given in the previous section. (His units here are $c = G = \hbar = k_B = 1$.) The divergence of these quantities occurs simultaneously at the zero of A_1. Although he did not distinguish the meaning of his K_{TQ} and K_{TJ}, they are in fact the susceptibilities in different modes and, respectively, proportional to our I_{TQ} and K_{TJ}.

Regarding these divergences as indications of the presence of a second-order phase transition in the black hole states, he calculated various critical exponents defined as follows:

$$C_{JQ} \sim \begin{cases} \epsilon^{-\alpha^*} & \text{for } \Omega = 0 \\ \Omega^{\varphi} & \text{for } \epsilon = 0, \end{cases} \tag{21}$$

$$I_{TQ}^{-1} \sim \begin{cases} \epsilon^{-\gamma} & \text{for } \Omega = 0 \\ \Omega^{1-\delta^{-1}} & \text{for } \epsilon = 0, \end{cases} \tag{22}$$

$$J \sim \begin{cases} \epsilon^{\beta^*} & \text{for } \Omega = 0 \\ \Omega^{\delta^{-1}} & \text{for } \epsilon = 0, \end{cases} \tag{23}$$

where $\epsilon = T - T_c$ with T_c being the critical temperature. Although it is difficult to trace all of his calculations (mainly due to the ambiguity in the identification of his H and M in general cases with Ω and J adopted here, respectively), the claimed results are

$$\alpha^* = 1, \quad \beta^* = 0, \quad \gamma = 1, \quad \delta^{-1} = 0. \tag{24}$$

These exponents satisfy the scaling laws of the "first kind",

$$\begin{cases} \alpha^* + 2\beta^* + \gamma = 2 & \text{(Essam and Fisher)}, \\ \beta^*(\delta - 1) = \gamma & \text{(Widom)}, \\ \varphi(\beta^* + \gamma) = \alpha^*, \end{cases} \tag{25}$$

provided that the additional condition $\beta^*\delta = 1$ holds. The justification of the last condition, however, has not been given.

Other critical exponents associated with spatial correlations are further defined as follows. Usually the two-point correlation function at a large distance is written in the form [46],

$$G(r) \sim \frac{\exp[-r/\xi]}{r^{d-2+\zeta}}, \tag{26}$$

where ζ is Fisher's exponent, d is the effective dimension of the system under consideration and ξ is the correlation length which also diverges like

$$\xi \sim \epsilon^{-\nu}, \qquad \text{therefore} \quad G(r) \sim r^{-(d-2+\zeta)}, \tag{27}$$

at the critical point. The scaling laws of the "second kind" contains the exponents related to the correlation:

$$\begin{cases} \nu(2 - \zeta) = \gamma & \text{(Kadanoff)}, \\ \nu d = 2 - \alpha^* & \text{(hyperscaling law)}, \\ \mu^*(\beta^* + \gamma) = \nu. \end{cases} \tag{28}$$

The last equation does not appear in Lousto's work since the exponent μ^* is not defined there. It will appear in the next section.

Lousto has used, as a correlation function, the Green function of a scalar field which is in thermal equilibrium with a Schwarzschild black hole. In the Hartle-Hawking vacuum state, it is written at a large distance as [47]

$$G_\omega \sim \frac{\omega}{2\pi\left[\exp\left(\frac{2\pi\omega}{k_B}\right) - 1\right]}, \tag{29}$$

where ω is the angular frequency of the mode considered. Since this is independent of the distance r, we have

$$d - 2 + \zeta = 0, \tag{30}$$

from Eq. (27). The exponent ν has been obtained from the above expression as

$$\xi^2 = -\frac{1}{2\pi}\left(\frac{\partial^2 G(\omega)}{\partial \omega^2}\right)_{\omega=0} \sim \epsilon^{-1}, \quad \text{i.e.,} \quad \nu = \frac{1}{2}. \tag{31}$$

Substituting these results into the above two of Eq. (28), we obtain

$$d = 2, \qquad \zeta = 0. \tag{32}$$

Thus the effective dimension of a black hole is derived to be 2 as expected from the membrane viewpoint.

Unfortunately, however, there are some problems in Lousto's treatments as well as the fact that the zero of A_1 is in fact not a critical point. He has not considered the critical exponents associated with K_{TJ}^{-1} which must satisfy other scaling laws. The argument for justifying the relation $\beta^*\delta = 1$ has not been given. The definition of any order parameter has not been given explicitly, which is however indispensable for the physical understanding of a second-order phase transition. Since the exponent β^* is usually introduced to describe the critical behavior of such a parameter, Eq. (23) means that the angular momentum is regarded as an order parameter. However, J is not necessarily be zero on either side of the zero of A_1.

5. Critical Behavior (Original Work)

In Lousto's work, the point where C_{JQ} diverges is considered as a critical of a second-order phase transition. The relevant thermodynamic environment (i.e., the boundary conditions) is called canonical in which the system is allowed to exchange heat with a heat reservoir but is inhibited to exchange energy with both of a mechanical and an electric reservoirs, so that J and Q are kept constant. The fact that there is no contact with mechanical and electric reservoirs is reflected also in the point that the relevant quantities (generally called the response coefficients) in the respective modes are I_{TQ}^{-1} and K_{TJ}^{-1} instead of susceptibilities themselves.

In the canonical environment, the intrinsic variables (expressed as x_i in general) are $\beta \equiv T^{-1}$, J and Q and their conjugates (expressed as X_i in general) are $-M$, $-\mu$ and $-\phi$. The definition of the response coefficients is $\chi_i(0) = (\partial X_i/\partial x_i)_{\bar{x}_i}$ where $\bar{x}_i$ represents

the set of intrinsic variables excluding x_i itself. Their explicit forms are

$$\chi_1(0) = -\left(\frac{\partial M}{\partial \beta}\right)_{JQ} = \frac{C_{JQ}}{\beta^2} = \frac{B_1}{A_1}\frac{M}{\beta},$$
$$\chi_2(0) = -\left(\frac{\partial \mu}{\partial J}\right)_{\beta Q} = -\frac{\beta}{I_{\beta Q}} = -\frac{A_2}{A_1}\frac{\mu}{J}, \tag{33}$$
$$\chi_3(0) = -\left(\frac{\partial \phi}{\partial Q}\right)_{\beta J} = -\frac{\beta}{K_{\beta J}} = -\frac{A_3}{A_1}\frac{\phi}{Q}.$$

Each coefficient gives the slope of a tangent to the equilibrium curve drawn in the two-dimensional plane relevant for each mode. Since these tangents change their signs through $\pm\infty$ at a continuous part of each equilibrium curve these points of divergent tangents are all turning points [20]. In fact these points correspond to the same one state where A_1 vanishes. Thus a change of stability occurs in each mode simultaneously. The nature of these instabilities are trivial. Namely, all are associated with the thermal instability in a heat bath of a system with negative heat capacity.

The stability of a series of equilibrium states changes from stable (the part of $C_{JQ} > 0$) to unstable (the part of $C_{JQ} < 0$) when J and/or Q are decreased for a given M from the extremal limit to the Schwarzschild limit. Corresponding to this change, the entropy of the equilibrium series changes from locus of local maxima to that of local minima. Thus, at the turning point, the entropy becomes locally flat in the directions of off-equilibrium displacements, resulting in the vanishment of restoring forces. This is the origin of the divergent fluctuations at a turning point mentioned in section 2-B. On the other hand, the origin of the divergent fluctuations at a critical point is slightly different. For example, suppose the coexistence curve of water and vapor in the temperature vs. pressure plane. Away from the critical point where this curve terminates, there are two local maxima of the same height in the relevant Massieu function (or equivalently, two minima in the Helmholtz free energy) representing the liquid and gas phases of water. As the critical point is approached, these maxima become closer to each other and finally merge together resulting in a locally flat configuration.

The critical point in Kerr-Newman states is located in fact at the extremal black hole states. This can be derived from a physical consideration. As shown in section 3, for a given set of (M, J, Q) there are two different series of thermodynamic states corresponding to the outer and inner horizons. According to Curir [27], we regard these series as different

phases of the black hole states with different intensive variables. For a given J and Q, the line of coexistence (i.e., the black hole solution) terminates at the extremal limit where $M = M_X$ with $M_X^2 \equiv (Q^2 + \sqrt{Q^4 + 4J^2})/2$. From the analogy with the case of water, one can guess the extremal limit is a critical point. Indeed, the differences of the two phases vanishes as the two horizons merge together.

The existence of a critical point at this limit has also been confirmed from the occurrence of diverging fluctuations at a point other than a turning point [26]. The intrinsic variables are now M, J and Q, and their conjugates are β, $-\mu$ and $-\phi$. Therfore, the response coefficients in this microcanonical environment are

$$\begin{aligned}
\chi_1(0) &= \left(\frac{\partial^2 S}{\partial M^2}\right)_{JQ} = -\frac{\beta^2}{C_{JQ}} = \frac{A_1}{B_1}\frac{\beta}{M}, \\
\chi_2(0) &= \left(\frac{\partial^2 S}{\partial J^2}\right)_{MQ} = -\frac{\beta}{I_{MQ}} = -\frac{B_2}{B_1}\frac{\mu}{J}, \\
\chi_3(0) &= \left(\frac{\partial^2 S}{\partial Q^2}\right)_{MJ} = -\frac{\beta}{K_{MJ}} = -\frac{B_3}{B_1}\frac{\phi}{Q}.
\end{aligned} \tag{34}$$

All these quantities diverge at $h = 1$ (the extremal states) where B_1 becomes zero.

The order parameter for the present case may be [26] $T_+ - T_-$, $\Omega_- - \Omega_+$ or $\Phi_- - \Phi_+$. These are all positive for non-extremal Kerr-Newman states (hence the ordered states) and become zero at the critical point. Since we have to discuss the scaling laws for black hole cases, however, the definition of the critical exponents should be given following a systematic scheme which is common to the cases of usual matter. The scheme we adopt here is that the variables common to the two phases are regarded as intrinsic variables and the difference of each conjugate variable between the two phases is defined as an order parameter. Then, the critical exponents are introduced to describe the behavior of the response coefficients and these order parameters near the critical point. Thus we have been led to adopt the order parameters, $\eta_M \equiv \beta_+ - \beta_-$, $\eta_J \equiv \mu_- - \mu_+$ and $\eta_Q \equiv \phi_- - \phi_+$, for the thermal, mechanical and electrostatic modes, respectively. It must be noted here that these parameters do not go to zero at the critical limit but diverge. This is due to the special circumstances of black holes that the temperature goes to zero at this limit. Reflecting this fact, β^*'s and δ's become negative for black holes.

Our definitions of various critical exponents are as follows:

$$\chi_1(0) \sim \begin{cases} \epsilon_M^{-\alpha^*} & (J=0 \text{ or } Q=0) \\ \epsilon_J^{-\varphi_1} & (Q \neq 0,\ \alpha = \text{const.}) \\ \epsilon_Q^{-\varphi_2} & (J \neq 0,\ \alpha = \text{const.}), \end{cases} \tag{35}$$

$$\chi_2(0) \sim \begin{cases} \epsilon_M^{-\gamma_1} & (Q=0) \\ \epsilon_J^{-\sigma_1} & (Q \neq 0), \end{cases} \tag{36}$$

$$\chi_3(0) \sim \begin{cases} \epsilon_M^{-\gamma_2} & (J=0) \\ \epsilon_J^{-\sigma_2} & (J \neq 0), \end{cases} \tag{37}$$

$$\eta_J \sim \begin{cases} \epsilon_M^{\beta_1^*} & (Q=0) \\ \epsilon_J^{\delta_1^{-1}} & (Q \neq 0), \end{cases} \tag{38}$$

$$\eta_Q \sim \begin{cases} \epsilon_M^{\beta_2^*} & (J=0) \\ \epsilon_Q^{\delta_2^{-1}} & (J \neq 0), \end{cases} \tag{39}$$

$$\xi_J \sim \begin{cases} \epsilon_M^{-\nu_1} & (Q=0) \\ \epsilon_J^{-\mu_1^*} & (Q \neq 0), \end{cases} \tag{40}$$

$$\xi_Q \sim \begin{cases} \epsilon_M^{-\nu_2} & (J=0) \\ \epsilon_Q^{-\mu_2^*} & (J \neq 0), \end{cases} \tag{41}$$

where ϵ's are defined by the relations $M = M_X(1+\epsilon_M)$, $J = J_X(1-\epsilon_J)$ and $Q = Q_X(1-\epsilon_Q)$, and are mutually related like $\epsilon_J = \epsilon_M/\cos^2\alpha$ and $\epsilon_Q = \epsilon_M/\sin^2\alpha$. When $\zeta = 0$ in particular, we have the relations [46],

$$\xi \propto \chi^{1/2} \quad \Longrightarrow \quad \nu = \frac{1}{2}\gamma, \quad \mu^* = \frac{1}{2}\sigma, \tag{42}$$

for each mode.

It is well expected that the scaling laws hold even for the case of black holes, because ([46]) the entropy as a function of M, J and Q is a generalized homogeneous function. We obtain for black holes

$$\begin{aligned} &\alpha^* = \tfrac{3}{2}, \quad \gamma_1 = \gamma_2 = \tfrac{3}{2}, \quad \varphi_1 = \varphi_2 = \tfrac{3}{2}, \\ &\beta_1^* = \beta_2^* = -\tfrac{1}{2}, \quad \delta_1 = \delta_2 = -2, \quad \sigma_1 = \sigma_2 = 1, \end{aligned} \tag{43}$$

and these values actually satisfy the scaling laws of the first kind. The validity of the laws of the second kind, however, cannot be confirmed directly because of lack of information. It can only be derived from the third relation in Eq. (28) that $\mu^* = \nu$. We therefore examine some possibilities individually.

1) $d-2+\zeta=0$

This is the case in which the correlation function has no r-dependence as expected from the exact thermality of evaporated particles. Different from the case of Lousto, however, the assumption is not compatible with the scaling laws. The first relation in Eq. (28) yields $\nu d = 3/2$ while the second does $\nu d = 1/2$. Thus the scaling laws predict the non-thermality of evaporating particles.

2) $\zeta = 0$

The first and second relations yield $\mu^* = \nu = 3/4$ and $d = 2/3$, respectively. However, Eq. (42) gives $\mu^* = 1/2$ and $\nu = 3/4$. This is a contradiction.

3) $d = 2$

From the standpoint of the membrane paradigm, it is rather trivial that the effective dimension of a black hole is 2. Then the scaling laws predict that $\mu^* = \nu = 1/4$, $\zeta = -4$, and therfore that

$$G(r) \propto r^4 \exp\left[-\frac{r}{\xi}\right]. \tag{44}$$

To summarize, the scaling laws predict the presence of a long-time correlation in the evaporating particles since a large r means a long time separation. This fact might shed new light on the problem of information loss. It must also be mentioned that the superradiance dominates over the Hawking radiation near the extremal limit.

References

[1] J. D. Bekenstein, Phys. Rev. D **7**, 949 (1973); *ibid.* D **7**, 233 (1973); *ibid.* D **9**, 3292 (1974).

[2] S. W. Hawking, Comm. Math. Phys. **25**, 152 (1972).

[3] J. M. Bardeen, B. Carter and S. W. Hawking, Comm. Math. Phys. **31**, 161 (1973).

[4] S. W. Hawking, Nature **248**, 30 (1974); Comm. Math. Phys. **43**, 199 (1974).

[5] R. Penrose, Nuovo Cim. **1**, (1969) Special No 252.

[6] D. W. Sciama, Vistas in Astronomy **19**, 385 (1976); P. C. W. Davies, Rep. Prog. Phys. **41** 1313 (1978).

[7] K. S. Thorne, R. H. Price and D. A. Macdonald, *Black Holes: The Membrane Paradigm* (Yale University Press, New Heaven, 1986).

[8] L. Susskind, L. Thorlacius and Uglum, Phys. Rev. D **48**, 3743 (1993).

[9] M. Maggiore, Phys. Rev. D **49**, 2918 (1994).

[10] L. Susskind and L. Thorlacius, Phys. Rev. D **49**, 966 (1994).

[11] V. de Sabbata and Z. Zhang (ed.), *Black Hole Physics* (Kluwer Academic, Dordrecht,1992).

[12] G. W. Gibbons and S. W. Hawking, Phys. Rev. D **15**, 2752 (1977).

[13] J. W. Yrok Jr., Phys. Rev. D **33**, 2092 (1986); B. F. Whiting and J. W. York Jr., Phys. Rev. Lett. **61**, 1336 (1988); B. F. Whiting, Class. Quantum Grav. **7**, 15 (1990).

[14] G. W. Gibbons and K. Maeda, Nucl. Phys. B **298**,741 (1988); D. Garfinkle, G. T. Horowitz and A. Strominger, Phys. Rev. D **43**, 3140 (1991).

[15] R. Bartinik and J. McKinnon, Phys. Rev. Lett. **61**, 141 (1988); M. S. Volkov and D. V. Galt'son, JETP Lett. **50**, 346 (1989).

[16] S. Droz, M. Heusler and N. Straumann, Phys. Lett. B **268**, (1991); P. Bizon and T. Chmaj, Phys. Lett. B **297**, 55 (1992).

[17] B. R. Greene, S. D. Mathur and C. M. O'Neill, Phys. Rev. D **47**, 2242 (1993).

[18] J. Preskill, P. Schwarz, A. Shapere, S. Trivedi and F. Wilczek, Mod. Phy. Lett. A **6**, 2353 (1991).

[19] L. M. Krauss and F. Wilczek, Phys. Rev. Lett. **62**, 1221 (1989); S. Coleman, J. Preskill and F. Wilczek, Phys. Rev. Lett. **67**, 1975 (1991).

[20] O. Kaburaki, I. Okamoto and J. Katz, Phys. Rev. D **47**, 2234 (1993); J. Katz, I. Okamoto and O. Kaburaki, Class. Quantum Grav. **10**, 1323 (1993).

[21] O. Kaburaki, Phys. Lett. A **185**, 21 (1994).

[22] K. Maeda, T. Tachizawa T. Torii and T. Maki, Phys. Rev. Lett. **72**, 450 (1994); T. Torii, K. Maeda and T. Tachizawa, Proc. of the Third Workshop on General Relativity and Gravitation, eds. K. Maeda et al. (Tokyo, 1994) p. 55; T. Tachizawa T. Torii, K. Maeda and T. Maki, *ibid.* p. 65.

[23] P. C. W. Davies, Proc. R. Soc. London A **353**, 499 (1977); Y. K. Lau, Phys. Lett. **186**, 41 (1994).

[24] C. O. Lousto, Nucl. Phys. B **410**, 155 (1993); in *Csmology and Particle Physics* (ed. de Sabbata and H. Tso-Hsiu, Kluwer Academic, Dordrecht, 1994) p. 183; Gen. Rel. Grav. **27**, 121 (1995).

[25] D. Pavón and Rubí, Phys. Rev. D **37**, 2052 (1988); D. Pavón, Phys. Rev. D **43**, 2495 (1991); R.-K. Su, R.-G. Cai and P. K. N. Yu, Phys. Rev. D **50**, 2932 (1994).

[26] O. Kaburaki, Gen. Rel. Grav. (1996) in press; Proc. of the Fourth Workshop on General Relativity and Gravitation (eds. K. Nakao et al., Tokyo, 1994) p. 219.

[27] A. Curir, Gen. Rel. Grav. **13**, 417 (1981); *ibid.* 1177 (1981).

[28] R. M. Wald, Phys. Rev. D **48**, R3427 (1993); V. Iyer and R. M. Wald, Phys. Rev. D **52**, 4430 (1995).

[29] M. Bañados, C. Teitelboim and J. Zanelli, Phys. Rev. Lett. **72**, 957 (1994).

[30] M. Slendnicki, Phys. Rev. Lett. **71**, 666 (1993).

[31] R. Balbinot and A. Barletta, Class. Quantum Grav. **6**, 195 (1989); D. Hochberg, Phys. Rev. D **51**, 5742 (1995).

[32] S. N. Soldukhin, Phys. Rev. D **51**, 609 (1995); *ibid.* 618 (1995); D. V. Fursaev, Phys. Rev. D **51**, R5352 (1995).

[33] W. H. Zurek and K. S. Thorne, Phys. Rev. Lett. **54**, 2171 (1985).

[34] G. 't Hooft, Nucl. Phys. B **256**, 727 (1985); M. Maggiore, Nucl. Phys. B **429**, 205 (1994); C. O. Lousto, Phys. Lett. B **352**, 228 (1995).

[35] S. W. Hawking, G. T. Horowitz and S. F. Ross, Phys. Rev. D **51**, 4302 (1995).

[36] G. W. Gibbons and R. E. Kollosh, Phys. Rev. D **51**, 2839 (1995).

[37] S. W. Hawking, Phys. Rev. D **14**, 2460 (1976).

[38] J. A. Harvey, in *String Theory and Quantum Gravity '92* (ed. J. H. Harvey et al., World Scientific, Singapore, 1993) p. 122.

[39] Y. Aharonov, A. Casher and S. Nussinov, Phys. Lett. B **191**, 51 (1987).

[40] D. Page, Phys. Rev. Lett. **44**, 301 (1980); G. 't Hooft, Nucl. Phys. B **335**, 138 (1990).

[41] P. Candelas and D. W. Sciama, Phys. Rev. Lett. **23**, 1372 (1977).

[42] M. Crescimanno, Phys. Rev. D **50**, 3954 (1994).

[43] O. Kaburaki, Phys. Rev. D (1995) submitted.

[44] D. Christodoulou, Phys. Rev. Lett. **25**, 1596 (1970).

[45] L. Smarr, Phys. Rev. Lett. **30**, 71 (1973).

[46] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford Univ. Press, Oxford, 1971); L. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, New York, 1980).

[47] D. W. Sciama, P. Candelas and D. Deutsch, Adv. Phys. **30**, 327 (1981).

On the ergoregion instability – a new and general method for numerical mode analysis of classical massless scalar field

Shin'ichirou Yoshida*
and
Yoshiharu Eriguchi†

Jan. 24, 1996

1 Introduction

In general relativity, systems with sufficiently strong self-gravity and considerably large angular momenta can have the *ergoregion* where the time-time component of the metric satisfies the condition $g_{00} > 0$[1].

The Kerr black hole is the well-known example of the spacetime having the ergoregion. Also rotating dense stellar models without an event horizon can have the ergoregion as numerically shown by Butterworth & Ipser [1]. In the ergoregion, the dragging of inertial frames is too strong for counter-rotating timelike or null geodesics against the system to exist. One interesting and important nature of the ergoregion is that particles or fields within that region can be in negative energy states viewed from a stationary observer at infinity. Friedman [7] showed that, because of the existence of negative energy states, stars with the ergoregion suffer an instability driven by radiation fields such as scalar or electromagnetic ones (and also gravitational fields, as commented by Friedman). The mechanism of this instability can be explained as follows: Negative energy perturbations of the field trapped in the ergoregion couple with a radiation field having a positive flux of energy at infinity. This coupling reduces the field's energy in the ergoregion and results in growth of amplitudes of the perturbations.

Comins & Schutz [5] studied this instability for classical massless scalar fields in the spacetime of rotating general relativistic stars. [2] Although there were several numerically exact solutions of rapidly rotating and general relativistic stars they used the slow-rotation approximation developed by Hartle & Thorne [9] because the problem can be reduced to simplified one dimensional one. As dynamics of the massless scalar perturbation is described by a scalar wave equation on the curved background, it was examined by using the WKBJ approximation. They found that the e-folding times of unstable modes are several orders of magnitude longer than the age of the universe. Consequently, even if the ergoregion arises in realistic compact objects such as neutron or quark stars, it has little effect on the star.

As for the instability of rotating stars, we need to consider another kind of instability different from that related to the ergoregion. Rapidly rotating stars, with or without the ergoregion, suffer an instability driven by gravitational radiation, which was found by Chandrasekhar [2] and proved to exist universally by Friedman & Schutz [8] and Friedman [6]. Since this instability is much stronger in its effect than the

*Department of Astronomy, School of Science, University of Tokyo (E-mail: yoshida@astron.s.u-tokyo.ac.jp)
†Department of Earth Science and Astronomy, Graduate School of Arts and Sciences, University of Tokyo

[1] In this paper we adopt the convention of Misner, Thorne & Wheeler [14] and units $c = G = 1$.

[2] Sato & Maeda [16] analyzed a similar problem with quantized massive scalar field on Kerr-Schild type interior metric.

ergoregion instability (Ipser & Lindblom [10];Yoshida & Eriguchi [18]), it is more important to investigate such instabilities. However, instability due to gravitational radiation has not been investigated in the framework of general relativity yet. Therefore if some schemes developed for analysis of the ergoregion instability could be applied to this instability, it would be very helpful to understand characters of rapidly rotating neutron stars. However the techniques thus far developed by several authors seem to be unable to handle this problem. The main obstacle is the multidimensional feature of the problem. All quantitative investigations about the stability of compact objects have been done either by assuming spherical symmetry or by reducing the problem to one dimensional one by using some approximations such as slow rotation approximation.

In this context we need a new scheme which can be applied both for the ergoregion instability in two dimension and for the gravitational radiation driven instability. Therefore we will develop in this paper the alternative technique to study ergoregion instability which can be extended to multidimensional cases with a little amendments. This paper is organized as follows. In §2 formulation of the problem is presented. It is assumed that the systems we consider have no horizons. Numerical results are given in §3 and concluding remarks in §4. In the appendix some numerical results for modes of relativistic stars are briefly discussed. In particular we show that axial modes of ultra-compact stars (Chandrasekhar & Ferrari [3, 4],;Kokkotas [11]) can be treated by the newly formulated scheme.

2 Formulation

2.1 Scalar perturbation equations

Unperturbed stellar models are constructed by the same procedure as that of Schutz & Comins [17]. Since we assume that the stars consist of uniform density and are slowly rotating, the background metric can be expressed as:

$$ds^2 = -e^{2\Phi}dt^2 + e^{2\Lambda}dr^2 + r^2d\theta^2 + r^2\sin^2\theta(d\varphi - \omega dt)^2, \tag{1}$$

where $\Phi(r), \Lambda(r)$ and $\omega(r)$ are three potentials. The first two functions are the same as those of the spherical stars and the last potential denotes "dragging of the inertial frame".

Here we restrict our attention to the massless scalar field because of its mathematical simplicity. Behaviour of the massless scalar field is described by a scalar wave equation on a curved background,

$$\nabla^\mu\nabla_\mu\psi = 0, \tag{2}$$

where ∇_μ denotes covariant derivative in terms of the metric (1). If we expand the function ψ in terms of scalar spherical harmonics and assume that ψ has a harmonic time-dependency ($\sim e^{i\sigma t}$), i.e.

$$\psi = \frac{\zeta(r)}{r}Y_{lm}(\theta,\varphi)e^{i\sigma t}, \tag{3}$$

Eq.(6) is reduced to the following one dimensional equation:[3]

$$\zeta'' + \left[(\sigma + m\omega)^2e^{2(\Lambda-\Phi)} - \frac{l(l+1)e^{2\Lambda}}{r^2} - \frac{\Phi' + \Lambda'}{r}\right]\zeta + [\Phi' + \Lambda']\zeta' = 0, \tag{4}$$

[3]The term "$\Phi' + \Lambda'$" is written as "$\Phi' - \Lambda'$" in Comins & Schutz [5]. This may be an error, so we have corrected this. However this correction has not changed the results seriously (they do not change by an order).

where prime means differentiation with respect to r-coordinate. By introducing a new variable χ defined as:

$$\zeta = \chi e^{\left[-\frac{1}{2}(\Phi+\Lambda)\right]}, \tag{5}$$

this equation is transformed to,

$$\chi'' + m^2 e^{2(\Lambda-\Phi)}(\Sigma - V_+)(\Sigma - V_-)\chi = 0, \tag{6}$$

where $\Sigma = \sigma/m$, and $V_\pm$ are factored potentials defined in Comins & Schutz. It should be noted that in the high-m limit, the last term in T can be neglected. This approximation gives us the basic equation for the WKBJ analysis (Appendix A).

In our new formulation we reorganize the equation to an inhomogeneous Helmholtz equation in one dimensional space as follows:

$$\left[\Delta_1 + \sigma^2\right]\zeta = -\rho, \tag{7}$$

where

$$\Delta_1 \equiv \frac{d^2}{dr^2}, \tag{8}$$

and this is essentially the Laplacian operator in one dimension and

$$\rho \equiv \left[(\sigma + m\omega)^2 e^{2(\Lambda-\Phi)} - \sigma^2 - \frac{l(l+1)e^{2\Lambda}}{r^2} - \frac{\Phi' + \Lambda'}{r}\right]\zeta + \left[\Phi' + \Lambda'\right]\zeta'. \tag{9}$$

2.2 New method for the eigenvalue problem

In our new method, we treat the inhomogeneous Helmholtz equation, Eq.(14), as our basic equation for the problem. At the region far from the stars, i.e. at the vacuum and flat spacetime, this equation describes simply propagation of the scalar wave. In this sense, the physical meaning of the solution becomes simple and moreover it is easy to treat the boundary condition for the wave at infinity. However, since the curved background spacetime (1) approaches to a flat spacetime only asymptotically, i.e. as $r \to \infty$, the existence of the source term of Eq.(14) in a region near the star distorts this simple view because this source term behaves in a complicated manner due to strong gravity even if we treat the vacuum region.

In order to get rid of this ill-behaviour of the source term, we divide the space into two regions in our present numerical treatment. In the inner region we use Eq.(8) as our basic equation, while in the outer region we transform Eq.(14) to an integral equation by taking suitable boundary conditions into account. The conditions which we need to impose are (i) the continuity of the scalar field at the boundary between the inner region and the outer region and (ii) the so-called outgoing radiation condition at infinity, i.e. no incoming wave from infinity exists. These two boundary conditions can be expressed as,

$$\alpha\zeta + \beta\frac{\partial\zeta}{\partial n} = \gamma, \tag{10}$$

where α, β and γ are, in general, functions of coordinates at the boundary and $\partial/\partial n$ denotes a directional derivative normal to the boundary.

Applying these conditions, we can express a formal integral "solution" of Eq.(14) as follows. First, let us introduce Green's function satisfying the following equation:

$$(\Delta_1 + \sigma^2)G^+(x, x') = -\delta(x, x'), \tag{11}$$

where $G^+(x,x')$ and $\delta(x,x')$ are Green's function and Dirac's delta function, respectively. Then Eq.(14) can be expressed as the following integral form,

$$\zeta(x) = \int_{\mathbf{V}} \rho(x')G^+(x,x')d\mathbf{V}' + \oint_{\partial\mathbf{V}} G^+(x,x')\frac{\gamma(x')}{\beta(x')}d\mathbf{S}'. \tag{12}$$

where $\mathbf{V}$ and $\partial\mathbf{V}$ denote the region covered by the integral, i.e. the whole space in general and its boundary, respectively. It should be noted that this form does not give a "solution", for the source term of Eq.(14) involves an unknown function ζ and its spatial derivatives. However this expression is useful because we can include the boundary conditions explicitly in this form.

The outgoing wave boundary condition can be expressed at infinity as,

$$\left(\frac{\partial}{\partial n} + i\sigma\right)\zeta \to 0, \qquad (x \to \infty). \tag{13}$$

Although strictly speaking this condition must be satisfied at infinity, we now impose this condition at some finite distance and regard that place as the "infinity". Therefore we can choose $\beta = 1$ and $\gamma = 0$. As seen from numerical results in §3 and the appendix, this approximation works fairly well despite of its simplicity.

Concerning the numerical scheme to solve eigenvalue problems, we adopt the method developed by Yoshida & Eriguchi [18] for oscillating newtonian rotating stars. According to their scheme, differential and integral equations of two regions are discretized on finite grid points, each equation of which is reduced to a linear equation. An eigenvalue and a corresponding eigenfunction are simultaneously calculated by Newton-Raphson method by imposing a normalization condition of the eigenfunction as a supplementary equation.

3 Results

3.1 Reliability of the new method

Since numerical analysis by Comins & Schutz [5] is the only work which has treated the same problem discussed in this paper, we have to compare our results here with theirs. However we could not obtain the same results as theirs. After we have tried to find the cause of the discrepancies, we have found that, in their paper, there are some strange treatments about unperturbed models if the parameters cited in their paper were correct.

As unperturbed stellar models are assumed to have a homogeneous density distribution and rotate uniformly, models can be specified by two parameters, i.e. a dimensionless rotational period P and strength of gravity μ. Here the dimensionless rotational period is defined by

$$\tilde{P} = P/M, \tag{14}$$

and strength of gravity is measured by the following quantity:

$$\mu = 2M/R, \tag{15}$$

where M and R are the mass and the radius of the model, respectively. Even if we choose exactly the same parameters as those used in Comins & Schutz's paper [5], we have apparently different potentials $V_\pm$ for which positions of extrema of potentials are different and the size of ergoregion is different and so on (However we can reproduce the same quantitative characters of the equilibrium ergoregion with those

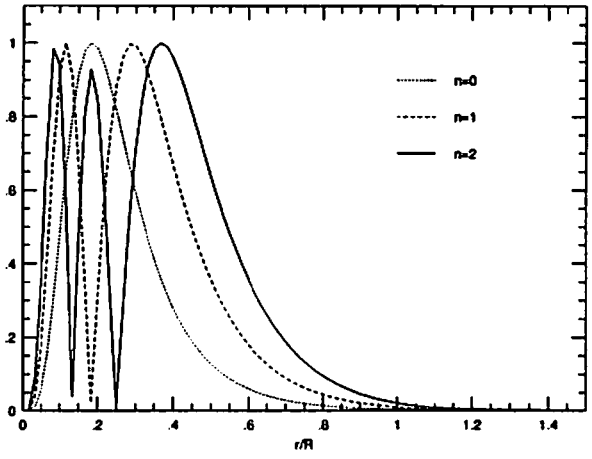


Figure 1: Three eigenfunctions of different modes for $m = 3$. Model parameters are $\tilde{P} = 30$ and $\mu = 0.88$. Horizontal axis denotes the nondimensional distance r/R, whereas the vertical axis is the absolute value of the complex eigenfunctions. The dotted, dashed and solid curves correspond to modes with no node, one node and two nodes, respectively.

by Schutz & Comins [17].). This discrepancy of the unperturbed models makes it difficult to compare our results obtained from our new method with their results. Therefore, in order to check our own results we have carried out the WKBJ analysis numerically by coding equations for the WKBJ approximation as mentioned in Appendix. [4]

In Table 1 results obtained by two different methods, i.e. our new scheme explained before and our WKBJ code, are given. It should be noted that in the WKBJ approximation the term in Eq.(8) which becomes very small in high-m approximation is neglected.

As seen from the table, results agree well each other, in particular for models with strong gravity and rapid rotation. It implies that our new method can give satisfactory values. In order to check our results further, we have tested our code as follows. As is explained in the previous chapter, we divide the space into two regions. Thus we can choose different mesh sizes for the two regions. Furthermore, as the outer boundary condition is imposed at a finite distance from the center of the star, we need to check the effect of this distance, i.e. the cutoff distance, on the eigenvalues. We have computed eigenvalues by changing these quantities for the model with $\tilde{P} = 30$ and $\mu = 0.88$. The investigated mode is $m = 3$.

When we double the cutoff distance by keeping the mesh size unchanged, the change of the eigenvalue is a fraction of 10^{-4} for the real part and a few percent for the imaginary part. The influence of the size of the region to be used for the integral representation, i.e. the size of the outer region is also very small to eigenvalues. It causes a very tiny change of the imaginary part, i.e. only a fraction of 10^{-4}. The most influential factor to results is the mesh size in the stellar interior. This can be expected from the behavior of the scalar wave in the ergoregion as seen in Fig.1. If we make the mesh size half, the change

[4]Numerical calculations here were performed mainly on a DEC 3000 AXP workstation at the Department of Astronomy and partially on a HP9000 workstation at the Department of Earth Science and Astronomy.

	WKBJ	our new method
Model 1	-0.138 - $i\cdot$ 2.60E-19	-0.138 - $i\cdot$ 3.66E-19
Model 2	-9.94E-2 - $i\cdot$ 7.64E-22	-9.92E-2 - $i\cdot$ 1.12E-21
Model 3	-5.61E-2 - $i\cdot$ 4.31E-26	-5.59E-2 - $i\cdot$ 6.22E-26

Table 1: Comparison of the results obtained by using the WKBJ analysis and our new method is shown here. Eigenvalues of $l = m = 7$ fundamental mode ($n = 0$ mode in eq.(A2)) calculated with two methods are compared. Each model has the same strength of gravity, $\mu = 0.88$, and rotational periods of them are $\tilde{P} = 30.0, 37.0, 50.0$ for model 1,2,3.

of eigenvalues amounts to a fraction of 10^{-3} for the real part and about ten percent for the imaginary part.

3.2 Modes with small azimuthal numbers m

In the WKBJ analysis by Comins & Schutz [5] e-folding times of unstable high m modes are shown to grow exponentially as m increases. Since these e-folding times are much longer than the age of the universe, the instability seems to be too weak to have any effect in this large m limit. They also have investigated the lowest unstable modes of particular models (as small as $m = 4$), the results of which show that the unstable modes grow with much longer time than the age of the universe. However lower modes, which are expected to have shorter e-folding times, cannot be investigated by the WKBJ analysis precisely. Thus it is important to study these lower modes by our new code exactly. We here have analyzed these lower modes with models having stronger gravity than theirs, where the effect of ergoregion would be more prominent.

Figure 2 shows how the real part of the eigenfrequency and the e-folding time of unstable modes change with azimuthal number m for three models with different parameters, i.e. ($\tilde{P} = 30, \mu = 0.88$), ($\tilde{P} = 40, \mu = 0.88$) and ($\tilde{P} = 30, \mu = 0.85$). It is clear that differences between the results of the WKBJ method and those of our method increases with decreasing m. We can see that e-folding times obtained by the WKBJ method are larger for the same m's. However the differences are at most in factors of ten. Thus, even if the inaccuracy of WKBJ method is taken into consideration, the results obtained here do not contradict qualitatively the conclusion by Comins & Schutz [5] that the unstable mode grow very slowly comparing the age of the universe, which assumed weaker gravity.

In Figure 3 the e-folding times of unstable modes for different values of m are plotted against the rotational period. These e-folding times are computed for the model with $\mu = 0.88$. We can see from this figure that the shorter the rotation period becomes, the more unstable the model becomes. This is because the ergoregion becomes wider and deeper as the rotation becomes faster and so the ergoregion instability works efficiently. Times are measured in a unit of M, so dimensionless time $\tilde{t}$ corresponds to:

$$\tilde{t}\cdot\left[\frac{GM}{c^3}\right] = \kappa\tilde{t}\cdot 4.9\times 10^{-6}(\mathrm{sec}) \tag{16}$$

for a κ solar mass model. For example the model with $\tilde{P} = 40$ rotates with period of 0.3msec for $1.4M_\odot$.

It is remarkable that as shown in Figure 3 lower unstable modes can have much shorter e-folding time than the age of the universe for stellar models with the mass range of ordinary compact stars.

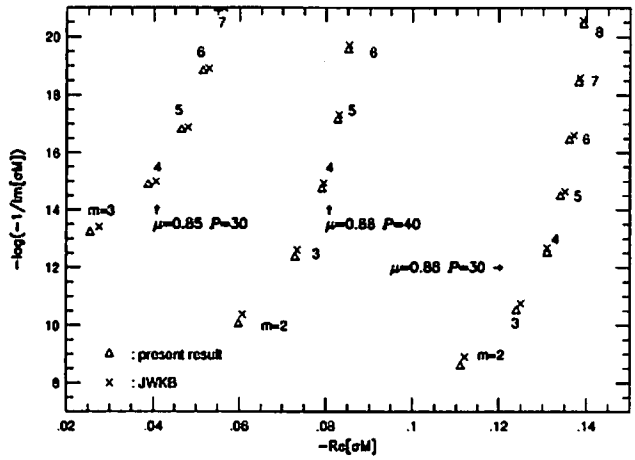


Figure 2: For comparison of WKBJ and our new method, e-folding times of unstable modes are plotted against the real part of the corresponding eigenfrequencies. Triangles denote the present results obtained by using our new method. Crosses are results of the WKBJ method. Three pairs of parameters $(\tilde{P}, \mu)$ are selected, i.e. $(\tilde{P} = 30, \mu = 0.88)$,$(\tilde{P} = 40, \mu = 0.88)$ and $(\tilde{P} = 30, \mu = 0.85)$, which form three sequences. Integers beside each points are the mode numbers m. Eigenfrequencies are normalized by using the stellar mass M.

4 Concluding remarks

We have reanalyzed the ergoregion instability of the scalar field by developing a new method totally different from the WKBJ analysis. Results obtained by our new method agree very well with those of the WKBJ analysis for high m's. Furthermore this method can also be used to investigate modes with low m's which give shorter e-folding times than the age of the universe.

Models investigated in this paper are, however, *not realistic* in two points. First, models have a homogeneous density and extremely strong gravity up to $\mu = 0.88$. It should be reminded that there exists no equilibrium state of the star with a uniform density for $\mu > 8/9$. Realistic stellar matters are inferred to be inhomogeneous and no equation of state currently known cannot support such strong gravity. Second, although the stars rotate rapidly, deformation from spherical shapes has not been taken into consideration at all.

Concerning the second point, the deformation is expected to have non-negligible effects on the eigenvalues. However, it is almost impossible to take this deformation into consideration by the WKBJ approach. Thus we need to investigate the problem by a full two dimensional formulation. Moreover it is shown that there are toroid-star systems with the ergoregion due to strong self gravity of toroids (Komatsu et al. [13], Nishida & Eriguchi [15]). Since these toroids with the ergoregion can consist of realistic nuclear matter, they are more interesting objects from the astrophysical point of view. Stability of these configurations have to be also treated as a full 2-D problem. Our new method has an advantage in this respect since the formulation does not depend on special nature of spherical symmetry. The main and essential point of our formulation is to reorganize basic equations into "wave equations" which becomes inhomogeneous Helmholtz equations asymptotically. Since this scheme is independent of the dimension of the space, our

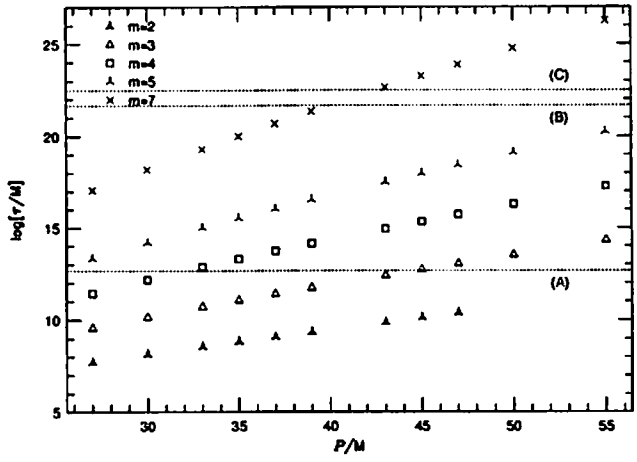


Figure 3: With m fixed dimensionless e-folding times of unstable modes are shown as a function of dimensionless rotation period P. Strength of gravity is fixed as $\mu = 0.88$. Dotted lines correspond to 1yr(A),1Gyr(B) and 10Gyr(C) for $1.4M_\odot$ model.

formulation can be easily extended to higher dimensions and will be applied to two dimensional problems in the near future.

Acknowledgments

We would like to thank Dr. Shijun Yoshida for his discussion. This work was supported by the Grant-in-Aid for Scientific Research of Japanese Ministry of Education, Science and Culture.

A WKBJ analysis

If an eigenvalue Σ satisfies the condition $V_+(r_+) < \Sigma < 0$ where r_+ denotes the radius at which the potential V_+ becomes minimum, this mode is unstable one which grows in time. In this case, the equation $V_+ = \Sigma$ has two roots, say $r = r_0$ and r_1, while $V_- = \Sigma$ has one root, say $r = r_2$. These values satisfy the relation $r_0 < r_1 < r_2$. According to the WKBJ analysis, the real part of this eigenvalue is determined by the so-called "quantization condition of Bohr-Sommerfeld",

$$m \int_{r_0}^{r_1} \sqrt{T} dr = \left(n + \frac{1}{2}\right)\pi. \qquad (n = 0, 1, 2 \ldots) \tag{17}$$

This condition implies that eigenstates are quasi-stationary states of the potential V_+. Furthermore imaginary parts $(1/\tau)$ are given by the following integral,

$$\alpha_n = \frac{d}{d\sigma}\left[m\int_{r_0}^{r_1}\sqrt{T}dr\right]\Big|_{\sigma=\sigma_n}, \tag{18}$$

and

$$\lambda = \exp\left[m\int_{r_1}^{r_2}\sqrt{|T|}dr\right]\Big|_{\sigma=\sigma_n}. \tag{19}$$

As $\tau \equiv 4\lambda^2\alpha_n$, λ is regarded as a wave transparency coefficient through the evanescent region ($r_1 < r < r_2$).

Thus if one unperturbed model of a star is specified, unstable modes can be calculated with these formulae.

B Axial modes of non-rotating ultra-compact stars

In general, nonradial perturbations of relativistic spherical stars can be decomposed into two classes of different parities by using tensor spherical harmonics. One is the *polar* or *even-parity* perturbation and the other is the *axial* or *odd-parity* perturbation. Perturbations of the latter class are disturbances of pure gravitational field without dynamical degrees of freedom for stellar matter. This perturbation has no counterpart in the Newtonian dynamics. This class of perturbations can be investigated with ease because of its mathematical and numerical simplicity. The equation describing this class of perturbations can be written as follows with an appropriate gauge choice and by expanding in terms of spherical harmonics (Kokkotas [11]):

$$\frac{d^2Z}{dr_*^2} + (\sigma^2 - V(r))Z(r) = 0 \tag{20}$$

where r_* is the "tortoise" coordinate defined as

$$r_* = \int_0^r \sqrt{\left|\frac{g_{11}}{g_{00}}\right|}dr, \tag{21}$$

and the function Z represents the gravitational field and the time-dependence is assumed as $\sim e^{i\sigma t}$.

An extremely dense and compact stellar model has a potential V with "dip". This dip enables "resonance modes"to exist similar to quasi-stable states of nuclei. This resonance has been analyzed by Chandrasekhar & Ferrari ([3, 4]) and Kokkotas [11]. We have investigated the same problem using the method developed in this paper.[5] As for the boundary conditions, we impose that the variable Z is regular at the origin and that it satisfies the outgoing radiation condition at infinity.

In Table A1 eigenfrequencies of axial modes with $l = 2$ of ultra-compact stars are given. Here K(1994) and C&F(1991) denote results of Kokkotas [11] and Chandrasekhar & Ferrari ([4]), respectively. Three results agree well.[6] In our formulation, since we must solve a system of linear equations whose dimension is [(number of variables) × (mesh number)]^2, it is much harder to increase mesh numbers than the former calculations. However, considering the agreement of our results to others, our new method can be said to give satisfactory results.

[5] In our numerical computation r coordinate is used instead of r_*.

[6] Kokkotas recently presented errratum of his results.[12] These values shown here are the corrected ones.

Table 2: Eigenfrequencies of axial modes with $l = 2$ of ultra-compact homogeneous stars.

R/M	K(1994)	C&F(1991)	Present
2.26	0.2139+i·2.430E-9	0.214+i·2.3E-9	0.214+i·2.58E-9
	0.2910+i·7.747E-8	—	0.290+i·8.04E-8
2.28	0.3690+i·1.231E-6	0.369+i·1.2E-6	0.369+i·1.27E-6
	0.5008+i·5.294E-5	—	0.500+i·5.37E-5
2.30	0.4735+i·2.503E-5	0.474+i·2.6E-5	0.473+i·2.56E-5
	0.6372+i·1.156E-3	—	0.637+i·1.12E-3

Eigenfrequencies of modes with different radial quantum numbers $n = 0$ (upper row for each stellar model) and $n = 1$ (lower row for each stellar model) are displayed. The eigenfrequencies are shown in units of $(3M/R^3)^{1/2}$.

References

[1] Butterworth E.M., Ipser J.R.,1975,ApJ, **204**,200

[2] Chandrasekhar S., 1970, Phys.Rev.Lett., **24**,611

[3] Chandrasekhar S., Ferrari V., 1991a, Proc.Roy.Soc.Lond., **A432**,247

[4] Chandrasekhar S., Ferrari V., 1991b, Proc.Roy.Soc.Lond., **A434**,449

[5] Comins N., Schutz B.F., 1978, Proc.Roy.Soc.Lond., **A364**,211

[6] Friedman J.L., 1978, Comm.Math.Phys., **62**,247

[7] Friedman J.L., 1978, Comm.Math.Phys., **63**,243

[8] Friedman J.L., Schutz B.F., 1978, ApJ, **222**,281

[9] Hartle J.B., Thorne K.S., 1968, ApJ, **153**,807

[10] Ipser J.R., Lindblom L., 1991, ApJ, **373**,213

[11] Kokkotas K.D., 1994, MNRAS, **268**,1015

[12] Kokkotas K.D., 1995, MNRAS, **277**,1599

[13] Komatsu H., Eriguchi Y., Hachisu I., 1989, MNRAS, **239**,153

[14] Misner C.W., Thorne K.S., Wheeler J.A., 1973, Gravitation, W.H.Freeman and Company, San Francisco

[15] Nishida S., Eriguchi Y., 1996, ApJ, in press

[16] Sato H., Maeda K., 1978, Prog.Theor.Phys., **59**,1173

[17] Schutz B.F., Comins N., 1978, MNRAS, **182**,69

[18] Yoshida S., Eriguchi Y., 1995, ApJ, **438**,830

Quasinormal modes of maximally charged black holes[1]

Hisashi Onozawa, Takashi Mishima, Takashi Okamura, and Hideki Ishihara

Department of Physics, Tokyo Institute of Technology,
Oh-okayama, Meguro, Tokyo 152, Japan

Email address: onozawa@phys.titech.ac.jp

The quasinormal modes of the extremal Reissner-Nordström black hole are accurately calculated by a continued fraction method and the curious behavior of the quasinormal frequencies in the extremal limit is reported.

Introduction

The determination of the quasinormal frequencies of black holes has for several years been of interest and is related to experiments which aim at detecting gravitational waves from supernovae or coalescences of binary neutron stars, which are thought to eventually form a black hole. In the late stage of the process of black hole formation, a certain mode of the gravitational wave dominates the emission. This is called the quasinormal mode of a black hole.

The quasinormal frequencies of the Schwarzschild black holes were first computed by Chandrasekhar and Detweiler[1], who treated it as a boundary value problem of the second order ordinary differential equation of Regge-Wheeler.

$$\left[\frac{d^2}{dr_*^2} + \omega^2 - V_s(r)\right] Z_s(r) = 0. \tag{1}$$

The boundary conditions are given by

$$Z_s \sim e^{-\rho r_*} \quad \text{as} \quad r_* \to \infty, \tag{2}$$

$$Z_s \sim e^{\rho r_*} \quad \text{as} \quad r_* \to -\infty, \tag{3}$$

which mean there only exists purely outgoing wave at the infinity and purely ingoing wave at the horizon. Integration of the Regge-Wheeler equation under these conditions is numerically unstable. To avoid the instability, Leaver[2] presents the numerically stable continued fraction method for

[1]This talk was based on the preprint "TIT/HEP-311/COSMO-61"

Schwarzschild and Kerr black holes. Here, a quasinormal mode function corresponds to a minimal solution of the recurrence relations which are satisfied by the coefficients of a series expansion of a wave function. The minimal solution of the three-term recurrence relation is obtained by calculating the corresponding continued fraction. The continued fraction method can give the values of frequencies with high numerical precision because it uses no approximation, as is the case in the semi-analytic WKB method. Leaver[3] also generalizes the continued fraction method to calculate accurate values of the quasinormal frequencies of Reissner-Nordström black holes.

In the extremal limit, however, the wave equation has an irregular singular point at the horizon, which makes the series expansion of a solution there invalid, and thus a continued fraction method does not seem to work. In this paper we show that a continued fraction method is applicable even to the extremal black holes if we expand a solution about a suitable ordinary point.

Continued Fraction Method for Extremal Cases

The idea is as follows: Since we can not expand a solution about the horizon, we expand it about the middle point between the horizon and the infinity to get a recurrence relation of the expansion coefficients (a_n). The expansion about the middle point makes us possible to simultaneously examine the convergences of the expansion at both boundaries (the horizon and infinity). The convergence at the infinity puts the condition $\Sigma a_n < \infty$ and that at the horizon puts the condition $\Sigma(-)^n a_n < \infty$. These two conditions are satisfied if both a_{2n} and a_{2n+1} are convergent. Then, we need to divide the sequence into an odd sequence and an even sequence to examine convergence of both series. The quasinormal boundary conditions are satisfied if each separated sequence is minimal. The ratio of successive terms of each sequence is given by a continued fraction. Consequently, the eigenvalue equation, which corresponds to the matching condition at the middle point, is constructed by substituting the continued fraction into the ratio of successive recurrence relations

The Results and Conclusion

The eigenvalue equation was solved using a nonlinear root search code. We compared our results with Leaver's results[3] of nearly maximally charged case and the third order WKB quasinormal frequencies of the extremal Reissner-Nordström black hole. The WKB quasinormal frequencies were

obtained using the same formula of Kokkotas and Schutz[4]. Our results were in agreement with these results within the accuracy of a few percent. The difference between WKB frequencies and ours is ascribed to the fact that the WKB method gives only the approximate values. Indeed, the tendency that the discrepancy grows with the mode number suggests the breakdown of the WKB approximation for higher n. The difference between Leaver's results and ours is from the numerical error caused by the breakdown of his series expansion at the limit of maximal charge. Thus we are sure that our quasinormal frequencies of the extremal Reissner-Nordström black hole are the most accurate.

We numerically find a remarkable coincidence in the quasinormal frequencies of the gravitational waves of multi-pole index l and those of electromagnetic waves of $l-1$ for a wide rage of multipole indices, $2 \leq l \leq 10$. This can be explicitly seen in Fig.1, where the trajectories of the first four lowest modes of gravitational quasinormal frequencies of $l = 2, 3, 4, 5$ and those of electromagnetic quasinormal frequencies of $l = 1, 2, 3, 4$ are plotted. The quasinormal frequencies of the gravitational waves surprisingly coincide with those of the electromagnetic waves in the limit of maximal charge.

These two modes are completely decoupled but they have the same resonant frequencies. The situation is very similar to the coincidence between the quasinormal frequencies of odd parity pertrubation and those of even parity perturbation. At present there is no easy way to understand the hitherto unobserved coincidence but it is interesting that it occurs only in the extremal limit, where the black hole may have an unknown symmetry.

References

[1] S. Chandrasekhar and S. Detweiler, *Proc. Roy. Soc. London* **A344** (1975) 441.

[2] E. Leaver, *Proc. Roy. Soc. London* **A402** (1985) 285.

[3] E. Leaver, *Phys. Rev. D* **41** (1990) 2986.

[4] K. Kokkotas and B. Schutz, *Phys. Rev. D* **37** (1988) 3378.

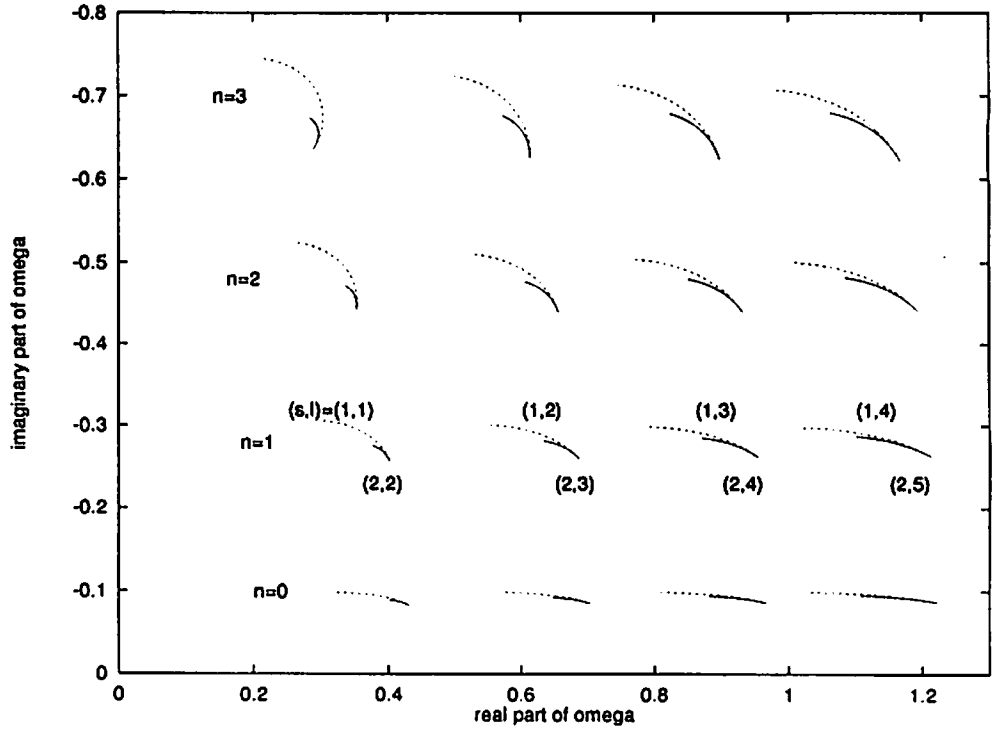


Figure 1: Solid lines and dashed lines are trajectories of the quasinormal frequencies of the gravitational and electromagnetic wave, respectively. Each left endpoint of lines corresponds to the quasinormal frequency of a charged black hole of $Q = 0.8$, and each right endpoint corresponds to the frequency in the limit of maximal charge. A trajectory of the gravitational quasinormal frequencies depicted in a solid line meets at the right end with a corresponding dashed line, which is a trajectory of the electromagnetic quasinormal frequencies belonging to lower multi-pole index by one.

The thermodynamic second law of 2D dilaton black hole

Kouichirou Hirata

Graduate School of Human and Environmental Studies,
Kyoto University, Kyoto 606, Japan

Yoshi Fujiwara

Yukawa Institute for Theoretical Physics,
Kyoto University, Kyoto 606, Japan

Jiro Soda

Department of Fundamental Sciences, FIHS,
Kyoto University, Kyoto 606, Japan

Abstract

The thermodynamic second law in the evaporating black hole space-time is examined in the context of two-dimensional dilatonic black hole. For most general semi-classical action which admit the linear dilatonic vacuum solution, the black hole entropy proposed by Wald is calculated by using the analytical perturbation method and numerical method. It is shown that the thermodynamic second law holds in the vicinity of infalling shell and in the region far from the infalling shell in the case of semi-classical CGHS model. The analysis of the general models is also presented and its implication is discussed.

1 Introduction

At present, the thorough understanding of quantum gravity has not been achieved yet. To reveal the nature of quantum gravity, it is important to figure out the quantum aspects of black hole. The existence of the non-trivial causal space-time structure such as black holes gives a profound insight into the character of gravity. Curiously, the geometry of black hole is intimately related to thermodynamics [?]. The discovery of Hawking radiation [?] stressed the role of the black hole thermodynamics. It also gives the picture of an evaporating black holes. Recently, the problem of information loss for the evaporating black hole has been discussed widely. Originally, the thermodynamic laws of black hole were proved in a stationary situation. Hawking's calculation was performed on a fixed

stationary black hole. However, Hawking radiation naturally leads us to the evaporating black hole picture. Hence, it is legitimate to ask how the thermodynamic laws are modified.

In this paper, we shall investigate these problems in the context of two-dimensional dilaton gravity. The two-dimensional CGHS model [?] provides an interesting toy-model for the study of black hole thermodynamics. For this model, it is easy to give the classical general solution. There exists a black hole solution. Moreover, the thermodynamic laws hold in this model[?] and the standard calculation yields the Hawking radiation with temperature $\frac{\lambda}{2\pi}$. It was shown that the effective action in which the backreaction of the Hawking radiations included gives the evaporating black hole picture. On the other hand, using the Noether charge technique invented by Wald [?], Myers has given a dynamical entropy and shows the existence of the thermodynamic second law in the RST model [?][?]. The question we would like to address is whether the proposed dynamical entropy is really "entropy" in the sense that the second law holds. Myers has shown that this is true in the special model, i.e., RST model. However, RST model is too special to conclude that the thermodynamic second law holds in general. We would like to consider the most general models which have freedom in choosing coefficients for local counter terms. Because we are considering the evaporating black hole picture, we restrict the models to have the linear dilaton vacuum solution.

2 Two dimensional dilaton gravity model

To make this paper self-contained, we start from classical CGHS model. The classical CGHS action is

$$I_0 = \frac{1}{2\pi}\int d^2x\sqrt{-g}\{e^{-2\phi}[R + 4(\nabla\phi)^2 + 4\lambda^2] - \frac{1}{2}\sum_{i=1}^{N}(\nabla f_i)^2\} \tag{1}$$

where ϕ is dilaton field, R is 2d Ricci scalar, λ is a positive constant, ∇ is the covariant derivative associated with 2d metric $g_{\mu\nu}$, and f_i are N massless scalar fields. In conformal gauge, $g_{+-} = -\frac{1}{2}e^{2\rho}$, $g_{++} = g_{--} = 0$, $x^{\pm} = t \pm r$, exact solution of this model can be obtained easily. Vacuum solution of this model in $\phi = \rho$ gauge is

$$e^{-2\rho} = e^{-2\phi} = M/\lambda - \lambda^2 x^+ x^-. \tag{2}$$

It can be easily shown that the Penrose diagram of this spacetime is the same as that of Schwarzschild's spacetime. This fact shows that CGHS model has a black hole solution. The solution for $M = 0$ is called Linear Dilaton Vacuum(LDV), which is flat solution.

Let us consider the matter falling along the null line, $x^+ = x_0^+$. Then the energy-momentum tensor of the scalar fields becomes

$$\frac{1}{2}\partial_+ f_i \partial_+ f_i = \frac{M}{\lambda x_0^+}\delta(x^+ - x_0^+), \tag{3}$$

$$\frac{1}{2}\partial_- f_i \partial_- f_i = \frac{1}{2}\partial_+ f_i \partial_- f_i = 0. \tag{4}$$

The solution for this situation is

$$e^{-2\phi} = e^{-2\rho} = -\frac{M}{\lambda x_0^+}(x^+ - x_0^+)\theta(x^+ - x_0^+) - \lambda^2 x^+ x^-, \tag{5}$$

where M is the ADM mass of black hole.

Now we turn to the quantum theory. Though we cannot perform the full quantization of gravity, we can calculate matter contribution and obtain effective action. In covariant non-local form, quantum correction which is called Liouville action is

$$I_1 = -\frac{\kappa}{8\pi}\int d^2x\sqrt{-g}R\frac{1}{\nabla^2}R \tag{6}$$

where $\kappa = N/12$. Other covariant local counter terms can be added to the action because of ambiguity in defining measures used in the functional integral [?]. Combinations of these counter terms make different models. Possible local counter terms which can be added to the action are

$$I_2 = -\frac{\kappa}{8\pi}\int d^2x\sqrt{-g}\{\alpha\phi R + \beta(\nabla\phi)^2\}, \tag{7}$$

$$I_3 = -\frac{\kappa}{8\pi}\int d^2x\sqrt{-g}\sum_{n=2}^{K}\{a_n\phi^n R + b_n\phi^{n-1}(\nabla\phi)^2\}. \tag{8}$$

Because we will consider the black hole solutions which are asymptotically flat, it is reasonable to require that models include the LDV solution. This requirement leads $a_n = b_n = 0$. Thus the action I_3 will not be considered below. As particular cases, it is known that the models are exactly soluble if $\alpha/2 + \beta/4 - 1 = 0$. For example, the case of $\alpha = 2$, $\beta = 0$ is known as RST model, and the case of $\alpha = 4$, $\beta = -4$ was studied by Bose *et.al* [?]. In the cases $\alpha/2 + \beta/4 - 1 > 0$, there is no singularity. As we would like to consider the entropy of the evaporating black holes, this case is excluded from our consideration.

3 Entropy as Noether charge

In this section, we briefly review the Noether charge approach to black hole entropy invented by Wald [?] and, following the work by Myers [?], apply it to the two-dimensional dilaton gravity models.

Wald presented a new way to look at the entropy of black hole as a Noether charge associated with diffeomorphism invariance of a theory. In particular, the first law of black hole mechanics is a general relation between "boundary terms" at event horizon and spatial infinity in a stationary asymptotically flat space-time.

Assumption made in Wald's analysis is the presence of a bifurcation surface on a Killing horizon in a stationary space-time. A Killing horizon is a null hypersurface whose null generators are orbits of a Killing field. A bifurcation surface is a spacelike cross section in the Killing horizon, where the Killing field vanishes. It is a set of fixed points of the isometry generated by the Killing vector and is at the intersection of two Killing horizons. Another essential ingredient is the assumption that the global charges such as mass and angular momentum are well defined. This is usually guaranteed by some appropriate boundary conditions of asymptotic flatness. Both of these assumptions are valid in the classical theory of the dilaton gravity models, as we can directly check for the exact solutions of black holes. On the other hand, *First law valid for quantum black holes?*

The first law of black hole mechanics takes the general form:

$$\frac{k}{2\pi}\delta S = \delta M - \Omega\,\delta J \tag{9}$$

where k is the surface gravity, and M, J and Ω are the mass, angular momentum and angular velocity of the black hole. The entropy S is given by a local geometrical expression over the bifurcation surface. Iyer and Wald derives an inductive algorithm for general diffeomorphism invariant theories [?]. Especially, for a Lagrangian of the form $L = L(f, \nabla_a f, g_{ab}, R_{abcd})$ involving only second derivatives of the metric g_{ab} and the first derivatives of the matter field f, the expression is given by

$$S = \frac{4\pi}{\sqrt{-g}}\frac{\partial L}{\partial R}. \tag{10}$$

The integral is evaluated over the bifurcation surface. (Actually it may be evaluated at any cross section of the Killing horizon and yields the same value provided that the bifurcation surface is regular.) It is understood that the integration is done with respect to the naturally induced volume element in the bifurcation surface. For the classical two-dimensional dilaton gravity, the action (??) gives the classical entropy expression as

$$S_0 = 2e^{-2\phi} \tag{11}$$

where the right hand side is evaluated on the event horizon (which is a killing horizon) for the black hole solution (??).

Let us now turn to the semi-classical theories. Since the local action fits into the above frame work, the correction term I_2 of the action yields

$$S_2 = -\frac{\kappa\alpha}{2}\phi \tag{12}$$

On the other hand, the nonlocal action needs careful argument in the Noether charge analysis. Myers showed that such a nonlocal term actually can be handled. The result is a nonlocal expression for the entropy:

$$S_1 = -\kappa \int d^2y\sqrt{-g(y)}\,G(x_H, y)R(y) \tag{13}$$

where $G(x, y)$ is the Green function, that is,

$$\nabla_x^2 G(x, y) = \delta^2(x, y) \tag{14}$$

We have to specify the boundary condition for the Green function $G(x, y)$ in order to make the expression (??) give definite value.

Thus entropy of the generalized model is

$$S = 2e^{-2\phi} + 2\kappa\rho - \frac{\kappa\alpha}{2}\phi + \kappa\log(-\lambda^2 x^+ x^-), \tag{15}$$

second line is written in conformal gauge, and last extra term came from boundary condition of (??).

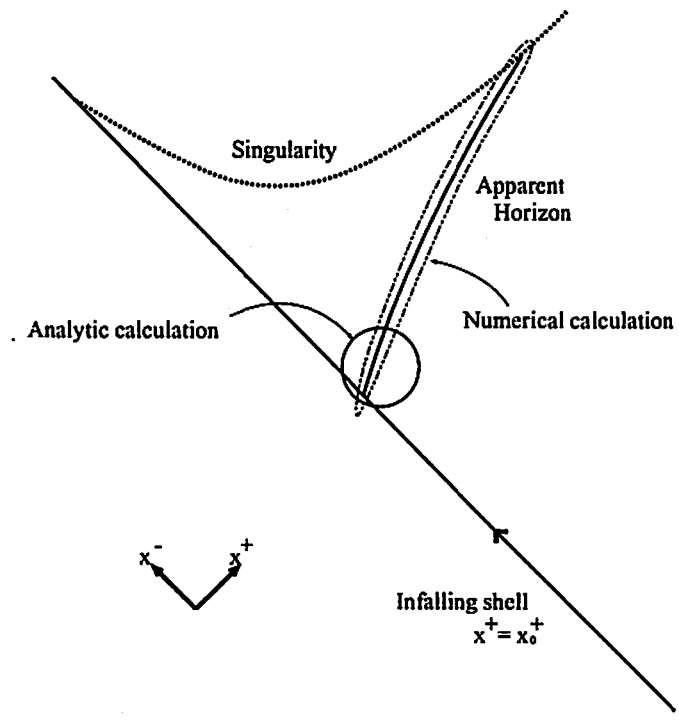


Figure 1: This figure shows schematic explanation of situation considered here.

4 Time evolution of entropy on apparent horizon

4.1 Analytic Approach

Time evolution of entropy on apparent horizon is studied in this section. We consider the dynamical situation in which infalling matters make black hole just like as classical situation, but also with the back reaction of Hawking radiation. We would like to evaluate the entropy along the apparent horizon. Unfortunately, in most cases of α and β, exact solution cannot be found, so that the behavior of entropy on apparent horizon is difficult to study. Instead we calculate the gradient of the entropy at the vicinity of the matter shell. At the point where apparent horizon starts, the derivative of entropy in x^+ direction is

$$\frac{dS}{dx^+} = \partial_+ S + \frac{dx^-}{dx^+}\Big|_{\text{A.H.}} \partial_- S, \tag{16}$$

where

$$\partial_+ S = 2\kappa\, \partial_+ \rho + \frac{\kappa}{x_0^+}, \tag{17}$$

$$\partial_- S = -\lambda^2 x_0^+ \left(2 + \frac{\kappa\alpha}{4e^{-2\phi}}\right), \tag{18}$$

on apparent horizon which is defined as $\partial_+\phi = 0$. $\frac{dx^-}{dx^+}|_{A.H.}$ is the gradient of the apparent horizon, and can be calculated from equations of motion[?],

$$\frac{dx^-}{dx^+}\Big|_{A.H.} = -\frac{\partial_+^2\phi}{\partial_+\partial_-\phi}\Big|_{A.H.}.$$

So, we must have $\partial_+\rho$, $\partial_+^2\phi$ and $\partial_+\partial_-\phi$ on the apparent horizon. First we lead $\partial_+\rho$ and evaluate $\partial_+ S$.

On the shell line ($x^+ = x_0^+$), we define $\Sigma = \partial_+\phi$, $\Xi = \partial_+\rho$, and $w = e^{-2\phi} = e^{-2\rho} = -\lambda^2 x_0^+ x^-$, which is LDV solution on $x^+ = x_0^+$. Boundary conditions for ϕ and ρ is such that they asymptotically coincide with the classical solution because quantum correction would be negligible in asymptotic region. The equations of motion in terms of Σ and Ξ are

$$\{2w - \kappa(\frac{\alpha}{4}+1)\}\partial_w\Xi - \{2w - \frac{\kappa}{4}(\alpha+\beta)\}\partial_w\Sigma = 0, \tag{19}$$

$$\kappa\partial_w\Xi - (2w + \frac{\alpha\kappa}{4})\partial_w\Sigma - 2\Sigma - \frac{1}{x_0^+} = 0, \tag{20}$$

By eliminating $\partial_w\Xi$, we have

$$\partial_w\Sigma\,[\{2w + \kappa(\frac{\alpha}{4}-1)\}^2 - \kappa^2(1 - \frac{\alpha}{2} - \frac{\beta}{4})] + \frac{1}{2}\Sigma\,[8w + 4\kappa(\frac{\alpha}{4}-1)] = -\frac{1}{4x_0^+}[8w + 4\kappa(\frac{\alpha}{4}-1)]. \tag{21}$$

The solution of this equation can be easily obtained as,

$$\Sigma = -\frac{1}{2x_0^+} + \text{const.}\frac{1}{\sqrt{A}}. \quad A = \{2w + \kappa(\frac{\alpha}{4}-1)\}^2 - \kappa^2(1 - \frac{\alpha}{2} - \frac{\beta}{4}). \tag{22}$$

Integration constant in (??) must be determined by the boundary condition that ϕ asymptotically approach to the classical solution. Thus

$$\Sigma = -\frac{1}{2x_0^+} + \frac{M}{\lambda x_0^+}\frac{1}{\sqrt{A}}. \tag{23}$$

In the case of $1 - \frac{\alpha}{2} - \frac{\beta}{4} \geq 0$ there are several solutions, but we are interested in the largest one because the other solutions are behind singularity. So, we take the largest one as the only solution, w_0. By solving $\Sigma = 0$, we have

$$w_0 = -\frac{\kappa}{2}(\frac{\alpha}{4}-1) + \frac{1}{2}\sqrt{\kappa^2(1 - \frac{\alpha}{2} - \frac{\beta}{4}) + (\frac{2M}{\lambda})^2}. \tag{24}$$

Another case is $1 - \frac{\alpha}{2} - \frac{\beta}{4} < 0$, which always gives positive A for every w. This fact implies that (??) is always regular, so singularity is absent even in our matter infalling situation. Detail of this case will not be analyzed here because we are now interested in black hole entropy.

Equation for Ξ is given by using (??), and the solution satisfying the boundary condition is

$$\Xi = (\frac{\alpha}{4} + \frac{2w}{\kappa})\Sigma + \frac{1}{2\kappa x_0^+}\{2w + \kappa(\frac{\alpha}{4}-1)\} - \frac{M}{\lambda\kappa x_0^+}.$$

Using this solution and $\Sigma = 0$, the evolution of the entropy in x^+ direction is given by

$$\partial_+ S|_{A.H.} = \frac{1}{x_0^+}\left\{\kappa - \frac{2M}{\lambda} + \sqrt{\kappa^2\left(1 - \frac{\alpha}{2} - \frac{\beta}{4}\right) + \left(\frac{2M}{\lambda}\right)^2}\right\}. \tag{25}$$

For the case $1 - \frac{\alpha}{2} - \frac{\beta}{4} > 0$, (??) is always equal to or larger than κ/x_0^+ (> 0). Particularly in the exactly soluble case of $1 - \frac{\alpha}{2} - \frac{\beta}{4} = 0$, $\partial_+ S|_{A.H.} = \kappa/x_0^+$.

Next, we calculate another part of dS/dx^+. The gradient of the apparent horizon contains $\partial_+\partial_-\phi$ and $\partial_+^2\phi$, the former can be estimate from equations of motion directly, and the latter can be done from derivatives of equations. Eliminating $\partial_+\partial_-\rho$ from the equations of motion, one can obtain the formula of $\partial_+\partial_+\phi$, and its value on the apparent horizon is

$$\partial_+\partial_-\phi|_{A.H.} = \frac{\lambda^4}{16M^2}A_0'. \tag{26}$$

where $'$ denotes ∂_w and the subscript letter '$_0$' denotes the evaluation on apparent horizon. Now, we define new variable $\mathcal{F} \equiv \partial_+^2\phi$. Combining above $\partial_+\partial_-\phi$ and $\partial_- S$, we have

$$\left.\frac{dx^-}{dx^+}\right|_{A.H.}\partial_- S = \frac{32M^2x_0^+}{\lambda^2}\frac{A_0' + 4\kappa}{A_0' - 4\kappa(\frac{\alpha}{4} - 1)}\frac{\mathcal{F}_0}{A_0'}. \tag{27}$$

Differentiating equations of motion and erasing $\partial_+^2\rho$, we have differential equation for $\mathcal{F}$,

$$A\mathcal{F}' + \frac{1}{2}A'\mathcal{F} = A'\partial_w(w\Sigma^2) - \frac{A'}{2x_0^+}(\Xi - \Sigma) + 4\kappa w\Sigma\partial_w(\Xi - \Sigma). \tag{28}$$

In the case of $1 - \frac{\alpha}{2} - \frac{\beta}{4} = 0$, which is exactly soluble case, we can take $\phi = \rho$ gauge even in first derivatives, so the latter two terms of above differential equation are zero. Thus, solution can be found easily,

$$\mathcal{F} = \frac{1}{2x_0^{+2}} - \frac{M}{\lambda x_0^{+2}}\frac{4w}{A} + \frac{4M^2}{\lambda^2 x_0^{+2}}\frac{w}{A^{3/2}}. \tag{29}$$

With this $\mathcal{F}$ evaluated at apparent horizon, we have time evolution of entropy as

$$\left.\frac{dS}{dx^+}\right|_{A.H.} = \frac{\kappa\alpha}{4x_0^+}\frac{\frac{2M}{\lambda}}{\frac{2M}{\lambda} - \kappa\left(\frac{\alpha}{4} - 1\right)}. \tag{30}$$

This takes positive value when the parameters satisfy

$$0 < \alpha < 4\left(\frac{2M}{\kappa\lambda} + 1\right), \text{ or } -4\left(\frac{4M}{\kappa\lambda} + 1\right) < \beta < 4. \tag{31}$$

In another case, $1 - \frac{\alpha}{2} - \frac{\beta}{4} > 0$, solution has more complex form. This result can not be reduced into (??), thus we must distinguish this result with one of exactly soluble case. Subtracting $\mathcal{F}|_{A.H.}$ into (??) and adding contribution of $\partial_+ S$, we have

$$\left.\frac{dS}{dx^+}\right|_{A.H.} = \frac{1}{x_0^+}\left\{\kappa - \frac{2M}{\lambda} + \sqrt{\kappa^2\left(1 - \frac{\alpha}{2} - \frac{\beta}{4}\right) + \left(\frac{2M}{\lambda}\right)^2}\right\} \tag{32}$$

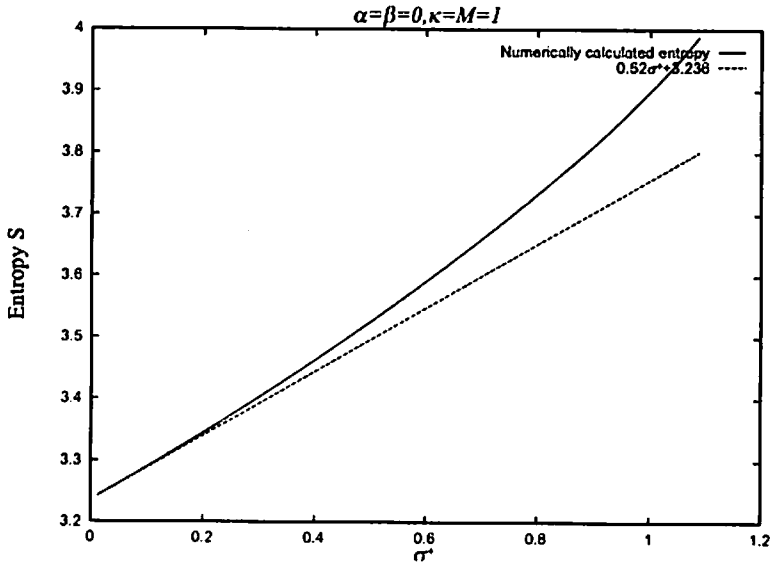


Figure 2: This figure shows the entropy of semi-classical CGHS model for parameters $\lambda = 1$, $\sigma_0^+ = 0\,(x_0^+ = 1)$, $\kappa = M = 1$. Solid curve is numerically calculated entropy, and dashed strait line is analytically calculated one.

$$
\begin{aligned}
&+\frac{A_0' + 4\kappa}{A_0' - 4\kappa\left(\frac{\alpha}{4} - 1\right)}\left\{ \frac{\kappa}{2x_0^+}\left\{3 - 2\left(\frac{\alpha}{4} - 1\right)\right\} - \frac{4M^2}{\kappa\lambda^2 x_0^+}\right. \\
&+\frac{32M^3}{\kappa\lambda^3 x_0^+}\frac{1}{A_0'} - \frac{2\kappa^2}{x_0^+}\left(1 - \frac{\alpha}{2} - \frac{\beta}{4}\right)\frac{1}{A_0'} \\
&\left.+\frac{4M^2}{\lambda^2 x_0^+}\frac{2\left(1 - \frac{\alpha}{2} - \frac{\beta}{4}\right) + \frac{\alpha}{4} - 1}{\sqrt{1 - \frac{\alpha}{2} - \frac{\beta}{4}}}\log\frac{A' + 4\kappa\sqrt{1 - \frac{\alpha}{2} - \frac{\beta}{4}}}{A' - 4\kappa\sqrt{1 - \frac{\alpha}{2} - \frac{\beta}{4}}}\right\}. \qquad (33)
\end{aligned}
$$

In the case of $\alpha = \beta = 0$ which is called semi-classical CGHS model, it can be shown that the gradient of the entropy is always positive at the vicinity of the infalling shell. In other case, the sign of dS/dx^+ surely depends on α and β. For example, $\alpha = 12$, $\beta = -20.1$ at $\lambda = 1$, $x_0^+ = 1$ leads negative value of dS/dx^+ near $\kappa = M$. Therefore there exist some models which have decreasing entropy with Wald's scheme in 2 dimension.

4.2 Numerical Approach

To know the behavior of entropy along the apparent horizon far from the shell, we must perform numerical calculation for some values of the parameters κ and M. Although we cannot cover all values of the parameters in numerical calculation, it is necessary to know information which cannot be obtained from the above analytic approach.

Numerical calculation was performed on $\sigma^\pm$ coordinate, and we fix the parameters, $\lambda = 1$, $\sigma_0^+ = 0$ $(x_0^+ = 1)$[?]. Equations for ϕ and ρ are second order partial differential equations. We have used the numerical schemes invented in [?]. The range of σ^+ is from

the infalling shell to the merging point of the apparent horizon and singularity. One of the typical numerical results is shown in Figure 2. The result is for $\kappa = M = 1$ in semi-classical CGHS model. Solid curve is numerically calculated entropy, and dashed strait line is analytically calculated one. In the latter one, the value of entropy on the infalling shell is evaluated by using the LDV solution, and gradient of entropy is calculated with analytic scheme presented in above subsection. This result shows that the entropy grows still even in the region far from the infalling shell. The range in which analytic approximation is good is not so wide in this case, it shows the higher order effects make the growing rate more rapidly. We also calculated the evolution of the entropy for other parameters. All of the results show the same feature.

5 Conclusion

In this paper, we have investigated the dynamical evolution of the entropy proposed by Wald along the apparent horizon in the context of the 2-dimensional dilaton gravitational theory. In particular, we have considered the quantum backreaction and then the evaporating spacetimes. In the case of the semi-classical CGHS model, we have shown analytically that the thermodynamic second law holds along the apparent horizon in the vicinity of the infalling shell. Furthermore, the numerical calculation is also presented, which suggest that the thermodynamic second law holds even in the region far from the infalling shell. Myers's result on the RST model and analysis on the semi-classical CGHS model are very impressive, however, they are yet special models. Hence, we have also studied the general models which seem physically natural in our judge. Analytic calculation shows that, even in the vicinity of the infalling shell, the thermodynamic law does not hold in certain models. There are two possible choices here. One is to use this fact for restrictions of models. Another is to modify the entropy formula. In the latter case, the statistical interpretation of the entropy would be necessary. However, this interesting subject is beyond the scope of this paper. Further implications of our analysis should be considered in the future.

Acknowledgments

The work is supported by Monbusho Grant-in-Aid for Scientific Research No. 07740216.

References

[1] J.M.Bardeen, B.Carter, and S.W. Hawking, *Comm. Math. Phys.* **31**(1973)161.

[2] S.W. Hawking, *Comm. Math. Phys.* **43**(1975)199.

[3] C.G. Callan, S.B. Giddings, J.A. Harvey, and A. Strominger, *Phys. Rev. D* **45**(1992) R1005.

[4] V.P.Frolov, *Phys. Rev. D* **46**(1992)5383.

[5] R.M. Wald, *Phys. Rev. D* **48**(1993) R3427.

[6] R.C. Myers, *Phys. Rev. D* **50**(1994) 6412.

[7] J.G. Russo, L. Susskind and L. Thorlacius, *Phys. Rev. D* **46**(1992) 3444.

[8] R.C. Myers, hep-th/9508063.

[9] S. Bose, L. Parker, and Y. Peleg, *Phys. Rev, D* **52**(1995) 3512.

[10] T.Jacobson, G.Kang, and R.C.Myers, *Phys. Rev. D* **49**(1994)6587.

[11] V. Iyer and R. M. Wald, *Phys. Rev.* **D50** (1994) 846.

[12] J.G. Russo, L. Susskind and L. Thorlacius, *Phys. Lett. B 292*(1992) 13.

[13] T. Piran and A. Strominger, *Phys. Rev. D* **48**(1993)4729.

[14] S.W. Hawking and J.M. Stewart, *Nucl. Phys. B* **400**(1993)393.

Field Theoretical Quantum Effects on the Kerr Geometry

Y. SATOH †

Institute of Physics, University of Tokyo,
Komaba, Tokyo 153 Japan

Abstract

We study some quantum aspects of the Einstein gravity with two commuting Killing vectors. The theory is formulated as a $SL(2,\mathbf{R})/U(1)$ nonlinear σ-model with a coordinate-dependent coupling coupled to gravity in two dimensions. The quantum analysis of the nonlinear σ-model part, which includes all the dynamical degrees of freedom, can be carried out consistently in a parallel way to ordinary nonlinear σ-models. We calculate the effective action, and obtain the forms of the beta functions to all orders up to numerical coefficients. As an application, we consider the quantum effects on the Kerr black hole, and find that the asymptotically flat region remains intact and stable, while, in a certain approximation, the inner geometry undergoes a considerable change however small the quantum effects may be.

1. INTRODUCTION

Quantum aspects of Einstein gravity have been investigated intensively for a long time. As is well known, there are, however, many difficulties both conceptually and technically. One of the most outstanding ones is its nonrenormalizability. Due to this, we have no consistent way to investigate its quantum theory. In spite of that, many attempts have been made, for example, by using semiclassical [1] or $1/N$ [2] approximation. Unfortunately, in these approaches, we encounter the instability of flat space-time. The models with higher derivative terms also have been investigated as low-energy effective models of more fundamental theory such as string theory [3]. In these cases, we perform weak-curvature expansion, and thus we cannot deal with the strong-curvature region where quantum effects may become important. In a much more simplified setting, the quantum mechanics of minisuperspace [4] or the Schwarzschild black hole [5] has been studied, but it seems still difficult to extract definite physical consequences.

†e-mail address: ysatoh@hep1.c.u-tokyo.ac.jp

The above argument indicates the necessity of understanding more fundamental theory of gravity. In this direction, intensive studies have been made. However, together with this, it seems to be still important to accumulate certain pieces of knowledge of quantum aspects of Einstein gravity even if in a simplified setting. In this talk, we shall investigate some quantum analysis of Einstein gravity from such a point of view.

The Einstein gravity with two commuting Killing vectors (for instance, the stationary and axisymmetric case) are known to be formulated as a $SL(2,\mathbf{R})/U(1)$ non-linear σ-model coupled to gravity in two dimensions. The quantum theory of non-linear σ-models is well studied, and we can expect to make use of such knowledge for studying quantum aspects of the Einstein gravity in our case.

Therefore, our aims here are (i) to study a quantum theory of the Einstein gravity with two commuting Killing vectors, (ii) to investigate its quantum effects on geometry, and (iii) get some insights into quantum aspects of Einstein gravity. For details, see [6].

2. Dimensionally Reduced Einstein Gravity

When there exist two commuting Killing vectors, all the fields in the theory become independent of these directions. We denote them by the coordinates x^0 and x^3, and, for definiteness, we assume that x^0 is time-like and x^3 is space-like as in the case of stationary and axisymmetric geometry. Consequently, the theory can be regarded as that in two dimensions represented by x^1 and x^2. There are several ways to perform this dimensional reduction, here we adopt the method using vielbein [7].

Let us begin with the following vierbein in a triangular gauge,

$$E^A_{\ M} \ = \ \begin{pmatrix} \Delta^{-1/2}\, e^a_{\ m} & \Delta^{1/2} A_m \\ 0 & \Delta^{1/2} \end{pmatrix}, \tag{1}$$

where $M(=0\text{ - }3)$ and $m(=1\text{ - }3)$ refer to the space-time indices , $A(=0\text{ - }3)$ and $a(=1\text{ - }3)$ to those of its tangent space and all the components are independent of x^0. Then we change the variables from A_m to B by a dual transformation,

$$\Delta^2 F_{mn} \ = \ -\epsilon_{mnl}\, \partial^l B\,, \tag{2}$$

where ϵ_{mnl} is the volume element defined by the dreibein $e^a_{\ m}$. Furthermore, we drop all the dependence of the fields on x^3 and set $e^a_{\ m}$ to be of the form,

$$e^a_{\ m} \ = \ \begin{pmatrix} \lambda \delta^a_\mu & 0 \\ 0 & \rho \end{pmatrix}, \tag{3}$$

where $\mu, \alpha = 1, 2$. We then get the following expression of the action,

$$\begin{aligned} \frac{1}{\hbar} S_{EH} &= \frac{1}{\hbar\kappa} \int d^4x \, E R^{(4)}(E) \\ &\to \frac{V}{l_p^2} \int d^2x \, \rho \left[e R^{(2)} - \frac{1}{2} \delta^{\mu\nu} g_{ij}(\phi) \partial_\mu \phi^i \partial_\nu \phi^j \right], \end{aligned} \tag{4}$$

where $\kappa = G/c^3$, V is the "volume" of (x^0, x^3) space-time and l_p is the Planck length. The curvature with respect to (x^1, x^2) space is given by $eR^{(2)} = -2\delta^{\mu\nu}\partial_\mu\partial_\nu \ln\lambda$, while ϕ^i and g_{ij} are by $\phi^i = (B, \Delta)$ and $g_{ij} = \Delta^{-2}\delta_{ij}$. They are identified with the coordinates and the metric of $SL(2,\mathbf{R})/U(1)$ manifold, respectively. Therefore, we see that the above action describes a $SL(2,\mathbf{R})/U(1)$ nonlinear σ-model coupled to two dimensional gravity. Note that, in the above expression, the model can be regarded as the one in flat space. Thus in the following, it is understood that the indices μ, ν are raised and lowered by the flat metric.

The independent equations of motion derived from (4) are

$$\partial_\mu \partial^\mu \rho = 0 \tag{5}$$

$$\Delta \partial^\mu (\rho \partial_\mu \mathcal{E}) = \rho \, \partial_\mu \mathcal{E} \partial^\mu \mathcal{E} \tag{6}$$

$$\partial_\zeta \rho \partial_\zeta \ln\lambda - \frac{1}{2}\partial_\zeta^2 \rho = \frac{1}{4} \rho \, \Delta^{-2} \partial_\zeta \mathcal{E} \partial_\zeta \bar{\mathcal{E}}, \tag{7}$$

where $\zeta = x^1 + ix^2$ and $\mathcal{E} = \Delta + iB$. $\mathcal{E}$ and the second equation, (6), are called the Ernst potential and the Ernst equation, respectively [8]. In our parametrization, the dual transformation (2) takes the form

$$\Delta^2 \partial_\zeta A = i\rho \partial_\zeta B, \tag{8}$$

and the line element becomes

$$ds^2 = \Delta^{-1} \left[\lambda^2 \left((dx^1)^2 + (dx^2)^2 \right) + \rho^2 (dx^3)^2 \right] - \Delta \left(dx^0 + A \, dx^3 \right)^2, \tag{9}$$

where we have set $A_{1,2} = 0$ and $A_3 \equiv A$ (this is possible because we are working in two dimensions).

The model has many interesting properties and has been investigated extensively mainly by relativists (see, e.g., [9, 7]). For example, it has an infinite dimensional symmetry called Geroch group [10, 11], and the similarity between this symmetry and those in dimensionally reduced supergravity has been recognized [12]. Recently, the relation between these symmetries and dual symmetries in string theory has been discussed [13].

Furthermore, the model is integrable [14, 15], and has a large class of solutions called the Tomimatsu-Sato class [16] (the Kerr solution is the simplest one in this class).

3. Quantum Theory

Now let us consider the quantum theory. From the action, we see that all the dynamical degrees of freedom are included in the nonlinear σ -model part. Thus we shall investigate the quantum theory of that part and its effects on geometry. In other words, we shall incorporate the quantum fluctuations depending only on x^1 and x^2 and study its physical effects. The quantization including the gravity part is the subject of two dimensional quantum gravity, but we do not yet have any way to deal with our model in this context because, for instance, the dynamical degrees of freedom exceeds 1. From a different point of view, a quantization of the entire system has been discussed [17].

Here a comment may be in order. Since the fluctuations to x^0 and x^3 directions are ignored, our analysis is not enough to know the full quantum properties of Einstein gravity, of course. However, we have at present no consistent way to investigate the full quantum theory of Einstein gravity because of its nonrenormalizability and various difficulties. Our attitude here is a modest one. Although only a part of the quantum fluctuations can be incorporated, in this simplified setting we can carry out a consistent quantum analysis of Einstein gravity and extract some quantum effects on geometry. We believe that our analysis gives some insights into quantum aspects of general relativity.

Please recall that the action of the nonlinear σ-model part was

$$\frac{1}{\hbar}S_{NL} = -\frac{1}{2}\int d^2x T_0^{-1}(x) g_{ij}(\phi)\partial_\mu\phi^i\partial^\mu\phi^j , \tag{10}$$

where $T_0 = e_0^2/\rho(x)$, $e_0^2 = l_p^2/V$, and $T_0(x)$ is regarded as a coordinate-dependent coupling like dilaton field. This is the difference between our model and ordinary nonlinear σ-models, where the couplings are constant.

First, we define the quantum theory by the above action and the $SL(2,\mathbf{R})/U(1)$ invariant measure maintaining the covariance of the target space. The model is renormalizable since it is regarded as the one in two dimensions. Next, we calculate the effective action perturbatively. In this calculation, we adopt the background field method [18] and the dimensional regularization. In spite of the existence of the anomalous coupling, $T_0(x)$, we find that we can deal with our model in a parallel way to ordinary cases by some change

of variables. However, since the actual calculation is rather technical and complicated, we display only the result here.

The calculation is quite simplified by making use of the facts that $SL(2,\mathrm{R})/U(1)$ has a negative constant curvature, i.e., $R^{SL(2,\mathrm{R})/U(1)} = -2$, and all the quantities with the indices of the target manifold are proportional to the metric, e.g., $R_{ij}^{SL(2,\mathrm{R})/U(1)} \propto g_{ij}$. We then get the effective action of the form

$$-\frac{1}{\hbar}\, S_{NL}^{eff}[\phi] \;=\; \frac{1}{2}\int d^2x\; T^{-1}(x;\mu) g_{ij}(\phi)\partial_\mu\phi^1\partial^\mu\phi^j + \frac{1}{2}\int d^2x\; U^{-1}(x;\mu)\;\partial_\mu \ln\rho\partial^\mu\ln\rho \;+\; (\mathrm{H.D.T})\,. \tag{11}$$

Here $T(x;\mu)$ is the renormalized coupling, and μ refers to the renormalization point. The term including $U = U(T(x;\mu))$ arises from the renormalization, which is absent in ordinary nonlinear σ-models, and (H.D.T) represents the higher derivative terms. Since we are interested in global geometry of space-time, their effects are expected to be small. Thus we shall henceforth focus on the quadratic derivative terms.

Furthermore, following the standard procedure, we get the forms of the beta functions for $T(x;\mu)$ and $U(x;\mu)$ to all orders :

$$\beta_T(T) \;\equiv\; \mu\frac{\partial}{\partial\mu}T \;=\; -\frac{1}{2\pi}\sum_{N=1} N a_N^{(0)} T^{N+1}\,, \tag{12}$$

$$\beta_U(T) \;\equiv\; \mu\frac{\partial}{\partial\mu}U \;=\; \frac{1}{2\pi}U^2\sum_{N=1}\left(N b_{N+1}^{(0)} + \partial_T \ln U\cdot b_N^{(0)}\right)T^N\,, \tag{13}$$

where $a_N^{(0)}$ and $b_N^{(0)}$ are some numbers determined by the explicit loop calculations. For example, they are $a_1^{(0)} = -1$, $a_2^{(0)} = 1/2\pi$ [18, 19], and $b_1^{(0)} = 1/2$.

4. Quantum Effects on the Kerr Geometry

In the previous section, we carried out the renormalization of our model. Then let us study its effects on geometry. The equations of motion derived from the effective action and the dual relation are given by

$$\partial_\mu\partial^\mu\rho \;=\; 0\,, \tag{14}$$

$$\Delta\partial_\mu\left(T^{-1}\,\partial^\mu\mathcal{E}\right) \;=\; T^{-1}\,\partial_\mu\mathcal{E}\partial^\mu\mathcal{E}\,, \tag{15}$$

$$\partial_\zeta\rho\partial_\zeta\ln\lambda - \frac{1}{2}\partial_\zeta^2\rho \;=\; \frac{1}{4}e^2(\mu_0)\left(T^{-1}\Delta^{-2}\partial_\zeta\mathcal{E}\partial_\zeta\bar{\mathcal{E}} + U^{-1}\partial_\zeta\ln\rho\partial_\zeta\ln\rho\right)\,, \tag{16}$$

$$\Delta^2\partial_\zeta A \;=\; i\rho\partial_\zeta B\,, \tag{17}$$

where μ_0 is the scale of the classical theory and $e^2(\mu_0) \equiv T(x;\mu_0)/\rho(x)$. Note that the quantum effects are represented by the renormalized coupling $T(x;\mu)$ and the additional term including $U(x;\mu)$. For later use, we adopt a specific coordinate system called Weyl's canonical coordinate system. Since $\rho(x)$ is a free field in two dimensions, we can take it as a coordinate. Then introducing another free field, z, conjugate to $\rho(x)$, we set

$$x^1 = l_p\rho, \quad x^2 = l_p z, \tag{18}$$

where, on dimensional grounds, we have multiplied ρ and z by l_p which is the only constant with dimension of length in our theory.

As an interesting example, in the following we shall consider the Kerr geometry. Introducing new coordinates, r and θ, (Boyer-Lindquest coordinates) defined by

$$l_p\rho = \sqrt{r^2 - 2mr + a^2}\,\sin\theta, \quad l_p z = (r-m)\cos\theta, \tag{19}$$

and denoting x^0 and x^3 by t and ω, respectively, the line element of the Kerr solution is written as

$$\begin{aligned} ds^2 = & -\left(1 - \frac{2mr}{\Sigma}\right)dt^2 + \Sigma\left(\frac{dr^2}{D} + d\theta^2\right) \\ & -\frac{4mar}{\Sigma}\sin^2\theta\, d\omega\, dt + \left(r^2 + a^2 + \frac{2ma^2r}{\Sigma}\sin^2\theta\right)\sin^2\theta\, d\omega^2, \end{aligned} \tag{20}$$

where $D = r^2 - 2mr + a^2$, $\Sigma = r^2 + a^2\cos^2\theta$, and m and a represent the mass and the angular momentum per unit mass of the Kerr black hole, respectively. Notice that zeros of D(we denote them by $r_\pm$) correspond to the horizons and the outer zero of $D - a^2\sin^2\theta$ to the outer boundary of the ergosphere.

Now let us discuss the quantum effects on the geometry. First it is easy to see the behavior of the renormalized coupling $T(x;\mu)$. It tends to blow up as we approach the horizons ($r \to r_\pm$) or the axis of the rotation of the black hole ($\sin\theta \to 0$), while it tends to vanish as we go to infinity ($r \to \infty$). Therefore we find that the asymptotically flat region has no quantum effects and remains stable.

Second, we find also that the effects of the term including $U(x;\mu)$ are absorbed into a factor of λ. Indeed, defining λ_T by $\lambda = f(\rho;\mu)\lambda_T$ and $f(\rho;\mu) = \exp\left[(e^2(\mu_0)/4)\int_\rho^\infty d\rho'\,\rho'^{-2}U^{-1}(T(\rho';\mu))\right]$, the equation (16) becomes

$$\partial_\zeta\rho\partial_\zeta \ln\lambda_T - \frac{1}{2}\partial_\zeta^2\rho = \frac{1}{4}e^2(\mu_0)T^{-1}\Delta^{-2}\partial_\zeta\mathcal{E}\partial_\zeta\bar{\mathcal{E}}, \tag{21}$$

in which U disappears.

Third, in order to get the geometry including the quantum effects, we have to solve Eqs. (15),(21) and (17). However this is a hard task at a generic loop order, and hence we concentrate on one loop order and consider only the deviation from the classical solution in the neighborhood $\rho(x) = \rho_0$. At one loop order, we have

$$e^2(\mu_0)T^{-1}(x;\mu) \;=\; \rho(x)\left\{1 - \frac{1}{2\pi}e^2(\mu_0)\rho^{-1}(x)\ln(\mu/\mu_0)\right\}. \tag{22}$$

Moreover we approximate this by $\alpha(\rho_0)\rho(x)$, where $\alpha(\rho_0)=1-(e^2(\mu_0)/2\pi)\rho_0^{-1}\ln(\mu/\mu_0) \sim$ const. . Note that at the classical scale, i.e., $\mu = \mu_0$, we have $\alpha(\rho_0) = 1$. Since $\partial_{x^1}\rho^{-1} = -\,l_p^{-1}\rho^{-2}$ and $\partial_{x^2}\rho^{-1} = 0$ hold, this approximation is valid where $\rho(x) = (\sqrt{r^2 - 2mr + a^2}\,\sin\theta)/l_p >> 1$, namely, in the region far from the horizons or the rotation axis measured by the Planck unit. In this approximation, the quantum effects on the geometry are represented by a factor in front of $\Sigma\,(dr^2/D + d\theta^2)$, and we get the following geometry,

$$\begin{aligned} ds^2 \;=\; & -\left(1 - \frac{2mr}{\Sigma}\right)dt^2 + f^2(\rho_0)\left(\frac{F_1}{F_2}\right)^{\alpha(\rho_0)-1}\Sigma\left(\frac{dr^2}{D} + d\theta^2\right) \\ & - \frac{4mar}{\Sigma}\sin^2\theta\; d\omega\; dt + \left(r^2 + a^2 + \frac{2ma^2r}{\Sigma}\sin^2\theta\right)\sin^2\theta\; d\omega^2, \end{aligned} \tag{23}$$

where $F_1 = D - a^2\sin^2\theta$ and $F_2 = D + (m^2 - a^2)\sin^2\theta$. Therefore we find that additional zeros and singularities appear in the metric where F_1 or F_2 vanishes unless $\alpha(\rho_0) = 1$. After some calculations, we find also that they are zeros or singularities of curvature invariants. Furthermore, we see that the most outside one among them develops at the outer boundary of the ergosphere in which several unusual phenomena can take place. We remark that our approximation still valid there as long as m and a are large enough. We do not know whether or not we take the existence of such zeros and singularities seriously. However, our result indicates that the inner geometry of the Kerr black hole undergoes a considerable change due to quantum effects no matter how small the quantum effects may be.

5. Discussion

In this talk, we dealt with the Einstein gravity with two commuting Killing vectors and formulated it as a $SL(2,\mathbf{R})/U(1)$ nonlinear σ-model coupled to gravity in two dimensions. We then quantized the nonlinear σ-model part, which included all the dynamical

degrees of freedom and renormalizable, and calculated the effective action and the beta functions. The quantum effects of our analysis on geometry, i.e., the effects of the quantum fluctuations depending only on two coordinates, were investigates by taking the Kerr geometry as an interesting example. As a result, we found that the asymptotically flat region had no quantum effects and remained stable, while the inner geometry underwent a considerable change however small the quantum effects might be.

Admittedly, our analysis is not enough to know quantum aspects of full Einstein gravity. We can not consider the effects of the truncated degrees of freedom and statistical aspects of Einstein gravity. In our simplified setting, however, we can carry out some quantum analysis consistently. Moreover, we can confirm that there certainly exist the quantum fluctuations by which the quantum effects become more and more important as we approach the inner region of the Kerr geometry, as expected. We believe that our results give useful insights into quantum aspects of Einstein gravity.

Finally, we make two comments. First, we can perform a similar analysis also in the Einstein-Maxwell system. In this case, the extension is straightforward and the target manifold of the nonlinear σ-model part becomes $SU(2;1)/(SU(2)\times U(1))$ or $SU(2;1)/(SU(1;1)\times U(1))$ according to the signatures of the Killing vectors [20, 9]. Second, Einstein gravity is already formulated as a $SL(2,\mathbf{R})/U(1)$ nonlinear σ-model coupled to gravity in three dimension just by dropping the dependence on one Killing direction. Thus it may be possible to investigate quantum aspects of the Einstein gravity with only one Killing vector by making use of the knowledge of three dimensional nonlinear σ-models.

References

[1] G.T. Horowitz and R.M. Wald, Phys. Rev. D **17**, 414 (1978).

[2] E.T. Tomboulis, Phys. Lett. **B70**, 361 (1977).

[3] D.G. Boulware and S. Desser, Phys. Rev. Lett. **55**, 2656 (1985); J.T. Wheeler, Nucl. Phys. **B268**, 737 (1986); R.C. Meyers, Nucl. Phys. **B289**, 701 (1987).

[4] B. Hartle and S.W. Hawking, Phys. Rev. D **28**, 2960 (1983).

[5] K.V. Kuchař, Phys. Rev. D **50**, 3961 (1994).

[6] Y. Satoh, preprint UT-Komaba 95-11, hep-th/9512019.

[7] H. Nicolai, in: *Recent Aspects of Quantum Fields*, edited by H. Mitter and G. Gausterer (Springer, Berlin, 1991).

[8] F. Ernst, Phys. Rev. **167**, 1175 (1968).

[9] *Solutions of Einstein's Equations: Techniques and Results*, edited by C. Hoenselaers and W. Dietz (Springer, Berlin, 1984).

[10] R. Geroch, J. Math. Phys. **12**, 918 (1971).

[11] P. Breitenlohner and D. Maison, Ann. Inst. Henri Poincarè **46** , 215 (1987).

[12] B. Julia, in: *Proceedings of the 5th Johns Hopkins Workshop on Current Problems in Particle Theory*, (Johns Hopkins University, Baltimore, MD, 1981).

[13] A. Sen, Nucl. Phys. **B447**, 62 (1995); A. Kumar and K. Ray, Phys. Lett. **B358**, 223 (1995); S. Mizoguchi, preprint hep-th/9506192 (1995).

[14] V.A. Bellinskii and V.E. Zakharov, Sov. Phys. JETP **48**, 985 (1978).

[15] D. Maison, Phys. Rev. Lett. **41**, 521 (1978).

[16] A. Tomimatsu and H. Sato, Prog. Theor. Phys. **50**, 95 (1975).

[17] D. Korotkin and H. Nicolai, Phys. Rev. Lett. **74**, 1272 (1995).

[18] L. Alvarez-Gaumé, D.Z. Freedman, and S. Mukhi, Ann. Phys. **134**, 85 (1981).

[19] D. Friedan, Phys. Rev. Lett. **45**, 1057 (1980); Ann. Phys. **163**, 318 (1985).

[20] W. Kinnersley, J. Math. Phys. **14**, 651 (1973).

Thermodynamic Properties of Two Dimensional R^2-Gravity *

Shoichi ICHINOSE †
Department of Physics, Universuty of Shizuoka,
Yada 52-1, Shizuoka 422, Japan

February, 1996

Abstract

Two dimensional quantum R^2-gravity is studied in the semiclassical method. The thermodynamic properties, such as the equation of state, the temperature and the entropy, are examined. The classical solutions (vacua) of R^2-Liouville equation are obtained by making use of the well-known solution of the ordinary Liouville equation. They are constant curvature solutions. The positive constant curvature solution and the negative one are, after proper infrared regularization, 'dual' each other. Each solution has two branches(±). We characterize all phases appearing in all solutions and branches. The topology constraint and the area constraint are properly taken into account. A total derivative term and an infrared regularization play important roles. The topology of a sphere is mainly considered.

1 Introduction

In the recent analysis of two dimensional(2d) quantum gravity(QG), the conformal approach or the matrix model approach have been intensively done. Those approaches are nonperturbative ones and it is expected that some non-perturbative features are important to understand the theory. At the same time, however, it is known that an orthodox perturbative approach, the semiclassical approach, is also useful in 2d QG[1, 2]. We present a close examination of 2d R^2-gravity using the latter approach. A key point in the analysis of 2d QG is how to treat the area constraint and the topology constraint. The regularization of infrared divergence (and ultraviolet divergence in the quantum evaluation) is also important. The semiclassical treatments of 2d QG so far are insufficient in these points. We present a new analysis.

Despite the long period of research, the physical picture of the Liouville theory, in relation to 2d QG, still remains obscure. It shows the delicacy or the subtlety of the theory and requires some other proper formalism and regularization. So far as the popular formalism based on the conformal field theory is taken, the barrier $c_m = 1$ does not seem to be overcome. The computer simulations, however, seem to no special difficulty for

*Talk at Workshop on General Relativity and Gravitation, Jan.22-25,1996, Nagoya Univ.,Japan.
†E-mail address:ichinose@u-shizuoka-ken.ac.jp

the prohibitted region of c_m[3, 4]. This conflict seems gradually serious. Although the problem is examined from different approaches, it is fair to say the true situation is not *known* at present. Recently it has been shown that the semiclassical results nicely explain the simulation data[5, 6].

It is well known, in the lattice QG, that the higher-derivative term, βR^2, regularizes the lattice system very well (for a proper sigh of β). It has the effect of smoothing or roughening the surface. We can expect a rich phase structure. The importance of the term is also stressed in the continuum context[7]. From the simple power-counting argument, we see the ultra-violet behaviour becomes well regularized. We can take two standpoints about the R^2-term: 1) We are considering the 2d R^2-QG as one gravitational model; 2) We regard R^2-term as a regularization to define the $\beta = 0$ theory and expect its effect disappears when some averaging (renormalization) procedure is taken. Although 1) is mainly taken in the present paper, both standpoints are important at this time of development.

In the semiclassical analysis of 2d R^2-QG in [5, 6], one of the classical solutions (positive curvature solution) is analysed. In the present paper we present the full structure of the classical vacua. We will see an interesting fact that the positive and negative solutions are, after doing infrared regularization differently for each solution, symmetric by a reflection in the coupling β-space. Each solution has two ($\pm$) branches. $-$branch solutions are continuous over all real β, whereas $+$branch ones live in the half real region of β. The explanation is self-contained.

It is well known, in gravitational systems, that the macroscopic quantities, such as the entropy, the volume and the temperature, obey the laws of thermodynamics. Those quantites are thermodynamic state variables of an equilibrium state. We will find those properties in the present 2d model of QG. Especially we can characterize all phases, appearing in the theory, by the β-dependence of the temperature.

2 Semiclassical Approach

We analyse the simulation data by the semiclassical approach. The R^2- gravity interacting with c_m-components scalar matter fields is described by

$$S = \int d^2x\sqrt{g}\left(\frac{1}{G}R - \beta R^2 - \mu - \frac{1}{2}\sum_{i=1}^{c_m}\partial_a\Phi_i \cdot g^{ab}\cdot\partial_b\Phi_i\right) \quad , \quad (a,b = 1,2) \quad , \qquad (1)$$

where G is the gravitaional coupling constant, μ is the cosmological constant , β is the coupling strength for R^2-term and Φ is the c_m- components scalar matter fields. The signature is Euclidean. The partition function , under the fixed area condition $A = \int d^2x\sqrt{g}$ and with the conformal-flat gauge $g_{ab} = e^{\varphi}\,\delta_{ab}$, is written as [8],

$$\bar{Z}[A] = \int \frac{\mathcal{D}g\mathcal{D}\Phi}{V_{GC}}\{exp\tfrac{1}{\hbar}S\}\;\delta(\int d^2x\sqrt{g} - A) = exp\tfrac{1}{\hbar}\left(\tfrac{8\pi(1-h)}{G} - \mu A\right)\times Z[A] \quad ,$$

$$Z[A] \equiv \int \mathcal{D}\varphi\; e^{+\frac{1}{\hbar}S_0[\varphi]}\;\delta(\int d^2x\; e^{\varphi} - A) \quad , \qquad (2)$$

$$S_0[\varphi] = \int d^2x\;\left(\tfrac{1}{2\gamma}\varphi\partial^2\varphi - \beta\, e^{-\varphi}(\partial^2\varphi)^2 + \tfrac{\xi}{2\gamma}\partial_a(\varphi\partial_a\varphi)\right) \quad , \quad \tfrac{1}{\gamma} = \tfrac{1}{48\pi}(26 - c_m) \quad , \qquad (3)$$

where h is the number of handles [1]. V_{GC} is the gauge volume due to the general coordinate invariance. ξ is a free parameter. The total derivative term generally appears when

[1] The sign for the action is different from the usual convention as seen in (2).

integrating out the anomaly equation $\delta S_{ind}[\varphi]/\delta\varphi = \frac{1}{\gamma}\partial^2\varphi$. This term turns out to be very important. [2] We consider the manifold of a fixed topology of the sphere ,$h = 0$ and the case $\gamma > 0$ ($c_m < 26$). $\hbar$ is Planck constant. [3]

$Z[A]$ is rewritten as[5], after the Laplace transformation and the inverse Laplace one,

$$\begin{aligned} Z[A] &= \int \frac{d\lambda}{\hbar}\int \mathcal{D}\varphi\, exp\, \frac{1}{\hbar}[\, S_0[\varphi] - \lambda(\int d^2x e^{\varphi} - A)] \\ &= \int \frac{d\lambda}{\hbar} e^{\frac{1}{\hbar}\lambda A}\int \mathcal{D}\varphi\, exp\, \{\frac{1}{\hbar}S_\lambda[\varphi]\} \quad , \\ S_\lambda[\varphi] &\equiv S_0[\varphi] - \lambda\int d^2x\, e^{\varphi} \\ &= \int d^2x\, (\frac{1}{2\gamma}\varphi\partial^2\varphi - \beta\, e^{-\varphi}(\partial^2\varphi)^2 + \frac{\xi}{2\gamma}\partial_a(\varphi\partial_a\varphi) - \lambda\, e^{\varphi}) \quad , \end{aligned} \tag{4}$$

where the λ-integral should be carried out along an appropriate contour parallel to the imaginary axis in the complex λ-plane. Note that the δ-function constraint (area condition) in (2) is substituted by the λ-integral. The leading order configuration is given by the stationary minimum.

$$\begin{aligned} \left.\frac{\delta S_\lambda[\varphi]}{\delta\varphi}\right|_{\varphi_c} &= \frac{1}{\gamma}\partial^2\varphi + \beta\{e^{-\varphi}(\partial^2\varphi)^2 - 2\partial^2(e^{-\varphi}\partial^2\varphi)\} - \lambda e^{\varphi}\Big|_{\varphi_c} = 0 \quad , \\ \left.\frac{d}{d\lambda}(\lambda A + S_\lambda[\varphi_c])\right|_{\lambda_c} &= 0 \quad , \\ Z[A] &\approx \frac{1}{\hbar} exp\, \frac{1}{\hbar}\{\lambda_c A + S_{\lambda_c}[\varphi_c]\} \equiv \frac{1}{\hbar} exp\, \frac{1}{\hbar}\Gamma_c^{eff} \quad . \end{aligned} \tag{5}$$

Generally this approximation is valid for a large system. In the present case, the system size is proportional to $\frac{4\pi}{\gamma} = \frac{26-c_m}{12}$ because all quantitative variables are proportinal to the factor. We expect the approximation is valid unless it is too near the region: $c_m \sim 26$. The valid region is the same as stated in Sect.1.

The solution φ_c and λ_c ,which describes the positive-constant curvature solution and is continuous at $\beta = 0$, are given by[5]

$$\begin{aligned} \varphi_c(r) &= -ln\, \{\frac{\alpha_c}{8}(1 + \frac{r^2}{A})^2\} \quad , \quad r^2 = (x^1)^2 + (x^2)^2 \quad , \\ \alpha_c &= \frac{4\pi}{w}\{w + 1 - \sqrt{w^2 + 1 - 2\xi w}\,\} \quad , \quad w = 16\pi\beta'\gamma \quad , \quad \beta' \equiv \frac{\beta}{A} \quad , \\ \gamma\lambda_c A &= \frac{w}{16\pi}(\alpha_c)^2 - \alpha_c \quad , \end{aligned} \tag{6}$$

where ξ must satisfy $-1 \le \xi \le +1$ for the realness of α_c. (x^1, x^2) are the flat coordinates of the plane on which the sphere , which the above classical configuration represents, is stereographically projected. The partition function at the classical level is given by

$$\begin{aligned} \Gamma_c^{eff} &= ln\, Z[A]|_{\hbar^0} = \lambda_c A + (1+\xi)\tfrac{4\pi}{\gamma}\, ln\tfrac{\alpha_c}{8} - \tfrac{\alpha_c}{\gamma}w + C(A) \quad , \\ C(A) &= \tfrac{8\pi(2+\xi)}{\gamma} + \tfrac{8\pi\xi}{\gamma}\{\, ln(L^2/A) - 1\,\} + O(A/L^2) \quad , \\ &\tfrac{L^2}{A} \gg 1 \quad , \end{aligned} \tag{7}$$

[2]The uniqueness of this term, among all possible total derivatives, is shown in [5].

[3]In this section only,we explicitly write $\hbar$ (Planck constant) in order to show the perturbation structure clearly.

where L is the *infrared cut-off* ($r^2 \leq L^2$) introduced for the divergent volume integral of the total derivative term. Note that $C(\Lambda)$ does not depend on β (or w) but on c_m (or γ) and Λ. In addition to the above branch (we call it $-$ branch), there is $+$ branch (for the positive-constant curvature solution) where the solution lives only in the positive real region of w[5].

Similarly we can obtain the negative-constant curvature solution[9].

3 Physical Quantities of R^2-Gravity

We list here the obtained result of important physical quantites for the positive curvature case. They are expreesed by three physical parameters β ,γ ,Λ and one free parameter ξ.

$$\begin{aligned}
\text{Curvature} \quad & \alpha_c^{\pm} = \tfrac{4\pi}{w}\{w+1 \pm \sqrt{w^2+1-2\xi w}\} \quad , \\
\text{Classical Action} \quad & S_\lambda[\varphi_c] = (1+\xi)\tfrac{4\pi}{\gamma}\, ln\, \tfrac{\alpha_c}{8} - \tfrac{w}{\gamma}\alpha_c + C(\Lambda) \quad , \\
\text{String Tension} \quad & \lambda_c \Lambda = \tfrac{1}{16\pi\gamma}({\alpha_c}^2 w - 16\pi\alpha_c) \quad , \\
\text{An Expect. Value} \quad & -\Lambda < \int d^2x\sqrt{g}R^2 >|_c = \tfrac{\partial \Gamma^{eff}[\varphi_c]}{\partial\beta'} \\
& = -16\pi\alpha_c + \alpha_c^2 + \Xi \times \tfrac{1}{\alpha_c}\tfrac{d\alpha_c}{d\beta'} \quad , \\
\text{Free Energy} \quad & -\Gamma^{eff}|_c = -S_\lambda[\varphi_c] - \lambda_c\Lambda \quad ,
\end{aligned} \tag{8}$$

where $w \equiv 16\pi\beta'\gamma$ and $C(\Lambda)$ is given in (7) [4] . Note that $\Xi = 0$.

The curvature$\times\Lambda$ ($= R(\varphi_c)\times\Lambda = \alpha(w)$), the classical value of $\Lambda < \int d^2x\sqrt{g}R^2 >$ ($= -\frac{\partial\Gamma^{eff}[\varphi_c]}{\partial\beta'}$) and the string tension$\times\gamma\Lambda = \gamma\lambda_c\Lambda$ are plotted, for the $-$branch solution, in Fig.1,2 and 3 respectively. In the figures we take $\xi = 0.99$ whose meaning is explained in [9, 10], and the curves for the negative curvature solution are also plotted. The asymptotic behaviours of these quantities will be listed in Sec.4.

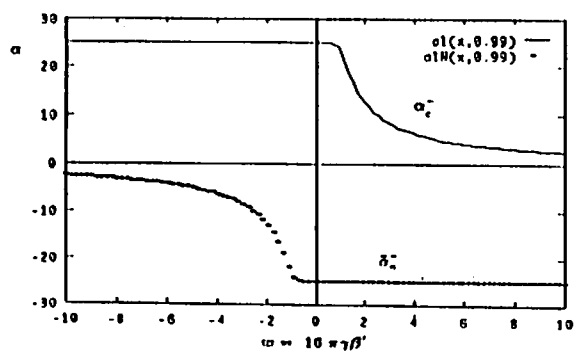


Fig.1 $A\times$ Curvature $= \alpha(w)$, Positive and Negative Curv. Sols., $-$Branch

[4]Because of the term $-\lambda(\int d^2x\, e^{\varphi} - A)$ in (4), we see λ can be interpreted as the string (surface) tension.

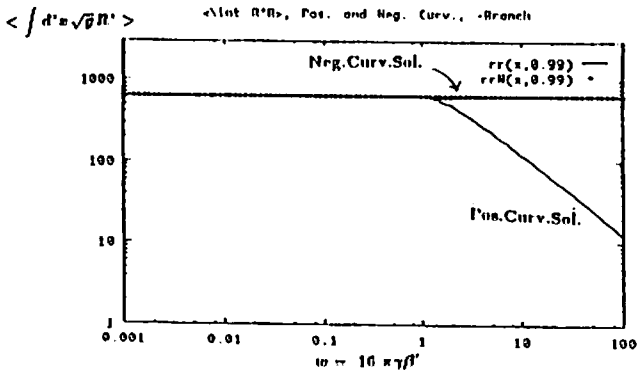


Fig.2a Log-Log Plot of $A < \int d^2x\sqrt{g}R^2 >|_c$, $w > 0$, Positive and Negative Curv. Sols., $-$Branch

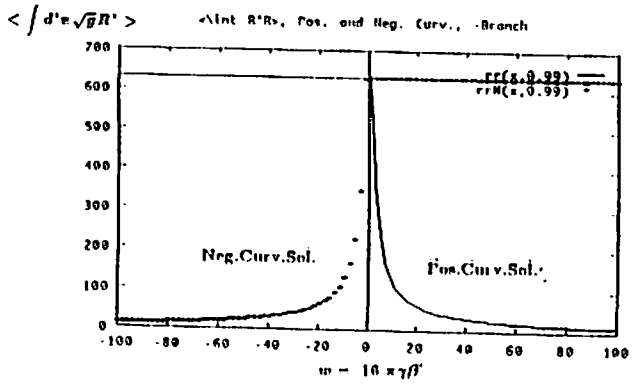


Fig.2b Linear Plot of $A < \int d^2x\sqrt{g}R^2 >|_c$, Positive and Negative Curv. Sols., $-$Branch

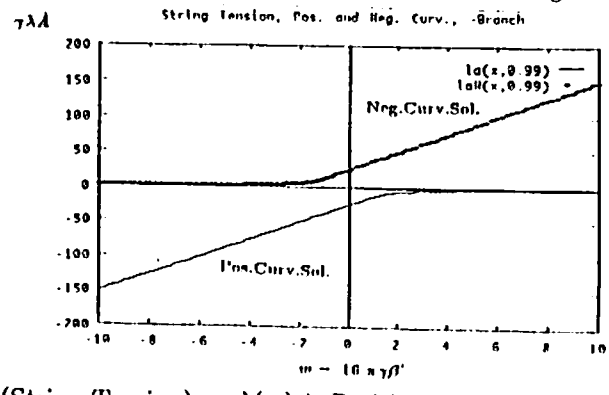


Fig.3 $\gamma A \times$(String Tension)= $\gamma\lambda(w)A$, Positive and Negative Curv. Sols., $-$Branch

We show the behaviours of the negative curvature solution in the dotted lines. The figures show the 'reflection symmetry' between the positive and negative-constant curvature solutions clearly. The asymptotic behaviours of the above physical quantities will be listed in Table 1 of Sec.4.

It is very interesting that we can define finite quantities in the open manifold (negative curvature case) by introducing a proper regularization and by a rescaling procedure [9]. It reminds us of the renormalization in the quantum field theory. The present case is, however, a procedure to absorb the infrared divergence due to the coordinate singularity

of the classical open manifold, not to absorb the ultraviolet divergence in the quantum theory. Note that the constant-curvature sign remains negative after the rescaling and the Euler number $\int d^2x\sqrt{g}R$ is *negatively divergent*. These facts make us envisage Fig.4 as the rescaled manifold. It describes a sphere punctured everywhere on the surface. Each puncture absorbs the infrared divergence.

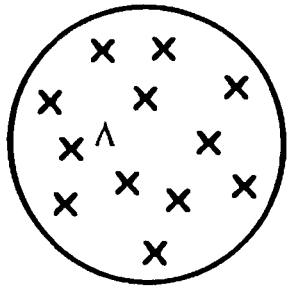


Fig.4 Punctured sphere absorbing infrared divergence

4 Phases and Asymptotic Behaviours

The asymptotic behaviours of the physical quantites, for the positive curvature case, are listed in Table 1.

		$w \ll -1$	$-1 \ll w < 0$	$0 < w \ll 1$	$1 \ll w$
+	Phase	/	/	(E)	(D)
	α_c^+	< 0 , not allowed	< 0 , not allowed	$\frac{4\pi}{w}\{2 + w(1-\xi) + O(w^2)\}$	$4\pi\{2 + \frac{1-\xi}{w} + O(w^{-2})\}$
	$-\frac{\partial\Gamma_+^{eff}}{\partial\beta'}$	/	/	$-\frac{64\pi^2}{w^2}\{1 - (1+\xi)w + O(w^2)\}$	$64\pi^2\{1 + \frac{0}{w} + O(w^{-2})\}$
	$\gamma\lambda_c^+ A$	/	/	$\frac{4\pi}{w}\{-1 + 0\cdot w + O(w^2)\}$	$4\pi w\{1 + O(w^{-1})\}$
	$-\gamma\Gamma_+^{eff}$	/	/	$\frac{4\pi}{w}\{1 + 2w + (1+\xi)w\ ln\ w + O(w^2)\} - \gamma C(A)$	$4\pi w\{1 + O(w^{-1})\} - \gamma C(A)$
−	Phase	(C)	(B)		(A)
	α_c^-	$8\pi + O(\lvert w\rvert^{-1})\}$	$4\pi(1+\xi)\{1 - \frac{1-\xi}{2}w\} + O(w^2)$		$\frac{4\pi(1+\xi)}{w} + O(w^{-2})$
	$-\frac{\partial\Gamma_-^{eff}}{\partial\beta'}$	$64\pi^2 + \frac{0}{\lvert w\rvert} + O(w^{-2})$	$16\pi^2(1+\xi)\{3 - \xi - (1-\xi)^2 w\} + O(w^2)$		$\frac{64\pi^2(1+\xi)}{w} + O(w^{-2})$
	$\gamma\lambda_c^- A$	$-4\pi\lvert w\rvert\{1 + O(\frac{1}{\lvert w\rvert})\}$	$4\pi(1+\xi)\{-1 + \frac{3-\xi}{4}w\} + O(w^2)$		$-\frac{\pi}{w}(1+\xi)(3-\xi) + O(w^{-2})$
	$-\gamma\Gamma_-^{eff}$	$-4\pi\lvert w\rvert\{1 + O(\frac{1}{\lvert w\rvert})\} - \gamma C(A)$	$4\pi(1+\xi)\{1 - ln\ \frac{1+\xi}{2} + \frac{3-\xi}{4}w\} + O(w^2) - \gamma C(A)$		$4\pi(1+\xi)\ ln\ w + \text{const} - \gamma C(A)$

Table 1 Asymp. behaviour of physical quantities.
$R > 0, w \equiv 16\pi\beta'\gamma, \gamma = \frac{48\pi}{26 - c_m} > 0$ ($c_m < 26$). $C(A)$ is given by (7).

Due to the 'reflection symmetry', each phase of the negative-curvature solution is characterzed in the similar way as in the positive-curvature case. We list the phase characterization in Table 2. ('Primes' in Table 2 mean modification due to the sign difference.)

	$w \ll -1$	$-1 \ll w < 0$	$0 < w \ll 1$	$1 \ll w$
+	(D')	(E')	/	/
-	(A')	(B')		(C')

Table 2 Phases of Negative Curvature Solution. $w \equiv 16\pi\tilde{\beta}'\gamma$.

All phases are explained in [5] using the above asymptotic behaviour. In the present paper we will characterize each phase by the equation of state in Sec.5.

5 Phases, Thermodynamic Properties and Equation of State

In this section we examine the thermodynamic properties of the system using the obtained analytic expression. We consider the positive curvature solution. The partition

function is given by

$$Z[A] = \int d\lambda \; exp\{ \hat{\Gamma}[\lambda] + \lambda A \} \approx exp\{\hat{\Gamma}[\lambda_c] + \lambda_c A\} \quad ,$$

$$\frac{d}{d\lambda}(\hat{\Gamma}[\lambda] + \lambda A)|_{\lambda_c} = \frac{d\hat{\Gamma}[\lambda_c]}{d\lambda_c} + A = 0 \quad . \tag{9}$$

Under the variation of the total area: $A \to A + \Delta A$, ln $Z[A]$ changes by $\Delta(\ln Z[A]) = \lambda_c \cdot \Delta A + \Delta A \cdot \frac{d\lambda_c}{dA} \cdot (\frac{d\hat{\Gamma}}{d\lambda} + A)|_{\lambda_c} = \lambda_c \cdot \Delta A$. Because the free energy F is given by $F = -ln\ Z[A]$, the pressure P is obtained as

$$P = -\frac{\partial}{\partial A} F = \frac{\partial}{\partial A} \ln\ Z[A] = \lambda_c \quad . \tag{10}$$

The pressure is the same as the string tension.

We define the temerature $T(w)$, imitating the Boyle-Charles' law, as follows.

$$P \cdot A \equiv \frac{4\pi}{\gamma} T(w) \quad , \quad T(w) = \frac{\gamma \lambda_c A}{4\pi} = \frac{1}{64\pi^2}(\alpha_c{}^2 w - 16\pi\alpha_c) \quad , \tag{11}$$

$$\alpha_c(w) = \begin{cases} \frac{4\pi}{w}\{w + 1 + |w-1|\} & \text{for + branch solution} \\ \frac{4\pi}{w}\{w + 1 - |w-1|\} & \text{for } -\text{branch solution} \end{cases}$$

$N \equiv 4\pi/\gamma = (26 - c_m)/12$ corresponds to the 'mol number'. The temperature is the (dimensionless) string tension per a unit mol. The final analytic form of the temperature is given by

$$\text{+branch solution} \qquad T(w) = \begin{cases} -\frac{1}{w} & \text{for } 0 < w \leq 1 \\ w - 2 & \text{for } 1 \leq w \end{cases}$$

$$-\text{branch solution} \qquad T(w) = \begin{cases} w - 2 & \text{for } w \leq 1 \\ -\frac{1}{w} & \text{for } 1 \leq w \end{cases} \tag{12}$$

The behaviour of $T(w)$ is plotted in Fig.5.

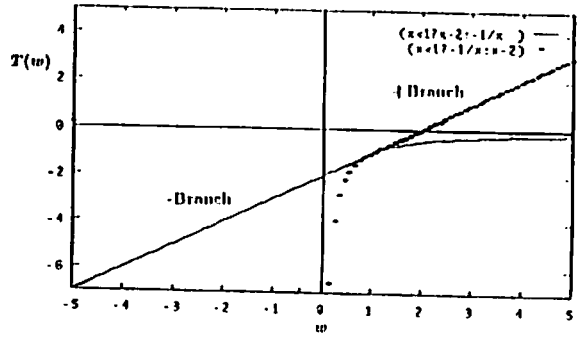


Fig.5 Temperature $T = T(w)$, Pos. Curv. Sol.

The asymptotic form of temperature in each phase is given by,

$$\begin{aligned}
&\text{Phase (A)} \quad w \gg 1 \quad , \quad P \cdot A = -\tfrac{4\pi}{\gamma}\tfrac{1}{w} \quad , \quad T_{(A)} = -\frac{1}{w} \quad , \\
&\text{Phase (B)} \quad |w| \ll 1 \quad , \quad P \cdot A = -\tfrac{8\pi}{\gamma}(1 + O(w)) \quad , \quad T_{(B)} \approx -2 \quad , \\
&\text{Phase (C)} \quad w \ll -1 \quad , \quad P \cdot A = \tfrac{4\pi}{\gamma} w(1 + O(w^{-1})) \quad , \quad T_{(C)} \approx w \quad , \\
&\text{Phase (D)} \quad w \gg 1 \quad , \quad P \cdot A = \tfrac{4\pi}{\gamma} w(1 + O(w^{-1})) \quad , \quad T_{(D)} \approx w \quad , \\
&\text{Phase (E)} \quad 0 < w \ll 1 \quad , \quad P \cdot A = -\tfrac{4\pi}{\gamma}\tfrac{1}{w} \quad , \quad T_{(E)} = -\frac{1}{w} \quad .
\end{aligned} \tag{13}$$

The negativeness of temperature, in Phase (A),(B) and (C), says the 'gass particles' attract each other. The small absolute value of $T_{(A)}$ says the 'gass particles' move almost freely. We can do the same analysis for the negative curvature solution. The corresponding temperture is obtained by 'reflecting' the graph of Fig.5. It is interesting that the 'gass particles' *attracting* each other on an *open* manifold can be regarded as the *repulsive* particles on a regularized *closed* manifold.

The entropy is similarly obtained. Using the relation: $\Delta w|_{A:fixed} = w \cdot \frac{\Delta\beta}{\beta}$, $\Delta T|_{A:fixed} = \frac{\partial T}{\partial w} \cdot w\frac{\Delta\beta}{\beta}$, it is given as

$$S_{ent} = -(\frac{\partial F}{\partial T})_A = +\frac{1}{w}\frac{\partial w}{\partial T} \cdot \beta\frac{\partial}{\partial\beta} \ln\ Z[A] = -\frac{1}{16\pi\gamma}\frac{\partial w}{\partial T} \cdot A < \int d^2x\sqrt{g}R^2 > \quad . \tag{14}$$

We see the entropy is related to the expectation value: $A < \int d^2x\sqrt{g}R^2 >$ considered in Sec.3, as above. Using the follwoing results from (8) ($\xi = 1$ is taken),

$$\text{+branch solution} \qquad A < \int d^2x\sqrt{g}R^2 >= \begin{cases} (8\pi)^2(\frac{2}{w} - \frac{1}{w^2}) & \text{for } 0 < w \leq 1 \\ +(8\pi)^2 & \text{for } 1 \leq w \end{cases} ,$$

$$\text{−branch solution} \qquad A < \int d^2x\sqrt{g}R^2 >= \begin{cases} +(8\pi)^2 & \text{for } w \leq 1 \\ (8\pi)^2(\frac{2}{w} - \frac{1}{w^2}) & \text{for } 1 \leq w \end{cases} , \tag{15}$$

we obtain the expression for the entropy.

$$\text{+branch solution} \qquad S_{ent} = \begin{cases} -\frac{4\pi}{\gamma}(2w - 1) & \text{for } 0 < w \leq 1 \\ -\frac{4\pi}{\gamma} & \text{for } 1 \leq w \end{cases} ,$$

$$\text{−branch solution} \qquad S_{ent} = \begin{cases} -\frac{4\pi}{\gamma} & \text{for } w \leq 1 \\ -\frac{4\pi}{\gamma}(2w - 1) & \text{for } 1 \leq w \end{cases} . \tag{16}$$

The graph of S_{ent} is plotted in Fig.6.

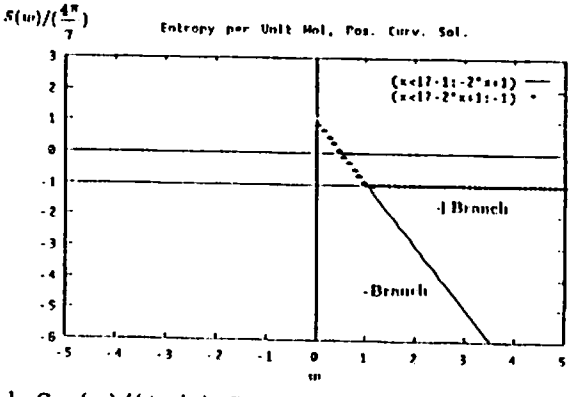


Fig.6 Entropy per unit mol, $S_{ent}(w)/(4\pi/\gamma)$, Pos.Curv.Sol.

The largeness of the absolute value of S_{ent} in Phase (A) shows the much amount of freedom of the system, whereas the fixed value in Phase (B) and (C) shows the possible configurations are restricted.

In Fig.7, all phases above are pictorially depicted.

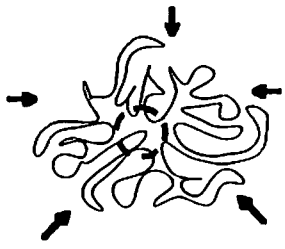


Fig. 14E

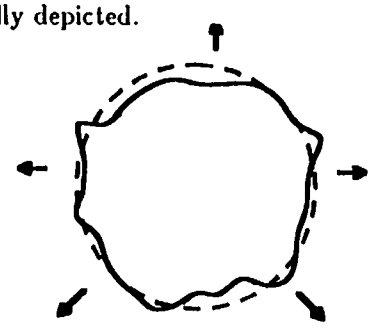


Fig. 14D

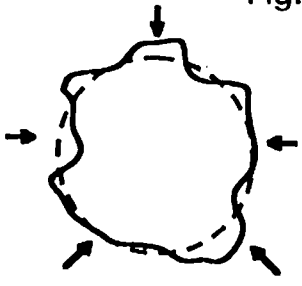


Fig. 14C

Fig. 14B

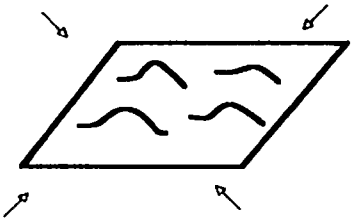


Fig. 14A

Fig.7 Schematic image of surface in each phase. The dotted sphere represents the effective sphere of (6).

6 Discussions and Conclusions

In conclusion, we have clarified the following properties of 2d R^2-gravity using the semiclassical formalism.

1). All classical vacua correponding to the constant curvature configuration are obtained.

2). All phases appearing in the obtained solutions are characterized by the asymptotic behaviours.

3). Thermodynamic properties are clarified. Especially the equation of states, the temperature and the entropy are obtained. They are macroscopic properties of the fully-quantum system.

We have some comments as follows.

- Among two branches, the $-$ branch (of the positive curvature) solution appears in the lattice simulation[5, 6]. It is consistent with the present analysis, where $-$ branch is energetically prefarable to $+$ branch. Some features of $+$ branch are the same as those obtained in [11] using the conformal field approach[5]. It seems important to analyse the relation between the present semiclassical approach and the conformal field approach.

- The semiclassical approach can easily provide the physical image of a system through its thermodynamic properties. The present system can be regarded as a thermodynamic system where many scalar-matter particles and gravitons move with gravitational interaction and whose macroscopic configuration is thermally in an equilibrium state. The temperature depends on the R^2 coupling, w (or β). The phase difference can be thermodynamically interpreted as the difference of w-dependence of the temperature.

The present approach is valid for the higher-dimensional quantum gravity. The 3 dim QG has been recently 'measured' in the Lattice simulation with a high statistics[12]. The semiclassical analysis of the data will soon become an urgent work to be done. The success of the perturbative 2d QG using the semiclassical method is strongly encouraging for the further progress of the perturbative quantum gravity in the realistic dimensions.

References

[1] E.D'Hoker and R.Jackiw,Phys.Rev.**D26**,3517(1982)

[2] N.Seiberg,Prog.Theor.Phys.Suppl.**102**,319(1990)

[3] J.Ambjørn,S.Jain and G.Thorleifsson,Phys.Lett.**307B**,34(1993)

[4] J.Ambjørn and G.Thorleifsson,Phys.Lett.**323B**,7(1994)

[5] S.Ichinose,N.Tsuda and T.Yukawa,Nucl.Phys.**B445**(1995)295

[6] S.Ichinose,N.Tsuda and T.Yukawa,Preprint of Univ.of Shizuoka,US-95-05, "MINBU Dstribution of Two Dimensional Quantum Gravity: Simulation Result and Semiclassical Analysis"

[7] V.G.Knizhnik,A.M.Polyakov and A.B.Zamolodchikov,Mod.Phys.Lett.**A3**,819(1988)

[8] A.M.Polyakov,Phys.Lett.**103B**,207(1981)

[9] S.Ichinose,Preprint of Univ. of Shizuoka, US-95-06,to be published in Int.Jour.Mod.Phys.A, "Thermodynamic Properties, Phases and Classical Vacua of Two Dimensional R^2-Gravity"

[10] S.Ichinose,Nucl.Phys.**B445**(1995)311;**B457**(1995)688(Erratum)

[11] H.Kawai and R.Nakayama,Phys.Lett.**306B**,224(1993)

[12] H.Hagura,N.Tsuda and T.Yukawa, Preprint of KEK,Tsukuba, KEK-CP-040, UTPP-45(1996), 'Phases and fractal structures of three-dimensional simplicial gravity'

Analysis of the Back-Reaction on the Global Degrees of Freedom in (2+1)-Dimensional Spacetime

Masafumi Seriu[†]

Yukawa Institute for Theoretical Physics
Kyoto University, Kyoto 606, Japan

Abstract

We investigate the back-reaction effect of a quantum scalar field on the global degrees of freedom in (2+1)-dimensional toroidal universe.

We construct a homogeneous model of the toroidal universe, $\mathcal{M} \simeq T^2 \times \mathbf{R}$, and examine explicitly the back-reaction effect of the Casimir energy of a massless, conformally coupled scalar field, with a conformal vacuum. The system reduces to the Hamiltonian system for three canonical pairs, (V, σ), (τ^1, p_1) and (τ^2, p_2), where V is the 2-volume (area) of the torus, and (τ^1, τ^2) are the Teichmüller parameters describing the global shape of the torus. The Casimir energy of the matter field is represented as a modular-invariant function of V and (τ^1, τ^2), which causes non-trivial back-reaction effects. The back-reaction causes an instability of the universe: The torus becomes thinner and thinner as it evolves, while its total 2-volume becomes smaller and smaller.

The back-reaction caused by the Casimir energy can be compared with the influence of the negative cosmological constant: Both of them make the system unstable and the torus becomes thinner and thinner in shape. On the other hand, the Casimir energy is a complicated function of the Teichmüller parameters (τ^1, τ^2) causing highly non-trivial dynamical evolutions, while the cosmological constant is simply a constant.

We also analyze the asymptotic properties of the canonical equations as $t \to \infty$, and show that the instability of the torus due to back-reaction is universal irrespective of the initial conditions.

[†]Research Fellow of the Japan Society for the Promotion of Science.

1 Introduction

General Relativity in the Hamiltonian form describes the dynamics of the geometry of a spatial hypersurface. The geometrical information possessed by the hypersurface can be classified into two categories: Local geometry and global geometry. Here, the local geometry is typically represented by the spatial metric h_{ab}, while the global geometry consists of topology and other global properties such as moduli[1] . So far, most of the efforts have been focused on the analysis of the dynamics of the local geometry, and the dynamics of the global geometry has rarely been investigated.

Let us look at the semiclassical gravity from this standpoint. One of the basic subjects to be investigated in semiclassical gravity is the analysis of the back-reaction from quantum matter fields on the spacetime structures. We are naturally led to the concept of the 'back-reaction on the global geometry', in view of the intensive investigations of the back-reaction effects on the local geometry. For instance, from the viewpoint of the scale-dependent topology (physical topology) [1], it is of great importance to describe the dynamics of a topological handle, possibly driven by the back-reaction from quantum matter: Whether the handle becomes bigger in size, whether it 'combines' with other handles to make another kind of handle, and so on. However, such a problem has seldom been analyzed so far. One reason for this situation is the difficulty in constructing models of topological handles as a solution of the initial value problem. Another reason is the lack of the analytical, quantitative representation of the global geometrical structures. It may be said that we have not found an appropriate language to speak in the arena of physics of global spacetime structures [1, 2].

As the first attempt in this direction, we set up and analyze a tractable model, which yields many interesting results [3]: We here consider a $(2+1)$-dimensional spacetime $\mathcal{M} \simeq \Sigma \times \mathbf{R}$, with $\Sigma \simeq T^2$, a torus. As a matter field, we take a massless conformally coupled scalar field with a conformal vacuum. Then, we investigate explicitly the back-reaction effect, resulting from the Casimir energy of matter, on the global degrees of freedom of the torus.

[1]It should be noted that such a classification cannot be done without any ambiguity. Needless to say, these two categories are deeply linked with each other. It is most clearly illustrated by the Gauss-Bonnet theorem, or more general index theorems. Causal structures, which are characteristic in pseudo-Riemannian geometry, may be on the border between two categories.

There are several advantages in reducing the number of spacetime dimension from 4 to 3: In the case of dimension 3, only the finite global modes and a spatial volume remain dynamical in pure gravity [4, 5, 6]. Choosing a matter field in a vacuum state preserves this nice property. Another advantage of the reduction of dimension is that the treatment of the quantum field becomes simplified. For instance, no trace anomaly appears in our model, which simplifies the manipulations. One of other advantages of dimension 3 is that we can fix the path-integral measure by the use of the techniques developed in string theories, though it is beyond the scope of this brief report [3].

2 The model

We consider a model of (2+1)-dimensional spacetime with topology $T^2 \times \mathbf{R}$, where the space $\Sigma \simeq T^2$ is locally flat. A metric is given as [2]

$$dl^2 = -dt^2 + V\hat{h}_{ab}d\xi^a d\xi^b \quad , \tag{1}$$

where

$$\hat{h}_{ab} = \frac{1}{\tau^2}\begin{pmatrix} 1 & \tau^1 \\ \tau^1 & |\tau|^2 \end{pmatrix} \quad . \tag{2}$$

Both of the coordinates ξ^1 and ξ^2 are identified with period 1 to get a torus. Here [3], (τ^1, τ^2) are the Teichmüller parameters independent of spatial coordinates (ξ^1, ξ^2), and $\tau := \tau^1 + i\tau^2$, $\tau^2 > 0$. In Eq.(1), V is also independent of (ξ^1, ξ^2), and it is interpreted as a 2-volume (area) of the torus.

Although our torus is a locally flat surface, its global geometry is non-trivial: The parameter τ^2 characterizes the thinness of the torus, and the limits $\tau^2 \downarrow 0$ or $\tau^2 \to \infty$ correspond to the thin-limit of the torus. On the other hand, the parameter τ^1 characterizes, in a sense, the 'twisting' of the torus, which can be represented by the number of intersections of two mutually-orthogonal geodesics.

As a background spacetime, we take a flat spacetime for which V and (τ^1, τ^2) are constants w.r.t. the time parameter t. Therefore, our background

[2]Throughout this report, we choose appropriate units such that $c = 1$, and that the Einstein-Hilbert action becomes $S = \int d^3x \sqrt{-g}\, R$.

[3]Throughout this report, τ^2 always indicates the second component of (τ^1, τ^2), and not the square of τ. The latter never appears in the formulas.

spacetime is conformally flat. What we shall observe is that the back-reaction of matter causes non-trivial evolutions of V and (τ^1, τ^2).

As a matter field, we choose a massless conformally coupled scalar field ψ,

$$S_m = -\frac{1}{2}\int (g^{\alpha\beta}\partial_\alpha\psi\partial_\beta\psi + \frac{1}{8}R\psi^2)\sqrt{-g}d^3x \quad . \tag{3}$$

We choose the conformal vacuum as a vacuum state for the matter field. Because our background spacetime is conformally flat, only we have to do is to compute $< T_{\alpha\beta}(\eta) >$ using the metric $ds^2 = -dt^2 + \hat{h}_{ab}d\xi^a d\xi^b$, which can most easily be done by the method of images. (Here enters into discussions the information of torus-topology, or the periodicity w.r.t. (ξ^1, ξ^2).) Then we obtain $< T_{\alpha\beta}(g) >$, which we really need, by the relation

$$< T_{\alpha\beta}(g) >= V^{-1/2} < T_{\alpha\beta}(\eta) > \quad . \tag{4}$$

Note that this simplification occurs, because the trace anomaly $< T^\alpha_\alpha(g) >$ vanishes when the spacetime dimension is odd. It is also useful to note that the basic modes of the field w.r.t. the metric $ds^2 = -dt^2 + \hat{h}_{ab}d\xi^a d\xi^b$ are given as

$$\omega_{n_1 n_2} = 2\pi(\tau^2)^{-1/2}(|\tau|^2 n_1^2 - 2\tau^1 n_1 n_2 + n_2^2)^{1/2} \quad (n_1, n_2 \in \mathbf{Z}) \quad . \tag{5}$$

At this stage, we should keep in mind the limitation of our analysis, which all of the back-reaction problems more or less share in common: Since the back-reaction causes the time-evolution of (ξ^1, ξ^2), the spacetime does not remain as conformally flat. Therefore, $< T^\alpha_\alpha(g) >$ calculated as above should be regarded as an approximate one. In other words, we should look at the results of the analysis in an adiabatic sense, which are valid when terms including $\dot{\tau}^1$ and $\dot{\tau}^2$ are not dominant in the formulas. We shall see in §4 that, in our present model, this adiabatic treatment is a good approximation because $\dot{\tau}^1$ and $\dot{\tau}^2$ turn out to be sufficiently small.

3 The back-reaction on the global geometry of a torus

The action describing the back-reaction effect is given by

$$S = \int dx^3(\pi^{ab}\dot{h}_{ab} - N\mathcal{H} - N^a\mathcal{H}_a) \quad , \tag{6}$$

where the Hamiltonian and momentum constraints are [4], respectively,

$$\mathcal{H} = \{(K_{ab}K^{ab} - K^2 - {}^{(2)}R) + < T_{\alpha\beta} > n^\alpha n^\beta\}\sqrt{h} \quad , \tag{7}$$

$$\mathcal{H}_a/\sqrt{h} = -2D_b(K_a{}^b - \delta_a{}^b K) - < T_{\alpha\beta} > n^\beta \quad . \tag{8}$$

It turns out that, in our background spacetime, $< T_{\alpha\beta} > n^\beta = 0$, and that $\sigma := -K$ =spatially constant (York's time slicing). Therefore, Eq.(8) boils down to

$$\mathcal{H}_a/\sqrt{h} = -2D_b\tilde{K}_a^b = 0 \quad , \tag{9}$$

where $\tilde{K}_a^b := K_a{}^b - \frac{1}{2}\delta_a{}^b K$. It is a standard result in the theory of Riemann surfaces that there are two independent solutions to Eq.(9) on $\Sigma \simeq T^2$, so that $\tilde{K}^{ab}$ can be represented as a linear combination of the solutions, $\{\Psi^{A\,ab}\}_{A=1,2}$ [6], $\tilde{K}^{ab} = \sum_{A=1}^{2} p_A \Psi^{A\,ab}$. It can also be shown that, on $\Sigma \simeq T^2$, the lapse function N can be chosen as $N = N(t)$ without any contradiction with the York's slice [6]. Now, it is straightforward to see that our system reduces to [3]

$$S = \int dt(\sigma\frac{dV}{dt} + \sum_{A=1}^{2} p_A \frac{d\tau^A}{dt} - N(t)H) \quad , \tag{10}$$

where

$$H = \frac{(\tau^2)^2}{2V}(p_1^2 + p_2^2) - \frac{1}{2}\sigma^2 V - \hbar\tau^{2\,3/2} f(\tau) V^{-1/2} = 0 \quad . \tag{11}$$

Here,

$$\begin{aligned} f(\tau^1, \tau^2) &:= \frac{1}{4\pi} \sum_{n_1, n_2 = -\infty}^{\infty\prime} \frac{1}{|n_1 + \tau n_2|^3} \\ &= \frac{1}{4\pi} \sum_{n_1, n_2 = -\infty}^{\infty\prime} \frac{1}{(n_1^2 + 2\tau^1 n_1 n_2 + |\tau|^2 n_2^2)^{3/2}} \quad . \end{aligned} \tag{12}$$

It is obvious that $f(-\tau^1, \tau^2) = f(\tau^1, \tau^2)$, $f(\tau^1 + n, \tau^2) = f(\tau^1, \tau^2)$, $f(n + a, \tau^2) = f(n - a, \tau^2)$ (n: integer, a: real) and that $f(\tau^1, \tau^2)$ is singular at $(\tau^1, \tau^2) = (n, 0)$. Furthermore, the combination $2\pi(\tau^2)^{3/2} f(\tau^1, \tau^2)$ on

[4]Here, $n^\alpha = (1/N, -N^a/N)$ is the normal unit vector of the spatial surface Σ, K_{ab} is the extrinsic curvature of Σ, $K := K_a^a$, ${}^{(2)}R$ is the scalar curvature on Σ, and D_a is the covariant derivative w.r.t. the spatial metric (see Eqs.(1) and (2)).

the L.H.S. of Eq.(11) is known as the non-holomorphic Eisenstein series $G(\tau, 3/2)$, whose modular invariance can be shown easily [5]. The first term on the L.H.S. of Eq.(11) is also modular invariant. Thus, the Hamiltonian constraint Eq.(11) is modular invariant as it should be. The last term in Eq.(11) comes from $< T_{\alpha\beta} > n^\alpha n^\beta$ in Eq.(7). This negative contribution to the Hamiltonian constraint is understood as the Casimir energy, which reflects the torus-topology of Σ. Note that it is $O(\hbar)$.

Now, we analyze the back-reaction effect. From Eqs.(10) and (7), we have

$$\dot{V} = -\sigma V \quad , \tag{13}$$

$$\dot{\sigma} = \frac{1}{2}\sigma^2 + \frac{(\tau^2)^2}{2V^2}(p_1^2 + p_2^2) - \frac{\hbar}{2}(\tau^2)^{3/2} f(\tau) V^{-3/2} \quad , \tag{14}$$

$$\dot{\tau}^1 = \frac{1}{V}(\tau^2)^2 p_1 \quad , \tag{15}$$

$$\dot{p}_1 = \hbar(\tau^2)^{3/2} \frac{\partial f(\tau)}{\partial \tau^1} V^{-1/2} \quad , \tag{16}$$

$$\dot{\tau}^2 = \frac{1}{V}(\tau^2)^2 p_2 \quad , \tag{17}$$

$$\dot{p}_2 = -\frac{1}{V}\tau^2(p_1^2 + p_2^2) + \frac{3\hbar}{2}(\tau^2)^{1/2} f(\tau) V^{-1/2} + \hbar(\tau^2)^{3/2} \frac{\partial f(\tau)}{\partial \tau^2} V^{-1/2} \quad . \tag{18}$$

We can see that the time evolution becomes trivial in a sense when there is no matter field, $f(\tau) \equiv 0$: In this case, Eqs.(11), (13) and (14) allow a solution, V =constant, $\sigma \equiv 0$, $p_1 = p_2 \equiv 0$. Then, Eqs.(11) and (15)-(18) do not allow any solution, compatible with V =constant, $\sigma \equiv 0$, other than τ^1 =constant, τ^2 =constant. The spacetime corresponding to this solution is the static one constructed from 3-dimensional Minkowski space in the standard coordinates (T, X^1, X^2) with suitable identifications in spatial section (X^1, X^2) described by (τ^1, τ^2). (Of course, the same static spacetime can also be represented in the Hamiltonian form using different time slices: The initial condition $\sigma \neq 0$ allows (τ^1, τ^2) to evolve in time.)

The back-reaction of the quantum field causes a non-trivial evolution of (τ^1, τ^2), i.e. global deformations of the torus. Due to the negativity of the

[5] The modular invariance is defined as the invariance under the transformation $\tau \mapsto \tau' = \frac{a\tau + b}{c\tau + d}$, where $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbf{Z})/\{\mathbf{1}, -\mathbf{1}\}$.

term $-\hbar(\tau^2)^{3/2}f(\tau)V^{-1/2}$ in Eq.(11), a non-trivial evolution of (τ^1, τ^2) occurs, even when $\sigma \approx 0$ initially so that the term $-\frac{1}{2}\sigma^2 V$ in Eq.(11) is negligible. The choice V =constant, $\sigma \equiv 0$ is not allowed as a solution any more, as is seen from Eqs.(13) and (14).

Let us see *Figures* 1,2,3 which show a typical example of the evolution of (τ^1, τ^2), (p_1, p_2) and (V, σ), respectively. Units s.t. $\hbar = 1$ have been chosen. The initial conditions are $V = 1.000$, $\sigma = 0.000$, $(\tau^1, \tau^2) = (0.500, 0.500)$ and $p_1 = 1.000$. The initial condition for p_2, $p_2 = 2.175$, has been determined by the constraint equation Eq.(11). The typical points A-E and Z have been shown in *Figures* to give an idea of the time-evolution. (The point A is the initial point and the point Z is the final point of the calculation. For instance, the point C in *Figure* 1 and the point C in *Figure* 2 correspond to the same instant of time.)

We observe the same asymptotic behavior of the system due to the back-reaction irrespective of the initial conditions: The back-reaction drives the system into the direction corresponding to a thinner torus, i.e. $\tau^2 \to 0$ while $\tau^1 \to$ finite. At the same time, the 2-volume V asymptotically approaches zero. We find out that this behavior is universal by setting various generic initial conditions. This universal behavior can also be understood by investigating the asymptotic characteristics of Eqs. (11)-(18) as $t \to \infty$, which shall be discussed in the next section.

4 The asymptotic analysis of the evolutions

Our treatment is based on the adiabatic approximation as is noted at the end of §2. Thus, the results should be taken with a caveat. In the present case, there are good reasons to regard the unstable behavior as a real one. Let us discuss on this point.

We can analyze the general asymptotic behavior of the system, by investigating Eqs.(11)-(18). We can show that [3],

(a) $V \to 0$, $\sigma \to \infty$, $\sigma V^n \to \infty$ $(n = 1, 2, 3, \cdots)$.

(b) $(\tau^2)^2(p_1^2 + p_2^2)$ increases, at least as [6] $\sigma^2 V^2$.

[6]Here, '$y(t)$ increases at least as $x(t)$' means that, $|x(t)/y(t)| \to c$, $0 \leq c < \infty$ when $t \to \infty$.

(c) $\tau^2 \downarrow 0$, or $\tau^2 \to \infty$.

(d) $p_1\dot{\tau}^1 + p_2\dot{\tau}^2$ increases at least as $\sigma^2 V$, and $\frac{1}{(\tau^2)^2}((\dot{\tau}^1)^2 + (\dot{\tau}^2)^2)$ increases at least as σ^2.

The asymptotic behavior of the generic trajectories shows that $\dot{\tau}^1 \to 0$, $\dot{\tau}^2 \to 0$ (and $\dot{V} \sim o(\sigma)$) although $p_2 \to -\infty$. This comes from the fact that $\dot{\tau}^1 \propto \frac{(\tau^2)^2 p_1}{V}$, $\dot{\tau}^2 \propto \frac{(\tau^2)^2 p_2}{V}$, and τ^2 becomes a strong suppression (stronger than $\exp\sigma t$), while $1/V$ is at most $\sim \sigma$. Thus, the adiabatic treatment for τ^1 and τ^2 becomes better and better as $\tau^2 \to 0$: From Eq.(5), $\dot{\omega}_{n_1 n_2}/\omega^2_{n_1 n_2} \sim (\tau^2)^{3/2} \cdot \frac{\dot{\tau}^2}{(\tau^2)^2} = (\tau^2)^{1/2} \cdot (\frac{\dot{\tau}^2}{\tau^2}) \to 0$. Furthermore, $\dot{V}$ does not harm the adiabatic treatment because of the conformal invariance of the matter field.

It is interesting to compare our model with the case of the negative cosmological constant without matter field, which also serves as another support for our results.

Looking at Eq.(7), the Λ-term can be introduced with ease:

$$H = \frac{(\tau^2)^2}{2V}(p_1^2 + p_2^2) - \frac{1}{2}\sigma^2 V - \Lambda V = 0 \quad , \tag{19}$$

$$\dot{V} = -\sigma V \quad , \tag{20}$$

$$\dot{\sigma} = \frac{1}{2}\sigma^2 + \frac{(\tau^2)^2}{2V^2}(p_1^2 + p_2^2) + \Lambda \quad , \tag{21}$$

$$\dot{\tau}^1 = \frac{1}{V}(\tau^2)^2 p_1 \quad , \tag{22}$$

$$\dot{p_1} = 0 \quad , \tag{23}$$

$$\dot{\tau}^2 = \frac{1}{V}(\tau^2)^2 p_2 \quad , \tag{24}$$

$$\dot{p_2} = -\frac{1}{V}\tau^2(p_1^2 + p_2^2) \quad . \tag{25}$$

Here, $-\Lambda$ corresponds to the cosmological constant ($\Lambda > 0$). The negativity of the last term in Eq.(19) causes the same kind of the evolution for (τ^1, τ^2) as in the case of the matter field. It strongly suggests that the instability in our model is a real one independent of the adiabatic treatment. We should also note the essential difference between our case and the case of the negative cosmological constant: The difference between (16) and (23) is prominent in particular. Furthermore, $\hbar(\tau^2)^{3/2} f(\tau) V^{-3/2}$, corresponding to Λ (compare (11) with (19)), depends on (τ^1, τ^2) and V, causing a highly non-trivial evolution.

5 Concluding remarks

We have focussed on the dynamics of the global geometry of a spatial surface. As the first step in this direction, we have posed the concept of the 'back-reaction on the global structures'. Setting a (2+1)-dimensional toroidal spacetime, $\Sigma \simeq T^2 \times \mathbf{R}$, we have investigated the back-reaction of a quantum matter field on the moduli of the torus. We have seen the non-trivial evolutions of the moduli caused by the back-reaction. Our results suggest that the back-reactions play a very important role to understand the semiclassical dynamics of the global structures of the universe.

At the same time, to go beyond the present analysis, we feel the need for a suitable representation of the global structures, which can be handled by the analytic schemes [2]. Only after doing this, we may get the whole view on the theory of the spacetime structures.

* * *

This work has been financially supported by the Japan Association for Mathematical Sciences, the Yukawa Memorial Foundation, and the Japan Society for the Promotion of Science.

References

[1] M. Seriu, Phys. Lett. **B319** (1993), 74.

[2] M. Seriu, to appear in Phys. Rev. **D** (YITP-96-9, gr-qc/9603002).

[3] For detailed discussions, see M. Seriu, Phys. Rev. **D53** (1996), 1889.

[4] E. Martinec, Phys. Rev. **D30** (1984), 1198.

[5] V. Moncrief, J. Math. Phys. **30** (1989), 2907.

[6] A. Hosoya and K. Nakao, Prog. Theo. Phys. **84** (1990), 739.

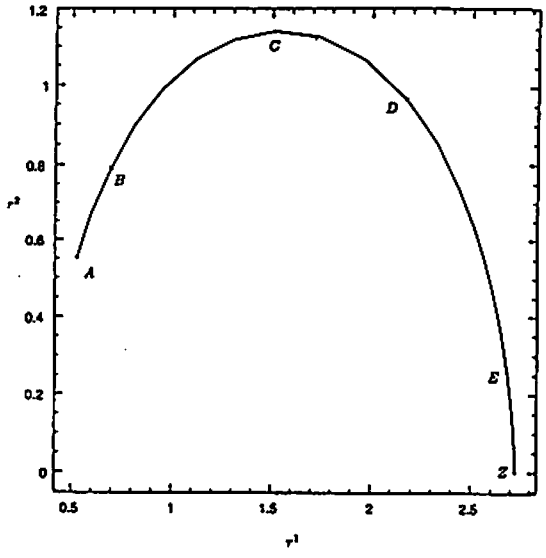


Figure 1

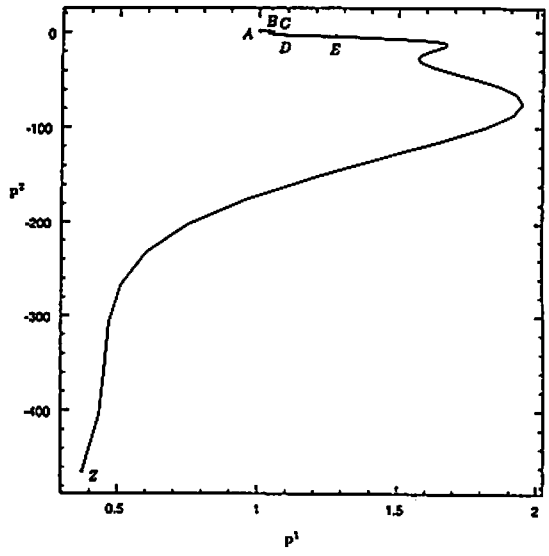


Figure 2

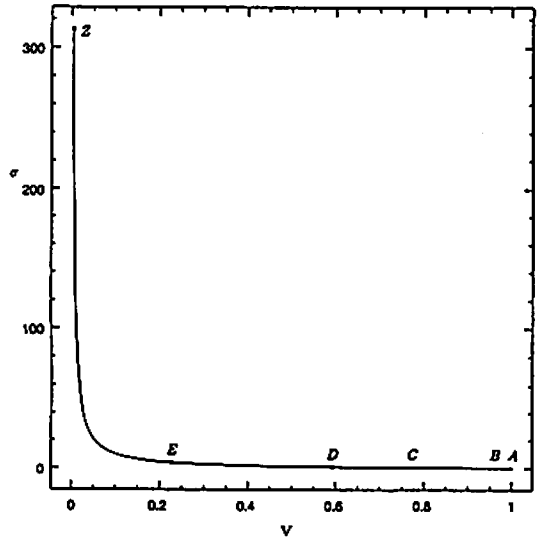


Figure 3

Quantum Harmonic Oscillators in the expanding universe

Atsushi TARUYA * and Yasusada NAMBU †

Department of Physics, Nagoya University, Chikusa-ku, Nagoya 464-01, Japan

(February 26, 1996)

Abstract

We comment on the decoherence of the density perturbation concerning the structure formation of our universe. We show that the cosmic expansion plays an important role in the quantum-classical transition. To clarify the effect of the cosmic expansion quantitatively, we use the two-coupled harmonic oscillators as a toy model. We derive the master equation by tracing out the environment and analyze the reduced density matrix for an arbitrary expansion rate. The preliminary analysis indicates that the cosmic expansion with the small expansion rate does not affect so much the quantum decoherence.

I. INTRODUCTION

The structure formation of our universe is one of the most important issue in cosmology. We believe that due to the gravitational instability, the small density inhomogeneities grew into the present large scale structure such as the voids and the filaments etc. Inflation explains the origin of such the density inhomogeneities although it is an idea which resolves the problems of the homogeneity and flatness. That is, the fluctuation of the quantum field (inflaton field) was produced during inflation and evolved into the classical density pertur-

*e-mail address: ataruya@allegro.phys.nagoya-u.ac.jp

†e-mail address: nambu@allegro.phys.nagoya-u.ac.jp

bation. We should remark that this process includes the classicalization of the quantum system. It has been argued that at least two condition are necessary for the classicalization, i.e,

- The classical correlation is realized,
- The quantum coherence is lost (decoherence) .

The former corresponds that the Wigner function has the strong peak along the classical trajectory. The latter means that the off-diagonal element (the interference part) of the density matrix vanishes. Because quantum mechanics itself satisfies the unitary evolution, we need a description of irreversible evolution for the quantum-classical transition of the density perturbation. In order to implement such the description, it is necessary to take a coarse-graining.

The coarse-graining is a method that is often used in the paradigm of the quantum brownian motion(QBM), which explains the friction and the random force kinematically. Consider the system interacting with the external degree of freedom which refers to the environment. Ignoring the environmental degree of freedom, the system could obey so-called master equation which represents the evolution of the reduced density matrix [1]-[3].

By introducing such the coarse-graining, several authors discuss the decoherence of the density perturbation at the inflationary era [5]- [9]. Projecting out the environment, similar results are obtained: when the scale of the fluctuation exceeds the Hubble horizon, the quantum coherence of the scalar field is lost. However, we should carefully treat the decoherence in the expanding universe. We can show that it is quite different from the decoherence on QBM. This means that the cosmic expansion strongly affects the decoherence of density perturbation. Recently, there is a remarkable discussion that due to the cosmic expansion, *the decoherence without environment* is also possible [10]. In order to discuss this issue, it is necessary to understand the role of the cosmic expansion clearly.

In this article, we treat the decoherence of the density perturbation and consider the effect of the cosmic expansion. In sec.II, we explain the coarse-graining in QBM and the necessary condition for the decoherence. We point out that the decoherence of the density perturbation is different from that on QBM. In sec.III, in order to clarify the role of the cosmic expansion, we use two-coupled harmonic oscillators and analyze the master equation. The last section IV is devoted to summary and discussion.

II. THE COARSE-GRAINING IN QBM AND THE DECOHERENCE OF THE DENSITY PERTURBATION

A. Quantum brownian motion

QBM is a paradigm of the quantum open system. The main motivation is to provide the quantum-mechanical representation for a brownian particle described by Langevin equation such as

$$\ddot{Q} + \Gamma\dot{Q} + V'(Q) = \xi(t), \quad \langle\xi(t)\rangle = 0, \quad \langle\xi(t)\xi(t')\rangle = \nu(t-t'), \tag{1}$$

where Γ, ν denote the friction and the noise coefficients, respectively. The essential point to represent the brownian motion is to introduce the external degrees of freedom (environment). When the system(Q) interacts with the environmental degrees of freedom($\{q_n\}$), the Lagrangian can be written in the form

$$L = L_{sys}(Q) + L_{env}(\{q_n\}) + L_{int}(Q;\{q_n\}). \tag{2}$$

To get the irreversibility of the brownian particle, it is necessary to coarse-grain the environment. In terms of the density matrix, it corresponds to trace out the environmental degrees of freedom:

$$\rho_{red}(Q,Q') = \prod_n \int dq_n dq_n' \delta(q_n - q_n')\rho(Q,Q';\{q_n,q_n'\}). \tag{3}$$

The reduced density matrix ρ_{red}, which describe the evolution of the system, obeys the master equation. We should remark that it contains some additional terms except for the Liouville term of the system :

$$\dot{\rho}_{red}(Q,Q') = \boldsymbol{L}\rho_{red} - \frac{F}{2}(Q-Q')\left(\frac{\partial}{\partial Q} - \frac{\partial}{\partial Q'}\right)\rho_{red} - D(Q-Q')^2\rho_{red} + \cdots. \qquad (4)$$

where $\boldsymbol{L}$ is Liouville operator and F , D are the dissipation and the diffusion coefficients respectively, which relate with Γ, ν in eq.(1). We observe that the third term of RHS in eq.(4) could lead to the decoherence.

However, we must require the further condition to the quantum decoherence: the environment must have *infinite degrees of freedom.* Unless this condition holds, the coherence of the system cannot be lost. In other words, in the case of the environment with finite degrees of freedom, information about the off-diagonal part of ρ_{red} does return in finite characteristic time even though it is transferred to the environment. [1]

B. decoherence of the density perturbation

Using the treatment of the coarse-graining in QBM (introducing and tracing out the environment), several authors use the toy models and discuss the docoherence of the density perturbation. Some of them pay an attention to the stochastic dynamics due to the short-wavelength mode of quantum field [5] [6] [7]. They treat the long-wavelength mode ($k/aH \lesssim 1$) of the scalar field as the system and the short-wavelength mode as the environment. Others consider the two-interacting scalar fields and regards one of them as the environment [8] [9]. Both authors analyze the density matrix in the inflationary era by tracing out the environment. Here, we specifically focus on the model discussed by Brandenberger *et al.* [9], which adopts the latter approach. They use the two massless scalar fields (ϕ, χ) with the

---

[1] There is an interesting discussion that the non-linearity(chaos) also plays a sufficient role to the diagonalization of ρ_{red}. [4]

gradient interaction($\partial_\mu\phi\partial^\mu\chi$). The important point is that this model is equivalent to the two-coupled harmonic oscillators. When we put the metric

$$ds^2 = a^2(\eta)[-d\eta^2 + d\vec{x}^2], \tag{5}$$

the action can be written in the Fourier representation as follows:

$$\begin{aligned} \mathcal{I} &= \int dt \int d^3k \left\{ \mathcal{L}^k_{sys.} + \mathcal{L}^k_{env.} + \mathcal{L}^k_{int.} \right\} \qquad (6) \\ \mathcal{L}^k_{sys} &= \frac{a^2}{2}\left\{ \left(\frac{d\phi_k}{d\eta}\right)^2 - k^2\phi_k^2 \right\}, \\ \mathcal{L}^k_{env} &= \frac{a^2}{2}\left\{ \left(\frac{d\chi_k}{d\eta}\right)^2 - k^2\chi_k^2 \right\}, \\ \mathcal{L}_{int} &= a^2\epsilon\left(\frac{d\phi_k}{d\eta}\frac{d\chi_k}{d\eta} - k^2\phi_k\chi_k \right), \end{aligned}$$

where ϵ is coupling constant. Although the model has infinite degrees of freedom, the system ϕ_k interacts only with the same k-mode of the environment χ_k. Accordingly this model can be regarded as infinite sets of the two-coupled oscillators. From the analysis of the density matrix in de Sitter space, they conclude that the coherence of the long-wavelength mode ϕ_k is rapidly lost. It seems that the decoherence may occur even when the environment has finite degrees of freedom. Remember that it is necessary for the decoherence in QBM (flat space) to have the infinite degrees of freedom. Comparing both results, we recognize that the cosmic expansion plays an important role in the decoherence of the density perturbation.

Recently, Polarski and Starobinsky use the single massless scalar field and consider *the decoherence without environment* taking into account the cosmic expansion [10]. Since the massless scalar field can be regarded as an upside-down harmonic oscillator, the solution of its mode function has the growing-mode and the decaying mode. They say that the decoherence without environment corresponds to neglecting the decaying mode. This coarse-graining is different from that in QBM. Therefore it is difficult to understand in terms of the density matrix and we think it needs another formulation. However, we should investigate further the nature of the decoherence in the expanding universe. Keeping in mind that most of the discussions about the decoherence are restricted in the stage of the inflation. In order

to investigate how the cosmic expansion affects the decoherence, we should analyze the quantum coherence under an arbitrary expansion rate.

III. MASTER EQUATION FOR THE QUANTUM HARMONIC OSCILLATORS

In sec.II, we saw that the cosmic expansion seems to play an important role for the quantum decoherence of the density perturbation. In this section, using a toy model, we will investigate more precisely how the cosmic expansion affect the decoherence. Following Brandenberger *et al.* [9], we adopt the two-coupled harmonic oscillators, which is the simplest model to study the nature of the decoherence in the expanding universe. Tracing out one-side degree of freedom, we derive the master equation. By varying the expansion rate, we shall analyze behavior of the reduced density matrix.

A. Master equation

Consider the two-coupled harmonic oscillators Q, q with qQ-interaction. When we choose the synchronous time-slicing ($ds^2 = -dt^2 + a^2(t)d\vec{x}^2$), the Lagrangian density takes the form

$$\mathcal{L} = a^3 \left[\frac{1}{2} \left(\dot{Q}^2 - \Omega^2 Q^2 \right) + \frac{1}{2} \left(\dot{q}^2 - \omega^2 q^2 \right) - \epsilon q Q \right], \tag{7}$$

where Ω, ω denote the constant frequency. Similar to Ref. [9], this model has only two degrees of freedom. Note that we can also regard (7) as the massive scalar fields in long-wavelength limit. The canonical momenta become

$$P = a^3 \dot{Q}, \quad p = a^3 \dot{q}. \tag{8}$$

Following the method described in sec.II A, we shall derive the master equation. We regard the variable Q as the system and q as the environment. The starting point is the evolution equation of the total density matrix $\rho(q, q'; Q, Q')$. Using the Schrödinger equation, we obtain

$$i\frac{\partial}{\partial t}\rho = \left[-\frac{1}{2a^3}\left(\frac{\partial^2}{\partial Q^2} - \frac{\partial^2}{\partial Q'^2}\right) + \frac{a^3}{2}\Omega^2(Q^2 - Q'^2)\right]\rho$$
$$+\left[-\frac{1}{2a^3}\left(\frac{\partial^2}{\partial q^2} - \frac{\partial^2}{\partial q'^2}\right) + \frac{a^3}{2}\omega^2(q^2 - q'^2)\right]\rho + \epsilon a^3(qQ - q'Q')\rho, \tag{9}$$

where the first and the second brackets of RHS represent the Liouville terms for the system and environment, respectively. The last term denotes the interaction. Because we are specifically interested in the time evolution of the system, we trace out the environment q (see eq.(3)). Assuming that the system and the environment are initially uncorrelated, we derive the following master equation (for detail derivation, see [2] [3]):

$$i\frac{\partial}{\partial t}\rho_{red} = \left[-\frac{1}{2a^3}\left(\frac{\partial^2}{\partial Q^2} - \frac{\partial^2}{\partial Q'^2}\right) + \frac{a^3}{2}\Omega^2(Q^2 - Q'^2)\right]\rho_{red}$$
$$+a^3\left[\frac{1}{2}\delta\Omega^2(t)(Q^2 - Q'^2) - D_1(t)(Q - Q')\left(\frac{\partial}{\partial Q} + \frac{\partial}{\partial Q'}\right)\right.$$
$$\left.-\frac{i}{2}F(t)(Q - Q')\left(\frac{\partial}{\partial Q} - \frac{\partial}{\partial Q'}\right) - iD_2(t)(Q - Q')^2\right]\rho_{red}. \tag{10}$$

The terms contained in the second bracket are appeared by the coarse-graining. Each of them are as follows: the first term means frequency-shift. The second term is t anomalous diffusion term. The third and forth terms are the dissipation and diffusion term respectively. We should keep in mind that these coefficients $\delta\Omega^2, F, D_1, D_2$ are time-dependent in general. Putting the initial state of the environment, we can get these coefficients by using the solution to the equation of motion of (7). For simplicity, we shall put the vacuum state of the flat case as the initial state of the environment. Writing down the solutions as

$$Q(t) = uQ(t_0) + vP(t_0) + wq(t_0) + xp(t_0), \tag{11}$$

$$q(t) = \bar{u}Q(t) + \bar{v}P(t) + \bar{w}q(t_0) + \bar{x}p(t_0), \tag{12}$$

where $u, v, w, x, \bar{u}, \bar{v}, \bar{w}, \bar{x}$ are the time-dependent variables and t_0 is an initial time. The coefficients of the master equation can be obtained uniquely (for more detail determination, see [2]):

$$\delta\Omega^2(t) = \epsilon\bar{u}, \qquad F(t) = \epsilon\bar{v}, \tag{13}$$

$$D_1(t) = -\frac{\epsilon}{2}\left(\frac{w\bar{w}}{\omega} + \omega x\bar{x}\right), \qquad D_2(t) = -\frac{\epsilon}{2}a^3(t)\left(\frac{\dot{w}\bar{w}}{\omega} + \omega\dot{x}\bar{x}\right). \tag{14}$$

B. Analysis

We proceed to investigate the nature of quantum decoherence by varying the expansion rate. We wish to evaluate whether the reduced density matrix $\rho_{red}(Q,Q')$ diagonalizes or not. To discuss quantitatively, it is convenient to introduce the degree of decoherence δ_{QD}. Taking the Gaussian ansatz, the density matrix can be written due to the symmetry of the master equation (10) as follows:

$$\rho_{red}(Q,Q') = N^{-1}\exp\left[-\alpha(Q-Q')^2 - \beta(Q+Q')^2 - i\gamma(Q-Q')(Q+Q')\right], \tag{15}$$

where N denotes normalization. The degree of decoherence δ_{QD} can be defined as

$$\delta_{QD} = \sqrt{\frac{\beta}{\alpha}}. \tag{16}$$

When δ_{QD} becomes 0, we find that the quantum coherence of the system is lost. As for the real functions α, β, γ, we get the evolution equations from (10):

$$\dot{\alpha} = \frac{4\alpha\gamma}{a^3} - a^3\left(2\alpha F(t) + 2\gamma D_1(t) - D_2(t)\right), \tag{17}$$

$$\dot{\beta} = \frac{4\beta\gamma}{a^3}, \tag{18}$$

$$\dot{\gamma} = \frac{2}{a^3}(4\alpha\beta - \gamma^2) - a^3\left\{\frac{1}{2}\left(\Omega^2 + \delta\Omega^2(t)\right) - \gamma F(t) + 4\beta D_1(t)\right\}. \tag{19}$$

Therefore we can solve α, β, γ after obtaining the coefficients $\delta\Omega^2, F, D_1, D_2$.

We shall analyze the decoherence in an arbitrary expansion rate. We put the form of the scale factor as

$$a(t) = \left(\frac{t}{t_0}\right)^p ;\ (p = constant), \tag{20}$$

where t_0 denotes an initial time and we set $t_0 = 1$. As a first step, we evaluate the degree of decoherence δ_{QD} in the non-inflating case, i.e, $0 \leq p \leq 1$. Because we want to study the effect of the cosmic expansion, we restrict the analysis to the weak-coupling case, i.e, $\Omega^2, \omega^2, |\Omega^2 - \omega^2| \gg \epsilon$.

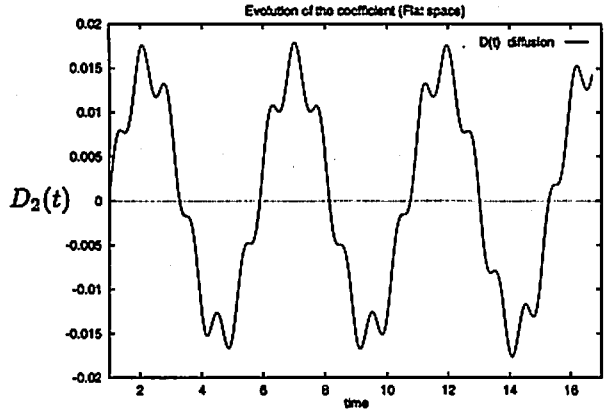


Fig.1-a

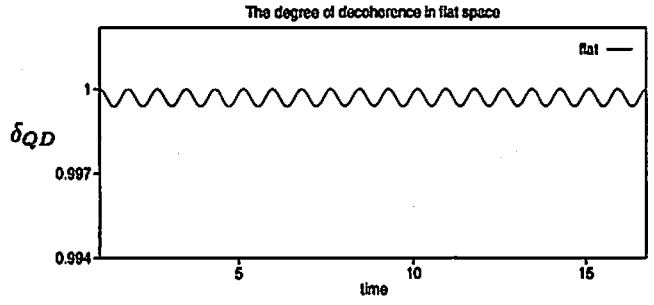


Fig.1-b

FIG.1: The numerical evaluation of the diffusion coefficient and the degree of decoherence in the flat space ($a(t) = 1$): (a) The diffusion coefficient $D_2(t)$; (b) The degree of decoherence δ_{QD}. We put the parameters $\omega = 3.14$, $\Omega = \sqrt{2} \cdot 3.14$, $\epsilon = 0.5$, which correspond to the weakly coupling condition. The calculations are started at $t_0 = 1$.

We firstly examine the flat case($p = 0$). In this case, we find that energy in the system is partially transferred into and recovered from the environment periodically. We can expect that information about the off-diagonal elements of ρ_{red} may also goes out and back periodically and never be lost. The diffusion coefficient D_2, which affects the diagonalization of ρ_{red} (see eq.(10)), helps us to understand the behavior of δ_{QD}. We see that it strongly relates with the energy flow from eq.(14). By the weakly coupling approximation, $D_2(t)$ can be evaluated as

$$D_2(t) \simeq \frac{\epsilon^2}{4\omega}\left[\frac{\sin(\Omega+\omega)(t-t_0)}{\Omega+\omega} + \frac{\sin(\Omega-\omega)(t-t_0)}{\Omega-\omega}\right]. \tag{21}$$

Accordingly the diffusion coefficient becomes oscillatory behavior and its recurrence time τ_{rec} can be estimated as

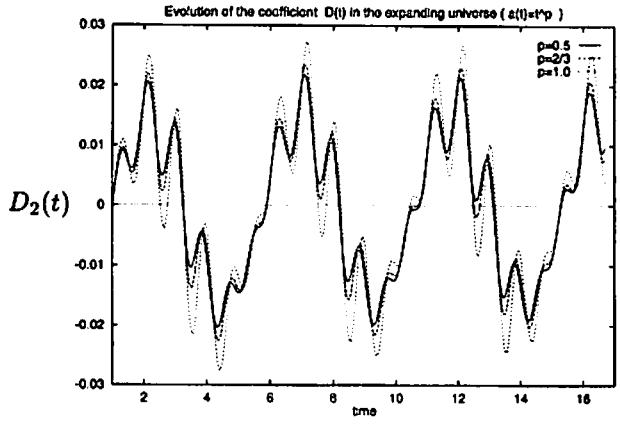


Fig.2-a

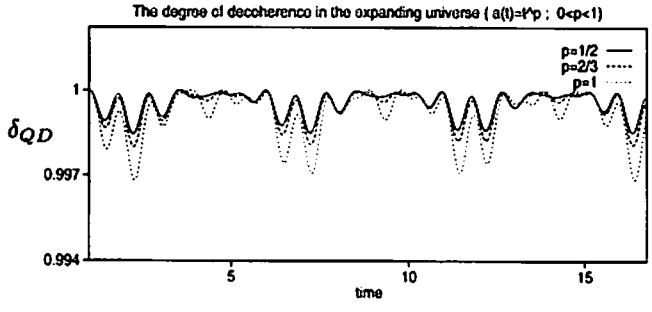


Fig.2-b

FIG.2: The numerical evaluation of the diffusion coefficient and the degree of decoherence in the expanding universe ($a \propto t^p; 0 < p \leq 1$)): (a) The diffusion coefficient $D_2(t)$; (b) The degree of decoherence δ_{QD}. We used the same parameters ω, Ω, ϵ in the flat case and calculated from an initial time $t_0 = 1$. We plotted the cases with the power indices $p = 1/2, 2/3$ and 1, which correspond to the real, dashed and dotted lines respectively.

$$\tau_{rec} \simeq \frac{2\pi}{|\Omega - \omega|}. \tag{22}$$

We also obtained $D_2(t)$ numerically (see Fig.1-a). The results is agreement with the approximation (21). δ_{QD} can be evaluated numerically by solving (17)(18)(19). The results is shown in Fig1-b, where we set the initial condition as $\delta_{QD} = 1$. The graph says that δ_{QD} oscillates and never vanishes. We conclude that in the flat case, the system cannot lose the quantum coherence.

Next, we investigate the decoherence in the expanding universe ($0 < p \leq 1$). Using the same parameters ω, Ω, ϵ in the flat case, we calculated D_2, δ_{QD} numerically (see Fig.2). The graphs are drawn in the cases with the power indices $p = 1/2$(radiation-dominated era),

2/3(matter-dominated era) and 1, which correspond to the real, dashed and dotted lines respectively. Compared with the flat case, we observe that there is not so great difference in the diffusion coefficients. We can also check the behavior by the weakly coupling approximation and obtain the same recurrence time as eq.(22). Thus we understand that the degree of decoherence δ_{QD} does also become the similar behavior to the flat case and that the reduced density matrix cannot diagonalize. This indicates that the cosmic expansion with the small expansion rate does not affect so much the decoherence of the density perturbation.

IV. SUMMARY AND DISCUSSION

In this article, we have discussed the decoherence of the density perturbation on the structure formation of our universe. In order to investigate the effect of the cosmic expansion to the quantum decoherence quantitatively, we have treated the two-coupled harmonic oscillators as a toy model. Applying the coarse-graining approach in QBM, we derived the master equation and investigated the reduced density matrix in the non-inflating case. We showed that there is not so remarkable difference between the flat and the expanding cases. It indicates that the cosmic expansion with the small expansion rate does not affect so much the decoherence of the density perturbation.

The most important point in the problem of decoherence of the density perturbation is to clarify what is essential for the quantum-classical transition in the expanding universe. According to Polarski and Starobinsky, who suggest *the decoherence without environment*, it seems to be essential for the decoherence that the system can be regarded as an upside-down oscillator. In the case of the harmonic oscillator such as the model (7), using the ratio of the frequency(Ω) to the Hubble parameter($H = \dot{a}/a$), we can examine whether the oscillator has the upside-down nature. That is, when $\Omega/H \lesssim 1$, the system can be reduced to the upside-down oscillator. In the non-inflating universe, i.e, $a \propto t^p; 0 \leq p \leq 1$, we can show that the nature of the upside-down oscillator is realized only during the early period. Accordingly we think that the expanding universe with the small expansion rate does not

play an important role to the decoherence of the density perturbation. This is consistent with our preliminary result. During the inflation($p > 1$), however, the system could have the upside-down nature for longer period as the power index p in the Hubble parameter becomes larger. We can expect that the inflating universe affects the quantum decoherence of the density perturbation effectively. We wish to confirm this thing quantitatively by varying the expansion rate. Remember that the self-reproducing fluctuation of the quantum field during inflation becomes the origin of the large-scale structure of our universe. It is also necessary to discuss how the quantum coherence of such the fluctuation evolves after the inflation. Therefore we should extend the analysis from the inflation to the post-inflation continuously. We think that it is possible to investigate the decoherence after inflation by modifying the model (7) to the scalar fields.

Keep in mind that the result obtained here is just preliminary and it is necessary to analyze another model with the different interaction in order to exclude model-dependence. To clarify the role of the cosmic expansion, we shall develop research further.

REFERENCES

[1] A.O.Caldeira and A.J.Leggett, Physica **121A**, 587 (1983).

[2] J.J.Halliwell and T.Yu, Phys.Rev.**D53**, 2012 (1996).

[3] B.L.Hu, J.P.Paz and Y.Zhang, Phys.Rev.**D45**, 2843 (1992).

[4] H.Kubotani, T.Okamura and M.Sakagami, Physica **214A**, 560 (1995).

[5] Y.Nambu, Phys.Lett.**B276**, 11 (1992).

[6] S.Habib, Phys.Rev.**D46**, 2408 (1992).

[7] F.Lombardo and F.D.Mazzitelli, Phys.Rev.**D53**, 2001 (1996).

[8] M.Sakagami, Prog.Theor.Phys.**79**, 442 (1988).

[9] R.Brandenberger, R.Laflamme and M.Mijić, Mod.Phys.Lett. **A5**, 2311 (1990); Physica Scripta. **T36**, 265 (1991)

[10] D.Polarski and A.A.Starobinsky, preprint gr-qc/ 9504030.

Low-Energy Interaction of a Cosmic String and an Extreme Dilatonic Black Hole * †

Kiyoshi Shiraishi‡ §

Akita Junior College

Shimokitade-sakura, Akita-shi, Akita 010, Japan

We construct the solitonic solution in Einstein-Maxwell-dilaton theory on the conical spacetime and investigate the motion of the solitonic object at a low velocity. In the following, using the method of images, we obtain a solitonic solution on a conical space and the low-energy interaction.[1]

I. WHAT IS A "COSMIC STRING"?

The spacetime outside of a straight, thin cosmic string is described by the metric [2]

$$ds^2 = -dt^2 + dr^2 + \frac{r^2}{\nu^2}d\theta^2 + dz^2 \,, \tag{1}$$

with $\nu = (1 - 4G\mu)^{-1}$, where μ is the mass density per unit length.

If we use a new coordinate

---

*Talk given at the fifth workshop on GRG, Nagoya University, 1/22-25/1996.

†This talk is based on the work gr-qc/9512001.

§Address after April 1, 1996: Faculty of Science, Yamaguchi Universiry, Yoshida, Yamaguchi-shi, Yamaguchi 753, Japan.

‡e-mail: g00345@sinet.ad.jp

[1]The multi-soliton system in Einstein-Maxwell-dilaton theory on the spacetime with torus compactification has been studied by obtaining exact solutions and low energy interactions in a similar method [1].

$$\vartheta = \frac{\theta}{\nu} \qquad 0 \leq \vartheta < \frac{2\pi}{\nu} \quad , \tag{2}$$

the metric (1) reduces to the one of the flat spacetime

$$ds^2 = -dt^2 + dr^2 + r^2 d\vartheta^2 + dz^2 \, . \tag{3}$$

Then the deficit angle Δ is defined as [2]

$$\Delta \equiv 2\pi \left(1 - \frac{1}{\nu}\right) = 8\pi G\mu \, . \tag{4}$$

II. WHAT IS AN "EXTREME DILATONIC BLACK HOLE"?

We consider a model described by the following action:

$$S = \int d^4x \, \frac{\sqrt{-g}}{16\pi G} \left[R - 2(\nabla\phi)^2 - e^{-2\phi}F^2\right] + (\text{surface terms}) \, . \tag{5}$$

where R is the scalar curvature, ϕ a dilaton, and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the abelian gauge field strength.

We assume that the electric solution in the Einstein-Maxwell-dilaton theory takes the form [3,4]

$$ds^2 = -\frac{1}{V}dt^2 + V d\boldsymbol{x}^2, \; A_\mu dx^\mu = \frac{1}{\sqrt{2}}\left(1 - \frac{1}{V}\right)dt, \; e^{-2\phi} = V \, . \tag{6}$$

Then V satisfies

$$\nabla^2 V = 0 \, , \tag{7}$$

up to a number of delta functions in the right hand side of Eq. (7).

For example, one can find a solution:

$$V = 1 + \sum_\alpha \frac{2Gm_\alpha}{|\boldsymbol{x} - \boldsymbol{x}_\alpha|} \, . \tag{8}$$

By studying the relation among the electric and dilatonic charge and mass of each solitonic object, we find that the solution describes the configuration that the α-th nonrotating, charged dilatonic black hole in the extreme limit with mass m_α located at $\boldsymbol{x} = \boldsymbol{x}_\alpha$ [3,4].

Now let us construct the black hole solution on a conical space. [5] [2]

When the string parameter ν is equal to an integer $p = 1, 2, 3, \ldots$, we can derive the solution for V, which corresponds to the static configuration of one extreme dilatonic black hole and a cosmic string at the origin, by the method of images [6]. The function $\Gamma_p(\boldsymbol{x}, \boldsymbol{x}')$ satisfying Eq. (7) and having periodicity $2\pi/p$ in ϑ and ϑ' can be expressed by the summation of p images of $\Gamma(\boldsymbol{x}, \boldsymbol{x}')$, which is a solution of Eq. (7) in the Minkowski space:

$$\Gamma_p(\boldsymbol{x}, \boldsymbol{x}') = \sum_{n=0}^{p-1} \Gamma(\boldsymbol{x}, \Lambda^n \boldsymbol{x}') = \sum_{n=0}^{p-1} \frac{1}{|\boldsymbol{x} - \Lambda^n \boldsymbol{x}'|} , \tag{9}$$

where Λ means the rotation by $2\pi/p$ around the origin. Using the integral representation, we then obtain the solution for V as follows:

$$\begin{aligned} \Gamma_p(\boldsymbol{x}, \boldsymbol{x}') &= \sum_{n=0}^{p-1} \frac{1}{r^2 + r'^2 - 2rr' \cos(\vartheta - \vartheta' - 2\pi n/p) + (z - z')^2} \\ &= \frac{1}{\pi\sqrt{2rr'}} \int_{u_0}^{\infty} \frac{du}{\sqrt{\cosh u - \cosh u_0}} \frac{p \sinh pu}{\cosh pu - \cos p(\vartheta - \vartheta')} , \end{aligned} \tag{10}$$

where

$$\cosh u_0 = \frac{r^2 + r'^2 + (z - z')^2}{2rr'} . \tag{11}$$

We have the solution of V which corresponds to asymptotically flat space:

$$V = 1 + 2GM\Gamma_p(\boldsymbol{x}, \boldsymbol{x}') , \tag{12}$$

where M is a constant. Note that V is a solution of Eq. (7) even if p is replaced by an arbitrary real number ν. Consequently, we have the solution of V corresponding to an extreme dilatonic black hole located at $\boldsymbol{x}' = (r', \theta', z')$ in the conical space:

$$V = V_{CS} = 1 + \frac{2GM}{\pi\sqrt{2rr'}} \int_{u_0}^{\infty} \frac{du}{\sqrt{\cosh u - \cosh u_0}} \frac{\nu \sinh \nu u}{\cosh \nu u - \cos(\theta - \theta')} . \tag{13}$$

---

[2] The cosmic string-extreme black hole solution in a general Einstein-Maxwell-dilaton model with an arbitrary dilaton coupling [4] can be constructed in the same manner.

This result can be extracted by using the mode expansion in terms of the Legendre functions. V_{CS} can be equivalently expressed as the sum of the Legendre functions:

$$V_{CS} = 1 + \frac{2GM\nu}{\pi\sqrt{rr'}} \sum_{n=-\infty}^{\infty} Q_{\nu|n|-1/2}(\cosh u_0)\, e^{in(\theta-\theta')} . \tag{14}$$

In the limit of $\nu \to 1$, V_{CS} is reduced to V in the flat space [3,4]. On the other hand, in the limit of $r' \to 0$, V_{CS} is reduced to

$$V = 1 + \frac{2GM\nu}{\sqrt{r^2 + (z-z')^2}} , \tag{15}$$

which gives the metric representing a cosmic string trapped by a black hole. Thus the constant M is identified to the black hole mass [7].

III. WHAT IS "LOW-ENERGY INTERACTION"?

Next we will present the low-energy interaction in the cosmic string-extreme black hole system.

The interaction energy of the maximally charged dilatonic black holes in R^4 at low velocities has been calculated without a long-distance approximation by making use of the exact, static solution (6) with (8) [8,9]. Since there are only two-body velocity-dependent forces in the multi-black hole system in this case, the general expression for the interaction energy of $O(v^2)$ of an arbitrary number of extreme dilatonic black holes can be obtained as: [8–10]

$$H = \frac{1}{2}\mathcal{M}V^2 + \sum_{\alpha\beta} \frac{m_\alpha m_\beta |\boldsymbol{v}_\alpha - \boldsymbol{v}_\beta|^2}{4\mathcal{M}} \left(1 + \frac{2G\mathcal{M}}{|\boldsymbol{x}_\alpha - \boldsymbol{x}_\beta|}\right) , \tag{16}$$

where $\boldsymbol{v}_\alpha$ is the velocity of the extreme dilatonic black hole with mass m_α. $\mathcal{M}$ is the total mass, $\mathcal{M} = \ _\alpha m_\alpha$. $\boldsymbol{V}$ is the velocity of the center of mass; $\boldsymbol{V} \equiv \ _\alpha m_\alpha \boldsymbol{v}_\alpha/\mathcal{M}$.

We use the method of images to obtain the interaction energy of extreme dilatonic black holes. We must note that the images of a certain black hole are located at $\Lambda^n\boldsymbol{x}$ and have the same mass M and the velocity $\Lambda^n\boldsymbol{v}$. We also note that the sum of the interaction energy

must be divided by the number of the images. The number p is extended to a real number ν after completing the calculation.

Finally we get the following interaction energy up to $O(v^2)$ of the extreme dilatonic black hole with the mass M and the velocity $\boldsymbol{v}$ on the conical space:[3]

$$H_{CS} = \frac{1}{2}Mv_{\parallel}^2 + \frac{1}{2}M\left[1 + \frac{2GM}{\rho}\tan\frac{\Delta}{4}\right]\boldsymbol{v}_{\perp}^2 , \tag{17}$$

where ρ is the distance between the cosmic string and the extreme black hole. $v_{\parallel}$ is the z component of the velocity of the extreme dilatonic black hole, while $\boldsymbol{v}_{\perp}$ is the component of the velocity perpendicular to the cosmic string. It is quite reasonable that the interaction energy diverges in the limit of $\Delta \to 2\pi$.

IV. SCATTERING OF AN EXTREME DILATONIC BLACK HOLE BY A COSMIC STRING

We consider the scattering of a maximally charged dilatonic black hole by a cosmic string by using the metric of moduli space [11].

We assume that the black holes move in a plane which is perpendicular to the string. Therefore the moduli space of this configuration is reduced to be a two-dimensional space parameterized by the distance ρ, the azimuthal angle φ. For this two-body system, the metric on moduli space which spanned by the coordinate ξ can be read from Eq. (17) as

$$ds_{MS}^2 = \gamma_{ij}d\xi^i d\xi^j = \gamma(\rho)\left(d\rho^2 + \frac{\rho^2}{\nu^2}d\varphi^2\right) , \tag{18}$$

with

$$\gamma(\rho) = 1 + \frac{2GM}{\rho}\tan\frac{\Delta}{4} = 1 + \frac{2\rho_0}{\rho} . \tag{19}$$

---

[3]For a case of a general dilaton coupling, since the interaction contains many-body, velocity-dependent forces [8–10], the expression will be more complicated.

Here φ stands for the azimuthal angle which has a range $0 \leq \varphi < 2\pi$.

The path of the moving extreme black hole is determined by the geodesic equation on the moduli space [11], because the geodesic on the metric (18, 19) realizes the path of minimal energy. We find that the scattering trajectory satisfies the following differential equation:

$$\left(\frac{du}{d\varphi}\right)^2 + \nu^2 u^2 = \frac{1}{b^2}\left(1 + 2\rho_0 u\right) , \tag{20}$$

where $u = 1/\rho$ and b is the impact parameter.

The scattering angle Θ can be obtained by solving the equation (20) and expressed as:

$$\frac{\Theta}{\nu} = \frac{\Delta}{2} + 2\arctan\left(\frac{\rho_0 \nu}{b}\right) . \tag{21}$$

The angle of deflection is written as the sum of the contributions from the deficit angle [2] and the Rutherford scattering [9].

The quantum mechanical approach to the study of the scattering process is also possible by using the moduli space metric [12,9]. The Schrödinger equation on the moduli space, which is spanned by the coordinates (ρ, φ, Z), can be read as [12,9]

$$i\hbar\frac{\partial}{\partial t}\Psi = -\frac{\hbar^2}{2M}\left(\frac{1}{\sqrt{\det\gamma_{ij}}}\partial_k\sqrt{\det\gamma_{ij}}\gamma^{k\ell}\partial_\ell + \frac{\partial^2}{\partial Z^2}\right)\Psi . \tag{22}$$

When we assume the wave function Ψ can be decomposed as

$$\Psi = \psi(\rho, \varphi)\, e^{-i\frac{E}{\hbar}t} e^{ikZ} , \tag{23}$$

we get the following wave equation:

$$\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho} + \frac{\nu^2}{\rho^2}\frac{\partial^2}{\partial\varphi^2} + q^2\left(1 + \frac{2\rho_0}{\rho}\right)\right]\psi = 0 , \tag{24}$$

where $q^2 \equiv 2ME/\hbar^2 - k^2$. The equation (24) is the same as Eq. (4.2) in the paper on the Coulomb problem on a cone by Gibbons, Ruiz and Vachaspati [13], if we replace $\rho \to r$, $\varphi \to \nu\phi$, $q \to k$ and $q^2\rho_0 \to \mu K$. Thus we do not repeat the calculations on the scattering problem here. One can find that the differential cross-section diverges in the forward direction, but the divergence is shifted due to the presence of deficit angle of the spacetime [13]. This nature agrees with that of the classical scattering.

[1] K. Shiraishi, "*Extreme Dilatonic Black Holes on a Torus*", gr-qc/9511005.

[2] A. Vilenkin and E. P. S. Shellard, *Cosmic Strings and other Topological Defects* (Cambridge University Press, 1994).

[3] G. W. Gibbons and K. Maeda, Nucl. Phys. **B298**, 741 (1988). D. Garfinkle, G. Horowitz and A. Strominger, Phys. Rev. **D43**, 3140 (1991); (E) **D45**, 3888 (1992).

[4] K. Shiraishi, J. Math. Phys. **34**, 1480 (1993).

[5] B. Linet, Classical and Quantum Gravity **4**, L 33-L35(1987).

[6] A. G. Smith, in: *The Formation and Evolution of Cosmic Strings*, G. W. Gibbons, S. W. Hawking and T. Vachaspati (eds.), Cambridge Univ, Press, 1990.

[7] M. Aryal, L. H. Ford and A. Vilenkin, Phys. Rev. **D34**, 2263 (1986). D. V. Gal'tsov, Fortschr. Phys. **38**, 945 (1990). D. V. Gal'tsov and E. Masár, Class. Quantum. Grav. **6**, 1313 (1989).

[8] K. Shiraishi, Nucl. Phys. **B402**, 399 (1993).

[9] K. Shiraishi, Int. J. Mod. Phys. **D2**, 59 (1993).

[10] K. Shiraishi, "*Many-body systems in Einstein-Maxwell-Dilaton theory*", gr-qc/9507029.

[11] N. Manton, Phys. Lett. **B110**, 54 (1982); *ibid.* **B154**, 397 (1985). R. S. Ward, Phys. Lett. **B158**, 424 (1985). M. F. Atiyah and N. J. Hitchin, *The Geometry and Dynamics of Magnetic Monopoles* (Princeton University Press, Princeton, 1988).

[12] J. Traschen and R. Ferrell, Phys. Rev. **D45**, 2628 (1992).

[13] G. W. Gibbons, F. R. Ruiz and T. Vachaspati, Commun. Math. Phys. **127**, 295 (1990).

BREMSSTRAHLUNG INDUCED BY THE COSMIC EXPANSION

Masahiro HOTTA
Department of Physics, Tohoku University, Sendai 980-77, JAPAN
in collaboration with T.Futamase, H.Inoue and M.Yamaguchi

Abstract

We shall argue that the cosmic expansion may enhance intereaction rates of elementary particles enormously. For example, 1/2 helicity component of gravitino decays with decay rate $\Gamma \geq m_{3/2}H(t = \textit{production time})/M_{pl}$.

1. Introduction

Particles in the expanding Universe are decelerated and lose their physical momentum. Thus it is surely expected that such particles can emanate radition or massless particles due to the bremsstrahlung process by the external gravitational field. We show that such geometric bremsstrahlung process works effectively especially for gravitino decay in the early Universe.

This paper is based on a work[1] and more detailed analysis can be seen in it.

2. Transition Probability in the Expanding Universe

Spacetime we consider is spacially flat Robertson-Walker Universe with Minkowskian in- and out- regions. The scale factor $a(t)$ is arbitrary function of time t except that $a(t \sim \infty) = 1$ and $a(t = -\infty) = b < 1$, where $b = a_i/a_f$.

Let us discuss a decay process of gravitino Ψ_μ with conformal momentum $\vec{p}$ into photon A_μ with $\vec{k}$ and photino λ with $\vec{q}$. The invariant amplitude is

$$Amp = -i \int d^4x \sqrt{-g} \frac{1}{8iM_{pl}} (\partial_\alpha A^*_\beta(\vec{k},x) - \partial_\beta A^*_\alpha(\vec{k},x)) \bar{\lambda}(\vec{q},x) \gamma^\mu [\gamma^\alpha, \gamma^\beta] \Psi_\mu(\vec{p},x)$$

where M_{pl} is the Planck mass. From this we get transition probability W such that

$$W = \frac{(2\pi)^3}{V} \int d^3q d^3k |Amp|^2, \tag{1}$$

where V is conformal volume of the space.

In the early Universe momentum of particles is much larger than the Hubble parameter and their masses. In such situation the WKB amplitudes are available and partly analytic calculations can be performed. For forward emission with $\theta \sim m/p$ we get the following useful estimate for W. For 1/2 helicity components of gravitino,

$$W\left(|\vec{p}| \gg m, 0 \leq \theta \leq \frac{m}{p}\right) \sim O\left(\left[\frac{m}{M_G}\right]^2\right) \frac{1}{b^4} \qquad (b \sim 0).$$

Notably this estimate does not depend on details of the function $a(t)$.

3. Conclusion

The decay occurs when $W(b) \sim 1$ and the radiation dominant Universe expands like

$$\frac{a_f}{a_i} = \frac{1}{b} = \left(\frac{\tau_f}{\tau_i}\right)^{\frac{1}{2}}, \tag{2}$$

where $\tau_i(\tau_f)$ is initial(final) proper time in the comoving frame. Therefore we get

$$\Gamma_{gb} = \frac{1}{\tau_f} \sim \frac{m}{M_G} H_i, \tag{3}$$

where $H_i = 1/(2\tau_i)$ is the Hubble parameter at the production time of the gravitino. On the other hand ordinary nongravitational decay process gives us its rate,

$$\Gamma_o \sim \left(\frac{m}{M_G}\right)^2 m. \tag{4}$$

Taking the mass $m \sim 1 - 0.1$ TeV, the lifetime $1/\Gamma_o$ is estimated as $10^5 - 10^8$ sec, while the geometrobremsstrahlung process gives much shorter lifetime $1/\Gamma_{gb} \leq 10^{-19}$ sec for the production temperature of the gravitino higher than 10^{15}GeV. Thus the gravitinos with 1/2 helicity decay during the epoch when they are relativistic, as opposed to arguments so far which insist that they decay after becoming dust matter.

4. References

1. T.Futamase, M.Hotta, H.Inoue and M.Yamaguchi,1995,
Geometrobremsstrahlung in the Early Universe, gr-qc/9508053,TU-490

On Perturbations in a Self-Similar Spacetime

Kenji TOMITA

Yukawa Institute for Theoretical Physics,

Kyoto University, Kyoto 606-01

E-mail : tomita@yukawa.kyoto-u.ac.jp

Abstract

Gravitational instability is studied in a self-similar spacetime, which appears in one of three regions in the super-horizon-scale inhomogeneous cosmological model (Tomita 1995). In this region the spacetime is spherically symmetric but locally anisotropic, so that three modes of perturbations are coupled, and their behaviors of these perturbations are similar to those in anisotropic homogeneous cosmological models.

1. Introduction

If we assume the original Einstein theory of gravitation without cosmological constant and the existence of inflation at the very early stage of our Universe, it is natural that its present mean status is represented by the Einstein-de Sitter model which is spatially flat ($\Omega_0 = 1.0$). The comparison of the observable CMB anisotropy with the theoretical CDM models supports the Einstein-de Sitter model as the background model (Sugiyama and Silk 1994; Bunn and Sugiyama 1995). On the other hand, it has been shown by the observational studies about the spatial distributions of galaxies and clusters (Bahcall and Cen 1993) that the Universe is locally represented well by the low-density CDM model ($\Omega_0 \approx 0.2$). This contradictory situation is not compatible with the homogeneous cosmological models which have only weak perturbations on large-scales.

To avoid this dilemma, a spherically-symmetric inhomogeneous model with nonlinear struc-

ture was recently proposed (Tomita 1995), in which the Universe consists of three regions — the inner low-density homogeneous region, the self-similar region, and the outer nearly-flat homogeneous region. The boundaries between first two regions and second two regions are specified by the redshifts z_1 ($= 1.5 \sim 2.0$) and z_2 (≥ 5.0), respectively. Since matter is assumed to be pressureless, the model is described by the Tolman-Bondi solution with the line-element

$$ds^2 = -c^2dt^2 + g_{rr}(t,r)dr^2 + S^2(t,r)r^2d\Omega^2. \tag{1.1}$$

In the self-similar region ($r_1 \leq r \leq r_2$), the scale-factor S depends only on $\xi \equiv ct/r$ and $g_{rr} = S^2(\xi)(1+\xi S'/S)^2/(1+\alpha_0 r_1{}^2)$, where α_0 is a constant, and (r_1, r_2) corresponds to (z_1, z_2). Using $x \equiv \ln r$, we can rewrite the line-element as

$$ds^2 = e^{2x}d\bar{s}^2, \tag{1.2}$$

$$d\bar{s}^2 = -d\xi^2 - 2\xi d\xi + M^2(\xi)dx^2 + S^2(\xi)d\Omega^2, \tag{1.3}$$

where

$$M^2 = -\xi^2 + g_{rr}(\xi). \tag{1.4}$$

These relations mean that this spacetime is conformal to a homogeneous spacetime $d\bar{s}^2$, which has a translational invariance $x \rightarrow x + const.$

The implications of the above inhomogeneous model in the problems of bulk motions and quasar lensing statistics were discussed in recent two papers (Tomita 1966a, b).

2. Gauge-invariant perturbations in general spherically-symmetric spacetimes

The perturbations in general spherically-symmetric spacetimes were treated by Gerlach and Sengupta (1979), extending the Regge-Wheeler formalism for the Schwarzschild spacetime.

When we put the background line-element and energy-momentum tensor in the forms:

$$ds^2 = g_{AB}dx^Adx^B + R^2(x^C)d\Omega^2, \tag{2.1}$$

$$t_{\mu\nu}dx^\mu dx^\nu = t_{AB}dx^Adx^B \tag{2.2}$$

with suffices $A, B = 0, 1$ and $a, b = 2, 3$, their perturbations in the forms $h_{\mu\nu}dx^\mu dx^\nu$ and $\Delta t_{\mu\nu}dx^\mu dx^\nu$ are classified into two types as follows according to their parities in the θ, ϕ transformation:

2.1. Odd parity

$$h_{\mu\nu}dx^\mu dx^\nu = 2h_A(x^C)S_{,a}(\theta,\phi)dx^A dx^a + 2h(x^C)S_{;a;b}dx^a dx^b, \tag{2.3}$$

$$\Delta t_{\mu\nu}dx^\mu dx^\nu = 2\Delta t_A(x^C)S_a(\theta,\phi)dx^A dx^a. \tag{2.4}$$

The corresponding gauge-invariant quantities are

$$k_A = h_A - R^2(h/R^2)_{,A}, \text{ and } L_A = \Delta t_A. \tag{2.5}$$

2.2. Even parity

$$\begin{aligned} h_{\mu\nu}dx^\mu dx^\nu &= h_{AB}(x^C)Y(\theta,\phi)dx^A dx^B + 2h_A Y_{,a}dx^A dx^a \\ &\quad + R^2[K(x^C)Y\gamma_{ab} + G(x^C)Y_{;a;b}]dx^a dx^b, \end{aligned} \tag{2.6}$$

$$\Delta t_{\mu\nu}dx^\mu dx^\nu = \Delta t_{AB}(x^C)Y dx^A dx^B + 2\Delta t_A\ Y_a dx^A dx^a. \tag{2.7}$$

The corresponding gauge-invariant quantities are

$$k_{AB} = h_{AB} - (p_{A|B} + p_{B|A}) \text{ and } k = K - 2v^A p_A, \tag{2.8}$$

$$T_{AB} = \Delta t_{AB} - t_{AB}{}^{|C}\ p_C - t_{CA}\ p^C_{|B} - t_{CB}\ p^C_{|A} \text{ and } T_A = \Delta t_A - t^C_A\ p_C, \tag{2.9}$$

where $p_A \equiv h_A - \frac{1}{2}R^2 G_{,A}$ and $v_A \equiv R_{,A}/R$.

3. Self-similar spacetime

Because of the conformality and the flatness in the x direction of the conformal space, the perturbed quantities are expanded using the factor e^{2x} and the Fourier component e^{ikx}, where k is a wave number.

3.1. Odd parity

The metric components k_A is expanded as

$$k_A = \mathrm{e}^{2x}\mathrm{e}^{ikx}\bar{k}_A(\xi), \tag{3.1}$$

and the matter components are $L = L^1 = 0$ and

$$L^0 = \mathrm{e}^{-2x}\bar{L}^0(\xi). \tag{3.2}$$

In this case the perturbations include rotational motions (vector perturbations) and gravitational waves (tensor perturbations), and L^0 represents the momentum of the rotational motion. The dynamical equations reduce to the following ordinary differential equations with respect to ξ:

$$\left(S^2N\bar{L}^0\right)^{\cdot} = 0, \tag{3.3}$$

$$\bar{k}^1 = -\frac{(N\bar{k}^0)^{\cdot}}{(2+ik)N}, \tag{3.4}$$

$$\begin{aligned}\left(\frac{\bar{k}^0}{S^2}\right)^{\cdot\cdot} = & -\left[2\frac{\dot{S}}{S} - \frac{\dot{N}}{N} + 2\frac{\dot{M}}{M} + \frac{2(1+ik)\xi}{M^2}\right]\left(\frac{\bar{k}^0}{S^2}\right)^{\cdot} \\ & -\left[\left(2\frac{\dot{S}}{S} - \frac{\dot{N}}{N}\right)^{\cdot} + \frac{2+k^2-ik}{M^2} + \left(\frac{ik\xi}{M^2} + \frac{2\dot{M}}{M}\right)\left(2\frac{\dot{S}}{S} - \frac{\dot{N}}{N}\right)\right. \\ & \left. +(l-1)(l+2)\left(\frac{N}{MS}\right)^2\right]\left(\frac{\bar{k}^0}{S^2}\right) + \left(\frac{N}{MS}\right)^2\kappa\bar{L}^0,\end{aligned} \tag{3.5}$$

where a dot denotes the derivative with respect to ξ, $N(\xi) \equiv (M^2(\xi)+\xi^2)^{1/2} = S(\xi)(1+\xi S'/S)/(1+\alpha_0 {r_1}^2)$, and $\kappa \equiv 16\pi G/c^4$. Eq.(3.3) can be integrated as

$$\bar{L}^0 = (\bar{L}^0)_i/(S^2N), \tag{3.6}$$

where $(\bar{L}^0)_i$ is an integration constant. If this momentum vanishes, we have only free gravitational waves. In the non-vanishing case there are coupled perturbations of rotational motions and gravitational waves. This situation is similar to that in the anisotropic and homogeneous cosmological model of Bianch type I, which was studied by Perko et al.(1972) in the synchronous coordinate condition and by Tomita and Den (1986) and Den (1988) in the gauge-invariant formalism.

3.2. Even parity

In this case we have density (scalar) perturbations, rotational (vector) perturbations, and gravitational-wave (tensor) perturbations. They are coupled in the same way as in the odd parity. Dynamical variables are expanded as

$$k = e^{2x}e^{ikx}\bar{k}(\xi), \quad k_{AB} = e^{2x}e^{ikx}\bar{k}_{AB}(\xi) \tag{3.7}$$

and the components of energy-momentum tensor are related to the density perturbation Δ and velocities V, U as

$$\bar{T}^{00} = \bar{\rho}\bar{\Delta}, \quad \bar{T}^{01} = \bar{\rho}\bar{V}, \quad \bar{T}^{0} = \bar{\rho}\bar{U}, \quad \bar{T}^{11} = \bar{T}^{1} = 0 \tag{3.8}$$

for the pressureless matter, where ρ is the unperturbed density, V and U are the velocities in the radial and transverse directions, respectively. The fluid-dynamical equations for $\bar{\Delta}, \bar{V}$ and $\bar{U}$ and the equations of gravitational potentials $\bar{k}$ and $\bar{k}_{AB}$ are derived from the Einstein equations. Their lengthy expressions are omitted here, but can be easily solvable because all of them are ordinary differential equations with respect to ξ.

4. Concluding remarks

There remain following several problems to be done: 1. In the above treatment the metric perturbations include two or three mixed modes. It is necessary to decompose them in an invariant manner, 2. Next we should solve the above equations and adjust the solutions with the solutions in other homogeneous regions, and 3. we should study the Sachs-Wolfe effect in the self-similar region to analyze the CMB anisotropy in the present cosmological model.

The observational validity of the present cosmological model will be clarified through various optical high-redshift observations by the digital sky survey, HST, large telescopes and so on in the future.

References

Bahcall, N.A. and Cen, R. 1993, APJ, 407, L49

Bunn, E. F. and Sugiyama, N. 1995, APJ, 446, 49

Den, M. 1988, Prog. Theor. Phys. 79, 1110

Gerlach, U.H. and Sengupta, U.K. 1979, Phys. Rev. D19, 2268

Perko, T.E., Matzner, R.A., and Shepley, L.C. 1972, Phys. Rev. D6, 969

Regge, T., and Wheeler, J.A. 1957, Phys. Rev. 108, 1063

Sugiyama, N., and Silk, J. 1994, PRL, 73, 509

Tomita, K. and M. Den, 1986, Phys. Rev. D34, 3570

Tomita, K. 1995, APJ 451, 1

Tomita, K. 1996a, APJ 461, No.2

Tomita, K. 1996b, PASJ 48, No. 2

Gravitational Collapse in Szekeres Universes

Hiraku Mutoh, Toshinari Hirai and Kei-ichi Maeda

Department of physics, Waseda University, Shinjuku-ku, Tokyo 169, Japan

Abstract

We study the stability of attractor spacetimes of *silent universes*, in which the magnetic part of the Weyl tensor H_{ab} is assumed to vanish. We find that a spindle attractor is unstable against the perturbations of H_{ab}, while a pancake attractor is stable. We conclude that the $H_{ab} = 0$ ansatz is not valid in generic dust gravitational collapse.

1 Introduction

The recent observations of the large scale structure in the universe show that quite nonlinear wall-like structures or filament-like structures seem to exist. Such structures may give us important information about the structure formation process in the universe. It may be important to study nonlinear dynamics in order to know what kind of structure is formed on large scales.

In Newtonian gravity, the Zel'dovich approximation [1], which describes the evolution of inhomogeneities well even into the weakly nonlinear regime, predicts a pancake collapse. The analysis of a homogeneous ellipsoid in an expanding universe also supports such a pancake collapse [2].

In general relativity, the analysis of nonlinear dynamics is quite difficult because the Einstein equations are nonlinear partial differential equations. It has been restricted to very simple cases such as the spherically symmetric Tolman-Bondi model ([3], [4]). Recently the nonlinear dynamics of irrotational dust with vanishing magnetic part of the Weyl tensor ($H_{ab} = 0$) has been examined ([5]–[10]). Under this condition, these systems are described by nonlinear ordinary differential equations, and each fluid element evolves independently. Such systems have no freedom of gravitational waves, and no information transfer. Therefore these models are called *silent universes*. Bruni *et al.* [9] found the attractor solutions in these systems and showed that the generic gravitational collapse leads to a triaxial spindle-like configuration.

However the assumption $H_{ab} = 0$ is controversial. It is not necessarily justified for the general cosmological case. Therefore we consider the perturbations of *silent universes*, taking into account the effect of non-vanishing H_{ab}. Analyzing the stability of the attractor solution (Kasner-type spacetime) and some particular Szekeres solution [11], we discuss whether the analysis of the *silent universe* [9] is valid or not and what kind of configuration is more probable in the nonlinear regime of structure formation in general relativity.

We take $c = 8\pi G = 1$, and the signature $(-, +, +, +)$.

2 Silent Universes

First, we give a short review of the *silent universe* [9]. The dynamics of an irrotational dust fluid is described by the following variables: the mass density ρ, the expansion scalar θ, the shear tensor σ_{ab} of the fluid, and the electric and magnetic parts of the Weyl tensor E_{ab}, H_{ab} [12].

The *silent universe* is obtained by assuming, (i) the fluid is collisionless dust ($p = 0$), (ii) it has no vorticity ($\omega_{ab} = 0$), and (iii) the magnetic part of the Weyl tensor vanishes ($H_{ab} = 0$). The traceless σ_{ab} and E_{ab} are diagonalized under this condition as shown by [5]. Their independent components are then σ_{11}, σ_{22} and E_{11} , E_{22}. The metric is written as

$$ds^2 = -dt^2 + \sum_{\alpha=1}^{3} l_\alpha^2(\boldsymbol{x}, t)(dx^\alpha)^2, \tag{1}$$

and,

$$\frac{\dot{l}_\alpha}{l_\alpha} = \sigma_{\alpha\alpha} + \frac{1}{3}\theta, \tag{2}$$

where the l_α's represent the scale factors in the α-direction ($\alpha = 1, 2, 3$).

Next, we define dimensionless variables, Ω, Σ_{ab}, and E_{ab}, ($\Sigma_\pm$, $E_\pm$) as,

$$\rho = \frac{1}{3}\Omega\theta^2, \ \sigma_{ab} = \Sigma_{ab}\theta, \ E_{ab} = \varepsilon_{ab}\theta^2, \tag{3}$$

and,

$$\Sigma_\pm = \frac{1}{2}(\Sigma_{11} \pm \Sigma_{22}), \ \varepsilon_\pm = \frac{1}{2}(\varepsilon_{11} \pm \varepsilon_{22}). \tag{4}$$

Then the evolution equations of *silent universes* (for collapse) are described by the following nonlinear ordinary differential equations.

$$\theta' = \theta\left[\frac{1}{3} + 6\Sigma_+^2 + 2\Sigma_-^2 + \frac{1}{6}\Omega\right], \tag{5}$$

$$\Omega' = -\frac{1}{3}\Omega\left[36\Sigma_+^2 - 1 + 12\Sigma_-^2 + \Omega\right], \tag{6}$$

$$\Sigma_+' = \Sigma_+\left[\frac{1}{3} - \Sigma_+(1+6\Sigma_+) - 2\Sigma_-^2 - \frac{1}{6}\Omega\right] + \frac{1}{3}\Sigma_-^2 + \varepsilon_+, \tag{7}$$

$$\Sigma_-' = \Sigma_-\left[\frac{1}{3} - 2\Sigma_+(3\Sigma_+ - 1) - 2\Sigma_-^2 - \frac{1}{6}\Omega\right] + \varepsilon_-, \tag{8}$$

$$\varepsilon_+' = \varepsilon_+\left[\frac{1}{3} - 3\Sigma_+(4\Sigma_+ - 1) - 4\Sigma_-^2 - \frac{1}{3}\Omega\right] - \Sigma_-\varepsilon_- + \frac{1}{6}\Sigma_+\Omega, \tag{9}$$

$$\varepsilon_-' = \varepsilon_-\left[\frac{1}{3} - 3\Sigma_+(4\Sigma_+ + 1) - 4\Sigma_-^2 - \frac{1}{3}\Omega\right] - 3\Sigma_-\varepsilon_+ + \frac{1}{6}\Sigma_-\Omega, \tag{10}$$

where prime ($'$) denotes the derivative with respect to τ, ($\tau = -\int \theta dt$). Because of the absence of spatial derivatives in the evolution equations, each fluid element evolves without influence of the environment. Thus, we call such a system a *silent universe*. It includes not only homogeneous Bianchi I spacetimes, but also inhomogeneous spacetimes such as spherically symmetric Tolman-Bondi spacetimes and Szekeres solutions.

We see the evolution equation for θ is decoupled from the others. Then the rest of the equations are 5 first-order differential equations. Bruni *et al.* [9] analyze the behavior of the solutions and find attractor solutions in 5-dimensional "phase" space $\{\Omega, \Sigma_+, \Sigma_-, E_+, E_-\}$.

The attractors in the collapsing case ($\theta < 0$), correspond to a family of vacuum Kasner spacetimes locally. The attractor family is the circle in Fig. 1; only three points of the attractors correspond to pancake collapse (two of the l_a stay finite but the other one vanishes), and the rest of the attractors correspond to spindle collapse (two of l_a's vanish but the other one diverges). Bruni *et al.* [9] have concluded that collapsing regions generally tend to be spindle-like configurations in *silent universes*.

3 Perturbations of Silent Universes

In order to generalize the analysis by Bruni *et al.* [9], we consider linear perturbations of a *silent universe* with non-zero H_{ab}. We study two cases, the background solutions are (i) the attractor solutions of collapsing *silent universes* (Kasner-type spacetimes), and (ii) inhomogeneous Szekeres solutions which approach the particular attractor solutions in the former case. Here the background quantities are represented with bar ($\bar{}$), and the perturbed ones with tilde ($\tilde{}$).

3.1 Perturbations of the Attractor Solutions

The attractor solutions of *silent universes* are homogeneous vacuum Kasner-type spacetimes, which are characterized by the value of Σ_+ ($-1/3 \leq \Sigma_+ \leq +1/3$), as

$$\Omega = 0, \tag{11}$$

$$\Sigma_- = \frac{1}{\sqrt{3}}\sqrt{1-9\Sigma_+^2}, \tag{12}$$

$$\varepsilon_+ = \frac{1}{3}\Sigma_+(6\Sigma_+ + 1) - \frac{1}{9}, \tag{13}$$

$$\varepsilon_- = -\frac{\sqrt{3}}{9}(6\Sigma_+ - 1)\sqrt{1-9\Sigma_+^2}. \tag{14}$$

We analyze linear perturbations of this solution with non-zero H_{ab}. Applying the Fourier expansion, we consider only a single plane wave mode with a wave number $\boldsymbol{k} = (k, 0, 0)$. The spatial gradient is replaced with,

$$\tilde{\Omega}_{,1} = ik\tilde{\Omega}. \tag{15}$$

Then we integrate the perturbation equations for various background quantities (Σ_+) and the scale of the perturbations (k^{-1}).

The characteristic behavior of the perturbations is shown in Fig. 2. The evolution of perturbations is quite different depending on whether the background is spindle-like or pancake-like. In the case of spindle-like background, the perturbations of the magnetic part grow larger and larger, i.e., the *silence* is broken. Other variables such as $\tilde{\Sigma}_\pm$, $\tilde{\varepsilon}_\pm$ also diverge. This behavior is qualitatively independent of the wave number and background parameters. On the other hand, in the case of pancake-like background, the perturbations do not grow for any wave number or any wave direction.

Therefore we conclude that for generic perturbation with H_{ab}, the pancake-like attractor is still stable, but the spindle-like attractor becomes unstable.

3.2 Perturbations of the Szekeres Solutions

In the above analysis, since the background spacetime is the homogeneous attractor solution, our result might not be generic. It is possible that the evolution of the perturbations in inhomogeneous backgrounds is different from that in the homogeneous background. Therefore we consider the following particular Szekeres solution ([11], [13]) as an inhomogeneous background *silent universe*:

$$ds^2 = -dt^2 + (t-t_*)^{4/3}\left\{dx^2 + dy^2 + \left[1-(t-t_*)^{-1}\beta_-(z)\right]^2 dz^2\right\} \tag{16}$$

where $\beta_-(z)$ represents an inhomogeneity in the z-direction and t_* is a constant. When $\beta_-(z)$=0, the spacetime is the flat Friedmann-Robertson-Walker (FRW) dust universe.

For these solutions, the dimensionless variables defined above are represented as,

$$\Omega = 1 - 9\Sigma_+^2, \tag{17}$$

$$\Sigma_- = 0, \tag{18}$$

$$\varepsilon_- = 0, \tag{19}$$

$$\varepsilon_+ = \Sigma_+(\Sigma_+ + \frac{1}{3}), \tag{20}$$

where,

$$\Sigma_+ = \frac{\beta_-}{3\,[\beta_- - 2(t - t_*)]}. \tag{21}$$

These Szekeres solutions, characterized by $\beta_-(z)$ and t_*, are not stationary points in the "phase" space (see Fig. 1). In the collapsing case, even if the Szekeres universe is located initially near the isotropic FRW point ($\Sigma_+ = 0$), according to the sign of Σ_+, the spacetime falls into a pancake-like attractor ($\Sigma_+ < 0$), or falls into a spindle-like attractor ($\Sigma_+ > 0$).

Since the z-dependent inhomogeneities are included in the background exact Szekeres solutions, we focus only on x, y-dependence of the perturbations. Hence, we expand x, y-dependent perturbations on a $t = const$, $z = const$ surface by plane waves in the same way as in the case of the Kasner-type background.

The results are shown in Fig. 3. For the spindle-like background ($\bar{\Sigma}_+ > 0$), the perturbation of the magnetic part of the Weyl tensor grows infinitely and the *silence* is broken, and the electric part and its shear also grow. But for the pancake-like background, the perturbations do not grow. Such characteristic behavior depends only on the collapsing configurations (the sign of $\bar{\Sigma}_+$ of the background), but does not depend on the other factors such as the spatial scale of perturbations.

For these Szekeres backgrounds, the behavior of the perturbations is very similar to the Kasner-type background case. Our analysis shows that the pancake-like collapse is stable while the spindle-like collapse is unstable. One can see that isotropic collapse (FRW) is also unstable against a perturbation with magnetic part H_{ab} (what is called the tensor mode).

4 Conclusion

We have studied the perturbations of a *silent universe* with non-zero H_{ab}. The analysis of the *silent universe* [9] shows the gravitational collapse tends to form spindle-like configurations. However, considering the more realistic case, the ansatz of $H_{ab} = 0$ may be too strong, then one should take into account the contribution of H_{ab}. We have considered two cases: the background solutions are (i) the attractors of the *silent universe*, and (ii) particular Szekeres solutions. In both cases the behavior of the evolution of perturbations are qualitatively the same. It is interesting that the stability clearly depends on whether the configurations of the background spacetime are pancake-like or spindle-like. In the case of a spindle-like background, whose perturbations grow, the *silence* is broken and the spacetime approaches the most generic homogeneous spacetime, that is, the locally Bianchi IX universe. On the other hand, the perturbations do not grow in the case of a pancake-like background. Neglecting H_{ab} is not justified in generic gravitational collapse.

In general relativity, taking into account the effect of H_{ab}, the spindle collapse is destabilized. This meets with the fact that the Newtonian collapse suggests a pancake-like configuration. The magnetic part of the Weyl tensor has influence on structure formation, and its effect cannot be ignored.

References

[1] Zel'dovich, Ya. B., 1970, A & A, **5**, 84.

[2] White, S. D. M., & Silk, J., 1979, ApJ, **231**, 1.

[3] Tolman, R. C., 1934, Proc. Natl. Acad. Sci., **20**, 169.

[4] Bondi, H., 1947, MNRAS, **107**, 410.

[5] Barnes, A., & Rowlingson, R. R., 1989, Class. Quantum Gravity, **6**, 949.

[6] Matarrese, S., Pantano, O., & Saez, D., 1993, Phys. Rev. D, **47**, 1311.

[7] Croudace, K. M., Parry, J., Salopek, D. S., & Stewart, J. M., 1994, ApJ, **431**, 22.

[8] Bertschinger, E., & Jain, B., 1994, ApJ, **431**, 486.

[9] Bruni, M., Matarrese, S., & Pantano, O., 1995, ApJ, **445**, 958.

[10] Bruni, M., Matarrese, S., & Pantano, O., 1995, Phys. Rev. Lett., **74**, 1916.

[11] Szekeres, P., 1975, Comm. Math. Phys., **41**, 55.

[12] Ellis, G. F. R. 1971, in General Relativity and Cosmology, ed. Sachs, R. K. (Academic, London).

[13] Goode, S. W., & Wainwright, J., 1982, Phys. Rev. D, **26**, 3315.

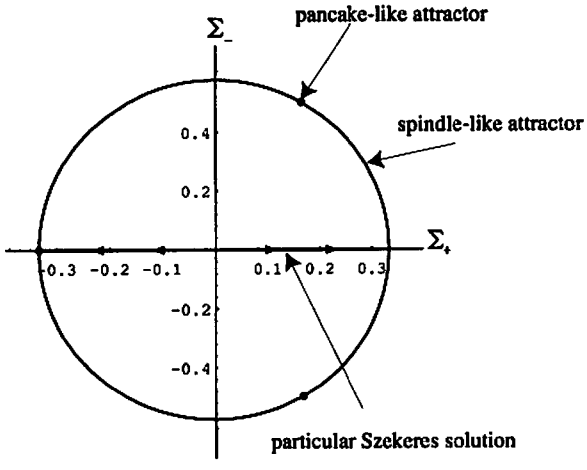


Fig. 1: The attractor family of *silent universes* (in collapse) in the $\{\Sigma_+, \Sigma_-\}$ plane; dots represent the pancake-like attractors. The line $\Sigma_- = 0$, $-1/3 < \Sigma_+ < 1/3$ shows the particular Szekeres solutions we use as the background solutions in §3.2.

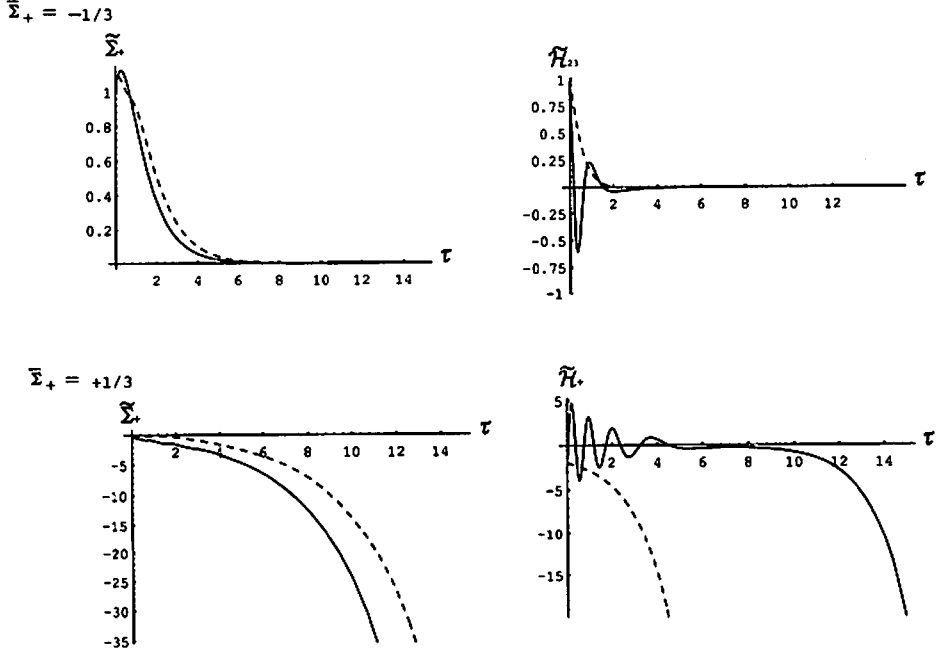


Fig. 2: The evolution of perturbations of the attractor solutions of *silent universes*. The upper panels are in the case of pancake background, the lower panels are in the case of spindle background. $\tilde{\mathcal{H}}_+ = (H_{11} + H_{22})/2\theta^2$, $\tilde{\mathcal{H}}_{23} = H_{23}/\theta^2$. The solid lines show the case of initial dimensionless wave number $k/l_1\theta = 10$, the dashed lines show the case of $k/l_1\theta = 1$. The initial value of the perturbations are taken to be of order unity.

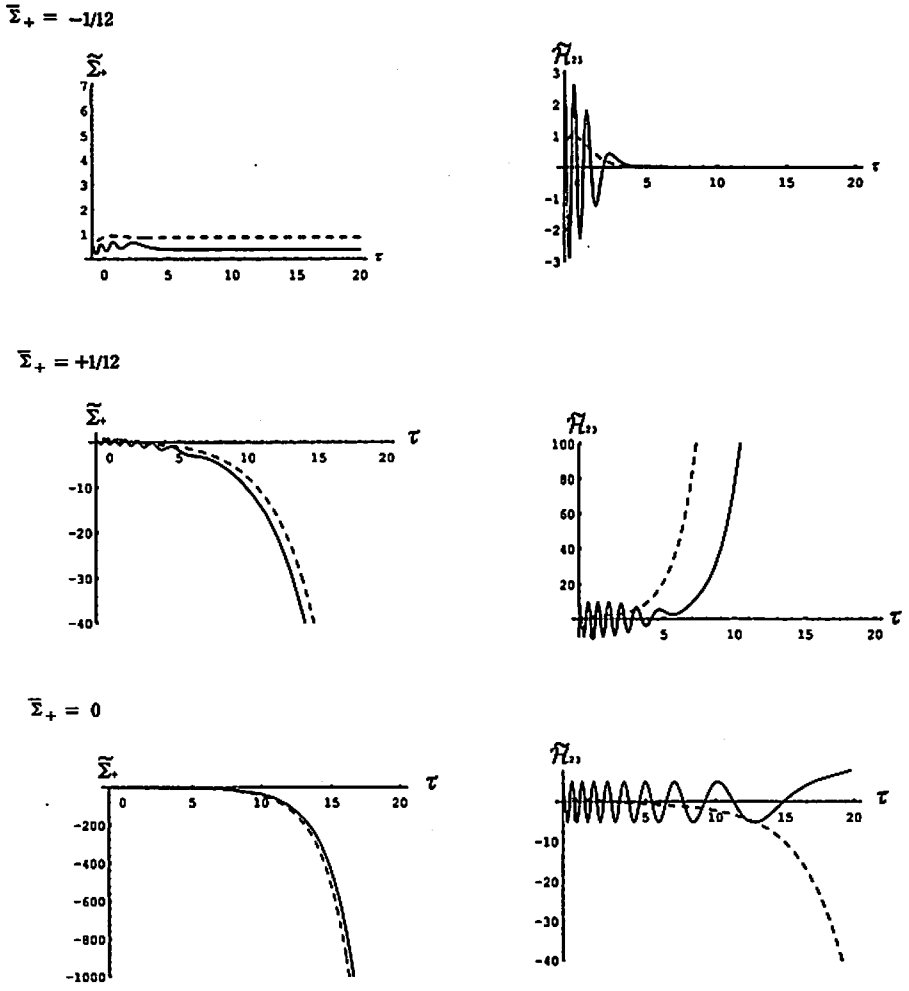


Fig. 3: The evolution of perturbations of the particular Szekeres solutions. The top panels are in the case of pancake background, the middle panels are in the case of spindle background, and the bottom are the case of isotropic background. Notation as in Fig. 2

ASCA Observations of Clusters of Galaxies

Koujun Yamashita
Department of Physics, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-01, Japan

X-ray images and spatially-resolved spectra of nearly 150 clusters of galaxies in the cosmological redshift up to 0.5 have been observed in the energy range of 0.5 - 10 keV by Japanese X-ray astronomy satellite ASCA. Those clusters include group of galaxies, poor clusters, nearby rich clusters($z<0.1$) and distant clusters($z>0.1$). Their plasma temperatures and abundances relative to the cosmic value were obtained to be 1 - 10 keV and 0.1-0.5 over whole clusters, respectively. Spatially-resolved spectra make it possible to derive their distribution in an intracluster medium(ICM), from which the merging and cooling flow are directly deduced. Detailed analysis of emission lines from highly ionized atoms allows to determine elemental abundance and ionization temperature. These results give a great insight into the origin of ICM, the formation and evolution of clusters. Some of S-Z effect and gravitational lensing clusters were also observed to investigate cosmological problems.

1. Introduction

Clusters of galaxies are the largest gravitationally-bound ensemble in the universe, containing hundreds of galaxies and hot intracluster gas. They are identified as isolated objects by X-ray observations. X-ray emissions from clusters are most likely originated from a thin hot plasma in a collisional ionization equilibrium. The optical depth of continuum component is order of 10^{-3}, whereas that of emission lines is around unity. Present emisssion models used for spectral fitting can not estimate this effect, so that the determination of elemental abundances includes large uncertainty. The spatially-resolved spectroscopy with ASCA gives a clue to investigate the physical state of hot intracluster gas and a impact to reconsider the basic atomic processes. This is an important issue to deeply understand the structure, formation and evolution of clusters, and the origin of intracluster gas.

Jones and Forman[1] classified X-ray images observed by Einstein into seven group, called single symmetric, offset center, elliptical, complex multiple structures, double, primary with small secondary and primarily galaxy emission. According to physical processes of cluster formation and evolution, their morphologies are roughly divided into two categories, they are, centrally concentrated and spherically symmetric clusters with small core radii having cD galaxy at the center associated with cooling flows and largely extended clusters with large core radii and merging substructures.

ROSAT all sky survey detected thousands of clusters with the high angular resolution[2]. These results are complimentary for ASCA data, since the angular resolution of ASCA is not good enough for detailed analysis of x-ray images, especially for distant clusters.

David et al. compiled temperatures of 104 clusters ever observed[3]. Ginga observations obtained temperatures and abundances of 45 clusters[4]. The X-ray properties of clusters are represented by the morphological parameters of core radii and radial gradient of gas density distribution, spectral parametes of temperature, abundance and luminosity, and the cosmological redshift. ASCA makes it possible to derive these parameters from the spatially-resolved spectroscopic observations[5].

The Hubble constant(H_0) was derived to be less than 50 km/sec/Mpc by combining x-ray image and spectra with Sunyaev-Zeldovich(hereafter S-Z) effect in microwave region. We have proposed new method for H_0 detemination making use of resonant scattering of emission lines in an intracluster gas

Here we present the observational results of ASCA and discuss the cooling flow and merging of clusters, and the Hubble constant determination.

2. Observations

ASCA has already observed nearly 150 clusters of galaxies in the energy range of 0.5 - 10 keV and in the redshift range up to 0.5, including group of galaxies, poor clusters, nearby rich clusters ($z<0.1$), distant clusters ($z>0.1$) and superclusters. Some of S-Z effect clusters and gravitational lensing clusters were also observed. ASCA puts on board four sets of multi-nested thin foil conical mirror X-ray telescope (XRT) incorporated with two X-ray CCDs(SIS) and imaging gas scintillation proportional counters(GIS) for each two sets[6]. The angular resolution and field of view are 3' (HPD), 22'x22' for SIS and 50' in diameter for GIS, respectively. The energy resolutions are 2% and 8% at 6keV for SIS and GIS, respectively. ASCA instrumentations are described in detail by Serlemitsos et al.[7]. The standard observation time is 40ksec/target corresponding to one day observation. Multiple pointings were carried out for largely extended clusters to completely cover the whole cluster. In table 1 there are listed 48 clusters in the redshift order for each category together with X-ray intensities, plasma temperatures and abundances derived from the spectral analysis.

3. Data analysis

The spectral analysis of observed data can be done by fitting model spectra of thermal emissions from a thin hot plasma in an collisional ionization equilibrium like Raymond-Smith(R-S) or MEKA model, taking into account the absorption of neutral hydrogen in the line of sight. The expected spectra of R-S model folded with the

Table 1 Clusters of galaxies observed with ASCA

Name	Redshift	kT(keV)	Abundance	Ix(c/s) (GIS)
		Group of Galaxies		
NGC1399(For)	0.00445	1.04 +/-0.05	0.45 +/-0.20	
NGC2300	0.0076	0.87 +0.05/-0.04	0.09 +0.06/-0.02	
WP23(NGC5044)	0.0087	1.0 +/-0.05	0.40 +/-0.20	
HCG42	0.0133	0.37 +/-0.04	-	
HCG62	0.0137	0.86 +/-0.01	0.30 +/-0.10	
HCG51	0.0258	1.1 +/-0.1	0.30 +/-0.10	
		Poor Clusters		
AWM7	0.0179	4.0 +/-0.3	0.40 +/-0.05	3.0
MKW4	0.0196	2.0 +/-0.2	0.37 +/-0.08	
3A0336+09	0.0349	3.2 +/-0.3	0.45 +/-0.05	
MKW9	0.03978	2.4 +/-0.2	0.29 +/-0.09	
MKW3s	0.0434	3.5 +/-0.3	0.33 +/-0.03	0.8
Hyd A	0.0522	3.8 +/-0.3	0.40 +/-0.04	0.8
		Nearby Rich Clusters (z<0.1)		
Vir(M87)(r<10')	0.0038	1.73 +/-0.02	0.63 +/-0.08	5.2
Vir(r=10'-20')		2.66 +/-0.05	0.34 +/-0.05	
Cen Cl	0.0107	3.8 +/-0.3	0.30 +/-0.05	3.0
A1060	0.0114	3.3 +/-0.2	0.37 +/-0.04	1.4
A262	0.0163	2.2 +/-0.2	0.32 +/-0.04	
A426(Per)(r<15')	0.0183	5.04 +/-0.05	0.42 +/-0.02	20
A1656(Coma)(r<10')	0.0235	8.8 +/-0.3	0.25 +/-0.05	5.1
Oph Cl(r<6')	0.028	9.8 +0.5/-0.4	0.28 +/-0.04	9.9
A2199	0.0305	4.5 +/-0.5	0.35 +/-0.05	2.0
A2634	0.0312	3.5 +/-0.2	0.30 +/-0.05	
A496	0.0316	4.03 +/-0.06	0.38 +/-0.03	3.0
PKS2354-35	0.046	3.8 +/-0.2	0.43 +/-0.05	
A2319	0.0529	9.1 +/-0.3	0.19 +/-0.04	2.0
A3266	0.0590	8.0 +/-0.4	0.22 +/-0.06	1.36
A2256	0.0601	7.0 +/-0.3	0.22 +/-0.04	1.4
A1795	0.0621	5.5 +/-0.5	0.35 +/-0.05	1.5
A644	0.0704	6.5 +/-0.2	0.30 +/-0.03	0.65
A3921	0.0936	5.7 +/-0.3	0.27 +/-0.07	0.32
A1750 (S-W)	0.094	4.3 +/-0.3	0.30 +/-0.13	
A1750 (N-E)	0.094	3.5 +/-0.2	0.46 +/-0.18	
A2440 (S-W)	0.093	4.2 +/-0.3	0.18 +/-0.12	
A2440 (N-E)	0.097	4.0 +/-0.3	0.39 +/-0.19	
		Distant Clusters (z>0.1) (r<8')		
A1413	0.1427	7.2 +/-0.3	0.26 +/-0.06	0.412
A1204	0.1710	4.0 +/-0.3	0.21 +0.12/-0.11	0.117
A2218	0.1710	6.8 +/-0.5	0.20 +/-0.08	0.191
A586	0.171	6.4 +/-0.6	0.26 +/-0.12	0.196
A1689	0.181	8.9 +/-0.4	0.26 +/-0.05	0.499
A665	0.1816	8.1 +/-0.6	0.24 +/-0.08	0.232
MS0440+0204	0.1900	5.1 +0.7/-0.6	0.11 +0.18/- 0.11	0.056
A2163	0.203	12.5 +1.0/-0.8	0.24 +/-0.08	0.509
A963	0.206	6.3 +0.5/-0.6	0.33 +/-0.11	0.134
A1246	0.216	6.0 +0.6/-0.3	0.22 +/-0.10	0.106
A773	0.2170	7.9 +/-0.5	0.21 +/-0.07	
A1942	0.224	3.9 +0.6/-0.4	0.12 +0.27/-0.12	0.045
Zw3146	0.2906	6.3 +/-0.4	0.26 +/-0.08	0.212
A1722	0.328	5.8 +0.8/-0.6	0.42 +0.19/-0.17	0.058
A370	0.373	6.5 +1.1/-0.9	0.33 +0.21/-0.20	0.049
3C295	0.461	5.83 - 7.02	0.28 - 0.53	
Cl0016+16	0.545	7.4 +1.2/-1.0	0.10 +0.18/-0.10	0.052

abundance relative to cosmic values($[Fe/H]=4.68\times10^{-5}$)

telescope response are shown for plasma temperatures(kT) of 1-10keV and abundance of 0.5 relative to cosmic values in Fig.1. They are less sensitive to kT above 10keV. Intensities of emission lines depend upon elemental abundance, ionization states and kT. There shows distinct difference of the energy resolution between SIS and GIS. The equivalent width (indicator of line intensity) of He-like and H-like emission lines of Si and Fe-K is shown against kT in Fig.2. Thus the plasma temperature(kT) and abundance of each cluster corrected for the cosmological redshift(z) were derived by fitting R-S model(Fig.1) to observed spectra integrated over whole angular extent of clusters, as shown in Fig.3 and summerized in table 1.

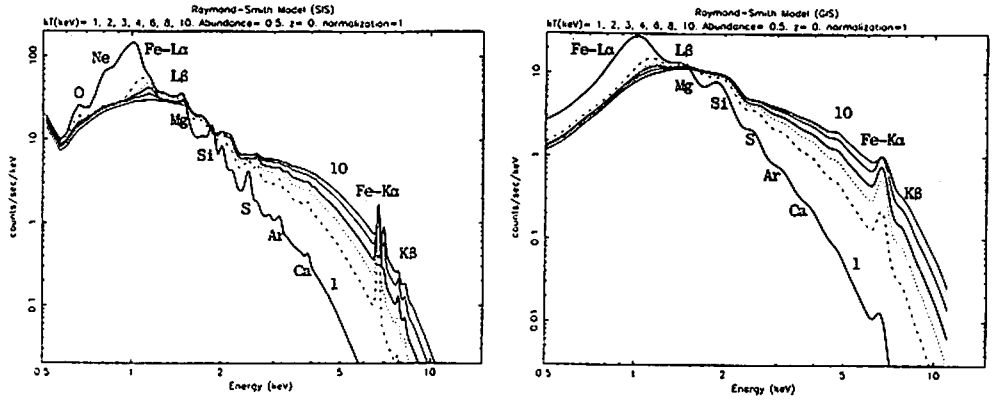


Fig.1 Expected spectra of R-S model folded with telescope response for SIS(left) and GIS(right)

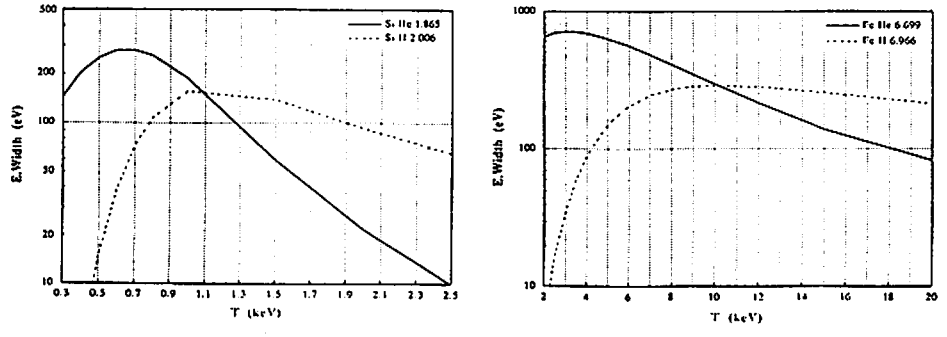


Fig.2 Equivalent width(eV) of He-like and H-like emission lines of Si(left) and Fe(right) against plasma temperature(keV)

It is also important to fit a model of thermal bremsstrahlung and Gaussian lines as free parameters of kT, line intensities and energies. In this case we can derive an electron temperature from the continuum component, an ionization temperature from intensity ratio of H-like and He-like ions, abundance from line intensities of each

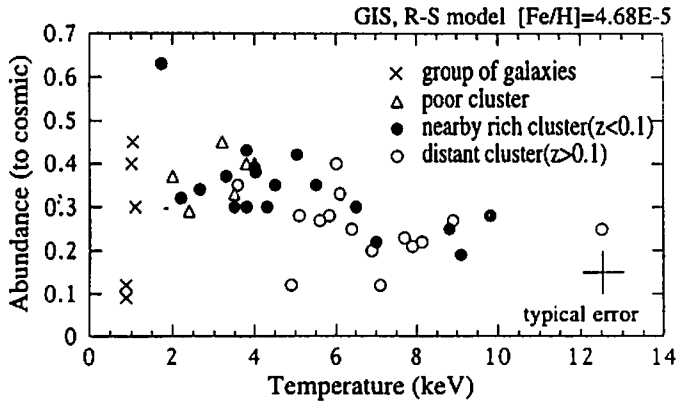


Fig.3 Abundance vs. plasma temperature (see Table 1)

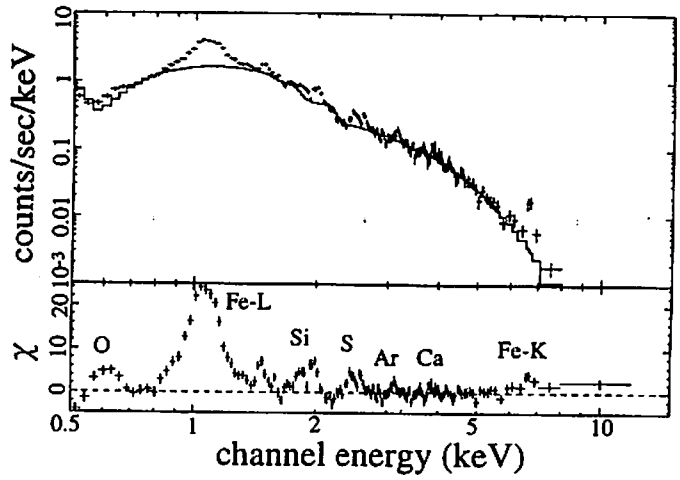


Fig.4 SIS spectrum of M87 within the radius of 2.5' in the Virgo cluster fitted with continuum component of kT=1.9keV

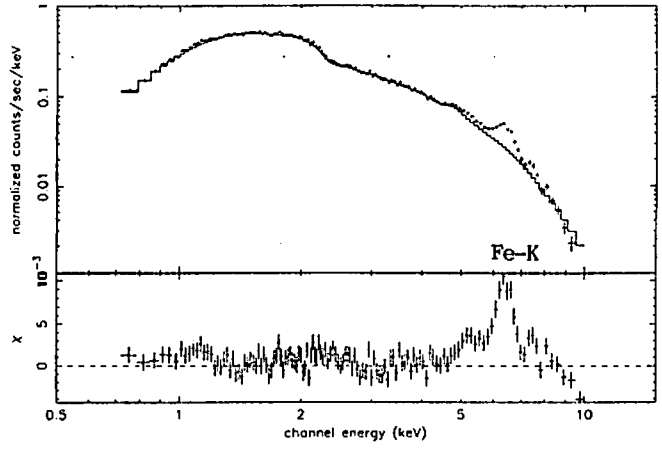


Fig.5 GIS spectrum of A644 fitted with continuum component of kT=6.5keV

element and the redshift from line energies. This procedure is very useful to qualify observed data. SIS has a capability to resolve emission lines of H-like and He-like ions of heavy elements, whereas GIS can distinguish $K\alpha$ and $K\beta$ emission lines. Furthermore the optical depth of each line in an intracluster medium can be estimated from the intensity ratio of $K\alpha$ and $K\beta$. In Fig.4 there is shown an SIS spectrum of M87(z=0.0038) within the radius of 2.5' fitted with a thermal bremsstrahlung of kT=1.9keV. Residuals clearly show the existence of several emission lines, such as O, Si, S, Ar, Ca and Fe. If we introduce number of emission lines at expected energies of each element, we can get satisfactory fitting. The intensity ratio of H- and He-like emission lines of Si and S gives consistent temperature to that of the continuum. Strong enhancement around 1keV is due to Fe-L line complex mixed with Ne and Mg lines. R-S model with a single temperature does not fit well to these data, so that we need at least two temperature components, otherwise we have to take into account uncertainty of calculated intensities of emission lines. If kT>4keV, only Fe-K lines are dominant and other elements are fully ionized. GIS spectrum of A644(z=0.0704) fitted with a thermal bremsstrahlung of kT=6.5keV is shown in Fig.5. Fe-$K\alpha$ and $K\beta$ are clearly recognized in the residuals.

The image processing can be done by folding the point spread function, vignetting and instrumental image distortion, which is rather complicated for extended objects like clusters. Raw image of A644 observed in 1-10keV for 50 ksec by GIS is shown in Fig.6, which is completely covered in the GIS field of view. The surface brightness distribution against the projected radius is fitted by a standard β-model, $I(r)=I(0)(1+(r/r_c)^2)^{-3\beta+1/2}$, where r_c is core radius, assuming spherical symmetric distribution. r_c and β of this cluster are obtained to be 3 arcmin and 0.68, respectively. The image deconvolution is still under way in order to apply to nearby clusters with large angular extent. Some of results for distant clusters with small radius have been published.

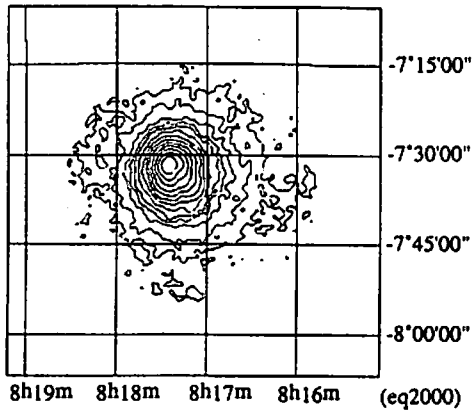


Fig.6 X-ray image of A644 observed by GIS

4. Cooling flow

Well relaxed clusters with a central dominant galaxy(cD) show an evidence of cooling flow which are observationally recognized as excess emissions and absorption in the central core region and radially decreasing temperature distribution to the cluster center. These clusters are shown in Fig.3, having temperature lower than 7keV and abundance higher than 0.2. Most of poor clusters have temperature in the range of 2-4keV. Spectro-imaging observations with ASCA clearly revealed the radial gradient of temperatures in many clusters as well as the abundance gradient in some cases. Abundance should be determined by taking into account resonant scattering effect of emission lines and physical process of recombining plasma in an intracluster gas. In this sense it seems that presently known abundance has large uncertainty. In Fig.7 the temperature distribution of M87/Vir, A496, the Per cluster and A644 is shown against the projected radius(Mpc) with H_0=50 km/sec/Mpc. Angular radius of these clusters are divided into annular ring of 2.5', 5,' 10' and 15'. Generally speaking, low temperature clusters show steep distribution which reflects the potential given by dark matter in the cluster. The radiative cooling is more effective in the central core region within the cosmic time scale, since the emission measure is proportional to square of electron density. The cooling radius could be defined by temperature gradient, though outer boundaries of clusters are not well defined due to the poor statistics.

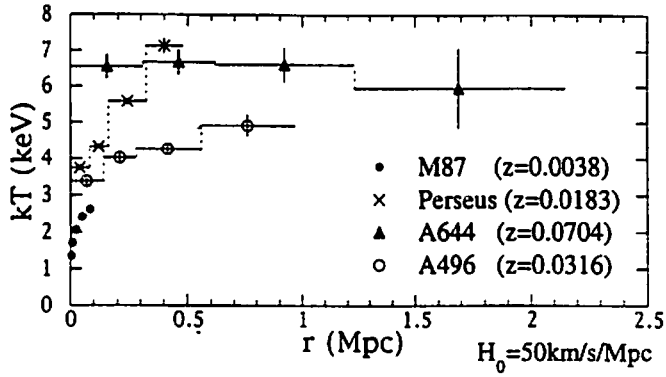


Fig.7 Plasma temperature vs. projected radius

5. Merging

Spherically symmetric clusters without a central dominant galaxy classified as nXD seem to be isothermal and in the hydrostatic equilibrium. These clusters have relatively high temperature, low abundance and less central concentration than cooling flow clusters. The temperature of the Coma cluster averaged over whole cluster was obtained to be 8.09keV by Ginga. ASCA observed spatially resolved spectra within the radius of 40 arcmin. They show inhomogeneous temperature distribution in the range of 5.7-12keV as shown in Fig.8. The Coma accompanies a subcluster which shows low

temperature[8]. This fact indicates that different temperature component corresponds to different substructure going to merge in a cluster potential. It is also possible to consider that shock heating during merging produces patchy temperature structure. This problem is related to top down or bottom up senario in the formation of galaxies and clusters, since the mass of intracluster gas is several times higher than that of all the member galaxies in this type of clusters. Low abundance suggests that the fraction of primordial gas is larger than that of processed gas probably ejected from galaxies. We have to wait for further detailed analysis of ASCA data combining with ROSAT data.

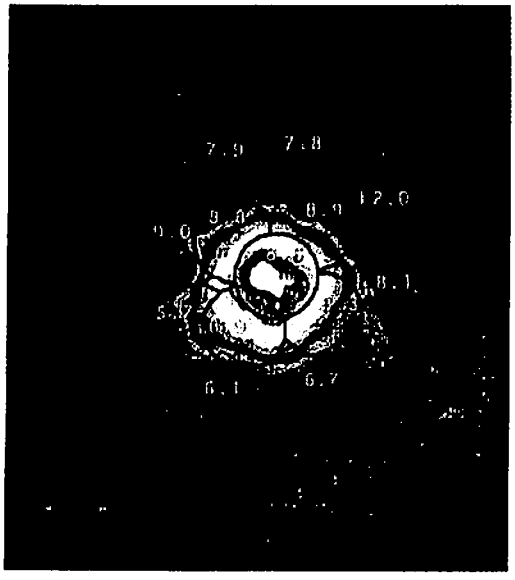


Fig.8 X-ray image of the Coma cluster made of 14 pointed observations by GIS. Temperarures(keV) of region by region are written in the figure

6. Determination of the Hubble constant

Combining the decrement of the brightness temperature of the cosmic microwave background through an intracluster gas with plasma temperature obtained by ASCA and the surface brightness profile observed by Einstein, we have derived H_0 to be less than 50 km/s/Mpc for A665(z=0.1816) and CL0016+16 (z=0.545)[5]. The Hubble space telescope observations of galaxies in the distance up to the Coma cluster (z=0.0235) gave the value to be 80(+/-17) km/s/Mpc[9]. These results are summerized in Fig.9. We have proposed another method by means of resonant scattering effect of Fe-K emission lines in an intracluster gas, as H_0 is expressed as,

$$H_0=73(\mathrm{km/sec/Mpc})(T/10^8\mathrm{K})^{-1/2}(\theta/\mathrm{mrad})(\langle n_e\rangle/5\times10^{-3}\mathrm{cm}^{-3})\,(\mathrm{Ab(Fe)}/\tau(\mathrm{K}\alpha))(z/0.1).$$

where T is plasma temperature, θ angular size, $<n_e>$ mean electron density, Ab(Fe) iron abundance and $\tau(K\alpha)$ optical depth of Fe-Kα emission lines. This method using ASCA data is applicable to spherically symmetric isothermal clusters in z=0.01-0.1. We expect that it would be possible to derive the z dependence of H_0. Data analysis is now going on to derive the definite value with selected clusters.

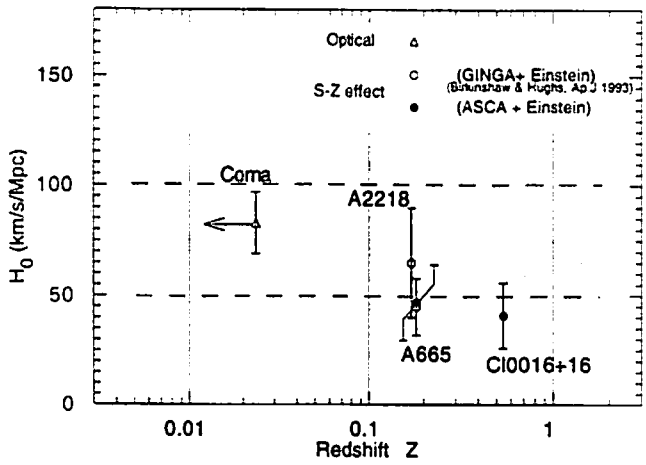


Fig.9 The Hubble constant vs. cosmological redshift

7. Summary

Spatially resolved X-ray spectra of clusters of galaxies have been observed in the energy range of 0.5 - 10 keV and in the redshift range up to 0.5 by ASCA. The evidence of cooling flow and merging in clusters are observationally deduced from the temperature and abundance distribution in an intracluster medium. It is suggested that the derivation of abundance is not so confirmative by means of present method and ASCA data seemingly gives an impact to basic atomic processes. The determination of the Hubble constant is discussed by means of Sunyaev-Zeldovich effect and resonant scattering effect of Fe-K emission lines.

Acknowledgements. I thank to our graduate students for their assistance of data analysis. This work was supported in part by Grant-in-aid for Scientific Research on Specially Promoted Rearch, contract No.07102007, from the Ministry of Education. Science. Sports and Culture, Japan.

References

1. Jones, C. and Forman, W., 1992, NATO ASI Series vol. 366 "Clusters and Superclusters of Galaxies" (ed. A.C. Fabian, Kluwer Acad. Publ.), p. 49.
2. Briel, U.G., and Henry, P.J., 1993, Astr. Astrophys., 278, 379.

3. David, L.P., et al., 1993, Astrophys. J., 412, 479.
4. Yamashita, K., 1992, Proc. Yamada Conference XXVIII "Frontiers of X-Ray Astronomy" eds. Y. Tanaka and K. Koyama, p.475.
5. Yamashita, K., 1994, Clusters of Galaxies, eds. F. Durret, A. Mazure and J. Tran Thanh Van, Editions Frontiers, p.153.
6. Tanka, Y., et al., 1994, Publ. Astron. Soc. Japan, 46, L37.
7. Serlemitsos, P.J., et al., 1994, Publ. Astron. Soc. Japan, 47, 105.
8. Briel, U.G., et al., 1992, Astr. Astrophys. 259, L31.
9. Freedman, W.L., et al., 1994, Nature 371, 757.

Thermal Inflation, the Moduli (Polonyi) Problem and Baryogenesis

Ewan D. Stewart
Research Center for the Early Universe
University of Tokyo

Supersymmetric Scalar Potentials

It is generally believed that the world is supersymmetric with the supersymmetry breaking scale

$$m_{\rm susy} \sim 10^2 \text{ to } 10^3 \,\text{GeV}$$

Scalar potentials of the form

$$V(\phi) = V_0 - m^2\phi^2 + \frac{\phi^{4+2n}}{M_X^{2n}} + \dots$$

with $m \sim m_{\rm susy}$, $M_X \sim M_{\rm GUT} \simeq 2 \times 10^{16}\,\text{GeV}$ or $M_X \sim M_{\rm Pl} = 2.4 \times 10^{18}\,\text{GeV}$, and $n \geq 1$, occur naturally in supersymmetric theories; a combination of supersymmetry and gauge symmetries forbidding a significant quartic term.

Such a potential has its minimum at

$$\phi = M \sim m^{1/(n+1)} M_X^{n/(n+1)} \gg m$$

and in order to cancel the cosmological constant we require

$$V_0 \sim m^2 M^2$$

For $n = 1$ or 2 we get

$$M \sim 10^9 \text{ to } 10^{13}\,\text{GeV}$$

and

$$V_0^{1/4} \sim 10^6 \text{ to } 10^8\,\text{GeV}$$

Thermal Inflation

The finite temperature effective potential has the form

$$V(T) \sim \left(V_0 + T^4\right) + \left(T^2 - m^2\right)\phi^2 + \ldots$$

for $\phi \lesssim T \ll M$ and couplings of order one. Thus, in the early Universe, ϕ can be trapped at $\phi = 0$ with V_0 dominating the energy density when

$$m \lesssim T \lesssim V_0^{1/4} \sim \sqrt{mM}$$

We therefore get

$$N \sim \ln\left(\frac{T_{\text{initial}}}{T_{\text{final}}}\right) \sim \ln\left(\frac{\sqrt{mM}}{m}\right) \sim 10$$

e-folds of 'thermal' inflation.

We still require ordinary inflation to occur before thermal inflation in order to solve the horizon, flatness and homogeneity problems and to generate density perturbations. So what does thermal inflation do?

Moduli

Moduli are predicted by superstring theory. They are scalar fields with gravitational strength interactions and, neglecting supersymmetry breaking, exactly flat potentials.

In the true vacuum with supersymmetry breaking they have potentials of the form

$$V_{\text{true}}(\Phi) = m_{\text{susy}}^2 M_{\text{Pl}}^2 f\left(\frac{\Phi}{M_{\text{Pl}}}\right) = m_\Phi^2 \left(\Phi - \Phi_{\text{true}}\right)^2 + \ldots$$

where f is some generic function and $m_\Phi \sim m_{\text{susy}}$.

In the early Universe, the energy density also breaks supersymmetry and gives an additional contribution

$$V_{\text{cosm}}(\Phi) = H^2 M_{\text{Pl}}^2 g\left(\frac{\Phi}{M_{\text{Pl}}}\right) = \alpha H^2 \left(\Phi - \Phi_{\text{cosm}}\right)^2 + \ldots$$

where $\alpha \sim 1$ and $\Phi_{\text{cosm}} - \Phi_{\text{true}} \sim M_{\text{Pl}}$. The precise values of α and Φ_{cosm} depend on the composition of the Universe.

The total potential is then very crudely given by

$$V(\delta\Phi) \sim \left(m_\Phi^2 + H^2\right)\left[\delta\Phi - \left(\frac{H^2}{m_\Phi^2 + H^2}\right)\Phi_0\right]^2 + \ldots$$

where $\delta\Phi = \Phi - \Phi_{\text{true}}$ and $\Phi_0 = \Phi_{\text{cosm}} - \Phi_{\text{true}} \sim M_{\text{Pl}}$.

The Moduli (Polonyi) Problem

Thus when $H \gg m_\Phi$

$$\delta\Phi = \Phi_0 \sim M_{\rm Pl}$$

and so when H falls below m_Φ,

$$V(\delta\Phi) \sim m_\Phi^2 \delta\Phi^2$$

with

$$\delta\Phi \sim \Phi_0 \sim M_{\rm Pl}$$

initially. This corresponds to an abundance of moduli particles

$$\frac{n_\Phi}{s} \sim 10^7$$

However, the nucleosynthesis constraint is

$$\frac{n_\Phi}{s} \lesssim 10^{-12} \text{ to } 10^{-15}$$

Thus about 10^{20} times too many moduli particles are produced at $H \sim m_\Phi$.

Thermal Inflation and the Moduli Problem

Ordinary inflation occurs at too high an energy scale to be able to dilute the moduli produced at $H \sim m_\Phi$. However, thermal inflation occurs at a relatively low energy scale corresponding to

$$H \sim \frac{mM}{M_{\rm Pl}} \ll m_\Phi$$

and so thermal inflation can dilute the moduli produced at $H \sim m_\Phi$. There is enough dilution if roughly

$$M \gtrsim 10^{10} \text{ to } 10^{12}\,\text{GeV}$$

However during thermal inflation

$$V(\delta\Phi) \sim m_\Phi^2 \left[\delta\Phi - \left(\frac{H^2}{m_\Phi^2}\right)\Phi_0\right]^2 + \ldots$$

and so $\delta\Phi \sim (H^2/m_\Phi^2)\Phi_0 \sim M^2/M_{\rm Pl}$. Thus moduli are produced at the end of thermal inflation. Their abundance is not too big if roughly

$$M \lesssim 10^{12} \text{ to } 10^{13}\,\text{GeV}$$

Thus thermal inflation can solve the moduli problem if

$$10^{10} \text{ to } 10^{12}\,\text{GeV} \lesssim M \lesssim 10^{12} \text{ to } 10^{13}\,\text{GeV}$$

Baryogenesis

Big Bang nucleosynthesis requires

$$\frac{n_B}{s} \sim 3 \times 10^{-11}$$

Any baryons formed before thermal inflation will be diluted to nothing.

The temperature after thermal inflation will be high enough for electroweak baryogenesis, but there will be entropy production after the electroweak phase transition and so electroweak baryogenesis will not work either.

However, if the scalar field that gives rise to thermal inflation also gives right-handed neutrinos their mass, then a lepton asymmetry can be generated at the end of thermal inflation which would then be transformed into a baryon asymmetry by electroweak processes.

Ordinary Inflation and Thermal Inflation

	Ordinary Inflation	Thermal Inflation
clock	slowly rolling scalar field	temperature
e-folds	$\gg 40$	~ 10
energy scale $V^{1/4}$/GeV H/GeV	 10^{16} to 10^{10} 10^{14} to 10^{2}	 10^{8} to 10^{6} 10^{-3} to 10^{-6}
theory	difficult	natural
horizon, flatness and homogeneity problems	yes	no
density perturbations	yes	no
moduli problem	no	yes
baryogenesis	no	yes

References

[1] D. H. Lyth and E. D. Stewart, hep-ph/9510204, to be published in Phys. Rev. D (1996).

Mexican Hat Gravitational Lenses

Paul Haines*, Waseda University
G.F.R. Ellis and Deon Solomons
Department of Mathematics and Applied Mathematics
University of Cape Town
Rondebosch 7000, South Africa

March 4, 1996

Abstract

The mexican hat function is used to model a compensated (in density) gravitational lens. First we discuss the need for compensation. Then we integrate photon orbits to construct wavefronts. One consequence of underdensities is an off-axis caustic. We report the structure of the critical curves and caustics using the thin lens approximation, both for the mexican hat and for a perturbed mexican hat. We discuss the nonlocal effect of shear.

1 Introduction

Gravitational lensing of distant galaxies by foreground clusters of galaxies is a recent and powerful new tool for observational cosmology. For example, the HST image of the Abell Cluster A2218 dramatically illustrates that images of background galaxies are distorted [8]. Lensing can be used either as a powerful telescope to observe very distant galaxies, or to investigate the distribution of matter around the lensing cluster. However, the current models for mass distribution of clusters are all uncompensated; for example, the most common model is the singular isothermal sphere, and the mass is formally infinite.

2 Metric

We will consider a metric of the form

$$ds^2 = -(1+2\phi)dt^2 + a^2(t)(1-2\phi)(dx^2+dy^2+dz^2). \tag{1}$$

This is probably adequate for discussions of gravitational lensing. We should assume that $|\phi| << 1$ and $\dot{\phi} = 0$. (This second assumption is not true in general; it is a good approximation for flat universes, but it breaks down for $\Omega < 1$, with important observational consequences for the CMB [7].) The density fluctuations are then found to be $\delta \propto \nabla^2\phi$, where $\rho = \rho_0(1+\delta)$, and δ need not be small.

Gravitational lens models usually consider only positive mass concentrations. In fact, the negative fluctuations must occupy more volume. "Voids" are easily identified in the galaxy redshift surveys. This issue of compensation is related to the more general problems of *fitting* and *averaging* [6]. If we naively add cluster-sized lenses to our "background" model, assuming 10% of the mass is bound in

*e-mail: haines@cfi.waseda.ac.jp

clusters, we would incidentally increase the mass of the universe by 10%, thus changing the background model. The new Hubble constant would be 5% larger.

On the other hand, if we want the average of metric (1) to be the background metric, then we should insist $< \phi >= 0$. However, even if we neglect the difficulties of averaging in general relativity, it is not possible to have $< \phi >= 0$ and $< \delta >= 0$ simultaneously. What does matter physically is $< \delta >$. This is what must be properly matched, because it is physically measurable (unlike the potential). Therefore to match the space-times we must have $< \delta >= 0$. We accept whatever this then implies for the metric and so for ϕ.

3 Mexican hats

The matter distribution for our model lens is

$$\delta(r) = C\left(3 - \frac{r^2}{r_c^2}\right)\exp\left(-\frac{r^2}{2r_c^2}\right) \tag{2}$$

which is a "mexican hat". This is a compensated perturbation: $\int r^2 dr\sqrt{-g}\delta(r) \approx 0$; r_c is the scale width of the distribution. The general relativistic correction for the volume element is $\sqrt{-g} \approx 1 - 2\phi$; as $\phi \to 0$, the lens becomes exactly compensated. In the thin lens approximation, the density is projected onto a plane and then scaled. The two dimensional surface density is

$$\kappa(x) = \kappa_0\,(2 - x^2)\,\exp(-x^2/2). \tag{3}$$

4 Gravitational lensing

First we integrate the equation of motion for a photon, following Weinberg [11], using the perturbed metric (1). The equation of motion for a photon is

$$\frac{du}{dt} \approx -2\nabla\phi + 4u(u\cdot\nabla\phi) \tag{4}$$

where $u^i = dx^i/dt$ is the coordinate photon velocity. The perturbed gravitational potential is

$$\phi = -4\pi C r_c^2 \exp\left(-\frac{r^2}{2r_c^2}\right). \tag{5}$$

Note $\nabla^2\phi \propto 4\pi\delta(r)$.

Now we integrate individual photon orbits numerically. We find that for positive values of C, which characterise our toy models for clusters, the wavefront can form a cusp, and an observer will see multiple images. For negative values of C, which characterise our toy models for voids, an interesting off-axis cusp forms. The source and lens are both on the x axis to the far left of the plots. The graph for $C = 0$ should be a section of a sphere. Only the endpoints for photons with positive initial angle are shown, so for $C > 0$ some paths are deflected below the x-axis. Please refer to Figures 1 and 2.

It is much quicker for most lensing applications to use the thin lens approximation. This is described in detail in the book by Schneider, Ehlers, and Falco [9]. For the projected mexican hat lens, the scaled velocity potential, deflection angle, and Fermat potential are

$$\Psi(x) = 2\kappa_0\left(1 - \exp(-x^2/2)\right), \tag{6}$$

$$\alpha(x) = 2\kappa_0 x\exp(-x^2/2), \tag{7}$$

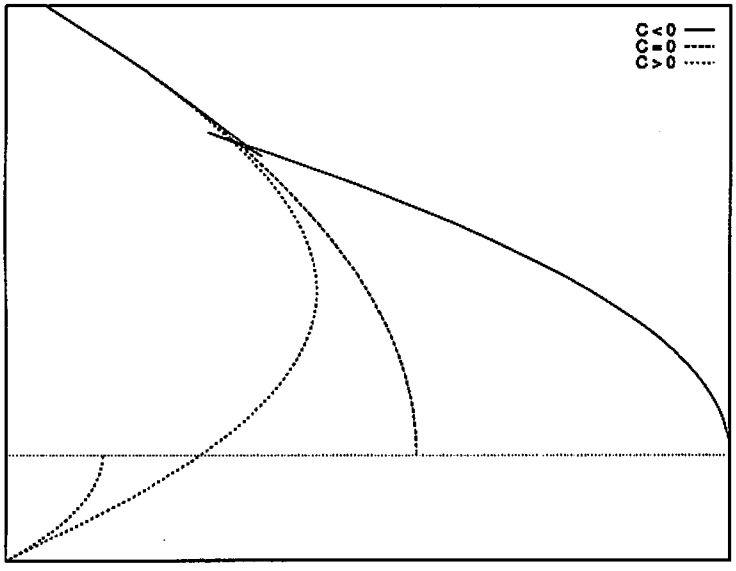


Fig. 1: Wavefronts.

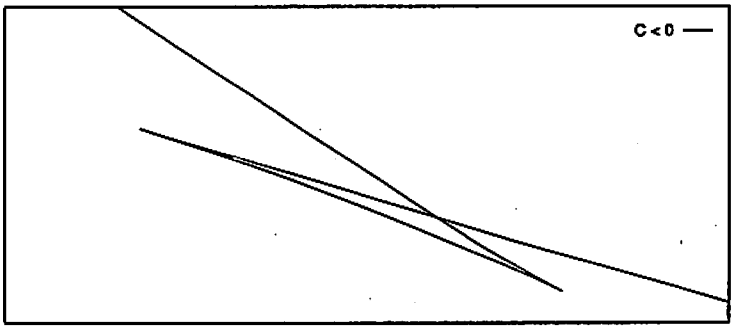


Fig. 2: Detail of the cusp due to the underdensity lens.

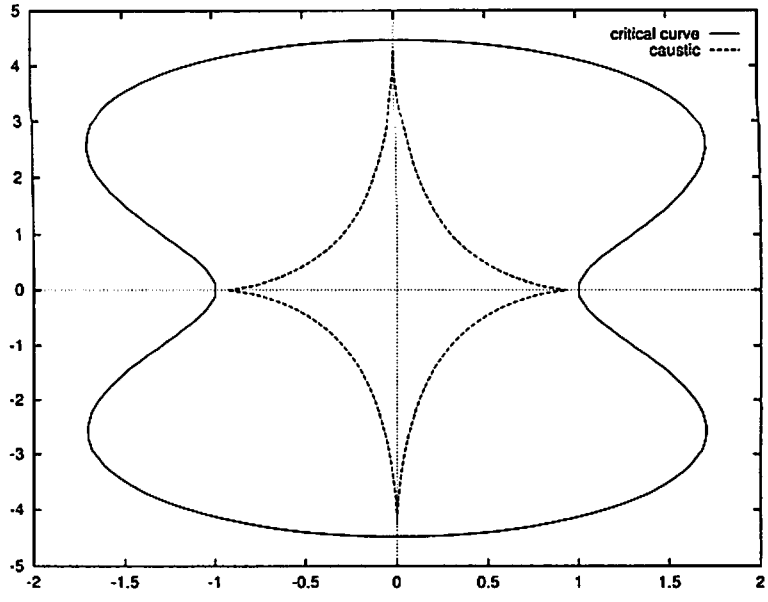


Fig. 3: Caustic and critical curve for the Chang-Refsdal Mexican hat lens.

$$\Phi(x,y) = \frac{1}{2}(x-y)^2 - \Psi(x). \tag{8}$$

The lens equation is then

$$x\left(1 - 2\kappa_0 \exp(-x^2/2)\right) - y = 0. \tag{9}$$

The source position is y and the image position is x. The caustics and critical curves are circles.

A more interesting model is based on the Chang-Refsdal lens system. In the CR system, the lens is a star, but the external influences of the host galaxy's surface mass and shear field break the symmetry and greatly enhance the lensing effect [3]. We treat the "cluster" as a large mexican hat, and then consider a single member galaxy "member" as the lens. For one particular configuration, the critical curves and caustics are seen in figure 3. We then model distant galaxies (in the source plane) as circles, and resulting distorted images are seen in figure 4. The distortion pattern is similar to that produced by the singular isothermal sphere model.

5 Shear

In the circularly symmetric lens case, there is no shear. However, once the lens is perturbed, shear is present. The general lens equation shows two components to the deflection, surface density and shear. The surface density contributes in a "local" way in the sense that the mass interior to a bundle of rays focuses them. But the source of the shear term includes the entire mass distribution – and there

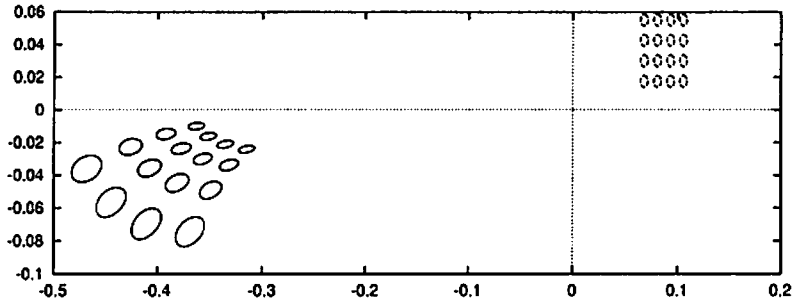


Fig. 4: Sources (on the right) and images seen through the Chang-Refsdal mexican hat lens (on the left).

is a component of shear due to the underdense regions. Recent papers show that the shear due to large scale structure cannot be neglected [1], [10]. Although shear from voids may be less important, perhaps it should not be ignored.

6 Acknowledgment

It is a pleasure to thank Kei-ichi Maeda and R. Easther for reading the manuscript.

References

[1] R Bar-Kenna (1995). *preprint*, astro-ph/9511056.

[2] M Bartelmann (1995). *A&A*, **303**, 643.

[3] K Chang, S Refsdal (1979). *Nature*, **282**, 561.

[4] J Ehlers (1960). *Abh. Mainz Akad. Wiss. u. Litt.*, Mat-Nat. Kl., Nr. 11.

[5] G F R Ellis (1971). In *General Relativity and Cosmology*, Proc Int School of Physics "Enrico Fermi" (Varenna), Course XLVII. Ed. R K Sachs (Academic Press), 104-179.

[6] G F R Ellis (1988). In *Proc. 2nd Canadian Conf. on GR and Astrophysics*, Ed. A Coley et al (World Scientific), p1.

[7] M Kamionkowski, D N Spergel (1994). *ApJ*, **432**, 7.

[8] J P Kneib, R S Ellis, I Smail, W J Couch, R M Sharples (1995). preprint astro-ph 9511015.

[9] P Schneider, J Ehlers, E Falco (1992). *Gravitational Lenses* (Heidelberg: Springer)

[10] J V Villumsen (1995). *preprint*, astro-ph/9507007.

[11] S Weinberg (1992). Gravitation and Cosmology (Wiley: New York).

Anisotropies of Cosmic Microwave Background and Observational Cosmology

Naoshi SUGIYAMA
Department of Physics & Research Center for the Early Universe, The University of Tokyo, Tokyo 113,
e-mail: sugiyama@utaphp1.phys.s.u-tokyo.ac.jp

Rapid increasing of observational data reveals the weakness of the paradigm of the cold dark matter (CDM) scenario. We briefly summarize the current status of the observational cosmology and its problem. On the other hand, Cosmic Microwave Background (CMB) anisotropies are becoming one of the most important observational quantities to understand the early stage of our universe. In this paper, it is shown that temperature fluctuations of CMB contain rich informations, i.e., cosmological parameters, geometry of the universe, initial condition of the density perturbations and more. In order to obtain intuitive insight, we develope simple toy models describing the acoustic oscillation of photon-baryon systems. Together with full numerical calculations, it will be able to reconstruct the power spectrum of temperature fluctuations if once we get a full sky map in fine resolution.

I. STATUS OF OBSERVATIONAL COSMOLOGY

One of the most important purposes of the observational cosmology is to determine the cosmological parameter of the Friedmann universe, i.e., density parameter Ω_0, Hubble constant h which is normalized by 100km/s/Mpc, cosmological constant Λ, baryon density Ω_B and so on. Recently, the number of observations are rapid increasing. Two new powerful telescopes, HST and Keck are providing marvelous data of faint deep sky objects. Several galaxy surveys, such as APM, IRAS, QDOT and CfA, have been already done. X-ray satellites ROSAT and ASCA show us new feature of the sky which is observed by X-ray. Several new observational quantities such as gravitational lensing and neutrino have been appeared. Moreover, many new projects are waiting in line. Keck II will take its first light in this year. Japanese 8-m telescope Subaru is now building. The Sloan Digital Sky Survey is a project to measure the redshift of millions of galaxies and hundreds of thousands QSOs. Super-Kamiokande is almost ready to take data. There are several projects to detect gravitational wave. It is almost certain that we are going to have huge amounts of new data in the next decade. Later in this paper, we discuss the possibility to use Cosmic Microwave Background (CMB) anisotropies as a new tool to obtain the information of the early universe.

However, the current status of the observational data is rather confusing. Let us show some of them here.

- Age of the universe: the age of globular clusters sets the lower bound on the cosmic age as $t_0 \geq 15.8 \pm 2.1$Gyr. (Bolte & Hogan 1995). From the luminosity function of the White Dwarf, we know the age of the disk of our galaxy as $10^{+2.5}_{-1.5}$Gyr. (Oswalt, Wmith & Wood unpublished). It is consistent with the age of globular clusters because the disk component was formed later than the halo. We should accept 13 or 14Gyr. as the lower bound of the cosmic age.

- Hubble constant: New measurements of Cepheid variables in M100 or Leo1 by HST provide fairly high value of the Hubble constant as $H_0 = 80 \pm 17$km/s/Mpc (Freedman et al 1994) or 69 ± 8km/s/Mpc (Tanvir et al 1995). Recently, a new method to obtain the distance to type Ia Supernovea is presented by Riess et al (1995). They take into account the fact that the declining of the light curve after the peak depends on the absolute maximum luminosity of the explosion. They succeed to reduce errors of measurements of distance by using this new method. And their value is again fairly high as $H_0 = 67 \pm 7$km/s/Mpc. These high values of the Hubble constant do not allow us to have an old universe without assuming the existence of the cosmological constant. For example, if we take $H_0 = 70$km/s/Mpc and $t_0 = 14$Gyr., Ω_0 has to be almost 0 in the Friedmann universe.

- Density parameter Ω_0: Measuring the peculiar velocity, we can obtain the mass distributions. Strauss & Willick (1995) summarize current status of peculiar velocity measurements. Since it is very difficulty to get the distance, there are still big deviations of the number of Ω_0 between each measurements: $\Omega_0^{0.6}/b = 0.4 - 1.0$, where b is the bias parameter.

- Cosmological constant Λ: The strongest and most reliable constraint on Λ is obtained by the statistical counts of the gravitational lensing event. The latest study by Kochanek (1995) shows $\Lambda < 0.66$.

- Deceleration parameter q: If we measure deep sky $z \geq 0.1$, observational quantities are sensitive to the deceleration parameter q (see e.g., Weinberg 1972). The relation between q and Ω_0 and Λ is : $q = \Omega_0/2 - \lambda$. Recently

two important measurements have been done. First one is the number count of galaxies in K band by the Keck telescope (Gardner et al. 1993 and Djorgovski et al 1995). Their favourite value is $q \sim 0.02$. They conclude neither a cosmological constant dominated universe ($\Lambda = 0.9$) nor an Einstein-De Sitter universe is consistent with their data. The second one is the measurements of distant SNIa ($z \geq 0.1$) by the group of Lawrence Berkeley Laboratory. Although they are still analysing data, it seems to be at least q is positive.

- Baryon density parameter Ω_B: It is known that $\Omega_B h^2 = 0.01 - 0.03$ from the argument of the standard big bang nucleosynthesis (see e.g., Walker et al 1991). Hata et al (1995) studied the statistics of observations of light elements. However, they couldn't find the allowed region if they assumed that the number of the neutrino family is three. Moreover the recent detection of the Deuterium Lyman alpha line (Sangaila et al 1994) suggests extreamly low value of Ω_B as $\Omega_B h^2 = 0.006$. Finally, White et al (1993) obtained the Ω_B in the coma cluster. They concluded that about 10% of the total mass is baryonic.

As is shown above, it is premature to give any definite conclusions about the cosmic parameters. It seems that low density universe is favourite from the observations of the Hubble constant and the cosmic age. On the other hand, it is unlikely to have a high value of the cosmological constant.

As for the large scale structure formation of the universe, the CDM model was the standard in the last decade. However recent deep observations reveal serious discrepancies between theoretical predictions of this model and observational data. There are two serious problems. First one is the mismatch of the shape of the matter power spectrum. The power spectrum of the CDM model has a turnover point in its shape which corresponds to the horizon scale of the matter radiation equality which is proportional to $\Omega_0 h^2$. Since the unit of the wave number is $h\mathrm{Mpc}^{-1}$, the CDM power spectrum is parametrized by a single parameter $\Gamma \equiv \Omega_0 h$. In order to match the observational data, $\Gamma \simeq 0.25$ (Peacock & Dodds 1994). For the standard CDM ($\Omega_0 = 1$ and $h = 0.5$), $\Gamma = 1 \times 0.5 = 0.5$. The second problem is that the power spectrum has too much power on small scales if we normalize it by COBE-DMR data. The number of σ_8 which is the fluctuations of mass distribution at $8h^{-1}$Mpc is 1.4 for the standard CDM model by using COBE normalization. It is unti-bias. However, the large scale structure simulations prefer lower value, $\sigma_8 \approx 0.6 - 0.7$. In order to avoid these problems, several alternative scenarios such as low density CDM models, mixed dark matter models and tilted initial power models, are considered. But non of them are as attractive as the simple CDM model because we need to introduce additional parameters.

II. CMB AND THERMAL HISTORY

In 1992, the Cosmic Background Explore (COBE) discovered temperature fluctuations on Cosmic Microwave Background (CMB) by using Differential Microwave Radiometer (DMR) (Smoot et al. 1992). After this sensational discovery, the number of observational data of temperature fluctuations have been rapidly increasing (see e.g., White, Scott and Silk 1994). Nowadays, CMB anisotropies are becoming one of the key observational object in the modern cosmology. CMB anisotropies provide us direct information at last scattering surface, i.e., redshift $z \approx 1000$. Since the shape of the angular power spectrum of CMB anisotropies is highly sensitive to geometry of the universe, cosmological models and cosmological parameters, i.e., density parameter Ω_0, Hubble constant h which is normalized by 100km/s/Mpc, cosmological constant Λ, baryon density Ω_B and so on, CMB anisotropies are expected to be a new tool to understand our universe.

In the early stage of the universe, there were at least three important epochs for the fate of CMB. At the redshift $z = 7 \times 10^6$, the expansion time and the time scale of the double Compton scattering became equal. In other word, the double Compton scattering was not effective anymore after this epoch. Before that, since the double Compton scattering changed the number of photons, the distribution function of photons could keep the black body shape without a chemical potential even if there existed an incident which produced photons such as the radiative decay of the massive neutrino. On the other hand, if such incident happened after this epoch, the chemical potential μ would be produced on the photon distribution function.

Second epoch was the one when the expansion time and the time scale of the (single) Compton scattering became equal. It was at $z = 10^5$. After this epoch, photons and electrons were not anymore in the thermal equilibrium. Therefore if photons obtained energy from moving electrons after this epoch, the black body shape was distorted. These moving electrons transfered the low energy photons to high energy parts. Eventually, on the Rayleigh Jeans regions, we expected negative temperature fluctuations as (Zeldovich & Sunyaev 1969)

$$\Delta T/T = -2y \tag{1}$$

$$y \equiv \int dt \frac{T_e - T_\gamma}{m_e} n_e \sigma_T, \tag{2}$$

where y is the Compton y-parameter, T_e and T_γ are electron and photon temperatures, m_e is the electron mass, n_e is the free electron number density and σ_T is the Thomson scattering cross section.

Almost the perfect black body shape of the observation of COBE-FIRAS detector sets the upper limits on μ and y as (Mather et al. 1994)

$$\mu < 3.3 \times 10^{-4} \tag{3}$$

$$y < 2.5 \times 10^{-5}. \tag{4}$$

These upper limits strongly constrain the energy injection in the early universe.

However these distortions have been already observed in the individual cluster of galaxies. In the cluster of galaxies, there exist hot gases $T_g \approx 10^{6-7}$K. These hot gases distort CMB. This effect is known as the Sunyaev-Zeldovich (SZ) effect (Sunyaev & Zeldovich 1970). If once we know the amount of temperature fluctuations, we can measure the distance to the cluster. In other word, we can get the number of Hubble constant from the SZ effect of clusters. It is because it is possible to obtain the size of the cluster if we individually measure the SZ effect from the radio telescope and the temperature of the hot gas from the X-ray satellite. Assuming the size of the cluster on the line of sight direction and the angular diameter size of the cluster are equal, we can get the angular diameter distance. Recently, the Ryle telescope group in the Cambridge University is working hard to obtain imaging of the SZ effect (Jones et al. 1993). They have already measured distances to several clusters. Their number of Hubble constant is relatively low, i.e., $H \approx 50$km/s/Mpc.

The final important epoch for CMB was the recombination of the hydrogen atom at $z \approx 1000$. After this epoch, the universe became transparent. Photons traveled toward us without being scattered. Therefore observing the CMB means observing the sky at $z \approx 1000$. The fluctuations on CMB which were produced during inflation epoch were frozen out at that epoch. We will discuss about these primary anisotropies in the rest of this paper.

III. PRIMARY CMB ANISOTROPIES

The modern cosmology tells us that the inflation in the very early stage of the universe produced the seeds of fluctuations both on the matter densities and radiation densities. Matter density fluctuations grew to be the large scale structure of the universe such that galaxies, clusters of galaxies, super-clusters and so on. Photon density fluctuations are observed as temperature anisotropies in CMB.

For the reason of simplicity, we treat density perturbations in the Fourier space. The temperature fluctuation Θ is defined as the energy integral of the linear perturbation of the temperature distribution function. We expand it into multipole components ℓ as

$$\Theta(\eta, \mu, k) = \sum_\ell (-i)^\ell \theta_\ell(k, \eta) P_\ell(\mu), \tag{5}$$

where η is conformal time, μ is direction cosign, P_ℓ is the Legendre polynomial. In order to get expected CMB anisotropies for each models with different cosmological parameters, we have to solve the Boltzmann equation for Θ. As end results, we get the temperature power spectrum in ℓ space as

$$\frac{2\ell+1}{4\pi} C_\ell = \frac{V}{2\pi^2} \int \frac{dk}{k} \frac{k^3 |\theta_\ell|^2}{2\ell+1}. \tag{6}$$

There are several efforts to solve the equation numerically (see e.g., Sugiyama 1995). (Most recent one is the COMBA project. Many experts in this field join to get power spectrum for many grids of models in very high accuracy.) It is possible to get C_ℓ in high accuracy from the numerical calculation. Although this procedure itself is perfectly reasonable, the final power spectrum contains a lot of different physical processes and it is very hard to separate each effects numerically. Moreover it takes long CPU time to calculate a single model. Recently we developed an analytic treatment of CMB anisotropies (Hu & Sugiyama 1995a) which makes us possible to separate each contributions. We can calculate CMB power spectrum by very short CPU time within about 10% accuracy in temperature. Short CPU time is important to generate spectrum for many models. Moreover this agreement convinces us that we properly take into account all physical contributions in our calculations. In order to get physical insight, we further developed a toy model based on the analogy of a harmonic oscillator. In this paper, we explain the dependence of CMB anisotropies on cosmological parameters by using this toy model. Throughout this paper, we employ cold dark matter (CDM) models as an example. Applications to other models are rather straight forward.

IV. PHYSICS ON CMB ANISOTROPIES AND THE TOY MODEL

There are several physical effects working on the evolution of CMB anisotropies (see e.g., Hu & Sugiyama 1995a,b). (a)Acoustic oscillations: Before the recombination, photons and baryons are tightly coupled each other. Once fluctuations cross the sound horizon, fluctuations of this tight coupled fluid oscillate as acoustic waves. These acoustic oscillations create peaks and wriggles on the power spectrum. Peaks are generated by not velocity fluctuations but density fluctuations. Therefore the naming of 'Doppler peak' is misleading. We would like to call them as 'Acoustic peaks'. (b)Diffusion Damping: Random walk of photons causes the exponential damping on fluctuations (Silk 1968). The damping length is smaller than the sound horizon before the recombination. During the recombination process, however, the damping length grows to infinity. On the other hand, the scattering probability between photons and baryons decreases and becomes negligibly small right after the recombination. Eventually the total damping rate becomes smaller on larger scales. (c)Gravitational redshift: After recombination, the coupling between photons and baryons are cut off. Without getting further scattering unless reionization of the vast region of the universe happens, temperature perturbations freely stream toward us. Photon perturbations are only disturbed by the gravitational potential. Climbing up the gravitational well, photons loose energy and are redshifted. This redshift effect caused by the gravitational potential at the recombination (last scattering surface) is called as the Sachs-Wolfe (SW) effect (Sachs & Wolfe 1967). For adiabatic perturbations, Gravitational potential stays constant during pure matter or radiation dominated regime in the flat $\Omega_0 = 1$ universe. However, it decays if one of these assumptions are broken. The decay of the gravitational potential also causes the redshift (or rather blue-shift) of photons. Here we refer to this effect as the Integrated Sachs-Wolfe (ISW) effect. The ISW contribution is separated into two parts. Right after the recombination, if the universe is still not purely matter dominated, the potential decays. We call this as the early ISW effect. It should be noticed that this effect is important even for the 'standard' model with $\Omega_0 = 1$ and $h = 0.5$. In case of the low density universe, the potential starts to decay very near present epoch when the curvature or the cosmological constant starts to become the dominant component. We refer to this effect as the late ISW effect. (d)Doppler effect: If the baryon velocity is different from photon's, the baryon velocity induces temperature fluctuations. During the tight coupling regime, baryon and photon velocities are identical. If the universe is transparent, there is no interaction between photons and baryons. Therefore this effect only appears in late time reionized models.

Our analytical treatment includes all above physics. We solve the acoustic oscillations inside the gravitational well which is generated by collisionless particles (dark matter) on small and intermediate scales by using WKB approximation. Before recombination, we can assume the compton scattering time is extremely short. This implies that the differential optical depth $\dot{\tau} \equiv d\tau/d\eta = a x_e n_e \sigma_T$ is very high, where x_e is the ionization fraction, n_e is the electron number density and σ_T is the Thomson scattering cross section. We expand the Boltzmann and Euler equations in terms of τ^{-1}. The zeroth order solution gives the tight coupling between photons and baryons. Ignoring cosmic expansion terms, the acoustic oscillations are obtained as the first order solution. The second order solution provides the diffusion damping. Then we join these solutions with the exact solution on large scales. We take into account all the details, i.e., time evolution of gravitational potential, recombination process and even neutrino quadrupole anisotropies.

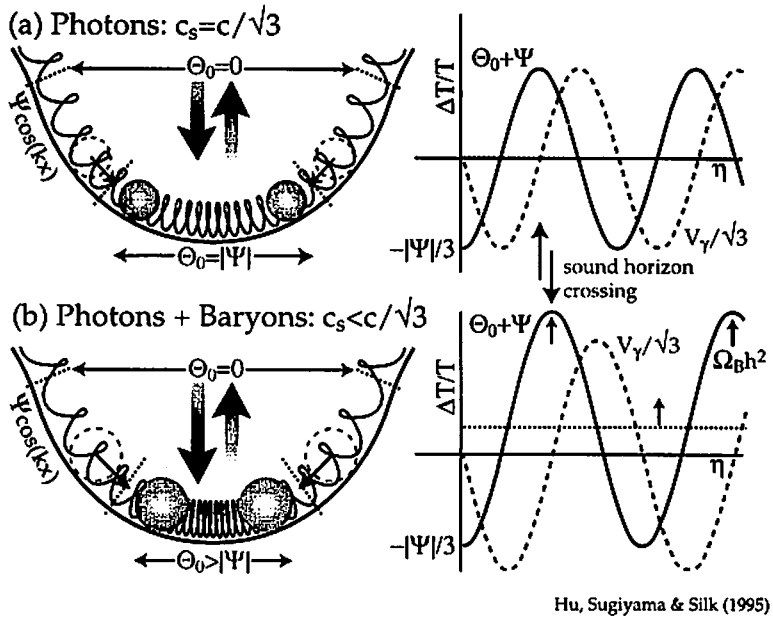


FIG. 1. A toy model of the acoustic oscillation. Θ_0 is temperature fluctuations, Ψ is gravitational potential, V_γ is velocity perturbations of photons, c_s is sound peed and c is the speed of light. (a): low Ω_B limit, i.e., $c_s^2 = c^2/3$. (b): high Ω_B. This figure is reprinted from Hu, Sugiyama & Silk (1995).

Our analogy of a harmonic oscillator is shown in figure 1. Balls connected by a spring are set inside the gravitational well. The deviation of the balls' location from the center of the oscillation corresponds to the amplitude of the fluctuations. The spring represents the pressure. The initial location is specified by the initial condition, i.e., adiabatic or isocurvature conditions. Balls are kept at this initial location until fluctuations cross the sound horizon. Once crossing the sound horizon, they start to oscillate. At the recombination epoch, all oscillations are frozen out. They climb up the potential well to transfer to observers at present. If fluctuations of which wave length is larger than the sound horizon at last scattering surface are considered, balls stay at initial location until recombination epoch. Assuming matter domination epoch, we get the location as $-(2/3)\Psi$ for adiabatic perturbations. Here Ψ is the gravitational potential. The redshift effect caused by climbing up the well is Ψ. Therefore we get the famous Sachs-Wolfe value $\Psi/3$ for long wave length fluctuations. Next let us consider shorter wave length fluctuations. The recombination occurs after oscillation starts. A first compression mode at the last scattering surface corresponds to the first peak, a first depression mode corresponds to the second peak and so on.

V. SHAPE OF CMB SPECTRUM

Using our simple analytic picture, we can easily explain the dependence of CMB spectrum on various cosmological parameters. First, we investigate the height of the peaks. General statements are followings. Pressure of photon-baryon fluid prevents fluctuations from growth. Hence less pressure (which corresponds to heavier balls in our analogy) means higher peaks. On the other hand, Gravitational potential induces adiabatic growth. This causes blue-shift. Deeper potential seems to generate higher peaks. However, deeper potential well means bigger SW effect which causes redshift. Therefore deeper potential doesn't necessarily produce higher peaks. It is known that the gravitational potential decays if fluctuations cross the horizon during radiation dominated era. This decay boosts acoustic oscillations. First reason of this boost is because of resonant oscillation since the decay of potential approximately synchronizes with the cycle of acoustic oscillations. Second reason is the time dilation effect. Gravitational potential stretches the geometry and causes time delay, i.e., redshift on CMB anisotropies. As potential decays, photons get blue-shifted. Eventually, the shallower (deeper) potential makes peaks higher (lower).

How about the dependence of each cosmological parameters? If we consider $\Omega_B = 0$ limit, the sound speed is $c/\sqrt{3}$ where c is the speed of light. In this case, the amplitudes of density and velocity peaks are identical (see Fig.1 (a)). Therefore there is no peak in C_ℓ simply because density and velocity perturbations are $\pi/2$ off phase. Let us consider more realistic case. The sound speed is always smaller than $c/\sqrt{3}$ if we take into account the baryon density. Increasing $\Omega_B h^2$ decreases the sound speed. Hence we can expect higher peaks (See Fig.1 (b)). Increasing $\Omega_0 h^2$ pushes the matter-radiation equality earlier. This causes potential to be deeper and peaks to be lower.

We can explain most of parameter dependence from above arguments. For example, if we increase Ω_B with fixing Ω_0 and h, the sound speed becomes smaller. Therefore we get larger peaks as shown in figure 2 (a). Because above two effects compensate each other, h dependence with fixing Ω_B and Ω_0 is complicated. However, if we fix $\Omega_B h^2$ which is determined by BBN, increasing h simply provides lower peaks (see figure 2 (b)).

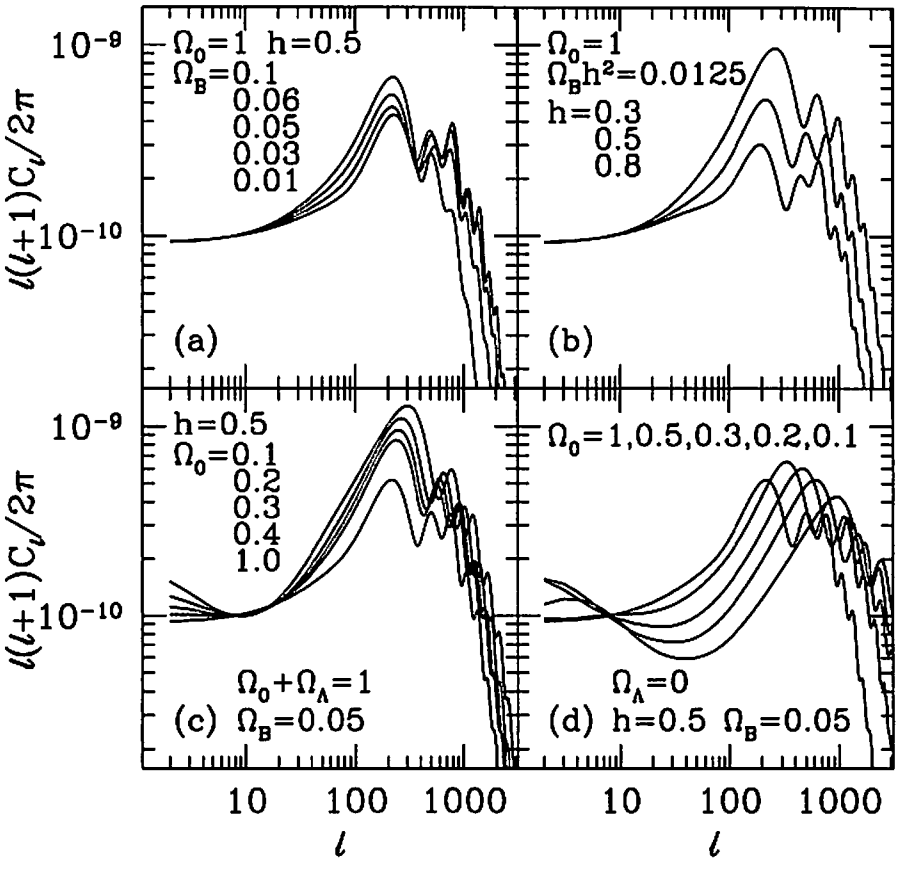


FIG. 2. Cosmological parameter dependence of CMB power spectrum C_ℓ's as function of multipole ℓ. C_ℓ's are normalized to COBE-DMR 2 year data. Panel (a) shows Ω_B dependence with fixing Ω_0 and h. From the top to the bottom, $\Omega_B = 0.1, 0.06, 0.05, 0.03$ and 0.01. Panel (b) is h dependence with fixing $\Omega_B h^2$. From the top to the bottom, $h = 0.3, 0.5$ and 0.8. Panel (c) and (d) show Ω_0 dependence for flat cosmological constant dominated models and open models, respectively. From the top to the bottom of panel (c), $\Omega_0 = 0.1, 0.2, 0.3, 0.4$ and 1.0. From the left to the right of the first peak location of panel (d), $\Omega_0 = 1.0, 0.5, 0.3, 0.2$ and 0.1. For open models (panel (d)), we assume the flat potential power spectrum for the shape of the initial power spectrum which is predicted by open inflationary models (Lyth and Stwart 1990, Ratra and Peebles 1995).

Secondly we show dependence of the location of peaks on cosmological parameters. The peak location is determined by the projection of the sound horizon at last scattering surface into observed angle. Because the sound horizon is a weak function of $\Omega_B h^2$, the peak location is almost independent on Ω_B and h. On the other hand, it strongly depends

on Ω_0. The low density universe has longer age. Therefore the last scattering surface is further than the one of the high density universe. This effect makes the sound horizon correspond to the smaller angular scale. Moreover, in the open universe, there is a geodesic effect which causes the same but bigger effect (Kamionkowski, Sugiyama & Spergel 1994). As is shown in figure 2 (c) and (d), peak locations are placed in order of $\Omega_0 = 1$, Λ and open models from large to small scales.

Finally, we would like to mention the peculiar behaviour of CMB power spectrum on large scales (small ℓ) for low density models which is shown in figure 2 (c) and (d). On large scales, the SW effect which produces the flat tail makes the dominant contribution for the $\Omega_0 = 1$ model. On the other hand, there is the late ISW contribution for low density models. As for the Λ model, the ISW effect dominates on large scales. Because this effect has the damping caused by the finite thickness of *gravitational last scattering* (Hu & Sugiyama 1994), it is only significant on very large scales (small ℓ). For the open model, there is another effect, i.e., the cutoff at the curvature scale (Wilson 1983, Sugiyama & Silk 1994, Hu & Sugiyama 1994). Because the fluctuations outside the curvature scale do not contribute on C_ℓ, there is a cutoff on very large scales.

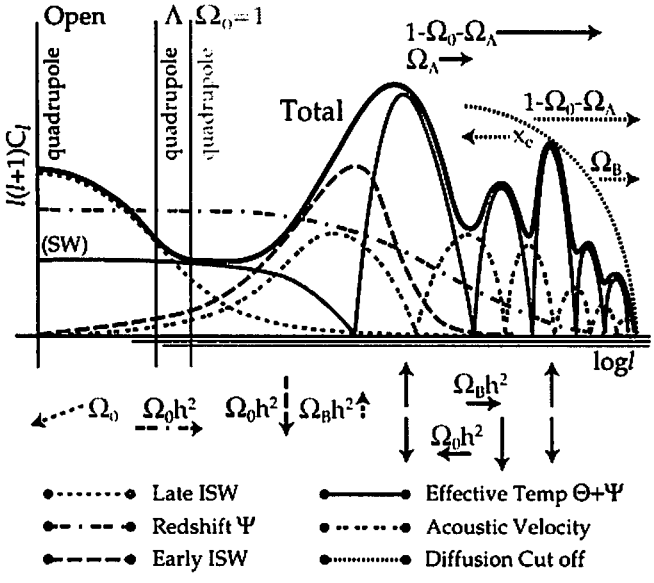


FIG. 3. Individual effects on C_ℓ and their dependence on cosmological parameters. This figure is reprinted from Hu et al. (1995).

VI. CONSTRAINTS FROM OBSERVATIONS

It is found that the power spectrum of CMB anisotropies C_ℓ contains rich informations in its shape. On very large scales ($\ell < 50$), C_ℓ's are sensitive to the initial condition of perturbations, Ω_0 and Λ. In fact, we did give an constraint on Λ by using two year COBE DMR data (Bunn & Sugiyama 1995). Employing Bayesian analysis, we got $\Lambda < 0.8$. On scales $50 < \ell < 200$, contributions of the early ISW effect on CMB anisotropies are large. Therefore the shape of C_ℓ in this region is sensitive to $\Omega_0 h^2$. Moreover, if the isocurvature perturbations are considered, the first peak location is around $\ell = 100$. In this case, since the second peak (around $\ell = 300$) is usually higher than the first one, we can easily distinguish it. Finally, if sufficient reionization happens after recombination, a new peak is generated at the place corresponding to the horizon scale of new last scattering surface which is typically $\ell \sim 50$ (Sugiyama, Silk & Vittorio 1993). If there is no peak smaller than this scale, it might be the evidence of the late time reionization. On scales larger than $\ell = 100$, the most prominent figure of C_ℓ which we expect is the existence of high peaks. The location of the first peak will reveal the geometry of the universe. The height and location of peaks tells us about $\Omega_B h^2$ and $\Omega_0 h^2$. The damping scale is mostly sensitive to Ω_B. In near future, we will be able to determine these cosmological parameters out of observations of CMB anisotropies. In figure 3, we summarize all dependence of C_ℓ on various cosmological parameters. In Figure 4, the numerical values of C_ℓ's for several cosmological models with the

current status of observations are shown. We might be able to get a rough idea how strong present constraints on model parameters from observations.

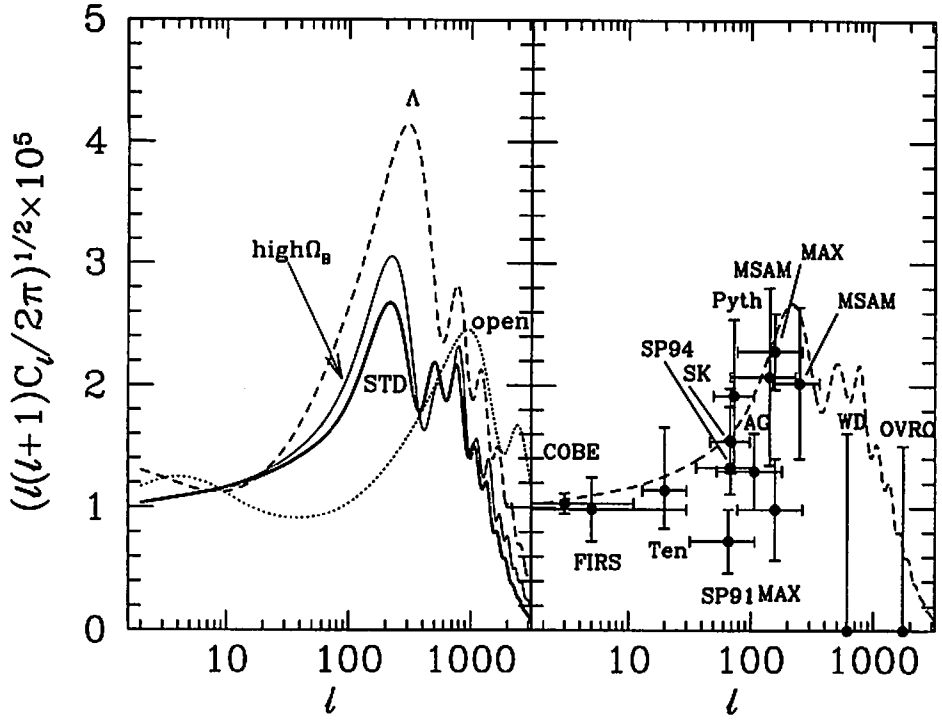


FIG. 4. Left: Square root of the CMB power spectrum as function of multipole ℓ. STD is a CDM model with $\Omega_0 = 1, h = 0.5$ and $\Omega_B = 0.05$. Λ (dashed) and open (dotted) are models with $\Omega_0 = 0.1$. High Ω_B (thin solid) is the one with $\Omega_0 = 1$ and $\Omega_B = 0.1$. Right: several recent observations. The STD model is also plotted.

VII. CONCLUDING REMARKS

We developed simple analytic treatment of the evolution of CMB anisotropies and a toy model based on the analogy of a harmonic oscillator. These are powerful tools to understand the dependence of C_ℓ on cosmological parameters intuitively. We show such dependence in this paper. Several exciting new observations of CMB anisotropies are being planned. New satellite missions by Europe and USA and long duration flight balloon experiments are discussed. If once we get all (or nearly all) sky maps with fine resolution (~ 10arcmin.), we will be able to reconstruct the CMB anisotropy spectrum C_ℓ up to the damping scale. We will soon see the nature of our universe.

Acknowledgments

This work is based on collaborations with W. Hu. The author is deeply grateful to him and other CMB gangs in Berkeley, E. Bunn, L. Cayon, D. Scott, M. White and especially professor J. Silk.

Bolte, M. & Hogan, C.J. 1995, *Nature* **376**, 399.
Bunn, E., & Sugiyama, N. 1995, *Astrophys. J.*, **446**, 49.

Djorgovski, S. et al. 1995, *Astrophys. J. Lett.*, **438**, L13.
Freedman, W. et al. 1994, *Nature* **371**, 757.
Gardner, J. P., Cowie, L. L., & Wainscoat, R. J. 1993, *Astrophys. J. Lett.*, **415**, L9.
Hata, N. et al. 1995, *Phys. Rev. Lett.*, **75**, 3977.
Hu, W., & Sugiyama, N. 1994, *Phys. Rev.* **D50**, 627.
Hu, W., & Sugiyama, N. 1995a, *Astrophys. J.*, **444**,489.
Hu, W., & Sugiyama, N. 1995b, *Phys. Rev.* **D51**, 2599.
Hu, W., Sugiyama, N., & Silk, J. 1995, *submitted to Nature.*
Jones, M. et al., 1993, *Nature* **365**, 320.
Kamionkowski, M., Spergel, D. N., & Sugiyama, N. 1994, *Astrophys. J. Lett.*, **426**, L57.
Kochanek, C. S. 1995, *Astro-ph* 9510077.
Lyth, D., & Stewart, E.D. 1990, *Phys. Lett.*, **B252**, 336.
Mather, J.C. et al. 1994, *Astrophys. J.*, **420**,439.
Peacock, J. A., & Dodds, S. J. 1994, *M.N.R.A.S* **267**, 1020.
Ratra, B., & Peebles, P.J.E. 1995, *Phys. Rev.* **D52**, 1837.
Riess, A. G., Press, W. H., & Kirshner, R. P. 1995, *Astrophys. J. Lett.*, **438**, L17.
Sachs, R. K., & Wolfe, A. M. 1967, *Astrophys. J.*, **147**, 73.
Sangaila, A., Cowie, L. L. and Hogan, C. J. 1994, *Nature* **368**, 599.
Silk, J. 1968, *Astrophys. J.*, **151**, 459.
Smoot, G. et al. 1992, *Astrophys. J. Lett.*, **396**, L1.
Strauss, M., & Willick, J. A. 1995, *Physics Reports*, **261**,271.
Sugiyama, N. 1995, *Astrophys. J. Suppl.*, **100**,281.
Sugiyama, N., & Silk, J., 1994, *Phys. Rev. Lett.*, **73**, 509.
Sugiyama, N., Silk, J., & Vittorio, N. 1993, *Astrophys. J. Lett.*, **419**, L1.
Sunyaev, R.A., & Zeldovich, Ya.B. 1970, *Ap. Space Sci.*, **9**, 378.
Tanvir, N.R. et al. 1995, *Nature* **377**, 27.
Walker, T. P. et al. 1991, *Astrophys. J.*, **376**, 51.
Weinberg, S. 1972, *Gravitaion and Cosmology; Wiley.*
White, M., Scott, D., & Silk, J. 1993, *ARA& A*, **32**, 319.
White, S. D. M. et al. 1993, *Nature* **366**, 429.
Wilson, M. L. 1983, *Astrophys. J.*, **273**, 2.
Zeldovich, Ya.B., & Sunyaev, R.A. 1969, *Ap. Space Sci.*, **4**, 301.

Department of Physics
Kyoto University

Evolution of Cosmological Perturbations during Reheating

Takashi Hamazaki[1] and Hideo Kodama[2]

[1] *Department of Physics, Kyoto University*
Kyoto 606-01, Japan
[2] *Yukawa Institute for Theoretical Physics, Kyoto University*
Kyoto 606-01, Japan

We discuss whether the Bardeen's quantity $\Phi - \Upsilon$ is conserved on superhorizon scales during reheating. Since the energy transfer from the inflaton to radiation is described as a sum of many local particle conversion reactions, we average the rapidly oscillating inflaton over many periods, and express its dynamical state by the adiabatic invariants. On the basis of the WKB approximation, we develop the gauge invariant perturbation theory of the multicomponent system consisting of the scalar field and radiation, and with a help of it, investigate the cosmological perturbations during reheating. We conclude that the energy transfer cannot generate superhorizon scale perturbations irrespectively of the detail of particle conversion reactions and the duration of reheating.

I. INTRODUCTION

In the inflationary paradigm, cosmological large scale structures, such as galaxies, clusters of galaxies are thought to be generated from seed density perturbations caused by quantum fluctuations of the inflaton. In the slow rolling phase, the behavior of the subhorizon perturbations of the inflaton is well described in terms of the quantum field theory in the (unperturbed) de Sitter space: The zero-point quantum fluctuations of the inflaton field generate the density perturbations $\delta\rho = V_\phi(\phi)\delta\phi$. On subhorizon scales $k/aH \gg 1$, cosmological perturbations show an oscillatory behavior, while as scales of the perturbations are stretched out into superhorizon scales $k/aH \ll 1$ (the initial Hubble-radius crossing; IH) by the expansion of the Universe, their amplitudes converge (freeze) into certain constant values. On superhorizon scales, the matter fluctuations interact with the metric perturbations such as the spatial curvature perturbation Φ, and the gravitational potential Ψ. Small metric perturbations are generated by the matter fluctuations according to the Einstein constraint equations, and they evolve according to the dynamical Einstein equations together. In the oscillating phase when the Universe Friedmann-expands, the superhorizon cosmological perturbations behave like those of a dust as stated later, and they enter into the horizon (the final Hubble-radius crossing; FH) as the adiabatic fluctuation with a scale invariant (Harrison Zel'dovich) spectrum, which is suited to explain very large scale structures of the Universe.

To estimate the density perturbation, Bardeen et al. [1] proposed to use the fact that the Bardeen's parameter ζ which is written as $\Phi - \Upsilon = \Phi - (aH/k)V$ in the notation of Kodama and Sasaki [5] is approximately conserved on superhorizon scales. In the radiation dominated stage, $\Phi = -[(n+1)/2(n-2)]\Upsilon \propto$ const for the growing mode, and $\Phi = \Upsilon \propto a^{-3(n-1)/2}$ for the decaying mode. If $\Phi - \Upsilon$ is conserved, the spatial curvature perturbations Φ at FH are given by

$$\Phi_{FH} = \frac{n+1}{3n-3}(\Phi-\Upsilon)_{FH} = \frac{n+1}{3n-3}(\Phi-\Upsilon)_{IH} = -\frac{n+1}{3n-3}\Upsilon_{IH} = -\frac{n+1}{3n-3}\frac{HX}{\dot\phi}_{IH}, \tag{1.1}$$

where X is the gauge invariant variable corresponding to $\delta\phi$ on the Newtonian slicing. [5]

Although this conservation law is used in almost all quantitative discussions based on the inflationary scenario, no exact proof has been given so far in a generic context as far as authors know. In the restricted case, there are some

discussions as follows. If reheating occurs instantaneously, in other words, it can be regarded as a phase transition from the coherently oscillating scalar field to radiation, this problem is reduced to that of the metric junction between the two different phases, ϕ, and radiation [2]. In this case, the conservation law can be shown almost trivially. However in realistic scenarios of reheating in which reheating occurs rather slowly, the inflaton and radiation coexist for a long time and the effect of the multi-component property of the system cannot be neglected in general. So we would like to investigate whether $\Phi - \Upsilon$ is conserved between IH and FH by treating the system as the multicomponent system consisting of a coherently oscillating scalar field and radiation generated from it.

Our strategy in the description of reheating is as follows. As for the coherently oscillating scalar field, two kinds of dynamical processes with different time scales occur simaltaneously. One is particle-physical processes containing the coherent oscillation around the potential minimum with a time scale m^{-1} and particle conversion reactions from the inflaton to particles which are constituents of radiation. The other is its gravitational evolution with a large time scale through its interaction with metric perturbations. We pick up variables which change in the largest time scale by averaging, and describe reheating by only them. The main point of our discussion is whether the perturbations of the energy transfer controled by local quantities generate superhorizon entropy perturbations between the scalar field and radiation.

This paper is organized as follows. In section 2, we construct the gauge invariant formalism of the coherently oscillating scalar field. Following the strategy proposed in this section, we show how to pick up a slowly changing part relevant to the gravitational evolution and the secular energy transfer from the rapidly changing ϕ in terms of the WKB approximation. We stress the analogy between the dynamical behavior of it and that of the perfect fluid. In section 3, we explicitly give the equations which describe reheating, and show that $\Phi - \Upsilon$ is conserved on superhorizon scales in most cases. Section 4 is devoted to discussions.

Thoughout this paper, we use the notations of the gauge invariant perturbation theory adopted by Kodama and Sasaki [5]. We imply the spatial dimension by n, which is 3 in our Universe. As the corection factors from the spatial curvature K, we use

$$I_K = 1 + \frac{K}{a^2 H^2} \quad C_K = 1 - \frac{nK}{k^2} \quad . \tag{1.2}$$

We imply the differentiation with respect to η by $'$, and f differentiated with respect to S by f_S, etc.

II. PERTURBATIONS IN THE COHERENTLY OSCILLATING SCALAR FIELD

During reheating, the inflaton field ϕ oscillates around the potential minimum rapidly. If we try to describe it in so great detail that we could see each particle conversion reaction minutely, the quantum theoretical uncertainty of energy (For example, the energy of radiation) becomes too large ($\Delta E \Delta t \geq \hbar$) to describe radiation in terms of the classical perfect fluid picture. Therefore we will give the energy transfer rate from the inflaton to radiation as a mean value. Each particle conversion reaction occurs non-adiabatically, but the energy transfer as the whole progresses adiabatically. This fact is favorable to such an averaging. In the treatment of the rapidly oscillating inflaton, we are also required to average over a time scale $\gg 1/m$. In the remarkable paper [4], authors have done such averaging partially in the action in order to get the nonrelativistic action from the general relativistic one of the axion field. We would like to do further thoroughly in the energy momentum tensor with doing the WKB approximation simaltaneously.

The energy momentum tensor of the scalar field is

$$T_\phi{}^\mu{}_\nu = \nabla^\mu \phi \nabla_\nu \phi - [\frac{1}{2} \nabla^\lambda \phi \nabla_\lambda \phi + U(\phi)] \delta^\mu{}_\nu \quad . \tag{2.1}$$

In the oscillating phase, we decompose the scalar field ϕ into the amplitude σ and the phase S (which corresponds to the Hamilton's principal function) as

$$\phi = \sigma f(S) \quad f(S) = f(S + 2\pi) \quad , \tag{2.2}$$

where $f(x)$ is a general periodic function with a period 2π. By substituting the equation above, we get

$$\begin{aligned} \nabla^\mu \phi \nabla_\nu \phi = {} & \nabla^\mu \sigma \nabla_\nu \sigma [f(S)]^2 + \sigma \nabla^\mu \sigma \nabla_\nu S [f(S) f_S(S)] \\ & + \sigma \nabla^\mu S \nabla_\nu \sigma [f(S) f_S(S)] + \sigma^2 \nabla^\mu S \nabla_\nu S [f_S(S)]^2 \quad . \end{aligned} \tag{2.3}$$

In general, an arbitrary periodic function $f(x)$ can be written in the form of Fourier series such as

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n e^{inx} \quad \left(\begin{array}{l} c_0; \text{real} \\ c_n = c_{-n}^*(n = 1, 2, \cdots) \end{array} \right) \quad . \tag{2.4}$$

We introduce an average $< A(x) >$ by

$$< A(x) > := \frac{1}{2\pi} \int_0^{2\pi} dx A(x) \quad . \tag{2.5}$$

By taking an average, we drop the rapidly oscillating part and keep the adiabatically changing part. ¿From the above definition, we get

$$\begin{aligned} < [f(x)]^2 > &= \sum_{n=-\infty}^{+\infty} |c_n|^2 =: a \quad , \\ < [f_x(x)]^2 > &= \sum_{n=-\infty}^{+\infty} n^2 |c_n|^2 =: b \quad , \\ < f(x) f_x(x) > &= 0 \quad . \end{aligned} \tag{2.6}$$

As for a quadratic form appearing in the energy momentum tensor, we obtain

$$< \nabla^\mu \phi \nabla_\nu \phi > = a \nabla^\mu \sigma \nabla_\nu \sigma + b \sigma^2 \nabla^\mu S \nabla_\nu S \quad . \tag{2.7}$$

By the (covariant) WKB approximation,

$$\nabla^\mu \sigma \ll \sigma \nabla^\mu S \quad , \tag{2.8}$$

which implies that a change of the amplitude σ during one period $1/\nabla^\mu S$ is negligibly small, we get

$$< \nabla^\mu \phi \nabla_\nu \phi >_{WKB} = b \sigma^2 \nabla^\mu S \nabla_\nu S \quad . \tag{2.9}$$

So the energy momentum tensor which is averaged, and WKB-approximated is

$$\begin{aligned} < T_\phi{}^\mu{}_\nu >_{WKB} &= b \sigma^2 \nabla^\mu S \nabla_\nu S - [\frac{1}{2} b \sigma^2 \nabla^\lambda S \nabla_\lambda S + V(\sigma)] \delta^\mu{}_\nu \\ V(\sigma) &= < U(\phi = \sigma f(S)) > \quad , \end{aligned} \tag{2.10}$$

and we supress $< \cdots >_{WKB}$ from now on.

The divergence of $T_\phi{}^\mu{}_\nu$ is

$$\begin{aligned} \nabla_\mu T_\phi{}^\mu{}_\nu &= b[2\sigma \nabla_\mu \sigma \nabla^\mu S + \sigma^2 \Box S] \nabla_\nu S \\ &\quad - [b \sigma \nabla_\mu S \nabla^\mu S + V_\sigma(\sigma)] \nabla_\nu \sigma \quad . \end{aligned} \tag{2.11}$$

We can get the Hamilton Jacobi equation by putting the second bracket on the right hand side of Eq. (2.11) at zero. Under such a semi-classical (WKB) approximation, it is natural to think that $\nabla_\nu S$ represents 4-momentum along which the energy momentum is transferred from the scalar field to other components. Therefore we can interpret that

$$T_\phi{}^\mu{}_{\nu;\mu} = Q_{\phi\,\nu} \quad , \tag{2.12}$$

implies that

$$\nabla^\mu [\sigma^2 \nabla_\mu S] = \frac{1}{b} \frac{Q_\phi^\nu \nabla_\nu S}{\nabla^\nu S \nabla_\nu S} \quad , \tag{2.13}$$

$$b \sigma \nabla^\mu S \nabla_\mu S + V_\sigma(\sigma) = 0 \quad , \tag{2.14}$$

where

$$Q_{\phi\,\nu} \propto \nabla_\nu S \quad . \tag{2.15}$$

It is to be remembered that we have decomposed one degree of freedom ϕ into two degrees of freedom such as the adiabatically changing σ and the rapidly increasing S, therefore the latters are not independent. So we must give

one constraint which determines (or 'gauge fix') how to decompose ϕ into σ and S and which restores a number of independent degrees of freedom from 2 to 1. We do this by Eq.(2.14). Eq.(2.13) is a substantial equation of motion and with $Q_{\phi\ \nu} = 0$, it reduces to the equation of continuity of the perfect fluid.

By comparing Eq. (2.10) with the general form of the energy momentum tensor as

$$T_{\alpha}{}^{\mu}{}_{\nu} = \rho_{\alpha} u_{\alpha}^{\mu} u_{\alpha\nu} + P_{\alpha}(\delta^{\mu}{}_{\nu} + u_{\alpha}^{\mu} u_{\alpha\nu}) + \pi_{\alpha}{}^{\mu}{}_{\nu} \quad , \tag{2.16}$$

we get correspondences as follows,

$$\begin{aligned} \rho_{\phi} &= -\frac{1}{2} b\sigma^2 \nabla^{\lambda} S \nabla_{\lambda} S + V(\sigma) \quad , \\ P_{\phi} &= -\frac{1}{2} b\sigma^2 \nabla^{\lambda} S \nabla_{\lambda} S - V(\sigma) \quad , \\ u_{\phi}^{\mu} &= \frac{\nabla^{\mu} S}{\sqrt{-\nabla^{\lambda} S \nabla_{\lambda} S}} \quad , \\ \pi_{\phi}{}^{\mu}{}_{\nu} &= 0 \quad . \end{aligned} \tag{2.17}$$

which implies that the coherently oscillating scalar field behaves in the same manner as the perfect fluid. It seems to be the well known fact, but we would like to point out that it has been derived without assuming the spatial homogeneity of the scalar field, so results obtained so far can be applied to not only background quantities but also perturbed ones.

We can describe the dynamical behavior of our coherently oscillating scalar field ϕ by using the gauge invariant variables Δ_{ϕ}, V_{ϕ} which is the same as those describe the dynamical behavior of the perfect fluid.

In particular, in case that the potential $V(\sigma)$ is a monomial of σ,

$$V(\sigma) = g\sigma^m \quad , \tag{2.18}$$

Eq.(2.14) gives the equation of state as,

$$w_{\phi} = c_{\phi}^2 = \frac{m-2}{m+2} \quad . \tag{2.19}$$

III. PERTURBATIONS DURING THE REHEATING

We will discuss the evolution of the cosmological perturbations during reheating. We consider the system consisting of the coherently oscillating scalar field and radiation generated from the former.

When the anisotropic stress perturbation Π can be neglected,

$$\mathcal{D}(\Phi - \Upsilon) = -\frac{K}{a^2 H^2}\Upsilon - \frac{c_S^2}{1+w}\frac{2}{n}\left(\frac{k}{aH}\right)^2 \frac{C_K}{I_K}\Phi - \frac{w}{1+w}\Gamma \quad , \tag{3.1}$$

where

$$\mathcal{D} := a\frac{d}{da} \quad . \tag{3.2}$$

$$\Upsilon := \frac{aH}{k} V \quad . \tag{3.3}$$

So we can conclude that $\Phi - \Upsilon$ is conserved on superhorizon scale $k/aH \ll 1$ if $\Phi - \Upsilon$ and Φ are of the same order, and if the spatial curvature K and the entropy perturbation Γ can be neglected. The term from the spatial curvature K is practically zero because of the inflationary expansion. But there are problems for Γ. If reheating can be regarded as a phase transition from a (coherent) scalar field to radiation, $\Phi - \Upsilon$ is conserved on superhorizon scales $k/aH \ll 1$ because the entropy perturbation Γ which is the intrinsic entropy perturbation for the coherently oscillating scalar field before the transition and for radiation after the transition, is exactly zero. But the recent studies of reheating suggest that the coherently oscillating scalar field and radiation coexist for a fairly long time. In this case, Γ is not zero. In general, the entropy perturbation Γ is decomposed into two parts.

$$P\Gamma = P\Gamma_{int} + P\Gamma_{rel} \quad , \tag{3.4}$$

$$P\Gamma_{int} = \sum_{\alpha} P_{\alpha}\Gamma_{\alpha} \quad , \tag{3.5}$$

$$P\Gamma_{rel} = \frac{1}{2}\sum_{\alpha}\sum_{\beta}\frac{h_{\alpha}h_{\beta}}{h}(1-q_{\alpha})(1-q_{\beta})(c_{\alpha}^2-c_{\beta}^2)\times\left[\frac{\Delta_{c\alpha}}{(1+w_{\alpha})(1-q_{\alpha})}-\frac{\Delta_{c\beta}}{(1+w_{\beta})(1-q_{\beta})}\right]$$
$$= \frac{1}{2}\sum_{\alpha}\sum_{\beta}\frac{h_{\alpha}h_{\beta}}{h}(c_{\alpha}^2-c_{\beta}^2)S_{\alpha\beta}+\frac{1}{1+w}\frac{\Delta}{nH}\sum_{\alpha}c_{\alpha}^2Q_{\alpha} \quad . \tag{3.6}$$

One is a part coming from the intrinsic entropy perturbation of each component, and the other is a part coming from the difference of the dynamical behavior of components. (Conventionally the latter is often called simply 'entropy perturbation' by contrast with the adiabatic perturbation in which all the components fluctuate in the same manner [3]) If the relative entropy perturbation Γ_{rel} is generated between the scalar field and radiation, in other words, radiation acquires another kind of perturbations than that which the scalar field has had with itself during reheating, $\Gamma = \Gamma_{rel}$ is non-zero, and the conservation law of $\Phi - \Upsilon$ is broken by the Γ term. We have to investigate whether such an entropy perturbation is generated during reheating or not.

For simplicity, we assume that $w_{\phi} = c_{\phi}^2 = \text{const}$ which is equivalent to the condition that $V(\sigma)$ is written as a monomial of σ. We describe the dynamical state of the system by 4 variables such as the density contrast and the velocity of the total system, Φ, V, the difference of the density contrasts, the velocities between the scalar field and radiation, Z, Ω.

$$Z = \frac{\rho_{\phi}\rho_r}{\rho^2}S_{\phi r} \quad , \quad \Omega = \frac{\rho_{\phi}\rho_r}{\rho^2}\frac{aH}{k}V_{\phi r} \quad , \tag{3.7}$$

where

$$S_{\phi r} = \frac{\Delta_{c\phi}}{1+w_{\phi}} - \frac{\Delta_{cr}}{1+w_r} \quad , \tag{3.8}$$

$$V_{\phi r} = V_{\phi} - V_r \quad . \tag{3.9}$$

The equation for the perturbed variables for the total system are

$$\mathcal{D}\Phi + (n-2)\Phi = -\frac{n}{2}I_K(1+w)\Upsilon \quad , \tag{3.10}$$

$$\mathcal{D}\Upsilon + \frac{n}{2}I_K(1+w)\Upsilon - \frac{K}{a^2H^2}\Upsilon = \{-(n-2)+\frac{2}{n}(\frac{k}{aH})^2\frac{C_K}{I_K}\frac{1}{1+w}\frac{1}{h}(h_{\phi}w_{\phi}+h_r\frac{1}{n})\}\Phi$$
$$+\frac{n+1}{n}\frac{1+w_{\phi}}{(1+w)^2}(w_{\phi}-\frac{1}{n})Z \quad . \tag{3.11}$$

The equation of the perturbed variables representing the difference of the dynamical behavior of ϕ and radiation are

$$\mathcal{D}Z + n(\frac{1}{n}+w_{\phi}-2w)Z = -(\frac{k}{aH})^2\Omega + \frac{n}{n+1}\frac{Qh}{H\rho^2}\frac{1}{1+w_{\phi}}E_{c\phi} - \frac{2Q}{nH\rho}\frac{1}{1+w_{\phi}}(1+\frac{n}{n+1}w_{\phi})\frac{C_K}{I_K}(\frac{k}{aH})^2\Phi \quad , \tag{3.12}$$

$$\mathcal{D}\Omega + [\frac{n}{2}(1+w)I_K - \frac{K}{a^2H^2} + n(w_{\phi}-2w) + (1-nw_{\phi})\frac{h_r}{h} - \frac{Q}{Hh}\{1+(w_{\phi}-\frac{1}{n}) + \frac{n+1}{n}\frac{1}{1+w_{\phi}}\frac{\rho_r}{\rho_{\phi}}\}]\Omega$$
$$= \frac{1}{h}[\frac{n+1}{n}w_{\phi}\rho_r + \frac{1}{n}(1+w_{\phi})\rho_{\phi}]Z + \frac{2}{n}(w_{\phi}-\frac{1}{n})\frac{C_K}{I_K}(\frac{k}{aH})^2\frac{\rho_{\phi}\rho_r}{\rho^2}\frac{\Phi}{1+w} \quad , \tag{3.13}$$

where Q, $E_{c\phi}$ are the background and the perturbed variable of the energy transfer from ϕ to radiation, respectively. Q is negative when the energy is transferred from ϕ to radiation.

For understanding the evolutions of Φ, Υ, Z, Ω from Eq. (3.10), (3.11), (3.12), (3.13), it is necessary to determine the initial conditions of Z, Ω, and to express $E_{c\phi}$ as linear combination of Φ, Υ, Z, Ω. We assume that at the initial moment of reheating

$$Z = 0 \quad \Omega = 0 \quad . \tag{3.14}$$

The reason is as follows. As for Z, since radiation does not exist at the start of reheating, $Z = 0$ yields from the definition of Z, Eq. (3.7). And as for Ω, since radiation is generated relative to the rest frame of the scalar field, $\Omega = 0$ yields.

¿From Eq. (3.12), (3.13), (3.14), we can conclude

$$Z \sim \Omega \sim (\frac{k}{aH})^2\Phi \quad , \tag{3.15}$$

if $E_{c\phi}$ is of the same order as Z, Ω, or at most of order $(k/aH)^2\Phi$.

Because of the general covariance, the energy transfer rate $\tilde{Q}$ which can be obtained after the averaging as stated in section 1, must be a function of scalar quantities which is represented by $\tilde{\theta}$ below.

$$\tilde{Q} = F(\tilde{\theta}) \quad , \tag{3.16}$$

$$\tilde{\theta} = \theta(1 + \delta_\theta Y) \quad . \tag{3.17}$$

In this case,

$$E_{c\phi} = \frac{F_\theta(\theta)\theta}{F(\theta)}[\delta_\theta - \frac{1}{k}\frac{\theta'}{\theta}(v - B)] \quad . \tag{3.18}$$

If each particle creation reaction is reduced to an ideally local process which does not suffer from the effect of the spacetime curvature at all, we can identify the energy tansfer rate in the local inertial frame with $\tilde{u}^\mu_\phi$ as the time axis, with the energy transfer rate calculated by the quantum field theory in the Minkowski spacetime with considering the scalar field oscillates (spatially) homogeneously, or quantum mechanically speaking, is in the coherent state in which macroscopically large number of particles occupy the zero-momentum mode, and all other finite-momentum modes are completely empty. In this case, the energy transfer rate $\tilde{Q}$ cannot be a function of (scalar) variables other than the oscillation amplitude $\tilde{\sigma}$, up to the coupling constant between the scalar field ϕ and particle species converted. By assuming that $\tilde{\theta} = \tilde{\sigma}$, we obtain

$$E_{c\phi} = \frac{F_\sigma(\sigma)\sigma}{F(\sigma)}\frac{1}{2}\frac{1 - c_\phi^2}{1 + w_\phi}\Delta_{c\phi} \quad . \tag{3.19}$$

Since

$$\begin{aligned}\frac{E_{c\phi}}{1 + w_\phi} \sim \frac{\Delta_{c\phi}}{1 + w_\phi} &= \frac{\Delta}{1 + w} + \frac{h_r}{h}S_{\phi r} \\ &= \frac{1}{1 + w}\frac{2}{n}\frac{C_K}{I_K}(\frac{k}{aH})^2\Phi + \frac{n + 1}{n}\frac{1}{1 + w}\frac{\rho}{\rho_\phi}Z \\ &= c_1(\frac{k}{aH})^2\Phi + c_2 Z \quad ,\end{aligned} \tag{3.20}$$

where c_1, c_2 are of order 1, we see that $E_{c\phi}$ which appears in Eq. (3.12) cannot enhance Z.

Further we consider the case that the particle conversion reaction is affected by the spacetime curvature, but that the time scale of it is sufficiantly small compared with the gravitational evolution of cosmological perturbations, and the adiabatic energy transfer as the whole. For example, in a recently proposed reheating scenario [6], the parametric resonance plays an important role, and it is thought to be affected by the expansion of the Universe. So we try to assume that $\tilde{Q}$ depends not only upon $\tilde{\sigma}$ but also upon $\tilde{\nabla}_\mu \tilde{u}^\mu_\phi$, which represents the expansion rate relative to the local inertial frame with $\tilde{u}^\mu_\phi$ as the time axis. When $\tilde{Q}$ is written as $F(\tilde{\sigma}, \tilde{\nabla}_\mu \tilde{u}^\mu_\phi)$, the background part of the energy transfer is

$$Q = F(\sigma, nH) \quad , \tag{3.21}$$

and for the perturbation of it, the terms as

$$\begin{aligned}\frac{F_{nH}\, nH}{F}(\frac{1}{n}\frac{k}{aH}V_\phi - \frac{K}{a^2H^2}\Upsilon) &= \frac{F_{nH}\, nH}{F}[\frac{1}{n}(\frac{k}{aH})^2\Upsilon + \frac{n + 1}{n^2}(\frac{k}{aH})^2\frac{\rho^2}{\rho_\phi h}\Omega - \frac{K}{a^2H^2}\Upsilon] \\ &= d_1(\frac{k}{aH})^2\Upsilon + d_2(\frac{k}{aH})^2\Omega + d_3\frac{K}{a^2H^2}\Upsilon \quad ,\end{aligned} \tag{3.22}$$

where d_1, d_2, d_3 is of order 1, is added to the right side hand of Eq. (3.19). Even in such a case, $E_{c\phi}$ cannot enhance Z. Since we derive that Z is at most of order $(k/aH)^2\Phi$, we can conclude that $\Phi - \Upsilon$ is conserved during reheating, from Eq. (3.10), (3.11).

IV. DISCUSSION

We can conclude that for the coherently scalar field with rapidly oscillating part removed, not only the background but also the perturbation have the same behavior as those of the perfect fluid, and the energy transfer which is governed by local processes cannot generate the large scale (superhorizon scale) coherence, so $\Phi - \Upsilon$ is conserved during reheating.

But we would like to point out that there are interesting exceptions. If $\tilde{Q}$ also depends upon the scalar quantities which have nothing to do with ϕ and radiation, for example, some other scalar field which does not dominate the energy of the Universe, new type of perturbations other than those the inflaton have had with itself, can be added to radiation during reheating. In most particle physics model, many parameters such as particle's masses, and coupling constants are often functions of the vacuum expectation values of the scalar fields which trigger the spontaneous symmetry breaking, such as Higgs fields. And not only the inflaton but also such vacuum expectation values fluctuate owing to the de Sitter expansion. (the fluctuation of the latter is well known to be the Nambu Goldstone mode.) So in the case that ϕ's mass or vertices which govern the particle conversion reactions are functions of such quantities, the conservation of $\Phi - \Upsilon$ does not hold. Such possibilities have been used as a mechanism to generate the baryon isocurvature perturbation. [7]

ACKNOWLEDGEMENTS

T. H. would like to thank Prof. H. Sato for continuous encouragements. T. H. would like to thank Y. Nambu, M. Sasaki, E. Stewart, J. Yokoyama for fruitful discussions. This work was supported by Grant-in-Aid for Scientific Reserch of the Ministry of Education, Science, and Culture of Japan (05640340)

---

[1] J. M. Bardeen, P. J. Steinhardt and M. S. Turner, Phys. Rev. **28**, 679, (1983).
[2] N. Deruelle and V. F. Mukhanov, Phys. Rev. **D52**, 5549 (1995).
[3] V. F. Mukhanov, H. A. Feldman and R. H. Brandenberger, Physics Reports **215**, 203 (1992)
[4] Y. Nambu and M.Sasaki, Phys. Rev. **D42**, 3918 (1990).
[5] H. Kodama and M. Sasaki, Prog. Theor. Phys. Suppl. **78**, 1 (1984).
[6] L. Kofman, A. Linde and A. A. Starobinsky, Phys. Rev. Lett. **73**, 3195 (1994). Y. Shtanov, J. Traschen and R. Brandenberger, Phys. Rev. **D51**, 5438 (1995). M. Yoshimura, Prog. Theor. Phys. **94**, 873, (1995)
[7] M. Yoshimura, Phys. Rev. Lett. **51**, 439, (1983). M. Sasaki and J. Yokoyama, Phys. Rev. **D44**, 970 (1991).

Theory of Force-Free Black Hole Magnetosphere

Toshio Uchida

Astronomical Institute, Faculty of Science, Tohoku University, Sendai 980, Japan

1. Introduction

In the relativistic magnetohydrodynamics (MHD), the force-free approximation is a powefull approximation when the electromagnetic energy is much larger than both the rest mass energy and the thermal energy of plasmas. In the magnetosphere of pulsars this approximation is considered farily good in its greater part, and probably also in the magnetosphere around the black hole. Further, it seems that this approximation is superior in the sence that the dynamics of the degenerate electromagnetic fields appeares its most transparent and simplest way. However most of foregoing works on the force-free approximation concerned only with the stationary and axisymmetric configuration and the systematic method to deal with the foce-free electromagetic fields is still absent. Further since the stationary and axisymmetric configuration is not so adequate to see the dynamics of the force-free electromagnetic fields, understanding on the dynamics of the force-free electromagnetic field is very poor untill now in my opinion. In the following, we will give a general method to deal with the force-free electromagnetic field and clarify its dynamics. Then we apply it to the black hole magnetosphere.

2. Basic Equation

The Euler equation for the relativistic magnetohydrodynamical plasmas is

$$mnu^{\nu}\nabla_{\nu}u_{\mu} = -\gamma_{\mu}{}^{\nu}\nabla_{\nu}p + F_{\mu\nu}J^{\nu}, \tag{1}$$

where n is the particle number density of the plasma, u^{ν} is the macroscopic four-velocity of the plasma, m is the rest mass of the plasma constituent particles, p is the thermal pressure, $\gamma_{\mu}{}^{\nu}$ is the projection tensor on the direction orthogonal to the velocity, $F_{\mu\nu}$ is the electromagnetic field and J^{μ} is the four-current. When the inertial force and the pressure force are much smaller than the Lorentz force, the left-hand side and the first term of the righ-thand side of the above equation can be neglected. This is the force-free approximation. Thus together with Maxwell equation, the basic equations for the force-free electromagnetic field are

$$\nabla_{\lambda}F_{\mu\nu} + \nabla_{\mu}F_{\nu\lambda} + \nabla_{\nu}F_{\lambda\mu} = 0, \tag{2}$$

$$\nabla_{\nu}F^{\mu\nu} = 4\pi J^{\mu}, \tag{3}$$

$$F_{\mu\nu}J^{\nu} = 0. \tag{4}$$

Here equation (3) gives the expression of the current by the electromagnetic field and hence equations (3) and (4) constitute one nonlinear equation for $F_{\mu\nu}$. Further, for physically meaningfull solutions, we must demand

$$F^{\mu\nu}F_{\mu\nu} > 0, \tag{5}$$

to ensure the existence of the velocity which satisfies

$$U^{\mu}U_{\mu} = -1, \qquad F_{\mu\nu}U^{\nu} = 0. \tag{6}$$

Since equation (2) is satisfied by introducing the vector potential, at first sight it seems that the force-free equation becomes closed set by the vector potential. However this is wrong. Because the force-feee equation effectively has only two components. Thus even adding the gauge condition on the vector potential, the number of variables are still exceed the number of equations. Thus the problem first we must tackle is to express the electromagnetic field by two variables.

3. Degenerate Electromagnetic Field

The force-free equation (4) implies that the electromagnetic field is degenerate, i.e. the electromagnetic field satisfies

$$*F^{\mu\nu}F_{\mu\nu} = 0. \tag{7}$$

This implies that $\vec{E}\cdot\vec{B} = 0$ in the non-covariant language. Carter (1979) pointed out that the degenerate electromagnetic field is written as

$$F_{\mu\nu} = \partial_{\mu}\phi_1\partial_{\nu}\phi_2 - \partial_{\mu}\phi_2\partial_{\nu}\phi_1, \tag{8}$$

by two scalars ϕ_1 and ϕ_2. Probaply such an expression has never been used in the relativistic MHD, but it is sataightforwad relativistic extention of the Euler potentials used in the non-relativistic MHD. Thus we call ϕ_1 and ϕ_2 also the Euler potentials. The surface on which ϕ_1 and ϕ_2 are constant is called the flux surface. The intersection of the flux surface and $t = const$ three-surface is the magnetic field line. That is, to the degenerate electromagnetic fields, the concept of the magnetic field line can be introduced without ambiguity. Since the electromagnetic field is expressed by two Euler potentials, we can close the force-free equation.

4. Basic Equation Written by Euler potentials

By inserting equation (8) to the force-free equation, we have

$$(\partial_{\mu}\phi_1\partial_{\nu}\phi_2 - \partial_{\mu}\phi_2\partial_{\nu}\phi_1)\nabla_{\lambda}\left(\partial^{\nu}\phi_1\partial^{\lambda}\phi_2 - \partial^{\nu}\phi_2\partial^{\lambda}\phi_1\right) = 0. \tag{9}$$

Since $\partial_\mu\phi_1$ and $\partial_\mu\phi_2$ are independent if $F_{\mu\nu}$ does not vanish, equation (9) is separated to two equations as

$$\partial_\nu\phi_1\partial_\lambda\left\{\sqrt{-g}\left(\partial^\nu\phi_1\partial^\lambda\phi_2-\partial^\nu\phi_2\partial^\lambda\phi_1\right)\right\}=0, \tag{10}$$

$$\partial_\nu\phi_2\partial_\lambda\left\{\sqrt{-g}\left(\partial^\nu\phi_1\partial^\lambda\phi_2-\partial^\nu\phi_2\partial^\lambda\phi_1\right)\right\}=0. \tag{11}$$

These are the most general form of the basic equation of the force-free approximation. Then it turnes out that the action which yields the basic equation is given by

$$I=-\frac{1}{16\pi}\int F^{\mu\nu}F_{\mu\nu}\sqrt{-g}d^4x, \tag{12}$$

where the Lagrangian density is

$$L=-\frac{\sqrt{-g}}{16\pi}F^{\mu\nu}F_{\mu\nu}=-\frac{\sqrt{-g}}{16\pi}\left(\partial^\nu\phi_1\partial^\lambda\phi_2-\partial^\nu\phi_2\partial^\lambda\phi_1\right)\left(\partial_\nu\phi_1\partial_\lambda\phi_2-\partial_\nu\phi_2\partial_\lambda\phi_1\right). \tag{13}$$

It is straightforwad to see this action indeed leads to our basic equations.

Since the basic equation is derived from the covariant action, it is easy to apply this formalism in the curved spacetime.

5. Time Evolution of Force-Free Electromagnetic field

In order to study time dependent problems, it is convenient to abandon covariant aproach and split time derivatives explicitly. Then the magnetic field $\vec{B}$ and the electric field $\vec{E}$ are respectively given by

$$\vec{B}=\nabla\phi_1\times\nabla\phi_2, \qquad \vec{E}=-(\dot{\phi}_1\nabla\phi_2-\dot{\phi}_2\nabla\phi_1). \tag{14}$$

Then after tedius manipulations, we can show the basic equations (10) and (11) become

$$\begin{aligned}
&\begin{pmatrix} \nabla\phi_2\cdot\nabla\phi_2 & -\nabla\phi_1\cdot\nabla\phi_2 \\ -\nabla\phi_1\cdot\nabla\phi_2 & \nabla\phi_1\cdot\nabla\phi_1 \end{pmatrix}\begin{pmatrix} \ddot{\phi}_1 \\ \ddot{\phi}_2 \end{pmatrix} \\
&+\begin{pmatrix} \nabla\dot{\phi}_2\cdot\nabla\phi_2 & \nabla\phi_1\cdot\nabla\dot{\phi}_2-2\nabla\dot{\phi}_1\cdot\nabla\phi_2 \\ \nabla\dot{\phi}_1\cdot\nabla\phi_2-2\nabla\phi_1\cdot\nabla\dot{\phi}_2 & \nabla\dot{\phi}_1\cdot\nabla\phi_1 \end{pmatrix}\begin{pmatrix} \dot{\phi}_1 \\ \dot{\phi}_2 \end{pmatrix} \\
&\pm(\dot{\phi}_2\nabla^2\phi_1-\dot{\phi}_1\nabla^2\phi_2)\begin{pmatrix} \dot{\phi}_2 \\ \dot{\phi}_1 \end{pmatrix}\pm\nabla\times(\nabla\phi_1\times\nabla\phi_2)\cdot\begin{pmatrix} \nabla\phi_2 \\ \nabla\phi_1 \end{pmatrix}=0,
\end{aligned} \tag{15}$$

where ∇^2 denotes Laplacian. Thus if

$$\begin{vmatrix} \nabla\phi_2\cdot\nabla\phi_2 & -\nabla\phi_1\cdot\nabla\phi_2 \\ -\nabla\phi_1\cdot\nabla\phi_2 & \nabla\phi_1\cdot\nabla\phi_1 \end{vmatrix}=(\nabla\phi_1\times\nabla\phi_2)^2=|\vec{B}|^2\neq 0 \tag{16}$$

is satisfied at everywhere, we can invert equation (15) to the form:

$$\begin{pmatrix} \ddot{\phi_1} \\ \ddot{\phi_2} \end{pmatrix} = -\frac{(\nabla\phi_1 \times \nabla\phi_2)}{(\nabla\phi_1 \times \nabla\phi_2)^2} \cdot \begin{pmatrix} \nabla\phi_1 \times \nabla\dot{\phi_2} & \nabla\phi_1 \times \nabla\dot{\phi_1} \\ \nabla\phi_2 \times \nabla\dot{\phi_2} & \nabla\phi_2 \times \nabla\dot{\phi_1} \end{pmatrix} \begin{pmatrix} \dot{\phi_1} \\ \dot{\phi_2} \end{pmatrix}$$
$$-\frac{\nabla \cdot (\dot{\phi_1}\nabla\phi_2 - \dot{\phi_2}\nabla\phi_1)}{(\nabla\phi_1 \times \nabla\phi_2)^2} \begin{pmatrix} \nabla\phi_1 \cdot \nabla\phi_2 & -\nabla\phi_1 \cdot \nabla\phi_1 \\ \nabla\phi_2 \cdot \nabla\phi_2 & -\nabla\phi_1 \cdot \nabla\phi_2 \end{pmatrix} \begin{pmatrix} \dot{\phi_1} \\ \dot{\phi_2} \end{pmatrix}$$
$$+\frac{1}{(\nabla\phi_1 \times \nabla\phi_2)^2} \begin{pmatrix} \nabla\phi_1 \times (\nabla\phi_1 \times \nabla\phi_2) \\ \nabla\phi_2 \times (\nabla\phi_1 \times \nabla\phi_2) \end{pmatrix} \cdot \nabla \times (\nabla\phi_1 \times \nabla\phi_2), \qquad (17)$$

where $\dot{\phi_i} = \partial_t\phi_i$. This equation of motion indicates that if we specify the initial data $\phi_1, \phi_2, \dot{\phi_1}, \dot{\phi_2}$ all over the region considered at an initial epoch (spatial derivative on the initial surface can also be obtained from these quantities), we can calculate time evolution of the Euler potentials untill $|\vec{B}| = 0$ occurs. Further by

$$\frac{1}{2}F^{\mu\nu}F_{\mu\nu} = B^2 - E^2 = (\nabla\phi_1 \times \nabla\phi_2)^2 - (\dot{\phi_1}\nabla\phi_2 - \dot{\phi_2}\nabla\phi_1)^2 \qquad (18)$$

$F^{\mu\nu}F_{\mu\nu} > 0$ will break down before $|\vec{B}|^2 = 0$ takes place, if simultaneously $|\vec{E}|^2$ also vanishes.

Unfortunately, at the present stage, we can say nothing certain about the way of break down of the force-free approxmation. However it is known that there arises the magnetic neutral line (or point, surface), where $|\vec{B}| = 0$, at the boundary of the open magnetic field lines and the closed field lines. And such places often relate to highly energetic phenomena. Thus it seems that to clarify the nature of the breakdown of the force-free approximation is very important.

6. Application to Black Hole

Now let us apply our formulation to the black hole magnetosphere. As the background spacetime, Kerr spacetime is assumed. The covariant basic equations are the same with equation (10) and (11). Here we make 3+1 split again. The magnetic field and the electric field are respectively given by

$$\vec{B} = \nabla\phi_1 \times \nabla\phi_2, \qquad \vec{E} = -(\phi_1'\nabla\phi_2 - \phi_2'\nabla\phi_1). \qquad (19)$$

where $'$ denotes the derivartive along the ZAMO, that is

$$\phi_i' = U_{ZAMO}^{\mu}\partial_\mu\phi_i = \pounds_{U_{ZAMO}}\phi_i, \qquad (20)$$

with

$$U_{ZAMO} = \frac{1}{\alpha}(\partial_t + \omega\partial_\phi). \qquad (21)$$

Then after some manipulation, we have

$$\begin{pmatrix} \phi_1'' \\ \phi_2'' \end{pmatrix} = -\frac{(\nabla\phi_1 \times \nabla\phi_2)}{\alpha(\nabla\phi_1 \times \nabla\phi_2)^2} \cdot \begin{pmatrix} [(\vec{m}\cdot\nabla\phi_2)\nabla\omega + (\nabla\omega\cdot\nabla\phi_2)\vec{m}] \times \nabla\phi_1 \\ [(\vec{m}\cdot\nabla\phi_2)\nabla\omega + (\nabla\omega\cdot\nabla\phi_2)\vec{m}] \times \nabla\phi_2 \end{pmatrix}$$
$$\begin{pmatrix} [(\vec{m}\cdot\nabla\phi_1)\nabla\omega + (\nabla\omega\cdot\nabla\phi_1)\vec{m}] \times \nabla\phi_1 \\ [(\vec{m}\cdot\nabla\phi_1)\nabla\omega + (\nabla\omega\cdot\nabla\phi_1)\vec{m}] \times \nabla\phi_2 \end{pmatrix} \begin{pmatrix} \phi_1' \\ \phi_2' \end{pmatrix}$$
$$-\frac{(\nabla\phi_1 \times \nabla\phi_2)}{(\nabla\phi_1 \times \nabla\phi_2)^2} \cdot \begin{pmatrix} \nabla\phi_1 \times \nabla\phi_2' & \nabla\phi_1 \times \nabla\phi_1' \\ \nabla\phi_2 \times \nabla\phi_2' & \nabla\phi_2 \times \nabla\phi_1' \end{pmatrix} \begin{pmatrix} \phi_1' \\ \phi_2' \end{pmatrix}$$
$$-\frac{\nabla\cdot(\phi_1'\nabla\phi_2 - \phi_2'\nabla\phi_1)}{(\nabla\phi_1 \times \nabla\phi_2)^2} \begin{pmatrix} \nabla\phi_1\cdot\nabla\phi_2 & -\nabla\phi_1\cdot\nabla\phi_1 \\ \nabla\phi_2\cdot\nabla\phi_2 & -\nabla\phi_1\cdot\nabla\phi_2 \end{pmatrix} \begin{pmatrix} \phi_1' \\ \phi_2' \end{pmatrix}$$
$$+\frac{1}{\alpha(\nabla\phi_1 \times \nabla\phi_2)^2} \begin{pmatrix} \nabla\phi_1 \times (\nabla\phi_1 \times \nabla\phi_2) \\ \nabla\phi_2 \times (\nabla\phi_1 \times \nabla\phi_2) \end{pmatrix} \cdot \nabla \times (\alpha\nabla\phi_1 \times \nabla\phi_2), \qquad (22)$$

where $\nabla\phi_i' = \nabla(\phi_i')$. The first term of the right-hand side represents the effect of the rotation of the spacetime and α represents the effect of the gravitation. It is easily seen that the structure of the equation is essentialy identical with the flat space case.

7. Application of Noether Identities

Since the Lagrangian density has a form like $L = L[\partial_\mu\phi_{(i)}, g_{\mu\nu}]$, by Lie differentiating the Lagrangian density by an arbitrary vector ξ, we have

$$0 = -\partial_\mu \frac{\partial L}{\partial(\partial_\mu\phi_{(i)})}\pounds_\xi\phi_{(i)} - \partial_\mu\left(\sqrt{-g}P^\mu[\xi]\right) + \frac{1}{2}\sqrt{-g}T_{\mu\nu}\pounds_\xi g_{\mu\nu}, \qquad (23)$$

where $P^\mu[\xi]$ is the canonical flux vector with respect to ξ, $T^{\mu\nu}$ is the metric energy-momentum tensor respectively given by

$$P^\mu[\xi] = -\frac{1}{\sqrt{-g}}\left\{\frac{\partial L}{\partial(\partial_\mu\phi_{(i)})}\pounds_\xi\phi_{(i)} - L\xi^\mu\right\}, \qquad (24)$$

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\partial L}{\partial g_{\mu\nu}}. \qquad (25)$$

Further we define the canonical energy-momentum tensor by

$$\mathcal{T}^\mu_\lambda = -\frac{1}{\sqrt{-g}}\left\{\frac{\partial L}{\partial(\partial_\mu\phi_{(i)})}\partial_\lambda\phi_{(i)} - L\delta^\mu_\lambda\right\}. \qquad (26)$$

Thus $P^\mu[\xi] = \mathcal{T}^\mu_\lambda\xi^\lambda$. When the equation of motion is satisfied, the first term of the right-hand side of equation (23) vanishes. Thus we have

$$(T^\mu_\lambda - \mathcal{T}^\mu_\lambda)\nabla_\mu\xi_\lambda - (\nabla_\mu\mathcal{T}^\mu_\lambda)\xi^\lambda = 0. \qquad (27)$$

Since ξ is arbitrary, ξ^λ and $\nabla_\mu\xi_\lambda$ can be independently specified. Thus we have

$$\nabla_\mu T^\mu_\lambda = 0, \qquad T'^\mu_\lambda = T^\mu_\lambda. \tag{28}$$

Thus we have also $\nabla_\mu T'^\mu_\lambda = 0$. Therefore the metric flux vector $\Pi^\mu[\xi] = T^{\mu\nu}\xi_\nu$ become identical with the canonical one, that is

$$\Pi^\mu[\xi] = P^\mu[\xi] = \frac{-1}{\sqrt{-g}}\left\{\frac{\partial L}{\partial(\partial_\mu\phi_{(1)})} + \frac{\partial L}{\partial(\partial_\mu\phi_{(2)})} - L\xi^\mu\right\}. \tag{29}$$

The reason why we have rewrite the metric flux by the canonical one is the latter is more convenient to deal with the symmetry properties when treating configurations with some symmetry. Here let us discuss the stationary and axisymmetric case, for example, by using equation (29). In case of the stationary and axisymmetric configuration, it can be shown that the Euler potentials take the form like

$$\phi_1 = \Psi(r,\theta), \qquad \phi_2 = \varphi - \Omega_F(\Psi)t + \psi_2(r,\theta), \tag{30}$$

where φ is the axial coordinate. This shows that even when the electromagnetic field is stationary, the magnetic field lines, on which ψ_1 and ψ_2 are constant, are moving. This is the reason why the energy and angular momentum are transported by the stationary and axisymmetric force-free electromagnetic field. Since we have

$$\pounds_{\mathbf{k}}\phi_1 = \pounds_{\mathbf{m}}\phi_1 = 0, \tag{31}$$

$$\pounds_{\mathbf{k}}\phi_2 = -\Omega_F, \qquad \pounds_{\mathbf{m}}\phi_2 = 1, \tag{32}$$

where $\mathbf{k}$ is the time Killing vector $= \partial_t$ and $\mathbf{m}$ is the axial Killing vector ∂_φ. Then the metric energy flux $\mathcal{E}^\mu = -\Pi^\mu[\mathrm{K}]$ and the metric angular momentum flux $\mathcal{J}^\mu = \Pi^\mu[\mathbf{m}]$ respectively become

$$\mathcal{E}^\mu = -\frac{1}{\sqrt{-g}}\left\{\Omega_F\frac{\partial L}{\partial(\partial_\mu\phi_2)} + L\mathrm{k}^\mu\right\}, \tag{33}$$

$$\mathcal{J}^\mu = -\frac{1}{\sqrt{-g}}\left\{\frac{\partial L}{\partial(\partial_\mu\phi_2)} - L\mathrm{m}^\mu\right\} \tag{34}$$

Thus their poloidal (r,θ) components relate as

$$\mathcal{E}^A = \Omega_F\mathcal{J}^A, \qquad A = r,\theta, \tag{35}$$

Further by using

$$\frac{\partial L}{\partial(\partial_\mu\phi_2)} = \frac{\sqrt{-g}}{4\pi}F^{\mu\nu}\partial_\nu\phi_1, \tag{36}$$

we have

$$\mathcal{E}^\mu\partial_\mu\phi_1 = \mathcal{E}^A\partial_A\Psi = 0, \qquad \mathcal{J}^\mu\partial_\mu\phi_1 = \mathcal{J}^A\partial_A\Psi = 0. \tag{37}$$

This implis that the energy and angular momentum flux flow along $\Psi = const$ surface, i.e, the poloidal surface. And we also have

$$\mathcal{J}^r = -\frac{1}{4\pi} F^{r\theta}\partial_\theta \Psi, \qquad \mathcal{J}^\theta = \frac{1}{4\pi} F^{r\theta}\partial_r \Psi, \tag{38}$$

and

$$\mathcal{E}^r = -\frac{\Omega_F}{4\pi} F^{r\theta}\partial_\theta \Psi, \qquad \mathcal{E}^\theta = \frac{\Omega_F}{4\pi} F^{r\theta}\partial_r \Psi. \tag{39}$$

Of course these are already well-known results. However, note that here we only use the symmetry, that is, the force-free equation is not explicitly used at all. Therefore derivation here manifests the fact that these properties come only from symmetry properties of the Euler potentials.

8. Conclusion

We have preseted a general theory to deal with the force-free electromagnetic field without assuming any symmetry. The electromagnetic field is expressed by the Euler potentials and gave the equations which determine time evolution of the Euler potentials. The theory is originaly written in the covariant way and hence it is straightforward to apply it to the black hole magnetosphere. We gave the equations which govern time evolution of the force-free electromagnetic field around the black hole. Further we show that the well-known properties of conserved fluxes in the stationary and axisymmetric configuration can be obtained only from the Lagrangian density and the symmetry.

Reference

Carter B.,1979 , in Hawking S.W. & Israel W. eds, General relativity: an Einstein Centenary Survay, Camblidge University Press, Camblige, p.294

Collimation of Magnetosonic Waves near a Black Hole

Kouichi HIROTANI
National Astronomical Observatory Japan
Osawa 2-21-1, Mitaka, Tokyo 181
E-mail: hirotani@ferio.mtk.nao.ac.jp

Abstract

We study the porpagation of fast-magnetosonic waves around a rotating (Kerr) black hole. Assuming that the magnetic field dominates the plasma accretion, we find that the accreting fluid suffers a large radial acceleration resulting from the Lorentz force, and becomes highly variable in a narrow region just beyond the fast-magnetosonic surface where the radial velocity of infalling plasma macthes the fast-magnetosonic speed. If this critical surface has an oblate shape for radial poloidal field lines, the energy of this large-amplitude fluid disturbance is carried in the form of outgoing fast-magnetosonic waves propagating from the equatorial super-fast-magnetosonic region to the polar sub-fast-magnetosonic region. It is further found that the waves are polewardly collimated and reach the pair creation region where both the inflows and the outflows originate. Thus, the effective acceleration of particles due to the outgoing fast-magnetosonic waves will work mainly in the polar direction as a mechanism of trigger of electron-positron pair beams in the nuclei of extragalactic radio sources.

1 Introduction

It is generally accepted that nuclei of extragalactic radio sources can eject plamsa relativistically. For example, radio observations of superluminal sources and the non detection of the inverse Compton catastrophe in variable compact radio sources give evidences for relativistic ejections. Furthermore, the discovery of some tens of strong and anisotropic γ-ray sources (e.g. Fichtel et al. 1994) associated with compact beamed radio sources provided new evidence for relativistic ejection from nuclei of powerful radio sources. More recently, Roland and Hermsen (1995) showed that the strong peaks in the 1-10 MeV range in the spectrum of GRO J0516-609 can be interpreted as the signature of a broad blue shifted $e^{\pm}$ annihilation line.

We are motivated by the need to investigate the formation of such relativistic $e^{\pm}$ beams in the nuclei of radio sources, or in general, in active galactic nuclei (AGNs). It seems reasonable to suppose that $e^{\pm}$ plasmas are supplied in the pair creation region where both inflows and outflows originate. At the pair creation region, the Goldreich-Julian charge density will vanish by analogy with the outer gap models for pulsar magnetospheres. In a black hole magnetosphere, the charge density possibly vanishes at the position where the angular velocity of space-time dragging (ω) due to black hole rotation roughly matches that of magnetic field lines (Ω_F). The position where

$\omega \sim \Omega_F$ holds is located close to the black hole, typically at $r \sim 1.6 r_H$ (Beskin et al. 1992), where r_H is the horizon radius in Boyer-Lindquist coordinate; therefore, we may expect plasma supply due to pair creation also at this radius.

In this context, the purpose here is to explore a little further the problem whether steady supply of energy is maintained in the pair creation region located near to the horizon. In particular, we shall concentrate on whether some portions of accretion energy distributed in the equatorial region will be conveyed to the polar pair creation region in the form of MHD waves. Adopting the magnetically dominated limit in which the rest-mass energy density of particles is negligible compared with the magnetic energy density, we will show that plasma accretion is intrinsically highly variable and that the energy of this disturbance will be collimated polewards and reach the pair creation region in the form of outgoing fast-magnetosonic waves.

2 Highly Variable Accretion

We will begin by considering basic equations describing a magnetosphere around a rotating black hole. Since the self-gravity of the electromagnetic field and plasma around the black hole is very weak, the background geometry of the magnetosphere surrounding such a black hole is described by the Kerr metric. We assume the ideal MHD conditions in which the electric field vanishes in the fluid rest frame. Furthermore, the motion of the fluid is described in the cold limit. The proper number density obeys the continuity equation.

In order to look closely at the problem of how the existence of the horizon affects the MHD interactions, we would like to focus attention on an analysis near to the horizon ($\Delta \equiv r^2 - 2Mr + a^2 \ll 1$), where r is the radial coordinate in Boyer-Lindquist coordinate, M the hole's mass, and a the hole's angular momentum per unit mass. Since the poloidal magnetic field near to the horizon tends to be bent into a radial shape, due to fluid inertia, even in the magnetically dominated limit (Hirotani et al. 1992), it may be safely assumed that the unperturbed field is radial. Because particles are frozen to the radial magnetic field lines, flow lines also become radial. Under this condition, the unperturbed electric field becomes meridional. Furthermore, we should notice that the fast-magnetosonic point ($r = r_F$) is located very close to the horizon ($r = r_H \equiv M + \sqrt{M^2 - a^2}$) in the magnetically dominated limit.

Let us next consider a non-stationary and axisymmetric perturbation superposed on the stationary and axisymmetric magnetically dominated magnetosphere. We consider the Fourier components of purturbed quantities and replace temporal derivatives into frequency (i.e., $\partial_t \rightarrow i\omega$) and adopt a local approximation in the meridional direction (i.e., $\partial_\theta \rightarrow -ik_\theta$). Solving linear perturbation equations, we obtain relations describing relative amplitude among perturbed quantities. To see the relative amplitude, it is convenient to take the short-wavelength limit not only in the meridional direction but also in the radial direction. Then we can see that the equipartition of energy between the electromagnetic field and the fluid is achieved near the fast-surface, although the perturbation energy is supplied mainly in the form of electromagnetic disturbances far from the horizon. In other words, a lot of perturbation energy is supplied to the fluid from the electromagnetic field, as a result of the large amplitude radial Lorentz force

resulting from large meridional current (J_θ). This effect is due to a redshift effect near the horizon, and disappears if the fast-magnetosonic point is located far from it (Hirotani et al. 1993).

3 Structure of Fluid Disturbance

We see in this section that the magnetically dominated accretion is intrinsically highly variable near the fast-surface located very close to the horizon. Let us next emamine the spatial structure of fluid disturbance. Combining linear perturbation equations, we obtain a single wave equation which fluid's radial velocity, energy, and angular momentum obey. Assuming that the perturbed quantities are proportional to $\exp(i\omega - ik_\theta)$, we can regard the wave equation as a 2nd order differential equation with respect to r. An approximated solution to this equation is expressed as

$$u^r = C_1 \frac{x^{i\sigma}}{[x - x_F(\theta) + (dx_F/d\theta)^2 + i\epsilon]^{1+2i\sigma}}, \tag{1}$$

where x is a non-dimensional radial coordinate defined by $x \equiv \Delta/\Sigma_H = 2(r_H - M)(r - r_H)/(r_H{}^2 + a^2\cos^2\theta)$; the imaginary part $i\epsilon$ is much smaller than the real part $(dx_F/d\theta)^2$. A sharp peak of the amplitude appears at the surface $x = x_F - (dx_F/d\theta)^2$ just beyond the fast-surface. The infalling motion of the plasma in the sub-fast region is mainly disturbed slightly outside the fast-surface, because the outgoing waves disperse and the amplitude decays like $| u^r | \sim x^{-1}$ for $x \gg x_F$. The steady supply of outgoing waves is due to the meridional propagation from the super-fast equatorial region to the sub-fast polar one, provided that the fast-surface is oblate ($dx_F/d\theta > 0$). A large part of the outgoing waves can escape from the narrow super-fast region in the range $x_F - (dx_F/d\theta)^2 < x < x_F$, before reaching the polar region. Hence, the peak amplitude of u^r at $x = x_F - (dx_F/d\theta)^2$ exponentially decreases as θ decreases. In the next section, we confirm the importance of meridional propagation in detail.

4 Poleward Collimation of Waves

To consider the meridional propagation, it is useful to calculate the characteistics of the wave equation. The characteristics of the wave equation in the super-fast region are presented by the non-linear differential equation

$$dx/d\theta = \mp\sqrt{x_F(\theta) - x}. \tag{2}$$

The minus sign corresponds to the wave propagating to lower lattitudes, while the plus sign to a polewardly propagating wave. Here, we must notice that $dx < 0$ holds near the horizon. Both waves are depicted in Figure 1. We are considering a situation in which accretion is mainly supplied from the equatorial disk. In this case, the magnetosphere has an oblate critical surface as described in Figure 1, because the fast-magnetosonic speed in the equatorial region is smaller than that in the polar regions (Hirotani et al. 1992).

If the fast-surface is oblate, only the waves propagating polewards can escape into a sub-fast region (for general discussion, see Bogovalov 1992). This could be clearly seen if we solved the equation (2) (Hirotani et al. 1994), but can also be understood intuitively from the figure. If the fast-surface were prolate, on the contrary, only the waves propagating into the equatorial region could escape.

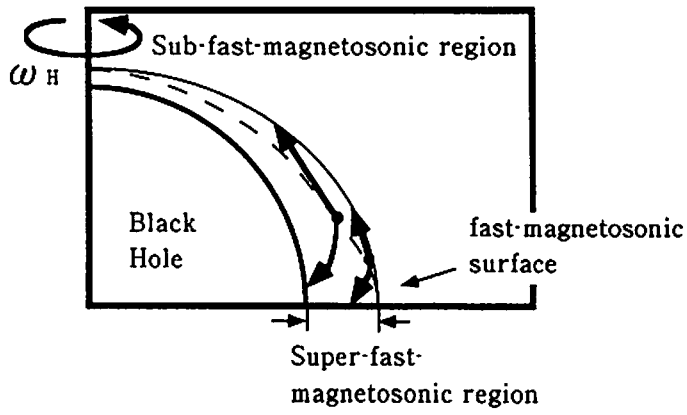


Fig. 1. Schematic figure of a black hole magnetosphere near to the horizon. A pair of thick solid curves denotes a pair of characteristics of equation (2). The maximum-amplitude surface $[x = x_F - (dx_F/d\theta)^2]$, which is indicated by the dashed line, coincides with the fast-magnetosonic separatrix surface.

5 Propagation of Waves in the Sub-fast Region

Now that we are sure that the large-amplitude fluid disturbance caused in the super-fast-magnetosonic region can escape to the sub-fast-magnetosonic region by propagating poleward, the next step is to investigate whether such disturbance can reach the pair creation region where not only inflows but also outflows are supplied. In this section, we shall confine our attention to this problem by analyzing the eikonal equations of a ray of a magnetosonic wave packet.

The equation of geometrical magneto-acoustics can be described in the covariant form

$$H(x^\alpha, \Psi_{,\beta}) = S^{\lambda\mu\nu\rho}\Psi_{,\lambda}\Psi_{,\mu}\Psi_{,\nu}\Psi_{,\rho} = 0, \tag{3}$$

where the magneto-acoustic metric $S^{\lambda\mu\nu\rho}$ is a function of the location of a wave packet,x^μ, and of its momentum, $p_\mu \equiv \Psi_{,\mu}$; the characteristic hypersurfaces which play the role of wave fronts in a stationary space-time, can be expressed as $\Psi(x^\mu) =$ constant. $S^{\lambda\mu\nu\rho}$ contains the terms describing the influence of the gravitational field and of the magneto-hydrodynamic flows, and is explicitly given by equation (35-28) of Lichnerowicz (1967).

The general theory of surfaces which satisfies first-order partial differential equation (3), asserts that such surfaces can be built up from a family of elements, so-called strips. The strip is characterized by $x^\mu(l)$ and $p_\mu(l)$, where l is the parameter of the strip. Instead of solving the

partial differential equations; however, we can build up the solution surface from its characteristic strips, which are solutions of ordinary differential equations (see e.g., Takahasi, Ishizuka, and Yokosara for the discussion of acoustic waves in Kerr space-time)

$$\dot{x}^{\mu}(l) = \frac{\partial H(x^{\alpha}, p_{\beta})}{\partial p_{\mu}}, \tag{4}$$

$$\dot{p}_{\mu}(l) = \frac{\partial H(x^{\alpha}, p_{\beta})}{\partial x^{\mu}}. \tag{5}$$

A solution to these equations is a characteristic of the original partial differential equations and gives the ray along which a magnetohydrodynamic perturbation propagates.

In a stationary and axisymmetric space-time, we can set p_t and p_ϕ to be constant. Thus, the r and θ components of the Hamilton equations (4) and (5), which hold not only in the vicinity of the horizon, but also far from it, describe the rays $r = r(l)$, $\theta = \theta(l)$ on the poloidal plane. As an initial condition, the position of a wave packet is chosen to be on the fast-surface; in addition, its momentum vector is set to be purely meridional (i.e., $p_r = 0$) in accordance with the results obtained in the last section. The meridional component of the initial momentum vector represents the amplitude of perturbation, on which the results does not depend much.

A typical example of rays around a Schwarzschild black hole is presented in figure 2. Here, the angle at which wave packets start is chosen to be 75°, and the azimuthal wave number $m = -p_\phi$ is chosen to be 0, 1, and 5. Since non-axisymmetric wave packets conserve non-zero angular momenta, they are reflected near the rotational axis of the magnetosphere. As the diagram indicates, fast-magnetosonic disturbance can reach the region $r \sim 1.6 r_{\rm H}$, where pair creation could be expected if the hole was rotating. We analyzed only Schwarzschild cases in this paper; however, we expect essentially the same results will be obtained for Kerr cases.

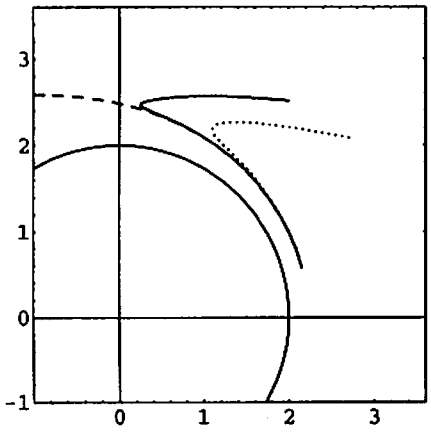


Fig. 2. Rays of fast-magnetosonic wave packets around a Schwarzschild black hole. Solid curve denotes $m = 1$ case, while dashed and dotted ones denote $m = 0$ and $m = 5$ cases, respectively.

6 Summary and Discussion

Let us summarize the main points that have been made in this paper. We have studied non-stationary and axisymmetric perturbations of magnetically dominated accretion onto a rotating black hole. A slight fluctuation of the electromagnetic field is accompanied by fluid motion which is highly disturbed in the super-fast region just beyond the fast-surface. Some portions of the energy of this fluid disturbance reach a sub-fast region by virtue of the meridional propagation in the form of outgoing fast-magnetosonic waves. After reaching the sub-fast region, the waves concentrate toward the polar pair creation region which will be located at $r \sim 1.6 r_{\rm H}$ for mildly rotating holes.

It would be possible to argue that the poleward collimation of waves may bring about the acceleration of particles in electron-positron beams. To understand how the wave energy is converted into particle's kinetic energy is an important problem for developing the MHD scenario of beam formation in the very central region of AGNs and/or galactic black hole candidates. For example, in addition to heating of the surrounding plasma due to the dissipation of wave energy into thermal energy, the wave stress exerted on the expanding atmosphere may result in a direct transfer of momentum and energy from the waves to the outgoing winds (Jacques 1977). This will be the subject of future investigations.

References

Beskin, V. S., Istomin, Ya. N., and Par'ev, V. I. 1992, Sov. Astron. 36, 642.

Bogovalov S. V. 1992, in Proceed. of 4th International Toki Conf. on Plasma Physics and Controlled Nuclear Fusion, ed Guyenne T. D., Hunt, J. J., (Europian Space Agency, Netherland) p. 317.

Hirotani K., Takahashi M., Nitta S., and Tomimatsu A. 1992, ApJ 386, 455.

Hirotani K., Tomimatsu A., and M. Takahashi 1993, PASJ 45, 431.

Hirotani K., and Tomimatsu A. 1994, PASJ 46, 643.

Jacques S. A. 1977, ApJ 215, 942

Roland, J., and Hermse, W. 1995, AA297, L9.