PROCEEDINGS OF THE 10TH WORKSHOP ON GENERAL RELATIVITY AND GRAVITATION IN JAPAN

OSAKA UNIVERSITY
SEPTEMBER 11-14, 2000

EDITED BY
MISAO SASAKI, JUN'ICHI YOKOYAMA
TAKASHI NAKAMURA, AND KENJI TOMITA

PREFACE

The tenth workshop on General Relativity and Gravitation (the 10th JGRG meeting) was held at the convention center, Osaka University, from September11th to 14th, 2000. The purpose of the workshop was to review the latest developments in general relativity and gravitation and related fields and to discuss future prospects. Unlike the previous workshops, this year we decided to make it more international, and invited a large fraction of review speakers from abroad. We also asked Japanese speakers to present their talks in English. The workshop was quite successful. It provided young Japanese participants a very good occasion to discuss topics of mutual interest with world-renowned experts in the field. We wish to thank all the participants for their contributions to the workshop.

We would like to thank Ms. Kiyoe Yokota for her devotion to secretarial duties of the workshop. Without her help, the workshop could not have been so successful. The workshop was supported by Monbukagakusho Grant-in-Aid for Scientific Research, No. 112440063 and by Grant-in-Aid for Creative Basic Research, No. 09NP0801.

JGRG10 Organizing committee

Takeshi Chiba (Kyoto University)
Yoshiharu Eriguchi (University of Tokyo)
Toshifumi Futamase (Tohoku University)
Akio Hosoya (Tokyo Inst. of Technology)
Masumi Kasai (Hirosaki University)
Hideo Kodama (YITP)
Kei-ichi Maeda (Waseda University)
Takashi Nakamura (YITP)
Ken-ichi Nakao (Osaka City University)
Yasusada Nanbu (Nagoya University)
Ken-ichi Oohara (Niigata University)
Misao Sasaki (Osaka University)
Naoshi Sugiyama (NAO)
Kenji Tomita (YITP)
Jun'ichi Yokoyama (Osaka University)

JGRG 10 PROGRAM

(ALL TALKS WILL BE GIVEN IN ENGLISH UNLESS OTHERWISE INDICATED.)

Monday, September 11

Morning session I (Chair: T. Tanaka)

9:10	M. Sasaki(Osaka)	Opening address
9:20	Y. Mino(Tokyo)	Gravitational radiation reaction
9:40	H. Tagoshi(Osaka)	Search for gravitational waves from inspiraling binaries using TAMA300 data
10:00	T. Harada(Waseda)	Gravitational memory? –A perturbative analysis—

10:20-10:40 Break

Morning session II (Chair: M. Seriu)

10:40	A. Ohashi(Kyoto)	Quasi-normal modes of massive fields
11:00	T. Kawai(Osaka CU)	Energy-momentum and angular momentum densities in gauge theories of gravity
11:20	M. Tanimoto(YITP)	Symmetries of the Gowdy equations and spatial topologies
11:40	A. Ishibashi(YITP)	Convex functions and spacetime geometry

12:00-1:00 Lunch

Invited lectures (Chair: H. Kodama)

1:00	B. Iyer(Raman)	Pade approximants for gravitational waves from inspiraling compact binaries
2:00	S. Iso(KEK)	Type IIB Matrix Model as a Nonperturbative Formulation of Superstring

3:00-3:20 Break

Afternoon session (Chair: K. Maeda)

3:20	K. Oohara(Niigata)	General relativistic simulation on coalescing binary neutron stars
3:40	H. Shinkai(Penn State)	Will hyperbolic formulations help numerical relativity?
4:00	K. Koyama(Kyoto)	Quantum fluctuations in the inflationary brane world
4:20	T. Shiromizu(Tokyo)	Black holes in brane world: Time symmetric initial data
4:40	O. Seto(YITP)	Gravitational waves in cosmological model of Horava-Witten theory
5:00	End	

Tuesday, September 12

Morning session I (Chair: K. Nakao)

9:00	R. Yamazaki(Kyoto)	Static solutions of Einstein-Born-Infeld dilaton system in 2+1 dimensions (Japanese)
9:20	K. Nakamura(Keio)	Initial condition of a self-gravitating loop string and linear perturbation (Japanese)
9:40	Y. Morisawa(YITP)	Thick domain walls around a black hole (Japanese)
10:00	H. Kudoh(YITP)	Birth of a timelike naked singularity (Japanese)

10:20-10:40 Break

Morning session II (Chair: H. Tagoshi)

10:40	S. Higaki(Kyoto)	Instability of a membrane intersecting a black hole
11:00	H. Iguchi(Osaka)	How does a naked singularity explode?
11:20	K. Nakao(Osaka CU)	Gravitational waves from naked singularity by Newtonian analysis

11:40-12:40 Lunch

Invited lectures (Chair: A. Hosoya)

12:40	T. Singh(Tata & YITP)	Comparing quantum black holes and quantum naked singularities
1:40	D. Page(Alberta)	Aspects of black hole entropy

2:40-3:00 Break

Afternoon session (Chair: A. Tomimatsu)

3:00	S. Mukohyama(Victoria)	Moving mirror and black hole entropy
3:20	A. Hosoya(TIT)	Generalized second law of black hole and quantum information theory
3:40	J. Koga(Waseda)	"Asymptotic" Killing horizon
4:00	D. Ida(Kyoto)	On black holes in 2+1 dimensions
4:20	T. Torii(Tokyo)	Scalar hair on the black holes in the presence of cosmological constant
4:40	H. Koyama(Nagoya)	Late-time behaviors of a massive scalar field in Reissner-Nordstrom background
5:00	End	

Wednesday, September 13

Morning session I (Chair: T. Chiba)

9:00	K. Yasuno(TIT)	Volume criticality and cosmological perturbations
9:20	K. Yamamoto(Hiroshima)	Cosmological test with the geometric distortion and evolution of bias
9:40	T. Tatekawa(Waseda)	Primordial fractal density perturbations and structure formation: 1-D collisionless sheet model

10:00-10:20 Break

Morning session II (Chair: K. Yamamoto)

10:20	O. Iguchi(Ochanomizu)	The origin of fractal distribution in self gravitating virialized system
10:40	M. Matsumiya(Osaka)	Cosmic inversion
11:00	T. Inoue(YITP)	CMB in closed multiply-connected universes
11:20	T. Chiba(Kyoto)	Reconstructing the quintessential potential in practice

11:40-12:40 Lunch

Invited lectures (Chair: J. Yokoyama)

12:40	R. Brandenberger(Brown)	Inflationary cosmology: Progress and problems
1:40	A. Starobinsky(Landau&Tokyo)	Determination of cosmological Lambda-term properties - an inverse dynamical problem

2:40-3:00 Break

Afternoon session (Chair: T. Shiromizu)

3:00	S. Tsujikawa(Waseda)	Multi-field fermionic preheating
3:20	M. Yamaguchi(Tokyo)	Chaotic inflation model in supergravity
3:40	N. Sakai(Yamagata)	Evolution of black hole mass in Brans-Dicke cosmology
4:00	T. Kanazawa(Tokyo)	Double inflation in supergravity and PBH formation
4:20	S. Kasuya(Tokyo)	Q-ball formation and cosmology
4:40	M. Seriu(Fukui)	Cognition of the Universe through a model
5:00	End	

5:30 Party at the Restaurant Rosen

Thursday, September 14

Morning session I (Chair: M. Takahashi)

9:00	T. Matsuoka(Nagoya)	Screw instability in black-hole magnetsphere (Japanese)
9:20	R. Takagi(TIT)	The number of peaks within GRB and internal shock (Japanese)
9:40	K. Asano(Osaka)	Neutrino pair annihilation above a Kerr black hole (Japanese)

10:00-10:20 Break

Morning session II (Chair: K. Oohara)

10:20	T. Ohta(Miyagi U Ed)	A generalized equilibrium condition in the electro-gravi dynamics in 1+1 dimensions
10:40	S. Kobayashi(Osaka)	Gamma-ray bursts and internal shocks
11:00	T. Nakamura(YITP)	Multiple sub-jet model of gamma ray bursts and possible origin of X-ray precursors and post-cursors
11:20	K. Ioka(Kyoto)	Gravitational waves from the giant flares of the soft gamma-ray repeaters

11:40-12:40 Lunch

Invited lectures (Chair: T. Nakamura)

12:40	T. Murakami(ISAS)	Origin of gamma-ray bursts
1:40	V. Kuzmin(INR&YITP)	Ultra-high energy cosmic rays and superheavy particles in the Universe

2:40-3:00 Break

Afternoon session (Chair: Y. Kojima)

3:00	K. Konno(Hiroshima)	Magnetic effects on the moment of inertia of a relativistic star
3:20	M. Takahashi(Aichi U Ed)	Fire ball generation by MHD jet
3:40	S. Koide(Toyama)	General relativistic simulations of jet formation in a rapidly rotating black hole
4:00	T. Nakamura(YITP)	Possible direct method to determine the radius of a star
4:20	Y. Kojima(Hiroshima)	Time evolution of purely axial displacements in isentropic neutron stars
4:40	S. Karino(Tokyo)	r-mode in a rapidly differentially rotating neutron star

5:00 End

PROGRAM (TABLE OF CONTENTS)

Gravitational Radiation Reaction
Yasushi Mino 1

Search for Gravitational Waves from Inspiraling Compact Binaries using TAMA300 data
Hideyuki Tagoshi and TAMA collaboration 8

Gravitational Memory? — a Perturbative Approach —
Tomohiro Harada, B. J. Carr and C. A. Goymer 14

Quasi-normal mode of massive scalar field: Analytic model approach
Akira Ohashi and Masa'aki Sakagami 21

Energy-Momentum and Angular Momentum Densities in Gauge Theories of Gravity
Toshiharu Kawai 27

Symmetries of the Gowdy Equations and Spatial Topologies
Masayuki Tanimoto 35

Convex Functions and Spacetime Geometry
Gary W. Gibbons and Akihiro Ishibashi 41

Padé Approximants for gravitational Waves from Inspiralling Compact Binaries
T. Damour, Bala R. Iyer and B.S. Sathyaprakash 49

Type IIB Matrix Model as a Constructive Formulation of Superstring
Satoshi Iso 60

A Maximal Slicing Code for Coalescencing Binary Neutron Stars
Ken-ichi Oohara and Takashi Nakamura 75

Will hyperbolic formulations help numerical relativity?
— Experiments using Ashtekar's connection variables —
Hisa-aki Shinkai and Gen Yoneda 80

Quantum Fluctuations of Bulk Inflation in Inflationary Brane World
Kazuya Koyama, Shinpei Kobayashi and Jiro Soda 87

Black holes in the brane world: Time symmetric initial data
Tetsuya Shiromizu and Masaru Shibata 95

Gravitational waves in cosmological models of Hořava-Witten theory
Osamu Seto 101

Static Solutions to Einstein-Born-Infeld-dilaton System in (2+1) Dimensions
Ryo Yamazaki and Daisuke Ida 108

Initial condition of a loop cosmic string and linear perturbation
Kouji Nakamura 115

Thick Domain Walls around a Black Hole
Yoshiyuki Morisawa, Daisuke Ida, Akihiro Ishibashi and Ken-ichi Nakao 123

An example of a timelike naked singularity formation
Hideaki Kudoh, Tomohiro Harada and Hideo Iguchi 131

Instability of a membrane intersecting a black hole
Susumu Higaki, Daisuke Ida and Akihiro Ishibashi 139

How does a naked singularity explode?
Hideo Iguchi and Tomohiro Harada 147

Gravitational Waves due to the Naked Singularity Formation by Newtonian Analysis
Ken-ichi Nakao, Hideo Iguchi and Tomohiro Harada 155

Comparing Quantum Black Holes and Naked Singularities
T. P. Singh 161

Thermodynamics of Near-Extreme Black Holes
Don N. Page 174

Moving mirror and black hole entropy — A proposal of moving-mirror entropy —
Shinji Mukohyama 186

— Quantum Information and Generalized Second Law of Black Hole —
Akio Hosoya, Alberto Carlini and Takeshi Shimomura 191

"Asymptotic" Killing horizon
Jun-ichirou Koga 198

On black holes in three-dimensions
Daisuke Ida 206

Scalar hair on the black hole in the presence of cosmological constant
Takashi Torii, Makoto Narita and Kengo Maeda 208

Late-time behaviors of a massive scalar field in Reissner-Nordström background
Hiroko Koyama and Akira Tomimatsu 216

Volume Criticality and Cosmological Perturbations
Katsuhito Yasuno and Masayuki Tanimoto 224

Cosmological Test with the Geometric Distortion and the Evolution of Bias
Kazuhiro Yamamoto 228

Primordial Fractal Density Perturbations and Structure Formation in the Universe: 1-D Collisionless Sheet Model
Takayuki Tatekawa and Kei-ichi Maeda 236

The Origin of Fractal Distribution in Self-Gravitating Virialized System
Osamu Iguchi, Yasuhide Sota, Masahiro Morikawa, Takayuki Tatekawa and Kei-ichi Maeda 244

Cosmic Inversion
Makoto Matsumiya, Misao Sasaki and Jun'ichi Yokoyama 252

Cosmic Microwave Background in Closed Multiply Connected Universes
Kaiki Taro Inoue 257

Reconstructing the Quintessential Potential in Practice
Takeshi Chiba 261

A Status Review of Inflationary Cosmology
Robert H. Brandenberger 268

Multi-field fermionic preheating
Shinji Tsujikawa, Bruce A. Bassett and Fermin Viniegra 288

Chaotic inflation model in supergravity
Masahide Yamaguchi, M. Kawasaki and T. Yanagida 296

Evolution of Black Holes in Brans-Dicke Cosmology
Nobuyuki Sakai and John D. Barrow 302

Double inflation in supergravity and the primordial black hole formation
Toshiyuki Kanazawa, M. Kawasaki and T. Yanagida 308

Q-ball formation and its cosmology
Shinta Kasuya 317

Cognition of the Universe through a Model
Masafumi Seriu 325

Screw Instability in Black-Hole Magnetospheres
Tomoya Matsuoka, Akira Tomimatsu and Masaaki Takahashi 332

Variability in GRBs and Internal Shocks
Ryo Takagi and Shiho Kobayashi 338

Neutrino Pair Annihilation above a Kerr Black Hole
Katsuaki Asano and Takeshi Fukuyama 343

A Generalized Equilibrium Condition in the Electro-Gravi Dynamics in 1+1 Dimensions
Tadayuki Ohta and Robert Mann 351

Gamma Ray Bursts and Internal Shocks
Shiho Kobayashi 356

Multiple Sub-Jet Model of Gamma Ray Bursts and Possible Origin of X-ray Precursors and Post-cursors
Takashi Nakamura 362

Gravitational Waves from the Giant Flares of the Soft Gamma-ray Repeaters
Kunihito Ioka 368

Origin of Gamma-Ray Bursts – From the viewpoint of iron line observations –
Toshio Murakami and Daisuke Yonetoku 373

Ultra-High Energy Cosmic Rays and Superheavy Particles in the Universe
V. A. Kuzmin 379

Magnetic Effects on the Moment of Inertia of a Relativistic Star
Kohkichi Konno 386

Fireball generation by MHD jets
Masaaki Takahashi 394

General Relativistic Simulations of Jet Formation in a Rapidly Rotating Black Holes
Shinji Koide, David L. Meier, Kazunari Shibata and Takahiro Kudoh 402

Possible Direct Method to Determine the Radius of a Star
Takashi Nakamura and Motoyuki Saijo 410

Time evolution of purely axial displacements in isentropic neutron stars
Yasufumi Kojima 417

R-mode oscillations of differentially and rapidly rotating Newtonian polytropic stars
Shigeyuki Karino, Yoshiharu Eriguchi and Shin'ichirou Yoshida 424

Gravitational Radiation Reaction

Yasushi Mino[1]

Department of Physics, Graduate School of Science,
University of Tokyo, Hongo 7-3-1, Bunkyo-ku 113-0033, Japan

Abstract

In this talk, we discuss an analytic method to derive the gravitational radiation reaction force of a point particle orbiting around a Schwarzschild/Kerr blackhole using a technique of linear metric perturbation. It is well-known that the gravitational potential diverges when we use a point source (Dirac's delta) as a source, and the physical information for the orbital evolution by its radiating self-gravity is buried in this divergent potential as was argued by Mino, Sasaki and Tanaka[1], and, later, by Quinn and Wald[2]. For the calculation of the orbital evolution, it is necessary to formulate a practical regularization method. We here point out that the gauge condition is very important in the regularization[1][2].

1 Introduction

With the present technological advance, we come to realize the gravitational wave observation by using the laser interferometry. Among various projects proposed world-wide, LISA project is remarkable for it is a space-based detector. Space-based gravitational wave detectors are free from the seisemic noise, and gravitational waves of several Hz are detectable by those detectors. One of the promising targets are solar-mass compact objects inspiralling into a supermassive blackhole at a galactic center, and the gravitational evolution of this system is our goal of investigation.

Because of the strong gravity of the system, we must take into account the general relativistic effect in this investigation. One of the approach is so called standard post-Newtonian expansion technique, however, the calculation is expected to be very complicated since the objects rotate for number of times around the supermassive blackhole. Therefore, we consider that the blackhole perturbation is the best technique to calculate the relativistic effect to the orbital evolution of this system.

A linear perturbation formalism of a blackhole is expected to be useful for this calculation of a high precision. In fact, a number of papers have been pulished, calculating the energy loss and angular loss of the particle motion around the blackhole[4]. In these papers, one calculate the asymptotic behavior of the metric perturbation induced by the partcle, and, using the energy conservation fomula, one derives the energy and angular momentum loss of the particle.

However, the energy and angular momentum are not sufficient to characterize the orbital motion around a Kerr blackhole. Since the Kerr blackhole is axial-symmetric, the orbit of a monopole particle has three constants of motion; energy, angular momentm and Carter constant. It is essential to know the adiabatic evolution of these three constants of motion in order to construct the theoretical templates for the gravitational wave observation. It is also expected to make a theoretical prediction of an adiabatic orbital evolution of a particle with various multipole structure, but, it is a far reaching problem if we can only use the conservation formula for the energy and angular momentum.

It is our goal to calculate the adiabatic orbital evolution of a general orbit of a particle with multi-pole structure around a Kerr blackhole, and here we propose an idea to implement this calculation.

It seems apparent that the idea using the conservation formula can no more be workable. The construction of conservation formula depends more or less on the symmetry of the background geometry. The energy and angular momentum conservation formula is derived by using the stationarity and axial-symmetry of the background geometry. There is some implication that the conservation formula of Carter constant might be related with a symmetry of the background geometry with a super-symmetric extension, however, this idea is not yet successfully implemented. The discussion about the spinning

[1] E-mail:mino@utap.phys.s.u-tokyo.ac.jp

particle is far from the practice since we do not have the constants of motion which can completely determine the motion.

For this reasoning, we consider the direct derivation of the equation of motion with the effect of the particle's self-gravity.

The equation of motion of a 'point particle' moving in a general curved background was argued by Mino, Sasaki and Tanaka[1], and later by Quinn and Wald[2]. In the former paper, the equation motion was derived by solving Einstein equation with a technique of so-called matched asymptotic expansion. In Ref.[1], it is assumed that the particle induces blackhole-like gravitational field around it with the scale much less than the background curvature scale at the location of the particle. This assumption is the case of the inspiralling stage of a binary system of a solar-mass compact object and a supermassive blackhole at a galactic center.

The discovery of Ref.[1] and Ref.[2] is that the equation of motion can be derived from the linear metric perturbation induced by the 'point particle' which is defined by Dirac's delta in the background spacetime. Though the result is natural if one believes the equivalence principle, it is striking as the linear metric perturbation around the 'point particle' diverges and the linear approximation can no more be correct. These papers also noted that the singular part of the metric perturbation irrelevant to the equation of motion can be derived by the generalized Hadamerd prescription[6], which is exactly similar to the electromagnetic case[6]. The rest of the metric perturbation is called tail part[6] since this part is generated from the source inside the lightcone.

Though the argument of the equation of motion in Ref.[1] and Ref.[2] seems useful in the present problem, the practical method to calculate the tail part of the metric perturbation was not discussed, and we give an possible idea on this problem.

The direct derivation of the tail part seems very difficult, and we consider to derive it by extracting the singular part from the full metric perturbation as

$$h^{tail}(z) = \lim_{x \to z} \left(h^{full}(x) - h^{sing.}(x) \right). \tag{1}$$

The well-known derivation of a full metric perturbation is Zerilli-Regge-Wheeler formalism for a Schwarzschild blackhole[7], or Chrzanovski's method for a Kerr blackhole[8]. Here we note two technical problems in this calculation.

One is *Extraction Problem.* A well-known formula derives the full metric perturbation which is decomposed into an infinite sum of Fourier-Harmonic series. Here each mode functions are finite, and we have no explicit divergence in this expression. On the other hand, the sigular part derived by the Hardmard prescription can be derived just around the location of the particle, and the divergence is explicit in the resulting expression. The possible strategies to extract the singular part from the full metric perturbation are; 1) to calculate the infinite series summation of the full metric perturbation, or 2) to decompose the singular part into a sum of Fourier-Harmonic series. Later, we discuss both of these possible strategies.

The other difficulty is *Gauge Problem.* As we discuss, the singular part of the metric perturbation irrelevant to the equation of motion is derived by the Hadamard prescription. The Hadamard prescription is well-defined only when the differential equation is hyperbolic[5], therefore, the possible condition for the metric perturbation is constrained, such as the harmonic gauge. In fact, we shall show that if we ignore the gauge condition for the metric perturbation, we have a residual divergence after the extraction of the singular part. The possible strategies here are; 1) to calculate the gauge transformation of the full metric perturbation, or 2) to calculate the gauge transformation of the singular part. We shall discuss both strategies in this paper.

2 Extraction Problem

2.1 Mode-by-Mode extraction

We discuss the extraction by decomposing the Fourier-Harmonic mode.

Besides the gauge problem we shall discuss in the next section, it is expected that the full metirc

perturbation is derived as a sum of Fourier-Harmonic series as

$$h^{full}_{\mu\nu}(x) = \int d\omega \sum_{lm} R_{lm\omega}(r)\mathcal{Y}_{lm\omega}(r,\theta,\phi)\frac{1}{\sqrt{2\pi}}e^{-i\omega t}, \tag{2}$$

where we adopt Boyer-Lindquist coordinates $x^\alpha = \{t, r, \theta, \phi\}$ and $\mathcal{Y}_{lm\omega}(r,\theta,\phi)$ is a tensor spheroidal harmonics.[2]

If the Fourier-Harmonic decomposotion of the singular part is possible as

$$h^{sing.}_{\mu\nu}(x) = \int d\omega \sum_{lm} R^{sing.}_{lm\omega}(r)\mathcal{Y}_{lm\omega}(r,\theta,\phi)\frac{1}{\sqrt{2\pi}}e^{-i\omega t}, \tag{3}$$

the extraction is implemented as

$$h^{tail}_{\mu\nu}(x) = \int d\omega \sum_{lm} \left(R^{full}_{lm\omega}(r) - R^{sing.}_{lm\omega}(r)\right)\mathcal{Y}_{lm\omega}(r,\theta,\phi)\frac{1}{\sqrt{2\pi}}e^{-i\omega t}. \tag{4}$$

But the problem in this strategy is the Fourier-Harmonic decomposition of the singular part is not defined exactly. The singular part which is irrelevant to the equation of motion is derived by using bi-tensor[1] as

$$h^{sing.}_{\mu\nu}(x) = \frac{\mu}{4\pi}\bar{g}_{\mu\alpha}(x, z_{ret.})\bar{g}_{\nu\beta}(x, z_{ret.})v^\alpha(\tau_{ret.})v^\beta(\tau_{ret.}). \tag{5}$$

where $\bar{g}_{\mu\alpha}(x,z)$ is a bi-tensor of paralell transport. $z_{ret.}$ is the particle's location at the retarded time $\tau_{ret.}$, which is defined as a function of x by the past lightcone condition. (See Ref.[1] in detail.) The explicit calculation of the related bi-tensors on the whole spacetime is impossible in fact since the these bi-tensors are not uniquely defined when the separation of two field point x and z is order of the curvature length. One can only evaluate these bi-tensors as an expansion of the curvature. Therefore the singular part of the metric perturbation can be derived locally around the particle's location. However, the exact Fourier-Harmonic decomposition is defined only when (5) is defined whole time t and each spheroidal shell (whole θ, ϕ).

The idea here is we only decompose the local singularity buried in (5). We consider to derive the force at the article's location z^α. We calculate the local singularity of (5) around this point. For the field point $x^\alpha = z^\alpha + y^\alpha$, (5) is evaluated as

$$h^{sing.}_{\mu\nu}(x) = A^{(-1,0)}_{\mu\nu}|y|^{-1} + A^{(-2,1)}_{\mu\nu|\alpha}|y|^{-2}y^\alpha + A^{(-3,2)}_{\mu\nu|\alpha\beta}|y|^{-3}y^\alpha y^\beta + A^{(0,0)}_{\mu\nu}|y|^{-1} + A^{(-1,1)}_{\mu\nu|\alpha}|y|^{-2}y^\alpha + \cdots \tag{6}$$

where $|y| = \sqrt{g_{\alpha\beta}(z)y^\alpha y^\beta}$. Since we take the coincidence limit $y^\alpha \to 0$ after the extraction (4), it is enought to calculate the expansion coefficients $A^{(a,b)}$ for all $a+b<2$, and we can ignore the terms of $a+b>1$. We find this calculation is possible, and we also find that the Fourier-Harmonic decomposition of these terms is possible. (In fact, one can recursively construct the formulae of the Fourier-Harmonic decomposition for arbitrary (a,b).)

2.2 Mode-Sum extraction

We consider the extraction by decomposing the power of r instead of the Fourier-Harmonic series. As was argued in Ref.[9]), the calculation is possible only when the weak field limit (slow motion approximation) is possible, however, by using the systematic calculation technique of the radial mode functions as in Ref.[10]), one can calculate to an arbitrary order of expansion.

By using the result of Ref.[10]), we can derive the full metric perturbation analytically as a power expansion as

$$h^{full}_{\mu\nu}(x) = \sum_n h^{full(n)}_{\mu\nu} r^n. \tag{7}$$

[2] 'Tensor spheroidal harmonics' are not well defined objects as tensor spherical harmonics or spheroidal harmonics. They appear when we use the Crzanovski's method in generating the homogeneous vector potential and tensor potential, therefore, they are objects which could be defined only in a Kerr background.

(In some case, we encounter the non-integer power. Then using the technique in Ref.[9], one can obtain the integer power.) Under the weak field limit, it is argued in Ref.[9] that the sum over the Fourier-Harmonic series is easy.

On the other hand, the singular part of the metric perturbation is locally obtained, and one can take the power expansion of the local singularity as

$$h_{\mu\nu}^{sing.}(x) = \sum_n h_{\mu\nu}^{sing.(n)} r^n . \tag{8}$$

Since the power expansion unique, we can extract the power coefficients as

$$h_{\mu\nu}^{tail}(x) = \sum_n \left(h_{\mu\nu}^{full(n)} - h_{\mu\nu}^{sing.(n)} \right) r^n , \tag{9}$$

and calculate the force.

3 Gauge Problem

3.1 Regge-Wheeler gauge

This strategy applies to the case that the background is a Schwarzschild blackhole.

The derivation of the metric perturbation was argued in Ref.[7], and the radial functions can be derived both numerically and analytically.

The problem in this approach is the derivation of the singular part. The derivation of the singular part in Ref.[1] and in Ref.[2] is not useful since the singular part was derived by a generalized Hadamard prescription in which the hyperbolicity of the equation is essential[5]. Though the Zerilli-Regge-Wheeler equation looks hyperbolic, it is not clear whether the resulting metric perturbation satisfies the hyperbolic equation.

In fact, there seems to be a reason to say it might be non-hyperbolic. Suppose we have a metric perturbation in the harmonic gauge condition $\mathbf{h}^{harmonic}$. One can see $\mathbf{h}^{harmonic}$ satisfies the hyperbolic equation. We define the gauge transformation to the Regge-Wheeler gauge as

$$\mathbf{h}^{Regge-Wheeler} = \mathbf{h}^{harmonic} + \nabla\xi .$$

The gauge equation for the gauge transformation was argued in Ref.[7], and satisfies a first differential equation

$$(\partial + V)\xi = \mathbf{S}[h^{harmonic}] .$$

Since the source term S is a function of the metric perturbation in the harmonic gauge, it is not apparent whether it has a hyperbolic structure or not as a whole. However, since this is a first differential eqation, we do not have the freedom as the hyperbolic system should have. Say, if one defines the no-ingoing-wave boundary condition at the infinity, one cannot put the no-ingoing-wave boundary condition at the horizon at the same time that one puts the continuity condition of the gauge equation. Because of the arbitrary character of the boundary condition, the gauge is not made causal, therefore, the gauge equation is not hyperbolic. As a whole, the metric perturbation in the Regge-Wheeler gauge does not satisfy the hyperbolic equation.

However the question arises that as the metric perturbation in the Regge-Wheeler gauge is derived from the Zerilli-Regge-Wheeler functions alone, the equation must be hyperbolic. The trick is that we do not solve the gauge equation with continuity condition at the radius where there is a matter. This non-continuity compensates the arbitrary character of the boundary condition in the gauge equation. Thus, we can conclude that the metric perturbation in the Regge-Wheeler gauge is made causal, but, the equation is not necessarily hyperbolic.

We find that there is a way to derive the hyperbolic equation for the metric perturbation in the Regge-Wheeler gauge with some local term. It seems possible to implement the regularization calculation with

this equation to derive the tail part of the metric perturbation, however, it is not guaranteed to apply the Hadamard prescription in an exact sense (though we did not discuss in this talk).

The other option is to use the Green function. We find that the gauge is continuous at the radius where there is no matter. The Green function satisfies

$$(\nabla^2 + R)\mathrm{G}(x,z) = \delta(x-z)\,.$$

One can see there is no source outside the coincidence $x^r = z^r$, and one can solve the gauge equation continuously at $x^r > z^r$ and at $x^r < z^r$, assuming the appropriate boundary condition. Using the singular part of the metric perturbation in the harmonic gauge in Ref.[1] and in Ref.[2], one can calculate the gauge transformation and derive the singular part of the metric perturbation in the Regge-Wheeler gauge.

3.2 Radiation gauge

This strategy applies to the case that the background is a Kerr blackhole.

The most serious problem in this approach is that we have no way to calculate the complete metric perturbation in the radiation gauge condition yet. The metric perturbation in a Kerr background was investigated in Ref.[8] in a very nice way. It was also discussed that, by using the Green function approach, the inhomogenous solution can be derived, however, this is true only at the vacuum. Ref.[8] found that the homogeneous metric perturbation can be derived by operating derivatives on the homogeneous Teukolsky functions

$$\mathrm{h}^{homo.} \;=\; \bar{\mathrm{h}}R^{homo.}\mathcal{Y}_{lm}\frac{1}{\sqrt{2\pi}}e^{-i\omega t}\,.$$

They consider that the tensor Green function can be derived as a mode-sum of the homogeneous solution in a usual way

$$\mathcal{G}(x,z) \;=\; \sum_{lm\omega}\left(\mathrm{h}^{homo.(up)}(x)\mathrm{h}^{homo.(in)}(z)\theta(x^r-z^r) + \mathrm{h}^{homo.(in)}(x)\mathrm{h}^{homo.(up)}(z)\theta(z^r-x^r)\right).$$

We see whether we can recover the Green function of the Teukolsky function by operating derivatives. We find that there are some extra terms with delta functions as

$$\tilde{R}\cdot\int dz\mathcal{G}(x,z)\mathcal{T}(\natural) = \int dzG(x,z)S(z) + (\mathcal{D}+V)T(x)\,,$$

where $G(x,z)$ is the Green function of the Teukolsky function.

The emergence of this terms owes to the gauge condition. As in the case of Regge-Wheeler gauge, one can see the gauge equation is not made hyperbolic, therefore, the gauge must be solved discontinuously. This discontinuity is related with the source term of the Einstein equation, which we do not have yet. In fact, in the Zerilli-Regge-Wheeler formalism of the metric perturbation[7], we have a local term in the tensor Green function

$$\begin{aligned}\mathcal{G}(x,z) \;=\; \sum_{lm\omega}\Big(&\mathrm{h}^{homo.(up)}(x)\mathrm{h}^{homo.(in)}(z)\theta(x^r-z^r) + \mathrm{h}^{homo.(in)}(x)\mathrm{h}^{homo.(up)}(z)\theta(z^r-x^r)\\ &+\mathcal{S}(x^r)\delta(x^r-z^r)\Big)\,.\end{aligned}$$

As for the singular part of the metric perturbation, the problem is almost same in the case of the Regge-Wheeler gauge. However, the first option is of no use since we do not have the way to derive the metric perturbation from the Teukolsky function.[3] The second option might be possible if one can derive the gauge equation.

[3] However, recently Ref.[11] discusses this approach.

3.3 Harmonic gauge

The harmonic gauge condition has a surprisingly nice property. Since the metric perturbation satisfies the hyperbolic equation and the continuity condition, the Green function can be constructed as a sum of homogeneous metric perturbation even in the case that the background is a Kerr blackhole.

The crucial problem is there is no effective way to derive the homogeneous metric perturbation in the harmonic gauge. We consider to derive the metric perturbation in the harmonic gauge condition by calculating the gauge equation, then the gauge equation here is a vector hyperbolic equation, and can be reduced into a set of Zerilli-Regge-Wheeler type equations or a set of Teukolsky equations. However, these sets of radial equations are inhomogeneous with extended source terms. Though these are just ordinary differential equations, the source terms are not guaranteed to damp out enough earlier, and we have no effective way to integrate with enough precision.

The possible idea is to integrate these radial equations around the particle's orbit in an adiabatic manner. Then the gauge is not harmonic in an exact sense, but the resulting metric perturbation satisfies the harmonic gauge condition just around the particle.

Though it is not exact, we consider that this is enough for the calculation since the singular part of the metric perturbation depends the local geometric quantities, not on the asymptotic structure. The ambiguity from the in-exact gauge condition is just a finite gauge, therefore, we expect that it does not modifies the orbital evolution after the long time integration.

4 Summary

Our final goal is to establish the calculation method of the local gravitational radiation reaction force to a point particle moving around a Kerr blackhole. We pointed out that there are two problems here; 'extaction problem' and 'gauge problem'. We consider that now the gauge problem of gravitational radiation reaction is far more serious than the extraction problem, and we point out several problems around the gauge below;

1) *no Einstein equation for the metric perturbation*

We have a linearized Einstein equation for the metric perturbation, however, it does not make sense as we do not have the gauge fixing term in the Regge-Wheeler gauge condition or the Radiation gauge condition. The problem is that without the equation for the equation of the metric perturbation, one can hardly establish the regularization scheme as in Ref.[1]. Axiomatic approach as in Ref.[2] might be promising, however, we have no way to confirm the resulting scheme.

One could consider the regularization with the Regge-Wheeler equation or the Teukolski equation, however, in this case, the formulation of the regularization scheme might be of the priority though we have another problem as follows.

2) *delicate decompling of the singular part and the tail*

One may consider that the Regge-Wheeler equation or Teukolsky equation would help since we can calculate the metric perturbation from its solution. The problem here is that the direct (tail) part of the Regge-Wheeler/Teukolski Green function has no relation with the direct (tail) part of the Green function of the metric perturbation.

The tail part is written such as

$$\int^{\tau_x} d\tau v(x, z(\tau)) s(\tau) ,$$

where $z(\tau)$ is the orbit of the point source with the parameter τ, τ_x is the orbital parameter with past light cone condition, hence it is a function of the field point x, $v(x, z)$ is from the tail part of Green function, and $s(\tau)$ is from the source term of the equation. If we operate x-derivative on this integration, we have a finite contribution from the light cone as

$$(\nabla_x \tau_x) v(x, z(\tau_x)) s(\tau_x) ,$$

which looks like a direct part.

3) *break down of hyperbolicity*

The crucial problem in formulating the regularization scheme in Regge-Wheeler gauge or Radiation gauge is that we are not guaranteed to apply the generalized Hadamard prescription as in Ref.[6] and in Ref.[1]. In the previous regularization formalism in gravity[1], we use the harmonic gauge condition, and the derivative operator in the linearized Einstein equation becomes hyperbolic, in which case we are allowed to use the generalized Hadamard prescription. On the other hand, the linearized Einstein equation in Regge-Wheeler gauge or Radiation gauge, is not hyperbolic though the resuting metric is made causal, therefore, we are not guaranteed to apply the Hadamard prescription.

Without Hadamard prescription, we have no way to calculate the metric perturbation in a local manner, satisfying the local retarded condition. It seems almost impossible to formulate the regularization scheme in Regge-Wheeler or Radiation gauge. Even if the regularization scheme is possible to be formulated, it would be something far different from the one in Ref.[1].

4) *no complete derivation of metric perturbation in a Kerr background*

Besides the difficulty in formulating the regularization scheme, we do not have a convenient way to calculate the metric perturbation in a Kerr background.

The possibility of solving the gauge problem is to use the quasi-harmonic gauge condition[4] One can easily find it is not inconsistent with the previous result using the energu balance equation[4].

The extraction strategy could be both mode-by-mode and mode-sum regularization, which seem equally efficient, but, in any way, we have great amount of calculation.

Acknowledgements

We thank all the members of the theoretical gravitational-wave physics group in Japan and all the members of the theoretical astrophysics group at Osaka University and Univeristy of Tokyo for fruitful conversations. Special thanks are due to Prof. K. Sato, Prof. H. Takahara, Prof. Y. Suto, Prof. S. Yamada, Prof. M.Shibata, Prof. T. Tanaka, Dr. H. Tagoshi, Dr. T. Shiromizu, Dr. T. Chiba, Dr. A. Taruya, G. Uchida, H. Nakano and Y. Himemoto for invaluable discussions and encouragement. We thank to Prof. K. Thorne for his hospitality and discussion during our stay at CalTech, We thank to Prof. B. Schutz for his hospitality and discussion during his stay at Einstein Institute at Potsdam.

The authors are supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists, No.0704.

References

[1] Y. Mino, M. Sasaki, and T. Tanaka, Phys. Rev.**D55**(1997)3457; Y. Mino, M. Sasaki, and T. Tanaka, Prog. Theor. Phys. Supp.**128**(1997)373.

[2] T. C. Quinn and R. M. Wald, Phys. Rev.**D56**(1997)3381.

[3] Y. Mino and H. Nakano, Prog. Theore. Phys.**100**(1998)507.

[4] Y. Mino, M. Sasaki, M. Shibata, H. Tagoshi and T. Tanaka, Prog. Theor. Phys. Supp.**128**(1997)1.

[5] J. Hadamard, "Lectures on Cauchy's Problem in Linear Partial Differential Equations." Yale Univ. Press, New Haven, Connecticut,1923.

[6] B. S. DeWitt and R. W. Brehme, Ann. Phys.(N.Y.)**9**(1960)220.

[7] F. J. Zerilli, Phys. Rev.**D2**(1970)2141.

[8] P. L. Chrzanowski, Phys. Rev.**D11**(1975)2042.

[9] Y. Mino andH. Nakano, Prog. Theore. Phys.**100**(1998)507.

[10] S. Mano, H. Suzuki and E. Takasugi, Prog. Theor. Phys. **95**(1996)1079.

[11] H. Nakano and M. Sasaki, gr-qc/0010036.

[4] we say 'quasi' since it is not exact, but is still harmonic around the particle.

Search for Gravitational Waves from Inspiraling Compact Binaries using TAMA300 data

Hideyuki Tagoshi

Department of Earth and Space Science,
Osaka University, Toyonaka, Osaka 560-0043, Japan

and TAMA collaboration

Abstract

We analyzed the data from the TAMA300 detector taken in September 2000. We searched for gravitational waves from inspiraling compact binaries using matched filtering. We obtained an upper limit of 0.59/hour (C.L.=90%) on the event rate of inspirals of compact binaries with mass between $0.3M_\odot$ and $10M_\odot$ and with signal-to-noise ratio greater than 7.2. The distance of $1.4M_\odot$ ($0.5M_\odot$) binaries which produce the signal-to-noise ratio 7.2 was estimated to be 6.2kpc (2.9kpc) when the position of the source on the sky and the inclination angle of the binaries were optimal.

1 Introduction

Several laser interferometric gravitational wave detectors are now under construction. These include LIGO[1], VIRGO[2], GEO600[3], and TAMA300[4]. The TAMA300 detector has been developed over the past five years. It is a power-recycled, Fabry-Perot-Michelson interferometer, which consists of mirrors that are suspended by vibration isolation systems. The differential armlength changes in the 300m Fabry-Perot arm cavities caused by propagation of gravitational waves are monitored by the Michelson interferometer.

The TAMA300 detector became ready to operate in the summer of 1999. Most of the designed systems (except power recycling) were installed by that time. The first test operation was done on August 6th 1999. The first long-term data were taken for three nights between September 17th and 20th 1999 [4][5]. The total data length amounted to about 30 hours, with the longest continuous lock time of the interferometer lasting nearly 8 hours. The best sensitivity was about $3 \times 10^{-20}/\sqrt{\mathrm{Hz}}$ around 900Hz. The data of the first two days contained many burst noises due to the instability of the interferometer system and it was very difficult to use those data for the gravitatioal wave event search. However, the quality of the data of the final day was good enough to perform the search for gravitational waves.

2 Inspiraling binaries

Gravitational waves from inspiraling compact binaries, consisting of neutron stars with mass $\sim 1.4M_\odot$ or black holes, have been considered to be the most promising target for laser interferometers. These compact binaries can be produced as a consequence of normal stellar evolution of binaries. It has been also suggested that MACHOs[6] in our Galactic halo may be primordial black holes with mass $\sim 0.5M_\odot$. If so, it is reasonable to expect that some of them are in binaries which coalesce due to the gravitational radiation reaction[7].

Inspiraling compact binaries are nearly ideal sources for data analysis, since their waveforms can be theoretically calculated to a very high accuracy. However, in this paper, we did not take account of the effect of spin angular momentum. Thus, we cannot deny the possibility that we lose signals from binaries with spin [8].

3 Matched filtering

In this paper, we do not explain the method to calibrate the data which was explained by Telada et al.[9] We denote the strain equivalent one-sided power spectrum density of the noise by $S_n(f)$. In order to calculate the expected wave forms, which are called *templates*, we used restricted post-Newtonian wave forms of order 2.5, in which the phase evolution was correctly taken into account up to the 2.5 post-Newtonian order, but the amplitude was evaluated by using the quadrupole formula. As for the 2.5 post-Newtonian phase evolution, we used formulas derived by Blanchet et al.[10]

When the gravitational wave passes through the interferometer, it produces a relative difference ΔL between the two armlengths L. The gravitational wave strain amplitude is defined by $\Delta L = Lh(t)$. The wave form $h(t)$ is calculated by combining two independent modes of the gravitational wave and the antenna pattern of the interferometer as

$$h(t) = \mathcal{A}[h_c(t - t_c)\ \cos\alpha + h_s(t - t_c)\ \sin\alpha], \tag{1}$$

where t_c is the coalescence time, and $h_c(t)$ and $h_s(t)$ are the two independent templates with the phase difference $\pi/2$. To construct filters, we need the Fourier transforms of $h_c(t)$ and $h_s(t)$. They were computed directly by using the stationary phase approximation, the validity of which was established by Droz et al.[11] The parameters to distinguish the wave forms are the amplitude $\mathcal{A}$, the two masses m_1, m_2, the coalescence time t_c and the phase α. We did not include spins of the stars in the parameters.

We denote the data from the detector as $s(t)$. We define a filtered unnormalized signal-to-noise ratio ρ after the maximization over α as

$$\rho = \sqrt{(s,h_c)^2 + (s,h_s)^2}, \tag{2}$$

$$(a,b) \equiv 2\int_{-\infty}^{\infty} df \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(|f|)}, \tag{3}$$

where $\tilde{a}(f)$ denotes the Fourier transform of $a(t)$ and the asterisk denotes the complex conjugation. This ρ has an expectation value $\sqrt{2}$ in the presence of only Gaussian noise. Thus, the signal-to-noise ratio, SNR, is given by SNR$= \rho/\sqrt{2}$.

Analyzing the real data from TAMA300, we found that the noise contained a large amount of non-stationary and non-Gaussian noise whose statistical properties have not been understood well yet. In order to remove the influence of such noise, we introduced a χ^2 test of the time-frequency behavior of the signal [12]. We divide each template into n mutually independent pieces in the frequency domain, chosen so that the expected contribution to ρ from each frequency band is approximately equal. For two template polarizations $h_c(t)$ and $h_s(t)$, we calculate χ^2 by summing the square of the deviation of each value of ρ from the expected value[13]. This quantity must satisfy the χ^2-statistics with $2n-2$ degrees of freedom, as long as the data consists of Gaussian noise plus chirp signals. However, there was a strong tendency that an event with large χ^2 has a large value of ρ. Thus, by applying a threshold to the χ^2 value, we can reduce the number of fake events without significantly losing the detectability of real events. For convenience, we renormalized χ^2 as $\chi^2/(2n-2)$. In the current analysis, we chose $n = 16$. This number was determined mainly by the limited memory of our computer.

We searched the parameter space of $0.3M_\odot \leq m_1, m_2 \leq 10M_\odot$ with total mass less than $10M_\odot$. The low mass limit was chosen so that it covers the estimated mass of MACHOs as much as possible. In this parameter space, we prepared a mesh. The mesh points define the templates used for search. The spacing of the mesh points was determined so as not to lose more than 2 % of signal-to-noise due to the mismatch between actual mass parameters and those at mesh points. Using geometrical arguments, we introduced a new parameterization of masses that simplifies the algorithm to determine the mesh points[15].

The parameter space defined by using our new mass parameters turned out to contain 2057 templates in the present analysis. Although this number is not too large to perform matched filtering with a simple algorithm, when the noise power spectrum of data is improved, the necessary number of templates will increase. Thus, we introduced a two-step hierarchical search algorithm and developed tools of matched filtering which can be used to analyze future much larger data streams.

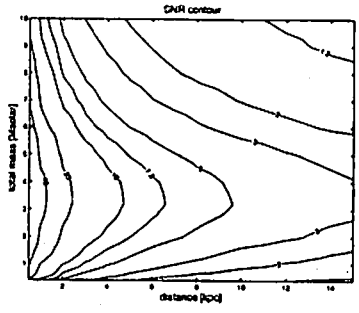


Figure 1: A contour plot of SNR as a function of distance and total mass of equal mass binaries in the case when the position of the source on the sky and the inclination angle of the binaries are optimal.

4 Two-step hierarchical search

The basic idea of the two-step search is simple. The templates for the first step search are chosen at the mesh points with coarse spacing. We perform the first step search with an appropriate threshold which must be chosen to be sufficiently low not to lose real events. When there are events that exceed the threshold, we search the points around them in detail.

Before starting the two-step search, we evaluated the values of ρ and χ^2 of all the data with a few selected templates to obtain an estimate of the background distribution of ρ and χ^2, assuming all of these events are due to noise. Based on this distribution, we looked for several combinations of (ρ^*, χ^{*2}) so that there were only a few fake events which satisfy $\rho > \rho^*$ and $\chi^2 < \chi^{*2}$ simultaneously. These thresholds became our preliminary second step thresholds. The first step threshold and the first step length of the mesh separation were determined to satisfy the following condition (C1): *When ρ and χ^2 of an event at the second step mesh point satisfies a second step threshold, such an event must also be detected at, at least, one of the first step mesh points around the second step point with more than 98 % probability.* By imposing this condition, it is guaranteed that the false dismissal rate of real detectable events due to the presence of the first step search is kept to be small[14]. Note that the preliminary second step thresholds, (ρ^*, χ^{*2}), which were used to decide the first step thresholds, determined the detection probability of events with given amplitude.

We performed simulations by adding signals to real data, and looked for combinations of the first step thresholds and the mesh separation which satisfy the criterion (C1). We chose one of them which minimizes the computation time in total. We introduced, at the first step, the χ^2 threshold as well as the ρ threshold to select the candidates for the second step, which reduced the computation time for the second step.

Note that we introduced several combinations of ρ and χ^2 threshold. For larger ρ threshold, we use larger χ^2 threshold. If the signal with large ρ was located at a point far from the mesh points, the value of χ^2 could become larger than unity. Thus, this method is useful to detect the signal with large ρ located between the mesh points.

Here, we also mention the cut off frequency f_c of templates. The post-Newtonian wave forms must be terminated at some frequency roughly corresponding to the inner most stable circular orbit, at which the binaries are expected to transit to plunge orbit to coalesce. However, it is difficult to determine, theoretically, the optimal value of f_c. At the first step search, the choice of f_c is important because the mesh separation is large. If the signal was located at a point far from the first step mesh points, the value of χ^2 could become much larger than unity due to the difference of f_c between the first step template and the signal. Thus, at the first step search, we searched for the value of f_c which produced the largest value of ρ. On the other hand, at the second step search, we simply adopted $f_c = c^3/(7^{3/2}\pi GM)$ where M is the total mass, G is the Newton's constant, and c is the speed of light. This corresponds to the orbital radius of $r \sim 7GM/c^2$.

In order to reduce the computation time, we further introduced various techniques, some of which are explained in Tanaka and Tagoshi.[15]

5 Results

Our analysis was done with 8 Compaq Alpha machines. Each of them can perform the FFT of data consisting of $\sim 10^6$ single precision real numbers with 140MFlops.

From the 30 hours of total data, we selected the data with better quality taken between 14:42 and 21:04 19th September 1999 (UTC time). The data consist of two separate stretches when the interferometer was continuously locked. An interval of 12.5 minutes around an unlocked part was not used for matched filtering. The total length of data used for matched filtering was 6.1663 hours.

In Fig.1, we show a contour plot of the signal-to-noise ratio as a function of the distance to the source and the total mass for equal mass binaries. We found that the signal-to-noise ratio becomes maximum around $1.6M_\odot$ binaries and decreases above this mass. This is because the noise level of the detector increases rapidly below several hundreds Hz.

In performing the two-step search, the data were divided into data segments of 265.42 seconds with an overlap of 29.9 seconds, which was longer than the longest template. The strain power spectrum density $S_n(f)$ was estimated using 530.84 seconds of data near the segment which was analyzed.

As a result of the second step search, we obtain ρ and χ^2 as functions of masses and the coalescing time t_c. In each small interval of coalescing time Δt_c, we looked for an event which had the maximum ρ and which satisfies certain χ^2 threshold. In Fig.2, we show the numbers of these events which survived after the maximization as a function of SNR for various χ^2 thresholds. The time interval Δt_c was chosen so that the numbers of correlated events were small enough, but at the same time there was a sufficiently large number of events to investigate the properties of its distribution. In Fig.2, we chose $\Delta t_c = 2$ seconds.

The computation time needed to perform the two-step search depends strongly on the first step threshold. If we had optimized the first step threshold by applying the above condition (C1), we would have obtained the threshold given in Table I(b) in the case when the unit area of the parameter space of the first step search is 72 times larger than that of the second step. In that case, the computation time would have been about 1.5 hours. However, we used a lower first step threshold as given in Table I(c) to obtain the result of Fig.2, in order to investigate the background distribution which was used for statistical analysis. The computation time thus needed to obtain the results in Fig.2 was about 5.5 hours. In this analysis, we used a fixed set of the first step thresholds and spacing for all of the data. Thus, the number of events which satisfy the first step threshold, and therefore total computation time, differs very much among the portion of data.

Using this result, we estimate an upper limit on the event rate. In Fig.2, we show a fitting to the number of events as a function of SNR^2. The value of the χ^2 threshold, 1.5, was chosen so that there was a 3.8% chance of rejecting real events in Gaussian noise. We chose the SNR threshold as SNR= 7.2. Thus, the observed number of events which exceed this threshold is 2. Using this analytic fitting, we estimate the expected number of background events N_{BG} which is larger than SNR= 7.2 as $N_{BG} = 2.5$. Thus, using Bayesian statistics, and assuming uniform prior probability for the real event rate and the Poisson distribution of real and background events, we estimate with 90% confidence that the expected number of real events which exceed SNR= 7.2 is smaller than 3.68 in this data[16]. Thus, we obtained the upper limit of the event rate as 0.59/hour (C.L.= 90%). Note that we did not assume that the observed 2 events were real or not. (The largest SNR event for $\chi^2 < 1.5$ has SNR=9.0 and $\chi^2 = 1.35$. Although this event seems to show a small excess from the background distribution for $\chi^2 < 1.5$, it also seems to be absorbed in the background for $\chi^2 < 2.5$.)

From Fig.1, we find the distance of a source that would produce SNR = 7.2 as 2.9kpc for $m_1 = m_2 = 0.5M_\odot$, and 6.2kpc for $m_1 = m_2 = 1.4M_\odot$ respectively. In this paper, we estimate the distance in the case when the position of the source on the sky and the inclination angle of the binaries were optimal. In order to evaluate the efficiency, we performed simulations. We added artificial signals to the data and performed the same two-step search which derived Fig.2. We then evaluated the detection probability of artificial signals with the final thresholds determined in Fig.2. The result is given in Fig.3. We found that TAMA300 could observe $1.4M_\odot$ events in several kpc. In other words, TAMA300 with the present sensitivity can give a direct observational limit of the coalescing rate within several kpc. Here, it is important to note again that a new detector, TAMA300, produces data which are good enough to perform data analysis and to give upper limit of the event rate, even if it is sensitive only to events within several kpc.

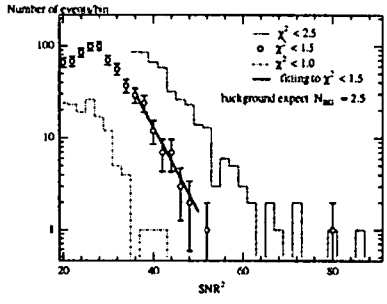


Figure 2: The number of events for $\chi^2 < 2.5$, $\chi^2 < 1.5$ and $\chi^2 < 1.0$ as a function of SNR^2, and an analytic fitting to $\chi^2 < 1.5$. The fitting was determined between SNR^2=35 and 50. This fitting gives the number of events larger than SNR=7.2 and $\chi^2 < 1.5$ as 2.5. The $\chi^2 < 2.5$, $\chi^2 < 1.5$, and $\chi^2 < 1.0$ correspond to 10^{-3}%, 3.8%, and 46% false dismissal rate in Gaussian noise, respectively.

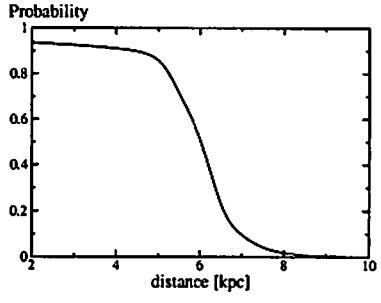


Figure 3: The detection probability of artificial signals of $1.4M_\odot$ binaries as a function of distance in the case when the final thresholds were (SNR^*, χ^{*2})= (7.2,1.5).

The sensitivity of TAMA300 detector is now being improved rapidly. In August 2000, 160 hours data with improved sensitivity was taken. With final goal sensitivity of TAMA300, $1.4M_\odot$ ($0.5M_\odot$) binaries at distance D = 820kpc (D = 340kpc) will produce a signal-to-noise ratio 10. By analyzing such data, we will be able to obtain a more stringent upper limit to the event rate. And if we are very lucky, we may find a plausible candidate for a real gravitational wave signal.

Acknowledgments

This work was supported by Monbusho Grant-in-Aid for Creative Basic Research 09NP0801. This work was also supported in part by Monbusho Grant-in-Aid 11740150, 12640269, and by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists.

References

[1] A.Abramovici et al., Science **256**, 325 (1992).

[2] B.Caron et al., in *Gravitational Wave Experiments*, edited by E. Coccia, G. Pizzella, and F. Ronga, (World Scientific, Singapore, 1995).

[3] K. Danzmann et al., in *Gravitational Wave Experiments*, edited by E. Coccia, G. Pizzella, and F. Ronga, (World Scientific, Singapore, 1995).

[4] For details about the first data taking of TAMA300 project, see *Gravitational Wave Detection II, (Proceedings of the 2nd TAMA workshop on Gravitational Wave Detection)*, (Univ. Acad. Press, Tokyo, 2000).

[5] M. Ando et al., in preparation.

[6] C. Alcock et al., Astrophys. J. **486**, 697 (1997); C. Alcock et al., *ibid.* **542**, 281 (2000).

[7] T. Nakamura, M. Sasaki, T. Tanaka, and K.S. Thorne, Astrophys. J. Lett. **487**, L139 (1997).

[8] T.A. Apostolatos, Phys. Rev. **D52**, 605 (1995).

[9] S.Telada et al., in ref.4.

[10] L. Blanchet, T. Damour, B. Iyer, C.M. Will, and A.G. Wiseman, Phys. Rev. Lett. **74**, 3515 (1995); L. Blanchet, Phys. Rev. **D54**, 1417 (1996).

[11] S. Droz, D.J. Knapp, E. Poisson, and B.J. Owen, Phys. Rev. **D59**, 124016 (1999).

[12] B. Allen et al. Phys. Rev. Lett. **83**, 1498 (1999).

[13] For more details of χ^2 test, see B. Allen et al. "GRASP software package", release 1.9.4, section 6.24, http://www.lsc-group.phys.uwm.edu/~ballen/grasp-distribution/ .

[14] Note the difference of this criterion to that discussed in S.D. Mohanty, Phys. Rev. **D57**, 630 (1998).

[15] T. Tanaka, and H. Tagoshi, Phys. Rev. **D62**, 082001 (2000).

[16] For example, Particle Data Group, *Review of Particle Physics*, Phys. Lett. **B204** (1988), 81; *ibid.* Euro. Phys. J. **C3**, (1998), 172-177, and references cited there.

Gravitational Memory ?
— a Perturbative Approach —

Tomohiro Harada[1] ,

Department of Physics, Waseda University,
Shinjuku, Tokyo 169-8555, Japan

B.J. Carr[2] , and C.A. Goymer[3]

Astronomy Unit, Queen Mary and Westfield College,
University of London, London E1 4NS, England

Abstract

It has been pointed out that the value of the gravitational "constant" in the early universe may be different from that at present. In that case, it was conjectured that primordial black holes may "remember" the value of the gravitational constant in the early universe. The present analysis shows that this is not the case, at least in certain contexts.

1 Introduction

In the early universe, black holes may have been formed due to inhomogeneous initial conditions, phase transitions, or other mechanisms. They are called primordial black holes (PBHs). The mass of the PBHs is of order the mass contained within the Hubble horizon at the formation epoch. Hawking [1] discovered that black holes emit radiation due to the quantum effects of the curved spacetime. PBHs therefore lose mass and those lighter than $\sim 10^{15}$g may contribute to the cosmic gamma-ray background and spoil the success of the big bang nucleosysthesis scenario. PBHs heavier than that may dominate the density of the universe at present. Therefore the fraction β of the universe which may have gone into PBHs can be constrained from various cosmological observations (Carr [2]).

It has been pointed out that gravity in the early universe may have been dilatonic, since this arises naturally as a low-energy limit of string theory. In such theories, the gravitational "constant" is given by a function of a scalar field which couples non-minimally with gravity. In particular, the gravitational constant in the early universe may have been different from that at the present epoch. The Brans-Dicke theory is the simplest such theory of gravity. Barrow [3] considered PBHs in Brans-Dicke theory and proposed two scenarios. In scenario A, the gravitational constant at the black hole event horizon is always the same as that at the cosmological particle horizon, i.e., $G_{EH}(t) = G_{PH}(t)$. In scenario B, the gravitational constant at the black hole event horizon is constant with time and therefore always the same as that at the formation epoch, i.e., $G_{EH}(t) = G_{EH}(t_f) = G_{PH}(t_f)$. Scenario B corresponds to what is called the *gravitational memory*, while scenario (A) entails no gravitational memory.

Barrow and Carr [4] obtained observational constraints on β in both scenarios. They assumed that the energy emission due to the Hawking radiation is determined by the temperature $T_H = (8\pi G_{EH}(t)M)^{-1}$, so that the mass loss rate is $\dot{M} \simeq -(G_{EH}(t)M)^{-2}$. The results for the two scenarios can be significantly different. Subsequently, Carr and Goymer [5] argued that the true situation may be intermediate between Scenario A and Scenario B and this allows other possibilities. Therefore, in order to deduce the history of the early universe from the present observations, it is very important to examine which scenario is realized.

Recently, Jacobson [6] has argued that there is no gravitational memory for a Schwarzschild black hole with a time-varying boundary condition which mimics the cosmological evolution of a scalar field.

[1] E-mail:harada@gravity.phys.waseda.ac.jp
[2] E-mail:B.J.Carr@qmw.ac.uk
[3] E-mail:C.A.Goymer@qmw.ac.uk

However, this argument applies only if the black hole is much smaller than the particle horizon, so we must consider more general situations. The purpose of this article is to solve the problem self-consistently for a Friedmann-Robertson-Walker (FRW) background.

2 Basic Equations

The field equations in Brans-Dicke theory are given by

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi}{\phi}T_{\mu\nu} + \frac{\omega}{\phi^2}\left(\phi_{,\mu}\phi_{,\nu} - \frac{1}{2}g_{\mu\nu}\phi_{,\alpha}\phi^{,\alpha}\right) + \frac{1}{\phi}\left(\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\Box\phi\right), \tag{1}$$

$$\nabla_\mu T^{\mu\nu} = 0, \tag{2}$$

$$\Box\phi = \frac{8\pi}{3+2\omega}T^\mu{}_\mu, \tag{3}$$

where the constant ω is the Brans-Dicke parameter and ϕ is the Brans-Dicke scalar field, which is related to the gravitational constant by $G \sim \phi^{-1}$. In the limit $\omega \to \infty$, the theory goes to general relativity with a minimally coupled scalar field. The constraint $\omega \gtrsim 3300$ is required by recent observations of light deflection [7]. However, if we consider more general scalar-tensor theories, ω also becomes a function of ϕ. In this case, ω may have been small in the early universe, even though it is large today.

Here we calculate the time evolution of the gravitational constant by using the following post-general-relativistic expansion. First we assume that the scalar field is constant, i.e., $\phi = \phi_0$. Then we find the general relativistic solution (g_{GR}, T_{GR}) which satisfies Eqs. (1) and (2) with $G = \phi_0^{-1}$. Next we put $(g, T) = (g_{GR}, T_{GR})$ in Eq. (3) and assume that $\delta\phi = \phi - \phi_0$ is of order ω^{-1} for $\omega \gg 1$. We can then obtain ϕ to this order, denoting it as $\phi^{(1)}$, by solving the following equation:

$$\Box_{GR}\phi^{(1)} = \frac{8\pi}{3+2\omega}T_{GR}{}^\mu{}_\mu, \tag{4}$$

where $\Box_{GR}$ denotes the d'Alembertian in the geometry given by g_{GR}. Substituting the solution $\phi^{(1)}$ of Eq. (4) into the right-hand side of Eq. (1), we can then determine $g^{(1)}$ and $T^{(1)}$ by solving Eqs. (1) and (2). Then, $\phi^{(2)}$, which is the solution up to $O(\omega^{-2})$, is determined by solving Eq. (3) for the background $(g^{(1)}, T^{(1)})$. Repeating this process, we can construct the Brans-Dicke solution from the general relativity solution order by order.

Here we truncate the expansion for the scalar field at $O(\omega^{-1})$. This implies that we only have to solve Eq. (4) and we can neglect the effect of the scalar field on the background curvature. This approximation was shown to be very good for the generation and propagation of gravitational waves from the collapse to a black hole in asymptotically flat spacetime for Brans-Dicke theory with $\omega \gtrsim 4$ [8].

3 Lemaitre-Tolman-Bondi solution

We adopt the Lemaitre-Tolman-Bondi model as the background solution (g_{GR}, T_{GR}) in which the scalar field evolves. This model is an exact solution of general relativity which describes a spherically symmetric inhomogeneous universe with dust. The line element in the synchronous comoving coordinates is given by

$$ds^2 = -dt^2 + A^2(t,r)dr^2 + R^2(t,r)d\Omega^2, \tag{5}$$

where

$$d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2. \tag{6}$$

The function $R(t,r)$ is given by

$$R = \left(\frac{9F(r)}{4}\right)^{1/3}(t - t_s(r))^{2/3} \quad \text{for} \quad f = 0, \tag{7}$$

$$\begin{cases} R = \dfrac{F(r)}{2f(r)}(\cosh\eta - 1) \\ t - t_s(r) = \dfrac{F(r)}{2f^{3/2}(r)}(\sinh\eta - 1) \end{cases} \quad \text{for} \quad f > 0, \tag{8}$$

$$\begin{cases} R = \dfrac{F(r)}{2(-f)(r)}(1 - \cos\eta) \\ t - t_s(r) = \dfrac{F(r)}{2(-f)^{3/2}(r)}(\eta - \sin\eta) \end{cases} \quad \text{for} \quad f < 0, \tag{9}$$

where the prime and dot denote the derivatives with respect to r and t, respectively. The function $A(t,r)$ is given by

$$A^2(t,r) = \frac{(R')^2}{1+f(r)}. \tag{10}$$

The functions $t_s(r)$, $f(r) > -1$ and $F(r) \geq 0$ are arbitrary and one of them corresponds to the gauge freedom. The first two relate to the big bang time and the total energy per unit mass for the shell at radius r. The third relates to the mass within radius r, the density of the dust being given by

$$\rho = \frac{F'}{8\pi R^2 R'}. \tag{11}$$

4 Method

We are interested in the behavior of the scalar field long after the black hole has formed. In this case, the characteristic method is suitable. This method was first applied to the Lemaitre-Tolman-Bondi background by Iguchi, Nakao and Harada [9]. If we introduce the retarded time u, the line element becomes

$$ds^2 = -\alpha(u,\bar{r})du^2 - 2\alpha(u,\bar{\alpha})A(u,\bar{r})du d\bar{r} + R^2(u,\bar{r})d\Omega^2, \tag{12}$$

where α must satisfy

$$\frac{\partial}{\partial\bar{r}}\alpha = \dot{A}\alpha. \tag{13}$$

In terms of the derivatives, d/du and $\partial/\partial\bar{r}$, the d'Alembertian is given by

$$\Box\phi = -\frac{2}{\alpha A R}\frac{d\varphi}{du} - \frac{A'}{A^3 R}\varphi + \frac{1}{AR}\left[(A\dot{R})^{\cdot} - \left(\frac{R'}{A}\right)'\right]\phi, \tag{14}$$

$$\varphi = \frac{\partial}{\partial\bar{r}}(R\phi). \tag{15}$$

We can then integrate Eq. (4) along the characteristic curves, i.e., null geodesics. With this choice of coordinate system, we can calculate the behavior of the scalar field long after black hole formation without having to impose boundary conditions on the event horizon.

5 Models

As we have seen, the Lemaitre-Tolman-Bondi solution has three arbitrary functions. We want to obtain models which describe a PBH in a flat FRW universe by fixing these functions appropriately. Here we adopt the following assumptions: the big bang occurs at the same time everywhere; the model is asymptotically flat FRW, which requires the overdense region to be compensated by a surrounding underdense region; the model is free from naked singularities. In order to satisfy these assumptions, we first choose

$$t_s(r) = 0, \tag{16}$$

and

$$f(r) = \begin{cases} -\left(\frac{r}{r_c}\right)^2 & \text{for} \quad r < r_w \\ -\left(\frac{r}{r_c}\right)^2 \exp\left[-\left(\frac{r-r_w}{r_w}\right)^4\right] & \text{for} \quad r \geq r_w \end{cases}, \tag{17}$$

where r_c and r_w are the curvature scale and the size of the overdense region. Using the arbitrary function $F(r)$, we then set the radial coordinate r to be

$$r = R(t_0, r) \tag{18}$$

where $t = t_0$ is the initial time.

Before integrating Eq. (4), we have to fix the initial data for ϕ, so we choose the cosmological homogeneous solution given by

$$\phi = \phi_0 \left(1 + \frac{1}{3+2\omega}\frac{4}{3}\ln\frac{t}{t_0}\right). \tag{19}$$

We set the initial null hypersurface as the null cone whose vertex is at $(t, r) = (t_0, 0)$ and regard the cosmological solution as the initial data on this hypersurface. Although the value of the scalar field at the cosmological particle horizon must be given by this solution, the value in the perturbed region may be different from this. We have therefore examined the effects of other initial data which are different from the cosmological solution in the perturbed region. However, the results are not changed much.

6 Results

We adopt units in which the asymptotic value of the Hubble parameter at $t = t_0$ is unity. We set the parameters which specify the initial data as $r_c = 2$ and $r_w = 1$, while the Brans-Dicke parameter is chosen as $\omega = 5$. It is noted that r_c must satisfy $r_c \gtrsim r_w$, else the overdense region is isolated from the rest of the universe, as Carr and Hawking [10] pointed out.

The results are shown in Figs. 1-3. In Figs. 1 (a) and (b), a set of initial data for the background geometry is shown. The initial density perturbation δ is defined as

$$\delta(t, r) \equiv \frac{\rho(t, r) - \rho(t, \infty)}{\rho(t, \infty)}. \tag{20}$$

In Fig. 2, the trajectories of outgoing light rays in this background geometry are shown. It is shown that the event horizon is formed at $R \simeq 10$.

The results of integrating Eq. (4) are shown in Figs. 3(a) and (b), where $\Delta\phi$ is defined as

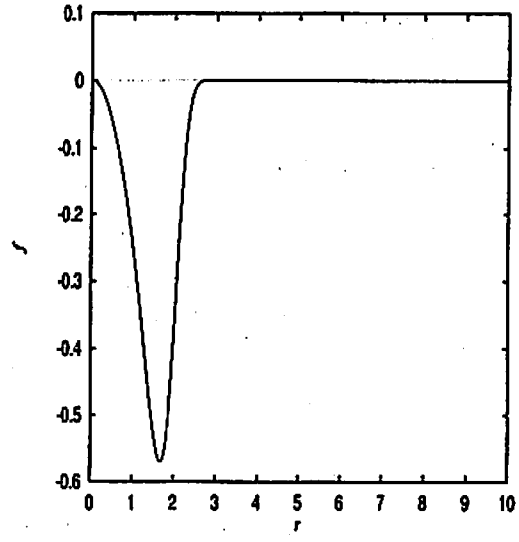
$$\Delta\phi \equiv \frac{\phi(t, r) - \phi_0}{\phi_0}. \tag{21}$$

It is found that the value of the scalar field around the black hole is slightly smaller than the asymptotic value because of the underdensity of the surrounding region. Nevertheless, the scalar field as a whole is almost spatially homogeneous at each moment, in spite of the existence of the black hole. The evolution of the scalar field near the event horizon is well described by the cosmological homogeneous solution.

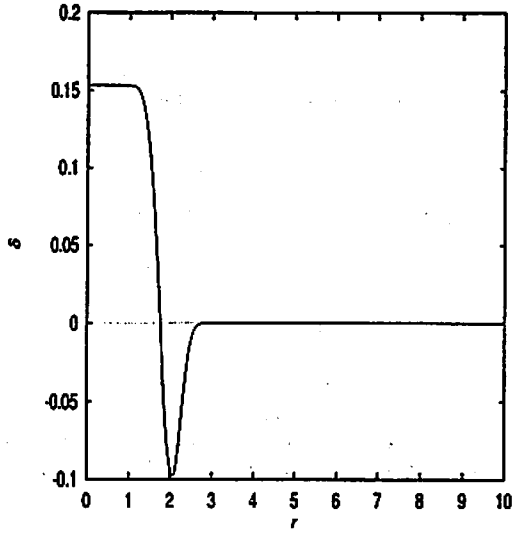
7 Summary

We have calculated the evolution of the Brans-Dicke scalar field in the presence of a primordial black hole formed in a flat FRW universe. We have found that the value of the scalar field at the event horizon almost maintains the cosmological value at each moment. This suggests that primordial black holes "forget" the value of the gravitational constant at their formation epoch. However, it should be stressed that this result has only been demonstrated for a dust universe in which the scalar field does not appreciably affect the background curvature. It remains to be seen whether the same conclusion applies when these assumptions are dropped.

We are grateful to T. Nakamura for helpful discussions and useful comments.



(a)



(b)

Figure 1: (a) Specific energy $f(r)$ and (b) initial density perturbation δ at $t = t_0$.

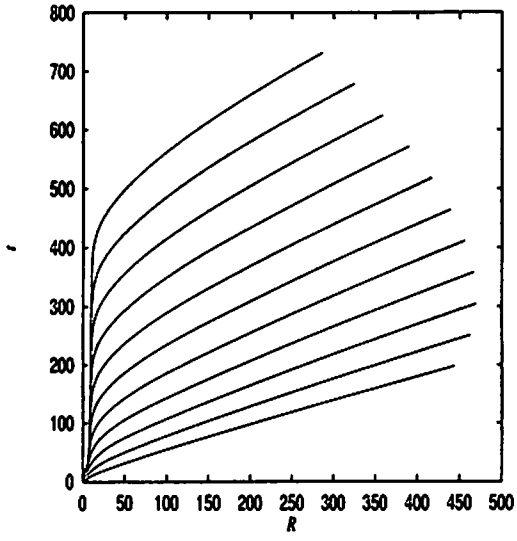


Figure 2: Trajectories of outgoing light rays in the background geometry.

References

[1] S.W. Hawking, Commun. Math. Phys. **43**, 199 (1975).

[2] B.J. Carr, Astrophys. J. **201**, 1 (1975).

[3] J.D. Barrow, Phys. Rev. D **46**, 3227 (1992).

[4] J.D. Barrow and B.J. Carr, Phys. Rev. D **54**, 3920 (1996).

[5] B.J. Carr and C.A. Goymer, Prog. Theor. Phys. **136**, 321 (1999); C.A. Goymer and B.J.Carr, Phys. Rev. D., in press.

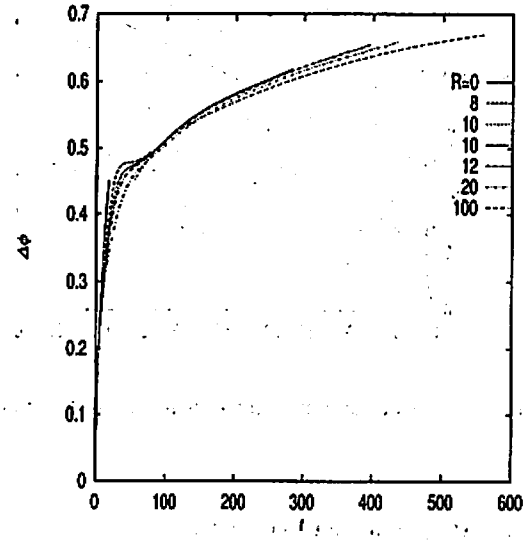
[6] T. Jacobson, Phys. Rev. Lett. **83**, 2699 (1999).

[7] T.M. Eubanks et al., Bull. Am. Phys. Soc., Abstract #K 11.05 (1997); C.M. Will, gr-qc/9811036.

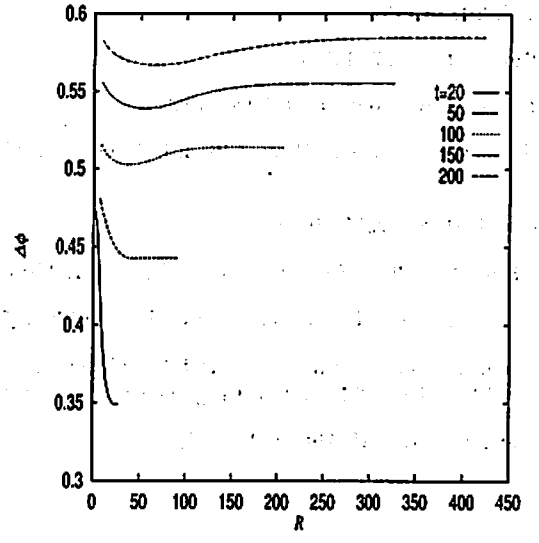
[8] T. Harada, T. Chiba, K. Nakao and T. Nakamura, Phys. Rev. D **55**, 2024 (1997).

[9] H. Iguchi, K. Nakao and T. Harada, Phys. Rev. D **57**, 7262 (1998); H. Iguchi, T. Harada and K. Nakao, Prog. Theor. Phys. **101**, 1235 (1999); Prog. Theor. Phys. **103**, 53 (2000).

[10] B.J. Carr and S.W. Hawking, MNRAS **168**, 399 (1974).



(a)



(b)

Figure 3: $\Delta\phi \equiv (\phi - \phi_0)/\phi_0$ (a) along the world lines $R = \text{const}$ and on the hypersurfaces $t = \text{const}$.

Quasi-normal mode of massive scalar field : Analytic model approach

OHASHI AKIRA[1] and SAKAGAMIi MASA'AKI[2]

Department of Physics, FIHS,
Kyoto University, Kyoto, 606-8501, Japan

Abstract

We study quasi-normal modes of massive scalar field. In this report, we analyze a model which can be solved analytically and we derive an analytical expression of QNMs frequencies. This analysis make us to be able to investigate behaviours of QNMs in detail.

1 Introduction

Quasi-normal mode (QNM) is one of the important topics on black hole physics. Phenomenologically, QNMs have been studied as important source of direct detection on the earth of gravitational waves emitted by coalescing compact object binaries which are one of main targets of detectors of gravitational waves. On theoretical sides, QNMs are closely connected with Green functions of fields and, therefore, QNMs also have been studied in relation to quantum physics on black hole spacetime such as Hawking radiation, vacuum fluctuation of fields and moreover AdS-CFT corresponding [1].

There are many studies on QNMs, but almost all of them were based on consideration of massless fields on black hole spacetime. Recently, however Tomimatsu and Koyama [2] reported an interesting phenomenon which seemed to be a resonance phenomenon of a quantum massive scalar field on a black hole spacetime in their paper on vacuum polarization of massive fields. Their analysis was based on quantum theory, but as written above, because of close connection between QNMs and Green function of fields, it is expected that there may be some phenomenon in classical physics like QNMs on black hole spacetime.

These above two points are the main motivations of our study. At first, we researched about QNMs of massive scalar fields on Reissner-Nordström black hole spacetime (RNBH) by a numerical method [3]. In that study, we found some interesting behaviours of QNMs. Next, in order to understand qualitatively those behaviours, we analyzed some models which can be solved analytically [3]. In this report, we explain the latter model analysis and we only refer some numerical results for comparison.

2 Quasi-normal mode of massive field

In this report, we consider mainly a Shrödinger-type equation of a field $\psi \sim e^{-i\omega t}\Psi(r)$ because the radial equation in case of RNBH can be transformed into the following form:

$$\left[\frac{d^2}{dr_*^2} + \omega^2 - V(r)\right]\Psi(r) = 0, \tag{1}$$

[1]E-mail:ohashi@phys.h.kyoto-u.ac.jp
[2]E-mail:sakagami@phys.h.kyoto-u.ac.jp

where r_* represents the tortoise coordinate.

The definition of QNM of massless fields is that 1)QNMs are damping modes with time, 2)QNMs have only outgoing wave components at the spatial infinity and 3)QNMs have only ingoing wave components at the horizon of the black hole. These conditions can be represent as

$$\mathrm{Im}\omega < 0, \tag{2}$$

$$R(r) \sim e^{i\omega r} \qquad (r \to \infty), \tag{3}$$

$$R(r) \sim e^{-i\omega r_*} \qquad (r_* \to -\infty), \tag{4}$$

respectively. Although these conditions are valid with no ambiguities in the case of massless field, they are no longer clear conditions when we consider massive fields. It is because 1)there is no definition of "outgoing" at the spatial infinity in massive field case and 2)there are another possible modes which are damping at the spatial infinity. In the rest of this section we explain these problems.

First, we consider what "outgoing" means. Because the potential have two different values at the horizon and the spatial infinity, $\Psi(r)$ behaves in different ways such as,

$$\Psi \sim e^{-i\omega r_*} \qquad \text{near the horizon,} \tag{5}$$

$$\Psi \sim e^{-ivr} \qquad \text{at the infinity,} \tag{6}$$

where $v^2 = \omega^2 - m^2$. If both ω and v are real number, we can define simply that waves are "outgoing" when ω and v have a same sign. However because QNMs are modes with complex frequencies, we cannot adopt the same definition. Moreover when we use v, it is not clear that we should choose which branch of v. Including this ambiguity, we define "outgoing" such that waves with complex frequencies are "outgoing" if the real parts of the frequencies is "outgoing" (see the above condition in the case of real frequencies). This condition can be written in the following form:

$$\mathrm{sgn}(\mathrm{Re}\omega) = \mathrm{sgn}(\mathrm{Re}v). \tag{7}$$

After a short consideration, it can be shown that this condition is equivalent to

$$\mathrm{Im}v < 0 \tag{8}$$

under the condition of $\mathrm{Im}\,\omega < 0$. We will see that this condition is completely complement with a condition of bounded solutions of the equation and thus the above condition is regarded as a natural definition.

Next problem is the existence of damping solutions. These solutions do not exist in the case of massless fields. Because there is an offset in the potential $V(r)$ at the spatial infinity, the existense of such damping modes is expected. These solutions may be also QNMs but they belong to a different series of the "usual" QNMs. Therefore for distinguishing the two kinds of QNMs, we call these new QNMs as BQNMs (Bounded QNMs) and we call the "usual" QNMs as NQNMs (Normal QNMs). The condition of BQNM is almost clear and it is

$$\mathrm{Im}v > 0. \tag{9}$$

We can see this condition is complmentary with the condition of NQNM (8).

In our previous numerical study on the QNM of a massive scalar field on the RNBH space-time, we found some interesting behaviour of QNMs. In this report, however, we will only refer the some results of the numerical study.

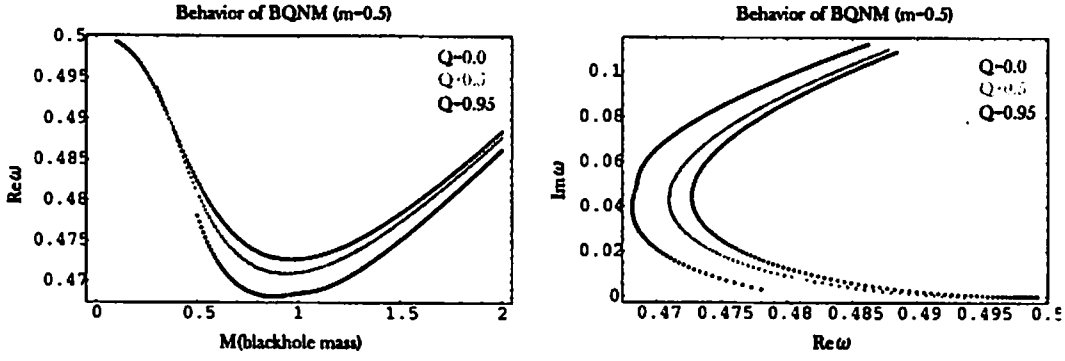


Figure 1: Behaviours of BQNM. These figures show the behaviour of BQNM under the condition that we fix the mass of the scalar field and change the mass of the black hole. Q is the charge of the black hole.

3 Hyperbolic tangent model

In our previous work, we found some interesting behaviour of QNMs, but because of the accuracy of numerical calculation we could not make all physics of the behaviour clear. One of the interesting behaviour is that when the mass of the scalar field is increasing, the imaginary part of NQNM frequencies is increasing (note that the imaginary part of QNM must be negative) and at a certain value of the mass the signature of the imaginary part seems to be changed (see the left figure of Fig.(3). This means that such frequencies do not satisfy the QNM condition, thus they are no longer QNM. However we could not check the such vanishing behaviour of NQNM by our numerical calculations. Another interesting behaviour is that the trajectory of the BQNM is not simple (Fig.1). However because of a certain problem we will not tell about BQNMs in this report.

For the purpose of understanding the above behaviour, we analyzed an analytically solvable potential model. In this analysis, we approximate the potential profile (Fig.2) by an resemble one and derive the QNMs frequencies analytically. This modeling approach was already adopted by Blome and Mashhoon in the case of massless fields and they obtained some adequate results [4].

In this report, our goal is to understand qualitatively the behaviour of QNMs and therefore we neglect the some detail of the profile of the potential but we consider an essentially similar model(Fig.2). Our model is the following:

$$\left[\frac{d^2}{dx^2} + \omega^2 - V(x)\right]\Psi(x) = 0, \tag{10}$$

$$V = \frac{1}{4}k^2 u(1-u) + m^2 u, \tag{11}$$

$$u = \frac{1}{2}(1 + \tanh\frac{d}{2}x). \tag{12}$$

We can solve this model analytically by using the hyper-geometrical function $F(\alpha, \beta, \gamma; u)$ such as,

$$\Psi(x) = u^a(1-u)^b F(\alpha, \beta, \gamma; u), \tag{13}$$

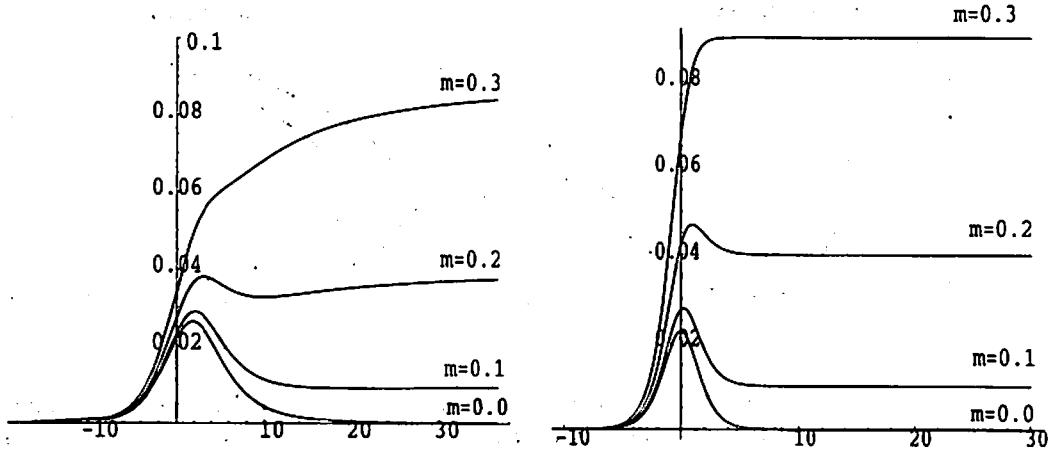


Figure 2: Profiles of potentials. The left graph shows the potential of a massive scalar field on the Schwartchild spacetime with mass $m = 0.0,\ 0.1,\ 0.2,\ 0.3$. The right one represents the potential of the analytically solvable model.

where α, β and γ are given by

$$\alpha = \frac{1}{2}\left\{1 + 2a + 2b + \sqrt{k^2 - \left(\frac{M}{d}\right)^2}\right\}, \tag{14}$$

$$\beta = \frac{1}{2}\left\{1 + 2a + 2b - \sqrt{k^2 - \left(\frac{M}{d}\right)^2}\right\}, \tag{15}$$

$$\gamma = 1 + 2a, \tag{16}$$

and a, b are defined by

$$a^2 d^2 + \omega^2 = 0, \tag{17}$$

$$b^2 d^2 + \omega^2 - m^2 = 0. \tag{18}$$

By using the Gauss's transition formula, we can represent the above solution around $u = 0$ by a linear combination of solutions around $u = 1$ like

$$\begin{aligned} & u^a(1-u)^b F(\alpha,\beta,\gamma;u) = \\ & \quad \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)} u^a(1-u)^b F(\alpha,\beta,1+\alpha+\beta-\gamma;1-u) \\ & \quad + \frac{\Gamma(\gamma)\Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha)\Gamma(\beta)} u^a(1-u)^{-b} F(\gamma-\alpha,\gamma-\beta,1+\gamma-\alpha-\beta;1-u). \end{aligned} \tag{19}$$

The QNM condition at the horizon tell us

$$a = -\frac{i\omega}{d} \tag{20}$$

(here "horizon" corresponds $u = 0$, that is, $x = -\infty$). On the other hand, according to the new outgoing condition, when we set

$$b = \sqrt{\omega^2 - m^2}, \tag{21}$$

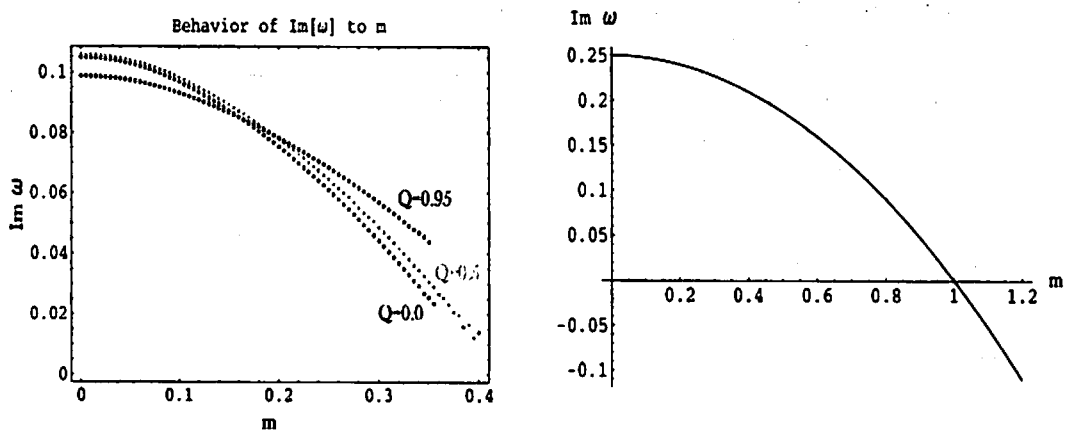


Figure 3: These figure show the behaviour of the imaginary part of NQNM frequencies. The left figure is one of the numerical results on RNBH QNM. It contains three cases: $Q = 0.0,\ 0.5,\ 0.95$ (Q is the charge of the BH) The right one is a result of analytically-solvable model analysis studied here. Both figure show a very similar behaviour and the analytical one makes the vanishing of NQNM clear.

(where the branch of $\sqrt{X}$ is chosen as $\sqrt{X} > 0$), we obtain that the second term of the equation (19) represents the outgoing wave and the first term corresponds to the bounded solutions. Therefore, NQNM and BQNM frequencies can be derived from

$$\frac{1}{\Gamma(\alpha)\Gamma(\beta)} = 0 \quad , \tag{22}$$

$$\frac{1}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)} = 0 \quad , \tag{23}$$

respectively. These equations can be written into one algebraic equation

$$\frac{1}{2}\left\{1 + 2a \pm 2b \pm \sqrt{k^2 - \left(\frac{M}{d}\right)^2}\right\} = -n, \qquad (n = 0, 1, 2, \cdots). \tag{24}$$

By solving this equation, we obtain the following analytic expression of QNM frequencies,

$$\omega = i\left(\frac{m^2}{k^2} - \frac{1}{4}\right) + \left(\frac{m^2}{k^2} + \frac{1}{4}\right)\sqrt{k^2 - 1}. \tag{25}$$

Here for simplicity we derive only the case of $n = 0$ and $d = 0$ Note that although four possibilities to choose the sign in the equation (24), we can obtain only one solution. This solution represnts the NQNM frequencies and there no solution corresponding to BQNM in this hyperbolic tangent model. We show the graphs of ω in the right figure in Fig.3.

4 Results and summary

In this report, we consider a potential model which can be solved analytically and calculate NQNM frequencies. As a result, we can confirm that NQNM will vanish when the mass of the

scalar field exceed a critical value. This phenomenon can be interpreted as like the following: when the mass of the field increases, the Compton wavelength of the field decreases smaller and then the field becomes to behave as particle rather than waves. Moreover we emphasize that there is a frequency that is not complex but real. At this frequency, the mode is not QNM. Moreover this mode exist without damping with time. Therefore we can regard this mode as a kind of resonance mode. Of course this mode does not exist in the case of massless field. This is a very interesting find of our work.

On the other hand, after some detailed calculations, we can show that there is no BQNMs in this analytically-solvable model. This fact seems to contradict the results of our numerical study (in the numerical calculation we have found BQNMs by adopting the real BH potential). However because the BQNM frequencies exist near the mass of the scalar field (this can be known from our previous numerical results), it is expected that the tail of the potential profiles may play an important role. Until now we have no exact proof that supports the above expectation, but we have some interesting evidences and will obtain a exact proof soon.

References

[1] G.T.Horowitz and V.E.Hubeny. Phys.Rev.**D62**,024027, (2000).

[2] A.Tomimatsu and H.Koyama. Phys.Rev.**D61**,124010, (2000).

[3] A.Ohashi and M.Sakagami. *in preparation*, (2001).

[4] H.Blome and B.Mashhoon. Phys.Lett.**100A**,231, (1984).

[5] E.W.Leaver. Proc.R.Soc.Lond.A **402**, 285.

[6] T.Nakamura and H.Sato. Phys.Lett.**61B**,371, (1976).

[7] B.F.Schutz and C.M.Will. ApJ.**291**,L33.

Energy-Momentum and Angular Momentum Densities in Gauge Theories of Gravity

Toshiharu Kawai[1]

Department of Physics, Osaka City Univeristy,
Sumiyoshi-ku, Osaka 558-8585, Japan

Abstract

In the Poincaré gauge theory of gravity, which has been formulated on the basis of a principal fiber bundle over the space-time manifold having the covering group of the proper orthochronous Poincaré group as the structure group, we examine the tensorial properties of the dynamical energy-momentum density ${}^{G}\mathbf{T}_{k}{}^{\mu}$ and the "spin" angular momentum density ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ of the gravitational field. They are both space-time vector densities, and transform as tensors under *global* $SL(2,C)$- transformations. Under *local* internal translation, ${}^{G}\mathbf{T}_{k}{}^{\mu}$ is invariant, while ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ transforms inhomogeneously. The dynamical energy-momentum density ${}^{M}\mathbf{T}_{k}{}^{\mu}$ and the "spin" angular momentum density ${}^{M}\mathbf{S}_{kl}{}^{\mu}$ of the matter field are also examined, and they are known to be space-time vector densities and to obey tensorial transformation rules under internal Poincaré gauge transformations. The corresponding discussions in extended new general relativity which is obtained as a teleparallel limit of Poincaré gauge theory are also given, and energy-momentum and "spin" angular momentum densities are are all space-time vector densities.

1 Introduction

The energy-momentum and angular momentum play central roles in modern theoretical physics. Local objects such as energy-momentum and angular momentum densities are well defined if gravitational fields are not present in the system in question.

In general relativity, however, the energy-momentum and angular momentum densities of the gravitational field so far proposed are not space-time tensor densities. Rather, it is usually asserted [1], on the basis of the equivalence principle, that well-behaved energy-momentum and angular momentum densities cannot be defined for the gravitational field, while total energy-momentum and total angular momentum are defined well for asymptotically flat space-time.

In the Poincaré gauge theory of gravity (PGT) [2], which has been formulated on the basis of principal fiber bundle over the space-time manifold having the covering group of the proper orthochronous Poincaré group as the structure group, we have defined dynamical energy-momentum and "spin" angular momentum densities. For the asymptotically flat space-time, the integration of the dynamical energy-momentum density over space-like surface σ is the generator of *internal* translation and gives the total energy-momentum of the system. Also, the integration of "spin" angular momentum density over σ is the generator of *internal* $SL(2,C)$-transformations and gives the *total* (=*spin*+*orbital*) angular momentum, when the Higgs-type field ψ^{k} is chosen to be $\psi^{k} = e^{(0)k}{}_{\mu}x^{\mu} + \psi^{(0)k} + O(1/r^{\beta})$ with constants $e^{(0)k}{}_{\mu}, \psi^{(0)k}$ [3, 4, 5]. In extended new general relativity (ENGR) which is obtained as a teleparallel limit of PGT, corresponding results have been obtained [6].

The purpose of this note is to examine, both in PGT and in ENGR, the transformation properties of the dynamical energy-momentum densities and "spin" angular momentum densities under general coordinate transformations and under Poincaré gauge transformations [7, 8]. The main result is that *all the dynamical energy-momentum densities and "spin" angular momentum densities in these theories are true space-time vector densities.* Discussions in "§20.4. WHY THE ENERGY OF THE GRAVITATIONAL FIELD CANNOT BE LOCALIZED" in Ref. [1] are not applicable to our theories.

[1] E-mail: kawai@sci.osaka-cu.ac.jp

2 Poincaré Gauge Theory

PGT is formulated on the basis of the principal fiber bundle $\mathcal{P}$ over the space-time M possessing the covering group $\bar{P}_0$ of the proper orthochronous Poincaré group as the structure group. The space-time is assumed to be a noncompact four-dimensional differentiable manifold having a countable base. The bundle $\mathcal{P}$ admits a connection Γ, whose translational and rotational parts of the coefficients will be written as $A^k{}_\mu$ and $A^k{}_{l\mu}$, respectively. The fundamental field variables are $A^k{}_\mu$, $A^k{}_{l\mu}$, the Higgs-type field $\psi = \{\psi^k\}$, and the matter field $\phi = \{\phi^A | A = 1, 2, 3, \ldots, N\}$ [9]. These fields transform according to [10]

$$\begin{aligned}
\psi'^k &= (\Lambda(a^{-1}))^k{}_l(\psi^l - t^l) , \\
A'^k{}_\mu &= (\Lambda(a^{-1}))^k{}_l(A^l{}_\mu + t^l{}_{,\mu} + A^l{}_{m\mu}t^m) , \\
A'^k{}_{l\mu} &= (\Lambda(a^{-1}))^k{}_m A^m{}_{n\mu}(\Lambda(a))^n{}_l + (\Lambda(a^{-1}))^k{}_m(\Lambda(a))^m{}_{l,\mu} , \\
\phi'^A &= (\rho((t,a)^{-1}))^A{}_B\phi^B ,
\end{aligned} \tag{1}$$

under the Poincaré gauge transformation

$$\sigma'(x) = \sigma(x) \cdot (t(x), a(x)) , \quad t(x) \in T^4 , \quad a(x) \in SL(2,C) . \tag{2}$$

Here, Λ is the covering map from $SL(2,C)$ to the proper orthochronous Lorentz group, and ρ stands for the representation of the Poincaré group to which the field ϕ^A belongs. Also, σ and σ' stand for local cross sections of $\mathcal{P}$. Dual components $e^k{}_\mu$ of vierbeins $e^\mu{}_k\partial/\partial x^\mu$ are related to the field ψ^k and the gauge potentials $A^k{}_\mu$ and $A^k{}_{l\mu}$ through the relation $e^k{}_\mu = \psi^k{}_{,\mu} + A^k{}_{l\mu}\psi^l + A^k{}_\mu$, and these transform according to $e'^k{}_\mu = (\Lambda(a^{-1}))^k{}_l e^l{}_\mu$, under the transformation (2). Also, they are related to the metric $g_{\mu\nu}dx^\mu \otimes dx^\nu$ of M through the relation $g_{\mu\nu} = \eta_{kl}e^k{}_\mu e^l{}_\nu$ with $(\eta_{kl}) \stackrel{\text{def}}{=} \text{diag}(-1,1,1,1)$.

There is a 2 to 1 bundle homomorphism F from $\mathcal{P}$ to affine frame bundle $\mathcal{A}(M)$ over M, and an extended spinor structure and a spinor structure exist associated with it [13]. The space-time M is orientable, which follows from its assumed noncompactness and from the fact that M has a spinor structure.

The affine frame bundle $\mathcal{A}(M)$ admits a connection Γ_A. The T^4-part $\Gamma^\mu{}_\nu$ and $GL(4,R)$-part $\Gamma^\mu{}_{\nu\lambda}$ of its connection coefficients are related to $A^k{}_{l\mu}$ and $e^k{}_\mu$ through the relations

$$\Gamma^\mu{}_\nu = \delta^\mu{}_\nu , \quad A^k{}_{l\mu} = e^k{}_\nu e^\lambda{}_l\Gamma^\nu{}_{\lambda\mu} + e^k{}_\nu e^\nu{}_{l,\mu} , \tag{3}$$

by the requirement that F maps the connection Γ into Γ_A, and the space-time M is of the Riemann-Cartan type.

The field strengths $R^k{}_{l\mu\nu}$, $R^k{}_{\mu\nu}$ and $T^k{}_{\mu\nu}$ of $A^k{}_{l\mu}$, $A^k{}_\mu$ and of $e^k{}_\mu$ are given by [11]

$$\begin{aligned}
R^k{}_{l\mu\nu} &\stackrel{\text{def}}{=} 2(A^k{}_{l[\nu,\mu]} + A^k{}_{m[\mu}A^m{}_{l\nu]}) , \\
R^k{}_{\mu\nu} &\stackrel{\text{def}}{=} 2(A^k{}_{[\nu,\mu]} + A^k{}_{l[\mu}A^l{}_{\nu]}) , \\
T^k{}_{\mu\nu} &\stackrel{\text{def}}{=} 2(e^k{}_{[\nu,\mu]} + A^k{}_{l[\mu}e^l{}_{\nu]}) ,
\end{aligned} \tag{4}$$

and we have the relation

$$T^k{}_{\mu\nu} = R^k{}_{\mu\nu} + R^k{}_{l\mu\nu}\psi^l . \tag{5}$$

The field strengths $T^k{}_{\mu\nu}$ and $R^k{}_{l\mu\nu}$ are both invariant under *internal* translations. The torsion is given by

$$T^\lambda{}_{\mu\nu} = 2\Gamma^\lambda{}_{[\nu\mu]} , \tag{6}$$

and the T^4- and $GL(4,R)$-parts of the curvature are given by

$$R^\lambda{}_{\mu\nu} = 2(\Gamma^\lambda{}_{[\nu,\mu]} + \Gamma^\lambda{}_{\rho[\mu}\Gamma^\rho{}_{\nu]}) , \quad R^\lambda{}_{\rho\mu\nu} = 2(\Gamma^\lambda{}_{\rho[\nu,\mu]} + \Gamma^\lambda{}_{\tau[\mu}\Gamma^\tau{}_{\rho\nu]}) , \tag{7}$$

respectively. Also, we have

$$T^k{}_{\mu\nu} = e^k{}_\lambda T^\lambda{}_{\mu\nu} = e^k{}_\lambda R^\lambda{}_{\mu\nu} , \quad R^k{}_{l\mu\nu} = e^k{}_\lambda e^\rho{}_l R^\lambda{}_{\rho\mu\nu} , \tag{8}$$

which follow from Eq. (3).

The covariant derivative of the matter field ϕ takes the form [12]

$$D_k\phi^A = e^\mu{}_k \left(\partial_\mu \phi^A + \frac{i}{2} A^{lm}{}_\mu (M_{lm}\phi)^A + i A^l{}_\mu (P_l \phi)^A \right) . \tag{9}$$

Here, M_{kl} and P_k are representation matrices of the standard basis of the group $\bar{P}_0$: $M_{kl} = -i\rho_*(\bar{M}_{kl})$, $P_k = -i\rho_*(\bar{P}_k)$. The matrix P_k represents the "intrinsic" energy-momentum of the field ϕ^A [13], and it is vanishing for all the observed fields.

The Lagrangian density

$$\bar{L} = L^M(e^k{}_\mu, \psi^k, D_k\phi^A, \phi^A) + \bar{L}^G(T^{klm}, R^{klmn}) \tag{10}$$

satisfies the requirement of $\bar{P}_0$ gauge invariance. Here, L^M is the Lagrangian density of the matter field $\phi = \{\phi^A\}$ and $\bar{L}^G$ is the Lagrangian density of the gauge potentials given by

$$\bar{L}^G \stackrel{\text{def}}{=} L^T + L^R + dR \tag{11}$$

with

$$L^T \stackrel{\text{def}}{=} c_1 t^{klm} t_{klm} + c_2 v^k v_k + c_3 a^k a_k , \tag{12}$$

$$L^R \stackrel{\text{def}}{=} d_1 A^{klmn} A_{klmn} + d_2 B^{klmn} B_{klmn} + d_3 C^{klmn} C_{klmn} + d_4 E^{kl} E_{kl} + d_5 I^{kl} I_{kl} + d_6 R^2 . \tag{13}$$

In the above, c_i, d_j $(i = 1,2,3, j = 1,2,3,\ldots,6)$ and d are all real constants, t_{klm}, v_k and a_k are the irreducible components of the field strength T_{klm}, and $A_{klmn}, B_{klmn}, C_{klmn}, E_{kl}, I_{kl}$, and R are the irreducible components of the field strength R_{klmn}. Their definitions are enumerated in Appendix.

The gravitational Lagrangian density $\bar{L}^G$ agrees with that in Poincaré gauge theory (PGT) [14], and hence *gravitational field equations in* $\bar{\text{P}}$GT *take the same forms as those in* PGT [15].

In order to get conserved generators, we have employed in Refs. [3, 4, 5] the Lagrangian density $L \stackrel{\text{def}}{=} L^G + L^M$, which leads to same field equations as $\bar{L}$ does. Here L^G is defined by

$$L^G \stackrel{\text{def}}{=} \bar{L}^G + \Delta/\sqrt{-g} \tag{14}$$

with

$$g \stackrel{\text{def}}{=} \det(g_{\mu\nu}) , \quad \Delta \stackrel{\text{def}}{=} (\mathrm{W}_{kl}{}^{\mu\nu} A^{kl}{}_\mu)_{,\nu} , \quad \mathrm{W}_{kl}{}^{\mu\nu} \stackrel{\text{def}}{=} 2d\sqrt{-g} e^\mu{}_{[k} e^\nu{}_{l]} . \tag{15}$$

Let us consider Poincaré gauge transformations (2) with infinitesimal functions t^k and with $a \in SL(2,C)$ such that $\Lambda(a)$ are represented as

$$(\Lambda(a))^k{}_l = \delta^k{}_l + \omega^k{}_l \tag{16}$$

with infinitesimal functions $\omega_{kl} = -\omega_{lk}$. Also, we consider the infinitesimal coordinate transformations

$$x'^\mu = x^\mu + \epsilon^\mu \tag{17}$$

with ϵ^μ being infinitesimal functions. Under the product transformations of Eqs. (16) and (17), the Lagrangian density L transforms according to

$$L' = L + \frac{1}{\sqrt{-g}} \partial_\nu \Lambda^\nu \tag{18}$$

with

$$\Lambda^\nu \stackrel{\text{def}}{=} \omega^{kl}{}_{,\mu} \mathrm{W}_{kl}{}^{\mu\nu} . \tag{19}$$

We see that L is invariant under the product transformations of Eq. (16) with *constant* ω_{kl} and Eq. (17), but it violates *local* $SL(2,C)$ invariance.

In considering energy-momentum and angular momentum, there are two possibilities in choosing the set of independent field variables [3, 4, 5], one is to choose the set $\{\psi^k, A^k{}_\mu, A^k{}_{l\mu}, \phi^A\}$ and the other is

to choose the set $\{\psi^k, e^k{}_\mu, A^k{}_{l\mu}, \phi^A\}$ instead. In the present note, we employ $\{\psi^k, A^k{}_\mu, A^k{}_{l\mu}, \phi^A\}$ as the set of independent field variables, because this choice is preferential to the other, as we have seen in Refs. [3, 4, 5].

From Eq. (18) and the transformation properties of the fundamental fields $\psi^k, A^k{}_\mu, A^k{}_{l\mu}$ and ϕ^A, we can get the following [2]:

(A) The total dynamical energy-momentum density ${}^{\rm tot}\mathbf{T}_k{}^\mu$ and the total "spin" angular momentum density ${}^{\rm tot}\mathbf{S}_{kl}{}^\mu$ are expressed as follows:

$$ {}^{\rm tot}\mathbf{T}_k{}^\mu \stackrel{\rm def}{=} {}^G\mathbf{T}_k{}^\mu + {}^M\mathbf{T}_k{}^\mu = \frac{\partial \mathbf{L}}{\partial A^k{}_\mu} , \tag{20} $$

$$ {}^{\rm tot}\mathbf{S}_{kl}{}^\mu \stackrel{\rm def}{=} {}^G\mathbf{S}_{kl}{}^\mu + {}^M\mathbf{S}_{kl}{}^\mu = -2\frac{\partial \mathbf{L}}{\partial A^{kl}{}_\mu} + 2\mathbf{W}_{kl}{}^{\mu\nu}{}_{,\nu} , \tag{21} $$

with

$$ \mathbf{L} \stackrel{\rm def}{=} \sqrt{-g}L , \quad \mathbf{L}^G \stackrel{\rm def}{=} \sqrt{-g}L^G , \quad \mathbf{L}^M \stackrel{\rm def}{=} \sqrt{-g}L^M , \tag{22} $$

$$ \mathbf{F}_k{}^\mu \stackrel{\rm def}{=} \frac{\partial \mathbf{L}}{\partial \psi^k{}_{,\mu}} , \quad \mathbf{F}_k{}^{\mu\nu} \stackrel{\rm def}{=} \frac{\partial \mathbf{L}}{\partial A^k{}_{\mu,\nu}} = \frac{\partial \mathbf{L}^G}{\partial A^k{}_{\mu,\nu}} = -\mathbf{F}_k{}^{\nu\mu} , \tag{23} $$

$$ \mathbf{F}_{kl}{}^{\mu\nu} \stackrel{\rm def}{=} \frac{\partial \mathbf{L}}{\partial A^{kl}{}_{\mu,\nu}} = \frac{\partial \mathbf{L}^G}{\partial A^{kl}{}_{\mu,\nu}} = -\mathbf{F}_{kl}{}^{\nu\mu} , \tag{24} $$

$$ {}^G\mathbf{T}_k{}^\mu \stackrel{\rm def}{=} \frac{\partial \mathbf{L}^G}{\partial \psi^k{}_{,\mu}} + \mathbf{F}_l{}^{\nu\mu}A_k{}^l{}_\nu = \frac{\partial \mathbf{L}^G}{\partial A^k{}_\mu} , \tag{25} $$

$$ {}^M\mathbf{T}_k{}^\mu \stackrel{\rm def}{=} \frac{\partial \mathbf{L}^M}{\partial \psi^k{}_{,\mu}} + i\frac{\partial \mathbf{L}^M}{\partial \phi^A{}_{,\mu}}(P_k\phi)^A = \frac{\partial \mathbf{L}^M}{\partial A^k{}_\mu} , \tag{26} $$

$$ \begin{aligned} {}^G\mathbf{S}_{kl}{}^\mu &\stackrel{\rm def}{=} -2\frac{\partial \mathbf{L}^G}{\partial \psi^{[k}{}_{,\mu}}\psi_{l]} - 2\mathbf{F}_{[k}{}^{\nu\mu}A_{l]\nu} - 4\mathbf{F}_{[km}{}^{\nu\mu}A_{l]}{}^m{}_\nu \\ &= -2\frac{\partial \mathbf{L}^G}{\partial A^{kl}{}_\mu} + 2\mathbf{W}_{kl}{}^{\mu\nu}{}_{,\nu} , \end{aligned} \tag{27} $$

$$ {}^M\mathbf{S}_{kl}{}^\mu \stackrel{\rm def}{=} -2\frac{\partial \mathbf{L}^M}{\partial \psi^{[k}{}_{,\mu}}\psi_{l]} - i\frac{\partial \mathbf{L}^M}{\partial \phi^A{}_{,\mu}}(M_{kl}\phi)^A = -2\frac{\partial \mathbf{L}^M}{\partial A^{kl}{}_\mu} . \tag{28} $$

Thus, ${}^{\rm tot}\mathbf{T}_k{}^\mu$ has the standard form of gauge current in Yang-Mills theories, while ${}^{\rm tot}\mathbf{S}_{kl}{}^\mu$ has not. There is an additional term $2\mathbf{W}_{kl}{}^{\mu\nu}{}_{,\nu}$ which originates from the term Δ violating the $SL(2,C)$-gauge invariance of the gravitational Lagrangian density.

(B) When the field equations $\delta\mathbf{L}/\delta A^k{}_\mu \stackrel{\rm def}{=} \partial\mathbf{L}/\partial A^k{}_\mu - \partial_\nu(\partial\mathbf{L}/\partial A^k{}_{\mu,\nu}) = 0$ and $\delta\mathbf{L}/\delta\phi^A = 0$ are both satisfied, we have the following:

(i) The differential conservation laws

$$ \partial_\mu {}^{\rm tot}\mathbf{T}_k{}^\mu = 0 , \quad \partial_\mu {}^{\rm tot}\mathbf{S}_{kl}{}^\mu = 0 . \tag{29} $$

(ii) The field equation $\delta\mathbf{L}/\delta\psi^k = 0$ is automatically satisfied, and hence ψ^k is not an independent dynamical variable.

The densities ${}^G\mathbf{T}_k{}^\mu$ and ${}^M\mathbf{T}_k{}^\mu$ are the dynamical energy-momentum densities of gravitational field and of the matter field ϕ^A, respectively, while ${}^G\mathbf{S}_{kl}{}^\mu$ and ${}^M\mathbf{S}_{kl}{}^\mu$ are "spin" angular momentum densities of gravitational field and of ϕ^A, respectively.

In Refs. [3, 4, 5], we have examined the integrations of ${}^{\rm tot}\mathbf{T}_k{}^\mu$ and ${}^{\rm tot}\mathbf{S}_{kl}{}^\mu$ for asymptotically flat space-time by choosing ψ^k as

$$ \psi^k = e^{(0)k}{}_\mu x^\mu + \psi^{(0)k} + O(1/r^\beta) , \quad \psi^k{}_{,\mu} = e^{(0)k}{}_\mu + O(1/r^{\beta+1}) , \quad (\beta > 0) , \tag{30} $$

where ${e^{(0)k}}_\mu$ is a constant satisfying ${e^{(0)k}}_\mu \eta_{kl} {e^{(0)l}}_\nu = \eta_{\mu\nu}$, and $\psi^{(0)k}$ and β are constants. Also, we have defined $r \stackrel{\text{def}}{=} \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$, and $O(1/r^\alpha)$ with positive α denotes a term for which $r^\alpha O(1/r^\alpha)$ remains finite for $r \to \infty$; a term $O(1/r^\alpha)$ may of course be zero. We have shown the following:

$$M_k \stackrel{\text{def}}{=} \int_\sigma {}^{\text{tot}}\mathbf{T}_k{}^\mu d\sigma_\mu = {e^{(0)}}_k{}^\mu M_\mu \,, \tag{31}$$

$$S_{kl} \stackrel{\text{def}}{=} \int_\sigma {}^{\text{tot}}\mathbf{S}_{kl}{}^\mu d\sigma_\mu = {e^{(0)}}_{k\mu} {e^{(0)}}_{l\nu} M^{\mu\nu} + 2\psi^{(0)}{}_{[k} M_{l]} \,, \tag{32}$$

$$\tag{33}$$

where $d\sigma_\mu$ denotes the surface element on a space-like surface σ. Also, we have defined

$$M_\mu \stackrel{\text{def}}{=} \eta_{\mu\lambda} \int_\sigma \theta^{\lambda\nu} d\sigma_\nu \,, \quad M^{\mu\nu} \stackrel{\text{def}}{=} \int_\sigma (x^\mu \theta^{\nu\lambda} - x^\nu \theta^{\mu\lambda}) d\sigma_\lambda \tag{34}$$

with

$$\theta^{\lambda\nu} \stackrel{\text{def}}{=} 2d\partial_\rho \partial_\sigma \{(-g) g^{\lambda[\nu} g^{\rho]\sigma}\} \,, \tag{35}$$

Actually, Eq. (31) has been obtained without using Eq. (30), but it is crucial in obtaining the expression (32). Also, the expression of $\theta^{\lambda\nu}$ agrees with that of the symmetric energy-momentum density proposed by Landau-Lifshitz in general relativity.

The dynamical energy-momentum M_k is the generator of internal translations and the total energy-momentum of the system. The "spin" angular momentum S_{kl} is the generator of internal $SL(2,C)$-transformations and the *total* (=*spin*+*orbital*) angular momentum of the system.

Under the product transformations of Eqs. (16) and (17), the densities ${}^G\mathbf{T}_k{}^\mu$, ${}^M\mathbf{T}_k{}^\mu$, ${}^G\mathbf{S}_{kl}{}^\mu$ and ${}^M\mathbf{S}_{kl}{}^\mu$ transform according to

$$\begin{aligned}
{}^G\mathbf{T}'_k{}^\mu &= {}^G\mathbf{T}_k{}^\mu - {\omega_k}^l {}^G\mathbf{T}_l{}^\mu + {\epsilon^\mu}_{,\nu} {}^G\mathbf{T}_k{}^\nu - {\epsilon^\lambda}_{,\lambda} {}^G\mathbf{T}_k{}^\mu - {\omega_k}^l{}_{,\nu} \mathbf{F}_l{}^{\mu\nu} + \frac{\partial(\partial_\nu \Lambda^\nu)}{\partial {A^k}_\mu} \,, && (36)\\
{}^M\mathbf{T}'_k{}^\mu &= {}^M\mathbf{T}_k{}^\mu - {\omega_k}^l {}^M\mathbf{T}_l{}^\mu + {\epsilon^\mu}_{,\nu} {}^M\mathbf{T}_k{}^\nu - {\epsilon^\lambda}_{,\lambda} {}^M\mathbf{T}_k{}^\mu \,, && (37)\\
{}^G\mathbf{S}'_{kl}{}^\mu &= {}^G\mathbf{S}_{kl}{}^\mu - {\omega_k}^m {}^G\mathbf{S}_{ml}{}^\mu - {\omega_l}^m {}^G\mathbf{S}_{km}{}^\mu - 2t_{[k} {}^G\mathbf{T}_{l]}{}^\mu + {\epsilon^\mu}_{,\nu} {}^G\mathbf{S}_{kl}{}^\nu - {\epsilon^\lambda}_{,\lambda} {}^G\mathbf{S}_{kl}{}^\mu \\
&\quad -2t_{[k,\nu} \mathbf{F}_{l]}{}^{\mu\nu} - 2\omega_{[k}{}^m{}_{,\nu} \Sigma_{l]m}{}^{\mu\nu} - 2\frac{\partial(\partial_\nu \Lambda^\nu)}{\partial A^{kl}{}_\mu} \,, && (38)\\
{}^M\mathbf{S}'_{kl}{}^\mu &= {}^M\mathbf{S}_{kl}{}^\mu - {\omega_k}^m {}^M\mathbf{S}_{ml}{}^\mu - {\omega_l}^m {}^M\mathbf{S}_{km}{}^\mu - 2t_{[k} {}^M\mathbf{T}_{l]}{}^\mu + {\epsilon^\mu}_{,\nu} {}^M\mathbf{S}_{kl}{}^\nu - {\epsilon^\lambda}_{,\lambda} {}^M\mathbf{S}_{kl}{}^\mu \,. && (39)
\end{aligned}$$

3 Extended New General Relativity

In $\bar{\text{P}}$GT, we consider the case in which the field strength ${R^k}_{l\mu\nu}$ vanishes identically, ${R^k}_{l\mu\nu} \equiv 0$, then, the curvature vanishes and we have a teleparallel theory.

We choose the $SL(2,C)$-gauge such that ${A^k}_{l\mu} \equiv 0$, which reduces the expressions of vierbeins ${e^k}_\mu$, affine connection coefficients ${\Gamma^\lambda}_{\mu\nu}$ and the covariant derivative $D_k\phi$ to

$$\begin{aligned}
{e^k}_\mu &= {\psi^k}_{,\mu} + {A^k}_\mu \,, && (40)\\
{\Gamma^\lambda}_{\mu\nu} &= {e^\lambda}_l {e^l}_{\mu,\nu} \,, && (41)\\
D_k\phi^A &= {e^\mu}_k (\partial_\mu \phi^A + i {A^l}_\mu (P_l \phi)^A) \,, && (42)
\end{aligned}$$

respectively.

Since $L^R = 0 = dR$, the gravitational Lagrangian density L^G is reduced to $L^G = L^T$, which agrees with the gravitational Lagrangian density in new general relativity (NGR) [16, 17]. Thus, *the gravitational field equations in* ENGR *take the same forms as those in* NGR.

As is the case of the Lagrangian density $L^G + L^M$ in $\bar{\text{P}}$GT, $L = L^T + L^M$ is invariant under the product transformations of Eq. (16) with *constant* ω_{kl} and Eq. (17), but it violates *local* $SL(2,C)$ invariance.

The expression of the dynamical energy-momentum densities and the "spin" angular momentum densities are reduced to

$$ {}^{tot}\mathbf{T}_k{}^{\mu} \stackrel{def}{=} {}^{G}\mathbf{T}_k{}^{\mu} + {}^{M}\mathbf{T}_k{}^{\mu} = \frac{\partial \mathbf{L}}{\partial A^k{}_{\mu}} , \tag{43} $$

$$ {}^{tot}\mathbf{S}_{kl}{}^{\mu} \stackrel{def}{=} {}^{G}\mathbf{S}_{kl}{}^{\mu} + {}^{M}\mathbf{S}_{kl}{}^{\mu} = -2\frac{\partial \mathbf{L}}{\partial A^{kl}{}_{\mu}} + 2\mathbf{W}_{kl}{}^{\mu\nu}{}_{,\nu} , \tag{44} $$

with

$$ {}^{G}\mathbf{T}_k{}^{\mu} = \frac{\partial \mathbf{L}^G}{\partial \psi^k{}_{,\mu}} = \frac{\partial \mathbf{L}^G}{\partial A^k{}_{\mu}} , \tag{45} $$

$$ {}^{G}\mathbf{S}_{kl}{}^{\mu} = -2\frac{\partial \mathbf{L}^G}{\partial \psi^{[k}{}_{,\mu}}\psi_{l]} - 2\mathbf{F}_{[k}{}^{\nu\mu}A_{l]\nu} , \quad {}^{M}\mathbf{S}_{kl}{}^{\mu} = -2\frac{\partial \mathbf{L}^M}{\partial \psi^{[k}{}_{,\mu}}\psi_{l]} - i\frac{\partial \mathbf{L}^M}{\partial \phi^A{}_{,\mu}}(M_{kl}\phi)^A , \tag{46} $$

and with ${}^{M}\mathbf{T}_k{}^{\mu}$ as given by (26). The statements (A) and (B) hold true also for ENGR.

In Ref. [6], we have examined, for the case with $c_1 = -1/(3\kappa) = -c_2$, M_k and S_{kl} defined in the same ways as in $\bar{\text{P}}$GT for asymptotically flat space-time by choosing ψ^k as given by Eq. (30) and assuming some additional conditions on asymptotic behaviors of field variables. The same expressions as Eqs. (31) and (32) hold also for this case.

Under the product transformations of Eqs. (16) and (17), these densities transform according to

$$ {}^{G}\mathbf{T'}_k{}^{\mu} = {}^{G}\mathbf{T}_k{}^{\mu} - \omega_k{}^{l}{}^{G}\mathbf{T}_l{}^{\mu} + \epsilon^{\mu}{}_{,\nu}{}^{G}\mathbf{T}_k{}^{\nu} - \epsilon^{\lambda}{}_{,\lambda}{}^{G}\mathbf{T}_k{}^{\mu} - \omega_k{}^{l}{}_{,\nu}\mathbf{F}_l{}^{\mu\nu} , \tag{47} $$

$$ {}^{M}\mathbf{T'}_k{}^{\mu} = {}^{M}\mathbf{T}_k{}^{\mu} - \omega_k{}^{l}{}^{M}\mathbf{T}_l{}^{\mu} + \epsilon^{\mu}{}_{,\nu}{}^{M}\mathbf{T}_k{}^{\nu} - \epsilon^{\lambda}{}_{,\lambda}{}^{M}\mathbf{T}_k{}^{\mu} , \tag{48} $$

$$ {}^{G}\mathbf{S'}_{kl}{}^{\mu} = {}^{G}\mathbf{S}_{kl}{}^{\mu} - \omega_k{}^{m}{}^{G}\mathbf{S}_{ml}{}^{\mu} - \omega_l{}^{m}{}^{G}\mathbf{S}_{km}{}^{\mu} - 2t_{[k}{}^{G}\mathbf{T}_{l]}{}^{\mu} + \epsilon^{\mu}{}_{,\nu}{}^{G}\mathbf{S}_{kl}{}^{\nu} - \epsilon^{\lambda}{}_{,\lambda}{}^{G}\mathbf{S}_{kl}{}^{\mu} - 2t_{[k,\nu}\mathbf{F}_{l]}{}^{\mu\nu} , \tag{49} $$

$$ {}^{M}\mathbf{S'}_{kl}{}^{\mu} = {}^{M}\mathbf{S}_{kl}{}^{\mu} - \omega_k{}^{m}{}^{M}\mathbf{S}_{ml}{}^{\mu} - \omega_l{}^{m}{}^{M}\mathbf{S}_{km}{}^{\mu} - 2t_{[k}{}^{M}\mathbf{T}_{l]}{}^{\mu} + \epsilon^{\mu}{}_{,\nu}{}^{M}\mathbf{S}_{kl}{}^{\nu} - \epsilon^{\lambda}{}_{,\lambda}{}^{M}\mathbf{S}_{kl}{}^{\mu} . \tag{50} $$

4 Summary and Discussions

We have examined the transformation properties of energy-momentum densities and of angular momentum densities both in $\bar{\text{P}}$GT and in ENGR

Results can be summarized as follows:

[1] Results in Poincaré gauge theory ($\bar{\text{P}}$GT):

From Eqs. (36)$\sim$ (39), we see that the densities ${}^{G}\mathbf{T}_k{}^{\mu}$, ${}^{M}\mathbf{T}_k{}^{\mu}$, ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ and ${}^{M}\mathbf{S}_{kl}{}^{\mu}$ are all space-time vector densities, i.e., they transform as vector densities under general coordinate transformations. Their transformation properties under internal Poincaré transformations are summarized as follows:

(a) The dynamical energy-momentum density ${}^{G}\mathbf{T}_k{}^{\mu}$ of the gravitational field is invariant under *local* translations. It transforms as a vector under *global* $SL(2,C)$-transformations. But, it is not vectorial under *local* $SL(2,C)$-transformations.

(b) The dynamical energy-momentum density ${}^{M}\mathbf{T}_k{}^{\mu}$ of the matter field ϕ^A is invariant under *local* translations. It transforms as a vector under *local* $SL(2,C)$-transformations.

(c) Under *global* translations, the "spin" angular momentum density ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ of the gravitational field receives transformations which correspond to translations in internal space-time, and it transforms as a tensor under *global* $SL(2,C)$-transformations. But, it is not tensorial under *local* Poincaré transformations.

(d) The "spin" angular momentum density ${}^{M}\mathbf{S}_{kl}{}^{\mu}$ of the matter field ϕ^A is tensorial under *local* Poincaré transformations.

[2] Results in extended new general relativity (ENGR):

All the densities ${}^{G}\mathbf{T}_k{}^{\mu}$, ${}^{M}\mathbf{T}_k{}^{\mu}$, ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ and ${}^{M}\mathbf{S}_{kl}{}^{\mu}$ are space-time vector densities. Also for the case of ENGR, the same statements as (a), (b), and (d) in (1A) hold true for ${}^{G}\mathbf{T}_k{}^{\mu}$, ${}^{M}\mathbf{T}_k{}^{\mu}$ and ${}^{M}\mathbf{S}_{kl}{}^{\mu}$. As for ${}^{G}\mathbf{S}_{kl}{}^{\mu}$, we have the following [19]:

(c′) The density ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ receives transformations which correspond to translations of the origin of internal space-time under *global* translations, and it transforms as a tensor under *local* $SL(2,C)$-transformations. But, it is not tensorial under *local* internal translations.

Since ${}^{G}\mathbf{T}_{k}{}^{\mu}$, ${}^{M}\mathbf{T}_{k}{}^{\mu}$, ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ *and* ${}^{M}\mathbf{S}_{kl}{}^{\mu}$ *are all space-time vector densities in both theories, the energy-momenta and angular momenta of the gravitational field and of the matter field are defined well, and independent of the coordinate system employed.* For example, for the energy-momentum ${}^{G}M_{k}$ of the gravitational field, we have

$$ {}^{G}M_{k} \stackrel{\text{def}}{=} \int_{\sigma} {}^{G}\mathbf{T}_{k}{}^{\mu} d\sigma_{\mu} = \int_{\sigma} {}^{G}\mathbf{T}'_{k}{}^{\mu} d\sigma'_{\mu} \,. \tag{51} $$

In both in $\bar{\text{P}}$GT and in ENGR, the densities ${}^{M}\mathbf{T}_{k}{}^{\mu}$ and ${}^{M}\mathbf{S}_{kl}{}^{\mu}$ are well behaved under *local* internal Poincaré transformations, while the energy-momentum density ${}^{G}\mathbf{T}_{k}{}^{\mu}$ of the gravitational field is well behaved under *local* internal translations. In ENGR, the "spin" angular momentum density ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ of the gravitational field is well behaved under *local* internal $SL(2,C)$-transformations.

It is worth mentioning here that the Lagrangian densities $L^{G} + L^{M}$ in $\bar{\text{P}}$GT and $L^{T} + L^{M}$ in ENGR are both invariant under *local* internal translations and under *global* internal $SL(2,C)$-transformations, but they violate the invariance under *local* internal $SL(2,C)$-transformations. Thus, one may claim that we need not bother about the fact that ${}^{G}\mathbf{T}_{k}{}^{\mu}$ in $\bar{\text{P}}$GT and in ENGR and ${}^{G}\mathbf{S}_{kl}{}^{\mu}$ in $\bar{\text{P}}$GT are not tensorial under *local* internal $SL(2,C)$-transformations. In ENGR, in particular, this can be strongly asserted, because *local* $SL(2,C)$-gauge invariance is rather accidental [17] in this theory due to the lack of $SL(2,C)$-gauge potential $A^{k}{}_{l\mu}$.

In general relativity, all the known energy-momentum and angular momentum densities of the gravitational field are not space-time tensor densities. Poincaré gauge theory and extended new general relativity are preferential to general relativity in the point that the former two theories have energy-momentum and angular momentum densities of the gravitational field which are true space-time vector densities.

A

The irreducible components t_{klm}, v_{k} and a_{k} of T_{klm} are defined by the following:

$$ \begin{aligned} t_{klm} &\stackrel{\text{def}}{=} \frac{1}{2}(T_{klm} + T_{lkm}) + \frac{1}{6}(\eta_{mk}v_{l} + \eta_{ml}v_{k}) - \frac{1}{3}\eta_{kl}v_{m} \,, \\ v_{k} &\stackrel{\text{def}}{=} T^{l}{}_{lk} \,, \quad a_{k} \stackrel{\text{def}}{=} \frac{1}{6}\varepsilon_{klmn}T^{lmn} \,, \end{aligned} \tag{52} $$

where the symbol ε_{klmn} stands for completely anti-symmetric Lorentz tensor with $\varepsilon_{(0)(1)(2)(3)} = -1$ [20].

The irreducible components $A_{klmn}, B_{klmn}, C_{klmn}, E_{kl}, I_{kl}, R$ of R_{klmn} are defined by the following:

$$ \begin{aligned} A_{klmn} &\stackrel{\text{def}}{=} \frac{1}{6}(R_{klmn} + R_{kmnl} + R_{knlm} + R_{lmkn} + R_{lnmk} + R_{mnkl}) \,, \\ B_{klmn} &\stackrel{\text{def}}{=} \frac{1}{4}(W_{klmn} + W_{mnkl} - W_{knlm} - W_{lmkn}) \,, \\ C_{klmn} &\stackrel{\text{def}}{=} \frac{1}{2}(W_{klmn} - W_{mnkl}) \,, \quad E_{kl} \stackrel{\text{def}}{=} \frac{1}{2}(R_{kl} - R_{lk}) \,, \\ I_{kl} &\stackrel{\text{def}}{=} \frac{1}{2}(R_{kl} + R_{lk}) - \frac{1}{4}\eta_{kl}R \,, \quad R \stackrel{\text{def}}{=} \eta^{kl}R_{kl} \end{aligned} \tag{53} $$

with

$$ W_{klmn} \stackrel{\text{def}}{=} R_{klmn} - \frac{1}{2}(\eta_{km}R_{ln} + \eta_{ln}R_{km} - \eta_{kn}R_{lm} - \eta_{lm}R_{kn}) + \frac{1}{6}(\eta_{km}\eta_{ln} - \eta_{lm}\eta_{kn})R \,, \tag{54} $$

$$ R_{kl} \stackrel{\text{def}}{=} \eta^{mn}R_{kmln} \,. \tag{55} $$

References

[1] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), p. 466.

[2] T. Kawai, Gen. Rel. Gravit. **18**, 995 (1986); **19**, 1285(E) (1987).

[3] T. Kawai, Prog. Theor. Phys. **79**, 920 (1988).

[4] T. Kawai and H. Saitoh, Prog. Theor. Phys. **81**, 280 (1989).

[5] T. Kawai and H. Saitoh, Prog. Theor. Phys. **81**, 1119 (1989).

[6] T. Kawai and N. Toma, Prog. Theor. Phys. **85**, 901 (1991).

[7] The covering group $\bar{P}_0$ of the proper orthochronous Poincaré group will be called the $\overline{\text{Poincaré}}$ group, and its elements will be called $\overline{\text{Poincaré}}$ transformations.

[8] For details, please see T. Kawai, Phys. Rev. D **62**, 10401 (2000).

[9] Unless otherwise stated, we use the following conventions for indices: The middle part of the Greek alphabet, $\mu, \nu, \lambda, \ldots$, refers to 0, 1, 2, and 3. In a similar way, the middle part of the Latin alphabet, $i, j, k, \ldots$, means 0, 1, 2, and 3, unless otherwise stated. The capital letters A and B are used for indices for components of the field ϕ, and N denotes the dimension of the representation ρ.

[10] For function f on M, we define $f_{,\mu} \stackrel{\text{def}}{=} \partial f/\partial x^{\mu}$.

[11] We define $A_{\ldots[\mu\cdots\nu]\ldots} \stackrel{\text{def}}{=} (A_{\ldots\mu\cdots\nu\ldots} - A_{\ldots\nu\cdots\mu\ldots})/2$.

[12] The field components ${e^k}_{\mu}$ and ${e^{\mu}}_k$ are used to convert Latin and Greek indices. Also, raising and lowering the indices $k, l, m, \ldots$ are accomplished with the aid of $(\eta^{kl}) \stackrel{\text{def}}{=} (\eta_{kl})^{-1}$ and (η_{kl}).

[13] T. Kawai, Prog. Theor. Phys. **82**, 850 (1989).

[14] K. Hayashi and T. Shirafuji, Prog. Theor. Phys. **64**, 866 (1980); **64**, 883 (1980); **64**, 1435 (1980).

[15] The requirement that the theory has the Newtonian limit restricts the parameter d as $d = 1/(2\kappa)$ [14]. Here, κ denotes the Einstein constant $\kappa = 8\pi G/c^4$ with G and c being the Newton gravitational constant and the light velocity in the vacuum, respectively. If $3d_2 + 2d_3 = 0 = d_5 + 12d_6$ and if the intrinsic spins of the source fields are negligible in addition, the Einstein equation is obtainable [14]. In the discussions below, however, we do not impose any restriction on the parameters c_i, d_j, d $(i = 1, 2, 3, j = 1, 2, 3, \ldots, 6)$.

[16] K. Hayashi and T. Shirafuji, Phys. Rev. D **19**, 3524 (1979).

[17] The case with $c_1 = -1/(3\kappa) = -c_2$ is quite favorable experimentally. Further, when $c_3 = -3/(4\kappa)$ in addition, the Lagrangian density L^T agrees with the gravitational Lagrangian density $R(\{\})/(2\kappa)$ in general relativity up to a total divergence, and the gravitational field equation agrees with the Einstein equation [16]. Here, $R(\{\})$ denotes the Riemann scalar curvature. Even for this case, L^T is not invariant under *local* $SL(2, C)$-gauge transformations, while $R(\{\})/(2\kappa)$ is. The discussions below will be developed for arbitrary values of the parameters c_1, c_2, and c_3.

[18] The set $\{\psi^k, {A^k}_{\mu}, \phi^A\}$ will be employed as the set of independent field variables.

[19] Note that ${}^G\mathrm{S}_{kl}{}^{\mu}$ in $\bar{\text{P}}$GT is *not* tensorial under *local* $SL(2, C)$-transformations, as has been stated in (c).

[20] Latin indices are put in parentheses to distinguish them from Greek indices.

Symmetries of the Gowdy Equations and Spatial Topologies

Masayuki TANIMOTO [1]

Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan
&
Department of Earth and Space Science, Graduate School of Science,
Osaka University, Toyonaka 560-0043, Japan.

Abstract

We examine some kinds of discrete symmetries which are dynamically preserved, using the (generalized) Gowdy models of the first kind.

1 Introduction

It seems like there are some connections between dynamical properties of a spacetime and its spatial topology. For example, we know the recollapsing conjecture [1, 2, 3]; As is well known, in a positive curvature (topologically S^3) homogeneous and isotropic cosmological model, the universe contracts and recollapses after an expanding era. The conjecture claims that this recollapsing property does not depend on the symmetry imposed on the model, that is, any spacetime recollapses if the space is *topologically* S^3 (and if only appropriate energy and pressure conditions are fulfilled). Thus we may think of the recollapsing property as a result of the spatial topology. (For another interesting examples related to asymptotic dynamics, see [4, 5].)

In this article we compare some dynamical properties of the (generalized) Gowdy models of the first kind. The spatial manifold of this kind is a T^2-bundle over S^1, and each T^2-fiber is generated by two commuting local Killing vectors. (Such a model was first considered by Rendall [6], and discussed extensively by the author [7, 8]. These models are generalizations of Gowdy's T^3 model [9].) There are infinite number of such bundles which are topologically distinct from each other. However, since each fiber is generated by (local) Killing vectors in our models the natural reduced (spatial) manifold is S^1 in all cases. Because of this reason we can compare the dynamical properties of each model in rather detail.

What we investigate concerns symmetries or sets of symmetric data which are preserved temporarily. More specifically, we consider reflection symmetries (in a generalized sense). For example, for the usual T^3 Gowdy model any set of initial data which are described by even (spatial) functions evolves in such a way that the data remain even functions at any time. Also, the same statement is true even if we replace "even" by "odd", and in general we will see there is much larger set of such "reflection" symmetries which are preserved. We can consider corresponding symmetric data for any other models, but the symmetries are not necessarily preserved. We may interpret this is a manifestation of the influence of spatial topology.

This article basically treats the same topic as Ref.[8], but the presentation is made in a somewhat different way.

2 The Gowdy Models of the 1st Kind

The metric of the Gowdy models can be written in the form

$$ds^2 = -e^{A}(d\tau^2 - dx^2) + R[e^{P}(dy + Qdz)^2 + e^{-P}dz^2], \tag{1}$$

where A, R, P, and Q are functions of τ and x. There are two commuting Killing vectors for this metric, $\partial/\partial y$ and $\partial/\partial z$. We may think of this metric as that on $\mathbf{R}^4$, but we can also compactify the spatial part $\mathbf{R}^3$ to a T^2-bundle over S^1 if imposing appropriate boundary conditions on the metric functions. In such a case, the model is called a (generalized) Gowdy model of the 1st kind, which we will consider. After the compactification the two Killing vectors, in general, descend to local Killing vectors, generating the T^2-fibers.

Note that each T^2-fiber is characterized by its volume *and* two moduli parameters $(X, Y) \equiv (Q, e^{-P})$.

[1] email: tanimoto@yukawa.kyoto-u.ac.jp

3 Symmetries

As well known [9, 8], the vacuum Einstein equation for metric function R is given by the simple wave equation $(\partial_{\tau\tau} - \partial_{xx})R = 0$, and we can set without loss of generality, $R = \tau$. With this choice, the equations for functions P and Q are found to form a closed set of equations, and the remaining function A is integrable once P and Q are solved. So, we are led to concentrate on the moduli parameters X and Y, and the equations for them are derivable from the following (reduced) Hamiltonian H,

$$H = \int \mathcal{H} dx, \tag{2}$$

$$\mathcal{H} = \frac{1}{2}(g^{AB}\Pi_A\Pi_B + e^{-2\tau}g_{AB}X^{A\prime}X^{B\prime}), \tag{3}$$

where scripts $A, B, \cdots$ run 1 to 2. We have set $(X^1, X^2) \equiv (X, Y)$, and (Π_1, Π_2) is the set of the conjugate momenta. The "metric" g_{AB} on the moduli space $\mathcal{M}$ (spanned by X and Y) is given by the following *hyperbolic metric*

$$dS^2 = \frac{dX^2 + dY^2}{Y^2}. \tag{4}$$

So, the moduli space is also a hyperbolic plane H^2 with three dimensional isometry group IsomH^2, of which connected component is $SL_2\mathbf{R}$. (Exactly speaking, $\mathcal{M}$ is not H^2 itself, but a quotient space of it. This point is explained in §.5.)

Note that the data for a moment (τ = constant), i.e., the configuration of the spatial manifold, is given by a loop $l(x)$ in the moduli space:

$$l : [0, 1] \longrightarrow \mathcal{M}, \quad l(0) = l(1), \tag{5}$$

and the whole spacetime is described by a one-parameter (τ) family of such loops: $l(x, \tau)$.

The isometry group IsomH^2 provides symmetries of motion, i.e., if the "spacetime" $l(x, \tau)$ is a solution for the vacuum Einstein equation, then the action $g \cdot l(x, \tau)$ is also another solution for $g \in$ IsomH^2. This is one of well-known properties of the Gowdy equations.

We also should point out a rather trivial but important symmetry, which is given by the following reparametrization:

$$R \cdot l(x, \tau) = l(-x, \tau). \tag{6}$$

As is easily seen, R inverts (or reflects) each spatial configuration about the origin $x = 0$. This also define a symmetry of motion.

4 The Reflection Operators

While we may think of the operator R itself as a reflection operator, it is also possible to generalize reflections using the isometries IsomH^2, i.e., we consider the composition

$$\mathcal{R} = R \cdot g, \quad g \in \text{Isom}H^2. \tag{7}$$

Any such an operator defines a symmetry of motion. We should furthermore impose the condition for $\mathcal{R}$ to form a $\mathbf{Z}_2$ group (together with the identity). For convenience, defining $\mathcal{R}$ as a complex function

$$\mathcal{R} : \ z = X + iY \to \mathcal{R}(z), \tag{8}$$

this condition is expressed by

$$\mathcal{R}^2(z) = z, \quad \text{for all } z \in \mathbf{C}. \tag{9}$$

From this we find two classes of operators; one is given by R itself

$$\mathcal{R}_I \equiv R, \tag{10}$$

and the other is given by a one-parameter family

$$\mathcal{R}_\theta \equiv R \cdot \gamma \cdot f_\theta, \quad \theta \in [0, \pi), \tag{11}$$

where

$$\gamma(z) = -\bar{z}, \tag{12}$$

$$f_\theta(z) = \frac{z\cos 2\theta - \sin 2\theta}{z\sin 2\theta + \cos 2\theta}. \tag{13}$$

However, these operators do not necessarily preserve the boundary condition imposed on the solution $X(x,\tau)$ and $Y(x,\tau)$, so not all these operators are appropriate for a given spatial topology. To account for this point for more detail, we should explain the four representative models, called the T^3, E^3, Nil, and Sol models.

5 The Four Representative Models

We have noticed any spatial configuration is represented by a loop l on the moduli space $\mathcal{M}$, but this does not mean the metric functions $X(x)$ and $Y(x)$ are periodic with respect to the spatial coordinate x. This is because $\mathcal{M}$ is a quotient space. It is a classical fact that $\mathcal{M}$ is represented by H^2 *modulo* $SL_2\mathbf{Z}$, $\mathcal{M} \simeq H^2/SL_2\mathbf{Z}$, where $SL_2\mathbf{Z}$ is the mapping class group for T^2. (This discrete group, $SL_2\mathbf{Z}$, is generated by the shift $T : z \to z+1$ and the inversion $S : z \to -1/z$.) H^2 is naturally represented by the upper half plane $\mathbf{C}^+ \equiv \{z \in \mathbf{C} | \mathrm{Im}(z) > 0\}$, and a fundamental domain $\mathcal{F}$ for $\mathcal{M}$ is given by $\mathcal{F} = \{z \in \mathbf{C}^+ | -\frac{1}{2} \le \mathrm{Re}(z) \le \frac{1}{2}, |z| \ge 1\}$. See Fig.1. For the boundary, the two vertical lines (at $\mathrm{Re}(z) = \pm\frac{1}{2}$) should be identified, and the arc (at $|z| \le 1$) is folded about $z = i$ and the opposite sides are identified, as well. So, the point $z = i$ becomes singular. (These are a standard fact about the geometric structure on T^2. See a standard text for a detail.) Let us denote a curve on $\mathbf{C}^+$ which descends to a loop l on $\mathcal{M}$, as $\tilde{l}$. It is $\tilde{l}$ that has a direct correspondence to the boundary condition imposed on $X(x)$ and $Y(x)$. Apparently, $\tilde{l}$ is not necessarily a loop, so what is said above comes. Note that since two (topologically) distinct T^2-bundle over S^1 can have the same configuration of moduli parameters, correspondences of a loop l (or curve $\tilde{l}$) to a topology are not unique, but there are not ambiguities between the four models.

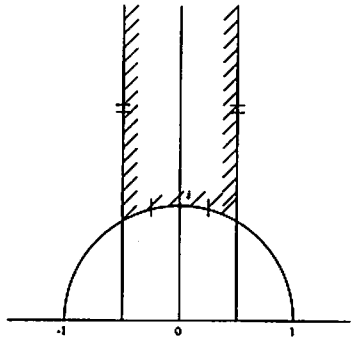


Figure 1: The fundamental domain $\mathcal{F}$ for the moduli space $\mathcal{M}$ is shown (the shaded region).

Now, we are in a position to describe the four models. (However, the following descriptions are not complete enough. See Ref.[8] for more complete ones, though they are less pictorial.)

(i) T^3 model: This model is the conventional one, for which the spatial manifold is topologically T^3. The boundary condition for the metric functions (moduli parameters) is simply given by the periodic boundary condition. A loop l for this model is also represented by a loop $\tilde{l}$ on $\mathbf{C}^+$. See Fig.2 (i). $\tilde{l}$ does not encircle the singular point $z = i$, so can smoothly contract to a point, which corresponds to a flat homogeneous space (Bianchi I space).

(ii) E^3 model: This model is the representative model for the models whose spatial part is T^3, $T^3/\mathbf{Z}_2$, $T^3/\mathbf{Z}_3$, $T^3/\mathbf{Z}_4$, or $T^3/\mathbf{Z}_6$. The boundary condition is periodic, but the loop $\tilde{l}$ on $\mathbf{C}^+$ must encircle the

singular point $z = i$, so *cannot* smoothly contract to a point. See Fig.2 (ii). The locally homogeneous limit corresponds to a Bianchi VII_0 space.

(iii) Nil model: $\tilde{l}$ is an open curve which is invariant under the shifting T. One end approaches to $\mathrm{Re}(z) = +\infty$, and the other to $\mathrm{Re}(z) = -\infty$. See Fig.2 (iii). $\tilde{l}$ descends to a loop on $\mathcal{M}$ which winds around the singular point $z = i$ only once. At a locally homogeneous limit, $\tilde{l}$ becomes a horizontal straight line, which corresponds to a Bianchi II space.

(iv) Sol model: $\tilde{l}$ is an open (arc-like) curve for which both ends approach to $\mathrm{Im}(z) = 0$; one end to $|z| < 1$ and the other end to $|z| > 1$. See Fig.2 (iv). $\tilde{l}$ descends to a loop on $\mathcal{M}$ which winds around the singular point $z = i$ twice or more. The locally homogeneous limit corresponds to a Bianchi VI_0 space.

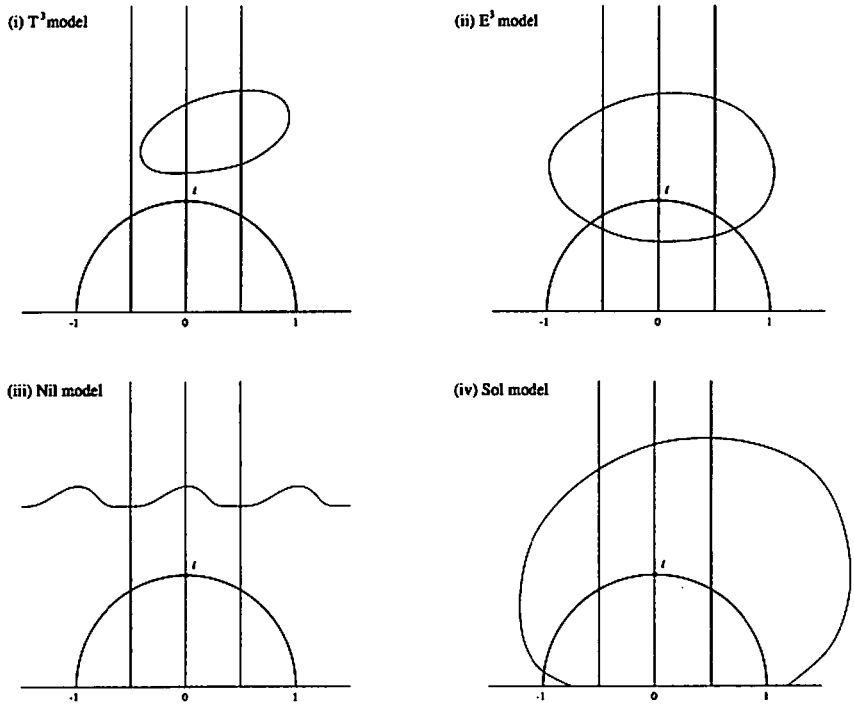


Figure 2: A morphological classification of the four models is shown. A possible curve $\tilde{l}$ is depicted for each model in the upper half plane $\mathbf{C}^+$. Every curve descends to a loop (in $\mathcal{M}$) after identifications generated by T and S.

6 The Invariant Sets

Table 1 shows the appropriate reflection operators for each model [8]. These operators are extracted by requiring they preserve the appropriate boundary condition for each model, i.e., $\tilde{l}$ belonging to a model be mapped by one of these operators to another $\tilde{l}'$ belonging to the same model. (More precisely, to get these reflection operators, we also require a compatibility with a translation operator [8].)

Note that once a reflection operator $\mathcal{R}$ is given, we can accordingly define a set of reflection symmetric solutions $\mathcal{S}(\mathcal{R})$:

$$\mathcal{S}(\mathcal{R}) \equiv \{l(x,\tau) |\ \mathcal{R}(l(x,\tau)) = l(x,\tau)\}. \tag{14}$$

These solutions are invariant under the action of the reflection operator. We call such a set of solutions an *invariant (sub)set.*

Model	Reflections
T^3	$\mathcal{R}_I$, $\mathcal{R}_\theta$
E^3	$\mathcal{R}_\theta$
Nil	$\mathcal{R}_0$
Sol	$\mathcal{R}_{\pi/4}$

Table 1: This shows which reflection operators are eligible for each model. The parameter θ in T^3 or E^3 model can take any value in the range $[0, \pi)$, whereas Nil and Sol models can take $\theta = 0$ and $\pi/4$, respectively. $\mathcal{R}_I$ is possible only for T^3 model.

Apparently, we have the largest union of such sets for the T^3 model, while that of the Nil or Sol model is the smallest. *This is a character coming from topological distinctions.*

7 Dynamical Interpretation

We interpret the above fact as follows.

Note first that the phase space $\mathcal{P}$ for the T^3 model is given by the space of the sets of four (smooth) periodic functions (X, Y, Π_X, Π_Y). The phase space for Nil or Sol model would of course be given by another space, since the metric functions for them are not periodic functions. However, we know it must still be possible to identify it with $\mathcal{P}$, since the base spaces of the bundles are all S^1. All we have to do to do this is to re-represent the metric functions in such a way that they represent deviations from the configuration of a locally homogeneous limit. Practically, this can be easily done by replacing the coordinate basis of the metric (1) by an invariant basis for the appropriate Bianchi type, which was given in §.5.

Next consider the reflection operators $\mathcal{R}_I$ and $\mathcal{R}_\theta$. Although they have been defined so they act on a configuration space, we can naturally extend them to act on the phase space using "Hamiltonian equations" $\Pi_X = Y^{-2}\dot{X}$, $\Pi_Y = Y^{-2}\dot{Y}$. Then those operators define a set of reflection symmetric data in $\mathcal{P}$ by requiring these data be invariant under the action of one of the operators. These sets are common for all the models, since the above two (Hamiltonian) equations are common even if the variables are re-represented in the way described above.

So, we have the same (underlying) phase space and the same reflection symmetric data sets for all the four models. The Hamiltonians are, however, different in this setting, and we can compare dynamical properties of the four models from this point of view. In fact we can interpret [8] the result of the previous section as telling that *only the reflection symmetries shown in Table.1 are preserved by the Hamiltonian flow.* This is a manifestation of the influence of topology to dynamics.

Remark: One might think that there were much more possibilities for another set of reflection symmetric data. However, since we are interested only in those preserved dynamically, it is sufficient to consider the operators $\mathcal{R}_I$ and $\mathcal{R}_\theta$.

References

[1] J. D. Barrow, G. J. Galloway, and F. J. Tipler, Mon. Not. R. Astron. Soc. **223**, 835 (1985); J. D. Barrow and F. J. Tipler, The Anthropic Cosmological Principle, (Oxford University Press, Oxford, 1988).

[2] X.-F. Lin and R. M. Wald, Phys. Rev. **D40**, 3280 (1989); **D41**, 2444 (1990).

[3] G. A. Burnett, Phys. Rev. **D51**, 1621 (1995).

[4] A. E. Fischer and V. Moncrief, Class. Quant. Gravit. **16**, L79, (1999).

[5] M. T. Anderson, preprint, (gr-qc/0006042).

[6] A. D. Rendall, Commun. Math. Phys. **189**, 145 (1997).

[7] M. Tanimoto, J. Math. Phys. 39, 4891 (1998).

[8] M. Tanimoto, to appear in Class. Quant. Gravit., (gr-qc/0003033).

[9] R. H. Gowdy, Ann. Phys. 83, 203 (1974).

Convex Functions and Spacetime Geometry

Gary W. Gibbons,[1]

DAMTP, Center for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, the United Kingdom

Akihiro Ishibashi[2]

Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

Abstract

Convexity and convex functions play an important role in theoretical physics. To initiate a study of the possible uses of convex functions in General Relativity, we discuss the consequences of a spacetime $(M, g_{\mu\nu})$ or an initial data set (Σ, h_{ij}, K_{ij}) admitting a suitably defined convex function. In particular, we focus on the relations between submanifolds such as hypersurfaces in a spacetime and the existence of convex functions in the ambient manifold.

1 Introduction

Convexity and convex functions play an important role in theoretical physics. For example, Gibbs's approach to thermodynamics [G] is based on the idea that the free energy should be a convex function. A closely related concept is that of a convex cone which also has numerous applications to physics. Perhaps the most familiar example is the lightcone of Minkowski spacetime. Equally important is the convex cone of mixed states of density matrices in quantum mechanics. The set of convex functions gives a simple example of a convex cone. Convexity and convex functions also have important applications to geometry including Riemannian geometry [U]. It is surprising therefore that, to our knowledge, that techniques making use of convexity and convex functions have played no great role in General Relativity. The purpose of this paper is to initiate a study of the possible uses of such techniques. As we shall see, the existence of a convex function, suitably defined, on a spacetime places important restrictions on the properties of the spacetime. For example it can contain no closed spacelike geodesics, and asymptotically flat initial data admitting a convex function looking like $\frac{1}{2}x^2$ near infinity must have negative ADM masses. We shall treat Riemannian as well as Lorentzian manifolds with an eye to applications in black hole physics, cosmology and quantum gravity. The detailed results and examples will be four-dimensional but extensions to other dimensions are immediate.

One way of viewing the ideas of this paper is in terms of a sort of duality between paths and particles on the one hand and functions waves on the other. Mathematically the duality corresponds to interchanging range and domain. A curve $x(\lambda)$ is a map $x : \mathbb{R} \to M$ while a function $f(x)$ is a map $f : M \to \mathbb{R}$. A path arises by considering invariance under diffeomorphisms of the domain (i.e. of the world volume) and special paths, for example geodesics have action functionals which are reparametrization invariant. Much effort, physical and mathematical has been expended on exploring the global properties of spacetimes using geodesics. Indeed there is a natural notion of convexity based on geodesics. Often a congruence of geodesics is used.

On the dual side, one might consider properties which are invariant under diffeomorphisms $f(x) \to g(f(x))$ of the range or target space. That is one may explore the global properties of spacetime using the foliations provided by the level sets of a suitable function. The analogue of the action functional for geodesics is one like

$$\int_M \sqrt{|\nabla_\mu f \nabla^\mu f|}, \tag{1}$$

[1] E-mail:gwg1@amtp.cam.ac.uk
[2] E-mail:akihiro@yukawa.kyoto-u.ac.jp

which is invariant under reparametrizations of the range. The level sets of the solutions of the Euler-Lagrange equations are then minimal surfaces. One may also consider foliations by totally umbilic surfaces or by "trace K equal constant" foliations, and this is often done in numerical relativity. The case of a convex function then corresponds to a foliation by totally expanding hypersurfaces that is hypersurfaces with positive definite second fundamental form.

In this article, we shall give definitions of convex functions on spacetime manifolds $(M, g_{\mu\nu})$ and some examples. We shall mainly discuss the relations between submanifolds such as hypersurfaces in a spacetime and the existence of convex functions in the ambient manifold. The proofs of Propositions and Theorems in this article, and further examples of spacetime convex functions are given in the paper [GI].

2 Convex functions on Riemannian manifolds

2.1 Definitions and a standard example

We shall briefly recapitulate standard definitions of convex functions on an n-dimensional Riemannian manifold (Σ, h_{ij}), which are regarded as spacelike submanifolds embedded in a spacetime or as Euclidean solutions of Einstein's equation.

Definition 1 Convex function:
A C^∞ function $\tilde{f} : \Sigma \to \mathbb{R}$ is said to be (i) convex, (ii) strictly convex, (iii) uniformly convex if the Hessian of $\tilde{f}$ satisfies, respectively,

$$(i) \quad D_i D_j \tilde{f} \geq 0 \,, \tag{2}$$

$$(ii) \quad D_i D_j \tilde{f} > 0 \,, \tag{3}$$

$$(iii) \quad {}^{\exists} c > 0; \, {}^{\forall} V \in T\Sigma, \; V^i V^j D_i D_j \tilde{f} \geq \tilde{c} h_{ij} V^i V^j \,, \tag{4}$$

where D_i is the Levi-Civita connection of the positive definite metric h_{ij}.

Properties of convex functions on Riemannian manifolds have been extensively studied [U].

The standard example is an $\mathbb{E}^n$ with coordinates $\{x^i\}$ and

$$\tilde{f} = \frac{1}{2} \left\{ (x^1)^2 + (x^2)^2 + \cdots + (x^n)^2 \right\} . \tag{5}$$

Note that the vector field $\tilde{D}^i := h^{ij} D_j \tilde{f}$ satisfies $\mathcal{L}_{\tilde{D}} h_{ij} = 2\tilde{c}\, h_{ij}$, and $\tilde{D}$ is the standard dilatation vector field. Thus, Σ admits a homothety.

2.2 Convex functions and submanifolds

Suppose (Σ, h_{ij}) admits a convex function of the type (ii), or (iii). Then we deduce the following proposition.

Proposition 1 *Σ cannot be closed, and cannot admit closed minimal submanifolds $S \subset \Sigma$.*

As a simple example, which also illustrates some of the subtleties involved, consider the following metric,

$$h_{ij} dx^i dx^j = \left(1 + \frac{M}{2\rho} \right)^4 (d\rho^2 + \rho^2 d\Omega^2) \,, \quad M > 0, \; \rho > 0 \,. \tag{6}$$

This is the Schwarzschild initial data in isotropic coordinates. The two-surface S defined as $\rho = M/2$ ($r = 2M$) is totally geodesic ($H_{pq} = 0$) and hence minimal ($k^{pq} H_{pq} = 0$). Thus Σ admits no convex function. Note that if we are concerned only with the region $\rho > M/2$ (outside the trapped region) the restriction of the function

$$\tilde{f} = \frac{1}{2} \left(1 + \frac{M}{2\rho} \right)^4 \rho^2 \tag{7}$$

is a strictly convex function. However it ceases to be *strictly* convex on the totally geodesic submanifold at $\rho = M/2$. Restricted to $\rho = M/2$, the function is a constant, and hence not (strictly) subharmonic. Because $\rho = M/2$ is closed and totally geodesic it contains closed geodesics which are also closed geodesics of the ambient manifold. In fact every such geodesic is a great circle on the 2-sphere $\rho = M/2$. Again this shows that the complete initial data set admits no strictly convex function.

From a spacetime view point, (Σ, h_{ij}) in the example above is a time-symmetric hypersurface and the two-surface S at $\rho = M/2$ is a marginally inner and outer trapped surface in the maximally extended Schwarzschild solution. As an application to time symmetric initial data sets, consider (Σ, h_{ij}, K_{ij}) with the vanishing second fundamental form $K_{ij} = 0$. Then $S \subset \Sigma$ would have been a marginally inner and outer trapped surface. We can observe that if Σ admits a convex function $\tilde{f}$ then no such apparent horizons can exist. This will be discussed further in the next section.

2.3 Mass formula

To get some insight into the physical significance of the existence of convex functions, we shall examine a mass-formula for a 3-dimensional asymptotically flat initial data (Σ, h_{ij}, K_{ij}).

Theorem 1 *Suppose (Σ, h_{ij}) has a vanishing Ricci scaler, i.e., $R_{ij}h^{ij} = 0$ and admits a convex function $\tilde{f}$ such that $D_iD_j\tilde{f} > \tilde{c}h_{ij}$ with $\tilde{c}$ being a positive constant and in the asymptotic region $\tilde{f}$ behaves as*

$$\tilde{f} \sim \frac{1}{2}x^i x^i \,, \tag{8}$$

where $\{x^i\}$ is an asymptotically Euclidean coordinate system in which

$$h_{ij} = \left(1 + \frac{2M}{r}\right)\delta_{ij} + O(r^{-2}) \,, \quad r^2 = x^i x^i \,. \tag{9}$$

Then, the ADM mass must be negative.

From this and the positive mass theorem [SY, W] one obtains the following:

Corollary 1 *A regular asymptotically flat initial data set with vanishing Ricci scaler cannot admit a convex function satisfying* (8).

It is often the case that spacetimes with negative ADM mass have timelike singularities. For example, in the Schwarzschild spacetime with a negative mass, which has a timelike singularity, the function

$$\tilde{f} = \frac{1}{2}r^2 \tag{10}$$

is convex. Thus one might anticipate that the existence of a convex function on timeslices is a signal of the existence of naked singularities.

3 Spacetime convex functions

3.1 Definitions and a canonical example

Now we shall define a convex function on an $(n+1)$-dimensional Lorentzian manifold $(M, g_{\mu\nu})$.

Definition 2 Spacetime convex function:
A smooth function $f : M \to \mathbb{R}$ is called a spacetime convex function *if the Hessian $\nabla_\mu\nabla_\nu f$ has Lorentzian signature and*

$$^\exists c > 0; ^\forall V^\mu \in TM, \nabla_\mu\nabla_\nu f \geq c g_{\mu\nu} \,. \tag{11}$$

One way of thinking about spacetime is in terms of the causal structure provided by the Lorentzian-metric. Since what we shall call a spacetime convex function has a Hessian with Lorentzian signature which defines a causal structure and the spacetimes admitting spacetime convex functions have particular types of causal structures; The forward light cone C_f defined by the Hessian $f_{\mu\nu} := \nabla_\mu \nabla_\nu f$ lies (strictly) inside the light cone C_g defined by the spacetime metric $g_{\mu\nu}$ as depicted in the figure 1. This means, in some intuitive sense, that the condition (11) prevents the "collapse" of the light cone defined by $g_{\mu\nu}$.

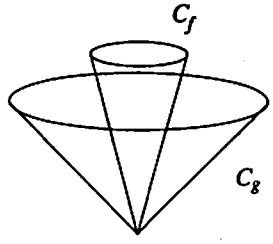


Figure 1: The light cones C_f and C_g.

One of the simplest examples of a spacetime convex function is the canonical one,

$$f = \frac{1}{2}\left(x^i x^i - \alpha t^2\right) \ , \quad (t, x^i) \in \mathbb{E}^{n,1} \ , \tag{12}$$

where α is a constant such that $0 < \alpha \leq 1$.

The Hessian is given by

$$f_{\mu\nu} = \eta_{\mu\nu} + (1-\alpha) t_\mu t_\nu \tag{13}$$

where $\eta_{\mu\nu}$ is the Minkowski metric and $t_\mu := \partial_\mu t$, therefore $f_{\mu\nu}$ actually has Lorentzian signature and can define a light cone C_f. In the particular case $\alpha = 1$, $f_{\mu\nu} = \eta_{\mu\nu}$ and c must be unit.

Note that our conventions mean that, considered as a function of time, f in (12) is what is conventionally called a "concave function of time." However it is a conventional convex function of x^i.

In the canonical example (12), $\mathbb{E}^{n,1} = (\mathbb{R}, -dt^2) \times \mathbb{E}^n$ and f may be regarded as a direct sum of a concave function $-\alpha t^2/2$ on $(\mathbb{R}, -dt^2)$ and a convex function $x^i x^i/2$ on $\mathbb{E}^n$. In general, when a spacetime (M, g) splits isometrically as a product $(\mathbb{R} \times \Sigma, -dt^2 \oplus h)$, one can construct a spacetime convex function by summing concave functions of time $t \in \mathbb{R}$ and convex functions on (Σ, h).

3.2 Spacetime convex functions and submanifolds

Suppose a spacetime convex function f exists in $(M, g_{\mu\nu})$. Then, we deduce

Proposition 2 *$(M, g_{\mu\nu})$ admits no closed spacelike geodesics.*

Proposition 3 *Consider a spacetime $(M, g_{\mu\nu})$ and a closed spacelike surface $S \subset M$ of co-dimension two. Suppose that in a neighbourhood of S the metric is written as*

$$g_{\mu\nu} dx^\mu dx^\nu = \gamma_{ab} dy^a dy^b + k_{pq} dz^p dz^q \ , \tag{14}$$

where $k_{pq} dz^p dz^q$ is the metric on S and the components of the two-dimensional Lorentzian metric γ_{ab} are independent of the coordinates z^p. Then, S cannot be a closed marginally inner and outer trapped surface.

This suggests that the existence of a spacetime convex function might be incompatible with a worm hole like structure.

Suppose $(M, g_{\mu\nu})$ contains an embedded spacelike hypersurface (Σ, h_{ij}) with the second fundamental form K_{ij}. Then we have

$$\nabla_i \nabla_j f = D_i D_j f + K_{ij} \frac{\partial f}{\partial t} , \tag{15}$$

where $\partial f/\partial t := t^\mu \partial_\mu f$ with t^μ being a unit timelike vector, $t^\mu t_\mu = -1$, normal to Σ. Assuming that $(M, g_{\mu\nu})$ admits a spacetime convex function f, we deduce

Proposition 4 *If Σ is totally geodesic, i.e., Σ is a surface of time symmetry, then (Σ, h_{ij}) admits a convex function.*

Proposition 5 *If Σ is maximal, i.e., $h^{ij}K_{ij} = 0$, it admits subharmonic function.*

Corollary 2 *(Σ, h_{ij}) cannot be closed.*

Corollary 3 *(Σ, h_{ij}) admits no closed geodesics nor minimal two surface in the case that $K_{ij} = 0$.*

4 Convex functions and black hole spacetimes

Now let us examine a function (12)

$$f = \frac{1}{2}(r^2 - \alpha t^2) , \tag{16}$$

on the Schwarzschild spacetime, whose metric is given by

$$ds^2 = -\Delta dt^2 + \frac{dr^2}{\Delta} + r^2 \Omega_{pq} dz^p dz^q , \quad \Delta := 1 - \frac{r_g}{r} , \tag{17}$$

where $r_g = 2M$ is the radius of the event horizon. We are concerned with outside the event horizon, where $\Delta > 0$.

For an arbitrary function $f = f(t, r)$, we have

$$\nabla_t \nabla_t f = \left(-\frac{1}{\Delta} \tilde{f} + \frac{\Delta'}{2} f' \right) g_{tt} , \tag{18}$$

$$\nabla_r \nabla_r f = \left(\Delta f'' + \frac{\Delta'}{2} f' \right) g_{rr} , \tag{19}$$

$$\nabla_p \nabla_q f = \frac{\Delta}{r} f' g_{pq} , \tag{20}$$

where the *dot* and the *prime* denote t and r derivatives, respectively.

We see from (18), (19), and (20) that f can be a spacetime convex function if there exists a constant $c > 0$ satisfying

$$0 < \frac{\alpha}{\Delta} + \frac{\Delta'}{2} r \le c , \quad \Delta + \frac{\Delta'}{2} r \ge c , \quad \Delta \ge c . \tag{21}$$

Since Δ and $\alpha/\Delta + \Delta' r/2$ are monotonically increasing and decreasing function, respectively, one can observe that, for $0 \le \alpha < 1$, there exists a radius $r_* (> r_g)$ for which

$$\Delta(r_*) = \frac{\alpha}{\Delta(r_*)} + \frac{\Delta'(r_*)}{2} r_* . \tag{22}$$

Then, if one chooses

$$c = \Delta(r_*) , \tag{23}$$

the condition (21) holds in the region $r_* \le r < \infty$, hence f becomes a spacetime convex function there.

Note that, in the black hole spacetime, the Hessian of (16) has Lorentzian signature even for $\alpha = 0$, while in the strictly flat spacetime, the Hessian becomes positive semi-definite for $\alpha = 0$.

One can also see that, for $\alpha < 0$, the condition (21) cannot hold in the asymptotic region, and for $1 \le \alpha$, (21) holds nowhere.

5 Convex functions and spacetime foliations

5.1 Level sets and foliations

Locally the level sets $\Sigma_c = \{x \in M | f(x) = \text{constant} = c\}$ of a function f have a unit normal given by

$$n_\mu = -\frac{\nabla_\mu f}{\sqrt{\epsilon \nabla_\nu f \nabla^\nu f}}, \tag{24}$$

where $\epsilon = -1$ if the normal is timelike and $\epsilon = 1$ if it is spacelike. Given two vectors X^μ and Y^ν tangent to Σ_c so that $X^\mu \partial_\mu f = 0 = Y^\mu \partial_\mu f$ one may evaluate $K_{\mu\nu} X^\mu Y^\mu$ in terms of the Hessian of f:

$$K_{\mu\nu} X^\mu Y^\mu = X^\mu Y^\nu \frac{\nabla_\mu \nabla_\nu f}{\sqrt{\epsilon \nabla_\nu f \nabla^\nu f}}. \tag{25}$$

Thus, in the Riemannian case (for which the metric $g_{\mu\nu}$ on M is positive definite and $\epsilon = 1$), a strictly convex function has a positive definite second fundamental form. Of course there is a convention here about the choice of direction of the normal. We have chosen f to decrease along n^μ. The converse is not necessarily true, since $f(x)$ and $g(f(x))$ have the same level sets, where g is a monotonic function of $\mathbb{R}$. Using this gauge freedom we may easily change the signature of the Hessian. However, given a hypersurface Σ_0 with positive definite second fundamental form, we may, locally, also use this gauge freedom to find a convex function whose level $f = 0$ coincides with Σ_0, If we have a foliation (often called a "slicing" by relativists) by hypersurfaces with positive definite fundamental form, we may locally represent the leaves as the levels sets of a convex function.

Similar remarks apply for Lorentzian metrics. The case of greatest interest is when the level sets have a timelike normal. For a classical strictly convex function, the second fundamental form will be positive definite and the hypersurface orthogonal timelike congruence given by the normals n^μ is an expanding one. For a spacetime convex function, our conventions also imply that the second fundamental form of a spacelike level set is positive definite. This can be illustrated by the canonical example (12) with $\alpha = 1$ in flat spacetime. The spacelike level sets foliate the interior of the future (or past) light cone. Each leaf is isometric to hyperbolic space. The expansion is homogeneous and isotropic, and if we introduce coordinates adapted to the foliation we obtain the flat metric in open FLRW form with scale factor $a(\tau) = \tau$. This is often called the Milne model. The convex function underlying this FLRW coordinate system is globally well defined but the coordinate system based on it breaks down at the origin of Minkowski spacetime.

From what has been said it is clear that there is a close relation between the existence of convex functions and the existence of foliations with positive definite second fundamental form. A special case, often encountered in practice, are totally umbilic foliations. These have $K_{ij} = \frac{1}{n} g_{ij} \mathrm{Tr} K$. In a 4-dimensional Einstein space for which $R_{\mu\nu} = \Lambda g_{\mu\nu}$ the Gauss-Codazzi equations imply that 3-dimensional umbilic hypersurface Σ_c is a constant curvature space with a sectional curvature C and the trace of K_{ij} is given by $\mathrm{Tr} K = 3\sqrt{\frac{\Lambda}{3} - C}$. The sectional curvature C could depend only upon "time function" f and thus totally umbilic hypersurfaces are a special case of "trace K = constant" hypersurfaces which are used in numerical relativity.

5.2 Barriers and Constant mean curvature foliation

In the theory of maximal hypersurfaces and more generally hypersurfaces of constant mean curvature, an important role is played by the idea of a "barriers." The basic idea is that, if two spacelike hypersurfaces Σ_1 and Σ_2 touch, then the one in the future, Σ_2 say, can have no smaller a mean curvature than the Σ_1, the hypersurface in the past, i.e., $\mathrm{Tr} K_1 \leq \mathrm{Tr} K_2$.

More formally, we have, using the Maximum Principle, the following theorem:

Theorem 2 Eschenburg [E]: *Let M_1 and M_2 be disjoint open domains with spacelike connected C^2-boundaries having a point in common. If the mean curvatures $\mathrm{Tr} K_1$ of ∂M_1 and $\mathrm{Tr} K_2$ of ∂M_2 satisfy*

$$\mathrm{Tr} K_1 \leq -a, \quad \mathrm{Tr} K_2 \leq a\,, \tag{26}$$

for some real number a, then $\partial M_1 = \partial M_2$, and $\mathrm{Tr} K_2 = -\mathrm{Tr} K_1 = a$.

As an example, let us suppose that Σ_1 is maximal, i.e., $\mathrm{Tr}K_1 = 0$, where K_1 is the second fundamental form of Σ_1. Now consider a function f whose level sets lie in the future of Σ_1 and one of which touches Σ_1. The second fundamental form K_f of the level set f =constant must have positive trace, $\mathrm{Tr}K_f > 0$. In particular, $\mathrm{Tr}K_f < 0$ is excluded. Now if f is a spacetime *concave* function, i.e.,

$$\nabla_\mu \nabla_\nu f < -c g_{\mu\nu} \,, \tag{27}$$

for $c > 0$, we have $\mathrm{Tr}K_f < 0$. It follows that the level sets of a concave function cannot penetrate the "barrier" produced by the maximal hypersurface Σ_1.

As an application of this idea, consider the interior region of the Schwarzschild solution. The hypersurface r = constant has

$$\mathrm{Tr}K(r) = -\frac{2}{r}\left(\frac{2M}{r} - 1\right)^{-1/2}\left(1 - \frac{3M}{2r}\right) \,. \tag{28}$$

The sign of $\mathrm{Tr}K(r)$ is determined by the fact that regarding r as a time coordinate, r decrease as time increases. Thus for $r > 3M/2$, $\mathrm{Tr}K$ is negative but for $r < 3M/2$, $\mathrm{Tr}K$ is positive. If $r = 3M/2$, we have $\mathrm{Tr}K = 0$, that is, $r = 3M/2$ is a maximal hypersurface.

Consider now attempting to foliate the interior region by the level sets of a concave function. Every level set must touch an r = constant hypersurface at some point. If $r = 3M/2$ this can, by the theorem, only happen if the level set lies in the past of the hypersurface $r = 3M/2$. Evidently therefore a foliation by level sets of a spacetime *concave* function can never penetrate the barrier at $r = 3M/2$. In particular, such a foliation can never extend to the singularity at $r = 0$.

These results are relevant to work in numerical relativity. One typically sets up a coordinate system in which the constant time surfaces are of constant mean curvature, i.e., constant trace K. It follows from our results that, if the constant is *negative*, then, the coordinate system can never penetrate the region $r < 3M/2$. In fact it could never penetrate any maximal hypersurface. On the other hand, the regular $\mathrm{Tr}K = \lambda > 0$ hypersurfaces can penetrate the barrier hypersurface $r = 3M/2$, and if the value of TrK is large enough, such hypersurfaces can approach the singularity at $r = 0$ arbitrarily closely [BCI].

Since a regular hypersurface which reach asymptotic region has a "throat" on it, or contains a closed minimal surface [GI], we can say from Proposition 1 that:

Remark 1 *A regular constant mean curvature hypersurface in a black hole spacetime do not admit a strictly or uniformly convex function which lives on the hypersurface.*

However, regular Σs which foliate the interior of a black hole do not intersect the closed marginally inner and outer trapped surface. Thus, there remains a possibility that $(M, g_{\mu\nu})$ admits a *spacetime convex* function whose level sets coincide with Σs. This suggests that spacetime convex functions might play more important role in the study of spacetime foliations.

We anticipate that further study of spacetime convex functions and foliations by convex hypersurfaces will provide further insights into global problems in general relativity and should have applications to numerical relativity.

Acknowledgments

G.W.G. would like to thank Profs. T. Nakamura and H. Kodama for their kind hospitality at the Yukawa Institute for Theoretical Physics where the main part of this work was done. A.I. would like to thank members of DAMTP for their kind hospitality. A.I. was supported by Japan Society for the Promotion of Science.

References

[BCI] Brill, D.R., Cavallo, J.M., and Isenberg, J.A.: J. Math. Phys.21, 2789–2796 (1980)

[G] Gibbs, J.W.: On the equilibrium of Heterogeneous substances. Trans. Conn. Acad., vol III, 108–248 (1876); 343–524 (1878); reprinted in *The Scientific Papers of J. Willard Gibbs Vol. One Thermodynamics.* pp. 55–353. Dover Publications, inc. New York New York

[GI] Gibbons, G.W., Ishibashi. A.: gr-qc/0011055

[E] Eschenburg, J.-H.: *Maximum principle for hypersurfaces.* Manuscripta math. 64 55–75 (1989)

[SY] Schoen, R., Yau, S.-T.: Commun. Math. Phys. 65, 45–76 (1979); *ibid* 79, 231–260 (1981)

[U] Udrişte, C.: *Convex Functions and Optimization Methods on Riemannian Manifolds.* Mathematics and its applications vol 297 Kluwer Academic Publishers, 1994

[W] Witten, E.: Commun. Math. Phys. 80, 381–402 (1981)

Padé Approximants for gravitational Waves from Inspiralling Compact Binaries

T. Damour[1] ,

IHES,
91440 Bur-Sur Yvette,
France.

Bala R. Iyer [2] ,

Raman Research Institute,
Bangalore, 560 080,
India

B.S. Sathyaprakash[3]

Department of Physics and Astronomy, Cardiff University, P.O. Box 913, Cardiff, CF2 3YB UK.

Abstract

Re-summation techniques using Padé approximants are introduced to improve the convergence of post-Newtonian expansions for the binding energy and the gravitational wave flux of a compact binary consisting of neutron stars and/or black holes. An effective one-body approach to the two-body dynamics and gravitational wave phasing also allows one to go beyond the adiabatic approach to inspiral. The re-summed versions of the post-Newtonian expansions are *effectual, faithful and efficient* for extracting the true inspiral signal from coalescing compact binaries.

1 Introduction

After forty years of development Gravitational Wave (GW) detectors are sensitive enough for the first detections. Massive binary systems of black holes with total mass greater than about $10M_{\odot}$ are likely to be the first potentially detectable events for these Laser Interferometric GW detectors [1]. Laser interferometers like TAMA are already taking data while others like LIGO, GEO and VIRGO should go online soon. The first events may have small signal-to-noise ratios (SNRs) and barely cross an a priori chosen detection threshold. Consequently, one requires the best signal analysis available. The optimal data analysis strategy is to employ a matched filter which looks for correlations of the noisy data with templates, *i.e.* theoretically predicted waveforms. Good templates exist for inspiralling compact binaries and probably for spinning non-spherical neutron stars, but not yet for supernovae and the merger phase of binary black holes.

Gravitational wave data analysis of inspiralling compact binaries places severe demands on both of the following:

- Data analysis hardware: Data will be collected at the rate of several Mbytes per second. The total volume of data will be about 10's of Tbytes per year; even the reduced data may amount to several Tbytes.

[1] E-mail:damour@ihes.fr
[2] E-mail:bri@rri.res.in
[3] B.Sathyaprakash@cardiff.ac.uk

- Theoretical prediction of waveforms: To build up sufficient SNR the template must not dephase with the signal by even a cycle in the thousands of cycles in the detector bandwidth. This requires, in principle, a 3PN accurate waveform. In the test-particle limit [2, 3, 4] the waveforms are available to 5.5PN orders beyond the standard quadrupole formula [5]. For binaries of comparable masses results are available to only 2.5PN orders beyond the quadrupole order at present [6]. The 3PN computations have been ongoing for the last few years [7, 8, 9]. Unfortunately, at present, we cannot count on the extension of the post-Newtonian (PN) perturbative calculations by another two orders, to order v^7, because currently used PN techniques leave undetermined a physically crucial parameter entering at the v^6 level [7, 8, 9].

To place gravitational wave data analysis in perspective, it is instructive to compare it with the radio observations of binary pulsars which are based on the timing formula

$$\phi_n^{\rm PSR} = F[t_n; p_i]\,, \tag{1}$$

where t_n are the arrival times of the radio pulses and p_i are the binary parameters. It is sufficient to work with the *lowest order* gravitational wave radiation damping in order to derive the accurate timing formula required. Similarly, gravitational wave observations of compact binaries is based on the phasing formula,

$$\phi^{\rm GW} = 2\Phi = F[t; p_i]\,. \tag{2}$$

However, unlike for pulsar timing, the lowest order radiation reaction is not sufficient; one needs crucially the *higher order* gravitational radiation reaction for phasing since the systems one is dealing with are highly relativistic.

The computation of gravitational radiation from an inspiralling compact binary requires the analysis of the following three aspects:

1. Problem of motion: Given a binary system, iterate the Einstein's equations to discuss the conservative motion of the system.
2. Generation problem: Given the motion of the binary on a fixed orbit, iterate Einstein's equations to compute the multipoles of the gravitational field and hence the far-zone flux of energy and angular momentum carried by the gravitational waves.
3. Radiation reaction problem: Given the flux of radiated energy and angular momentum, use the heuristic balance argument to compute the reaction of the emitted flux on the orbit and hence the phasing of the gravitational waves.

Once we have the three modules above in control, we can explicitly write down the gravitational waveform and its phase evolution under gravitational radiation reaction. The relevant expressions are usually given in terms of the invariant velocity $v = (\pi m F)^{1/3}$ and the chirp mass $\mathcal{M} = \nu^{3/5} m$, where $\nu = m_1 m_2/m^2$ is the symmetric mass ratio and m is the total mass. We obtain,

$$h(t) = 2a(t)\cos\phi(t) = a(t)e^{-i\phi(t)} + a(t)e^{i\phi(t)}\,, \tag{3}$$

$$\text{where,}\quad \frac{d\phi(t)}{dt} \equiv 2\pi F(t) > 0\,, \tag{4}$$

$$a(t) = C_{\mathcal{M}}(\pi\mathcal{M}F(t))^{2/3}\,, \tag{5}$$

$$\phi(t) = \phi_c - 2\left[\frac{(t_c - t)}{5\mathcal{M}}\right]^{5/8}\,, \tag{6}$$

$$\pi\mathcal{M}F(t) = \left[\frac{5\mathcal{M}}{256(t_c - t)}\right]^{3/8}\,. \tag{7}$$

The associated phasing formula is given by,

$$t(v_F) = t_c + m\int_{v_F}^{v_{\rm ISO}} dv\,\frac{E'(v)}{F(v)}\,, \tag{8}$$

$$\phi(v_F) = \phi_c + 2\int_{v_F}^{v_{\rm ISO}} dv\, v^3\,\frac{E'(v)}{F(v)}\,. \tag{9}$$

The energy function appearing in the above phasing formulas may be written as,

$$E(v) = \nu \left(\frac{1 - 2v^2}{\sqrt{1 - 3v^2}} - 1 \right) , \tag{10}$$

in the test-mass limit, where $\nu \to 0$. In the finite mass case on the other hand,

$$E_{2PN}(v) = -\frac{1}{2} \nu v^2 \left[1 - \frac{1}{12} (9 + \nu) v^2 - \frac{1}{8} \left(27 - 19\nu + \frac{\nu^2}{3} \right) v^4 \right] . \tag{11}$$

The flux function on the other hand is computed to be:

$$\begin{aligned} F(v,\nu) &= \frac{32}{5} \nu^2 v^{10} \left[1 - \left(\frac{1247}{336} + \frac{35}{12} \nu \right) v^2 + 4\pi v^3 + \left(-\frac{44711}{9072} + \frac{9271}{504} \nu + \frac{65}{18} \nu^2 \right) v^4 - \right. \\ &\left. \left(\frac{8191}{672} + \frac{535}{24} \nu \right) \pi v^5 \right] . \end{aligned} \tag{12}$$

We are now in a position to formulate the *data analysis problem*: Approximation methods yield the energy function E_{T_n} and the far zone flux function F_{T_n} and hence an approximate waveform $h_n^A(t;\mu_k)$ where μ_k are parameters like masses m_1 and m_2 labelling the waveform. The true signal is $h_X(t;\lambda_k)$ where λ_k are the parameters labelling the signal. In general the number of parameters in the two need not be the same. In the data analysis problem we are interested in knowing how well does $h_A(t;\mu_k)$ represent $h_X(t;\mu_k)$. A measure of this is the overlap of the two functions given by

$$\mathcal{O}(k,h) \equiv \frac{\langle h_X, h_A \rangle}{\sqrt{\langle h_X, h_X \rangle \langle h_A, h_A \rangle}} , \tag{13}$$

where the Wiener scalar product is defined by

$$\langle g, h \rangle \equiv 2 \int_{-\infty}^{\infty} \frac{df}{S_n(f)} \tilde{g}^*(f) \tilde{h}(f) . \tag{14}$$

In the above $S_n(f)$ denotes the one-sided noise power spectral density of the detector. For initial LIGO [10], for instance, we have,

$$S_n(f) = S_0 \left[2 + 2 \left(\frac{f}{f_0} \right)^2 + \left(\frac{f}{f_0} \right)^{-4} \right] , \; f \geq f_s , \tag{15}$$

$$= \infty , \; f < f_s , \tag{16}$$

with $f_s = 40$ Hz, $f_0 = 200$ Hz and $S_0 = 1.47 \times 10^{-46}$ Hz^{-1}. The minimum of $S_n(f)$ is at $f = f_0$ and is equal to $S_{\rm min} = 5S_0$. However, a physically more relevant quantity is the minimum of the dimensionless quantity $h_n^2(f) \equiv f S_n(f)$ (effective GW noise). This is reached at the *characteristic detection frequency* $f = f_{\rm det} = 0.8347 f_0$, and is equal to $(h_n^{\rm min})^2 = 4.6 f_0 S_0$. The above numerical value for f_0 and S_0 leads to $f_{\rm det} = 167$ Hz and corresponds to $h_n^{\rm min} = 2.6 \times 10^{-22}$. For VIRGO on the other hand, the corresponding noise curve is given by

$$S_n(f) = S_0 \left[10^3 \left(\frac{f_s}{f} \right)^5 + 2 \left(\frac{f_0}{f} \right) + 1 + \left(\frac{f}{f_0} \right)^2 \right] , \; f \geq f_s , \tag{17}$$

$$= \infty , \; f < f_s . \tag{18}$$

In this case, $f_s = 20$ Hz, $f_0 = 500$ Hz while $S_0 = 3.24 \times 10^{-46}$ Hz^{-1}[11]. The minimum of $h_n(f) = \sqrt{f S_n(f)}$ is reached at $f = 103$ Hz and is equal to $h_n^{\rm min} = 4.3 \times 10^{-22}$. The secret of extracting physics from the detectors lies in our ability to 'tune' the noise curves to the sources we are interested in.

It should be noted that GW data analysis is compute intensive since it depends on multiple parameters like masses and spins. Therefore, one needs to have the best algorithmic implementation of the search routines. The algorithm must be optimal, *i.e* it should return the maximum SNR and efficient, *i.e.* it

should take the least search time. The best strategy seems to be the hierarchical search proposed by Mohanty and Dhurandhar [12]. In this method one would make a coarse search for binaries of component masses in the range 0.25–50$M_\odot$ with about 50,000 templates stored in memory, requiring about 500 Gbytes. This would be followed by a finer search that would require ten times more the number of templates in the coarse grid. But these fine grid templates cannot be stored and must be computed online.

The normal route to the construction of templates is via Taylor expansions:

$$(E_{T_n}\,,\,\mathcal{F}_{T_n}) \xrightarrow{T} h_n^T(t,\lambda_k)\,. \tag{19}$$

However Post-Newtonian expansions are very slowly convergent. They are bad since they are oscillatory. For instance, the 2.5PN approximant performs worse compared to the 2PN order approximant. This led Damour, Iyer and Sathyaprakash (DIS) to the proposal of the P-filters [13]. The suggestion was to use Padé techniques to improve the convergence of the poorly converging Taylor approximants. The trick is rather well-known. However, in practice some impediments must be overcome before one can proceed. They are (1) presense of logarithmic terms in the flux for PN orders greater than 6, (2) the energy function has a branch cut, and (3) the energy flux blows up near the light ring. DIS introduced new energy and flux functions to get around the above obstacles.

Before defining new "re-summed" energy and flux functions with improved performances, we digress for a brief reminder of Padé re-summation, which is a standard mathematical technique used to accelerate the convergence of poorly converging power series. Let $S_n(v) = a_0 + a_1\,v + \cdots + a_n\,v^n$, be a truncated Taylor series of a function $S(v)$. A Padé approximant of the function, whose Taylor approximant to order v^n is S_n, is defined by two integers m, k, such that $m+k = n$; if $T_n[\cdots]$ denotes the operation of expanding a function in Taylor series and truncating it to accuracy v^n (included), the P_k^m Padé approximant of S_n is defined by

$$P_k^m(v) = \frac{N_m(v)}{D_k(v)};\;\; T_n[P_m^k(v)] \equiv S_n(v), \tag{20}$$

where N_m and D_k are *polynomials* in v of order m and k, respectively. If one assumes that $D_k(v)$ is normalised so that $D_k(0) = 1$; i.e. $D_k(v) = 1 + q_1\,v + \cdots$, one shows that Padé approximants are uniquely defined by Eq. (20). In many cases the most useful[4] Padé approximants are the ones near the "diagonal", $m = k$, i.e. P_m^m, if $n = 2m$ is even, and P_m^{m+1} or P_{m+1}^m, if $n = 2m+1$ is odd. In this work we shall use the diagonal (P_m^m) and the "sub-diagonal" (P_{m+1}^m) approximants which can be conveniently[5] written in a continued fraction form [14]. For example, given $S_2(v) = a_0 + a_1\,v + a_2\,v^2$ one looks for

$$P_1^1(v) = \frac{c_0}{1 + \frac{c_1 v}{1+c_2 v}} = \frac{c_0(1+c_2 v)}{1 + (c_1+c_2)v}. \tag{21}$$

The continued fraction Padé coefficients of a series containing six terms, i.e. $S_5(v)$, are given by

$$\begin{aligned}
c_0 &= a_0,\;\; c_1 = -\frac{a_1}{a_0},\;\; c_2 = -\frac{a_2}{a_1} + \frac{a_1}{a_0},\;\; c_3 = \frac{a_0(a_1 a_3 - a_2^2)}{a_1(a_1^2 - a_2 a_0)} \\
c_4 &= -\frac{c_0 c_1 (c_2+c_1)^3 + c_0 c_1 c_2 c_3 (c_3 + 2c_2 + 2c_1) - a_4}{c_0 c_1 c_2 c_3}, \\
c_5 &= -\frac{((c_2+c_1)^2 + c_2 c_3)^2}{c_2 c_3 c_4} - \frac{(c_4 + c_3 + c_2 + c_1)^2}{c_4} - \frac{a_5}{c_0 c_1 c_2 c_3 c_4}\,.
\end{aligned} \tag{22}$$

In [13] DIS provide a ν-dependent estimate of the rate of loss of angular momentum under gravitational radiation damping for comparable mass binaries in quasi-circular orbits using the following inputs. (a) Work only with invariant functions like conserved energy and far-zone energy flux of invariant arguments like the invariant velocity related to the gravitational wave frequency, (b) Inject plausible information

[4] The rare theorems dealing with the Padé technique concern the convergence of "near-diagonal" Padé approximants, i.e. $m \to \infty$ with $|m-k|$ fixed.

[5] A convenience of this form is that the n-th continued-fraction coefficient c_n (see below) depends only on the knowledge of the PN coefficients of order $\leq n$.

about the meromorphic structure of the these functions one wishes to to obtain, (c) Employ Padé approximants to estimate the functions from the first few known terms in the perturbative expansion of these functions. The new energy and flux functions proposed in DIS1 are:

$$e(x) \equiv \left(\frac{E_{\rm tot}^2 - m_1^2 - m_2^2}{2m_1 m_2}\right)^2 - 1, \tag{23}$$

$$\hat{f}(v;\nu) \equiv (1 - v/v_{\rm pole})\,\hat{\mathcal{F}}(v;\nu), \tag{24}$$

where $\hat{\mathcal{F}}(v) \equiv \mathcal{F}(v)/\mathcal{F}_N(v) = 5\mathcal{F}(v)/(32\nu^2 v^{10})$, is the Newton-normalised flux. In terms of them the construction of the P-filters may be schematically shown as:

$$(E_{T_n}, \mathcal{F}_{T_n}) \to (e_{T_n}, f_{T_n}) \to (e_{P_n}, f_{P_n}) \to (E[e_{P_n}], \mathcal{F}[e_{P_n}, f_{P_n}]) \to h_n^P(t, \lambda_k). \tag{25}$$

In DIS1 [13] it was shown that the P-approximants are *effectual*. Their overlap with fiducial exact signals is ≥ 0.965 after maximisation over all the parameters. This result is relevant for the choice of search templates. Moreover, the P-approximants are *faithful*. They provide good overlaps for small differences in parameters. They are accurate and this aspect is relevant for parameter estimation. However, there is a point of worry. The *time-domain* P-approximants discussed above are computationally expensive. They are not fast enough since they have to be Fourier transformed for use in filtering. To overcome this handicap in DIS2 [15] we constructed *frequency-domain* P-filters.

To proceed further let us recall that Fourier domain phasing formulas are based on the stationary phase approximation (SPA).

$$\tilde{h}^{\rm uspa}(f) = \frac{a(t_f)}{\sqrt{\dot{F}(t_f)}} e^{i[\psi_f(t_f) - \pi/4]}, \quad \psi_f(t) \equiv 2\pi f t - \phi(t), \tag{26}$$

where t_f is the saddle point defined by solving $d\psi_f(t)/dt = 0$ for t, i.e. the time t_f when the GW frequency $F(t)$ becomes equal to the Fourier variable f. In the (adiabatic) approximation the value of t_f is given by the following integral:

$$t_f = t_{\rm ref} + m \int_{v_f}^{v_{\rm ref}} \frac{E'(v)}{\mathcal{F}(v)} dv, \tag{27}$$

where $v_f \equiv (\pi m f)^{1/3}$. Using t_f from the above equation and $\phi(t_f)$ in the phasing formula one finds that

$$\psi_f(t_f) = 2\pi f t_{\rm ref} - \phi_{\rm ref} + 2\int_{v_f}^{v_{\rm ref}} (v_f^3 - v^3) \frac{E'(v)}{\mathcal{F}(v)} dv. \tag{28}$$

In the above equation the Fourier phase is given as an explicit function of frequency and hence faster than the *time-domain* procedure. Recently, in [16] there was an examination of the adequacy of the SPA for GW data analysis via matched filtering. This work estimated that the higher order corrections to the SPA via the method of steepest descent and showed that the higher order corrections were small. It also studied the effect of windowing on the overlaps and concluded that these were *always* small and consequently that the SPA was *always* adequate. DIS2, on the other hand, showed that if the *time-domain* signal is a time-truncated chirp, *i.e.* a chirp abruptly terminated at the last stable orbit (lso)

$$h(t) = 2a(t)\cos\phi(t)\theta(t_{\rm max} - t), \tag{29}$$

where θ denotes the Heaviside step function, ($\theta(x) = 1$, for $x > 0$ and $\theta(x) = 0$, for $x < 0$), there exist edge corrections. The time-windowing has three effects: (a) It induces oscillations in the usually considered frequency-domain $f < F_{max}$, (b) it generates a tail in the usually disregarded frequency-domain $f > F_{max}$ and finally, (c) it renders non-negligible the 'non-resonant' contribution $\tilde{h}_+(f)$. This was consistent with the findings of DIS1 that the SPA became bad as compared to the FFT both as the total mass or the PN order increased. This was because of the smaller number of cycles in the effective detector bandwidth. They had warned of the inadequacy of the SPA and in [13] had worked in the time-domain to avoid uncontrolled numerical errors in the comparison of the Taylor approximants versus the Padé approximants. This was reconfirmed in [15]. It was shown that the effect of windowing was

significant on overlaps for total mass greater than about $20M_\odot$. Recall that binary black holes in the mass range of $(28-30M_\odot)$ have the best SNR and may be the most promising sources for ground based laser interferometers. However, for such systems $F_{lso} \sim F_{det}$. Thus the most promising sources $(5-40M_\odot)$ are in the regime where the window oscillations are important. For such systems one needs a more careful treatment of the Fourier transform. DIS2 [15] constructs the improved SPA via complementary error functions to model the edge oscillations as follows:

$$f \leq F_{\max}: \quad \tilde{h}^{\rm inspa}_{-<} = \mathcal{C}(\zeta_<(f)) \frac{a(t_f)}{\sqrt{\dot{F}(t_f)}} e^{i(\psi_f(t_f)-\pi/4)}, \tag{30}$$

$$f \geq F_{\max}: \tilde{h}^{\rm inspa}_{->}(f) = \mathcal{C}\left(\zeta_>(f)\right) \frac{a(t_{\max})}{\sqrt{\dot{F}(t_{\max})}} \exp i \left[\psi_f(t_{\max}) + \frac{\pi(f-F_{\max})^2}{\dot{F}(t_{\max})} - \pi/4\right], \tag{31}$$

$$\mathcal{C}(\zeta) \equiv \frac{1}{2}\mathrm{erfc}\left(e^{\frac{i\pi}{4}}\zeta\right), \tag{32}$$

$$\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^{+\infty} e^{-x^2}\,dx, \tag{33}$$

$$\zeta_< \equiv -\sqrt{\psi_f(t_f) - \psi_f(t_{\max})}, \tag{34}$$

$$\zeta_>(f) = \frac{\sqrt{\pi}(f-F_{\max})}{\sqrt{\dot{F}(t_{\max})}}. \tag{35}$$

Using this one can marry the Padé approximants to the SPA and provide the Stationary Phase Padé (SPP) approximants. Assuming the adiabatic approximation to the inspiral all the way up to the LSO leads to $F(t) \to \infty$. We can construct an analytic representation called the improved relativistic SPA for this case. For details see [15].

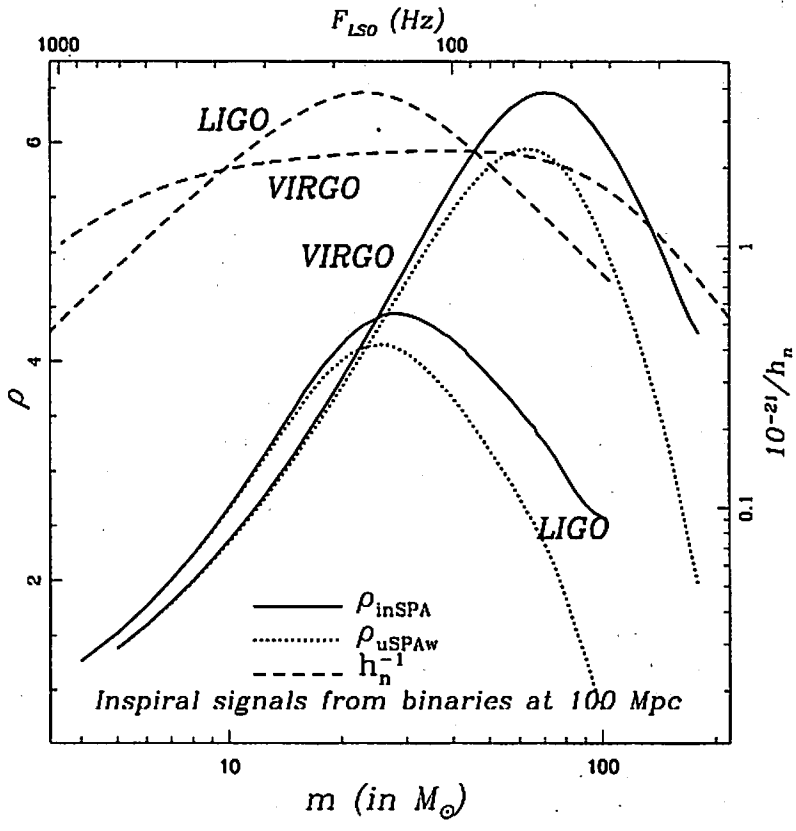
It is interesting to investigate the dependence of the above analysis on the noise curve of the detectors by comparing the results for LIGO and VIRGO. The summary of this study is presented in Fig. 1. The effective sensitivity of LIGO I peaks near 167 Hz, which is the lso frequency of a binary of total mass $27M_\odot$. The effective sensitivity of VIRGO peaks near 104 Hz. Lighter binaries with total mass $< 30M_\odot$ integrate longer in the low frequency. The correction to the usual SPA is less important and the usual SPA follows the improved SPA for total mass $< 30M_\odot$. For LIGO on the other hand it is important to correct the usual SPA for masses beyond $15M_\odot$; the usual SPA loses about 40% of the events. Unlike LIGO even for low and medium mass binaries VIRGO will require P-filters to phase better.

The performance of the SPP approximants are as illustrated in the Tables 1, 2, and 3, which explore their accuracy, faithfulness and effectualness respectively. From these results and others on computing costs we can see that SPP approximants are *accurate, effectual, faithful and efficient.*

A crucial quantity in these discussions above is the location of the last stable orbit (lso). Earliar works to determine the lso include the perturbative approach due to Lincoln and Will [17] and a partial resummation approach called the hybrid method due to Kidder, Will and Wiseman [18]. The latter analysis predicts a lso in the comparable mass case higher than $6M$ corresponding to the test mass case. The orbital frequency at the lso is lower relative to the test particle motion in a Schwarzschild background. It also leads to ν-dependent corrections $> 100\%$ and hence one cannot be sure about the robustness of the result or the consistency of the approach. It should also be noted that in all the treatment upto now, we restricted ourselves to the standard "adiabatic approximation", where one estimates the phasing by combining the energy-balance equation $dE_{\rm tot}/dt = -\mathcal{F}$ with some resummed estimates for the energy and flux as functions of the instantaneous circular orbital frequency. To make further progress, it is imperative to attempt an answer to the following question: Is the adiabatic approximation adequate? The answer is in the negative and one feels the need for an improved knowledge of the transition between the inspiral phase and the plunge phase. To go beyond the adiabatic approximation one notes that the plunge at the lso is determined by the conserved part of the non-linear relativistic force law and corresponds to $R = 6M$ for a test particle in Schwarzschild background. Recently, Buonnano and Damour [19] have introduced a new approach to GW from coalescing binaries which is no longer limited to the adiabatic

approximation, and which is expected to describe rather accurately the transition between the inspiral and the plunge, and to give also a first estimate of the following plunge signal. For this analysis they combined two nonperturbative resummation techniques [13, 20] with information about the allowed forms for gravitational radiation reaction in generic gauges [21].

Figure 1: This Figure compares the signal-to-noise ratio (SNR) for an inspiral signal searched by means of two different filters, in two different interferometers [initial LIGO, Eq. (15) or VIRGO, Eq. (17)]. The exact signal h is assumed to be a time-truncated Newtonian chirp (sufficiently well approximated by the inSPA). The solid lines represent the optimal SNR, obtained when the filter k is identical with the exact signal h. The dotted lines represent the sub-optimal SNR, obtained when the filter k is the frequency-windowed usual SPA (uSPAw). In the case of LIGO, for low mass binaries uSPAw extracts the full SNR, but at higher masses, which are the most crucial ones for initial interferometers, it loses SNR up to a factor of 1.5 leading to a loss in the number of detectable events up to 70%. In the case of VIRGO, uSPAw performs quite well when the total mass $m < 50M_\odot$, and requires inSPA for heavier binaries. The loss in the number of events caused by uSPAw for heavier mass binaries in the range $m > 50M_\odot$ is unacceptable.



The first resummation technique is due to Buonanno and Damour [20] which provided a non-perturbative ν-dependent estimate of the conservative non-linear force law by mapping the actual two-body problem

Table 1: Accuracy of the various approximations to the Fourier transform of a chirp signal computed in the test mass limit for P-approximants at different post-Newtonian orders v^n (column 1) is measured by their overlaps with a waveform of exactly the same parameters but computed in the time-domain and then Fourier transformed using an FFT algorithm. The approximations considered are the usual uSPAw (column 2), inSPAw (column 3) and irSPA (column 4). Both uSPA and inSPA are windowed in frequency beyond $F_{\rm LSO}$. Only the irSPA has a spectrum extending beyond $F_{\rm LSO}$. For more details see DIS2.

n	uSPAw	inSPAw	irSPA
	$\left\langle \tilde{h}^{\rm FFT}_{\rm P_n}, \tilde{h}^{\rm uSPAw}_{\rm P_n} \right\rangle$	$\left\langle \tilde{h}^{\rm FFT}_{\rm P_n}, \tilde{h}^{\rm inSPA}_{\rm P_n} \right\rangle$	$\left\langle \tilde{h}^{\rm FFT}_{\rm P_n}, \tilde{h}^{\rm irSPA}_{\rm P_n} \right\rangle$
	$m_1 = 1.4M_\odot,\ m_2 = 10M_\odot$		
4	0.9967	0.9986	0.9994
5	0.9965	0.9990	0.9997
6	0.9965	0.9986	0.9993
	$m_1 = m_2 = 10M_\odot$		
4	0.9764	0.9762	0.9951
5	0.9771	0.9775	0.9955
6	0.9751	0.9786	0.9953
	$m_1 = m_2 = 20M_\odot$		
4	0.8613	0.8613	0.9891
5	0.8680	0.8773	0.9819
6	0.8897	0.9038	0.9829

Table 2: *Faithfulness* of the Fourier domain P-approximants in the formal *test mass* case. Values quoted are the *minimax* overlap (see DIS1) of an approximant waveform generated directly in the frequency-domain with the exact waveform generated in the time-domain and Fourier transformed using the FFT algorithm. The approximations considered are the usual uSPAw (column 2) and irSPA (column 3). For compactness of presentation, in this Table and 3 we use for the column headings the abbreviation uSPAw to denote $\left\langle \tilde{h}^{\rm FFT}_{\rm X}, \tilde{h}^{\rm uSPAw}_{\rm P_n} \right\rangle$ irSPA to denote $\left\langle \tilde{h}^{\rm FFT}_{\rm X}, \tilde{h}^{\rm irSPA}_{\rm P_n} \right\rangle$.

n	(1.4,10)		(10,10)		(13,13)		(20,20)	
	uSPAw	irSPA	uSPAw	irSPA	uSPAw	irSPA	uSPAw	irSPA
4	0.8360	0.8316	0.9596	0.9707	0.9513	0.9965	0.8248	0.9785
5	0.9755	0.9729	0.9727	0.9914	0.9473	0.9973	0.8283	0.9855
6	0.9921	0.9903	0.9739	0.9938	0.9479	0.9972	0.8285	0.9856

Table 3: *Effectualness* of the Fourier domain P-approximants in the formal *test mass* case. We list *minimax* overlaps of the approximate waveforms generated directly in the frequency-domain with the exact waveform generated in the time-domain and Fourier transformed using an FFT algorithm. However, in addition to maximisation over the lag parameter τ and the initial phase of the approximate waveform we also maximize over the two masses m_1 and m_2. The percentage bias in the estimation of the total mass $100(1 - m_A/m)$ are listed in parenthesis below the overlaps.

n	(1.4,10)		(10,10)		(13,13)		(20,20)	
	uSPAw	irSPA	uSPAw	irSPA	uSPAw	irSPA	uSPAw	irSPA
4	0.9942	0.9963	0.9798	0.9984	0.9651	0.9979	0.9166	0.9966
	(−2.300)	(−2.744)	(−0.664)	(−0.858)	(−0.962)	(−1.081)	(−2.500)	(−1.250)
5	0.9882	0.9987	0.9795	0.9979	0.9653	0.9968	0.9174	0.9965
	(+0.480)	(−1.301)	(−0.626)	(−0.467)	(−0.631)	(−0.481)	(−1.875)	(−0.340)
6	0.9970	0.9982	0.9805	0.9964	0.9649	0.9963	0.9174	0.9966
	(−0.311)	(−1.166)	(0.000)	(−0.626)	(−0.662)	(0.000)	(−1.250)	(−0.340)

to an effective one-body motion of the reduced mass in an effective background metric $g^{eff}_{\mu\nu}(x)$. The ν-dependent terms are small near the lso in this analysis. This provides an alternative and improved proposal to DIS1 to estimate the energy function.

The other input one requires is the acceleration induced by gravitational radiation reaction in binary systems. The most general form for relative acceleration of the binary in general orbits due to radiation reaction was derived by Iyer and Will [21]. They assumed balance equations for energy and angular momentum that relate the time variation of the 2PN conserved energy and angular momentum to far-zone fluxes of energy and angular-momentum. The relative acceleration is determined up to a gauge. The residual gauge freedom is related to a shift of the conservative coordinate system by radiative corrections. Then, using post-Newtonian expressions for the energy and the angular momentum flux at infinity, and assuming energy and angular momentum balance, they obtained at lowest (Newtonian) fractional order:

$$\mathcal{F}_i^{\rm Lag} = \mu \left[\alpha(R,v)\,\dot{R}\,n^i + \beta(R,v)\,v^i\right], \tag{36}$$

$$\alpha(R,v) = \frac{8}{5}\,\nu\,\frac{M}{R^2}\,\frac{M}{R}\,(A_{5/2} + A'_{7/2} + \cdots), \tag{37}$$

$$\beta(R,v) = -\frac{8}{5}\,\nu\,\frac{M}{R^2}\,\frac{M}{R}\,(B_{5/2} + B'_{7/2} + \cdots), \tag{38}$$

$$A_{5/2} = 3\,(1+\overline{\beta})\,v^2 + \frac{1}{3}\,(23 + 6\,\overline{\alpha} - 9\,\overline{\beta})\,\frac{M}{R} - 5\,\overline{\beta}\,\dot{R}^2, \tag{39}$$

$$B_{5/2} = (2+\overline{\alpha})\,v^2 + (2-\overline{\alpha})\,\frac{M}{R} - 3\,(1+\overline{\alpha})\,\dot{R}^2, \tag{40}$$

where R is the relative radius and v is the relative (coordinate)- velocity. See Ref. [21] for the expressions of the 1PN-accurate radiation damping terms $A_{7/2}$ and $B_{7/2}$ in the equations of motion (equivalent, after some reshuffling, with the Lagrangian contributions $A'_{7/2}$, $B'_{7/2}$ in Eqs. (37), (38)) and [22] for the corresponding 4.5PN terms.

The final input to go beyond the adiabatic approximation is the second resummation technique adapted from DIS1 which provided a ν-dependent estimate of the rate of loss of angular momentum under gravitational radiation damping.

The basic equations to go beyond the adiabatic approximation are the following:

$$\frac{dr}{d\hat{t}} = \frac{\partial \widehat{H}}{\partial p_r}(r, p_r, p_\varphi), \tag{41}$$

$$\frac{d\varphi}{d\hat{t}} = \widehat{\omega} \equiv \frac{\partial \widehat{H}}{\partial p_\varphi}(r, p_r, p_\varphi), \tag{42}$$

$$\frac{dp_r}{d\hat{t}} + \frac{\partial \widehat{H}}{\partial r}(r, p_r, p_\varphi) = 0, \tag{43}$$

$$\frac{dp_\varphi}{d\hat{t}} = \widehat{\mathcal{F}}_\varphi(\widehat{\omega}(r, p_r, p_\varphi)), \tag{44}$$

with

$$\widehat{H} = \frac{1}{\nu}\sqrt{1 + 2\nu\left[\sqrt{A(r)\left(1 + \frac{p_r^2}{B(r)} + \frac{p_\varphi^2}{r^2}\right)} - 1\right]}, \tag{45}$$

$$\widehat{\mathcal{F}}_\varphi(v_\omega \equiv \widehat{\omega}^{1/3}) = \frac{\mathcal{F}_\varphi}{\mu} = -\frac{32}{5}\,\nu\,v_\omega^7\,\frac{\widehat{f}_{\rm DIS}(v_\omega;\nu)}{1 - v_\omega/v_{\rm pole}^{\rm DIS}(\nu)}, \tag{46}$$

and where in Eq. (45) we use the scaled versions of our current best estimate of the effective metric coefficients $A(r)$, $B(r)$, (see [20])

$$A(r) \equiv 1 - \frac{2}{r} + \frac{2\nu}{r^3}, \qquad B(r) \equiv \frac{1}{A(r)}\left(1 - \frac{6\nu}{r^2}\right). \tag{47}$$

The analysis in [19] provides the initial dynamical data, *i.e.* position and momenta for binary black holes that have started to plunge. Numerical relativity will probably not be able to accurately evolve binary systems over even ten orbits for many years. Buonnano and Damour propose a method to compute accurate initial data at an epoch so late in the evolution that there remains less than an orbit, (about $\sim .6$ of an orbit) to evolve. The gravitational data is still incomplete. From the point of view of gravitational wave data analysis the paper provides a 10% accurate complete waveform emitted by coalescence of two black holes. The adiabatic inspiral is followed by plunge and eventual coalescence in a stationary state. Details of the transition to plunge is important for construction of faithful templates for masses $5 < M < 40 M_{\odot}$, to test the performance of the SPP templates in [15] and further attempts to improve them.

To conclude: the Stationary Phase Padé (SPP) filters are effectual, *i.e.*, they return bigger overlaps and this is important for detection templates. They are faithful, *i.e.*, they introduce small biases in parameter estimation and this is important to extract astrophysical information. They are efficient, *i.e.*, computationally inexpensive and this is relevant for online implementation.

A resource paper DIS3 [23] comparing search templates for gravitational waves from binary inspiral is now available. This compares the performance of the SPP filters relative to the new and more complete *effective one body waveforms* of Buonanno and Damour. DIS3 summarises the situation and provides a detailed comparison of the Taylor and P- approximants as also the route to go beyond the usual adiabatic approximation by use of the effective one body templates.

References

[1] V.M. Lipunov, K.A. Postnov, and M.E. Prokhorov, New Astron. **2**, 43 (1997);

[2] C. Cutler, et. al. Phys. Rev. Lett. **70**, 2984 (1993).

[3] E. Poisson, Phys. Rev. D **47**, 1497 (1993).

[4] E. Poisson, Phys. Rev. D **52**, 5719 (1995).

[5] T.Tanaka, H.Tagoshi and M.Sasaki, Prog. Theor. Phys. **96**, 1087 (1996).

[6] L. Blanchet, T. Damour, B.R. Iyer, C.M. Will and A. Wiseman, Phys. Rev. Lett. **74**, 3515 (1995); L. Blanchet, T. Damour and B.R. Iyer, Phys. Rev. D **51**, 5360 (1995); C.M. Will and A.G. Wiseman, Phys. Rev. D **54**, 4813 (1996); L. Blanchet, B.R. Iyer, C.M. Will and A. Wiseman, Class. and Quantum Grav. **13**, 575 (1996); L. Blanchet, Phys. Rev. D **54**, 1417 (1996).

[7] P. Jaranowski and G. Schäfer, Phys. Rev. D **57**, 5948 (1998); *ibid* D **57**, 7274 (1998); *ibid* D **60**, 1240003 (1999).

[8] T. Damour, P. Jaranowski and G. Schäfer, Phys. Rev. D **62**, 021501; *ibid* D **62**, 044024; *ibid* D **62**, 084011 (2000).

[9] L. Blanchet and G. Faye, Phys. Lett., **271**, 58 (2000); L. Blanchet and G. Faye, gr-qc/0004009; gr-qc/0006100; gr-qc/7051.

[10] We use the functional form used in C. Cutler and É.É. Flanagan, Phys. Rev. D **49**, 2658 (1994). The value of S_0 is read off using the curves in K.S. Thorne, in *Proceedings of the Snowmass 95 summer study on particle and nuclear astrophysics*, Eds. E.W. Kolb and R. Peccei (World Scientific, Singapore, 1995) (gr-qc 9506086); and B.F. Schutz, in, *Relativistic Gravitation and Gravitational Radiation*, Eds. J.A. Marck, J.P. Lasota. Cambridge Univ. Press (1997), 447-475.

[11] M. Beccaria, E. Cuoco et G. Curci, VIRGO Internal Note, VIR-NOT-PIS-1390-096, (1996).

[12] S. Mohanty and S.V. Dhurandhar, Phys. Rev. D **54**, 7108, (1996); S. Mohanty, Phys. Rev. D **57**, 630, (1998).

[13] T. Damour, B.R. Iyer, and B.S. Sathyaprakash, Phys. Rev. D **57**, 885 (1998) referred to as DIS1.

[14] C.M. Bender and S.A. Orszag, *Advanced mathematical methods for scientists and engineers* (McGraw Hill, Singapore, 1984).

[15] T. Damour, B.R. Iyer, and B.S. Sathyaprakash, Phys. Rev. D **62**, 084036 (2000) referred to as DIS2.

[16] S. Droz, D.J. Knapp, E. Poisson and B.J. Owen, Phys. Rev. D **59**, 124016 (1999).

[17] C.W. Lincoln and C.M. Will, Phys. Rev. **D42**, 1123 (1990).

[18] L.E. Kidder, C.M. Will and A.G. Wiseman, Class. Quantum Grav. **9**, L127 (1992); Phys. Rev. **D47**, 3281 (1993).

[19] A. Buonanno and T. Damour, Phys. Rev. D **62**, 064015 (2000).

[20] A. Buonanno and T. Damour, Phys.Rev. D **59**, 084006 (1999).

[21] B.R. Iyer and C.M. Will, Phys. Rev. Lett. **70**, 113 (1993); Phys. Rev. **D52**, 6882 (1995).

[22] A. Gopakumar, B.R. Iyer and S. Iyer, Phys. Rev. **D55**, 6030 (1997); and Erratum: ibid. **D57**, 6562 (1998).

[23] T. Damour, B.R. Iyer and B.S. Sathyaprakash, *A comparison of search templates for gravitational waves from binary inspiral*, Phys. Rev. D (in press) gr-qc/0010009 (2000) referred to as DIS3.

Type IIB Matrix Model as a Constructive Formulation of Superstring [1]

Satoshi Iso,[2]

High Energy Accelerator Research Organization (KEK),
Tsukuba, Ibaraki, 305-0801, Japan

Abstract

Type IIB matrix model has been proposed as a constructive formulation of superstring. In this note, I briefly review several properties of the model and explain why this model can be considered as a constructive formulation of superstring. In particular, I review where diffeomorphism invariance is hidden in this model.

1 Introduction

Several proposals have been made as constructive definitions of superstring theory [1, 5]. The type-IIB matrix model [1, 2, 3], a large N reduced model of maximally supersymmetric Yang-Mills theory, is one of those proposals. It is defined by the following action:

$$S = -\frac{1}{g^2} Tr(\frac{1}{4}[A_\mu, A_\nu][A^\mu, A^\nu] + \frac{1}{2}\bar{\psi}\Gamma^\mu[A_\mu, \psi]), \tag{1}$$

where A_μ and ψ are $N \times N$ Hermitian matrices, the former is a ten-dimensional vector and the latter is a ten-dimensional Majorana-Weyl spinor field respectively. It is formulated in a manifestly covariant way, which is suitable for studying nonperturbative issues of superstring theory. Since it is a simple model of matrices in zero dimension, it does not possess degenerate vacua unlike its higher dimensinal cousins. It is possible that the model possesses a unique vacuum, namely our space-time. If so, we can in principle predict the dimensionality of the space-time, low-energy gauge group and matter contents by solving this model. In such an endeavor, this model can be studied by numerical simulations effectively. In this note, we review the IIB matrix model and explain how the space-time appears dynamically and how the low energy gauge symmetry and diffeomorphism invariance emerge microscopically.

We first list several important properties of the IIB matrix model. This model can be regarded as a large N reduced model of ten-dimensional $\mathcal{N} = 1$ supersymmetric $SU(N)$ Yang-Mills theory. It was shown [6] that a large N gauge theory can be equivalently described by its reduce model, namely a model defined on a single point. In this reduction procedure, a space-time translation is represented in the color $SU(N)$ space, and the eigenvalues of the matrices are interpreted as the momenta of fields. Therefore, the basic assumption in this identification is that the eigenvalues are uniformly distributed. As a constructive definition of a superstring, on the other hand, we will see that we need to interpret the eigenvalues of matrices as the coordinates of space-time points, the interpretation of which is T-dual to the above.

Since our IIB matrix model is defined on a single point, the commutator of the supersymmetry which we inherit from the ten-dimensional $\mathcal{N} = 1$ supersymmetric $SU(N)$ Yang-Mills theory ,

$$\begin{cases} \delta^{(1)} A_\mu &= i\epsilon_1 \Gamma_\mu \psi \\ \delta^{(1)} \psi &= \frac{i}{2}\Gamma^{\mu\nu}[A^\mu, A^\nu]\epsilon_1, \end{cases} \tag{2}$$

vanishes up to a field-dependent gauge transformation, and we can no longer interpret this supersymmetry as space-time supersymmetry in the original sense. However, after the reduction, we acquire an extra bosonic symmetry,

$$\delta A_\mu = c_\mu \mathbf{1}, \tag{3}$$

[1] This note is based on several collaborations with H. Aoki, N. Ishibashi, H. Kawai, Y. Kitazawa and T. Tada.
[2] E-mail:satoshi.iso@kek.jp

whose transformation is proportional to the unit matrix 1 and an extra supersymmetry,

$$\begin{cases} \delta^{(2)} A_\mu &= 0 \\ \delta^{(2)} \psi &= \epsilon_2. \end{cases} \tag{4}$$

The linear combinations of these two supersymmetries (2) and (4),

$$\tilde{Q}^{(1)} = Q^{(1)} + Q^{(2)}, \ \tilde{Q}^{(2)} = i(Q^{(1)} - Q^{(2)}), \tag{5}$$

satisfy the following commutation relations,

$$[\bar{\epsilon_1} \tilde{Q^{(i)}}, \bar{\epsilon_2} \tilde{Q^{(j)}}] = -2\bar{\epsilon_1} \gamma_\mu \epsilon_2 p^\mu \delta^{(ij)}, \tag{6}$$

where p^μ is the generator of the translation (3) and $i, j = 1, 2$. Therefore, if we interpret the eigenvalues of the matrices A_μ as our space-time coordinates, the above symmetries can be regarded as ten-dimensional $\mathcal{N} = 2$ space-time supersymmetry. Since the maximal space-time supersymmetry guarantees the existence of gravitons if the theory permits the massles spectrum , it supports our conjecture that the IIB matrix model is a constructive definition of superstring. This is one of the major reasons to interpret the eigenvalues of A_μ as being the coordinates of the space-time which has emerged out of the matrices.

The second important and confusing property is that the model has the same action as the low-energy effective action of D-instantons. We should emphasize here the differences between these two theories, since we are led to different interpretations of space-time. From an effective-theory point of view, the eigenvalues represent the coordinates of D-instantons in ten-dimensional bulk space-time, which we have assumed a priori from the beginning of constructing the effective action. On the other hand, from a constructive point of view, we cannot assume such a bulk space-time, in which matrices live. This is because not only fields, but also the space-time, should be dynamically generated as a result of the dynamics of the matrices. The space-time should be constructed only from the matrices. The most natural interpretation is that space-time consists of N discretized points, and that the eigenvalues represent their space-time coordinates. Here we need to assume that the dynamics of the IIB matrix model is such that the resulting eigenvalue distributions are smooth enough to be interpreted as Riemannian geometry.

Another important property is that the type-IIB matrix model has no free parameters. The coupling constant g can always be absorbed by field redefinitions:

$$\begin{cases} A_\mu \to g^{1/2} A_\mu \\ \psi \to g^{3/4} \psi. \end{cases} \tag{7}$$

This is reminiscent of string theory where a shift of the string coupling constant is always absorbed to that of the dilaton vacuum expectation value (vev). In an analysis of the Schwinger-Dyson equation of the IIB matrix model [2], we introduced an infrared cut-off ϵ, which gives a string coupling constant, $g_{st} = 1/N\epsilon^2$. However, through a more careful analysis of the dynamics of the eigenvalues [3], we have shown that there is no such infrared divergences associated with infinitely separated eigenvalues, and that the infrared cutoff ϵ which we have introduced by hand can be determined dynamically in terms of N and g. (The Schwinger-Dyson equation and the double scaling limit are discussed in [7].)

We then explain several reasons we believe that the IIB matrix model is a constructive definition of type-IIB superstring in addition to the symmetry argument. First, this action can be related to the Green-Schwarz action of a superstring by using the semiclassical correspondence in the large N limit:

$$\begin{aligned} -i[\,,\,] &\to \{\,,\,\}, \\ Tr &\to \int d^2\sigma \sqrt{\hat{g}}. \end{aligned} \tag{8}$$

In fact, eq.(1) is related to the Green-Schwarz action in the Schild gauge[8]:

$$S_{\text{Schild}} = \int d^2\sigma [\sqrt{\hat{g}} \alpha (\frac{1}{4} \{X^\mu, X^\nu\}^2 - \frac{i}{2} \bar{\psi} \Gamma^\mu \{X^\mu, \psi\}) + \beta \sqrt{\hat{g}}]. \tag{9}$$

We need to integrate over the scale factor of the metric $\sqrt{g}$ in order to quantize the Schild action

$$Z = \int D\sqrt{g} DX D\psi e^{-S_{\text{Schild}}}. \tag{10}$$

The matrix analog is the following grand canonical ensemble:

$$Z = \sum_{n=0}^{\infty} \int dA d\psi e^{-S(\beta)},$$
$$S(\beta) = \alpha(-\frac{1}{4}Tr[A_\mu, A_\nu]^2 - \frac{1}{2}Tr(\bar{\psi}\Gamma^\mu[A_\mu, \psi])) + \beta Tr 1. \quad (11)$$

If the large N limit is smooth, we expect that the $\beta \to \beta_C$ limit is identical to consider the microcanonical ensemble with fixed N and take N large.

The correspondence can go farther beyond the above identification of the model with a matrix regularization of the first quantized superstring. Namely, we can describe an arbitrary number of interacting D-strings and anti-D-strings as blocks of matrices, each of which corresponds to the matrix regularization of a string. Off-diagonal blocks induce interactions between these strings [1]. Thus, it must be clear that the IIB matrix model is definitely not the first quantized theory of a D-string, but a full second quantized theory.

It has also been shown [2] that Wilson loops satisfy the string field equations of motion for type-IIB superstring in the light-cone gauge, which is a second evidence for the conjecture that the IIB matrix model is a constructive definition of superstring. We consider the following regularized Wilson loop [2]:

$$w(C) = Tr[\prod_{n=1}^{M} exp\{i\epsilon(k_n^\mu A_\mu + \bar{\lambda}_n \psi)\}]. \quad (12)$$

Here, k_n^μ are the momentum densities distributed along a loop C; we have also introduced fermionic sources, λ_n . ϵ is a short-distance cutoff of string world sheet. In the large N limit, ϵ should go to 0 so as to satisfy the double scaling limit. In the paper [2], it was proved, once we have taken the correct scaling limit, that the $\mathcal{N} = 2$ supersymmetry is enough to reproduce the lightcone field equation of type IIB superstring from the IIB matrix model. In order to resolve the problem of the double scaling limit, we need to evaluate several quantities (i.e. an expectation value of the Wilson loop with almost zero total momentum) and there is still some subtly as to how to take this double scaling limit in which we obtain an interacting string theory.

Considered as a matrix regularization of the Green-Schwarz IIB superstring, the IIB matrix model describes interacting D-strings. On the other hand, in an analysis of the Wilson loops, the IIB matrix model describes joining and splitting interactions of fundamental IIB superstrings created by the Wilson loops. From these considerations, it is plausible to conclude that if we can take the correct double scaling limit, the IIB matrix model could become a constructive definition of type-IIB superstring. Furthermore, we believe that all string theories are connected by duality transformations, and once we construct a nonperturbative definition of any one of them, we can describe the vacua of any other strings, particularly the true vacuum in which we live.

The dynamics of eigenvalues, that is, dynamical generation of space-time was first discussed in the paper [3]. An effective action of eigenvalues can be obtained by integrating all of the off-diagonal bosonic and fermionic components, and then the diagonal fermionic coordinates (which we call fermion zeromodes). If we quench the bosonic diagonal components, x_μ^i $(i = 1...N)$, and neglect the fermion zeromodes ξ^i, the effective action for x_μ^i coincides with that of supersymmetric Yang Mills theory with maximal supersymmetry and vanishes respecting the stability of the supersymmetric moduli. The inclusion of fermion zeromodes as well as the non-planar contributions lifts the degeneracy, and we can obtain a nontrivial effective action for the space-time dynamics. In the paper [3] we estimated this effective action by perturbation at one loop, which is valid when all eigenvalues are far from one another, $|\mathbf{x}_i - \mathbf{x}_j| >> \sqrt{g}$. Of course, this one-loop effective action is not sufficient to determine the full space-time structure, but we expect that it captures some of the essential points concerning the formation of space-time. One of important properties of the effective action is that, as a result of grassmannian integration of the fermion zeromodes, space-time points make a network connected locally by bond interactions. This feature becomes important when we extract diffeomorphism symmetry from our matrix model, which is discussed in section 2.

Once we are convinced that IIB matrix model is a constructive definition of superstring, we then have to give natural interpretation of low energy dynamics. That is, we need to show how we can obtain local

field theory in a low energy approximation and the origin of local gauge symmetry in our space-time generated dynamically from matrices. We also have to show how the background metric is encoded in a low energy field theory in the space-time, especially the origin of diffeomorphism invariance. In the paper [4], we have shown that, if we suppose that the eigenvalue distribution consists of small clusters of size n, the low-energy theory acquires $SU(n)$ local space-time gauge symmetry. This gauge invariance assures the existence of a gauge field propagating in the space-time of distributed eigenvalues. Also, we have obtained a low energy effective action (a gauge-invariant kinetic action) for a fermion in the adjoint representation of $SU(n)$, which becomes massless. The low-energy behavior for these fields is formulated as a lattice gauge theory on a dynamically generated random lattice, and hence supports our interpretation of space-time. We have also shown in [4] that the diffeomorphism invariance of our model originates in invariance under permutations of the eigenvalues. Our model realizes the invariance in an interesting way by summing all possible graphs connecting the space-time points. The diffeomorphism invariance restricts the low-energy behavior of the model, and indicates the existence of a massless graviton in the low energy effective field theory on dynamically generated space-time. The background metric for propagating fields is shown to be encoded in the density correlation of the eigenvalues, while the dilaton vacuum expectation value is encoded in the eigenvalue density. A curved background can be described as a nontrivial distribution of eigenvalues whose density correlation behaves inhomogeneously.

Both of these fundamental symmetries, local gauge symmetry and the diffeomorphism symmetry, originates in the $SU(N)$ invariance of the matrix model. It is quite interesting that these fundamental symmetries can arise from a very simple matrix model defined on a single point. These are discussed in section 3.

In the final section, we will explain recent developments on the relation with matrix models and field theory in noncommutative space. As is well known, field theoretic approach to quantize gravity encounters serious difficulties at short distances. String theory contains the notion of the minimal length (string scale) which is expected to cure such difficulties. Another school of thought which also introduces the minimum length scale advocates to replace Riemannian geometry by noncommutative geometry[9]. Remarkably, these various schools of thought have converged recently. We have pointed out that noncommutative Yang-Mills theory is naturally obtained in IIB matrix model by expanding the theory around a noncommutative space-time[10]. We have further studied the Wilson loops. These investigations have shown that the theory contains not only point like field theoretic degrees of freedom but also open string like extended objects [11]. Then we can expect an existence of higher spin fields even if we start from gauge theory since the above nonlocal fields can be expanded in terms of higher spin local fields. Furthermore it is expected that even diffeomorphism invariance might be hidden in gauge theory in noncommutative space-time. We discuss these points in section 4.

2 Dynamics of Eigenvalues and Space-time Generation

In this section we analyze the structure of space-time, and in particular, try to explain why our space-time is four-dimensional [3]. As we mentioned in the introduction, the diagonal elements of the bosonic matrices A_μ can be interpreted as space-time itself. For example, if the diagonal elements distribute within a manifold which extends in four dimensions but shrinks in six dimensions, then a natural interpretation is that the space-time is four-dimensional. We thus derive an effective theory for the diagonal elements, and analyze their distribution.

We decompose A_μ into diagonal part X_μ and off-diagonal part $\tilde{A}_\mu$. We also decompose ψ into ξ and $\tilde{\psi}$:

$$A_\mu = X_\mu + \tilde{A}_\mu;\ X_\mu = \begin{pmatrix} x_\mu^1 & & & \\ & x_\mu^2 & & \\ & & \ddots & \\ & & & x_\mu^N \end{pmatrix},$$

$$\psi = \xi + \tilde{\psi};\ \xi = \begin{pmatrix} \xi^1 & & & \\ & \xi^2 & & \\ & & \ddots & \\ & & & \xi^N \end{pmatrix}, \tag{1}$$

where x_μ^i and ξ_α^i satisfy the constraints $\sum_{i=1}^N x_\mu^i = 0$ and $\sum_{i=1}^N \xi_\alpha^i = 0$, respectively, since we may fix the $U(1)$ part by translation invariance. We then integrate out the off-diagonal parts $\tilde{A}_\mu$ and $\tilde{\psi}$ and obtain the effective action for supercoordinates of space-time $S_{\rm eff}[X,\xi]$. The effective action for the space-time coordinates $S_{\rm eff}[X]$ can be obtained by further integrating out ξ:

$$\begin{aligned} \int dAd\psi e^{-S[A,\psi]} &= \int dXd\xi e^{-S_{\rm eff}[X,\xi]} \\ &= \int dX e^{-S_{\rm eff}[X]}, \end{aligned} \tag{2}$$

where dX and $d\xi$ stand for $\prod_{i=1}^{N-1}\prod_{\mu=0}^{9} dx_\mu^i$ and $\prod_{i=1}^{N-1}\prod_{\alpha=1}^{16} d\xi_\alpha^i$, respectively.

We perform integrations over off-diagonal parts $\tilde{A}_\mu$ and $\tilde{\psi}$ by the perturbative expansion in g^2, which is valid when all of the diagonal elements are widely separated from one another: $|x^i - x^j| >> g^{1/2}$. After adding a gauge fixing and the Fadeev-Popov ghost term

$$S_{\rm g.f.} + S_{\rm F.P.} = -\frac{1}{2g^2} Tr([X_\mu, A^\mu]^2) - \frac{1}{g^2} Tr([X_\mu, b][A^\mu, c]), \tag{3}$$

the action can be expanded up to the second order of the off-diagonal elements $\tilde{A}_\mu, \tilde{\psi}$ as

$$\begin{aligned} S_2 + S_{\rm g.f.} &= \frac{1}{2g^2}\sum_{i\neq j}\Big((x_\nu^i - x_\nu^j)^2 \tilde{A}_\mu^{ij\,*}\tilde{A}^{ij\mu} - \bar{\tilde{\psi}}^{ji}\Gamma^\mu(x_\mu^i - x_\mu^j)\tilde{\psi}^{ij} \\ &\quad + (\bar{\xi}^i - \bar{\xi}^j)\Gamma^\mu\tilde{\psi}^{ij}\tilde{A}_\mu^{ij\,*} + \bar{\tilde{\psi}}^{ji}\Gamma^\mu(\xi^i - \xi^j)\tilde{A}_\mu^{ij}\Big). \end{aligned} \tag{4}$$

The first and the second terms are the kinetic terms for $\tilde{A}$ and $\tilde{\psi}$ respectively, while the last two terms are $\tilde{A}\tilde{\psi}\xi$ vertices. A bosonic off-diagonal element $\tilde{A}_\mu^{ij}$ is transmuted to a fermionic off-diagonal element $\tilde{\psi}^{ij}$ emitting a fermion zeromode ξ^i or ξ^j. This vertex conserves indices for space-time points, i, j. Note that the propagators for $\tilde{A}_\mu$ and $\tilde{\psi}$ behave as $\eta_{\mu\nu}/(x^i - x^j)^2$ and $(x_\mu^i - x_\mu^j)\Gamma^\mu/(x^i - x^j)^2$ respectively, thus they decrease in the long distances.

We obtain the effective action for the zeromodes, x_μ^i and ξ^i, at one-loop level:

$$\begin{aligned} \int d\tilde{A}d\tilde{\psi}dbdc\, e^{-(S_2 + S_{\rm g.f.} + S_{\rm F.P.})} &= \prod_{i<j} \det{}_{\mu\nu}\big(\eta^{\mu\nu} + S_{(ij)}^{\mu\nu}\big)^{-1} \\ &\equiv e^{-S_{\rm eff}^{1-\rm loop}[X,\xi]}, \end{aligned} \tag{5}$$

where

$$S_{(ij)}^{\mu\nu} = \bar{\xi}^{ij}\Gamma^{\mu\alpha\nu}\xi^{ij}\frac{x_\alpha^{ij}}{(x^{ij})^4}. \tag{6}$$

Here ξ^{ij} and x_μ^{ij} are abbreviations for $\xi^i - \xi^j$ and $x_\mu^i - x_\mu^j$. The effective action can be expanded as

$$\begin{aligned} S_{\rm eff}^{1-\rm loop}[X,\xi] &= \sum_{i<j} tr\ln(\eta^{\mu\nu} + S_{(ij)}^{\mu\nu}) \\ &= -\sum_{i<j} tr\Big(\frac{S_{(ij)}^4}{4} + \frac{S_{(ij)}^8}{8}\Big), \end{aligned} \tag{7}$$

which is a sum of all pairs (ij) of space-time points. Here the symbol tr in the lower case stands for the trace over Lorentz indices, μ, ν. Other terms in the expansion vanish due to the properties of Majorana-Weyl fermions in ten dimensions.

Note that the one-loop effective action $S_{\rm eff}^{\rm 1-loop}[X,\xi]$ has $\mathcal{N}=2$ supersymmetry,

$$\begin{cases} \delta^{(1)}x_\mu^i &= i\bar{\epsilon}_1\Gamma_\mu\xi^i \\ \delta^{(1)}\xi^i &= 0 \end{cases}, \quad \begin{cases} \delta^{(2)}x_\mu^i &= 0 \\ \delta^{(2)}\xi^i &= \epsilon_2 \end{cases}, \tag{8}$$

which is a remnant of the one in the original theory,

$$\begin{cases} \delta^{(1)}A_\mu &= i\bar{\epsilon}_1\Gamma_\mu\psi \\ \delta^{(1)}\psi &= \frac{i}{2}\Gamma^{\mu\nu}[A_\mu,A_\nu]\epsilon_1 \end{cases}, \quad \begin{cases} \delta^{(2)}A_\mu &= 0 \\ \delta^{(2)}\psi &= \epsilon_2 \end{cases}. \tag{9}$$

Transformations for $\epsilon_1=\epsilon_2$ and $\epsilon_1=-\epsilon_2$ correspond to those generated by $\mathcal{N}=1$ supersymmetry generator Q and its covariant derivative D. In this sense zeromodes of x_i and ξ_i may be viewed as supercoordinates of $\mathcal{N}=1$ superspace.

The effective action for the space-time coordinates x_μ^i is given by further integrating out the fermion zeromodes ξ_i:

$$\begin{aligned} \int dX e^{-S_{\rm eff}^{\rm 1-loop}[X]} &= \int dX d\xi e^{-S_{\rm eff}^{\rm 1-loop}[X,\xi]} \\ &= \int dX d\xi \prod_{i<j}[1+\frac{tr(S_{(ij)}^4)}{4}+(\frac{1}{2}(\frac{tr(S_{(ij)}^4)}{4})^2+\frac{tr(S_{(ij)}^8)}{8})]. \end{aligned} \tag{10}$$

Here the products are taken over all possible different pairs (ij) of the space-time points. When we expand the multi-products, we select one of the three different factors, 1, $tr(S_{(ij)}^4)/4$ or $(tr(S_{(ij)}^8)/8+(tr(S_{(ij)}^4))^2/32)$ for each pair of (ij). Since the last two factors are functions of $(x_\mu^i-x_\mu^j)$, they can be visualized by bonds that connect the space-time points x_μ^i and x_μ^j. Since the factors $tr(S_{(ij)}^4)/4$ and $(tr(S_{(ij)}^8)/8+(tr(S_{(ij)}^4))^2/32)$ contain 8 and 16 spinor components of ξ^{ij}, we call them an 8-fold bond and a 16-fold bond, respectively. We don't assign any bond to the factor 1. In this way we can associate each term in the expansion of multi-products in eq. (10) with a graph connecting the space-time points by 8-fold bonds and 16-fold bonds. Therefore the multi-products in eq. (10) can be replaced by a summation over all possible graphs:

$$\begin{aligned} \int dX e^{-S_{\rm eff}^{\rm 1-loop}[X]} &= \int dX d\xi \sum_{G:\text{graph}} \prod_{(ij):\text{bond of } G} \\ &\quad [(\frac{tr(S_{(ij)}^4)}{4}) \text{ or } (\frac{1}{2}(\frac{tr(S_{(ij)}^4)}{4})^2+\frac{tr(S_{(ij)}^8)}{8})]. \end{aligned} \tag{11}$$

Here we sum over all possible graphs consisting of 8-fold and 16-fold bonds. For each bond (ij) of G, we assign the first or the second factor depending on whether it is an 8-fold or a 16-fold bond.

In order to saturate the grassmann integration $d\xi$, we need $16(N-1)$ fermion zeromodes. Since an 8-fold bond and a 16-fold bond contain 8 and 16 fermion zeromodes respectively, graphs remaining after the ξ integrations are those where the sum of the number of 8-fold bonds and twice the number of 16-fold bonds is equal to $2(N-1)$. Thus, the number of bonds is of order N, which is much smaller than possible number of pairs $N(N-1)/2$. In this sense N space-time points are weakly bound.

Since 8-fold bond and 16-fold bond terms behave as $(x^{ij})^{-12}$ and $(x^{ij})^{-24}$, both are strong attractive interactions. Hence, only the closer points can be connected by the bonds, and these interactions become local. On the other hand, since all possible graphs must be summed up, this system has a permutaion invariance among the points. This reminds us of summation over all triangulations in the dynamical triangulation approach to quantum gravity. We come back to this analogy in section 3.

In order to see some important features of the system (11), let's take an approximation. If there were only 16-fold bonds, considerable simplifications take place and the ξ integrations can be performed exactly. Since a 16-fold bond term contains 16 fermion zeromodes, it is proportional to delta function of the grassmann variables as

$$tr(S_{(ij)}^8) \sim \delta^{(16)}(\xi^i-\xi^j)\frac{1}{(x^i-x^j)^{24}}. \tag{12}$$

If there is a loop in the graph, the contribution vanishes, since the product of delta functions of grassmann variables on the loop vanishes:

$$\delta^{(16)}(\xi^{i_1 i_2})\delta^{(16)}(\xi^{i_2 i_3})\cdots\delta^{(16)}(\xi^{i_k i_1}) = 0. \tag{13}$$

Also, as we mentioned above, the remaining graphs have $(N-1)$ number of 16-fold bonds. Hence, the remaining graphs are tree graphs which connect all N points. Such type of graphs are called "maximal trees". We also note that all maximal trees contribute equally as we can see by performing ξ integrations from the end points of each maximal tree. Therefore, only with 16-fold bonds, the distribution of x^i_μ becomes "branched polymer" type:

$$\int dX e^{-S_{\rm eff}^{\rm 16-fold\ bond}[X]} = \int dX \sum_{G:\text{maximal graph}} \prod_{(ij):\text{bond of } G} \frac{1}{(x^i - x^j)^{24}} \tag{14}$$

Note that all points are connected by the bonds, and each x^{ij} integration can be performed independently and converges for large x^{ij} on each bond. Thus this system is infrared convergent. As we proved rigorously in [3], this feature of IR convergence holds even with 8-fold bonds, and also, to all orders in perturbation expansion. This is consistent with the explicit calculations of the partition function. This shows that all points are gathered as a single bunch and hence space-time is inseparable. Thus, the size and the dimensionality of the space-time can be determined dynamically. Note also that the dynamics of branched polymer is well known and its Hausdorff dimension is four. It is conceivable that the smooth four dimensional space-time emerges by taking account of the effects from 8-fold bonds as we argue in what follows. Therefore, the model (11) constitutes a candidate of models for dynamical generation of four dimensional space-time. Numerical simulation have been done to investigate this possibility but so far we have not reached the final answer.

3 Symmetries in the low energy theory

3.1 Local Gauge Invariance

Once we describe the space-time as a dynamically generated distribution of the eigenvalues, low-energy effective theory in the space-time can be obtained by solving the dynamics of the fluctuations around the background X_μ. Both of the space-time X_μ and matter $\tilde{A}_\mu$ are unified in the same matrices A_μ and should be determined dynamically. Low-energy fluctuations are generally composites of A_μ and ψ, and it is natural from the analysis of the Schwinger-Dyson equation for the Wilson loops that a local operator in space-time is given by a microscopic limit of the Wilson loop operators, such as

$$w(k;O) = Tr[O(A,\psi)\ exp(ik^\mu A_\mu)]. \tag{1}$$

Here, $O(A,\psi)$ is some operator made of A_μ and ψ. In order to identify the total momentum of this operator with k, the operator $O(A,\psi)$ should be invariant under a constant shift of A_μ, that is, a translation in the space-time coordinates.

In the first approximation around the diagonal background X_μ, the coordinate representation of this operator is given by

$$\begin{aligned}\hat{w}(x;O) &= \int \frac{d^{10}k}{(2\pi)^{10}}\ exp(-ik^\mu x_\mu)\ w(k;O) \\ &\sim \sum_{i=1}^{N} O_{ii}\delta^{(10)}(\mathbf{x}-\mathbf{x}_i).\end{aligned} \tag{2}$$

Here, we have replaced A_μ by $X_\mu + \tilde{A}_\mu$ and taken the leading term. O_{ii} is the (ii) component of the operator O. Due to the delta function, the operator has support only in the domain where the eigenvalues distribute. Vanishing of the operator $\hat{w}(x;O)$ outside of the domain of the distributed eigenvalues implies that space-time simply does not exist outside the domain. This fact supports our interpretation of the space-time in IIB matrix model.

We can apply a similar analysis to strings which propagate in the space-time. In the $1/N$ expansion, the correlation between Wilson loop operators can be evaluated by summing over all surfaces made of Feynman diagrams connecting the Wilson loops at the boundary. This surface is interpreted as a string world sheet connecting strings at the boundaries. Each eigenvalue (x_μ^i) associated to a loop in the diagrams represents a coordinate on the world sheet, and it takes a value in the eigenvalue distribution in the leading approximation around the diagonal background X_μ. Hence, a string world sheet evolves only in the space-time of the eigenvalue distribution, which again supports our interpretation of space-time.

It is generally difficult to obtain how fluctuations propagate in the eigenvalue distribution, which is reminiscent of the QCD effective theory: In QCD excitations are expressed as composite operators of microscopic variables, and their low-energy dynamics can be discussed only through a symmetry argument, namely an argument based on chiral symmetry. Also, in our case, we will show that there are eigenvalue distributions around which symmetry arguments allow us to discuss the low-energy dynamics for some excitations. Suppose that the eigenvalue distribution forms clusters consisting of n eigenvalues. At a length scale much larger than the size of each cluster, the $SU(N)$ symmetry is broken down to $SU(n)^m$, where $m = N/n$.

We can expand A_μ and ψ around such a background X_μ similarly to the analysis in the previous section. First, write A_μ and ψ in block forms:

$$A_\mu = \begin{pmatrix} A_\mu^{11} & A_\mu^{12} & \cdots & \\ A_\mu^{21} & A_\mu^{22} & & \\ \vdots & & \ddots & \\ & & & A_\mu^{mm} \end{pmatrix},$$

$$\psi = \begin{pmatrix} \psi^{11} & \psi^{12} & \cdots & \\ \psi^{21} & \psi^{22} & & \\ \vdots & & \ddots & \\ & & & \psi^{mm} \end{pmatrix}. \tag{3}$$

Each block, A_μ^{ij} or ψ^{ij}, is an $n \times n$ matrix, and the diagonal blocks can be further decomposed:

$$\begin{aligned} A_\mu^{ii} &= x_\mu^i 1 + \tilde{A}_\mu^{ii}, \\ \psi^{ii} &= \xi^i 1 + \tilde{\psi}^{ii}, \end{aligned} \tag{4}$$

where 1 is an $n \times n$ unit matrix and $tr\tilde{A}_\mu^{ii} = 0$. Here, tr means the trace for a submatrix of $n \times n$. We interpret each cluster of the eigenvalues as being a space-time point with an internal structure $SU(n)$. Since each $SU(n)$ symmetry acts on the variables at position i independently, the unbroken $SU(n)^m$ symmetry can be regarded as being local gauge symmetry. Indeed, under a gauge transformation of the unbroken $SU(n)^m$ symmetry,

$$g = \begin{pmatrix} g_1 & & & \\ & g_2 & & \\ & & \ddots & \\ & & & g_m \end{pmatrix} \in SU(n)^m \subset SU(N), \tag{5}$$

the diagonal block fields, $\tilde{A}_\mu^{ii}$ and $\tilde{\psi}^{ii}$, transform as adjoint matters (i.e. site variables in the lattice gauge theory), while the off-diagonal block fields, A_μ^{ij} and ψ^{ij}, transform as gauge connections (i.e. link variables):

$$\begin{aligned} \tilde{A}_\mu^{ii} &\to g_i \tilde{A}_\mu^{ii} g_i^\dagger, \\ \tilde{\psi}^{ii} &\to g_i \tilde{\psi}^{ii} g_i^\dagger, \\ A_\mu^{ij} &\to g_i A_\mu^{ij} g_j^\dagger, \\ \psi^{ij} &\to g_i \psi^{ij} g_j^\dagger. \end{aligned} \tag{6}$$

Some of the dynamics for low-energy excitations is governed by this local gauge invariance. Gauge fields live on the links and transform as the link variables in lattice gauge theory. In our case, we have too many such fields (at least 10 boson fields A_μ^{ij} for a link (ij)), but only one unitary link variable is assured to be massless by the gauge symmetry, and the others acquire mass dynamically. Therefore, in deriving low-energy effective theory, we first apply a polar decomposition to A_μ^{ij} into unitary and hermitian degrees of freedom, and identify all of the unitary components of A_μ^{ij} by setting them to be one common field U^{ij} on each link. We have to integrate those massive off-diagonal block fields A_μ^{ij} while keeping the unitary components U^{ij} and this procedure is performed by the following replacements:

$$\begin{aligned} \overline{A_\mu^{ij}} &= 0, \\ \overline{A_\mu^{ij} \otimes A_\nu^{ji}} &= g^2 \frac{\delta_{\mu\nu}}{(x^{ij})^2} U^{ij} \otimes U^{ji}. \end{aligned} \tag{7}$$

In the second equation, the factor, $1/(x^{ij})^2$, corresponds to the propagation of the hermitian degrees of freedom while the appearance of the link variable U^{ij} corresponds to keeping the unitary degrees of freedom. Higher order correlations can be obtained by using Wick theorem in general except ten-body correlation function for A_μ^{ij}. Due to the chiralness of the ten-dimensional fermion, we obtain an extra term proportional to $\epsilon_{\mu_1 \ldots \mu_{10}}$, coming from fermion one-loop integral.

In order to derive the effective theory for fluctuations around the assumed background, we integrate massive fields first and obtain effective action for other fields as we have obtained the effective action for the diagonal components (i.e. space-time coordinates) in the previous section. Generally speaking, we can expect any terms which are not forbidden by symmetries; supersymmetry and local $SU(n)$ gauge symmetry.

A plaquette action for gauge fields U^{ij} can be generated as follows. The relevant terms in the action are

$$S_4 = \frac{1}{g^2} \sum_{i \neq j \neq k \neq l} tr(A_\mu^{ij} A_\nu^{jk} - A_\nu^{ij} A_\mu^{jk}) A_\mu^{kl} A_\nu^{li}. \tag{8}$$

By integrating out the hermitian degrees of freedom of the off-diagonal blocks with the procedure (7), this action vanishes; $\bar{S}_4 = 0$. However, interactions generated by $(S_4)^2$ induce a kinetic term for the gauge field:

$$\overline{(S_4)^2} \sim \sum_{i \neq j \neq k \neq l} \frac{g^4}{(x^{ij})^2 (x^{jk})^2 (x^{kl})^2 (x^{li})^2} tr(U^{ij} U^{jk} U^{kl} U^{li}) tr(U^{il} U^{lk} U^{kj} U^{ji}). \tag{9}$$

This is the plaquette action generated by a Wilson loop for the adjoint representation, and hence the gauge field U^{ij} indeed propagates in the space-time of the eigenvalue distribution. The gauge field can hop between any pair of space-time points, but the hopping is suppressed by $1/x^8$ for distant points and we will recover locality in the continuum limit. Similarly we can obtain a gauge invariant hopping term for adjoint fermion ψ^{ij}.

To summarize this subsection, supposing that the distribution of the eigenvalues consists of small clusters of size n, we have shown that the low-energy effective theory contains several massless fields, such as the gauge field associated with the local $SU(n)$ gauge symmetry and fermion field in the adjoint representation of $SU(n)$ gauge symmetry. Gauge-invariant kinetic terms were also derived. Our findings are reminiscent of those which are obtained by considering n coincident D9 branes. Presumably our argument here is related to the standard argument of coincident D branes.

Our system is a lattice gauge theory on a dynamically generated random lattice. It is invariant under a permutation for the set of the m discrete space-time points, since the permutation group S_m is a subgroup of the original $SU(N)$ symmetry. The existence of the permutation symmetry is the crucial difference from the ordinary lattice gauge theory on a fixed lattice, which becomes important in deriving the diffeomorphism invariance of our model. We will come back to this point in the next subsection. Although the permutation invariance requires that all space-time points are equivalent, locality in the space-time will be assured due to suppression of the hopping terms between distant points. In general, however, we need a sufficient power for the damping of the hopping terms in order to assure locality in the continuum limit. Though we do not yet know the real condition for locality, we expect that terms with lower powers are canceled due to supersymmetry or by averaging over gauge fields.

3.2 Diffeomorphism Invariance

As shown in section 2, the one-loop effective action for the space-time points is described as a statistical system of N points whose coordinates are x^i_μ. Integration over the fermion zeromodes ξ gives the Boltzmann weight, which depends on a graph (or network) connecting the space-time points locally by the bond interactions:

$$Z = \sum_{G:\text{graph}} \int dX \; W[X;G]. \tag{10}$$

$W[X;G]$ is a complicated function of a configuration X and a graph G. An important property is that the weight is suppressed at least by a damping factor of $1/(x^i - x^j)^{12}$, when two points, i and j, are connected. This system is, of course, invariant under permutations S_N [3] of N space-time points, which is a subgroup of the original symmetry $SU(N)$, while the Boltzmann weight for each graph G is not. The invariance is realized by summing over all possible graphs. In other words, the system becomes permutation invariant by rearrangements of the bonds in the network of the space-time points. This reminds us of the dynamical triangulation approach to quantum gravity, where diffeomorphism invariance is believed to arise from summing all possible triangulations. Our system satisfies both the locality and permutation invariance simultaneously by summing over all possible graphs whose points are connected through the local interactions.

Now we see that the permutation invariance of our system actually leads to diffeomorphism invariance. To see how the background metric is encoded in the effective action for low-energy excitations, let us consider, as an example, a scalar field ϕ^i propagating in distributed eigenvalues. The effective action will be given by

$$S = \sum_{i,j} \frac{(\phi^i - \phi^j)^2}{2} f(x^i - x^j) + \sum_i m(\phi^i)^2, \tag{11}$$

where $f(x)$ is a function decreasing sufficiently fast at infinity to assure locality in the space-time. Introducing the density function of the eigenvalues,

$$\rho(x) = \sum_i \delta^{(10)}(x - x^i), \tag{12}$$

and a field $\phi(x)$ which satisfies $\phi(x^i) = \phi^i$, the action can be rewritten as

$$S = \int dxdy\langle\rho(x)\rho(y)\rangle \frac{(\phi(x) - \phi(y))^2}{2} f(x-y) + m\int dx\langle\rho(x)\rangle\phi(x)^2. \tag{13}$$

Here, the expectation, $\langle ... \rangle$, for the density and the density correlation means that we have taken average over configurations X and networks G of the space-time points.

Normalizing the density correlation in terms of the density,

$$\langle\rho(x)\rho(y)\rangle = \langle\rho(x)\rangle\langle\rho(y)\rangle r(x,y), \tag{14}$$

and expanding $\phi(x) - \phi(y) = (x-y)_\mu \partial^\mu \phi(x) + \cdots$, the action becomes

$$\begin{aligned} S &= \frac{1}{2}\int dx\langle\rho(x)\rangle \left[\int dy\langle\rho(y)\rangle (x-y)_\mu (x-y)_\nu f(x-y) r(x,y)\right] \\ &\quad \partial^\mu\phi(x)\partial^\nu\phi(x) + m\int dx\langle\rho(x)\rangle\phi(x)^2 \cdots . \end{aligned} \tag{15}$$

This expansion shows that the field $\phi(x)$ propagating in the eigenvalue distribution feels the density correlation as the background metric, while the density itself as vacuum expectation value of the dilaton field. Namely, we can identify

$$g_{\mu\nu}(x) \sim \int dy\langle\rho(y)\rangle (x-y)_\mu (x-y)_\nu f(x-y) r(x,y) \tag{16}$$

$$\sqrt{g}e^{-\Phi(x)} \sim \langle\rho(x)\rangle. \tag{17}$$

[3] In this section we consider general eigenvalue distributions in which all eigenvalues have nondegenerate space-time coordinates. If we take the cluster type distribution considered in the previous subsection, the permutation symmetry responsible for the diffeomorphism invariance should be $S_{N/n}$.

If the density correlation respects the original translational and rotational symmetry, that is, if they are not spontaneously broken, the metric becomes flat, $g_{\mu\nu} \sim \eta_{\mu\nu}$. (Normalization can be absorbed by the dilaton vev.) The fact that the background metric is encoded in the density correlations indicates that our system is general covariant, even though the IIB matrix model action (1) defined in flat ten dimensions does not have a manifest general covariance.

Then, let us see how the diffeomorphism invariance is realized in our model. The action (1) is invariant under the permutation S_N of the eigenvalues, which is a subgroup of $SU(N)$. Under a permutation,

$$x^i \to x^{\sigma(i)} \quad \text{for} \quad \sigma \in S_N, \tag{18}$$

the field ϕ^i transforms into $\phi^{\sigma(i)}$. Then, from the definition of the field $\phi(x)$, we should extend the transformation (18) into x,

$$x \to \xi(x), \tag{19}$$

such that $\xi(x^i) = x^{\sigma(i)}$. Under this transformation, the eigenvalue density transforms as a scalar density and the field $\phi(x)$ as a scalar field. On the other hand, the metric transforms as a second-rank tensor, if the function $f(x-y)$ decreases rapidly around $x = y$ and the y integral in (16) has support only near $y = x$. The tensor property of the metric is also required from the invariance of the action under transformation (19). In this way, the invariance under a permutation of the eigenvalues leads to the invariance of the low-energy effective action under general coordinate transformations.

The background metric is encoded in the density correlation of the eigenvalues. Since we have started from the IIB matrix model action (1) which is Poincare-invariant, the density correlation is expected to be translational and rotational invariant, and we may obtain a low-energy effective action in a flat background. A nontrivial background can be induced dynamically if the Lorentz symmetry is spontaneously broken and the eigenvalues are non-trivially distributed.

A nontrivial background can be also described by condensing a graviton operator. Bosonic parts of graviton and dilaton operators are given by

$$S_{\mu\nu}(k) \sim Tr(F_{\mu\lambda}F^{\lambda}{}_{\nu}e^{ik\cdot A}) + (\mu \leftrightarrow \nu), \tag{20}$$

$$D(k) \sim Tr(F^2 e^{ik\cdot A}). \tag{21}$$

Their condensation induces extra terms in the IIB matrix model action,

$$S_{\text{cond}} = \int dk (\sum h^{\mu\nu}(k) S_{\mu\nu}(k) + h(k) D(k)). \tag{22}$$

We can similarly obtain an effective action for fluctuations around a diagonal background from this modified matrix model action. Condensation of dilaton changes the Yang-Mills coupling constant g locally in space-time. Since g is the only dimensionful constant in our model, and thus determines the fundamental length scale, a local change in g leads to a local change in the eigenvalue density. This is consistent with our earlier discussion that the dilaton expectation value is encoded in the eigenvalue density. On the other hand, the condensation of graviton induces an asymmetry of space-time. For a condensation of the $k = 0$ graviton mode, it is obvious that the condensation can be compensated by a field redefinition of matrices A_μ,

$$A_\mu \to (\delta_\mu^\nu + h_\mu^\nu) A_\nu, \tag{23}$$

and the two models, the original IIB matrix model and the modified one with the $k = 0$ graviton condensation, are directly related through the above field redefinition. The density of the eigenvalues is mapped accordingly, and the density correlation is expected to become asymmetric in the modified matrix model. For a more general condensation, if the graviton operator $\hat{S}_{\mu\nu}(x)$ (coordinate representation of (20)) changes only the local property of the dynamics of the eigenvalues, the density correlation will become asymmetric locally in space-time around x, and therefore induces a local change in the background metric.

Our low-energy effective action is formulated as a lattice gauge theory on a dynamically generated random lattice. Since the lattice itself is generated dynamically from matrices, we must sum over all possible graphs. In this way, our system is permutation S_N invariant, which is responsible for the diffeomorphism invariance. The background metric is encoded in the density correlation of the eigenvalues, and the

low-energy effective action becomes manifestly general covariant. The graviton operator is represented as fluctuation around the background space-time, and is constructed from the off-diagonal components of the matrices. A microscopic derivation of the propagation of the graviton is difficult to obtain, but once we have clarified the underlying diffeomorphism symmetry, it is natural that the low-energy effective action for the graviton is described by the Einstein Hilbert action. By employing this diffeomorphism invariance and the supersymmetry, we will be able to derive the low-energy behavior of the graviton multiplets, which will be reported in a separate paper.

4 Diffeomorphism in noncommutative field theories?

In this section, we first review how noncommutative Yang Mills is obtained from matrix models and then investigate special properties in the local gauge transformations in noncommutative background. Especially we study special types of gauge transformations which can be interpreted as local coordinate transformations.

First we give a brief review on matrix model description of noncommutative field theories. The noncommutative background $\hat{x}$ satisfying

$$[\hat{x}^\mu, \hat{x}^\nu] = -iC^{\mu,\nu}. \tag{24}$$

is a classical solution of the type IIB matrix model. $C^{\mu\nu}$ are c-numbers. We assume the rank of $B_{\mu\nu}$ to be $\tilde{d}$ and define its inverse $B_{\mu\nu}$ in $\tilde{d}$ dimensional subspace. $\hat{x}^\mu$ satisfy the canonical commutation relations and they span the $\tilde{d}$ dimensional phase space. Therefore momentum operators are proportional to the coordinates as

$$p_\mu = B_{\mu\nu}x^\nu. \tag{25}$$

The semiclassical correspondence shows that the volume of the phase space (measured in the coordinate space of x^μ) is $V = N(2\pi)^{\tilde{d}/2}\sqrt{detC}$.

We therefore expand the bosonic matrices A^μ around $\hat{x}^\mu = C^{\mu\nu}p_\nu$

$$A^\mu = C^{\mu\nu}(p_\nu + \tilde{a}_\nu). \tag{26}$$

If we can assume that all fields can be expanded in terms of noncommutative plane wave $exp(ik\cdot\hat{x})$ (which is, of course, correct for finite N), we obtain a map from a matrix

$$\hat{a} = \sum_k \tilde{a}(k)exp(ik\cdot\hat{x}) \tag{27}$$

to a function

$$a(x) = \sum_k \tilde{a}(k)exp(ik\cdot x) \tag{28}$$

on the $\tilde{d}$-dimensional noncummutative plane. By this construction, a product of matrices is mapped to the $\star$ product of functions

$$\begin{aligned} \hat{a}\hat{b} &\rightarrow a(x)\star b(x), \\ a(x)\star b(x) &\equiv exp(\frac{C^{\mu\nu}}{2i}\frac{\partial^2}{\partial\xi^\mu\partial\eta^\nu})a(x+\xi)b(x+\eta)|_{\xi=\eta=0} \end{aligned} \tag{29}$$

and the operation Tr over matrices can be exactly mapped onto the integration over functions as

$$Tr[\hat{a}] = \sqrt{detB}(\frac{1}{2\pi})^{\frac{\tilde{d}}{2}}\int d^{\tilde{d}}x a(x). \tag{30}$$

The reduced model can be shown to be equivalent to noncommutative Yang-Mills by the the following map from matrices onto functions

$$\begin{aligned} \hat{a} &\rightarrow a(x), \\ \hat{a}\hat{b} &\rightarrow a(x)\star b(x), \\ Tr &\rightarrow \sqrt{detB}(\frac{1}{2\pi})^{\frac{\tilde{d}}{2}}\int d^{\tilde{d}}x. \end{aligned} \tag{31}$$

Applying the rule eq.(31), we can obtain $U(1)$ gauge theory on $\tilde{d}$-dimensional noncummutative space. The bosonic action becomes

$$\begin{aligned} & -\frac{1}{4g^2}Tr[A_\mu, A_\nu][A_\mu, A_\nu] \\ = & \frac{\tilde{d}nB^2}{4g^2} - \sqrt{detB}(\frac{1}{2\pi})^{\frac{\tilde{d}}{2}}\int d^{\tilde{d}}x\frac{1}{g^2}(\frac{1}{4}[D_\alpha, D_\beta][D_\alpha, D_\beta] \\ & +\frac{1}{2}[D_\alpha, \varphi_\nu][D_\alpha, \varphi_\nu] + \frac{1}{4}[\varphi_\nu, \varphi_\rho][\varphi_\nu, \varphi_\rho])_\star. \end{aligned} \tag{32}$$

In this expression, the indices α, β run over $\tilde{d}$ dimensional world volume directions and ν, ρ over the transverse directions. We have replaced $a_\nu \to \varphi_\nu$ in the transverse directions. Inside $(\ )_\star$, the products should be understood as $\star$ products and hence commutators do not vanish. The fermionic action becomes

$$\begin{aligned} & \frac{1}{g^2}Tr\bar{\psi}\Gamma_\mu[A_\mu, \psi] \\ = & \sqrt{detB}(\frac{1}{2\pi})^{\frac{\tilde{d}}{2}}\int d^{\tilde{d}}x\frac{1}{g^2}(\bar{\psi}\Gamma_\alpha[D_\alpha, \psi] + \bar{\psi}\Gamma_\nu[\varphi_\nu, \psi])_\star. \end{aligned} \tag{33}$$

We therefore find noncommutative U(1) gauge theory. Noncommutative $U(m)$ gauge theory can be similarly obtained by expanding around $x^\mu \otimes 1_m$.

We may interpret the newly emerged coordinate space as the semiclassical limit of $\hat{x}^\mu = C^{\mu\nu}\hat{p}_\nu$. The space-time translation is realized by the following unitary operator:

$$exp(i\hat{p}\cdot\epsilon)\hat{x}^\mu exp(-i\hat{p}\cdot\epsilon) = \hat{x}^\mu + \epsilon^\mu. \tag{34}$$

It is amusing to note that translation in the noncommutative space is realized by $U(N)$ gauge invariance in matrix models. This realization has been known as Parisi prescription in old reduced models [12] and investigated [13] from the noncommutative point of view to study extended gauge invariant operators (open Wilson lines). The local gauge symmetry of noncommutative Yang-Mills is originated in the invariance under $U(N)$ invariance of the type IIB matrix model

$$A_\mu \to UA_\mu U^\dagger. \tag{35}$$

Indeed, if we expand $U = exp(i\hat{\lambda})$ and parameterize

$$\hat{\lambda} = \sum_k \tilde{\lambda}(k)exp(ik\cdot\hat{x}), \tag{36}$$

we find the fluctuations around the fixed background transform as

$$\hat{a}_\mu \to \hat{a}_\mu + i[\hat{p}_\mu, \hat{\lambda}] - i[\hat{a}_\mu, \hat{\lambda}] \tag{37}$$

and therefore after mapping this transformation onto the gauge transformation in noncommutative Yang-Mills, we have

$$\begin{aligned} & a_\alpha(x) \to a_\alpha(x) + \frac{\partial}{\partial x^\alpha}\lambda(x) - i[a_\alpha(x), \lambda(x)]_\star, \\ & a_i \to a_i - i[a_i(x), \lambda(x)]_\star, \\ & \psi \to \psi - i[\psi(x), \lambda(x)]_\star \end{aligned} \tag{38}$$

where $1 \le \alpha \le \tilde{d}$ and $i > \tilde{d}$. If we take as in (34)

$$\hat{\lambda} = \epsilon^\alpha \hat{p}_\alpha, \tag{39}$$

(38) becomes translation in the noncommutative space up to a constant shift of the gauge field;

$$\begin{aligned} & a_\alpha(x) \to a_\alpha(x) - B_{\alpha\beta}\epsilon^\beta + \epsilon^\beta\partial_\beta a_\alpha(x), \\ & a_i(x) \to a_i(x) + \epsilon^\beta\partial_\beta a_i(x), \\ & \psi(x) \to \psi(x) + \epsilon^\beta\partial_\beta\psi(x). \end{aligned} \tag{40}$$

As the above example shows, local gauge symmetries in noncommutative gauge theories are very different from those in the ordinary gauge theories and even space-time translation is included. Hence all gauge invariant operators should be constructed by integrating over space-time, which is reminiscent of the gravity. We are, therefore, tempted to generalize the above discussion to local transformations. The reason gauge transformations in noncomutative space-time are much larger than those in commutative space is understood as such that gauge transformation parameters λ contain not only ordinary functions but differential operators in the semiclassical limit in the following sense. Functions in noncommutative space are expanded in terms of noncommutative plane waves as in (36). For finite N, the momenta of plane waves take values of

$$k_n = \sqrt{2\pi B} N^{-1/\tilde{d}}\, n \tag{41}$$

where $n = 0, 1, , , N^{2/\tilde{d}}$. This can be seen from the explicit construction of the plane waves in terms of 't Hooft matrices U and V satisfying $UV = exp(i2\pi/N)VU$. The number of independent plane waves is N^2, which is the same as the number of degrees of freedom of a Hermitian matrix. Since the natural cut off in the noncommutative plane (24)is $l_0 = \sqrt{2\pi C} = \sqrt{2\pi/B}$, natural cut off of momenta should be $2\pi/l_0 = \sqrt{2\pi B}$. However some plane waves with momenta (41) exceed this natural bound and they become very nonlocal objects since such high momentum plane waves generate translation in space with $l_0 N^{-1/\tilde{d}} n$. Hence only N out of N^2 plane waves can be interperted as ordinary plane waves in the semiclassical limit and others should be interpreted as differential operators that can generate translation in noncommutative direction in space-time. From matrix model point of view, such nonlocal waves correspond to off-diagonal elements while local ones to diagonal elements.

Now let us consider gauge transformations related to local coordinate transformations. A natural generalization of the global translation (39) will be

$$\lambda = \frac{1}{2}(\hat{p}_\alpha \hat{\epsilon}^\alpha + \hat{\epsilon}^\alpha \hat{p}_\alpha) \tag{42}$$

or, if we want to include both of the gauge transformations and local coordinate transformations in the semiclassical limit, we can expand as

$$\lambda = \hat{\lambda}_0 + \frac{1}{2}(\hat{p}_\alpha \hat{\epsilon}^\alpha + \hat{\epsilon}^\alpha \hat{p}_\alpha). \tag{43}$$

Similarly we expand bosonic matrices as

$$A_\mu = \tilde{a}_\mu + \frac{1}{2}(\hat{p}_\alpha \hat{e}^\alpha_\mu + \hat{e}^\alpha_\mu \hat{p}_\alpha) \tag{44}$$

and assume that the field $\hat{e}^\alpha_\mu$ is close to δ^α_μ;

$$\hat{e}^\alpha_\mu = \delta^\alpha_\mu + h^\alpha_\mu. \tag{45}$$

This will be a natural generalization of the expansion (26). Applying Unitary transformation generated by (42) to the bosonic field expanded as above, we have the following transformation law;

$$\begin{aligned} \delta h^\alpha_\mu(x) &= e^\beta_\mu(x)\partial_\beta \epsilon^\alpha + \epsilon^\beta \partial_\beta h^\alpha_\mu, \\ \delta a(x)_\mu &= \epsilon^\beta \partial_\beta a(x)_\mu. \end{aligned} \tag{46}$$

In addition to the transformations of fields, we need to transform the background

$$\delta \hat{p}_\alpha = B_{\alpha\beta}\hat{\epsilon}^\beta \tag{47}$$

and accordingly

$$[p_\alpha + B_{\alpha\beta}\hat{\epsilon}^\beta, p_{\alpha'} + B_{\alpha'\beta'}\hat{\epsilon}^{\beta'}] = i(B_{\alpha\alpha'} - B_{\beta\alpha'}\partial_\alpha \epsilon^\beta + B_{\alpha\beta}\partial_{\alpha'}\epsilon^\beta). \tag{48}$$

Therefore the shift of the background can be interpreted as the transfomation of $B_{\alpha\alpha'}$.

The transformation (46) indicates that local coordinate transformations can be realized by gauge transformations in noncommutative space. But as it shows the gauge field or fermion field transforms

as a scalar and h_μ^a as a vector field. Or in another word, the $SO(10)$ index has nothing to do with this coordinate transformation as it should be since $SU(N)$ transformations and $SO(10)$ transformations are decoupled from the beginning. The IIB matrix model is not invariant under local $SO(10)$, that is, local Lorentz transformations. Hence we may hope that there is a generalized model with larger symmetries, which becomes IIB matrix model after gauge fixing[14].

References

[1] N. Ishibashi, H. Kawai, Y. Kitazawa and A. Tsuchiya, Nucl. Phys. **B498** (1997) 467, hep-th/9612115.

[2] M. Fukuma, H. Kawai. Y. Kitazawa and A. Tsuchiya, Nucl.Phys.B510 (1998)158, hep-th/9705128.

[3] H. Aoki, S. Iso, H. Kawai, Y. Kitazawa and T. Tada, Prog. Theor. Phys. **99** (1998) 713, hep-th/9802085.

[4] S. Iso and H. Kawai, *Space-Time and Matter in IIB Matrix Model: Gauge symmetry and Diffeomorphism*, hep-th/9903217

[5] T. Banks, W. Fischler, S.H. Shenker and L. Susskind, Phys. Rev. **D55** (1997) 5112, hep-th/9610043.

[6] T. Eguchi and H. Kawai, Phys. Rev. Lett. **48** (1982) 1063.
G. Parisi, Phys. Lett. **112B** (1982) 463.
D. Gross and Y. Kitazawa, Nucl. Phys. **B206** (1982) 440.
G. Bhanot, U. Heller and H. Neuberger, Phys. Lett. **113B** (1982) 47.

[7] H. Aoki, S. Iso, H. Kawai, Y. Kitazawa, T. Tada and T. Tsuchiya, *IIB matrix model* Prog. Theor. Phys. Suppl. 134(1999)47 ;hep-th/9908038

[8] A. Schild, Phys. Rev. **D16** (1977) 1722.

[9] A. Connes, Comm. Math. Phys. 182 (1996) 155; hep-th/9603053.

[10] H. Aoki, N. Ishibashi, S. Iso, H. Kawai, Y. Kitazawa and T. Tada, *Noncommutative Yang-Mills in IIB matrix model*; hep-th/9908141, Nuclear Physics B565 (2000) 176-192.

[11] N. Ishibashi, S. Iso, H. Kawai and Y. Kitazawa, *Wilson loops in noncommutative Yang-Mills*; hep-th/9910004, Nuclear Physics B573 (2000) 573-593.

[12] G. Parisi, Phys. Lett. **112B** (1982) 463.

[13] S. R. Das and S. Rey, Nucl. Phys. **B590**, 453 (2000)
D. J. Gross, A. Hashimoto and N. Itzhaki, hep-th/0008075.

[14] T. Azuma, S. Iso, H. Kawai and Y. Ohwashi, in preparation

A Maximal Slicing Code for Coalescencing Binary Neutron Stars

Ken-ichi Oohara[1] and Takashi Nakamura[2]

[1] *Department of Physics, Niigata University, Ikarashi, Niigata 950-2181*
[2] *Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan*

Abstract

We are developing 3 dimensional simulation codes for coalescing binary neutron stars. Now a stable code using maximal slicing condition is obtained. We present some results on our simulations including gravitational radiation from a coalescing neutron star binary.

1 Introduction

We showed our computer code development on 3D numerical relativity a few years ago at this series of workshop [1]. We used a conformal slicing condition and we were faced with instability that is contained in this slicing. Then we started to construct a new code using maximal slicing code. Disadvantage to maximal slicing is that as extraction of gravitational waves is not straightforward as well as that more CPU hours are required since an elliptic partial differential equation should be solved for the lapse function α. However, we found that the gauge dependent modes in a spatial metric perturbation is so small at the vacuum exterior region that a good estimate of gravitational waves is given easily.

In the followings, we show some results of 3D relativistic simulation of coalescence of binary neutron stars and estimation of gravitational waves.

2 Basic Equations

In order to follow time evolution of space-time and matter, the Einstein equation should be reduced to a system of evolution equations. Then we use (3+1)-formalism of the Einstein equation. In the (3+1)-formalism of the Einstein equation, the line element takes the form

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt), \tag{1}$$

where α, β^i and γ_{ij} are the lapse function, the shift vector and the intrinsic metric of 3-space, respectively. The evolution equations for the metric and the hydrodynamics variables are essentially the same as the code using conformal slicing [1], [2].

The maximal slicing condition demands the trace of the extrinsic curvature K vanish. The time derivative of K is given by

$$\partial_t K - \beta^\ell \partial_\ell K = \alpha \left(\hat{K}_{ij} \hat{K}^{ij} + \frac{1}{3} K^2 + 4\pi (\rho_H + S) \right) - D^m D_m \alpha, \tag{2}$$

where

$$\hat{K}_{ij} = \phi^{-4} K_{ij} - \frac{1}{3} \gamma_{ij} \gamma^{k\ell} K_{k\ell}, \tag{3}$$

$$\phi \equiv (\det(\gamma_{ij}))^{\frac{1}{12}} \tag{4}$$

[1] E-mail:oohara@astro.sc.niigata-u.ac.jp
[2] E-mail:takashi@yukawa.kyoto-u.ac.jp

and D_i denotes the covariant derivative with respect to γ_{ij}. The quantities ρ_H and S_{ij} is the energy density and the stress tensor of the matter, respectively, measured by the normal line observer and $S = \gamma^{ij} S_{ij}$. Then $\partial_t K = 0$ gives the equation of α for the maximal slicing.

$$D^m D_m \alpha = \alpha \left(\hat{K}_{ij} \hat{K}^{ij} + \frac{1}{3} K^2 + 4\pi (\rho_H + S) \right) \tag{5}$$

The shift vector is determined from the pseudo-minimal distortion condition, which demands the time derivative of $\tilde{\gamma}_{ij} \equiv \phi^{-4} \gamma_{ij}$ be the transverse-traceless; we have

$$\nabla^2 \beta^i + \frac{1}{3} \partial_i \left(\partial_t \beta^t \right) = \partial_j \left(2\alpha \hat{K} \right) - \partial_j \left[h_{jt} \partial_i \beta^t + h_{it} \partial_j \beta^t - \frac{2}{3} h_{ij} \partial_t \beta^t + \beta^t \partial_t h_{ij} \right], \tag{6}$$

where ∇^2 is the simple flat-space Laplacian and $h_{ij} = \tilde{\gamma}_{ij} - \delta_{ij}$.

3 Numerical Results

We performed numerical simulations for coalescing binary neutron stars. Calculations are performed on Hitachi SR8000/F1 at KEK and Fujitsu VPP300/16R at NAO. The typical grid size is $187 \times 187 \times 94$ with $\Delta x = \Delta y = \Delta z = 1 M_\odot$ assuming the symmetry with respect to the equatorial plane. It requires the memory of 5GBytes. Numerical results obtained from one of them is shown here. As the initial condition, we put two spherical stars of rest mass $1.5 M_\odot$ and radius $7.7 M_\odot = 11.6$km. The separation between the center of each mass is $30.2 M_\odot = 46.2$km. The initial rotational velocity is given so that the circulation of the system vanishes;

$$\vec{V}_a(\vec{r}) = \vec{\Omega} \times \vec{r} - \vec{\Omega} \times (\vec{r} - \vec{r}_a), \tag{7}$$

where $\vec{\Omega}$ is the orbital angular velocity and $\vec{r}_a$ $(a = 1, 2)$ is the location of the center of each star. Now we set $\Omega = 0.010/M_\odot$, then the angular momentum J becomes $6.7 M_\odot^2$. The total gravitational mass M_g of the system is $2.76 M_\odot$ and $q \equiv J/M_g^2 = 0.89$. It takes approximately 200 CPU hours to follow evolution up to 45000 time steps on SR8000 using one processor node or VPP300 using 15 processor elements.

Figure 1 shows the evolution of the density distribution on the x-y plane as gray scale figures. Our grid is so coarse that the precision in calculation at the central region becomes very low at the last stage. However we can go on with the calculation, while we were faced with numerical instability in case of a conformal slicing [2].

In the conformal slicing, $h_{ij} = \tilde{\gamma}_{ij} - \delta_{ij}$ can be considered as the gravitational waves [3]. In the maximal slicing, we need a gauge-invariant wave extraction[4], although h_{ij} is still transverse-traceless. In order to compared h_{ij} and its time derivatives between the conformal slicing and the maximal slicing, we calculated them with the same initial data containing two $1.0 M_\odot$ stars, and found the difference is within the numerical errors. For example, Fig.2 shows 'wave forms' h_+ and $h_\times$ along z-axis, that is, at $\theta = \phi = 0$ obtained using the conformal slicing code and the maximal slicing code, where h_+ and $h_\times$ are defined by

$$h_+ = \frac{1}{r} \left(h_{\hat{\theta}\hat{\theta}} - h_{\hat{\phi}\hat{\phi}} \right) \tag{8}$$

$$h_\times = \frac{2}{r} \left(h_{\hat{\theta}\hat{\phi}} \right) \tag{9}$$

and

$$\begin{aligned} h_{\hat{\theta}\hat{\theta}} &= \left(h_{xx} \cos^2\phi + h_{yy} \sin^2\phi + 2 h_{xy} \sin\phi \cos\phi \right) \cos^2\theta \\ &\quad + h_{zz} \sin^2\theta - 2 \left(h_{xz} \cos\phi + h_{yz} \sin\phi \right) \sin\theta \cos\theta, \end{aligned} \tag{10}$$

$$h_{\hat{\phi}\hat{\phi}} = h_{xx} \sin^2\phi + h_{yy} \cos^2\phi - 2 h_{xy} \sin\phi \cos\phi, \tag{11}$$

$$\begin{aligned} h_{\hat{\theta}\hat{\phi}} &= (h_{yy} - h_{xx}) \cos\theta \sin\phi \cos\phi + h_{xy} \cos\theta (\cos^2\phi - \sin^2\phi) \\ &\quad + h_{xz} \sin\theta \sin\phi - h_{yz} \sin\theta \cos\phi. \end{aligned} \tag{12}$$

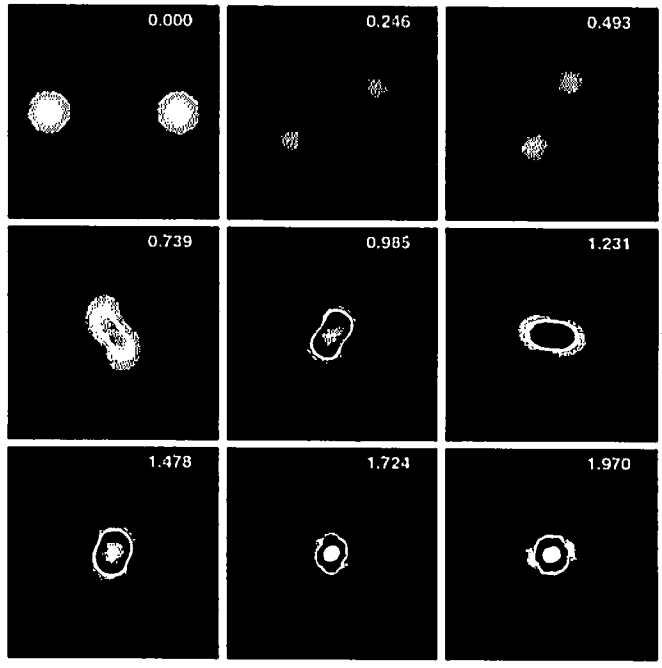


Figure 1: Mass density ρ_N on the x-y plane. The density profiles are shown as gray scale figures.

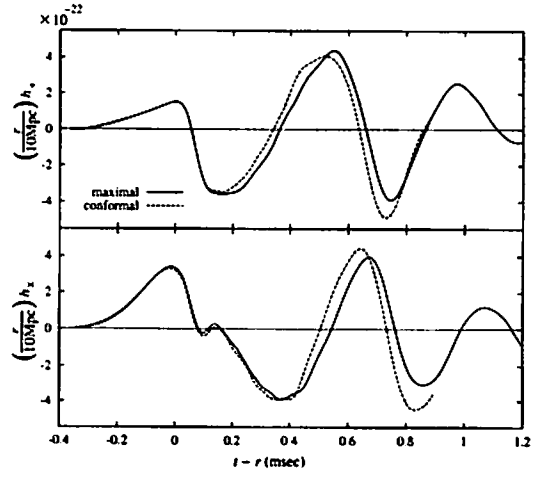


Figure 2: Comparison of 'wave forms' h_+ and $h_\times$ along z-axis between the conformal slicing code and the maximal slicing code. In each calculation, the initial data contains two $1.0M_\odot$ stars.

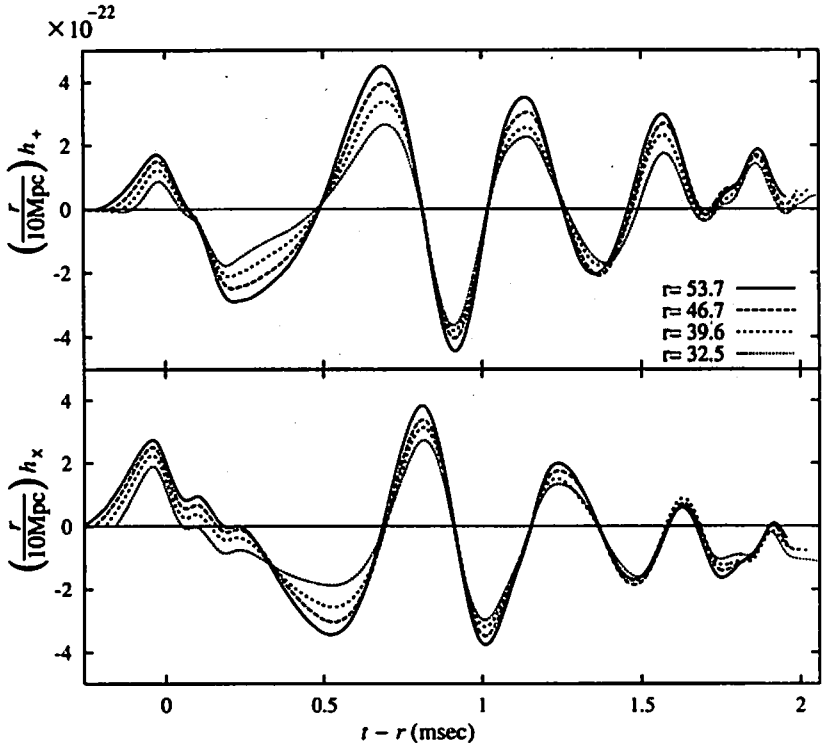


Figure 3: The gravitational waves h_+ and $h_\times$ at $\theta = \pi/4$ and $\phi = 0$ as functions of the retarded time $t - r$.

Since h_+ and $h_\times$ defined by Eq.(9) give two polarizations of gravitational waves in the conformal slicing, they must give a good estimation in gravitational waves even in the maximal slicing.

Figure 3 shows the wave form h_+ and $h_\times$ for $1.5M_\odot$ star case observed at $\theta = \pi/4$ and $\phi = 0$ as functions of the retarded time $t - r$. The wave forms are calculated at various grid point and they practically coincide with each other. It means that the waves certainly travel at the light velocity and h decreases in proportion to r^{-1}.

We have not yet searched the apparent horizon, but the value of the lapse function α at the stellar surface is less than 0.3 and its gradient is very steep at t =1.724 msec, while it is about 0.7 at t =0 as shown in Fig.4. It might suggest the formation of a black hole.

4 Conclusion

Now we constructed the first version of 3D code of numerical relativity including matter using maximal slicing. The calculation presented here used very coarse grid and only 6 or 7 grid points are included from the center through the surface of the central star at the final stage of our calculation. Then we should perform simulation using finer grid to study the final destiny of the coalescing binary neutron stars. But in this simulation, we obtained some information for the gravitational radiation. The maximal amplitude is approximately

$$h \approx 4 \times 10^{-22} \left(\frac{10\mathrm{Mpc}}{r} \right) \tag{13}$$

and the total energy radiated up to 2msec is nearly 0.15% of the initial gravitational mass.

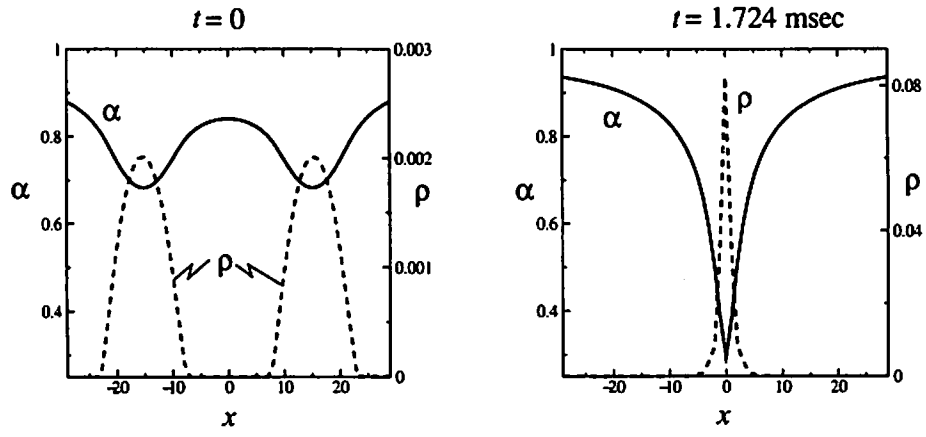


Figure 4: Density ρ and lapse function α along the x-axis at $t = 0$ and t =1.724 msec.

Acknowledgment

Numerical computations were carried out on SR8000/F1 at Hight Energy Accelerator Research Organization(KEK) as the Supercomputer Project No.00-63(FY2000) and on VPP300/16R at the Astronomical Data Analysis Center of the National Astronomical Observatory Japan(NAO). In addition, this work was in part supported by Monbusho Grand-in-Aid (09NP0801, 10640227).

References

[1] K. Oohara and T. Nakamura, *Proceedings of the seventh workshop on General Relativity and Gravitation* ed. by Y. Eriguchi et.al. (1997), 179.

[2] K. Oohara and T. Nakamura, Prog. Theor. Phys. Suppl., No.136 (1999), 270.

[3] M. Shibata and T. Nakamura, Prog. Theor. Phys. **88** (1992), 317.

[4] A. Abrahams at.al, Phys. Rev. **D45**, (1992), 3544.

Will hyperbolic formulations help numerical relativity?
– Experiments using Ashtekar's connection variables –

HISA-AKI SHINKAI[†] and GEN YONEDA[‡]

shinkai@gravity.phys.psu.edu yoneda@mn.waseda.ac.jp

[†] *Centre for Gravitational Physics and Geometry, 104 Davey Lab., Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802-6300, USA*

[‡] *Dept. of Mathematical Science, Waseda University, Shinjuku, Tokyo 169-8555, Japan*

In order to perform accurate and stable long-term numerical integration of the Einstein equations, several hyperbolic systems have been proposed. We here report our numerical comparisons between weakly hyperbolic, strongly hyperbolic, and symmetric hyperbolic systems based on Ashtekar's connection variables. The primary advantage for using this connection formulation is that we can keep using the same dynamical variables for all levels of hyperbolicity. Our numerical code demonstrates gravitational wave propagation in plane symmetric spacetimes, and we compare the accuracy of the simulation by monitoring the violation of the constraints. By comparing with results obtained from the weakly hyperbolic system, we observe the strongly and symmetric hyperbolic system show better numerical performance (yield less constraint violation), but not so much difference between the latter two.

We also study asymptotically constrained systems for numerical integration of the Einstein equations, which are intended to be robust against perturbative errors for the free evolution of the initial data. First, we examine the previously proposed "λ-system", which introduces artificial flows to constraint surfaces based on the symmetric hyperbolic formulation. We show that this system works as expected for the wave propagation problem in the Maxwell system and in general relativity using Ashtekar's connection formulation. Second, we propose a new mechanism to control the stability, which we call the "adjusted system". This is simply obtained by adding constraint terms in the dynamical equations and adjusting its multipliers. We explain why a particular choice of multiplier reduces the numerical errors from non-positive or pure-imaginary eigenvalues of the adjusted constraint propagation equations. This "adjusted system" is also tested in the Maxwell system and in the Ashtekar's system. This mechanism affects more than the system's symmetric hyperbolicity.

1 Introduction

Numerical relativity - solving the Einstein equations numerically - is now an essential field in gravity research. The current mainstream of numerical relativity is to demonstrate the final phase of compact binary objects related to gravitational wave observations [3], and these efforts are now again shedding light on the mathematical structure of the Einstein equations.

Up to a couple of years ago, the standard Arnowitt-Deser-Misner (ADM) decomposition of the Einstein equations was taken as the standard formulation for numerical relativists. Difficulties in accurate/stable long-term evolutions were supposed to be overcome by choosing proper gauge conditions and boundary conditions. Recently, however, several numerical experiments show that the standard ADM is not the best formulation for numerics, and finding a better formulation has become one of the main research topics.

One direction in the community is to apply conformally decoupled and tracefree re-formulation of ADM system which were first used by Nakamura *et al.* [4]. Although there is an effort to show why this re-formulation is better than ADM, we do not yet know this method is robust for all situations.

Another alternative approach to ADM is to formulate the Einstein equations to reveal hyperbolicity [5]. A certain kind of hyperbolicity of the dynamical equations is essential to analyze their propagation features mathematically, and are known to be useful in numerical approximations. Several hyperbolic formulations have been proposed to re-express the Einstein equations, with different levels: weakly, strongly and symmetric hyperbolic systems. Several numerical tests were also performed in this direction.

The following questions, therefore, naturally present themselves (cf. [6]):

(1) Does hyperbolicity actually contribute to the numerical accuracy/stability?

(2) If so, which level of hyperbolic formulation is practically useful for numerical applications? (or does the symmetric hyperbolicity solve all the difficulties?)

(3) Are there any other approaches to improve the accuracy/stability of the system?

In this report, we try to answer these questions with our simple numerical experiments. Such comparisons are appropriate when the fundamental equations are cast in the same interface, and that is possible using Ashtekar's connection variables [7, 14]. More precisely, the authors' recent studies showed the following:

(a) the original set of dynamical equations proposed by Ashtekar already forms a weakly hyperbolic system [9],

(b) by requiring additional gauge conditions *or* adding constraints to the dynamical equations, we can obtain a strongly hyperbolic system [9],

(c) by requiring additional gauge conditions *and* adding constraints to the dynamical equations, we can obtain a symmetric hyperbolic system [9, 10], and finally

(d) based on the above symmetric hyperbolic system, we can construct a set of dynamical systems which is robust against perturbative errors for constraints and reality conditions [11] (*aka.* λ-system [12]).

Based on the above results (a)-(c), we developed a numerical code which handles gravitational wave propagation in the plane symmetric spacetime. We performed the time evolutions using the above three levels of Ashtekar's dynamical equations together with the standard ADM equation. We compare these for accuracy and stability by monitoring the violation of the constraints. We also show the demonstrations of our λ-system (above (d)), together with new proposal for controlling the stability.

It is worth remarking that this study is the first one which shows full numerical simulations of Lorentzian spacetime using Ashtekar's connection variables. This research direction was suggested [13] soon after Ashtekar completed his formulation, but has not yet been completed. Historically, an application to numerical relativity of the connection formulation was also suggested [14, 15] using Capovilla-Dell-Jacobson's version of the connection variables [16], which produce an direct relation to Newman-Penrose's Ψs. However here we apply Ashtekar's original formulation, because we know how to treat its reality conditions in detail, and how they form hyperbolicities. We will also describe the basic numerical procedures in this paper.

Due to the limited space, we have to omit details in some part in this report. More complete discussion can be found in our recent articles [1, 2].

2 Field equations to be compared

2.1 Hyperbolic formulations

We begin by providing our definitions of hyperbolic systems.

Definition 1 *We assume a certain set of (complex) variables* u_α $(\alpha = 1, \cdots, n)$ *forms a first-order (quasi-linear) partial differential equation system,*

$$\partial_t u_\alpha = J^{l\beta}{}_\alpha(u)\,\partial_l u_\beta + K_\alpha(u), \tag{1}$$

where J *(the characteristic matrix) and* K *are functions of* u *but do not include any derivatives of* u*. We say that the system (1) is:*

(I). **weakly hyperbolic,** *if all the eigenvalues of the characteristic matrix are real.*

(II). **strongly (diagonalizable) hyperbolic,** *if the characteristic matrix is diagonalizable and has all real eigenvalues.*

(III). **symmetric hyperbolic,** *if the characteristic matrix is a Hermitian matrix.*

We treat $J^{l\beta}{}_\alpha$ as a $n \times n$ matrix when the l-index is fixed. We say λ^l is an eigenvalue of $J^{l\beta}{}_\alpha$ when the characteristic equation, $\det(J^{l\beta}{}_\alpha - \lambda^l \delta^\beta{}_\alpha) = 0$, is satisfied. The eigenvectors, p^α, are given by solving $J^{l\alpha}{}_\beta\, p^{l\beta} = \lambda^l p^{l\alpha}$.

These three classes have the relation (III) $\in$ (II) $\in$ (I). The symmetric hyperbolic system gives us the energy integral inequalities, which are the primary tools for studying well-posedness of the system.

For more concrete descriptions for each systems, please refer our paper [9]. It might be worth remarking that the standard ADM formulation cannot be a first-order hyperbolic form, since there are curvature terms in the r.h.s. of dynamical equations.

2.2 Ashtekar's dynamical equations: three levels of hyperbolicity

The key feature of Ashtekar's formulation of general relativity [7] is the introduction of a self-dual connection as one of the basic dynamical variables. The new basic variables are the densitized inverse triad, $\tilde{E}^i_a$, and the SO(3,C) self-dual connection, $\mathcal{A}^a_i$, where the indices $i, j, \cdots$ indicate the 3-spacetime, and $a, b, \cdots$ are for SO(3) space. The total four-dimensional spacetime is described together with the gauge variables $\underset{\sim}{N}, N^i, \mathcal{A}^a_0$, which we call the densitized lapse function, shift vector and the triad lapse function. The system has three constraint equations,

$$C^{\rm ASH}_H := (i/2)\epsilon^{ab}{}_c \tilde{E}^i_a \tilde{E}^j_b F^c_{ij} \approx 0, \tag{2}$$

$$C^{\rm ASH}_{Mi} := -F^a_{ij}\tilde{E}^j_a \approx 0, \tag{3}$$

$$C^{\rm ASH}_{Ga} := \mathcal{D}_i \tilde{E}^i_a \approx 0, \tag{4}$$

which are called the Hamiltonian, momentum, and Gauss constraints equation, respectively. The dynamical equations for a set of $(\tilde{E}^i_a, \mathcal{A}^a_i)$ are

$$\partial_t \tilde{E}^i_a = -i\mathcal{D}_j(\epsilon^{cb}{}_a \underset{\sim}{N}\tilde{E}^j_c \tilde{E}^i_b) + 2\mathcal{D}_j(N^{[j}\tilde{E}^{i]}_a) + i\mathcal{A}^b_0 \epsilon_{ab}{}^c \tilde{E}^i_c, \tag{5}$$

$$\partial_t \mathcal{A}^a_i = -i\epsilon^{ab}{}_c \underset{\sim}{N}\tilde{E}^j_b F^c_{ij} + N^j F^a_{ji} + \mathcal{D}_i \mathcal{A}^a_0, \tag{6}$$

where $F^a_{ij} := 2\partial_{[i}\mathcal{A}^a_{j]} - i\epsilon^a{}_{bc}\mathcal{A}^b_i \mathcal{A}^c_j$ is the curvature 2-form.

We have to consider the reality conditions when we use this formalism to describe the classical Lorentzian spacetime. Fortunately, the metric will remain on its real-valued constraint surface during time evolution automatically if we prepare initial data which satisfies the reality condition. More practically, we further require that triad is real-valued. But again this reality condition appears as a gauge restriction on $\mathcal{A}^a_0$[17], which can be imposed at every time step. In our actual simulation, we prepare our initial data using the standard ADM approach, so that we have no difficulties in maintaining these reality conditions.

The set of dynamical equations (5) and (6) [the *original* equations] does have a weakly hyperbolic form [9], so that we regard the mathematical structure of the original equations as one step advanced from the standard ADM. Further, we can construct higher levels of hyperbolic systems by restricting the gauge condition and/or by adding constraint terms, $C^{\rm ASH}_H, C^{\rm ASH}_{Mi}$ and $C^{\rm ASH}_{Ga}$, to the original equations [9]. We extract only the final expressions here.

In order to obtain a symmetric hyperbolic system [1)], we add constraint terms to the right-hand-side of (5) and (6). The adjusted dynamical equations,

$$\partial_t \tilde{E}^i_a = -i\mathcal{D}_j(\epsilon^{cb}{}_a \underset{\sim}{N}\tilde{E}^j_c \tilde{E}^i_b) + 2\mathcal{D}_j(N^{[j}\tilde{E}^{i]}_a) + i\mathcal{A}^b_0\epsilon_{ab}{}^c \tilde{E}^i_c + \kappa_1 P^i{}_{ab} C^{{\rm ASH}b}_G, \tag{7}$$

$$\text{where} \quad P^i{}_{ab} \equiv N^i \delta_{ab} + i\underset{\sim}{N}\epsilon_{ab}{}^c \tilde{E}^i_c,$$

$$\partial_t \mathcal{A}^a_i = -i\epsilon^{ab}{}_c \underset{\sim}{N}\tilde{E}^j_b F^c_{ij} + N^j F^a_{ji} + \mathcal{D}_i \mathcal{A}^a_0 + \kappa_2 Q^a_i C^{\rm ASH}_H + \kappa_3 R_i{}^{ja} C^{\rm ASH}_{Mj}, \tag{8}$$

$$\text{where} \quad Q^a_i \equiv e^{-2}\underset{\sim}{N}\tilde{E}^a_i, \qquad R_i{}^{ja} \equiv ie^{-2}\underset{\sim}{N}\epsilon^{ac}{}_b \tilde{E}^b_i \tilde{E}^j_c$$

form a symmetric hyperbolicity if we further require $\kappa_1 = \kappa_2 = \kappa_3 = 1$ and the gauge conditions,

$$\mathcal{A}^a_0 = \mathcal{A}^a_i N^i, \qquad \partial_i N = 0. \tag{9}$$

We remark that the adjusted coefficients, $P^i{}_{ab}, Q^a_i, R_i{}^{ja}$, for constructing the symmetric hyperbolic system are uniquely determined, and there are no other additional terms (say, no $C^{\rm ASH}_H, C^{\rm ASH}_M$ for $\partial_t \tilde{E}^i_a$, no $C^{\rm ASH}_G$ for $\partial_t \mathcal{A}^a_i$) [9]. The gauge conditions, (9), are consequences of the consistency with (triad) reality conditions.

We can also construct a strongly hyperbolic system by restricting to a gauge $N^i \neq 0, \pm N\sqrt{\gamma^{ii}}$ (where γ^{ii} is the three-metric and we do not sum indices here) for the original equations (5), (6). Or we can also construct from the adjusted equations, (7) and (8), together with the gauge condition

$$\mathcal{A}^a_0 = \mathcal{A}^a_i N^i. \tag{10}$$

As for the strongly hyperbolic system, we hereafter take the latter expression.

[1)] Iriondo et al [18] presented a symmetric hyperbolic expression in a different form. The differences between ours and theirs are discussed in [9, 10].

	system	variables	Eqs of motion	remark
I	Ashtekar (weakly hyp.)	$(\tilde{E}^i_a, \mathcal{A}^a_i)$	(5), (6) (original)	"original Ashtekar"
II	Ashtekar (strongly hyp.)	$(\tilde{E}^i_a, \mathcal{A}^a_i)$	(7), (8) (adjusted with $\kappa = 1$)	(10) required
III	Ashtekar (symmetric hyp.)	$(\tilde{E}^i_a, \mathcal{A}^a_i)$	(7), (8) (adjusted with $\kappa = 1$)	(9) required
	Ashtekar (lambda)	$(\tilde{E}^i_a, \mathcal{A}^a_i, \lambda s)$	(17), (11) - (13)	
	Ashtekar (adjusted)	$(\tilde{E}^i_a, \mathcal{A}^a_i)$	(7), (8) (adjusted with different κ)	

Table 1: List of systems that we compare in this article.

2.3 Alternative approaches to obtain robust evolution 1: "λ-system"

Based on the symmetric hyperbolic feature of the system, Brodbeck, Frittelli, Hübner and Reula (BFHR) proposed an alternative dynamical system to obtain stable evolutions, which they named "λ-system" [12]. The idea of this approach is to introduce additional variables, λ, which indicates the violation of the constraints, and to construct a symmetric hyperbolic system for both the original variables and λs together with imposing dissipative dynamical equations for λs. BFHR constructed their λ-system based on Frittelli-Reula's symmetric hyperbolic formulation of the Einstein equations [19], and we [11] have also presented the similar system for the Ashtekar's connection formulation based on the above symmetric hyperbolic expression. Here we present our system which evolves the spacetime to the constraint surface, $\mathcal{C}_H \approx \mathcal{C}_{Mi} \approx \mathcal{C}_{Ga} \approx 0$ as the attractor. In [11], we also present a system which controls the perturvative violation of the reality condition.

We introduce new variables $(\lambda, \lambda_i, \lambda_a)$, as they obey the dissipative evolution equations

$$\partial_t \lambda = \alpha_1 \mathcal{C}_H - \beta_1 \lambda, \tag{11}$$

$$\partial_t \lambda_i = \alpha_2 \tilde{\mathcal{C}}_{Mi} - \beta_2 \lambda_i, \tag{12}$$

$$\partial_t \lambda_a = \alpha_3 \mathcal{C}_{Ga} - \beta_3 \lambda_a, \tag{13}$$

where $\alpha_i \neq 0$ (allowed to be complex numbers) and $\beta_i > 0$ (real numbers) are constants.

If we take $u_a^{(DL)} = (\tilde{E}^i_a, \mathcal{A}^a_i, \lambda, \lambda_i, \lambda_a)$ as a set of dynamical variables, then the principal part of (11)-(13) can be written as

$$\partial_t \lambda \cong -i\alpha_1 \epsilon^{bcd} \tilde{E}^j_c \tilde{E}^l_d (\partial_t \mathcal{A}^b_j), \tag{14}$$

$$\partial_t \lambda_i \cong \alpha_2 [-e\delta^l_i \tilde{E}^j_b + e\delta^j_i \tilde{E}^l_b](\partial_t \mathcal{A}^b_j), \tag{15}$$

$$\partial_t \lambda_a \cong \alpha_3 \partial_t \tilde{E}^l_a. \tag{16}$$

The characteristic matrix of the system $u_a^{(DL)}$ does not form a Hermitian matrix. However, if we modify the right-hand-side of the evolution equation of $(\tilde{E}^i_a, \mathcal{A}^a_i)$, then the set becomes a symmetric hyperbolic system. This is done by adding $\bar{\alpha}_3 \gamma^{il}(\partial_t \lambda_a)$ to the equation of $\partial_t \tilde{E}^i_a$, and by adding $i\bar{\alpha}_1 \epsilon^a{}_c{}^d \tilde{E}^c_i \tilde{E}^l_d (\partial_t \lambda) + \bar{\alpha}_2(-e\gamma^{lm}\tilde{E}^a_i + e\delta^m_i \tilde{E}^{la})(\partial_t \lambda_m)$ to the equation of $\partial_t \mathcal{A}^a_i$. The final principal part, then, is written as

$$\partial_t \begin{pmatrix} \tilde{E}^i_a \\ \mathcal{A}^a_i \\ \lambda \\ \lambda_i \\ \lambda_a \end{pmatrix} \cong \begin{pmatrix} A^l{}_a{}^{bi}{}_m & 0 & 0 & 0 & \bar{\alpha}_3 \gamma^{il} \delta_a{}^b \\ 0 & D^{la}{}_{ib}{}^m & i\bar{\alpha}_1 \epsilon^a{}_c{}^d \tilde{E}^c_i \tilde{E}^l_d & \bar{\alpha}_2 e(\delta^m_i \tilde{E}^{la} - \gamma^{lm} \tilde{E}^a_i) & 0 \\ 0 & -i\alpha_1 \epsilon_b{}^{cd} \tilde{E}^m_c \tilde{E}^l_d & 0 & 0 & 0 \\ 0 & \alpha_2 e(\delta^m_i \tilde{E}^l_b - \delta^l_i \tilde{E}^m_b) & 0 & 0 & 0 \\ \alpha_3 \delta_a{}^b \delta^l{}_m & 0 & 0 & 0 & 0 \end{pmatrix} \partial_t \begin{pmatrix} \tilde{E}^m_b \\ \mathcal{A}^b_m \\ \lambda \\ \lambda_m \\ \lambda_b \end{pmatrix}. \tag{17}$$

Clearly, the solution $(\tilde{E}^i_a, \mathcal{A}^a_i, \lambda, \lambda_i, \lambda_a) = (\tilde{E}^i_a, \mathcal{A}^a_i, 0, 0, 0)$ represents the original solution of the Ashtekar system. If the λs decay to zero after the evolution, then the solution also describes the original solution of the Ashtekar system in that stage. Since the dynamical system of $u_a^{(DL)}$, (17), constitutes a symmetric hyperbolic form, the solutions to the λ-system are unique. BFHR showed analytically that such a decay of λs can be seen for λs sufficiently close to zero with a choice of appropriate combination of αs and βs, and that statement can be also applied to our system. Therefore, the dynamical system, (17), is useful for stabilizing numerical simulations from the point that it recovers the constraint surface automatically.

2.4 Alternative approaches to obtain robust evolution 2: "adjusted system"

We also try to compare a set of evolution system, which we propose as "adjusted-system". The essential procedures are to add constraint terms in the right-hand-side of the dynamical equations with multipliers, and to choose the

multipliers so as to decrease the violation of constraint equation. This second step will be explained by obtaining non-positive (or non-zero) eigenvalues of the adjusted constraint propagation equations. We remark that this eigenvalue criterion is also the core part of the theoretical support of the above λ-system.

The fundamental equations that we will demonstrate in this report, are the same with (7) and (8), but here the real-valued constant multipliers κs are not necessary equals to unity. We set $\kappa \equiv \kappa_1 = \kappa_2 = \kappa_3$ for simplicity. Apparently the set of (7) and (8) becomes the original weakly hyperbolic system if $\kappa = 0$, becomes the symmetric hyperbolic system if $\kappa = 1$ and $N = const.$. The set remains strongly hyperbolic systems for other choices of κ except $\kappa = 1/2$ which only forms weakly hyperbolic system.

3 Comparing numerical performance

3.1 Model and Numerical method

The model we present here is gravitational wave propagation in a planar spacetime under periodic boundary condition. We perform a full numerical simulation using Ashtekar's variables. We prepare two +-mode strong pulse waves initially by solving the ADM Hamiltonian constraint equation, using York-O'Murchadha's conformal approach. Then we transform the initial Cauchy data (3-metric and extrinsic curvature) into the connection variables, $(\tilde{E}^i_a, \mathcal{A}^a_i)$, and evolve them using the dynamical equations. For the presentation in this article, we apply the geodesic slicing condition (ADM lapse $N = 1$ or densitized lapse $\underset{\sim}{N} = 1$, with zero shift and zero triad lapse). We have used both the Brailovskaya integration scheme, which is a second order predictor-corrector method, and the so-called iterative Crank-Nicholson integration scheme for numerical time evolutions. The details of the numerical method are described in [1], where we also described how our code shows second order convergence behaviour.

More specifically, we set our initial guess 3-metric as

$$\tilde{\gamma}_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ sym. & 1 + K(e^{-(x-L)^2} + e^{-(x+L)^2}) & 0 \\ sym. & sym. & 1 - K(e^{-(x-L)^2} + e^{-(x+L)^2}) \end{pmatrix}, \tag{18}$$

in the periodically bounded region $x = [-5, +5]$. Here K and L are constants and we set $K = 0.3$ and $L = 2.5$ for the plots.

In order to show the expected "stabilization behaviour" clearly, we artificially add an error in the middle of the time evolution. We added an artificial inconsistent rescaling once at time $t = 6$ for the $\mathcal{A}^2_y$ component as $\mathcal{A}^2_y \to \mathcal{A}^2_y(1 + \text{error})$.

3.2 Differences between three levels of hyperbolicity

We have performed comparisons of stability and/or accuracy between weakly and strongly hyperbolic systems, and between weakly and symmetric hyperbolic systems[1]. (We can not compare strongly and symmetric hyperbolic systems directly, because these two requires different gauge conditions.)

We omit figures in this report, but one can see a part of results in Fig.1 and Fig.2. We may conclude that higher level hyperbolic system gives us slightly accurate evolutions. However, if we evaluate the magnitude of L2 norms, then we also conclude that there is no measurable differences between strongly and symmetric hyperbolicities. This last fact will be supported more affirmatively in the next experiments.

3.3 Demonstrating "λ-system"

Next, we show a result of the "λ-system" [2]. Fig.1 (a) shows how the violation of the Hamiltonian constraint equation, $\mathcal{C}_H$, become worse depending on the term error. The oscillation of the L2 norm $\mathcal{C}_H$ in the figure due to the pulse waves collide periodically in the numerical region. We, then, fix the error term as a 20% spike, and try to evolve the same data in different equations of motion, i.e., the original Ashtekar's equation [solid line in Fig.1 (b)], strongly hyperbolic version of Ashtekar's equation (dotted line) and the above λ-system equation (other lines) with different βs but the same α. As we expected, all the λ-system cases result in reducing the Hamiltonian constraint errors.

3.4 Demonstrating "adjusted system"

We here examine how the adjusted multipliers contribute to the system's stability [2]. In Fig.2, we show the results of this experiment. We plot the violation of the constraint equations both $\mathcal{C}_H$ and $\mathcal{C}_{M_x}$. An artificial error

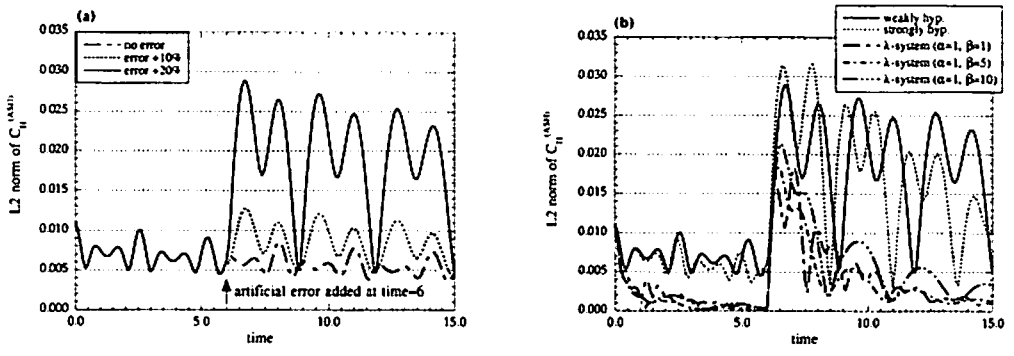


Figure 1: Demonstration of the λ-system in the Ashtekar equation. We plot the violation of the constraint (L2 norm of the Hamiltonian constraint equation, C_H) for the cases of plane wave propagation under the periodic boundary. To see the effect more clearly, we added artificial error at $t = 6$. Fig. (a) shows how the system goes bad depending on the amplitude of artificial error. The error was of the form $\mathcal{A}_y^2 \to \mathcal{A}_y^2(1 + \text{error})$. All the lines are of the evolution of Ashtekar's original equation (no λ-system). Fig. (b) shows the effect of λ-system. All the lines are 20% error amplitude, but shows the difference of evolution equations. The solid line is for Ashtekar's original equation (the same as in Fig.(a)), the dotted line is for the strongly hyperbolic Ashtekar's equation. Other lines are of λ-systems, which produces better performance than that of the strongly hyperbolic system.

term was added in the same way as above. The solid line is the case of $\kappa = 0$, that is the case of "no adjusted" original Ashtekar equation (weakly hyperbolic system). The dotted line is for $\kappa = 1$, equivalent to the symmetric hyperbolic system. We see other line ($\kappa = 2.0$) shows better performance than the symmetric hyperbolic case.

4 Why adjusted equations have better performance?

The remaining question is: why we can get the better performance by adding constraint terms in the dynamical equations? The added terms are basically *error* terms during the evolution for its original dynamical equations. Nevertheless, these terms improve the accuracy of the evolution. This question applies to both higher level of hyperbolicities and also for our proposing "adjusted systems". Here we introduce our plausible explanation schematically. The detail explanations and numerical experiments are in [2].

Suppose we have constraint equations, $C_1 \approx 0, C_2 \approx 0, \cdots$, in a system. We, normally, monitor the error of the evolution by evaluating these constraint equations on the each constant-time hypersurface. Such monitoring, on the other hand, can be performed also by checking the evolution equations of the constraint, which we denote constraint propagation equations (cf. [21]). If this set of constraint propagation equations could be written in a first order form, then we may predict the evolution behavior by its characteristic matrix, M, which is expressed by

$$\partial_t \begin{pmatrix} C_1 \\ C_2 \\ \vdots \end{pmatrix} \simeq M \partial_i \begin{pmatrix} C_1 \\ C_2 \\ \vdots \end{pmatrix}, \tag{19}$$

where $\simeq$ means we are extracting only the characteristic part.

The idea here is to estimate the eigenvalues of the characteristic matrix, M, after we took the Fourier transformation on the both sides of (19).

1. Clearly, if all the real part of the eigenvalues are negative, then all constraints decays to zero along to the system's evolution.
2. Alternatively, if all the eigenvalues are pure-imaginary, then all constraints evolve similar to wave equations, and diffusion due to differenciation may preserve from its unstable evolution.

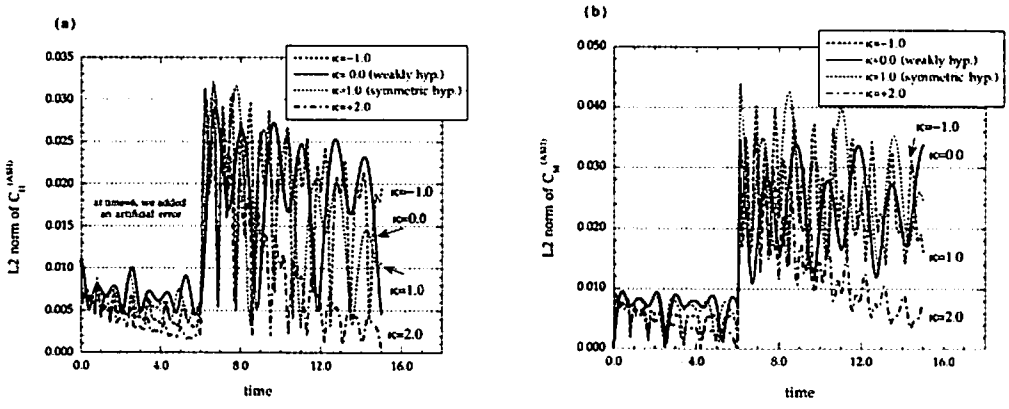


Figure 2: Demonstration of the adjusted system in the Ashtekar equation. An artificial error term was added at $t = 6$, in the form of $\mathcal{A}_y^2 \to \mathcal{A}_y^2(1 + \text{error})$, where error is +20% as before. Fig. (a) and (b) are L2 norm of the Hamiltonian constraint equation, C_H, and momentum constraint equation, C_{Mx}, respectively. The solid line is the case of $\kappa = 0$, that is the case of "no adjusted" original Ashtekar equation (weakly hyperbolic system). The dotted line is for $\kappa = 1$, equivalent to the symmetric hyperbolic system. We see other line ($\kappa = 2.0$) shows better performance than the symmetric hyperbolic case.

These guidelines are supported by von Neumann's stability analysis. For the adjusted Ashtekar equations, the eigenvalues of the amplicication matrix in von Neumann's method for the set of $(\mathcal{C}_H, \mathcal{C}_{Mi}, \mathcal{C}_{Ga})$ are $\{1, 1-\nu^2\kappa\sin\theta \pm i\nu\kappa|\sin\theta|$(multiplicity 2 each), $1-\nu^2(1-2\kappa)^2\sin^2\theta \pm i\nu|1-2\kappa||\sin\theta|\}$, where $\nu = \Delta t/\Delta x$, θ is the parameter for the frequency in grid spacing domain, and κ is the adjusted parameter (§2.4). We observe that von Neumann's necessary condition for the stability can be obtained in larger non-zero κ.

In [2], we show that such a case can be obtained by adding 'adjusted terms' both for Ashtekar's and Maxwell's systems. There we also show examples of unstable evolution by choosing adjusted terms which produce positive eigenvalues of M.

In this point, we can say that adjusted terms are responsible for obtaining the stable and/or accurate evolution system, and this is a way to control the stability of simulation, which effects more than the system's hyperbolicity.

5 Concluding remarks

We presented numerical comparisons on the accuracy/stability of the systems, mainly between three mathematical levels of hyperbolicity: weakly hyperbolic, strongly hyperbolic, and symmetric hyperbolic systems. We apply Ashtekar's connection formulation, because this is the only known system in which we can compare three hyperbolic levels with the same interface. We also tested two another approaches with a purpose of constructing an "asymptotically constrained" system, "λ-system" and "adjusted system", which can be robust against perturbative errors for the free evolution of the initial data.

We may conclude that higher levels of hyperbolic formulations help the numerics more, though its differences are small [22]. Two other approaches work as desired. These show quite good performance than those of the original and its symmetric hyperbolic form. This indicates, in turn, that the symmetric hyperbolic system is not always the best for controlling accuracy or stability.

The reason why higher hyperbolicity shows better stability may be explained by the eigenvalue analysis of the propagation equation of the constraints. Our two guidelines, negative-real eigenvalues or pure-imaginary eigenvalues, give us clear indications whether the constraints decay (i.e. stable system) or not for perturbative errors, at least to all our numerical experiments.

We think these results open a new direction to numerical relativists for future treatment of the Einstein equations.

To conclude, we are glad to announce that Ashtekar's connection variables have finally been applied in

Quantum Fluctuations of Bulk Inflaton in Inflationary Brane World

Kazuya Koyama,[1] Shinpei Kobayashi[2]

Graduate School of Human and Environment Studies, Kyoto University, Kyoto 606-8501, Japan ,

Jiro Soda[3]

Department of Fundamental Sciences, FIHS, Kyoto University, Kyoto, 606-8501, Japan

Abstract

The inflationary scenario for the brane world driven by the bulk inflaton is proposed. The quantum fluctuations of the inflaton is calculated and compared to those of the conventional 4-dimensional inflationary scenario. It is shown that the deviation of the primordial spectrum of this model from that of the conventional one is too small to be observed even if AdS radius is very large. Hence, it turns out that the inflation caused by the bulk inflaton is viable in the context of brane world cosmology. (This talk is based on our article [1].)

1 Introduction

The possibility that our universe is a four-dimensional hypersurface (3-brane) embedded in the higher-dimensional spacetime was known before and forgotten for a long time [2, 3]. Recently, stimulated by M-theory, this possibility comes to be argued eagerly. In such a current, 'An alternative to compactification' was suggested by Randall and Sundrum (RS) [4]. It is fascinating because the internal dimension is not compact and hence there is no stability problem. The bulk of this model consists of two five-dimensional anti-de Sitter spacetimes (AdS_5). On the brane, there are 0-mode and massive modes of graviton in the linearlized gravity as in the Kaluza-Klein compactification. The Newtonian gravity on the brane is realized by 0-mode [5, 6], on the other hand, massive modes give the correction to Newtonian gravity . This is the outline of the RS model.

Since the RS model is proposed, the cosmological consequences of the model have been studied and found that it is completely consistent with the present cosmological observations [7]-[12]. However the validity of this model in the context of the very early universe is not investigated completely. Especially, the detailed analysis of the primordial spectrum of the fluctuation during the inflation is lacked.

In the RS model, the tension is balanced with the bulk cosmological constant. This has allowed the Minkowski spacetime as a ground state. In the presence of the cosmological matter, the evolution of the spacetime can be described by the Friedmann equation at low energy. If we assume the scalar field as a cosmological matter, the inflationary solution will arise because the effective tension of the brane exceeds the bulk cosmological constant. This is a naive extension of the conventional inflationary scenario [13]-[21].

But considering theories of particle physics (e.g.String Theory), many scalar fields live in the higher-dimensional spacetime (bulk). So it is natural to consider inflatons in the bulk - AdS_5. Suppose that a scalar field lives in the bulk and has a potential. When it slow-rolls, what will happen? The potential works as the effective cosmological constant and it makes non-zero Hubble constant H on the brane. In other words, the de Sitter expanding universe can be realized by changing the bulk cosmological constant. Furthermore, this scalar field continues rolling down, oscillates around some vacuum and get to some stable state with the reheating process. This model does have all properties which the conventional

[1] E-mail:kazuya@phys.h.kyoto-u.ac.jp
[2] E-mail:shinpei@phys.h.kyoto-u.ac.jp
[3] E-mail:jiro@phys.h.kyoto-u.ac.jp

inflationary scenario has and it has the extra dimension. The purpose of this letter is to study the nature of the quantum fluctuations in this inflationary brane world. As a result, we will show that our model is observationally acceptable as an inflationary scenario in the brane world cosmology.

The organization of this paper is as follows: In Section 2 we shall present the inflation model driven by the bulk inflaton. In Section 3 we shall calculate the quantum fluctuations of the bulk inflaton. Final section is devoted to the conclusion.

2 Inflation Driven by the Bulk Inflaton

How can we make the brane which expands exponentially? One of the most popular and natural ideas to make the de Sitter brane is to put an inflaton on the brane. This model is not different from the ordinary four-dimensional inflationary scenario in respect of using the four-dimensional matter. The models using the tension of the brane also exist [13]-[22]. But our model is different from any other models proposed so far [23]. In our model, the de Sitter brane is realized by the potential of the scalar field in the bulk. We see this mechanism in detail.

The action of this system can be divided into three parts,

$$S = S_{grav.} + S_{brane} + S_{scalar} \ . \tag{1}$$

The first term of the action is the usual Einstein-Hilbert action with a negative cosmological constant:

$$S_{grav.} = \frac{1}{16\pi G_5} \int d^5x\sqrt{-g_5}(R - 2\Lambda_5) \ , \tag{2}$$

where g_5 and G_5 are the five-dimensional metric and gravitational constant, R denotes the five-dimensional Ricci scalar. The five-dimensional cosmological constant Λ_5 is related to the AdS radius $\bar{l}$ by the following equation,

$$\Lambda_5 = -\frac{6}{\bar{l}^2} \ . \tag{3}$$

Note that $\bar{l}$ is the AdS radius without a scalar field. The second term of the action is the brane action:

$$S_{brane} = -\sigma \int d^4x\sqrt{-g} \ . \tag{4}$$

Here σ works as the tension of the brane. The last term of S is the action of the scalar field living in the bulk:

$$S_{scalar} = -\int d^5x\sqrt{-g_5}\Big(\frac{1}{2}(\nabla\phi)^2 + U(\phi)\Big) \ , \tag{5}$$

where ∇ is the covariant derivative with respect to g_5, and $U(\phi)$ is the potential of the scalar field, which plays an important role.

Now turn to find the solution of this system. If there is no scalar field, the geometry of the bulk is represented by AdS_5 and the metric of AdS_5 becomes

$$ds^2 = dy^2 + (\bar{l}H)^2 \ \sinh^2(y/\bar{l})(-dt^2 + H^{-2}\exp(2Ht)\delta_{ij}dx^i dx^j) \ . \tag{6}$$

Here H is the four-dimensional Hubble constant and y denotes the extra dimension. We consider a four-dimensional hypersurface (3-brane) at $y = y_0$ and the location of the brane is related with H as follows,

$$H = \frac{1}{\bar{l}\sinh(y_0/\bar{l})}. \tag{7}$$

Following the Randall-Sundrum model, we glue this AdS_5 and the copy of it at $y = y_0$. So the bulk has Z_2 symmetry along the y-direction. In this coordinate system, the range of y is $0 \le y \le y_0$. Calculating Einstein equations and junction conditions at $y = y_0$, we obtain the condition for the Hubble constant

$$H^2 = \Big(\frac{4\pi G_5}{3}\sigma\Big)^2 + \frac{\Lambda_5}{6} \ . \tag{8}$$

In the RS model, the tension is tuned to

$$\sigma = \frac{3}{4\pi G_5 l} . \tag{9}$$

From (3), we can see that H becomes zero, then the brane is flat. But in this paper, we want to make the de Sitter expanding brane, that is, the brane having the positive Hubble constant.

Here we have two ways to realize $H > 0$. One of them is to increase σ and another is to decrease $|\Lambda_5|$. Our model is the latter. We consider the potential $U(\phi)$ to decrease $|\Lambda_5|$. This potential can be regarded as the effective cosmological constant, that is, we can define the effective cosmological constant Λ_{eff},

$$\Lambda_{eff} \sim \Lambda_5 + U(\phi) . \tag{10}$$

Here we neglect some coefficients. Note that we consider the situation where the slow-roll conditions are satisfied, so we neglect the kinetic term of the scalar field. Λ_{eff} is related to the new AdS radius l as

$$\Lambda_{eff} = -\frac{6}{l^2}. \tag{11}$$

With the change of AdS radius, the metric changes as

$$ds^2 = dy^2 + (lH)^2 \sinh^2(y/l)(-dt^2 + H^{-2}\exp(2Ht)\delta_{ij}dx^i dx^j) . \tag{12}$$

Furthermore, we can rewrite (8) as follows

$$H^2 = \left(\frac{4\pi G_5}{3}\sigma\right)^2 + \frac{\Lambda_{eff}}{6} . \tag{13}$$

We consider the following conditions

$$\Lambda_5 < \Lambda_{eff}, \tag{14}$$

$$\Lambda_{eff} < 0. \tag{15}$$

Eq.(14) is the condition to realize the de Sitter brane and eq.(15) is the condition to make the bulk AdS_5. Thus we have AdS_5 with the de Sitter brane. This is a novel inflationary scenario. When a scalar field starts slow-rolling, observers living on the brane see inflation. Our model is different from other models which treat the inflatons confined on the brane. After the scalar field slow-rolls along the potential, it oscillates around the vacuum and gets to a stable state. If we tune the minimum of the potential to the Randall-Sundrum value, the system finally gets to the Randall-Sundrum system (flat brane scenario).

3 Quantum Fluctuations

We want to investigate the fluctuations of the scalar field ϕ living in the bulk (AdS_5) and evaluate it on the brane. We consider the zero-th order of the slow-roll approximation, that is, $U(\phi) \sim const.$. Therefore the spacetime of the brane is exact de Sitter. The equation for the fluctuations is Klein-Gordon equation in AdS_5:

$$\left[\partial_y^2 + \frac{4}{l}\coth(y/l)\partial_y - \frac{1}{(lH)^2\sinh^2(y/l)}(\partial_t^2 + 3H\partial_t) + \frac{1}{l^2\sinh^2(y/l)e^{2Ht}}\partial_{\vec{x}}^2\right]\delta\phi = 0. \tag{16}$$

As AdS_5 is highly-symmetric, we can use the separation of variables. The fluctuation $\delta\phi$ can be expanded as

$$\delta\phi(y,t,x) = \int d^3p\, dm \quad (a_{pm}f_m(y)g_{pm}(t)e^{ipx} + (h.c.)) . \tag{17}$$

Here, a_{pm} represents the annihilation operator of the field theory in AdS_5 and a_{pm} satisfies usual commutation relation,

$$\left[a_{pm}, a^\dagger_{p'm'}\right] = \delta(p-p')\delta(m-m') . \tag{18}$$

The vacuum can be defined with $\{a_{pm}\}$ as

$$a_{pm}|0> = 0, \quad for\ \forall p, m. \tag{19}$$

Now the equations for $f_m(y)$ and $g_{pm}(t)$ become

$$f_m''(y) + \frac{4}{l} f_m'(y) \coth(y/l) + \frac{m^2}{(lH)^2 \sinh^2(y/l)} f_m(y) = 0 , \tag{20}$$

$$\ddot{g}_{pm}(t) + 3H\dot{g}_{pm}(t) + \left(p^2 e^{-2Ht} + m^2\right) g_{pm}(t) = 0 , \tag{21}$$

where prime denotes derivative with respect to y, dot denotes derivative with respect to t and m corresponds to the Kaluza-Klein mass which appears in the Kaluza-Klein compactification.

As the background has Z_2 symmetry along the y-direction, it is natural to impose the Z_2 symmetry on the scalar field also. Then, the junction condition becomes

$$\partial_y \delta\phi(y,t,x) \,|_{y_0} = 0 . \tag{22}$$

So we solve the equations under Neumann boundary condition. As to the condition of $g(t)$, we choose the Bunch-Davis vacuum, which coincides with the Minkowski vacuum at high-frequency ($p/aH \to \infty$, a is a scale factor).

Under these conditions, we can solve (20) and (21) uniquely. We get 0-mode ($m = 0$) and massive modes ($m > 3H/2$). The massive modes with $0 < m < 3/2$ are not normalizable. So we exclude these modes from solutions. The solution for 0-mode is

$$f_0(y) = \frac{1}{lH\sqrt{l \sinh(y_0/l)\cosh(y_0/l) - y_0}} , \tag{23}$$

$$g_{p0}(t) = \frac{\sqrt{\pi}}{2} H(-\eta)^{3/2} H^{(1)}_{3/2}(-p\eta) , \tag{24}$$

where $H^{(1)}$ is the first Hankel function and η is conformal time defined as

$$\eta \equiv -e^{-Ht}. \tag{25}$$

Similarly, the solutions for massive modes are

$$f_m(y) = \frac{1}{lH\sqrt{l(\zeta(\beta) + \xi(\beta))}} \sinh^{-2}(y/l) \times \left(P^{-2}_{-1/2+i\beta}(\coth(y/l)) - \alpha(y_0/l) Q^{-2}_{-1/2+i\beta}(\coth(y/l))\right) , \tag{26}$$

$$g_{pm}(t) = \frac{\sqrt{\pi}}{2} H(-\eta)^{3/2} e^{-\pi\beta/2} H^{(1)}_{i\beta}(-p\eta) , \tag{27}$$

where

$$\beta = \sqrt{\frac{m^2}{H^2} - \frac{9}{4}} , \tag{28}$$

$$\zeta(\beta) = \left| \frac{\Gamma(i\beta)}{\Gamma(5/2 + i\beta)} \right|^2 , \tag{29}$$

$$\xi(\beta) = \left| \frac{\Gamma(-i\beta)}{\Gamma(5/2 - i\beta)} - \pi\alpha(y_0/l) \frac{\Gamma(-3/2 + i\beta)}{\Gamma(1 + i\beta)} \right|^2 , \tag{30}$$

$$\alpha(y_0/l) = \frac{(5/2 + i\beta) P^{-2}_{1/2+i\beta}(\coth(y_0/l)) + (3/2 - i\beta)\coth(y_0/l) P^{-2}_{-1/2+i\beta}(\coth(y_0/l))}{(5/2 + i\beta) Q^{-2}_{1/2+i\beta}(\coth(y_0/l)) + (3/2 - i\beta)\coth(y_0/l) Q^{-2}_{-1/2+i\beta}(\coth(y_0/l))} . \tag{31}$$

We can calculate the vacuum expectation value of $\delta\phi^2$ using these solutions, but note that $\delta\phi$ is the five-dimensional scalar field. The four-dimensional scalar field ϕ_4 is related to ϕ by the following equation,

$$\phi_4 = l^{1/2}\phi. \tag{32}$$

Consequently, the four-dimensional vacuum expectation value of $\delta\phi_4^2$ on the brane becomes

$$\begin{aligned} < \delta\phi_4(y_0,t,x)^2 > &= l < \delta\phi(y_0,t,x)^2 > \\ &= l\int d^3p\Big(|f_0|^2|g_{p0}|^2 + \int_{3/2}^{\infty} dm\ |f_m|^2|g_{pm}|^2\Big) \qquad (33) \\ &\equiv \int d^3p P(p) \qquad (34) \end{aligned}$$

Here, we define $P(p)$ as the power spectrum. We are interested in the fluctuations on the brane, so we evaluate it at $y = y_0$. And we sum up m, since m is the momentum of the extra dimension. We can not see each m but the summation.

Here is the point we have to notice. It is that in the large m and p limit, $|g_{pm}|^2$ behaves

$$|g_{pm}|^2 \sim \frac{H^2}{2\sqrt{m^2+(-p\eta)^2}}. \qquad (35)$$

Here we remove the conformal factor $(-\eta)^{3/2}$ to compare the field theory in Minkowski spacetime. And Hubble constant H disappears when $|g_{pm}|^2$ is multiplied by $|f_{pm}|^2$. The equation (35) shows that the summation has the logarithmic divergence. This ultra-violet divergence appears in the five-dimensional field theory in Minkowski spacetime as well. For this reason, this divergence is not the real divergence in the sense that we can not observe it. After all, we can conclude that we have to subtract this divergence to get the correct expectation value.

The subtraction of $|g_{pm}|^2$ can be carried out as

$$\frac{|g_{pm}|^2}{(-\eta)^{3/2}} - \frac{H^2}{2\omega}. \qquad (36)$$

Here $\omega = \sqrt{(-p\eta)^2+m^2}$ and we inserted H for the reason as mentioned before. Thus subtracted $|g_{pm}|^2$ behaves as $\mathcal{O}(\frac{1}{m^3})$ for large m and, hence, no divergence arises in the integration with respect to m. Finally we get the power spectrum of the physical expectation value as

$$P(p) = l\left(|f_0|^2\left[|g_{p0}|^2 - (-\eta)^3\frac{H^2}{2\sqrt{(-p\eta)^2}}\right] + \int_{3/2}^{\infty} dm\ |f_m|^2\left[|g_{pm}|^2 - (-\eta)^3\frac{H^2}{2\omega}\right]\right). \qquad (37)$$

Next, we analyze the $P(p)$ on the brane, especially the effect of massive modes. The reason why we are interested in massive modes is that massive modes appear due to the extra dimension. In the four-dimensional theory, Kaluza-Klein mass does not actually exist. So it is massive modes that show the difference between the Einstein theory and the brane-world theory.

Figure 1 is the lH-dependence of the ratio of massive modes amplitude to the 0-mode amplitude We evaluate it at $-p\eta = 1$ (horizon-crossing). As l is the AdS radius, it can be said that l defines the scale of the compactification. In the RS model, the effect of massive modes dominates as l becomes large. When the radius of the compactification becomes large, we come to be able to see the extra dimension and to feel the effect of the massive modes. In our case, the same phenomenon occurs at $lH < 1$. But $lH \gg 1$, the ratio approaches to the constant value. Since the lH-dependence of the ratio is determined only by $f(y)$, we pay attention to the asymptotic form of $f(y)$ at large lH. Using (7), $|f_m|^2/|f_0|^2$ can be written as

$$\frac{|f_m(y_0)|^2}{|f_0(y_0)|^2} \overset{lH\to\infty}{\sim} \frac{2}{3\pi}\frac{m^2-\frac{9}{4}}{m^2} \qquad (38)$$

This equation shows that if we determine the value of m, $|f_m|^2/|f_0|^2$ approaches to constant at large lH.

This result can be understood in the following way. Following [18, 22], we rewrite (20) by using $\tilde{y}$ defined as

$$\sinh\tilde{y} \equiv \sinh^{-1}(y/l). \qquad (39)$$

In this relation we can see that $y \to 0$ corresponds to $\tilde{y} \to \infty$. Setting $f(\tilde{y}) = (\sinh\tilde{y})^{3/2}F(\tilde{y})$, the equation for $F(\tilde{y})$ becomes Schrödinger-like form,

$$-\frac{d^2F(\tilde{y})}{d\tilde{y}^2} + V(\tilde{y})F(\tilde{y}) = m^2F(\tilde{y}) \qquad (40)$$

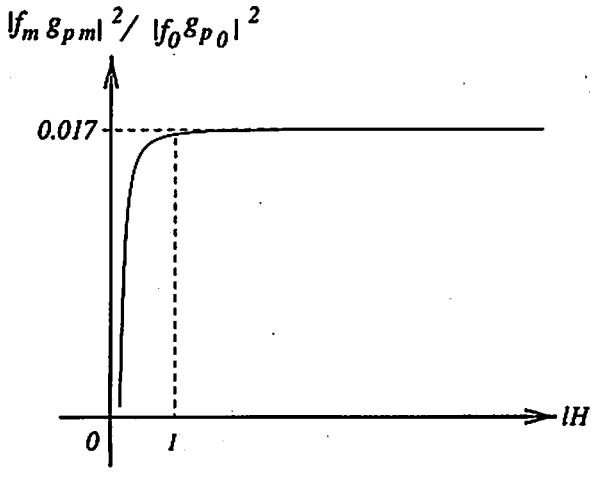


Figure 1: The ratio of the massive mode and 0-mode. $|f_m g_m|^2$ takes the maximum value around $m = 1.9$ so we use $m = 1.9$.

where $V(\tilde{y})$ is so-called 'volcano' potential,

$$V(\tilde{y}) = \frac{15}{4\sinh\tilde{y}} + \frac{9}{4} - 3\coth\tilde{y}_0\delta(\tilde{y} - \tilde{y}_0). \tag{41}$$

This potential is shown in Figure 2.

Since the subtracted $g_{pm}(t)$ behaves as $\mathcal{O}(\frac{1}{m^3})$, only the massive modes around $m = 3H/2$ are relevant. Such massive modes are able to penetrate easily as lH becomes large, because lH becomes large, $\tilde{y}_0$ becomes large and the 'depth' of the volcano becomes shallow.

The difference between our model and RS model is the asymptotic value. In RS model, volcano potential asymptotically approaches to 0. So when lH becomes large, 'volcano' disappears completely and it breaks the localization of 0-mode. In the end, Newtonian gravity on the brane breaks down. In contrast, there continues to exist the volcano in our model because of the non-zero asymptotic value. Numerical calculation and (38) shows that the contribution of massive modes to the spectrum is enough small compared with 0-mode even at large lH. As mentioned before, $|f_m(y_0)|^2/|f_0(y_0)|^2 \sim 0.080$ at $m = 1.9H$. Thus the ratio of massive modes to 0-mode at large lH and $m = 1.9H$ becomes about 0.017. This is the maximum value of the ratio. If we integrate about m at large lH, the ratio becomes 0.03, namely, the contribution of 0-mode is more than 30 times larger than that of massive modes.

At last, we refer to 0-mode. 0-mode is different from the mode which appears in the usual four-dimensional theory by some coefficients. So the amplitude of 0-mode has changed, but p-dependence of 0-mode does not change at all.

4 Conclusion

We have proposed an inflationary scenario in the brane world, i.e., inflation driven by the bulk inflation. The fluctuations of the bulk inflation has been also calculated. The contribution of massive modes is too small to change the spectrum at super horizon scale. So we can conclude that massive modes can not be observed even if lH is very large.

Here we refer to the density perturbation. In the conventional (i.e., four-dimensional) inflationary scenario, the relation between the density perturbation and $\delta\phi_4$ is well known. In contrast, the scalar field living in the bulk has both t and y-dependence. There may exist the non-trivial background solution

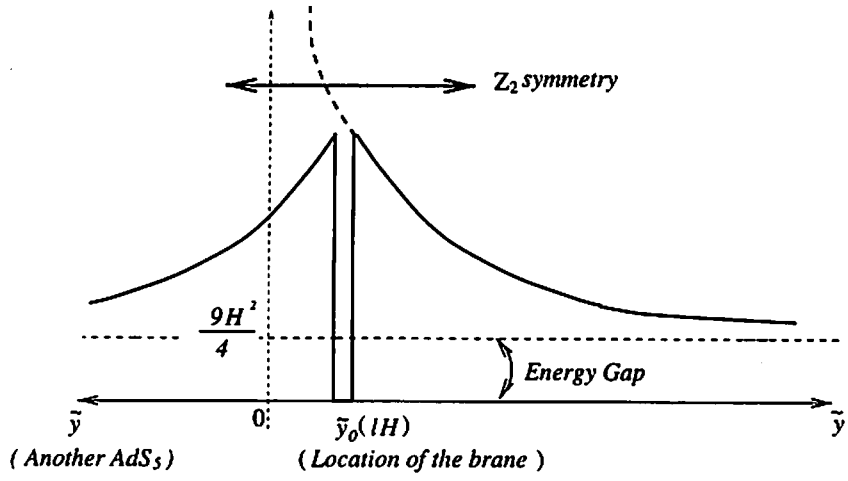


Figure 2: Volcano potential is shown. $\tilde{y_0}$ denotes the location of the brane and two AdS_5s are glued together at $\tilde{y} = \tilde{y_0}$. If $\tilde{y_0}$ becomes large (i.e. lH becomes large), the height of the volcano becomes low.

due to this y-dependence. If so, the density perturbation in this case may change drastically compared with that of the four-dimensional case. But if the mass of the scalar field is enough small, we can use the slow-roll approximation and we expect to get the almost same result as the four-dimensional case.

Now, there is a point we do not refer in this research. That is the evolution after the inflation in our context. Cosmological evolution has been investigated in [24]-[29]. Considering those results, there is the possibility that massive modes start dominating in the evolution after the inflation even if massive modes do not dominate at the end of the inflation. So we can say that concerning the initial spectrum, the brane-world model has survived and our scenario is one possible candidate where 'inflation' and 'brane' are compatible with each other.

Acknowledgements

We would like to thank M.Sasaki and Y.Himemoto for discussions. The authors are also grateful to H.Kodama for useful comments. The work of K.K. was supported by JSPS Research Fellowships for Young Scientist No.4687.

References

[1] S. Kobayashi, K. Koyama and J. Soda, hep-th/0009160, submitted to Phys. Lett. B.

[2] V. A. Rubakov and M. E. Shaposhnikov, Phys. Lett. **125B**, 136 (1983).

[3] K. Akama, "Pregeometry" in Lecture Notes in Physics, 176, Gauge Theory and Gravitation, Proceedings, Nara, 1982, (Springer-Verlag), edited by K. Kikkawa, N. Nakanishi and H. Nariai, 267-271, hep-th/0001113.

[4] L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 4690 (1999). See also L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999).

[5] J.Garriga and T.Tanaka, Phys. Rev. Lett. **84** 2778 (2000).

[6] S.B.Giddings, E.Katz and L.Randall JHEP 0003 (2000) 023.

[7] T.Shiromizu, K. Maeda and M. Sasaki, Phys. Rev. **D62** 024012 (2000).

[8] P. Binétruy, C. Deffayet, U. Ellwanger and D. Langlois, Phys. Lett.**B477**, 285 (2000).

[9] E. E. Flanagan, S. H. H. Tye and I. Wasserman, Phys. Rev. **D62** 044039 (2000).

[10] S.Mukohyama, Phys. Lett. **B473** 241 (2000).

[11] D.Ida, gr-qc/9912002.

[12] P. Kraus, JHEP 9912 (1999) 011.

[13] H. A. Chamblin and H. S. Reall, Nucl. Phys. **B562**, 133 (1999).

[14] N. Kaloper, Phys. Rev. **D60**, 123506 (1999).

[15] T. Nihei, Phys. Lett. **B465**, 81 (1999).

[16] H.B. Kim, H.D. Kim, Phys. Rev. **D61**, 064003 (2000).

[17] K. Koyama and J. Soda, Phys. Lett. **B483**, 432 (2000).

[18] J. Garriga and M. Sasaki, Phys. Rev. **D62**, 043523 (2000).

[19] S. W. Hawking, T. Hertog, H. S. Reall, Phys. Rev. **D62**, 043501 (2000).

[20] S.Nojiri and D.Odintsov, Phys. Lett. **B484**, 119 (2000).

[21] R. Maartens, D. Wands, B. Bassett and I. Hear, Phys. Rev. **D62**, 041301 (2000).

[22] D.Langlois, R.Maartens and D.Wands, hep-th/0006007 (2000).

[23] After almost completing this work, we were informed that the similar idea has been developed independently by M.Sasaki and his colabolators. M.Sasaki, talk at "The Ninth Marcel Grossmann Meeting", Rome (2000).

[24] K. Koyama and J. Soda, hep-th/0005239, Phys. Rev. **D**, in press.

[25] H. Kodama, A. Ishibashi, O. Seto, Phys. Rev. **D62**, 064022 (2000).

[26] R. Maartens, hep-th/0004166.

[27] D. Langlois, hep-th/0005025.

[28] S.Mukohyama, hep-th/0004067.

[29] C. van de Bruck, M. Dorca, R. Brandenberger and A. Lukas, hep-th/0005032.

Black holes in the brane world : Time symmetric initial data

Tetsuya Shiromizu

RESCEU, University of Tokyo, Tokyo 113-0033, Japan

Masaru Shibata

Graduate School of Arts and Sciences,
University of Tokyo, Tokyo 153-8902, Japan

Abstract

We numerically construct time-symmetric initial data sets of a black hole in the Randall-Sundrum brane world model, assuming that the black hole is spherical on the brane. We find that the apparent horizon is cigar-shaped in the 5D spacetime.

1 Introduction

Motivated by Horava-Witten model [1], the so called brane world model has been actively investigated [2]. Among several models, a simple, but very attractive model was recently proposed by Randall and Sundrum [3, 4]. According to their scenario, we are living in a 4D domain wall in 5D bulk spacetime. The noteworthy features of their model are that in the linearized theory, the conventional gravity can be recovered on the brane [3, 4, 5, 6, 7] and that a homogeneous, isotropic universe can be simply described if we consider a 4D domain wall moving in the 5D Schwarzschild-anti de Sitter spacetime [8].

One of the most non-linear objects in the theory of gravity is a black hole, which should be also investigated to understand the nature of the models in strong fields. However, because of the complexity of the equations, any realistic, exact solutions for black holes have not been discovered in the brane world model, even with help of numerical computation so far. We only know that the effective 4D gravitational equation on the brane is different from the Einstein equation [9] (see Appendix A), so that the static solution for a non-rotating black hole should not be identical with the 4D Schwarzschild solution. Indeed, a linear perturbation analysis [5, 7] shows that a solution of gravitational field outside self-gravitating bodies on the brane is slightly different from the 4D Schwarzschild solution. Chamblin et al. [10] conjecture that the topology of black hole event horizons would be spherical with the cigar-shaped surface in the 5D spacetime. However, nothing has been clarified substantially.

In this paper, as a first step toward self-consistent studies for black holes in the brane world, we numerically compute a black hole space using a time symmetric initial value formulation; namely we solve the 5D Einstein equation only on a spacelike 4D hypersurface. Thus, the black hole obtained here is not static nor the exact solution for the 5D Einstein equation, implying that we cannot identify the event horizon. However, we can investigate the property of the horizon determining the apparent horizon which could give us an insight on the black hole in the brane world. We focus on the Randall-Sundrum's second model [4], and assume that the black hole is spherical on the brane, but the shape of the horizon is non-trivial in the bulk. We will determine the apparent horizon on the brane and show that the black hole is cigar-shaped as conjectured in [10].

2 formulation and results

We consider time symmetric, spacelike hypersurfaces, Σ_t, in the brane world model assuming the vanishing extrinsic curvature; i.e.,

$$H_{\mu\nu} \equiv (\delta^\alpha_\mu + t_\mu t^\alpha)^{(4)}\nabla_\alpha t_\nu = 0, \tag{1}$$

where t^μ is the unit normal timelike vector to Σ_t and ${}^{(4)}\nabla_\alpha$ is the covariant derivative with respect to the 4D metric on Σ_t. In this case, the momentum constraint is satisfied trivially, and the equation of the Hamiltonian constraint becomes

$$ {}^{(4)}R = 16\pi G_5(\Lambda + {}^{(5)}T_{\mu\nu}t^\mu t^\nu), \tag{2} $$

where ${}^{(4)}R$ is the Ricci scalar on Σ_t, and $G_5(=\kappa_5^2/8\pi)$, Λ and ${}^{(5)}T_{\mu\nu}$ denote the gravitational constant, negative cosmological constant, and energy-momentum tensor in 5D spacetime [cf., Eq. (10)]. We choose the line element on Σ_t in the form

$$ dl^2 = \frac{1}{z^2}\Big[\ell^2 dz^2 + \psi^4(dr^2 + r^2 d\Omega)\Big], \tag{3} $$

where $\ell = \sqrt{-\kappa_5^2\Lambda/6}$, z (≥ 1) denotes the coordinate orthogonal to the brane and r (≥ 0) is the radial coordinate on the brane. We assume that the brane is located at $z = 1$. Note that we simply choose this line element for convenience of the analysis. In this paper, we focus on a black hole which is spherical on the brane, i.e., $\psi = \psi(r, z)$. Then, the explicit form of the Hamiltonian constraint in the bulk (for $z > 1$) is written in the form

$$ \psi'' + \frac{2}{r}\psi' + \frac{3}{2\ell^2}\Big[\Big(\partial_z^2\psi - \frac{3}{z}\partial_z\psi\Big)\psi^4 + 3(\partial_z\psi)^2\psi^3\Big] = -\frac{\kappa_5^2}{4}{}^{(5)}\tau_{\mu\nu}t^\mu t^\nu. \tag{4} $$

where $' = \partial/\partial r$, and ${}^{(5)}\tau_{\mu\nu}$ is the energy-momentum tensor in the bulk, which is introduced for numerical convenience.

Equation (4) is an elliptic type equation and should be solved imposing boundary conditions at $z = 1$, $z \gg 1$, $r = 0$, and $r \gg \ell$. The boundary condition at $z = 1$ is derived from Israel's junction condition [11] as (see Appendix A for the derivation)

$$ \partial_z\psi|_{z=1} = 0. \tag{5} $$

The boundary conditions at $z \gg 1$ and $r \gg \ell$ are obtained from the linear perturbation analysis (see Appendix B). For $r \gg \ell$ and $r > \ell z$, it becomes

$$ \psi \simeq 1 + \frac{MG_4}{2r}\Big[1 + \frac{1}{2}\Big(\frac{R}{r}\Big)^2 + O\Big((\ell/r)^4\Big)\Big], \tag{6} $$

where $G_4 = G_5/\ell$, M is the gravitational mass on the brane, and $R = (2/3)^{1/2}\ell$. For $z \gg 1$,

$$ \psi \simeq 1 + \frac{3}{4}\frac{G_4 M}{Rz}\Big(1 + \frac{r^2}{z^2R^2}\Big)^{-3/2}. \tag{7} $$

To determine the existence of a black hole, we search for the apparent horizon. Here, we determine two horizons [12]. One is defined to be the spherical two-surface on the brane on which the expansion of the null geodesic congruence confined on the brane is zero[13], i.e.,

$$ \theta_3 = \frac{2}{\psi^3}\Big(2\psi' + \frac{1}{r}\psi\Big) = 0. \tag{8} $$

The other is the apparent horizon in full 4D space, which is defined with respect to the null geodesic congruence in full 5D spacetime and satisfies[13]

$$ \theta_4 = {}^{(4)}\nabla_i s^i = 0, \tag{9} $$

where s^i is a unit normal to the surface of the apparent horizon. Explicit equation for determining this apparent horizon is shown in Appendix C.

The procedure of numerical analysis is as follows. First, we artificially put the matter of $\rho_h \equiv {}^{(5)}\tau_{\mu\nu}t^\mu t^\nu > 0$ in the bulk. This method is employed because we do not have to consider the inner

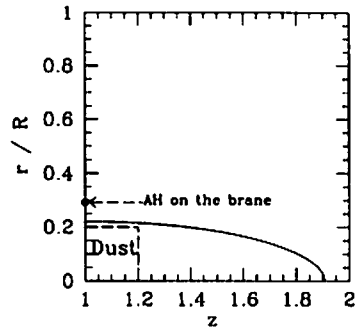


Figure 1: Location of the apparent horizons on the brane (filled circle) and in the 4D space (solid line). Artificial matter is confined in the region shown by the dashed line.

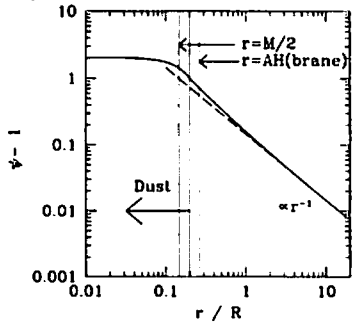


Figure 2: Profile of $\psi - 1$ on the brane (solid line). Location of the apparent horizon on the brane is shown. The dashed line denotes $\psi - 1 = M/2r$

boundary condition of black holes with this treatment. As long as ρ_h is confined around the brane and inside the horizon, it does not significantly affect the geometry outside the horizon. Then, we solve Eq. (4), and try to find the apparent horizon both on the brane and in the bulk. When the distribution of ρ_h is sufficiently compact, the apparent horizons exist. It should be noted that two horizons do not coincidently appear. In some cases, the apparent horizon on the brane exists although that in the bulk does not.

Here, we show one example of numerical results. We set $G_4 = 1$. In this example, an artificial matter is put for $0 \leq r \leq 0.2R$ and $1 \leq z \leq 1.2$. Equation (4) is solved using a uniform grid with grid size 1200×1200 for r and z directions, which covers a domain with $0 \leq r/R \leq 17.1$ and $1 \leq z \leq 18.1$. In this case, the gravitational mass on the brane is $M \simeq 0.29R$, and both apparent horizons on the brane and in the bulk exist. We note that the results are essentially the same for $0.25 \leq M/R \leq 0.5$. In Fig. 1, we show the location of apparent horizons in the bulk and on the brane. The apparent horizon in the bulk is apparently cigar-shaped. Due to this cigar-shape the circumferential radius of the apparent horizon is different depending on the choice of the circumference in the bulk. In Fig. 2, we show that the profile of $\psi - 1$ on the brane. For $r \gg R$, $\psi - 1$ behaves as $M/2r$, implying that the solution approximately agrees with that in the 4D Einstein gravity, i.e., the bulk effect is small. However, the existence of the bulk is significant for $r \sim R$ as expected. Indeed, $\psi - 1$ deviates from $M/2r$ with decreasing r. This effect is in particular important for the location and area of the apparent horizon on the brane: In the case of 4D gravity without bulk, the apparent horizon is located at $r_{AH} = M/2$ with the area $A_{AH} = 16\pi M^2$. However, in the brane world model, they take different values in general. (In this example, $r_{AH} \simeq 0.9M$ and $A_{AH} \simeq 88.6M^2$, and the coefficients converge to well-know 4D values (0.5 and 16π) with increasing M, implying that the effect of the existence of the bulk becomes less important.)

3 Summary

We numerically computed time symmetric initial data sets of a black hole in the brane world model, assuming that the black hole is spherical on the brane. As has been expected, the black hole (apparent horizon) is cigar-shaped in the bulk [10].

We remind that we only present time symmetric initial data of a black hole space. This implies that the black hole is not static and will evolve to other state with time evolution. The quantitative features of the final fate could be different from the present result. Self-consistent analysis for static black holes should be carried out for future to obtain a definite answer with regard to black holes in the brane world. However, we believe that the present result provides us a guideline for such future works.

We thank B. Carter, N. Dadhich, D. Langlois, R. Maartens, K. Maeda, M. Sasaki and T. Tanaka for discussions. TS is grateful to thank Relativity Group for their hospitality at Golm near by Potsdam. MS gratefully acknowledges support by JSPS and the hospitality at the Department of Physics of University of Illinois at Urbana-Champaign.

A The essence of the brane world

We briefly review the covariant formalism of the brane world[9]. For the matter source of the 5D Einstein equation, ${}^{(5)}G_{\mu\nu} = \kappa_5^2({}^{(5)}T_{\mu\nu} - \Lambda {}^{(5)}g_{\mu\nu})$, we choose the energy-momentum tensor as

$$ {}^{(5)}T_{\mu\nu} = \delta(\chi)[-\lambda q_{\mu\nu} + {}^{(4)}T_{\mu\nu}] + {}^{(5)}\tau_{\mu\nu}, \tag{10} $$

where $\chi = \ell \ln z$, λ is the tension of the brane, $q_{\mu\nu}$ is the induced metric on the brane, and ${}^{(4)}T_{\mu\nu}$ is the energy momentum tensor on the brane. Due to the singular source at $\chi = 0$ and the Z_2 symmetry, we can derive the Israel's junction condition at $\chi = 0$ as

$$ K_{\mu\nu} = -\frac{1}{6}\kappa_5^2 \lambda q_{\mu\nu} - \frac{1}{2}\kappa_5^2 \Big({}^{(4)}T_{\mu\nu} - \frac{1}{3} q_{\mu\nu} {}^{(4)}T_{\sigma}{}^{\sigma} \Big), \tag{11} $$

where $K_{\mu\nu} = q_\mu^\sigma q_\nu^\lambda D_\sigma n_\lambda$, and D_σ and n^μ are the covariant derivative with respect to $q_{\mu\nu}$ and the unit spacelike normal vector to the brane. In the text, we consider the cases in which ${}^{(4)}T_{\mu\nu} = 0$. Using (4+1) formalism, the effective 4D equation on the brane has the form

$$ {}^{(4)}G_{\mu\nu} = -\Lambda_4 q_{\mu\nu} - E_{\mu\nu}, \tag{12} $$

where ${}^{(4)}G_{\mu\nu}$ is the 4D Einstein tensor on the brane,

$$ \Lambda_4 = \frac{1}{2}\kappa_5^2 \Big(\Lambda + \frac{1}{6}\kappa_5^2 \lambda^2 \Big) \quad \text{and} \quad E_{\mu\nu} = {}^{(5)}C_{\mu\rho\nu\sigma} n^\rho n^\sigma, \tag{13} $$

where ${}^{(5)}C_{\mu\rho\nu\sigma}$ is 5D Weyl tensor. In the above, for simplicity, we set ${}^{(5)}\tau_{\mu\nu} = 0$. Equation (12) implies that we can consider $E_{\mu\nu}$ as the effective source term of the 4D Einstein equation on the brane, and as long as $E_{\mu\nu}$ is not vanishing, the geometry on the brane is different from that in the 4D gravity even in the vacuum case. Only for very special case such as for the black string solution [10, 14], $E_{\mu\nu} = 0$ holds.

From Eq. (12), we find that the Minkowski spacetime is realized on the brane when $E_{\mu\nu} = 0$ and $\Lambda_4 = 0$. In this paper, we set $\Lambda_4 = 0$ to focus on asymptotically flat brane. Then, the junction condition at $\chi = 0$ is rewritten to $K_{\mu\nu} = -\frac{1}{\ell} q_{\mu\nu}$. In the case when we choose the line element as Eq. (3), the junction condition reduces to Eq. (5).

B Asymptotic boundary conditions

To specify the boundary condition at infinities, we investigate the linearized equation of Eq. (4):

$$ \varphi'' + \frac{2}{r}\varphi' + \frac{1}{R^2}\Big(\partial_z^2 \varphi - \frac{3}{z}\partial_z \varphi \Big) = -\frac{\kappa_5^2}{4}\rho_h, \tag{14} $$

where $\psi = 1 + \varphi$ and $\varphi \ll 1$. We can obtain the formal solution with aid of the Green function $G(x, z; x', z')$ as

$$\varphi \simeq -2\pi G_4 \ell \int d^3x' dz' G(x, z; x', z') \rho_h(x', z'). \tag{15}$$

Assuming that ρ_h is non-zero only in the small region around the brane, we can derive the relevant Green function as[5]

$$\begin{aligned} G(x, z; x', z') = & -\int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \\ & \times \left[\frac{1}{\ell \mathrm{k}^2} + \int_0^\infty dm \frac{u_m(z) u_m(1)}{\mathrm{k}^2 + m^2}\right] \\ = & \; G_0 + G_{\mathrm{KK}}, \end{aligned} \tag{16}$$

where $u_m(z)$ is the mode function given from the Bessel functions J_n and N_n as

$$\begin{aligned} u_m(z) = & \; z^2 \sqrt{\frac{mR^2}{2\ell}} \\ & \times \frac{J_1(mR) N_2(mRz) - N_1(mR) J_2(mRz)}{\sqrt{(J_1(mR))^2 + (N_1(mR))^2}}, \end{aligned} \tag{17}$$

where $R = (2/3)^{1/2}\ell$. G_0 and G_{KK} are the Green function of zero and KK modes, respectively. From Eq. (15) we can derive the asymptotic boundary conditions shown in the text.

C Apparent horizon in the bulk

We derive the equation for the apparent horizon in the bulk. After we perform the coordinate transformation from (r, z) to (x, θ) as $z = 1 + x|\cos\theta|$ and $r = \ell x \sin\theta$, the surface of the apparent horizon is denoted by $x = h(\theta)$. Then, the non-zero components of s_i is written as

$$s_x = C \quad \text{and} \quad s_\theta = -C h_{,\theta}, \tag{18}$$

where $C[\equiv \psi^2 \hat{C}/(1 + x|\cos\theta|)]$ is a normalization constant calculated from $s^i s_i = 1$, and $h_{,\theta} = dh/d\theta$. Then, the equation for h can be written to the following ordinary differential equation of second order

$$\begin{aligned} \frac{d^2h}{d\theta^2} = & \; \frac{h^2}{\psi^4 \hat{C}^2}\left[\left(4\frac{\partial_x \psi}{\psi} + \frac{3}{h(1 + h|\cos\theta|)} + \frac{\partial_x \hat{C}}{\hat{C}}\right)\right. \\ & \times \left(\sin^2\theta + \psi^4 \cos^2\theta - (1 - \psi^4)\sin\theta\cos\theta \frac{h_{,\theta}}{h}\right) \\ & + h^{-1}\left(4\frac{\partial_\theta \psi}{\psi} + 3\frac{h\sin\theta}{1 + h|\cos\theta|} + 2\cot\theta + D\right) \\ & \times \left((1 - \psi^4)\sin\theta\cos\theta - (\cos^2\theta + \psi^4 \sin^2\theta)\frac{h_{,\theta}}{h}\right) \\ & + 4\psi^3 \partial_x \psi (\cos^2\theta + h^{-1}\sin\theta\cos\theta h_{,\theta}) \\ & + h^{-2}(1 - \psi^4)\sin\theta\cos\theta h_{,\theta} \\ & + h^{-1}(1 - \psi^4)\cos(2\theta) - 4h^{-1}\sin\theta\cos\theta\psi^3\partial_\theta\psi \\ & \left. + \{(1 - \psi^4)\sin(2\theta) - 4\sin^2\theta\psi^3\partial_\theta\psi\}\frac{h_{,\theta}}{h^2}\right], \end{aligned} \tag{19}$$

where

$$\begin{aligned} D = -\hat{C}^2[& \; (1 - \psi^4)\{1 - h^{-2}h_{,\theta}^2\}\sin\theta\cos\theta \\ & - h^{-1}(1 - \psi^4)\cos(2\theta)h_{,\theta} \\ & + 2\psi^3\partial_\theta\psi(\cos\theta + h^{-1}\sin\theta h_{,\theta})^2]. \end{aligned} \tag{20}$$

Eq. (19) is solved imposing boundary conditions at $\theta = 0$ and $\pi/2$. In the limit $\theta \to 0$, we impose the following boundary condition,

$$h = h_0 + h_2\theta^2 + O(\theta^3), \tag{21}$$

where h_2 is evaluated at $x = h_0$ and $\theta = 0$ from the following equation;

$$h_2 = \frac{h_0^2}{6}\left[\frac{8\partial_x\psi}{\psi} + \frac{3}{h_0(1+h_0)} + \frac{\partial_x\hat{C}}{\hat{C}} + \frac{3}{h_0}(1-\psi^4)\right]. \tag{22}$$

At $\theta = \pi/2$, the boundary condition is imposed as $h_{,\theta} = 0$.

Note that in the limit $\theta \to \pi/2$ (i.e., on the brane), Eq. (19) is written in the form

$$\frac{d^2h}{d\theta^2} = h + \frac{\ell^2h^2}{\psi^4}\left(\frac{4\partial_x\psi}{\psi} + \frac{2}{h}\right), \tag{23}$$

where we use $h_{,\theta} = 0$ and the relation $\partial_\theta\psi = D = \partial_x\hat{C} = 0$. Note that the equation which the apparent horizon on the brane satisfies is $4\partial_x\psi/\psi + 2/h = 0$ [cf., Eq. (2.8)]. Thus, unless $d^2h/d\theta^2 = h$ at $\theta = \pi/2$, the apparent horizon on the brane cannot coincide with that in 4D space. Note that the black string solution [10, 14] exceptionally satisfies $d^2h/d\theta^2 = h$ at $\theta = \pi/2$.

References

[1] P. Horava and E. Witten, Nucl. Phys. **B460**, 506 (1996); *ibid* **B475**, 94 (1996).

[2] For earlier work on this topic, see V. A. Rubakov and M. E. Shaposhinikov, Phys. Lett. **152B**, 136 (1983); M. Visser, Phys. Lett. **159B**, 22 (1985); M. Gogberashvili, Mod. Phys. Lett. **A14**, 2025 (1999).

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

[4] L. Randall and R. Sundrum, Phys. Rev. Lett. **83**,4690 (1999).

[5] J. Garriga and T. Tanaka, Phys. Rev. Lett. **84**, 2778 (2000).

[6] M. Sasaki, T. Shiromizu and K. Maeda, Phys. Rev. D **62**, 024008 (2000).

[7] S. B. Giddings, E. Katz, and L. Randall, JHEP **0003**, 023 (2000).

[8] A. Chamblin and H. S. Reall, Nucl. Phys. **B562**,133 (1999); T. Nihei, Phys. Lett. **B465**, 81 (1999); N. Kaloper, Phys. Rev. D **60**, 123506 (1999); H. B. Kim and H. D. Kim, Phys. Rev. D **61**, 064003 (2000); P. Binetruy, C. Deffayet, U. Ellwanger and D. Langlois, Phys. Lett. **B477**, 285 (2000); P. Kraus, JHEP **9912**, 01 (1999); S. Mukohyama, Phys. Lett. **B473**, 241 (2000); D. Ida, gr-qc/9912002; J. Garriga and M. Sasaki, Phys. Rev. **D62**, 043523 (2000); S. Mukohyama, T. Shiromizu and K. Maeda, Phys. Rev. D **62**, 024028 (2000); P. Bowcock, C. Charmousis and R. Gregory, hep-th/0007177.

[9] T. Shiromizu, K. Maeda and M. Sasaki, Phys. Rev. D **62**, 024012 (2000).

[10] A. Chamblin, S. W. Hawking and H. S. Reall, Phys. Rev. D **61**, 065007 (2000).

[11] W. Israel, Nuovo. Cimento. **44B**, 1 (1996).

[12] We note that in the brane world model, photons can propagate only on the brane, so that the apparent horizon on the brane is with regard to the photons. On the other hand, the apparent horizon in full 4D space is with regard to gravitons which can propagate in full 5D spacetime.

[13] For example, M. Shibata, Phys. Rev. D **55**, 2002 (1997).

[14] R. Gregory, hep-th/0004101.

Gravitational waves in cosmological models of Hořava - Witten theory

Osamu Seto [1]

Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

Abstract

We study the behavior of gravitational waves in cosmological solutions of the five-dimensional Hořava-Witten theory. As a dynamical background, we consider two cosmological models in which flat FRW branes are embedded. We show that, on these dynamical backgrounds, the tensor perturbations can be decomposed into two families of modes: essentially massive modes and the modes which describe massless graviton on the brane when the orbifold space is static. Then, we explicitly solve the evolution equations for the tensor perturbations on these dynamical backgrounds. We give the relation between the dynamics of the background brane-worlds and the late time behavior of the tensor perturbations; the evolution of the massive modes depends on the difference between the expansion rate of the four-dimensional FRW brane and that of the orbifold space.

1 Introduction

Over the past few years, a considerable number of studies have been made on brane-world scenario in which our universe is realized as a boundary of a higher dimensional spacetime [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]. In particular, inspired by recent progress of heterotic M-theory [2], five-dimensional brane-world models in which 3-branes are embedded in an effective five-dimensional spacetime compactified on S^1/Z_2 [3, 4] have attracted much attention. For example, the solutions of cosmological 3-branes in the five-dimensional Hořava-Witten theory have been discovered [5, 6, 7, 8, 9]. The five-dimensional models of Randall and Sundrum, which are proposed to solve the Hierarchy problem [12] and to demonstrate an alternative to compactification [13], also have many similarities to the five-dimensional Hořava-Witten theory.

In the brane-world scenario, all ordinary matter fields are confined on the brane, while graviton can propagate in the fifth-dimension. Hence, in order to test the idea of the brane-world, one needs to study the nature of the gravity in this scenario. The behavior of linearized gravity in the Randall - Sundrum models has recently been studied in detail. It has been shown that massless modes of the metric perturbations are decoupled from the massive modes and the Einstein gravity is recovered at low energy scales (See, *e.g.* [12, 13, 14, 15]). However these investigations have been done only in the highly symmetric background models such that four-dimensional maximally symmetric branes, *i.e.*, Minkowski brane [12, 13] and de Sitter brane [15], are embedded in a five-dimensional anti-de Sitter spacetime, which also is locally maximally symmetric in five-dimensions.

On the other hand, many people have discussed the possibility of a homogeneous and isotropic Friedmann-Robertson-Walker (FRW) brane-world in the Hořava-Witten theory [5, 6, 7] and in Randall and Sundrum's scenario [16]. However, there have so far been few attempts to investigate the perturbations on such a dynamical cosmological brane-world, although the four-dimensional real universe has inhomogeneous fluctuations as is shown by the observations of the CMB anisotropy.

As pointed out in Ref. [17], it is expected that all gravitons on the brane become massive in a dynamical brane-world model. Then, the excitation of massive graviton would become a crucial defect in the brane-world scenario, or provide a new model of dark matter in the brane-world cosmology. It is therefore important to study the evolution of perturbations on a dynamical brane model in order to explore the cosmological consequences of the brane-world idea.

[1] E-mail:oseto@yukawa.kyoto-u.ac.jp

Recently, only the formalism for cosmological perturbations on brane-world has been developed by several authors [17, 18, 19, 20, 21, 22]. For general cosmological brane-world models, the evolution equations for the metric and matter perturbations in the bulk and the boundary conditions for them at the brane have been established by Kodama *et al.*[17] in terms of gauge-invariant variables. The perturbations on a brane are inevitably coupled to the perturbations on the bulk. The evolution equations and boundary conditions for cosmological perturbations, in particular, the scalar and vector perturbations on the cosmological brane are too complicated to solve. On the other hand, as shown in [17], when the anisotropic stress perturbation vanishes, the tensor perturbations decouple from the matter perturbation and the boundary condition becomes a Neumann type. Hence, as far as the tensor perturbations are concerned, the problem is easier to deal with.

In this paper, as a first step to investigate the perturbations on a dynamical brane-world, we shall study the behavior of the tensor perturbations on the cosmological brane-world models in five-dimensional Hořava-Witten theory. In general, the equations for perturbations are no longer separable for a dynamical 3-brane which is not maximally symmetric as a hypersurface. Fortunately, for the cosmological Hořava-Witten models found in [5, 6], we will see that the evolution equation of tensor perturbations becomes separable. However the evolution equation is still rather complicated to solve exactly. So, in this paper, we shall consider the late time behavior of the tensor perturbation by using WKB approximation to analyse the evolution equation. We show that the evolution of massive modes depends on the difference between the expansion rate of the four-dimensional FRW brane and that of the orbifold space. We see that, in one of the background models, the amplitude of the massive modes decreases faster than that of the modes which correspond to the massless modes when the orbifold space is static.

This paper is organized as follows. In the next section we briefly recapitulate the five-dimensional Hořava-Witten theory and the cosmological solutions which are used as the background for perturbation. In Sec. III we give the equations of motion for tensor perturbation and the boundary condition for them. Then, we analyse the behavior of the tensor perturbations by using WKB method. Section IV is devote to conclusion and discussion.

2 Cosmological solutions of Hořava - Witten theory

Hořava and Witten have shown that the strongly coupled limit of the $E_8 \times E_8$ heterotic string theory has been identified with M-theory compactified on a S^1/Z_2 orbifold with E_8 gauge fields on each orbifold fixed plane [2]. After a compactification on a Calabi-Yau three-fold, the fields of the standard model can be confined to 3-brane [3]. Thus one has an effective model in which our four-dimensional universe is a 3-brane embedded in an effective five-dimensional spacetime compactified on S^1/Z_2. The five-dimensional effective action of Hořava - Witten theory is then given by [4]:

$$\begin{aligned} S &= \frac{1}{2\kappa_5^2}\left[\int_{\mathcal{M}_5}\sqrt{-g}\left(\frac{1}{2}R - \frac{1}{2}(\partial\phi)^2 - e^{-\sqrt{2}\phi}\partial\xi\cdot\partial\bar{\xi} - \frac{1}{6}\alpha^2 e^{-2\sqrt{2}\phi}\right)\right. \\ &\left. +\sqrt{2}\int_{M_4^{(1)}}\sqrt{-g}\alpha e^{-\sqrt{2}\phi} - \sqrt{2}\int_{M_4^{(2)}}\sqrt{-g}\alpha e^{-\sqrt{2}\phi}\right], \end{aligned} \tag{1}$$

where κ_5 is the five-dimensional gravitational constant, α is a constant, ϕ is defined so that $e^{\sqrt{2}\phi}$ measures the deformation of Calabi-Yau manifold, ξ is a scalar field related to the components of the three form on the Calabi-Yau space, and $\mathcal{M}_5$ is the five dimensional spacetime bounded by the branes $\mathcal{M}_4^{(1)}$ and $\mathcal{M}_4^{(2)}$

The cosmological solutions with flat FRW branes have been constructed by Lukas *et.al.* [5] and been generalized to include closed and open FRW branes by Reall [6]. In the present paper, as a simple case, we shall consider only the model that expanding flat FRW branes are embedded in the five-dimensional bulk with vanishing ξ field.

Let y be a coordinate in the orbifold direction with $y \in [-\pi\rho, \pi\rho]$ and Z_2 acting on S^1 by $y \to -y$. The orbifold fixed planes are located at $y = 0, \pi\rho$. Then, starting from the effective five-dimensional action (1) with the metric ansatz:

$$ds^2 = a_0^2 e^{2A(t)}C(y)(-dt^2 + \delta_{ij}dx^i dx^j) + e^{2B(t)}D(y)dy^2\,, \tag{2}$$

one can obtain the cosmological solutions of five-dimensional Hořava - Witten theory with flat FRW branes [5]:

$$C(y) = D(y)^{1/4} = \frac{\sqrt{2}}{3}\alpha|y| + c_0\,, \tag{3}$$

$$e^{2A} = t^{1-\delta}, \quad e^{2B} = t^{2\delta}, \tag{4}$$

where a_0 and c_0 are constants and $\delta = \pm\sqrt{3}/2$. Hereafter we shall refer to the upper and lower choices of sign as the $(\uparrow)$ and $(\downarrow)$ solutions, respectively.

The $(\downarrow)$ solution describes the model that the four-dimentional FRW universe expands while the orbifold space shrinks. On the other hand, the $(\uparrow)$ solution describes the model in which both the four-dimentional FRW universe and the orbifold space expand, and the latter expands faster than the former does.

For a while, we shall focus our attention on the four-dimensional brane at $y = 0$. In terms of the cosmic proper time τ defined by

$$\tau := a_0\sqrt{c_0}\int e^A dt = \frac{2a_0\sqrt{c_0}}{3-\delta}t^{(3-\delta)/2}. \tag{5}$$

The Hubble parameter is given by

$$H(t) := \frac{1}{a}\frac{da}{d\tau} = \frac{1-\delta}{2a_0\sqrt{c_0}}t^{-(3-\delta)/2}\,. \tag{6}$$

Then the wave-number k_H whose wavelength corresponds to the horizon scale is given by

$$k_H := aH = \frac{1-\delta}{2t}\,. \tag{7}$$

The scale factor of the four-dimensional FRW brane is written as

$$a(\tau) = a_0\sqrt{c_0}\left(\frac{3-\delta}{2a_0\sqrt{c_0}}\tau\right)^{(1-\delta)/(3-\delta)}\,. \tag{8}$$

Comparing this scale factor with that of the no-extra-dimension cosmology, we can say that the brane expands as if it were a standard four-dimensional flat FRW universe which contains a perfect fluid obeying the equation of state $p = w\rho$,

$$w = \frac{3+\delta}{3(1-\delta)}\,, \tag{9}$$

although the brane-world solutions considered here are vacuum solutions. Considering in this picture, we can say that the FRW brane in the $(\uparrow)$ solution looks like an unphysical universe because the dominant energy condition violates, $w > 1$. On the other hand, in the $(\downarrow)$ case, the FRW brane describes a physical universe in the sense that the dominant energy condition holds, $0 < w \approx 0.37 < 1$.

3 Behavior of tensor perturbations

3.1 Equations of motion for tensor perturbations

The gravitational perturbations are decomposed into components of the three types, scalar, vector and tensor, and each component can be expanded by tensor harmonics of the same type on the 3-space of constant curvature. Then, the tensor perturbations represent gravitatinal wave modes in the four-dimentional FRW brane.

As far as the action (1) is considered, the anisotropic stress perturbations both in the bulk and in the brane vanish. Then, for the background metric Eq.(2) and (3), the equation of motion for the tensor perturbations and the boundary condition are given, respectively, by [17]

$$\ddot{H}_T + \left(2\dot{A}(t) + \dot{B}(t)\right)\dot{H}_T - \frac{a_0^2 e^{2(A(t)-B(t))}}{C(y)^3}H_T'' + k^2 H_T = 0, \tag{10}$$

$$H_T{}' = 0, \quad at\ y = 0, \pi\rho, \tag{11}$$

where H_T is the expansion coefficient of the metric perturbations, $\delta g_{ij} = 2a_0^2 e^{2A} C(y) H_T T_{ij}$, of tensor type by harmonic tensor T_{ij} on the flat 3-space [17]. Here, $-k^2$ is an eigenvalue of the d'Alembertian on the flat 3-space, and dots and primes denote derivatives with respect to t and y, respectively.

Note that the boundary condition (11) is simply written as the derivative with respect to y. On the other hand, the evolution equation (10) contains an additional friction term $\dot{B}\dot{H}_T$, which ceases to exist when the background orbifold space is static, and the y-derivative term (the third term), which gives graviton's mass as we shall see below. So, these two terms of Eq.(10) reflect the effect of the dynamics of the background branes model on the gravitational waves.

Provided that $H_T(t, y) = T(t)Y(y)$, the equations of motion (10) are reduced to the set of equations for $Y(y)$ and $T(t)$, respectively:

$$Y_l'' + \left(\frac{m}{a_0}\right)^2 C(y)^3 Y_l = 0, \tag{12}$$

$$\ddot{T}_l + \left(2\dot{A}(t) + \dot{B}(t)\right)\dot{T}_l + k^2 T_l + m^2 e^{2(A(t)-B(t))} T_l = 0, \tag{13}$$

where the dimensionless constant m is defined as follows:

$$m^2 := \left(\frac{\sqrt{2}\alpha l a_0}{3}\right)^2 . \tag{14}$$

Here, l represents the revel of inhomogeneity of the orbifold direction and m takes the discreate values labelled by l, as we shall see in the next subsection. Note that the signature of the second term in LHS of Eq.(12) is chosen so that the solutions of Eq.(12) satisfy the boundary condition (11).

In the case of $\dot{B} = 0$, the last term in Eq.(13) provides the eigenvalue (times e^{2A}/a_0^2) of the d'Alembertian on the four-dimensional brane. Therefore the $m = 0$ modes behave the massless mode in the brane when the orbifold space is static. In this sense, we shall refer to $m = 0$ ($l = 0$) and $m \neq 0$ ($l \neq 0$) modes as "massless" and "massive" modes, respectively.

3.2 Solutions of y-dependent part

In this subsection, we shall give the solutions of Eq.(12) which satisfy the boundary condition (11).

For $m = 0$ modes, Eq.(12) reads

$$Y_l(y) = C_1 + C_2\, y, \tag{15}$$

where C_1 and C_2 are integration constants. From the boundary condition (11), we find that C_2 must vanish. Therefore, the zero-mode($l = 0$) solution for y-dependent part is

$$Y_l(y) = C_1. \tag{16}$$

For $m \neq 0$ modes, the solutions of Eq.(12) for $y \geq 0$ are given by

$$Y_l(y) = \left(\frac{\sqrt{2}}{3}\alpha y + c_0\right)^{1/2} \left[C_1 H^{(1)}_{1/5}\left(\frac{2l}{5}\left(\frac{\sqrt{2}}{3}\alpha y + c_0\right)^{5/2}\right) + C_2 H^{(2)}_{1/5}\left(\frac{2l}{5}\left(\frac{\sqrt{2}}{3}\alpha y + c_0\right)^{5/2}\right)\right], \tag{17}$$

where C_1 and C_2 are constants. The solutions for the whole region is given by replacing y with $|y|$ according to the Z_2 symmetry.

From the boundary condition at $y = 0$, we find that the ratio of C_1 and C_2 becomes

$$\frac{C_2}{C_1} = -\left.\frac{H^{(1)}_{-4/5}(z)}{H^{(2)}_{-4/5}(z)}\right|_{z=2lc_0^{5/2}/5} \tag{18}$$

for each l. From the boundary condition at $y = \pi\rho$, we obtain

$$\left.\frac{H^{(1)}_{-4/5}(z)}{H^{(2)}_{-4/5}(z)}\right|_{z=2lc_0^{5/2}/5} = \left.\frac{H^{(1)}_{-4/5}(z)}{H^{(2)}_{-4/5}(z)}\right|_{z=2l(\sqrt{2}\alpha\pi\rho/3+c_0)^{5/2}/5} . \tag{19}$$

This gives the value of l for each excited mode.

For the case of $2lc_0^{5/2}/5 \gg 1$, l can be writen by

$$l \approx \frac{5}{2} \frac{n\pi}{\left(\sqrt{2}\alpha\pi\rho/3 + c_0\right)^{5/2} - c_0^{5/2}} \quad , \tag{20}$$

where $n = 1, 2, ...$. In the case of $c_0 \gtrsim \mathcal{O}(1)$, Eq.(20) reads

$$m \approx \frac{a_0 n}{c_0^{3/2}\rho} \quad , \tag{21}$$

because it is assumed that $\alpha\pi\rho \ll 1$ is satisfied in the context of five-dimensional Hořava-Witten theory, except for inflationary epoch [8]. We can see that the reduced KK "mass" defined by,

$$m_{KK,l} := \frac{m}{a_0 e^{B(t)}\sqrt{C(0)}}, \tag{22}$$

takes a typical value of the orbifold energy scale.

The norm squared of $Y_l(y)$ is then given by

$$|Y_l(y)|^2 \approx \cos^2 \frac{2l}{5} \left(\left(\sqrt{2}\alpha|y|/3 + c_0\right)^{5/2} - c_0^{5/2} \right). \tag{23}$$

Therefore $|Y_l(y)|^2$ has peak at each boundary (at $y = 0$ and $y = \pi\rho$).

3.3 Solutions of t-dependent parts

In this subsection, we shall deal with Eq.(13). For simplicity, hereafter we omit the saffix l.

For the $m = 0$ modes, we have the general solutions to Eq.(13),

$$T = C_1 H_0^{(1)}(kt) + C_2 H_0^{(2)}(kt), \tag{24}$$

where C_1 and C_2 are constants.

For $m \neq 0$ modes we analyse Eq.(13) by means of the WKB methods because it is too hard to solve Eq.(13) exactly. The equation of motion (13) is rewritten as

$$\left(t^{1/2}T\right)^{\cdot\cdot} + S(t)^2 t^{1/2} T = 0, \tag{25}$$

where $S(t)$ is defined by,

$$S(t) := \sqrt{k^2 + \frac{1}{4t^2} + m^2 t^{1-3\delta}}. \tag{26}$$

In the region where $|\dot{S}| \ll S^2$ holds, we can use WKB method and then obtain,

$$T(t) \approx C_1(k)(tS(t))^{-1/2} \exp\left[-i\int^t S(t')dt'\right] + C_2(k)(tS(t))^{-1/2} \exp\left[i\int^t S(t')dt'\right]. \tag{27}$$

Since

$$\frac{|\dot{S}|}{S^2} = \frac{2\left|-1 + 2(1-3\delta)m^2 t^{3(1-\delta)}\right|}{\left(1 + 4(kt)^2 + 4m^2 t^{3(1-\delta)}\right)^{3/2}}, \tag{28}$$

the WKB approximation is valid in the region where $t \gg m^{-2/(3(1-\delta))}$ or $t \gg 1/(2k)$. The former relation is satisfied when the time scale is larger than the scale of the orbifold radius and the latter is satisfied when the wavelength is shorter than the Horizon radius *i.e.* $k \gg k_H$.

In the ($\uparrow$) background case, when $t \gg 1/(2k)$ and $k^2 \gg m^2 t^{1-3\delta}$ hold, namely when we consider sub-horizon scale $k \gg k_H$, we see from Eq.(24) and (27) that the behavior of the $m \neq 0$ modes becomes similar to the behavior of the $m = 0$ modes.

In the (↓) background case, when we consider the region where $t \gg 1/(2k)$ hold, we have $S(t) \cong \sqrt{k^2 + m^2 t^{1-3\delta}}$ and its integration,

$$\int^t S(t')dt' = \frac{kt}{3(1-\delta)}\left[2\sqrt{1+\left(\frac{m}{k}\right)^2 t^{1-3\delta}}\right.$$
$$\left.+(1-3\delta){}_2F_1\left((1-3\delta)^{-1}, \frac{1}{2}; 1+(1-3\delta)^{-1}; -\left(\frac{m}{k}\right)^2 t^{1-3\delta}\right)\right]. \quad (29)$$

Then we obtain for $k^2 \ll m^2 t^{1-3\delta}$

$$T \approx (mt^{3(1-\delta)/2})^{-1/2}\left(C_1(k)\exp\left[-i\frac{2mt^{3(1-\delta)/2}}{3(1-\delta)} - i\theta\right] + C_2(k)\exp\left[i\frac{2mt^{3(1-\delta)/2}}{3(1-\delta)} + i\theta\right]\right), \quad (30)$$

where the phase θ arises from the hypergeometrical function in Eq.(29). Note that the late time solution (30) does not depend on the wavenumber k and the variable of (30) is propotional to $m_{KK}\tau$.

From Eq.(26) and Eq.(27), we can see that the ratio of the amplitude of the massless modes and that of the massive modes has the following time dependence,

$$\frac{T_{m\neq 0}}{T_{m=0}} \propto S(t)^{-1/2} \sim e^{(B(t)-A(t))/2} \quad (31)$$

for $m_{KK} \gg k/a$.

4 Conclusion and Discussion

We have studied the evolution of gravitational wave perturbations in the dynamical FRW brane-world models of five-dimensional Hořava - Witten theory. As the background spacetime, we have used two cosmological solutions, *i.e.* the (↑) and (↓) solutions, which contain the flat FRW branes with vanishing ξ. We showed that the "massless"($m = 0$) modes, which correspond to the massless graviton on the brane if the orbifold space is static, are decoupled from the discontinuous $m \neq 0$ modes in these background spacetimes. The $m \neq 0$ modes have peak at each brane, while that of the $m = 0$ modes are constant.

We have obtained the exact solution of the equation of the t-dependent part, for the $m = 0$ modes. The time evolution of the $m = 0$ modes in the present brane-world models is similar to that of gravitational waves in the no-extra-dimensional case. However, one may notice that the difference of the expansion of scale factor and the additional friction which is due to the dynamics of orbifold space produce the different adiabatic damping rate and the behavior of the long-wavelength limit.We analysed the late time behavior of the $m \neq 0$ modes by means of WKB method. We showed that the late time behavior of the $m \neq 0$ modes in the (↑) background solution becomes similar to that of the $m = 0$ modes. On the other hand, the behavior of the $m \neq 0$ modes in the (↓) background solution is quite different from that of the $m = 0$ modes; the "mass" of the $m \neq 0$ modes increases and the ratio of the massive and massless modes decreases as a power of conformal time.

We can understand the behavior of the $m \neq 0$ modes by observing the behavior of $A(t) - B(t)$: the difference between the expansion rate of the four-dimensional FRW brane and that of the orbifold space. We have $\lambda m_{KK,l} \propto \exp[2(A(t) - B(t))]$, where $\lambda \equiv a_0 e^{A(t)}\sqrt{C(y)}k^{-1}$ is the the reduced wavelength. When $\lambda m_{KK,l}$ is sufficiently small, we can say that the modes we consider is "massless". In the case of (↑) background, $A(t) - B(t)$ is a decreaseing function, hence in the late time $\lambda m_{KK,l} \longrightarrow 0$, and the $m \neq 0$ modes behave as if it were $m = 0$ modes. On the other hand, in the case of the (↓) background, $A(t) - B(t)$ is an increaseing function, thus the "mass" of the $m \neq 0$ modes increase, and the amplitude of the $m \neq 0$ modes are damped faster than that of the $m = 0$ modes.

acknowledgements

This work is based on the collabolation with H. Kodama. The author would like to thank A. Ishibashi for discussion and reading the manuscript.

References

[1] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, Phys. Lett. **B429**, 263 (1998); Phys. Rev. **D59**, 086004 (1999) ; I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali Phys. Lett. **B436**, 257 (1998) ; T. Nihei, *ibid.* **B465**, 81 (1999) ; Z. Kakushadze and S. H. H. Tye, Nucl. Phys. **B548**, 180 (1999) ; H. A. Chamblin and H. S. Reall, *ibid.* **B562**, 133 (1999) ; A. Chamblin and G. W. Gibbons, Phys. Rev. Lett. **84**, 1090 (2000) ; J. D. Lykken, Phys. Rev. **D54**, 3693 (1996) ; N. Kaloper, *ibid.* **D60**, 123506 (1999) ; H. B. Kim and H. D. Kim, *ibid.* **D61**, 064003 (2000) ; A. Chamblin, S. W. Hawking and H. S. Reall, *ibid.* **D61**, 065007 (2000) ; M. Sasaki, T. Shiromizu and K. Maeda, *ibid.* **D62**, 024008 (2000) ;T. Shiromizu, K. Maeda and M. Sasaki, *ibid.* **D62**, 024012 (2000) ; R. Emparan, G. T. Horowitz and R. C. Myers, JHEP **0001**, 007 (2000) ;021, (2000).

[2] P. Hořava and E. Witten, Nucl. Phys. **B460**, 506 (1996); **B475**, 94 (1996).

[3] E. Witten, Nucl. Phys. **B471**, 135 (1996).

[4] A. Lukas, B. A. Ovrut, K. S. Stelle and D. Waldram, Nucl. Phys. **B552**, 246 ; Phys. Rev. **D59**, 086001 (1999).

[5] A. Lukas, B. A. Ovrut and D. Waldram, Phys. Rev. **D60**, 086001 (1999).

[6] H. S. Reall, Phys. Rev.**D59**, 103506 (1999).

[7] U. Ellwanger, hep-th/0001126.

[8] A. Lukas, B. A. Ovrut and D. Waldram, Phys. Rev. **D61**, 023506 (1999).

[9] J. E. Lidsey, Class. Quantum. Grav. **17**, L39 (2000).

[10] M. P. Dabrowski, Phys. Lett. **B474**, 52 (2000)

[11] K. A. Meissner and M. Olechowski, Phys. Lett. **B480**, 312 (2000)

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

[13] L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 4670 (1999).

[14] J. Garriga and T. Tanaka, Phys .Rev. Lett. **84**, 2778 (2000) ; S. B. Giddings, E. Katz and L. Randall, JHEP **0003**, 023 (2000).

[15] J. Garriga and M. Sasaki, Phys. Rev. **D62**, 043523 (2000) ; S. W. Hawking, T. Hertog, and H. S. Reall, *ibid.* **D62**, 043501 (2000) ; R. Maartens, D. Wands, B. A. Bassett and I. Heard, *ibid.* **D62**, 041301 (2000) ; D. Langlois, R. Maartens and D. Wands, Phys. Lett. **B489**, ; 259 (2000).

[16] J. M. Cline, C. Grojean and G. Servant, Phys. Rev. Lett. **83**, 4245 (1999) ; D. J. H. Chung and K. Freese, Phys. Rev. **D61**, 023511 (2000) ; E. E. Flanagan, S. H. H. Tye and I. Wasserman, *ibid.* **D62**, 044039 (2000) ; C. Csaki, M. Graesser, L. Randall and J. Terning, *ibid.* **D62**, 045015 (2000) ; C. Csaki, M. Graesser, C. Kolda and J. Terning, Phys. Lett. **B462**, 34 (1999) ; P. J. Steinhardt, *ibid.* **B462**, 41 (1999) ; P. Kanti, I. I. Kogan, K. A. Olive and M. Pospelov, *ibid.* **B468**, 31 (1999) ; P. Binetruy, C. Deffayet, U. Ellwanger and D. Langlois, *ibid.* **B477**,285 (2000) ; P. Binetruy, C. Deffayet and D. Langlois, Nucl. Phys. **B565**,269 (2000) ; S. Mukohyama, *ibid.* **B473**, 241 (2000) ; D. N. Vollick, hep-th/9911181 ; P. Kraus, JHEP **9912**, 011 (1999) ; D. Ida, *ibid* **0009**, 014 (2000).

[17] H. Kodama, A. Ishibashi and O. Seto, Phys. Rev. **D62**, 064022 (2000).

[18] S. Mukohyama, Phys. Rev. **D62**, 084015 (2000).

[19] R. Maartens, Phys. Rev. **D62**, 084023 (2000).

[20] D. Langlois, hep-th/0005025.

[21] C. van de Bruck, M. Dorca, R. H. Brandenberger and A. Lukas, hep-th/0005032.

[22] K. Koyama and J. Soda, Phys. Rev. **D62**, 123502 (2000).

Static Solutions to Einstein-Born-Infeld-dilaton System in (2+1) Dimensions

Ryo Yamazaki [1] and Daisuke Ida[2]

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

Abstract

(2+1) dimensional static and circularly symmetric solutions to Einstein-Born-Infeld-dilaton system with general couplings are derived. The solutions to the string action are investigated in detail. When cosmological constant is negative, the black hole horizons exist in some parameter regions. Some differences between Born-Infeld nonlinear electrodynamics and Maxwell electrodynamics are revealed.

1 Introduction

In string theory, conformal invariance condition leads the low energy effective action [1]. Neglecting Kolb-Ramond anti-symmetric tensor field, the proposed action describes Einstein-Born-Infeld-dilaton system. As stated in [1], cosmological constant term arises when one considers 1-loop correction. So, Einstein-Born-Infeld-dilaton system with cosmological term naturally arises in string theory. Born-Infeld Lagrangian, which represents nonlinear electrodynamics[2], reduces in the low energy limit, the Lagrangian which leads ordinary Maxwell electrodynamics. So, we can see Einstein-Born-Infeld-dilaton system as a generalization of Einstein-Maxwell-dilaton system.

We investigate stringy black holes which are the solutions to Einstein-Born-Infeld-dilaton system. The mass of stringy black holes should be much smaller than that of astrophysical black holes which are formed after gravitational collapse. Kerr-Newman black hole is the unique stationary black hole solution to the Einstein-Maxwell equation. However, the situation of stringy black holes is rather different, since the black hole uniqueness theorem does not hold any longer, or, various class of black hole solutions may exist. This seems to show that it is hard to reveal the general properties of such stringy black holes. We want to manage to approach this difficult problem.

Recently, numerical studies of Einstein-Born-Infeld-dilaton system have been done in (3+1) dimensions [16]. In (3+1) dimensional spacetime, no analytic solutions to Einstein-Born-Infeld-dilaton system exist even in the simplest situations such as static and spherically symmetric case. Using numerical study, one can only investigate individual cases which are often highly symmetric, and therefore cannot reveal general properties of black holes. We investigate (2+1) dimensional Einstein-Born-Infeld-dilaton system in order to reveal the properties of stringy black holes analytically. (2+1) dimensional Einstein-Born-Infeld-dilaton system gives the model in which analytic solutions exist as shown in next section. The (2+1) dimensional gravity is a simple version of general relativity and gives one useful approach to the black hole physics, since the discovery of BTZ black hole[5]. Three dimensional black hole also arises in some higher dimensional theories. For example, the near horizon geometry of five dimensional Reissner-Nordström solution is locally $AdS_3 \times S^3 \times T^4$ [3]. We shall study the property of Born-Infeld theory by investigating (2+1) dimensional black hole spacetime as a simplest example.

This paper is organized as follows. In Sec.2, we derive basic equations for Einstein-Born-Infeld-dilaton system, and solve them assuming some conditions. In Sec 3, we analize the string solutions in detail, and investigate the differences between Born-Infeld electrodynamics and ordinary Maxwell electrodynamics. Finally, we summarize the results in Sec 4.

[1] E-mail:yamazaki@tap.scphys.kyoto-u.ac.jp
[2] E-mail:ida@tap.scphys.kyoto-u.ac.jp

2 Basic Equations and Solutions

We first consider general form of Lagrangian which describes nonlinear electrodynamics. We adopt the action which is written in string frame as

$$S = \int d^3x\sqrt{-g}e^{-2\phi}\left(R - 2\Lambda + 4(\partial\phi)^2 + \mathcal{L}[F^2]\right), \tag{1}$$

where $F^2 = F_{ij}F^{ij}$ and $\mathcal{L}$ is its functional. The action in Einstein frame is given via the conformal transformation $g_{ij} \mapsto e^{4\phi}g_{ij}$ by

$$S = \int d^3x\sqrt{-g}\left(R - 2e^{4\phi}\Lambda - 4(\partial\phi)^2 + e^{4\phi}\mathcal{L}[e^{-8\phi}F^2]\right). \tag{2}$$

We consider an action with general dilaton coupling constants α, β and $\gamma > 0$;

$$S = \int d^3x\sqrt{-g}\left(R - 2e^{\alpha\phi}\Lambda - \gamma(\partial\phi)^2 + e^{\beta\phi}\mathcal{L}[e^{-2\beta\phi}F^2]\right). \tag{3}$$

The equations of motion of the dilaton field ϕ and the electromagnetic field F are

$$2\gamma\phi_{,i}{}^{;i} - 2\alpha\Lambda e^{\alpha\phi} + \beta e^{\beta\phi}\mathcal{L} - 2\beta e^{-\beta\phi}F^2\dot{\mathcal{L}} = 0, \tag{4}$$

$$\left(e^{-\beta\phi}\dot{\mathcal{L}}F^{ij}\right)_{;j} = 0, \tag{5}$$

respectively, where $\dot{\mathcal{L}}[x] = \delta\mathcal{L}/\delta x$. The Einstein equation is

$$R_{ij} = 4\phi_{,i}\phi_{,j} + \left(2e^{\alpha\phi}\Lambda - e^{\beta\phi}\mathcal{L}\right)g_{ij} - 2e^{-\beta\phi}\dot{\mathcal{L}}(F_{ik}F_j{}^k - F^2 g_{ij}), \tag{6}$$

We consider the Born-Infeld nonlinear electrodynamics, which corresponds to

$$\mathcal{L}[x] = 4b^2\left[1 - \left(1 + \frac{x}{2b^2}\right)^{\frac{1}{2}}\right], \tag{7}$$

where $b > 0$ is the Born-Infeld parameter. In the limit of large b, this gives Maxwell Lagrangian; $\mathcal{L} \to -x$ The three-dimensional static and circularly symmetric space-time can be written in the form

$$ds^2 = -f(r)dt^2 + \frac{e^{2\delta(r)}}{f(r)}dr^2 + r^2d\varphi^2. \tag{8}$$

The electromagnetic field is assumed to have the following pure electrostatic form

$$F = e^{\beta\phi(r)+\delta(r)}E(r)dt \wedge dr. \tag{9}$$

Then, Eq. (5) becomes

$$\left(\frac{rE}{\sqrt{1 - E^2/b^2}}\right)' = 0, \tag{10}$$

where prime denotes derivative with respect to r. This can be integrated to give

$$E = h^{-\frac{1}{2}}\frac{Q}{r}, \tag{11}$$

where

$$h = 1 + \frac{Q^2}{b^2r^2} \tag{12}$$

and Q is an integration constant. The constant Q is the electric charge, namely

$$Q = -\frac{1}{4\pi}\oint_\Gamma e^{-\beta\phi}\dot{\mathcal{L}} * F \tag{13}$$

for any smooth closed space-like curve Γ enclosing $r = 0$. The equation of motion of the dilaton field (4) becomes

$$\frac{\gamma e^{-\delta}}{r}(rfe^{-\delta}\phi')' - \alpha\Lambda e^{\alpha\phi} + 2\beta b^2 e^{\beta\phi}\left(1 - h^{\frac{1}{2}}\right) = 0 \tag{14}$$

The Einstein equations are

$$\left(f'' + \frac{f'}{r} - f'\delta'\right)e^{-2\delta} = -4\Lambda e^{\alpha\phi} + 8b^2 e^{\beta\phi}\left(1 - h^{\frac{1}{2}} + \frac{Q^2}{2b^2r^2}h^{-\frac{1}{2}}\right), \tag{15}$$

$$\delta' = \gamma r(\phi')^2, \tag{16}$$

$$\frac{f' - f\delta'}{r}e^{-2\delta} = -2\Lambda e^{\alpha\phi} + 4b^2 e^{\beta\phi}(1 - h^{\frac{1}{2}}). \tag{17}$$

Here, we assume the following form of the metric function δ;

$$\delta = n\ln\left(\frac{r}{r_0}\right), \quad \text{for } r > 0 \tag{18}$$

where n and $r_0 > 0$ are constants. From Eq.(16), it turns out that n must be positive, and then the dilaton field ϕ becomes

$$\phi = \pm\left(\frac{n}{\gamma}\right)^{\frac{1}{2}}\ln\left(\frac{r}{r_1}\right) \tag{19}$$

with a positive constant r_1. Then, Eqs. (14) and (17) become

$$\left(\frac{f}{r^n}\right)' = \pm\frac{r^{n+1}}{(\gamma n)^{\frac{1}{2}}r_0{}^{2n}}\left[\alpha\Lambda\left(\frac{r}{r_1}\right)^{\pm\alpha\sqrt{n/\gamma}} - 2\beta b^2\left(\frac{r}{r_1}\right)^{\pm\beta\sqrt{n/\gamma}}(1 - h^{\frac{1}{2}})\right] \tag{20}$$

and

$$\left(\frac{f}{r^n}\right)' = \frac{r^{n+1}}{r_0{}^{2n}}\left[-2\Lambda\left(\frac{r}{r_1}\right)^{\pm\alpha\sqrt{n/\gamma}} + 4b^2\left(\frac{r}{r_1}\right)^{\pm\beta\sqrt{n/\gamma}}(1 - h^{\frac{1}{2}})\right], \tag{21}$$

respectively. These two equations (20) and (21) are consistent if and only if

$$\alpha = \beta = \mp\sqrt{4n\gamma}. \tag{22}$$

Eq. (22) includes the string case ($n = 1$, $\alpha = \beta = \gamma = 4$). When the condition (22) is satisfied, Eqs. (20) and (21) become

$$\left(\frac{f}{r^n}\right)' = \left(\frac{r_1}{r_0}\right)^{2n}\left[(4b^2 - 2\Lambda) - 4b^2h^{\frac{1}{2}}\right]r^{1-n}. \tag{23}$$

The solutions of Eq. (23) are

$$f = \left(\frac{r_1}{r_0}\right)^{2n}\left[\frac{4b^2 - 2\Lambda}{2 - n}r^2 - 4b^2r^n\int^r r^{1-n}\left(1 + \frac{Q^2}{b^2r^2}\right)^{\frac{1}{2}}dr\right]. \tag{24}$$

for $n \neq 2$, and

$$f = \left(\frac{r_1}{r_0}\right)^4\left[(4b^2 - 2\Lambda)r^2\ln\left(\frac{r}{r_2}\right) - 4b^2r^2\int^r r^{-1}\left(1 + \frac{Q^2}{b^2r^2}\right)^{\frac{1}{2}}dr\right]. \tag{25}$$

for $n = 2$.

The above solutions include known solutions as special cases. When $n = 0$, the dilaton field vanishes, and Eq. (24) becomes

$$f = -M - (\Lambda - 2b^2)r^2 - 2b^2r\sqrt{r^2 + Q^2/b^2} - 2Q^2\ln\left(r + \sqrt{r^2 + Q^2/b^2}\right), \tag{26}$$

which has been given by Cataldo and A.García [6]. While in the limit of $b \to \infty$, the Born-Infeld field reduces just to the Maxwell field, then Eq. (24) gives

$$f = \left(\frac{r_1}{r_0}\right)^{2n} \left[-Ar^n - \frac{2\Lambda}{2-n}r^2 + \frac{2Q^2}{n}\right] \tag{27}$$

for $n \neq 0, 2$,

$$f = \left(\frac{r_1}{r_0}\right)^{4} \left[-Ar^2 - 2\Lambda r^2 \ln\left(\frac{r}{r_2}\right) + Q^2\right] \tag{28}$$

for $n = 2$, and

$$f = -M - \Lambda r^2 - 2Q^2 \ln\left(\frac{r}{r_3}\right) \tag{29}$$

for $n = 0$. Eqs. (27) coincide with the solution obtined by Chan and Mann [7], and Eq. (29) corresponds to the charged BTZ solution [4]. Other solutions can be obtained by taking special values of parameters in Eq. (24) or (25). Table 1 shows the classifications of solutions found previously [4, 5, 6, 7, 8, 9, 10, 11, 12, 13]

Reference	b	Λ	α, β, γ	n	M	Q	System
Eq.(24) or Eq.(25)			$\alpha = \beta = \mp 2\sqrt{n\gamma}$				EBIDA
Eq.(30)			$\alpha = \beta = \gamma = 4$	1			EBIDA (String)
Cataldo and García(1999)			$\alpha = \beta = 0$	0			EBIA
Chan and Mann(1994)	∞		$\alpha = \beta = \mp 2\sqrt{n\gamma}$, $n \neq 2$				EMDA
McGuigan *et al.*(1992)	∞		$\alpha = \beta = \gamma = 4$	1			EMDA (String)
Mandal *et al.*(1991)	∞		$\alpha = \beta = \gamma = 4$	1		0	EDA (String)
Sa *et al.*(1996)	∞	< 0	$\alpha = \beta = n\gamma = 4$, $n \neq 1, 2$			0	EDA
Sa *et al.*(1996)	∞	< 0	$\alpha = \beta = 4, \gamma = 2$	2	0	0	EDA
Barrow *et al.*(1986)	∞	< 0	$\alpha = -\sqrt{2}$, $\gamma = 1$	1/2	0	0	EDA
Bañados *et al.*(1992)	∞		$\alpha = \beta = 0$	0			EMA
Barrow *et al.*(1986)	∞	0	$\gamma = 1$	$\neq 0$		0	ED
Deser and Mazur(1985)	∞	0	$\alpha = \beta = 0$	0			EM
Gott,III *et al.*(1986)	∞	0	$\alpha = \beta = 0$	0	0		EM
Bañados *et al.*(1993)	∞		$\alpha = \beta = 0$	0		0	EA

Table 1: Classification of solutions to (2+1) dimensional Einstein equation. Abbreviations "E","BI","M","D" and "Λ" stand for "Einstein","Born-Infeld","Maxwell", "dilaton" and "cosmological constant",respectively.

3 Solutions to the String Action

The case $n = 1$, $\alpha = \beta = \gamma = 4$ is particulary important, since this yields the string action. We shall investigate this case in detail. In this case, Eq. (24) is expressed in terms of elementary functions as

$$f(r) = \left(\frac{r}{r_0}\right)\left\{-M - \frac{2\Lambda {r_1}^2}{r_0}r + \frac{4b^2{r_1}^2}{r_0}\left[1 - \left(1 + \frac{Q^2}{b^2r^2}\right)^{\frac{1}{2}}\right]r \right.$$
$$\left. + \frac{4b|Q|{r_1}^2}{r_0}\ln\left[\left(1 + \frac{Q^2}{b^2r^2}\right)^{\frac{1}{2}} + \frac{|Q|}{br}\right]\right\}. \tag{30}$$

The integration constant M can be regarded as the total mass of the system. In fact, according to Brown and York [14][15], the (quasi-local) mass $M(r)$ within the circle of radius r becomes

$$M(r) = 2f^{\frac{1}{2}}\left(f_0^{\frac{1}{2}} - f^{\frac{1}{2}}\right)e^{-\delta}, \tag{31}$$

where $f_0 = f_0(r)$ is a function defining the zero of mass corresponding to some background space-time. It will be natural to choose

$$f_0(r) = -\frac{2\Lambda {r_1}^2}{{r_0}^2} r^2, \tag{32}$$

which implies that the Born-Infeld field vanishes in the background space-time. Then, M is the total mass in the sense $M = \lim_{r\to+\infty} M(r)$.

There is a curvature singularity at the center $r = 0$, where the scalar curvature and the energy density diverge. The asymptotic behavior near $r = 0$ of these quantities are shown in Table 2. In addition, the cases of the Einstein-Maxwell-dilaton-Λ system [7], the Einstein-Maxwell-Λ system (charged BTZ) [4], the Einstein-Born-Infeld system [6], and the Einstein-dilaton-Λ system ($Q = 0$) are also shown for comparison. It can be seen that nonlinearity of electrodynamics weakens the divergence of curvature scalars at least in the case of dilaton-coupled systems. Curvature scalars also diverge in string frame. Table 3 shows the behavior in string frame.

System	R	ρ_{EM}	ρ_D	ρ_Λ
EBID($Q \neq 0$)	$r^{-3}\ln r$	r^{-3}	$r^{-3}\ln r$	r^{-2}
EMD($Q \neq 0$)	r^{-4}	r^{-4}	r^{-4}	r^{-2}
EBI($Q \neq 0$)	r^{-1}	r^{-1}	0	r^0
charged BTZ	r^{-2}	r^{-2}	0	r^0
ED($Q = 0$)	r^{-3}	0	r^{-3}	r^{-2}

Table 2: Leading orders of scalar curvature and energy density when $r \to 0$ and $\Lambda \neq 0$ (Einstein frame).

System	R	ρ_{EM}	ρ_D	ρ_Λ
EBID($Q \neq 0$)	$r^{-1}\ln r$	r^0	$r^{-1}\ln r$	r^0
EMD($Q \neq 0$)	r^{-2}	r^0	r^{-2}	r^0
ED($Q = 0$)	r^{-1}	r^0	r^{-1}	r^0

Table 3: Leading orders of scalar curvature and energy density when $r \to 0$ and $\Lambda \neq 0$ (string frame).

A null hypersurface on which the horizon function $f(r)$ vanishes is a Killing horizon. The behavior of the function $f(r)$ determines the causal structure of the space-time. Introduce the following dimensionless quantities ($Q \neq 0$)

$$x \equiv \frac{b}{|Q|} r, \quad q \equiv \frac{2r_1 Q}{r_0}, \quad m \equiv \frac{r_0 M}{4br_1^2|Q|}, \quad \lambda \equiv \frac{\Lambda}{2b^2}. \tag{33}$$

Then, Eq. (30) can be rewritten as

$$f/q^2 = x\left\{-m + (1-\lambda)x - (1+x^2)^{\frac{1}{2}} + \ln\left[\frac{1+(1+x^2)^{\frac{1}{2}}}{x}\right]\right\}. \tag{34}$$

It can be seen that the causal structure is determined by essentially two parameters m and λ.

For any m and λ, the behavior of f/q^2 near $x = 0$ is

$$f/q^2 \sim -x\ln x (> 0) \quad as \quad x \to 0 \tag{35}$$

Therefore, $r_* \to finite$ as $x \to 0$, which implies the curvature singularity $r = 0$ is timelike. On the other hand, it can be seen that $r_* \to \infty$ as $x \to \infty$ for any m and λ. This implies that spatial infinity $r = \infty$ is null.

The behavior of f/q^2 at spatial infinity depends on m and λ. So, the causal structure are classified into several cases.

1. $\lambda > 0$
   In this case, for any m, $f/q^2 \sim -\lambda x^2$ as $x \to \infty$. Furthermore, $f/q^2 = 0$ has one positive root. So there exists cosmological horizon. These properties do not depend on m.

2. $\lambda = 0$
   In this case, the causal structure depend on m. If m is negative, there exist cosmological horizon, and causal structure is same as $\lambda > 0$ case. However, when m is positive or zero, $f > 0$ for any $x > 0$. Therefore, the singularity $r = 0$ is naked.

3. $\lambda < 0$
   In this case, too, the causal structure depend on m. The behavior of f/q^2 at spatial infinity is same as $\lambda > 0$ case except for the signature. When m satisfies

$$m = m_*(\lambda) \equiv \ln\left(1 - \lambda + \sqrt{\lambda(\lambda - 2)}\right), \tag{36}$$

   the curve $f = f(x)$ touches x-axis. Whether the curvature singularity $r = 0$ is covered by black hole horizon depends on whether m exceeds m_* or not.

   (a) $m < m_*$
   In this case, $f > 0$ for any $x > 0$. This implies that $r = 0$ is naked.

   (b) $m = m_*$
   This is the extremal case. The black hole horizon are degenerated, and causal structure is same as (3+1) dimensional extremal Reissner-Nordström spacetime. The horizon radius can be calculated as $x_{ext} = [\lambda(\lambda - 2)]^{-\frac{1}{2}}$.

   (c) $m > m_*$
   $f/q^2 = 0$ has two positive roots. These correspond inner Cauchy horizon and black hole horizon, respectively. The causal structure is same as (3+1) dimensional non-extremal Reissner-Nordström spacetime.

We now show the effect of nonlinear electrodynamics. We rewrite Eq.(36) in terms of original physical quantities. Black hole horizons can exist if and only if $\Lambda < 0$ and

$$\frac{|Q|}{M} \leq \left(\frac{|Q|}{M}\right)_* \equiv \frac{r_0}{4r_1^2 b}\left(\ln\left[1 - \frac{\Lambda}{2b^2} + \sqrt{\frac{\Lambda}{2b^2}\left(\frac{\Lambda}{2b^2} - 2\right)}\right]\right)^{-1} \tag{37}$$

To compare with Einstein-Maxwell-dilaton system [7], we take Maxwell limit of (37) with the other parameters fixed ;

$$\left(\frac{|Q|}{M}\right)_* \to \frac{r_0}{4r_1^2}|\Lambda|^{-\frac{1}{2}} \quad (b \to \infty). \tag{38}$$

For finite value of $b(> 0)$, $\left(\frac{|Q|}{M}\right)_*$ is always larger than this limit. Therefore, there is a certain set of parameters M and Q for which black hole horizons exist in Einstein-Born-Infeld-dilaton system but do not exist in Einstein-Maxwell-dilaton system.

4 Conclusion

We have found static, circularly symmetric, and pure electrically charged solutions of (2+1) dimensional Einstein-Born-Infeld-dilaton system which is inspired by string theory. Electric field E does not diverge, which is one of the main effects of Born-Infeld nonlinear electrodynamics.

Taking a certain parameter limit, we have reproduced the solutions which have already been obtained. If we take $b \to \infty$ limit with the other parameters fixed, our solutions reduce to the solutions to Einstein-Maxwell-dilaton system [7]. And we have reproduced the solutions of Einstein-Born-Infeld system [6] if we impose the condition that the dilaton field vanishes.

We have investigated in detail the solutions to the string theory ($n = 1$ case). The solutions have six parameters r_0, r_1, Q, M, b and Λ. However, the solutions are governed by two non-dimensional parameters m and λ, which mean normalized charge-to-mass ratio and cosmological constant, respectively. The solutions have no black hole horizons if $\lambda \geq 0$, while in this case cosmological horizons exist. When $\lambda < 0$ and $m \geq m_*$, there exists a black hole horizon. When $\lambda < 0$ and $m < m_*$, black hole horizon does not exist. Causal structures of charged black hole solutions ($\lambda < 0$ case) are similar to those of (3+1) dimensional Reissner-Nordström one. As stated in [7], this is an effect of existence of a dilaton field which does not approach a constant at infinity.

We revealed two differences between Born-Infeld nonlinear electrodynamics and linear Maxwell electrodynamics. The first is about curvature singularity. Curvature scalar and energy density diverge as $r \to 0$ in both Einstein frame and string frame. We showed that in the dilaton-coupled system, the divergence of curvature scalars is weaker in the case of Born-Infeld electrodynamics than of Maxwell electrodunamics. Secondly, the effect of nonlinearity of Born-Infeld electrodynamics makes the black hole horizon form. For given electric charge, there is a lower bound of mass for which black hole horizons exist. This lower bound mass is lower in the case of Born-Infeld electrodynamics than of Maxwell electrodynamics.

Acknowledgments

We would like to acknowledge many helpful discussions with Prof. Humitaka Sato and Prof. Takeshi Chiba. The author D.I. is supported by Research Fellowships of the Japan Society for the Promotion of Science of Young Scientists.

References

[1] C.G.Callan, C.Lovelace, C.R.Nappi and S.A.Yost, Nucl.Phys.B **288**, 525 (1987)

[2] M.Born and L.Infeld, Proc.Roy.Soc.A144,425 (1934)

[3] S.Hyun, hep-th/9704005

[4] M.Bañados, C.Teitelboim and J.Zanelli, Phys.Rev.lett.**69**,1849 (1992)

[5] M.Bañados, M.Henneaux, C.Teitelboim and J.Zanelli, Phys.Rev.D **48**,1506 (1993)

[6] M.Cataldo and A.García, Phys.Lett.B **456**,28 (1999)

[7] K.C.K.Chan and R.B.Mann, Phys.Rev.D **50**, 6385 (1994) ; (Erratum) Phys.Rev.D **52**, 2600 (1995)

[8] M.D.McGuigan, C.R.Nappi and S.A.Yost, Nucl.Phys.B **375**, 421 (1992)

[9] G.Mandal, A.Sengupta and S.Wadia, Mod.Phys.Lett.A **6**, 1685 (1991)

[10] P.M.Sa, A.Kleber and J.P.S.Lemos, Class.Quantum Grav. **13**, 125 (1996)

[11] J.D.Barrow, A.B.Burd and D.Lancaster, Class.Quantum Grav. **3**, 551 (1986)

[12] S.Deser and P.O.Mazur, Class.Quantum Grav. **2**, L51 (1985)

[13] J.R.Gott,III, J.Z.Simon and M.Alpert Gen.Rel.Grav. **18**, 1019 (1986)

[14] J.D.Brown and J.W.York,Jr., Phys.Rev.D **33**, 1407 (1993)

[15] J.D.Brown, J.Creighton and R.B.Mann Phys.Rev.D **50**, 6394 (1994)

[16] T.Tamaki and T.Torii, Phys.Rev.D **62**, 061501 (2000)

Initial condition of a loop cosmic string and linear perturbation

Kouji Nakamura[1]

Department of Physics, Keio University,
Hiyoshi Yokohama, 223-8521, Japan

Abstract

The initial data for the gravitational field produced by a loop thick string is considered. We show that the loop is not a geodesic on the initial hypersurface, while a loop conical singularity is. This suggests that there is the *"critical thickness"* of a string at which the linear perturbation theory around Minkowski spacetime fails to describe the gravity produced by cosmic strings. Using the above initial data, we show that the linear perturbation around flat Euclid space is plausible if the string thickness is larger than $\sim 5 \times 10^{-3} a$, where a is the curvature radius of the string.

1 Introduction

Cosmic strings are topological defects associated with the symmetry breaking in unified theories[1, 2]. These spatially one-dimensional objects present a particular interest, because it has been thought that they radiate gravitational waves by their rapid oscillations and lead to presently observable effects[3]. Detection of the gravitational waves from the strings, which are remnants of the early universe, might be a crucial evidence of the spontaneous symmetry breaking in the early universe. Thus, it is important to study the precise dynamics of cosmic stings and their gravitational effects.

In the simplest case, the dynamics of cosmic strings is idealized by the Nambu-Goto string. If the self-gravity of the string is ignored, we call test string case, the Nambu-Goto action admits oscillatory solutions. Then it is considered that these oscillations lead gravitational wave emission and the strings gradually lose their kinetic energy by this emission.

However, in [4], we have considered the dynamics of an infinite self-gravitating Nambu-Goto string within the first order with respect to its oscillation amplitude. It is shown that the string displacement is directly given by the variables of gravitational waves and there is no dynamical degree of freedom concerning the free oscillations of an infinite Nambu-Goto string. This shows that the dynamics of a test string and a self-gravitating one is quite different and suggests that the above decaying scenario of cosmic string into gravitational wave needs to be reexamined.

We also see in [4] that a string with zero thickness never bend within the linear perturbation analysis. This is the similar situation to that pointed out by Unruh et.al[6]. They conclude that the worldsheet of a conical singularity is totally geodesic and confirmed their result using a time symmetric initial data of gravity produced by a loop conical singularity[5]. If their result is true for cosmic strings in the early universe, oscillating strings bend the spacetime around them so that they are geodesics even if the deficit angle produced by the string is extremely small. However, the zero thickness of a string is essential to the above result, while cosmic strings in the early universe do have their own thickness. So these results obtained by Unruh et.al[6] and us[4] should not be taken seriously.

Though the generic world sheet of cosmic strings in the early universe will not be geodesics, the above results also suggest that there is the *"critical thickness"* of a string at which the linear perturbation theory around Minkowski spacetime fails to describe the gravity of cosmic strings. Because geodesics in the Minkowski spacetime are not circles but straight lines. To discuss this criterion, the above initial data of a loop conical singularity is a good example and we derive this criterion using this initial data.

In this article, we first show that a loop string *with non-zero thickness* in a momentarily static initial data is not geodesic in the initial hypersurface using a particular example. We only consider the case in which the matter distribution is given by the torus shell. Though this matter distribution is too artificial for a cosmic string loop, our conclusion is true in generic matter distribution. This suggests generic

[1] E-mail:kouchan@phys-h.keio.ac.jp

loop cosmic strings are not geodesics and the linear perturbation theory around Minkowski spacetime will be plausible if the strings have enough thickness. Second, we derive the above critical thickness by comparing the solution to the linear perturbation theory with flat space background and the initial data for a loop conical singularity obtained by Frolov et.al[5]. Though the initial data for the gravity in [5] are classified into two version: one is the conformally non-flat one and the other is conformally flat one, in this article, we only consider the conformally flat version, for simplicity. Our conclusion is also true in the conformally non-flat version. The discussion for another version will be seen elsewhere[8].

2 Initial condition of a loop string

Here, we consider the momentarily static initial data for the gravitational field produced by a loop string. Since a spacetime containing a straight gauge string has asymptotically conical structure[7], one may regard a cosmic string loop as a conical singularity when the radius of the loop is much larger than the string thickness. Then, we assume that the vacuum region outside the loop is given by the initial data for a loop conical singularity discussed by Frolov et.al[5]. So we first review a initial data for this loop conical singularity. Next, we consider the gravity produced by the matter energy distribution. We assume that the matter energy density is concentrated at a torus shell, for simplicity.

In a 4-dimensional spacetime $(\mathcal{M}, g_{\mu\nu})$, we consider a zero extrinsic curvature 3-dimensional spacelike submanifold (Σ, q_{ab}) with the unit normal vector u^μ. The zero extrinsic curvature means the spacetime is momentarily static at (Σ, q_{ab}). The ADM initial-value constraints for (Σ, q_{ab}) are reduced to

$$\tilde{D}^a\tilde{D}_a\phi - \frac{1}{8}\phi^{(3)}\tilde{R} = -2\pi G\phi^5 T_{\perp\perp}, \tag{1}$$

where $\tilde{D}_a$ is the covariant derivative and ${}^{(3)}\tilde{R}$ is the scalar curvature associated with the conformally transformed three metric defined by

$$\tilde{q}_{ab} := \phi^{-4} q_{ab}$$

and $T_{\perp\perp} := T_{\mu\nu}u^\mu u^\nu$ is the energy density of the matter field. Throughout this article, we concentrate only on the constraint (1).

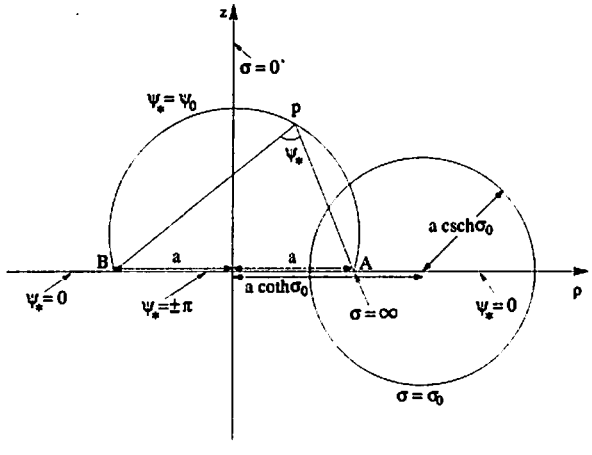


Figure 1: Toroidal coordinates σ, ψ_* is defined on the extension of a φ=constant half-plane in Euclidean space. A and B are the points at which the equatorial circle $\rho = a$ intersects this plane. For any point $P(\sigma, \psi_*, \varphi)$, $\sigma = \ln(AP/PB)$, and $\psi_* = \pm(\text{angle } APB)$, with the sign equal to the sign of z.

To discuss the initial condition for a loop string, we introduce the toroidal coordinate system. First, we consider the Euclidean metric

$$ds_E^2 = d\rho^2 + dz^2 + \rho^2 d\varphi^2. \tag{2}$$

Defining the functions σ and ψ_* by

$$\rho = a\frac{\sinh\sigma}{N^2}, \quad z = a\frac{\sin\psi_*}{N^2},$$

where $N(\sigma,\psi_*) := \sqrt{\cosh\sigma - \cos\psi_*}$, $0 \le \sigma \le \infty$ and $-\pi \le \psi_* \le \pi$, Eq. (2) becomes

$$ds_E^2 = \frac{1}{N^4}ds_*^2 := \frac{a^2}{N^4}\left(d\sigma^2 + d\psi_*^2 + \sinh^2\sigma d\varphi^2\right). \tag{3}$$

The geometrical meaning of the coordinates σ and ψ_* depicted in Fig.1. A surface of constant σ is a torus with circular cross sections of radius $a\mathrm{csch}\sigma$; the central axis of the tube forms a circle of radius $\rho = a\coth\sigma$ in the plane $z = 0$. The ring $\rho = a$, $z = 0$ is the limiting torus $\sigma = \infty$.

Since we assume that the matter distribution of the string is a δ-function distribution at a torus shell, (Σ, q_{ab}) divided into two parts: the outside the matter torus (Σ_-, q_{ab-}) and the inside (Σ_+, q_{ab+}). (Σ_-, q_{ab-}) have asymptotically flat region. We also denote the torus matter shell on $\Sigma_\pm$ by $\partial\Sigma_\pm$, respectively. We identify $\mathcal{S} = \partial\Sigma_- = \partial\Sigma_+$. We also assume that $(\Sigma_\pm, q_{ab\pm})$ are vacuum, (Σ_-, q_{ab-}) have the deficit angle around $\partial\Sigma_-$ but (Σ_+, q_{ab+}) has no deficit angle at the limit torus. The whole initial surface is given by $\Sigma = \Sigma_- \cup \mathcal{S} \cup \Sigma_+$.

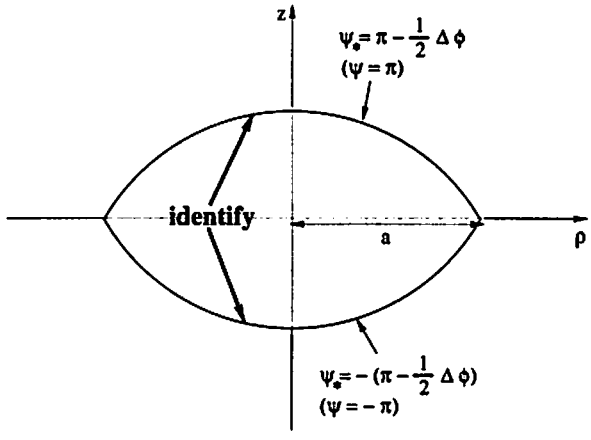


Figure 2: The conformally flat geometry of the circular string loop, located at $\rho = a$ in this Euclidean map, is constructed by conformally squeezing the two spherical caps $\psi_* = \pm(\pi - \Delta\phi/2)$ to become in effect, equatorial disks spanning the loop, and topologically identifying them.

To construct the initial data on (Σ_-, q_{ab-}), we extend the surface Σ_- to the limit torus $\sigma = \infty$ and we remove a pair of spherical caps spanning the loop as shown in Fig.2[5]. First, we fix an arbitrary angular deficit $\Delta\phi = 2\pi\alpha < 2\pi$. We restrict ψ_* to the range $-(\pi - \frac{1}{2}\Delta\phi) < \psi_* < \pi - \frac{1}{2}\Delta\phi$, identify end points of this interval and applying a conformal factor $(\phi N)^4$ to Eq.(3), the physical metric q_{ab} becomes

$$ds^2 = (\phi N)^4 ds_E^2 = \phi^4 ds_*^2, \tag{4}$$

We may regard that $ds_*^2 = \tilde{q}_{ab}dx^a dx^b$. Second, we rescale the coordinate ψ_* by

$$\psi = \eta\psi_*, \quad -\pi < \psi < \pi,$$

where $\eta = 1/(1-\alpha)$. The conformally transformed metric $\tilde{q}_{ab}$ is given by

$$\tilde{q}_{ab}dx^a dx^b = a^2\left(d\sigma^2 + \frac{1}{\eta^2}d\psi^2 + \sinh^2\sigma d\varphi^2\right).$$

Though the exposed spherical faces have the same intrinsic geometry, their extrinsic curvatures have opposite sign in $(\Sigma, N^{-4}\tilde{q}_{ab})$. This discontinuity disappears in (Σ, q_{ab}) if the conformal factor $\phi(\sigma,\psi)$ satisfies the boundary conditions as follows: $\phi(\sigma,\psi)$ is (i) even in the inversion $\psi \to -\psi$ and smooth at $\psi = \pm\pi$; (ii) $\phi N \to 1$ at $\sigma \to \infty$; (iii) $\phi N \to$ constant at $\sigma \to 0$. By virtue of the condition (i), the extrinsic curvature of the surface $\psi = \pm\pi$ in (Σ, q_{ab}) vanishes and no discontinuity is there. The condition (ii) guarantees that (Σ_-, q_{ab-}) has the conical structure with the deficit angle α at the limit tours $\sigma = \infty$. (Σ_-, q_{ab-}) becomes asymptotically flat by virtue of the final condition (iii).

Assuming $\varphi^a := (\partial/\partial\varphi)^a$ is a Killing vector, Eq. (1) for the vacuum region Σ_- in the coordinate system (σ,ψ,φ) is given by [5]

$$\frac{1}{\sinh\sigma}\partial_\sigma(\sinh\sigma\partial_\sigma\phi) + \eta^2\partial_\psi^2\phi + \frac{1}{4}\phi = 0. \tag{5}$$

The general solution to this equation with the boundary conditions (i) and (ii) has the form

$$\phi = \frac{\sqrt{2}}{\pi}\sum_{n=0}^{\infty} a_n\epsilon_n Q_{\eta n-\frac{1}{2}}(\cosh\sigma)\cos n\psi, \tag{6}$$

where $Q_\nu(*)$ is the second class Legendre function and $\epsilon_n := 2-\delta_n^0$ is the Neumann factor. By comparing this with the expansion

$$N^{-1} = \frac{\sqrt{2}}{\pi}\sum_{n=0}^{\infty}\epsilon_n Q_{n-\frac{1}{2}}(\cosh\sigma)\cos n\psi_*, \tag{7}$$

one can easily see that

$$\phi = \frac{\sqrt{2}}{\pi}\sum_{n=0}^{\infty}\epsilon_n Q_{\eta n-\frac{1}{2}}(\cosh\sigma)\cos n\psi, \tag{8}$$

is one of the simplest choice which continuously reduce to (7) in the limit $\alpha \to 0$ and satisfies the boundary conditions (i) and (ii).

From the asymptotic behavior of (8) in the limit $r := \sqrt{\rho^2+z^2} = a\sqrt{\cosh\sigma+\cos\psi}/N \to \infty$, we can easily confirm that (Σ, q_{ab}) is asymptotically flat, i.e., (8) satisfies the boundary condition (iii). Since we can show that

$$\eta\phi N \sim 1 + \frac{2a}{\pi}\frac{I(\eta)}{\sqrt{\rho^2+z^2}} = 1 + \frac{2a}{\pi}\frac{I(\eta)}{r}, \quad (r\to\infty), \tag{9}$$

where

$$I(\eta) := \int_0^\infty \mathrm{csch}x(\eta\coth\eta x - \coth x)dx. \tag{10}$$

Then the physical metric (4) is given by

$$ds^2 = (\eta\phi N)^4\frac{ds_E^2}{\eta^4} \sim \left(1+\frac{8a}{\pi}\frac{I(\eta)}{r}\right)\left\{\left(\frac{dr}{\eta^2}\right)^2 + \left(\frac{r}{\eta^2}\right)^2 d\Omega_2^2\right\}, \tag{11}$$

This shows that (Σ_-, q_{ab-}) is asymptotically flat with the ADM mass

$$M_{ADM} = \frac{4a}{\pi\eta^2}I(\eta). \tag{12}$$

If one extend to the solution (8) to the limit torus $\sigma = \infty$, (Σ_-, q_{ab-}) contains a loop conical singularity at the limit torus and the conical singularity is a geodesic on (Σ, q_{ab}). This is the point in [6]. To see this, we consider the orbit of the Killing vector φ^a and pay attention to the magnitude of the "acceleration" of this Killing orbit. In the neighborhood of the limit torus $\sigma = \infty$, the norm V of φ^a behaves

$$\frac{V^2}{a^2} = \phi^4\sinh^2\sigma \sim 1 + \frac{8}{\sqrt{\pi}}e^{-\eta\sigma}\frac{\Gamma(\eta+\frac{1}{2})}{\Gamma(\eta+1)}\cos\psi + O(e^{-2\sigma}). \tag{13}$$

Then the magnitude κ of the acceleration is given by

$$\kappa := \frac{\sqrt{(D^a V)(D_a V)}}{V} \sim \frac{2\eta}{\sqrt{\pi}a}\frac{\Gamma(\eta+\frac{1}{2})}{\Gamma(\eta+1)}e^{-(\eta-1)\sigma} \sim \frac{2\eta}{\sqrt{\pi}a}\frac{\Gamma(\eta+\frac{1}{2})}{\Gamma(\eta+1)}\left(\frac{d}{2a}\right)^{\eta-1}, \tag{14}$$

where $d \sim 2ae^{-\sigma}$ is the geodesic distance from the Killing orbits of φ^a with a finite σ to the limit torus $\sigma = \infty$. Then we conclude that the above loop conical singularity at the limit torus is a geodesic on (Σ, q_{ab}). We also note that this is true in the conformally non-flat initial data obtained in Ref.[5]. More generic discussion was also given by Unruh et.al[6] and they conclude that generic conical singularities in a spacetime are geodesic under appropriate situations.

Next, we consider the initial data for (Σ_+, q_{ab+}). Since we assume that (Σ_+, q_{ab+}) is also vacuum and has no conical singularity at the limit torus $\sigma = \infty$, we may consider the solution to (5) with $\eta = 1$. We also impose the above boundary conditions (i) and (ii) on the conformal factor ϕ and the general solution is given by (6) with $\eta = 1$. a_n in (6) are determined by the boundary condition $\partial\Sigma_+$.

Here, we consider the boundary conditions at a torus shell $\mathcal{S} = \partial\Sigma_\pm$. We assume that φ^a is a Killing vector on the whole initial surface Σ and identify the orbits of φ^a at $\mathcal{S} = \partial\Sigma_\pm$. By this identification, the coordinate function φ can be extended to the smooth function on Σ. Since ψ is the periodic coordinate with the period 2π on each $\Sigma_\pm$, we may identify the orbits of $\psi^a = (\partial/\partial\psi)^a$ at $\mathcal{S} = \partial\Sigma_\pm$ and the function ψ can be also extended to the smooth function on Σ. Finally, we assume that $\sigma = \sigma_{1\pm}$ surfaces in the toroidal coordinate system on $\Sigma_\pm$ as the boundary $\partial\Sigma_\pm$, respectively, where $\sigma_{1\pm}$ are constants. Then the intrinsic metrics $h_{ab\pm}$ on $\partial\Sigma_\pm$ are given by

$$\begin{aligned} h_{ab-} &= a^2\phi_-^4(\sigma_{1-},\psi)\left(\frac{1}{\eta^2}d\psi^2 + \sinh^2\sigma_{1-}d\varphi^2\right), \\ h_{ab+} &= a^2\phi_+^4(\sigma_{1+},\psi)\left(d\psi^2 + \sinh^2\sigma_{1+}d\varphi^2\right), \end{aligned} \tag{15}$$

where we denote the conformal factor ϕ on $\Sigma_\pm$ by $\phi_\pm$, respectively. Since $\partial\Sigma_\pm$ are diffeomorphic each other, the intrinsic metrics on $\partial\Sigma_\pm$ should be identical. Then, we obtain

$$\eta\sinh\sigma_{1-} = \sinh\sigma_{1+}, \quad \phi_+ = \phi_-\eta^{-\frac{1}{2}}. \tag{16}$$

The first condition in (16) gives the relation of $\sigma_{1\pm}$ and the second condition gives the coefficients a_n in ϕ_+ as follows:

$$a_n = \frac{\eta^{-\frac{1}{2}}Q_{n\eta-\frac{1}{2}}(\cosh\sigma_{1-})}{Q_{n-\frac{1}{2}}(\cosh\sigma_{1+})} = \frac{\eta^{-\frac{1}{2}}Q_{n\eta-\frac{1}{2}}(\cosh\sigma_{1-})}{Q_{n-\frac{1}{2}}(\sqrt{1+\eta^2\sinh\sigma_{1-}})} = \frac{\eta^{-\frac{1}{2}}Q_{n\eta-\frac{1}{2}}(\sqrt{1+\eta^{-2}\sinh\sigma_{1+}})}{Q_{n-\frac{1}{2}}(\cosh\sigma_{1+})}. \tag{17}$$

Since we assume that the matter energy density is given by the δ-function which have support only on $\mathcal{S}$,

$$T_{\perp\perp} = \rho(\psi)\delta(\chi), \tag{18}$$

where $\delta(\chi)$ is the Dirac δ-function and χ is the Gaussian normal coordinate near $\mathcal{S}$ defined by $d\chi = a\phi_\pm^2 d\sigma_\pm$, the Hamiltonian constraint (1) gives the condition for the trace of the extrinsic curvature $\kappa_{a\pm}^{\ a}$ of $\partial\Sigma_\pm$ in $\Sigma_\pm$:

$$\kappa_{a+}^{\ a} - \kappa_{a-}^{\ a} = 8\pi G\rho, \tag{19}$$

where the extrinsic curvature κ_{ab} is defined by $\kappa_{ab\pm} := -h_{a\pm}^{\ c}h_{b\pm}^{\ d}(D_c n_d)_\pm$ with the unit normal vector $n_a = a\phi_\pm^2(d\sigma_\pm)_a$ of $\partial\Sigma_\pm$, respectively. This condition determines the energy density ρ as a function of ψ:

$$8\pi G\rho(\psi) = -\frac{4}{a\phi_+^3(\sigma_{1+},\psi)}\partial_\sigma\phi_+(\sigma_{1+},\psi) - \frac{\coth\sigma_{1+}}{a\phi_+^2(\sigma_{1+},\psi)} + \frac{4}{a\phi_-^3(\sigma_{1-},\psi)}\partial_\sigma\phi_-(\sigma_{1-},\psi) + \frac{\coth\sigma_{1-}}{a\phi_-^2(\sigma_{1-},\psi)}. \tag{20}$$

In the limit $\sigma_{1\pm} \to \infty$, i.e., for the sufficiently thin matter distribution, $a\phi_+^2\rho(\psi)$ has the asymptotic form

$$a\phi_+^2(\sigma_{1+},\psi)\rho(\psi) = \frac{1}{8\pi G}\left[\left(1-\frac{1}{\eta}\right) + O(e^{-2\sigma_{1\pm}})\right]. \tag{21}$$

This shows that the relation of the energy density and the deficit angle coincides with that of a straight string in the leading order.

Here, we show that the limit torus $\sigma = \infty$ in this model is not geodesic on (Σ, q_{ab}). In the above model, the "acceleration" κ of the Killing orbit of the Killing vector φ^a has the asymptotic form

$$\kappa = \frac{a_1}{a a_0^3} + O(e^{-\sigma}) \tag{22}$$

in the limit $\sigma \to \infty$. Here a_n are constants which depends on the locus of the torus matter shell. Then, the limit torus $\sigma = \infty$ in this model is not a geodesic while that is a geodesic in the initial data (8).

When the torus shell of the matter field is sufficiently thin ($\sigma_{1\pm} \to \infty$), κ behaves

$$\kappa = \frac{1}{a}\frac{2}{\sqrt{\pi}}\frac{\Gamma(\eta+1/2)}{\Gamma(\eta+1)}\eta^{\eta} e^{-(\eta-1)\sigma_{1+}} + O(e^{-\sigma_{1\pm}}). \tag{23}$$

This shows that the acceleration of the limit torus approaches to zero as the thickness of the string becomes to zero. Then we recover the situation in the initial data given by (8). Thus, the inside matter torus shell, the Killing orbit of $(\partial/\partial\varphi)^a$ is not a geodesic on the whole spacetime if the string has its finite thickness.

Though the matter distribution considered here is too artificial, the above behavior of the magnitude κ of the "acceleration" is unchanged if the loop string has a finite length and the matter distribution does not diverge at the limit torus $\sigma = \infty$. From the metric (4), the proper length of the Killing orbit of φ^a is given by $2\pi a\phi^2 \sinh\sigma$. The limit torus $\sigma = \infty$, which is interior of the string, have a finite proper length only when $\phi = O(e^{-\sigma/2})$. In this case, the right hand side of (1) behaves

$$-2\pi G\phi^5 T_{\perp\perp} = O(e^{-\frac{5}{2}\sigma}) \times T_{\perp\perp}. \tag{24}$$

This shows that the right hand side of (1) contributes to neither $O(e^{-\sigma/2})$ nor $O(e^{-3\sigma/2})$ in ϕ but to $O(e^{-\frac{5}{2}\sigma})$ or higher order in ϕ if the energy density does not diverge at the limit torus. Then, the coefficients of $O(e^{-\sigma/2})$ and $O(e^{-3\sigma/2})$ in ϕ should be constant by virtue of (1) and the regularity at the limit torus $\sigma = \infty$. On the other hand, the behavior of the magnitude of the "acceleration" is completely determined by the order of $O(e^{-\sigma/2})$ and $O(e^{-3\sigma/2})$ in ϕ. Then, even when the distribution of $T_{\perp\perp}$ is smooth, κ must be constant if the loop string have the finite proper length and $T_{\perp\perp}$ does not diverge at $\sigma = \infty$.

3 Loop cosmic string and linear perturbation

In the last section, we have seen a thick loop energy distribution is not a geodesic in the initial surface, while a loop conical singularity is. Then we may say that the zero thickness is the essential reason why a loop conical singularity is a geodesic. On the other hand, cosmic strings in the early universe have their own thickness. Then it should not be taken seriously.

The above consideration also suggests that there is a *"critical thickness"* of a cosmic string at which the linear perturbation theory around Minkowski spacetime fails to describe the gravity produced by a string. Because the geodesics in the Minkowski spacetime are not circles but straight lines. Here, we derive this criterion.

Since there exist solutions to the Hamiltonian constraint of a thick string with the outside vacuum geometry given by (8), we discuss the above critical thickness by comparing the conformal factor (8) and the Newton potential ϕ_N produced by a loop matter distribution with zero thickness which is given by

$$\phi_N = -\frac{\alpha(1-\alpha)a}{\sqrt{(\rho+a)^2+z^2}}K(k),$$

where $K(k)$ is the first class complete elliptic integral with $k^2 := 4\rho a/((\rho+a)^2+z^2)$. This is the solution to the linear perturbation of the Hamiltonian constraint (1) with the flat Euclid space background. From (11), we can easily see that $2(1-\eta\phi N)$ coincides with ϕ_N in the asymptotically flat region $r \to \infty$.

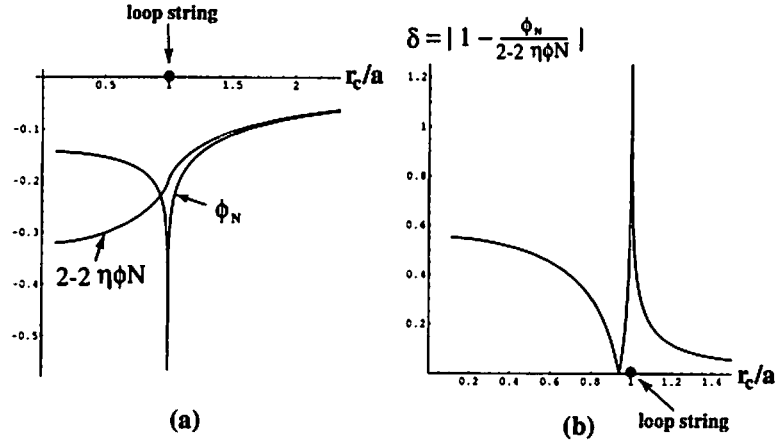


Figure 3: (a) $2(1-\eta\phi N)$ and ϕ_N in the equatorial plane with $\alpha = 0.1$ are drawn at the same circumference radius r_c of the Killing orbit of φ^a normalized by the loop radius a. (b) From $2(1-\eta\phi N)$ and ϕ_N, we can draw the profile of $\delta = |1 - \phi_N/(2-2\eta\phi N)|$.

To derive the *"critical thickness"*, we evaluate $\delta = |1 - \phi_N/(2-2\eta\phi N)|$ at the same circumference radius of the orbit of φ^a. For example, in the equatorial plane, $2(1-\eta\phi N)$ and ϕ_N with $\alpha = 0.1$ are drawn in Fig.3. δ diverge at the locus of the loop due to the divergence of ϕ_N, while $\delta \to 0$ at the asymptotically flat region. Then we calculate the circumference radii where $\delta = 1$ for each deficit angles and evaluate the critical thickness using these radii. The result is given by Fig.4. Our result shows that the linear perturbation theory around Minkowski background is plausible when the thickness is larger than $\sim 5 \times 10^{-3} a$ for a string with the deficit angle $\alpha = 10^{-6}$, which corresponds to the GUT string, where a is the loop radius.

4 Discussion

We have considered the initial data for the gravitational field produced by a loop string and show that a thick string is not a geodesic on the initial surface, while the loop conical singularity is. In [4], we discussed the fluctuating string and showed that the displacement of the fluctuating string vanishes when the thickness of the string is zero, i.e., the dynamics of strings with zero thickness are determined without their equation of motion. These results in [4] will be due to the same situation discussed here. This also suggests that the thickness must remain finite when we investigate the dynamics of self-gravitating strings and their gravitational wave emission.

Further, we also discussed the *"critical thickness"* of a cosmic string, at which the linear perturbation theory around the Minkowski spacetime fails to describe the gravitational field produced by a string. We have derived this criterion by comparing the Newton potential and the conformal factor obtained by Frolov et.al[5] as seen in Fig.3. Cosmic strings in the early universe do have their own thickness which depends on the model parameter of the phase transition[1]. Our conclusion suggests that large loops of cosmic strings will be described by the linear perturbation theory around the Minkowski spacetime.

Fig.3(a) also shows that there is the discrepancy of these potentials in the region near loop axis $z = 0$. This might be caused by the degree of freedom of gravitational waves on the initial surface. To clarify this, we must solve the time evolution from this initial data. However, δ in this region is smaller than unity and this discrepancy will be described by the linear perturbation around the Minkowski spacetime. So it is not essential to our discussion about the "critical thickness."

We must note that our conclusion here does not contradict to our conclusion in [4]; there is no dynamical degree of freedom of the free oscillations of a Nambu-Goto string within the first order with

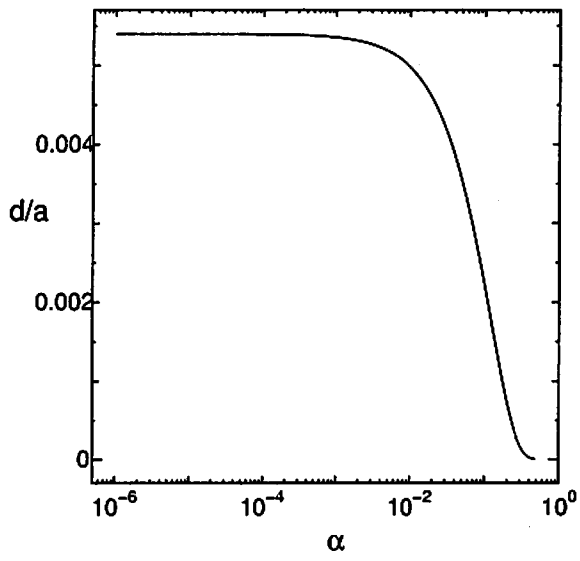


Figure 4: The *"critical thickness" d* of a cosmic string for each deficit angle α is shown. a is the loop radius. We may say that the linear perturbation theory around flat Euclid space fails to describe the gravity of a string when the string thickness is smaller than the critical thickness.

respect to its oscillation amplitude. We have carefully excluded the gauge freedom of the perturbation in [4]. Then our result will suggest that we may use the linear perturbation theory with Minkowski background if the string thickness is much larger than the above critical thickness but we have to exclude gauge freedom of the perturbation, carefully. This will be our future work.

Finally, we must comment that our arguments here and in [4] are that for the string in the situation where its bending scale is much larger that its thickness. So, we cannot say anything about the geometry of near cusps which may arise in the complex dynamics of cosmic strings.

References

[1] A. Vilenkin and E.P.S Shellard, *"Cosmic Strings and Other Topological Defects"*, (Cambridge University Press, Cambridge, England, 1994) and the references therein.

[2] E.W. Kolb and M.S. Turner, *The Early Universe*, (Addison-Wesly Publishing Company, 1993) and the references therein.

[3] A. Vilenkin, and A.E. Everett Phys. Rev. Lett. **48** 1867, (1982). T. Vachaspati, A.E. Everett and A. Vilenkin, Phys. Rev. **D30** 2046, (1984).

[4] K. Nakamura A. Ishibashi and H. Ishihara, Phys. Rev. **D62**, 101502R, (2000). K. Nakamura and H. Ishihara, preprint (gr-qc/0011035).

[5] V.P. Frolov, W. Israel and W.G. Unruh, Phys. Rev., **D39**, 1084, (1989).

[6] W.G. Unruh, G. Hayward, W. Israel and D. McManus, Phys. Rev. Lett. **62**, 2897, (1989).

[7] D. Garfinkle, Phys. Rev. **D32**, 1323, (1985). T. Futamase and D. Garfinkle, *ibid.*, **D37**, 2086, (1988).

[8] K. Nakamura in preparation.

Thick Domain Walls around a Black Hole

Yoshiyuki Morisawa,[1] Daisuke Ida,[2] Akihiro Ishibashi,[3] and Ken-ichi Nakao[4]

[1,3] *Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan,*

[2] *Department of Physics, Kyoto University, Kyoto 606-8502, Japan,*

[4] *Department of Physics, Osaka City University, Osaka 558-8585, Japan*

Abstract

We discuss the gravitationally interacting system of a thick domain wall and a black hole. We numerically solve the scalar field equation in the Schwarzschild spacetime and show that there exist scalar field configurations representing thick domain walls located around the black hole.

1 Introduction

Topological defects arise during spontaneous symmetry breaking associated with phase transitions, and cosmological evolution of them is considered to have played an important role in cosmology (see *e.g.*, [1]). Topological defects produced in the early universe might give us some information on high energy phenomena, which cannot be reached by accelerator experiments, and those produced at the late time phase transitions [2] also have attracted attention as a potential source of the cosmic structures. To study the topological defects is therefore crucially important in cosmology and elementary particle physics.

In general relativity, topological defects have interesting features. Topological defects are extended and relativistic objects due to their large tension. In particular, the gravitational field produced by an infinitely thin domain wall shows a repulsive nature [3, 4]. On the other hand, although there are many studies about the properties of thick domain walls in the flat and de Sitter spacetime backgrounds [1], little is known about the existence of thick wall configurations on an inhomogeneous, strongly curved background such as a black hole spacetime. Thus it is intriguing to study the gravitational interaction between two extended relativistic objects: a topological defect and a black hole.

In most studies of a defect-black-hole system, topological defects have been approximately treated as infinitely thin and non-gravitating objects whose dynamics are governed by Nambu-Goto action. Within this context, the scattering problem of a Nambu-Goto string by a background black hole has been studied in detail [5, 6, 7, 8, 9]. In the domain wall case, Christensen, Frolov, and Larsen [10, 11] considered Nambu-Goto walls embedded in the Schwarzschild black hole spacetime and found the static axisymmetric solutions. They showed that there exist a family of infinitely thin walls which intersect the black hole event horizon. For the case where the wall is located at the equatorial plane of the Reissner-Nordstrøm-de Sitter black hole, the stability against the perturbation is investigated by Higaki, Ishibashi, and Ida [12]. They showed that such the wall is unstable except the Schwarzschild case. However, little attention has been given to the system of a thick defect interacting with a black hole although a defect as a topologically stable configuration of a scalar field has a finite thickness.

Now we shall consider the validity of a thin-wall approximation in the system of a topological defect and a black hole. In such a system, there are two characteristic scales: the thickness w of the defect and the black hole radius R_g. In the case of the system of an astrophysical black hole with the mass $\sim M_\odot$ and a defect formed during a GUT phase transition, the thickness of the defect is much smaller than the black hole radius and therefore thin-wall approximation would be valid. However, it is not so hard to consider the situation where the thickness of defects cannot be negligible as compared with the size of a

[1] E-mail:morisawa@yukawa.kyoto-u.ac.jp
[2] E-mail:ida@tap.scphys.kyoto-u.ac.jp
[3] E-mail:akihiro@yukawa.kyoto-u.ac.jp
[4] E-mail:knakao@sci.osaka-cu.ac.jp

small black hole. Over the last few decades, many people have studied the formation of small black holes called primordial black holes (PBH's) and their cosmological implications. For example, studying the contribution of PBH's to cosmic rays enables one to place limits on the spectrum of density fluctuations in the early universe (see *e.g.* [13, 14, 15]). On the other hand, the possibility of thick defects and their roles in cosmology, has been discussed *e.g.*, as a source of large-scale structure in the universe [2], or as a candidate for some kind of dark matter [16]. For example, it is thought that the typical mass of PBH's which evaporate at the present epoch is 10^{15}g, so $R_g \sim 10^{-13}$cm. When one considers the topological defects formed during a phase transition at $\lesssim$ 100MeV, such defects become thicker than the size of PBH's and thin-wall approximation is no longer valid.

In our previous work [17], we investigated the gravitational interaction between a domain wall and a black hole taking the thickness of the wall into account. We numerically solved scalar fields equation in the Schwarzschild black hole spacetime with potentials which have a discrete set of degenerate minima. It was shown that the thick domain wall intersecting a black hole can exist as the static configuration of scalar fields for the case where the core of the wall is located at the equatorial plane of the black hole. In order to answer the question why scalar field could have such the static configuration without falling into the horizon, we speculated that the domain wall is suspended from the asymptotic region and supported against falling into the black hole by its large tension. To examine this speculation, in this work, we remove the limitation that the core of wall is located at the equatorial plane of the black hole. Then we evaluate extrinsic curvature of the wall to investigate interaction between the wall and the black hole. We consider a non-gravitating domain wall for simplicity. This test wall assumption might be valid when the symmetry braking scale of the scalar field is much lower than Planck scale (as shown in [17] by dimensional analysis).

This paper is organized as follows. In section 2, we derive the basic equation and discuss the boundary conditions which represent the situation we want to study. In section 3, we show the numerical result. We summarize our work in section 4. Throughout this paper, we use units such that $c = \hbar = G = 1$ unless otherwise stated.

2 The basic equation and the boundary conditions

We consider a static thick domain wall in a black hole spacetime. The domain wall is constructed by a scalar field with self-interaction in a given curved spacetime. As a background spacetime, we consider the Schwarzschild black hole

$$g = -\left(1 - \frac{2M}{R}\right) \mathrm{d}t^2 + \left(1 - \frac{2M}{R}\right)^{-1} \mathrm{d}R^2 + R^2(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2). \tag{1}$$

For our purpose, we find that it is more convenient to work in the isotropic coordinates $\{t, r, \vartheta, \varphi\}$, where the new radial coordinate r is defined by

$$R = r\left(1 + \frac{M}{2r}\right)^2. \tag{2}$$

We mainly consider the region outside the event horizon in this paper, which corresponds to $r > M/2$. In this coordinate system, the metric has a spatially conformally flat form

$$g = -\left(\frac{2r - M}{2r + M}\right)^2 \mathrm{d}t^2 + \left(1 + \frac{M}{2r}\right)^4 \left[\mathrm{d}r^2 + r^2(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2)\right]. \tag{3}$$

Let us consider a real scalar field ϕ with a potential $V[\phi]$, of which Lagrangian is given by

$$\mathcal{L} = -(-\det g)^{1/2}\left(\frac{1}{2}\nabla\phi \cdot \nabla\phi + V[\phi]\right). \tag{4}$$

The equation of motion for ϕ is

$$\nabla^2\phi - \partial V/\partial\phi = 0. \tag{5}$$

In this paper, we consider familiar type of potential which has a discrete set of degenerate minima; the ϕ^4 potential

$$V[\phi] = \frac{\lambda}{4}(\phi^2 - \eta^2)^2. \tag{6}$$

Since the Schwarzschild spacetime is asymptotically flat, the asymptotic boundary condition for Eq. (5) would be given by the solution in the flat spacetime. The relevant solution in the flat spacetime $g = -dt^2 + dx^2 + dy^2 + dz^2$ is the static and plane-symmetric solution

$$\phi(z) = \eta \tanh \sqrt{\lambda/2}\eta(z - z_c) \tag{7}$$

for the phi^4 potential. This solution represents domain wall in the flat spacetime characterized by the thickness of the wall

$$w = 1/\sqrt{\lambda}\eta, \tag{8}$$

and the position of the wall's core z_c. In the Schwarzschild background, the solution compatible with the above asymptotic boundary condition would have a static and axisymmetric form $\phi = \phi(r, \vartheta)$. Then, the explicit form of the equation of motion (5) becomes

$$\left(\frac{2r}{2r+M}\right)^4 \left[\frac{\partial^2}{\partial r^2} + \frac{8r}{(4r^2 - M^2)}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2}{\partial \vartheta^2} + \cot\vartheta\frac{\partial}{\partial \vartheta}\right)\right]\phi = \frac{\partial V}{\partial \phi}. \tag{9}$$

This equation can be parametrized by a single dimensionless parameter

$$\epsilon = M/2w, \tag{10}$$

by introducing dimensionless variables

$$\rho = 2r/M, \quad \Phi(\rho, \vartheta) = \phi(r, \vartheta)/\eta. \tag{11}$$

In terms of these variables, Eq. (9) becomes

$$\left(\frac{\rho}{\rho+1}\right)^4 \left[\frac{\partial^2}{\partial \rho^2} + \frac{2\rho}{(\rho^2 - 1)}\frac{\partial}{\partial \rho} + \frac{1}{\rho^2}\left(\frac{\partial^2}{\partial \vartheta^2} + \cot\vartheta\frac{\partial}{\partial \vartheta}\right)\right]\Phi = \epsilon^2\frac{\partial U}{\partial \Phi}, \tag{12}$$

where the dimensionless potential $U[\Phi] = V[\phi]/\lambda\eta^4$ is defined. U has minima at $\Phi = \pm 1$. The parameter ϵ is just a ratio of the horizon radius to the wall thickness measured in the asymptotic region, namely if ϵ is smaller (larger) than unity, then the wall is said to be thick (thin) as compared to the size of the black hole.

We now consider the boundary conditions. First, the regularity of the scalar field at the symmetric axis is given by the Neumann boundary conditions

$$\frac{\partial \Phi}{\partial \vartheta}|_{\vartheta=0} = \frac{\partial \Phi}{\partial \vartheta}|_{\vartheta=\pi} = 0. \tag{13}$$

Second, the boundary condition at the event horizon $\{\rho = 1\}$ is given by the Neumann boundary condition

$$\frac{\partial \Phi}{\partial \rho}|_{\rho=1} = 0. \tag{14}$$

This condition is the consequence of a natural requirement that the energy density observed by a freely falling observer remains finite at the event horizon. Finally, in practice, the region of numerical integration should be finite, so that we need an asymptotic boundary condition at $\rho = \rho_{\max}$ for $\rho_{\max} \gg 1$. Taking into account the flat background solution (7), we impose the Dirichlet boundary condition

$$\Phi|_{\rho=\rho_{\max}} = \tanh[2^{-1/2}\epsilon(\rho_{\max}\cos\vartheta - z_c)]. \tag{15}$$

It is limited to the case of $z_c = 0$ in the previous work.

In the next section, we numerically integrate the field equation (12) using relaxation method under these boundary conditions at the symmetric axis (13), at the event horizon (14) and in the asymptotic region (15).

ϵ^{-1}	1		2		3		4	
z_c	z_0	h	z_0	h	z_0	h	z_0	h
2	4.020381	.086934	4.644266	.070539	4.525091	.080955	4.242473	.102803
3	5.058451	.056269	5.542292	.050223	5.504168	.054564	5.362395	.061395
4	6.038380	.041349	6.413395	.038354	6.414705	.039931	6.339115	.043156
5	6.969985	.031687	7.278119	.030451	7.303108	.030968	7.264447	.033042
6	7.853169	.025660	8.144259	.024804	8.185429	.024857	8.169399	.025911
7	8.651215	.021679	9.015162	.020570	9.068382	.020427	9.067069	.020989
8	9.408504	.018431	9.892277	.017282	9.954960	.017089	9.963745	.017399

Table 1: The values of z_0 and h for several sets of parameters (ϵ, z_c). z_0 represents the position of the wall's core on the axis in z-coordinate value. h is extrinsic curvature of the wall's core evaluated on the axis.

3 Numerical results

We have done the numerical integration for the parameters $\epsilon^{-1} = 1,2,3,4$ and $z_c = 2,3,4,5,6,7,8$. Then, we get domain wall configurations in all the cases. The scalar field configuration $\Phi(x,z)$ for the typical case $\epsilon^{-1} = 2, z_c = 4$ in Fig. 1, where x and z are the Cartesian coordinates $x = \rho\sin\vartheta$, $z = \rho\cos\vartheta$. We also show the energy density E of the scalar field given by

$$E \equiv \frac{T_t^{\ t}}{\lambda\eta^4} = \frac{1}{2\epsilon^2}\left(\frac{\rho}{\rho+1}\right)^4\left[\left(\frac{\partial\Phi}{\partial\rho}\right)^2 + \frac{1}{\rho^2}\left(\frac{\partial\Phi}{\partial\vartheta}\right)^2\right] + U[\Phi] \tag{16}$$

in Fig. 2 corresponding to Fig. 1. One can see that the configuration actually has a wall-like structure, namely the energy density is localized with a certain thickness corresponding to ϵ.

We call the $\phi = 0$ surface a core of the wall. For fixed ϵ, the core of the wall with smaller $|z_c|$ exists closer to the equatorial plane, whereas the core of the wall with z_c does not exist on $z = z_c$ plane. Accordingly, we define the z-coordinate value of the core on axis as z_0 and regard z_0 as a physical parameter rather than z_c. Extrinsic curvature of the core surface shows the wall's response to the gravitational force by the black hole. Especially, we evaluate trace of extrinsic curvature k at the nearest point from black hole, *i.e.* on the axis. The value of z_0 and k are shown in Table 1. Fig. 3 shows the relation between k and z_0 for each ϵ value. The positive value of k means that the core surface is convex to black hole side contrary to appearance in the isotropic coordinates. Thicker wall (smaller ϵ) has larger value of extrinsic curvature. It is consistent with the fact that the Nambu-Goto wall (*i.e.* infinitely thin wall) is minimal surface (*i.e.* extrinsic curvature is 0). The k-lines fall down to 0 for large distance z_0. Thus it is expected that the thin wall approximation is valid in asymptotic region. However, this implies the possibility that such the approximation breaks down in inhomogeneous, strongly curved background.

We shall comment on the computational domain and the grid spacing taken in our calculation. In order that the asymptotic boundary condition (15) make sense, $\rho_{\max}$ must be large enough. In the above calculation, we take the computational domain which is fifty times as large as the horizon radius (*i.e.* $\rho_{\max} = 51$), and we carry out the integration on 400×128 grid (the grid spacing in ρ- and ϑ- direction are horizon-radius/2^3 and $\pi/2^7$ rad, respectively).

4 Summary

We have numerically solved the equation of motion for real scalar fields with ϕ^4 potential, which have a discrete set of degenerate minima, in the Schwarzschild black hole background. We showed that there exist the static axi-symmetric field configurations which represent thick domain walls located around and attracted by the black hole. We evaluated extrinsic curvature of wall's core and investigated its dependence on thickness and position of walls.

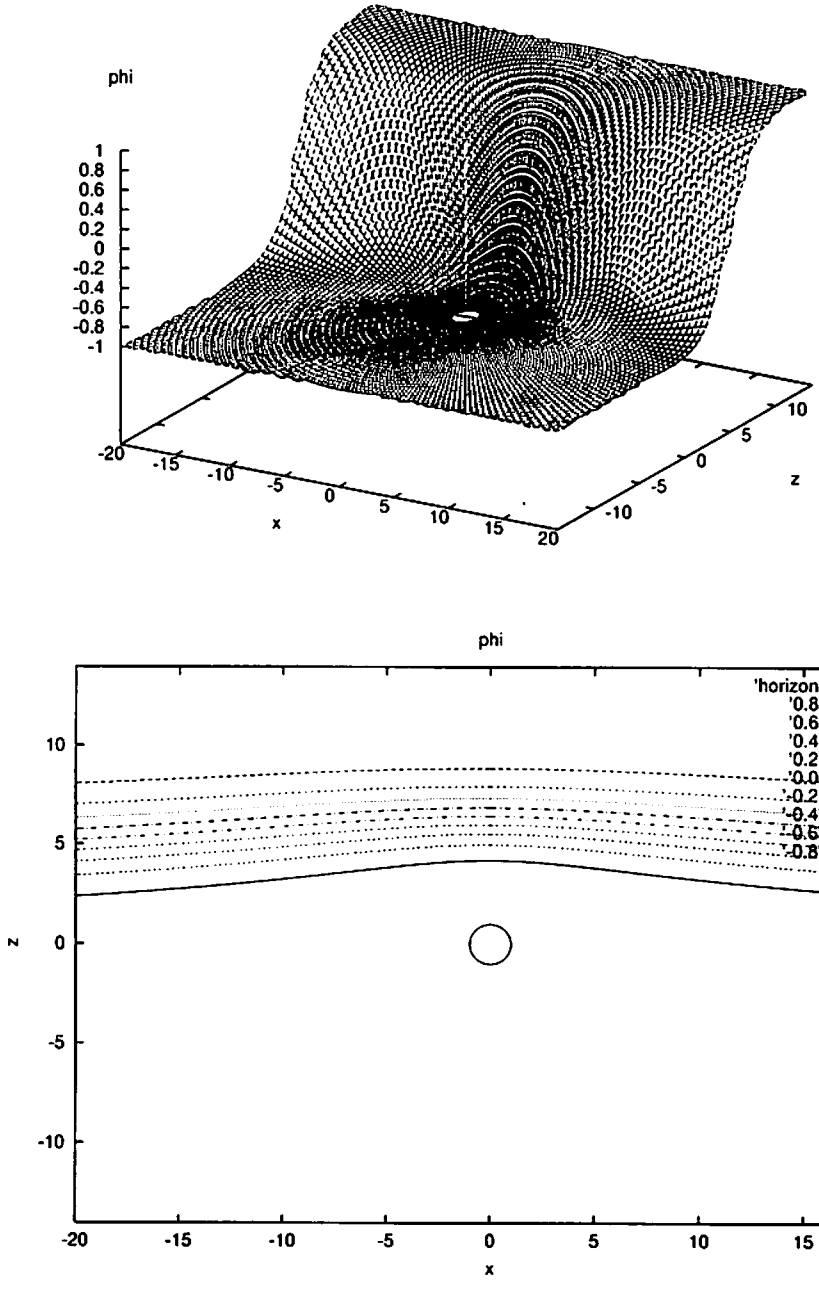


Figure 1: Numerical solution of Eq. (12) for $\epsilon^{-1} = 2, z_c = 4$. This solution expresses a domain wall configuration whose thickness is twice as the Schwarzschild radius. The core ($\phi = 0$) surface exists at $z = 6.413395$ on axis

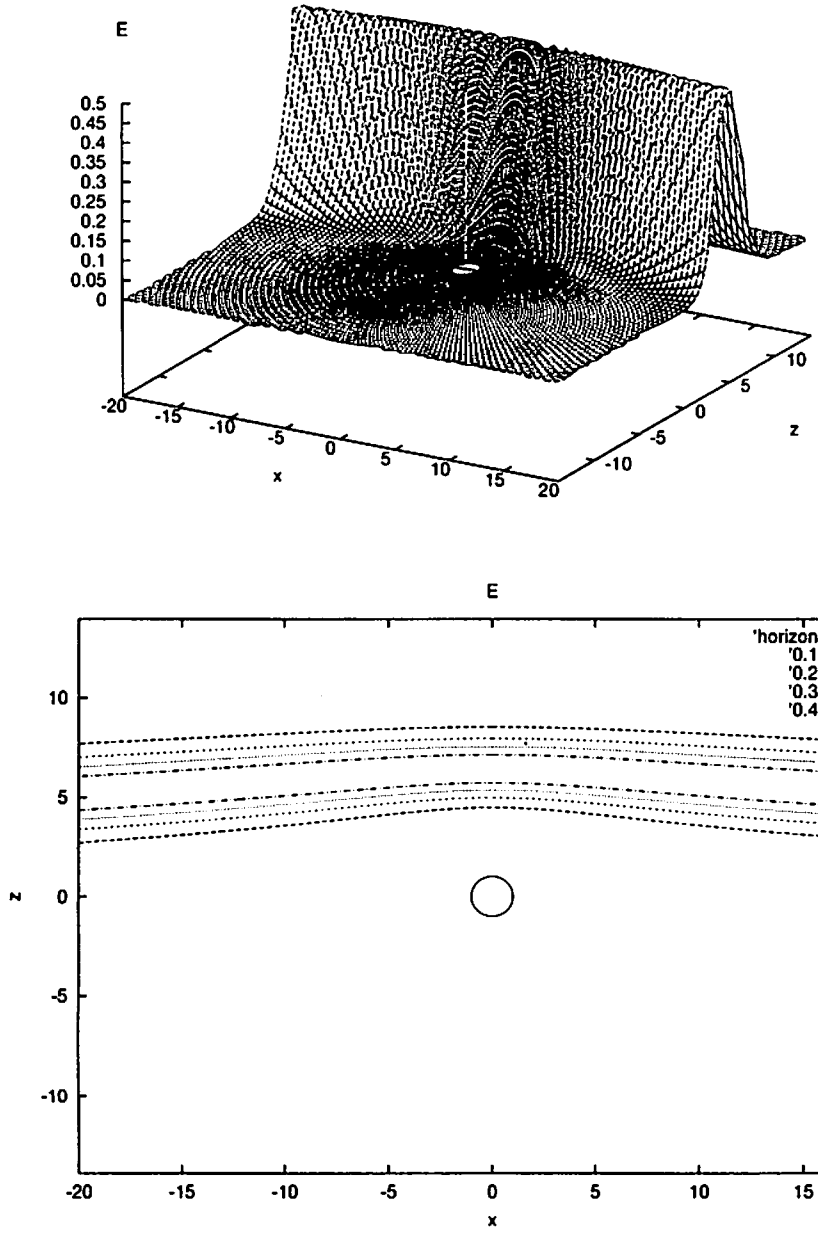


Figure 2: The energy density $E(x,z)$ of ϕ^4 scalar field for $\epsilon^{-1} = 2, z_c = 4$. This is calculated from the numerical solution shown in Fig. 1

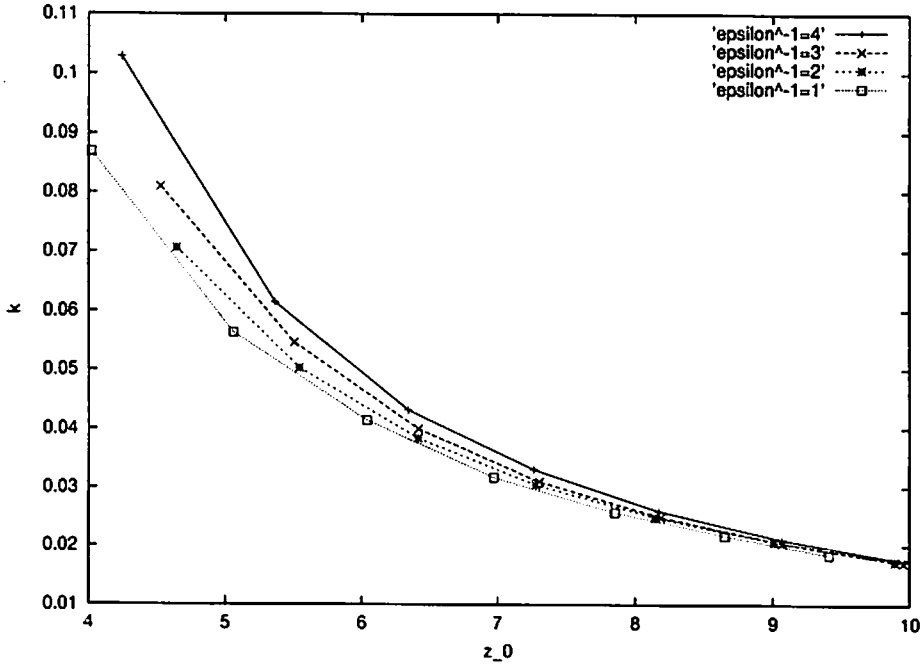


Figure 3: The relation between k and z_0 for each ϵ value.

Acknowledgments

We would like to thank Profs. Hideki Ishihara, Hideo Kodama and Takashi Nakamura for many useful suggestions and comments. We also thank Ryo Yamazaki and Susumu Higaki for stimulating discussions. D.I. and A.I. were supported by the JSPS. This work was supported in part by the Grant-in-Aid for Scientific Research Fund (D.I., No. 4318).

References

[1] A. Vilenkin and E.P.S. Shellard, *Cosmic Strings and other Topological Defects* (Cambridge University Press, New York, 1994).

[2] C. T. Hill, D. N. Schramm and J. N. Fry, Comments Nucl. Part. Phys. **19**, 25 (1989).

[3] A. Vilenkin, Phys. Lett. **B133**, 177 (1983).

[4] J. Ipser and P. Sikivie, Phys. Rev. **D30**, 712 (1984).

[5] J. De Villiers and V. Frolov, Int. J. Mod. Phys. **D7**, 957 (1998).

[6] J. De Villiers and V. Frolov, Phys. Rev. **D58**, 105018 (1998).

[7] D. N. Page, Phys. Rev. **D58**, 105026 (1998).

[8] J. De Villiers and V. Frolov, Class. Quant. Grav. **16**, 2403 (1999).

[9] D. N. Page, Phys. Rev. **D60**, 023510 (1999).

[10] M. Christensen, V. P. Frolov and A. L. Larsen, Phys. Rev. **D58**, 085008 (1998).

[11] V. P. Frolov, A. L. Larsen and M. Christensen, Phys. Rev. **D59**, 125008 (1999).

[12] S. Higaki, A. Ishibashi and D. Ida, [gr-qc/0008063].

[13] For an overview, *Proceedings, 3rd International Symposium, and Workshop on Primordial Black Holes and Hawking Radiation, 1998*, edited by D. B. Cline, Phys. Rept. **307**, 1 (1998).

[14] B. Carr and C. Goymer, Prog. Theor. Phys. Suppl. **136**, 321 (1999).

[15] J. Yokoyama, Prog. Theor. Phys. Suppl. **136**, 338 (1999).

[16] R. A. Battye, M. Bucher and D. Spergel, [astro-ph/9908047].

[17] Y. Morisawa, R. Yamazaki, D. Ida, A. Ishibashi and K. Nakao, Phys. Rev. **D62**, 084022 (2000).

An example of a timelike naked singularity formation

Hideaki Kudoh, [1]

Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

Tomohiro Harada, [2]

Department of Physics, Waseda University, Shinjuku, Tokyo 169-8555, Japan

Hideo Iguchi, [3]

Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

Abstract

The Harada-Iguchi-Nakao (HIN)'s exact solution is one of the spacetimes which describe the dynamical formation of a naked singularity. The HIN spacetime is constructed as the collapse of a regular spherical cluster of counterrotating particles. There are three kinds of radial null geodesics in the HIN spacetime. One is the regular null geodesics and the other two are the null geodesics which terminate at the singularity. The central massless singularity is a timelike naked singularity and satisfies the strong curvature condition along the null geodesics except for the instant of singularity formation. The cluster dynamically asymptotes to the singular static Einstein cluster in which the centrifugal force is balanced with gravity. The HIN solution provides an interesting example of a timelike naked singularity.

1 Introduction

The cosmic censorship hypothesis (CCH), which roughly states that singularities forming in gravitational collapse must be hidden behind event horizons, is one of the most important open problems in general relativity since it plays an important role in theories of black hole physics. Many types of gravitational collapse have been studied so far in the context of the CCH. Some of them produce globally naked singularities, but they cannot immediately be counterexamples to the CCH because the CCH requires suitable matter and appropriate initial conditions.

The well-known example of a spacetime which describes the dynamical formation of naked singularity is the Lemaître-Tolman-Bondi (LTB) spacetime [1, 2]. The LTB spacetimes describe the gravitational collapse of a spherically symmetric dust cloud and have an explicit form of the entire spacetime metric. It has been proved that the LTB spacetimes have ingoing null naked singularity from the generic initial data (see Fig. 1) [3, 4, 5]. However, this model is rather simplified because pressure is not taken into account.

Some aspects of the effect of pressure on the naked singularity formation have been investigated and several important results have been obtained. Ori and Piran investigated the spherically symmetric collapse of a perfect fluid numerically under the assumption of self-similarity [6], and analytic discussions based on self-similarity followed it [7]. Harada also solved numerically the non-self-similar spherically symmetric collapse of a perfect fluid [8]. Their results are summarized as that naked singularity can occur in the spherically symmetric collapse of a perfect fluid for a sufficiently soft equation of state. In

[1] E-mail:kudoh@yukawa.kyoto-u.ac.jp
[2] E-mail: harada@gravity.phys.waseda.ac.jp
[3] E-mail:iguchi@yukawa.kyoto-u.ac.jp

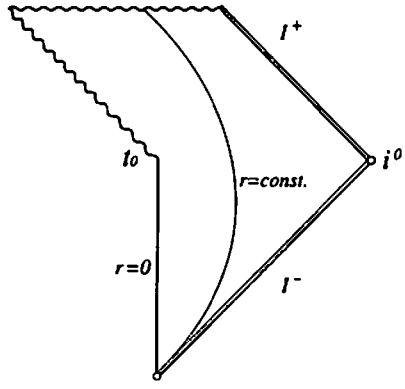


Figure 1: The conformal diagram of the LTB spacetime with naked singularity. There is one parameter family of outgoing null geodesics which emanate from the naked singularity but only one ingoing null geodesic which terminates at the naked singularity.

contrast to isotropic pressure, to include only tangential pressure is turned out to be more tractable [9, 10]. Among the solutions with vanishing radial pressure, there is a system of a spherical cloud of counterrotating particles, in which the physical origin of tangential pressure is clear [11, 12]. The each particle in the cluster has its angular momentum so that the average effect of all particles is a nonvanishing tangential pressure. This model is particularly interesting from a physical point of view because it gives insights into rotational effects on the gravitational collapse, without raising terrible difficulties. Harada, Iguchi and Nakao (HIN) have recently analyzed singularity occurrence in this model [13, 14]. Using the mass-area coordinates, which were first introduced by Ori [15] and were applied in this system by Magli [16], HIN found a new exact solution that describes dynamical formation of massless naked singularity. The results show that the counter-rotation undresses the covered singularity, and in particular tangential pressure and rotation may induce the formation of naked singularity. Since the HIN solution is given in the mass-area coordinates, the motion of each shell and the global properties of the spacetime are not trivial. We study the HIN solution in detail to make them clear [22].

2 The HIN spacetime

The HIN solution is the spherical cloud of counterrotating particles which is marginally bound. The specific angular momentum $L(r)$ of each particle at comoving radius r equals to $4F(r)$ [13], where $F(r)$ is the conserved Misner-Sharp mass function [18]. Using comoving coordinates, the line element for this system is reduced to

$$ds^2 = -e^{2\nu(t,r)}dt^2 + R'^2(t,r)\left(1+\frac{16F^2(r)}{R^2(t,r)}\right)dr^2 + R^2(t,r)(d\theta^2+\sin^2\theta d\phi^2), \tag{1}$$

where $\nu(t,r)$ and $R(t,r)$ satisfy the following set of coupled partial derivative equations:

$$\nu'(t,r) = \frac{16F^2}{R(R^2+16F^2)}R'(t,r), \tag{2}$$

$$V \equiv e^{-\nu}\dot{R}(t,r) = -\frac{|R-4F|\sqrt{2F}}{\sqrt{R(R^2+16F^2)}}. \tag{3}$$

The prime and overdot denote the partial derivatives with respect to r and t, respectively. The energy density $\epsilon(t,r) \equiv -T^t_t$ and the tangential pressure $\Pi(t,r) \equiv T^\theta_\theta = T^\phi_\phi$ are given by $\epsilon(t,r) = F'/4\pi R^2 R'$,

$\Pi(t,r) = 8F^2\epsilon/(R^2+16F^2)$. Using the coordinate transformation $m = F(r)$ and $R = R(t,r)$, the HIN solution is given in the mass-area coordinates (m,R) by

$$ds^2 = -Adm^2 - 2BdRdm - CdR^2 + R^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{4}$$

where A, B and C are given by

$$A = H\left(1-\frac{2m}{R}\right), \tag{5}$$

$$B = -\frac{R}{|R-4m|}\sqrt{\frac{RH}{2m}}, \tag{6}$$

$$C = \frac{1}{V^2} = \frac{R(R^2+16m^2)}{2m(R-4m)^2}, \tag{7}$$

$$\begin{aligned}\sqrt{H(m,R)} \equiv \sqrt{B^2-AC} &= \frac{(F^{-1})_{,m}((F^{-1})^2+16m^2)}{|F^{-1}-4m|\sqrt{2mF^{-1}}} \\ &+ \mathrm{sign}(F^{-1}-4m)\left[\left(\frac{R^2-16mR+144m^2}{3\sqrt{2}m(R-4m)}\sqrt{\frac{R}{m}}+4\sqrt{2}\ln\frac{\sqrt{R}+2\sqrt{m}}{|\sqrt{R}-2\sqrt{m}|}\right)\right. \\ &\left.-\left(\frac{(F^{-1})^2-16mF^{-1}+144m^2}{3\sqrt{2}m(F^{-1}-4m)}\sqrt{\frac{F^{-1}}{m}}+4\sqrt{2}\ln\frac{\sqrt{F^{-1}}+2\sqrt{m}}{|\sqrt{F^{-1}}-2\sqrt{m}|}\right)\right].\end{aligned} \tag{8}$$

We denote the inverse function of $F(r)$ as $F^{-1} = F^{-1}(m)$ and have set $R = F^{-1}(m)$ on the initial spacelike hypersurface, which corresponds to $R(0,r) = r$.

We can assume that the metric functions are C^∞ class with respect to the local Cartesian coordinates at least in the neighborhood of the center $r = 0$ before encountering a central singularity. By this assumption, the metric variables in the comoving coordinates are expanded near the center as

$$\nu(t,r) = \nu_0(t) + \nu_2(t)r^2 + \nu_4(t)r^4 + O(r^6), \tag{9}$$

$$R(t,r) = R_1(t)r + R_3(t)r^3 + R_5(t)r^5 + O(r^7). \tag{10}$$

Then from Eq. (3), the arbitrary mass function $F(r)$ is also expanded as

$$F(r) = F_3r^3 + F_5r^5 + F_7r^7 + O(r^9). \tag{11}$$

Substituting Eq. (9), (10) and (11) into Eq. (2) and Eq. (3), we obtain

$$R(t,r) = \left[1-\frac{t}{t_0}\right]^{2/3} r + O(r^3), \tag{12}$$

where t_0 is given by

$$t_0 = \frac{1}{3}\sqrt{\frac{2}{F_3}}. \tag{13}$$

The break down of these expansions means the appearance of central singularity. Since the break down is given by $R_1(t) = 0$, the singularity formation time is $t = t_0$.

On the initial spacelike hypersurface, the regularity requires $R(t=0,r) = r > 4F = O(r^3)$ in a sufficiently small, but finite region around $r = 0$. From Eq. (3), $R(t,r)$ is a monotonically decreasing function with respect to t, and then each initially collapsing shell approaches $R = 4F$. Using proper time $\tau(t,r) = \int e^\nu dt$ of each shell, the behavior of this approach is $R - 4F \propto \exp\left(-\frac{\tau}{6F}\right)$. This behavior shows that R approaches $4F$ asymptotically and thus R is always $R \geq 4F$ if it initially holds. Since the trapped region is given by $0 \leq R < 2F$, the region around the center is not trapped eternally. Hence the central singularity is globally naked.

3 Null geodesics in the HIN spacetime

3.1 Algebraic root equation method

The nakedness of the central singularity in the HIN spacetime is also confirmed by examining the algebraic root equation, which probes the existence of outgoing radial null geodesics from the center [16]. To obtain the root equation, consider the radial null geodesic equation in mass-area coordinates

$$\frac{dR}{dm} = \frac{-B \mp \sqrt{H}}{C} = \frac{|z-1|}{(z^2+1)\sqrt{2z}} \left(z \mp \frac{|z-1|}{\sqrt{2z}} \right), \tag{14}$$

where we have defined $z \equiv R/4m (\geq 1)$ for later convenience. The upper and lower signs refer to outgoing and ingoing null geodesics in the collapse phase, respectively. Hereafter we use this sign convention. To investigate the behavior of null geodesics near the center, we define

$$x(m) = \frac{R}{2m^\alpha}, \tag{15}$$

where α is a constant. Applying l'Hospital theorem to Eq. (15), we obtain the root equation

$$x_0 = \lim_{m \to 0} \frac{m^{3(1-\alpha)/2}(x_0 - 2m^{1-\alpha})}{2\alpha x_0 \left({x_0}^2 + 4m^{2(1-\alpha)} \right)} \sqrt{H(m, 2x_0 m^\alpha)} \left[x_0^{3/2} \mp m^{(1-\alpha)/2}(x_0 - 2m^{1-\alpha}) \right], \tag{16}$$

where $x_0 \equiv x(0)$ is introduced. If we find a consistent set of x_0 and α, we have null geodesics which behave as Eq. (15) terminating at the center $m = 0$. The root equation method picks up only the geodesics behaving as Eq. (15). To find the another possible null geodesics, we must solve the null geodesic equation.

From the detailed analysis of Eq. (16), we can prove that all possible values of α for positive finite x_0 are 1/3, 7/9 and 1. In reality, these three values of α correspond to three different types of null geodesics in the HIN spacetime, and it will be proved in this section that there are no other null geodesics in the HIN spacetime. We will see below the properties of these three types of null geodesics in detail.

3.2 Regular null geodesics : $\alpha = 1/3$

The null geodesics with $\alpha = 1/3$ correspond to regular null geodesics. For this case, the value of x_0 for $\alpha = 1/3$ cannot be determined by the root equation method. We will see that the regular null geodesics actually have $\alpha = 1/3$. From Eqs. (1), (9), (10) and (12), the radial null geodesic equation for lowest order is

$$\frac{dt}{dr} \approx \pm \left[1 - \frac{t}{t_0} \right]^{2/3}. \tag{17}$$

Inserting the solution $t = t(r)$ of Eq. (17) to Eq. (12), we obtain the regular null geodesics in the mass-area coordinates

$$R(m) \approx \left(1 - \frac{t(0)}{t_0} \right)^{2/3} \left(\frac{m}{F_3} \right)^{1/3}, \tag{18}$$

where $t(0) < t_0$ is the time when the null geodesics arrive at or emanate from the regular center $r = 0$. The arbitrariness of x_0 for $\alpha = 1/3$ comes from the arbitrary constant $t(0)$ of Eq. (18).

We can find that the scalar curvature is finite at the center along these null geodesics. In fact, we can prove that all the null geodesics with $\alpha = 1/3$ terminate at or emanate from the regular center.

3.3 The earliest singular null geodesic: $\alpha = 7/9$

The behavior of the null geodesic for $\alpha = 7/9$ is found as

$$R \approx 2x_0 m^{7/9}, \tag{19}$$

$$x_0 = \left(\frac{24F_3^2 - F_5}{4\sqrt{2}F_3^{13/6}} \right)^{2/3}. \tag{20}$$

It does not behave as regular and thus it terminates at the central singularity. In fact, the scalar curvature diverges along this null geodesic.

We will see below that there is only one ingoing or outgoing null geodesic with $\alpha = 7/9$ and that this null geodesic is the first one which arrives at or emanates from the singularity at $t = t_0$. The null geodesic equation in mass-area coordinates has a singular point at $(m, R) = (0, 0)$. To make the singular point tractable, we introduce new coordinates as

$$\chi \equiv m^{1/9}, \tag{21}$$

$$\vartheta \equiv \left(\frac{R}{m^{7/9}}\right)^{3/2}, \tag{22}$$

and then Eq. (14) becomes

$$\frac{d\vartheta}{d\chi} + \frac{6}{\chi}(\vartheta - \lambda) = \lambda\Psi_{\mp}(\chi, \vartheta), \tag{23}$$

where $\Psi_{\mp}$ is given by

$$\Psi_{\mp}(\chi, \vartheta) = \frac{1}{\chi}\left(\frac{\psi_{\mp}(\chi, \vartheta)}{\lambda} - 6\right), \tag{24}$$

$$\psi_{\mp}(\chi, \vartheta) = \frac{9}{2}\left(3\chi^2\vartheta^{1/3}\sqrt{H(\chi, \vartheta)}J_{\mp}(\chi, \vartheta) - \vartheta\right), \tag{25}$$

and we have introduced a parameter $0 < \lambda < \infty$.

The form of Eq. (23) is similar to the form of the null geodesic equation of the LTB spacetime given in [4, 5]. One may now follow to Christodoulou's argument to show the existence and uniqueness of a continuous solution of Eq. (23) [4]. As it is shown in [22], the boundary condition which the unique solution satisfies is $\vartheta(0) = \lambda_0 \equiv (2x_0)^{3/2}$, and thus the unique solution exactly agrees with the geodesics of Eq. (19). It is also shown that there exists no solution with the boundary condition $0 < \vartheta(0) < \lambda_0$ and $\lambda_0 < \vartheta(0) < \infty$. In other words, another possible solution must be $\vartheta(0) = 0$ or ∞.

We consider possible solutions with $\vartheta(0) = \infty$. When $\vartheta(0) = \infty$, Eq. (23) is approximated around $\chi = 0$:

$$\frac{d\vartheta}{d\chi} \approx -\frac{6\vartheta}{\chi}. \tag{26}$$

The integration of this equation gives $\vartheta(\chi) \propto 1/\chi^6$. This null geodesics coincide with the regular null geodesics of Eq. (18) because all the geodesics behaving as $R \propto m^{1/3}$ terminate at the regular center as we have mentioned before.

It is important that $R(t, r)$ is a monotonically decreasing function with respect to t, and thus that ϑ decreases as t increases when χ is fixed. Because of this time direction in the mass-area coordinates and the fact that there are no other null geodesics with $(2x_0)^{3/2} < \vartheta(0) < \infty$, the geodesic with $\vartheta(0) = (2x_0)^{3/2}$ is the first null geodesic which arrives at or emanates from the appeared singularity. Hence we conclude that the arrival or emanational time in comoving coordinates is the singularity formation time $t = t_0$.

3.4 Later singular null geodesics: $\alpha = 1$

There are also null geodesics with $\alpha = 1$. $x(m)$ is expressed for $\alpha = 1$ by a nonanalytic function

$$x(m) \approx 2 + 2\exp\left(-\frac{D}{\sqrt{m}}\right), \tag{27}$$

where D is the positive constant which parameterizes the null geodesics. We find that the scalar curvature diverges at the center along these null geodesics.

In the coordinate ϑ, these null geodesics are described by solutions with $\vartheta(0) = 0$ if they exist. We are actually able to show that there exists the solution whose boundary condition is $z(0) = 1$. To study this solution, we introduce the new coordinate $y \equiv z - 1$. By expanding Eq. (14) around $y = 0$, we obtain the consistent solution with $y(0) = 0$ in the lowest order approximation. We give here the null geodesic

equation which is expanded up to the enough order so that the difference between ingoing and outgoing null geodesics appears in the expansion:

$$\frac{dy}{dm} \approx \frac{y}{2m}\left(-\ln y + \frac{\beta}{m^{1/3}} + \delta \mp \sqrt{2} + \gamma m^{1/3}\right). \tag{28}$$

It can be integrated to

$$y \approx \exp\left(-\frac{D}{\sqrt{m}} + \frac{3\beta}{m^{1/3}} + \delta \mp \sqrt{2} + \frac{3\gamma}{5} m^{1/3}\right), \tag{29}$$

where $D > 0$ is an integration constant and δ is defined as $\delta \equiv 2\ln 2 - \frac{8}{3}$. It is shown in [22] that the constant D is related to the time when the null geodesics terminate at the center $m = 0$. The geodesics of Eq. (29) satisfy $y(0) = 0$ and they are classified into the geodesics with $\alpha = 1$.

We should note the curvature strength of the singularity in the HIN spacetime. According to Tipler and Królak, we classify the curvature strength whether the singularity satisfies the strong curvature condition (SCC) or the limiting focusing condition (LFC) [19, 20]. Applying theorems of [21] to the HIN spacetime, we can know the curvature strength along each null geodesic. The geodesic with $\alpha = 7/9$ satisfy the LFC. For $\alpha = 1$, the SCC is satisfied along the null geodesics. In conclusion, the naked singularity is relatively weak at the formation but becomes strong after that.

4 Causal structure of the HIN spacetime

It is important that Eq. (29) includes all the null geodesics emanating from the center after the singularity appears. We construct double null coordinates by using Eq. (29) and study the causal structure of the spacetime which is covered by these null geodesics. From Eq. (29), we introduce double null coordinates (u, v) which satisfy around $m = 0$

$$2\sqrt{2}u \approx \sqrt{m}\left(-\ln y + \frac{3\beta}{m^{1/3}} + \delta - \sqrt{2} + \frac{3\gamma}{5} m^{1/3}\right), \tag{30}$$

$$2\sqrt{2}v \approx \sqrt{m}\left(-\ln y + \frac{3\beta}{m^{1/3}} + \delta + \sqrt{2} + \frac{3\gamma}{5} m^{1/3}\right). \tag{31}$$

Then, m and y are expressed in the null coordinates by

$$\sqrt{m} \approx v - u, \tag{32}$$

$$y \approx \exp\left(-\frac{\sqrt{2}(v+u)}{v-u} + \frac{3\beta}{(v-u)^{2/3}} + \delta + \frac{3\gamma}{5}(v-u)^{2/3}\right). \tag{33}$$

We are considering a sufficiently small, but finite region around $m = 0$. Then Eq. (32) restricts u and v to $u \approx v$. From Eqs. (32) and (33), dm and dR are given by

$$dm \approx 2(v-u)(dv-du), \tag{34}$$

$$\begin{aligned} dR - 4dm \approx & -4\sqrt{2}y\left(v + u - \sqrt{2}\beta(v-u)^{1/12} + (\sqrt{2}+1)(v-u) + \frac{\sqrt{2}}{5}\gamma(v-u)^{5/12}\right)du \\ & +4\sqrt{2}y\left(v + u - \sqrt{2}\beta(v-u)^{1/12} + (\sqrt{2}-1)(v-u) + \frac{\sqrt{2}}{5}\gamma(v-u)^{5/12}\right)dv. \end{aligned} \tag{35}$$

Inserting Eqs. (34) and (35) into Eq. (4) and expanding the metric functions A, B and C around $(m, y) = (0, 0)$, we obtain the line element in the double null coordinates as

$$ds^2 \approx -512(v-u)^2 dvdu + R^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{36}$$

In the null coordinates, the central singularity is represented at $u = v$ which corresponds to the center $m = 0$ because of Eq. (32). It is found that the world line of $u = v$ is timelike. Hence the central singularity in the spacetime which is covered by the null geodesics given by Eq. (29) is timelike.

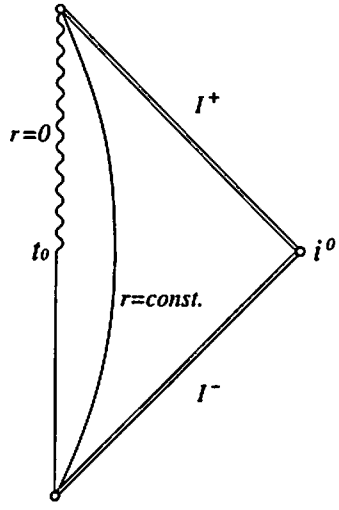


Figure 2: The conformal diagram of the HIN spacetime. A timelike naked singularity emerges at the center at $t = t_0$. There is one parameter family of both ingoing and outgoing null geodesics which terminate at and emanate from the naked singularity.

Here we summarize the obtained results. The whole of the spacetime is covered by three types of null geodesics which can be classified by the value of α into $\alpha = 1/3$, $7/9$ and 1. The null geodesics with $\alpha = 1/3$ emanate from or terminate at the regular center and they are parametrized by one parameter. There is only one null geodesic with $\alpha = 7/9$ and it is the earliest singular null geodesic which arrives at or emanates from the singularity at its formation time t_0. The null geodesics with $\alpha = 1$ emanate from or terminate at the timelike singularity and they are parametrized by one parameter. From these results, it is now possible to draw the conformal diagram of the HIN spacetime (see Fig. 2). It is found by comparison with the LTB model that the effect of counterrotation makes the singularity timelike in this model.

5 Conclusions

We have studied the causal structure of the HIN spacetime and it was shown that the central massless singularity of this spacetime is timelike. To show this fact, we have investigated the radial null geodesics in detail. The null geodesics are classified to three types, $\alpha = 1/3$, $7/9$ and 1, by their dependence of R on m near the center. We have shown for the singular null geodesics with $\alpha = 1$ that there is only one set of ingoing and outgoing geodesics for each parameter. This fact shows that the central singularity has timelike property. To show the timelikeness of the singularity completely, we have constructed double null coordinates around the central singularity from the null geodesics with $\alpha = 1$. The line element in these double null coordinates shows that there is timelike singularity in this spacetime. The curvature strength of this singularity was also investigated. The LFC is satisfied along the null geodesics with $\alpha = 7/9$ and the SCC is satisfied for $\alpha = 1$. The curvature strength of the naked singularity is relatively weak at the formation and it becomes strong after that.

In summary, the HIN solution describes a dynamical formation of timelike naked singularity, that is, the birth of timelike singularity. The solution dynamically tends to the static singular Einstein cluster as time proceeds. Though the HIN system is simply composed of collisionless particles, the collapse leads to the nontrivial causal structure. It implies that the effects of rotation and tangential pressure play important roles in the final stage of collapse, particularly the singularity formation.

References

[1] R. C. Tolman, Proc. Natl. Acad. Sci. U.S.A. **20**, 169 (1934).

[2] H. Bondi, Mon. Not. R. Astron. Soc. **107**, 410 (1947).

[3] D.M. Eardly and L. Smarr, Phys. Rev. D **19**, 2239 (1979).

[4] D. Christodoulou, Commun. Math. Phys. **93**, 171 (1984).

[5] R. P. A. C. Newman, Class. Quantum Grav. **3**, 527 (1986).

[6] A. Ori and T. Piran, Phys. Rev. Lett. **59**, 2137 (1987); Gen. Relativ. Gravit. **20**, 7 (1988); Phys. Rev. D **42**, 1068 (1990).

[7] P. S. Joshi and I. H. Dwivedi, Commun. Math. Phys. **146**, 333 (1992); Lett. Math. Phys. **27**, 235 (1993).

[8] T. Harada, Phys. Rev. D **58**, 104015 (1998).

[9] T. P. Singh and L. Witten, Class. Quantum Grav. **14**, 3489 (1997).

[10] G. Magli, Class. Quantum Grav. **14**, 1937 (1997).

[11] B. K. Datta, Gen. Relativ. Gravit. **1**, 19 (1970).

[12] H. Bondi, Gen. Relativ. Gravit. **2**, 321 (1971).

[13] T. Harada, H. Iguchi, and K. Nakao, Phys. Rev. D **58**, 041502 (1998).

[14] S. Jhingan and G. Magli, Phys. Rev. D **61**, 124006 (2000).

[15] A. Ori, Class. Quantum Grav. **7**, 985, (1990).

[16] G. Magli, Class. Quantum Grav. **15**, 3215, (1998).

[17] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[18] C. W. Misner and D. H. Sharp, Phys. Rev. **136**, B571 (1964).

[19] F. J. Tipler, Phys. Lett. **64A**, 8 (1977).

[20] A. Królak, J. Math. Phys. **28**, 138 (1987).

[21] T. Harada, K. Nakao, and H. Iguchi, Class. Quantum Grav. **16**, 2785 (1999).

[22] H. Kudoh, T. Harada and H. Iguchi, Phys. Rev. **D62**, 104016 (2000)

Instability of a membrane intersecting a black hole

Susumu Higaki,[1] Daisuke Ida[2]

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

Akihiro Ishibashi[3]

Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan

Abstract

The stability of a Nambu-Goto membrane at the equatorial plane of the Reissner-Nordstrøm-de Sitter spacetime is studied. The covariant perturbation formalism is applied to study the behavior of the perturbation of the membrane. The perturbation equation is solved numerically. It is shown that a membrane intersecting a charged black hole, including an extremely charged one, is unstable and that the positive cosmological constant strengthens the instability.

1 Introduction

It is believed that in the early universe a series of vacuum phase transitions led to several types of topological defects [1]. Topological defects are relics of the early universe and are expected to convey some information on physics of very high energy scales beyond our reach with ground-based accelerators. On the other hand, topological defects are candidates for the seeds of the observed large scale structure of the universe such as sheetlike or filamentary structures or voids. Thus topological defects are attractive examples which connect high energy physics and cosmology and we might be able to have some information on high energy physics through cosmological observations such as gravitational wave detection in the future. Topological defects, if they existed, would interact with other strong gravitational sources such as black holes, and then they might have experienced large deformation and emitted some information on themselves as gravitational waves. If we succeed in detecting such gravitational waves and identifying them, we will be able to confirm the existence of topological defects and hence the occurrence of a vacuum phase transition.

Topological defects with a finite extent such as a cosmic string or domain wall are known to produce an unusual gravitational field [2, 3] and their dynamics is slightly complicated. There are some studies on the interaction between a cosmic string and a black hole with the hope of detecting gravitational waves from cosmic strings. De Villiers and Frolov studied [4, 5] the dynamics of the scattering and capturing process of an infinitely thin test string by a Schwarzschild black hole. However less is known on the interaction between a domain wall and a black hole; only the static configurations have been studied so far. Morisawa *et al.* [6] showed the existence of static configurations of a thick domain wall intersecting a Schwarzschild black hole in its equatorial plane. Christensen *et al.* [7] numerically found a series of static configurations of a domain wall as a Nambu-Goto membrane in the Schwarzschild spacetime, some of which represent an intersecting pair of a domain wall and a black hole. When arranged in sequence, these configurations seem to represent a scattering and capturing process of a domain wall by a black hole. In most situations of cosmological interest, the thickness of a domain wall will be much smaller than the horizon radius of a solar-mass or primordial black hole, so that the infinitely thin wall (membrane) approximation will be valid in such cases. When we further neglect the gravitational effect of a domain wall, its spacetime history, i.e., the world sheet, is governed by the Nambu-Goto action and it is called the Nambu-Goto membrane.

[1] E-mail:higaki@tap.scphys.kyoto-u.ac.jp
[2] E-mail:ida@tap.scphys.kyoto-u.ac.jp
[3] E-mail:akihiro@yukawa.kyoto-u.ac.jp

A membrane lying in the equatorial plane of the black hole spacetime has the highest symmetry among the configurations which represent an intersecting pair of a membrane and a black hole. This simple configuration is a possible candidate for a final state of the scattering and capturing process provided such a state stably exists. Having this in mind, we study the stability of a Nambu-Goto membrane at the equatorial plane of the Reissner-Nordstrøm-de Sitter (RNdS) spacetime. The charge of the black hole and the positive cosmological constant are included to see their effect on the stability. This mimics the situation where a domain wall resulting from a vacuum phase transition is intersecting a charged primordial black hole in the early universe. We consider the linear perturbation of the membrane by means of the covariant perturbation formalism [8, 9]. The perturbation equation reduces to a wave equation on the world sheet with the mass term, which is negative in the presence of the charge of the black hole or the positive cosmological constant. Accordingly instability is naively expected. Solving the perturbation equation reduces to solving a two point boundary value problem. We numerically solve it by the shooting and the relaxation methods. We find that a membrane intersecting a RNdS black hole is unstable. We find the instability at least in the range of $0.001 \leq |e|/M \leq 1$ for $\Lambda > 0$, and $0.75 \leq |e|/M \leq 1$ for $\Lambda = 0$, whereas the membrane in the Schwarzschild background is stable.

2 Basic Equations

2.1 Equation of motion

The history of a membrane (world sheet) is described by the timelike hypersurface (Σ, γ_{ab}) embedded in the four-dimensional spacetime $(M, g_{\mu\nu})$. The embedding is given by

$$x^\mu = X^\mu(\xi^a), \qquad \mu = 0, \cdots, 3, \quad a = 0, \cdots, 2, \tag{1}$$

where x^μ's are the spacetime coordinates and ξ^a's the world sheet coordinates. The induced metric γ_{ab} on Σ is given by

$$\gamma_{ab} = X^\mu_{,a} X^\nu_{,b} g_{\mu\nu}. \tag{2}$$

The dynamics of the membrane is described by the Nambu-Goto action

$$S[X^\mu, X^\mu_{,a}] = -\sigma \int_\Sigma d^3\xi \sqrt{-\gamma}, \tag{3}$$

where σ represents the surface energy density of the membrane in its rest frame.

We consider the variation of the action with respect to X_μ,

$$\frac{\delta S}{\delta X_\mu} = \sigma(\Gamma^\mu_{\alpha\beta} \gamma^{ab} X^\alpha_{,a} X^\beta_{,b} + \Box X^\mu), \tag{4}$$

where $\Gamma^\mu_{\alpha\beta}$ is the spacetime affine connection and $\Box$ is the d'Alembertian on the world sheet

$$\Box = \frac{1}{\sqrt{-\gamma}} \frac{\partial}{\partial \xi^a} \left(\sqrt{-\gamma} \gamma^{ab} \frac{\partial}{\partial \xi^b} \right). \tag{5}$$

Recall the Gauss-Weingarten equation

$$D_b X^\mu_{,a} + \Gamma^\mu_{\alpha\beta} X^\alpha_{,a} X^\beta_{,b} = K_{ab} n^\mu, \tag{6}$$

where K_{ab} is the extrinsic curvature defined as

$$K_{ab} = -X^\mu_{,a} X^\nu_{,b} \nabla_\nu n_\mu, \tag{7}$$

n_μ is the normal to the world sheet, D_a is the world sheet covariant derivative and ∇_μ is the spacetime covariant derivative. Contracting Eq. (6) with γ^{ab}, we obtain

$$\Box X^\mu + \Gamma^\mu_{\alpha\beta} \gamma^{ab} X^\alpha_{,a} X^\beta_{,b} = \gamma^{ab} K_{ab} n^\mu. \tag{8}$$

With this equation, Eq. (4) becomes

$$\frac{\delta S}{\delta X_\mu} = \sigma \gamma^{ab} K_{ab} n^\mu = 0. \tag{9}$$

Equation (9) has only the component perpendicular to the world sheet. Hence the variation parallel to the world sheet has no physical meaning. Finally the equation of motion of the membrane is

$$K = \gamma^{ab} K_{ab} = 0. \tag{10}$$

This is the equation of minimal surfaces. In general, Eq. (10) cannot be solved analytically.

2.2 Perturbation equation

As noted above, the physically meaningful measure of the perturbation δX^μ is the scalar

$$\Phi = n_\mu \delta X^\mu. \tag{11}$$

Taking the variation of Eq. (10) with respect to Φ and setting it to zero, we obtain the perturbation equation

$$\Box\Phi - \left({}^3R - {}^4R_{\mu\nu} h^{\mu\nu}\right)\Phi = 0, \tag{12}$$

where $h_{\mu\nu} \equiv g_{\mu\nu} - n_\mu n_\nu$ is the projection operator onto the world sheet.

2.3 Perturbation of a membrane at the equatorial plane of the RNdS spacetime

A membrane lying in the equatorial plane of the black hole spacetime has the highest symmetry among the configurations found by [7] which represent an intersecting pair of a membrane and a Schwarzschild black hole. This simple configuration is a possible candidate for a final state of the scattering and capturing process of a membrane and a black hole provided such a state stably exists.

In general, spherically symmetric black hole spacetimes have the discrete symmetry about the equatorial plane, i.e., the equatorial plane consists of fixed points of this symmetry. Hence the equatorial plane is totally geodesic and $K = 0$.

The most general spherically symmetric black hole spacetime is the Reissner-Nordstrøm-de Sitter spacetime

$$g_{\mu\nu} dx^\mu dx^\nu = -f dt^2 + f^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2, \tag{13}$$

where

$$f = 1 - \frac{2M}{r} + \frac{e^2}{r^2} - \frac{\Lambda}{3} r^2, \tag{14}$$

Λ is the cosmological constant and M and e are the mass and charge of the black hole, respectively. The relation between the spacetime coordinates $\{t, r, \theta, \varphi\}$ and the world sheet coordinates $\{T, R, \Psi\}$ of the membrane at the equatorial plane is simply

$$t = T, \quad r = R, \quad \theta = \frac{\pi}{2}, \quad \varphi = \Psi. \tag{15}$$

Then the induced metric (2) is

$$\gamma_{ab} d\xi^a d\xi^b = -f dT^2 + f^{-1} dR^2 + R^2 d\Psi^2. \tag{16}$$

This means that the geometry of the membrane is a three-dimensional black hole spacetime. The perturbation equation (12) becomes

$$\Box\Phi + \left(\frac{e^2}{r^4} + \Lambda\right)\Phi = 0. \tag{17}$$

The mass term becomes negative when the charge of the black hole or the positive cosmological constant is present, which implies instability. This is why we consider the equatorial plane of the Reissner-Nordstrøm-de Sitter spacetime. In addition, it mimics a cosmologically interesting situation where a domain wall

resulting from a vacuum phase transition is intersecting a charged primordial black hole in the early universe.

By a separation of variables, $\Phi = (\chi(r)/\sqrt{r})\exp(i\omega t + im\varphi)$, Eq. (17) is transformed into a Schrödinger type equation in the three-dimensional black hole spacetime

$$-\frac{d^2\chi}{dr_*^2} + V(r)\chi = \omega^2\chi, \tag{18}$$

$$V(r) = f\left[\left(\frac{m^2}{r^2} - \frac{e^2}{r^4} - \Lambda\right) - \frac{f - 2f'r}{4r^2}\right], \tag{19}$$

where $r_* = \int^r dr/f$ is the usual tortoise coordinate. For the cases $M \neq 0$ we normalize r, e, Λ, ω as

$$r/M \to r, \quad e/M \to e, \quad \Lambda M^2 \to \Lambda, \quad \omega M \to \omega. \tag{20}$$

2.4 Boundary conditions

The present problem reduces to solving a Schrödinger type equation, and so does the problem of the metric perturbation of a four-dimensional black hole spacetime. Hence we will follow the stability analysis of the latter.

We consider the cases $M \neq 0$. Then the tortoise coordinate r_* goes from $-\infty$ to $+\infty$. $r_* = -\infty$ corresponds to the event horizon and $r_* = +\infty$ to infinity ($\Lambda = 0$) or the cosmological horizon ($\Lambda > 0$). The potential (19) approaches zero as r_* goes to $\pm\infty$. Therefore the asymptotic solution to Eq. (18) is a linear combination of $e^{+i\omega r_*}$ and $e^{-i\omega r_*}$. With the time dependence of the solution of the form $e^{i\omega t}$, the solutions $e^{+i\omega r_*}$ and $e^{-i\omega r_*}$ represent an ingoing wave and an outgoing wave, respectively. At infinity or the cosmological horizon we admit an outgoing wave, and at the event horizon an ingoing wave. This is because we assume no external source of perturbation of the membrane during its evolution. Hence we set the boundary conditions as follows:

$$\chi \longrightarrow \begin{cases} C_{out}\exp(-i\omega r_*) & (r_* \to +\infty), \\ C_{in}\exp(+i\omega r_*) & (r_* \to -\infty), \end{cases} \tag{21}$$

with $C_{in/out}$ constant. This form of boundary conditions appears in the standard analysis of a quasinormal mode of a black hole [10].

With the time dependence $e^{i\omega t}$, $Im(\omega) < 0$ corresponds to unstable modes. Then from the boundary conditions (21), an unstable solution to Eq. (18) decays as $e^{-|Im(\omega)|r_*}$ when r_* goes to $+\infty$ and as $e^{|Im(\omega)|r_*}$ when r_* goes to $-\infty$. Multiplying Eq. (18) by $\bar{\chi}$ (upper bar means the complex conjugate) and integrating by parts, we obtain

$$\int_{-\infty}^{+\infty}\left(\left|\frac{d\chi}{dr_*}\right|^2 + V|\chi|^2\right)dr_* + \left[-\bar{\chi}\frac{d\chi}{dr_*}\right]_{-\infty}^{+\infty} = \omega^2\int_{-\infty}^{+\infty}|\chi|^2 dr_*. \tag{22}$$

As long as we consider unstable modes, the surface term on the left-hand side (LHS) vanishes and the integrals on both sides converge. Therefore we obtain

$$\omega^2 = \int_{-\infty}^{+\infty}\left(\left|\frac{d\chi}{dr_*}\right|^2 + V|\chi|^2\right)dr_* \Big/ \int_{-\infty}^{+\infty}|\chi|^2 dr_*. \tag{23}$$

Since the RHS of Eq. (23) is real, ω is real or pure imaginary. Here we seek for unstable modes [$Im(\omega) < 0$]. So we can set $\omega = i\sigma$ ($\sigma < 0$). Then the boundary conditions (21) read

$$\chi \longrightarrow \begin{cases} C_{out}\exp(+\sigma r_*) & (r_* \to +\infty), \\ C_{in}\exp(-\sigma r_*) & (r_* \to -\infty). \end{cases} \tag{24}$$

We shall examine the eigenvalue problem of Eq. (18) subject to the boundary conditions (24).

2.5 Algorithm

The eigenvalue problem considered reduces to a two point boundary value problem. We define four dependent variables as

$$y_1 = \chi,\ y_2 = \frac{d\chi}{dr_*},\ y_3 = \sigma,\ y_4 = A, \tag{25}$$

where all y_i's are functions of r_*, and σ and A are constants. A is the ratio of C_{in} to C_{out} (or C_{out} to C_{in}). The evolution equations are

$$y_1' = y_2,\ y_2' = [V + (y_3)^2]y_1,\ y_3' = 0,\ y_4' = 0, \tag{26}$$

where the prime denotes the differentiation with respect to r_*. We impose two sets of boundary conditions. The set 1 ($y_4 = C_{out}/C_{in} \equiv A_1$) is

$$y_1 = \exp(-y_3 r_{*1}), \quad y_2 = -y_3 y_1 \quad (\text{at } r_* = r_{*1}), \tag{27}$$

$$y_1 = y_4 \exp(y_3 r_{*2}), \quad y_2 = y_3 y_1 \quad (\text{at } r_* = r_{*2}). \tag{28}$$

The set 2 ($y_4 = C_{in}/C_{out} \equiv A_2$) is

$$y_1 = y_4 \exp(-y_3 r_{*1}), \quad y_2 = -y_3 y_1 \quad (\text{at } r_* = r_{*1}) \tag{29}$$

$$y_1 = \exp(y_3 r_{*2}), \quad y_2 = y_3 y_1 \quad (\text{at } r_* = r_{*2}). \tag{30}$$

The condition $A_1 \times A_2 = 1$ is used for checking the reliability of calculations. Note that the boundary conditions (27)-(30) leave y_3, y_4 arbitrary. In general an arbitrary choice of y_3, y_4 at one boundary and the subsequent integration of the dependent variables do not ensure that the boundary conditions at the other boundary are satisfied. We solve this problem by the shooting method. We also perform the calculation by the relaxation method to confirm the results of the shooting method.

3 Results

3.1 Schwarzschild case

For pure Reissner-Nordstrøm cases the potential (19) is positive semidefinite for $m \geq 1$:

$$V(r) = \left(1 - \frac{2}{r} + \frac{e^2}{r^2}\right)\left[\left(m^2 - \frac{1}{4}\right)\frac{1}{r^2} + \frac{3}{2r^3} - \frac{9e^2}{4r^4}\right] \geq 0 \qquad (r \geq r_+), \tag{31}$$

where $r_+ = 1 + \sqrt{1 - e^2}$ is the radius of the outer event horizon. The equality is satisfied only if $m = e = 1$. Therefore, by the argument of the energy integral [10], there is no unstable mode subject to the boundary conditions (24) when $m \geq 1$.

For the Schwarzschild case with $m = 0$ the radial function $R(r) = \chi(r)/\sqrt{r}$ satisfies

$$\left(1 - \frac{2}{r}\right)\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr} - \frac{\sigma^2}{1 - 2/r}R = 0. \tag{32}$$

In terms of a new variable $\zeta = \ln(r - 2)$ this equation becomes

$$-\frac{d^2R}{d\zeta^2} + \sigma^2 r^2 R = 0. \tag{33}$$

The solution to this equation is regarded as the zero energy eigenfunction for the potential $\sigma^2 r^2$. This potential is positive definite since σ is real and negative. Therefore we have no solution which decays when ζ approaches $\pm\infty$ (the event horizon and infinity) as required by the boundary conditions (24)

$$R(r) = \frac{\chi(r)}{\sqrt{r}} \longrightarrow \begin{cases} e^{\sigma r}/\sqrt{r} & [r \to \infty\ (\zeta \to +\infty)], \\ (r-2)^{-2\sigma}/\sqrt{2} & [r \to r_+\ (\zeta \to -\infty)]. \end{cases} \tag{34}$$

Hence there exists no unstable mode for the Schwarzschild case.

3.2 RNdS case

Before describing the numerical results, we comment on the upper bound of m for which unstable modes may exisit in the Reissner-Nordstrøm-de Sitter case. The potential (19) takes the form

$$V(r) = \begin{cases} fF(r)/(4r^3) & (e=0), \\ fG(r)/(4r^4) & (e \neq 0), \end{cases} \tag{35}$$

where f is defined by Eq. (14),

$$F(r) = -5\Lambda r^3 + (4m^2 - 1)r + 6 \tag{36}$$

and

$$G(r) = -5\Lambda r^4 + (4m^2 - 1)r^2 + 6r - 9e^2. \tag{37}$$

First we consider $e = 0$ case. $F(r)$ is a monotonously decreasing function of r and has the only one root for $r \geq 0$. Therefore, by the argument of the energy integral [10], the sufficient condition for the stability of perturbation is

$$F(r_c) \geq 0, \tag{38}$$

where r_c is the radius of the cosmological horizon. Equation (38) reads

$$m^2 \geq \frac{r_c - 6 + 5\Lambda {r_c}^3}{4r_c} = 4 - \frac{9}{r_c}, \tag{39}$$

where $f(r_c) = 0$ is used. When $e = 0$, $r_c \geq 3$. Therefore Eq. (39) shows that at least for $m \geq 2$ the potential (19) is positive definite and that a membrane is stable against such perturbation.

Next we consider $e \neq 0$ case. For $r \geq 0$, $G(r)$ either stays negative, or becomes positive for an interval (r_1, r_2) but otherwise negative. Then the sufficient condition for the stability of perturbation is

$$G(r_+) \geq 0 \text{ and } G(r_c) \geq 0, \tag{40}$$

which excludes the possibility that $G(r)$ stays negative, and ensures $r_1 \leq r_+ \leq r_c \leq r_2$. When e^2 increases, r_+ and r_2 decrease and r_1 and r_c increase. Hence it is sufficient to consider $e^2 = 1$ case. In this case Eq. (40) reads

$$m^2 \geq \max\left\{ \frac{16r_c - 3r_c + 24}{4{r_c}^2}, \frac{16r_+ - 3r_+ + 24}{4{r_+}^2} \right\} = 4 + \max\left\{ -\frac{3}{4r_c} + \frac{6}{{r_c}^2}, -\frac{3}{4r_+} + \frac{6}{{r_+}^2} \right\}, \tag{41}$$

where again $f(r_+) = f(r_c) = 0$ are used. When $\Lambda > 0$, $r_+ > 1 + \sqrt{1 - e^2} = 1$ (now $e^2 = 1$). For $r > 1$

$$-\frac{3}{128} \leq -\frac{3}{4r} + \frac{6}{r^2} < 21/4. \tag{42}$$

Therefore Eq. (41) shows that at least for $m \geq 4$ there exists no unstable mode.

The condition that the Reissner-Nordstrøm-de Sitter spacetime has a static region is $\Lambda < \Lambda_{max}$, where

$$\Lambda_{max}(e^2) = \frac{3(1 + \sqrt{9 - 8e^2})}{(18 - 4e^2)(3 + \sqrt{9 - 8e^2}) - 24e^2}. \tag{43}$$

This is a monotonously increasing function of e^2 and has the minimum value $\Lambda_{max}(0) = 1/9$. So we numerically investigate the cases $\Lambda M^2 = 0, 1/1000, 1/100, 1/10$. We found unstable modes for various parameter sets $(\Lambda M^2, |e|)$. The eigenvalues of $m = 0$ modes are shown in Fig. 1. The results of the shooting method and the relaxation method agree very well, the error being within 5%.

Figure.1 shows that when $|e|/M$ becomes larger, we have a larger $|\sigma M|$ and hence a higher growth rate of the perturbations. Therefore the charge of the black hole indeed destabilizes the membrane at the equatorial plane. Since the membrane is electrically neutral, this instability should be understood as the curvature effect of the charge of the black hole. We plotted some of the eigenfunctions for $\Lambda M^2 = 0$ cases in Fig. 2.

As to the effect of the cosmological constant Λ, we find no simple tendency: for a given $|e|/M$, a larger ΛM^2 does not necessarily lead to a larger $|\sigma M|$. However with a non-vanishing cosmological constant

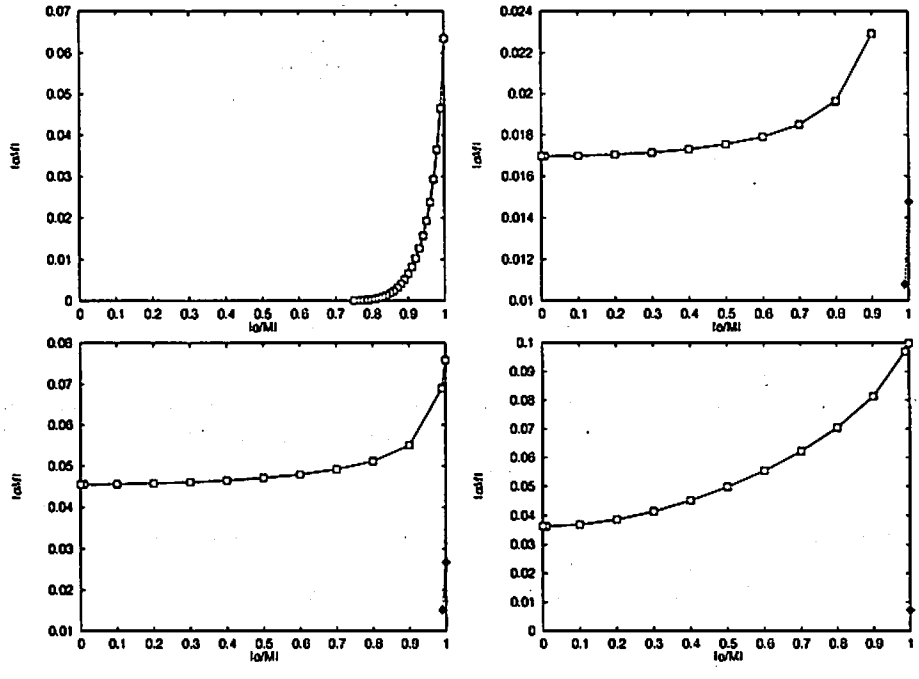


Figure 1: The plots of the eigenfrequency, $|\sigma M|$, versus the specific charge, $|e|/M$, of the black hole for the $m = 0$ cases. The top left panel shows the plot for the $\Lambda M^2 = 0$ case, the top right for the $\Lambda M^2 = 1/1000$, the bottom left for the $\Lambda M^2 = 1/100$ and the bottom right for the $\Lambda M^2 = 1/10$. We obtained two sequences of the eigenvalues except for the $\Lambda M^2 = 0$ case.

the magnitude of the eigenvalue is typically 10^{-2} for $0.001 \leq |e|/M \leq 1$ in contrast to the RN cases, in which the magnitude of the eigenvalue almost exponentially decreases with decreasing specific charge and is far below 10^{-2} for $|e|/M \lesssim 0.80$. Therefore we can at least say that the cosmological constant also destabilizes the membranes. This expectation is supported by the consideration of the de Sitter background case [11].

4 Discussion

In this paper we numerically studied the stability of a Nambu-Goto membrane at the equatorial plane of the Reissner-Nordstrøm-de Sitter spacetime and found that in general such a membrane is unstable when the black hole is charged.

A membrane is a two-dimensional extended object in the three-space. So when it moves towards a black hole, it is expected to be inevitably captured by the black hole. A series of configurations found by Christensen *et al.* seem to represent such situations. Among these configurations a membrane lying in the equatorial plane of the black hole spacetime has the highest symmetry and is most likely to be the final state of the scattering and capturing process. On seeing our result, however, we expect that it is not the case. In particular, Fig. 2 allows us to imagine that the membrane moves away from the equatorial plane. Then what eventually happens to that membrane? Causality prohibits the membrane from escaping from the event horizon. By analogy with the scattering process of a cosmic string off a black hole [4, 5] the membrane might experience large deformation and the topology of the membrane might change. Here are some speculations on the fate of the membrane; it may settle down to some other configuration than the equatorial plane, or it may break up into two parts with one swallowed by the black hole and the other escaping to infinity. We need to perform a full dynamical computation to resolve this issue. However, that is beyond the scope of this paper.

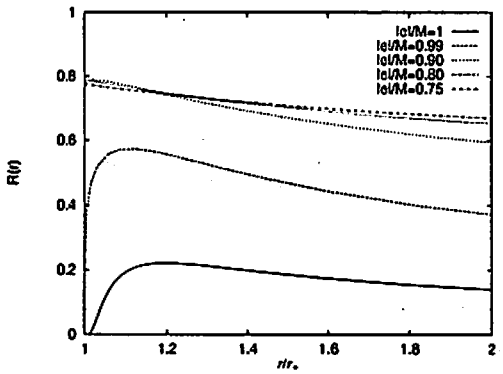


Figure 2: The radial profiles $R(r) = \chi(r)/\sqrt{r}$ of the $m = 0$ mode eigenfunctions for the pure Reissner-Nordstrøm cases ($|e|/M = 1, 0.99, 0.90, 0.80, 0.75$). The abscissa is normalized in the unit of the horizon radii r_+.

We ignored the gravitational effect of a domain wall whereas a gravitating domain wall is known to make a repulsive gravitational field [2], which is opposite to the strong attractive gravity of a black hole. It is quite intriguing to find out the consequences of the competition of these opposite forces.

Acknowledgements

We would like to thank Professor H. Sato, Professor T. Nakamura, Professor H. Kodama and Professor T. Chiba for many useful comments. We also thank Y. Morisawa and R. Yamazaki for fruitful discussion. We are grateful to Dr. T. Harada for the discussion, especially, on the boundary conditions. A.I. and D.I. were supported by the JSPS. This work was supported in part by the Grant-in-Aid for Scientific Research Fund (D.I., No. 4318).

References

[1] for example, E.W. Kolb and M.S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990); A. Vilenkin and E.P.S. Shellard, *Cosmic Strings and Other Topological Defects* (Cambridge University Press, Cambridge, England, 1994).

[2] J. Ipser and P. Sikivie, Phys. Rev. D **30**, 712 (1984).

[3] A. Vilenkin, Phys. Rev. D **23**, 852 (1981).

[4] J.P. De Villiers and V.P. Frolov, Phys. Rev. D **58**, 105018 (1998).

[5] J.P. De Villiers and V.P. Frolov, Class. Quantum Grav. **16**, 2403 (1999).

[6] Y. Morisawa, R. Yamazaki, D. Ida, A. Ishibashi, and K. Nakao, Phys. Rev. D **62**, 084022 (2000).

[7] M. Christensen, V.P. Frolov, and A.L. Larsen, Phys. Rev. D **58**, 085008 (1998).

[8] J. Garriga and A. Vilenkin, Phys. Rev. D **44**, 1007 (1991).

[9] J. Guven, Phys. Rev. D **48**, 4604 (1993).

[10] S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Oxford University Press, New York, 1992).

[11] S. Higaki, A. Ishibashi, and D. Ida, Phys. Rev. D **63**, 025002 (2001).

How does a naked singularity explode?

Hideo Iguchi [1]

Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan

Tomohiro Harada [2]

Department of Physics, Waseda University, Shinjuku, Tokyo 169-8555, Japan

Abstract

It is known that the gravitational collapse of a spherical dust cloud results in naked singularity formation. It was shown that explosive radiation is emitted during the formation process of the naked singularity. The expectation value of the quantum stress tensor for the massless scalar field on the classical background of spherical dust collapse is considered. The results show that the back reaction does not become important during the semiclassical phase.

1 Introduction

It is known that several models of spherical gravitational collapse end in central shell-focusing naked singularities from generic initial data. If we consider a quantized massless scalar field on the classical background of such collapse, we found that the out-going quantum flux diverges in the approach to the Cauchy horizon [1, 2, 3, 4]. The interpretation for this diverging flux is important for the nature of these naked singularities and for the proof or disproof of the cosmic censorship hypothesis. For example, it can be expected that the back reaction is important and prevents the naked singularity formation. Alternatively, it can enhance the naked singularity formation. We can also expect the intense radiation from a naked singularity and get information about strongly curved spacetime.

The back reaction is calculated in the semiclassical treatment as follows. First we calculate the vacuum expectation values of the energy momentum tensor of a quantized matter field. Next we use this expectation values as a source term of Einstein equations to caliculate the back reaction. From the investigation the expectation values of the energy momentum tensor we can discuss the generation of the diverging flux locally, i.e., we can conclude where this flux generated. And then we can argue that the effect of the quantum flux for the naked singularity formation.

However, in this semiclassical treatment we assume that the gravitational field can be treated classically. Whereas the appearance of naked singularity in the background spacetime means that its final evolution should be treated with quantum gravity. Therefore it is the problem whether or not the back reaction in the semi classical treatment becomes important before the quantum gravity phase. There are two criterion for deciding as to when the back reaction becomes important. One is that the total flux received at infinity becomes comparable to the mass of the collapsing star. In [5] it has been shown that, in the spherical dust model, the emitted flux is at most order of Planck energy before the quantum gravity phase. If the mass of the collapsing star is much greater than the Planck mass, then the back reaction does not become important. The second criterion is that the energy density of the quantum field becomes comparable to the background energy density inside the star. To estimate the energy density of the quantum field we should calculate the vacuum expectation value of the stress tensor for the quantum field. In a four-dimensional model, it is difficult to establish this issue. However one can study the quantum stress tensor for the 2D model which is obtained by suppressing the angular part of the 4D model [7]. For the self-similar dust collapse, the exact expression of it was computed [6]. Using this model, we investigate the exact behaviors of the quantum stress tensor. Also we show that the back reaction does not become significant during the semiclassical evolution inside the collapsing star. This report is based on the results of [8].

[1]E-mail:iguchi@yukawa.kyoto-u.ac.jp
[2]E-mail:harada@gravity.phys.waseda.ac.jp

2 Exact expression for quantum stress tensor in self-similar dust collapse

We consider the exact expression for the quantum stress tensor for the massless scalar fields in a 2D collapsing LTB spacetime. This was first done by Barve *et al.* [6]. Here we briefly introduce their results.

2.1 Self-similar spherical dust collapse

A spherically symmetric dust collapse is represented by the Lemaitre-Tolman-Bondi (LTB) solution. Using the synchronous comoving coordinate system, the line element of the LTB spacetime can be expressed in the form

$$ds^2 = dt^2 - \tilde{A}^2(t,r)dr^2 - R^2(t,r)d\Omega^2. \tag{1}$$

The energy-momentum tensor for the dust fluid is

$$T^{\mu\nu} = \rho(t,r)u^\mu u^\nu, \tag{2}$$

where $\rho(t,r)$ is the rest mass density and u^μ is the 4-velocity of the dust fluid.

Then the Einstein equations and the equation of motion for the dust fluid reduce to the following simple equations

$$\tilde{A} = \frac{R'}{\sqrt{1+f(r)}}, \tag{3}$$

$$\rho(t,r) = \frac{1}{8\pi}\frac{1}{R^2R'}\frac{dF(r)}{dr}, \tag{4}$$

$$\dot{R}^2 - \frac{F(r)}{R} = f(r), \tag{5}$$

where $f(r)$ and $F(r)$ are arbitrary functions of the radial coordinate r, and the overdot and prime denote partial derivatives with respect to t and r, respectively. From Eq. (4), $F(r)$ is related to the Misner-Sharp mass function, $m(r)$, of the dust cloud in the manner

$$m(r) = 4\pi\int_0^{R(t,r)}\rho(t,r)R^2dR = 4\pi\int_0^r\rho(t,r)R^2R'dr = \frac{F(r)}{2}. \tag{6}$$

Hence Eq. (5) might be regarded as the energy equation per unit mass. This means that the other arbitrary function, $f(r)$, is recognized as the specific energy of the dust fluid. The motion of the dust cloud is completely specified by the function, $F(r)$, and the specific energy, $f(r)$.

The self-similar solution corresponds to the choice of $2m(r) = \lambda r$, where λ is a constant, and $f(r) = 0$. Following Barve et al. [6], we assume that the central singularity appear at $t = 0$ and rescale R as $R(0,r) = (3/2)^{2/3}\lambda^{1/3}r$ at $t = 0$. The self-similar collapse solution is then

$$R = (2m(r))^{1/3}\left(\frac{3}{2}(r-t)\right)^{2/3}. \tag{7}$$

The derivative of R with respect to r is

$$R' = \left(\frac{9\lambda}{4}\right)^{1/3}\frac{1-z/3}{(1-z)^{1/3}} \tag{8}$$

where $z = t/r$. And the first and second derivatives of R' with respect to z are given by

$$\frac{dR'}{dz} = \frac{2}{9}\left(\frac{9\lambda}{4}\right)^{1/3}\frac{z}{(1-z)^{4/3}}, \tag{9}$$

and

$$\frac{d^2R'}{dz^2} = \frac{2}{9}\left(\frac{9\lambda}{4}\right)^{1/3}\frac{1+z/3}{(1-z)^{7/3}}. \tag{10}$$

Using this solution, the density function $\rho(t,r)$ is expressed as

$$\rho(t,r) = \frac{1}{6\pi(r-t)(3r-t)}. \tag{11}$$

At the center the density diverge as

$$\rho(t,0) = \frac{1}{6\pi t^2}, \tag{12}$$

to the singularity.

From the analysis of the null geodesic equations, the condition for the central singularity to be naked is obtained as [9]

$$\lambda \leq 156 - 90\sqrt{3}. \tag{13}$$

If the singularity is naked, there exist z which satisfies

$$z - R'(z) = z - \left(\frac{9\lambda}{4}\right)^{1/3} \frac{1-z/3}{(1-z)^{1/3}} = 0, \tag{14}$$

or

$$f(x) \equiv x^4 + \frac{a}{3}x^3 - x + \frac{2}{3}a = 0, \tag{15}$$

where $x = (1-z)^{1/3}$ and $a = \left(\frac{9\lambda}{4}\right)^{1/3}$. The Cauchy horizon determined by $z = z_{CH}$ which satisfies this equation. It can be shown that this singularity is locally naked as well as globally naked.

To obtain the exact expression for the quantum stress tensor, we need double-null coordinates for the 2D part of the metric (1). In the interior region, we can introduce null coordinates (η, ζ)

$$\eta = \begin{cases} re^{\int dz/z-R'} & (z-R'>0) \\ -re^{\int dz/z-R'} & (z-R'<0), \end{cases} \tag{16}$$

and

$$\zeta = \begin{cases} re^{\int dz/z+R'} & (z+R'>0) \\ -re^{\int dz/z+R'} & (z+R'<0). \end{cases} \tag{17}$$

For the convenience of the calculation, we introduce another null coordinates (U,V)

$$U = \begin{cases} \ln\eta & (z-R'>0) \\ -\ln|\eta| & (z-R'<0) \end{cases} \tag{18}$$

$$V = \begin{cases} \ln\zeta & (z+R'>0) \\ -\ln|\zeta| & (z+R'<0). \end{cases} \tag{19}$$

Then 2D metric is expressed as

$$ds^2 = A^2(U,V)dUdV, \tag{20}$$

where

$$A^2(U,V) = \begin{cases} -r^2(z^2-R'^2) & (z+R'>0) \\ r^2(z^2-R'^2) & (z+R'<0). \end{cases} \tag{21}$$

In the exterior region, the metric can be expressed as

$$ds^2 = D^2(u,v)dudv = \left(1-\frac{2M}{R}\right)dudv, \tag{22}$$

using the Eddinton-Finkelstein double null coordinates given by

$$u = T - R^*, \qquad v = T + R^*, \tag{23}$$

where

$$R^* = R + 2M\ln\left(\frac{R}{2M} - 1\right), \tag{24}$$

and

$$T = t - 2\sqrt{2MR} - 2M\ln\frac{\sqrt{R}-\sqrt{2M}}{\sqrt{R}+\sqrt{2M}}. \tag{25}$$

2.2 Exact expression for quantum stress tensor

It is known that any two dimensional spacetime is conformally flat. Then its metric can be expressed by double null coordinates $(\hat{u}, \hat{v})$ as

$$ds^2 = C^2(\hat{u}, \hat{v}) d\hat{u} d\hat{v}. \tag{26}$$

Here these null coordinates are chosen so that the initial quantum state of the scalar field is the vacuum state of the normal modes in these coordinate. The expectation value of the stress tensor of the scalar fields is then [7]

$$\langle T_{\hat{u}\hat{u}} \rangle = -\frac{1}{12\pi} C \left(\frac{1}{C}\right)_{,\hat{u},\hat{u}}, \tag{27}$$

$$\langle T_{\hat{v}\hat{v}} \rangle = -\frac{1}{12\pi} C \left(\frac{1}{C}\right)_{,\hat{v},\hat{v}}, \tag{28}$$

$$\langle T_{\hat{u}\hat{v}} \rangle = \frac{\mathcal{R} C^2}{96\pi} \tag{29}$$

where $\mathcal{R}$ is the two dimensional scalar curvature.

We are now in the situation to compute the exact expression for the quantum stress tensor in the two-dimensional self-similar collapse. Assuming the relation between internal and external null coordinates,

$$U = \alpha(u) \qquad v = \beta(V), \tag{30}$$

we can obtain all the coordinate relations as

$$\hat{v} = v = \beta(V) \qquad \hat{u} = \beta(U + 2R_0) = \beta[\alpha(u) + 2R_0]. \tag{31}$$

Then we can transform the quantum stress tensor given by $\hat{u}$ $\hat{v}$ coordinates to the one in the interior and the exterior double null coordinates.

The components of the interior (U, V) coordinates is given by

$$\langle T_{UU} \rangle = F_U(\beta') - F_U(A^2), \tag{32}$$

$$\langle T_{VV} \rangle = F_V(\beta') - F_V(A^2), \tag{33}$$

$$\langle T_{UV} \rangle = -\frac{1}{24\pi} (\ln A^2)_{,U,V}, \tag{34}$$

$$= \pm \frac{1}{24\pi} \frac{z^2 - R'^2}{2R'} \frac{d^2 R'}{dz^2}$$

where

$$F_x(y) = \frac{1}{12\pi} \sqrt{y} \left(\frac{1}{\sqrt{y}}\right)_{,x,x}. \tag{35}$$

The β' in Eqs. (32) should be considered as $\beta'(U + 2R_0)$. After the rather long calculation, we obtain

$$F_V(\beta') = \frac{1}{12\pi} \left\{ \frac{3}{4} \left(\frac{(\beta')_{,V}}{\beta'} \right)^2 - \frac{1}{2} \frac{(\beta')_{,V,V}}{\beta'} \right\} \tag{36}$$

$$= \frac{1}{12\pi} \frac{1}{4(1+x)^2} \left\{ (1 + x - \frac{x^2}{2})^2 + x^5 (1 + \frac{x}{2}) \left(\frac{1}{\lambda} - \frac{2}{3x^3} + \frac{x}{\lambda} + \frac{1}{3x^2} \right) \right\}, \tag{37}$$

where x is determined by the equation

$$\frac{r_b - v}{2M} = \frac{2}{3x^2} + \frac{2}{x} - \frac{1}{x^2} - 2 \ln \frac{1+x}{x}. \tag{38}$$

$F_U(\beta')$ can be expressed by the same equation but with x which is determined by the equation related to the retarded time as

$$\frac{r_b - \beta(\alpha(u) + 2R_0)}{2M} = \frac{2}{3x^2} + \frac{2}{x} - \frac{1}{x^2} - 2 \ln \frac{1+x}{x}. \tag{39}$$

$F_U(A^2)$ and $F_V(A^2)$ become

$$F_U(A^2) = \frac{1}{12\pi}\left\{\frac{1}{4}\left(\frac{dR'}{dz}-1\right)^2+\frac{1}{2}\frac{z^2-R'^2}{2R'}\frac{d^2R'}{dz^2}\right\}, \tag{40}$$

$$F_V(A^2) = \frac{1}{12\pi}\left\{\frac{1}{4}\left(\frac{dR'}{dz}+1\right)^2+\frac{1}{2}\frac{z^2-R'^2}{2R'}\frac{d^2R'}{dz^2}\right\}. \tag{41}$$

In the exterior Schwarzschild region we obtain

$$\langle T_{uu}\rangle = -F_u(D^2)+\alpha'^2F_U(\beta')+F_u(\alpha'), \tag{42}$$

$$\langle T_{vv}\rangle = -F_v(D^2), \tag{43}$$

$$\langle T_{uv}\rangle = -\frac{1}{24\pi}(\ln D^2)_{,u,v}, \tag{44}$$

$$= \frac{1}{24\pi}\left(2\frac{M^2}{R^4}-\frac{M}{R^3}\right)$$

The β' in Eqs. (42) also should be considered as $\beta'(U+2R_0)$.

In this region we obtain

$$F_u(D^2)=F_v(D^2)=-\frac{1}{24\pi}\left(\frac{3}{2}\frac{M^2}{R^4}-\frac{M}{R^3}\right), \tag{45}$$

this values are always positive if $R>\frac{3}{2}M$. Also we obtain

$$F_u(\alpha')=\frac{1}{12\pi}\frac{3w^6\lambda(-(5w^4+12w^3-8w^2-24w+12)\lambda-6w^7+18w^6-12w^5)}{16(2M)^2(3w^3-2\lambda-3w^4-w\lambda)^2}, \tag{46}$$

where w is determined by the equation

$$\frac{r_b-u}{2M}=\frac{2}{3w^3}+\frac{2}{w}+\frac{1}{w^2}+2\ln\frac{1-w}{w}. \tag{47}$$

The relation between u and U is obtained as

$$\frac{1}{\alpha'(u)}=-\frac{2M}{1-w}\left(\frac{1}{\lambda}-\frac{2}{3w^3}-\frac{w}{\lambda}-\frac{1}{3w^2}\right). \tag{48}$$

3 Exact behavior of quantum stress tensor

Using the above exact expressions, we can now investigate the exact behaviors of the quantum stress tensor for the massless scalar fields. The key processes are to get x which satisfy Eqs. (38) and (39) and w in Eq. (47).

From Eqs. (23), (24) and (25) we obtain

$$\frac{r-v}{2M} = \frac{2}{3}\left(\frac{R}{2M}\right)^{3/2}+2\sqrt{\frac{R}{2M}}-\frac{R}{2M}+2\ln\left(\frac{1}{\sqrt{\frac{R}{2M}}+1}\right), \tag{49}$$

$$\frac{r-u}{2M} = \frac{2}{3}\left(\frac{R}{2M}\right)^{3/2}+2\sqrt{\frac{R}{2M}}+\frac{R}{2M}-2\ln\left(\frac{1}{\sqrt{\frac{R}{2M}}-1}\right). \tag{50}$$

To obtain the value of x in Eq. (38) at some world point we need R at the surface on the line $V=$ const which go through this point. For Eq. (39) we consider the line $V=$ const which intersect the line Uconst which go through the world point. We use R at the surface on this line $V=$ const. Then we set x for both cases as follows

$$x=\sqrt{\frac{2M}{R}}. \tag{51}$$

We perform this procces by integrating the null geodesic equation numerically. At the surface u is obtained from eqs. (23), (24) and (25). Therefore we can obtain w for this u as

$$w = \sqrt{\frac{2M}{R}}, \tag{52}$$

using surface value R.

To determine what an observer would actually measure, the world line of the apparatus must be specified. For the observer with four velocity u^μ, the energy density $\langle T_{\mu\nu}\rangle u^\mu u^\nu$ and energy flux $\langle T_{\mu\nu}\rangle u^\mu n^\nu$ are measured, where $u^\mu n_\mu = 0$.

The energy density observed by comoving observer becomes

$$\rho_{\rm qst} \equiv \langle T_{\mu\nu}\rangle u^\mu u^\nu = \frac{\langle T_{UU}\rangle}{r^2(z-R')^2} + \frac{\langle T_{VV}\rangle}{r^2(z+R')^2} \mp 2\frac{\langle T_{UV}\rangle}{r^2(z^2-R'^2)}. \tag{53}$$

To investigate the importance of the back reaction, we compare this value to the background energy density. The results are shown in Fig. 1. We can see that near the center the background density is larger than the energy density of the scalar field. Then we conclude that the back reaction does not become significant during the semiclassical evolution.

Next we consider the total flux received at infinity. The radiated power of quantum flux can be expressed as

$$P = \langle T_{\mu\nu}\rangle u^\mu n^\nu = \alpha'^2 F_U(\beta') + F_u(\alpha'), \tag{54}$$

where u^μ is considered as four velocity of static observer. We plot this in Fig. 2. We reproduce the results of the geometric optics estimates for the 4D model.

The total flux is obtained from the integration

$$E = \int (\alpha'^2 F_U(\beta') + F_u(\alpha'))du \tag{55}$$

at the future null infinity. The results are shown in Fig. 3. We can see that the emitted semiclassical flux is much smaller than the mass of the original star in the range that the semiclassical approximation can be trusted.

4 Summary

We have investigated the expectation value of the quantum stress tensor for the massless scalar field on the 2D self-similar spherical dust collapse. Near the center the energy density of the background dust field is larger than the energy density of quantized scalar field. The total flux radiated during the semiclassical phase is much smaller than the mass of the background dust cloud. This result is consictant with the result of [5]. It can be said that quantum gravitational effects are more important than the back reaction in the whole course of the collapse.

References

[1] L. H. Ford and L. Parker, Phys. Rev. **D17**, 1485 (1978).

[2] W. A. Hiscock, L. G. Williams and D. M. Eardley, Phys. Rev. **D26**, 751 (1982).

[3] S. Barve, T. P. Singh, C. Vaz and L. Witten, Nucl. Phys. **B532**, 361 (1998).

[4] T. Harada, H. Iguchi and K. Nakao, Phys. Rev. D61, 101502 (2000); Phys. Rev. **D62**, 084037 (2000).

[5] T. Harada, H. Iguchi, K. Nakao, T. P. Singh, T. Tanaka and C. Vaz, gr-qc/0010101.

[6] S. Barve, T. P. Singh, C. Vaz and L. Witten, Phys. Rev. **D58**, 104018 (1998).

[7] P. C. Davies, S. A. Fulling and W. G. Unruh, Phys. Rev. **D13**, 2720 (1976).

[8] H. Iguchi and T. Harada, in preperation

[9] P. S. Joshi and T. P. Singh, Phys. Rev. D51, 6778 (1995)

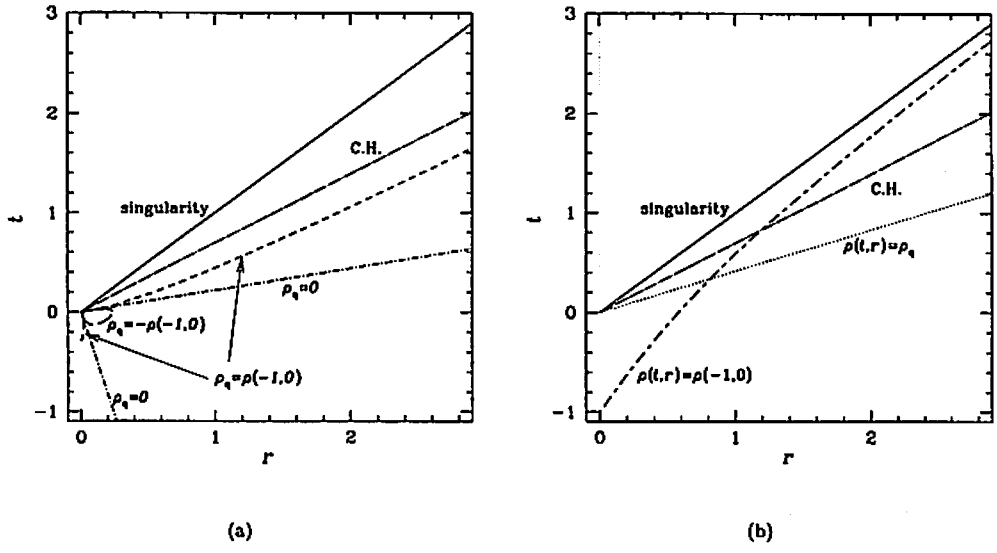


Figure 1: Plots of the lines ρ_q = const in (a) and $\rho(t,r)$ = const and $\rho_q = \rho(t,r)$ in (b). Each line has the following meaning. Solid line: singularity, long-dashed-dotted line: Cauchy horizon, dashed line: $\rho_q = \rho(-1,0)$, dashed-dotted line: $\rho_q = 0$, long-dashed line: $\rho_q = -\rho(-1,0)$, long-dashed-short-dashed line: $\rho(t,r) = \rho(-1,0)$, dotted line: $\rho_q = \rho(t,r)$.

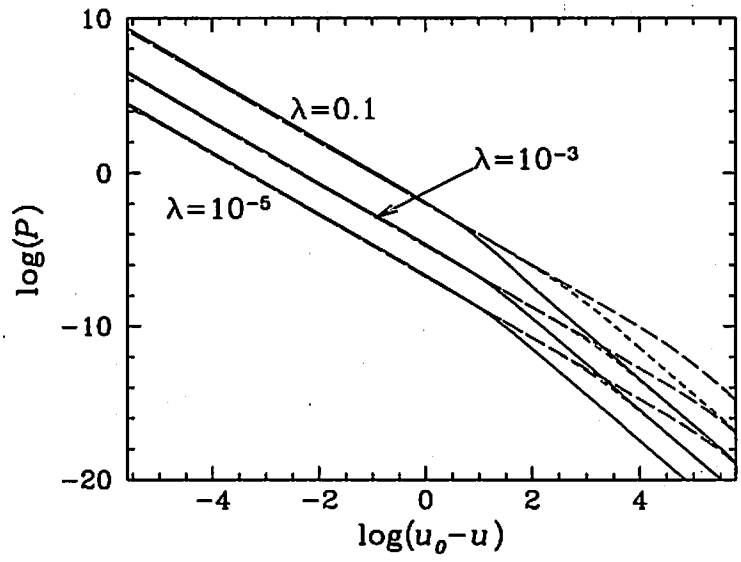


Figure 2: Radiated power received at the infinity. Upper lines correspond to $\lambda = 0.1$, , and middle lines correspond to 10^{-3}, and lower lines correspond to 10^{-5}. Solid, dashed, and long-dashed lines correspond to $r_b = 10, 10^3, 10^5$, respectively.

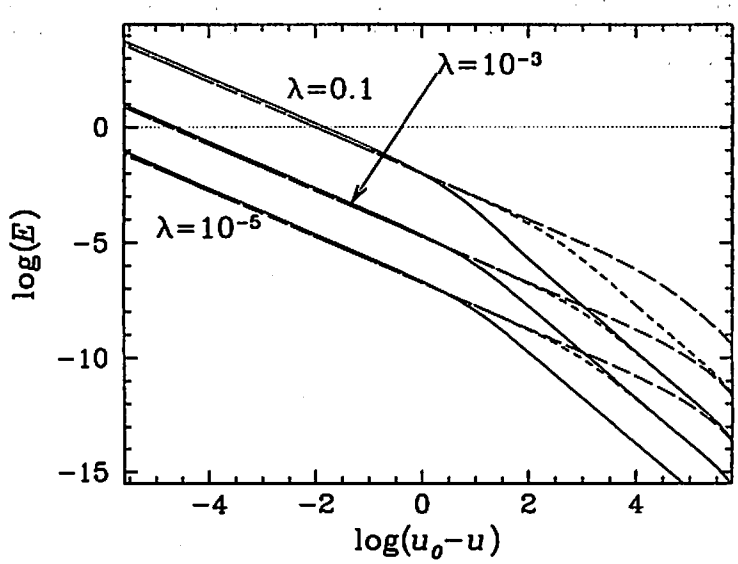


Figure 3: Integrated radiated power. Upper lines correspond to $\lambda = 0.1$, , and middle lines correspond to 10^{-3}, and lower lines correspond to 10^{-5}. Solid, dashed, and long-dashed lines correspond to $r_b = 10, 10^3, 10^5$, respectively.

Gravitational Waves due to the Naked Singularity Formation by Newtonian Analysis

Ken-ichi Nakao[1]

Department of Physics, Osaka City University, Sumiyoshi-ku, Osaka, 558-8585, Japan

Hideo Iguchi[2]

Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

Tomohiro Harada[3]

Department of Physics, Waseda University, Ohkubo, Shinjuku, Tokyo, 169-8555, Japan

Abstract

Spherical dust collapse generally forms a shell focusing naked singularity at the symmetric center. However, the Newtonian gravitational potential and speed of the dust fluid elements are everywhere much smaller than unity until the formation of the Cauchy horizon associated with the shell focusing naked singularity if an appropriate initial condition is set up. Hence the analysis by the Newtonian approximation scheme is available even in the vicinity of the space-time singularity as long as we focus on the chronological past of the Cauchy horizon. This remarkable feature makes the analysis of such singularity formation process very easy. We investigate non-spherical even-parity matter perturbations in this scheme and estimate linear gravitational waves generated just before the formation of the central shell focusing naked singularity by the quadrupole formula. The result shows good agreement with the relativistic perturbation analysis recently performed by Iguchi et al. The energy flux of the gravitational waves is finite but the space-time curvature carried by them diverges.

1 Introduction

The formation of the spacetime singularity is one of the important predictions of general relativity. However due to the non-linearity of the Einstein equations and the equations of motion for the matter and radiation, there are several unsolved problems about it. One of those is whether the spacetime singularity formed from the gravitational collapse of massive objects is visible for distant observers or not[1].

The visible spacetime singularity for distant observers is called the globally naked singularity. This globally naked singularity formed by the gravitational collapse will accompany the divergence of the energy density, pressure and the spacetime curvature. Then the classical general relativity will be inappropriate to describe the dynamics in the neighborhood of this singularity and we need the quantum theory of gravity. Hence the formation of the globally naked singularity might be called a quantum gravitational astrophysical phenomenon if it really occurs in our universe.

Unfortunately, the theory of quantum gravity has not yet been constructed, we can not definitively predict what happens in the quantum regime of the naked singularity formation. So we shall focus on

[1] E-mail:knakao@sci.osaka-cu.ac.jp
[2] E-mail:iguchi@yukawa.kyoto-u.ac.jp
[3] E-mail:harada@gravity.phys.waseda.ac.jp

what signals we get from the singularity forming region just before the quantum gravitational effects are efficient. About this issue, Nakamura et al. proposed a conjecture which states that the naked singularity formation will be a strong source of the gravitational radiation[2]. After that, there were several researches in connection with this conjecture, but it is still an open question whether this conjecture is true or not.

Recently, Iguchi et al. investigated the behavior of non-spherical perturbations of a spherically symmetric dust collapse which is an important example of the globally naked singularity formation by the gravitational collapse[3]. Their study revealed that the energy flux of the gravitational radiation generated in the neighborhood of the naked singularity is finite but the spacetime curvature carried by those gravitational waves indefinitely grow in approaching the Cauchy horizon associated with this singularity. This result does not support the conjecture by Nakamura et al. but also does not reject this conjecture since the study of the gravitational collapse of a perturbed dust sphere does not seem to lead a conclusion with generality. However, although the study by Iguchi et al. was the linear perturbation analysis, this study seems to have a special importance since they performed numerical integrations with high precision by taking into account the global structure of the spacetime which is crucial in the naked singular spacetime. Then the theme of this article is the reanalysis of the non-spherical perturbations of the spherically symmetric dust by using the Newtonian approximation scheme.

2 The Spherically Symmetric Dust Collapse by the Newtonian Approximation

The globally naked singularity formation by the gravitational collapse of a spherically symmetric dust ball is well described by the Newtonian approximation if the initial data is appropriately set up. Here we will not discuss about the validity of the Newtonian approximation in more detail and please see Ref.[4].

The Lagrangian coordinate system is appropriate to describe the dynamics of the dust ball. Hence hereafter, we adopt this coordinate system. The Newtonian line element in this coordinate system is written as

$$ds^2 = -\left(1 + 2\Phi_{\rm N} - \dot{R}^2\right) d\tau^2 + 2\dot{R}R' d\tau dx + R'^2 dx^2 + R^2(d\theta^2 + \sin^2\theta d\varphi^2), \tag{1}$$

where dot and dash mean the partial derivatives with respect to τ and x, respectively. The coordinates, x, θ and φ are the co-moving with respect to the fluid element of the dust. $\Phi_{\rm N}$ is the Newton's gravitational potential. The equation of motion for the dust fluid and the equation for the Newton's gravitational potential are given as

$$F(x) = 8\pi \int_0^x \bar{\rho} R' R^2 dx, \tag{2}$$

$$\dot{R}^2 = f(x) + \frac{F(x)}{R}, \tag{3}$$

$$\Phi'_{\rm N} = \frac{R'}{2R^2} F(x), \tag{4}$$

where $\bar{\rho}$ is the rest mass density of the dust, while $f(x)$ and $F(x)$ are arbitrary functions. In the following discussion, we shall focus on the case of $f(r) = 0$. Then the solution of Eq.(3) is obtained as

$$R = \left(\frac{9F}{4}\right)^{\frac{1}{3}} [\tau_R(x) - \tau]^{\frac{2}{3}}, \tag{5}$$

where $\tau_R(x)$ is an arbitrary function which determines the moment of the singularity formation. Hereafter we assume that this function is expressed in the following Taylor series in the neighborhood of $x = 0$;

$$\tau_R(x) = 1 + x^2 \sum_{n=1}^{\infty} \tau_{R(n)} x^{2n}. \tag{6}$$

By this assumption, the naked singularity forms at the origin $x = 0$ at the moment of $\tau = 1$.

3 Even Mode Perturbation

Here we consider non-spherical linear perturbations in the system considered in the previous system. The line element with the perturbations is given by

$$ds_{\rm L}^2 = -\left(1+2\Phi_{\rm N}+2\delta\Phi_{\rm N}-\dot{R}^2\right)d\tau^2+2\dot{R}R'd\tau dx+R'^2dx^2+R^2d\Omega^2, \tag{7}$$

where $\delta\Phi_{\rm N}$ is the potential perturbation. We write the rest mass density ρ and 4-velocity u^μ in the following forms

$$\rho=\bar{\rho}(1+\delta_\rho), \qquad \text{and} \qquad u^\mu=\bar{u}^\mu+\delta u^\mu. \tag{8}$$

By the definition of the Lagrangian coordinate system, the background 4-velocity is $(\bar{u}^\mu)=\left(\bar{u}^0,0,0,0\right)$.

We will specialize the axi-symmetric even mode perturbations. It is worthy to notice that this restriction does not loose any generality because any non-axi-symmetric perturbations can be constructed by appropriate linear combinations of rotated axi-symmetric modes. Then the perturbation quantities are expressed in the form,

$$\delta_\rho = \Delta_\rho(\tau,x)P_2(\cos\theta), \quad \delta\Phi_N=\Delta_\Phi(\tau,x)P_2(\cos\theta), \tag{9}$$

$$\delta u_1 = U_x(\tau,x)P_2(\cos\theta), \quad \delta u_2=U_\theta(\tau,x)\frac{d}{d\theta}P_2(\cos\theta), \quad \delta u_3=0. \tag{10}$$

The equations for these quantities are given as

$$\dot{\Delta}_\rho+\frac{1}{F'}\left(\frac{F'}{R'^2}U_x\right)'-\frac{6U_\theta}{R^2} = 0, \tag{11}$$

$$\dot{U}_x+\Delta_\Phi'=0, \quad \dot{U}_\theta+\Delta_\Phi = 0, \tag{12}$$

$$\frac{1}{R'R^2}\left(\frac{R^2}{R'}\Delta_\Phi'\right)'-\frac{6\Delta_\Phi}{R^2}-4\pi\bar{\rho}\Delta_\rho = 0. \tag{13}$$

The gravitational waves at the null infinity can be estimated by the so-called quadrupole formula in the framework of the Newtonian approximation. In order to so, we need the mass quadrupole moment $Q_{\ell m}$ and its temporal derivatives. Using the perturbation quantities defined in the above, the mass quadrupole moment is expressed as

$$Q_{\ell m}\equiv\int\rho\left(X_\ell X_m-\frac{1}{3}R^2\delta_{\ell m}\right)d^3X=\frac{4\pi}{15}Q(T)\text{diag}[-1,-1,2], \tag{14}$$

where T is the Eulerian time coordinate and

$$Q(T)\equiv\int_0^\infty\bar{\rho}\Delta_\rho R^4dR. \tag{15}$$

The temporal derivatives of $Q(T)$ are given by

$$\frac{d^mQ(T)}{dT^m}=\frac{1}{8\pi}\int_0^\infty F'\frac{\partial^m}{\partial\tau^m}\left(\Delta_\rho R^2\right)dx. \tag{16}$$

The power carried by the gravitational radiation to infinity is given by

$$L_{\rm GW}=\frac{32\pi^2}{375}\left(\frac{d^3Q(u)}{du^3}\right)^2, \tag{17}$$

where $u\equiv T-R$. On the other hand, the spacetime curvature carried by the gravitational waves to infinity, i.e., the Weyl scalar Ψ_4 at the null infinity is estimated as

$$\Psi_4=-\frac{3\pi}{5}\frac{d^4Q(u)}{du^4}\sin^2\theta. \tag{18}$$

To summarize; We can estimate the gravitational waves at the null infinity by integrating Eqs. (11)–(13) and then by using Eqs.(15)–(18). This procedure is much simpler than the relativistic calculations and hence we can easily get the result in less efforts.

4 Asymptotic Behavior

We are interested in the gravitational waves generated in the neighborhood of the globally naked singularity. Hence we will introduce a new variable defined by

$$w \equiv \delta\tau^{-1/2}x, \tag{19}$$

where $\delta\tau \equiv (1-\tau)/\tau_{R(1)}$. Then we consider the limit of $\tau \to 1$ with fixed w. Here it should be noted that x also approaches to zero in this limit.

In this limit, all background quantities behave in the manner[4],

$$Z(\tau, x) \longrightarrow z(w)\delta\tau^j + \text{constant}, \tag{20}$$

where $z(w)$ is some function of w and j is constant. Hence it is expected that the perturbation quantities also behave in the same manner,

$$x^{-2}\Delta_\rho = \delta^*_\rho(w)\delta\tau^{-p}, \tag{21}$$

$$x^{-1}U_x = \frac{1}{x}\partial_x(x^2U^*_\theta) = \tau^{1/3}_{R(1)}\left(w\frac{du^*_\theta}{dw} + 2u^*_\theta\right)\delta\tau^{-q}, \quad x^{-2}U_\theta = \tau^{1/3}_{R(1)}u^*_\theta(w)\delta\tau^{-q}, \tag{22}$$

$$x^{-2}\Delta_\Phi = \tau^{-2/3}_{R(1)}\delta^*_\Phi(w)\delta\tau^{-r}. \tag{23}$$

In order to confirm this expectation, we investigate the perturbation quantities normalized by those values at the origin $x = 0$.

$$\Delta^\dagger_\rho \equiv \Delta^*_\rho(\tau,x)/\Delta^*_\rho(\tau,0), \quad U^\dagger_\theta \equiv U^*_\theta(\tau,x)/U^*_\theta(\tau,0) \quad \text{and} \quad \Delta^\dagger_\Phi \equiv \Delta^*_\Phi(\tau,x)/\Delta^*_\Phi(\tau,0). \tag{24}$$

If we regard these quantities as functions of w, these quantities might not have temporal dependence. If so, our assumption (21)–(23) is supported. We show the numerical result of $\Delta^\dagger_\rho$ in Fig.1. In this figure, the data of $2.0\times10^{-4} \le 1-\tau \le 6.6\times10^{-3}$ are plotted as a function of w. There are seven curves but nobody can distinguish those from each other. During this period, the density perturbation $\Delta^*_\rho(\tau,0)$ at $1-\tau = 2.0\times10^{-4}$ becomes 381 times larger than that at $1-\tau = 6.6\times10^{-3}$. This results strongly support the above assumption on the asymptotic forms of the perturbation quantities.

Further by numerical calculations, we confirmed that the temporal dependence of the perturbation quantities follows the power law with respect to $\delta\tau$ assumed in the above. The power indices p, q and r are, respectively,

$$p = 1.701, \quad q = 0.3672 \quad \text{and} \quad r = 1.367. \tag{25}$$

Note that all perturbation quantities grow.

Using these asymptotic forms, let us estimate the contribution near the globally naked singularity to the mass quadrupole moment. The integrand in Eq.(16) behaves in the limit of $\delta\tau \to 0$ with fixed w as

$$I^{(m)}(\tau,x) \equiv \frac{\partial^m}{\partial\tau^m}\left(F'\Delta_\rho R^2\right) \longrightarrow \frac{2^{2-m}}{3}\tau^{4/3-m}_{R(1)}\delta\tau^{13/3-p-m}w^{26/3-2p-2m} \times \left(w^3\frac{d}{dw}\right)^m\left\{w^{2(p-4/3)}(1+w^2)^{4/3}\delta^*_\rho(w)\right\}. \tag{26}$$

Thus the contribution of $0 \le x \le x_o$ (x_o is sufficiently small constant) to the integral in Eq.(16) is given by

$$\int_0^{x_o} I^{(m)}(\tau,x)dx = \delta\tau^{1/2}\int_0^{w_o} I^{(m)}(\tau,\delta\tau^{1/2}w)dw \longrightarrow \frac{2^{2-m}}{3}\tau^{4/3-m}_{R(1)}\delta\tau^{29/6-p-m} \times \int_0^\infty w^{26/3-2p-2m}\left(w^3\frac{d}{dw}\right)^m\left\{w^{2(p-4/3)}(1+w^2)^{4/3}\delta^*_\rho(w)\right\}dw. \tag{27}$$

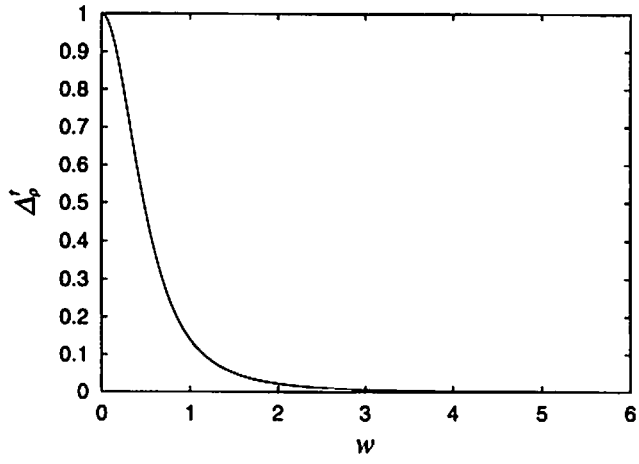


Figure 1: Normalized density perturbation $\Delta_\rho^\dagger$ is plotted as a function of w.

This result means that the contribution near the globally naked singularity to d^mQ/dT^m diverges in the limit of $\tau \to 1$ only if $m \geq 4$. Hence even if the globally naked singularity forms, the power $L_{\rm GW}$ of the gravitational waves generated in the course of the formation of the naked singularity is finite. However the spacetime curvature carried by these gravitational waves becomes infinite in this limit because it is proportional to d^4Q/dT^4. Using the above equation, we find that the Weyl scalar Ψ_4 behaves in the limit of $\delta\tau \to 0$ as

$$\Psi_4 \propto \delta\tau^{5/6-p} \propto (1-\tau)^{-0.867}. \tag{28}$$

This result quite well agrees with the result obtained by the relativistic perturbation analysis performed by Iguchi et al[3].

5 Summary

By using the Newtonian approximation, we re-analyzed the dynamics of the non-spherical perturbations of the spherically symmetric dust collapse with the formation of the globally naked singularity at the symmetric center. The result quite well agrees with that obtained by the relativistic perturbation analysis performed by Iguchi et al.

Intuitively, the gravity near the globally naked singularity is weak and hence the horizon does not form and further the singularity is visible for the distant observers. Thus even for the case with non-vanishing pressure or for less symmetric non-linear systems (eg. axi-symmetric system), the Newtonian approximation scheme seems to be available.

The relativistic treatment of a less symmetric non-linear system is very complicated. On the other hand, in the framework of the Newtonian approximation, the treatment of it is much simpler than the relativistic one, although the analysis is also difficult. Hence the strategy using the Newtonian approximation seems to be very effective to investigate the process of the naked singularity formation. It is worthy to notice that if the Newtonian approximation does not break down in the course of the singularity formation, this singularity will be globally naked since there is no horizon in the Newtonian situation.

References

[1] R. Penrose, Riv. Nuovo Cim. **1**, 252 (1969).

[2] T. Nakamura, M. Shibata and K. Nakao, Prog. Theor. Phys. **89**, 821 (1993).

[3] H. Iguchi, T. Harada and K. Nakao, Prog. Theor. Phys. 103, 53 (2000); other important references their in

[4] K. Nakao, H. Iguchi and T. Harada, gr-qc/0006057, Phys.Rev.D in press.

Comparing Quantum Black Holes and Naked Singularities

T. P. Singh,[1]

Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India
Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan

Abstract

There are models of gravitational collapse in classical general relativity which admit the formation of naked singularities as well as black holes. These include fluid models as well as models with scalar fields as matter. Even if fluid models were to be regarded as unphysical in their matter content, the remaining class of models (based on scalar fields) generically admit the formation of visible regions of finite but arbitrarily high curvature. Hence it is of interest to ask, from the point of view of astrophysics, as to what a stellar collapse leading to a naked singularity (or to a visible region of very high curvature) will look like, to a far away observer. The emission of energy during such a process may be divided into three phases - (i) the classical phase, during which matter and gravity can both be treated according to the laws of classical physics, (ii) the semiclassical phase, when gravity is treated classically but matter behaves as a quantum field, and (iii) the quantum gravitational phase. In this review, we first give a summary of the status of naked singularities in classical relativity, and then report some recent results comparing the semiclassical phase of black holes with the semiclassical phase of spherical collapse leading to a naked singularity. In particular, we ask how the quantum particle creation during the collapse leading to a naked singularity compares with the Hawking radiation from a star collapsing to form a black hole. It turns out that there is a fundamental difference between the two cases. A spherical naked star emits only about one Planck energy during its semiclassical phase, and the further evolution can only be determined by the laws of quantum gravity. This contrasts with the semiclassical evaporation of a black hole wherein gravity can be treated classically all the way till the final stages of evaporation until a Planck mass remnant remains. Hence spherical collapse leading to a naked singularity provides an interesting and promising system for testing our understanding of quantum gravity. The results on the semiclassical phase of naked collapse reviewed here have been obtained in collaboration with Barve, Harada, Iguchi, Nakao, Tanaka, Vaz and Witten

1 Cosmic Censorship and Classical General Relativity

Put in broad physical terms, the Cosmic Censorship Hypothesis states that *generic* gravitational collapse of *physically reasonable* matter does not end in the formation of a *naked singularity*. A naked singularity is a singularity which is visible to a far away observer, i.e. outgoing light rays starting from the singularity terminate on the singularity in the past. By 'physically reasonable' we mean matter which satisfies one or more positive energy conditions - matter which can in principle be prepared in the laboratory. By 'generic' we mean that the initial conditions leading to the formation of the naked singularity are not special (of zero measure).

The Censorship Hypothesis is important in classical general relativity because there are theorems, for instance the black hole area theorem, whose proof assumes the validity of the Hypothesis. However, it is not obvious a priori that the results of these theorems cannot be proved without assuming Censorship. This by itself is an interesting direction for research in the classical theory.

The issue of naked singularities is not trivialized by quantum gravity, even though it might be true that quantum gravity will remove the singularities of classical general relativity, by replacing them by a smeared out region of Planck scale curvature. If the classical singularity resides inside an astrophysical black hole, such a smeared region will be invisible to a distant observer. However, if the classical singularity is naked,

[1] E-mail:tpsingh@nagaum.tifr.res.in

the Planck curvature region by which it will be replaced will be visible to the distant observer and will have physical consequences very different from that of the black hole. For instance, it could give rise to an intense particle production which could differ in character from the Hawking radiation from a black hole.

If we think of classical general relativity as a limit of quantum gravity, then we should think of Cosmic Censorship in a broader perspective, and take the Hypothesis to mean: 'Effects of quantum gravity cannot be observed in gravitational collapse'. Looked at in this way, it would be more useful for physics if Censorship were to be violated, as then the Universe could consist of naked stars - stars which according to the classical theory end as naked singularities, and in whose collapse we may witness the effects of quantum gravity.

Over the last twenty years or so, there have been many investigations of Cosmic Censorship, largely using spherical models of gravitational collapse of physically reasonable matter. The simplest of these models is dust collapse, which has been studied in great detail and it has been shown that both black holes and naked singularities form from generic initial conditions. Eardley and Smarr [1] used numerical methods to study the dust model. Christodoulou [2] analysed dust collapse with C^∞ initial data and Newman [3] showed that the consequent naked singularities were gravitationally weak. Dwivedi and Joshi developed a roots analysis to check naked singularity formation [4] and demonstrated the occurrence of strong curvature naked singularities in dust collapse. Waugh and Lake [6] showed the occurrence of strong curvature naked singularity in self-similar dust collapse. The initial data leading to these strong curvature naked singularities was worked out in [7] and a unified derivation of strong and weak naked singularities was obtained in [8]. The structure of the apparent horizon in dust collapse was studied in [9] and a simplified derivation of the dust naked singularity was given in [10]. Harada et al. [11, 12] investigated the stability of spherical dust collapse against non-spherical perturbations. The strength of the singularity has been studied by [13], the nature of non-spacelike geodesics in [14] and the role of initial data in [15, 16]. A cosmological constant has been included in [17, 18].

Similar results have been obtained for null dust collapse described by the Vaidya model [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30].

The next class of models is spherical perfect fluid collapse, which has been examined numerically [31] and analytically [32, 33] with and without the assumption of self-similarity [34, 35] and again it has been found that both black hole and naked singular solutions arise from generic initial data.

Another tractable class of models is the spherical collapse of fluids which have only tangential pressure, and for which the radial pressure is assumed to be negligible. For these models also it is observed that covered as well as naked solutions exist [36, 37, 38, 39, 40, 41, 42].

One might consider dividing physically reasonable matter models into two classes - fluids and fields. The former may be regarded as a phenomenological description of matter while the latter may be regarded as fundamental. Could it be that naked singularities do not arise generically in the collapse of matter which is treated as a field? The work of Christodoulou [43, 44, 45, 46, 47, 48, 49] on massless scalar fields suggests this could be the case. It should be said though that this issue of phenomenological versus fundamental needs further investigation, and it is not entirely clear why fluids should admit naked singularities but fields should not.

Even if one were to accept that Cosmic Censorship should be investigated with only fundamental matter, there is a broader outlook which in my view seals the case against Censorship. This has to do with the result that in spherical scalar collapse one can form black holes of arbitrarily small mass [50]. Hence, although the naked singularity itself forms from non-generic initial data, visible regions of arbitrarily high curvature (the curvature on the horizon of the arbitrarily small black hole) form generically.

So it may turn out to be true that the collapse of fundamental matter does not generically form naked singularities. Yet it appears to be more physical to think of Cosmic Censorship as a statement not about singularities, but about regions of extremely high curvature (approaching Planck scale) and in this sense censorship does not appear to hold in spherical general relativity, even if only fundamental matter is considered.

Attempts have also been made to study the nature of singularities in spherical collapse without having to restrict to a specific form of matter [51, 52, 53, 54, 55]

If we accept that Censorship does not hold, in the sense described in the previous paragraph, are there

any interesting implications for astrophysics? This is far from clear. We have to address the following questions: if a star collapses to form a naked singularity (or a black hole of very small mass) what does the process look like to a far off observer? How much energy is emitted and over what duration and with what spectrum? How does the back reaction from the emission modify the collapse? In this paper, we address some of these questions.

Let us broadly divide 'censorship violating' spherical collapse into two classes: one which results in a naked singularity at the center, and the second which results in a black hole of very small mass while the rest of the mass disperses to infinity. In this article we focus attention on the first of these two classes.

The emission from a star collapsing to form a naked singularity may be divided into three phases - classical, semiclassical and quantum gravitational. The classical phase would consist of the radiation which is emitted while both matter and gravity can be treated by the laws of classical physics. Could most of the star blow itself up during this phase? We do not know for sure. This phase has been investigated by Joshi, Dadhich and Maartens [56]. In the next phase of the collapse, matter will have to be treated as a quantum field whereas gravity can be treated classically (the semiclassical phase). This is the analog of the black hole phase in which the black hole emits Hawking radiation. The final collapse phase is quantum gravitational.

The main purpose of the present article is to report some recent results of Harada et al. [57] on the semiclassical phase of a spherical star forming a naked singularity. We show that the nature of particle creation in this phase is very different from the Hawking radiation from a black hole and that quantum gravity is essential for understanding the gravitational collapse to a naked singularity.

Some of the earlier reviews on Cosmic Censorship are listed in [58, 59, 60, 61, 62, 63, 64, 65]. The reader might find it useful to read the present article in conjunction with some of the earlier reviews.

2 Quantum Black Holes Versus Quantum Naked Singularities

Consider a spherical star which collapses to form a black hole and another spherical star which collapses to form a naked singularity. Let us quantize a massless scalar field on the background of either star, and prepare the field to be initially in the Minkowski vacuum, when the gravitational field of the star is weak enough so that spacetime is nearly flat. As is well known, the star which settles into the black hole state emits Hawking radiation, which has a thermal spectrum with temperature inversely proportional to the mass of the black hole.

We are interested in finding out the nature of the corresponding quantum radiation emitted in the naked singularity case. A comparison of the representative Penrose diagrams (Fig. 1) in the two cases

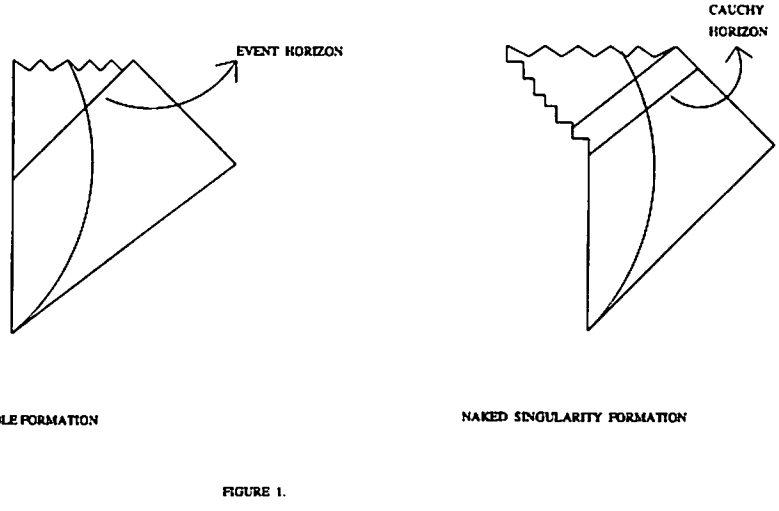


FIGURE 1.

shows the immediate obstacle to repeating Hawking's Bogoliubov transformation calculation in the naked case. In the latter case, a part of the asymptotic future null infinity is exposed to the naked null singularity,

making it impossible to define a complete set of modes on $\mathcal{I}^+$ unless some boundary conditions are prescribed on the naked singularity. Hence it is not possible to carry out the Bogoliubov transformations and a particle creation calculation the way it is done for the black hole case.

However, there is in principle no problem in calculating the expectation value of the quantum stress energy tensor of the scalar field in the region of the Penrose diagram prior to the Cauchy horizon. From this, one can obtain the expectation value of the outgoing quantum flux everywhere in that part of spacetime which is not exposed to the naked singularity, and in particular on that part of $\mathcal{I}^+$ which is prior to its intersection with the Cauchy horizon. Understandably, for the black hole case, such a calculation yields the standard Hawking thermal flux. We are now asking for the result of this calculation in the case when the star collapses to a naked singularity.

In the case of a four-dimensional spherical collapse Ford and Parker [66] obtained a formula for the outgoing quantum flux on the Cauchy horizon in the geometric optics approximation, using point-splitting regularisation. Suppose an ingoing ray $V = constant$ coming in from $\mathcal{I}^-$ gets mapped into an outgoing ray $U = \mathcal{F}(V)$ when it reaches $\mathcal{I}^+$. Then the outgoing quantum flux, for a given low angular momentum mode, is

$$P = \frac{\hbar}{24\pi}\left[\frac{\mathcal{F}'''}{\mathcal{F}'^3} - \frac{3}{2}\left(\frac{\mathcal{F}''}{\mathcal{F}'^2}\right)^2\right]. \tag{1}$$

In terms of the inverse map $V = \mathcal{G}(U)$ the power is given by

$$P = \frac{\hbar}{24\pi}\left[\frac{3}{2}\left(\frac{\mathcal{G}''}{\mathcal{G}'^2}\right)^2 - \frac{\mathcal{G}'''}{\mathcal{G}'}\right]. \tag{2}$$

For a two-dimensional spacetime model, the outgoing quantum flux can be calculated exactly. Suppose the spacetime is given by double null coordinates u, v as

$$ds^2 = C^2(u,v)dudv. \tag{3}$$

Then, the use of the trace anomaly and the divergencelessness condition

$$Tr < in|T_{\mu\nu}|in >= \frac{\mathcal{R}}{24\pi}, \qquad < T^{\mu\nu} >_{;\nu} = 0 \tag{4}$$

allows the outgoing flux to be written as [67]

$$< T_{uu} >= -\frac{1}{12\pi}C\frac{\partial^2}{\partial u^2}(1/C). \tag{5}$$

Here, $\mathcal{R}$ is the curvature scalar for the two dimensional spacetime. In the four dimensional case as well as for the 2-d approximation, one needs to be able to solve for null rays in order to calculate the flux. We now discuss these two approximations one by one. Before doing that, we give a brief outline of the naked singularity in the classical spherical dust collapse model [10].

Spherical Dust Collapse

In comoving coordinates (t, r, θ, ϕ) the spacetime metric for marginally bound spherical dust collapse is given by

$$ds^2 = dt^2 - R'^2dr^2 - R^2d\Omega^2 \tag{6}$$

where $R(t,r)$ is the area radius at time t of the shell having the comoving coordinate r. A prime denotes partial derivative w.r.t. r. The energy-momentum tensor for dust has only one non-zero component $T^0_0 = \epsilon(t,r)$, which is the energy density. The Einstein equations for the collapsing cloud are

$$\frac{8\pi G}{c^4}\epsilon(t,r) = \frac{F'}{R^2R'}, \quad \dot{R}^2 = \frac{F(r)}{R}. \tag{7}$$

A dot denotes partial derivative w.r.t. time t. The function $F(r)$ results from the integration of the second order equations. Henceforth we shall set $8\pi G/c^4 = 1$.

The second of these equations can be easily solved to get

$$R^{3/2}(t,r) = r^{3/2} - \frac{3}{2}\sqrt{F}t \tag{8}$$

where we have used the freedom in the scaling of the comoving coordinate r to set $R(0,r) = r$ at the starting epoch of collapse, $t = 0$. It follows from the first equation in (7) that the function $F(r)$ gets fixed once the initial density distribution $\epsilon(0,r) = \rho(r)$ is given, i.e.

$$F(r) = \int \rho(r) r^2 dr. \tag{9}$$

Hence $F(r)$ has the interpretation of being twice the mass to the interior of the shell labeled r. If the initial density $\rho(r)$ has a series expansion

$$\rho(r) = \rho_0 + \rho_1 r + \frac{1}{2!}\rho_2 r^2 + \frac{1}{3!}\rho_3 r^3 + ... \tag{10}$$

near the center $r = 0$, the resulting series expansion for the mass function $F(r)$ is

$$F(r) = F_0 r^3 + F_1 r^4 + F_2 r^5 + F_3 r^6 + ... \tag{11}$$

where $F_q = \rho_q/q!(q+3)$, and $q = 0, 1, 2, 3..$We note that we could set $\rho_1 = 0$ without in any way affecting the conclusions of this paper. Further, the first non-vanishing derivative in the series expansion in (10) should be negative, as we will consider only density functions which decrease as one moves out from the center.

According to (8) the area radius of the shell r shrinks to zero at the time $t_c(r)$ given by

$$t_c(r) = \frac{2r^{3/2}}{3\sqrt{F(r)}}. \tag{12}$$

At $t = t_c(r)$ the Kretschmann scalar

$$K = 12\frac{F'^2}{R^4 R'^2} - 32\frac{FF'}{R^5 R'} + 48\frac{F^2}{R^6} \tag{13}$$

diverges at the shell labeled r and hence this represents the formation of a curvature singularity at r. In particular, the central singularity, i.e. the one at $r = 0$, forms at the time

$$t_0 = \frac{2}{3\sqrt{F_0}} = \frac{2}{\sqrt{3\rho_0}}. \tag{14}$$

At $t = t_0$ the Kretschmann scalar diverges at $r = 0$. Near $r = 0$, we can expand $F(r)$ and approximately write for the singularity curve

$$t_c(r) = t_0 - \frac{F_n}{3F_0^{3/2}} r^n. \tag{15}$$

Here, F_n is the first non-vanishing term beyond F_0 in the expansion (11). We note that $t_c(r) > t_0$, since F_n is negative.

We wish to investigate if the singularity at $t = t_0, r = 0$ is naked, i.e. are there one or more outgoing null geodesics which terminate in the past at the central singularity. We restrict attention to radial null geodesics. Let us start by assuming that one or more such geodesics exist, and then checking if this assumption is correct. Let us take the geodesic to have the form

$$t = t_0 + ar^\alpha \tag{16}$$

to leading order, in the $t - r$ plane, where $a > 0$, $\alpha > 0$. In order for this geodesic to lie in the spacetime, we conclude by comparing with (15) that $\alpha \geq n$, and in addition, if $\alpha = n$, then $a < -F_n/3F_0^{3/2}$.

As is evident from the form (6) of the metric, an outgoing null geodesic must satisfy the equation

$$\frac{dt}{dr} = R'. \tag{17}$$

In order to calculate R' near $r = 0$ we first write the solution (8) with only the leading term F_n retained in $F(r)$ in (11). This gives

$$R = r\left(1 - \frac{3}{2}\sqrt{F_0}\left[1 + \frac{F_n}{2F_0}r^n\right]t\right)^{2/3}. \tag{18}$$

Differentiating this w.r.t. r gives

$$R' = \left(1 - \frac{3}{2}\sqrt{F_0}\left[1 + \frac{F_n}{2F_0}r^n\right]t\right)^{-1/3}\left(1 - \frac{3}{2}\sqrt{F_0}t - \frac{(2n+3)F_n}{4\sqrt{F_0}}r^n t\right). \tag{19}$$

Along the assumed geodesic, t is given by (16). Substituting this in R' and equating the resulting R' to $dt/dr = \alpha a r^{\alpha-1}$ gives

$$\alpha a r^{\alpha-1} = \frac{\left(1 - \frac{3}{2}\sqrt{F_0}\left[t_0 + ar^\alpha\right] - \frac{(2n+3)F_n}{4\sqrt{F_0}}r^n\left[t_0 + ar^\alpha\right]\right)}{\left(1 - \frac{3}{2}\sqrt{F_0}\left[1 + \frac{F_n}{2F_0}r^n\right]\left[t_0 + ar^\alpha\right]\right)^{1/3}}. \tag{20}$$

This is the key equation. If it admits a self-consistent solution then the singularity will be naked (i.e. at least one outgoing null geodesic will terminate at the singularity), otherwise not. We simplify this equation by putting in the requirement mentioned earlier, that $\alpha \geq n$. Consider first $\alpha > n$. In this case we get, to leading order

$$\alpha a r^{\alpha-1} = \left(1 + \frac{2n}{3}\right)\left(-\frac{F_n}{2F_0}\right)^{2/3} r^{2n/3} \tag{21}$$

which implies that $\alpha = 1 + 2n/3$, and $a = (-F_n/2F_0)^{2/3}$. By substituting integral values for n we find that only for $n = 1$ and $n = 2$ the condition $\alpha > n$ is satisfied. Hence the singularity is naked for $n = 1$ and $n = 2$, i.e. for the models $\rho_1 < 0$ and for $\rho_1 = 0, \rho_2 < 0$. There is at least one outgoing geodesic given by (16) which terminates in the central singularity in the past. If $n > 3$ then the condition $\alpha > n$ cannot be satisfied and the singularity is not naked. This is the case $\rho_1 = \rho_2 = \rho_3 = 0$. It includes as a special case the homogeneous black hole model of Oppenheimer and Snyder.

Consider next that $\alpha = n$. In this case we get from (20) that

$$n a r^{n-1} = \frac{-\frac{3}{2}a\sqrt{F_0} - \frac{(2n+3)F_n}{6\sqrt{F_0}}}{\left(-\frac{F_n}{2F_0} - \frac{3a}{2}\sqrt{F_0}\right)^{1/3}} r^{2n/3} \tag{22}$$

which implies that $n = 3$ and gives an implicit expression for a in terms of F_3 and F_0. This expression for a can be simplified to get the following quartic for a:

$$12\sqrt{F_0}a^4 - a^3(-4F_3/F_0 + F_0^{3/2}) - 3F_3a^2 - 3F_3^2/F_0^{3/2}a - (F_3/F_0)^3 = 0. \tag{23}$$

By defining $b = a/F_0$ and $\xi = F_3/F_0^{5/2}$ this quartic can be written as

$$4b^3(3b + \xi) - (b + \xi)^3 = 0. \tag{24}$$

The singularity will be naked if this equation admits one or more positive roots for b which satisfy the constraint $b < -\xi/3$. This last inequality is the same as the condition $a < -F_n/3F_0^{3/2}$ given below equation (16). We note that ξ is negative. This quartic can be made amenable to further analysis by substituting $Y = -2b/\xi$, and then $\eta = -1/6\xi$, so as to get

$$Y^3(Y - 2/3) - \eta(Y - 2)^3 = 0. \tag{25}$$

This quartic has two positive real roots provided $\eta \geq \eta_1$ or $\eta \leq \eta_2$ where

$$\eta_1 = \frac{26}{3} + 5\sqrt{3},\ \eta_2 = \frac{26}{3} - 5\sqrt{3}. \tag{26}$$

We also require that $Y < 2/3$. By examining the quartic (25) one can see that if $\eta \geq \eta_1$ then $Y \geq 2$; hence this range of η is ruled out. Thus the singularity is naked provided $\eta \leq \eta_2$, or equivalently $\xi \leq -25.9904$.

The above results give the conditions on the initial data for which the singularity is at least locally naked, i.e. light manages to escape from the singularity. Further constraints have to be put on the initial distribution in order to obtain a globally naked singularity, i.e. one in which light rays starting from the singularity escape the dust cloud entirely. In general a choice of initial conditions is available for which a singularity which is at least locally naked is also globally naked. The spacetime with a globally naked singularity is described by the Penrose diagram in Figure 1. From the point of view of the present article the globally naked singularities are the interesting ones; they are the ones whose asymptotic quantum appearance differs from those of black holes.

A useful special case of the class $\alpha = n = 3$ is the self-similar model, one which admits a homothetic Killing vector field [31]. The condition for nakedness is the same as for the whole class $\alpha = n = 3$ and the singularity is necessarily globally naked.

We recall next the quantisation of a massless scalar field on the background spacetime provided by dust collapse.

Geometric Optics Approximation

We now consider the application of the Ford-Parker formula (2) for calculating the quantum flux of a massless scalar field on the dust background. For the case in which the collapse ends in a black hole, the map $\mathcal{G}(U)$ is given by

$$\mathcal{G}(U) = A - B\exp(-U/4M). \tag{27}$$

Using this map in (2) gives the expected thermal Hawking flux

$$P = \frac{1}{768\pi M^2}. \tag{28}$$

The Hawking flux for the black hole is found using the map $\mathcal{F}(v)$ which in turn depends on the behaviour of null rays in the vicinity of the event horizon. In this sense the map is universal for black holes - it does not depend on the interior geometry of the star.

On the other hand the map $\mathcal{G}(U)$ very much depends on the internal geometry in the naked case. Hence one has to first specify the internal solution and calculate the propagation of null rays inside the star. There are only a limited number of cases where this has been done so far. The first example to be considered was that of shell-crossing dust naked singularity [66] for which the outgoing flux was calculated and shown to be finite.

Of greater interest is the shell-focusing naked central singularity in spherical dust collapse, described above. This was considered by Barve et al. [68] for the self-similar dust collapse and by Singh and Vaz [69] for self-similar null dust collapse. The map $\mathcal{G}(U)$ was calculated and shown to be

$$\mathcal{G}(U) = A - B(V_0 - V)^\gamma \tag{29}$$

where γ is a positive constant determined by the collapse model. The flux can be calculated using (2) and near the Cauchy horizon $U = U_0$ it behaves as

$$P(U) \sim \frac{\hbar}{(U_0 - U)^2}. \tag{30}$$

It thus diverges in the approach to the Cauchy horizon. Such a divergence appears to be a universal feature of collapse leading to a naked singularity. It has also been observed in dust collapse models which do not assume self-similarity [70, 71], for which the divergence is of the form

$$P(U) \sim \frac{\hbar}{(U_0 - U)^{3/2}}. \tag{31}$$

What is the meaning of this divergence? At first sight, it suggests that as a result of the intense particle creation accompanying the formation of the naked singularity, a huge amount of energy is emitted - essentially the whole star blows up. However, we realize that in inferring this divergence we have extended the semiclassical approximation all the way up to the epoch of naked singularity formation, when the curvature of the background spacetime diverges. But we should not a priori trust the results of the semiclassical theory beyond the epoch when the curvature in the central region of the star approaches Planck scale. This happens about one Planck time before the singular epoch, and this can be shown [57] to translate to $U_0 - U \approx t_{Planck}$ in (30). It then follows from (30) that during the domain of validity of the semiclassical approximation the emitted energy is only about one Planck energy, even if the total initial mass of the star is much larger, say a solar mass. Hence this system is very different from the black hole, for which essentially the entire mass of the collapsing object evaporates during the semiclassical phase and Planck scale curvatures are realised only when a Planck mass remnant remains.

The result that the semiclassical energy emission is very small appears to be supported by the observation that the map $\mathcal{F}(V)$ in the naked case is determined by examining a small spacetime region near the central singularity [72].

We mentioned earlier that the calculation of the spectrum of created particles, which is based on Bogoliubov transformations, cannot be performed in the naked case without specifying boundary conditions on the singularity. If the semiclassical approximation were assumed to be valid all the way up to the formation of the naked singularity, the resulting divergence on the Cauchy horizon would suggest a complete evaporation of the star. Under such a situation it appears reasonable to assume that the singularity does not form in the first place, and an appropriate boundary condition would be that of no emission from the null singularity. Under these assumptions it has been shown that the emitted spectrum is non-thermal [73, 74, 71]. It however remains to be understood how these results on the spectrum are modified by the quantum gravity phase which comes into play about a Planck time prior to the singularity.

2-d collapse models

The 2-d models that we consider are obtained by suppressing the angular coordinates in a 4-d spherical model. If one applies the 2-d quantum flux result to the Schwarzschild geometry, one again obtains the Hawking result (28) for the power. The first application to a naked singularity model was by Hiscock et al. [19] who calculated the outgoing flux for the naked singularity in self-similar Vaidya null dust collapse and showed it to diverge on the Cauchy horizon. A similar divergence was found by Barve et al. [75] for the case of self-similar Tolman-Bondi dust collapse:

$$< T_{uu} > \sim \frac{\hbar}{(z - z_-)^2} \tag{32}$$

where z is the self-similarity parameter and $z = z_-$ is the Cauchy horizon. It was shown in [76] that the divergence on the Cauchy horizon persists in non-self-similar 2-d dust collapse.

Once again, it is apparent that the semiclassical evolution should be stopped about a Planck time before the singular epoch. It then turns out that in the 2-d model as well the energy emitted during the semiclassical phase is about one Planck energy.

An important aspect of the above analysis is that we have ignored the back-reaction, and one could well ask if our conclusions could be affected by its inclusion. We now give arguments as to why the back reaction cannot be important during the semiclassical evolution, in the present case. One could propose two different criteria for deciding as to when the back reaction becomes important. One is that the total flux received at infinity becomes comparable to the mass of the collapsing star. As we have seen above, if the mass of the collapsing star is much greater than the Planck mass (as is of course usually the case) then the back reaction does not become important during the semiclassical phase.

The second criterion could be that the energy density of the quantum field becomes comparable to the background energy density, *inside* the star. In a four-dimensional model, it is difficult to establish whether this happens. However, one can study the evolution of the quantum field inside the star using the 2-d model obtained by suppressing the angular part of the 4-d Tolman-Bondi model [75]. Using this model, Iguchi and Harada [77] have shown that the back reaction does not become significant during the semiclassical evolution, inside the star. It is plausible that this result holds for the four dimensional

stellar dust model as well, in which case we can conclude that quantum gravitational effects are more important than the back reaction in deciding the evolution of the star.

Again we see that the semiclassical evolution is very different from the black hole case, since the back reaction plays a crucial role during the semiclassical history of the black hole.

3 Discussion and Unresolved Issues

We see that when a spherical dust star collapses to form a naked singularity, the character of the semiclassical emission (resulting from the quantization of a massless scalar field on the dust background) is very different from the Hawking emission from a dust star collapsing to form a black hole. The duration of the semiclassical phase for the naked case is of the order of the collapse time $\sim 1/\sqrt{G\rho}$, in the comoving frame as well as for the asymptotic observer. (Since there is no event horizon preceding the formation of the singularity, there is no significant time-dilation caused by the gravitational redshift on the boundary). This time scale is of course extremely short, compared to the astronomical time scale for the evaporation of an astrophysical black hole.

With hindsight, it is not difficult to see why the semiclassical phase of a naked collapse is so different from that of collapse leading to a black hole. The central naked singularity forms extremely quickly, as seen by a far away observer, and this naked central spacetime region can only be dealt with using quantum gravity. In contrast, the asymptotic observer never gets to see the singularity inside the black hole, except during the final stages of black hole evaporation. Hence the black hole semiclassical epoch is extremely long.

Starting at about one Planck time before singularity formation, we must analyse the further evolution of the naked star using the laws of the (yet unknown) quantum theory of gravity. We thus have a novel situation where a spherical dust star (say of one solar mass) enters the quantum gravity phase with essentially all its mass intact, and unlike the black hole, it does so very much within the lifetime of the Universe. This is perhaps the first time that physicists have explicitly encountered a dynamical system whose further evolution, within the age of the Universe, cannot be understood without a knowledge of quantum gravity! For an alternative discussion of Cosmic Censorship and quantum gravity see [78]. Naked singularities in a dilaton gravity context are discussed in [79, 80].

What happens to this spherical dust star next? How does the apparently 'quantum gravitational' region near the center interact with the apparently classical regions which are away from the center? Does the entire star evaporate explosively, as suggested by the semiclassical theory, or does it settle into a black hole and then emit slowly by the standard Hawking radiation? Is there a way in the quantum theory to separate the evolution of the central region from that of the outer region? These are fascinating and challenging conceptual questions to which there are no obvious answers, and which perhaps cannot be addressed without quantum gravity. Nonetheless, they provide a valuable testbed for examining candidate theories of quantum gravity.

Motivated by such issues, Vaz et al. [81] have begun an investigation of quantum gravitational effects in spherical dust collapse leading to a naked singularity. As a midisuperspace quantum gravity model, one sets up a Wheeler-deWitt equation for the collapse, in terms of physically well-motivated canonical variables. This approach is inspired by Kuchar's canonical formalism for the Schwarzschild black hole [82]. By investigating the solutions of the Wheeler-deWitt equation for this system one hopes to address the questions raised in the previous paragraph.

As an aside, one could marvel at the richness of the dust collapse system, whose studies were first initiated by Oppenheimer and Snyder, way back in 1939. Their work on homogeneous collapse was over the years generalised to include shell-crossings and inhomogeneities, and then used as a background model in quantum field theory in curved space, and now as a quantum gravity model! We can be sure that we have not yet learnt all that dust collapse has to teach us.

Does this picture of semiclassical naked collapse persist for spherical naked singularities that form with other kinds of matter, say fluids? We think it does, because the semiclassical properties are related to the occurrence of the naked singularity in the small central region - this is a property generic to the formation of a naked singularity in spherical collapse. It is not restricted to the choice of dust as matter. Hence we believe that the choice of dust as such is irrelevant to the importance of quantum gravity in spherical naked collapse.

The other class of censorship violating spherical models which we mentioned at the beginning of this article are those which form small mass black holes, and for which the remaining mass escapes to infinity. What kind of results does one expect for the semiclassical quantization of such systems? We would like to conjecture that the semiclassical phase will be very similar to the dust naked singularity system. The small mass black hole represents a region of extremely high (Planck scale) curvature which cannot be dealt with semiclassically, and must be analysed using quantum gravity. Furthermore, this region (unlike for an evaporating black hole) is realised on a very short gravitational collapse time scale. The only physical difference from the dust case is the dispersion of the outer region, which does not constitute a significant difference in so far as the semiclassical quantisation is concerned.

Hence we can say with confidence that the picture of semiclassical spherical collapse which we have seen in the dust model is representative of censorship violating spherical collapse. The semiclassical phase of such collapse emits an insignificant amount of energy and the full evolution can only be understood in a quantum gravity framework.

Throughout the article, we have carefully emphasized that these semiclassical results apply only to naked singularities forming in spherical collapse. The picture could well be very different if naked singularities were to form in non-spherical collapse, for instance in the collapse to a elongated cigar shaped naked singularity. Under such a situation an entire region (and not just a limited central region) develops visible high curvature, and the particle creation and energy release during the process could be enormous and explosive. This is an important outstanding problem - is the semiclassical phase of highly non-spherical naked collapse similar to or very different from the semiclassical phase of naked spherical collapse? Also, the nature of gravity wave emission during the formation of aspherical naked singularities [83, 84] is a promising research problem. Of course, it should be added that our knowledge about formation of non-spherical naked singularities is at present much more restricted compared to spherical ones [85, 86, 87, 88, 89].

Finally, one could well ask whether its worth anyone's effort to investigate the possible occurrence and astrophysical properties of naked singularities. An instructive answer can be found in the history of black hole physics. Many decades ago, when a handful of relativists were trying to understand the physical properties of black holes, very few astronomers and astrophysicists took them seriously - and their efforts were often regarded as exercises in mathematical physics, irrelevant to the real world of astronomical observations. Needless to say, all this has changed with the detection of black holes in the centers of galaxies and in x-ray binaries. But its fair to add that an understanding of these detections would still be long in coming had relativists not prepared the ground by working out what black holes would look like, if they were to exist.

With regard to research on naked singularities and naked stars, we are where black hole physics was, say seventy years ago. Only further research can tell whether naked stars occur in nature. Considering the potential role of such systems as experimental testbeds for quantum gravity it seems worth the effort to explore their consequences. The possibility that gamma-ray bursts and/or some yet undiscovered astronomical object can be identified with naked stars cannot be ruled out at present [90, 91]. Even if the results of such searches turn out to be negative, one will at least have found a route to proving Cosmic Censorship.

Acknowledgments: It is a pleasure to thank the Yukawa Institute for Theoretical Physics, Kyoto for its warm hospitality and the organisers of JGRG10 for a lively and fruitful meeting. I would also like to thank my collaborators Sukratu Barve, Tomohiro Harada, Hideo Iguchi, Sanjay Jhingan, Pankaj Joshi, Ken-ichi Nakao, Takahiro Tanaka, Cenalo Vaz and Louis Witten. I am grateful to Hideo Kodama and Takashi Nakamura for their support and encouragement. I acknowledge the partial support of the Fundação para a Ciência e a Tecnologia (FCT), Portugal under contract number SAPIENS/32694/99.

A note on the references: No claim is made that the references are exhaustive. If there are papers references to which have been inadvertently overlooked, I would like to apologise for the same.

References

[1] D. M. Eardley and L. Smarr (1979) Phys. Rev. D **19**, 2239.

[2] D. Christodoulou (1984) Commun. Math. Phys. **93** 171.

[3] R. P. A. C. Newman, Class. Quantum Grav. **3**, 527 (1986).

[4] P. S. Joshi and I. H. Dwivedi (1993) Phys. Rev. **D47** 5357.

[5] I. H. Dwivedi and P. S. Joshi (1992) Class. Quantum Grav. **9** L69.

[6] B. Waugh and K. Lake (1988) Phys. Rev. D **38** 1315.

[7] P. S. Joshi and T. P. Singh (1995) Phys. Rev. **D51** 6778 [gr-qc/9405036].

[8] T. P. Singh and P. S. Joshi (1996) Class. Quantum Grav. **13** 559 [gr-qc/9409062].

[9] S. Jhingan, P. S. Joshi and T. P. Singh (1996) Class. Quantum Grav. **13** 3057 [gr-qc/9604046].

[10] S. Barve, T. P. Singh, C. Vaz and L. Witten (1999) Class. Quantum Grav. **16** 1727 [gr-qc/9901054].

[11] H. Iguchi, K. Nakao and T. Harada (1998), Phys. Rev. **D57** 7262 [gr-qc/9804015].

[12] H. Ken-ichi Nakao, T. Harada, (1998) Phys. Rev. **D57** 7262 [gr-qc/9804015].

[13] S.S. Deshingkar, I.H. Dwivedi and P.S. Joshi, (1999) Phys. Rev. **D59** 044018 [gr-qc/9805055].

[14] S.S. Deshingkar and P.S. Joshi, (2001) Phys.Rev. **D63** 024007 [gr-qc/0010015].

[15] I.H. Dwivedi and P.S. Joshi (1997) Class. Quantum Grav **14** 1223 [gr-qc/9612023].

[16] F. C. Mena, R. Tavakol and P. S. Joshi, (2000) Phys. Rev. **D62** 044001 [gr-qc/0002062].

[17] S.S. Deshingkar, A. Chamorro, S. Jhingan and P.S. Joshi, (2000) [gr-qc/0010027].

[18] Sergio M.C.V. Goncalves (2000) [gr-qc/0012032].

[19] W. A. Hiscock, L. G. Williams and D. M. Eardley (1982) Phys. Rev. **D26** 751.

[20] A. Papapetrou (1985) in A Random Walk in General Relativity, Eds. N. Dadhich, J. K. Rao, J. V. Narlikar and C. V. Vishveshwara (Wiley Eastern, New Delhi).

[21] G. P. Hollier (1986) Class. Quantum Grav. **3** L111.

[22] W. Israel (1986) Can. Jour. Phys. **64** 120.

[23] K. Rajagopal and K. Lake (1987) Phys. Rev. **D35** 1531.

[24] I. H Dwivedi and P. S. Joshi (1989) Class. Quantum Grav. **6** 1599.

[25] I. H Dwivedi and P. S. Joshi (1991) Class. Quantum Grav. **8** 1339.

[26] P. S. Joshi and I. H. Dwivedi (1992) Gen. Rel. Gravn. **24** 129.

[27] J. Lemos (1992) Phys. Rev. Lett. **68** 1447.

[28] S.M. Wagh and S.D. Maharaj (1999) Gen. Rel. Grav. **31** 975 [gr-qc/9903083].

[29] S.G. Ghosh and A. Beesham (2000) Phys. Rev. **D61** 067502.

[30] S. G. Ghosh and R. V. Saraykar (2000) Phys. Rev. **D62** 107502.

[31] A. Ori and T. Piran,(1990) Phys. Rev. **D42** 1068.

[32] P. S. Joshi and I. H. Dwivedi (1992) Commun. Math. Phys. **146** 333.

[33] P. S. Joshi and I. H. Dwivedi (1993) Lett. Math. Phys. **27** 235.

[34] T. Harada (1998) Phys. Rev. **D58** 104015 [gr-qc/9807038]

[35] F.I. Cooperstock, S. Jhingan, P.S. Joshi, T.P. Singh (1997) Class. Quantum Grav. **14** 2195 [gr-qc/9609051].

[36] G. Magli (1997) Class. Quantum Grav. **14** 1937.

[37] T.P. Singh and Louis Witten (1997) Class. Quantum Grav. **14** 3489 [gr-qc/9701002].

[38] T. Harada, H. Iguchi and K. Nakao (1998) Phys. Rev. **D58** 041502 [gr-qc/9805071].

[39] T. Harada, K. Nakao and H. Iguchi (1999) Class. Quantum Grav. **16** 2785 [gr-qc/9904073].

[40] H. Kudoh, T. Harada, H. Iguchi, (2000) Phys. Rev. **D62** 104016 [gr-qc/0006089].

[41] S. Jhingan and G. Magli (2000) Phys. Rev. **D61** 124006.

[42] S. Barve, T.P. Singh and Louis Witten (2000) Gen. Rel. Grav. **32** 697 [gr-qc/9901080].

[43] D. Christodoulou (1986) Commun. Math. Phys. **105** 337.

[44] D. Christodoulou (1986) Commun. Math. Phys. **106** 587.

[45] D. Christodoulou (1987) Commun. Math. Phys. **109** 591.

[46] D. Christodoulou (1991) Commun. Pure Appl. Math. **XLIV** 339.

[47] D. Christodoulou (1993) Commun. Pure Appl. Math. **XLVI** 1131.

[48] D. Christodoulou (1994) Ann. Math. **140** 607.

[49] D. Christodoulou, (1999) Annals of Mathematics **149** 183.

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

[51] K. Lake (1992) Phys. Rev. Lett. **68** 3129.

[52] I. H. Dwivedi and P. S. Joshi (1994) Commun. Math. Phys. **166** 117 [gr-qc/9405049].

[53] T.P. Singh (1998) Phys. Rev. **D58** 024004 [gr-qc/9711049].

[54] T.P. Singh (1999) Class. Quantum. Grav. **16** 3307 [gr-qc/9808003].

[55] Brien C. Nolan (2000) [gr-qc/0010032].

[56] P. S. Joshi, N. K. Dadhich and R. Maartens (2000) Mod. Phys. Lett. **A15** 991 [gr-qc/0005080].

[57] T. Harada, H. Iguchi, K. Nakao, T. P. Singh, T. Tanaka and C. Vaz (2000) [gr-qc/0010101].

[58] P. S. Joshi (1993) Global Aspects in Gravitation and Cosmology, Clarendon Press, OUP, Oxford.

[59] T. P. Singh in *Classical and Quantum Aspects of Gravitation and Cosmology* ed. G. Date and B. R. Iyer (Inst. of Math. Sc., Madras, 1996), [gr-qc/9606016].

[60] P. S. Joshi in *Singularities, Black Holes and Cosmic Censorship* ed. P. S. Joshi (IUCAA, Pune, 1997), [gr-qc/9702036].

[61] R. Penrose in *Black Holes and Relativistic Stars* ed. R. M. Wald (Chicago University Press, 1998).

[62] R. M. Wald (1999) in *Black Holes, Gravitational Radiation and the Universe* ed. B. R. Iyer and B. Bhawal [gr-qc/9710068].

[63] Andrzej Krolak (1999) Prog. Theor. Phys. Suppl. **136** 45 [gr-qc/9910108].

[64] T. P. Singh (1999) J. Astrophys. Astr. **20** 221 [gr-qc/9805066].

[65] P. S. Joshi (2000) Pramana **55** 529 [gr-qc/0006101].

[66] L. H. Ford and L. Parker (1978) Phys. Rev. **D17** 1485.

[67] P. C. W. Davies, S. A. Fulling and W. G. Unruh (1976) Phys. Rev. **D13** 2720.

[68] S. Barve, T.P. Singh, C. Vaz and Louis Witten (1998) Nucl. Phys. **B532** [gr-qc/9802035].

[69] T.P. Singh and C. Vaz (2000) Phys. Lett. **B481** 74 [gr-qc/0002018].

[70] T. Harada. H. Iguchi and K. Nakao, (2000) Phys. Rev. **D61** 101502 [gr-qc/0003036].

[71] T. Harada. H. Iguchi and K. Nakao, (2000) Phys. Rev. **D62** 084037 [gr-qc/0005114].

[72] T. Tanaka and T.P. Singh, (2000) [gr-qc/0010110].

[73] C. Vaz and L. Witten, (1998) Phys. Lett. B442 90 [gr-qc/9804001].

[74] T. P. Singh and C. Vaz, (2000) Phys. Rev. **D61** 124005 [gr-qc/9912063].

[75] S. Barve, T.P. Singh, C. Vaz and L. Witten (1998) Phys. Rev. **D58** 104018 [gr-qc/9805095].

[76] S. Barve, T.P. Singh and C. Vaz (2000) Phys. Rev. **D62** 084021 [gr-qc/0001085].

[77] H. Iguchi and T. Harada, in preparation.

[78] Shahar Hod and Tsvi Piran (2000) [gr-qc/0011003].

[79] Cenalo Vaz and Louis Witten (1996) Class. Quantum. Grav. **13** L59 [gr-qc/9511018].

[80] Cenalo Vaz and Louis Witten (1997) Nucl. Phys. **B487** 409 [hep-th/9604064].

[81] C. Vaz, L. Witten and T.P. Singh (2000) [gr-qc/0012053].

[82] K. V. Kuchar, (1994) Phys. Rev. **D50** 3961.

[83] T. Nakamura, M. Shibata and K. Nakao, (1993) Prog. Theo. Phys. **89** 821.

[84] K. Thorne in *Relativistic Astrophysics* ed. B. J. T. Jones and D. Markovic (Cambridge University Press, 1997).

[85] S. L. Shapiro and S. A. Teukolsky (1991) Phys. Rev. Lett. **66** 994.

[86] C. Barrabès, W. Israel and P. S. Letelier (1991) Phys. Lett. **A160** 41.

[87] P. Szekeres (1975) Phys. Rev. **D12** 2941.

[88] P. S. Joshi and A. Krolak (1996) Class. Quantum Grav. **13** 3069 [gr-qc/9605033].

[89] S.S. Deshingkar, S. Jhingan and P.S. Joshi (1998) Gen. Rel. Grav. **30** 1477 [gr-qc/9806055].

[90] S. K. Chakrabarti and P. S. Joshi (1994) Int. Jour. Mod. Phys. **D3** 647 [hep-th/9208060].

[91] T. P. Singh, (1998) Gen. Rel. Grav. **30** 1563-1567 [gr-qc/9805062].

Thermodynamics of Near-Extreme Black Holes

Don N. Page[1]

CIAR Cosmology Program
Department of Physics
University of Alberta
Edmonton, Alberta
Canada T6G 2J1

Abstract

The thermodynamics of nearly-extreme charged black holes depends upon the number of ground states at fixed large charge and upon the distribution of excited energy states. Here three possibilities are examined: (1) Ground state highly degenerate (as suggested by the large semiclassical Hawking entropy of an extreme Reissner-Nordstrom black hole), excited states not. (2) All energy levels highly degenerate, with macroscopic energy gaps between them. (3) All states nondegenerate (or with low degeneracy), separated by exponentially tiny energy gaps. I suggest that in our world with broken supersymmetry, this last possibility seems most plausible. An experiment is proposed to distinguish between these possibilities, but it would take a time that is here calculated to be more than about 10^{837} years.

1 Introduction

What are the energy levels of a charged black hole? For fixed electric charge $Q \gg 1$, the minimum mass of a classical electrically charged black hole (given by the Kerr-Newman metric, or, in the nonrotating case, by the Reissner-Nordstrom metric) is $M = Q$. In this paper, we shall examine the consequences of various assumption for the levels of the excess energy

$$E \equiv M - Q. \tag{1}$$

(Strictly speaking, vacuum polarization is likely to shift the minimum mass for fixed charge Q slightly away from Q if the theory does not have unbroken supersymmetry, in which case in Eq. (1), and in many equations below, one should replace Q by the minimum for the mass M for the fixed charge. However, for simplicity I shall here and henceforth simply call this minimum mass Q.)

At $E = 0$, one has (classically) an extreme-charged Reissner-Nordstrom black hole with event horizon radius $r_+ = M = Q$ and hence Bekenstein-Hawking entropy

$$S_0 \equiv S_H(E = 0) = A/4 = \pi Q^2 \gg 1. \tag{2}$$

This formula for the entropy of an extreme black hole is also supported by D-brane calculations for black holes in superstring theory with unbroken supersymmetry. It suggests that the ground state of a black hole with large fixed charge is highly degenerate. On the other hand, there have been counter-arguments [1, 2] suggesting that the entropy of an extreme black hole is zero. Which idea is correct?

There seem to be three extreme possibilities in our world with broken supersymmetry, though of course intermediate possibilities also exist: (1) Ground state highly degenerate, excited states not. (2) All states highly degenerate, with macroscopic mass gaps. (3) All states nondegenerate, separately microscopically. Here I wish to examine some of the thermodynamic consequences of these possibilities.

[1] E-mail:don@phys.ualberta.ca

2 Semiclassical Thermodynamics

For classical Kerr-Newman black holes with given Q and E (or $M = Q + E$), the maximum Bekenstein-Hawking entropy is obtained if the black holes have zero angular momentum and are hence Reissner-Nordstrom black holes. Therefore, states with zero (or at least relatively small) angular momentum might be expected to dominate in number for a given Q and range of E, and so we shall focus on such black holes.

The classical Reissner-Nordstrom metric is

$$ds^2 = -\frac{\Delta}{r^2}dt^2 + \frac{r^2}{\Delta}dr^2 + r^2 d\Omega^2, \tag{3}$$

where $d\Omega^2$ is the metric on the unit S^2 and where

$$\Delta \equiv r^2 - 2Mr + Q^2 \equiv (r - r_+)(r - r_-), \tag{4}$$

with

$$r_\pm = M \pm \sqrt{M^2 - Q^2} \tag{5}$$

being the radii of the two horizons ($r = r_+$ at the outer or event horizon, and $r = r_-$ at the inner or Cauchy horizon).

In terms of r_+ and r_-, or in terms of the event horizon radius r_+ and the charge Q, which are two convenient parameters for describing the Reissner-Nordstrom metrics, the black hole mass is

$$M = \frac{1}{2}(r_+ + r_-) = \frac{1}{2}r_+ + \frac{Q^2}{2r_+}, \tag{6}$$

and so the energy excess over the minimum for that Q is

$$E \equiv M - Q = \frac{(r_+ - Q)^2}{2r_+}. \tag{7}$$

The Bekenstein-Hawking entropy is

$$\begin{aligned} S &= \frac{1}{4}A = \pi r_+^2 = \pi(M + \sqrt{M^2 - Q^2})^2 \\ &= \pi(Q + E + \sqrt{2QE + E^2})^2 \\ &\approx \begin{cases} \pi Q^2 + \pi\sqrt{8Q^3E} & \text{for } E \ll Q \\ 4\pi E^2 & \text{for } E \gg Q \end{cases} \end{aligned} \tag{8}$$

From these relations, one can easily derive that the Hawking temperature is

$$\begin{aligned} T_H &= \left(\frac{\partial E}{\partial S}\right)_Q = \frac{r_+ - r_-}{4\pi r_+^2} = \frac{r_+^2 - Q^2}{4\pi r_+^3} = \frac{Mr_+ - Q^2}{2\pi r_+^3} \\ &\approx \begin{cases} \sqrt{\frac{E}{2\pi^2 Q^3}} & \text{for } E \ll Q \Rightarrow T_H \sim E \text{ for } E \sim \frac{1}{Q^3} \\ \frac{1}{8\pi E} & \text{for } E \gg Q \end{cases} \end{aligned} \tag{9}$$

and that the heat capacity (at fixed charge) is

$$C_Q = \left(\frac{\partial E}{\partial T_H}\right)_Q = \frac{2\pi r_+^2(r_+^2 - Q^2)}{3Q^2 - r_+^2}. \tag{10}$$

The heat capacity is thus positive for $r_+ < \sqrt{3}Q$ or $M < 2Q/\sqrt{3}$ or $E < Q/(2\sqrt{3} + 3) \approx 0.1547\,Q$.

The free energy of a Reissner-Nordstrom black hole is

$$F = E - TS = \frac{(r_+ - Q)^2}{2r_+} - \pi T r_+^2, \tag{11}$$

which has no global minimum, as a function of r_+ (or of E), for fixed Q and T. Therefore, even though black holes with $C_Q > 0$ are locally stable against small energy increases or decreases from absorbing or emitting small amounts of thermal radiation, they are globally unstable to absorbing an arbitrarily large amount of radiation from the hypothetical infinite heat bath. This is a case of the general result that in asymptotically flat spacetime with an infinitely large heat bath at infinity, any black hole is unstable to growing indefinitely with negative heat capacity (once the energy gets sufficiently larger than the charge) and entropy $S \sim 4\pi E^2$. In other words, the canonical ensemble, weighting all states by $e^{-E/T}$, is divergent and hence not defined for gravitational systems in asymptotically flat spacetime, with T being the temperature at asymptotic infinity (the only place where one can have an infinite heat bath) [3].

To define a canonical ensemble artificially, one may introduce a negative cosmological constant, $\Lambda \equiv -3/b^2 < 0$ [4]. Then the metric for a charged black hole with zero angular momentum has the same form as Eq. (3), but now with

$$\Delta \equiv r^2 - 2Mr + Q^2 + r_+^4/b^2. \tag{12}$$

The explicit formula for the event horizon radius r_+, the greatest radius where Δ vanishes, now involves the solution of a quartic polynomial and hence is rather complicated when written in terms of M, Q, and b (or Λ), but the explicit expressions for various quantities remain simple when written in terms of r_+ and Q (and in terms of b, which is assumed to be kept fixed as part of the theory). In particular,

$$M = \frac{1}{2}r_+ + \frac{Q^2}{2r_+} + \frac{r_+^3}{2b^2}, \tag{13}$$

and

$$S = \pi r_+^2 \sim \pi(2b^2M)^{2/3} \text{ for } M \gg b, \tag{14}$$

from which one may derive that

$$T_H = \left(\frac{\partial M}{\partial S}\right)_Q = \frac{1}{4\pi r_+}\left(\frac{3r_+^2}{b^2} + 1 - \frac{Q^2}{r_+^2}\right) \tag{15}$$

and

$$C_Q = \left(\frac{\partial M}{\partial T_H}\right)_Q = 2\pi r_+^2\left(\frac{3r_+^4 + b^2r_+^2 - b^2Q^2}{3r_+^4 - b^2r_+^2 + 3b^2Q^2}\right). \tag{16}$$

The fact that the entropy does not rise so fast as linearly with the mass for $M \gg b$ means that with the negative cosmological constant, the canonical ensemble is well defined [4], as opposed to the asymptotically flat case (zero cosmological constant) in which the entropy rises faster than linearly, $S \sim 4\pi M^2$. Alternatively, the free energy (using M in place of E, since otherwise the expression becomes rather unwieldy) is

$$F = M - TS = \frac{r_+^3}{2b^2} - \pi T r_+^2 + \frac{r_+}{2} + \frac{Q^2}{2r_+}, \tag{17}$$

which is now bounded below for fixed Q, T, and b, having one or two minima.

There are two minima if $b > 6Q$, and indeed we shall be interested in the case that $b \gg Q$, so that the negative cosmological constant gives only a tiny perturbation of the metric near the black hole. Then if (but not only if) $b < 1/(\pi T)$, the minimum at least positive r_+ is the global minimum. For $Q \ll b < 1/(\pi T)$, this global minimum occurs when T is the Hawking temperature of the black hole (with T defined with respect to a timelike Killing vector that is normalized to have unit magnitude at the location where $g_{00} = -1$),

$$T = T_H = \frac{1}{4\pi r_+}\left(\frac{3r_+^2}{b^2} + 1 - \frac{Q^2}{r_+^2}\right), \tag{18}$$

which is very nearly that given by Eq. (9) for a Reissner-Nordstrom black hole with zero cosmological constant, obtained by taking b to infinity in Eq. (18).

For example, the minimum that gives $E \sim T_H \sim 1/Q^3 \ll 1$ is a global minimum of the free energy F if $b \lesssim Q^3$, which for $Q \gg 1$ (as we are assuming) allows the length scale b of the cosmological constant to be much larger than the length scale $r_+ \approx Q$ of the nearly extreme black hole. To take a value of Q that arises

below, $Q \sim 3 \times 10^{46} \sim 4 \times 10^8 M_\odot$ gives the very weak restriction $b \lesssim Q^3 \sim 3 \times 10^{82}$ Mpc in order for this configuration to be a global minimum of the free energy. Therefore, for $T \ll 1/Q$, take $b \sim 1/(\pi T) \gg Q$ to stabilize the canonical ensemble but then ignore correction terms like $r_+^2/b^2 \sim Q^2 T^2 \ll 1$ from the tiny cosmological constant.

Some strange thermodynamic behavior occurs when $E \lesssim T$. Preskill *et al.* [5] argued that the thermal description of a black hole breaks down when the emission of a particle of energy T changes T by a significant fraction of itself. For a charged Reissner-Nordstrom black hole with $E = M - Q \ll Q$, this occurs for $E \lesssim 1/Q^3$ or $T \lesssim 1/Q^3$, where $E \lesssim T$. Such a black hole does not have sufficient energy to emit a particle of energy T.

Another argument for strange thermodynamic behavior at $E \lesssim 1/Q^3$ came from a derivation by Maldacena and Susskind [6] of a mass gap $E_{\rm gap} = 1/(8Q^3)$ from string theory. Maldacena and Strominger [7] showed that it could be derived semiclassically by implicitly assuming that the first excited state of an extreme Reissner-Nordstrom black hole is an extreme Kerr-Newman black hole with angular momentum $J = 1/2$:

$$E \equiv M - Q = \sqrt{Q^2 + a^2} - Q \approx \frac{a^2}{2Q} = \frac{J^2}{2QM^2} \approx \frac{J^2}{2Q^3} = \frac{1}{8Q^3}. \tag{19}$$

Of course, this derivation fails if the lowest excited states have $J = 0$, as would occur for an ordinary chunk of solid in the idealization of zero baryon decay.

3 Evolving a Black Hole to Energy $E \sim 1/Q^3$

To get a black hole to $E = M - Q \lesssim 1/Q^3$ by emitting its excess energy, we need it to have a charge Q so huge that it Hawking radiates to this point without emitting charge. Assume that only photons and gravitons are massless, and that all other particles (e.g., neutrinos) have rest masses much larger than the temperature of the hole. Then the primary emission will be to photons with unit angular momentum. The three possible values of the axial angular momentum, and the two helicities for each, then give an energy emission rate

$$\frac{dM}{dt} \approx -\frac{6}{2\pi} \int_0^\infty \frac{P \omega d\omega}{e^{\omega/T} - 1}. \tag{20}$$

Here $P = P(Q/M, M\omega)$ is the absorption probability for a photon of frequency ω and unit angular momentum by a Reissner-Nordstrom black hole of mass M and charge Q.

This absorption probability can be obtained from the solution to the radial equation [8, 9]

$$\left(\frac{d^2}{dr_*^2} + \omega^2 - \frac{2\Delta}{r^4} - \frac{4Q^2\Delta}{r^6} \right) Z(r) = 0 \tag{21}$$

obtained after separating variables for the photon mode, where

$$\frac{d}{dr_*} = \frac{\Delta}{r^2} \frac{d}{dr} = \frac{r^2 - 2Mr + Q^2}{r^2} \frac{d}{dr}. \tag{22}$$

In particular, the boundary conditions for the radial mode function $Z(r)$ are

$$r_* \to +\infty: \quad Z \sim Z_{\rm in} e^{-i\omega r_*} + Z_{\rm out} e^{i\omega r_*}, \tag{23}$$

$$r_* \to -\infty: \quad Z \sim Z_{\rm down} e^{-i\omega r_*}. \tag{24}$$

Then the absorption probability is

$$P = \left| \frac{Z_{\rm down}}{Z_{\rm in}} \right|^2 = 1 - \left| \frac{Z_{\rm out}}{Z_{\rm in}} \right|^2. \tag{25}$$

Then the goal is to evaluate $P(Q/M, M\omega)$ from Eq. (25) and use it in Eq. (20) to find the energy emission rate and thereby find a minimum charge Q_* such that if and only iff $Q > Q_*$, the time to emit energy in uncharged particles (mainly photons) to go to $E = 1/(8Q^3) \ll 1$ is less than the time to emit

a positron. Even a single positron emission would spoil the process of getting to $E = 1/(8Q^3) \ll 1$, since it would reduce the black hole charge Q by the positron charge $e = \sqrt{\alpha} \approx 0.085424543$, and reduce the black hole mass M by the much smaller amount of roughly the positron mass, $m \approx 4.18 \times 10^{-23}$ in the Planck units we are using, so that during the positron emission, $E \equiv M - Q$ would *increase* by roughly $0.08542 \gg 1/(8Q^3)$.

First, we look for an approximate solution to the radial Eq. (21) for the frequencies that dominate the thermal emission for a nearly extreme black hole. These are frequencies ω that have roughly the same order of magnitude as the Hawking temperature T_H. Let

$$\epsilon \equiv \frac{r_+ - r_-}{r_+} \ll 1 \tag{26}$$

be the tiny quantity that parametrizes how close to extremality the black hole is. Then

$$T_H = \frac{\epsilon}{4\pi r_+}, \tag{27}$$

where, as discussed above, we are here and henceforth neglecting the effect of the miniscule cosmological constant needed to stabilize the canonical ensemble, and we are also neglecting the effect of the angular momentum of the hole, which is also expected to be relatively small for most states of fixed Q for $E = M - Q \ll Q$. Therefore, we are considering just pure Reissner-Nordstrom black holes that are very near their critical charge.

Eq. (27) implies that for frequencies ω that have roughly the same order of magnitude as the Hawking temperature T_H, the quantity $r_+\omega$ is of the order of ϵ and hence is very small. Therefore, we shall solve the radial Eq. (21) assuming both $\epsilon \ll 1$ and $r_+\omega \ll 1$. This then allows

$$r_+\omega \ll \nu \equiv \frac{r_+\omega}{\epsilon} = \frac{\omega}{4\pi T_H} \ll \frac{1}{\epsilon}. \tag{28}$$

That is, the dimensionless frequency factor ν (or $4\pi\nu = \omega/T_H$) that occurs in the thermal Hawking emission formula can range from zero (when $r_+\omega = 0$) to a very large upper limit (much smaller than, but within a few orders of magnitude of, $1/\epsilon$), thereby spanning effectively all of the thermal emission range to give an excellent approximation to the mass emission rate given by Eq. (20).

To solve Eq. (21) using the approximations of Eq. (28), split the region external to the black hole horizon into three overlapping regions, solve Eq. (21) in each region, and match in the two overlap regions.

In the near-horizon Region I defined by $r - r_+ \ll r_+$,

$$Z \approx Z_{\rm down} \frac{r}{r_+} \left(\frac{r - r_-}{r_+ - r_-}\right)^2 \left(\frac{r - r_+}{r - r_-}\right)^{-i\nu} \left[1 + \frac{4(1+i\nu)}{1-2i\nu}\left(\frac{r-r_+}{r-r_-}\right) + \frac{(1+i\nu)(1+2i\nu)}{(1-i\nu)(1-2i\nu)}\left(\frac{r-r_+}{r-r_-}\right)^2\right]. \tag{29}$$

In the intermediate Region II defined by $r_= - r_- \ll r - r_+ \ll 1/\omega$,

$$Z \approx A\frac{(r+2r_+)(r-r_+)^2}{r} + B\frac{r^4(r-3r_+)}{(r-r_+)^3}. \tag{30}$$

And in the far-field Region III defined by $r_+ \ll r$,

$$Z \sim Z_{\rm in} e^{-i\omega r_*}\left(1 - \frac{i}{\omega r_+}\right) + Z_{\rm out} e^{i\omega r_*}\left(1 + \frac{i}{\omega r_+}\right). \tag{31}$$

Then we match the solutions across Regions I and II for $r_+ - r_- \ll r - r_+ \ll r_+$, using $\epsilon \ll 1$ to guarantee the existence of this overlap region, to determine the constants A and B of Eq. (30) in terms $Z_{\rm down}$. Similarly, we match the solutions across Regions II and III for $r_+ \ll r \ll 1/\omega$, using $r_+\omega \ll 1$ to guarantee the existence of this overlap region, to determine the constants $Z_{\rm in}$ and $Z_{\rm out}$ of Eq. (31) in terms of A and B (and hence in terms of $Z_{\rm down}$ from the results of the previous overlap calculation).

The result is the absorption probability for photons of unit angular momentum over the entire dominant part of the Hawking emission spectrum:

$$P = \left|\frac{Z_{\rm down}}{Z_{\rm in}}\right|^2 \approx \frac{1}{9}\epsilon^8\nu^4(1+\nu^2)(1+4\nu^2)$$
$$= \frac{1}{9}r_+^4\omega^4(\epsilon^2+r_+^2\omega^2)(\epsilon^2+4r_+^2\omega^2). \tag{32}$$

Again, this approximate result is valid for $\epsilon = (r_+ - r_-)/r_+ \ll 1$ and for $\nu \ll 1/\epsilon$ or $\omega \ll 1/r_+ = 4\pi T_H/\epsilon$, and within this approximation, the emission of unit angular momentum photons overwhelmingly dominates over the emission of higher angular momentum photons and gravitons, by several factors of the large number $1/\epsilon$.

One can use this result to determine that the absorption cross section of a nearly-extreme Reissner-Nordstrom black hole to photons with $r_+\omega \ll 1$ is approximately

$$\sigma = \frac{(2\ell+1)\pi}{\omega^2}P \approx \frac{\pi}{3}r_+^4\omega^2(\epsilon^2+r_+^2\omega^2)(\epsilon^2+4r_+^2\omega^2). \tag{33}$$

It is interesting that although this formula was derived only for nearly-extreme black holes, $\epsilon \ll 1$, it gives the correct low-frequency limit for the uncharged Schwarzschild metric, which has $\epsilon = 1$, namely $\sigma \approx (16\pi/3)M^4\omega^2$. On the other hand, for a precisely extreme Reissner-Nordstrom black hole, $\epsilon = 0$ and $r_+ = Q$, we get that the low-frequency photons ($Q\omega \ll 1$) have $\sigma \approx (4\pi/3)Q^8\omega^6$.

Now we can insert Eq. (32) into Eq. (20) and integrate to determine the approximate time for a black hole with sufficiently large charge to evolve to $E \equiv M - Q = 1/(8Q^3) \ll 1$. In particular, if the charge is sufficiently large that $|dQ/dt| \ll |dM/dt|$ (and how large is sufficient will be discussed below), and if $E \equiv M - Q \ll Q$, then $\epsilon \approx \sqrt{8E/Q}$ and $\omega \approx \epsilon\nu/Q \approx \nu\sqrt{8E/Q^3}$, so

$$\frac{dE}{dt} \approx \frac{dM}{dt} \approx -\frac{6}{2\pi}\int_0^\infty \frac{P\omega d\omega}{e^{\omega/T}-1} = -\frac{3\epsilon^2}{\pi Q^2}\int_0^\infty \frac{P\nu d\nu}{e^{4\pi\nu}-1}$$
$$\approx -\frac{\epsilon^{10}}{3\pi Q^2}\int_0^\infty \frac{(\nu^5+5\nu^7+4\nu^9)d\nu}{e^{4\pi\nu}-1} = -\frac{509\,\epsilon^{10}}{42\,577\,920\,\pi Q^2} \approx -\frac{4\,072\,E^5}{10\,395\,\pi Q^7}. \tag{34}$$

The time to Hawking radiate to $E = 1/(8Q^3)$ is then

$$t \approx \frac{10\,395\,\pi Q^7}{16\,288\,E^4} \approx \frac{1\,330\,560}{509}\pi Q^{19} \approx 8\,212.333\,047\,46\,Q^{19}. \tag{35}$$

Next, we need to find how large Q needs to be, $Q \gtrsim Q_*$, in order that the probability be low that the charged black hole will emit a positron during this time and hence spoil the process of slowly evolving to $E = 1/(8Q^3)$. For a black hole with $1 \ll Mm \ll eQ \ll (Mm)^2$, where m and e are the mass and charge of a positron, the Hawking emission rate of positrons is well approximated by integrating the Schwinger pair production rate over the exterior of the black hole, where the electric field strength is $\mathcal{E} = Q/r^2$:

$$\frac{dN}{dt} \approx \int_{r_+}^\infty 4\pi r^2 dr \frac{e^2\mathcal{E}^2}{4\pi^3}\sum_{n=1}^\infty \frac{1}{n^2}\exp\left(-\frac{n\pi m^2}{e\mathcal{E}}\right)$$
$$\approx \int_{r_+}^\infty \frac{e^2Q^2}{\pi^2 r^2}dr \exp\left(-\frac{\pi m^2 r^2}{eQ}\right)$$
$$\approx \frac{e^3Q^3}{2\pi^3 m^2 r_+^3}\exp\left(-\frac{r_+^2}{Q_0 Q}\right) \approx \frac{e^3}{2\pi^3 m^2}\exp\left(-\frac{Q}{Q_0}\right), \tag{36}$$

where the last approximate equality assumes $\epsilon \ll 1$ so that $r_+ \approx Q$, and where

$$Q_0 \equiv \frac{e}{\pi m^2} = \frac{em_p^2}{\pi m^2}m_p^{-2} \approx 1.553 \times 10^{43}$$
$$\approx 1.8 \times 10^{44}e \approx 1.7 \times M_\odot \approx 0.84\,s. \tag{37}$$

We need that this positron emission rate be less than the reciprocal of the time to Hawking radiate to $E = 1/(8Q^3)$ that is given by Eq. (35). This implies that we need

$$Q > Q_* = Q_0 x \tag{38}$$

with x the solution of

$$x - 19 \ln x = y \equiv \ln\left(\frac{665\,280\, e^{22}}{509\,\pi^{21} m^{40}}\right) \approx 1990. \tag{39}$$

This implies that

$$x \approx y + 19 \ln y + \frac{19^2}{y} \ln y \approx 2136. \tag{40}$$

Then one gets the minimum charge Q_* of a black hole that would Hawking radiate to $E = 1/(8Q^3)$ before it is likely to emit a positron and thereby have $E \equiv M - Q$ jump upward by much more than $1/(8Q^3)$ from the reduction of Q by the emitted positron charge e:

$$\begin{aligned} Q_* &\approx 3.316 \times 10^{46} \approx 3.88 \times 10^{47} e \approx 3.6 \times 10^8 M_\odot \approx 3.6\,\mathrm{AU} \\ &\approx 5.4 \times 10^{11}\ \text{meters} \approx 1788\,\text{seconds} \approx 0.497\,\text{hour}. \end{aligned} \tag{41}$$

The magnitude of this last quantity, applied for the size of a silent object that might exist in the heavens, reminded me of a verse in the last book of the Bible, Revelation 8:1: "When the Lamb opened the seventh seal, there was silence in heaven for about half an hour."

Now if we go back and calculate the times given by Eq. (35) and by the reciprocal of the rate given by Eq. (36), we find that for $Q > Q_*$ the time to Hawking radiate to $E = 1/(8Q^3)$ is

$$t \approx (1.1 \times 10^{837}\mathrm{yr}) \left(\frac{Q}{Q_*}\right)^{19}, \tag{42}$$

which is less than the expected time to emit a positron,

$$t \approx (1.1 \times 10^{837}\mathrm{yr}) \exp\left[\frac{2136(Q - Q_*)}{Q_*}\right]. \tag{43}$$

Q_* is the charge of 6.5×10^{20} kg of protons, 0.88% of the lunar protons, or 65% of the asteroid Ceres' protons. If superclusters of mass $M_S \sim 10^{16} M_\odot$ have charge $Q_S \sim m_p M_S/e \sim 9.00 \times 10^{-19} M_S \sim 8 \times 10^{35} \sim 2.5 \times 10^{-11} Q_*$, we would need $\sim 4 \times 10^{10}$ such superclusters to collapse into the black hole without the surrounding negative plasma to get a black hole of charge Q_*. If a sphere of N_S such superclusters collapses, with the $\sim N_S^{2/3}$ on the surface falling in unneutralized, one needs $N_S \sim (4 \times 10^{10})^{3/2} = 8 \times 10^{15}$ such superclusters. This would initially form a black hole of mass $M \sim N_S M_S \sim 10^{32} M_\odot \sim 10^{70} \sim 10^{19}$ yr, with initial charge-to-mass ratio $Q/M \sim 10^{-23}$.

Using the results of [10] for the emission of photons and gravitons (here assumed to be the only particles of mass less than $1/M \sim 10^{-70} \sim 10^{-42} eV$, as discussed above) from a nearly Schwarzschild black hole, the time to evaporate to the point where the charge becomes important for the geometry and makes the Hawking temperature attain its maximum value of $1/(6\sqrt{3}\pi Q_*) \approx 9.24 \times 10^{-49} \approx 1.31 \times 10^{-16} K \approx 1.13 \times 10^{-20} eV$ is roughly $8895\, M^3 \sim 10^{163}$ years. Then the time to evaporate to $E = 1/(8Q^3) = 1/(8Q_*^3) \approx 3.43 \times 10^{-141} \approx 4.19 \times 10^{-113}\, eV$ (as the temperature plummets to $1/(4\pi Q_*^3) \approx 2.18 \times 10^{-141} \approx 3.09 \times 10^{-109} K$) would be roughly the 10^{837} years discussed above. One can see that even if the universe expands as slowly as a matter-dominated $k = 0$ Friedmann-Robertson-Walker universe, with scale size going as the time to the 2/3 power, the current $3K$ microwave background would have by then cooled below $10^{-551} K$ and so would be negligible in comparison with the roughly $10^{-109} K$ temperature of the nearly extreme black hole.

4 Density of States

Let us consider various possibilities for the density of states (energy levels) of a nearly extreme charged black hole. First, let us make a semiclassical estimate. As we found above for nearly extreme black holes,

$E \equiv M - Q \ll Q$ (neglecting the effect of the tiny cosmological constant that was introduced to stabilize the canonical ensemble), we get that the Hawking temperature T_H (here denoted simply as T) is

$$T = \frac{r_+ - r_-}{4\pi r_+^2} \approx \sqrt{\frac{E}{2\pi^2 Q^3}}, \tag{44}$$

and the entropy is

$$S = \pi r_+^2 \approx \pi Q^2 + \sqrt{8\pi^2 Q^3 E}. \tag{45}$$

Expressed in terms of the temperature, this gives

$$E \approx 2\pi^2 Q^3 T^2 \tag{46}$$

and

$$S \approx \pi Q^2 + 4\pi^2 Q^3 T. \tag{47}$$

If E is reinterpreted to be $\langle E \rangle = -d \ln Z/d\beta$ with $\beta = 1/T$ and partition function

$$Z(\beta) = \sum_{n=1}^{\infty} e^{-\beta E_n} \tag{48}$$

from energy levels E_n, then

$$\ln Z = S - \beta \langle E \rangle \approx \pi Q^2 + 2\pi^2 Q^3 \beta^{-1} \tag{49}$$

or

$$Z \approx e^{\pi Q^2 + 2\pi^2 Q^3 T}. \tag{50}$$

The inverse Laplace transform gives the density of states

$$\rho(E) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} Z(\beta) e^{\beta E} d\beta \approx e^{\pi Q^2} \delta(E) + e^{\pi Q^2} \sqrt{\frac{2\pi^2 Q^3}{E}} I_1 \left(\sqrt{8\pi^2 Q^3 E} \right) \theta(E), \tag{51}$$

with I_1 (and I_0 below) being Bessel functions of imaginary argument. The number of states $\leq E$ for $E \geq 0$ is then

$$\begin{aligned} N(E) &= \int_{-\infty}^{E} \rho(E') dE' \approx e^{\pi Q^2} I_0 \left(\sqrt{8\pi^2 Q^3 E} \right) \\ &= e^{\pi Q^2} \left[1 + \frac{2\pi^2 Q^3 E}{(1!)^2} + \frac{(2\pi^2 Q^3 E)^2}{(2!)^2} + \cdots \right]. \end{aligned} \tag{52}$$

For $1/Q^3 \ll E \ll Q$,

$$N(E) \approx \frac{e^{\pi Q^2 + \sqrt{8\pi^2 Q^3 E}}}{2\pi (2Q^3 E)^{1/4}} \approx \frac{e^{S(E)}}{2\pi (2Q^3 E)^{1/4}}. \tag{53}$$

If $S_0 \equiv S(E = 0) = \pi Q^2$, perhaps a good approximation for all $E \geq 0$ is that the number of states of lower E is, by this semiclassical estimate,

$$\begin{aligned} N(E) &\approx e^{S_0} I_0[S(E) - S_0] = e^{\pi Q^2} I_0[\pi r_+^2(E) - \pi Q^2] \\ &= e^{\pi Q^2} I_0[\pi (Q + E + \sqrt{2QE + E^2})^2 - \pi Q^2]. \end{aligned} \tag{54}$$

This gives a huge $e^{\pi Q^2}$ degeneracy of states at $E = 0$ and then a very dense quasi-continuum of states. For $E \ll 1/(2\pi^2 Q^3)$, $N(E) \approx e^{\pi Q^2}(1 + 2\pi^2 Q^3 E)$, giving a density of states

$$\rho(E) = \frac{dN(E)}{dE} \approx e^{\pi Q^2} [\delta(E) + 2\pi^2 Q^3 \theta(E)]. \tag{55}$$

The quasi-continuum of excited states has an average energy separation

$$\delta E = \frac{1}{\rho} \approx \frac{e^{-\pi Q^2}}{2\pi^2 Q^3}, \tag{56}$$

exponentially small.

For $e^{-\pi Q^2} \ll 2\pi^2 Q^3 T \ll 1$, we get $Z \approx e^{\pi Q^2}(1 + 2\pi^2 Q^3 T)$, which implies that

$$\langle E \rangle = -\frac{d \ln Z}{d\beta} = \frac{T^2}{Z}\frac{dZ}{dT} \approx 2\pi^2 Q^3 T^2 = (2\pi^2 Q^3 T)T \ll T. \tag{57}$$

Preskill *et al.* [5] argue, "The statistical treatment of the radiation is inappropriate if the ensemble of states from which it is drawn is small. In estimating the size of this ensemble, we should not include the residual entropy at zero temperature, since this is unavailable to the radiation."

However, contrary to this claim, there are actually about $2\pi^2 Q^3 T e^{\pi Q^2} \gg 1$ excited states of energy $E < T$, which is huge in absolute number. For example, it is larger than a googolplex ($10^{10^{100}}$) if $Q \gtrsim \sqrt{\ln 50/\pi}\, 10^{50} \approx 8.6 \times 10^{49} \approx 2582 Q_*$. $\approx 9.37 \times 10^{11}\, M_\odot \approx 53$ days, corresponding to a nearly extreme black hole with the mass of a large galaxy. The true reason why $\langle E \rangle \ll T$ for $T \ll 1/Q^3$ is that the number of ground states is even much higher than the enormous number of excited states with $E < T \ll 1/Q^3$, so that the canonical ensemble is dominated by the huge number of degenerate ground states and has only a tiny probability for excitation.

In particular, there are $d_0 \approx e^{\pi Q^2}$ ground states with $E = 0$ for the density of states given above, giving the total probability to be in one or another of these ground states as $P_{\text{ground}} = d_0/Z \approx 1 - 2\pi^2 Q^3 T$. Then the total probability for all excitations is $P_{\text{excite}} = 1 - P_{\text{ground}} \approx 2\pi^2 Q^3 T \ll 1$. Thus

$$\langle E \rangle = P_{\text{ground}} 0 + P_{\text{excite}} \langle E \text{ if excited} \rangle, \tag{58}$$

combined with the value of $\langle E \rangle$ given by Eq. (57), implies that

$$\langle E \text{ if excited} \rangle = \frac{\langle E \rangle}{P_{\text{excite}}} \approx T, \tag{59}$$

which is precisely what one would expect if there were a very dense uniform distribution of states, when T is greater than the average separation between the levels.

If the lowest excited state is nondegenerate and has $E = \delta E \sim e^{-\pi Q^2}/(2\pi^2 Q^3)$, which is the average separation between energy levels for the density of states given by Eq. (55), and if the next excited state has $E = (1 + f)\delta E$, etc., then for $T \ll f\delta E$, $P_{\text{excite}} \approx e^{-\pi Q^2 - \delta E/T}$ and $\langle E \rangle \approx (\delta E) e^{-\pi Q^2 - \delta E/T}$. Thus it is only for $T \lesssim 1/(\delta E)$ that the absolute scarcity of excited states with $E < T$ has a significant effect in changing the form of $\langle E \rangle$; for $1/(\delta E) \ll T \lesssim 1/Q^3$, the number of excited states with $E < T$ is huge in absolute number, and it is only the fact that it is less than the even larger number of degenerate ground states that makes $\langle E \rangle \ll T$.

Thus the semiclassical estimate gives a ground state degeneracy of $d_0 \approx e^{\pi Q^2}$ and then a quasi-continuum with an exponentially large density of states $\rho(E) \approx 2\pi^2 Q^3 e^{\pi Q^2}$ for $E \lesssim 1/Q^3$.

One alternative for the energy states of a nearly extreme charged black hole is that there is a whole series of highly degenerate levels, separated by macroscopic mass gaps. For example, the classical formula for extreme black holes with angular momentum $J = n/2$ for nonnegative integers n is

$$E_n = \sqrt{\frac{1}{2}[Q^2 + \sqrt{Q^4 + 4J^2}]} - Q = Q\left\{\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{n^2}{Q^4}}} - 1\right\} = \frac{n^2}{8Q^3} - \frac{5n^4}{128Q^7} + \cdots. \tag{60}$$

One might postulate that each of these levels has degeneracy

$$d_n \approx e^{S(E,J)} = e^{\pi r_+^2} = e^{\pi(Q+E_n)^2} \approx e^{\pi Q^2}\left(1 + \frac{\pi n^2}{4Q^2}\right), \tag{61}$$

possibly multiplied by some function of Q and n that is not exponentially large in Q^2.

If $E_n \approx E_1 n^a$ and $d_n \propto n^b$ for some constants a and b, then for $T \equiv 1/\beta \gg E_1$,

$$Z = \sum_{n=0}^{\infty} d_n e^{-\beta E_n} \propto \int_0^\infty dn\, n^b e^{-\beta E_1 n^a} \propto \beta^{-(1+b)/a}, \tag{62}$$

which then implies that

$$\langle E \rangle \approx \frac{1+b}{a} T. \tag{63}$$

For example, the alternative above, in which $E_n \approx n^2/(8Q^3)$ and d_n is approximately independent of n, corresponds to $a = 2$ and $b = 0$ and hence gives $\langle E \rangle \approx T/2$ for $E_1 \approx 1/(8Q^3) \ll T \ll 1/Q$. On the other hand, for $T \ll E_1$, $Z \approx d_0 + d_1 e^{-\beta E_1}$, which implies that $\langle E \rangle \approx (d_1/d_0) e^{-E_1/T}$, which is exponentially suppressed when there are no states in an initial mass gap.

There is an opposite alternative for the energy states of a nearly extreme charged black hole that to me seems more plausible in our real world with broken supersymmetry. This alternative is that there is no huge degeneracy anywhere, even for the ground state, but instead only an exponentially dense quasi-continuum of states.

For example, if one takes the semiclassical estimate above and then changes the number of ground states from the value above of $d_0 \approx e^{\pi Q^2}$ to $d_0 = 1$, then one just subtracts roughly $e^{\pi Q^2} - 1$ from the semiclassical estimate for the partition function. This also reduces by roughly $e^{\pi Q^2} - 1$ the number of states of energy $\leq E$ (for $E \geq 0$; for $E < 0$ there are no states) from the approximate value given by Eq. (54) to what now becomes, after inserting the operator Int to take the integer part of what follows,

$$\begin{aligned} N(E) &\approx \text{Int}\left\{e^{S_0}[I_0(S(E) - S_0) - 1] + 1\right\} \\ &= \text{Int}\left\{e^{\pi Q^2}[I_0(\pi r_+^2(E) - \pi Q^2) - 1] + 1\right\} \\ &= \text{Int}\left\{e^{\pi Q^2}[I_0(\pi(Q + E + \sqrt{2QE + E^2})^2 - \pi Q^2) - 1] + 1\right\} \\ &\approx \text{Int}\left\{e^{\pi Q^2}[I_0(\sqrt{8\pi^2 Q^3 E}) - 1] + 1\right\} \\ &\approx \text{Int}\left\{2\pi^2 Q^3 e^{\pi Q^2} E + 1\right\}. \end{aligned} \tag{64}$$

Here the penultimate expression on the right hand side (after the second $\approx$ sign) applies for $E \ll Q$, any nearly extreme charged black hole, and the final expression on the right hand side applies for $E \ll 1/Q^3 \ll 1$, where the density of states is nearly uniform, with average spacing given by Eq. (54).

If the temperature is much greater than this average spacing, $e^{-\pi Q^2}/(2\pi^2 Q^3)$, but much less than the maximum temperature of a Reissner-Nordstrom black hole of fixed charge Q, $1/(6\sqrt{3}\pi Q)$, so that then the canonical ensemble is dominated by a large number of excited states that are all nearly extreme black holes (using a tiny cosmological constant to suppress the contribution of the much larger nearly Schwarzschild-anti-deSitter black holes that would have similar or lower temperatures if the cosmological constant were zero), then the partition function has the form given by subtracting roughly $e^{\pi Q^2}$ from that given by Eq. (50):

$$Z \approx e^{\pi Q^2 + 2\pi^2 Q^3 T} - e^{\pi Q^2}. \tag{65}$$

In this temperature range, one thus gets

$$\begin{aligned} \langle E \rangle &= \frac{T^2}{Z}\frac{dZ}{dT} \approx \frac{2\pi^2 Q^3 T^2}{1 - e^{-2\pi^2 Q^3 T}} \\ &\approx \begin{cases} T & \text{for } e^{-\pi Q^2} \ll 2\pi^2 Q^3 T \ll 1 \\ 2\pi^2 Q^3 T^2 & \text{for } 1 \ll 2\pi^2 Q^3 T \end{cases} \end{aligned} \tag{66}$$

Furthermore, for a fixed temperature in this range (and with a tiny cosmological constant to stabilize the canonical ensemble, as discussed above), the von Neumann entropy of the thermodynamic canonical ensemble with the presently-assumed density of states is

$$\begin{aligned} S &= \ln Z + \frac{\langle E \rangle}{T} \approx \pi Q^2 + 2\pi^2 Q^3 T \left(\frac{2 - e^{-2\pi^2 Q^3 T}}{1 - e^{-2\pi^2 Q^3 T}}\right) + \ln\left(1 - e^{-2\pi^2 Q^3 T}\right) \\ &\approx \begin{cases} \pi Q^2 + 1 + \ln(2\pi^2 Q^3 T) & \text{for } e^{-\pi Q^2} \ll 2\pi^2 Q^3 T \ll 1 \\ \pi Q^2 + 4\pi^2 Q^3 T & \text{for } 1 \ll 2\pi^2 Q^3 T \end{cases} \end{aligned} \tag{67}$$

On the other hand, for a temperature comparable to or smaller than the spacing between energy levels, $T \lesssim e^{-\pi Q^2}/(2\pi^2 Q^3)$, the discreteness of the energy levels is important. If the levels are indeed evenly spaced with a separation of $e^{-\pi Q^2}/(2\pi^2 Q^3)$ for $E \lesssim 1/Q^3$, then for all $T \lesssim 1/Q^3$ one gets the thermodynamics of a single harmonic oscillator of frequency given by the energy spacing,

$$Z \approx \left[1 - \exp\left(-\frac{1}{2\pi^2 Q^3 e^{\pi Q^2} T}\right)\right]^{-1}, \tag{68}$$

$$\langle E \rangle \approx \left\{2\pi^2 Q^3 e^{\pi Q^2}\left[\exp\left(\frac{1}{2\pi^2 Q^3 e^{\pi Q^2} T}\right) - 1\right]\right\}^{-1}, \tag{69}$$

$$S \approx \left\{2\pi^2 Q^3 e^{\pi Q^2} T\left[\exp\left(\frac{1}{2\pi^2 Q^3 e^{\pi Q^2} T}\right) - 1\right]\right\}^{-1} - \ln\left[1 - \exp\left(-\frac{1}{2\pi^2 Q^3 e^{\pi Q^2} T}\right)\right]. \tag{70}$$

The second and third of these reduce to the expressions given in Eqs. (66) and (67) for $e^{-\pi Q^2} \ll 2\pi^2 Q^3 T \ll 1$, but for $T \ll e^{-\pi Q^2}/(2\pi^2 Q^3)$, $\langle E \rangle$ and S (as well as $Z - 1$) are exponentially suppressed.

These expressions show that if the ground state of a black hole of fixed charge Q is not degenerate (or has a degeneracy much less than the exponential of one-quarter its area of $4\pi Q^2$), but the density of excited states is roughly that given by the semiclassical estimate, then for temperatures T and excess energies E (i.e., energies above that of the ground state energy, which is roughly Q for the fixed charge Q) greater than about $1/Q^3$, the entropy and other thermodynamic properties are very near what one would calculate for a classical black hole with mass given by the charge plus the expectation value of the excitation energy. However, for temperatures and excess energies less than about $1/Q^3$, then the absence of a huge ground-state degeneracy changes the thermodynamic values from what a classical calculation would indicate. In particular, as one lowers the temperature, the thermodynamic entropy is reduced below one-quarter the area of the corresponding classical Reissner-Nordstrom black hole with the same classical temperature or same energy as that of the thermal expectation value. Indeed, when the temperature and excess energy expectation value are reduced to zero, the entropy reduces to the logarithm of the number of degenerate ground states, which is zero if the ground state is not degenerate.

These deduced thermodynamic properties at very low temperatures and energies are different from what a naïve classical calculation would give, but they seem to be a plausible possibility of what happens for charged gravitational systems. Of course, the simple modification of the semiclassical calculation by dropping the degeneracy of the ground state is also naïve, so its detailed predictions are not likely to be correct (e.g., a fairly uniform density of states all the way down to a nondegenerate ground state), but they seem to be a reasonable zeroth-order guess and plausibly are qualitatively correct.

How might one be able to tell in principle whether the ground state of a charged black hole has a huge degeneracy or not? One way is to examine the time for the hole to become extreme.

If the ground state has a huge degeneracy $\sim e^{\pi Q^2}$, comparable to the number of excited states with excitation energies up to about $1/Q^3$, then after the black hole radiates down to $E \sim 1/Q^3$, a reasonable fraction of the lower-energy states are ground states, so it seems plausible that the hole may go to $E = 0$ in one more transition with reasonable probability. Then the expected time to go to one of the ground states would be comparable to (e.g., say a factor of two greater than) the expected time to go to some energy of the order of $1/Q^3$, which we estimated above to be about $10^4\, Q^{19}$ and greater than the expected time to emit a positron for $Q \gtrsim Q_\bullet \sim 4 \times 10^8 M_\odot$, which implies a time equal to or greater than about 10^{837} years for a black hole of charge $Q \gtrsim Q_\bullet$ to become extreme and stop emitting (except for emitting positrons, which we found occurs at a slower rate that is suppressed by an exponential of $Q/Q_\bullet$).

However, if the ground state of a charged black hole is has only a small degeneracy, say smaller by a factor of roughly $e^{\pi Q^2}$, then one might expect that the time to decay through all the excited states and reach one of the small number (e.g., one) of ground states might be larger by this same factor (assuming that Q stayed constant). But a time of order $e^{\pi Q^2}$, whose logarithm grows quadratically with Q, is always larger than the time to emit a positron, which goes as $e^{Q/Q_\bullet}$, whose logarithm rises only linearly with Q. So if this is indeed the case, one would not expect ever to reach a ground state for nonzero charge; the charge would almost always decrease from any fixed value before one reached the ground state for that charge.

Therefore, to tell whether a charged black hole has a ground state degeneracy $\sim e^{\pi Q^2}$ for $Q \gg Q_\bullet \sim 4\times 10^8 M_\odot$, just wait several times $10^4\, Q^{19} \sim 10^{837}\mathrm{yr}(Q/Q_\bullet)^{19}$ and see whether the black hole has stopped radiating. Unfortunately, this is slightly longer than the time over which a Ph.D. student generally hopes to complete his or her thesis research and graduate.

This work was supported in part by the Natural Sciences and Engineering Council of Canada.

References

[1] G.W. Gibbons and R.E. Kallosh, Physical Review **D51** (1995) 2839.

[2] S.W. Hawking, G.T. Horowitz, and S.F. Ross, Physical Review **D51** (1995) 4302.

[3] S.W. Hawking, Physical Review **D13** (1976) 191.

[4] S.W. Hawking and D.N. Page, Communications in Mathematical Physics **87** (1983) 577.

[5] J. Preskill, P. Schwarz, A. Shapere, S. Trivedi, and F. Wilczek, Modern Physics Letters **A6** (1991) 2353.

[6] J.M. Maldacena and L. Susskind, Nuclear Physics **B475** (1996) 679.

[7] J.M. Maldacena and A. Strominger, Physical Review **D56** (1997) 4975.

[8] F.J. Zerilli, Physical Review **9** (1974) 860.

[9] V. Moncrief, Physical Review **9** (1974) 2707, **10** (1974) 1057, **12** (1975) 1526.

[10] D.N. Page, Physical Review **D13** (1976) 198.

Moving mirror and black hole entropy - A proposal of moving-mirror entropy - [1]

Shinji Mukohyama

Department of Physics and Astronomy, University of Victoria
Victoria, BC, Canada V8W 3P6

Abstract

We consider a quantized scalar field in a two-dimensional Minkowski spacetime with a moving mirror and propose a definition of moving-mirror entropy associated with temporarily inaccessible information about the future.

1 Introduction

One of the most embarrassing problems in gravitational physics is the so called information loss problem. When Hawking discovered a thermal radiation from black holes (Hawking radiation) in 1974 by using semiclassical and quasi-stationary approximations, he argued that formation and evaporation of black holes would introduce a non-unitarity in quantum mechanics [2]. Actually, his calculation shows that an initial pure state evolves to a totally uncorrelated thermal state, which is impossible in a unitarily evolving system. In general within these approximations, the evolution is described by a non-unitary superscattering matrix which maps initial mixed states to final mixed states, and superscattering matrix elements can be calculated explicitly [3, 4]. Hence, as far as the approximations are justified, some information with respect to the initial state seems to be lost.

On the other hand, Page [5] argued that, if the superscattering matrix describing the whole system including quantum gravity is CPT invariant, the description can be reduced to an S matrix which maps pure initial states into pure final states. In other words, if quantum gravity is CPT invariant then there is no loss of information in the process of formation and evaporation of black holes. In this case information can be lost only temporarily and the temporarily missing information should be completely recovered. Although his arguments might seem to contradict the semiclassical result that emission from a black hole is uncorrelated, he also argued later that information may come out initially so slowly, or else be so spread out, that it would never show up in a perturbative analysis [6]. Hence, it seems very difficult to prove or disprove his arguments by any perturbative analysis, in particular the semiclassical calculation.

Page's arguments seem to be consistent with the interpretation of entanglement entropy proposed in [7, 8]. Entanglement entropy of a pure state with respect to a division of a Hilbert space into two subspaces 1 and 2 was interpreted as an amount of information which can be transmitted through 1 and 2 from a system interacting with 1 to another system interacting with 2. In this interpretation, the transmitting medium is the quantum entanglement between 1 and 2, and the entanglement entropy may be a quantity which can in principle cancel the black-hole entropy, the amount of the temporarily missing information, to restore information loss. However, so far we do not have any concrete models to realize this idea.

Recent progress in string theory suggests the existence of a unitary S-matrix description for the Hawking process. (For reviews of black holes in superstring theory see, for example, [9, 10, 11] and references therein.) In the D-brane picture, an extremal or near extremal black hole is described by stacked D-branes on which open strings are attached and moving. Hawking radiation can be described as a decay process of two open strings into a closed string away from the D-branes. Since the evolution of the whole system in the D-brane picture including closed and open strings is manifestly unitary, the Hawking process should be described as a unitary process at least for near extremal black holes. Moreover, the AdS/CFT correspondence [12, 13] might be considered as a conjecture about a unitary description for arbitrary gravitational processes including formation and evaporation of black holes.

[1] This talk is based on collaboration with Werner Israel [1].

The D-brane picture also provides us an interpretation of black-hole entropy of an extremal or near extremal black hole as the logarithm of the number of different states of open strings on stacked D-branes [14]. Hence, in some sense the black-hole entropy seems to be an amount of temporarily inaccessible information. Namely, if the black hole evaporates completely then information about the quantum state of the black hole should come out as correlations among closed strings, provided that a unitary description similar to the D-brane picture holds for evaporating black holes.

However, it seems fair to say that the above consideration about the restoration of missing information is no more than speculation. For example, see Hawking's objections [15]. Moreover, because of the Page's arguments that the recovery of information would never show up in a perturbative analysis, it seems very difficult to prove or disprove it.

On the other hand, one might consider a different approach: it might be relevant to investigate whether it is possible to assign entropy to a manifestly unitary process. In this talk, we shall consider the so called moving-mirror effect [16, 17], which is manifestly unitary but is often referred as an analogue of the Hawking process.

2 Moving-mirror entropy

For simplicity, consider a minimally-coupled scalar field in two-dimensional Minkowski spacetime ($ds^2 = -dt^2 + dx^2 = -dudv$) with a moving-mirror. Denote the trajectory of the mirror by $x^\mu = X^\mu(\tau)$, where τ is the proper time along the mirror trajectory and assume that both sides of the mirror are perfectly reflecting. For this situation, suppose that the functions $X^\mu(\tau)$ for $\tau < \tau_0$ are known and that the remaining $\tau > \tau_0$ part of the trajectory is unknown. Of course, provided that the initial state at $\mathcal{I}^-$ is given, we can determine the quantum state of wave-packet modes which are reflected by the mirror before τ_0. However, we cannot predict the quantum state of wave-packet modes which will be reflected by the mirror after τ_0 since that depends on the unknown future trajectory of the mirror. In the following, we shall propose a definition of moving-mirror entropy associated with the lack of information about the future trajectory of the mirror.

It is expected that this lack of information should be related to uncertainty about the quantum state of wave-packet modes which will be reflected by the mirror in the future since the latter is determined by the former, if the initial quantum state at $\mathcal{I}^-$ is fixed. In order to represent such uncertainty quantitatively, it is convenient to consider a quantum-mechanical generalization of conditional entropy, the so called von Neumann conditional entropy.

Classically, the conditional entropy of an experiment A relative to another experiment B is defined by $H(A|B) = -\sum_{a,b} p(a,b) \ln p(a|b)$, where a and b represent outcomes of A and B, respectively, $p(a,b)$ is the joint probability of a and b, $p(a|b) = p(a,b)/p(b)$ is the conditional probability of a for a given outcome b, and $p(b)$ is the probability of b. The conditional entropy corresponds to uncertainty about the outcome of A after the experiment B has been performed. The quantum analogue of conditional entropy was considered in references [18, 19]. Consider a Hilbert space $\mathcal{F}$ of the form $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$ and let ρ be a density matrix on $\mathcal{F}$. Given ρ, the conditional entropy of the subsystem 1 relative to 2 is defined by

$$S_{1|2} = \mathrm{Tr}\left[\rho\sigma_{1|2}\right], \tag{1}$$

where $\sigma_{1|2} = \mathbf{1}_1 \otimes \ln\rho_2 - \ln\rho$, and

$$\rho_2 = \mathrm{Tr}_1\rho. \tag{2}$$

In our case, $\mathcal{F}_1$ and $\mathcal{F}_2$ are symmetric Fock spaces constructed from Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ of mode functions which are reflected by the mirror after and before τ_0, respectively. It seems natural to assume that the initial quantum state is the vacuum $|0,-\rangle$ determined by positive-frequency mode functions at $\mathcal{I}^-$. Correspondingly, the density matrix ρ represents a pure state given by $\rho = |0,-\rangle\langle 0,-|$. It is easy to see that, if ρ is pure,

$$S_{1|2} = -S_{ent}, \tag{3}$$

where S_{ent} is entanglement entropy defined by

$$S_{ent} = -\mathrm{Tr}_2[\rho_2 \ln\rho_2]. \tag{4}$$

Moreover, in this case, the quantum state $|0,-\rangle$ is the direct product of quantum states $|0,-R\rangle$ and $|0,-L\rangle$ in the right and left sectors R and L separated by the mirror. Hence, the entanglement entropy can be calculated separately in each of the two regions, and the total entanglement entropy is the sum of these two contributions.

To calculate the entanglement entropy in the R region, it is convenient to write the mirror trajectory as $v = p(u)$ and to introduce a new coordinate $\tilde{u}$ by $\tilde{u} \equiv p(u)$. In terms of this coordinate, the mirror trajectory becomes $v = \tilde{u}$ and the R region is $v \geq \tilde{u}$. Thence, as we shall briefly illustrate below, the entanglement entropy in the R region is easily calculated and given by

$$S_{entR} = \frac{1}{12} \ln[L_{\tilde{u}}/l_{\tilde{u}}], \tag{5}$$

where $L_{\tilde{u}}$ and $l_{\tilde{u}}$ are infrared and ultraviolet cutoffs in the $\tilde{u}$-coordinate [20].

Now let us verify this expression for the case of Dirichlet boundary conditions at the mirror. For this purpose note first that the entanglement entropy does not depend on the bases in $\mathcal{F}_1$ and $\mathcal{F}_2$ and so we can choose these bases in a convenient way. Hence, following Unruh [21], we introduce Rindler coordinates $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$ defined by

$$\begin{aligned} \tilde{u}^{(1)} &= -\ln(\tilde{u} - \tilde{u}_0) \quad \text{for} \quad \tilde{u} > \tilde{u}_0, \\ \tilde{u}^{(2)} &= -\ln(-\tilde{u} + \tilde{u}_0) \quad \text{for} \quad \tilde{u} < \tilde{u}_0, \end{aligned} \tag{6}$$

where $\tilde{u}_0 = p(X^u(\tau_0))$. Next, we can construct symmetric Fock spaces $\mathcal{F}_{1R}$ and $\mathcal{F}_{2R}$ from the following mode functions $\varphi_\omega^{(1)}$ and $\varphi_\omega^{(2)}$, respectively.

$$\begin{aligned} \varphi_\omega^{(1)} &= \theta(\tilde{u} - \tilde{u}_0) e^{i\omega\tilde{u}^{(1)}} - \theta(v - \tilde{u}_0) e^{i\omega v^{(1)}}, \\ \varphi_\omega^{(2)} &= \theta(-\tilde{u} + \tilde{u}_0) e^{-i\omega\tilde{u}^{(2)}} - \theta(-v + \tilde{u}_0) e^{-i\omega v^{(2)}}, \end{aligned} \tag{7}$$

where $v^{(1)}$ and $v^{(2)}$ are defined by

$$\begin{aligned} v^{(1)} &= -\ln(v - \tilde{u}_0) \quad \text{for} \quad v > \tilde{u}_0, \\ v^{(2)} &= -\ln(-v + \tilde{u}_0) \quad \text{for} \quad v < \tilde{u}_0. \end{aligned} \tag{8}$$

Evidently, the Fock space $\mathcal{F}_{1R}$ (and $\mathcal{F}_{2R}$) is the subspace of $\mathcal{F}_1$ (and $\mathcal{F}_2$, respectively) which is relevant to the R region. Thirdly, the vacuum $|0,-R\rangle$ at $\mathcal{I}^-$ in the R region can be defined by the following set of mode functions $\Phi_\omega^{(1)}$ and $\Phi_\omega^{(2)}$ which are analytic in the lower complex v-plane.

$$\begin{aligned} \Phi_\omega^{(1)} &= N_\omega(\varphi_\omega^{(1)} + e^{-\pi\omega}\varphi_\omega^{(2)*}), \\ \Phi_\omega^{(2)} &= N_\omega(\varphi_\omega^{(2)} + e^{-\pi\omega}\varphi_\omega^{(1)*}), \end{aligned} \tag{9}$$

where N_ω is a normalization constant. Thence, following [22], it is now easy to expand the vacuum $|0,-R\rangle$ in terms of Fock space states constructed from $\{\varphi_\omega^{(1,2)}\}$ and to trace over the degrees of freedom in $\mathcal{F}_{1R}$. The resulting reduced density matrix on $\mathcal{F}_{2R}$ is the thermal density matrix with temperature $T = 1/2\pi$, provided that it is written in terms of Fock space states constructed from $\{\varphi_\omega^{(2)}\}$. Therefore the corresponding entropy density, or entropy per unit interval of the $\tilde{u}^{(2)}$ coordinate, is $S = (\pi/6)T = 1/12$. Finally, by introducing the ultraviolet and infrared cutoffs $l_{\tilde{u}}$ and $L_{\tilde{u}}$, we obtain

$$S_{entR} = \int_{\tilde{u}=\tilde{u}_0-L_{\tilde{u}}}^{\tilde{u}=\tilde{u}_0-l_{\tilde{u}}} S d\tilde{u}^{(2)} = \frac{1}{12} \ln[L_{\tilde{u}}/l_{\tilde{u}}]. \tag{10}$$

Similarly, to calculate the entanglement entropy in the L region, we write the mirror trajectory as $u = q(v)$ and introduce a new coordinate $\tilde{v}$ by $\tilde{v} \equiv q(v)$. Using this coordinate, the entanglement entropy in the L region is given by

$$S_{entL} = \frac{1}{12} \ln[L_{\tilde{v}}/l_{\tilde{v}}], \tag{11}$$

where $L_{\tilde{v}}$ and $l_{\tilde{v}}$ are infrared and ultraviolet cutoffs in the $\tilde{v}$-coordinate.

The ultraviolet cutoffs $l_{\tilde{u}}$ and $l_{\tilde{v}}$ can be related to the ultraviolet cutoff l in the proper time τ as follows.

$$\begin{aligned} l_{\tilde{u}} &= \frac{d\tilde{u}}{d\tau}l + \frac{1}{2}\frac{d^2\tilde{u}}{d\tau^2}l^2 + \frac{1}{6}\frac{d^3\tilde{u}}{d\tau^3}l^3 + O(l^4), \\ l_{\tilde{v}} &= \frac{d\tilde{v}}{d\tau}l + \frac{1}{2}\frac{d^2\tilde{v}}{d\tau^2}l^2 + \frac{1}{6}\frac{d^3\tilde{v}}{d\tau^3}l^3 + O(l^4), \end{aligned} \tag{12}$$

or

$$\begin{aligned} l_{\tilde{u}} &= (p'(u))^{1/2}l\left[1 + \frac{1}{2}la^\mu n_\mu + \frac{l^2}{6}\left(\frac{d}{d\tau}(a^\mu n_\mu) + a^2\right) + O(l^3a^3, l^3a\partial_\tau a)\right], \\ l_{\tilde{v}} &= (q'(v))^{1/2}l\left[1 - \frac{1}{2}la^\mu n_\mu + \frac{l^2}{6}\left(-\frac{d}{d\tau}(a^\mu n_\mu) + a^2\right) + O(l^3a^3, l^3a\partial_\tau a)\right], \end{aligned} \tag{13}$$

where $a^\mu = d^2X^\mu/d\tau^2$, $a^2 = a^\mu a_\mu$, and n^μ is the unit normal to the trajectory directed toward the R region. Since $p'q' = 1$, we obtain

$$l_{\tilde{u}}l_{\tilde{v}} = l^2\left[1 + \frac{1}{12}l^2a^2 + O(l^3a^3, l^3a\partial_\tau a)\right] \tag{14}$$

as a function of the proper time τ along the mirror trajectory. On the other hand, provided that the mirror is asymptotically inertial, we obtain $L_{\tilde{u}}L_{\tilde{v}} = L^2$, where L is the infrared cutoff in the proper time τ. Hence, we obtain the total entanglement entropy.

$$S_{ent} = S_{entR} + S_{entL} = -\frac{1}{144}l^2a^2\left[1 + O(la, l\partial_\tau \ln a)\right] + \frac{1}{6}\ln[L/l]. \tag{15}$$

Thus,

$$S_{1|2} = \frac{1}{144}l^2a^2\left[1 + O(la, l\partial_\tau \ln a)\right] - \frac{1}{6}\ln[L/l]. \tag{16}$$

Now we define *moving-mirror entropy* so that it measures how much the motion of the mirror increases the uncertainty of the quantum state of wave-packet modes which are reflected by the mirror after τ_0. The simplest procedure is to subtract $S_{1|2}$ for a non-accelerating mirror from $S_{1|2}$ for the actual trajectory:

$$S_{MM}(\tau_0) \equiv S_{1|2} - S_{1|2}|_{a=0} = \frac{1}{144}l^2a^2(\tau_0)\left[1 + O(la, l\partial_\tau \ln a)\right]. \tag{17}$$

3 Summary and discussion

In summary we have calculated von Neumann conditional entropy for a quantized scalar field in a two-dimensional Minkowski spacetime with a moving mirror and proposed a definition of moving-mirror entropy S_{MM} as the difference between von Neumann conditional entropy for the actual mirror trajectory and the inertial trajectory. We expect that S_{MM} is related to uncertainty about the future or temporarily inaccessible information. Since black-hole entropy has many interpretations [23, 24, 25], probably it is also worth while investigating other possible interpretations of the moving-mirror entropy.

References

[1] S. Mukohyama and W. Israel, Phys. Rev. **D62**, 121501 (2000).

[2] S. W. Hawking, Nature, **248**, 30 (1974); Commun. Math. Phys. **43**, 199 (1975).

[3] P. Panangaden and R. M. Wald, Phys. Rev. **D16**, 929 (1977).

[4] S. Mukohyama, Phys. Rev. **D56**, 2192 (1997).

[5] D. N. Page, Phys. Rev. Lett. **44**, 301 (1980).

[6] D. N. Page, Phys. Rev. Lett. **71**, 3743 (1993).

[7] S. Mukohyama, Phys. Rev. **D58**, 104023 (1998).

[8] S. Mukohyama, Phys. Rev. **D61**, 064015 (2000).

[9] G. T. Horowitz, gr-qc/9604051; gr-qc/9704072.

[10] J. M. Maldacena, Ph. D. thesis, hep-th/9607235.

[11] A. W. Peet, hep-th/0008241.

[12] J. Maldacena, Adv. Theor. Math. Phys., **2**, 231(1998).

[13] O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri and Y. Oz, hep-th/9905111 and reference therein.

[14] A. Strominger and C. Vafa, Phys. Rev. Lett. **43**, 199 (1996).

[15] S. W. Hawking, in *Black Holes and Relativistic Stars*, edited by R. M. Wald, (University of Chicago Press 1998).

[16] S. A. Fulling, Phys. Rev. **D7**, 2860 (1973).

[17] P. C. W. Davies, J. Phys. **A8**, 609 (1975).

[18] N. J. Cerf, and C. Adami, Phys. Rev. Lett. **79**, 5194 (1997).

[19] N. J. Cerf, and C. Adami, Phys. Rev. **A55**, 3371 (1997).

[20] T. M. Fiola, J. Preskill, A. Strominger, and S. P. Trivedi, Phys. Rev. **D50**, 3987 (1994).

[21] W. G. Unruh, Phys. Rev. **D14**, 870 (1976).

[22] W. Israel, Phys. Lett. **57A**, 107 (1976).

[23] V. Frolov and D. Fursaev, Class. Quant. Grav. **15**, 2041 (1998).

[24] V. Iyer and R. M. Wald, Phys. Rev. **D52**, 4430 (1995).

[25] S. Mukohyama, Ph. D. thesis, gr-qc/9812079.

—Quantum Information and Generalized Second Law of Black Hole —

Akio HOSOYA[1] ,Alberto Carlini[2] ,and Takeshi Shimomura[3]

Department of Physics, Tokyo Institute of Technology
Oh-Okayama, Meguro-ku, Tokyo 152
Japan

Abstract

We study the generalized second law of black hole in the light of quantum information theory. In a quantum gedanken experiment we compute the sum of black hole entropy, matter entropy and entropy of detector and show that the total entropy increases.

1 Introduction

The generalized second law of black hole physics states that the total of Bekenstein-Hawking black hole entropy[1][2], which is a quarter of the area A of the event horizon, and the ordinary thermodynamical entropy $S^{\rm matter}$

$$S^{\rm total} = \frac{1}{4}A + S^{\rm matter} \tag{1}$$

always increases.

An essentially classical gedanken experiment initiated by Bekenstein[3] and discussed by many people. A box of mass E^{box} and entropy S^{box} is lowered by a string from infinity to a point near the horizon. The box is then thrown into the black hole. Apparently the entropy contained in the box is gone for an observer in the outside region of the black hole. This decrease of entropy of the outside region of the black hole is compensated by the increase of the Bekenstein-Hawking entropy of the black hole due to the increase of a mass of the black hole which is ultimately attributed to the work done by the observer at the other end of the string located far away. Unruh and Wald [4] [5]studied the work taking into account the buoyancy force by the thermal atmosphere created by the quantum field theoretical effect in an accelerated system and concluded that the generalized second law is not violated in this gedanken experiment.

In the present work we pursue this direction of research so that we elucidate the role of the quantum entanglement of the states for inside and outside region of a black hole. We shall apply a quantum information theoretical approach[6] in a gedanken quantum experiment in the black hole spacetime. We shall see that a quantum measurement reduces the entanglement entropy but the increase of the uncertainty of the measurement outcomes and the increase of the Bekenstein-Hawking entropy by increase of the mass which comes from the work needed for the measurement together give the increase of the total entropy. The entanglement was first discussed by Frolov and Page[7] in the context of black hole, where they considered the change of generalized entropy before and after collapse of a star to a black hole.

2 Quantum Maxwellian demon

In this section we first see the decrease of the entanglement entropy by measurement, a quantum version of Maxwellian demon problem. We then see that the entropy acquired by the measurement apparatus is more than that decrease.

Consider an entangled state:

$$|\psi> = \sum_n \sqrt{c_n}|n>_B |n>_A |\Phi_0>, \quad c_0 \geq c_1 \ldots \geq 0 \tag{2}$$

[1] E-mail:ahosoya@th.phys.titech.ac.jp
[2] E-mail:carlini@th.phys.titech.ac.jp
[3] E-mail:tshimomu@th.phys.titech.ac.jp

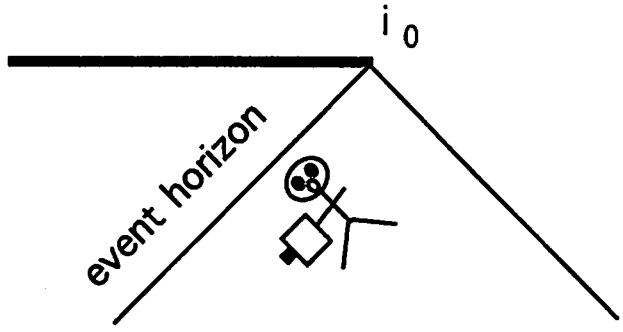


Figure 1: Quantum measurement near the horizon

in a Schmidt decomposition form. Here $|n>_A$ is the Alice state and $|n>_B$ is the Bob state, which are entangled.We have also introduced a detector state, which is initially $|\Phi_0>$. For simplicity we assume that the detector state is either $|\Phi_0>$ or $|\Phi_1>$.

Suppose further that we are completely ignorant of the Bob state $|>_B$ so that we trace over the Bob states to obtain a density matrix for the Alice states and the detector state;

$$\rho = Tr_B|\psi><\psi| = \sum_n c_n|n>_A|\Phi_0>_A<n|<\Phi_0| \tag{3}$$

so that the von Neumann entropy is given by

$$S^{(\text{von Neumann})} = -Tr(\rho log\rho) = -\sum_n c_n log c_n \tag{4}$$

,which is called the entanglement entropy.

Let us switch on the apparatus so that the Alice state and the detector state undergoes an unitary evolution to obtain

$$|\psi>\rightarrow|\psi'>$$
$$|\psi'>=\sum_n\sqrt{c_n}|n>_B\sum_{m,\alpha}S^\alpha_{nm}|m>_A|\Phi_\alpha>,$$

where the S^α_{nm} is the matrix element of the unitary transformation.

The measurement process involves also a split of a wave packet into macroscopically separated ones corresponding to the "collapse of wave functions":

$$|\psi'>\rightarrow|\psi'^\alpha>$$
$$|\psi'^\alpha>=\sum_n\sqrt{\frac{c_n}{K_\alpha}}|n>_B\sum_m S^\alpha_{nm}|m>_A|\Phi_\alpha>,$$

where the normalization, K_α:

$$K_\alpha = \sum_{m,n} c_n|S^\alpha_{nm}|^2, \quad (\sum_\alpha K_\alpha = 1) \tag{5}$$

is the probability to obtain the state $|\psi'^\alpha>$ after the measurement.

For the splitting of the wave packet, one may imagine the archetypical Stern-Gerlach experiment.

We then do a unitary transformation[4] T^α to the Bob state $|n>_B$ knowing the outcome α of the experiment via classical communication so that the resultant state is of the Schmidt form,

$$\begin{aligned}|\tilde{\psi}'^\alpha> &= \sum_n\sqrt{\frac{c_n}{K_\alpha}}(\sum_l T^\alpha_{nl}|l>_B)\sum_m S^\alpha_{nm}|m>_A|\Phi_\alpha> \\ &= \sum_m\frac{1}{\sqrt{K_\alpha}}W^\alpha_m|m>_B|m>_A|\Phi_\alpha>.\end{aligned}$$

[4] This unitary transformation does no affect the entropy S'^α. This process can be skipped.

Here the unitary matrix T^α is arranged so that

$$\sum_n \sqrt{c_n} T^\alpha_{nl} S^\alpha_{nm} = W^\alpha_m \delta_{l,m} \tag{6}$$

holds. For each α the density matrix is

$$\begin{aligned} \rho'^\alpha &= Tr_B |\tilde{\psi}'^\alpha >< \tilde{\psi}'^\alpha| \\ &= \sum_m \frac{|W^\alpha_m|^2}{K_\alpha} |m>_A \, <m|_A \end{aligned}$$

and the corresponding conditional entropy is

$$\begin{aligned} S'^\alpha = S^{(\text{von Neumann})}(\rho'^\alpha) &= -Tr(\rho'^\alpha log \rho'^\alpha), \\ &= -\sum_m \frac{|W^\alpha_m|^2}{K_\alpha} log \frac{|W^\alpha_m|^2}{K_\alpha}. \end{aligned}$$

Therefore the weighted mean entropy S'^{ave} is

$$S'^{ave} = \sum_\alpha K_\alpha S'^\alpha. \tag{7}$$

This is the entropy we expect knowing that the experiment has been done but without knowing its outcome.

We can show by the cocavity of the von Neumann entropy that

$$S'^{ave} = \sum_\alpha K_\alpha S'^\alpha \leq S^{(\text{von Neumann})}(\sum_\alpha K_\alpha \rho'^\alpha) = S^{(\text{von Neumann})}(\rho). \tag{8}$$

Therefore the average entropy decreases by the measurement process.

In general Nielsen[8] showed that by local operation which consists of unitary transformation of the Alice and detector states and the measurement causes a transition:

$$\begin{aligned} |\psi> &= \sum_n \sqrt{c_n} |n>_B \, |n>_A \rightarrow \\ |\psi'> &= \sum_n \sqrt{c'_n} |n>_B \, |n>_A, \end{aligned}$$

with $c'_0 \geq c_0, c'_0 + c'_1 \geq c_0 + c_1, \ldots c'_0 + c'_1 + \cdots = 1 = c_0 + c_1 + \cdots$, i.e. $\{c'\}$ "majorizes" $\{c\}$. [5] In the present case,

$$\begin{aligned} c'_m &= \frac{|W^\alpha_m|^2}{K_\alpha} \\ |W^\alpha_m|^2 &= \sum_n c_n |S^\alpha_{nm}|^2. \end{aligned}$$

The majorization implies that

$$-\sum_n c'_n log c'_n \leq -\sum_n c_n log c_n \tag{9}$$

from the concavity of the entropy. At any rate the average entropy decreases by the observation. This is not surprising because the wave function collapse accompanied by any measurement cuts off the entanglement and thus reduces the entanglement entropy.

However, the total entropy of the detector entropy

$$S^{det} = -\sum_\alpha K_\alpha log K_\alpha \tag{10}$$

and average entropy is shown to increase,

$$S^{tot} = S^{det} + S^{ave} = -\sum_m |W^\alpha_m|^2 log |W^\alpha_m|^2 \geq -\sum_n c_n log c_n = S. \tag{11}$$

This increase of total entropy is sometimes called Landauer's principle.[9]

[5] For readers not familiar with quantum information theory, the following example may be helpful to grasp the majorization. Consider an entangled state: $\frac{1}{\sqrt{2}}(|0>_A \, |0>_B + |1>_A \, |1>_B)$, i.e. $c_0 = c_1 = \frac{1}{\sqrt{2}}$. Upon measurement of the Bob sate, the whole state will collapse to either $|0>_A \, |0>_B$ or $|1>_A \, |1>_B$,i.e. $c'_0 = 1, c'_1 = 0$ or $c'_0 = 0, c'_1 = 1$. The lesson is that the measurement tends to concentrate the coefficients.

3 The generalized second law in the black hole thermodynamics

3.1 statement of the problem

As is by now well-known, the radiation state of an eternal black hole can be the Hartle-Hawking state:

$$|\psi>=\Pi_\omega \sum_n \sqrt{c_n}|n>_B |n>_A \tag{12}$$

with $c_n = e^{-\frac{n\omega}{T_{BH}}}/Z$. Here T_{BH} is the Hawking temperature of the black hole and Z is the partition function.$|n>_B$is the n particle states of the inside of the event horizon while $|n>_A$is the n particle states of the outside of the event horizon. (Alice is in the out region while Bob is in the in region.) It should be stressed that the entanglement is naturally produced by a black hole. Hereafter to simplify the presentation we consider a particular mode of the angular frequency ω.

Imagine we do a local operation by observing the Alice states. Tracing over the Bob's states we see that the entanglememt entropy decreases as we have seen in the previous section. Of course, this does not mean the violation of the second law. We take into account the detector entropy and the totality will increase as we have seen before. However, what happens if their detector carrying the entropy acquired by the measurement is dropped into the black hole? It appears that the decreased radiation entropy by the local operation has no way to be compensated by the detector entropy, which is gone into the black hole.

3.2 Bekenstein's classical gedanken experiment

To elucidate the meaning of the generalized second law, a gedanken experiment was initiated by Bekenstein and discussed by Geroch and Unruh and Wald. They considered a box with entropy S^{box} of energy E^{box},which is lowered by a string from infinity to a point near the event horizon. The apparent violation of the second law $\Delta S^{matter} = -S^{box}$ is more than compensated by the increase of the black hole entropy increase $\Delta S^{BH} = \frac{\Delta E}{T_{BH}}$. More precisely, $\Delta E = E^{box} + \Delta W$, where ΔW is the work done by the observer at the other end of the string which is infinitely far away from the horizon. The lower bound of the work in the isothermal process can be obtained by the ordinary second law of the thermodynamics:

$$\Delta W \geq F_{after} - F_{before}, \tag{13}$$

where $F_{after}(F_{before})$ is the Helmholz free energy measured by the distant observer after(before) cutting off the rope which has suspended the box. The equality holds for quasiequlibrium processes. Clearly,

$$F_{after} = (E^{box} - T^{box}S^{box})\chi$$
$$F_{before} = (E^{box} - T^{box}_{\infty}S^{box}) + (E^{rad} - \tilde{T}S^{rad})\chi,$$

where E^{rad} and S^{rad} are the energy and the entropy of the radiation. χ is the red-shift factor of the floating point of the box, where the gravity and the buoyancy force to the box balances. By choosing the floating point we see that

$$\Delta S^{tot} = \Delta S^{BH} + \Delta S^{matter} \geq S^{rad} - S^{box} \geq 0, \tag{14}$$

because the box entropy is maximized at the equilibrium distribution.

We would like to emphasize that the generalized second law makes sense only when we consider the entropy of the outside region and can be shown through the ordinary second law of thermodynamics. The decrease of the matter entropy by the dropping box is more than compensated by the increase of the black hole entropy which arises from the work needed for the dropping of the box.

4 The generalized second law and quantum information

Consider a situation that an observer performs a quantum measurement in the outside region of a black hole. The measurement is arranged so that it splits the wave packet prepared on the hypersurface Σ

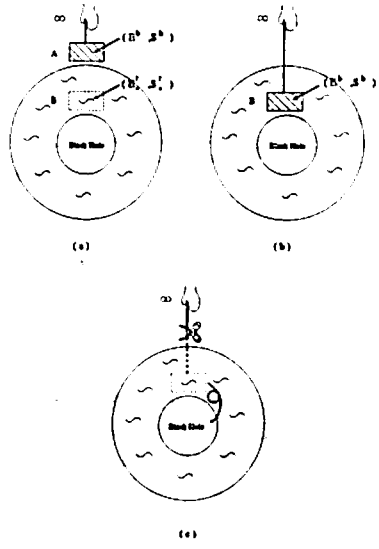


Figure 2: The box is initially located outside the cavity and then lowered down towards the region near the horizon. We then cut the string and the box is thrown into the black hole.

according to the detector states $|\Phi_0>$ and $|\Phi_1>$ on the hypersurface Σ'. On the final hypersurface Σ'', only one of the packets, $|\Phi_0>$ remains at the end of the string, while the other one in the state $|\Phi_1>$ is freely falling into the black hole.

On the hypersurface Σ', the reduced density matrix ρ'^0 obtained by tracing over the Bob states

$$\rho'^{\alpha} = Tr_B|\tilde{\psi}'^{\alpha}><\tilde{\psi}'\alpha|$$
$$= \sum_m \frac{|W_m^{\alpha}|^2}{K_{\alpha}}|m>_A <m|_A$$

and the corresponding entropy is $S'^{\alpha} = -Tr_A(\tilde{\rho}'^{\alpha} log\tilde{\rho}'^{\alpha}), (\alpha = 0, 1)$.

By Nielsen's inequality we have an apparent violation of the second law

$$S'^{\alpha} \leq S = -\sum_n c_n log c_n \quad (\alpha = 0, 1) \tag{15}$$

which is caused by the local operation involving the measurement.

In Sec.2, we have seen that

$$K_0 S'^0 + K_1 S'^1 - K_0 log K_0 - K_1 log K_1 \geq S$$

Let us consider first the total entropy change $\Delta S^{tot} = \frac{\Delta W}{T_{BH}} + \Delta S^{matter}$ by the measurement before the conditional cutting-off of the string. The change of matter entropy is

$$\Delta S^{matter} = (K_0 S'^0 + K_1 S'^1 - K_0 log K_0 - K_1 log K_1) - S^{rad}.$$

Note that this is the entropy change by the quantum measurement but without knowing the outcomes. On the other hand, the work done by the measurement is bounded from below by

$$\Delta W \geq \frac{F_{after} - F_{before}}{T_{BH}}$$

with

$$F_{after} = K_0(E_0 - \tilde{T}S'^0)\chi + K_1(E_1 - \tilde{T}S'^1)\chi$$
$$F_{before} = (E - \tilde{T}S^{rad})\chi.$$

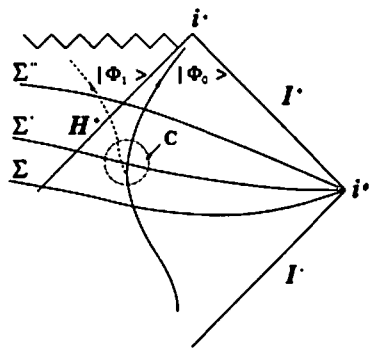


Figure 3: We prepare the wave packet on the hypersurface Σ, and then perform a quantum measurement which splits the wave packet into macroscopically distant ones, according to the detector states $|\Phi_0>$ and $|\Phi_1>$ on the hypersurface Σ'. On the final hypersurface Σ'' only one of the packets, $|\Phi_0>$, remains at the end of the string, while the other, $|\Phi_1>$, is freely falling into the black hole.

By using the energy conservation,

$$E = K_0 E_0 + K_1 E_1$$

we see that

$$\Delta' S^{tot} \geq -K_0 log K_0 - K_1 log K_1 \geq 0.$$

Now proceed to the entropy change by the cutting-off process. The matter entropy change is

$$\begin{aligned} \Delta S^{matter} &= (K_0 S'_0 - K_0 log K_0 - K_1 log K_1) - (K_0 S'^0 + K_1 S'^1 - K_0 log K_0 - K_1 log K_1) \\ &= -K_1 S'^1 \ (\geq 0), \end{aligned}$$

while the work needed is bounded by

$$\begin{aligned} \Delta W \ &\geq F_{after} - F_{before} = [K_0(E_0 - \tilde{T}S'^0)\chi + K_1 E_1 \chi] - [K_0(E_0 - \tilde{T}S'^0)\chi + K_1(E_1 - \tilde{T}S'^1)\chi] \\ &= K_1 S'^1 T_{BH}. \end{aligned}$$

Hence we see that also in the final step

$$\Delta'' S^{tot} \geq 0.$$

Note that for the free falling $|\Phi_1>$ the temperature is zero so that the entropy part is absent and that $\tilde{T}\chi = T_{BH}$ as before.

Putting all the results together we conclude that the whole change of the entropy is

$$\Delta S^{tot} \geq -K_0 log K_0 - K_1 log K_1 \geq 0.$$

The decrease of the matter entropy is more than compensated by the increase of the black hole entropy via the increase of the black hole mass which is ultimately attributed to the work done by the measurement. The lower bound is nothing but the entropy corresponding to the uncertainty of the outcome of the experiment.

5 Summary and outlook

We have proposed a quantum version of gedanken experiment which supports the generalized second law of black hole physics. A quantum measurement of particles in the outside region of event horizon decreases the entropy of the outside matter by the entanglement of the inside and outside particle states. As far as the detector remains outside the sum of the detector entropy acquired by the measurement and

the change of the particle entropy is positive; the ordinary second law holds. We discuss the case that the detector is conditionally dropped into the black hole depending on the experimental outcome. For the outside region of a black hole the generalized entropy increases, while the matter entropy decreases when the detector is conditionally dropped into the black hole. The decrease of the matter entropy is more than compensated by the increase of the black hole entropy via the increase of the black hole mass which is ultimately attributed to the work done by the measurement.

We have only discussed the case in which we know that experiment has been done near the horizon but not its outcomes for the observer at infinity. Of course we can easily see by letting K_0 or K_1 to unity that the generalized entropy increases even if the outcomes are revieled. We can generalize our discussion to a mixed state for the initial state so that the initial detector entropy is no-zero. This case may have more direct correspondence with Bekenstein's gedanken experiment.

Many people have talked about so-called "information loss" in the evaporation process of a black hole. However, the "information" has not yet been quantitatively defined in the black hole physics. It is our belief that physical information should be properly defined in black hole physics and the well-developed machinery of the so-called quantum information will be useful. The present work is just its first step in this direction.

References

[1] J.D.Bekenstein,*Phys.Rev.* **D7** 2333 (1973),**D9** 3292 (1974),**D12** 3077(1975).

[2] S.W. Hawking,*Commun.Math.Phys.* **43** 199 (1975).

[3] J.D.Bekenstein,*Phys.Rev.* **D23** 287 (1981).

[4] W.G. Unruh and R.M. Wald,*Phys.Rev.* **D25** 942 (1982).

[5] W.G. Unruh and R.M. Wald,*Phys.Rev.* **D27** 2271 (1983).

[6] A. Peres, "Quantum Theory: Concepts and Methods", (1995) Kluwer Academic Publishers.

[7] V.P.Frolov and D.Page, *Phys.Rev.Lett.* **71**(24) 3902 (1993).

[8] M.A.Nielsen,*Phys.Rev.Lett.* **83** 436 (1999).

[9] W.H. Zurek, Maxwell's demon,Szilard's engine and quntum measurements in Frontiers of Non-equilibrium Statistical Physics,(Plenum Press,1984)

"Asymptotic" Killing horizon

Jun-ichirou Koga,[1]

Advanced Research Institute for Science and Engineering, Waseda University, Shinjuku, Tokyo 169-8555, Japan

Abstract

We consider the asymptotic symmetry near the cosmological horizon in the de Sitter spacetime. We first derive the asymptotic Killing vectors, along which the Lie derivative of the perturbed metric vanishes on the horizon, so that they generate the asymptotic symmetry. We show that the $O(3)$ rotations of the horizon two-sphere and the supertranslations along the null direction are allowed as the asymptotic symmetry. Then, it is shown that the local properties of the Killing horizon are preserved under the transformations generated by the asymptotic symmetry. We also analyze the Poisson bracket algebra of the conserved quantities conjugate to the asymptotic Killing vector, and show that it does not have non-trivial central extension. Finally, we briefly discuss the role of the asymptotic symmetry near the horizon in the investigation of the microscopic states of the horizon.

1 Introduction

Due to the recent developments of the technique of the D-branes[1] and the AdS/CFT correspondence[2, 3], the string theory also has provided the successful methods to count the microscopic states of the black hole and derive the Bekenstein-Hawking formula, where the conformal field theory and its algebra, the so-called Virasoro algebra, play the central role. However, these derivations of the black hole entropy heavily rely on the features particular to the black hole solutions in the string theory; supersymmetry or constancy of the curvature tensor. We cannot apply these methods to the generic black hole, even the simplest one, the Schwarzschild black hole. Furthermore, it is not clear where those microscopic states of the black hole live; on, inside or outside the horizon. In order to resolve these issues, it will be important to understand whether the microscopic states of the arbitrary black hole are described locally *on the horizon* by use of the conformal field theory.

One of the possible approaches will be to consider the asymptotic symmetry near the horizon. The well-known examples of the asymptotic symmetry in the four-dimensional asymptotically flat spacetime are the B.M.S. group[4] and the Spi group[5] at the null and spacelike infinity, respectively. Another important example is the conformal group at the spacelike infinity in the three-dimensional anti-de Sitter spacetime[6]. Indeed, the Bekenstein-Hawking formula of the B.T.Z. black hole has been derived by using the asymptotic symmetry in the three-dimensional anti-de Sitter spacetime[2], although it is the asymptotic symmetry at infinity, not on the horizon. The asymptotic symmetries mentioned above are closely related to the asymptotic behavior of the spacetime geometries. Therefore, if the microscopic states of the black hole live on the horizon, we may expect that they will be described by a kind of asymptotic symmetry near the horizon. In addition, a remarkable feature of those asymptotic symmetries is that they are generated by much larger groups than that of the maximal symmetry. Recall that a large number of degrees of freedom will be required for the entropy of a sufficiently large black hole, which we will be able to analyze semi-classically. The asymptotic symmetry is able to associate a large number of degrees of freedom to the local geometry near the horizon.

Then, in this contribution, we apply the idea of the asymptotic symmetry to the horizon. Although some works have been done in this context, where the microscopic states of the horizon are described by the conformal field theory[7, 8], they have not given a satisfactory solution, unfortunately; their formulations contain some failures and the physical meaning of the microscopic states has not become clear. One of the reasons seems that the asymptotic behavior of the metric near the horizon is analyzed,

[1] E-mail:koga@gravity.phys.waseda.ac.jp

explicitly or implicitly, in a singular coordinate system. It is still to be clarified even that there exists a well-defined asymptotic symmetry near the horizon.

To analyze this issue, we focus on the Killing horizon in the four-dimensional spacetime. Although we specifically work in the case of the cosmological horizon in the de Sitter spacetime for simplicity, it is straightforward to generalize to the black hole horizon of the Schwarzschild spacetime.

Another crucial point worth investigation is the local structures of the horizon. If the microscopic states of the black hole reside on the horizon, it may be possible that those are described geometrically by the local structures of the horizon. Then, it should be clarified how the asymptotic symmetry near the horizon, if any, is related to the local structures of the horizon. Although the event horizon is defined globally and hence its local description seems impossible, it is known that the event horizon coincides with the Killing horizon, if the spacetime is sufficiently symmetric. The properties of the Killing horizon are expressed in the local forms, which suits to our purpose. We then analyze the relation between the asymptotic symmetry and the local structures of the Killing horizon.

In addition, it is also important to clarify whether the asymptotic symmetry is described by the Virasoro algebra, as in the AdS/CFT correspondence. In the case of the AdS/CFT correspondence, the Poisson bracket algebra of the conserved charges associated to the asymptotic symmetry, which generate the transformations of the canonical variables in the phase space, is given by the Virasoro algebra. Thus, we also calculate the Poisson bracket algebra of the asymptotic symmetry near the horizon, and discuss whether or not the asymptotic symmetry is responsible for the black hole entropy.

2 Asymptotic symmetry near the horizon

We first analyze existence and features of an asymptotic symmetry near the Killing horizon in the four-dimensional spacetime by focusing on the cosmological horizon in the de Sitter spacetime, in a similar way to the usual asymptotic symmetries, such as the B.M.S. group[4] and the conformal group in the anti-de Sitter spacetime[6]. The wording "asymptotic" will not be appropriate here, because the horizon locates at the finite physical distance from the center of the de Sitter spacetime, while it is infinitely far in terms of the tortoise coordinate. However, here we call it an "asymptotic symmetry", because of its similarity to the usual, genuine asymptotic symmetries.

In the usual asymptotic symmetries, we consider metric perturbations with the boundary condition that the perturbed metric coincides with the background metric on the boundary surface that we focus on, e.g., the null infinity in the case of the B.M.S. group. In the present case, the background corresponds to the de Sitter spacetime and the boundary surface to the cosmological horizon. In the analysis of the asymptotic symmetry, however, we need not to consider the whole of the boundary surface. Here, we are concerned only with a local region of the cosmological horizon, which is spanned by a horizon two-sphere (a cross section of the horizon and a Cauchy surface) during a small interval of the null affine parameter on the horizon.

In order to impose the boundary condition on the cosmological horizon appropriately, we work in the regular coordinate system on the horizon, where the background metric ${}^0g_{\mu\nu}$ is written as

$$ds^2 = {}^0g_{\mu\nu}\,dx^\mu dx^\nu = -\frac{4a^4}{(uv-a^2)^2}\,du\,dv + a^2\frac{(uv+a^2)^2}{(uv-a^2)^2}\,d\Omega^2\,. \tag{1}$$

Here, a is related to the cosmological constant Λ by $a^2 = 3/\Lambda$, and Kruskal coordinates u and v are defined by

$$u \equiv a\exp\left[\frac{U}{a}\right]\,, \qquad v \equiv -a\exp\left[-\frac{V}{a}\right]\,, \tag{2}$$

where the retarded and advanced null coordinates U and V are given by $U \equiv t - r^*$ and $V \equiv t + r^*$ respectively, in terms of the Killing time coordinate t and the tortoise coordinate r^*.

Then, the future and past cosmological horizon locate at $v = 0$ and $u = 0$, respectively. In what follows, we will specifically focus on the future cosmological horizon, whereas we can analyze the past horizon in the same manner.

As in the usual asymptotic symmetries, we first perturb the metric ${}^0g_{\mu\nu}$ of the background spacetime, so that the components of the perturbed metric $g_{\mu\nu}$ in the coordinate system of Eq.(1) coincide on the

future horizon ($v = 0$) with those of the background metric ${}^0g_{\mu\nu}$. By assuming suitable differentiability of $g_{\mu\nu}$ in the local region of the horizon under consideration, the asymptotic form of $g_{\mu\nu}$ near the future horizon will be written as

$$g_{\mu\nu} = {}^0g_{\mu\nu} + v\,{}^1g_{\mu\nu}(u,\theta,\phi) + \mathcal{O}(v^2)\,, \tag{3}$$

where the perturbation of the metric, described by ${}^1g_{\mu\nu}$, depends arbitrarily on u, θ and ϕ.

Now we consider a class of the metrics with the asymptotic forms of Eq.(3) and the coordinate transformations which transform any metric in this class into another (possibly the same) one in this class. This defines a symmetry between the metrics in this class and a sub-group of the general diffeomorphism, which we call the asymptotic symmetry and the asymptotic group, respectively.

The infinitesimal transformation of the asymptotic group will be generated by the infinitesimal coordinate transformation, $x^\mu \to x^\mu - \zeta^\mu$, where $g_{\mu\nu}$ is transformed by the Lie derivative along the vector ζ^μ as $g_{\mu\nu} \to \bar{g}_{\mu\nu} = g_{\mu\nu} + \mathcal{L}_\zeta g_{\mu\nu}$, with the components of both metrics satisfying the boundary condition near the horizon, i.e., $g_{\mu\nu} = {}^0g_{\mu\nu} + \mathcal{O}(v)$ and $\bar{g}_{\mu\nu} = {}^0g_{\mu\nu} + \mathcal{O}(v) = g_{\mu\nu} + \mathcal{O}(v)$. Thus, the infinitesimal transformation of the asymptotic group is generated by *the asymptotic Killing vector* ζ^μ, which satisfies *the asymptotic Killing equation*

$$\mathcal{L}_\zeta\, g_{\mu\nu} = \mathcal{O}(v)\,, \tag{4}$$

for the arbitrarily perturbed metric $g_{\mu\nu}$ that satisfies the boundary condition Eq.(3).

Then, we derive the asymptotic Killing vector ζ^μ, so that it solves the asymptotic Killing equation, Eq.(4), for any perturbations of the metric ${}^1g_{\mu\nu}$. We find that the solution of ζ^μ exists and its general form is written as

$$\zeta^\mu = \omega(u,\theta,\phi)\,\xi^\mu_{(h)} + a^{(i)}\,\xi^\mu_{(i)} + v\,X^\mu(\omega) + \mathcal{O}(v^2)\,. \tag{5}$$

Here, $\xi^\mu_{(h)}$ is the timelike Killing vector in the de Sitter spacetime that becomes null on the cosmological horizon, and $\xi^\mu_{(i)}$ denotes three Killing vectors which generate the $O(3)$ rotations of the horizon two-sphere. (Summation over the index (i) is understood.) The function $\omega(u,\theta,\phi)$ depends arbitrarily on u, θ and ϕ, the coefficients $a^{(i)}$ are arbitrary constants and $X^\mu(\omega)$ is defined as

$$X^\mu(\omega) \equiv \frac{2u}{a}\,\nabla^\mu\omega(u,\theta,\phi)\,. \tag{6}$$

We see that $\zeta^\mu\,\nabla_\mu v = \mathcal{O}(v)$ near the horizon. Thus, the asymptotic Killing vector ζ^μ is always tangent to the horizon ($v = 0$), and so it never pushes or pulls the horizon. We also see that ζ^μ consists of the $O(3)$ rotations and the position-dependent translations in the null direction on the horizon described by the arbitrary function $\omega(u,\theta,\phi)$. Those position-dependent translations are very similar to the supertranslations in the B.M.S. group, which is also the asymptotic symmetry near the null hypersurface, but the supertranslations in the B.M.S. group have the angular dependence only. It is a distinguishing feature that ζ^μ can depend also on u in the arbitrary manner, and this results in interesting consequences, as we will see below. Thus, the rigid rotations of the horizon two-sphere and the supertranslations along the null direction are allowed as the asymptotic symmetry near the cosmological horizon.

The Lie bracket between two asymptotic Killing vectors, ζ^μ_1 and ζ^μ_2, is given as

$$\left[\zeta_1\,,\,\zeta_2\right]^\mu = \mathcal{L}_{\zeta_1}\zeta^\mu_2 = \omega_{12}\,\xi^\mu_{(h)} + a^{(i)}_{12}\,\xi^\mu_{(i)} + v\,X^\mu(\omega_{12}) + \mathcal{O}(v^2)\,, \tag{7}$$

where $\omega_{12} \equiv \zeta^\nu_1\,\nabla_\nu\omega_2 - \zeta^\nu_2\,\nabla_\nu\omega_1$ and $a^{(i)}_{12} \equiv a^{(j)}_1 a^{(k)}_2 C_{(j)(k)}{}^{(i)}$ with $C_{(i)(j)}{}^{(k)}$ being the structure constant of the $O(3)$ group. This has the form of the asymptotic Killing vector, Eq.(5). Therefore, Eq.(7) explicitly shows that the transformations generated by the asymptotic Killing vectors close, and hence they form the Lie bracket algebra of the asymptotic group, as they should.

3 Asymptotic Killing horizon

The asymptotic symmetry we have derived above leaves invariant the asymptotic behavior of the metric near the horizon. However, it does not indicate that the local structures of the horizon are also unaffected. In this section, we analyze how the the transformations generated by the asymptotic symmetry are related to the properties of the horizon which characterize its local structures.

We first specify the local properties of the horizon, which must be the generalization of those of the usual Killing horizon, so that the null hypersurface under consideration locally looks like a Killing horizon.

The usual, global Killing horizon in the background spacetime ${}^0g_{\mu\nu}$ is generated by the timelike Killing vector $\xi^\mu_{(h)}$, which becomes null on the horizon and is normal (and so tangent) to the horizon. It then follows that, on the horizon, $\xi^\mu_{(h)}$ is hypersurface orthogonal and its derivative along itself, $\xi^\nu_{(h)}\,{}^0\nabla_\nu\,\xi^\mu_{(h)}$, is proportional to itself, with the proportional factor being the surface gravity κ.

However, if we focus only on the horizon, we need not take care that these properties are satisfied in far regions from the horizon. When the Killing vector $\xi^\mu_{(h)}$ in the background spacetime satisfies

$$\mathcal{L}_{\xi_{(h)}} g_{\mu\nu} = \mathcal{O}(v)\,, \tag{8}$$

$$g_{\mu\nu}\,\xi^\mu_{(h)}\,\xi^\nu_{(h)} = \mathcal{O}(v)\,, \tag{9}$$

$$\xi^{[\mu}_{(h)}\nabla^\nu\xi^{\rho]}_{(h)} = \mathcal{O}(v)\,, \tag{10}$$

$$\xi^\nu_{(h)}\nabla_\nu\,\xi^\mu_{(h)} = \kappa\,\xi^\mu_{(h)} + \mathcal{O}(v)\,, \tag{11}$$

in the local region of the hypersurface $v = 0$, where the Killing horizon in the background spacetime ${}^0g_{\mu\nu}$ is located, it locally looks like a part of a Killing horizon even in the perturbed spacetime $g_{\mu\nu}$. Then we call such a region of the hypersurface *an asymptotic Killing horizon*. Thus, an asymptotic Killing horizon shares the same local properties with the usual Killing horizon.

Next, we consider the transformations generated by the asymptotic symmetry. In the previous section, the asymptotic group has been derived as a sub-group of diffeomorphism, where the metric is transformed as $g_{\mu\nu} \to g_{\mu\nu} + \mathcal{L}_\zeta g_{\mu\nu}$. If we consider the transformation generated by diffeomorphism, the vector $\xi^\mu_{(h)}$, which plays the central role in the properties of the asymptotic Killing horizon, is also transformed by the Lie derivative as $\xi^\mu_{(h)} \to \xi^\mu_{(h)} + \mathcal{L}_\zeta\xi^\mu_{(h)}$. Then, the four properties of the asymptotic Killing horizon, Eqs.(8)–(11), are preserved under the transformation, since those are defined covariantly and we have $\zeta^\mu\,\nabla_\mu v = \mathcal{O}(v)$. However, this is a trivial transformation. We look at the same horizon in the same spacetime from a different position on the horizon. (The asymptotic Killing vector ζ^μ is tangent to the horizon, and hence it moves a point on the horizon to another point that is still on the horizon.)

As we mentioned before, we are interested in the microscopic states of the horizon. Although those states have not been identified so far, it may be possible that they are represented by the different geometries, at least when the semi-classical analysis is valid. Then, they will be related to each other by the non-trivial transformations, rather than diffeomorphism.

Notice here that even when we transform the metric by the Lie derivative, it does not necessarily mean that the transformation under consideration is diffeomorphism. The metric $g_{\mu\nu}$ (dynamical variable) and the vector $\xi^\mu_{(h)}$ (non-dynamical variable independent of the metric) can be separately transformed, in general. This will be well illustrated, when we consider the transformation, described by

$$g_{\mu\nu} \longrightarrow \bar{g}_{\mu\nu} = g_{\mu\nu} + \mathcal{L}_\zeta\, g_{\mu\nu} \equiv g_{\mu\nu} + v\,\gamma_{\mu\nu} + \mathcal{O}(v^2)\,, \tag{12}$$

$$\xi^\mu_{(h)} \longrightarrow \bar{\xi}^\mu_{(h)} = \xi^\mu_{(h)}\,. \tag{13}$$

Here, the metric is transformed by the Lie derivative, but the vector $\xi^\mu_{(h)}$ is held fixed. We call this transformation the dynamical field transformation.

Although the dynamical field transformation changes the metric into another one, we can pull it back by diffeomorphism to the original one, since the metric is transformed by the Lie derivative under the dynamical field transformation. When we apply the dynamical field transformation followed by such diffeomorphism, the result is the spacetime with the same metric as the original one, but the vector $\xi^\mu_{(h)}$ transformed. It will be appropriate to call this transformation the horizon deformation, since the vector $\xi^\mu_{(h)}$ characterizes the horizon. Now the metric $g_{\mu\nu}$ is fixed, but the vector $\xi^\mu_{(h)}$ is transformed by the Lie derivative along ζ^μ, as

$$g_{\mu\nu} \longrightarrow \tilde{g}_{\mu\nu} = g_{\mu\nu}\,, \tag{14}$$

$$\xi^\mu_{(h)} \longrightarrow \tilde{\xi}^\mu_{(h)} = \xi^\mu_{(h)} + \mathcal{L}_\zeta\xi^\mu_{(h)}\,. \tag{15}$$

where $\mathcal{L}_\zeta\xi^\mu_{(h)}$ is calculated, by using Eq.(7), as

$$\mathcal{L}_\zeta\xi^\mu_{(h)} = -\,\Omega\,\xi^\mu_{(h)} + v\,X^\mu(-\Omega) + \mathcal{O}(v^2)\,, \tag{16}$$

with Ω defined by

$$\Omega \equiv \xi^{\nu}_{(h)} \nabla_{\nu} \omega \,. \tag{17}$$

The transformations of the properties of the asymptotic Killing horizon, Eqs.(8)-(11), are found to be

$$\mathcal{L}_{\xi_{(h)}} g_{\mu\nu} = \mathcal{O}(v) \quad \longrightarrow \quad \mathcal{L}_{\tilde{\xi}_{(h)}} g_{\mu\nu} = \mathcal{O}(v) \,, \tag{18}$$

$$g_{\mu\nu}\, \xi^{\mu}_{(h)}\, \xi^{\nu}_{(h)} = \mathcal{O}(v) \quad \longrightarrow \quad g_{\mu\nu}\, \tilde{\xi}^{\mu}_{(h)}\, \tilde{\xi}^{\nu}_{(h)} = \mathcal{O}(v) \,, \tag{19}$$

$$\xi^{[\mu}_{(h)} \nabla^{\nu} \xi^{\rho]}_{(h)} = \mathcal{O}(v) \quad \longrightarrow \quad \tilde{\xi}^{[\mu}_{(h)} \nabla^{\nu} \tilde{\xi}^{\rho]}_{(h)} = \mathcal{O}(v) \,, \tag{20}$$

$$\xi^{\nu}_{(h)} \nabla_{\nu}\, \xi^{\mu}_{(h)} = \kappa + \mathcal{O}(v)\; \xi^{\mu}_{(h)} \quad \longrightarrow \quad \tilde{\xi}^{\nu}_{(h)} \nabla_{\nu}\, \tilde{\xi}^{\mu}_{(h)} = \tilde{\kappa}\; \tilde{\xi}^{\mu}_{(h)} + \mathcal{O}(v) \,, \tag{21}$$

respectively, where $\tilde{\kappa}$ is defined by

$$\tilde{\kappa} = \kappa - \kappa\Omega - \xi^{\nu}_{(h)} \nabla_{\nu} \Omega \,. \tag{22}$$

Therefore, the local properties of the asymptotic Killing horizon are preserved under the horizon deformation. In other words, the asymptotic symmetry generates the deformations of the horizon, under which the horizon keeps its local structures. Since the horizon deformation is equivalent to the dynamical field transformation upto diffeomorphism and the local properties of the asymptotic Killing horizon are preserved under diffeomorphism, those properties are preserved also under the dynamical transformation. Notice here that the null coordinate dependence of the arbitrary function ω in the expression Eq.(5), which describes the supertranslations, generates the non-trivial transformations of the vector $\xi^{\mu}_{(h)}$, i.e., $\Omega \neq 0$ in Eq.(16).

We also immediately see that any covariant tensor equations, which depend only on the metric but not on the non-dynamical variables (such as the vector $\xi^{\mu}_{(h)}$), are left invariant under both the dynamical field transformation and the horizon deformation. Thus, the Einstein equation holds in the perturbed spacetime.

4 Poisson bracket algebra

In this section, we investigate the Poisson bracket algebra of the conserved charges conjugate to the asymptotic Killing vectors, which generate the transformations of the canonical variables in the phase space. To do this, we employ the covariant phase space formulation developed by Wald and his collaborators[9].

In the covariant phase space formulation, the conserved charge $\mathcal{H}[g;\zeta]$ conjugate to the asymptotic Killing vector ζ^{μ} on the Cauchy surface $\mathcal{C}$ is written as

$$\bar{\delta}\, \mathcal{H}[g;\zeta] = \bar{\delta} \int_{\partial \mathcal{C}} \frac{1}{2}\, \epsilon_{\beta\alpha\mu\nu}\, Q^{\beta\alpha}(g;\zeta) + \int_{\partial \mathcal{C}} \epsilon_{\beta\alpha\mu\nu}\, \zeta^{\beta}\, \Theta^{\alpha}(g;\bar{\delta} g) \,, \tag{23}$$

when the Einstein equation for $g_{\mu\nu}$ and the linearized Einstein equation for $\bar{\delta} g_{\mu\nu}$ are satisfied. Here, $\Theta^{\alpha}(g;\bar{\delta} g)$ is defined by the variation of the Lagrangian density $\epsilon_{\mu\nu\rho\sigma}\, L(g)$, as

$$\bar{\delta} \Big(\epsilon_{\mu\nu\rho\sigma}\, L(g) \Big) = \epsilon_{\mu\nu\rho\sigma}\, E_{\mu\nu}\, \bar{\delta} g^{\mu\nu} + \epsilon_{\mu\nu\rho\sigma}\, \nabla_{\beta}\, \Theta^{\beta}(g;\bar{\delta} g) \,, \tag{24}$$

with $E_{\mu\nu} \equiv R_{\mu\nu} - g_{\mu\nu} R/2 + g_{\mu\nu}\Lambda = 0$ being the Einstein equation. Then, we find that the vector defined by $J^{\beta}(g;\zeta) \equiv \Theta^{\beta}(g;\mathcal{L}_{\zeta} g) - \zeta^{\beta}\, L(g)$ is divergence free on shell, and the anti-symmetric tensor $Q^{\beta\alpha}(g;\zeta)$ is defined by $J^{\beta}(g;\zeta) \equiv \nabla_{\alpha} Q^{\beta\alpha}(g;\zeta)$.

As we see from Eq.(23), $\bar{\delta}\, \mathcal{H}[g;\zeta]$ is not defined as a total variation. Hence, it is not ensured that $\mathcal{H}[g;\zeta]$ always exists. Indeed, Wald and Zoupas[10] showed that the necessary and sufficient condition for the existence of $\mathcal{H}[g;\zeta]$ is given by

$$\int_{\partial \mathcal{C}} \zeta^{\beta}\, \omega_{\beta\mu\nu}(g;\bar{\delta}_1 g, \bar{\delta}_2 g) = 0 \,, \tag{25}$$

where

$$\omega_{\mu\nu\rho}(g;\bar{\delta}_1 g, \bar{\delta}_2 g) \equiv \bar{\delta}_2 \Big(\epsilon_{\beta\mu\nu\rho}\, \Theta^{\beta}(g;\bar{\delta}_1 g) \Big) - \bar{\delta}_1 \Big(\epsilon_{\beta\mu\nu\rho}\, \Theta^{\beta}(g;\bar{\delta}_2 g) \Big) \,. \tag{26}$$

For the Einstein theory with the cosmological constant, $\Theta^\beta(g;\bar{\delta}g)$ is given by

$$\Theta^\beta(g;\bar{\delta}g) = \frac{1}{16\pi}\left[g_{\mu\nu}\nabla^\beta\bar{\delta}\,g^{\mu\nu} - \nabla_\nu\,\bar{\delta}\,g^{\nu\beta}\right], \tag{27}$$

and it is easy to show that Eq.(25) is satisfied under the boundary condition that we imposed to derive the asymptotic symmetry, i.e., $\bar{\delta}\,g_{\mu\nu} = \mathcal{O}(v)$ near the horizon. Therefore, we see that $\mathcal{H}[g;\zeta]$ exists.

We now analyze the Poisson bracket between $\mathcal{H}[g;\zeta_1]$ and $\mathcal{H}[g;\zeta_2]$, conjugate to two asymptotic Killing vectors ζ_1^μ and ζ_2^μ, respectively. Recall here that the Poisson bracket algebra of the asymptotic symmetry in the three-dimensional anti-de Sitter spacetime takes the form of

$$\left\{\mathcal{H}[g;\zeta_1]\,,\ \mathcal{H}[g;\zeta_2]\right\} = \mathcal{H}[g;\mathcal{L}_{\zeta_1}\zeta_2] + K[\zeta_1,\zeta_2]\,, \tag{28}$$

where $K[\zeta_1,\zeta_2]$ is constant under the variation $\bar{\delta}$ and is called the central term. Then, we may expect that the Poisson bracket algebra in the present case also takes the same form. Indeed, when the linearized Einstein equation for $\bar{\delta}g^{\mu\nu}$ and Eq.(25) are satisfied, we can show

$$\bar{\delta}\left(\bar{\delta}_{\zeta_2}\mathcal{H}[g;\zeta_1] - \mathcal{H}[g;\mathcal{L}_{\zeta_1}\zeta_2]\right) = 0\,. \tag{29}$$

Thus, by integrating Eq.(29), we immediately obtain

$$\bar{\delta}_{\zeta_2}\mathcal{H}[g;\zeta_1] = \mathcal{H}[g;\mathcal{L}_{\zeta_1}\zeta_2] + K[\zeta_1,\zeta_2]\,. \tag{30}$$

Since $\mathcal{H}[g;\zeta_2]$ will generate the variation $\bar{\delta}_{\zeta_2}\mathcal{H}[g;\zeta_1]$ of $\mathcal{H}[g;\zeta_1]$ through the Poisson bracket, we see that the Poisson bracket algebra is given by Eq.(28).

In order to obtain the central term $K[\zeta_1,\zeta_2]$, which does not depend on the metric $g^{\mu\nu}$, we calculate the value of $\bar{\delta}_{\zeta_2}\mathcal{H}[g;\zeta_1] - \mathcal{H}[g;\mathcal{L}_{\zeta_1}\zeta_2]$ on the background spacetime ${}^0g^{\mu\nu}$. By the straightforward calculation, we have

$$\begin{aligned}
&\bar{\delta}_{\zeta_2}\mathcal{H}[{}^0g;\zeta_1] - \mathcal{H}[{}^0g;\mathcal{L}_{\zeta_1}\zeta_2] \\
&= -\mathcal{H}[{}^0g;\mathcal{L}_{\zeta_1}\zeta_2] + \int_{\partial\mathcal{C}}\frac{3}{2}\,{}^0\varepsilon_{\beta\alpha\mu\nu}\,{}^0\nabla_\gamma\left(\zeta_1^{[\beta}\,Q^{\alpha\gamma]}({}^0g;\zeta_2)\right) \\
&\quad + \int_{\partial\mathcal{C}}\frac{1}{2}\,{}^0\varepsilon_{\beta\alpha\mu\nu}\Big[\mathcal{L}_{\zeta_2}Q^{\beta\alpha}({}^0g;\zeta_1) - \mathcal{L}_{\zeta_1}Q^{\beta\alpha}({}^0g;\zeta_2) - Q^{\beta\alpha}({}^0g;\mathcal{L}_{\zeta_2}\zeta_1) \\
&\quad + Q^{\beta\alpha}({}^0g;\zeta_1)\left({}^0\nabla_\gamma\,\zeta_2^\gamma\right) - Q^{\beta\alpha}({}^0g;\zeta_2)\left({}^0\nabla_\gamma\,\zeta_1^\gamma\right) + 2\,\zeta_1^{[\beta}\,\zeta_2^{\alpha]}\,L({}^0g)\Big]\,.
\end{aligned} \tag{31}$$

The second term in the righthand side of Eq.(31) is an integral of a total divergence, and it makes no contribution, since the horizon two-sphere $\partial\mathcal{C}$ has no boundary.

In the Einstein theory, $Q^{\beta\alpha}(g;\zeta)$ is given by

$$Q^{\beta\alpha}(g;\zeta) = \frac{1}{16\pi}\left[\nabla^\alpha\zeta^\beta - \nabla^\beta\zeta^\alpha\right]. \tag{32}$$

Then, we calculate as

$$\begin{aligned}
&\bar{\delta}_{\zeta_2}\mathcal{H}[{}^0g;\zeta_1] - \mathcal{H}[{}^0g;\mathcal{L}_{\zeta_1}\zeta_2] = -\mathcal{H}[{}^0g;\mathcal{L}_{\zeta_1}\zeta_2] \\
&+ \int_{\partial\mathcal{C}}\frac{1}{16\pi}\,{}^0\varepsilon_{\beta\alpha\mu\nu}\left[\left({}^0\nabla^\gamma\zeta_1^\alpha\right)\left({}^0\nabla_\gamma\,\zeta_2^\beta\right) - \left({}^0\nabla^\gamma\zeta_1^\beta\right)\left({}^0\nabla_\gamma\,\zeta_2^\alpha\right) - {}^0R^{\beta\alpha}{}_{\gamma\lambda}\,\zeta_1^\gamma\,\zeta_2^\lambda + \zeta_1^{[\beta}\,\zeta_2^{\alpha]}\left({}^0R - 2\,\Lambda\right)\right] \\
&+ \int_{\partial\mathcal{C}}\frac{1}{2}\,{}^0\varepsilon_{\beta\alpha\mu\nu}\left[Q^{\beta\alpha}({}^0g;\zeta_1)\left({}^0\nabla_\gamma\,\zeta_2^\gamma\right) - Q^{\beta\alpha}({}^0g;\zeta_2)\left({}^0\nabla_\gamma\,\zeta_1^\gamma\right)\right] + (\text{surface term})\,,
\end{aligned} \tag{33}$$

where we substituted $L = \frac{1}{16\pi}\,(R - 2\Lambda)$. We immediately see that the second integral (the third term) in Eq.(33) does not contribute, by recalling the fact that ζ^μ is the asymptotic Killing vector, and hence

${}^0\nabla_\lambda \zeta^\lambda$ goes to zero on the horizon. Furthermore, we can show that the the first integral (the second term) also vanishes on the horizon, when we use the Einstein equation, the asymptotic Killing equation for ζ^μ and the hypersurface orthogonality of the Killing vector $\xi^\mu_{(h)}$.

Therefore, we obtain

$$\bar{\delta}_{\zeta_2}\mathcal{H}[{}^0g;\zeta_1] - \mathcal{H}[{}^0g;\mathcal{L}_{\zeta_1}\zeta_2] = -\mathcal{H}[{}^0g;\mathcal{L}_{\zeta_1}\zeta_2]\,, \tag{34}$$

or, $\bar{\delta}_{\zeta_2}\mathcal{H}[{}^0g;\zeta_1] = 0$. The Poisson bracket algebra is then expressed as

$$\Big\{\mathcal{H}[g;\zeta_1]\,,\ \mathcal{H}[g;\zeta_2]\Big\} = \mathcal{H}[g;\mathcal{L}_{\zeta_1}\zeta_2] - \mathcal{H}[{}^0g;\mathcal{L}_{\zeta_1}\zeta_2]\,. \tag{35}$$

The last term of Eq.(35) looks like a central term. However, we should check whether or not the central term can be eliminated by the constant shift of $\mathcal{H}[g;\zeta]$, since it is defined only by the variation as Eq.(23). Although the Poisson bracket does not change under the shift of the value of the conserved charge, the righthand side of Eq.(35) does. If we redefine $\mathcal{H}[g;\zeta]$ as $\mathcal{H}[g;\zeta] \to \mathcal{H}'[g;\zeta] = \mathcal{H}[g;\zeta] - \mathcal{H}[{}^0g;\zeta]$, the Poisson bracket algebra becomes

$$\Big\{\mathcal{H}'[g;\zeta_1]\,,\ \mathcal{H}'[g;\zeta_2]\Big\} = \mathcal{H}'[g;\mathcal{L}_{\zeta_1}\zeta_2]\,, \tag{36}$$

where the central term has disappeared.

This result is contrasted with the case of the Virasoro algebra, where the central term cannot be eliminated by the shift of the value of the conserved charge.

5 Summary and discussion

We have analyzed the asymptotic symmetry near the cosmological horizon in the four-dimensional de Sitter spacetime, which is generated by the asymptotic Killing vector. The asymptotic Killing vector is shown to consist of the rigid $O(3)$ rotations of the horizon two-sphere and the supertranslations along the null direction, which depend not only on the angular coordinates, but also on the null coordinate, in contrast to the supertranslations in the B.M.S. group.

Then, we have found that the properties that characterize the local structures of the Killing horizon are left invariant under the transformations generated by the asymptotic symmetry. We considered two types of transformations, which are equivalent to each other upto diffeomorphism. One is the dynamical field transformation, where the metric is transformed by the Lie derivative along the asymptotic Killing vector, while the horizon is fixed. The other is the horizon deformation, where the metric is held fixed but the timelike Killing vector in the background spacetime is subject to the transformation generated by the Lie derivative along the asymptotic Killing vector. Thus, we have seen that the asymptotic symmetry generates the deformations of the horizon, which keep the local structures of the horizon.

Although we have shown these results in the case of the cosmological horizon in the de Sitter spacetime, we can readily derive the same results for the black hole horizon of the Schwarzschild spacetime. Then, we emphasize that the results obtained here are not specific to the cosmological horizon.

We also showed that the Poisson bracket algebra of the conserved charges, conjugate to the asymptotic Killing vectors, does not have non-trivial central extension, whereas the Virasoro algebra does. Thus, the asymptotic group does not contain, as a sub-group, the conformal group described by the Virasoro algebra. This indicates that, if the entropy associated to the horizon of the arbitrary black hole is described by the Virasoro algebra, as in the AdS/CFT correspondence, the asymptotic symmetry derived here is not responsible for the black hole entropy.

Nevertheless, the results obtained here will provide the useful foundations, based on which we may extend the asymptotic symmetry by weakening the boundary condition on the horizon. It will be expected that, by such an extended asymptotic symmetry, we can describe the microscopic states of the horizon responsible for the black hole entropy. Since the asymptotic symmetry we have found is not responsible for the black hole entropy, when we find an extended asymptotic symmetry whose algebra contains the Virasoro algebra as a sub-algebra, we will understand the geometrical origin of the black hole entropy by analyzing the behavior of such extension.

References

[1] A. Strominger and C. Vafa, Phys. Lett. **B379**, 99 (1996); see, for review, G. T. Horowitz, preprint gr-qc/9604051.

[2] A. Strominger, JHEP **02**, 009 (1998).

[3] J. Maldacena, Adv. Theor. Math. Phys. **2**, 231 (1998); S. S. Gubser, I.R. Klebanov and A. M. Polyakov, Phys. Lett. **B428**, 105 (1998); E. Witten, Adv. Theor. Math. Phys. **2**, 253 (1998); O. Aharony, S. S. Gubser and J. Maldacena, H. Ooguri and Y. Oz, Phys. Rep. **323**, 183 (1998).

[4] H. Bondi, M. G. J. van der Burg and A. W. K. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962); R. K. Sachs, Proc. Roy. Soc. (London) **A270**, 103 (1962); R. K. Sachs, Phys. Rev. **128**, 2851 (1962).

[5] A. Ashtekar and R. O. Hansen, J. Math. Phys. **19**, 1542 (1978); A. Ashtekar, in *General Relativity and Gravitation, Vol.2*, ed. A. Held (New York, Plenum).

[6] J. D. Brown and M. Henneaux, Comm. Math. Phys. **104**, 207 (1986).

[7] S. Carlip, Phys. Rev. Lett. **82**, 2828 (1999).

[8] S. Carlip, Class. Quantum Grav. **16**, 3327 (1999).

[9] J. Lee and R. M. Wald, J. Math. Phys. **31**, 725 (1990); R. M. Wald, Phys. Rev. D **48**, 3427 (1993); V. Iyer and R. M. Wald, Phys. Rev. D **50**, 846 (1994).

[10] R. M. Wald and A. Zoupas, Phys. Rev. D **61**, 084027 (2000).

On black holes in three-dimensions

Daisuke Ida[1]

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

Abstract

The black holes in three-dimensional gravity seem to exist only when there is a negative cosmological constant. This is explained by the non-existence theorem of black holes for three-dimensional theories satisfying the dominant energy condition.

The (2+1)-dimensional theory provides us with one of useful approaches to more complicated (3+1)-dimensional classical gravity or conceptual problems in quantum gravity [1]. At first sight, the (2+1)-dimensional gravity looks trivial. In particular, the vacuum Einstein equation implies that the space-time is locally flat, corresponding to the absence of the gravitational radiation (Weyl tensor) in three dimensions. However, the local distribution of matter fields has a global effect on the outer empty space; for instance, the gravitational field of a point particle is described by a conical space with its deficit angle corresponding to the mass of the particle [2], which causes the gravitational lens effect. One should also note that the triviality of local geometry does not necessarily imply the triviality of the theory itself; namely, the topological degrees of freedom plays an important role in the theory of gravitation [3, 4]. The triviality of local geometry in the (2+1)-gravity theory holds even if the cosmological term is taken into account. The Einstein space is simply a space of constant curvature in three dimensions, so that educated relativists would not imagine that there is a black hole solution in this theory until in 1992 Bañados *et al.* show that there actually exists a black hole in the locally anti-de Sitter space [5, 6]. This black hole space-time, called BTZ black hole, is obtained by identifying certain points of (the covering manifold of) the anti-de Sitter space. A different identification makes a space-time representing the BTZ black hole in a closed universe [9], multiple BTZ black holes [10] or a creation of the BTZ black hole [11]. The BTZ black hole is characterized by the mass, angular momentum and cosmological constant, and has almost all features of the Kerr-anti-de Sitter black hole in the conventional four-dimensional Einstein gravity. The BTZ black hole was shown to be also the solution of a low energy string theory [7, 8].

It might be interesting to note that all the black hole solutions listed above require a negative cosmological constant, otherwise a certain kind of energy conditions is violated. A typical example might be the BTZ black hole. As already mentioned, the BTZ black hole may be constructed by making identifications in the anti-de Sitter space. We may also consider a similar construction in the de Sitter space. In this case, a natural procedure might be identifying two geodesic circles in each Poincaré disk associated with the open chart of the de Sitter space. The resultant space-time represents an inflating universe rather than a black hole. The absence of black hole in this example might be due to the difference in the causal structure of conformal infinity.

In Ref. [12], I have given a reason for this situation. In particular, we will be able to answer the question: *"Why the BTZ black hole requires a negative cosmological constant?"* by showing the following theorem.

Theorem 1: *Let (M, g) be a three-dimensional space-time subject to the Einstein equation* $\mathrm{Ric} - (R/2)g + \Lambda g = -8\pi T$ *with* $\Lambda > 0$. *If the stress-energy tensor T satisfies the dominant energy condition, then (M, g) contains no apparent horizons.*

This explains why black hole solutions require a negative cosmological constant. Strictly speaking, we can only say that there is no non-degenerate apparent horizon ($\rho = 0$, $\mathcal{L}_X \rho \neq 0$) in the case of $\Lambda = 0$, however, the presence of matter fields such as the dilaton or Maxwell field will exclude even degenerate horizons.

Thus, a black hole in (2+1)-gravity requires negative energy such as a negative cosmological constant. This implies the breakdown of the predictability in certain three-dimensional theories. As in four space-time dimensions, we may consider the Oppenheimer-Snyder model of the gravitational collapse. The

[1] E-mail:ida@tap.scphys.kyoto-u.ac.jp

homogeneous disk of dust will collapse to a central point and a naked conical singularity will be left. This picture of gravitational collapses will remain unchanged unless the negative cosmological constant is added. Even in the case of the non-symmetric gravitational collapse of gauge fields or scalar fields, there will not form a black hole, so that when a singularity is formed, such a singularity will be naked.

We have disscussed the existence problem of apparent horizons, while the black hole is often defined by the event horizon. Since the theorem 1 relies on the local analysis, we cannot argue the global structure of space-time such as an event horizon. An important exception is the stationary case; we can replace "apparent horizons" with "stationary event horizons" in the theorem 1, since it is known that these coincide in this case.

The author was supported by JSPS Research Fellowships for Young Scientists, and this research was supported in part by the Grant-in-Aid for Scientific Research Fund (No. 4318).

References

[1] S. Carlip, *Quantum gravity in (2+1)-dimensions*, (Cambridge, Cambridge University Press, 1998).

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

[3] E. Witten, Nucl. Phys. B **311**, 46 (1988);

[4] A. Hosoya and K. Nakao, Class. Quantum Grav. **7**, 163 (1990).

[5] M. Bañados, C. Teitelboim and J. Zanelli, Phys. Rev. Lett. **69**, 1849 (1992).

[6] M. Bañados, M. Henneaux, C. Teitelboim and J. Zanelli, Phys. Rev. D **48**, 1506 (1993).

[7] G.T. Horowitz and D.L. Welch, Phys. Rev. Lett. **71**, 328 (1993).

[8] N. Kaloper, Phys. Rev. D **48**, 2598 (1993).

[9] M. Siino, Class. Quantum Grav. **11**, 1995 (1994).

[10] D.R. Brill, Phys. Rev. D **53**, R4133 (1996).

[11] H-J. Matschull, Class. Quantum Grav. **16**, 1069 (1999).

[12] D. Ida, Phys. Rev. Lett. **85**, 3758 (2000).

Scalar hair on the black hole in the presence of cosmological constant

Takashi Torii,[1]

Research Center for the Early Universe,
University of Tokyo, Bunkyo-ku, Tokyo, 113-0033, Japan, and
Advanced Research Institute for Science and Engineering,
Waseda University, Shinjuku-ku, Tokyo 169-8555, Japan

Makoto Narita[2]

Department of Physics, Rikkyo University, Toshima-ku, Tokyo 171-8501, Japan

Kengo Maeda[3]

Department of Physics, Waseda University, Shinjuku-ku, Tokyo 169-8555, Japan

Abstract

We examined the no-hair conjecture in the asymptotically anti-de Sitter (AdS) spacetime. For the first step, we consider the real scalar field as the matter field and assume the static spherically symmetric spacetime. By analyzing the asymptotic behavior of the scalar field, we find that the scalar field must approaches the extremum of its potential. By this fact, it is proved that there is no regular black hole solution when the scalar field is massless or has a "convex" potential. While the scalar field has a growing mode around the local minimum potential, there is no such mode around the local maximum. It implies that the local maximum is a kind of "attractor" of the asymptotic scalar field. We give an example of the new black hole solutions with nontrivial scalar field configuration numerically in the double well potential model. We study its stability in order to examine the scalar hair is physical or not. We find that the potential function of the perturbation equation is positive semi-definite in some wide parameter range and that the new solution is stable. As a result, the black hole no-hair conjecture is violated in asymptotically AdS spacetime.

1 Introduction

The exterior gravitational field of a stationary source may have a large number of independent multipole moments. But when the source lies within an black hole event horizon (BEH), a radical simplification occurs as proposed by Ruffini and Wheeler[1]: after the gravitational collapse of the matter field, the resultant black hole approaches stationary spacetime and that all its multipole moments are then uniquely determined by two parameters, M and a, which are physically interpreted as the mass and the angular momentum of the black hole. When the source has a net charge Q, then of course the parameter is also requested to uniquely determine its (electric and gravitational) multipole moments. This statement is called black hole no hair conjecture.

In order to examine whether the black hole no-hair conjecture is true or not, someone tried to prove this conjecture and someone tried to construct a counter-example after the proposal. In the former approach, several no-hair theorems established. For example, the black hole uniqueness theorems in electrovacuum theories[2] strongly support the conjecture. The works of Ref. [3] show that stationary black hole solutions are hairless in a variety of theories coupling classical fields to Einstein gravity.

[1] E-mail:torii@resceu.s.u-tokyo.ac.jp
[2] E-mail:narita@se.rikkyo.ac.jp
[3] E-mail:g_maeda@gravity.phys.waseda.ac.jp

In the proof of these theories, one usually assume matter fields, symmetry of spacetime and asymptotic condition. In the latter approach, some of these assumptions were relaxed. Considering different matter fields, we have several kind of black hole solutions. One of the most impressive solutions is colored black hole solution which is the solution in the Einstein-Yang-Mills system[4]. Although this solution was found to be unstable[5], non-Abelian hair is generic and many other non-Abelian black holes were discovered after the colored black hole[6].

Most of the proofs of the no-hair theorems impose asymptotic flatness as the asymptotic condition. Hence the following natural question arises: *Can we extend no-hair theorems to spacetimes with different asymptotic structure?* The authors first studied the scalar hair in the asymptotically de Sitter spacetime[7]. In the asymptotically flat case, Bekenstein[8] and Sudarsky[?] provided simple proofs for the no-scalar hair theorem in spherically symmetric spacetime in the case where the matter consists of multiple scalar fields with an arbitrary positive semi-definite potential. In the asymptotically de Sitter case, if the scalar field is massless or has a "convex" potential such as a mass term, it was proved that there is no regular black hole solution. Here the "convex" stand for that if $dV/d\phi = 0$ for $\exists\phi$ then $d^2V/d\phi^2 > 0$. For a general positive semi-definite potential, we searched for black hole solutions which support the scalar field with a double well potential, and found them by numerical calculations. When we take the zero horizon radius limit, the solution becomes a boson star like solution[9]. These black hole solutions are, however, unstable against the linear perturbations. As a result, we can conclude that the no-scalar hair conjecture holds in the case of scalar fields with a "convex" or a double well potential. We expect that this no-scalar hair theorem is extended to general positive semi-definite potential.

Then what happens if we consider the system with the negative cosmological constant, especially the asymptotically anti-de Sitter (AdS) spacetime. Recently, a tremendous amount of interest has evolved in several issues related to the AdS spacetime, such as the AdS/CFT (conformal field theory) correspondence[10] and the brane world scenarios[11]. It shoud be noted that the colored black hole solution is stabilized by putting the negative cosmological constant into the system[12]. Moreover, it was shown that there is a black hole solution with dyonic hair which can not be exist in asymptotically flat spacetime[13]. Hence we expect that the negative cosmological constant will affect the stability and existence of black hole solutions.

2 Model and Basic Equations

We will consider the model given by the action

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{16\pi G}(R - 2\Lambda) - \frac{1}{2}(\nabla\phi)^2 - V(\phi)\right], \tag{1}$$

where ϕ is the real scalar field and $V(\phi)$ is its potential. We shall assume the cosmological constant Λ is negative. The metric of a spherically symmetric spacetime can be expressed in the Schwarzschild type form,

$$ds^2 = -fe^{-2\delta}dt^2 + f^{-1}dr^2 + r^2d\Omega^2, \quad f = 1 - \frac{2Gm}{r} - \frac{\Lambda}{3}r^2. \tag{2}$$

The mass function m and the lapse function δ depend on both the time coordinate t and the radial coordinate r.

Varying the action (1) and substituting *Ansatz* (2), we derive the field equations

$$m' = 4\pi r^2\left[\frac{1}{2}f^{-1}e^{2\delta}\dot{\phi}^2 + \frac{1}{2}f\phi'^2 + V(\phi)\right], \tag{3}$$

$$\delta' = -4\pi r\left[f^{-2}e^{2\delta}\dot{\phi}^2 + \phi'^2\right], \tag{4}$$

$$\dot{m} = 4\pi r^2 f\dot{\phi}\phi', \tag{5}$$

$$-\left[e^{\delta}f^{-1}\dot{\phi}\right]^{\cdot} + \frac{1}{r^2}\left[r^2e^{-\delta}f\phi'\right]' = e^{-\delta}\frac{dV(\phi)}{d\phi}. \tag{6}$$

Here, we have used the dimensionless variables, $\sqrt{|\Lambda|}t \to t$, $\sqrt{|\Lambda|}r \to r$, $\sqrt{|\Lambda|}Gm \to m$, $\sqrt{G}\phi \to \phi$ and $GV/|\Lambda| \to V$. By using these variables, f is expressed as

$$f = 1 - \frac{2m}{r} + \frac{r^2}{3}. \tag{7}$$

We will use these dimensionless variables below. A dot and a prime in the field equations denote derivatives with respect to the dimensionless variables t and r, respectively.

For the boundary conditions of the metric functions, we first impose the existence of a regular black hole event horizon (BEH) at $r = r_B$. Secondary, we impose the non-existence of singularity outside of the BEH. As for the asymptotic behavior, we can expect roughly three possibilities depending on the form of the potential $V(\phi)$. (i) If $\lambda_{eff} < 0$, where $\lambda_{eff} := \lambda_{add} - 1 := 8\pi V(\phi_\infty) - 1$ and $\phi_\infty = \phi(\infty)$, the spacetime will approaches AdS one. (ii) If $\lambda_{eff} > 0$, the spacetime will approaches de Sitter spacetime. In this case, there should appear the cosmological event horizon (CEH) at $r = r_C$. (iii) If $\lambda_{eff} = 0$, we expect that the asymptotically flat spacetime is realized. However, some fine tuning mechanism (dynamical one or just by hand) should be needed for this case. Later we will mainly investigate the case (ii).

3 Asympotic behavior and no scalar hair theorem in "convex" potential case

In this paper, we focus on the static solution whose asymptotic structure is AdS. First of all, we have to give the definition of the asymptotically AdS spacetime. In our metric ansatz, it is reasonable to define it as (i) $f \sim Ar^2$ where $A > 0$ and (ii) $\delta \to \delta_\infty$. From the condition (i), the mass function behaves $m \lesssim Br^3$ [$B < 1/6$ (or $B < |\Lambda|/6$ in the dimensional variables) in the $m \sim Br^3$ case]. By using Eq. (3),

$$m \sim 4\pi \int_{r_B}^{r} \left[\frac{1}{2} A r^4 \phi'^2 + V(\phi) r^2 \right] dr + m_B, \tag{8}$$

in the static case. Hence, $\phi' \lesssim O(r^{-1})$ and $V(\phi) \lesssim$ const. If $\phi' \sim O(r^{-1})$, however, the scalar field diverges as $\sim \ln r$. This is incompatible with the boundary condition of the scalar field, and we find $\phi' \sim O(r^{-1})$. From the condition (ii) and Eq. (4), we can obtain the same bahavior.

We study the asymptotic behavior of the scalar field by using its field equation (6) in detail under the static *Ansatz*. Substituting the asymptotic form of the metric functions, we obtain

$$Ar^2\phi'' + 4Ar\phi' - \frac{dV}{d\phi} = 0. \tag{9}$$

This is easily integrated as

$$Ar^4\phi' = \int r^2 \frac{dV}{d\phi} dr. \tag{10}$$

Since the left hand side of this equation behaves as $\sim O(r^3)$, $dV/d\phi \lesssim O(r^{-\epsilon})$ and $dV/d\phi \to 0$ as $r \to \infty$. This implies that ϕ_∞ takes the extremum value of the potential. If $V(\phi_\infty) > 0$, the mass function behaves $m \sim 4\pi V(\phi_\infty) r^3/3$. Hence $V(\phi_\infty) < 1/8\pi$. Otherwise the spacetime will approach de Sitter or other exotic one.

By these asymptotic behaviors, we can prove the no-hair theorem for the "convex" potential case. The case when the extremum is realized at $\phi = \infty$ is also included. We assume $\phi_{min} = 0$ where $V(\phi_{min}) = min_\phi V(\phi)$ If $\phi(r_B) = 0$, the scalar field becomes trivial. Hence, we can assume $\phi(r_B) > 0$ without loss of generality. By Eq. (6), $f'\phi' = dV/d\phi$ on the BEH. Since $f' > 0$ and $dV/d\phi >$ around the BEH, $\phi'(r_B) > 0$. At the extrema of the scalar field, i.e., $\phi' = 0$, $f\phi'' = dV/d\phi$. Hence $\phi'' > 0$ for $\phi > 0$. These imply that the scalar field must increase monotonically. Once the scalar field becomes $\phi > \phi_0$ at $r = r_0$ where $V(\phi_0) = 1/8\pi$, $V > 1/8\pi$ for $r > r_0$. Hence the CEH appears at a finite radius. Eq. (??) also hold, however, at the CEH and $f' < 0$ around the CEH. By these, ϕ must decrease at the CEH. This is a contradiction. If $\phi < \phi_0$ for $r > r_B$, ϕ must approach a constant ϕ_∞ ($< \phi_0$). However, $dV/d\phi$ must vanish there by the asymptotically AdS condition. This implies that $\phi_\infty = 0$ in the "convex" potential. This is a contradiction. The remaining possibility is that the scalar field approaches $\phi \to \phi_0$. In this

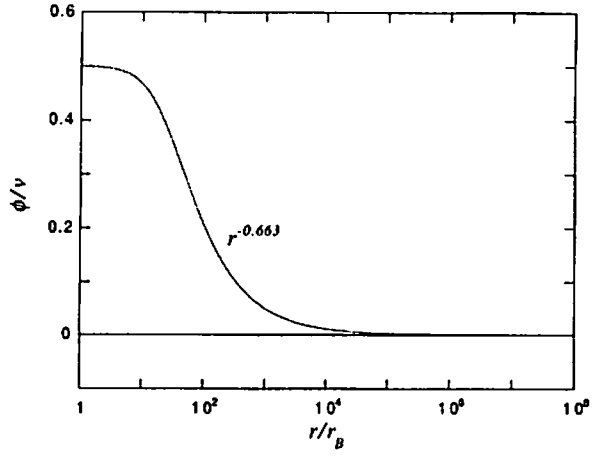


Figure 1: The configuration of the scalar field ϕ in the symmetric double wel potential model. We set the parameters $\lambda = 50$, $v = 0.1$, $r_B = 0.1$ and $\phi(r_B) = 0.5v$. The scalar field decays with the power $\lambda_+ = -0.663$

case, the spacetime becomes asymptotically flat since $\lambda_{eff} = 0$. Hence, while the left hand side of the Eq. (6) becomes zero, the right hand side is non-zero at $\phi = \phi_0$. This is a contradiction. As a result, there is no solution for the "convex" potential except for the trivial one. Note that we did not assume the asymptotically AdS condition in the proof. So the scalar hair can not be put on the black hole in any reasonable asymptotic condition if the cosmological constant is negative and the potential have "convex" shape in general.

Let us further examine the asymptotic behavior of the scalar field. Since we have known that the asymptotic value of the scalar field is the extremum of the potential, we next study how the scalar field approaches it. We assume $\phi \to \phi_\infty$ as $r \to \infty$. Defining $\varphi := \phi - \phi_\infty$ and expanding the Eq. (6), we obtain the linear equation

$$\varphi'' + 2\mu\varphi' + \omega^2\varphi = 0, \qquad \mu := \frac{1}{r} + \frac{f'}{2f}, \quad \omega^2 := -\frac{1}{f}\frac{d^2V}{d\phi^2}\bigg|_{\phi=\phi_\infty}. \tag{11}$$

Since the metric function f behaves as

$$f \to \frac{-\lambda_{eff}}{3}r^2, \tag{12}$$

we obtain

$$\varphi'' + \frac{4}{r}\varphi' + \frac{\alpha}{r^2}\varphi = 0. \tag{13}$$

The solution of this equation is

$$\varphi = C_+ r^{\lambda_+} + C_- r^{\lambda_-}, \quad \lambda_\pm := \frac{-3 \pm \sqrt{9 - 4\alpha}}{2}. \tag{14}$$

where $C_\pm$ are constants. If $\alpha < 0$, λ_+ becomes positive and the growing mode is dominant. Since negative α implies that $d^2V/d\phi^2|_{\phi=\phi_\infty} > 0$ by definition, the scalar field goes away off the local minimum of the potential but for fine tuning with which the growing mode disappears. If $0 < \alpha < 9/4$, both of $\lambda_\pm$ are negative and the scalar field approaches the local maximum with the power λ_+ in general. If $\alpha = 9/4$, the scalar field approaches the local maximum with the power $-3/2$. This corresponds to the critical damping. If $\alpha > 9/4$, there is no growing mode and the scalar field behaves

$$\varphi \sim r^{-3/2}\cos\left(\frac{\sqrt{4\alpha - 9}}{2}\ln r\right). \tag{15}$$

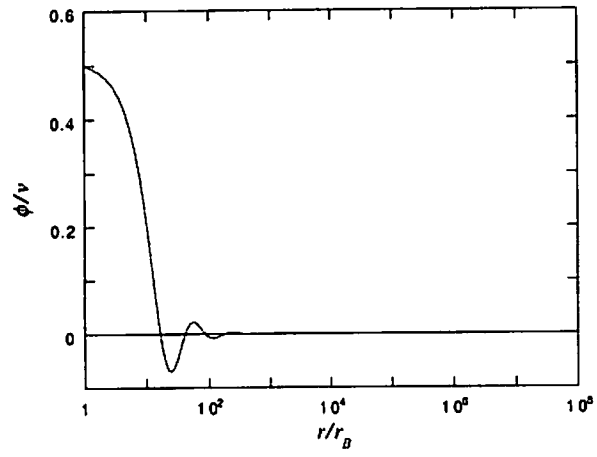


Figure 2: The configuration of the scalar field ϕ in the symmetric double wel potential model. We set the parameters $\lambda = 500$, $v = 0.1$, $r_B = 0.1$ and $\phi(r_B) = 0.5v$. The scalar field oscillates with damping. We find that the envelope of the oscillation decays with power $-3/2$ and that the wavelength is proportional to $\ln r$.

Namely the scalar field oscillate around the local maximum with the frequency $\nu = \exp(-4\pi/\sqrt{4\alpha - 9})$ and with its envelope decreasing with the power $-3/2$. As a result the local maximum of the potential is a kind of attractor of the scalar field.

4 Black hole solution with scalar hair

In this section we make sure the discussion in the previous section by using concrete models and derive nontrivial black hole solutions with a scalar hair. Here we will adopt the system including a double well potential

$$V(\phi) = \frac{\lambda}{4}(\phi^2 - v^2)^2 \tag{16}$$

as an example of a positive semi-definite potential. If we normalize the constants as $\lambda/G|\Lambda| \to \lambda$ and $\sqrt{G}v \to v$, they become dimensionless variables.

Now we search for non-trivial static solutions using numerical analysis. We drop the time derivative term of the field equations and integrate them from the BEH. We can restrict $\phi(r_B) > 0$ without loss of generality. Furthermore, we restrict $\phi(r_B) < v$, unless the scalar field diverges as in the convex potential case. Since the equation of the scalar field (6) becomes singular on the BEH, we expand it and variables by power series of $(r - r_B)$ to guarantee the regularity on the BEH, and use their analytic solutions for the first step of integration.

We found the black hole solutions with nontrivial scalar field configuration for any boundary value of the scalar field in the range $0 < \phi(r_B) < v$. This is one of the different property from the positive cosmological constant case[7], where we have to choose the suitable boundary value at the BEH by the shooting method in order to satisfy the boundary condition at the CEH.

We show the configuration of the scalar field in Fig. 1. We choose the value of the parameters as $\lambda = 50$, $v = 0.1$, $r_B = 0.1$ and $\phi(r_B) = 0.5v$. In this case, $\alpha = 1.55$. As discussed in the previous section, we find that the scalar field decays with the power $\lambda_+ = -0.663$. We show the configuration of another solution with the different parameter $\lambda = 500$, $v = 0.1$, $r_B = 0.1$ and $\phi(r_B) = 0.5v$. in Fig. 2. In this case, $\alpha = 21.9$. The scalar field oscillates with damping.

It should be noted that there are actually black hole solutions with the non-trivial scalar field in the asymptotically AdS spacetime in spite that there are no counterparts in asymptotically flat spacetime.

5 Stability analysis by linear perturbation method

In this section we investigate the stability of new solutions by using a linear perturbation method in order to check whether the scalar hair is really physical or not.

First we expand the field functions around the static solution ϕ_0, m_0 and δ_0 as follows:

$$\phi(t,r) = \phi_0(r) + \frac{\phi_1(t,r)}{r}\epsilon, \quad m(t,r) = m_0(r) + m_1(t,r)\epsilon, \quad \delta(t,r) = \delta_0(r) + \delta_1(t,r)\epsilon. \tag{17}$$

Here ϵ is an infinitesimal parameter. Substituting them into the field functions (3)-(6) and dropping the second and higher order terms of ϵ, we find the perturbation equation of the scalar field equation

$$-\frac{d^2\xi}{dr^{*2}} + U(r)\xi = \sigma^2\xi, \tag{18}$$

where we set $\phi_1 = \xi(r)e^{i\sigma t}$ and $m_1 = \eta(r)e^{i\sigma t}$. If σ^2 is positive, ϕ oscillates around the static solution and the solution is stable. On the other hand, if it is negative, the perturbation ϕ_1 and m_1 diverges exponentially with time and then the solution is unstable. r^* is the tortoise coordinate defined by $dr^*/dr = e^{\delta_0} f_0^{-1}$, where $f_0 := 1 - 2m_0/r + 1/3$. The potential function is

$$U(r) = e^{-2\delta_0} f_0 \left[(1 - 8\pi r^2 \phi_0'^2) \left(\frac{f_0'}{r} + 4\pi f_0 \phi_0'^2 \right) + 16\pi r \phi_0' \left. \frac{dV}{d\phi} \right|_0 + \left. \frac{d^2V}{d\phi^2} \right|_0 \right]. \tag{19}$$

We examine the asymptotic behavior of the potential function U. In the leading order, it behaves as

$$U(r) = \frac{2-\alpha}{9} \lambda_{eff}{}^2 e^{-2\phi_0} r^2. \tag{20}$$

Hence, when $\alpha > 2$, the potential diverges to minus infinity. This means that there exist infinite number of unstable modes and the black hole solution is unstable. On the other hand, when $\alpha < 2$, the potential diverges to plus infinity. Although it seem that there are no unstable mode in this case, there is a possibility that there appears a deep well of the potential in the central region enough to produce negative eigenmodes. This depends on the potential of the scalar field $V(\phi)$ we employ.

We plot the potential function $U(r)$ of the double well potential model in Fig. 3. The value of the parameters are $\lambda = 50$, $v = 0.1$, $r_B = 0.1$ and $\phi(r_B) = 0.5v$ ($\alpha = 1.55$). We find that the potential diverges to plus infinity as $\sim r^2$ as we analyzed. Note that it is positive semi-definite. This means that there is no unstable mode and that the black hole solution is stable. This is very important result because the no-scalar hair conjecture violates in the asymptotically AdS spacetime. The criteria of the stability is exactly $\alpha < 2$ in the double well model as obtained by the asymptotic analysis.

6 Conclusion

We examined the no-hair conjecture under the asymptotically AdS condition. For the first step, we consider the real scalar field as the matter field and assume the static spherically symmetric spacetime. By analyzing the asymptotic behavior of the scalar field, we find that the scalar field must approaches the extremum of its potential. By this fact, it is proved that there is no regular black hole solution when the scalar field is massless or has a "convex" potential like mass term. While the scalar field has a growing mode around the local minimum of the potential, there is no growing mode around the local maximum. It implies that the local maximum is a kind of "attractor" of the asymptotic scalar field. When the variable α defined in Sec. 3 satisfies $\alpha > 9/4$, the scalar field oscillates with damping around its asymptotic value ϕ_∞. When $0 < \alpha < 9/4$, it decays without oscillation numerically. We give the example of the new black hole solutions with nontrivial scalar field configuration in the double well potential model. We can find the new solution without tuning the value of the scalar field at the BEH. It shows exactly same behavior as those obtained in the asymptotic analysis. We study stability of the new solution by using the linear perturbation method in order to examine whether the scalar hair is physical or not. While for the potential with $\alpha > 2$ the new solution is unstable, the asymptotic analysis shows that the solution may be stable for the $\alpha < 2$ case. In the double well potential model, we find that the potential function

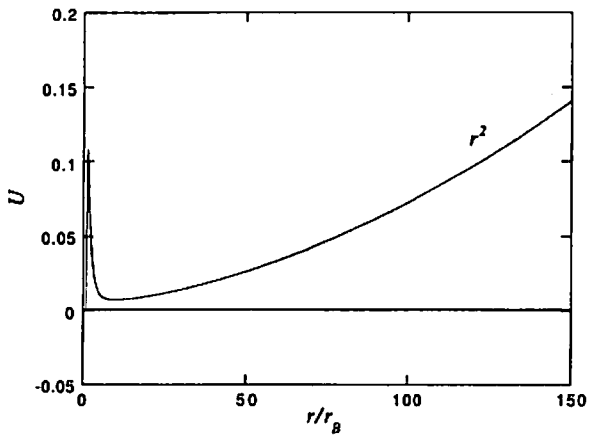


Figure 3: The profile of the potential function U of the linear perturbation equation for the double well potential model. We set the parameters $\lambda = 500$, $v = 0.1$, $r_B = 0.1$ and $\phi(r_B) = 0.5v$. We find that the potential is positive semi-definite, which means that there is no unstable mode.

of the perturbation equation is positive semi-definite in the damping solution with $\alpha < 2$ and the new solution is stable. This implies that the black hole no-hair conjecture is violated in asymptotically AdS spacetime.

In our present analysis, we show the black hole can have a scalar hair explicitly only in a particular model, i.e., in the double well potential. It is expected, however, that there are similar solutions in different models, and we confirmed it. Furthermore, we can include the gauge field in our system as in the positive cosmological case[14]. Then the scalar hair is independent on the gauge hair unlike the dilaton hair which can not exist without gauge field[15]. In this sense our scalar hair is classified into the primary hair.

The new solution has singularity at the origin which is hidden by the BEH. There will be also an everywhere regular solution, i.e., boson star solution. In the positive cosmological constant case, there exists such solutions by assuming the regular boundary condition at the origin in stead of the black hole boundary condition, and they have interesting properties[9]. Unfortunately, they were found out to be unstable against the perturbations. In the negative Λ case, however, it is expected that the boson star solution is stable since the black hole counterpart is stable for the model with $\alpha < 2$. If so and if the universe experienced the period with negative Λ at least effectively, this boson star solution would affect the history of the universe.

Although our analysis is restricted in the 4-dimensional spacetime, it can be extended to the higher dimensional case. In particular we are interested in the 5-dimensional brane world scenario, which consists of 4-dimensional thin/thick wall where we live and the bulk spacetime with negative cosmological constant. The static black hole solution has not been obtained yet in this context because the brane world have nontrivial S^1/Z_2 symmetry. However, our analysis shows the possibility that the scalar field can be a new primary hair of the black hole on brane world. The black hole in asymptotically AdS spacetime has more variety than the asymptotically flat case.

- Acknowledgments -

This work was supported partially by the Grant-in-Aid for Scientific Research Fund of the Ministry of Education, Science and Culture (No. 10640286), and by the Grant-in-Aid for JSPS (No. 199906147).

References

[1] R. Ruffini and J. A. Wheeler, Phys. Today **B24** 30 (1971).

[2] W. Israel, Phys. Rev. **164**, 1776 (1967); Commun. Math. Phys. **8**, 245 (1971); B. Carter, Phys. Rev. Lett. **26**, 331 (1971); R. M. Wald, Phys. Rev. Lett. **26**, 1653 (1971); D. C. Robinson, Phys. Rev. Lett. **34**, 905 (1977); P. O. Mazur, J. Phys. A **15**, 3173 (1982); Phys. Lett. A **100**, 341 (1984).

[3] J. E. Chase, Commun. Math. Phys. **19**, 276 (1970); J. D. Bekenstein, Phys. Rev. **D5**, 1239 (1972); J. B. Hartle, in *Magic without Magic*, edited by J. Klauder (Freeman, San Francisco, 1972); C. Teitelboim, Phys. Rev. **D5**, 2941 (1972).

[4] M. S. Volkov and D. V. Gal'tsov, Yad. Fiz. **51**, 1171 (1990) [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).

[5] N. Straumann and Z.-H. Zhou, Phys. Lett. B **243**, 33 (1990); M. S. Volkov, et al., Phys. Lett. B **349**, 438 (1995).

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

[7] T. Torii, K. Maeda and M. Narita, Phys. Rev. **D59**, 064027 (1999).

[8] J. D. Bekenstein, Phys. Rev. **D51**, R6608 (1995); D. Sudarsky, Class. Quantum Grav. **12**, 579 (1995).

[9] T. Torii, K. Maeda and M. Narita, Phys. Rev. **D59**, 104002 (1999).

[10] J. Maldacena, Adv. Theor. Math. Phys. **2**, 231 (1998); E. Witten, *ibid.* **2**, 253 (1998); **2**, 505 (1998).

[11] L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999); *ibid* **83**, 4690 (1999).

[12] E. Winstanly, Class. Quantum. Grav. **16**, 1963 (1999).

[13] J. Bjoraker and Y. Hosotani, Phys. Rev. Lett. **84**, 1852 (2000); Phys. Rev. D **62**, 043513 (2000).

[14] T. Torii, K. Maeda and M. Narita, to be appeard in Phys. Rev. D.

[15] G. W. Gibbons and K. Maeda, Nucl. Phys. **B298**, 41 (1988); D. Garfinkle, G. T. Horowitz and A. Strominger, Phys. Rev. D **43**, 3140 (1991).

Late-time behaviors of a massive scalar field in Reissner-Nordström background

Hiroko Koyama,[1] Akira Tomimatsu[2]

Department of Physics, Nagoya University, Nagoya, 464-8602, Japan

Abstract

Late-time behaviors of massive scalar fields in Reissner-Nordström background are investigated. It is suggested that massive scalar fields are dominated by tail behaviors whose decay rate is $t^{-5/6}$ at asymptotically late times. We discuss relations between field mass and the time scale when $t^{-5/6}$ tails begin to dominate.

1 Introduction

Various interactions of black holes with scalar fields have been extensively studied for a long time. In particular, it has been found that massive scalar fields can cause phenomena qualitatively different from massless fields. The interesting example is the vacuum polarization $\langle\phi^2\rangle$ of a quantum massive scalar field ϕ, which is in thermal equilibrium with a nearly extreme Reissner-Nordström black hole [1]. Because the black-hole temperature is very low, the mass-induced excitation of $\langle\phi^2\rangle$ results in the resonant amplification at $mM \simeq O(1)$ for the field mass m and the black-hole mass M.

Such a resonance behavior due to the mass of a field interacting with a black hole may appear in various processes as a basic feature of black-hole geometry. For a step to support this conjecture, in this paper, we turn our attention to a problem on time evolution of classical fields.

The late-time evolution of classical fields outside a collapsing star has important implications for the mass-inflation scenario. The asymptotic late-time tails along the outer horizon of a rotating or a charged black hole are used as the initial input for the ingoing perturbations which penetrate into the black hole. These perturbations are the physical cause for the well-known phenomena of mass-inflation [2]. Recently, Hod and Piran [3] have pointed out that the late-time tails of massive scalar fields are quite different from massless fields. If the field mass is small, namely $mM \ll 1$, the oscillatory inverse power-law behavior

$$\psi \sim t^{-l-\frac{3}{2}} \sin(mt), \tag{1}$$

dominates as the intermediate late-time tails. Note that massless fields decay more rapidly without any oscillation, as was first studied by [4]. Though the analytical approximation (1) was numerically verified at intermediate late times, Hod and Piran [3] also suggested that another wave pattern dominates at very late times, namely the intermediate tail (1) is not a final asymptotic behavior.

Our purpose of this paper is to clarify what kind of mass-induced behaviors dominates in the asymptotic late-time tails as a result of interaction of massive scalar fields with a black hole. In Sec. 2 we introduce the black-hole Green's function using the spectral decomposition method [5]. In Sec. 3 we consider a nearly extreme Reissner-Nordström background, motivated by the fact that the resonance phenomena in vacuum polarization of quantum scalar fields become clear. Then, based on the procedure of asymptotic matching, we construct approximate solutions in the nearly extreme limit. In Sec. 4 we study the intermediate tails which appear in the case of small mass field, confirming our result in comparison with (1). In Sec. 5 we derive the asymptotic late-time tails which is the main result in this paper. The final section is devoted to summary and discussion.

[1] E-mail:hiroko@allegro.phys.nagoya-u.ac.jp
[2] E-mail:atomi@allegro.phys.nagoya-u.ac.jp

2 The black-hole Green's function

We consider time evolution of a massive scalar field in Reissner-Nordström background with mass M and charge Q. The metric is

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right) dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{-1} dr^2 + r^2 d\Omega^2, \tag{2}$$

and the scalar field ϕ with mass m satisfies the wave equation

$$\left(\Box - m^2\right) \phi = 0. \tag{3}$$

Resolving the field into spherical harmonics

$$\phi = \sum_{l,m} \frac{\psi_l(t,r)}{r} Y_l^m(\theta,\varphi), \tag{4}$$

we obtain a wave equation for each multiple moment;

$$\psi_{,tt} - \psi_{,r_* r_*} + V\psi = 0, \tag{5}$$

where r_* is the tortoise coordinate defined by

$$dr_* = \frac{dr}{1 - \frac{2M}{r} + \frac{Q^2}{r^2}}, \tag{6}$$

and the potential V is

$$V = \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)\left[\frac{l(l+1)}{r^2} + \frac{2M}{r^3} - \frac{2Q^2}{r^4}\right]. \tag{7}$$

The time evolution of ψ is given by

$$\psi(r_*,t) = \int \left[G(r_*,r'_*;t)\psi_t(r',0) + G_t(r_*,r'_*;t)\psi(r'_*,0)\right] dr'_* \tag{8}$$

for $t > 0$, where $G(r,r';t)$ is the (retarded) Green's function satisfying

$$\left[\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial r_*^2} + V\right] G(r_*,r'_*;t) = \delta(t)\delta(r_* - r'_*) \tag{9}$$

with the initial condition $G(r,r';t) = 0$ for $t \leq 0$. We calculate the Green's function through the Fourier transform;

$$G(r'_*,r_*;t) = -\frac{1}{2\pi}\int_{-\infty+ic}^{\infty+ic} \tilde{G}(r_*,r'_*;\omega) e^{-i\omega t} d\omega \tag{10}$$

where c is some positive constant. The usual procedure is to close the contour of integration into the lower half of the complex frequency plane shown in Fig. 1. Then, the late-time tail behaviors which are our main concern should be given by the integral along the branch cut placed along the interval $-m \leq \omega \leq m$.

The Fourier component $\tilde{G}(r_*,r'_*;\omega)$ in the range $-m \leq \omega \leq m$ can be expressed in terms of two linearly independent solutions $\tilde{\psi}_1$ and $\tilde{\psi}_2$ for the homogeneous equation;

$$\left(\frac{d^2}{dr_*^2} + \omega^2 - V\right)\tilde{\psi}_i = 0, \quad i = 1,2. \tag{11}$$

The boundary condition for the basic solution $\tilde{\psi}_1$ is to describe purely ingoing waves crossing the event horizon, i.e.,

$$\tilde{\psi}_1 \simeq e^{-i\omega r_*}, \tag{12}$$

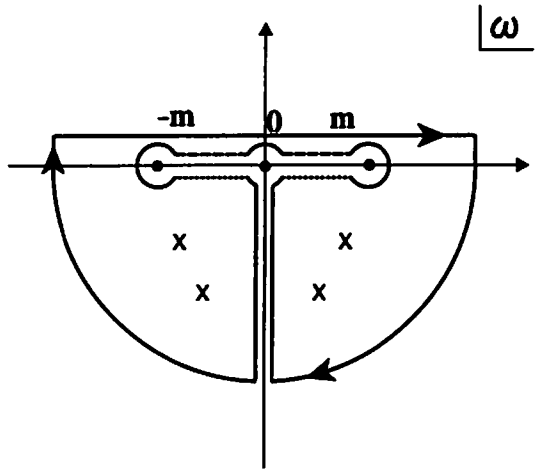


Figure 1: Integration contours in the complex frequency plane. The original integration contour for the Green's function lies above the real frequency axis. We choose the value of ϖ on the dashed line to be $\varpi = |\varpi|$ and that on the dotted lines to be $\varpi = e^{\pm i\pi}|\varpi|$. The poles in $G(r_*, r_{*'}; \omega)$ are also shown, which give the quasinormal modes.

as $r_* \to -\infty$, while $\tilde{\psi}_2$ is required to damp exponentially at spatial infinity, i.e.,

$$\tilde{\psi}_2 \simeq e^{-\varpi r_*}, \tag{13}$$

as $r_* \to \infty$, where $\varpi \equiv \sqrt{m^2 - \omega^2}$. Because the complex conjugate $\tilde{\psi}_1^*$ is also a solution for (11), $\tilde{\psi}_2$ can be written by the linear superposition

$$\tilde{\psi}_2 = \alpha\tilde{\psi}_1 + \beta\tilde{\psi}_1^*, \tag{14}$$

and the Wronskian is estimated to be

$$W(\omega) = \tilde{\psi}_1\tilde{\psi}_{2,r_*} - \tilde{\psi}_{1,r_*}\tilde{\psi}_2 = 2i\omega\beta. \tag{15}$$

Using these two solutions, the Green's function can be written by

$$\tilde{G}(r_*, r'_*; \omega) = -\frac{1}{2i\omega\beta}\begin{cases} \tilde{\psi}_1(r'_*, \omega)\tilde{\psi}_2(r_*, \omega), & r' > r, \\ \tilde{\psi}_1(r_*, \omega)\tilde{\psi}_2(r'_*, \omega), & r' < r. \end{cases} \tag{16}$$

The contribution G^C from the branch cut to the Green's function is reduced to

$$G^C(r_*, r'_*; t) = -\frac{1}{4\pi i}\int_{\rm cut}\frac{1}{\omega}\frac{\alpha}{\beta}\tilde{\psi}_1(r'_*, \omega)\tilde{\psi}_1(r_*, \omega)e^{-i\omega t}d\omega. \tag{17}$$

Then the main task to evaluate G^C is to derive the coefficients α and β.

3 Extreme-limit approximation

It is difficult in general to obtain the coefficients α and β, since exact solutions for the wave equation (11) cannot be expressed by any elementary function or any trancedental function already known. Fortunately, in nearly extreme Reissner-Nordström geometry, the procedure of asymptotic matching turns out to be very useful. Let us change the variable r to z defined as

$$\frac{z-1}{2} = \frac{r - r_+}{2\kappa r_+^2}, \tag{18}$$

where r_+ and r_- are the outer and inner horizon radii, respectively, and κ is the surface gravity defined as $\kappa \equiv (r_+ - r_-)/2r_+^2$. Then we can rewrite the wave equation (11) into

$$(z^2-1)\frac{d^2\tilde{\psi}}{dz^2} + \frac{\{2\kappa r_+^2 + z(r_+ + r_-)\}}{r_+\{(z-1)\kappa r_+ + 1\}}\frac{d\tilde{\psi}}{dz} + \left[\frac{\omega^2}{\kappa^2}\frac{\{(z-1)\kappa r_+ + 1\}^4}{z^2-1} - l(l+1)\right.$$
$$\left. -m^2 r_+^2\{(z-1)\kappa r_+ + 1\}^2 + \frac{\kappa r_+^2\{2\kappa r_+^2 + z(r_+ + r_-)\}}{r_+^2\{(z-1)\kappa r_+ + 1\}^2}\right]\tilde{\psi} = 0. \qquad (19)$$

For the nearly extreme case such that

$$\kappa M \ll 1, \qquad (20)$$

we can derive the approximate solutions valid in the region $z \gg 1$ or $z \ll 1/\kappa M$ in order to determine coefficients α and β in (17). Our strategy will be formulated in [7] in detail. As a result we obtain

$$\alpha(|\omega|,\varpi) = \beta(e^{i\pi}|\omega|,\varpi)$$
$$= \left[\frac{\Gamma(2\mu)\Gamma(2\mu+1)(2\varpi M)^{-\mu+\frac{1}{2}}e^{i\frac{\pi}{2}(-\mu-\frac{1}{2})}}{\Gamma(\frac{1}{2}+\mu-\sigma_2)(2|\omega|M)^{\mu+\frac{1}{2}}\Gamma(\frac{1}{2}+\mu+2i|\omega|M)}\right.$$
$$\left. + \frac{\Gamma(-2\mu)\Gamma(-2\mu+1)(2\varpi M)^{\mu+\frac{1}{2}}e^{i\frac{\pi}{2}(\mu-\frac{1}{2})}}{\Gamma(\frac{1}{2}-\mu-\sigma_2)(2|\omega|M)^{-\mu+\frac{1}{2}}\Gamma(\frac{1}{2}-\mu+2i|\omega|M)}\right]\left(\frac{\omega}{\kappa}\right)^{2i|\omega|M} e^{-\pi|\omega|M}, \qquad (21)$$

and

$$\beta(|\omega|,\varpi) = \alpha(e^{i\pi}|\omega|,\varpi)$$
$$= \left[\frac{\Gamma(2\mu)\Gamma(2\mu+1)(2\varpi M)^{-\mu+\frac{1}{2}}e^{i\frac{\pi}{2}(\mu+\frac{1}{2})}}{\Gamma(\frac{1}{2}+\mu-\sigma_2)(2|\omega|M)^{\mu+\frac{1}{2}}\Gamma(\frac{1}{2}+\mu-2i|\omega|M)}\right.$$
$$\left. + \frac{\Gamma(-2\mu)\Gamma(-2\mu+1)(2\varpi M)^{\mu+\frac{1}{2}}e^{i\frac{\pi}{2}(-\mu+\frac{1}{2})}}{\Gamma(\frac{1}{2}-\mu-\sigma_2)(2|\omega|M)^{-\mu+\frac{1}{2}}\Gamma(\frac{1}{2}-\mu-2i|\omega|M)}\right]\left(\frac{\omega}{\kappa}\right)^{-2i|\omega|M} e^{-\pi|\omega|M}, \qquad (22)$$

where

$$\mu \equiv \sqrt{\left(l+\frac{1}{2}\right)^2 + m^2M^2 - 6\omega^2M^2} \qquad (23)$$

and

$$\sigma_2 = -2\varpi M + \frac{m^2M}{\varpi}. \qquad (24)$$

4 Intermediate late-time tails

We consider the late-time behavior of G^C at the timescale

$$mt \gg 1, \qquad (25)$$

when the decaying tail will dominate. Hod and Piran [3] pointed out that for the scalar field with small mass, namely $mM \ll 1$, the dominant behavior is given by Eq.(1) at the intermediate late times in the range

$$mM \ll mt \ll \frac{1}{(mM)^2} \qquad (26)$$

In this section we check the validity of Eqs. (21) and (22), by deriving the intermediate tail behavior. Following Hod and Piran [3], the effective contribution to the integral in Eq. (17) is claimed to be limited to the range $|\omega - m| = O(1/t)$ or equivalently $\varpi = O(\sqrt{m/t})$. This is due to the rapidly oscillating term $e^{-i\omega t}$ which leads to a mutual cancellation between the positive and the negative parts of the integrand.

Then, in the time scale given by Eq. (26) we note that the frequency range $\varpi = O(\sqrt{m/t})$ leads to the inequality

$$\sigma_2 \ll 1, \tag{27}$$

which allows us to obtain the approximations of α and β as follows,

$$\begin{aligned}\alpha(|\omega|,\varpi) \sim & \left[\frac{\Gamma(2l+1+2\epsilon_\mu)\Gamma(2l+2+2\epsilon_\mu)}{\Gamma(l+1+\epsilon_\mu-\sigma_2)\Gamma(l+1+\epsilon_\mu+2i|\omega|M)}(2\varpi M)^{-l-\epsilon_\mu}(2|\omega|M)^{-l-1-\epsilon_\mu}e^{i\frac{\pi}{2}(-l-1-\epsilon_\mu)}\right. \\ & \left.+\frac{\Gamma(-2l-1-2\epsilon_\mu)\Gamma(-2l-2\epsilon_\mu)}{\Gamma(-l-\epsilon_\mu-\sigma_2)\Gamma(-l-\epsilon_\mu+2i|\omega|M)}(2\varpi M)^{l+1+\epsilon_\mu}(2|\omega|M)^{l+\epsilon_\mu}e^{i\frac{\pi}{2}(l+\epsilon_\mu)}\right] \\ & \times\left(\frac{|\omega|}{\kappa}\right)^{2i|\omega|M}e^{-|\omega|M},\end{aligned} \tag{28}$$

and

$$\begin{aligned}\beta(|\omega|,\varpi) \sim & \left[\frac{\Gamma(2l+1+2\epsilon_\mu)\Gamma(2l+2+2\epsilon_\mu)}{\Gamma(l+1+\epsilon_\mu-\sigma_2)\Gamma(l+1+\epsilon_\mu-2i|\omega|M)}(2\varpi M)^{-l-\epsilon_\mu}(2|\omega|M)^{-l-1-\epsilon_\mu}e^{i\frac{\pi}{2}(l+1+\epsilon_\mu)}\right. \\ & \left.+\frac{\Gamma(-2l-1-2\epsilon_\mu)\Gamma(-2l-2\epsilon_\mu)}{\Gamma(-l-\epsilon_\mu-\sigma_2)\Gamma(-l-\epsilon_\mu-2i|\omega|M)}(2\varpi M)^{l+1+\epsilon_\mu}(2|\omega|M)^{l+\epsilon_\mu}e^{i\frac{\pi}{2}(-l-\epsilon_\mu)}\right] \\ & \times\left(\frac{|\omega|}{\kappa}\right)^{-2i|\omega|M}e^{-|\omega|M},\end{aligned} \tag{29}$$

where

$$\epsilon_\mu \equiv \mu - \left(l+\frac{1}{2}\right) \ll 1. \tag{30}$$

Expanding the ratio α/β as a power series in mM, we can approximate it as follows;

$$\frac{\alpha(|\omega|,\varpi)}{\beta(|\omega|,\varpi)} - \frac{\alpha(|\omega|,e^{i\pi}\varpi)}{\beta(|\omega|,e^{i\pi}\varpi)} \sim \frac{l!^4}{(2l)!^2(2l+1)!^2}(2M)^{4l+2}|\omega|^{2l+1}\varpi^{2l+1}2i. \tag{31}$$

Substituting Eq. (31) into Eq. (17), we obtain

$$\begin{aligned}G^C(r_*,r'_*;\omega) = & \frac{l!^4}{(2l)!^2(2l+1)!^2}(2M)^{4l+2}\int_0^m \omega^{2l+1}\varpi^{2l+1}\tilde{\psi}_1(r_*,\omega)\tilde{\psi}_1(r'_*,\omega)e^{-i\omega t}d\omega \\ & +(\text{complex conjugate}),\end{aligned} \tag{32}$$

which is similar to Eq. (29) in [3], giving the damping exponent in Eq. (1).

Different from [3], our analytical calculation is not based on the flat space approximation. The intermediate tail dominates in the range (26), when the integrand can be approximated by Eq. (27). It is easy to find that the larger the field's mass is, the sooner it leaves the intermediate tails, and the phase does not appear in the case of $mM \gtrsim 1$.

5 Asymptotic late-time tails

It is obvious from our calculation that the intermediate tail is not a final pattern of decay but should be replaced by another one, which was also numerically suggested in [3]. Physically the change of the tail behavior will be a result of dominant backscattering by spacetime curvature, which is the effect beyond the flat space approximation. What kind of wave pattern dominates at very late times? In addition, we must reveal late-time tails in the $mM \gtrsim 1$ case of large field mass, for which the intermediate tails do not appear. In this section we study a tail behavior dominant at asymptotic late times,

$$mt \gg \frac{1}{m^2M^2} \tag{33}$$

when for the parameter σ_2 we obtain

$$\sigma_2 \gg 1, \tag{34}$$

different from Eq. (27) at intermediate late times. Then, the coefficients $\alpha(\omega, e^{\pm i\pi}\varpi)$ and $\beta(\omega, e^{\pm i\pi}\varpi)$ are approximated by

$$\begin{aligned}
\alpha(|\omega|, e^{-i\pi}\varpi) &= \beta(e^{i\pi}|\omega|, e^{i\pi}\varpi) \\
&\simeq \frac{1}{\sqrt{2\pi}} e^{\frac{m^2M}{\varpi}} (2m^2M^2)^{-\frac{m^2M}{\varpi}+2\varpi M} (2\varpi M)^{-2\varpi M+\frac{m^2M}{\varpi}+\frac{1}{2}} \\
&\quad \left[\frac{\Gamma(2\mu)\Gamma(2\mu+1)(2m^2M^2)^{\mu}(2|\omega|M)^{-\mu-\frac{1}{2}} e^{i\frac{\pi}{2}(\mu-\frac{3}{2})}}{\Gamma(\frac{1}{2}+\mu+2i|\omega|M)} \right. \\
&\quad \left. + \frac{\Gamma(-2\mu)\Gamma(-2\mu+1)(2m^2M^2)^{-\mu}(2|\omega|M)^{\mu-\frac{1}{2}} e^{i\frac{\pi}{2}(-\mu-\frac{3}{2})}}{\Gamma(\frac{1}{2}-\mu+2i|\omega|M)} \right] \left(\frac{|\omega|}{\kappa} \right)^{2i|\omega|M} e^{-\pi|\omega|M}
\end{aligned} \tag{35}$$

and

$$\begin{aligned}
\beta(|\omega|, e^{-i\pi}\varpi) &= \alpha(e^{i\pi}|\omega|, e^{i\pi}\varpi) \\
&\simeq \frac{1}{\sqrt{2\pi}} e^{\frac{m^2M}{\varpi}} (2m^2M^2)^{-\frac{m^2M}{\varpi}+2\varpi M} (2\varpi M)^{-2\varpi M+\frac{m^2M}{\varpi}+\frac{1}{2}} \\
&\quad \left[\frac{\Gamma(2\mu)\Gamma(2\mu+1)(2m^2M^2)^{\mu}(2|\omega|M)^{-\mu-\frac{1}{2}} e^{i\frac{\pi}{2}(3\mu-\frac{1}{2})}}{\Gamma(\frac{1}{2}+\mu-2i|\omega|M)} \right. \\
&\quad \left. + \frac{\Gamma(-2\mu)\Gamma(-2\mu+1)(2m^2M^2)^{-\mu}(2|\omega|M)^{\mu-\frac{1}{2}} e^{i\frac{\pi}{2}(-3\mu-\frac{1}{2})}}{\Gamma(\frac{1}{2}-\mu-2i|\omega|M)} \right] \left(\frac{|\omega|}{\kappa} \right)^{-2i|\omega|M} e^{-\pi|\omega|M}
\end{aligned} \tag{36}$$

because of Eq. (34). Thus the contribution from this part corresponding to the dotted lines in Fig. 1 to the Green's function given by

$$\frac{1}{4\pi i} \int_{\text{dotted lines}} \frac{1}{\omega} \frac{\alpha(\omega, \varpi)}{\beta(\omega, \varpi)} \tilde{\psi}_1(r_*, \omega) \tilde{\psi}_1(r'_*, \omega) e^{-i\omega t} d\omega \tag{37}$$

is $O(t^{-1})$ at most, because $\alpha(\omega, e^{\pm i\pi}\varpi)$ and $\beta(\omega, e^{\pm i\pi}\varpi)$ converge in the limit $|\omega| \to m$. On the other hand, $\alpha(\omega, \varpi)$ and $\beta(\omega, \varpi)$ are reduced to

$$\frac{\alpha(|\omega|, \varpi)}{\beta(|\omega|, \varpi)} = \frac{\beta(e^{i\pi}|\omega|, \varpi)}{\alpha(e^{i\pi}|\omega|, \varpi)} \simeq \frac{\gamma^* e^{-i\pi\sigma_2} + \eta^* e^{i\pi\sigma_2}}{\gamma e^{i\pi\sigma_2} + \eta e^{-i\pi\sigma_2}}, \tag{38}$$

for the limit (34), where

$$\begin{aligned}
\gamma &= \frac{\Gamma(-2\mu)\Gamma(-2\mu+1)(2m^2M^2)^{-\mu}(2\omega M)^{\mu-\frac{1}{2}}}{\Gamma(\frac{1}{2}-\mu-2i|\omega|M)} e^{i\frac{\pi}{2}(\mu+\frac{1}{2})} \\
&\quad + \frac{\Gamma(2\mu)\Gamma(2\mu+1)(2m^2M^2)^{\mu}(2|\omega|M)^{-\mu-\frac{1}{2}}}{\Gamma(\frac{1}{2}+\mu-2i|\omega|M)} e^{i\frac{\pi}{2}(-\mu+\frac{1}{2})},
\end{aligned} \tag{39}$$

and

$$\begin{aligned}
\eta &= \frac{\Gamma(-2\mu)\Gamma(-2\mu+1)(2m^2M^2)^{-\mu}(2|\omega|M)^{\mu-\frac{1}{2}}}{\Gamma(\frac{1}{2}-\mu-2i|\omega|M)} e^{i\frac{\pi}{2}(-3\mu+\frac{1}{2})} \\
&\quad + \frac{\Gamma(2\mu)\Gamma(2\mu+1)(2m^2M^2)^{\mu}(2|\omega|M)^{-\mu-\frac{1}{2}}}{\Gamma(\frac{1}{2}+\mu-2i|\omega|M)} e^{i\frac{\pi}{2}(3\mu+\frac{1}{2})}.
\end{aligned} \tag{40}$$

Therefore we find the contribution from this part corresponding to the dashed line in Fig. 1 to the Green's function to be approximated by

$$\begin{aligned}
&\frac{1}{4\pi i} \int_0^m \frac{1}{\omega} \frac{\alpha(\omega, \varpi)}{\beta(\omega, \varpi)} e^{-i\omega t} \tilde{\psi}_1(r_*, \omega) \tilde{\psi}_1(r'_*, \omega) d\omega + (\text{complex conjugate}) \\
&\simeq \frac{1}{4\pi m i} \tilde{\psi}_1(r_*, m) \tilde{\psi}_1(r'_*, m) \int_0^m e^{i(2\pi\sigma_2 - \omega t)} e^{i\varphi} d\omega + (\text{complex conjugate}),
\end{aligned} \tag{41}$$

where the phase φ is defined by

$$e^{i\varphi} = \frac{\eta^* + \gamma^* e^{-2i\pi\sigma_2}}{\eta + \gamma e^{2i\pi\sigma_2}} \tag{42}$$

and it remains in the range $0 \leq \varphi \leq 2\pi$, even if σ_2 becomes very large, since we have

$$|\eta|^2 - |\gamma|^2 = \frac{\pi}{|\omega|M} e^{2\pi|\omega|M} > 0. \tag{43}$$

At late times when both (25) and (34) are satisfied, both terms of $e^{i\omega t}$ and $e^{2i\pi\sigma_2}$ are rapidly oscillating. In physical meaning, scalar waves are mixed states with multiple phases backscattered by spacetime curvature. Most of these wave are canceled out each other by those of the inverse phase. As a result, particular waves of which frequency satisfies

$$\frac{d}{d\omega}(2\pi\sigma_2 - \omega t) = 0, \tag{44}$$

remain without cancellation. Let us evaluate the dominant contribution from such waves to the late-time tails. We can find a solution for Eq. (44)

$$\varpi_0 \equiv \sqrt{m^2 - \omega_0^2} \simeq \left(\frac{2\pi m^3 M}{t}\right)^{\frac{1}{3}} \tag{45}$$

in the limit $\omega_0 \simeq m$. In order for the saddle point (45) to exist in the region (34), where $e^{2i\pi\sigma_2}$ are rapidly oscillating as a function of ω, we need the additional relation

$$mt \gg mM. \tag{46}$$

By the method of steepest descent, Eq. (41) can be approximated by

$$\frac{1}{4\pi i m}\int e^{i\frac{d^2}{d\omega^2}(2\pi\sigma_2 - \omega t)\big|_{\omega=\omega_0}(\omega-\omega_0)^2} e^{i\varphi_0} d\omega \sim \frac{i}{4\sqrt{3}}(2\pi)^{\frac{2}{3}}(mM)^{\frac{1}{3}}(mt)^{-\frac{2}{3}} e^{imt} e^{i\varphi_0}, \tag{47}$$

through the formula

$$\int_{-\infty}^{\infty} \cos x^2 dx = \int_{-\infty}^{\infty} \sin x^2 dx = \sqrt{\frac{\pi}{2}}. \tag{48}$$

Taking the decay rate into account, we can neglect the contribution from (37) to G^C in comparison with that from (47). Therefore tails such as (47) dominate at asymptotic late times. Finally we arrive asymptotic late-time tails as

$$G^C(r_*, r'_*; t) \simeq \frac{1}{2\sqrt{3}}(2\pi)^{\frac{2}{3}}(mM)^{\frac{1}{3}}(mt)^{-\frac{2}{3}} \sin(mt)\tilde{\psi}_1(r_*, m)\tilde{\psi}_1(r'_*, m). \tag{49}$$

It is found that the decay rate is independent of the location.

6 Summary

In this paper we have investigated mass-induced behaviors which appear in late-time tails of classical massive scalar fields in nearly extreme Reissner-Nordström background. If the field mass is small, namely $mM \ll 1$, the intermediate tails given by Eq. (1) have been shown to dominate at the intermediate late-time $mM \ll mt \ll 1/(mM)^2$, consistently with the result in [3]. Our main result is the asymptotic tails with the decay rate of $t^{-5/6}$, which is interestingly independent of the field mass m and the angular momentum parameter l.

We can discuss the timescale when the $t^{-5/6}$ tail dominates. The basic condition for the tail dominance is given by Eq. (25). However, for $mM \ll 1$, the $t^{5/6}$ tail requires the additional condition (34), which gives $mt \gg 1/m^2M^2$. On the other hand, for large field mass, $mM \gg 1$, the larger mM is, the later the tail begins to dominate, because the tail requires the further condition $\varpi_0/m \ll 1$ for Eq. (45) in

addition to the conditions (25) and (34), which gives $mt \gg mM$. Therefore the timescale when the $t^{-5/6}$ tail dominates will become minimum at $mM \simeq O(1)$, which means that the most effective backscattering occurs for such scalar fields at late times.

We note that the timescale when the $t^{-5/6}$ tail dominates is determined by the black-hole radius M. It is conjectured that the slow decay of $t^{-5/6}$ originates from the existence of resonant enhancement of massive scalar waves with the frequency $\omega \simeq m$ near the horizon. In this sense we claim that resonance behaviors are not peculiar to quantum scalar fields but manifest also in classical scalar fields.

In this paper we have calculated the tail behaviors in nearly extremal limit, because we are motivated by the previous work [1]. The extension of our calculation into non-extremal case remains in future works, to investigate whether our result in this paper is a special feature of the extremal case or not.

acknowledgments

The authors thank M. Sakagami, A. Ohashi and colleagues of CG laboratory in Nagoya University for useful discussions and comments. H.K. also thanks A. Yamada for drawing a figure in this paper.

References

[1] A. Tomimatsu and H. Koyama, Phys. Rev. D **61**, 124010 (2000).

[2] E. Poisson and W. Israel, Phys. Rev. D **41**, 1796 (1990).

[3] S. Hod and T. Piran, Phys. Rev. D **58**, 044018 (1998).

[4] R. H. Price, Phys. Rev. D **5**, 2419 (1972).

[5] E. W. Leaver, Phys. Rev. D **34**, 384 (1986).

[6] *Handbook of Mathematical Functions*, editted by M. Abramowitz and I.A. Stegun (Dover, New York, 1970).

[7] H. Koyama and A. Tomimatsu, in preparation

Volume Criticality and Cosmological Perturbations

Katsuhito Yasuno[1]

Department of Physics, Tokyo Institute of Technology,
Oh-Okayama Meguro-ku, Tokyo 152-0033, Japan

Masayuki Tanimoto[2]

Yukawa Institute for Theoretical Physics, Kyoto University,
Sakyo-ku, Kyoto 606-8502, Japan

Abstract

We examine that the criticality of the volume functional at a Friedmann-Lemaître-Robertson-Walker solution does not depend on the choice of gauge by introducing the spherical harmonics often used in the cosmological perturbation theory. In this approach, we find that the criticality is understood from the fact that there is no mixing of modes in the linearized Einstein equations.

1 Introduction

The volume criticality is a property pertaining to a Friedmann-Lemaître-Robertson-Walker (FLRW) solution. Roughly speaking, it says [1] that variations of a volume functional at an FLRW solution in the space of solutions give the *zero* function of time; We assume that every solution is equipped with a time slicing and time coordinate, and the volume functional is accordingly defined so as to give a function of time giving the spatial total volume. The spatial manifold is assumed to be compact without boundary, i.e., closed.

In [1], however, the consideration was restricted to the case where variations are evaluated in the synchronous and comoving gauge, and the matter content was assumed to be dust. Accordingly, gauge dependence and matter dependence was unclear. In this article, we shall consider these issues in the framework of cosmological perturbation theory.

In the next section, we show that the criticality of the volume functional does not depend on choices of gauge and matter content by introducing the spherical harmonics [2, 3]. In the final section, we summarize the result and discuss possible applications.

Throughout this article, early Roman indices $a, b, \cdots$ are abstract indices. Middle Roman indices $i, j, \cdots$ label spatial components of tensors and Greek indices $\mu, \nu \cdots$ label spacetime components of tensors.

2 Derivation of the criticality of volume functional

In this section, we show that an FLRW solution is a critical point with respect to the volume functional in the space of solutions of Einstein's equation.

Let $({}^4M, g_{ab})$ be a spatially closed spacetime satisfying the Einstein equation,

$$ {}^4G_{ab} = \kappa T_{ab}, \tag{1}$$

where ${}^4G_{ab}$ is the 4-dimensional Einstein tensor of g_{ab}, T_{ab} is the stress-energy tensor and $\kappa = 8\pi G$. Let (M_t, q_{ab}) be a closed spacelike hypersurface of $({}^4M, g_{ab})$ labelled by t.

[1]E-mail:kyasuno@th.phys.titech.ac.jp
[2]E-mail:tanimoto@yukawa.kyoto-u.ac.jp

The *volume functional*, of which value is a function of t, is defined by

$$V[q](t) := \int_{M_t} \sqrt{\det q_{ij}}\, d^3x, \tag{2}$$

where $q_{ij} = q_{ab}e^a_i e^b_j$ denote components of the induced metric with respect to a given frame $\{e^a_i\}$ on M_t. The first-order perturbation of $V[q]$ evaluated at the background metric $q^{(\mathrm{B})}_{ab}$ is given by

$$\delta V[q] = \frac{1}{2}\int_{M_t} q^{ij}_{(\mathrm{B})}\delta q_{ij}\sqrt{\det q^{(\mathrm{B})}_{ij}}\, d^3x. \tag{3}$$

We first show that *the first-order perturbations of the volume functional evaluated at the FLRW solution are represented by the (conformal) scalar zero-mode perturbations.* The Robertson-Walker metric is given by

$$g^{(\mathrm{B})}_{\mu\nu}dx^\mu dx^\nu = -d\tau^2 + q^{(\mathrm{B})}_{ij}dx^i dx^j = a^2(-d\eta^2 + \tilde{q}^{(\mathrm{B})}_{ij}dx^i dx^j), \tag{4}$$

where η is the conformal time defined by $d\eta = d\tau/a(\tau)$, $a(\eta)$ is the scale factor with respect to η and $\tilde{q}^{(\mathrm{B})}_{ab}$ is the metric of a 3-dimensional constant curvature space whose Riemann tensor is given by

$$\tilde{R}^{(\mathrm{B})}_{abcd} = \mathrm{K}(\tilde{q}^{(\mathrm{B})}_{ac}\tilde{q}^{(\mathrm{B})}_{bd} - \tilde{q}^{(\mathrm{B})}_{ad}\tilde{q}^{(\mathrm{B})}_{bc}), \tag{5}$$

where K takes values $+1$, 0 and -1 for $\mathbb{S}^3$, $\mathbb{E}^3$ and $\mathbb{H}^3$, respectively. The spatial components of the first-order perturbations of $q^{(\mathrm{B})}_{ab}$ are given by

$$\delta q_{ij} = 2a^2(\eta)[C(\eta,x)\tilde{q}^{(\mathrm{B})}_{ij} + E_{ij}(\eta,x)], \tag{6}$$

where E_{ij} is trace-free. Inserting these into (3), we have

$$\delta V[q] = 3a^3(\eta)\int_{M_\eta} C(\eta,x)\sqrt{\det \tilde{q}^{(\mathrm{B})}_{ij}}\, d^3x. \tag{7}$$

Thus, only the conformal scalar part of the perturbations appears in $\delta V[q]$.

Let us then introduce the generalized scalar spherical harmonics [2], which is a function on a 3-dimensional constant curvature space satisfying the generalized Helmholtz equation:

$$(\Delta + k^2_{(n)})Q_{(n)} = 0, \tag{8}$$

where Δ is the Laplace-Beltorami operator defined on the constant curvature space. Using these eigenfunctions, we can expand the scalar variable $C(\eta,x)$ such as

$$C(\eta,x) = C_{(n)}(\eta)Q_{(n)}(x), \tag{9}$$

where the summation taken over allowed all modes is understood. Then (7) becomes

$$\delta V[q] = 3a^3(\eta)C_{(n)}(\eta)\int_{M_\eta} Q_{(n)}(x)\sqrt{\det \tilde{q}^{(\mathrm{B})}_{ij}}\, d^3x =: 3a^3(\eta)C_{(n)}(\eta)\mathcal{Q}_{(n)}. \tag{10}$$

However, the integral, $\mathcal{Q}_{(n)}$, of the scalar spherical harmonics over M_η vanishes except for $k_{(n)} = 0$. This is a direct consequence of the orthogonality among the eigenfunctions $Q_{(n)}$,

$$\int_{M_\eta} Q_{(n)}Q_{(m)}\sqrt{\det \tilde{q}^{(\mathrm{B})}_{ij}}\, d^3x = \delta_{nm}, \tag{11}$$

and the fact that the zero-mode eigenfunction $Q_{(0)}$ (for $k_{(n)} = 0$) is a constant function. Thus, $\delta V[q]$ reduces to

$$\delta V[q] = 3a^3(\eta)C_{(0)}(\eta)\mathcal{Q}_{(0)}, \tag{12}$$

where $Q_{(0)}$ is a constant since $\tilde{q}^{(\mathrm{B})}_{ij}$ is independent of η.

Now, recall the well-known fact: the linearized Einstein equations around a FLRW solution can be decomposed into three parts of perturbations (namely, scalar, vector, and tensor part) and can be solved for each type and each mode independently [3]. In particular, the mode independence is important for us. Due to this, solutions for the function $C_{(0)}(\eta)$ contain finite number of integration constants. We see this explicitly and give implications below.

To this, we need the scalar zero-mode linearized Einstein equations. For zero-mode perturbations, the perturbed metric is given by

$$g_{(0)\mu\nu}dx^\mu dx^\nu = a^2(\eta)\big[-(1+2A_{(0)}Q_{(0)})d\eta^2 + (1+2C_{(0)}Q_{(0)})\tilde{q}_{ij}dx^i dx^j\big], \tag{13}$$

and the perturbed stress-energy tensor is given by

$$T_{(0)\mu\nu}dx^\mu dx^\nu = a^2(\eta)\big[\rho_{(\mathrm{B})}(1+\delta_{(0)}Q_{(0)}+2A_{(0)}Q_{(0)})d\eta^2 + p_{(\mathrm{B})}(1+\pi_{L(0)}Q_{(0)}+2C_{(0)}Q_{(0)})\tilde{q}^{(\mathrm{B})}_{ij}dx^i dx^j\big], \tag{14}$$

where $\rho_{(\mathrm{B})}$ and $p_{(\mathrm{B})}$ are the energy density and the pressure of the background FLRW spacetime. The zero-mode linearized Einstein equations are

$$6\big[-\mathrm{K}C_{(0)} + h(-hA_{(0)} + \partial_\eta C_{(0)})\big] = \kappa\rho_{(\mathrm{B})}a^2\delta_{(0)}, \tag{15}$$

$$\partial_\eta(-hA_{(0)} + \partial_\eta C_{(0)}) - \Big(\frac{a''}{a} - h^2\Big)A_{(0)} + h\partial_\eta C_{(0)} = -\frac{\kappa}{6}a^2(\rho_{(\mathrm{B})}\delta_{(0)} + 3p_{(\mathrm{B})}\pi_{L(0)}), \tag{16}$$

where a prime denotes ∂_η and $h = a'/a$. To solve these equations, we need to assume an equation of state of matter and impose a gauge condition. As for the equation of state, for simplicity, we take

$$p_{(\mathrm{B})} = 0, \tag{17}$$

i.e. the matter is dust. We also need to an equation for the background spacetime containing dust,

$$2\frac{a''}{a} - h^2 + \mathrm{K} = 0, \tag{18}$$

If we employ the synchronous gauge, we have $A_{(0)} = 0$. In this case we have

$$C''_{(0)} + 2hC'_{(0)} - \mathrm{K}C_{(0)} = 0. \tag{19}$$

Its solution, obtained by combining with (18), is

$$C_{(0)} = \frac{h}{a}\Big(c_1\int\frac{d\eta}{h^2} + c_2\Big), \tag{20}$$

where c_1 and c_2 are integration constants, yields exactly the same expression for $\delta V[q]$ as given in [1]. This means that the two independent modes are those caused by the variations taken in the set spanned by the two-parameter FLRW solution; One of the two parameters corresponds to rescalings of the scale factor, and the other to translations of the initial time (big-bang time).

Here we represent another interesting example. Let us assume $-hA_{(0)} + C'_{(0)} = 0$. This condition comes from the uniform-Hubble-slicing gauge condition [3]. In this case, we have

$$\Big(2h - \frac{a''}{a'}\Big)C'_{(0)} = \mathrm{K}C_{(0)}. \tag{21}$$

Its solution is given by

$$C_{(0)} = c_1\exp\Big[\mathrm{K}\int\Big(2h - \frac{a''}{a'}\Big)^{-1}d\eta\Big], \tag{22}$$

where c_1 is an integration constant. Since the equation (21) is a first order linear ODE, there arises only one mode with an integration constant. This mode is also that caused by the variations taken in the FLRW solution. (Be careful that the usual (one or two-parameter) FLRW solution is given so as to preserve the synchronous gauge, but not the uniform-Hubble-slicing gauge.)

Thus, we have found that $\delta V[q] = 0$ for all inhomogeneous perturbations that is "orthogonal" to the space of the FLRW solution, irrespective of the gauge choice. Every FLRW solution (with the parameters fixed) is therefore a critical point of the volume functional in the "orthogonally decomposed" space of solutions.

3 Summary and Discussions

We have seen that to determine the first order perturbations (or variations) of the volume functional around an FLRW solution, it is sufficient to see the behavior of the (conformal) scalar zero-mode perturbations, described by the variable $C_{(0)}(\eta)$. Since this variable is governed by a closed ODE, there are only a finite number of degrees of freedom in the general solution. It is therefore possible to find directions in the space of solutions such that $\delta V[q] = 0$. Moreover, as we have seen, we can think of these directions as "orthogonal" to the space of the FLRW solution. The criticality holds for the variations taken in these directions. This is the case in any gauge choice.

We should note, in particular, that this result does not mean that one can always choose a *gauge* yielding $\delta V[q] = 0$. Our discussion should be understood as follows. That is, since the space of solutions is well-defined only if a gauge is fixed, perturbed solutions and the background solution should be compared within the *same gauge*.

Although we have only discussed the first order effects, one may be (possibly more) interested in higher order effects, for which a partial result has already been obtained in [1]. In general, the second order perturbations of the volume functional do not vanish. Moreover, it is not clear whether an FLRW solution is the local minimum, maximum, or saddle point. This may require further investigations.

Also, we could ask whether the volume criticality holds at other points than the FLRW solutions. As we have seen, the crucial point is whether the perturbation variable(s) corresponding to $C_{(0)}(\eta)$ is governed by a (set of) closed equation(s) or not. Another point may be the correspondence of the number of the degrees of freedom in the general solution to that in a background solution. This is work in progress.

Acknowledgments

We wish to thank T. Inoue for helpful discussions improving a part of the proof. M.T. is grateful to Y. Nambu for enlightening comments. K.Y. thanks A. Hosoya, H. Ishihara and M. Siino for continuous encouragements. He also acknowledges financial support from the Japan Society for the Promotion of Science and the Ministry of Education, Science, Sports and Culture.

References

[1] M. Tanimoto, Prog. Theor. Phys. **102** (1999) 1001.

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

[3] J. M. Bardeen, Phys. Rev. **D22** (1980) 1882.
H. Kodama and M. Sasaki, Prog. Theor. Phys. Suppl. **78** (1984).
J. M. Stewart, Class. Quantum Grav. **7** (1990) 1169.
V. F. Mukhanov, H. A. Feldman and R. H. Brandenberger, Phys. Rep. **215** (1992) 203.

Cosmological Test with the Geometric Distortion and the Evolution of Bias

Kazuhiro Yamamoto,[1]

Department of Physics, Hiroshima University, Kagamiyama, Higashi-Hiroshima, 739-8526, Japan

Abstract

Feasibility of the geometric test as a probe of the cosmological model is discussed assuming the future 2dF QSO sample. Starting from defining estimators for the spatial two-point correlation function and the power spectrum, in a rigorous manner, we derive useful formulas for the two-point clustering statistics incorporating the light-cone effect and the redshift distortions, as well as the nonlinear effect, simultaneously. Then we examine sensitivity of the QSO two-point correlation functions to a bias model whose evolution is phenomenologically parameterized. It is shown that the correlation functions depend sensitively on a mean amplitude of the bias and not to the speed of the redshift evolution. We will also demonstrate that an optimistic geometric test could suffer from confusion that a signal from the cosmological model can be confused with that from a stochastic character of the bias.

1 Introduction

The wide-field redshift survey projects, Two-degree Field (2dF) and Sloan Digital Sky Survey (SDSS), are upcoming, and large numbers of quasars and galaxies are expected to be detected. In particular, the 2dF group has recently reported their preliminary results about the evolution of the quasar luminosity function and the spatial correlation function based on the several thousands quasars detected from the survey so far (Shanks et al. 2000; Boyle et al. 2000). The SDSS survey project will also provide a set of high precision data for the distribution of quasars in near future. These surveys will provide very important clues on the origin and the evolution of the quasars, and will put constraints on theoretical models of the structure formation of the universe. However, there seems to be remained rooms for discussions how we compare the observational data with theoretical models and what information is drawn out from them. Concerning the latter subject, various cosmological tests with the future data have been proposed. Among such cosmological tests, the geometric test using the cosmological redshift-space (geometric) distortion is quite unique.

Alcock and Paczynski first pointed out the possibility of the geometric test (1979). They pointed out that if spherically distributed structures of cosmological objects were in the universe, the observation of such structures at various redshift in redshift space offers a plausible test for the cosmological model of the universe, which is sensitive to a cosmological constant. This test is based on the fact that the shape of the spherically distributed structure is observed distorted in the cosmological redshift-space in which an incorrect distance-redshift relation is assumed. We refer to the distortion as the cosmological redshift-space (geometric) distortion. The geometric distortion is traced back to a coordinate transformation of the distance between real space and redshift space. Several authors have extensively discussed about possible tests for cosmological models with the geometric distortion with the clustering statistics of high-redshift objects (Matsubara & Suto 1996; Ballinger et al. 1996; Popowski et al. 1998; Yamamoto & Nishioka 2000, references therein).

In principle, the geometric test is a unique test for the cosmological model, however, a measurement of the geometric distortion suffers from various other observational effects. In particular the peculiar motion of the cosmological objects causes additional distortion in the distribution in redshift space. Especially the linear distortion, which is due to the bulk motion of the cosmological objects, can be influential even at high-redshift on large scales. The net of the effect depends on the amplitude of the bias. Furthermore, the light-cone effect would become an important effect when the statistical quantities like the correlation function and the power spectrum are computed using the data in a wide range of the redshift. In this

[1] e-mail:kazuhiro@hiroshima-u.ac.jp

case, the evolution of the bias is an important but ambiguous factor in predicting the correlation function and the power spectrum. The theoretical studies based on the Press-Schechter theory and the extensive numerical works recently show that the time evolution of clustering bias for the galaxies and the clusters of galaxies strongly depends on the cosmological models, however, the quantitative prediction for the evolution of galaxy clustering is still under consideration. As for the quasars, the uncertainty of the formation mechanism makes it further difficult to predict the evolution of bias.

In this paper we consider the details of the theoretical predictions for quasar two-point statistics focusing on the evolution of the bias factor. We discuss the points how we can draw out information about the evolution of the bias from the two-point statistics of cosmological objects and how the evolution of bias alters the prediction for the two-point statistics, which determines feasibility of the cosmological test with the geometric distortion (see also Yamamoto et al. 2000; Yamamoto & Nishioka 2000). This paper is organized as follows: In section 2, we summarize theoretical formulas for the two-point correlation function and the power spectrum which incorporate the light-cone effect and the redshift-space distortion effects. Then we investigate the feasibility of the geometric distortion with the correlation functions assuming quasar samples which roughly match the 2dF survey in section 3. Section 4 is devoted to summary and discussions. Throughout this paper we use the unit in which the light velocity c equals 1.

2 Formulation

In this section we summarize the formulas for the two-point statistics of the distribution of cosmological objects in redshift space incorporating the light-cone effect. We here present a brief review for deriving the formula starting from estimators for the anisotropic correlation function and the power spectrum in cosmological redshift-space. We start from reviewing an idealized data processing to compute the two-point correlation function and the power spectrum from a map of cosmological objects. Assuming a redshift survey, data is given in terms of a set of pairs of the direction vector γ and the redshift z of the cosmological objects. When constructing the three dimensional map of the cosmological objects, we must introduce a radial coordinate s by assuming some distance-redshift relation $s = s(z)$. Because the cosmological model of our universe is not definitely determined, we must *assume* the distance-redshift relation $s = s(z)$. Frequently the distance extrapolating the Hubble law, $s = s_{\rm Hb}(z) = z/H_0$, or the comoving distance in the Einstein de Sitter universe,

$$s_{\rm Ed}(z) = \frac{2}{H_0}(1 - 1/\sqrt{1+z}), \tag{1}$$

have been used for the relation, where $H_0 = 100h$km/s/Mpc is the Hubble constant. Throughout the present paper we adopt $s = s_{\rm Ed}(z)$, and refer to the three dimensional space (s, γ) as the cosmological redshift-space.

Next we assume that the number density field $n(\mathbf{s})$ with $\mathbf{s} = s\gamma$ can be constructed from the map. By introducing a synthetic catalogue $n_{\rm syn}(\mathbf{s})$, we construct the density field:

$$F(\mathbf{s}) = \frac{n(\mathbf{s}) - n_{\rm syn}(\mathbf{s})}{[\int ds \bar{n}(s)^2]^{1/2}}, \tag{2}$$

where $\bar{n}(s)$ denotes the mean number density at the distance s. With the use of $\bar{n}(s)$, we write

$$n(\mathbf{s}) = \bar{n}(s)\big(1 + \Delta(\mathbf{s})\big), \tag{3}$$

$$n_{\rm syn}(\mathbf{s}) = \bar{n}(s)\big(1 + \Delta_{\rm syn}(\mathbf{s})\big). \tag{4}$$

The synthetic field $n_{\rm syn}(\mathbf{s})$ satisfies $\langle n_{\rm syn}(\mathbf{s}_1) n_{\rm syn}(\mathbf{s}_2)\rangle = \langle n(\mathbf{s}_1) n_{\rm syn}(\mathbf{s}_2)\rangle = \bar{n}(s_1)\bar{n}(s_2)$, which lead

$$\langle \Delta_{\rm syn}(\mathbf{s}_1)\Delta_{\rm syn}(\mathbf{s}_2)\rangle = \langle \Delta_{\rm syn}(\mathbf{s}_1)\Delta(\mathbf{s}_2)\rangle = 0. \tag{5}$$

Then the two-point correlation function can be computed by averaging the products of $F(\mathbf{s})$ over all the possible pairs of two-points separated by the distance R. A mathematical expression for this process

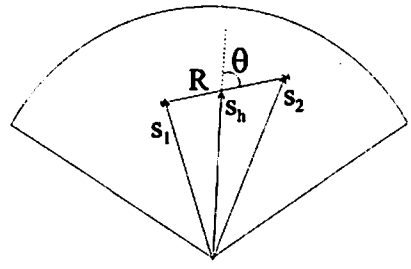


Figure 1: A sketch to illustrate variables defined in this paper.

can be explicitly written

$$\xi_l(R) = \int \frac{d\Omega_{\hat{\mathbf{R}}}}{4\pi} \int ds_1 \int ds_2 F(\mathbf{s}_1) F(\mathbf{s}_2) \delta^{(3)}(\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{R}) L_l(\hat{\mathbf{s}}_h \cdot \hat{\mathbf{R}})(2l+1), \tag{6}$$

where $R = |\mathbf{R}|$, $\hat{\mathbf{R}} = \mathbf{R}/R$, $\mathbf{s}_h = (\mathbf{s}_1 + \mathbf{s}_2)/2$, $s_h = |\mathbf{s}_h|$, $\hat{\mathbf{s}}_h = \mathbf{s}_h/s_h$, $\mu = \cos\theta$, and $L_l(\mu)$ is the Legendre polynomial of the l-th order. (See Figure 1). In the case $l = 0$, where $L_0(\mu) = 1$, $\xi_0(R)$ represents the isotropic (angular averaged) correlation function. On the other hand, other cases measure the anisotropies of the correlation function owing to the factor $L_l(\hat{\mathbf{s}}_h \cdot \hat{\mathbf{R}})$. Especially in the case $l = 2$, $\xi_2(R)$ expresses the quadru-pole anisotropy of the correlation function. We adopt this estimator for the isotropic and the anisotropic correlation functions.

Similar to the case of the correlation function we define the following estimator for the power spectrum (c.f., Yamamoto et al. 1999)

$$P_l(k) = \int \frac{d\Omega_{\hat{\mathbf{k}}}}{4\pi} \int ds_1 \int ds_2 F(\mathbf{s}_1) F(\mathbf{s}_2) e^{i\mathbf{k}\cdot(\mathbf{s}_1 - \mathbf{s}_2)} L_l(\hat{\mathbf{s}}_h \cdot \hat{\mathbf{k}})(2l+1), \tag{7}$$

where $k = |\mathbf{k}|$, $\hat{\mathbf{k}} = \mathbf{k}/k$, and $\mathbf{k}$ is the wave number vector. When adopting the above estimators for the correlation function and the power spectrum, which we have defined independently, we can show that a familiar relation holds between them. Namely, by using the mathematical formulas,

$$\delta^{(3)}(\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{R}) = \frac{1}{(2\pi)^3} \int d\mathbf{k} e^{i\mathbf{k}\cdot(\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{R})}, \tag{8}$$

and

$$e^{i\mathbf{k}\cdot\mathbf{R}} = 4\pi \sum_{l,m} i^l j_l(kR) Y_l^{m*}(\Omega_{\hat{\mathbf{k}}}) Y_l^m(\Omega_{\hat{\mathbf{R}}}), \tag{9}$$

where $Y_l^m(\Omega)$ and $j_l(x)$ denote the spherical harmonics on an unit sphere and the spherical Bessel function, respectively, we can verify that the following relation holds

$$\xi_l(R) = \frac{1}{2\pi^2 i^l} \int dk k^2 j_l(kR) P_l(k). \tag{10}$$

The inverse transformation leads

$$P_l(k) = 4\pi i^l \int dR R^2 j_l(kR) \xi_l(R). \tag{11}$$

As is well recognized as the problem of the cosmic variance, expectation values of the estimators are only predictable theoretically. Then we consider the ensemble average of the estimators, i.e.,

$$\xi_l^{\rm LC}(R) = \langle \xi_l(R) \rangle, \tag{12}$$

$$P_l^{\rm LC}(k) = \langle P_l(k) \rangle, \tag{13}$$

which define the correlation function and the power spectrum on a light cone. Applying the small angle approximation, $P_l^{\rm LC}(k)$ reduces to

$$P_l^{\rm LC}(k) = \frac{\int dz \left(\frac{dN}{dz}\right)^2 \left(\frac{s^2 ds}{dz}\right)^{-1} P_l^{\rm crd}(k,z)}{\int dz \left(\frac{dN}{dz}\right)^2 \left(\frac{s^2 ds}{dz}\right)^{-1}}, \tag{14}$$

where dN/dz denotes the number count of the objects per unit redshift and per unit solid angle, $P_l^{\rm crd}(k,z)$ is defined:

$$P_l^{\rm crs}(k,z) = \frac{2l+1}{c_\perp^2 c_\parallel} \int_0^1 d\mu L_l(\mu) P_{\rm QSO}\left(q_\parallel \to \frac{k\mu}{c_\parallel},\ |\mathbf{q}_\perp| \to \frac{k\sqrt{1-\mu^2}}{c_\perp}, z\right), \tag{15}$$

where $P_{\rm QSO}(q_\parallel, |\mathbf{q}_\perp|, z)$ is the QSO power spectrum, $q_\parallel$ ($\mathbf{q}_\perp$) is the wave number component parallel (perpendicular) to the line-of-sight direction in the real space. In equation (15), we defined

$$c_\perp = \frac{r(z)}{s(z)}, \quad c_\parallel = \frac{dr(z)}{ds(z)}, \quad \omega = \frac{c_\perp}{c_\parallel}. \tag{16}$$

The expression $P_l^{\rm crd}(k,z)$ is well-known as a power spectrum on a constant time hypersurface in the cosmological redshift-space (e.g., Ballinger et al. 1996; Magira, Jing, & Suto 2000). The essence of the geometric distortion is only the coordinate transformation between $(s(z),\gamma)$ and $(r(z),\gamma)$, which is described by the coefficients $c_\perp$ and $c_\parallel$. Expression (16) is correct only for the case of a spatially flat (see Suto et al. 2000; Yamamoto et al. 2000 for generalization to an open universe).

We model the power spectrum of the QSO distribution by introducing the bias factor $b(z)$,

$$P_{\rm QSO}(q_\parallel, |\mathbf{q}_\perp|, z) = \left\{ b(z)^2 + 2b(z)f(z)\mathcal{R}(z)\left(\frac{q_\parallel}{q}\right)^2 + f(z)^2\left(\frac{q_\parallel}{q}\right)^4 \right\} P_{\rm mass}(q,z) D\left[q_\parallel \sigma_P(z)\right], \tag{17}$$

where $q = \sqrt{q_\parallel^2 + |\mathbf{q}_\perp|^2}$, $P_{\rm mass}(q,z)$ is the CDM mass power spectrum, and we defined

$$f(z) = \frac{d\ln D_1(z)}{d\ln a(z)}, \tag{18}$$

with the linear growth rate $D_1(z)$ and the scale factor $a(z)$. The terms in proportion to $f(z)$ in (17) is traced back to the linear distortion effect (Kaiser 1987). To describe a cross correlation between the QSO distribution and the CDM mass distribution, we here introduced the cross correlation coefficient $\mathcal{R}(z)$. In the case of the deterministic bias, the cross correlation coefficient can be set $\mathcal{R}(z) = 1$, however, in the case of the stochastic bias, the cross correlation coefficient is allowed to deviate from unity (e.g., Pen 1998).

In equation (17), $D[q_\parallel \sigma_P(z)]$ is the damping factor due to the Finger-of-God effect, for which we adopt the exponential model for the distribution of the pairwise velocity dispersion σ_P. In this case we have the damping factor

$$D[\mu k \sigma_p(z)/c_\parallel] = \frac{1}{1+\mu^2 k^2 \sigma_p(z)^2/2c_\parallel^2}. \tag{19}$$

For σ_P we adopt an approximate formula whose validity is investigated by several authors (see e.g., Magira et al. 2000). The nonlinear effect is not a dominant effect on large length scales, however, it can not be neglected. Therefore we here take the nonlinear effects into account for definiteness.

The correlation function $\xi_l^{\rm LC}(R)$ can be computed from $P_l^{\rm LC}(k)$ with the use of equation (10). Equation (14) indicates that the light-cone effect is incorporated by averaging the local power spectrum $P_l^{\rm crd}(k,z)$ by weighting the factor in the formula:

$$W(z) = \left(\frac{dN}{dz}\right)^2 \left(s^2 \frac{ds}{dz}\right)^{-1}. \tag{20}$$

The formula (14) is obtained under the small angle approximation. Hence the validity of the use is limited to small scales. For large scales, $R \gtrsim s_{\max}$, where $s_{\max}$ is the size of a survey area, the finite size effect of the survey area affects the proper estimation of the correlation function and the power spectrum. (e.g., Yamamoto & Suto 1999; Nishioka & Yamamoto 2000).

Motivated from the recent report by Shanks et al. (2000), we assume a sample which roughly corresponds with the 2dF quasar redshift (2QZ) survey. Recently, Boyle et al. (2000) have reported on the evolution of QSO optical luminosity function using their preliminary result of the 2QZ survey. The evolution of the luminosity function is compared with analytic fitting formulas. For theoretical predictions which correspond to the on-going 2QZ survey, we adopt their quasar luminosity function with best-fitted parameters of a power-low polynomial evolution of luminosity under the assumption of the Einstein de Sitter universe. Then the number count dN/dz brighter than the limiting magnitude $B \leq 20.85$ can be obtained by integrating the luminosity function. [2]

3 QSO Two-point Statistics and the Evolution of Bias

In this section we discuss the detail of the two-point correlation function focusing on how the quasar two-point statistics depend on the evolution of bias. Recently bias models for the quasar distribution have been proposed (e.g., Martini & Weinberg 2000; Haiman & Hui 2000; Fang & Jing 1998). Construction of theoretical models for the quasar bias is a challenging problem. However, the bias mechanism of the quasar distribution is not well-understood at present. In this paper we assume one of the simplest phenomenological model of the bias with the form,

$$b(z) = \alpha + (b(z_*) - \alpha)\left(\frac{1+z}{1+z_*}\right)^{\beta}, \tag{21}$$

where α, β, and $b(z_*)$ are the free parameters (c.f., Matarrese et al. 1997). Here β specifies the rate of evolution of bias, $b(z_*)$ does the amplitude of bias at some redshift z_*, and α corresponds to the amplitude of bias at $z = 0$ in the limit $z_* \gg 1$. We have assumed the bias model independent of scales k. Throughout the present paper we fix $\alpha = 0.5$, however, this does not change the results for $0 \lesssim \alpha \lesssim 1$. We determined z_* of the bias model by the formula

$$z_* = \frac{\int dz z W(z)}{\int dz W(z)}, \tag{22}$$

which describes a mean redshift. We have $z_* = 1.2$ when assuming the quasar sample with $B_{\rm lim} = 20.85$ in the range $0.3 < z < 2.2$.

For the geometric test, the ratio of the correlation functions $\xi_2(R)/\xi_0(R)$ will be an important quantity to characterize the geometric distortion effect. Hereafter we omit the superscript 'LC' for simplicity. Figure 2 shows the contour of $\xi_2(R)/\xi_0(R)$ to discuss the feasibility of the geometric test. The solid lines show the contour of $\xi_2(R)/\xi_0(R)$ at $R = 20h^{-1}$Mpc for various cosmological models (a) ΛCDM model with $\Omega_0 = 0.3$; (b) ΛCDM model with $\Omega_0 = 0.4$; (c) open CDM model with $\Omega_0 = 0.3$; (d) open CDM model with $\Omega_0 = 0.4$. In each model we adopted $h = 0.7$ and σ_8 normalized by the cluster abundance (Kitayama & Suto 1997).

Recently it is reported that the QSO correlation function is consistent with being $\xi = (r/r_0)^{-1.8}$ with $r_0 = 4\ h^{-1}$Mpc from the preliminary result of the 2QZ survey (Shanks et al. 2000). Previous result of the QSO surveys similar to the 2QZ survey reported $r_0 = 6\ h^{-1}$Mpc (Croom & Shanks 1996). The dashed lines on the panels show the contours satisfying $R_c = 4\ h^{-1}$Mpc and $R_c = 6\ h^{-1}$Mpc, where the characteristic correlation length R_c is defined by $\xi_0(R_c) = 1$. The observational constraint is not strict at present because of statistical errors. Then Figure 2 should be regarded as a demonstration. However, it is instructive. Taking the constraint $4\ h^{-1}$Mpc $\leq R_c \leq 6\ h^{-1}$Mpc into account, the ratio ξ_2/ξ_0 at $R = 20h^{-1}$Mpc is in the range $-2.0 \sim -1.2$ ($-2.0 \sim -1.4$) for the ΛCDM with $\Omega_0 = 0.3$ ($\Omega_0 = 0.4$). While $\xi_2/\xi_0 \sim -1.2 \sim -0.8$ ($-1.4 \sim -1.0$) is predicted in the open CDM model with $\Omega_0 = 0.3$ ($\Omega_0 = 0.4$).

Similar feature appears in the quintessential CDM (QCDM) model. Figure 3 shows contours of ξ_2/ξ_0 on the $(\beta - b(z_*))$ plane with the separation R fixed as $20h^{-1}$Mpc for various cosmological models. Panels

[2] For the K-correction we assume the quasar energy spectrum $L_\nu \propto \nu^{-0.5}$.

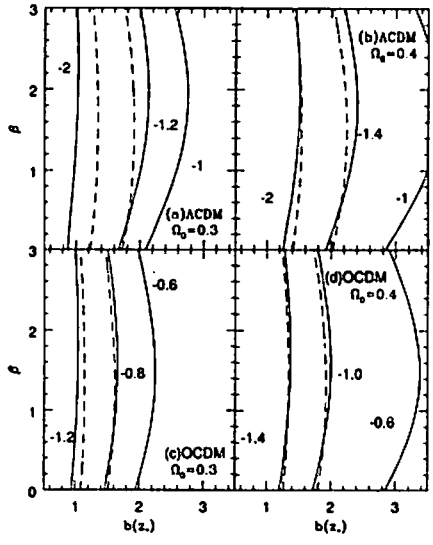


Figure 2: Contour of $\xi_2(R)/\xi_0(R)$ with fixed $R = 20h^{-1}$Mpc for various cosmological models. (a) ΛCDM model with $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$; (b) ΛCDM model with $\Omega_0 = 0.4$, $\Omega_\Lambda = 0.6$; (c) open CDM model with $\Omega_0 = 0.3$, $\Omega_\Lambda = 0$; (d) open CDM model with $\Omega_0 = 0.4$, $\Omega_\Lambda = 0$. In each model we adopted $h = 0.7$, $\Omega_b h^2 = 0.015$ and $\alpha = 0.5$. The contour levels of the solid lines are shown in each panel. The region between the dashed lines in each panel satisfies $4h^{-1}\text{Mpc} \leq R_e \leq 6h^{-1}\text{Mpc}$. We assumed the cosmological redshift space with $s = s_{Ed}(z)$. Here σ_8 is determined from the cluster abundance (Kitayama & Suto 1997), and the fitting formula by Peacock & Dodds (1996) for the nonlinear mass power spectrum is used.

(a) and (b) show the cases of the ΛCDM model ($\Omega_0 = 0.3$, $w_Q = -1$) and the QCDM model (Ma et al. 1999) ($\Omega_0 = 0.3$, $w_Q = -1/2$), respectively. In both panels we show the case of the deterministic bias. For the QCDM model we adopted the simple fitting formulas for the linear transfer function and the nonlinear mass perturbation power spectrum are presented by Ma et al. (1999). The formulas are applicable for the QCDM model with the constant equation of state $-1 \leq w_Q \lesssim -1/3$ and up to $z \sim 4$ with errors $\lesssim 10\,\%$. In addition we adopted the normalization by cluster abundance (Wang & Steinhardt 1998; Wang et al. 2000) for the QCDM model. Figures 2 and 3 show that the value of ξ_2/ξ_0 is sensitive only to the parameter $b(z_*)$, and not to the speed of evolution β.

Let us discuss a possible effect from stochastic character of the bias by considering the case $\mathcal{R}(z) \neq 1$ in (17). For simplicity we consider the case $\mathcal{R}(z)$ is constant. The panel (c) in Figure 3 plots the contour of ξ_2/ξ_0 of the ΛCDM model same as the panel (a) but with $\mathcal{R}(z) = 0.6$. As is clear from the panels (b) and (c), the both panel becomes almost same. That is, the ratio ξ_2/ξ_0 takes the same values in the cases ($w_Q = -1/2$, $\mathcal{R} = 1$) and ($w_Q = -1$, $\mathcal{R} = 0.6$). Thus the signal from the cosmological model can be confused with that from the stochastic character of the bias. This fact suggests that an optimistic geometric test suffers from degeneracy between the cosmological model parameter and the stochasticity parameter of the bias. The panel (d) in Figure 3 is same as the panel (b) but with $\Omega_0 = 0.35$.

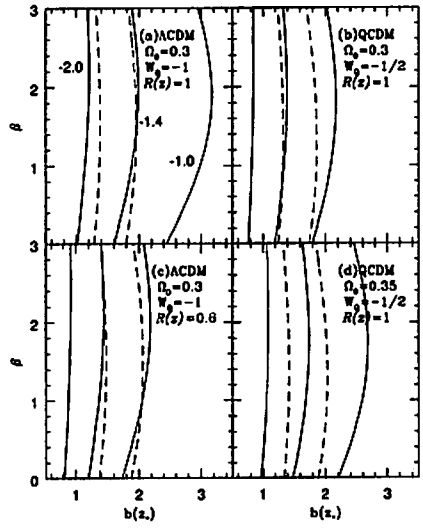


Figure 3: (*Solid line*) Contours of $\xi_2(R)/\xi_0(R)$ at $R = 20h^{-1}$Mpc on the $(\beta - b(z_*))$ plane for various cosmological models: (a) ΛCDM model ($\Omega_0 = 0.3, w_Q = -1.0$); (b) QCDM model ($\Omega_0 = 0.3, w_Q = -1/2$); (c) same as the panel (a) but with the stochastic bias case with $\mathcal{R}(z) = 0.6$; (d) QCDM model ($\Omega_0 = 0.35, w_Q = -1/2$). The contour levels of the solid line are indicated on the figure, and the same contour levels are adopted for each panel. In each panel we adopted $h = 0.7$, $\Omega_b h^2 = 0.015$, and $\alpha = 0.5$ for the bias model. Except for the panel (c), the deterministic bias is considered. The left (right) *dashed line* shows the contour that satisfies the condition of the characteristic correlation length $R_c = 4h^{-1}$Mpc ($R_c = 6h^{-1}$Mpc). Here σ_8 is determined from the cluster abundance (Wang & Steinhardt 1998; Wang et al. 2000), and the fitting formula by Ma et al. (1999) for the nonlinear mass power spectrum is used.

4 Summary and Conclusions

In this paper we have revisited the feasibility of the geometric test focusing on the ambiguous factor of the evolution of bias. We have started from defining the estimators for the anisotropic correlation function and the power spectrum. Then we have derived the useful formulas for the two-point statistics which incorporate the light-cone effect and the redshift-space distortions, simultaneously, in a rigorous manner. Our investigation makes clear how the theoretical two-point correlation function and the power spectrum in our paper are defined according to estimators in data processing. With this formula we have investigated the theoretical predictions assuming a simple model of quasar samples which roughly match with the 2dF survey. Adopting the model for the evolution of bias (21), which is phenomenologically parameterized in a simple form, we have examined the theoretical predictions for the two-point correlation functions.

The amplitude of the correlation functions is sensitive to the mean amplitude of the bias, and is rather insensitive to the speed of evolution due to the light-cone effect. We have found that the systematic difference appears in the ratio of the correlation functions ξ_2/ξ_0 depending on the cosmological model, due to the geometric distortion effect. We have also shown that, if the QSO bias has a stochastic character, the signal from the cosmological model can be confused with that from the stochastic bias. Hence the simple geometric test with only ξ_2/ξ_0 suffers from the degeneracy between the cosmological parameter and the bias parameter unless the stochasticity character is clarified. For example, the ΛCDM model

with $\mathcal{R}(z) = 0.6$ and the QCDM model with $w_Q = -1/2$ and $\mathcal{R}(z) = 1$ predict almost same value of ξ_2/ξ_0 at $R = 20h^{-1}$Mpc. Therefore other cosmological information is required, e.g., the higher order multipole moment of the correlation function ξ_4/ξ_0, which might be useful to break the degeneracy. However, the amplitude of the signal ξ_4/ξ_0 is rather small, which is of order of 10% compared with ξ_2/ξ_0, and the signal from ξ_4 seems to become noisy.

ACKNOWLEDGMENTS

This work is based on the collaboration with H.Nishioka and A.Taruya. This work is supported by the Inamori foundation and in part by the Grants-in-Aid program (11640280) by the Ministry of Education, Science, Sports and Culture of Japan.

References

[1] Alcock C., Paczyński B., 1979, Nature, 281, 358

[2] Ballinger W.E., Peacock J.A., Heavens A.F., 1996, MNRAS, 282, 877

[3] Boyle B.J., Shanks T., Croom S.M., Smith R.J., Miller L., Loaring N., Heymans C., 2000, MNRAS in press, astro-ph/0005368

[4] Croom S.M., Shanks T., 1996, MNRAS, 281, 893

[5] Fang L. Z., Jing Y. P., 1998, ApJ, 502, L95

[6] Haiman Z., Hui L., 2000, submitted to ApJ, astro-ph/0002190

[7] Kaiser N., 1987, MNRAS, 227, 1

[8] Kitayama T., Suto Y., 1997, ApJ, 490, 557

[9] Ma, C-P., Caldwell, R. R., Bode, P., & Wang, L. 1999, ApJ, 521, L1

[10] Magira H., Jing Y.P., Suto Y., 2000, ApJ, 528, 30

[11] Martin P., Weinberg D H., 2000, submitted to ApJ, astro-ph/0002384

[12] Matarrese S., Coles P., Lucchin F., & Moscardini L., 1997, MNRAS, 286, 115

[13] Matsubara T., Suto Y., 1996, ApJ, 470, L1

[14] Nishioka H., Yamamoto K., 2000, ApJS, 128, 1

[15] Peacock J. A., Dodds S. J. 1996, MNRAS, 280, L19

[16] Pen, U. L. 1998, ApJ, 504, 601

[17] Popowski P.A., Weinberg D.H., Ryden S., Osmer P., 1998, ApJ, 498, 11

[18] Shanks T., Boyle B.J., Croom S.M., Loaring N., Miller L., and Smith R.J., proceeding in *Clustering at High Redshift*, eds. Mazure A., and LeFevre O., 2000, in press, astro-ph/0003206

[19] Suto Y., Magira H., Yamamoto K, 2000, PASJ, 52, 249

[20] Wang, L., Steinhardt, P. J. 1998, ApJ, 508, 483

[21] Wang, L., Caldwell, R. R., Ostriker, J. P., & Steinhardt, P. J. 2000, ApJ, 530, 17

[22] Yamamoto K., Nishioka H., Suto Y., 1999, ApJ, 527, 488

[23] Yamamoto K., Nishioka H., Taruya A., 2000, MNRAS, submitted

[24] Yamamoto K., Nishioka H., 2000, ApJ, submitted

[25] Yamamoto K., Suto Y., 1999, ApJ, 517, 1

Primordial Fractal Density Perturbations and Structure Formation in the Universe : 1-D Collisionless Sheet Model

Takayuki Tatekawa[1] and Kei-ichi Maeda[2]

Department of Physics, Waseda University, 3-4-1 Okubo, Shinjuku, Tokyo 169-8555, JAPAN

Abstract

Two-point correlation function of galaxy distribution shows that the structure in the present Universe is scale-free up to a certain scale (at least several tens Mpc), which suggests that a fractal structure may exist. If small primordial density fluctuations have a fractal structure, the present fractal-like nonlinear structure below the horizon scale could be naturally explained. We analyze the time evolution of fractal density perturbations in Einstein-de Sitter universe, and study how the perturbation evolves and what kind of nonlinear structure will come out. We assume a one-dimensional collisionless sheet model with initial Cantor-type fractal perturbations. The nonlinear structure seems to approach some attractor with a unique fractal dimension, which is independent of the fractal dimensions of initial perturbations.

1 Introduction

The present Universe shows a variety of structures. The galaxies are not distributed randomly in the Universe. Totsuji and Kihara [1] and Peebles [2] showed that the observed two-point correlation function $\xi(r)$ is given by a power law with respect to a distance r as $\xi(r) \sim r^{-\gamma}$ with $\gamma \sim 1.8$. The recent galaxy surveys also agree with this result, i.e. the power γ is nearly equal to 1.8 (CfA [3], LCRS [4], and ESP [5], [6]). This may imply that the present distribution of galaxies is fractal. However, the observation of Cosmic Microwave Background Radiation (CMBR) has revealed that the Universe in the recombination era is homogeneous and isotropic at least in very large scale. Although CMBR observation seems to be more reliable, we should not decide yet whether the large scale structure of the Universe is really fractal up to the horizon scale or not.

However, since it seems true that the galaxy distribution is really fractal up to a certain scale, one may ask how such a structure is formed in the evolution of the Universe. One of the most plausible explanations is that the nonlinear dynamics of the perturbations will provide such a scale-free structure during the evolution of the Universe. The pioneering work to explain the power-law behavior in nonlinear stage has been done by Davis and Peebles [8]. They showed a relation between the power index γ of two-point correlation function and that of initial power spectrum n as $\gamma = 3(n+3)/(n+5)$. If we have $n = 0$, then we find that $\gamma = 1.8$. Since their additional condition is not trivial and might not be appropriate, Padmanabhan [9] and Yano and Gouda [10] extended their model. Since we do not know the stability of those solutions, in order to find which value of γ is most likely, we should study the dynamics of density fluctuations in other methods.

As for a fractal structure in the Universe, one may ask another question. Did the Universe not have any non-trivial structure such as a fractal in the initial density fluctuations? In the conventional approaches, initial density perturbations are usually assumed to be given by a power-law (or a power-law-like) spectrum with random Gaussian phase. Although such initial conditions may provide the presently observed nonlinear scale-free structure via nonlinear dynamics, no one has shown whether such a structure is fractal or not, and if yes, what kind of fractal structure comes out. To provide a fractal structure in the present Universe, we may adopt an alternative scenario, in which primordial density fluctuations have already a fractal-like structure in the beginning. Note that a background spacetime is assumed to be a smooth universe, which is described by the Einstein-de Sitter universe, but not a fractal universe. The properties of an initial fractal may be preserved during the evolution of the Universe, then nonlinear

[1] E-mail:tatekawa@gravity.phys.waseda.ac.jp
[2] E-mail:maeda@gravity.phys.waseda.ac.jp

fractal structure will be formed. In this scenario, several natural questions may arise. How does such a primordial fractal perturbation evolve into nonlinear regime? Will any properties of the initial fractal be preserved during the evolution of the Universe, or not? If not, what kind of nonlinear structure will come out at present? Is there any fundamental difference in the structure formation process between a conventional density perturbation and the present fractal one? In order to answer those questions, we study the time evolution of the initial density fluctuations with a fractal structure in Einstein de-Sitter universe.

Since we are interested in a fractal structure, a quite high resolution is required in our calculation. As we discussed, N-body simulation may not have enough resolution in the present state of computer development, unless we develop some skillful method. So we consider only a very simple toy model, which is a one-dimensional (1-D) sheet model, in order to get some insight into the questions raised in the above. To set up primordial fractal density perturbations, we distribute N sheets initially by some systematic rule, i.e. we apply a Cantor set or random Cantor-type set. Since we study a 1-D sheet-model, the motion of each sheet is described by an analytic solution [11], which guarantees enough resolution to analyze a fractal structure.

2 FORMALISM AND INITIAL DATA

2.1 Dynamical Equations

In order to study the structure formation of the expanding Universe, there are so far three approaches: N-body simulation, the Eulerian perturbation approach and the Lagrangian one. Although the final answer for structure formation would be obtained by N-body simulation, it may not be possible to so far answer the questions about a fractal structure. As for the perturbation approaches, these are just an approximation and will break down in nonlinear regime, although the Lagrangian approach would be better if we are interested in the density perturbations. This is just because a density fluctuation δ and a peculiar velocity $\mathbf{v}$ are perturbed quantities in the Eulerian approach [12], while displacement of particles is assumed to be small in the Lagrangian approach [11],[13],[14],[15],[16]. Its first order solution is the so-called Zel'dovich approximation [11]. The Lagrangian approach is confirmed to be better than the Eulerian approach by comparison of these results in several cases [17], [18], [19].

Even though the Lagrangian approach is better, it is still an approximation and then is not suitable to discuss a highly nonlinear structure formation. However, there is one exceptional case. If the distribution is plane symmetric, the system is one-dimensional, and then the Zel'dovich approximation turns out to be an exact solution. We discuss only one-dimensional problem, which Lagrangian perturbation is given by

$$x = q + S(q,t) \tag{1}$$

where x and q are the one-dimensional Eulerian and Lagrangian comoving coordinates, respectively. For Einstein de-Sitter model, the solution is given by ([14], [20]):

$$S(q,t) = a(t)S_1(q) + a(t)^{-3/2}S_2(q)\,, \tag{2}$$

where the scale factor a changes as $a = (t/t_0)^{2/3}$. Then we find the position and its velocity of a dust particle at a scale factor $a(t)$ with respect to the Lagrangian coordinate q as

$$x(q,a) = q + a(t)S_1(q) + a(t)^{-3/2}S_2(q)\,, \tag{3}$$

$$\tilde{v}(q,a) = S_1(q) - \frac{3}{2}a(t)^{-5/2}S_2(q), \tag{4}$$

where we have introduced a new peculiar velocity $\tilde{v} \equiv \partial x/\partial a$. In what follows, we use a scale factor $a(t)$ as a time coordinate instead of the physical time t.

Although the Zel'dovich solution is exact, in studying a formation of nonlinear structure, a serious problem will soon arise. As a density fluctuation grows, we find a shell crossing. If instead of a usual matter fluid, we have a collisionless particle such as some dark matter, we can go beyond the shell crossing. The particles, which are described by plane parallel sheets, will pass through each other without collision.

Then after this crossing, we rediscover that the Zel'dovich solution is again exact. Therefore, we have a series of exact solutions, which is almost analytic [20], [21]. We shall call it the 1-D collisionless sheet model.

In this model, Gouda and Nakamura [20] investigated a time evolution for the density perturbations with a scale-free initial power spectrum. They showed that the power spectrum will approach some characteristic value independent of the power index of the initial spectrum. This characteristic power index -1 is predicted by a catastrophe theory. Recently, Yano and Gouda [21] investigated a time evolution of the density perturbations for an initial power spectrum with a cutoff. In this case, they found that a self-similarity in all scales is no longer valid. The spectrum is classified into five ranges by its power index. Some spectra coincide with the above one for some scale ranges, but another power index, which is independent of the initial power spectrum, appears just beyond the cut-off scale.

Since this 1-D sheet model is powerful enough to see the fine structure, we shall use this model to analyze a time evolution of the primordial fractal density perturbations.

2.2 Setting Up Initial Data

Since we are interested in initial density fluctuations with a fractal distribution, we have to construct such an initial data. A Cantor set is given by the following procedure [22]. We first divide a line with the length L by some integer n_D and then remove one line segment at the center. If n_D is an even integer, not dividing a line by n_D, we just remove a line segment with the length L/n_D from the center of the line. We then repeat this procedure for the remained line segments. In mathematical definition, the removal procedure must be repeated infinite times. However we believe that a fractal structure, even if it exists in the Universe, is not a mathematical one but its self-similarity may end at some scale. Then we also stop our procedure after a finite number of repetitions. We regard the remaining line segment as a region of positive density fluctuations (δ_+ region), while the removed part corresponds to a region of negative density fluctuations (δ_- region). $\delta_+(>0)$ and $\delta_-(<0)$ are chosen to be uniform, i.e. both δ_+ and δ_- are some constants in all regions such that $\delta_+ \simeq |\delta_-| \simeq 10^{-3}$. Although this is very artificial and the realistic perturbations may depend on each scale just like conventional density perturbations, we analyze only this simplest case here.

We constructed 7 initial data by removing central line segments with a fixed ratio ($1/n_D$ ($n_D = 3, 6, 8, 10, 12, 15, 20$)) from the remained parts. Each density fluctuation has a different fractal dimension given by

$$D_0 = \frac{\log 2}{\log[2n_D/(n_D-1)]}. \tag{5}$$

We assume that $\delta_+ \sim |\delta_-|$. With this ansatz, we have fixed the repetition number of removal procedure (N_R). Although the number N_R can be different for each initial data, in order to keep the same resolution for each model, (i.e. for the ratio of the smallest line segment ℓ to that of the whole region L (our calculation space) to be almost same), we set $N_R = 5 \sim 7$. We also fix $\delta_+ = 10^{-3}$, which determines the negative density perturbation δ_- such that the mean of fluctuations must vanish (see Table 1).

n_D	D_0	N_R	ℓ/L	δ_-
3	0.631	5	1/243	-0.152×10^{-3}
6	0.792	6	1/191.10	-0.504×10^{-3}
8	0.838	7	1/325.95	-0.647×10^{-3}
10	0.868	7	1/267.62	-0.917×10^{-3}
12	0.888	7	1/235.36	-1.19×10^{-3}
15	0.909	7	1/207.47	-1.61×10^{-3}
20	0.931	7	1/183.29	-2.31×10^{-3}

Table 1 :
The number of division n_D, its fractal dimension D_0, the repetition number of removal procedure N_R, the resolution (the ratio of the length of the shortest line segment to that of the whole system ℓ/L), and each negative density perturbation δ_-.

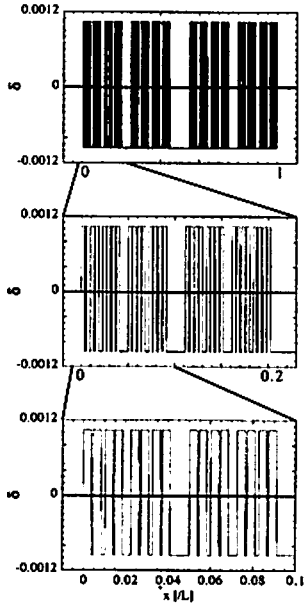


Figure 1: The initial fractal density fluctuation for the regular Cantor set with $n_D = 10$. Enlarging the picture, we find the same pattern up to the repetition number of removal procedure $N_R = 7$.

We show one example of initial data for $n_D = 10$ in Fig.1. We put 2^{17} sheets in our calculation space L.

Since the Cantor set is quite systematically constructed, one may wonder that our results strongly depend on such a special setting and may not be universal. To answer such a question, we also analyze two different initial data settings: one is a random Cantor-type set and the other is just a white noise.

3 TIME EVOLUTION OF PRIMORDIAL FRACTAL DENSITY PERTURBATIONS

In order to see how structures are formed, we have to describe the distribution by the Eulerian coordinate. It is convenient to compare the distribution at each time by use of the comoving coordinate x. We set the initial scale factor $a_0 = 1$. Since our initial density fluctuation is $\sim 10^{-3}$, we find a first shell crossing at $a = a_{\rm cross} \simeq 10^3$. We perform our calculation until $a = (2 \sim 3) \times 10^4$.

In order to see the detail to resolve a fractal structure, we put 2^{17} sheets in our calculation scale L. Using a box-counting method, i.e, counting the number of boxes which contain the region with density perturbation δ larger than 1, we determine the fractal dimension of the nonlinear structures. The size of the box ranges from $2^{-16}L$ to $10^{-1}L$.

First, we show the results for the regular Cantor set with $n_D = 10$. Before a shell crossing, the pattern of density fluctuations remains similar to that of initial distributions, although each separation is going to change through the gravitational interaction. After the shell crossing, the trace of initial Cantor set gradually disappears, because of the exchange of the shells by a crossing. Although many sheets cross each other, the peculiar velocities do not vanish immediately because of collisionless sheets, and then the pattern of nonlinear structure will change continuously. After sufficient evolution of the nonlinear structure, we find a self-similarity in the structure, which seems to be fractal. In fact, enlarging some regions, we find similar density distributions.

In order to judge whether such a structure is fractal or not, we use a box-counting method, which gives a fractal dimension D_F. Before a shell crossing, we find that the dimension D_F decreases in time from the initial value $D_0 = 0.868$, but the error in the estimation increases with time Fig.2 . This is because although the pattern of initial density fluctuations remains even in a nonlinear stage before a shell crossing, change of each separation breaks the initial fractal distribution. Then, the initial fractal distribution seems to disappear.

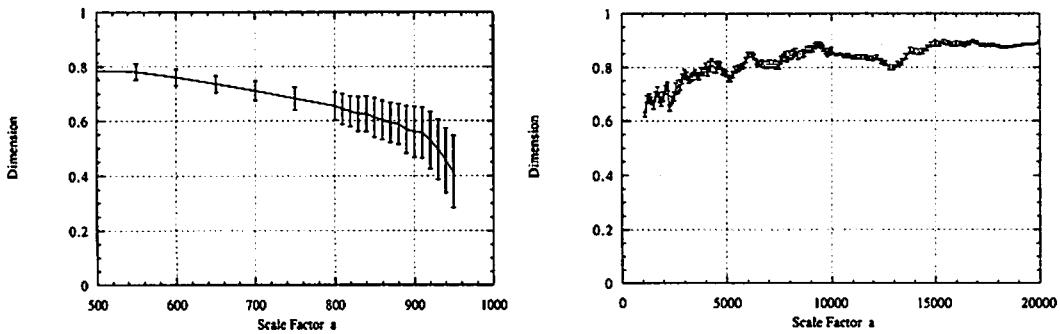


Figure 2: (a) The evolution of the "fractal" dimension D_F of nonlinear structure for the case of $n_D = 10$. Before a shell crossing ($a < a_{\rm cross} \sim 10^3$). (b) is the same as (a) but after a shell crossing ($a \geq a_{\rm cross} \sim 10^3$).

We find a fractal structure after a shell crossing. The fractal dimension is close to the fractal dimension of the initial density distributions. Does this dimension reflect that of the initial distributions? If this is so, why does it disappear once around a first shell crossing time and recover at very late time? In order to answer these question, we have looked for other initial conditions with different n_Ds, i.e. $n_D = 3, 6, 8, 12, 15, 20$. The fractal dimensions D_0 of the initial density distributions are given in Table 1. As for the time evolution, we find similar behaviour for all models. The evolution of "fractal" dimension D_F is shown in Fig. 10. Surprisingly, for all models, all D_F approach about 0.9 at $a = (2 \sim 3) \times 10^4$, which is the end of our calculation. To confirm our result, we have checked the size-number (x-$N(x)$) relation by a box-counting method, which shows almost a straight line as in Fig. 2. Since the initial dimensions of primordial fluctuations were different, we would conclude that the fractal dimension obtained after nonlinear evolution is universal within our numerical accuracy [23].

4 Other expansion rate cases

We calcurated fractal dimension of nonlinear structure in Einstein de-Sitter Universe. In this model, the scale factor is given as

$$a(t) \propto t^{2/3}. \tag{6}$$

In previous section, we find a kind of attractor with a universal fractal dimension (~ 0.9) as the fluctuations evolve into nonlinear regime.

Is the universal dimension of nonlinear fractal structures due to the expansion rate of background Universe model? In order to answer this question, we shall analyze nonlinear structure in other Universe model. Accoording to recent observation, our Universe also seems open or A-flat model. However we calcurate nonlinear evolution in these model, the density perturbation will freeze during evolution.

In order to answer the question, we take other situation; we assume that the scale factor obeys power law of time:

$$a(t) \propto t^n. \tag{7}$$

In this case, we also assume pressureless ($P = 0$) matter. And we assume this relation:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3}. \tag{8}$$

This model doesn't satisfy Einstein's equation, we analyze only the relation between spatial expantion rate and fractal dimension of nonlinear structure. We substitute these relation for the equation of growing rate b of fluctuation as follows:

$$\ddot{b} + 2\frac{\dot{a}}{a}\dot{b} - 4\pi G\bar{\rho}b = 0. \tag{9}$$

When we set faster expansion rate, for example $n = 1$, nonlinear structure doesn't become fractal. On the other hand, when we set slower expansion rate, i.e. $n = 1/2$ or $1/3$, the fractal dimension of nonlinear structure decreases. In these case, the dimension approaches to ~ 0.8.

Recently, Koyama and Konishi [24] reported about fractal structure formation in 1-D sheet model without cosmic expantion. In their case, even if primordial density fluctuation is not fractal, nonlinear fractal structure was formed. We can also analyze expansionless model with our equation. However our model is different from their model, because our equation is obtained from perturbation theory. On the other hand, Newtonian 1-D sheet model was obtained from equation of motion [25]. So when we analyze expansionless model using our equation, the result may different from their result.

5 CONCLUSIONS AND DISCUSSION

We have studied the nonlinear evolution of primordial fractal fluctuations by using a 1-D sheet model. We have analyzed 7 models with initial fluctuations constructed by a regular Cantor set. For all models except for the case with white noise, we find a kind of attractor with a universal fractal dimension (~ 0.9) as the fluctuations evolve into nonlinear regime.

Why is the fractal dimension close to 0.9? Is it really universal? Is the present fractal structure really an attractor? Although we need more analysis to answer this question, we have some hints in previous work. Gouda and Nakamura studied the present 1-D sheet model for the initial power law spectrum. They found two types of generic singularities when we have a shell crossing [20], [26]. When a first shell crossing appears, the relation between Eulerian and Lagrangian coordinates must be

$$x = q_c + \beta(q - q_c)^3 + \cdots, \tag{10}$$

while that after a shell crossing turns out to be

$$x = q_c + \beta(q - q_c)^2 + \cdots. \tag{11}$$

Following Arnold's classification, the former and latter cases are classified into A3 and A2, respectively. A3 is structurally unstable and may appear transiently in the expanding Universe. A2 is structurally stable and appears universally for the initial power-law spectrum. The latter case gives

$$\begin{aligned} \delta_k &= \int \delta(x) e^{ikx} dx \\ &\propto \int (\beta x)^{-1/2} e^{ikx} dx \\ &= (\beta k)^{-1/2} \int \eta^{-1/2} e^{i\eta} d\eta, \end{aligned} \tag{12}$$

i.e. $P(k) \sim k^{-1}$. This predicts $\gamma = 0$, i.e. $D_F = 1$, which is rather close to our "universal" dimension 0.9. Although one may wonder that these are essentially the same, we have another result which suggests that there seems to exist a new type of stable phase. Recently, Yano and Gouda analyzed a more realistic case, i.e. the initial power law spectrum with a cut-off and found 5 characteristic regions in Fourier space [21]. The regime 1 is the linear one and then it is just an initial power spectrum. In the regime 2, they found $P(k) \sim k^{-1}$, which is the single-caustic regime [20]. The regime 3 is called the multi-caustic regime, in

which the power spectrum depends on the initial power-law index. Beyond the cut-off scale, two regimes appear, one gives $P(k) \sim k^{-1}$ (regime 5), which may correspond to A2 type stable solution. In the intermediate wave number k between the regime 3 and regime 5, they find k^{ν}, which ν is independent of initial power index and close to 1, but a little less. They called it the virialized regime. This seems to be a new transient region, which may appear in some specific initial conditions. We would conjecture that the fractal structure with a universal dimension 0.9 corresponds to this virialized regime (regime 4) and the dimension 0.7 found in the case with white noise would be the regime 3. By reanalyzing the Yano-Gauda model in the case of $k = 0$, we have confirmed that $\nu =\sim 0.9$. We also find a small tail with index 0.7 in the size-number relation in the Cantor set model with $n_D = 15, 20$ [23].

This conjecture is also supported by the analysis for a self-gravitating 1-D sheet model without the background expansion of the Universe [27]. They found two time scales; one is a micro relaxation time ($t_{\rm micro} = Nt_c$) and the other is a global relaxation time ($t_{\rm global} = 4 \times 10^4 Nt_c$), where $t_c = \sqrt{L/4\pi GNm}$ is a crossing time. After $t_{\rm micro}$, some equilibrium state is reached by exchanging particle energy, but the global relaxation is not achieved, i.e. the partition function is not yet described by an equilibrium state such as an ergodic state [27]. In the present model, we can speculate that the fractal structure is obtained after this micro relaxation time but before the global relaxation time. In fact, if we estimate the above time scales in the present models, we find that $t_{\rm micro}$ corresponds to $a = 5 \times 10^3$, while $t_{\rm global}$ corresponds to $a = 5 \times 10^6$. The time when we find a stable fractal structure ($a = (1 \sim 3) \times 10^4$) is between those two time scales. If this speculation is true, our fractal structure is temporal. In the future of the Universe, it will evolve into more relaxed and ergodic state.

We also analyzed other expantion rate cases. In these model, fractal dimension of nonlinear structure approaches other value (~ 0.8). Although our model doesn't satisfy Einstein's equation, we think our result shows that expantion rate plays an important role in fratal structure formation.

Since we analyze the simplest case, we have to extend our analysis to more generic cases. First, we should study different types of fractal in order to check whether the present results are universal for any fractal distributions or not. Secondly, we need to analyze the case with scale-dependent fluctuations. In the present analysis, we set $\delta_+ =$ constant and $\delta_- =$ constant. In the realistic case, there must be a scale dependence to the fluctuations. In the conventional perturbations, we usually assume a power-law spectrum with some cutoff. Even if the primordial fluctuations contain a fractal structure, their amplitude may depend on the scale. Its dependence may change the present results. In particular, in the present model, the scenario of structure formation could be different from either TOP-DOWN or BOTTOM-UP for some scale-dependence. The primordial fractal fluctuations will evolve directly into a hierarchical nonlinear structure. But, it will definitely depend on the scale dependence of the fluctuations.

For more realistic cases, we must study either the 2-D or 3-D model. Since the Zel'dovich solution is no longer exact, we have to explore a new method. In order to preserve a high resolution, we may develop a kind of renormalization method in N-body simulation as Couchman and Peebles [28].

Finally, it would also be interesting to look for the origin of such a primordial fractal density perturbation. The inflationary scenario may provide the origin of primordial fluctuations. One may wonder whether such a fractal primordial fluctuation is expected in some inflationary models. If we have more than two scalar fields, then the system is not integrable and may show a chaotic behavior or a fractal property [29]. Such a model might show up a kind of fractal density perturbation.

6 acknowledgements

We would like to thank N. Gouda, P. Haines, O. Iguchi, T. Kurokawa, V. Lukash, M. Morikawa, A. Nakamichi, Y. Sota, T. Yano, and A. Yoshisato for many useful discussions. Part of this work was done while KM was participating the program, "Structure Formation in the Universe", at the Newton Institute, University of Cambridge. KM is grateful to the Newton Institute for their hospitality. Our numerical computation was carried out by Yukawa Institute Computer Faculty. This work was supported partially by a Grant-in-Aid for Scientific Research Fund of the Ministry of Education, Science and Culture (Specially Promoted Research No. 08102010), and by the Waseda University Grant for Special Research Projects.

References

[1] H. Totsuji and T. Kihara, Pub.Astron.Soc.Japan **21** (1969) 221.

[2] P.J.E. Peebles, Astron. Astrophys. **32** (1974) 197.

[3] M.J. Geller and J.P. Huchra, Science **246** (1989) 897.

[4] Y.P. Jing, H.J. Mo, and G. Börner, Astrophys. J. **494** (1998) 1.

[5] L. Guzzo et al. (The ESP Team) in *Proceedings of XIV IAP Colloquium "Wide Field Surveys in Cosmology"* (eds. S. Colombi and Y. Mellier, Paris, May 26-30 1998).

[6] L. Guzzo et al., Astron. Astrophys. **355** (2000) 1.

[7] F. Sylos Labini, M. Montuori, and L. Pietronero, Phys.Rep. **293** (1998) 61.

[8] M. Davis, and P.J.E. Peebles, Astrophys. J. Supp. **34** (1977) 425.

[9] T. Padmanabhan, Mon. Not. R. Astr. Soc. **278** (1996) L29.

[10] T. Yano and N. Gouda, Astrophys. J. **495** (1998) 533.

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

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

[13] F.R. Bouchet et al., Astrophys. J. **394**L51992.

[14] F.R. Bouchet et al., Astron. Astrophys. **296** (1995) 575B.

[15] P. Coles and F. Lucchin, *Cosmology* (John Wiley & Sons, 1995).

[16] P. Catelan, Mon. Not. R. Astr. Soc. **276** (1995) 115.

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

[18] V. Sahni and S. Shandarin, Mon. Not. R. Astr. Soc. **282** (1996) 641.

[19] A. Yoshisato, T. Matsubara, and M. Morikawa, Astrophys. J. **498** (1998) 49.

[20] N. Gouda and T. Nakamura, Prog.Theor.Phys. **81** (1989) 633.

[21] T. Yano and N. Gouda, Astrophys. J. Sup. **118** (1998) 267.

[22] K. Falconer, *Fractal Geometry* (John Wiley & Sons, 1990).

[23] T. Tatekawa and K. Maeda, Astrophys. J. in press.

[24] H. Koyama and T. Konishi, preprint astro-ph/0008208; preprint astro-ph/0008507; preprint astro-ph/0010030.

[25] G.L. Camm, Mon. Not. R. Astr. Soc. **110** (1950) 305.

[26] N. Gouda and T. Nakamura, Prog.Theor.Phys. **79** (1998) 765.

[27] T. Tsuchiya, T. Konishi, and N. Gouda, Phys.Rev. **E50** (1994) 2607.

[28] H.M.P. Couchman and P.J.E. Peebles, Astrophys. J. **497** (1998) 499.

[29] R. Easther and K. Maeda, Class. Quant. Grav. **16** (1999) 1637.

The Origin of Fractal Distribution in Self–Gravitating Virialized System

Osamu Iguchi[a] [1], Yasuhide Sota[a,c] [2], Masahiro Morikawa[a] [3],
Takayuki Tatekawa[b] [4], and Kei-ichi Maeda[b,c,d] [5]

[a] *Department of Physics, Ochanomizu University, 2-1-1 Ohtuka, Bunkyo, Tokyo,112-8610 Japan*

[b] *Department of Physics, Waseda University, Shinjuku, Tokyo 169-8555, Japan*

[c] *Advanced Research Institute for Science and Engineering, Waseda University, Shinjuku, Tokyo 169-8555, Japan*

[d] *Advanced Institute for Complex Systems, Waseda University, Shinjuku, Tokyo 169-8555, Japan*

Abstract

Using one-dimensional ring model in which N particles constrained on a circular ring are interacting by three-dimensional (3D) gravitating force, we examine the origin of fractal structure. Classifying the particles in each state into three phases (*core* , *halo* , and *gas* phases), we can characterize this virialized state by non-thermal property of the particles in *halo* phase. Non-Gaussian velocity distribution and fractal structure are found for *halo* particles. The fractal dimension is independent of the total energy. These results suggest that a scale free criticality in the gravitational interaction could be the main origin to cause the universal properties such as fractal structure and non-Gaussian velocity distribution in the 3D self-gravitating system.

1 Introduction

Astrophysical objects in our universe such as clusters of galaxies, galaxies, globular clusters, and molecular clouds are composed of gravitationally interacting elements and are regarded as a self-gravitating system. Their statistical properties are characterized by the non-Gaussian velocity (or pair-wise velocity) distribution [1], the fractal structure in a configuration space [2], and the scaling relation between mass density and its size [3]. Since most of them are gravitationally virialized, gravity seems to play an important role in causing such statistical properties.

As for molecular clouds, the statistical properties such as Larson's scaling law [4], non-Gaussian velocity distribution and fractal structure in the iso-thermal contour[5] seem to be connected with the turbulence theory [6]. However, the turbulence theory itself is unsatisfactory, since there is no dynamical model leading to such a peculiar scaling law. In particular, when the component gas is strongly compressed and its self-gravity becomes important, we do not have any experimental data in contrast with the turbulence flow in the laboratory.

It may be important to invoke another possibility. There are so far several approaches to explain the fractal phenomena in self-gravitating system from the viewpoint of criticality and phase transition in gravito-thermo dynamics [7]. Strictly speaking, 3D self-gravitational system cannot attain a stable equilibrium, because not only gravitational force is long range but also its potential diverges at origin. This causes the gravothermal catastrophe in the self-gravitating gas system enclosed by a solid adiabatic wall [8, 9, 10]. However, there may be a quasi-equilibrium state, which provide us a quasi-stable and important self-gravitating astrophysical object in the universe. In order to discuss such an equilibrium state, it is necessary to include a cutoff at small scale as well as at large scale. It is well known that 3D gravitational system with the small scale cutoff has the state with a negative specific heat in the

[1] E-mail:osamu@phys.ocha.ac.jp
[2] E-mail:sota@skyrose.phys.ocha.ac.jp
[3] E-mail:hiro@phys.ocha.ac.jp
[4] E-mail:tatekawa@gravity.phys.waseda.ac.jp
[5] E-mail:maeda@mse.waseda.ac.jp

microcanonical state [11]. Although no gravothermal catastrophe occurs against the case without a cutoff, this system is highly unstable in the canonical state where the system is immersed in the heat bath. The system jumps from a gas state to a cluster state as the temperature goes down, which is just the first order phase transition.

So far several toy models have been proposed to study such statistical properties of 3D gravitational system[12]. For example, the cell model and its extended version are among the simplest models to possess the property of phase transition, since the pair potential for the interaction takes the constant values both inside and outside the cells [13, 14]. It has the characteristic length scale of the cell size, against the scale-free 3D gravitational system. On the other hand, one-dimensional self-gravitating sheet model [15, 16], which has no characteristic length but no phase transition, is the antidote of this model. In these respects, these toy models are complementary, each other, but both of them are still missing to grasp the full nature of 3D self-gravitating system. There are also several numerical analysis with the different type interactions including the effect corresponding to the lower and higher cutoffs[14, 17]. However, the value of the cutoff among these models seems too large to see the intrinsic properties of real 3D gravitational model.

Here, we pay our attention to the dynamical aspects of the 3D gravitational force on such a phase transition and examine the one-dimensional N-body ring model where particles constrained on a ring interact each other through 3D gravitational force. We examine the case with the small cutoff and study the cutoff dependence on the statistical properties.

In section 2, we will introduce our ring model, then classify the quasi-equilibrium states into three including the case with negative specific heat in section 3. In sections 4 and 5, we analyze the particle motions from statistical point of view, and study the fractal structure of a formed cluster. Finally, we will discuss our results in section 6.

2 Self-gravitating ring model

In this section, we introduce a self-gravitating ring (SGR) model, which is one-dimensional self-gravitating N particles with mass m, whose motions are constrained on a circular ring with a fixed radius R. The force between each pair of particles is just the three dimensional gravitational interaction, whose distance is measured by a straight line in two dimensional Euclidean plane containing the ring.

We start with the physical Hamiltonian,

$$H_{\rm phys} = \frac{1}{2mR^2}\sum_{i=1}^{N} P_i^2 - \sum_{i<j}^{N} \frac{Gm^2}{\sqrt{2}R\sqrt{1-\cos(\theta_i-\theta_j)+\epsilon}}. \tag{1}$$

where the position of i-th particle constrained on the ring is described only by an angular variable θ_i as $\boldsymbol{r}_i = (R\sin\theta_i, R\cos\theta_i)$, and ϵ is a cutoff parameter and the momentum P_i conjugate to θ_i is given by $P_i = mR^2 d\theta_i/dt$. The gravitational interaction vanishes because of this cutoff, if the distance of two particles is smaller than the scale length $\xi := \sqrt{2\epsilon}R$.

For our numerical simulation, we shall introduce dimensionless variables as follows: The physical Hamiltonian (1) can be rescaled as

$$H_{\rm phys} = Gm^2\frac{N}{R}H, \tag{2}$$

where

$$H = \frac{1}{2}\sum_{i=1}^{N} p_i^2 - \sum_{i<j}^{N} \frac{1}{\sqrt{2}N\sqrt{1-\cos\theta_{ij}+\epsilon}}. \tag{3}$$

Here a dimensionless momentum p_i is defined by $p_i = d\theta/d\tau$ and a new time coordinate τ is introduced as $\tau := t/t_{\rm dyn}$. $t_{\rm dyn}$ is a characteristic dynamical time of the present system defined by

$$t_{\rm dyn} := \sqrt{\frac{R^3}{GNm}} \sim \frac{1}{\sqrt{G\bar{\rho}}}, \tag{4}$$

where $\bar{\rho}$ is a mean density of our system. $t_{\rm dyn}$ denotes a time scale of a particle going around the ring. Here we have numerically solved the present model with a simplectic integrator [18], by which the

total Hamiltonian is conserved with a quite high accuracy, even after a long time-interval of numerical simulation.

When a cluster is formed, we find another typical time scale $t_{\text{cl-dyn}}$, which is a dynamical time scale of the formed cluster and is different from t_{dyn}. This is just because the mass density in the cluster is much larger than that of a gaseous phase, i.e. $t_{\text{cl-dyn}}$ is estimated by

$$t_{\text{cl-dyn}} = \sqrt{\frac{1}{G\bar{\rho}_{\text{cl}}}} \sim \sqrt{\frac{d_{\text{cl}}^3}{GNm}}, \tag{5}$$

where ρ_{cl} and d_{cl} are a mass density of a cluster and a cluster size, respectively. Although t_{dyn} is independent of a cutoff parameter ξ, the dynamical time $t_{\text{cl-dyn}}$ may depend implicitly on ξ because d_{cl} is related to ξ.

As we see from (2), if we do not change the number density in the ring, i.e. if N/R is fixed, the present model is equivalent to the model defined by the Hamiltonian (3), then the properties are essentially determined by a particle number N and a cutoff parameter ϵ. It should also be stressed that the present model (3) keeps extensiveness of the system. In this paper, we fix the total number of particles to $N = 100$. Then, only the cutoff parameter ϵ is a parameter which determines all properties.

3 Classification of particles

Here, we study the properties of "equilibrium" state in SGR model. Since the present system includes a state with a negative specific heat as is shown below, there is no thermal equilibrium state in a usual sense. Hence we use "equilibrium" for a stationary or quasi-equilibrium state, which is sometimes found as a kind of transient attractor for a certain time period.

First we present a relation between the temperature and the internal energy, i.e. the T-U relation in Fig.1. The temperature T and the internal energy per one particle U of the present system are defined by $T := 2\langle K(t)\rangle/N$ and $U := H/N$, where $K(t)$ is the kinetic energy of the system at time t and $\langle K(t)\rangle$ denotes its time averaging from initial time to t. The internal energy U is conserved in the present microcanonical system.

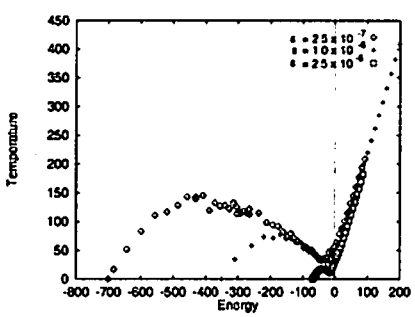


Figure 1: Temperature T vs energy per one particle U in the case of $\epsilon = 2.5 \times 10^{-5}, 1.0 \times 10^{-6}$, and 2.5×10^{-7}. In each T-U curve, there are two critical energy scales; $U_{\text{c-low}}(\epsilon)$ and $U_{\text{c-high}}(\epsilon)$, between which a negative specific heat appears. In this range, there seems to exist an asymptotic line of $T = -0.38U$ in the limit of $\epsilon \to 0$.

From Fig.1, we find two critical energy scales; $U_{\text{c-low}}(\epsilon)$ and $U_{\text{c-high}}(\epsilon)$ in this T-U curve. The energy scale $U_{\text{c-high}}(\epsilon)$ corresponds to a mean gravitational binding energy per one particle, which is estimated as $\left\langle -1/(\sqrt{2}N\sqrt{1-\cos\theta+\epsilon})\right\rangle \times (N(N-1)/2)/N \sim -0.5$, while, $U_{\text{c-low}}(\epsilon)$ strongly depends on a cutoff ϵ, e.g. $U_{\text{c-low}}(\epsilon) \sim -400$ and -40 for $\epsilon = 2.5 \times 10^{-7}$ and $\epsilon = 2.5 \times 10^{-5}$, respectively.

In the energy range between these two energy scale, i.e. $U_{\text{c-low}}(\epsilon) < U < U_{\text{c-high}}(\epsilon)$, the system has a negative specific heat. This existence of a negative specific heat may correspond to an appearance of a

	$U < U_{\text{c-low}}(\epsilon)$	$U_{\text{c-low}}(\epsilon) < U < U_{\text{c-high}}(\epsilon)$	$U_{\text{c-high}}(\epsilon) < U$
system	$\mathcal{C}$ -state	$\mathcal{V}$ -state	$\mathcal{G}$ -state
	$E_i < U_{\text{c-low}}(\epsilon)$	$U_{\text{c-low}}(\epsilon) < E_i < U_{\text{c-high}}(\epsilon)$	$U_{\text{c-high}}(\epsilon) < E_i$
particle	*core* phase	*halo* phase	*gas* phase

Table 1: Classification of "equilibrium" state of system by its total energy U and a particle in one "equilibrium" state by its total energy E_i.

kind of phase transition [11, 13, 14, 17]. Such a phase transition from a gaseous state to a cluster state in fact appears and a variety structure will be realized in the intermediate energy range with a negative specific heat (see below).

Using the above two critical energy scales, $U_{\text{c-low}}(\epsilon)$ and $U_{\text{c-high}}(\epsilon)$, we shall classify the present equilibrium state into three and define three phases in each "equilibrium" state in Table.1. In Table.1, the total energy of i'th particle is defined as

$$E_i = \frac{1}{2}p_i^2 - \sum_{j \neq i}^{N} \frac{1}{\sqrt{2}N\sqrt{1-\cos\theta_{ij}+\epsilon}}. \tag{6}$$

Note that even for the $\mathcal{C}$ -state, there exist a few particles in the *gas* phase. A group of particles either in *core* phase or in *halo* phase is called a cluster in this paper because those make one bound state.

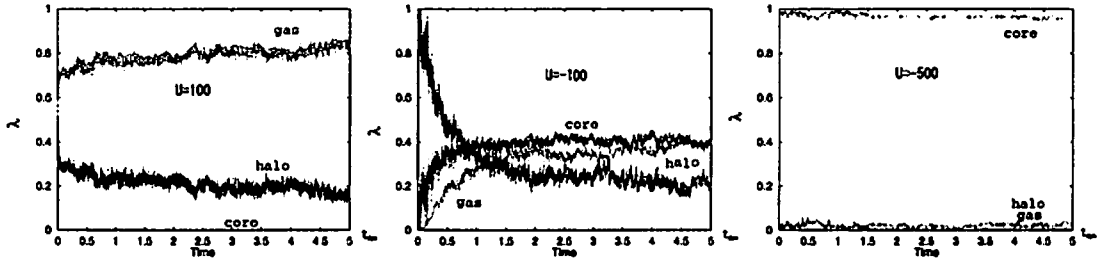


Figure 2: The time evolution of the ratio of each phase; λ_{core}, λ_{halo}, and λ_{gas} in $\mathcal{G}$ -state (left), in $\mathcal{V}$ -state (center), and in $\mathcal{C}$ -state (right) for $\epsilon = 2.5 \times 10^{-7}$. In $\mathcal{G}$ -state, gas particles dominate, although there is still small part of *halo* particles. While, in $\mathcal{C}$ -state, most particles are in *core* phase. In the $\mathcal{V}$ -state, however, many particles experience at least two phases. Some particles wander from one phase to another in three phases.

We describe the relative ratios of particle number in each phase by λ_{core}, λ_{halo}, and λ_{gas}, which satisfies the particle number conservation condition, $\lambda_{core} + \lambda_{halo} + \lambda_{gas} = 1$. We show the time evolution of λ_{core}, λ_{halo}, and λ_{gas} in Fig.2 for the case of $\epsilon = 2.5 \times 10^{-7}$. From these figures, these ratios approach the asymptotic value beyond $t \sim t_{\text{dyn}}$.

4 Particle motions

First we depict the velocity distributions for the clusters both in the $\mathcal{C}$ -state and $\mathcal{V}$ -state. The distribution is almost Gaussian for the $\mathcal{C}$ -state, but that for $\mathcal{V}$ -state is not exactly Gaussian but has an apparent tail. In order to clarify the origin of this tail, we show the distributions of the *halo* and *core* particles separately in Fig.3. We find that the distribution of the *core* particles is always Gaussian with the dispersion $\sigma := \langle v_{core}^2 \rangle^{1/2}$, while the distribution of the *halo* particles is non-Gaussian and has a v^{-2}-tail in Fig. 3.

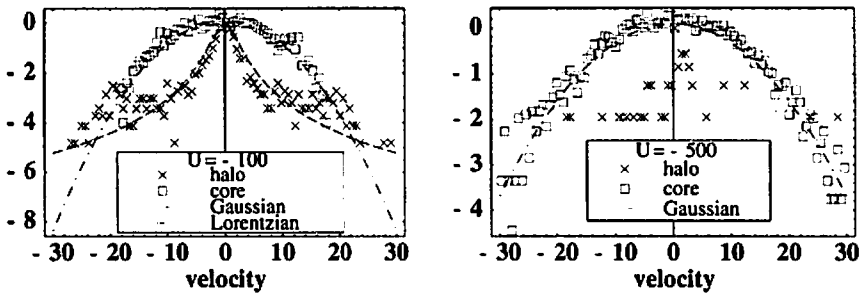


Figure 3: The linear-log plot of velocity distribution of particles in each phase by use of data taken every unit time ($t_{\rm dyn}$) from $1t_{\rm dyn}$ to $5t_{\rm dyn}$ and superposed. For the *core* particles in $\mathcal{V}$ -state and $\mathcal{C}$ -state, the Gaussian distribution with $\sigma = 7.29$ and 10.9 fit well, respectively. On the other hand, for the *halo* particles in $\mathcal{V}$ -state, we find the Lorentzian distribution ($P(v) = s/(\pi(v^2+s^2))$) with $s = 2.2$ is consistent with our numerical result.

Next, we will pay our attention to the recurrent motion of the particles, especially of the *halo* particles and will reveal a universal property in the motion of *halo* particles in $\mathcal{V}$ -state (or even in $\mathcal{C}$ -state).

We show that the particles in the *core* phase move regularly around the center of the cluster, whose period of the oscillation is approximated by the free-fall time determined by the density of the cluster, $t_{\text{cl-dyn}}$. On the other hand, the *halo* particles seem to move around the cluster without any typical time scale and amplitude, as we can see in Fig.4(a).

Enlarging the figure of the *halo* particle motions by ten times larger than the original one (Fig. 4(b)), we find the similar recurrent motions. The same pattern of motions is also found in the 100-times magnified figure (Fig.4(c)), although the lower cutoff effect is already observed in the recurrent time period. The repetition of the similar pattern suggests us a self-similar nested structure.

In order to analyze this behavior more quantitatively, we examine the frequency distribution of the recurrent time $t_{\rm rec} := t_{\rm in} - t_{\rm out}$, that is the time period from the moment $t_{\rm in}$ when one particle leaves the barycenter of the core to the moment $t_{\rm out}$ when it returns to the core and reaches the barycenter again. Here we defined the location of the barycenter of the core as $\theta_{bc} := \sum_{i_c=1}^{N_c} \theta_{i_c}/N_c$, where θ_{i_c} is the location of the i_c'th particle in the *core* phase and N_c is the total number of the core particles at each moment.

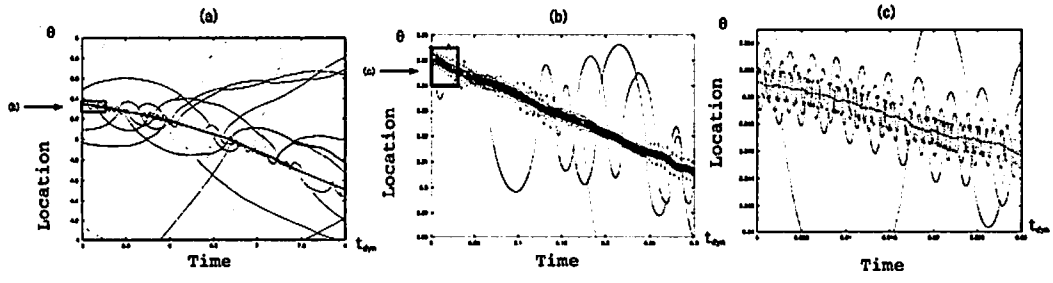


Figure 4: Trajectories of *halo* particles for the time interval (a) $t = (5 \sim 8)\ t_{\rm dyn}$, (b) $t = (5 \sim 5.3)\ t_{\rm dyn}$ and (c) $t = (5 \sim 5.03)\ t_{\rm dyn}$. (b) and (c) are magnified figures of the squared region in (a) and (b), respectively. A self-similar nested structure is found.

In Fig.5, we depict the frequency distribution of recurrent time for the *core* particles and for the *halo* particles. For *core* particles, as we expected, we find a peak in the distribution and a Gaussian-like

fluctuation around its peak. The time scale at this peak position is related with the free-fall time $t_{\text{cl-dyn}}$ defined in (5). On the other hand, for the halo particles, although a peak is found in the distribution, the distribution spreads very widely starting from the peak time interval and a long tail characterized by the scaling-law $f(t_{\text{rec}}) \sim t_{\text{rec}}^{-p} (p \approx 2.0)$ is found. This result is consistent with the motion of *halo* particles shown in Fig.4, i.e. the self-similar nested structure. So we cannot identify the recurrent time scale for the halo particles with any characteristic time period against the case of the *core* particles. This makes it difficult to estimate the relaxation time not only for the halo particles but also for the whole system.

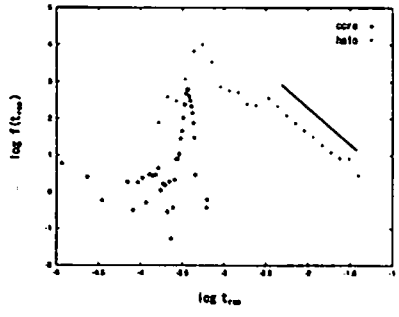


Figure 5: Log-log plots of frequency distribution $f(t_{\text{rec}})$ for *core* particles (cross) [(i)] and for *halo* particles (square) [(ii)], where t_{rec} is the recurrent time of each particle. The peak appears at $t_{\text{rec}} \sim 10^{-3.5} t_{\text{dyn}} \approx 0.0003 t_{\text{dyn}}$ and at $t_{\text{rec}} \sim 10^{-3.15} t_{\text{dyn}} \approx 0.0007 t_{\text{dyn}}$ in (i) and (ii), respectively, while the constant slope is seen only in (ii). The straight solid line represents the slope of -2.0.

5 Fractal analysis

In the previous section, we found that the velocity distribution of *halo* particles is non-Gaussian with the v^{-2}-tail and each *halo* particle moves recurrently without any typical time scale of oscillation. This recurrent motion also shows up the self-similar nested structure. We then expect that the *halo* particles may lead to the fractal structure in the particle distribution as well. In this section, we shall calculate the fractal dimension of the distribution with the box-counting method [19].

For this purpose, we first divide the total one-dimensional periodic distance $2\pi R$ by many boxes with the size ℓ and count the number $N(\ell)$ of the box which contains at least one *halo* particle. Then we define the following quantity.

$$D(\ell) := -\left\langle \frac{\partial(\log N(\ell))}{\partial(\log \ell)} \right\rangle_{\text{ens}}, \tag{7}$$

where, $\langle \cdots \rangle_{\text{ens}}$ refers to the ensemble average over the samples with the same energy U. We have used the data of particle distributions on 10 time slices which divide the period between $t = 2000 t_{\text{cl-dyn}}$ and $5000 t_{\text{cl-dyn}}$ by equal time interval. In this period, the phase ratios λ_{core}, λ_{halo}, and λ_{gas} almost attain to constant values.

If fractal structure exists, $D(\ell)$ will become constant for a finite range of $\ell_1 \le \ell \le \ell_2$. As for the bounds ℓ_1, ℓ_2, we have theoretical limits. The averaged distance between *gas* particles is estimated as $d_{\text{gas}}(\epsilon) \approx \frac{2\pi R}{N_{\text{gas}}(\epsilon)}$, where $N_{\text{gas}}(\epsilon)$ is the number of the *gas* particles which move independently of a cluster. If the box size ℓ is larger than $d_{\text{gas}}(\epsilon)$, we expect $D \approx 1.0$, because almost all of the boxes contain at least one particle. On the other hand, if ℓ is smaller than the cutoff distance ξ, we may loose the information about gravitational force, because in such a small scale, the *core* particles dominate and the interaction is no longer gravitational. Then the fractal structure caused by gravitational interaction, if it exists, will be found in some finite range of $\ell_1 \le \ell \le \ell_2$ with $\ell_1 > \xi$ and $\ell_2 < d_{\text{gas}}(\epsilon)$. In fact, in the case with $\epsilon = 2.5 \times 10^{-7}$ and $U = -100$, we can see the wide range of flat plateau corresponding to the small value of fractal dimension in the range of $10^{-3} R < \ell < 10^{-1} R$.

The energy and cutoff dependencies of fractal dimension D are shown in Fig.6. Here we show the fractal dimension of *halo* particles in the energy range of both $\mathcal{C}$ -state and $\mathcal{V}$ -state. We find that the fractal dimension D is almost independent of the energy U, and as the cutoff ϵ goes down, D gets lower but the range where D is constant becomes wider. On the other hand, in the limit of $\epsilon \to 1$, the fractal disappears, since the range of $\xi < \ell < d_{\rm gas}(\epsilon)$ becomes too narrow to exhibit the fractal structure.

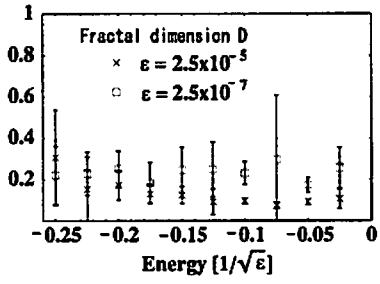


Figure 6: The fractal dimension D vs the energy per particle U. For the convenience, U is rescaled by the $\epsilon^{-1/2}$. The case of $\epsilon = 1.0 \times 10^{-5}$ (cross with error bars) and the case of $\epsilon = 1.0 \times 10^{-7}$ (square with error bars) are plotted. It seems that the dimension D does not depend on U. Since the number of halo particles is small in the low energy, the error becomes large.

6 Summary and discussion

In this paper we have studied the SGR model which is a one-dimensional toy model of a self-gravitating system. Although the system is one-dimensional, the force acting on the particles constrained on a circular ring is the conventional 3D gravitational interaction. As a result, in the microcanonical system, a negative specific heat appears and a phase transition occurs at low temperature. We introduce a short-range cutoff ϵ so as to prevent the system from collapsing.

Using this SGR model, we have examined characteristic properties of a self-gravitating system such as a fractal structure and non-Gaussian velocity distribution, and analyzed the origin of these characteristic properties from the viewpoint of the gravitational phase transition.

We find that a virialized state with negative specific heat appears at the intermediate energy scale, where a phase transition occurs and a cluster is formed. Classifying the particles in each state into three phases (*core* , *halo* , and *gas* phases), we can characterize this virialized state by non-thermal property of the particles in *halo* phase. We have then analyzed the motion of *halo* particles and its fractal structure. We find that the velocity distribution of a cluster in a virialized state is non-Gaussian with the v^{-2} tail. The *halo* particles also move recurrently without any typical time scale of oscillation. We also find the fractal structure of the *halo* particles in the $\mathcal{V}$ -state with a box-counting method. The fractal dimension is rather small and depends on the cutoff. In the gravitational limit, $\epsilon \to 0$, the energy range with negative specific heat becomes broad, and the fractal structure becomes more prominent, although the dimension decreases.

In the limit of the real 3D gravitational force ($\epsilon \to 0$), we find many universal features, the state with negative specific heat, non-Gaussian velocity distribution, and the fractal structure of the *halo* particles. Those are likely to be independent of the temperature or the energy of the system. The first order phase transition occurs even in the system with short range interactions as is seen in [13]. However, the above universal character could be the peculiar one which distinguishes the systems with the long-range interactions such as the gravitational force from those with the short range interactions. In two or three dimensional cases, it would be easy to compare the case with short-range interactions to the case with long-range ones, since the first order phase transition occurs in the former case. This would be our future work.

7 acknowledgments

We would like to thank A. Nakamichi, I. Joichi, K. Nakamura and M. Hotta for useful discussions and comments. This work was supported partially by a Grant-in-Aid for Scientific Research Fund of the Ministry of Education, Science and Culture (Specially Promoted Research No. 08102010), and by the Waseda University Grant for Special Research Projects.

References

[1] W.H. Zurek, P.J. Quinn, J.K. Salmon, and M.S. Warren, Astrophys. J. **431** (1994) 559.

[2] F. Sylos Labini, M. Montuori and L. Pietronero, Phys. Rep. **293** (1998) 61.

[3] G. de Vaucouleurs, Science **167** (1970) 1203.

[4] R.B. Larson, Mon. Not. R. astr. Soc. **194** (1981) 809.

[5] E. Falgarone, T.G. Phillips and C.K. Walker, Astrophys. J. **378** (1991) 186.

[6] T. Kambe and P.G. Drazin, *Fluid Dynamics Stability and Turbulence* (University of Tokyo Press, 1998).

[7] H.J. de Vega, N. Sánchez, and F. Combes, Phys. Rev. **D54** (1996) 6008.

[8] V. A. Antonov, Vest. leningr. gos. univ. **7** (1962) 135.

[9] D. Lynden-Bell and R. Wood, Mon. Not. R. astr. Soc. **138** (1968) 495.

[10] I. Hachisu and D. Sugimoto, Prog. Theor. Phys. **60** (1978) 123.

[11] E. B. Aronson and C. J. Hansen, Astrophys. J. **177** (1972) 145.

[12] T. Padmanabhan, Phys. Rep. **188** (1990) 285.

[13] P. Hertel and W. Thirring, Annals of physic. **63** (1971) 520.

[14] A. Compagner, C. Bruin, and A. Roelse, Phys. Rev. **A39**(1989) 5989.

[15] G.L. Camm, Mon. Not. R. astr. Soc. **110** (1950) 305.

[16] T. Tsuchiya, N. Gouda, and T. Konishi, Phys. Rev. **E53** (1996) 2210.

[17] Y. Aizawa, K. Sato, and K. Ito, Prog. Theor. Phys. **103** (2000) 519.

[18] M. Suzuki, Phys. Lett. A **146** (1990) 319.

[19] K. Falconer, *Fractal Geometry* (John Wiley & Sons, 1990).

Cosmic Inversion

Makoto Matsumiya,[1] Misao Sasaki[2] and Jun'ichi Yokoyama[3]

Department of Earth and Space Science, Graduate School of Science, Osaka University, Toyonaka, 560-0043, Japan

Abstract

We investigate the possibility of reconstructing the initial spectrum of metric fluctuations from the cosmic microwave background(CMB) anisotropy analytically. We show that the initial spectrum can be determined as a solution of a second-order ordinary differential equation. We find that this method works quite well in case of power law spectrum.

1 Introduction

Determining the primordial spectrum of the density fluctuations is one of the important problems of modern cosmology in understanding the structure formation of the Universe. The cosmic microwave background (CMB) anisotropy directly provides us with information on the primordial fluctuations [3, 11]. Although the angular power spectrum of the CMB has been calculated for various initial spectra and many experiments measured them [7, 10], the shape of the initial spectrum is still controversial. From the point of view that the initial spectrum should be determined without theoretical believes [2], it would be an interesting approach to extract information on the initial spectrum from observational data of the CMB.

In this paper, we shall present an analytic method to solve the problem of reconstructing the initial spectrum of metric fluctuations $P(k)$, which is related to density fluctuations through Poisson equation, in the Einstein-de Sitter universe. This is an inverse problem from two dimensional space to three dimensional space, but if we consider the angular correlation function $C(r)$, the information on $P(k)$ is given as Fourier sine transform of $C(r)$. This is a trivial issue if only the ordinary Sachs-Wolfe effect (OSW) [9] is considered. For practical use, however, we must further take account of other effects which make the problem extremely difficult, and to do so a new method is required. Thus, we solve the problem when intrinsic photon fluctuations, the OSW effect and the Doppler effect are considered. We fix cosmological parameters and do not study the dependence of $P(k)$ upon them, since our primary concern at the moment is on the possibility of solving the problem and making a formulation for it.

2 The Basic Equations

The evolution of the temperature perturbations $\Theta(\mathbf{x}, \gamma, \eta)$ in the Newtonian representation is governed by the first-order Boltzmann equation [6],

$$\frac{d}{d\eta}[\Theta + \Psi](\mathbf{x}, \gamma, \eta) = \dot{\Psi} - \dot{\Phi} + \dot{\tau}(\Theta_0 - \Theta + \gamma^i v_i + \frac{1}{16}\gamma^i\gamma^j\Pi_{ij}), \tag{1}$$

where v_i is the baryon velocity, Π_{ij} is the anisotropic stress perturbation , γ^i represents the direction cosine for the photon momentum, $\dot{\tau}$ is Thomson differential optical depth and overdot represents derivative with respect to the conformal time $\eta = \int dt/a$. Here Θ_0 is the monopole moment of Θ, Ψ and Φ are the gauge-invariant metric perturbations, which represent gravitational potential and curvature perturbation in

[1] E-mail:matumiya@vega.ess.sci.osaka-u.ac.jp
[2] E-mail:misao@vega.ess.sci.osaka-u.ac.jp
[3] E-mail:yokoyama@vega.ess.sci.osaka-u.ac.jp

Newtonian gauge, respectively. Ψ and Φ are related to the density perturbation through the generalized Poisson equation [5] as

$$\begin{aligned}(k^2 - 3K)\Phi &= 4\pi G a^2 \rho \Delta_T, \\ \Phi + \Psi &= -\frac{8\pi G}{k^2} a^2 p \Pi,\end{aligned} \tag{2}$$

where Δ_T is the total matter density perturbation in its rest frame, ρ and p are the density and the pressure of the matter respectively. If anisotropic stress is negligible, we find $\Psi = -\Phi$ from Eq.(2).

Before recombination, photons are tightly coupled with baryons through Thomson scattering. With the tight coupling approximation, we can expand Eq.(1) together with continuity and Euler equations of baryons in the Thomson scattering time [8]. To first-order, the dynamics of photon-baryon fluid is described by the following simple equation for each Fourier mode [3].

$$\ddot{\Theta}_0 + \frac{\dot{a}}{a}\frac{R}{1+R}\dot{\Theta}_0 + k^2 c_s^2 \Theta_0 = F(\eta), \tag{3}$$

where $F(\eta)$ is the driving term due to gravitational potential,

$$F(\eta) = -\ddot{\Phi} - \frac{\dot{a}}{a}\frac{R}{1+R}\dot{\Psi} - \frac{k^2}{3}\Psi, \tag{4}$$

c_s is the sound speed of the fluid and $R \equiv 3\rho_b/4\rho_\gamma$. The dipole moment can be obtained from the continuity equation,

$$k\Theta_1 = -3(\dot{\Theta}_0 + \dot{\Psi}). \tag{5}$$

The equations given above are only applicable to scales larger than the photon diffusion scale at that time.

Now we adopt some approximations to solve Eq.(3) and Eq.(5); (i) the gravitational potential Ψ is always independent of time *i.e.* $\dot{\Psi} = \dot{\Phi} = 0$, (ii) $R = 0$, (iii) the initial condition of the perturbations is adiabatic. Assumptions (i) and (ii) are valid for the scale larger than first Doppler peak, while (iii) is a natural consequence of inflationary cosmology. Then, the solutions of Eq.(3) and Eq.(5) are given by

$$\begin{aligned}[\Theta_0 + \Psi](\eta) &= \frac{1}{3}\Psi \cos(k c_s \eta), \\ \Theta_1(\eta) &= \frac{c_s}{3}\Psi \sin(k c_s \eta),\end{aligned} \tag{6}$$

where $c_s = \frac{1}{\sqrt{3}}$ in this case.

After the last scattering, photons free-stream to the present creating temperature anisotropies which is observed on the sky today. If the last scattering surface is very thin, we can drop off the collisional term on the right hand side of Eq.(1). In this case, with the assumption $\dot{\Psi} = \dot{\Phi} = 0$, the solution of Eq.(1) is expressed for each Fourier mode and angular mode [4] as

$$\Theta_l(\eta_0) = [\Theta_0 + \Psi](\eta_*)(2l+1) j_l(kd) + \Theta_1(2l+1) j_l'(kd), \tag{7}$$

where suffices $*$ and 0 denote the epoch of last scattering and the present, respectively, $j_l(z)$ is a spherical Bessel function of the l-th order and $d = \eta_0 - \eta_*$. Eq.(7) shows that there are three sources of temperature anisotropy, namely, intrinsic temperature variation, ordinary Sachs-Wolfe effect due to the static gravitational potential and Doppler effect due to bulk velocity of the fluid.

Then, the angular correlation function is generally written [1] as

$$C(\theta) = \langle \Theta(\eta, x, \gamma)\Theta(\eta, x, \gamma') \rangle = \sum_l \frac{2l+1}{4\pi} C_l P_l(\cos\theta),$$

$$\frac{2l+1}{4\pi} C_l = \frac{1}{2\pi^2}\int_0^\infty dk\, k^2 \frac{|\Theta_l(\eta, k)|^2}{2l+1}. \tag{8}$$

Here C_l is the angular power spectrum and $P_l(z)$ is Legendre polinomial of the l-th order.

Using the solutions given by Eq.(6), we obtain the expression of the angular power spectrum

$$C_l = \frac{2}{9\pi}\int_0^\infty dk\, k^2 P(k)\left[\cos^2 kr_* j_l^2(kd) + 2k\eta_* \frac{\sin 2kr_*}{2kr_*} j_l(kd) j_l'(kd) + (k\eta_*)^2 \left(\frac{\sin kr_*}{kr_*}\right)^2 j_l'^2(kd)\right], \quad (9)$$

where $r_* = c_s\eta$ and $P(k)$ is the power spectrum of $\Psi(\eta, \mathbf{x})$. Using the relation

$$\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta) j_l^2(kd) = \frac{\sin kr}{kr}, \quad (10)$$

where r is defined by

$$r = 2d\sin\frac{\theta}{2}, \quad (11)$$

and the angular correlation function is given by the following expression

$$\begin{aligned} C(r) &= \frac{1}{18\pi^2}\int_0^\infty dk\, k^2 P(k)\left[\cos^2 kr_* - r\frac{\eta_*}{d}\frac{\sin 2kr_*}{2kr_*}\frac{\partial}{\partial r} + \frac{1}{4}r^2\left(\frac{\eta_*}{d}\right)^2\left(\frac{\sin kr_*}{kr_*}\right)^2\left(\frac{\partial}{\partial r}\right)^2\right]\frac{\sin kr}{kr} \\ &= \frac{1}{18\pi^2}\int_0^\infty dk\, k^2 P(k)\left\{\left[\frac{1}{kr}\cos^2 kr_* - \frac{\eta_*}{d}\frac{1}{kr}\frac{\sin 2kr_*}{2kr_*} + \left(\frac{\eta_*}{d}\right)^2\left(\frac{\sin kr_*}{kr_*}\right)^2\left(\frac{1}{2kr} - \frac{kr}{4}\right)\right]\sin kr \right. \\ &\quad \left. + \left[\frac{\eta_*}{d}\frac{\sin 2kr_*}{2kr_*} - \frac{1}{2}\left(\frac{\eta_*}{d}\right)^2\left(\frac{\sin kr_*}{kr_*}\right)^2\right]\cos kr\right\}. \end{aligned} \quad (12)$$

Here, r denotes the distance between two points on the last scattering surface.

3 Reconstruction of $P(k)$

We subtract the monopole contribution from $C(r)$ and multiply Eq.(12) by r to define

$$\tilde{C}(r) = \tilde{C}(r) = rC(r) - rC_0P_0, \quad (13)$$

$$\begin{aligned} &= \frac{1}{18\pi^2}\int_0^\infty dk\, k^2 P(k)\left\{\left[\frac{1}{k}\cos^2 kr_* - \frac{\eta_*}{d}\frac{1}{k}\frac{\sin 2kr_*}{2kr_*} + \left(\frac{\eta_*}{d}\right)^2\left(\frac{\sin kr_*}{kr_*}\right)^2\left(\frac{1}{2k} - \frac{kr^2}{4}\right)\right]\sin kr \right. \\ &\quad \left. + \left[\frac{\eta_*}{d}\frac{\sin 2kr_*}{2kr_*} - \frac{1}{2}\left(\frac{\eta_*}{d}\right)^2\left(\frac{\sin kr_*}{kr_*}\right)^2\right] r\cos kr\right\} - \frac{1}{4\pi}rC_0P_0. \end{aligned} \quad (14)$$

Differentiating twice, we obtain

$$\begin{aligned} \frac{\partial^2}{\partial r^2}\tilde{C}(r) &= \frac{1}{36\pi^2}\int_0^\infty dk\, P(k)\left\{-2\left(\cos^2 kr_* - x\frac{\sin 2kr_*}{2kr_*}\right)k^3\sin kr \right. \\ &\quad \left. - \left[2x\frac{\sin 2kr_*}{2kr_*} + x^2\left(\frac{\sin kr_*}{kr_*}\right)^2\right]k^4 r\cos kr + \frac{x^2}{2}\left(\frac{\sin kr_*}{kr_*}\right)^2 k^5 r^2\sin kr\right\} \end{aligned} \quad (15)$$

where $x \equiv \eta_*/d \simeq 1/30$. After integrating by parts twice in order to eliminate r on the right-hand side of Eq.(15), assuming that the boundary terms vanish, we find

$$\begin{aligned} \frac{\partial^2}{\partial r^2}\tilde{C}(r) &= \frac{1}{36\pi^2}\int_0^\infty \left\{-\frac{1}{4}x^2\frac{1-\cos 2kr_*}{(kr_*)^2}k^5 P''(k) + \left[(x-x^2)\frac{\sin 2kr_*}{kr_*} - x^2\frac{1-\cos 2kr_*}{(kr_*)^2}\right]k^4 P'(k)\right. \\ &\quad \left. + \left[-1-(x-1)^2\cos 2kr_* + 2(x-x^2)\frac{\sin 2kr_*}{kr_*} - \frac{1}{2}x^2\frac{1-\cos 2kr_*}{(kr_*)^2}\right]k^3 P(k)\right\}\sin kr\, dr. \end{aligned} \quad (16)$$

From Fourier sine formula, we obtain

$$-\frac{1}{4}x^2\frac{1-\cos 2kr_*}{(kr_*)^2}k^5P''(k)+\left[(x-x^2)\frac{\sin 2kr_*}{kr_*}-x^2\frac{1-\cos 2kr_*}{(kr_*)^2}\right]k^4P'(k)$$
$$+\left[-1-(x-1)^2\cos 2kr_*+2(x-x^2)\frac{\sin 2kr_*}{kr_*}-\frac{1}{2}x^2\frac{1-\cos 2kr_*}{(kr_*)^2}\right]k^3P(k)=\frac{2}{\pi}\int_0^\infty\frac{\partial^2}{\partial r^2}\tilde{C}(r)\sin kr\,dr \tag{17}$$

where the factor $1/36\pi^2$ is included in $P(k)$. Hence, the problem of determining $P(k)$ is reduced to solving the second-order ordinary differential equation of $P(k)$.

Let us solve Eq.(17) under the assumption that the initial spectrum is given by scale invariant spectrum *i.e.* $P(k)=Ak^{-3}$. In this case, the left hand side of Eq.(17) can be obtained analytically from Eq.(15),

$$\frac{2}{\pi}\int_0^\infty\frac{\partial^2}{\partial r^2}\tilde{C}(r)\sin krdr=A\left(-1-(x-1)^2\cos 2kr_*-\frac{1}{2}x^2\frac{1-\cos 2kr_*}{(kr_*)^2}-(x-x^2)\frac{\sin 2kr_*}{kr_*}\right) \tag{18}$$

Since Eq.(17) has a number of singular points, it is difficult to solve it numerically. Hence we propose a perturbative method.

Note that $x=\eta_*/d\ll 1$. It is therefore appropriate to expand Eq.(17) in x if we assume that $P(k), kP'(k), k^2P''(k), S(k)$ have the same order values. However, the zeroth order term $(-1-\cos 2kr_*)P(k)$ vanishes at $kr_*=(n+\frac{1}{2})\pi$, thus we take $(-1-\cos 2kr_*+2x\cos 2kr_*)P(k)$ as zeroth order term. As we show in Fig.1, the solution obtained by this technique reproduces the initial spectrum well except for the points where the iteration becomes inappropriate.

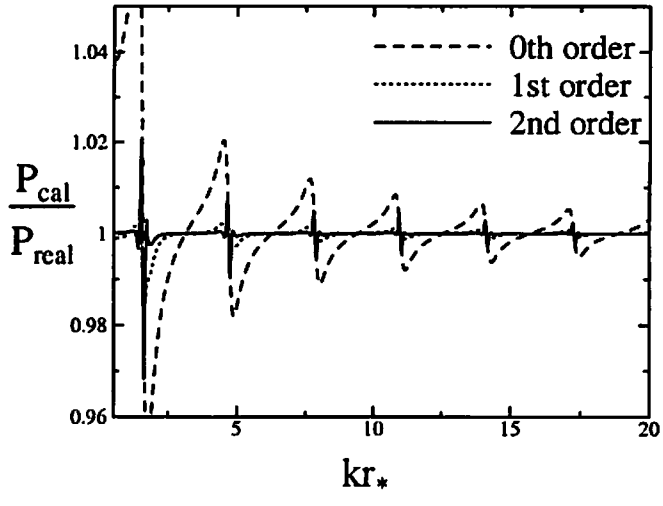


Figure 1: Solutions at each order of x. P_{cal} is an obtained solution and P_{real} is the scale invariant spectrum. The solution is getting unstable around the points where $kr_*=(n+\frac{1}{2})\pi$.

4 Conclusion

We have considered the problem of reconstruction of $P(k)$ when C_l is given. We have shown that mathematically it corresponds to the Fourier sine transform of $C(r)$, and $P(k)$ can be reconstructed as a solution of the second-order differential equation. The differential equation can be solved perturbatively

with a small expansion parameter x. Although we have demonstrated the power of this method only in case of the scale invariant spectrum in this article, we can verify the method for various types of the initial spectrum [12].

References

[1] Bond, J. R. & Efstathiou, G. Mon. Not. R. Astron. Soc. 226, 655 (1987).

[2] Harrison, H.R. PRD 1, 2726 (1970).

[3] Hu, W. & Sugiyama, N. ApJ 444, 489 (1995).

[4] Hu, W. & Sugiyama, N. PRD 51, 2599 (1995).

[5] Kodama, H. & Sasaki, M. Prog. Theor. Phys. Supp. 78,1 (1984).

[6] Kodama, H. & Sasaki, M. Int. J. Mod. Phys. A1, 265 (1986).

[7] Mauskopf, P. D.*et al.* ApJ 536, 59 (2000).

[8] Peebles, P. J. E. & Yu, J.T. ApJ 162, 815 (1970).

[9] Sachs, R. K. & Wolfe, A. M. ApJ 147, 73 (1967).

[10] Smoot, G. *et al.* ApJL 396, L1 (1992).

[11] White, M., Scott, D. & Silk, J. ARA&A 32, 319 (1994).

[12] Matsumiya, M. Master Thesis, Osaka University (2001).

Cosmic Microwave Background in Closed Multiply Connected Universes

Kaiki Taro Inoue,[1]

Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan

Abstract

We have investigated the cosmic microwave background (CMB) anisotropy in closed multiply connected universes (flat and hyperbolic) with low matter density. We show that the COBE constraints on these low matter density models with non-trivial topology are less stringent since a large amount of CMB anisotropy on large angular scales can be produced due to the decay of the gravitational potential at late time.

1 Introduction

For a long time, cosmologists have assumed the simply connectivity of the spatial hypersurface of the universe. If it is the case, the topology of closed 3-spaces is limited to that of a 3-sphere if Poincaré's conjecture is correct. However, there is no particular reason for assuming the simply connectivity since the Einstein equations do not specify the boundary conditions. If we allow the spatial hypersurface being multiply connected then the spatial geometry of closed models can be flat or hyperbolic as well. Then we should be able to observe the imprint of the "finiteness" of the spatial geometry if it is multiply connected on scales of the order of the particle horizon or less, in other words, if we live in a "small universe".

However, it has been claimed by several authors that the "small universes" have already been ruled out observationally. For a flat 3-torus model without the cosmological constant, the COBE data constrains the topological identification scale L (the minimum length of periodic geodesics) to be $L \geq 0.4 \times 2R_*$ where R_* is the comoving radius of the last scattering surface [1, 2, 3, 4]. The suppression of the fluctuations on scales beyond L leads to a decrease of the angular power spectrum C_l of the CMB temperature fluctuations on large angular scales.

In contrast, for low matter density models, the constraint could be considerably less stringent since a bulk of large-angle CMB fluctuations can be produced by the so-called (late) integrated Sachs-Wolfe effect [5] which is the gravitational blueshift effect of the free streaming photons caused by the decay of the gravitational potential[6]. If the background geometry is either flat or hyperbolic, then the angular size of a fluctuation becomes small as it approaches to the observation point. Although fluctuations beyond the size of the fundamental domain are suppressed, no significant suppression in large-angle power occurs if they are produced at place well after the last scattering. In other words, large-angle fluctuations can be generated when the fluctuations enter the topological identification scale at late time.

2 Closed Flat Models

Let us first consider closed models in which the spatial geometry is represented as a flat 3-torus obtained by gluing the opposite faces of a cube with sides L by three translations. Then the wave numbers of the square-integrable eigenmodes of the Laplacian are restricted to the discrete values $k_i = 2\pi n_i/L$, $(i = 1, 2, 3)$ where n_i's run over all integers. Assuming adiabatic initial perturbation, the angular power spectrum is written as

$$\begin{aligned} C_l &= \sum_{\mathbf{k}\neq 0} \frac{8\pi^3 \mathcal{P}_\Phi(k) F_{kl}^2}{k^3 L^3}, \\ F_{kl} &= \frac{1}{3}\Phi(\eta_*) j_l(k(\eta_o - \eta_*)) + 2\int_{\eta_*}^{\eta_o} d\eta \, \frac{d\Phi}{d\eta} j_l(k(\eta_o - \eta)), \end{aligned} \tag{1}$$

[1] E-mail:tinoue@yukawa.kyoto-u.ac.jp

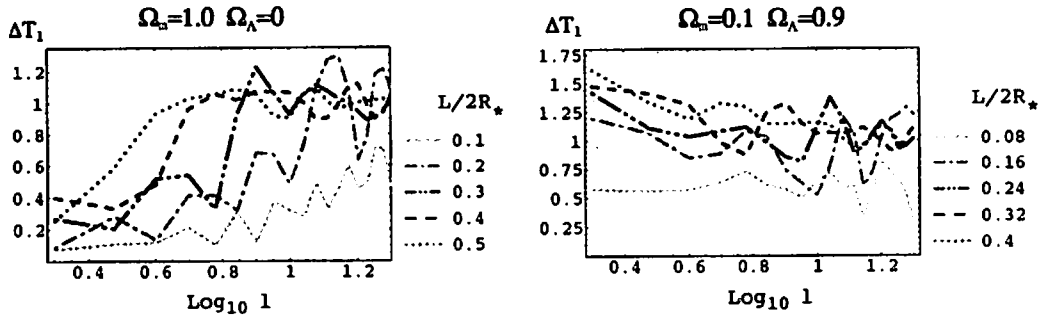


Figure 1: Suppression in large-angle power for 3-torus models with or without the cosmological constant. The significant large-angle suppression in $\Delta T_l \equiv \sqrt{l(l+1)C_l/(2\pi)}$ (all the plotted values are normalized by ΔT_{20} with infinite volume) occurs for the "standard" 3-torus model $(\Omega_m, \Omega_\Lambda) = (1.0, 0)$ at $l < l_{cut} \sim 2\pi R_*/L - 1$ while such prominent suppression is *not* observed for the 3-torus model with $(\Omega_m, \Omega_\Lambda)$=(0.1,0.9).

where η_* and η_0 correspond to the last scattering and the present conformal time, $\mathcal{P}_\Phi(k)$ is the initial power spectrum for the Newtonian curvature perturbation Φ and $k \equiv \sqrt{k_1^2 + k_2^2 + k_3^2}$. From now on we assume the scale-invariant Harrison-Zel'dovich spectrum ($\mathcal{P}_\Phi(k) = const.$) as the initial power spectrum. The angular scale that gives the suppression scale l_{cut} is determined by the lowest eigenmode on the last scattering surface. The oscillation scale for $l < l_{cut}$ is also determined by the first eigenmode corresponding to the contribution from the ordinary Sachs-Wolfe effect and the integrated Sachs-Wolfe effect. On smaller angular scales $l > l_{cut}$, each peak in the power corresponds to the fluctuation scale of the second and higher eigenmodes at the last scattering. This behavior is analogous to the acoustic oscillation where the oscillation scale is determined by the sound horizon at the last scattering. As shown in figure 1, the angular power for a model with $(\Omega_m, \Omega_\Lambda) = (0.1, 0.9)$ is jagged in l but strong suppression is not observed for even very small models. Surprisingly, in low matter density models, the slight excess power due to the integrated Sachs-Wolfe effect is cancelled out by the moderate suppression owing to the mode-cutoff which leads to a nearly flat spectrum. However, as observed in the "standard" 3-torus model, the power spectra have prominent oscillating features.

We have carried out Bayesian analyses using the COBE-DMR 4-year data and obtained the constraints in the size of the fundamental domain $L \geq 1.6 H_0^{-1}$ and $L \geq 2.2 H_0^{-1}$ for the "standard" 3-torus model $(\Omega_m, \Omega_\Lambda) = (1.0, 0)$ and the 3-torus model with low matter density $(\Omega_m, \Omega_\Lambda) = (0.1, 0.9)$, respectively. The maximum number N of images of the cell within the observable region at present is 8 and 49 for the former and the latter, respectively[7].

3 Closed Hyperbolic Models

Next we consider closed hyperbolic models with small volume. The interesting property of closed (compact) hyperbolic manifolds is the existence of the lower bound for the volume $V > 0.16668...$ (in unit of cube of curvature radius). If the creation of the universe with smaller volume is more likely then it gives the reason why the topological identification scale is comparable to the present horizon scale. For flat models, there is no reason for the "coincidence" in the scale since one can choose the volume arbitrarily. The smallest known one is called the Weeks manifold with volume 0.94 (in unit of cube of curvature radius). Today, the number of known examples of closed hyperbolic manifolds is more than 10000 which have been stored in a computer program "SnapPea" by J. Weeks[8].

In fact, we have found that the suppression in the angular power owing to the mode-cutoff for closed hyperbolic models is also very weak (figure 2). As is the case for flat 3-torus models, the suppression

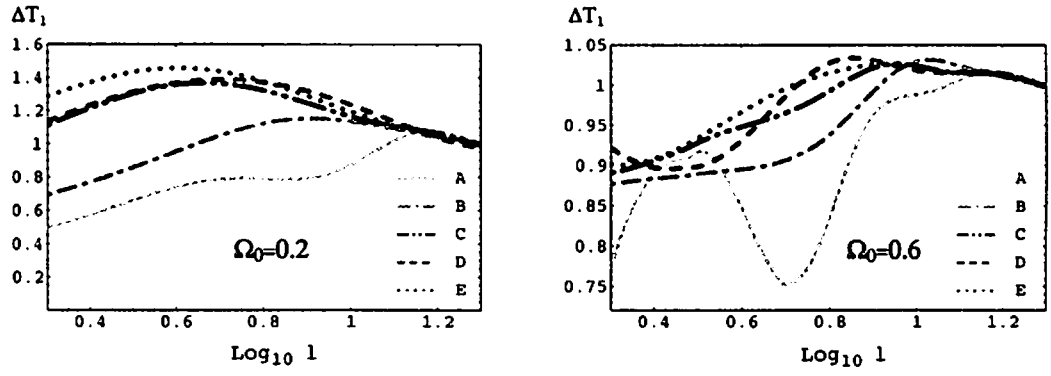


Figure 2: Suppression in large-angle power $\Delta T_l \equiv \sqrt{l(l+1)C_l/(2\pi)}$ for five closed hyperbolic models (name,volume)=A:(m003(3,-1),0.94), B:(m010(-1,3),1.9), C:(m082(-2,3),2.9), D:(m288(-5,1),3.9) and E:(s873(-4,1), 4.9).

scale l_{cut} is determined by the lowest eigenmode on the last scattering surface. Beyond the scale l_{cut} the ordinary Sachs-Wolfe contribution is strongly suppressed as in compact flat models without the Λ term. The weak suppressions imply the significant contribution from the integrated Sachs-Wolfe effect on large-angle scales $l > l_{cut}$. Because the curvature dominant epoch comes earlier in time than the Λ dominant epoch, the contribution from the integrated Sachs-Wolfe effect is much greater than that for flat Λ models. Below the scale l_{cut}, a considerable amount of contribution comes from the second lowest eigenmode. As the number of modes which contribute to C_l grows, C_l converges to the value for the infinite counterpart. The obtained constraint for the Weeks model (volume=0.94) is $\Omega_0 \geq 0.1$ and the maximum expected number of copies of the fundamental domain inside the present observable region is approximately 1560. Note that the obtained result agrees with other recent works[12, 13, 14]. Furthermore, for some closed hyperbolic models, it is found that the computed angular powers give a much better fit to the COBE data since the quadrapole is very low.

However, one might argue that the constraint using only the power spectrum is not sufficient since it contains only isotropic information for 2-point correlations[9]. In fact there is a gap between the likelihood using only the power spectrum and that using full covariance elements for the 3-torus models[7]. For globally anisotropic models, the fluctuations are anisotropic Gaussian for a given axis but are non-Gaussian if the likelihoods are marginalized over the axis. For locally closed Friedmann-Robertson-Walker models, the skewness is zero but the kurtosis is non-zero assuming that the initial fluctuations are Gaussian. Another feature is the correlation between the expansion coefficient a_{lm}'s of the temperature fluctuations in the sky which are independent random numbers for the standard infinite counterparts in which the initial fluctuations are homogeneous and isotropic Gaussian. In the case of flat 3-torus models, a_{lm}'s are written in terms of spherical harmonics $Y_{lm}(\vec{k})$ which correspond to the expansion coefficients of the eigenmodes in the 3-torus in terms of the eigenmodes of the 3-dimensional Euclidean space. Therefore, if the number of the term in the sum that gives a_{lm} is small then a_{lm}'s are no longer independent. This is the main reason for the gap in the likelihoods for the 3-torus models. In contrast, for closed hyperbolic models, the expansion coefficients of the eigenmodes are well approximated by Gaussian pseudo random numbers [10, 11]. Therefore for homogeneous ensembles, a_{lm} can be described as independent "random" numbers although they are non Gaussian (since they are written in terms of a sum of products of two independent Gaussian numbers). This implies that the value of the likelihood is highly *dependent on the place* of the observer in the manifold. In order to constrain closed hyperbolic models which are globally inhomogeneous, one should calculate the likelihood everywhere in the manifold. If one marginalizes the likelihood over the orientation (axis) and the position of the observer then a_{lm}'s become independent non-Gaussian random numbers. Thus we expect that the constraints using only the power spectrum give

better estimates for closed hyperbolic models than closed flat models in which the correlations in a_{lm}'s are prominent.

4 Conclusion

We have investigated the CMB anisotropy in closed multiply connected universes (flat and hyperbolic) with low matter density. We have seen that the COBE constraints for these models are less stringent compared with the simplest "standard" 3-torus model ($\Omega_m = 1.0$). On the other hand recent observations of distant supernova Ia [15, 16] and of the CMB on small angular scales [17, 18] imply that our universe is nearly flat with the cosmological constant (or "quintessence", X-matter etc.) which dominates the present universe. It seems that the closed hyperbolic models are ruled out but it is still not conclusive. If one includes the cosmological constant for a fixed curvature radius, the comoving radius of the last scattering surface in unit of curvature radius becomes large. Therefore the observable effects of the non-trivial topology become much prominent. For instance, the number N of copies of the fundamental domains inside the observable region at present is approximately 28 for the Weeks model with $\Omega_\Lambda = 0.6$ and $\Omega_m = 0.2$ whereas $N = 4$ if $\Omega_\Lambda = 0$ and $\Omega_m = 0.8$ If one allows hyperbolic orbifold models then the number can be increased as much as $N = 52$ for $\Omega_\Lambda = 0.75$ and $\Omega_m = 0.2$ since the volume of the smallest orbifold (arithmetic) is very small (=0.039). At least at the classical level it seems that there is no reason to exclude any orbifold models although they have singular points where the curvature diverges.

References

[1] I.Y. Sokolov, JETP Lett. **57** (1993) 617.

[2] A.A. Starobinsky, JETP Lett. **57** (1993) 622.

[3] D. Stevens, D. Scott and J. Silk, Phys. Rev. Lett. **71** (1993) 20.

[4] A. de Oliveira-Costa, G.F. Smoot, Astrophys. J. **448** (1995) 447.

[5] N. Cornish and D. Spergel, Phys. Rev. D **57** (1998) 5982.

[6] W. Hu, N. Sugiyama and J. Silk, Nature **386** (1997) 37.

[7] K.T. Inoue, astro-ph/0011462.

[8] J.R. Weeks, SnapPea, *A Computer Program for Creating and Studying Hyperbolic 3-manifolds*, available at: http://www.northnet.org/weeks

[9] J.R. Bond, D. Pogosyan and T. Souradeep, Phys. Rev. D **62** (2000) 043006.

[10] K.T. Inoue, Class. Quantum Grav. **16** (1999) 3071.

[11] K.T. Inoue, Phys. Rev. D **62**, (2000) 103001.

[12] R. Aurich, Astrophys. J. **524** (1999) 497.

[13] N.J. Cornish and D.N. Spergel, Phys. Rev. D **62** (2000) 087304.

[14] R. Aurich and F. Steiner, astro-ph/0007264.

[15] S. Perlmutter *et al*, Astrophys. J. **517** (1999) 565.

[16] A. Riess *et al*, Astron. J. **117** (1999) 707.

[17] A. Melchiorri *et al*, Astrophys. J. **536** Issue 2 (2000) L63.

[18] A. Balbi *et al*, astro-ph/0005124 acceped in Astrophys. J. Letters (2000).

Reconstructing the Quintessential Potential in Practice[1]

Takeshi Chiba [2]

Department of Physics, Kyoto University,
Kyoto 606-8502, Japan

Abstract

We investigate the feasibility of the method for reconstructing the equation of state and the effective potential of the quintessence field from SNIa data. We introduce a useful functional form to fit the luminosity distance with good accuracy (the relative error is less than 0.1%). We assess the ambiguity in reconstructing the equation of state and the effective potential which originates from the uncertainty in Ω_M. We find that the equation of state is sensitive to the assumed Ω_M, while the shape of the effective potential is not. We also demonstrate the actual reconstruction procedure using the data created by Monte-Carlo simulation.

1 Introduction

Recent various observations [2, 3], in particular distance measurements to SNIa [4, 5], strongly suggest that the universe is currently dominated by a positive vacuum energy density with negative pressure. The smallness of the vacuum energy density $\sim (10^{-12}\mathrm{GeV})^4$ has revived the idea that the cosmological "constant" is not really a constant but rather decaying. The idea of quintessence [6] (see also [7, 8] and references therein) is that vacuum energy density is played by a scalar field rolling down the almost flat potential similar to cosmological inflation, and a lot of models have been proposed so far. However, there is currently no clear guidance from particle physics as to which quintessence models may be suitable. Then it should be the observations that decide which model is correct or not. As a bottom-up approach, we have proposed that distance measurements to SNIa may allow one to reconstruct the equation of state of the dark energy or the effective potential of the quintessence field [9].

Future observational project, such as SNAP(SuperNova/Acceleration Probe)[3] could gather 2,000 SNIa in a single year and could put significant constraint on the cosmological parameters (including the equation of state of the dark energy w). In view of the future prospect for high-z SNIa search, we investigate in detail the feasibility of the method for reconstructing the equation of state and the effective potential of the quintessence field from SNIa data [9, 10, 11].

2 Parameterizing the Luminosity Distance

Considering the future prospect for high-z SNIa search, we believe it particularly useful to fit the observed luminosity distance $d_L(z)$ to a function of z which has the following properties: (1)the good convergence (the relative error is hopefully less than 0.1% because the distance error expected from SNAP will be less than a few percent) for $0 < z < 10$; (2) the correct asymptotic behavior for $z \gg 1$ ($H(z) \propto (1+z)^{3/2}$). We present a fitting function for $d_L(z)$.

We restrict ourselves to a flat FRW universe henceforth and assume Einstein gravity. In a flat model, the luminosity distance $d_L(z)$ is written in terms of the coordinate distance $r(z)$ to an object at z as

$$\frac{d_L(z)}{(1+z)} = r(z) = \int_t^{t_0} \frac{dt'}{a(t')} = \int_0^z \frac{dz'}{H(z')} = \int_y^1 \frac{2dy'}{H(y')y'^3}, \tag{1}$$

[1] This work is based on [1]
[2] E-mail:chiba@tap.scphys.kyoto-u.ac.jp
[3] http://snap.lbl.gov

where t_0 is the present time and $y = 1/\sqrt{1+z}$. It is interesting to note that in terms of y, $r(y)$ is a linear function of y if $\Omega_M = 1$. Therefore, the $d^2r(y)/dy^2$ contains the information of the non-zero pressure (see Eq.(5) below). We will elaborate on the advantage of using y and r in a separate paper.

In analogy with Pen's powerful fitting formula for $r(z)$ for a flat FRW universe with a cosmological constant [12], we propose to fit $r(z)$ in the following functional form:

$$H_0 r(z) = \eta(1) - \eta(y) \tag{2}$$

$$\eta(y) = 2\alpha \left[y^{-8} + \beta y^{-6} + \gamma y^{-4} + \delta y^{-2} + \sigma \right]^{-1/8} \tag{3}$$

The requirement $H(z)/H_0 \rightarrow 1$ for $z \rightarrow 0$ imply that σ is found to be a dependent parameter: $\sigma = (\alpha(1 + 3\beta/4 + \gamma/2 + \delta/4))^{8/9} - 1 - \beta - \gamma - \delta$. However, we shall treat σ as a free parameter for simplicity in numerical calculations.

2.1 demonstrating the goodness of fit

We shall demonstrate that the fitting function Eq.(3) indeed does a good job. For this purpose, we calculate the maximum relative error in $r(y)$ between the actual value and that calculated from the fitting function Eq.(3) in the range $0.3 < y < 1$ ($10 > z > 0$). We consider cosmological models consisting of matter and dark energy of constant equation of state $w \equiv p_X/\rho_X = 0, -1/3, -2/3, -1$ with $0.1 \leq \Omega_M \leq 1$. We fit each template luminosity distance, r_i, by the functional form Eq.(3) using the Davidon-Fletcher-Powell method. We minimize $\sum_{y_i}(r_i - r(y_i))^2/r_i^2$ with N=30 data points. The result is shown in Fig. 1. We find the maximum relative error is less than 0.05%. Our fitting function thus seems more powerful by order of magnitude than the one proposed in [11].

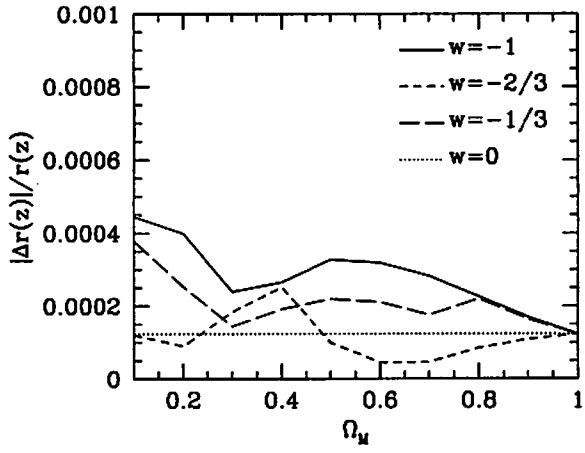


Fig. 1

Figure 1: The maximum deviation in the range $0.3 < y < 1$ ($10 > z > 0$) between the actual value and that calculated from our fitting function is shown as a function of Ω_M.

3 Reconstructing the Equation of State and the Effective Potential

One of the prime interest in the reconstruction issue is whether the effective equation of state of the x-component $w \equiv p_X/\rho_X$ is different from -1, which is the unique signature of the dynamical vacuum energy. If observations would suggest $w \neq -1$, then the next urgent project would be the real reconstruction of the effective potential, which should have profound implications for particle physics as well as for cosmology.

In this section, we reconstruct the equation of state of dark energy and the effective potential of the quintessence field. Regarding the energy density of the Universe, we assume two components: nonrelativistic matter with its present density parameter Ω_M and x-component with $\Omega_X = 1-\Omega_M$. An extension to include the radiation component is immediate but with negligible effect.

3.1 reconstructing the equation of state

In terms of a dimensionless variable $\hat{r} \equiv H_0 r$, the equation-of-state of x-component w is written as

$$w(z) = \frac{3d\hat{r}(z)/dz + 2(1+z)d^2\hat{r}(z)/dz^2}{3\left(d\hat{r}(z)/dz\right)\left(\Omega_M \left(d\hat{r}(z)/dz\right)^2 (1+z)^3 - 1\right)} \tag{4}$$

$$= \frac{-4yd^2\hat{r}/dy^2}{3\left(d\hat{r}/dy\right)\left(\Omega_M \left(d\hat{r}/dy\right)^2 - 4\right)}. \tag{5}$$

w thus depends on the second derivative of the luminosity function. Put another way, the luminosity distance depends on w through a multiple integral relation. Whether $w = -1$ or not will clearly signify whether the dark energy is constant in time or not.

3.2 reconstructing the effective potential

Reconstructing equations are

$$\hat{V}(z) = \frac{3}{\left(d\hat{r}(z)/dz\right)^2} + (1+z)\frac{d^2\hat{r}(z)/dz^2}{\left(dr(z)/dz\right)^3} - \frac{3}{2}\Omega_M(1+z)^3, \tag{6}$$

$$\left(\frac{d\hat{\phi}(z)}{dz}\right)^2 = \frac{\left(d\hat{r}(z)/dz\right)^2}{(1+z)^2}\left[-2(1+z)\frac{d^2\hat{r}(z)/dz^2}{\left(d\hat{r}(z)/dz\right)^3} - 3\Omega_M(1+z)^3\right], \tag{7}$$

where $\hat{\phi} \equiv \kappa\phi, \hat{V} \equiv \kappa^2 V/H_0^2$ with $\kappa^2 = 8\pi G$. Alternatively, in terms of $y = 1/\sqrt{1+z}$

$$\hat{V}(y) = \frac{6}{y^6(d\hat{r}/dy)^2} - \frac{2d^2\hat{r}/dy^2}{y^5(d\hat{r}/dy)^3} - \frac{3\Omega_M}{2y^6}, \tag{8}$$

$$\left(\frac{d\hat{\phi}(y)}{dy}\right)^2 = \frac{12}{y^2} + \frac{4d^2\hat{r}/dy^2}{y(d\hat{r}/dy)} - 3\Omega_M\frac{(d\hat{r}/dy)^2}{y^2}. \tag{9}$$

In order to demonstrate the effectiveness of the fitting function Eq.(3), we reconstruct the effective potential of the quintessence field. We consider three kinds of potentials which have some theoretical backgrounds; (a) cosine type [13], $\hat{V}(\hat{\phi}) = M^4(\cos(\hat{\phi})+1)$; (b) inverse power law type [14], $\hat{V}(\hat{\phi}) = M^4\hat{\phi}^{-\alpha}$; (c) exponential type [15], $\hat{V}(\hat{\phi}) = M^4\exp(-\lambda\hat{\phi})$. M and λ are fixed to give $\Omega_X = 1-\Omega_M$. $\alpha = 4$ is assumed hereafter.

The results are shown in Fig. 2. The dotted curves are numerically reconstructed $w(y)$ and $\hat{V}(\hat{\phi})$ with $\Omega_M = 0.30$ being assumed, while the solid curves are the original ones up to $y = 0.6$. We fix the present value of $\hat{\phi}$ to unity.

We also plot the dashed curves assuming $\Omega_M = 0.25, 0.27, 0.33, 0.35$. Since the smaller y, the larger the error in $w(y)$, the range of $\hat{\phi}$ significantly depends on the assumed Ω_M. We note that for $\Omega_M > 0.30$,

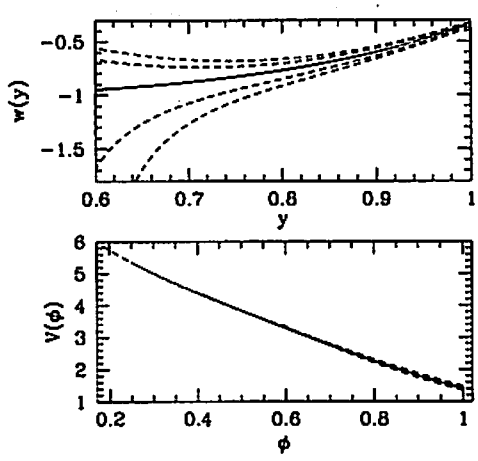


Fig. 2(a)

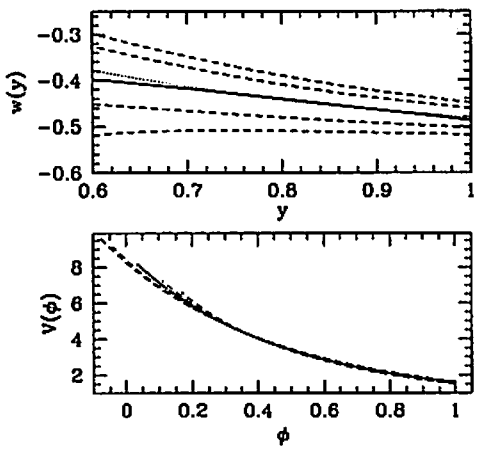


Fig. 2(b)

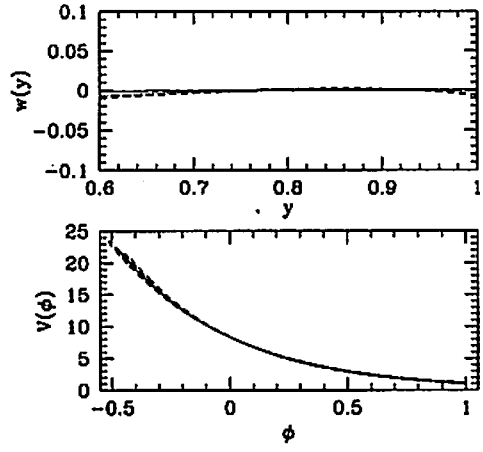


Fig. 2(c)

Figure 2: The reconstructed equation of state and the effective potential: (a) cosine potential, (b) inverse power law potential, (c) exponential potential. Solid curves are the original equations of state or effective potentials. Dotted curves are for $\Omega_M = 0.30$. The curves assuming $\Omega_M = 0.25, 0.27, 0.33, 0.35$ (from top to bottom for $w(y)$ or from left to right for V) are shown by dashed ones.

w becomes less than -1 and thus the right hand side of Eq.(9) turns negative at some y. So we stop the reconstruction of $\widehat{V}(\widehat{\phi})$ there. The range of $\widehat{\phi}$ is larger (smaller) for smaller (larger) Ω_M. For example, in the case of cosine potential, the true range of $\widehat{\phi}$ is $0.41 < \widehat{\phi} \leq 1.00$, while the reconstructed range is $0.41 \leq \widehat{\phi} \leq 1.00$ for $\Omega_M = 0.30$, $0.17 < \widehat{\phi} \leq 1.00$ for $\Omega_M = 0.25$, $0.25 < \widehat{\phi} \leq 1.00$ for $\Omega_M = 0.27$, $0.59 < \widehat{\phi} \leq 1.00$ for $\Omega_M = 0.33$, and $0.66 < \widehat{\phi} \leq 1.00$ for $\Omega_M = 0.35$. However, it is interesting that the whole shape of $\widehat{V}(\widehat{\phi})$ is less sensitive to the uncertainty in Ω_M, although w and the range of $\widehat{\phi}$ are dependent on the assumed value of Ω_M. If we assume smaller Ω_M, the resulting potential energy is larger via Eq.(8). On the other hand, $\widehat{\phi}$ decreases more rapidly back in time via Eq.(9). The opposite is the case for larger Ω_M. Both effects make the reconstructed shape of $\widehat{V}(\widehat{\phi})$ converge to the true one.

4 Reconstructing the Equation of State and the Effective Potential from Simulated Data

We simulate the actual reconstruction procedure using numerically generated data $r_i = r(y_i) + \delta r_i$ $(i = 1, \ldots, N)$ with δr_i being Gaussian distributed (zero mean and variance $\sigma r(y_i)$). The simulated data assumes a cosmological model with $\Omega_M = 0.30$. We consider $N = 30$ data and take $\sigma = 0.005$. The error is the distance error of the binned data expected from observations of 6000 supernovae by SNAP. We only consider statistical uncertainties.

We distribute the data uniformly in y from $y = 0.95$ ($z = 0.11$) to $y = 0.60$ ($z = 1.78$). We perform thousands of Monte-Carlo realizations. The 68% confidence intervals of the reconstructed potentials are shown in Fig. 3. The horizontal axis is the averaged value of $\widehat{\phi}$, while the vertical axis is the averaged value of $\widehat{V}$. We allow for 10% error in Ω_M to assess the ambiguity in the reconstruction of the potential. Obervations of very distant supernovae at $z \geq 3$ by NGST (Next Generation Space Telescope) and/or observations of galaxy cluster abundance by SZE (Sunyaev-Zel'dovich Effect) survey will determine Ω_M to a few % [17].

It should be noted that the fitted $\widehat{r}(y)$ does not necessarily satisfy the positivity of the right-hand-side of Eq.(9), that is, the weak energy condition. We perform the reconstruction using only the data which satisfy the weak energy condition.[4] That induces the bias toward larger $\dot{\phi}^2$. This is why the range of $\widehat{\phi}$ is larger than that of the true one and why the intervals of the reconstructed potential is distributed downward. Such an effect is particularly significant for the cosine type potential because in this model the equation of state is almost $w = -1$ at higher redshift. Therefore, good estimate of the *upper* bound of Ω_M is crucial for the success of the reconstruction.

5 Summary

We have studied the feasibility of reconstructing the equation of state of dark energy and the effective potential of the quintessence field from SNIa data by taking into account the uncertainty in Ω_M as well as the error in the luminosity distance. We have found that w and the range of $\widehat{\phi}$ are dependent on the assumed value of Ω_M, while the whole shape of $\widehat{V}(\widehat{\phi})$ is less sensitive to the uncertainty in Ω_M. If Ω_M could be constrained to some 10% accuracy by other observations, which may not be unrealistic expectation [17], then future high precision measurements of distances to thousands of SNIa could reveal the shape of the quintessential potential.

[4] Conversely, Eq.(9) might provide an upper bound on Ω_M, if we assume that the dark energy respects the weak energy condition, although it is not always the case [16].

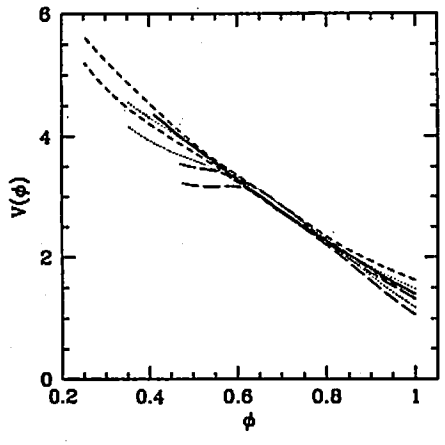


Fig. 3(a)

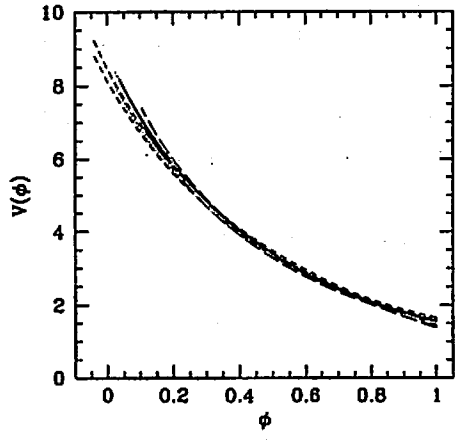


Fig. 3(b)

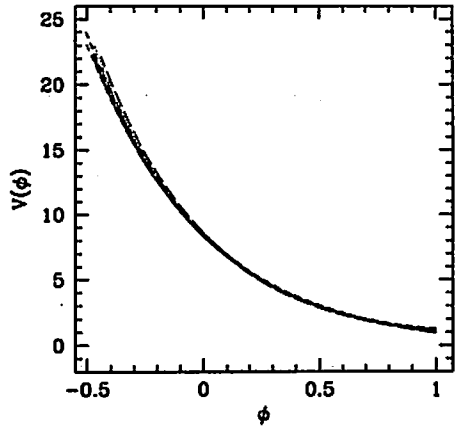


Fig. 3(c)

Figure 3: One sigma intervals for the reconstructed quintessential potential assuming luminosity distance error of 0.5%: (a) cosine potential, (b) inverse power law potential, (c) exponential potential. Solid curves are the original potentials. Dotted curves are for $\Omega_M = 0.30$, short dashed curves for $\Omega_M = 0.27$, long dashed curves for $\Omega_M = 0.33$

References

[1] T. Chiba and T. Nakamura, Phys. Rev. D **62**, 121301 (2000).

[2] P. de Bernardis, et al., Nature **404**, 955 (2000); A.E. Lange, et al., astro-ph/0005004.

[3] V. Eke, S. Cole, C.S. Frenk, Mon. Not. Roy. Astron. Soc. **282**, 263 (1996); N.A. Bahcall and X. Fan, Astrophys.J. **504**, 1 (1998);

[4] B.P. Schmidt *et al.*, Astrophys.J. **507**, 46 (1998); A.G. Riess *et al.*, Astron. J. **116**, 1009 (1998).

[5] S. Perlmutter *et al.*, Astrophys.J. **517**, 565 (1999).

[6] B. Ratra and P.J.E. Peebles, Phys. Rev. D **37**, 3406 (1988); R.R. Caldwell, R. Dave, and P.J. Steinhardt, Phys. Rev. Lett. **80**, 1582 (1998).

[7] T. Chiba, N. Sugiyama, and T. Nakamura, Mon. Not. Roy. Astron. Soc. **289**, L5 (1997); T. Chiba, N. Sugiyama, and T. Nakamura, Mon. Not. Roy. Astron. Soc. **301**, 72 (1998).

[8] M.S. Turner and M. White, Phys. Rev. D **56**, 4439 (1997).

[9] T. Nakamura and T. Chiba, Mon. Not. Roy. Astron. Soc. **306**, 696 (1999).

[10] A. Starobinsky, JETP Lett. **68**, 757 (1998); D. Huterer and M.S. Turner, Phys. Rev. D **60**, 081301 (1999).

[11] T.D. Saini, S. Raychaudhury, V. Sahni, and A.A. Starobinsky, Phys. Rev. Lett. **85**, 1162 (2000).

[12] U.-L. Pen, Astrophys.J.Suppl. **120**, 49 (1999).

[13] J.A. Frieman, C.T. Hill, and R. Watkins, Phys. Rev. D **46**, 1226 (1992); K. Choi, Phys. Rev. D **62**, 043509 (2000); J.E. Kim, JHEP **0006**, 016 (2000); Y. Nomura, T. Watari, and T. Yanagida, Phys. Lett. B **484**, 103 (2000).

[14] P. Binétruy, Phys. Rev. D **60**, 063502 (1999); A. Masiero, M. Pietroni, and F. Rosati, Phys. Rev. D **61**, 023504 (1999).

[15] C. Wetterich, Nucl. Phys. **302**, 645 (1988); Y. Fujii and T. Nishioka, Phys. Rev. D **42**, 361 (1990); P. Binétruy, M.K. Gaillard, and Y.-Y. Wu, Nucl. Phys. **493**, 27 (1997).

[16] R. Caldwell, astro-ph/9908168; T. Chiba, T. Okabe, and M. Yamaguchi, Phys. Rev. D **62**, 023511 (2000); A. Riazuelo and J.-P. Uzan, astro-ph/0004156.

[17] G. Efstathiou, Mon. Not. Roy. Astron. Soc. **310**, 842 (1999); Z. Haiman, J. Mohr, and G.P. Holder, astro-ph/002336.

A Status Review of Inflationary Cosmology

Robert H. Brandenberger,[1]

Physics Department, Brown University,
Providence, R.I., USA

Abstract

The first aim of this lecture is to highlight two areas of recent progress in inflationary cosmology, namely reheating and the quantum theory of cosmological perturbations. The second aim is to discuss important conceptual problems for the current realizations of inflation based on fundamental scalar matter fields, and to present some new approaches at solving these problems.

1 Introduction

Inflationary cosmology [1] has become one of the cornerstones of modern cosmology. Inflation was the first theory within which it was possible to make predictions about the structure of the Universe on large scales based on causal physics. The development of the inflationary Universe scenario has opened up a new and extremely promising avenue for connecting fundamental physics with experiment.

In this lecture I wish to highlight two areas in which there have been key improvements in our understanding, namely the theory of reheating (Section 3), and the quantum theory of cosmological perturbations (Section 4), the cornerstone on which the recent precision comparisons between theory and observations rest.

In spite of the remarkable success of the inflationary Universe paradigm, there are several serious conceptual problems for current models in which inflation is generated by the potential energy of a scalar matter field. These problems are discussed in Section 5.

Section 6 is a summary of some new approaches to solving the above-mentioned problems. An attempt to obtain inflation from condensates is discussed, a nonsingular Universe construction making use of higher derivative terms in the gravitational action is explained, and a framework for calculating the back-reaction of cosmological perturbations is summarized.

This short review focuses on areas of progress and on problems of inflationary cosmology. For comprehensive reviews of inflation, the reader is referred to [2, 3, 4, 5]. A recent review focusing on inflationary model building in the context of supersymmetric models can be found in [6]. An extended and more pedagogical version of these notes is [7].

2 Basics of Inflationary Cosmology

Most current models of inflation are based on Einstein's theory of General Relativity with a matter source given by a scalar field φ. Based on the cosmological principle, the metric of space-time on large distance scales can be written in Friedmann-Robertson-Walker (FRW) form:

$$ds^2 = dt^2 - a(t)^2 \left[\frac{dr^2}{1-kr^2} + r^2(d\vartheta^2 + \sin^2\vartheta d\varphi^2) \right] , \tag{1}$$

where the constant k determines the topology of the spatial sections. In the following, we shall set $k = 0$, i.e. consider a spatially flat Universe. In this case, we can without loss of generality take the scale factor $a(t)$ to be equal to 1 at the present time t_0, i.e. $a(t_0) = 1$. The coordinates r, ϑ and φ are comoving spherical coordinates.

[1] E-mail:rhb@het.brown.edu

For a homogeneous and isotropic Universe and setting the cosmological constant to zero, the Einstein equations reduce to the FRW equations

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho \tag{2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p), \tag{3}$$

where p and ρ denote the pressure and energy density, respectively. These equations can be combined to yield the continuity equation (with Hubble constant $H = \dot{a}/a$)

$$\dot{\rho} = -3H(\rho + p). \tag{4}$$

The equation of state of matter is described by a number w defined by

$$p = w\rho. \tag{5}$$

The idea of inflation [1] is very simple. We assume there is a time interval beginning at t_i and ending at t_R (the "reheating time") during which the Universe is exponentially expanding, i.e.,

$$a(t) \sim e^{Ht}, \quad t \epsilon [t_i, t_R] \tag{6}$$

with constant Hubble expansion parameter H. Such a period is called "de Sitter" or "inflationary." The success of Big Bang nucleosynthesis sets an upper limit to the time t_R of reheating, $t_R \ll t_{NS}$, t_{NS} being the time of nucleosynthesis.

During the inflationary phase, the number density of any particles initially present at $t = t_i$ decays exponentially. At $t = t_R$, all of the energy which is responsible for inflation is released (see later) as thermal energy. This is a non-adiabatic process during which the entropy increases by a large factor.

A period of inflation can solve the homogeneity problem of standard cosmology, the reason being that during inflation the physical size of the forward light cone exponentially expands and thus can easily become larger than the physical size of the past light cone at t_{rec}, the time of last scattering, thus explaining the near isotropy of the cosmic microwave background (CMB). Inflation also solves the flatness problem [8, 1].

Most importantly, inflation provides a mechanism which in a causal way generates the primordial perturbations required for galaxies, clusters and even larger objects. In inflationary Universe models, the Hubble radius ("apparent" horizon) and the ("actual") horizon (the forward light cone) do not coincide at late times. Provided that the duration of inflation is sufficiently long, then all scales within our present apparent horizon were inside the horizon since t_i. Thus, it is in principle possible to have a causal generation mechanism for perturbations [9, 10, 11, 12].

As will be discussed in Section 4, the density perturbations produced during inflation are due to quantum fluctuations in the matter and gravitational fields [10, 11]. The amplitude of these inhomogeneities corresponds to a temperature $T_H \sim H$, the Hawking temperature of the de Sitter phase. This leads one to expect that at all times t during inflation, perturbations with a fixed physical wavelength $\sim H^{-1}$ will be produced. Subsequently, the length of the waves is stretched with the expansion of space, and soon becomes much larger than the Hubble radius $\ell_H(t) = H^{-1}(t)$. The phases of the inhomogeneities are random. Thus, the inflationary Universe scenario predicts perturbations on all scales ranging from the comoving Hubble radius at the beginning of inflation to the corresponding quantity at the time of reheating. In particular, provided that inflation lasts sufficiently long, perturbations on scales of galaxies and beyond will be generated. Note, however, that it is very dangerous to interpret de Sitter Hawking radiation as thermal radiation. In fact, the equation of state of this "radiation" is not thermal [13].

In most current models of inflation, the exponential expansion is driven by the potential energy density $V(\varphi)$ of a fundamental scalar matter field φ with standard action

$$S_m = \int d^4x\sqrt{-g}\mathcal{L}_m \tag{7}$$

$$\mathcal{L}_m(\varphi) = \frac{1}{2}D_\mu\varphi D^\mu\varphi - V(\varphi), \tag{8}$$

where D_μ denotes the covariant derivative, and g is the determinant of the metric tensor. The resulting energy-momentum tensor yields the following expressions for the energy density ρ and the pressure p:

$$\rho(\varphi) = \frac{1}{2}\dot{\varphi}^2 + \frac{1}{2}a^{-2}(\nabla\varphi)^2 + V(\varphi) \tag{9}$$

$$p(\varphi) = \frac{1}{2}\dot{\varphi}^2 - \frac{1}{6}a^{-2}(\nabla\varphi)^2 - V(\varphi)\,. \tag{10}$$

It thus follows that if the scalar field is homogeneous and static, but the potential energy positive, then the equation of state $p = -\rho$ necessary for exponential inflation results (see (4)).

Most of the current realizations of potential-driven inflation are based on satisfying the conditions

$$\dot{\varphi}^2, a^{-2}(\nabla\varphi)^2 \ll V(\varphi)\,, \tag{11}$$

via the idea of slow rolling [14, 15]. Consider the equation of motion of the scalar field φ:

$$\ddot{\varphi} + 3H\dot{\varphi} - a^{-2}\,\nabla^2\,\varphi = -V'(\varphi)\,. \tag{12}$$

If the scalar field starts out almost homogeneous and at rest, if the Hubble damping term (the second term on the l.h.s. of (12) is large, and if the potential is quite flat (so that the term on the r.h.s. of (12) is small), then $\dot{\varphi}^2$ may remain small compared to $V(\varphi)$, in which case exponential inflation will result. Note that if the spatial gradient terms are initially negligible, they will remain negligible since they redshift.

Chaotic inflation [16] is a prototypical inflationary scenario. Consider a scalar field φ which is very weakly coupled to itself and other fields. In this case, φ need not be in thermal equilibrium at the Planck time, and most of the phase space for φ will correspond to large values of $|\varphi|$ (typically $|\varphi| \gg m_{pl}$). Consider now a region in space where at the initial time $\varphi(x)$ is very large, and approximately homogeneous and static. In this case, the energy-momentum tensor will be immediately dominated by the large potential energy term and induce an equation of state $p \simeq -\rho$ which leads to inflation. Due to the large Hubble damping term in the scalar field equation of motion, $\varphi(x)$ will only roll very slowly towards $\varphi = 0$ (we are making the assumption that $V(\varphi)$ has a global minimum at a finite value of φ which can then be chosen to be $\varphi = 0$). The kinetic energy contribution to ρ and p will remain small, the spatial gradient contribution will be exponentially suppressed due to the expansion of the Universe, and thus inflation persists. Note that the precise form of $V(\varphi)$ is irrelevant to the mechanism.

It is difficult to realize chaotic inflation in conventional supergravity models since gravitational corrections to the potential of scalar fields typically render the potential steep for values of $|\varphi|$ of the order of m_{pl} and larger. This prevents the slow rolling condition (11) from being realizable. Even if this condition can be satisfied, there are constraints from the amplitude of produced density fluctuations which are much harder to satisfy (see Section 5).

Hybrid inflation [17] is a solution to the above-mentioned problem of chaotic inflation. Hybrid inflation requires at least two scalar fields to play an important role in the dynamics of the Universe. As a toy model, consider the potential of a theory with two scalar fields φ and ψ:

$$V(\varphi,\psi) = \frac{1}{4}\lambda(M^2 - \psi^2)^2 + \frac{1}{2}m^2\varphi^2 + \frac{1}{2}\lambda'\psi^2\varphi^2\,. \tag{13}$$

For values of $|\varphi|$ larger than φ_c

$$\varphi_c = (\frac{\lambda}{\lambda'}M^2)^{1/2} \tag{14}$$

the minimum of ψ is $\psi = 0$, whereas for smaller values of φ the symmetry $\psi \to -\psi$ is broken and the ground state value of $|\psi|$ tends to M. The idea of hybrid inflation is that φ is slowly rolling like the inflaton field in chaotic inflation, but that the energy density of the Universe is dominated by ψ. Inflation terminates once $|\varphi|$ drops below the critical value φ_c, at which point ψ starts to move.

Note that in hybrid inflation φ_c can be much smaller than m_{pl} and hence inflation without super-Planck scale values of the fields is possible. It is possible to implement hybrid inflation in the context of supergravity (see e.g. [18]).

At the present time there are many realizations of potential-driven inflation, but there is no canonical theory. A lot of attention is being devoted to implementing inflation in the context of unified theories,

the prime candidate being superstring theory or M-theory. String theory or M-theory live in 10 or 11 space-time dimensions, respectively. When compactified to 4 space-time dimensions, there exist many *moduli* fields, scalar fields which describe flat directions in the complicated vacuum manifold of the theory. A lot of attention is now devoted to attempts at implementing inflation using moduli fields (see e.g. [19] and references therein).

Recently, it has been suggested that our space-time is a brane in a higher-dimensional space-time (see [20] for the basic construction). Ways of obtaining inflation on the brane are also under active investigation (see e.g. [21]).

It should also not be forgotten that inflation can arise from the purely gravitational sector of the theory, as in the original model of Starobinsky [22] (see also Section 6), or that it may arise from kinetic terms in an effective action as in pre-big-bang cosmology [23] or in k-inflation [24].

Theories with (almost) exponential inflation generically predict an (almost) scale-invariant spectrum of density fluctuations, as was first realized in [9, 10, 11, 12] and then studied more quantitatively in [25, 26, 27]. Via the Sachs-Wolfe effect [28], these density perturbations induce CMB anisotropies with a spectrum which is also scale-invariant on large angular scales.

The heuristic picture is as follows. If the inflationary period which lasts from t_i to t_R is almost exponential, then the physical effects which are independent of the small deviations from exponential expansion (an example of something which does depend on these deviations is effects connected with the remnant radiation density during inflation) are time-translation-invariant. This implies, for example, that quantum fluctuations at all times have the same strength when measured on the same physical length scale.

If the inhomogeneities are small, they can described by linear theory, which implies that all Fourier modes k evolve independently. The exponential expansion inflates the wavelength of any perturbation. Thus, the wavelength of perturbations generated early in the inflationary phase on length scales smaller than the Hubble radius soon becomes equal to the ("exits") Hubble radius (this happens at the time $t_i(k)$) and continues to increase exponentially. After inflation, the Hubble radius increases as t while the physical wavelength of a fluctuation increases only as $a(t)$. Thus, eventually the wavelength will cross the Hubble radius again (it will "enter" the Hubble radius) at time $t_f(k)$. Thus, it is possible for inflation to generate fluctuations on cosmological scales by causal physics.

Any physical process which obeys the symmetry of the inflationary phase and which generates perturbations will generate fluctuations of equal strength when measured when they cross the Hubble radius (see, however, Section 5.2):

$$\frac{\delta M}{M}(k, t_i(k)) = \text{const} \tag{15}$$

(independent of k). Here, $\delta M(k,t)$ denotes the r.m.s. mass fluctuation on a length scale k^{-1} at time t.

It is generally assumed that causal physics cannot affect the amplitude of fluctuations on super-Hubble scales (see, however, the comments at the end of Section 4.1). Therefore, the magnitude of $\frac{\delta M}{M}$ can change only by a factor independent of k, and hence it follows that

$$\frac{\delta M}{M}(k, t_f(k)) = \text{const}\,, \tag{16}$$

which is the definition of a scale-invariant spectrum [29].

3 Parametric Resonance and Reheating

Reheating is an important stage in inflationary cosmology. It determines the state of the Universe after inflation and has consequences for baryogenesis, defect formation and other aspects of cosmology.

After slow rolling, the inflaton field begins to oscillate uniformly in space about the true vacuum state. Quantum mechanically, this corresponds to a coherent state of $k = 0$ inflaton particles. Due to interactions of the inflaton with itself and with other fields, the coherent state will decay into quanta of elementary particles. This corresponds to post-inflationary particle production.

Reheating is usually studied using simple scalar field toy models. The one we will adopt here consists of two real scalar fields, the inflaton φ interacting with a massless scalar field χ representing ordinary

matter. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 + \frac{1}{2}\partial_\mu\chi\partial^\mu\chi - \frac{1}{2}g^2\varphi^2\chi^2\,, \tag{17}$$

with $m \sim 10^{13}$GeV (see Section 4 for a justification of this choice), and g^2 denoting the interaction coupling constant. The bare mass and self interactions of χ are neglected.

In the *elementary theory of reheating* (see e.g. [30] and [31]), the decay of the inflaton was calculated using first order perturbation theory. The decay rate Γ_B of φ typically turns out to be much smaller than the Hubble expansion rate at the end of inflation (see [7] for a worked example). The decay leads to a decrease in the amplitude of φ which can be approximated by adding an extra damping term to the equation of motion for φ:

$$\ddot{\varphi} + 3H\dot{\varphi} + \Gamma_B\dot{\varphi} = -V'(\varphi)\,. \tag{18}$$

From the above equation it follows that as long as $H > \Gamma_B$, particle production is negligible. During the phase of coherent oscillation of φ, the energy density and hence H are decreasing. Thus, eventually $H = \Gamma_B$, and at that point reheating occurs (the remaining energy density in φ is very quickly transferred to χ particles). However, at this point the matter temperature is much smaller than the energy scale of inflation (reheating is a slow process). This would imply no GUT baryogenesis, no GUT-scale defect production, and no gravitino problems in supersymmetric models with $m_{3/2} > T_R$, where $m_{3/2}$ is the gravitino mass. As we shall see, these conclusions change radically if we adopt an improved analysis of reheating.

As was first realized in [32], the above analysis misses an essential point. To see this, we focus on the equation of motion for the matter field χ. The equations for the Fourier modes χ_k in the presence of a coherent inflaton field oscillating with amplitude A,

$$\varphi(t) = Acos(mt)\,, \tag{19}$$

is

$$\ddot{\chi}_k + 3H\dot{\chi}_k + (k_p^2 + m_\chi^2 + \frac{1}{2}g^2A^2cos(2mt))\chi_k = 0, \tag{20}$$

where $k_p = k/a$ is the time-dependent physical wavenumber, and $m_\chi^2 = \frac{1}{2}A^2$ (for other toy models a similar equation is obtained, but with a different relationship between the mass and the coefficient of the oscillating term).

Let us for the moment neglect the expansion of the Universe. In this case, the friction term in (20) drops out and k_p is time-independent, and Equation (20) becomes a harmonic oscillator equation with a periodically varying mass. In the mathematics literature, this equation is called the Mathieu equation. It is well known that there is an instability. In physics, the effect is known as **parametric resonance** (see e.g. [33]). At frequencies ω_n corresponding to half integer multiples of the frequency ω of the variation of the mass, i.e.

$$\omega_k^2 = k_p^2 + m_\chi^2 = (\frac{n}{2}\omega)^2 \quad n = 1, 2, ..., \tag{21}$$

there are instability bands with widths $\Delta\omega_n$. For values of ω_k within the instability band, the value of χ_k increases exponentially:

$$\chi_k \sim e^{\mu t} \text{ with } \mu \sim \frac{g^2A^2}{\omega}\,. \tag{22}$$

In models of chaotic inflation $A \sim m_{pl}$. Hence, unless g^2 is unnaturally small (a typical value is $g^2 \sim m/m_{pl}$), it follows that $\mu \gg H$.

Since the widths of the instability bands decrease as a power of the (small) coupling constant g^2 with increasing n, for practical purposes only the lowest instability band is important. Its width is

$$\Delta\omega_k \sim gA\,. \tag{23}$$

Note, in particular, that there is no ultraviolet divergence in computing the total energy transfer from the φ to the χ field due to parametric resonance [32].

It is easy to include the effects of the expansion of the Universe (see e.g. [32, 34, 35]). The main effect is that the value of ω_k becomes time-dependent. Thus, a mode slowly enters and leaves the resonance bands. As a consequence, any mode lies in the resonance band for only a finite time.

The rate of energy transfer is given by the phase space volume of the lowest instability band multiplied by the rate of growth of the mode function χ_k. Using as an initial condition for χ_k the value $\chi_k \sim H$ given by the magnitude of the expected quantum fluctuations, we obtain

$$\dot{\rho} \sim \mu(\frac{\omega}{2})^2 \Delta\omega_k H e^{\mu t} . \tag{24}$$

Hence, the energy transfer will proceed fast on the time scale of the expansion of the Universe. There will be explosive particle production, and the energy density in matter at the end of reheating will be approximately equal to the energy density at the end of inflation.

The above is a summary of the main physics of the modern theory of reheating. The actual analysis can be refined in many ways (see e.g. [34, 35, 36], and, in the toy model considered here, [37]). First of all, it is easy to take the expansion of the Universe into account explicitly (by means of a transformation of variables), to employ an exact solution of the background model and to reduce the mode equation for χ_k to an equation which also admits exponential instabilities.

The next improvement consists of treating the χ field quantum mechanically (keeping φ as a classical background field). At this point, the techniques of quantum field theory in a curved background can be applied. There is no need to impose artificial classical initial conditions for χ_k. Instead, we may assume that χ starts in its initial vacuum state. The Bogoliubov mode mixing technique can be used to compute the number of particles at late times.

Note that the state of χ after parametric resonance is **not** a thermal state. The spectrum consists of high peaks in distinct wave bands. An important question is how this state thermalizes. For some recent progress on this issue see [38, 39]. Since the state after explosive particle production is not a thermal state, it is useful to follow [34] and call this process "preheating" instead of reheating.

Note that the details of the analysis of preheating are quite model-dependent. In fact [34, 36], in most models one does not get the kind of "narrow-band" resonance discussed here, but "broad-band" resonance. In this case, the energy transfer is even more efficient.

Recently [40] it has been argued that parametric resonance may lead to resonant amplification of super-Hubble-scale cosmological perturbations. The point is that in the presence of an oscillating inflaton field, the equation of motion for the cosmological perturbations takes on a similar form to the Mathieu equation discussed above (20). In some models of inflation, the first resonance band includes modes with wavelength larger than the Hubble radius, leading to the apparent amplification of super-Hubble-scale modes. Such a process does not violate causality [41] since it is driven by the inflaton field which is coherent on super-Hubble scales at the end of inflation as a consequence of the causal dynamics of an inflationary Universe. However, careful analyses for simple single-field [41, 42] models demonstrated that there is no net growth of the physical amplitude of gravitational fluctuations beyond what the usual theory of cosmological perturbations (see the following section) predicts. There is increasing evidence that this conclusion holds in general for models with purely adiabatic perturbations [43, 44]. In the case of many two field models it was shown that the perturbation mode which could undergo parametric amplification during reheating is exponentially suppressed during inflation [45, 46]. The criterion for models (such as the one proposed in [47]) to have exponential growth of the physical amplitude of cosmological perturbations during inflation is that there is an entropy mode which is not suppressed during inflation [48]. The resulting exponential amplification of fluctuations renders these models incompatible with the observational constraints, even including back-reaction effects [49].

4 Quantum Theory of Cosmological Perturbations

On scales larger than the Hubble radius the Newtonian theory of cosmological perturbations is inapplicable, and a general relativistic analysis is needed. On these scales, matter is essentially frozen in comoving coordinates. However, space-time fluctuations can still increase in amplitude.

In principle, it is straightforward to work out the general relativistic theory of linear fluctuations [50].

We linearize the Einstein equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{25}$$

(where $G_{\mu\nu}$ is the Einstein tensor associated with the space-time metric $g_{\mu\nu}$, and $T_{\mu\nu}$ is the energy-momentum tensor of matter) about an expanding FRW background $(g^{(0)}_{\mu\nu}, \varphi^{(0)})$:

$$g_{\mu\nu}(\underline{x}, t) = g^{(0)}_{\mu\nu}(t) + h_{\mu\nu}(\underline{x}, t) \tag{26}$$

$$\varphi(\underline{x}, t) = \varphi^{(0)}(t) + \delta\varphi(\underline{x}, t) \tag{27}$$

and pick out the terms linear in $h_{\mu\nu}$ and $\delta\varphi$ to obtain

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu} \,. \tag{28}$$

In the above, $h_{\mu\nu}$ is the perturbation in the metric and $\delta\varphi$ is the fluctuation of the matter field φ.

In practice, there are many complications which make this analysis highly nontrivial. The first problem is "gauge invariance" [51]. Imagine starting with a homogeneous FRW cosmology and introducing new coordinates which mix $\underline{x}$ and t. In terms of the new coordinates, the metric now looks inhomogeneous. The inhomogeneous piece of the metric, however, must be a pure coordinate (or "gauge") artifact. Thus, when analyzing relativistic perturbations, care must be taken to factor out effects due to coordinate transformations.

There are various methods of dealing with gauge artifacts. The simplest and most physical approach is to focus on gauge invariant variables, i.e., combinations of the metric and matter perturbations which are invariant under linear coordinate transformations.

The gauge invariant theory of cosmological perturbations is in principle straightforward, although technically rather tedious. In the following I will summarize the main steps and refer the reader to [52] for the details and further references (see also [53] for a pedagogical introduction and [54, 55, 56, 57, 58, 59, 60, 61] for other approaches).

We consider perturbations about a spatially flat Friedmann-Robertson-Walker metric

$$ds^2 = a^2(\eta)(d\eta^2 - d\underline{x}^2) \tag{29}$$

where η is conformal time (related to cosmic time t by $a(\eta)d\eta = dt$). At the linear level, metric perturbations can be decomposed into scalar modes, vector modes and tensor modes (gravitational waves). In the following, we will focus on the scalar modes since they are the only ones which couple to energy density and pressure. A scalar metric perturbation (see [62] for a precise definition) can be written in terms of four free functions of space and time:

$$\delta g_{\mu\nu} = a^2(\eta) \begin{pmatrix} 2\phi & -B_{,i} \\ -B_{,i} & 2(\psi\delta_{ij} + E_{,ij}) \end{pmatrix} . \tag{30}$$

The next step is to consider infinitesimal coordinate transformations which preserve the scalar nature of $\delta g_{\mu\nu}$, and to calculate the induced transformations of ϕ, ψ, B and E. Then we find invariant combinations to linear order. (Note that there are in general no combinations which are invariant to all orders [63].) After some algebra, it follows that

$$\Phi = \phi + a^{-1}[(B - E')a]' \tag{31}$$

$$\Psi = \psi - \frac{a'}{a}(B - E') \tag{32}$$

are two invariant combinations (a prime denotes differentiation with respect to η).

Perhaps the simplest way [52] to derive the equations of motion for gauge invariant variables is to consider the linearized Einstein equations (28) and to write them out in the longitudinal gauge defined by $B = E = 0$, in which $\Phi = \phi$ and $\Psi = \psi$, to directly obtain gauge invariant equations.

For several types of matter, in particular for scalar field matter, $\delta T^i_j \sim \delta^i_j$ which implies $\Phi = \Psi$. Hence, the scalar-type cosmological perturbations can in this case be described by a single gauge invariant variable. The equation of motion takes the form [27, 64, 58, 65, 66]

$$\ddot{\xi} = O\left(\frac{k}{aH}\right)^2 H\xi \tag{33}$$

where

$$\xi = \frac{2}{3}\frac{H^{-1}\dot{\Phi} + \Phi}{1+w} + \Phi\,. \tag{34}$$

The variable $w = p/\rho$ (with p and ρ background pressure and energy density respectively) is a measure of the background equation of state. In particular, on scales larger than the Hubble radius, the right hand side of (33) is negligible, and hence ξ is constant. The above equation (33) is only true if the system has purely adiabatic perturbations. In the case of several matter fields, entropy perturbations can lead to terms on the right hand side of (33) which are important even on length scales much larger than the Hubble radius (see e.g. [48] and references therein). This is the basic reason which allows the amplification of super-Hubble modes during reheating in such models as discussed at the end of the previous section.

If the equation of state of matter is constant, *i.e.*, $w = \text{const}$, then $\dot{\xi} = 0$ implies that the relativistic potential is time-independent on scales larger than the Hubble radius, i.e. $\Phi(t) = \text{const}$. During a transition from an initial phase with $w = w_i$ to a phase with $w = w_f$, Φ changes. In many cases, a good approximation to the dynamics given by (33) is

$$\frac{\Phi}{1+w}(t_i) = \frac{\Phi}{1+w}(t_f)\,, \tag{35}$$

To make contact with late time matter perturbations and with the Newtonian intuition, it is useful to note that, as a consequence of the Einstein constraint equations, at Hubble radius crossing Φ is a measure of the fractional density fluctuations:

$$\Phi(k, t_H(k)) \sim \frac{\delta\rho}{\rho}(k,\, t_H(k))\,. \tag{36}$$

The primordial perturbations in an inflationary cosmology are generated by quantum fluctuations (see also [67, 68]). What follows is a very brief description of the unified analysis of the quantum generation and evolution of perturbations in an inflationary Universe (for a detailed review see [52]). The basic point is that at the linearized level, the equations describing both gravitational and matter perturbations can be quantized in a consistent way. The use of gauge invariant variables makes the analysis both physically clear and computationally simple. This unified description of the generation and evolution of cosmological fluctuations goes back to the pioneering work of Sasaki [69] and Mukhanov [70] but has only fairly recently gained widespread recognition.

The first step of this analysis is to consider the action for the linear perturbations about a homogeneous and isotropic background, i.e. to expand the gravitational and matter action $S(g_{\mu\nu}, \varphi)$ to quadratic order in the fluctuation variables $h_{\mu\nu}, \delta\varphi$

$$S(g_{\mu\nu}, \varphi) = S_0(g^{(0)}_{\mu\nu}, \varphi^{(0)}) + S_2(h_{\mu\nu}, \delta\varphi; g^{(0)}_{\mu\nu}, \varphi^{(0)})\,, \tag{37}$$

where S_2 is quadratic in the perturbation variables. Focusing on the scalar perturbations, it turns out that one can express the resulting S_2 in terms of a single gauge invariant variable v which is a combination of metric and matter perturbations.

$$v = a\big(\delta\varphi + \frac{\varphi^{(0)'}}{\mathcal{H}}\Phi\big)\,. \tag{38}$$

In the above, a prime denotes the derivative with respect to conformal time, and $\mathcal{H} = a'/a$. It turns out that, after a lot of algebra, the action S_2 reduces to the action of a single gauge invariant free scalar field (namely v) with a time dependent mass [70, 69] (the time dependence reflects the expansion of the background space-time)

$$S_2 = \frac{1}{2}\int dt d^3x \big(v'^2 - (\nabla v)^2 + \frac{z''}{z}v^2\big)\,, \tag{39}$$

with

$$z = \frac{a\varphi_0'}{\mathcal{H}}\,. \tag{40}$$

This result is not surprising. Based on the study of classical cosmological perturbations, we know that there is only one field degree of freedom for the scalar perturbations. Since at the linearized level there

are no mode interactions, the action for this field must be that of a free scalar field. We can thus use standard methods to quantize this theory. If we employ canonical quantization, then the mode functions of the field operator obey the same equations as we derived in the gauge-invariant analysis of classical relativistic perturbations.

The time dependence of the mass leads to equations which have growing modes which correspond to particle production or equivalently to the generation and amplification of fluctuations. Since inflation exponentially dilutes the density of pre-existing matter, it is reasonable to assume that the perturbations start off (e.g. at the beginning of inflation) in the vacuum state (defined as a state with no particles with respect to a local comoving observer). The state defined this way will not be the vacuum state from the point of view of an observer at a later time. The Bogoliubov mode mixing technique can be used to calculate the number density of particles at a later time. In particular, expectation values of field operators such as the power spectrum can be computed.

The resulting power spectrum gives the following result for the mass perturbations at time $t_i(k)$:

$$\left(\frac{\delta M}{M}\right)^2 (k, t_i(k)) \sim k^3 \left(\frac{V'(\varphi_0)\delta\tilde{\varphi}(k, t_i(k))}{\rho_0}\right)^2 \sim \left(\frac{V'(\varphi_0)H}{\rho_0}\right)^2 . \tag{41}$$

If the background scalar field is rolling slowly, then

$$V'(\varphi_0(t_i(k))) = 3H|\dot{\varphi}_0(t_i(k))| . \tag{42}$$

and

$$(1 + p/\rho)(t_i(k)) \simeq \rho_0^{-1}\dot{\varphi}_0^2(t_i(k)) . \tag{43}$$

Combining (35), (41), (42) and (43) and we get

$$\frac{\delta M}{M}(k, t_f(k)) \sim \frac{3H^2|\dot{\varphi}_0(t_i(k))|}{\dot{\varphi}_0^2(t_i(k))} = \frac{3H^2}{|\dot{\varphi}_0(t_i(k))|} \tag{44}$$

This result can now be evaluated for specific models of inflation to find the conditions on the particle physics parameters which give a value

$$\frac{\delta M}{M}(k, t_f(k)) \sim 10^{-5} \tag{45}$$

which is required if quantum fluctuations from inflation are to provide the seeds for galaxy formation and agree with the CMB anisotropy limits.

For chaotic inflation with a potential

$$V(\varphi) = \frac{1}{2}m^2\varphi^2 , \tag{46}$$

we can solve the slow rolling equations for the inflaton to obtain

$$\frac{\delta M}{M}(k, t_f(k)) \sim 10^2 \frac{m}{m_{pl}} \tag{47}$$

which implies that $m \sim 10^{13}\,\mathrm{GeV}$ to agree with (45). Similarly, for a quartic potential with coupling constant λ, the condition $\lambda \leq 10^{-12}$ is required in order not to conflict with observations. Thus, in both examples one needs a very small parameter in the particle physics model. It has been shown quite generally [71] that small parameters are required if inflation is to solve the fluctuation problem.

To summarize the main results of the analysis of density fluctuations in inflationary cosmology:

1. Quantum vacuum fluctuations in the de Sitter phase of an inflationary Universe are the source of perturbations.

2. As a consequence of the change in the background equation of state, the evolution outside the Hubble radius produces a large amplification of the perturbations. In fact, unless the particle physics model contains very small coupling constants, the predicted fluctuations are in excess of those allowed by the bounds on cosmic microwave anisotropies.

3. The quantum generation and classical evolution of fluctuations can be treated in a unified manner. The formalism is no more complicated that the study of a free scalar field in a time dependent background.

4. Inflationary Universe models generically produce an approximately scale invariant Harrison-Zel'dovich spectrum

$$\frac{\delta M}{M}(k, t_f(k)) \simeq \text{const.} \tag{48}$$

5 Problems of Inflationary Cosmology

5.1 Fluctuation Problem

A generic problem for all realizations of potential-driven inflation studied up to now concerns the amplitude of the density perturbations which are induced by quantum fluctuations during the period of exponential expansion [26, 27]. From the amplitude of CMB anisotropies measured by COBE, and from the present amplitude of density inhomogeneities on scales of clusters of galaxies, it follows that the amplitude of the mass fluctuations $\delta M/M$ on a length scale given by the comoving wavenumber k at the time $t_f(k)$ when that scale crosses the Hubble radius in the FRW period is of the order 10^{-5}.

However, as was discussed in detail in the previous section, the present realizations of inflation based on scalar quantum field matter generically [71] predict a much larger value of these fluctuations, unless a parameter in the scalar field potential takes on a very small value. For example, as discussed at the end of the previous section, in a single field chaotic inflationary model with quartic potential the mass fluctuations generated are of the order $10^2\lambda^{1/2}$. Thus, in order not to conflict with observations, a value of λ smaller than 10^{-12} is required. There have been many attempts to justify such small parameters based on specific particle physics models, but no single convincing model has emerged.

With the recent discovery [40, 41] that long wavelength gravitational fluctuations may be amplified exponentially during reheating, a new aspect of the fluctuation problem has emerged. All models in which such amplification occurs (see e.g. [48] for a discussion of the required criteria) are ruled out because the amplitude of the fluctuations after back-reaction has set in is too large, independent of the value of the coupling constant [49].

5.2 Trans-Planckian Problem

In many models of inflation, in particular in chaotic inflation, the period of inflation is so long that comoving scales of cosmological interest today corresponded to a physical wavelength much smaller than the Planck length at the beginning of inflation. In extrapolating the evolution of cosmological perturbations according to linear theory to these very early times, we are implicitly making the assumptions that the theory remains perturbative to arbitrarily high energies and that it can be described by classical general relativity. Both of these assumptions break down on super-Planck scales. Thus the question arises as to whether the predictions of the theory are robust against modifications of the model on super-Planck scales.

A similar problem occurs in black hole physics [72]. The mixing between the modes falling towards the black hole from past infinity and the Hawking radiation modes emanating to future infinity takes place in the extreme ultraviolet region, and could be sensitive to super-Planck physics. However, in the case of black holes it has been shown that for sub-Planck wavelengths at future infinity, the predictions do not change under a class of drastic modifications of the physics described by changes in the dispersion relation of a free field [73, 74].

As was recently [75, 76] discovered, the result in the case of inflationary cosmology is different: the spectrum of fluctuations may depend quite sensitively on the dispersion relation on super-Planck scales. If we take the initial state of the fluctuations at the beginning of inflation to be given by the state which minimizes the energy density, then for certain of the dispersion relations considered in [74], the spectrum of fluctuations changes quite radically. The index of the spectrum can change (i.e. the spectrum is no longer scale-invariant) and oscillations in the spectrum may be induced. Note that for the class of dispersion relations considered in [73] the predictions are the standard ones.

The above results may be bad news for people who would like to consider scalar-field driven inflationary models as the ultimate theory. However, the positive interpretation of the results is that the spectrum of fluctuations may provide a window on super-Planck-scale physics. The present observations can already be interpreted in the sense [77] that the dispersion relation of the effective field theory which emerges from string theory cannot differ too drastically from the dispersion relation of a free scalar field.

5.3 Singularity Problem

Scalar field-driven inflation does not eliminate singularities from cosmology. Although the standard assumptions of the Penrose-Hawking theorems break down if matter has an equation of state with negative pressure, as is the case during inflation, nevertheless it can be shown that an initial singularity persists in inflationary cosmology [78]. This implies that the theory is incomplete. In particular, the physical initial value problem is not defined.

5.4 Cosmological Constant Problem

Since the cosmological constant acts as an effective energy density, its value is bounded from above by the present energy density of the Universe. In Planck units, the constraint on the effective cosmological constant Λ_{eff} is (see e.g. [79])

$$\frac{\Lambda_{eff}}{m_{pl}^4} \leq 10^{-122}. \tag{49}$$

This constraint applies both to the bare cosmological constant and to any matter contribution which acts as an effective cosmological constant.

The true vacuum value (taken on to be specific at $\varphi = 0$) of the potential $V(\varphi)$ acts as an effective cosmological constant. Its value is not constrained by any particle physics requirements (in the absence of special symmetries). Thus there must be some as yet unknown mechanism which cancels (or at least almost completely cancels) the gravitational effects of any vacuum potential energy of any scalar field. However, scalar field-driven inflation relies on the almost constant potential energy $V(\varphi)$ during the slow-rolling period acting gravitationally. How can one be sure that the unknown mechanism which cancels the gravitational effects of $V(0)$ does not also cancel the gravitational effects of $V(\varphi)$ during the slow-rolling phase? This problem is the Achilles heel of any scalar field-driven inflationary model.

Supersymmetry alone cannot provide a resolution of this problem. It is true that unbroken supersymmetry forces $V(\varphi) = 0$ in the supersymmetric vacuum. However, supersymmetry breaking will induce a non-vanishing $V(\varphi)$ in the true vacuum after supersymmetry breaking, and a cosmological constant problem of at least 60 orders of magnitude remains even with the lowest scale of supersymmetry breaking compatible with experiments.

6 New Avenues

In the light of the problems of potential-driven inflation discussed in the previous sections, it is of utmost importance to study realizations of inflation which do not require fundamental scalar fields, or new avenues towards early Universe cosmology altogether which, while maintaining (some of) the successes of inflation, address and resolve some of its difficulties.

Two examples of new avenues to early Universe cosmology which do not involve conventional inflation are the pre-big-bang scenario [23], and the varying speed of light postulate [80, 81]. The pre-big-bang scenario is a model in which the Universe starts in an empty and flat dilaton-dominated phase which leads to pole-law inflation. A nice feature of this theory is that the mechanism of super-inflationary expansion is completely independent of a potential and thus independent of the cosmological constant issue. The scenario, however, is confronted with a graceful exit problem [82], and the initial conditions need to be very special [83] (see, however, the discussion in [84]).

It is also easy to realize that a theory in which the speed of light is much larger in the early Universe than at the present time can lead to a solution of the horizon and flatness problems of standard cosmology and thus can provide an alternative to inflation for addressing them. For realizations of this scenario in the context of the brane world ideas see e.g. [85, 86, 87].

String theory may lead to a natural resolution of some of the puzzles of inflationary cosmology. This is an area of active research. The reader is referred to [19] for a review of recent studies of obtaining inflation with moduli fields, and to [21] for attempts to obtain inflation with branes. Below, three more conventional approaches to addressing some of the problems of inflation will be summarized.

6.1 Inflation from Condensates

At the present time there is no direct observational evidence for the existence of fundamental scalar fields in nature (in spite of the fact that most attractive unified theories of nature require the existence of scalar fields in the low energy effective Lagrangian). Scalar fields were initially introduced to particle physics to yield an order parameter for the symmetry breaking phase transition. Many phase transitions exist in nature; however, in all cases, the order parameter is a condensate. Hence, it is useful to consider the possibility of obtaining inflation using condensates, and in particular to ask if this would yield a different inflationary scenario.

The analysis of a theory with condensates is intrinsically non-perturbative. The expectation value of the Hamiltonian $\langle H \rangle$ of the theory contains terms with arbitrarily high powers of the expectation value $\langle \varphi \rangle$ of the condensate. A recent study of the possibility of obtaining inflation in a theory with condensates was undertaken in [88] (see also [89] for some earlier work). Instead of truncating the expansion of $\langle H \rangle$ at some arbitrary order, the assumption was made that the expansion of $\langle H \rangle$ in powers of $\langle \varphi \rangle$ is asymptotic and, specifically, Borel summable (on general grounds one expects that the expansion will be asymptotic - see e.g. [90])

$$\langle H \rangle = \sum_{n=0}^{\infty} n!(-1)^n a_n \langle \varphi^n \rangle \tag{50}$$

$$= \int_0^{\infty} ds \frac{f(s)}{s(sm_{pl} + \langle \varphi \rangle)} e^{-1/s} . \tag{51}$$

The cosmological scenario is as follows: the expectation value $\langle \varphi \rangle$ vanishes at times before the phase transition when the condensate forms. Afterwards, $\langle \varphi \rangle$ evolves according to the classical equations of motion with the potential given by (50) (assuming that the kinetic term assumes its standard form). It can easily be checked that the slow rolling conditions are satisfied. However, the slow roll conditions remain satisfied for all values of $\langle \varphi \rangle$, thus leading to a graceful exit problem - inflation will never terminate.

However, correlation functions, in particular $\langle \phi^2 \rangle$, are in general infrared divergent in the de Sitter phase of an expanding Universe. One must therefore introduce a phenomenological cutoff parameter $\epsilon(t)$ into the vacuum expectation value (VEV), and replace $\langle \varphi \rangle$ by $\langle \varphi \rangle / \epsilon$. It is natural to take $\epsilon(t) \sim H(t)$ (see e.g. [91, 92]). Hence, the dynamical system consists of two coupled functions of time $\langle \varphi \rangle$ and ϵ. A careful analysis shows that a graceful exit from inflation occurs precisely if $\langle H \rangle$ tends to zero when $\langle \varphi \rangle$ tends to large values.

As is evident, the scenario for inflation in this composite field model is very different from the standard potential-driven inflationary scenario. It is particularly interesting that the graceful exit problem from inflation is linked to the cosmological constant problem. Note that models of inflation based on composites presumably do not suffer from the trans-Planckian problem, the reason being that the effective field theory which describes the strongly interacting system is time-translation-invariant during the de Sitter phase. The physical picture is that mode interactions are essential, and are responsible for generating the fluctuations on a scale k when this scale leaves the Hubble radius at time $t_i(k)$.

6.2 Nonsingular Universe Construction

Another possibility of obtaining inflation is by making use of a modified gravity sector. More specifically, we can add to the usual gravitational action higher derivative curvature terms. These terms become important only at high curvatures. As realized a long time ago [22], specific choices of the higher derivative terms can lead to inflation. It is well motivated to consider effective gravitational actions with higher derivative terms when studying the properties of space-time at large curvatures, since any effective action for classical gravity obtained from string theory, quantum gravity, or by integrating out

matter fields, contains such terms. In our context, the interesting question is whether one can obtain a version of inflation avoiding some of the problems discussed in the previous section, specifically whether one can obtain nonsingular cosmological models.

Most higher derivative gravity theories have much worse singularity problems than Einstein's theory. However, it is not unreasonable to expect that in the fundamental theory of nature, be it string theory or some other theory, the curvature of space-time is limited. In Ref. [93] the hypothesis was made that when the limiting curvature is reached, the geometry must approach that of a maximally symmetric space-time, namely de Sitter space. The question now becomes whether it is possible to find a class of higher derivative effective actions for gravity which have the property that at large curvatures the solutions approach de Sitter space. A *nonsingular Universe construction* which achieves this goal was proposed in Refs. [94, 95]. It is based on adding to the Einstein action a particular combination of quadratic invariants of the Riemann tensor chosen such that the invariant vanishes only in de Sitter space-times. This invariant is coupled to the Einstein action via a Lagrange multiplier field in a way that the Lagrange multiplier constraint equation forces the invariant to zero at high curvatures. Thus, the metric becomes de Sitter and hence explicitly nonsingular.

If successful, the above construction will have some very appealing consequences. Consider, for example, a collapsing spatially homogeneous Universe. According to Einstein's theory, this Universe will collapse in a finite proper time to a final "big crunch" singularity. In the new theory, however, the Universe will approach a de Sitter model as the curvature increases. If the Universe is closed, there will be a de Sitter bounce followed by re-expansion. Similarly, spherically symmetric vacuum solutions of the new equations of motion will presumably be nonsingular, i.e., black holes would have no singularities in their centers. In two dimensions, this construction has been successfully realized [96].

The *nonsingular Universe construction* of [94, 95] and its applications to dilaton cosmology [97, 98] are reviewed in a recent proceedings article [99]. What follows is just a very brief summary of the points relevant to the problems listed in Section 5.

The procedure for obtaining a nonsingular Universe theory [94] is based on a Lagrange multiplier construction. Starting from the Einstein action, one can introduce Lagrange multipliers fields φ_i coupled to selected curvature invariants I_i and with potentials $V_i(\varphi_i)$ chosen such that at low curvature the theory reduces to Einstein's theory, whereas at high curvatures the solutions are manifestly nonsingular. The minimal requirements for a nonsingular theory are that *all* curvature invariants remain bounded and the space-time manifold is geodesically complete.

It is possible to achieve this by a two-step procedure. First, we choose a curvature invariant $I_1(g_{\mu\nu})$ (e.g. $I_1 = R$) and demand that it be explicitly bounded by the φ_1 constraint equation. In a second step, we demand that as $I_1(g_{\mu\nu})$ approaches its limiting value, the metric $g_{\mu\nu}$ approach the de Sitter metric $g^{DS}_{\mu\nu}$, a definite nonsingular metric with maximal symmetry. In this case, all curvature invariants are automatically bounded (they approach their de Sitter values), and the space-time can be extended to be geodesically complete. The second step can be implemented by [94] choosing a curvature invariant $I_2(g_{\mu\nu})$ with the property that

$$I_2(g_{\mu\nu}) = 0 \;\Leftrightarrow\; g_{\mu\nu} = g^{DS}_{\mu\nu}\,, \tag{52}$$

introducing a second Lagrange multiplier field φ_2 which couples to I_2, and choosing a potential $V_2(\varphi_2)$ which forces I_2 to zero at large $|\varphi_2|$:

$$S = \int d^4x\sqrt{-g}[R + \varphi_1 I_1 + V_1(\varphi_1) + \varphi_2 I_2 + V_2(\varphi_2)]\,, \tag{53}$$

with asymptotic conditions

$$V_1(\varphi_1) \;\sim\; \varphi_1 \text{ as } |\varphi_1| \to \infty \tag{54}$$

$$V_2(\varphi_2) \;\sim\; \text{const as } |\varphi_2| \to \infty \tag{55}$$

$$V_i(\varphi_i) \;\sim\; \varphi_i^2 \text{ as } |\varphi_i| \to 0 \quad i = 1, 2\,. \tag{56}$$

The first constraint renders R finite, the second forces I_2 to zero, and the third is required in order to obtain the correct low curvature limit.

The invariant

$$I_2 = (4R_{\mu\nu}R^{\mu\nu} - R^2 + C^2)^{1/2}\,, \tag{57}$$

singles out the de Sitter metric among all homogeneous and isotropic metrics (in which case adding C^2, the Weyl tensor square, is superfluous), all homogeneous and anisotropic metrics, and all radially symmetric metrics.

As a specific example one can consider the action [94, 95]

$$S = \int d^4x\sqrt{-g}\left[R + \varphi_1 R - (\varphi_2 + \frac{3}{\sqrt{2}}\varphi_1)I_2^{1/2} + V_1(\varphi_1) + V_2(\varphi_2)\right] \tag{58}$$

with

$$V_1(\varphi_1) = 12H_0^2\frac{\varphi_1^2}{1+\varphi_1}\left(1 - \frac{\ln(1+\varphi_1)}{1+\varphi_1}\right) \tag{59}$$

$$V_2(\varphi_2) = -2\sqrt{3}H_0^2\frac{\varphi_2^2}{1+\varphi_2^2}. \tag{60}$$

It can be shown that all solutions of the equations of motion which follow from this action are nonsingular [94, 95]. They are either periodic about Minkowski space-time $(\varphi_1, \varphi_2) = (0,0)$ or else asymptotically approach de Sitter space ($|\varphi_2| \to \infty$).

One of the most interesting properties of this theory is asymptotic freedom [95], i.e., the coupling between matter and gravity goes to zero at high curvatures. It is easy to add matter (e.g., dust, radiation or a scalar field) to the gravitational action in the standard way. One finds that in the asymptotic de Sitter regions, the trajectories of the solutions projected onto the (φ_1, φ_2) plane are unchanged by adding matter. This applies, for example, in a phase of de Sitter contraction when the matter energy density is increasing exponentially but does not affect the metric. The physical reason for asymptotic freedom is obvious: in the asymptotic regions of phase space, the space-time curvature approaches its maximal value and thus cannot be changed even by adding an arbitrarily high matter energy density. Hence, there is the possibility that this theory will admit a natural suppression mechanism for cosmological fluctuations. If this were the case, then the solution of the singularity problem would at the same time help resolve the fluctuation problem of potential-driven inflationary cosmology.

6.3 Back-Reaction of Cosmological Perturbations

The linear theory of cosmological perturbations in inflationary cosmology is well studied. However, effects beyond linear order have received very little attention. Beyond linear order, perturbations can effect the background in which they propagate, an effect well known from early studies [100] of gravitational waves. As will be summarized below, the back-reaction of cosmological perturbations in an exponentially expanding Universe acts like a negative cosmological constant, as first realized in the context of studies of gravitational waves in de Sitter space in [101].

Gravitational back-reaction of cosmological perturbations concerns itself with the evolution of space-times which consist of small fluctuations about a symmetric Friedmann-Robertson-Walker space-time with metric $g_{\mu\nu}^{(0)}$. The goal is to study the evolution of spatial averages of observables in the perturbed space-time. In linear theory, such averaged quantities evolve like the corresponding variables in the background space-time. However, beyond linear theory perturbations have an effect on the averaged quantities. In the case of gravitational waves, this effect is well-known [100]: gravitational waves carry energy and momentum which affects the background in which they propagate. Here, we shall focus on scalar metric perturbations.

The idea behind the analysis of gravitational back-reaction [102]is to expand the Einstein equations to second order in the perturbations, to assume that the first order terms satisfy the equations of motion for linearized cosmological perturbations [52] (hence these terms cancel), to take the spatial average of the remaining terms, and to regard the resulting equations as equations for a new homogeneous metric $g_{\mu\nu}^{(0,br)}$ which includes the effect of the perturbations to quadratic order:

$$G_{\mu\nu}(g_{\alpha\beta}^{(0,br)}) = 8\pi G\left[T_{\mu\nu}^{(0)} + \tau_{\mu\nu}\right] \tag{61}$$

where the effective energy-momentum tensor $\tau_{\mu\nu}$ of gravitational back-reaction contains the terms resulting from spatial averaging of the second order metric and matter perturbations:

$$\tau_{\mu\nu} = < T^{(2)}_{\mu\nu} - \frac{1}{8\pi G} G^{(2)}_{\mu\nu} > , \tag{62}$$

where pointed brackets stand for spatial averaging, and the superscripts indicate the order in perturbations.

As formulated in (61) and (62), the back-reaction problem is not independent of the coordinatization of space-time and hence is not well defined. It is possible to take a homogeneous and isotropic space-time, choose different coordinates, and obtain a non-vanishing $\tau_{\mu\nu}$. This "gauge" problem is related to the fact that in the above prescription, the hypersurface over which the average is taken depends on the choice of coordinates.

The key to resolving the gauge problem is to realize that to second order in perturbations, the background variables change. A gauge independent form of the back-reaction equation (61) can hence be derived [102] by defining background and perturbation variables $Q = Q^{(0)} + \delta Q$ which do not change under linear coordinate transformations. Here, Q represents collectively both metric and matter variables. The gauge-invariant form of the back-reaction equation then looks formally identical to (61), except that all variables are replaced by the corresponding gauge-invariant ones. We will follow the notation of [52], and use as gauge-invariant perturbation variables the Bardeen potentials [54] ϕ and Ψ which in longitudinal gauge coincide with the actual metric perturbations $\delta g_{\mu\nu}$. Calculations hence simplify greatly if we work directly in longitudinal gauge. These calculations have been confirmed [103] by working in a completely different gauge, making use of the covariant approach.

In [104], the effective energy-momentum tensor $\tau_{\mu\nu}$ of gravitational back-reaction was evaluated for long wavelength fluctuations in an inflationary Universe in which the matter responsible for inflation is a scalar field φ with the potential

$$V(\varphi) = \frac{1}{2} m^2 \varphi^2 . \tag{63}$$

Since in this model there is no anisotropic stress, the perturbed metric in longitudinal gauge can be written [52] in terms of a single gravitational potential ϕ

$$ds^2 = (1 + 2\phi) dt^2 - a(t)^2 (1 - 2\phi) \delta_{ij} dx^i dx^j , \tag{64}$$

where $a(t)$ is the cosmological scale factor.

It is now straightforward to compute $G^{(2)}_{\mu\nu}$ and $T^{(2)}_{\mu\nu}$ in terms of the background fields and the metric and matter fluctuations ϕ and $\delta\varphi$, By taking averages and making use of (62), the effective energy-momentum tensor $\tau_{\mu\nu}$ can be computed [104].

The general expressions for the effective energy density $\rho^{(2)} = \tau^0_0$ and effective pressure $p^{(2)} = -\frac{1}{3}\tau^i_i$ involve many terms. However, they greatly simplify if we consider perturbations with wavelength greater than the Hubble radius. In this case, all terms involving spatial gradients are negligible. From the theory of linear cosmological perturbations (see e.g. [52]) it follows that on scales larger than the Hubble radius the time derivative of ϕ is also negligible as long as the equation of state of the background does not change. The Einstein constraint equations relate the two perturbation variables ϕ and $\delta\varphi$, enabling scalar metric and matter fluctuations to be described in terms of a single gauge-invariant potential ϕ. During the slow-rolling period of the inflationary Universe, the constraint equation takes on a very simple form and implies that ϕ and $\delta\varphi$ are proportional. The upshot of these considerations is that $\tau_{\mu\nu}$ is proportional to the two point function $< \phi^2 >$, with a coefficient tensor which depends on the background dynamics. In the slow-rolling approximation we obtain [104]

$$\rho^{(2)} \simeq -4V < \phi^2 > \tag{65}$$

and

$$p^{(2)} = -\rho^{(2)} . \tag{66}$$

This demonstrates that the effective energy-momentum tensor of long-wavelength cosmological perturbations has the same form as a negative cosmological constant.

Note that during inflation, the phase space of infrared modes is growing. Hence, the magnitude of $|\rho^{(2)}|$ is also increasing. Hence, the back-reaction mechanism may lead to a dynamical relaxation mechanism for a bare cosmological constant driving inflation [105]. A similar effect holds for pure gravity at two loop order in the presence of a bare cosmological constant [101].

The interpretation of $\rho^{(2)}$ as a local density has been criticized, e.g. in [106]. Instead of computing physical observables from a spatially averaged metric, one should compute the spatial average of physical invariants corrected to quadratic order in perturbation theory. Work on this issue is in progress [107] (see also [43]).

7 Conclusions

Inflationary cosmology is an attractive *scenario*. It solves some problems of standard cosmology and leads to the possibility of a causal theory of structure formation. The specific predictions of an inflationary model of structure formation, however, depend on the specific realization of inflation, which makes the idea of inflation hard to verify or falsify. Many models of inflation have been suggested, but at the present time none are sufficiently distinguished to form a "standard" inflationary theory.

There has been a lot of recent progress in inflationary cosmology. As explained in Section 3, a new theory of inflationary reheating (preheating) has been developed based on parametric resonance. Preheating leads to a rapid energy transfer between the inflaton field and matter at the end of inflation, with important cosmological consequences. A recent development in this area is the realization that long wavelength gravitational fluctuations may be amplified exponentially in models with an entropy perturbation mode which is not suppressed during inflation.

As discussed in Section 4, there is now a consistent quantum theory of the generation and evolution of linear cosmological perturbations which describes the origin of fluctuations from an initial vacuum state of the fluctuation modes at the beginning of inflation, and which forms the basis for the precision calculations of the power spectrum of density fluctuations and of CMB anisotropies which allow detailed comparisons with current and upcoming observations.

However, there are important conceptual problems for scalar field-driven inflationary models. Four such problems discussed in this lecture (in Section 5) are the ***fluctuation problem***, the ***trans-Planckian problem***, the ***singularity problem*** and the ***cosmological constant problem***, the last of which is the Achilles heel of these inflationary models.

It may be that a convincing realization of inflation will have to wait for an improvement in our understanding of fundamental physics. Some promising but incomplete avenues which address some of the problems mentioned above and which yield inflation exist are discussed in Section 6.

Acknowledgements

I wish to thank Profs. Jun-ichi Yokoyama and Misao Sasaki for the invitation to present this lecture, and for their wonderful hospitality during the meeting. I also wish to thank the Prof. Richard MacKenzie for hospitality at the Université de Montreal and Profs. Jim Cline and Rob Myers for hospitality at McGill University during the time this manuscript was completed. The author is supported in part by the US Department of Energy under Contract DE-FG02-91ER40688, Task A.

References

[1] A. Guth, *Phys. Rev.* **D23**, 347 (1981).

[2] A. Linde, 'Particle Physics and Inflationary Cosmology' (Harwood, Chur, 1990).

[3] S. Blau and A. Guth, 'Inflationary Cosmology,' in '300 Years of Gravitation' ed. by S. Hawking and W. Israel (Cambridge Univ. Press, Cambridge, 1987).

[4] K. Olive, *Phys. Rep.* **190**, 307 (1990).

[5] A. Liddle and D. Lyth, 'Cosmological Inflation and Large-Scale Structure' (Cambridge Univ. Press, Cambridge, 2000).

[6] D. Lyth and A. Riotto, *Phys. Rept.* **314**, 1 (1999).

[7] R. Brandenberger, 'Inflationary Cosmology: Progress and Problems', hep-ph/9910410.

[8] D. Kazanas, *Ap. J.* **241**, L59 (1980).

[9] W. Press, *Phys. Scr.* **21**, 702 (1980).

[10] G. Chibisov and V. Mukhanov, 'Galaxy Formation and Phonons,' Lebedev Physical Institute Preprint No. 162 (1980);
G. Chibisov and V. Mukhanov, *Mon. Not. R. Astron. Soc.* **200**, 535 (1982).

[11] V. Lukash, *Pis'ma Zh. Eksp. Teor. Fiz.* **31**, 631 (1980).

[12] K. Sato, *Mon. Not. R. Astron. Soc.* **195**, 467 (1981).

[13] R. Brandenberger, *Phys. Lett.* **129B**, 397 (1983).

[14] A. Linde, *Phys. Lett.* **108B**, 389 (1982).

[15] A. Albrecht and P. Steinhardt, *Phys. Rev. Lett.* **48**, 1220 (1982).

[16] A. Linde, *Phys. Lett.* **129B**, 177 (1983).

[17] A. Linde, *Phys. Rev.* **D49**, 748 (1994).

[18] P. Binetruy and G. Dvali, *Phys. Lett.* **D388**, 241 (1996);
E. Halyo, *Phys. Lett.* **B387**, 43 (1996);
A. Linde and A. Riotto, *Phys. Rev.* **D56**, R1844 (1997).

[19] T. Banks, 'Remarks on M Theoretic Cosmology', hep-th/9906126.

[20] P. Horava and E. Witten, *Nucl. Phys.* **B460**, 506 (1996).

[21] G. Dvali and S. Tye, *Phys. Lett.* **B450**, 72 (1999);
N. Kaloper and A. Linde, *Phys. Rev.* **D59**, 101303 (1999).

[22] A. Starobinski, *Phys. Lett.* **91B**, 99 (1980).

[23] M. Gasperini and G. Veneziano, *Astropart. Phys.* **1**, 317 (1993).

[24] T. Damour and V. Mukhanov, *Phys. Rev. Lett.* **80**, 3440 (1998).

[25] V. Mukhanov and G. Chibisov, *JETP Lett.* **33**, 532 (1981).

[26] A. Guth and S.-Y. Pi, *Phys. Rev. Lett.* **49**, 110 (1982);
S. Hawking, *Phys. Lett.* **115B**, 295 (1982);
A. Starobinski, *Phys. Lett.* **117B**, 175 (1982);
V. Mukhanov, *JETP Lett.* **41**, 493 (1985).

[27] J. Bardeen, P. Steinhardt and M. Turner, *Phys. Rev.* **D28**, 1809 (1983).

[28] R. Sachs and A. Wolfe, *Ap. J.* **147**, 73 (1967).

[29] E. Harrison, *Phys. Rev.* **D1**, 2726 (1970);
Ya.B. Zel'dovich, *Mon. Not. R. astron. Soc.* **160**, 1p (1972).

[30] A. Dolgov and A. Linde, *Phys. Lett.* **116B**, 329 (1982).

[31] L. Abbott, E. Farhi and M. Wise, *Phys. Lett.* **117B**, 29 (1982).

[32] J. Traschen and R. Brandenberger, *Phys. Rev.* **D42**, 2491 (1990).

[33] L. Landau and E. Lifshitz, 'Mechanics' (Pergamon, Oxford, 1960);
V. Arnold, 'Mathematical Methods of Classical Mechanics' (Springer, New York, 1978).

[34] L. Kofman, A. Linde and A. Starobinski, *Phys. Rev. Lett.* **73**, 3195 (1994).

[35] Y. Shtanov, J. Traschen and R. Brandenberger, *Phys. Rev.* **D51**, 5438 (1995).

[36] L. Kofman, A. Linde and A. Starobinski, *Phys. Rev.* **D56**, 3258 (1997).

[37] P. Greene, L. Kofman, A. Linde and A. Starobinski, *Phys. Rev.* **D56**, 6175 (1997).

[38] S. Khlebnikov and I. Tkachev, *Phys. Rev. Lett.* **77**, 219 (1996);
S. Khlebnikov and I. Tkachev, *Phys. Lett.* **B390**, 80 (1997);
T. Prokopec and T. Roos, *Phys. Rev.* **D55**, 3768 (1997);
B. Greene, T. Prokopec and T. Roos, *Phys. Rev.* **D56**, 6484 (1997).

[39] G. Felder and L. Kofman, 'The Development of Equilibrium after Preheating', hep-ph/0011160.

[40] B. Bassett, D. Kaiser and R. Maartens, *Phys. Lett.* **B455**, 84 (1999).

[41] F. Finelli and R. Brandenberger, *Phys. Rev. Lett.* **82**, 1362 (1999).

[42] M. Parry and R. Easther, *Phys. Rev.* **D59**, 061301 (1999).

[43] N. Afshordi and R. Brandenberger, 'Super-Hubble Nonlinear Perturbations During Inflation', gr-qc/0011075.

[44] W. Lin, X. Meng and X. Zhang, *Phys. Rev.* **D61**, 121301 (2000), hep-ph/9912510.

[45] K. Jedamzik and G. Sigl, *Phys. Rev.* **D61**, 023519 (2000), hep-ph/9906287.

[46] P. Ivanov, *Phys. Rev.* **D61**, 023505 (2000), astro-ph/9906415.

[47] B. Bassett and F. Viniegra, *Phys. Rev.* **D62**, 043507 (2000), hep-ph/9909353.

[48] F. Finelli and R. Brandenberger, *Phys. Rev.* **D62**, 083502 (2000), hep-ph/0003172.

[49] J. Zibin, R. Brandenberger and D. Scott, 'Back Reaction and the Parametric Resonance of Cosmological Fluctuations', *Phys. Rev. D* (in press) (2001), hep-ph/0007219.

[50] E. Lifshitz, *Zh. Eksp. Teor. Fiz.* **16**, 587 (1946);
E. Lifshitz and I. Khalatnikov, *Adv. Phys.* **12**, 185 (1963).

[51] W. Press and E. Vishniac, *Ap. J.* **239**, 1 (1980).

[52] V. Mukhanov, H. Feldman and R. Brandenberger, *Phys. Rep.* **215**, 203 (1992).

[53] R. Brandenberger, H. Feldman, V. Mukhanov and T. Prokopec, 'Gauge Invariant Cosmological Perturbations: Theory and Applications,' publ. in "The Origin of Structure in the Universe," eds. E. Gunzig and P. Nardone (Kluwer, Dordrecht, 1993).

[54] J. Bardeen, *Phys. Rev.* **D22**, 1882 (1980).

[55] R. Brandenberger, R. Kahn and W. Press, *Phys. Rev.* **D28**, 1809 (1983).

[56] H. Kodama and M. Sasaki, *Prog. Theor. Phys. Suppl.* No. 78, 1 (1984).

[57] R. Durrer and N. Straumann, *Helvet. Phys. Acta* **61**, 1027 (1988);
R. Durrer, *Fund. Cosmic Phys.* **15**, 209 (1994), astro-ph/9311040.

[58] D. Lyth, *Phys. Rev.* **D31**, 1792 (1985);
D. Lyth and M. Mukherjee, *Phys. Rev.* **D38**, 485 (1988).

[59] J. Hwang and E. Vishniac, *Ap. J.* **353**, 1 (1990).

[60] G.F.R. Ellis and M. Bruni, *Phys. Rev.* **D40**, 1804 (1989);
G.F.R. Ellis, J. Hwang and M. Bruni, *Phys. Rev.* **D40**, 1819 (1989).

[61] D. Salopek and J. Stewart, *Phys. Rev.* **D51**, 517 (1995).

[62] J. Stewart, *Class. Quantum Grav.* **7**, 1169 (1990).

[63] J. Stewart and M. Walker, *Proc. R. Soc.* **A341**, 49 (1974).

[64] R. Brandenberger and R. Kahn, *Phys. Rev.* **D28**, 2172 (1984).

[65] R. Brandenberger, "Modern Cosmology and Structure Formation", in 'CP Violation and the Limits of the Standard Model (TASI 94)', ed. J. Donoghue (World Scientific, Singapore, 1995), astro-ph/9411049;
R. Brandenberger, in 'Physics of the Early Universe,' proc. of the 1989 Scottish Univ. Summer School in Physics, ed. by J. Peacock, A. Heavens and A. Davies (SUSSP Publ., Edinburgh, 1990);
R. Brandenberger, 'Lectures on Modern Cosmology and Structure Formation', in 'Particles and Fields', ed. by O. Eboli and V. Ribelles (World Scientific, Singapore 1994).

[66] J. Martin and D. Schwarz, *Phys. Rev.* **D57**, 3302 (1998);
M. Gotz, *Mon. Not. Roy. Astron. Soc.* **295**, 873 (1998), astro-ph/9704271.

[67] J. Bardeen, unpublished (1984).

[68] R. Brandenberger, *Nucl. Phys.* **B245**, 328 (1984).

[69] M. Sasaki, *Prog. Theor. Phys.* **76**, 1036 (1986).

[70] V. Mukhanov, *Zh. Eksp. Teor. Fiz.* **94**, 1 (1988).

[71] F. Adams, K. Freese and A. Guth, *Phys. Rev.* **D43**, 965 (1991).

[72] T. Jacobson, 'Black Hole Evaporation: An Open Question', in *Brighton 1990, Relativistic Astrophysics, Cosmology, and Fundamental Physics*, Proc. of the 1990 TEXAS/ESO-CERN symposium, ed. by J. Barrow, L. Mestel and P. Thomas (New York Acad. Sci., New York, 1991).

[73] W. Unruh, *Phys. Rev.* **D51**, 2827 (1995).

[74] S. Corley and T. Jacobson, *Phys. Rev.* **D54**, 1568 (1996);
S. Corley, *Phys. Rev.* **D57**, 6280 (1998).

[75] R. Brandenberger and J. Martin, 'The Robustness of Inflation to Changes in Super-Planck-Scale Physics', astro-ph/0005432.

[76] J. Martin and R. Brandenberger, 'The Transplanckian Problem of Inflationary Cosmology', hep-th/0005209.

[77] J. Martin and R. Brandenberger, 'A Cosmological Window on Trans-Planckian Physics', astro-ph/0012301.

[78] A. Borde and A. Vilenkin, *Phys. Rev. Lett.* **72**, 3305 (1993).

[79] S. Weinberg, *Rev. Mod. Phys.* **61**, 1 (1989);
S. Carroll, W. Press and E. Turner, *Ann. Rev. Astron. Astrophys.* **30**, 499 (1992);
S. Carroll, 'The Cosmological Constant', astro-ph/0004075.

[80] J. Moffatt, *Int. J. Mod. Phys.* **D2**, 351 (1995).

[81] A. Albrecht and J. Magueijo, *Phys. Rev.* **D59**, 0433516 (1999).

[82] R. Brustein and G. Veneziano, *Phys. Lett.* **B329**, 429 (1994);
R. Easther, K. Maeda and D. Wands, *Phys. Rev.* **D53**, 4247 (1996);
N. Kaloper, R. Madden and K. Olive, *Phys. Lett.* **B371**, 34 (1995).

[83] N. Kaloper, A. Linde and R. Bousso, *Phys. Rev.* **D59**, 043508 (1999).

[84] A. Buonanno, T. Damour and G. Veneziano, *Nucl. Phys.* **B543**, 275 (1999).

[85] E. Kiritsis, *JHEP* **9910**:010 (1999), hep-th/9906206.

[86] A. Kehagias and E. Kiritsis, *JHEP* **9911**:022 (1999), hep-th/9910174.

[87] S. Alexander, *JHEP* **0011**:017 (2000), hep-th/9912037.

[88] R. Brandenberger and A. Zhitnitsky, *Phys. Rev.* **D55**, 4640 (1997).

[89] R. Ball and A. Matheson, *Phys. Rev.* **D45**, 2647 (1992);
A. Matheson, R. Ball, A.-C. Davis and R. Brandenberger, *Nucl. Phys.* **B328**, 223 (1989).

[90] A. Zhitnitsky, *Phys. Rev.* **D54**, 5148 (1996).

[91] A. Linde, *Phys. Lett.* **116B**, 335 (1982).

[92] A. Vilenkin and L. Ford, *Phys. Rev.* **D25**, 2569 (1982).

[93] M. Markov, *Pis'ma Zh. Eksp. Theor. Fiz.* **36**, 214 (1982);
M. Markov, *Pis'ma Zh. Eksp. Theor. Fiz.* **46**, 342 (1987);
V. Ginsburg, V. Mukhanov and V. Frolov, *Pis'ma Zh. Eksp. Theor. Fiz.* **94**, 3 (1988);
V. Frolov, M. Markov and V. Mukhanov, *Phys. Rev.* **D41**, 383 (1990).

[94] V. Mukhanov and R. Brandenberger, *Phys. Rev. Lett.* **68**, 1969 (1992).

[95] R. Brandenberger, V. Mukhanov and A. Sornborger, *Phys. Rev.* **D48**, 1629 (1993).

[96] M. Trodden, V. Mukhanov and R. Brandenberger, *Phys. Lett.* **B316**, 483 (1993).

[97] R. Brandenberger, R. Easther and J. Maia, *JHEP* **9808**, 7 (1998), gr-qc/9806111.

[98] D. Easson and R. Brandenberger, *JHEP* **9909**:003 (1999), hep-th/9905175.

[99] R. Brandenberger and J. Magueijo, 'Imaginative cosmology', hep-ph/9912247.

[100] D. Brill and J. Hartle, *Phys. Rev.* **135** B271-278 (1964); R. Isaacson, *Phys. Rev.* **166**, 1263 and 1272 (1968).

[101] N. Tsamis and R. Woodard, *Phys. Lett.* **B292**, 269 (1993); *Phys. Rev.* **D54**, 2621 (1996); *Nucl. Phys.* **B474**, 235 (1996); *Ann. Phys.* **253**, 1 (1997).

[102] V. Mukhanov, L.R.W. Abramo and R. Brandenberger, *Phys. Rev. Lett.* **78**, 1624 (1997).

[103] L.R.W. Abramo and R. Woodard, *Phys. Rev.* **D60**, 044010 (1999), astro-ph/9811430.

[104] L.R.W. Abramo, R. Brandenberger and V. Mukhanov, *Phys. Rev.* **D56**, 3248 (1997), gr-qc/9704037.

[105] R. Brandenberger, 'Back Reaction of Cosmological Perturbations', hep-th/0004016.

[106] W. Unruh, 'Cosmological Long Wavelength Perturbations', astro-ph/9802323.

[107] L.R.W. Abramo and R. Woodard, in preparation (2001).

Multi-field fermionic preheating

Shinji Tsujikawa[1]

Department of Physics, Waseda University, 3-4-1 Ohkubo, Shinjuku-ku, Tokyo 169-8555, Japan

Bruce A. Bassett[2]

Relativity and Cosmology Group (RCG), University of Portsmouth, Mercantile House, Portsmouth, PO1 2EG, England

Fermin Viniegra[3]

Department of Theoretical Physics, Oxford University, Oxford OX1 3NP, England

Abstract

We present non-perturbative analysis of fermion production during preheating in the presence of multiple scalar fields in an expanding background paying particular attention to the interplay between instant preheating (χ-ψ) and direct fermion preheating (ϕ-ψ). In the broad resonance regime we find that instant fermion production is sensitive to suppression of the long wavelength χ modes during inflation. Further, the standard scenario of resonant fermionic preheating through inflaton decay can be significantly modified by the χ-ψ coupling, and may even lead to a decrease in the number of fermions produced.

1 Introduction

The evolution of spin 1/2 (and higher) quantum fields propagating on curved backgrounds has traditionally been considered an esoteric domain of little cosmological interest [1]. In the face of the power and simplicity of the inflationary paradigm the added complexities associated with their renormalization, regularization and solution in the presence of a non-trivial metric $g_{\mu\nu}$ made them unpopular topics of study.

The last two years have, however, seen a reversal in the fortune of higher-spin fields with the growing belief that they may be relevant for cosmology in a number of rather profound ways. Decay of the inflaton to massive states during inflation [2] can lead to non-trivial features in the primordial power spectrum and cosmic microwave background (CMB) [3]. In inflationary scenarios it is during reheating and preheating [4, 5, 6, 7] – the non-equilibrium end to many inflationary models – that fermions first became important for the subsequent evolution and nature of the universe.

In this work we will consider fermion production during preheating in the realistic multi-field case where fermion ψ is coupled to both inflaton ϕ and second scalar field χ interacting with ϕ. The following issues are investigated in detail.

- What is the interplay between direct (ϕ-ψ) and instant (ϕ-χ) [8] fermion preheating ?
- How is instant preheating sensitive to the χ vacuum expectation value (VEV) during inflation ? This will strongly alter the effective mass of fermions and therefore the nature of stochastic fermion production.

To answer these issues requires a 6-dimensional parameter space. We then study the evolution of the spectrum of massive spin-1/2 fermions over this "moduli" space of fermionic preheating.

[1] E-mail:shinji@gravity.phys.waseda.ac.jp
[2] E-mail:bruce.bassett@port.ac.uk
[3] E-mail:fermin@thphys.ox.ac.uk

2 The model and basic equations

Let us consider the massless chaotic inflation model

$$V = \frac{1}{4}\lambda\phi^4 + \frac{1}{2}g^2\phi^2\chi^2 + \tilde{g}^2\phi^3\chi + (m_\psi + h_1\phi + h_2\chi)\overline{\psi}\psi, \tag{1}$$

where ψ is the fermion field of interest with bare mass m_ψ, which is coupled to the inflaton, ϕ, and a scalar field χ through the couplings h_1 and h_2 respectively. Note that the effective mass of fermion is given by

$$m_{\rm eff} = m_\psi + h_1\phi + h_2\chi. \tag{2}$$

Typically we adopt the value $\lambda = 10^{-12}$ for the self-coupling which comes from CMB constraints. Conformally rescaled fields will be capped with a $\tilde{}$, i.e. $\tilde{\phi} \equiv a\phi$, $\tilde{\chi} \equiv a\chi$, and $\tilde{\psi} \equiv a^{3/2}\psi$ with a the scale factor.

Greene & Kofman [9] gave analytical insight into resonant fermion production and recent work has brought out the strong analogy with the scalar stochastic resonance regime [10, 11, 12, 13, 14, 15, 16]. We extend these works to the multi scalar field case where the χ field is coupled to both of ψ and ϕ fields. We have also included the $\tilde{g}^2\phi^3\chi$ coupling which may for example, appear in SUGRA models [17] in addition to the usual $\frac{1}{2}g^2\phi^2\chi^2$ coupling. Due to the existence of this term, $\langle\chi\rangle \neq 0$ during inflation and the exponential suppression for the χ field during inflation is avoided [18, 19]. We will see later that $\tilde{g}$ can significantly enhance the χ-ψ decays.

We work around a flat Friedmann-Lemaître-Robertson-Walker (FLRW) background with perturbations in the longitudinal gauge

$$ds^2 = -(1+2\Phi)dt^2 + a^2(1-2\Phi)\delta_{ij}dx^i dx^j, \tag{3}$$

where Φ is a gauge-invariant potential [20, 21, 22]. Decomposing the scalar fields into homogeneous parts and fluctuations $\phi(t,\mathbf{x}) \to \phi(t) + \delta\phi(t,\mathbf{x})$, $\chi(t,\mathbf{x}) \to \chi(t) + \delta\chi(t,\mathbf{x})$, one obtains the equations for the Fourier modes of field and metric perturbations:

$$\begin{aligned} &\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \left(\frac{k^2}{a^2} + 3\lambda\phi^2 + g^2\chi^2 + 6\tilde{g}^2\phi\chi\right)\delta\phi_k \\ =\ & 4\dot{\phi}\dot{\Phi}_k + 2(\ddot{\phi} + 3H\dot{\phi})\Phi_k - (2g^2\phi\chi + 3\tilde{g}^2\phi^2)\delta\chi_k, \end{aligned} \tag{4}$$

$$\begin{aligned} &\delta\ddot{\chi}_k + 3H\delta\dot{\chi}_k + \left(\frac{k^2}{a^2} + g^2\phi^2\right)\delta\chi_k \\ =\ & 4\dot{\chi}\dot{\Phi}_k + 2(\ddot{\chi} + 3H\dot{\chi})\Phi_k - (2g^2\phi\chi + 3\tilde{g}^2\phi^2)\delta\phi_k, \end{aligned} \tag{5}$$

$$\dot{\Phi}_k + H\Phi_k = 4\pi G(\dot{\phi}\delta\phi_k + \dot{\chi}\delta\chi_k), \tag{6}$$

where $G \equiv m_{\rm pl}^{-2}$ is Newton's gravitational constant. We include the second order field fluctuations in the background equations via k-space integrals. The equations for the Hubble parameter ($H \equiv \dot{a}/a$) and homogeneous components of the scalar fields are thus given by [23]

$$\begin{aligned} H^2 =\ & \frac{8\pi G}{3}\left[\frac{1}{2}\dot{\phi}^2 + \frac{1}{2}\langle\delta\dot{\phi}^2\rangle + \frac{1}{2a^2}\langle(\nabla\delta\phi)^2\rangle + \frac{1}{4}\lambda\langle\phi^4 + 6\phi^2\langle\delta\phi^2\rangle + 3\langle\delta\phi^2\rangle^2\rangle\right. \\ &+ \left.\frac{1}{2}\dot{\chi}^2 + \frac{1}{2}\langle\delta\dot{\chi}^2\rangle + \frac{1}{2a^2}\langle(\nabla\delta\chi)^2\rangle + \frac{g^2}{2}\langle\phi^2\chi^2\rangle + \tilde{g}^2\langle\phi^3\chi\rangle\right], \end{aligned} \tag{7}$$

$$\ddot{\phi} + 3H\dot{\phi} + \lambda\phi(\phi^2 + 3\langle\delta\phi^2\rangle) + g^2(\chi^2 + \langle\delta\chi^2\rangle)\phi + 3\tilde{g}^2(\phi^2 + \langle\delta\phi^2\rangle)\chi = 0, \tag{8}$$

$$\ddot{\chi} + 3H\dot{\chi} + g^2(\phi^2 + \langle\delta\phi^2\rangle)\chi + \tilde{g}^2(\phi^3 + 3\phi\langle\delta\phi^2\rangle) = 0, \tag{9}$$

where the expectation values of $\delta\phi^2$ and $\delta\chi^2$ are defined as

$$\langle\delta\phi^2\rangle = \frac{1}{2\pi^2}\int k^2|\delta\phi_k|^2 dk, \quad \langle\delta\chi^2\rangle = \frac{1}{2\pi^2}\int k^2|\delta\chi_k|^2 dk. \tag{10}$$

Since the Dirac spinor ψ has no homogeneous component, writing the general Dirac equation to first order in the perturbed metric (3) does not yield any terms containing Φ_k, and simply gives

$$(i\gamma^\mu\partial_\mu - m_{\rm eff}a)\tilde{\psi} = 0\,, \tag{11}$$

where ∂_0 denotes the derivative with respect to conformal time $\eta = \int a^{-1}dt$. We decompose the $\tilde{\psi}$ field into Fourier components as

$$\tilde{\psi} = \frac{1}{(2\pi)^{3/2}}\int d^3k \sum_s [a_s(k)\tilde{\mathrm{u}}_s(k,\eta)e^{+i\mathbf{k}\cdot\mathbf{x}} + b_s^\dagger(k)\tilde{\mathrm{v}}_s(k,\eta)e^{-i\mathbf{k}\cdot\mathbf{x}}]. \tag{12}$$

Imposing the following standard ansatz [9]: $\tilde{\mathrm{u}}_s(k,\eta) = (-i\gamma^\mu\partial_\mu - m_{\rm eff}a)\tilde{\psi}_k(t)R_\pm(k)$, where $R_\pm(k)$ are the eigenvectors of the helicity operator, which satisfy the relation $\gamma^0 R_\pm(k) = 1$ and $\mathbf{k}\cdot\sum R_\pm(k) = \pm 1$, we obtain the mode equation for the $\tilde{\psi}_k$:

$$\left[\frac{d^2}{dx^2} + \kappa^2 + f^2 - i\frac{df}{dx}\right]\tilde{\psi}_k = 0, \tag{13}$$

where $f \equiv \tilde{m}_{\rm eff}/(\sqrt{\lambda}\phi(0)), \kappa^2 \equiv k^2/(\lambda\phi^2(0)), x \equiv \sqrt{\lambda}\phi(0)\eta$, and $\tilde{m}_{\rm eff} \equiv am_{\rm eff} = am_\psi + h_1\tilde{\phi} + h_2\tilde{\chi}$. Eq. (13) has a WKB-form solution given by

$$\tilde{\psi}_k = \alpha_k N_+ e^{-i\int_0^t \Omega_k dx} + \beta_k N_- e^{+i\int_0^t \Omega_k dx}, \tag{14}$$

where $\Omega_k^2 \equiv \kappa^2 + f^2$ and $N_\pm \equiv 1/\sqrt{2\Omega_k(\Omega_k \pm f)}$. The comoving number density of fermions is given in terms of the Bogoliubov coefficients by

$$n_k \equiv |\beta_k|^2 = \frac{1}{2} - \frac{\kappa^2}{\Omega_k}\mathrm{Im}\left(\tilde{\psi}_k\frac{d\tilde{\psi}_k^*}{dx}\right) - \frac{f}{2\Omega_k}. \tag{15}$$

The initial conditions are chosen to be $\alpha_k(0) = 1$, $\beta_k(0) = 0$, which corresponds to $n_k(0) = 0$. The Bogoliubov coefficients satisfy the relation $|\alpha_k(t)|^2 + |\beta_k(t)|^2 = 1$, which means that the exclusion principle restricts the number density of fermions to below unity, $n_k(t) \leq 1$.

3 Fermion production during preheating

3.1 No χ decay: $h_1 \neq 0$, $h_2 = 0$

Let us consider the massive fermion production which may be relevant for leptogenesis scenario[4] . We first review the case of $h_1 \neq 0$ and $h_2 = 0$. Fermion production takes place around the region where the effective mass $m_{\rm eff} = m_\psi + h_1\phi$ vanishes, which corresponds to

$$\phi_* = -m_\psi/h_1. \tag{16}$$

As long as the relation $m_\psi < |h_1\phi|$ is satisfied, fermions are periodically created when the inflaton passes through ϕ_*. After the amplitude of inflaton decreases under the critical value ϕ_*, fermion particle creation ends.

[4] As for the massless fermion production, see Ref. [24] for details.

This allows an analytical estimate of the maximum mass of fermions produced through the h_1 coupling. Taking note that the conformal inflaton field and the scale factor evolve as [23] $\tilde{\phi} = a\phi \approx \tilde{\phi}_{\rm co}\cos(0.8472\sqrt{\lambda}\tilde{\phi}_{\rm co}\eta)$ and $a \approx \sqrt{2\pi\lambda/3}\tilde{\phi}_{\rm co}^2/m_{\rm pl}\eta$, the effective mass of the fermion field is approximately

$$m_{\rm eff} = m_\psi + h_1\sqrt{\frac{3}{2\pi\lambda}}\frac{m_{\rm pl}}{\tilde{\phi}_{\rm co}}\frac{1}{\eta}\cos(0.8472\sqrt{\lambda}\tilde{\phi}_{\rm co}\eta). \tag{17}$$

Since the minimum effective mass is achieved for $0.8472\sqrt{\lambda}\tilde{\phi}_{\rm co}\eta = \pi$, the upper limit of the fermion mass production is

$$m_\psi \lesssim (h_1/5)m_{\rm pl}. \tag{18}$$

This indicates that super-heavy fermions whose masses are of order 10^{17}-10^{18} GeV can be produced if we choose large couplings with h_1 close to unity [3].

In the vicinity of $m_{\rm eff} = 0$ where particles are non-adiabatically created, we can approximately express the effective mass of fermion as $m_{\rm eff} = h_1\phi_*'(\eta - \eta_*)$, where a prime denotes the derivative with respect to η. Since the expansion rate of the universe is typically smaller than the creation rate of fermions around $m_{\rm eff} = 0$, we can safely set $a(\eta_*) = 1$, $a'(\eta_*) = 0$ in Eq. (13). Introducing new variables $\tau \equiv \sqrt{h_1\phi_*'}(\eta - \eta_*)$, $p \equiv k/\sqrt{h_1\phi_*'}$, $d\tau \equiv \sqrt{h_1\phi_*'}d\eta$, one obtains [3]

$$\left[\frac{d^2}{d\tau^2} + p^2 + \tau^2 - i\right]\tilde{\psi}_k = 0. \tag{19}$$

The solution of this equation is described by the parabolic cylinder function, and the comoving number density of fermions is analytically derived as [3]

$$n_k \approx \exp\left(-\frac{\pi^2 k^2}{\lambda|\phi_*'|}\right) = \exp\left[-\frac{\pi\phi(0)}{|d\phi_*/dx|}\frac{\kappa^2}{\sqrt{q_1}}\right]. \tag{20}$$

This relation means that larger values of q_1 and $|d\phi_*/dx|$ lead to the production of massive particles with higher momenta. Massive fermions can be instantly created even after one oscillation of inflaton. Their occupation number increases to of order unity for low momentum modes $\kappa \sim 0$.

In Fig. 1, we depict the spectrum n_k after one oscillation of the inflaton with bare fermion mass $m_\psi = 10^{-5}m_{\rm pl}$ and couplings $q_1 \equiv h_1^2/\lambda = 10^3, 10^4, 10^5$ with $q_2 \equiv h_2^2/\lambda = 0$, $g^2/\lambda = 2$, and $\tilde{g} = 0$. When $q_1 = 10^4$ and 10^5 where fermions are created around $x \approx 3$, the phase space density of produced particle is described by Eq. (20). The larger coupling h_1 leads to the enhancement of higher momentum modes, which we can confirm in Fig. 1. Since nonadiabatic fermion production does not occur for $q_1 \lesssim 2.5 \times 10^3$ by Eq. (18) in this case, the spectrum for $q_1 = 10^3$ does not obey the analytic estimate of Eq. (20) (see the inset of Fig. 1). Note that the fermion spectra change with the passage of time. However, for a short period where fermions are created in the vicinity of $m_{\rm eff} = 0$, the spectrum is well described by Eq. (20).

Super-heavy fermions with masses greater than the GUT scale $m_\psi \sim 10^{16}$ GeV can be produced if the Yukawa-coupling is larger than $h_1 \approx 5 \times 10^{-3}$. These massive fermions may have played an important role in leptogenesis scenario or be candidates of the super-heavy dark matter [13].

3.2 Instant preheating considered: $h_1 = 0$, $h_2 \neq 0$

Let us next consider the case with $h_1 = 0$ and $h_2 \neq 0$. When $1 < g^2/\lambda < 3$, since long-wavelength modes of the χ field are resonantly amplified [23], one might assume that massive fermions are efficiently produced from the beginning of preheating. However, for instant production of super-massive fermions, the whole $h_2\chi$ term in the effective fermion mass must be comparable to the bare mass m_ψ at the start of preheating.

This condition is rather difficult to achieve unless both the coupling h_2 and the initial value of χ are large, which indicates that super-massive fermions heavier than $m_\psi \sim 10^{-6}m_{\rm pl}$ cannot be produced at the initial stage. In this case however, once the amplitude of the χ field grows through parametric resonance and the $|h_2\chi|$ becomes larger than m_ψ, χ-ψ production begins to be relevant.

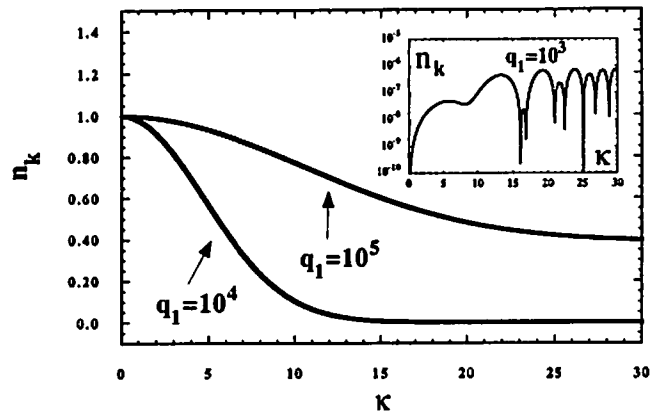


Fig. 1: The spectra of the comoving number density of fermions n_k after one oscillation of the inflaton with bare fermion mass $m_\psi = 10^{-5} m_{\rm pl}$ and $q_2 = 0$. For this mass, simple analytical estimates imply a threshold for instant production at around $q_1 = 2.5 \times 10^3$. The main figure shows n_k for $q_1 = 10^4, 10^5$ with $g^2/\lambda = 2$, and $\tilde{g} = 0$. **Inset:** n_k vs κ for q_1 below the threshold, $q_1 = 10^3$. The initial values of scalar fields are chosen to be $\phi(0) = 0.5 m_{\rm pl}$ and $\chi(0) = 10^{-9} m_{\rm pl}$.

When $g^2/\lambda \gg 1$ with the $\tilde{g}$ coupling, χ is estimated as $\chi \approx -(\tilde{g}/g)^2 \phi$ during inflation by Eq. (9) [18], since we can neglect the derivative term in the equation of χ. Then the effective mass of fermions at the end of inflation is approximately given by

$$m_{\rm eff} \approx m_\psi - h_2 \, (\tilde{g}/g)^2 \, \phi(0). \tag{21}$$

When h_2 and $\phi(0)$ are positive, $m_{\rm eff}$ can take negative values at the end of inflation for large couplings h_2 and ratio $\tilde{g}/g$. Since the amplitude of the χ field decreases at the initial stage of preheating due to cosmic expansion, instant fermion production takes place when $m_{\rm eff}$ vanishes before the χ fluctuation is amplified through resonance. This case corresponds to

$$h_2 \, (\tilde{g}/g)^2 \gtrsim 2 m_\psi / m_{\rm pl}, \tag{22}$$

where we used the value $\phi(0) = 0.5 m_{\rm pl}$.

For example, for a coupling of $h_2 = 1$ and $\tilde{g}/g = 10^{-2}$, fermions with bare masses of order $m_\psi \sim 5 \times 10^{-5} m_{\rm pl} \sim 5 \times 10^{14}$ GeV are produced at the initial stage. For instant creation of GUT scale fermions $m_\psi \sim 10^{16}$ GeV, we require $h_2 (\tilde{g}/g)^2 \gtrsim 10^{-3}$. Due to the existence of the $\tilde{g}/g$ term, which is required to be smaller than unity for successful inflation, massive fermion production via the h_2 coupling is generally hard to realize relative to the h_1 case. It is important to recognize that Eq. (22) is the condition of instant fermion production at the initial stage of preheating. Even when $h_2 \chi(0) \ll m_\psi$ initially, subsequent exponential growth of χ leads to fermion production at the later stages of preheating.

In Fig. 2, we show the maximum number density of fermions as a function of $\tilde{g}/g$ in two cases of $m_\psi = 10^{-6} m_{\rm pl}$ and $m_\psi = 10^{-8} m_{\rm pl}$ with $g^2/\lambda = 5 \times 10^3$, $h_1 = 0$, and $h_2 = 1$. We find that large values of $\tilde{g}/g$ are needed as one increases the fermion bare mass. We require strong coupling $\tilde{g}/g$ in order to produce GUT scale fermions $m_\psi \sim 10^{16}$ GeV.

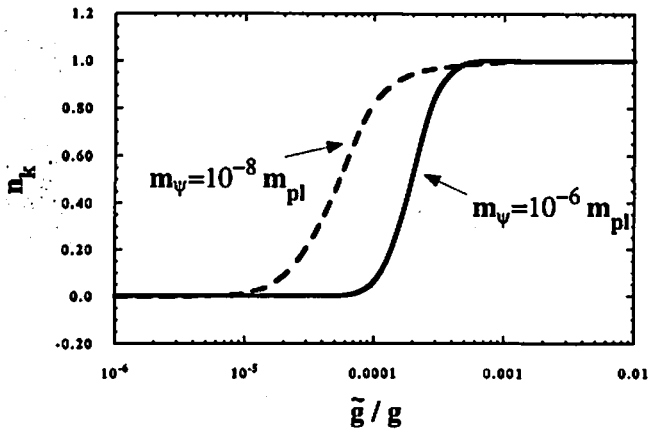


Fig. 2: The maximum comoving number density of fermions n_k for the Hubble-scale mode $\kappa = 1$ as a function of $\tilde{g}/g$ for $g^2/\lambda = 5.0 \times 10^3$, $h_1 = 0$, and $h_2 = 1$ with mass $m_\psi = 10^{-6} m_{\rm pl}$ and $m_\psi = 10^{-8} m_{\rm pl}$.

3.3 Both couplings taken into account: $h_1 \neq 0$, $h_2 \neq 0$

If both h_1 and h_2 are non-zero, fermionic preheating strongly depends both upon the absolute and relative strengths of the couplings. In the presence of the $\tilde{g}$ coupling, the effective mass of fermion at the initial stage of preheating is approximately given by

$$m_{\rm eff} \approx m_\psi + \left[h_1 - h_2 \left(\tilde{g}/g\right)^2\right] \phi. \tag{23}$$

This relation indicates that the negative coupling h_2 will assist fermion production caused by the positive h_1 coupling. Conversely, when h_1 and h_2 are both positive, including the h_2 coupling generally works as to *suppress* fermion production as long as $h_1 \geq h_2(\tilde{g}/g)^2$.

Let us consider concrete cases for the couplings: $h_1 = 10^{-5}$, $h_2 = 1$, $g^2/\lambda = 5.0 \times 10^3$, and $\tilde{g}/g = 3.0 \times 10^{-3}$ with $m_\psi = 5.0 \times 10^{-7} m_{\rm pl}$. When $h_2 = 0$, fermions are stochastically produced at the initial phase of preheating when the $m_{\rm eff}$ vanishes. However, the h_2 coupling makes the second term in Eq. (23) smaller than the bare mass m_ψ. Hence instant fermion production does not take place in this case, although n_k increases for $x \gtrsim 30$ due to the resonance from the h_2 coupling (see Fig. 3). This suppression is relevant for large coupling h_2 relative to h_1. When h_1 and h_2 are of the same order, fermion production mainly proceeds by the h_1 coupling, because the ratio $\tilde{g}/g$ is typically smaller than unity.

When $h_1 = 0$ and $h_2 \neq 0$, massive fermions which satisfy the relation (22) are instantly produced. This typically requires rather large coupling h_2 for producing massive fermions whose masses are greater than $m_\psi \sim 10^{-6} m_{\rm pl}$. However, introducing the h_1 coupling even much smaller than h_2 can potentially alter the dynamics of fermion production.

For example, in the case of $h_1 = 0$, $h_2 = 1$, $g^2/\lambda = 5.0 \times 10^3$, and $\tilde{g}/g = 3.0 \times 10^{-3}$ with $m_\psi = 10^{-5} m_{\rm pl}$, we can not expect stochastic amplification of fermions at the initial stage while n_k increases for $x \gtrsim 40$ after the χ field is sufficiently produced. Including the h_1 coupling greater than 10^{-4}, fermions are periodically created from the beginning of preheating. This means that the instant preheating scenario discussed in Ref. [8] can be further strengthened by taking into account the h_1 coupling.

4 Conclusions

We have studied spin-1/2 fermion production during preheating over a 6-dimensional coupling constant parameter space. This parameter space spans both the direct resonant decay of the inflaton to fermions

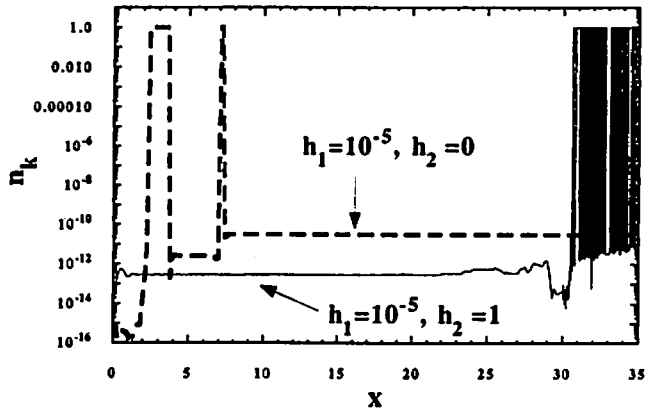


Fig. 3: h_2-suppression of instant fermion production. Number density of fermions n_k vs x in the cases $h_1 = 10^{-5}$, $h_2 = 1$ (solid), and $h_1 = 10^{-5}$, $h_2 = 0$ (dotted), with $g^2/\lambda = 5.0 \times 10^3$, $\tilde{g}/g = 3.0 \times 10^{-3}$, and $m_\psi = 5.0 \times 10^{-7} m_{\rm pl}$ for the mode $\kappa = 1$.

($\phi \to \psi$) and the indirect decays of instant preheating ($\phi \to \chi \to \psi$). We have verified the general expectation that increasing the ϕ-ψ (h_1) and χ-ψ (h_2) couplings typically leads to enhanced fermion production though with an interesting exception. In the case where $\langle\chi\rangle > 0$ and both h_1 and h_2 are non-negative, increasing h_2 leads to a *decrease* in the production of fermions.

Further, if $\langle\chi\rangle = 0$ during inflation (i.e. $\tilde{g} = 0$) instant preheating is initially sub-dominant relative to direct decays of the inflaton but later becomes important due to resonant growth of the χ field. If $\langle\chi\rangle \neq 0$ during inflation an initial burst of instant preheating is typically followed by a quiescent period and then a sudden chaotic burst of fermion production when the χ field becomes large due to resonance.

ACKNOWLEDGEMENTS

We thank Jürgen Baacke, Mar Bastero-Gil, Christopher Gordon, David Kaiser, Kei-ichi Maeda, and Takashi Torii for useful discussions. ST was supported partially by by the Waseda University Grant for Special Research Projects.

References

[1] N. D. Birrel and P. C. W. Davies, *Quantum fields in curved space* (1980).

[2] A. Berera and T. Kephart, Phys. Rev. Lett. **83**, 1084 (1999).

[3] D. Chung, E. Kolb, A. Riotto, and I. I. Tkachev, Phys. Rev. D **62**, 043508 (2000).

[4] J. Traschen and R. H. Brandenberger, Phys. Rev. D **42**, 2491 (1990); Y. Shtanov, J. Trashen, and R. H. Brandenberger, Phys. Rev. D **51**, 5438 (1995).

[5] L. Kofman, A. Linde, and A. A. Starobinsky, Phys. Rev. Lett. **73**, 3195 (1994); L. Kofman, A. Linde, and A. A. Starobinsky, Phys. Rev. D **56**, 3258 (1997).

[6] D. Boyanovsky, H. J. de Vega, R. Holman, and J. F. J. Salgado, Phys. Rev. D **54** 7570 (1996).

[7] S. Khlebnikov and I. I. Tkachev, Phys. Lett. **B390**, 80 (1997); S. Khlebnikov and I. I. Tkachev, Phys. Rev. Lett. **79**, 1607 (1997).

[8] G. Felder, L. Kofman, A. Linde, Phys. Rev. D**59**, 123523 (1999).

[9] P. B. Greene and L. Kofman, Phys. Lett. **B448**, 6 (1999).

[10] J. Baacke, K. Heitman, and C. Patzold, Phys. Rev. D **58**, 125013 (1998).

[11] D. Boyanovsky, M. D'attanasio, H. J. de Vega, R. Holman, and D.-S. Lee, Phys. Rev. D **52** 6805 (1995); D. Boyanovsky, H. J. de Vega, Y. J. Ng, D.-S. Lee, and S.-Y. Wang, Phys. Rev. D **59** 105001 (1999).

[12] S. A. Ramsey, B. L. Hu, and A. M. Stylianopoulos, Phys. Rev. D **57** 6003 (1998).

[13] G. F. Giudice, M. Peloso, A. Riotto, and I. I. Tkachev, JHEP **9908**, 014 (1999).

[14] J. Garcia-Bellido, S. Mollerach, and E. Roulet, JHEP **0002**, 034 (2000).

[15] P. B. Greene and L. Kofman, hep-ph/0003018.

[16] M. Peloso and L. Sorbo, hep-ph/0003045.

[17] G. G. Ross and S. Sarkar, Nucl. Phys. **B461**, 597 (1996).

[18] B. A. Bassett, C. Gordon, R. Maartens, and D. I. Kaiser, Phys. Rev. D **61**, 061302 (2000).

[19] S. Tsujikawa, JHEP **07**, 024 (2000).

[20] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rep. **215**, 293 (1992).

[21] H. Kodama and T. Hamazaki, Prog. Theor. Phys. **96**, 949 (1996); Y. Nambu and A. Taruya, Prog. Theor. Phys. **97**, 83 (1997); F. Finelli and R. Brandenberger, Phys. Rev. Lett. **82** 1362 (1999).

[22] B. A. Bassett, F. Tamburini, D. I. Kaiser, and R. Maartens, Nucl. Phys. **B561**, 188 (1999).

[23] P. B. Greene, L. Kofman, A. Linde, and A. A. Starobinsky, Phys. Rev. D **56**, 6175 (1997); D. I. Kaiser, Phys. Rev. D **56** 706 (1997).

[24] S. Tsujikawa, B. A. Bassett, and F. Viniegra, JHEP **08**, 019 (2000).

Chaotic inflation model in supergravity

Masahide Yamaguchi,[1]

Research Center for the Early Universe, University of Tokyo, Tokyo, 113-0033, Japan

M. Kawasaki,[2]

Research Center for the Early Universe, University of Tokyo, Tokyo, 113-0033, Japan

T. Yanagida,[3]

Department of Physics, University of Tokyo, Tokyo 113-0033, Japan
and
Research Center for the Early Universe, University of Tokyo, Tokyo, 113-0033, Japan

Abstract

We propose a chaotic inflation model in supergravity. In the model the Kähler potential has a Nambu-Goldstone-like shift symmetry of the inflaton chiral multiplet which ensures the flatness of the inflaton potential beyond the Planck scale. We show that the chaotic inflation naturally takes place by introducing a small breaking term of the shift symmetry in the superpotential. This may open a new branch of model building for inflationary universe in the framework of supergravity.

1 Introduction

The inflationary expansion of the early universe [1] is the most attractive ingredient in modern cosmology. This is not only because it naturally solves the longstanding problems in cosmology, that is the horizon and flatness problems, but also because it accounts for the origin of density fluctuations [2] as observed by the Comic Background Explorer(COBE) satellite [3]. Among various types of inflation models proposed so far, chaotic inflation model [4] is the most attractive since it can realize an inflationary expansion even in the presence of large quantum fluctuations at the Planck time. In fact, many authors have used the chaotic inflation model to discuss a number of interesting phenomena such as preheating [5], superheavy particle production [6], and primordial gravitational waves [7] in the inflationary cosmology [1].

On the other hand, supersymmetry (SUSY) [8] is widely discussed as the most interesting candidate for the physics beyond the standard model since it ensures the stability of the large hierarchy between the electroweak and the Planck scales against radiative corrections. This kind of stability is also very important to keep the flatness of inflaton potential at the quantum level. Therefore, it is quite natural to consider the inflation model in the framework of supergravity.

However, the above two ideas, i.e. chaotic inflation and supergravity, have not been naturally realized simultaneously. The main reason is that the minimal supergravity potential has an exponential factor, $\exp(\frac{\varphi_i^* \varphi_i}{M_G^2})$, which prevents any scalar fields φ_i from having values larger than the gravitational scale $M_G \simeq 2.4 \times 10^{18}$GeV. However, the inflaton φ is supposed to have a value much larger than M_G at the Planck time to cause the chaotic inflation. Thus, the above effect makes it very difficult to incorporate the chaotic inflation in the framework of supergravity. In fact, all of the existing models [9, 10] for chaotic inflation use rather specific Kähler potential, and one needs a fine tuning in the Kähler potential

[1] E-mail:gucci@resceu.s.u-tokyo.ac.jp
[2] E-mail:kawasaki@resceu.s.u-tokyo.ac.jp
[3] E-mail:yanagida@www-hep.phys.s.u-tokyo.ac.jp

since there is no symmetry reason for having such specific forms of Kähler potentials. Thus, it is very important to find a natural chaotic inflation model without any fine tuning.

Recently we proposed a natural chaotic inflation model in supergravity. In this talk, we explain our chaotic inflation model [11]. The form of Kähler potential in the model is determined by a symmetry. With this Kähler potential the inflaton φ may have a large value $\varphi \gg M_G$ to begin the chaotic inflation. Our models, in fact, need two small parameters for successful inflation. However, we emphasize that the smallness of these parameters is justified by symmetries and hence the model is natural in 't Hooft's sense [12].

The existence of a natural chaotic inflation model may open a new branch of inflation-model building in supergravity, since most of the model building in supergravity has been concentrated on other types of inflation models (e.g. hybrid inflation model etc. [13]). Furthermore, we consider that future astrophysical observations [14, 15] will be able to select types of inflation models.

Our model is based on the Nambu-Goldstone-like shift symmetry of the inflaton chiral multiplet $\Phi(x,\theta)$. Namely, we assume that the Kähler potential $K(\Phi,\Phi^*)$ is invariant under the shift of Φ,

$$\Phi \to \Phi + i\, CM_G, \tag{1}$$

where C is a dimensionless real parameter. Thus, the Kähler potential is a function of $\Phi+\Phi^*$, $K(\Phi,\Phi^*) = K(\Phi+\Phi^*)$. It is now clear that the supergravity effect $e^{K(\Phi+\Phi^*)}$ discussed above does not prevent the imaginary part of the scalar components of Φ from having a larger value than M_G. We identify it with the inflaton field φ. We also stress that the present model overcomes the so-called η problem [13] and it is an alternative to other inflation models such as D-term inflation models [16] and running inflaton mass models [17]. However, as long as the shift symmetry is exact, the inflaton φ never has a potential and hence it never causes the inflation. Therefore, we have to introduce a small breaking term of the shift symmetry in the theory. The simplest choice is to introduce a small mass term for Φ in the superpotential,

$$W = m\Phi^2. \tag{2}$$

Then, we have the potential,

$$V = e^K \left\{ \left(\frac{\partial^2 K}{\partial\Phi\partial\Phi^*} \right)^{-1} D_\Phi W D_{\Phi^*} W^* - 3|W|^2 \right\}, \tag{3}$$

with

$$D_\Phi W = \frac{\partial W}{\partial \Phi} + \frac{\partial K(\Phi+\Phi^*)}{\partial \Phi} W. \tag{4}$$

Here, Φ denotes the scalar component of the superfield Φ and we have set M_G to be unity. We easily see that $V \to -\infty$ as $|\varphi| \to \infty$ with $\Phi + \Phi^* = 0$ and the chaotic inflation does not take place, where $\varphi = -i(\Phi - \Phi^*)/\sqrt{2}$.

In this talk, we propose instead the following small mass term in the superpotential introducing a new chiral multiplet $X(x,\theta)$,

$$W = mX\Phi. \tag{5}$$

Notice that the present model possesses $U(1)_R$ symmetry under which

$$\begin{aligned} X(\theta) &\to e^{-2i\alpha} X(\theta e^{i\alpha}), \\ \Phi(\theta) &\to \Phi(\theta e^{i\alpha}), \end{aligned} \tag{6}$$

and Z_2 symmetry under which

$$\begin{aligned} X(\theta) &\to -X(\theta e^{i\alpha}), \\ \Phi(\theta) &\to -\Phi(\theta e^{i\alpha}). \end{aligned} \tag{7}$$

The above superpotential is not invariant under the shift symmetry of Φ. However, we should stress that the present model is completely natural in 't Hooft's sense [12], since we have an enhanced symmetry (the shift symmetry) in the limit $m \to 0$. That is, we consider that the small parameter m is originated

from small breaking of the shift symmetry in a more fundamental theory. We consider that as long as $m \ll \mathcal{O}(1)$, the corrections from the breaking term eq.(5) to the Kähler potential are negligibly small.[4] Then, we assume that the Kähler potential has the shift symmetry eq.(1) and the above $U(1)_R \times Z_2$ symmetry neglecting the breaking effects,

$$K(\Phi, \Phi^*, X, X^*) = K[(\Phi + \Phi^*)^2, XX^*]. \tag{8}$$

In the following analysis we take, for simplicity,

$$K = \frac{1}{2}(\Phi + \Phi^*)^2 + XX^* + \cdots. \tag{9}$$

2 Dynamics of inflation

The Lagrangian density $L(\Phi, X)$ is now given by

$$L(\Phi, X) = \partial_\mu \Phi \partial^\mu \Phi^* + \partial_\mu X \partial^\mu X^* - V(\Phi, X), \tag{10}$$

with the potential $V(\Phi, X)$ given by

$$V(\Phi, X) = m^2 e^K \left[|\Phi|^2(1 + |X|^4) + |X|^2 \left\{ 1 - |\Phi|^2 + (\Phi + \Phi^*)^2(1 + |\Phi|^2) \right\} \right], \tag{11}$$

where we have neglected higher order terms in the Kähler potential eq.(9) whose effects will be discussed later. Here, X denotes also the scalar component of the superfield X. Now, we decompose the complex scalar field Φ into two real scalar fields as,

$$\Phi = \frac{1}{\sqrt{2}}(\eta + i\varphi). \tag{12}$$

Then, the Lagrangian density $L(\eta, \varphi, X)$ is given by

$$L(\eta, \varphi, X) = \frac{1}{2}\partial_\mu \eta \partial^\mu \eta + \frac{1}{2}\partial_\mu \varphi \partial^\mu \varphi + \partial_\mu X \partial^\mu X^* - V(\eta, \varphi, X), \tag{13}$$

with the potential $V(\eta, \varphi, X)$ given by

$$\begin{aligned} V(\eta, \varphi, X) &= m^2 \exp\left(\eta^2 + |X|^2\right) \\ &\times \left[\frac{1}{2}(\eta^2 + \varphi^2)(1 + |X|^4) + |X|^2 \left\{ 1 - \frac{1}{2}(\eta^2 + \varphi^2) + 2\eta^2 \left(1 + \frac{1}{2}(\eta^2 + \varphi^2) \right) \right\} \right]. \end{aligned} \tag{14}$$

Note that η and $|X|$ should be taken as $|\eta|, |X| \lesssim \mathcal{O}(1)$ because of the presence of e^K factor. On the other hand, φ can take a value much larger than $\mathcal{O}(1)$ since e^K does not contain φ. For $\eta, |X| \ll \mathcal{O}(1)$, we can rewrite the potential as

$$V(\eta, \varphi, X) \simeq \frac{1}{2}m^2\varphi^2(1 + \eta^2) + m^2|X|^2. \tag{15}$$

Thus, the term proportional to φ^2 becomes dominant and the chaotic inflation takes place if the initial value $\varphi \gg 1$. The potential minimum for η during the inflation is given by the minimum of the Kähler potential, that is, zero. Then, during the chaotic inflation, the effective mass squared of η, m_η^2, becomes

$$m_\eta^2 \simeq m^2\varphi^2 \simeq 6H^2, \tag{16}$$

where $H(\simeq \frac{1}{\sqrt{6}}m\varphi)$ is the Hubble parameter. Because m_η^2 is much larger than $\frac{9}{4}H^2$, the field η rapidly oscillates around the origin with its amplitude damped in proportion to $a^{-3/2}$, where a is the scale factor. Thus, the field η settles down to the origin very quickly. On the other hand, the effective mass of X,

[4] The Kähler potential may have also the induced breaking terms such as $K \simeq |m\Phi|^2 + \cdots$. However, these breaking terms are negligible in the present analysis as long as $|\varphi| \lesssim m^{-1}$.

m_X, is m, which is smaller than the Hubble scale so that it does not oscillate but only slow-rolls. Using the slow-roll approximation, the classical equations of motion for both φ and X fields are given by

$$3H\dot{\varphi} \simeq -m^2\varphi, \tag{17}$$

$$3H\dot{X} \simeq -m^2 X, \tag{18}$$

where the dot represents the time derivative. Also, here and hereafter, we assume that X is real and positive. Then, we obtain the relation between φ and X fields,

$$\left(\frac{X}{X(0)}\right) \simeq \left(\frac{\varphi}{\varphi(0)}\right), \tag{19}$$

where $\varphi(0)$ and $X(0)$ are the initial values of φ and X fields. But, one should note that this relation is valid if and only if quantum fluctuations are unimportant for both φ and X fields. Therefore, we need to clarify when the classical description is feasible. For this purpose, first, we compare quantum fluctuations with classical changes for the field φ. During one expansion time, by use of eqs.(16) and (17), the classical change $\delta\varphi_c$ becomes

$$\delta\varphi_c \simeq |\dot{\varphi}|H^{-1} \simeq \frac{2}{\varphi}. \tag{20}$$

On the other hand, the amplitude of quantum fluctuations $\delta\varphi_q \simeq H/(2\pi)$. Thus, the above classical equation of motion for φ is valid only if $\varphi \ll \varphi_i \equiv \sqrt{4\pi\sqrt{6}/m}$. Otherwise, the universe is in a self-reproduction stage of eternal inflation [18, 19] and current horizon scale is contained in a domain where φ got smaller than φ_i and the classical description of φ with the above classical equation of motion became feasible. Therefore, we consider only the region $\varphi \ll m^{-1/2}$. Next, in order to estimate the amplitude of quantum fluctuations of X, we use the Fokker-Planck equation for the statistical distribution function of X, $P[X,t]$,

$$\frac{\partial}{\partial t}P[X,t] = \frac{1}{3H(t)}\frac{\partial}{\partial X}\left(m^2 X P[X,t]\right) + \frac{H^3(t)}{8\pi^2}\frac{\partial^2}{\partial X^2}P[X,t], \tag{21}$$

which is obtained through the Langevin equation based on eq.(18) with use of the stochastic inflation method of Starobinsky [20]. Then, the time evolution of the root mean square(RMS) of fluctuations of X is given by

$$\frac{d}{dt}\left\langle(\Delta X)^2\right\rangle = -\frac{2m^2}{3H}\left\langle(\Delta X)^2\right\rangle + \frac{H^3}{4\pi^2}. \tag{22}$$

Taking φ as a time variable in eq.(22) by virtue of eq.(17), we find that the RMS fluctuations of X in an initially homogeneous domain at $\varphi = \varphi_i$ are given by

$$\left\langle(\Delta X)^2\right\rangle = \frac{m^2}{96\pi^2}\left(\varphi_i^2\varphi^2 - \varphi^4\right), \tag{23}$$

at the epoch φ. Taking $\varphi_i \simeq \sqrt{4\pi\sqrt{6}/m}$, $\left\langle(\Delta X)^2\right\rangle$ asymptotically approaches

$$\left\langle(\Delta X)^2\right\rangle \simeq \frac{\sqrt{6}m}{24\pi}\varphi^2. \tag{24}$$

On the other hand, from eq.(19), the classical value of X, X_c is at most $\varphi/\varphi_i \simeq \sqrt{m}\varphi/\sqrt[4]{4\pi\sqrt{6}}$. Thus, during the chaotic inflation,

$$\sqrt{\left\langle(\Delta X)^2\right\rangle} \lesssim X_c \sim \sqrt{m}\varphi \ll 1 \ll \varphi, \tag{25}$$

because $m \ll 1$ as shown later. Thus, for X, quantum fluctuations are smaller than the classical value, and moreover our approximation that both η and X are much smaller than unity is consistent throughout the chaotic inflation.

Now that preparations are complete, the density fluctuations produced by this chaotic inflation is estimated as [21]

$$\frac{\delta\rho}{\rho} \simeq \frac{1}{5\sqrt{3}\pi}\frac{m}{2\sqrt{2}}(\varphi^2 + X^2). \tag{26}$$

Since $X \ll \varphi$, the amplitude of the density fluctuations is determined only by the φ field and the normalization at the COBE scale ($\delta\rho/\rho \simeq 2 \times 10^{-5}$ for $\varphi_{\rm COBE} \simeq 14$ [3]) gives [5]

$$m \simeq 10^{13}\ \text{GeV}. \tag{27}$$

After the inflation ends, an inflaton field φ begins to oscillate and its successive decays cause reheating of the universe. In the present model the reheating takes place efficiently if we introduce the following superpotential:

$$W = \lambda X H \bar{H}, \tag{28}$$

where H and $\bar{H}$ are a pair of Higgs doublets whose R-charge are assumed to be zero and λ is a constant.[6] Then, we have the coupling of the inflaton φ to the Higgs doublets as

$$L \sim \lambda m \varphi H \bar{H}, \tag{29}$$

which gives the reheating temperature

$$T_R \sim 10^9\ \text{GeV} \left(\frac{\lambda}{10^{-5}}\right) \left(\frac{m}{10^{13}\text{GeV}}\right)^{1/2}. \tag{30}$$

In order to avoid the overproduction of gravitinos, the reheating temperature T_R must be lower than 10^9GeV [22], which requires the small coupling $\lambda \lesssim 10^{-5}$. The small coupling λ is naturally understood in 't Hooft's sense [12] provided that $H\bar{H}$ is even under the Z_2 symmetry in eq.(7).

Recently, non-thermal production at the preheating stage was found to be important in some inflation models [23]. For the present model, as shown by Kallosh *et al.* [23], non-thermal production of gravitinos at the preheating phase is roughly estimated as

$$\left(\frac{n_{3/2}}{s}\right)_{\text{non-TH}} \sim \frac{m^3}{m^2/T_R} \lesssim 10^{-14} \left(\frac{T_R}{10^9\text{GeV}}\right) \left(\frac{m}{10^{13}\text{GeV}}\right), \tag{31}$$

where $n_{3/2}$ and s are the number density of gravitinos and entropy density. This is much less than the thermal production given by $(n_{3/2}/s)_{\rm TH} \sim 10^{-12}(T_R/10^9\text{GeV})$ and hence we can neglect the non-thermal production of gravitinos.

So far we have taken the minimal Kähler potential and neglected higher order terms like $(\Phi + \Phi^*)^4, |X|^4$, and $\cdots$. Here, we make a comment on the higher terms in the Kähler potential. Since the leading quadratic terms make the expectation values of η and X fields less than 1, the inflation dynamics is almost unchanged in the presence of the higher terms. The only relevant difference comes from the $\zeta|X|^4$ (ζ: constant) term which induces the effective mass of X given by

$$m_X^2 = 2m^2 - 2\zeta m^2 \varphi^2 \simeq -2\zeta m^2 \varphi^2. \tag{32}$$

Thus, ζ should be negative to ensure the positiveness of m_X^2. If $|\zeta| \gtrsim 1$, the effective mass becomes larger than the Hubble parameter and the X quickly settles down to $X = 0$ without slow-roll.

3 Conclusion

We have shown that a chaotic inflation naturally takes place if we assume that the Kähler potential has the Nambu-Goldstone-like shift symmetry of the inflaton chiral multiplet Φ and introduce a small breaking term of the shift symmetry in the superpotential eq.(5). Unlike other inflation models the chaotic inflation model has no initial value problem and hence it is the most attractive. However, it had been difficult to construct a natural chaotic inflation model in the framework of supergravity because the supergravity potential generally becomes very steep beyond the Planck scale. Therefore, the existence of a natural chaotic inflation model may open a new branch of inflation-model building in supergravity. Furthermore, the chaotic inflation is known to produce gravitational waves (tensor metric perturbations) [7] which might be detectable in future astrophysical observations [14, 15].

[5] The spectral index $n_s \simeq 0.96$ for $\varphi_{\rm COBE} \simeq 14$.

[6] $D_X W$ is changed to $D_X W = (m\Phi + \lambda H\bar{H})(1 + |X|^2)$. $|H|$ and $|\bar{H}|$ take values $\lesssim \mathcal{O}(1)$ due to the factor of $e^{K(H,\bar{H})}$ as the X field. Therefore, the $m\Phi$ term dominates $D_X W$ unless $\lambda \gtrsim \mathcal{O}(1)$, since $|\Phi(0)| \sim m^{-1}$ at the beginning of the universe, and the chaotic inflation begins. Once the inflation takes place, H and $\bar{H}$ acquire masses of the order the Hubble scale and rapidly go to zero. Thus, the above superpotential eq.(28) does not affect the dynamics of the inflation.

Acknowledgments

M.Y. is grateful to J. Yokoyama for useful discussions. M.K. and T.Y. are supported in part by the Grant-in-Aid, Priority Area "Supersymmetry and Unified Theory of Elementary Particles"(#707). M.Y. is partially supported by the Japanese Society for the Promotion of Science.

References

[1] See, for example, A.D. Linde, Particle Physics and Inflationary Cosmology, (Harwood, Chur, Switzerland, 1990).

[2] A. Guth and So-Y. Pi, Phys. Rev. Lett. **49**, 1110 (1982); S. W. Hawking, Phys. Lett. **115B**, 295 (1982); A. A. Starobinsky, Phys. Lett. **117B**, 175 (1982).

[3] C.L. Bennett *et al.*, Astrophys. J. **464**, 1 (1996).

[4] A. D. Linde, Phys. Lett. **129B**, 177 (1983).

[5] L. Kofman, A. D. Linde, and A. A. Starobinsky, Phys. Rev. Lett. **73**, 3185 (1994).

[6] D. J. H. Chung, E. W. Kolb, and A. Riotto, Phys. Rev. Lett. **81**, 4048 (1998); Phys. Rev. D **59**, 023501 (1999).

[7] A. A. Starobinsky, Quantum Gravity, Proc. 2nd Seminar Quantum Theory of Gravitation [in Russian], Inst. Nucl. Res. USSR Acad. Sci., Moscow, 1982, p. 58; Sov. Astron. Lett. **11**, 133 (1985).

[8] See, for a review, H. P. Nilles, Phys. Rep. **110**, 1 (1984).

[9] A. S. Goncharov and A. D. Linde, Phys. Lett. **139B**, 27 (1984); Class. Quantum Grav. **1**, L75 (1984).

[10] H. Murayama, H, Suzuki, T. Yanagida, and J. Yokoyama, Phys. Rev. D **50**, R2356 (1994).

[11] M. Kawasaki, M. Yamaguchi, and T. Yanagida, Phys. Rev. Lett. **85**, 3572 (2000).

[12] G. 't Hooft, in *Recent developments in gauge theories*, edited by G. 't Hooft *et al.* (Plenum Press, Cargèse, 1980).

[13] See, for a review, D. H. Lyth and A. Riotto, Phys. Rep. **314**, 1 (1999).

[14] `http://map.gsfc.nasa.gov`

[15] `http://astro.estec.esa.nl/SA-general/Projects/Planck/`

[16] E. D. Stewart, Phys. Rev. D **51**, 6847 (1995); P. Binetruy and G. Dvali, Phys. Lett. B **388**, 241 (1996); E. Halyo, Phys. Lett. B **387**, 43 (1996).

[17] E. D. Stewart, Phys. Lett. B **391**, 34 (1997); Phys. Rev. D **56**, 2019 (1997).

[18] A. D. Linde, Phys. Lett. B **175**, 395 (1986); Mod. Phys. Lett. **1**, 81 (1986).

[19] A. Vilenkin, Phys. Rev. D **27**, 2848 (1983).

[20] A. A. Starobinsky, In "Current Topics in Field Theory, Quantum Gravity, and Strings" Lecture Notes in Physics (Springer, 1986) **246**, 107.

[21] D. Polarski and A. A. Starobinsky, Nucl. Phys. **B385**, 623 (1992).

[22] M. Yu. Khlopov and A.D. Linde, Phys. Lett. **138B**, 265 (1984); J. Ellis, G.B. Gelmini, J.L. Lopez, D.V. Nanopoulos and S. Sarker, Nucl. Phys. **B373**, 399 (1992); M. Kawasaki and T. Moroi, Prog. Theor. Phys. **93**, 879 (1995).

[23] R. Kallosh, L. Kofman, A. Linde and A.V. Proeyen, Phys. Rev. D **61**, 103503 (2000); hep-th/0006179; G.F. Giudice, A. Riotto, and I. Tkachev J. High Energy Phys. **08**, 009 (1999); **11**, 036 (1999).

Evolution of Black Holes in Brans-Dicke Cosmology

Nobuyuki Sakai[1]

Department of Science, Faculty of Education, Yamagata University, Yamagata 990-8560, Japan

John D. Barrow[2]

DAMTP, Centre for Mathematical Science, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

Abstract

We consider a modified "Swiss cheese" model in the Brans-Dicke theory, and discuss the evolution of black holes in the expanding universe. We define the black hole radius by the Misner-Sharp mass and find the time evolution for dust and vacuum universes.

1 Introduction

The evolution of primordial black holes in scalar-tensor theories has been studied in the literature [1, 2, 3, 4]. The question of what happens to a black hole in an expanding universe in these theories was posed by Barrow, who discussed two scenarios [1]: (a) The effective gravitational "constant" $G(t)$ at a black hole changes along with its cosmological evolution so that the size of a black hole is approximated by $R = 2G(t)M$. (b) G remains constant at the black hole event horizon while it evolves on larger scales; a large inhomogeneity in G is therefore generated. The case (b) was called "gravitational memory" because the black hole remembers the value of G at its formation time. In either case constraints on primordial black holes would be modified [2].

Later Jacobson claimed that there is no "gravitational memory" effect, analyzing the evolution of a scalar field $\phi(t, r)$ in Schwarzschild background [3]. His analytic solution showed that ϕ at the event horizon evolves along with its asymptotic value $\phi(t, \infty)$. However, the solution used has a special form with very fast variation of $G \propto t^{-1}$ in the background universe. It was also argued that even if the black hole mass in the Einstein frame is constant the mass in the Jordan frame is time-dependent.

In order to investigate the time-dependence of the black hole mass, Saida and Soda constructed a "cell lattice" universe in the Brans-Dicke (BD) theory [4]. In their model the universe are tessellated by identical polyhedrons, which are replaced by Schwarzschild-like black holes. It was shown that the black hole mass has adiabatic time dependence, and that its time dependence is qualitatively different according to the sign of the curvature of the universe.

As an extension of Saida and Soda's work, we consider a "Swiss cheese" model [5] in the BD theory and discuss the evolution of the radius and mass of black holes. The ordinary Swiss cheese model refers to a cosmological model in which spherical regions in the Friedmann-Robertson-Walker (FRW) universe are replaced by Schwarzschild black spacetimes. Here we construct such a model in BD cosmology.

2 Background Universe in Brans-Dicke Theory

The BD theory is described by the action,

$$S = \int d^4x\sqrt{-g}\left[\frac{\phi}{16\pi}\mathcal{R} - \frac{\omega}{16\pi\phi}(\nabla_\mu\phi)^2 + \mathcal{L}_m\right], \tag{1}$$

[1] E-mail: sakai@ke-sci.kj.yamagata-u.ac.jp
[2] E-mail: J.D.Barrow@damtp.cam.ac.uk

where ϕ is the BD field, ω is the BD parameter, and $\mathcal{L}_m$ is the matter Lagrangian. The variation of Eq.(1) with respect to $g_{\mu\nu}$ and ϕ yield the field equations:

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = \frac{8\pi}{\phi}T_{\mu\nu} + \frac{\omega}{\phi}\left[\nabla_\mu\phi\nabla_\nu\phi - \frac{1}{2}g_{\mu\nu}(\nabla\phi)^2\right] + \nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\Box\phi, \tag{2}$$

$$\Box\phi = \frac{8\pi}{2\omega+3}\mathrm{Tr}T. \tag{3}$$

As a background universe we assume the FRW spacetime:

$$ds^2 = -dt^2 + a(t)^2\left\{\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2)\right\}, \tag{4}$$

where k is the sign of the spatial curvature. As an energy-momentum tensor we consider a dust fluid:

$$T_{\mu\nu} = \rho u_\mu u_\nu, \tag{5}$$

where ρ and u_μ are the density and the four velocity of dust, respectively. Then the field equations (2) and (3) reduce the following equations for the background universe:

$$H^2 + \frac{k}{a^2} = \frac{8\pi}{3\phi} - H\frac{\dot\phi}{\phi} + \frac{\omega}{6}\left(\frac{\dot\phi}{\phi}\right)^2 \tag{6}$$

$$\ddot\phi + 3H\dot\phi = -\frac{8\pi\rho}{2\omega+3}, \tag{7}$$

where an overdot denotes d/dt and $H \equiv \dot a/a$ is the Hubble parameter.

In the case of $k = 0$ we know the analytic solution for the dust universe:

$$a(t) = a_0(t-t_+)^{\lambda_+}(t-t_-)^{\lambda_-}, \quad \phi(t) = \phi_0(t-t_+)^{\kappa_+}(t-t_-)^{\kappa_-}, \tag{8}$$

with

$$\lambda_\pm = \frac{\omega+1\pm\sqrt{1+2\omega/3}}{3\omega+4}, \quad \kappa_\pm = \frac{1\pm 3\sqrt{1+2\omega/3}}{3\omega+4}, \tag{9}$$

where a_0, ϕ_0 and $t_\pm$ are arbitrary constants. If we take $t_+ = t_-$, the general solution (8) reduces to the special power-law solution,

$$a(t) = a_0(t-t_0)^{\frac{2\omega+2}{3\omega+4}}, \quad \phi(t) = \phi_0(t-t_0)^{\frac{2}{3\omega+4}}. \tag{10}$$

In the limit of $t \to \infty$, the general solution (8) converges to the special power-law solution (10). If the present cosmic age t is large enough, observations cannot constrain the relation between t_+ and t_-. Therefore, we keep $t_+ - t_-$ a free parameter.

In vacuum case ($\rho = 0$) we have the analytic solutions for all k. The vacuum solution for $k = 0$ is expressed as

$$a(t) = a_0 t^{\frac{1}{3(1+\alpha)}}, \quad \phi = \phi_0 t^{\frac{\alpha}{1+\alpha}}, \tag{11}$$

with

$$\alpha = \frac{1\pm\sqrt{1+2\omega/3}}{\omega}, \tag{12}$$

where we have omitted the arbitrary constant t_0 by fixing the origin of the time coordinate t. Introducing the conformal time $\eta = \int dt/a$, the vacuum solutions for $k = \pm 1$ are expressed as

$$k = +1: \quad a(\eta) = (\sin\eta)^{\frac{1-\lambda}{2}}(\cos\eta)^{\frac{1+\lambda}{2}}, \quad \phi(\eta) = (\tan\eta)^\lambda, \tag{13}$$

$$k = -1: \quad a(\eta) = (\sinh\eta)^{\frac{1-\lambda}{2}}(\cosh\eta)^{\frac{1+\lambda}{2}}, \quad \phi(\eta) = (\tanh\eta)^\lambda, \tag{14}$$

with

$$\lambda = \pm\frac{3}{3+2\omega}. \tag{15}$$

3 Modified Swiss-Cheese Model

We now consider the model of a black hole embedded in the FRW universe. We replace a sphere in the FRW spacetime with a vacuum region which contains a black hole. Here "vacuum" means $T_{\mu\nu} = 0$, and does not imply that $\mathcal{R}_{\mu\nu} = 0$ due to the existence of the BD field.

Extending Israel's junction conditions for a singular (or regular) hypersurface [6], Sakai and Maeda have studied bubble dynamics in the inflationary universe [7]. It was found that one can solve the equations of motion for the boundary without knowing the interior metric if the interior is vacuum, $T_{\mu\nu} = 0$, or has only vacuum energy (a cosmological constant), $T_{\mu\nu} = -\rho g_{\mu\nu}$. Applying this method to the present model, we find the mass and the radius of a black hole without specifying the interior metric, as we shall show below.

Let us consider a spherical hypersurface Σ which divides a spacetime into two regions, V^+ (outside) and V^- (inside). We define a unit space-like vector, N_μ, which is orthogonal to Σ and points from V^- to V^+. To describe the behavior of the boundary, we introduce a Gaussian normal coordinate system, $(n, x^i) = (n, \tau, \theta, \varphi)$, where τ is chosen to be the proper time on the boundary. Hereafter, we denote by $\Psi^\pm$ the value of any field variable Ψ defined on Σ by taking limits from $V^\pm$.

For the matter field, we consider dust (or vacuum) for V^+ and vacuum for V^-:

$$T_{\mu\nu}{}^+ = \rho u_\mu u_\nu, \quad T_{\mu\nu}{}^- = 0. \tag{16}$$

Although we assume a smooth boundary at which there is no surface density, it is not obvious that this matching is possible at all times. Therefore, we introduce the surface energy-momentum tensor on the boundary surface,

$$S_{ij} \equiv \lim_{\epsilon\to 0}\int_{-\epsilon}^{\epsilon} dn\, T_{ij} = \mathrm{diag}(-\sigma,\ \varpi,\ \varpi), \tag{17}$$

where σ and ϖ are the surface energy density and the surface pressure of Σ, respectively.

Introducing the extrinsic curvature tensor of the world hypersurface Σ, $K_{ij} \equiv N_{i;j}$, we can write the junction conditions on Σ as [7]

$$[K_{ij}]^\pm = -\frac{4\pi}{\phi}\left(S_{ij} - \frac{\omega}{2\omega+3}\mathrm{Tr}S\gamma_{ij}\right), \tag{18}$$

$$-S^j_{i\ |j} = [T^n_i]^\pm, \tag{19}$$

$$K^+_{ij}S^j_i + \frac{2\pi}{\phi}\left\{S^i_jS^j_i - \frac{\omega}{2\omega+3}(\mathrm{Tr}S)^2\right\} = [T^n_n]^\pm, \tag{20}$$

where we have defined the bracket as $[\Psi]^\pm \equiv \Psi^+ - \Psi^-$ and the vertical bar $|$ as the three-dimensional covariant derivative. The junction condition for the BD field is derived from Eq.(3) as

$$[\phi_{,n}]^\pm = -\frac{24\pi}{3+2\omega}\mathrm{Tr}S, \quad \phi^+ = \phi^-, \tag{21}$$

which implies that ϕ is continuous at Σ and inhomogeneous in V_-.

The extrinsic curvature tensor of Σ in the homogeneous side V^+ is expressed as [7]

$$K^\tau_\tau = \gamma^3\frac{dv}{dt} + \gamma v H, \tag{22}$$

$$K^\theta_\theta = \frac{\gamma(1+vHR)}{R} = \frac{\epsilon}{R}\sqrt{1+\left(\frac{dR}{d\tau}\right)^2 - R^2\left(H^2+\frac{k}{a^2}\right)}, \tag{23}$$

where

$$R = a(t)r|_\Sigma, \quad v \equiv a\frac{dr}{dt}\Big|_\Sigma, \quad \gamma \equiv \frac{\partial t}{\partial\tau}\Big|_\Sigma = \frac{1}{\sqrt{1-v^2}}, \quad \text{and} \quad \epsilon \equiv \mathrm{sign}(K^\theta_\theta) = \mathrm{sign}\left(\frac{\partial R}{\partial n}\right). \tag{24}$$

From Eqs.(16), (17), (19), (20), (22)-(24), we obtain the equations of motion:

$$\frac{dR}{dt} = \frac{dr}{d\chi}v + HR, \tag{25}$$

$$\gamma^3\frac{dv}{dt} = -\gamma\left\{\left(1-2w\right)vH - \frac{2w}{R}\frac{dr}{d\chi}\right\} + \frac{2\pi\sigma}{\phi}\left\{1+4w+\frac{(1-2w)^2}{(2\omega+3)}\right\} - \frac{\gamma^2v^2\rho}{\sigma}, \tag{26}$$

$$\frac{d\sigma}{dt} = -\frac{2\sigma(1+w)}{R}\frac{dR}{dt} + \gamma v\rho, \tag{27}$$

where $w \equiv \varpi/\sigma$.

Once initial values of R, v, and σ are given, the equations of motion (25)-(27) determine their evolution. As discussed in Ref.[7], initial values should satisfy the angular component of (18),

$$\gamma(1+vHR) - \epsilon^-\sqrt{1+\left(\frac{dR}{d\tau}\right)^2 - \frac{R_{MS}}{R}} = -\frac{8\pi\sigma R}{\phi}\frac{\omega+1+w}{2\omega+3}, \tag{28}$$

where we choose $\epsilon^- = +1$. R_{MS} is defined as

$$R_{MS} \equiv R^-(1 - g^{\mu\nu}R^-_{,\mu}R^-_{,\nu}), \tag{29}$$

where R^- is defined as $R^- \equiv \sqrt{g_{\theta\theta}}$ at Σ on the side of V^-. Because the Misner-Sharp mass is defined as [8]

$$M_{MS} \equiv \frac{R^-}{2G}(1 - g^{\mu\nu} = R^-_{,\mu} = R^-_{,\nu}), \tag{30}$$

we call R_{MS} the "Misner-Sharp radius". Note that R_{MS} is a purely geometrical quantity and independent of theories of gravitation.

If we considered a bubble in which there is no black hole (or singularity), we would have to solve the field equations with regularity condition at the centre and the boundary condition (28), as done in Ref.[7]. However, because we are interested in black hole solutions, we do not have to take such a regularity condition into account. Thus we can use Eq.(28) to determine R_{MS}.

At the initial time we suppose $v = 0$ and $\sigma = \varpi = 0$. Then R_{MS} is expressed as

$$R_{MS} = R^3\left(H^2 + \frac{k}{a^2}\right). \tag{31}$$

Let us now discuss whether v and σ remain zero during the ensuing evolution. Suppose $w = 0$, then the only nontrivial term in Eq.(26) is $\gamma^2v^2\rho/\sigma$. If v and σ evolved from zero, Eq.(28) shows $vH \sim \sigma/\phi$, so that $\rho v^2/\sigma \sim \rho v/H\phi$. Therefore, Eqs.(26) and (27) guarantee that, if $v = \sigma = 0$ at a certain time, $v = \sigma = 0$ at all time. Interestingly, this result is true only for the dust case, $\varpi/\sigma = 0$; otherwise $(2w/R)(dr/d\chi)$ in Eq.(26) would shift v from zero.

In the case of Schwarzschild spacetime, the Misner-Sharp radius coincides with the event horizon. Although it is not always true for general spacetimes, we speculate that the Misner-Sharp radius is a well-defined measure of the size of a black hole. In the next section, we show the evolution of R_{MS} for black holes in several background cosmological models.

4 Evolution of Black Holes

The evolution of the Misner-Sharp radius for the $k = 0$ dust universe is given by Eqs.(8) and (31),

$$R_{MS} = a_0^3r_0^3\left(\frac{\lambda_+}{t-t_+} + \frac{\lambda_-}{t-t_-}\right)^2(t-t_+)^{3\lambda_+}(t-t_-)^{3\lambda_-}, \tag{32}$$

where r_0 is the comoving radius of the vacuum region. Equation (32) shows that the black hole size decreases with time.

If we define the black hole mass as

$$M_{MS} \equiv \frac{\phi R_{MS}}{2}, \tag{33}$$

it coincides with the mass defined by Saida and Soda [4]. For the $k = 0$ dust universe, we obtain

$$M_{MS} = \frac{a_0^3 \phi_0}{2} \left(\frac{\lambda_+}{t - t_+} + \frac{\lambda_-}{t - t_-} \right)^2 (t - t_+)^{3\lambda_+ + \kappa_+} (t - t_-)^{3\lambda_- + \kappa_-}. \tag{34}$$

It is easy to see that Eq.(34) reduces to $M_{MS} =$ constant, if we choose $t_+ = t_-$, which is the same result as that found by Saida and Soda [4]. They also showed M_{MS} increases for $k = +1$ and decreases for $k = -1$, and concluded that the evolution of the mass depends qualitatively on the sign of the curvature of the universe. We should note, however, that their conclusion is true only for the special case $t_+ = t_-$, or equivalently, only for the asymptotic behavior of M_{MS} at $t \to \infty$.

Next, let us consider the scalar-field dominated (vacuum) universe. The evolution of R_{MS} and M_{MS} in flat, open, and closed universes are given by

$$k = 0: \qquad R_{MS} = \frac{a_0^3 r_0^3}{9(1+\alpha)^2} t^{\frac{-1-2\alpha}{1+\alpha}}, \tag{35}$$

$$M_{MS} = \frac{a_0^3 r_0^3 \phi_0}{18(1+\alpha)^2} t^{-1}, \tag{36}$$

$$k = +1: \qquad R_{MS} = \frac{r_0^3}{2} (\cos\eta)^{-\frac{3+\lambda}{2}} (\sin\eta)^{-\frac{1+\lambda}{2}} (\cos 2\eta - \sin 2\eta - \lambda), \tag{37}$$

$$M_{MS} = \frac{r_0^3}{4} (\tan\eta)^{\lambda} (\cos\eta)^{-\frac{3+\lambda}{2}} (\sin\eta)^{-\frac{1+\lambda}{2}} (\cos 2\eta - \sin 2\eta - \lambda), \tag{38}$$

$$k = -1: \qquad R_{MS} = \frac{r_0^3}{2} (\cosh\eta)^{-\frac{3+\lambda}{2}} (\sinh\eta)^{-\frac{1+\lambda}{2}} (\cosh 2\eta - \sinh 2\eta - \lambda), \tag{39}$$

$$M_{MS} = \frac{r_0^3}{4} (\tanh\eta)^{\lambda} (\cosh\eta)^{-\frac{3+\lambda}{2}} (\sinh\eta)^{-\frac{1+\lambda}{2}} (\cosh 2\eta - \sinh 2\eta - \lambda). \tag{40}$$

If we take $\omega > 500$, both α and λ are negligible. We see that R_{MS} and M_{MS} both decrease with increasing time, except in the contracting phase of the $k = +1$ universe.

5 Discussion

We have constructed a modified "Swiss cheese" model in the Brans-Dicke theory, and discussed the evolution of black holes for dust and vacuum universes. We have defined the size of a black hole by the Misner-Sharp mass, and found that it always decreases as long as the universe is in an expanding phase.

Although we have not specified the metric around a black hole, the mass and radius of a black hole we have obtained coincide with those of Saida and Soda [4], who assumed a Schwarzschild-like metric. This means that their ansatz of the Schwarzschild-like metric does not introduce a specialization of the problem. One fundamental problem remains. Although the radius R_{MS} represents a typical size of a black hole, the relation between it and the event horizon has not been demonstrated explicitly and this problem will be considered further in future work.

Acknowledgments

This work was partially supported by JSPS Programs for Research Abroad (the 1999 financial year).

References

[1] J.D. Barrow, Phys. Rev. D 46 (1992) R3227.

[2] J.D. Barrow and B.J. Carr, Phys. Rev. D 54 (1996) 3920.

[3] T. Jacobson, Phys. Rev. Lett. 83 (1999) 2699.

[4] H. Saida and J. Soda, preprint gr-qc/0006058.

[5] C.C. Dyer and C. Oliwa, preprint astro-ph/0004090.

[6] W. Israel, Nuovo Cim. 44B (1966) 1;
V.A. Berezin, V.A. Kuzmin and I.I. Tkachev, Phys. Rev. D 36 (1987) 2919.

[7] N. Sakai and K. Maeda, Prog. Theor. Phys. 90 (1993); Phys. Rev. D 48 (1993).

[8] C.W. Misner and D.H. Sharp, Phys. Rev. 136 (1964) B571;
S.A. Hayward, Phys. Rev. D 53 (1996) 1938.

Double inflation in supergravity and the primordial black hole formation

Toshiyuki Kanazawa[1]

Department of Physics, University of Tokyo, Tokyo, 113-0033, Japan

M. Kawasaki[2]

Research Center for the Early Universe (RESCEU), University of Tokyo, Tokyo, 113-0033, Japan

T. Yanagida[3]

Department of Physics and RESCEU, University of Tokyo, Tokyo 113-0033, Japan

Abstract

We study a double inflation model (a hybrid inflation + a new inflation) in supergravity and discuss the formation of primordial black holes (PBHs) with mass $\sim 10^{-20} - 10^{5} M_{\odot}$. We find that in a wide range of parameter space, we obtain PBHs which amount to $\Omega \simeq 1$, i.e., PBH dark matter. Also, we find a set of inflation parameters which produces PBHs evaporating now. Those PBHs may be responsible for antiproton fluxes observed by the BESS experiment.

1 Introduction

In the framework of supergravity the reheating temperature of inflation should be low enough to avoid overproduction of gravitinos [1, 2]. The new inflation model [3] generally predicts a very low reheating temperature and hence it is the most attractive among many inflation models [4]. However, the new inflation suffers from a fine-tuning problem about the initial condition; i.e., for a successful new inflation, the initial value of the inflaton should be very close to the local maximum of the potential in a large region whose size is much longer than the horizon of the universe.

A framework of a double inflation [5] was proposed to solve the initial value problem of the new inflation model[4] . It was shown that the above serious problem is naturally solved by supergravity effects if there exists a preinflation (e.g., hybrid inflation [6]) with a sufficiently large Hubble parameter before the new inflation [5].

In this double inflation model, density fluctuations produced by both inflations are cosmologically relevant if the e-fold number of the new inflation is smaller than ~ 60 (the total e-fold number ~ 60 is required to solve flatness and horizon problems in the standard big bang cosmology [8]). In this case, the preinflation should account for the density fluctuations on large cosmological scales [including the Cosmic Background Explorer (COBE) scales] while the new inflation produces density fluctuations on smaller scales. Although the amplitude of the fluctuations on large scales should be normalized to the COBE data [9], fluctuations on small scales are free from the COBE normalization and can have arbitrary power matched to observations. In Ref.[10], a cosmological implication of the double inflation for the large-scale structure formation was discussed. In this paper, we study primordial black hole (PBH) formation in the double inflation model. In Refs. [11, 12], the production of black hole MACHO was investigated in the double inflation model for a special case[5] . Here, we consider a wide range of parameter space where

[1] E-mail:kanazawa@utap.phys.s.u-tokyo.ac.jp
[2] E-mail:kawasaki@resceu.s.u-tokyo.ac.jp
[3] E-mail:yanagida@phys.s.u-tokyo.ac.jp
[4] Different models of double inflation were studied by various authors [7].
[5] Different models for the PBH formation have been studied in Ref. [13].

PBHs are formed[6] . In particular, we show that the double inflation creates small PBHs evaporating now if those PBHs are produced during matter-dominated (MD) era, i.e., before the end of reheating process after the new inflation. We stress that these evaporating PBHs may account for antiproton fluxes observed by the BESS experiment [14].

Throughout this paper the gravitational scale ($\simeq 2.4 \times 10^{18}$ GeV) is taken to be unity.

2 Black hole formation

In a radiation-dominated (RD) universe, PBHs are formed if the density fluctuations δ at horizon crossing satisfy a condition $1/3 \leq \delta \leq 1$ [15, 16], where δ is the over density at the horizon scale. Masses of the black holes M are roughly equal to the horizon mass,

$$M \simeq 4\pi\sqrt{\frac{3}{\rho}} \simeq 0.066 M_\odot \left(\frac{T}{\mathrm{GeV}}\right)^{-2} \left(\frac{g_*}{50}\right)^{-1/2}, \tag{1}$$

where ρ, T, and g_* are the total cosmic density, temperature, and statistical degrees of freedom at the horizon crossing, respectively.

The horizon length at the black hole formation epoch ($T = T_*$) corresponds to the scale L_* in the present universe given by

$$L_* \simeq \frac{a(T_0)}{a(T_*)} H^{-1}(T_*) \simeq 0.064 \text{ pc} \left(\frac{T_*}{\mathrm{GeV}}\right)^{-1} \left(\frac{g_*}{50}\right)^{-1/6}, \tag{2}$$

where T_0 is the temperature of the present universe. The comoving wave number corresponding to this length scale, $k_* \equiv 2\pi/L_*$, is

$$k_* \simeq 1.0 \times 10^8 \mathrm{Mpc}^{-1} \left(\frac{g_*}{50}\right)^{1/6} \left(\frac{T_*}{\mathrm{GeV}}\right). \tag{3}$$

Thus, we can write the PBH mass as a function of comoving wave number as

$$M_* \simeq 6.4 \times 10^{14} M_\odot \left(\frac{g_*}{50}\right)^{-1/6} \left(\frac{k_*}{\mathrm{Mpc}^{-1}}\right)^{-2}. \tag{4}$$

The mass fraction $\beta_*(= \rho_{BH}/\rho)$ of PBHs of mass M_* is given by [16]

$$\beta_*(M_*) = \int_{1/3}^{1} \frac{d\delta}{\sqrt{2\pi}\sigma(M_*)} \exp\left(-\frac{\delta^2}{2\sigma^2(M_*)}\right) \simeq \sigma(M_*) \exp\left(-\frac{1}{18\sigma^2(M_*)}\right), \tag{5}$$

where $\sigma(M_*)$ is the mass variance at the horizon crossing. Notice that the mass fraction $\beta_*(M_*)$ drops off sharply as $\sigma(M)$ decreases. The density of the black holes of mass M_*, $\rho_{BH}(M_*)$, is given by

$$\frac{\rho_{BH}(M_*)}{s} \simeq \frac{3}{4}\beta_*(M_*)T_*, \tag{6}$$

where s is the entropy density. Since ρ_{BH}/s is constant at $T < T_*$, we estimate the density parameter $\Omega_{BH}(M_*)$ of the black holes in the present universe as

$$\Omega_{BH}(M_*)h^2 \simeq 2.1 \times 10^8 \beta_*(M_*) \left(\frac{T_*}{\mathrm{GeV}}\right), \tag{7}$$

[6] In this paper we investigate the PBHs with mass $\sim 10^{-20} - 10^5 M_\odot$. The upper bound of $10^5 M_\odot$ is not rigorous but much heavier black holes can be produced in the present model if we take the appropriate model parameters. However, the mass of the black holes should be less than the galactic mass $\sim 10^{12} M_\odot$, otherwise the power spectrum conflicts the observations (e.g. distribution of galaxies). The lower bound $\sim 10^{-20} M_\odot$ comes from the requirement that e-fold number of the new inflation should be larger than 0.

where h is the present Hubble constant in units of 100 km/sec/Mpc. We write it as a function of PBH mass or PBH scale as

$$\Omega_{BH}(M_*)h^2 \simeq 5.4 \times 10^7 \beta_*(M_*) \left(\frac{g_*}{50}\right)^{-1/4} \left(\frac{M_*}{M_\odot}\right)^{-1/2}, \tag{8}$$

or

$$\Omega_{BH}(M_*)h^2 \simeq 2.1 \beta_*(M_*) \left(\frac{g_*}{50}\right)^{-1/6} \left(\frac{k_*}{\mathrm{Mpc}^{-1}}\right). \tag{9}$$

As for the mass of the PBHs produced during the RD era, we have a lower limit M_R. The mass of PBH produced at the reheating epoch is given by [see Eq.(1)]

$$M_R \simeq 0.066 M_\odot \left(\frac{T_R}{\mathrm{GeV}}\right)^{-2} \left(\frac{g_*}{50}\right)^{-1/2}, \tag{10}$$

where T_R is the reheating temperature. As seen later T_R is less than 10^6GeV in our inflation model, and hence M_R is larger than $\sim 10^{-13} M_\odot$. Thus, the PBHs lighter than M_R should be produced during the MD era.

In a MD universe, a relation between the comoving scale L_* and the horizon mass M_* is

$$L_* = L_R \left(\frac{M_*}{M_R}\right)^{1/3}, \tag{11}$$

where L_R is the comoving scale of the horizon at the reheating epoch. Thus, the mass M_* of the PBH produced during the MD epoch is given by

$$M_* \simeq 6.3 \times 10^{22} M_\odot \left(\frac{T_R}{\mathrm{GeV}}\right) \left(\frac{k_*}{\mathrm{Mpc}^{-1}}\right)^{-3}. \tag{12}$$

We see that small PBHs of mass, $M_* \sim 10^{-19} M_\odot$ for example, would be produced for $T_R \sim 10^6$GeV and $k_* \sim 10^{16}\mathrm{Mpc}^{-1}$. The condition for the PBH formation in a MD universe is discussed in Ref. [17], where the mass fraction of PBHs of mass M_* is estimated as [17]

$$\beta_*(M_*) \simeq 2 \times 10^{-2} \sigma(M_*)^{13/2}. \tag{13}$$

Notice that the mass fraction has a weeker dependence on σ than in the RD case [see Eq.(5)]

During the MD era, the mass fraction β_* stays constant and hence the density of the black holes of mass M_*, $\rho_{BH}(M_*)$, is given by

$$\frac{\rho_{BH}(M_*)}{s} \simeq \frac{3}{4} \beta_*(M_*) T_R. \tag{14}$$

We can write the present density parameter $\Omega_{BH}(M_*)$ for the black holes of mass M_* as

$$\Omega_{BH}(M_*)h^2 \simeq 2.1 \times 10^8 \beta_*(M_*) \left(\frac{T_R}{\mathrm{GeV}}\right). \tag{15}$$

3 Double Inflation Model

We adopt a double inflation model proposed in Ref.[5]. The model consists of two inflationary stages; the first one is called preinflation and we take a hybrid inflation [18] (see also Ref.[19]) as the preinflation. We also assume that the second inflationary stage is realized by a new inflation model [20] and its e-fold number is smaller than ~ 60. Thus, the density fluctuations on large scales are produced during the preinflation and their amplitude should be normalized to the COBE data [9]. On the other hand, the new inflation produces fluctuations on small scales. Since the amplitude of small scale fluctuations is free from the COBE normalization, we expect that the new inflation can produce large density fluctuations enough to form PBHs. We choose the predicted power spectrum to be almost scale invariant ($n_* \simeq 1$) on large cosmological scales which is favored for the structure formation of the universe [21]. On the other

hand, the new inflation gives the power spectrum which has large amplitude and shallow slope ($n_s < 1$) on small scales. Thus, this power spectrum has a large and sharp peak on the scale corresponding to a turning epoch from the preinflation to the new inflation, and we expect that PBHs are produced at that scale.

As for the detailed argument of the dynamics of our model, see Refs.[10, 11, 12].

3.1 Preinflation

First, let us briefly discuss a hybrid inflation model [18]. The hybrid inflation model contains two kinds of superfields: one is $S(x,\theta)$ and the others are a pair of $\Psi(x,\theta)$ and $\bar{\Psi}(x,\theta)$. Here θ is the Grassmann number denoting superspace. The model is based on the $U(1)_R$ symmetry under which $S(\theta) \to e^{2i\alpha}S(\theta e^{-i\alpha})$ and $\Psi(\theta)\bar{\Psi}(\theta) \to \Psi(\theta e^{-i\alpha})\bar{\Psi}(\theta e^{-i\alpha})$. The superpotential is given by [18]

$$W(S,\Psi,\bar{\Psi}) = -\mu^2 S + \lambda S\bar{\Psi}\Psi. \tag{16}$$

The R-invariant Kähler potential is given by

$$K(S,\Psi,\bar{\Psi}) = |S|^2 + |\Psi|^2 + |\bar{\Psi}|^2 + \cdots, \tag{17}$$

where the ellipsis denotes higher-order terms which we neglect in the present analysis for simplicity. We gauge the U(1) phase rotation:$\Psi \to e^{i\delta}\Psi$ and $\bar{\Psi} \to e^{-i\delta}\bar{\Psi}$. To satisfy the D-term flatness condition we take always $\Psi = \bar{\Psi}$ in our analysis.

We define $N_{\rm COBE}$ as the e-fold number corresponding to the COBE scale and the COBE normalization leads to a condition for the inflaton potential,

$$\left|\frac{V^{3/2}}{V'}\right|_{N_{\rm COBE}} \simeq 5.3 \times 10^{-4}, \tag{18}$$

where V is the inflaton potential obtained from Eqs.(16) and (17). In the hybrid inflation model, density fluctuations are almost scale invariant;

$$n_{\rm pre} \simeq 1 + 2\left(\frac{V''}{V}\right) - 3\left(\frac{V'}{V}\right)^2\Bigg|_{N_{\rm COBE}} \simeq 1 - \frac{1}{N_{\rm COBE}} \simeq 1, \tag{19}$$

where $n_{\rm pre}$ is a spectral index for a power spectrum of density fluctuations.

3.2 New inflation

Now, we consider a new inflation model. We adopt an inflation model proposed in Ref. [20]. The inflaton superfield $\phi(x,\theta)$ is assumed to have an R charge $2/(n+1)$ and $U(1)_R$ is dynamically broken down to a discrete Z_{2nR} at a scale v, which generates an effective superpotential [20, 5],

$$W(\phi) = v^2\phi - \frac{g}{n+1}\phi^{n+1}. \tag{20}$$

The R-invariant effective Kähler potential is given by

$$K(\phi,\chi) = |\phi|^2 + \frac{\kappa}{4}|\phi|^4 + \cdots, \tag{21}$$

where κ is a constant of order 1. We require that supersymmetry breaking effects make the potential energy at a vacuum vanish, and we have a relation between v and the gravitino mass $m_{3/2}$ as (for details, see Ref. [20])

$$m_{3/2} \simeq \left(\frac{n}{n+1}\right)|v|^2\left|\frac{v^2}{g}\right|^{\frac{1}{n}}. \tag{22}$$

The inflaton $\phi(x)$ (the scalar component of $\phi(x,\theta)$) has a mass m_ϕ in the vacuum with

$$m_\phi \simeq n|g|^{1/n}|v|^{2-2/n}. \tag{23}$$

The inflaton ϕ may decay into ordinary particles through gravitationally suppressed interactions, which yields reheating temperature T_R given by[7]

$$T_R \simeq 0.1 m_\phi^{3/2} \simeq 2.4 \times 10^{17}\mathrm{GeV} n^{3/2} |g|^{3/2n} |v|^{3(1-1/n)} \quad (24)$$

$$\lesssim 10^6\mathrm{GeV} \quad \text{for} \quad m_{3/2} \lesssim 1\mathrm{TeV}, n \geq 3. \quad (25)$$

An important point on the above density fluctuations is that it results in a tilted spectrum with spectral index $n_{\rm new}$ given by (see Refs. [5, 20])

$$n_{\rm new} \simeq 1 - 2\kappa. \quad (26)$$

3.3 Initial value and fluctuations of the inflaton φ

The crucial point observed in Ref. [5] is that the preinflation sets dynamically the initial condition for the new inflation. We identify the inflaton field $\varphi(x)/\sqrt{2}$ with the real part of the field $\phi(x)$. It gets an effective mass $m_{\rm eff} \sim \mu^2$ during the preinflation [5]. Thus, this inflaton φ tends to the potential minimum,

$$\varphi_{\rm min} \simeq -\frac{\sqrt{2}}{\sqrt{\lambda}} v \left(\frac{v}{\mu}\right). \quad (27)$$

Notice that $\varphi_{\rm min}$ deviates from zero due to the presence of a linear term $v^2\mu^2 S\varphi$ (see Ref. [5]). Thus, at the end of the preinflation the φ settles down to this $\varphi_{\rm min}$.

After the preinflation, the universe becomes MD because of the oscillation of the inflaton for preinflation. During the MD era between the two inflations, the energy density scales as $\propto a^{-3}$, and the new inflaton oscillates around $\varphi = 0$ with its amplitude decreasing as $\propto a^{-3/4}$. Since the scale factor increases by a factor $(\mu/v)^{4/3}$ during this era, the mean initial value φ_b of φ at the beginning of the new inflation is written as

$$\varphi_b \simeq \frac{\sqrt{2}}{\sqrt{\lambda}} v \left(\frac{v}{\mu}\right)^2. \quad (28)$$

Therefore, the amplitude of fluctuations with comoving wavelength corresponding to the horizon scale at the beginning of the new inflation is given by

$$\delta\varphi \simeq \frac{H_{\rm pre}}{2\pi}\left(\frac{H_{\rm pre}}{m_{\rm eff}}\right)^{\frac{1}{2}}\left[\left(\frac{\mu}{v}\right)^{2/3}\right]^{-3/2}\cdot\left[\left(\frac{\mu}{v}\right)^{4/3}\right]^{-3/4} \simeq \frac{H_{\rm pre}}{3^{1/4}2\pi}\left(\frac{v}{\mu}\right)^2, \quad (29)$$

where $H_{\rm pre}$ is the Hubble parameter during the hybrid inflation, $H_{\rm pre}^2 \simeq \mu^4/3$. The fluctuations given by Eq. (29) are a little less than newly induced fluctuations at the beginning of the new inflation [$\delta\varphi_{\rm new} \simeq v^2/(2\pi\sqrt{3})$]. Moreover, the fluctuations produced during the preinflation are more suppressed for smaller wavelength. Thus, we assume that the fluctuations of φ induced in the preinflation are negligible compared with fluctuations produced by the new inflation. As mentioned before, the new inflation gives the tilted spectrum on small scales [see Eq. (26)] and hence the fluctuations at the scale corresponding to the beginning of the new inflation is dominant.

Now let us estimate e-fold number which corresponds to our current horizon. The e-fold number is given by [22]

$$N_{\rm tot} = 62 - \ln\frac{k}{a_0 H_0} - \ln\frac{10^{16}\mathrm{GeV}}{V_k^{1/4}} + \ln\frac{V_k^{1/4}}{V_{\rm end}^{1/4}} - \frac{1}{3}\ln\frac{V_{\rm end}^{1/4}}{\rho_{\rm reh}^{1/4}}, \quad (30)$$

where V_k is a potential energy when a given scale k leaves the horizon, $V_{\rm end}$ that when the inflation ends, and $\rho_{\rm reh}$ energy density at the time of reheating.

We take $V_k \simeq V_{\rm end}$, and $\rho_{\rm reh}^{1/4} \simeq$ a few $\times\, T_{\rm reh}$. For $k = a_0 H_0$ (i.e., the present horizon scale), we have

$$N_{\rm tot} \simeq 67.1 + \left(\frac{5}{3} - \frac{1}{n}\right)\ln v + \frac{1}{2}\ln n + \frac{1}{2n}\ln g. \quad (31)$$

[7] The decay rate of the inflaton ϕ is discussed in Ref.[5]

In estimating $N_{\rm COBE}$ we must take into account the fact that the fluctuations induced at e-fold number less than $(2/3)\ln(\mu/v)$ before the end of the hybrid inflation reenter the horizon before the new inflation starts. Such fluctuations are cosmologically irrelevant since the new inflation produce much larger fluctuations [10]. Thus, $N_{\rm COBE}$ is given by

$$\begin{aligned} N_{\rm COBE} &= N_{\rm tot} - N_{\rm new} + \frac{2}{3}\ln\frac{\mu}{v} \\ &\simeq 67.1 + \left(\frac{5}{3} - \frac{1}{n}\right)\ln v + \frac{1}{2}\ln n + \frac{1}{2n}\ln g - N_{\rm new} + \frac{2}{3}\ln\frac{\mu}{v}. \end{aligned} \tag{32}$$

The COBE normalization in Eq. (18) should be imposed by using this $N_{\rm COBE}$.

3.4 Numerical Results

We estimate density fluctuations in the double inflation by calculating evolution of φ and σ numerically. For given parameters κ and λ, we obtain the break scale k_b and the amplitude of density fluctuations produced at the beginning of new inflation δ_b. Here, k_b^{-1} is the comoving scale corresponding to the Hubble radius at the beginning of the new inflation (a turning epoch). We can understand the qualitative dependence of (k_b, δ_b) on (κ, λ) as follows: When κ is large, the slope of the potential for the new inflation is too steep, and the new inflation cannot last for a long time. Therefore, the break occurs at smaller scales. As for δ_b, we can see from Eq.(18) that as λ gets larger, μ also gets large. In addition, we can show that

$$\delta_b \equiv \left(\frac{\delta\rho}{\rho}\right)_{{\rm new},k_b} \propto \frac{\sqrt{\lambda}\mu^2}{\kappa} \sim \frac{\lambda^{3/2}}{\kappa}, \tag{33}$$

for a given v (see Ref. [10]). Thus, we have larger δ_b for larger λ.

4 Mass Variance and Density Fluctuations

Our double inflation model predicts the amplitude of density fluctuations δ_b as a function of inflation parameters. On the other hand, the black hole abundance is expressed as a function of mass variance σ at the time when the corresponding scale enters the horizon. Therefore, when we compare the observations with the prediction of our model, we need a relation between the mass variance σ and the fluctuations δ_b.

For the power spectrum with the break scale k_b^{-1} which enters the horizon during the RD epoch, we have a relation between the mass variance and the amplitude of fluctuations as[8]

$$\delta_b \simeq \sigma_b/0.65, \tag{34}$$

(the numerical factor depends on the tilted spectral index n_s, and within the parameter range we consider, this factor lies between $0.62 \sim 0.67$.). For the power spectrum with the break scale k_b^{-1} which enters the horizon during the MD epoch, we have a relation between the mass variance and the amplitude of fluctuations as

$$\delta_b \simeq \sigma_b/2.3, \tag{35}$$

(again, the numerical factor depends on the tilted spectral index n_s, and within the parameter range we consider, this factor lies between $2.1 \sim 2.8$.).

First, let us consider PBH dark matter with $\Omega_{BH} \sim 1$ which are produced during the RD epoch. Since the density fluctuations at the break scale is dominant, and the mass fraction β_* has a sharp peak at that scale, only the PBHs of mass corresponding to the break scale are formed. For PBHs produced during the RD epoch (after reheating process), we have, from Eq. (8),

$$\Omega_{BH}h^2 \simeq 5.4\times10^7\beta_*\left(\frac{g_*}{50}\right)^{-1/4}\left(\frac{M_*}{M_\odot}\right)^{-1/2}. \tag{36}$$

[8] In Refs. [11, 12], an incorrect relation $\delta_b \simeq \sigma_b/6$ was used.

For example, if we require that the black holes with mass $\sim M_\odot$ (=MACHOs) be dark matter in the present universe, i.e. $\Omega_{BH}h^2 \sim 0.25$, we obtain $\beta_* \sim 5 \times 10^{-9}$, and from Eq.(5) we obtain

$$\sigma(M_\odot) \simeq 0.06. \tag{37}$$

From Eq.(34), we see that the break amplitude is $\delta_b \simeq 0.06/0.65 \simeq 0.092$, and we find from Eq.(4) that the break scale is $2.5 \times 10^7 \mathrm{Mpc}^{-1}$. In Fig.1, we plot the numerical results of our double inflation model for $n = 4, g = 1$, and $v = 10^{-7}$. We see a wide range of parameter space which may account for DM ($\Omega_{BH} \simeq 0.1 - 1$).

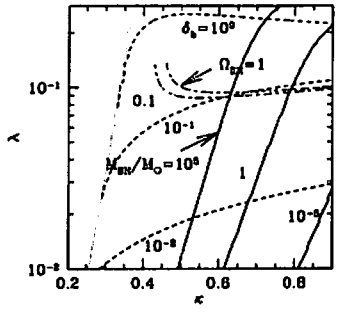


Figure 1: The amplitude of density fluctuations (δ_b) and PBH masses. Here we take $n = 4, g = 1, v = 10^{-7}$. The region on left hand side is irrelevant since a break scale exceeds our current horizon. The thick solid lines correspond to $M_{\rm BH}/M_\odot = 10^5, 1$, and 10^{-5}, from left to right. The short dash - long dash lines show $\Omega_{\rm BH} \simeq 1$ (top) and 0.1 (bottom).

Next, let us consider another interesting mass range of PBHs which are evaporating now ($M_{\rm evap} \sim 3 \times 10^{-19} M_\odot$). The PBHs of such light mass are produced during the MD epoch and from Eq.(15) we find

$$\Omega_{BH}(M_*)h^2 \simeq 4.2 \times 10^6 \sigma(M_*)^{13/2} \left(\frac{T_R}{\mathrm{GeV}}\right). \tag{38}$$

It has been reported, recently, that the BESS experiment [14] has observed antiproton fluxes, which may be explained by the evaporation of PBHs if $\Omega_{BH}h^2 \simeq 2 \times 10^{-9}$ [14]. In order to explain the BESS result by evaporating PBHs we need

$$\sigma(M_{\rm evap}) \simeq 4.4 \times 10^{-3} \left(\frac{T_R}{\mathrm{GeV}}\right)^{-2/13}. \tag{39}$$

From Eq.(24), we estimate the required fraction of the evaporating PBHs as

$$\sigma(M_{\rm evap}) \simeq 9.3 \times 10^{-6} n^{-3/13} |g|^{-3/13n} v^{-6(1-1/n)/13}. \tag{40}$$

Since the mass variance $\sigma(M)$ scales as $\sigma(M) \propto M^{(1-n_s)/6}$ [16] during the MD epoch, we obtain the mass variance at the break scale as

$$\sigma_b \simeq 9.3 \times 10^{-6} n^{-3/13} |g|^{-3/13n} v^{-6(1-1/n)/13} \left(\frac{M_b}{M_{\rm evap}}\right)^{(1-n_s)/6}, \tag{41}$$

and the amplitude δ_b is $\delta_b \simeq \sigma_b/2.3$ and $n_s \simeq 1 - 2\kappa$. In Fig.2, we plot an example of the numerical results of our double inflation model[9] for $n = 3, g = 10^{-4}$, and $v = 10^{-6}$. As shown in the figure, we have a set of inflation parameters (κ, λ) which may account for the BESS experiment.

[9] For the case of $n = 4$, we do not find a consistent parameter region with the BESS experiment.

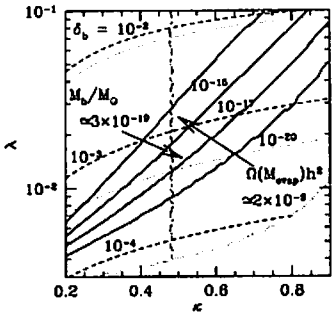


Figure 2: The amplitude of density fluctuations (δ_b) and PBH masses. Here we take $n = 3, g = 10^{-4}, v = 10^{-6}$. The region on down right corner is irrelevant since a break scale is too small and does not cross horizon during inflation. The thick solid lines correspond to the break masses $M_b/M_\odot = 10^{-15}, 10^{-17}, 3\times 10^{-19}$, and 10^{-20}, from left to right. The short dash - long dash line shows $\Omega(M_{\rm evap})h^2 \simeq 2\times 10^{-9}$. The two thin solid lines show the region $\Omega(M_b)h^2 \simeq 1$ (top) and $\simeq 2\times 10^{-9}$ (bottom), respectively.

5 Conclusions and Discussions

In this paper we have studied the formation of PBHs by taking a double inflation model in supergravity. We have shown that in a wide range of parameter space PBHs are produced of various masses. These PBHs are interesting since, for example, they may be identified with MACHOs ($M \sim M_\odot$) in the halo of our galaxy. Or, they may be PBHs which are just evaporating now ($M \sim 10^{-19} M_\odot$). Such black holes are one of the interesting candidates for the sources of antiproton fluxes recently observed in the BESS detector [14].

The dark matter PBHs play a role of the cold dark matter on the large scale structure formation. The scales of the fluctuations for PBH formation themselves are much smaller than the galactic scale and thus we cannot see any signals for such fluctuations in $\delta T/T$ measurements. However, the PBHs may be a source of gravitational waves. If the PBHs dominate dark matter of the present universe, some of them likely form binaries. Such binary black holes coalesce and produce significant gravitational waves [23] which may be observable in future detectors.

Acknowledgment

T. K. is grateful to K. Sato for his continuous encouragement. A part of work is supported by Grant-in-Aid of the Ministry of Education and by Grant-in-Aid, Priority Area "Supersymmetry and Unified Theory of Elementary Particles" (#707).

References

[1] S. Weinberg, Phys. Rev. Lett. **48**,1303 (1982).

[2] M. Yu. Khlopov and A.D. Linde, Phys. Lett. **B138**, 265 (1984) ; J. Ellis, E. Kim and D.V. Nanopoulos, Phys. Lett. **B145**, 181 (1984); J. Ellis, G.B. Gelmini, J.L. Lopez, D.V. Nanopoulos and S. Sarker, Nucl. Phys. **B373**, 399 (1992); M. Kawasaki and T. Moroi, Prog. Theor. Phys. **93**, 879 (1995).

[3] A. Albrecht and P.J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982); A.D. Linde, Phys. Lett. **B108**, 389 (1982).

[4] See for example, A.D. Linde, *Particle Physics and Inflationary Cosmology* (Harwood, Chur, Switzerland, 1990).

[5] K.I. Izawa, M. Kawasaki and T. Yanagida, Phys. Lett. **B411**, 249 (1997).

[6] A. D. Linde, Phys. Rev. **D49**, 748 (1994).

[7] G. Lazarides and Q. Shafi, Phys. Lett. **B148**, 35 (1984); L.A. Kofman, A.D. Linde and A.A. Starobinsky, Phys. Lett. B **157**, 361 (1985); J. Silk and M.S. Turner, Phys. Rev. D **35**, 419 (1987); D. Polarski and A.A. Starobinsky, Nucl. Phys. B **385**, 623 (1992); P. Peter, D. Polarski and A.A. Starobinsky, Phys. Rev. D **50**, 4827 (1994); S. Gottlöber, J. P. Mücket and A. A. Starobinsky, Astrophys. J. **434**, 417 (1994); R. Kates, V. Müller, S. Gottlöber, J. P. Mücket and J. Retzlaff, Mon. Not. Roy. Astron. Soc. **277**, 1254 (1995); J.A. Adams, G.G. Ross and S. Sarkar, Nucl. Phys. B **503**, 405 (1997); J. Lesgourgues and D. Polarski, Phys. Rev. D **56**, 6425 (1997); J. Lesgourgues D. Polarski and A.A. Starobinsky, Mon. Not. Roy. Astron. Soc. **297**, 769 (1998); M. Sakellariadou and N. Tetradis, hep-ph/9806461; G. Lazarides and N. Tetradis, Phys. Rev. **D58**, 123502 (1998); C. Panagiotakopoulos and N. Tetradis, Phys. Rev. **D59**, 083502 (1999).

[8] See for example, E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Reading MA, 1990).

[9] C.L. Bennett et al., Astrophys. J. **464**, L1 (1996).

[10] T. Kanazawa, M. Kawasaki, N. Sugiyama, and T. Yanagida, Phys. Rev. **D61**, 023517 (2000).

[11] M. Kawasaki, N. Sugiyama and T. Yanagida, Phys. Rev. **D57**, 6050 (1998).

[12] M. Kawasaki and T. Yanagida, Phys. Rev. **D59**, 043512 (1999).

[13] J. Garcia-Bellido, A. Linde and D. Wands, *ibid* **54**, 6040 (1996); J. Yokoyama, Astron. Astrophys. **318**, 673 (1997); Phys. Rev. **D58**, 083510 (1998); P. Ivatov, P. Naselsky and I. Novikov, Phys. Rev. **D50**, 7173 (1994); P. Ivatov, Phys. Rev. **D57**, 7145 (1998); E. Kotok and P. Naselsky, Phys.Rev. **D58**, 103517 (1998).

[14] K. Yoshimura et al. Phys. Rev. Lett. **75**, 3792 (1995).

[15] B. J. Carr and S. W. Hawking, Mon. Not. Roy. Astron. Soc. **168**, 399 (1974); B.J. Carr, Astrophys. J. **201**, 1 (1975); B.J. Carr, J.H. Gilbert and J.E. Lidsey, Phys. Rev. **D50**, 4853 (1994).

[16] A.M. Green and A.R. Liddle, Phys. Rev. **D56**, 6166 (1997); A.M. Green and A.R. Liddle and A. Riotto, *ibid* **D56**, 7554 (1997).

[17] M. Y. Khlopov and A. G. Polnarev, Phys. Lett. **97B**, 383 (1980); A. G. Polnarev and M. Y. Khlopov, Sov. Astron. **26**, 391 (1983); Sov. Phys. Usp. **28**, 213 (1985).

[18] G. Dvali, Q. Shafi, and R. K. Shaefer, Phys. Rev. Lett. **73**, 1886 (1994); E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart, and D. Wands, Phys. Rev. **D49**, 6410 (1994).

[19] C. Panagiotakopolous, Phys. Rev. **D55**, 7335 (1997); A. Linde and A. Riotto, Phys. Rev. **D56**, R1841 (1997).

[20] K.I. Izawa and T. Yanagida, Phys. Lett. **B393**, 331 (1997); K. Kumekawa, T. Moroi and T. Yanagida, Prog. Theor. Phys. **92**, 437 (1994).

[21] M. White, D. Scott, J. Silk and M. Davis, Mon. Not. R. Astron. Soc. **276**, L69 (1995).

[22] A. R. Liddle and D. H. Lyth, Phys. Rep. **231**, 1 (1995).

[23] T. Nakamura, M. Sasaki, T. Tanaka and K.S. Thorne, Astrophys. J. **487**, L139 (1997).

Q-ball formation and its cosmology

Shinta Kasuya,[1]

Research Center for the Early Universe, University of Tokyo, Tokyo 560-0043, Japan

Abstract

We consider the Q-ball formation, and its implication to cosmology in the context of gauge-mediated SUSY breaking in minimal supersymmetric standard model. In addition to the usual stable Q ball investigated in the literature, we obtain a new type of a stable Q ball. It is so-called gravity-mediation type of Q ball, but also stable against the decay into nucleons, since the energy per unit charge is equal to gravitino mass $m_{3/2}$, which can be smaller than nucleon mass in the gauge-mediation mechanism. We consider the cosmological consequences in this new Q-ball scenario, and find that this new type of the Q ball can be considered as the dark matter and the source for the baryon number of the universe simultaneously.

1 Introduction

In the supersymmetric standard models, Affleck-Dine (AD) mechanism [1] is the most promising procedure for the baryogenesis. In the minimal supersymmetric standard model (MSSM), there are many flat directions consist of squarks and sleptons [2], which can be identified as the AD field. Its potential is almost flat but slightly lifted up by effects of supersymmetry (SUSY) breaking. For the mechanism of SUSY breaking, there are two famous scenarios: the gravity- and gauge-mediated SUSY breakings. It was believed that the AD field stays at large field value at the inflationary stage, and, when the Hubble parameter becomes as small as the AD scalar mass after inflation, rolls down homogeneously its potential with rotation, making the baryon number of the universe.

Recently, however, it was revealed that the AD field does not evolve homogeneously, but feels spatial instabilities, which grow nonlinear and form into Q balls [3]. A Q ball is a kind of the nontopological soliton, whose stability is guaranteed by the nonzero charge Q [4, 5]. In the context of the AD baryogenesis, the charge Q is the baryon number B. In the gauge-mediated SUSY breaking, a Q ball is stable against the decay into nucleons, provided that its charge is large enough so that its energy per unit charge is less than nucleon mass: $E_Q/Q \simeq m_\phi Q^{-1/4} < 1$ GeV [6, 3], where $m_\phi \sim 1$ TeV, is the mass of AD field. Therefore, the Q ball itself can be a candidate for the dark matter. On the other hand, in the gravity-mediated SUSY breaking, the energy of a Q ball per unit charge is essentially constant: $E_Q/Q \simeq m_\phi > 1$ GeV [7]. Thus, it should decay into nucleons, and the dark matter will be lightest supersymmetric particles (neutralinos) produced in Q-ball decays. In either case, the dark matter and the baryon number of the universe can be explained simultaneously by the Q-ball formation through the AD mechanism.

In all the previous studies of Q balls in the context of SUSY breaking, the effects of gauge- and gravity-mediations are considered separately. However, it is natural to have both effects in the gauge-mediated SUSY breaking scenario, since the gravity-mediation effects will dominate over the gauge ones at the large field value. Cosmology including AD baryogenesis in such more realistic SUSY breaking scenario was considered in Ref. 8. There, AD field is regarded as a homogeneously rotating condensate, but we notice that it will form Q balls due to the instabilities of the field. Particular interest is the smallness of the gravitino mass comparing with that in the gravity-mediation scenario. It usually ranges between 100 keV and 1 GeV. Therefore, we can imagine a new type of a stable Q ball: the profile is the same as that in the gravity-mediation, but its energy per unit charge is less that 1 GeV because of the small gravitino mass. Here, we study the cosmological consequences of Q balls (baryogenesis and the dark matter) in the gauge-mediated SUSY breaking, taking into account the gravity-mediation effects at large field value.

[1] E-mail:kasuya@resceu.s.u-tokyo.ac.jp

2 Properties of Q balls in the gauge-mediated SUSY breaking

To be concrete, let us assume the following potential for the AD field,

$$V(\Phi) = m_\phi^4 \log\left(1 + \frac{|\Phi|^2}{m_\phi^2}\right) + m_{3/2}^2 |\Phi|^2 \left[1 + K \log\left(\frac{|\Phi|^2}{M_*^2}\right)\right], \tag{1}$$

where $m_{3/2}$ is the gravitino mass, $K(<0)$ term a one-loop correction, M_* the renormalization scale, and we assume that the second term should be neglected for very small field value. This is nothing but the sum of the potentials for the gauge- and gravity-mediation mechanisms studied previously [3, 7, 9, 10]. However, as we mentioned earlier, the gravitino mass is considerably smaller. The second term will dominate the potential when

$$\phi \gtrsim \phi_{\rm eq} \equiv \sqrt{2}\frac{m_\phi^2}{m_{3/2}}, \tag{2}$$

where $\Phi = \phi \exp(i\omega t)/\sqrt{2}$. In this case, a new type of a stable Q ball is produced [11]. Its property is very similar to that in the gravity-mediation, such as [7, 12]

$$\begin{aligned} R_Q &\simeq |K|^{-1/2} m_{3/2}^{-1}, & \omega &\simeq m_{3/2}, \\ \phi &\simeq |K|^{3/4} m_{3/2} Q^{1/2}, & E_Q &\simeq m_{3/2} Q, \end{aligned} \tag{3}$$

but, as can be seen from the last equation, it is stable against the decay into nucleons. In the opposite case, the Q-ball properties are the same as in the gauge-mediation only [3]:

$$\begin{aligned} R_Q &\simeq m_\phi^{-1} Q^{1/4}, & \omega &\simeq m_\phi Q^{-1/4}, \\ \phi &\simeq m_\phi Q^{1/4}, & E_Q &\simeq m_\phi Q^{3/4}. \end{aligned} \tag{4}$$

The energy per unit charge can be treated from unified viewpoint [11]. The largest charge of the Q ball formed depends linearly on the initial charge density of the AD field as [9, 10]

$$Q \simeq \beta \frac{q(0)}{m_{3/2}^3} \simeq \beta \left(\frac{\phi(0)}{m_{3/2}}\right)^2, \tag{5}$$

where $\beta \lesssim 1$ [12], and we use $q = \omega\phi^2 \simeq m_{3/2}\phi^2$. It can be understood by estimating Q-ball charge in the standard way. The charge is given by

$$Q = \int d^3x \omega \phi^2 = \left(\frac{\pi}{2}\right)^{3/2} \omega \phi_0^2 R^3 = \beta' \left(\frac{\phi_0}{m_{3/2}}\right)^2, \tag{6}$$

where we assume the Gaussian profile ansatz [7], $\phi = \phi_0 \exp(-r^2/R^2)$, which is a very good approximation, and use relations (3), and $\beta' \simeq 2 \times 10^3 (|K|/0.01)^{-3/2}$. The discrepancy of coefficients β and β' comes from the fact that $\phi(0) \neq \phi_0$ and there are more than one Q balls with charges of the same order of magnitude as the largest.

Inserting it into Eq.(2), we obtain

$$m_{3/2} \gtrsim (2\beta)^{1/4} m_\phi Q^{-1/4}. \tag{7}$$

The right hand side of this equation is identical to the expression for the energy per unit charge of the gauge-mediation besides the factor of order unity. The energy per unit charge of the Q ball is written as

$$\frac{E_Q}{Q} = \begin{cases} m_\phi Q^{-1/4} & \phi \lesssim \phi_{\rm eq} \\ m_{3/2} & \phi \gtrsim \phi_{\rm eq} \end{cases}, \tag{8}$$

which is shown in Fig. 1. The gap on the boundary should disappear and both sides of the curves will be smoothly connected because Q balls formed in this region are not the exact type of either (3) or (4), but will show properties between them.

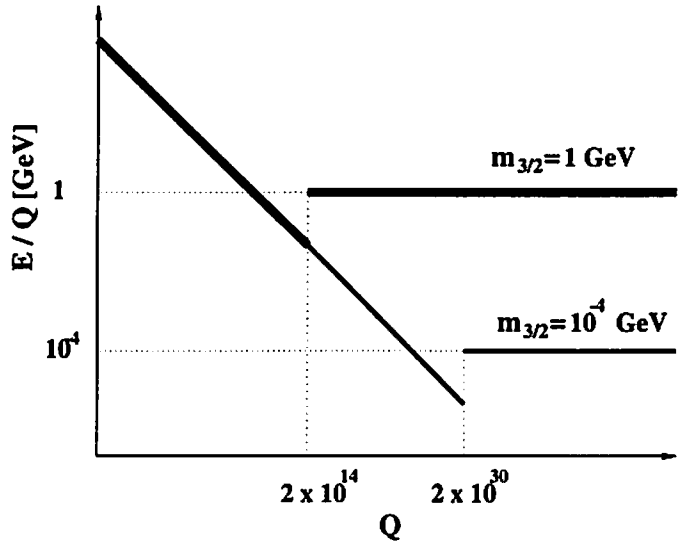


Figure 1: Dependence of the energy per unit charge on the charge of the Q ball. Gauge-mediation type of Q ball are formed with less charge, while gravity-mediation type has larger with E/Q fixed.

3 Q-ball formation through Affleck-Dine mechanism

We consider the full nonlinear equations of motion of the Affleck-Dine scalar field in order to see the formation of the Q-ball through the Affleck-Dine mechanism by numerical lattice simulations [9, 10, 12]. It is shown that large Q-balls are really produced by the fragmentation of the condensate of a scalar field whose potential is very flat, as in the supersymmetric standard theory.

The most important result is that, once Q-balls are formed, almost all the charges are absorbed into them, and only a tiny fraction of the charge is carried by the relic AD field, but its amplitude is very small and fluctuates so that it may not be possible to regard it as a condensate.

Also, we find that the charges stored within Q-balls grows as the initial charge density of the AD condensate increases. In other words, the larger the initial amplitude of the AD condensate is, the larger charge Q balls have.

4 Baryogenesis by the charge evaporation from Q-ball surfaces

Since Q balls are stable even for $\phi \gtrsim \phi_{\rm eq}$, where the gravity-mediation effect is crucial, the baryon number in the universe should be explained by the baryons evaporated from Q balls, as is the same as for the gauge-mediation type [13], because almost all the charges are absorbed into Q-balls during their formation. The evaporation rate of the Q ball is [13]

$$\Gamma_{\rm evap} \equiv \frac{dQ}{dt} = -\alpha \mu T^2 4\pi R_Q^2, \tag{9}$$

where μ is a chemical potential of the Q ball, which is estimated as $\mu \simeq \omega$ because ω is energy of ϕ-field inside the Q ball. Although the mass of the (free) AD particle m_ϕ is affected by thermal corrections, which should be changed as $m_\phi \to m_\phi(T) \sim T$, at $T \gg m_\phi$, the gravitino mass is not affected, since particles coupled to the AD field are decoupled from thermal bath when the AD field has a large vacuum expectation value. At $T \gtrsim m_\phi$, large numbers of ϕ-particles are in thermal bath outside Q balls, so $\alpha \sim 1$. On the other hand, since only light quarks are in thermal bath at $T \lesssim m_\phi$, the corresponding cross section is highly suppressed by the heavy gluino exchanges, and $\alpha \simeq (T/m_\phi)^2$.

However, if the rate of the charge diffusion from the "atmosphere" of the Q ball is smaller than the evaporation rate, chemical equilibrium will established there, which results in the suppression of the evaporation [14]. The diffusion rate is [14]

$$\Gamma_{\rm diff} \equiv \frac{dQ}{dt} = -4\pi\zeta R_Q D \mu_Q T^2, \tag{10}$$

where $D = a/T$ with $a \simeq 4-6$, and $\zeta \sim 1$.

The time scale of charge transportation is determined by the diffusion when $\Gamma_{\rm diff} \lesssim \Gamma_{\rm evap}$. It holds for

$$T \gtrsim T_* \equiv \begin{cases} a^{1/3} m_\phi Q^{-1/12} & \phi \lesssim \phi_{\rm eq} \\ a^{1/3} |K|^{1/6} (m_{3/2} m_\phi^2)^{1/3} & \phi \gtrsim \phi_{\rm eq} \end{cases}. \tag{11}$$

In the opposite case, the diffusion effect is negligible, and the charge transportation rate is completely determined by the evaporation rate [11].

Therefore, total amount of the charge evaporated from the Q ball is

$$\Delta Q \simeq \begin{cases} 12\pi A M a^{2/3} m_\phi^{-1} Q^{1/12} & \phi \lesssim \phi_{\rm eq} \\ 12\pi A M a^{2/3} |K|^{-2/3} (m_{3/2} m_\phi^2)^{-1/3} & \phi \gtrsim \phi_{\rm eq} \end{cases}, \tag{12}$$

where $A \approx 0.2$ and $M \simeq 2.4 \times 10^{18}$ GeV.

Provided that the initial charge of the Q ball is larger than the evaporated charge, we regard that the Q ball survives from evaporation, and contributes to the dark matter of the universe: $Q_{\rm init} \gtrsim \Delta Q$, that is,

$$Q_{\rm init} \gtrsim \begin{cases} 10^{18} \left(\frac{m_\phi}{\rm TeV}\right)^{-12/11} & \phi \lesssim \phi_{\rm eq} \\ 10^{19} \left(\frac{m_{3/2}}{\rm GeV}\right)^{-1/3} \left(\frac{m_\phi}{\rm TeV}\right)^{-2/3} & \phi \gtrsim \phi_{\rm eq} \end{cases}, \tag{13}$$

where we set $a = 4$ and $|K| = 0.01$.

Now we can relate the baryon number and the amount of the dark matter in the universe. As mentioned above, the baryon number of the universe should be explained by the amount of the charge evaporated from Q balls, ΔQ, and the survived Q balls become the dark matter. If we assume that Q balls do not exceed the critical density of the universe, i.e., $\Omega_Q \lesssim 1$, and the baryon-to-photon ratio as $\eta_B \sim 10^{-10}$,

$$Q \lesssim \begin{cases} 10^{23} \left(\frac{m_\phi}{\rm TeV}\right)^{-3} & \phi \lesssim \phi_{\rm eq} \\ 10^{21} \left(\frac{m_{3/2}}{\rm GeV}\right)^{-4/3} \left(\frac{m_\phi}{\rm TeV}\right)^{-2/3} & \phi \gtrsim \phi_{\rm eq} \end{cases}, \tag{14}$$

In order for the Q ball to be stable against the decay into nucleons,

$$\begin{cases} Q \gtrsim 10^{12} \left(\frac{m_\phi}{\rm TeV}\right)^4 & \phi \lesssim \phi_{\rm eq} \\ m_{3/2} \lesssim 1 \text{ GeV} & \phi \gtrsim \phi_{\rm eq} \end{cases}. \tag{15}$$

For the *usual* gauge-mediation type of the Q ball, we obtain the allowed region for explaining the baryon number of the universe, using three constraints, i.e., the first lines of Eqs. (13), (14), and (15). Figure 2 shows the allowed region on (Q, m_ϕ) plane. The shaded regions represent that this type stable Q balls are created, and the baryon number of the universe can be explained by the mechanism mentioned above. Furthermore, this type of stable Q balls contribute crucially to the dark matter of the universe at present, if the Q balls have the charge given by the thick line in the figure [12].

For the *new* type of the Q ball, namely gravity-mediation type, there is an additional constraint. Rewriting Eq.(7), we have

$$Q \gtrsim 2\beta \times 10^{12} \left(\frac{m_\phi}{\rm TeV}\right)^4 \left(\frac{m_{3/2}}{\rm GeV}\right)^{-4}. \tag{16}$$

Combining this constraint with the second lines of Eqs.(13), (14), and (15), we obtain the allowed region for the new type of the stable Q ball explaining the baryon number of the universe. Figure 3 shows the

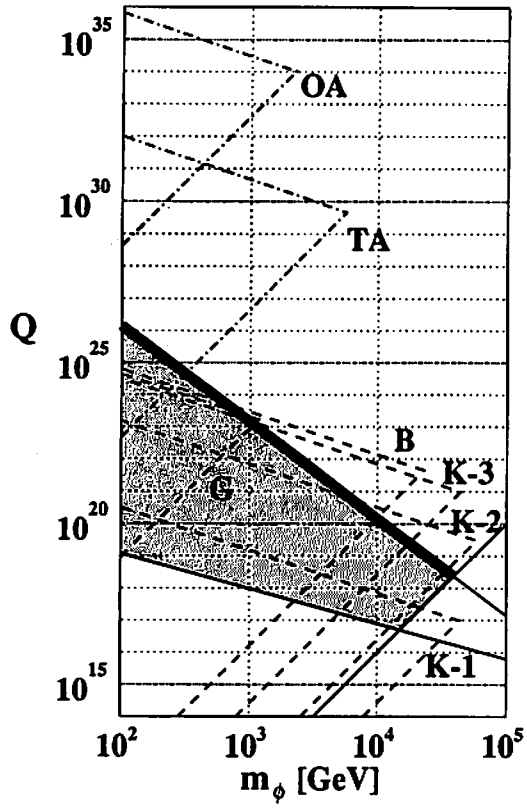


Figure 2: Summary of constraints on (Q, m_ϕ) plain for the *old* type of the Q ball. We also show the regions currently excluded by BAKSAN (B), Gyrlyand (G), and Kamiokande (K-1, K-2, K-3), and to be searched by the Telescope Array Project (TA) and OWL-AIRWATCH (OA) in the future.

allowed region on $(Q, m_{3/2})$ plane for $m_\phi = 1$ TeV. The shaded regions represent that the new type of stable Q balls are created, and the baryon number of the universe can be explained by the mechanism mentioned above. Furthermore, the new type of stable Q balls contribute crucially to the dark matter of the universe at present, if the Q balls have the charge given by the thick line in the figure.

5 Q-ball detection

One may wonder if these stable Q balls can be detected. When a Q ball collides with nucleons, they enter the surface layer of the Q ball, and dissociate into quarks, which are converted into squarks. In this process, Q balls release $\sim$ 1 GeV energy per collision by emitting soft pions. This process is the basis for the Q-ball detections [15, 16], which is called (Kusenko-Kuzmin-Shaposhnikov-Tinyakov) KKST process in the literature. It occurs for electrically neutral Q balls (ENQB). For electrically positively charged Q balls (EPCQB), the KKST process is strongly suppressed by Coulomb repulsion, and only electromagnetic processes will take place. For electrically negatively charged Q balls (ENCQB), the both KKST and electromagnetic processes occur, but the former is dominant, which is essentially the same as for ENQBs.

With the discussions similar to Ref. [16], we can restrict the parameter spaces (Q, m_ϕ) or $(Q, m_{3/2})$ by the several experiments. Figures 2 and 4 show the results for ENQBs. Lower left regions are excluded by the various experiments. For both cases, the allowed charges are $Q \sim 10^{25}$, where the initial amplitude of the AD field should be larger than ϕ_{eq}, which implies that the new type will be produced. Thus, we

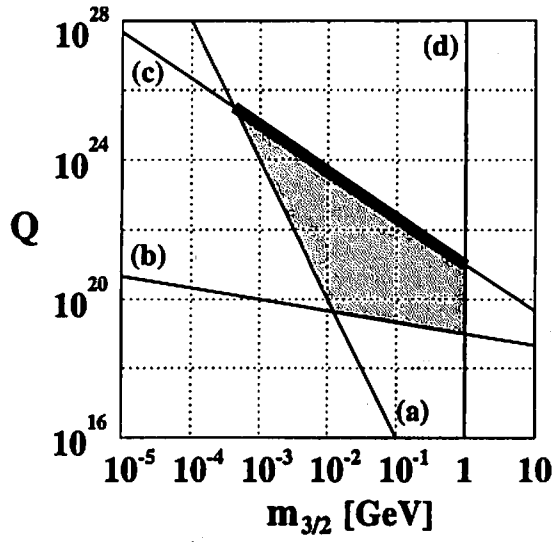


Figure 3: Summary of constraints on $(Q, m_{3/2})$ plain for the *new* type of the Q ball.

conclude that the stable gravity-mediation type Q balls with small gravitino mass can be both the dark matter and the source of the baryons, and future experiments such as the Telescope Array Project or the OWL-AIRWATCH detector may detect this type of Q balls with an interesting gravitino mass ~ 100 keV.

For EPCQBs with $Z = 1$, similar constraints as for ENQBs are put by the MACRO experiment [12].

6 Reheating temperature after inflation and thermal bath before reheating

In order for this scenario to work, the reheating temperature after inflation should be very low such as $10^3 - 10^4$ GeV [12]. However, this is the advantage because we can avoid the gravitino problem. In addition, thermal effects on the dynamics of the AD condensate may be less effective for such low temperatures. As is well known, the maximum temperature after inflation is achieved just after the inflaton oscillation begins, far before the reheating time ($T = T_{RH}$). If we take into account these thermal effects, the inflation scale should be very low: $V^{1/4} \lesssim 10^{12}$ GeV for the safe Q-ball formation [12]. In fact, the new inflation scenario can provide both very low scale inflation and very low reheating temperature.

7 Conclusion

We have considered the Q-ball formation and its cosmological consequences in the gauge-mediated SUSY breaking scenario. We have also showed the existence of the new type of stable Q balls. Many properties are the same as the gravity-mediation type of Q ball, but it is stable against the decay into nucleons, since the energy per unit charge is equal to the gravitino mass $m_{3/2}$, which can be smaller than nucleon mass of 1 GeV in the gauge-mediation mechanism. In either type, because of its stability, it can be a nice candidate for the dark matter of the universe. The baryons are produced only by evaporation from Q balls, since (almost) all the baryons are trapped in Q balls during their formation. We have found that the the *new* type of the Q ball with $Q \sim 10^{25}$ for $m_{3/2} \sim 10^{-3}$ GeV and $m_\phi = 1$ TeV, can account for both the dark matter and the baryon number of the present universe, and such Q balls may be detected by the future experiments.

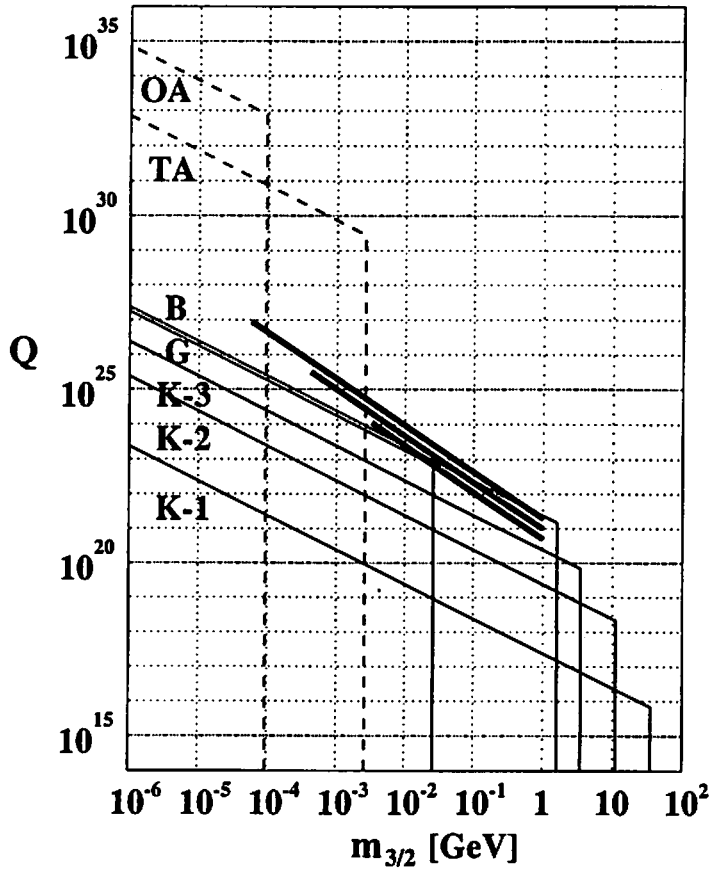


Figure 4: Restrictions by several experiments on $(Q, m_{3/2})$ plain for $|K| = 0.01$. We show the excluded region by the same experiment as for the *usual* gauge-mediation type.

References

[1] I. Affleck and M. Dine, Nucl. Phys. **B249**, 361 (1985).

[2] M. Dine, L. Randall, and S. Thomas, Nucl. Phys. **B458**, 291 (1996).

[3] A. Kusenko and M. Shaposhnikov, Phys. Lett. **B418**, 46 (1998).

[4] S. Coleman, Nucl. Phys. **B262**, 263 (1985).

[5] A. Kusenko, Phys. Lett. **B404**, 285 (1997).

[6] G. Dvali, A. Kusenko, and M. Shaposhnikov, Phys. Lett. **B417**, 99 (1998).

[7] K. Enqvist and J. McDonald, Phys. Lett. **B425**, 309 (1998); Nucl. Phys. **B538**, 321 (1999); Phys. Lett. **B440**, 59 (1998).

[8] A. de Gouvêa, T. Moroi, and H. Murayama, Phys. Rev. **D56**, 1281 (1997).

[9] S. Kasuya and M. Kawasaki, Phys. Rev. **D61**, 041301 (2000).

[10] S. Kasuya and M. Kawasaki, Phys. Rev. **D62**, 023512 (2000).

[11] S. Kasuya and M. Kawasaki, Phys. Rev. Lett. **85**, 2677 (2000).

[12] S. Kasuya and M. Kawasaki (in preparation).

[13] M. Laine and M. Shaposhnikov, Nucl. Phys. **B532**, 376 (1998).

[14] R. Banerjee and K. Jedamzik, Phys. Lett. **B484**, 278 (2000).

[15] A. Kusenko, V. Kuzmin, M. Shaposhnikov, and P.G. Tinyakov, Phys. Rev. Lett. **80**, 3185 (1998).

[16] J. Arafune, T. Yoshida, S. Nakamura, and K. Ogure, Phys. Rev. **D62**, 105013 (2000), and references therein.

Cognition of the Universe through a Model

Masafumi Seriu[1]

Department of Physics, Faculty of Engineering,
Fukui University, Fukui, 910-8507, Japan

Abstract

We analyze a fundamental issue in cosmology: To what extent we can rely on a cosmological model to infer the real history of the Universe. The spectral-representation scheme provides us with a suitable framework for studying this type of problems. Simulating the relation between the real Universe and its model, we focus on two spatial geometries that are very close to each other and investigate the time evolution of the spectral distance between them with the help of the spectral evolution equations. We then formulate the criteria for a model to be a suitable one.

1 Introduction

We here study one of the fundamental issues in cosmology: To what extent a certain cosmological model traces the real history of the Universe. This significant problem has been raised and studied for a long time and it is often called the "model-fitting problem in cosmology" or the "averaging problem in cosmology" [1]. We have not yet reached a satisfactory understanding of this issue.

This problem is rooted in the very nature of cosmology: The total quantity of the geometrical information contained in our Universe is too enormous to be perceived by human intelligence, and we are forced to truncate the information to suit our brain capacity. In a sense this is what cosmology is all about; it is our attempt to grasp the reality, which is full of complexity, through a format of cognition, viz. a model. In cosmology, thus, there is in principle an inevitable discrepancy between the real Universe and its model.

Now the true problem arises when we attempt to guess the past or future of the Universe: There is no *a priori* guarantee that the model remains a valid one when it is evolved in time backward or forward according to the Einstein equation, a set of highly non-linear differential equations. We need a guarantee that the inevitable discrepancy between the reality and its model does not develop so much within the time scale of concern. Unless this uneasiness is cleared we cannot safely state anything meaningful for the past or future of our Universe. One might symbolically state the problem depicted here as *"Whether cosmology is possible?"*

There are several points to be clarified when tackling this problem. Among them, we start our analysis by focusing on

(a) How to quantify "discrepancy" or "closeness".

(b) How to take into account the influence of the "coarseness" of truncating information.

(c) How to take into account the types of observations performed for collecting information.

In a series of work, we have constructed a formalism for analyzing geometrical structures which is especially suitable for studying the type of problems such as the issue of reality and its model. The formalism naturally takes care of the above (a)-(c) also. In Section 2, we will summarize the basic setting of this formalism, or the *spectral scheme* for simplicity. In Section 3, the spectral evolution equations, one of the fundamental ingredients of the spectral scheme, are introduced. In Section 4, the spectral scheme is applied to analyze the issue of reality and model. We derive the discrepancy equations for two nearby geometries and formulate the criteria for a model to be a suitable one. Section 5 is devoted for discussions.

[1] E-mail:mseriu@edu00.f-edu.fukui-u.ac.jp

2 Spectral scheme

We here summarize the spectral representation scheme for discussing the issue of models in cosmology in later sections.

The basic idea of the spectral scheme is easily understand: To characterize the geometry of a space, we use the vibration modes of a field on the space. One may associate this idea with the well-known phrase "Can one hear the shape of the drum?" [2] Imitating this phrase, our idea may be symbolically stated as "Let us hear the shape of the Universe!" [3]. However, it requires a series of non-trivial constructions to realize this primitive idea in the form of a scheme. The spectral scheme mainly consists of the following three elements:

[1] Spectral distance d_N [3, 4]

[2] Spectral space (*"Space of Spaces"*) $\mathcal{S}_N$ [5]

[3] Spectral evolution equations [6]

Now they are explained one by one.

To be specific, let us consider an eigenvalue problem of the Laplacian Δ on a compact Riemannian space (Σ, h), $\Delta f = -\lambda f$.[2] We then get the *spectra*, namely the set of eigenvalues $\{\lambda_n\}_{n=0,1,2,\cdots} := \{0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq \cdots\}$. We also obtain the set of real-valued eigenfunctions corresponding to the spectra, $\{f_n\}_{n=0,1,2,\cdots}$, normalized in a usual manner.

It can be said that *Issue (a)* in Sec.1 had been the main obstacle to make progress in the analysis of the relation of the reality and its model; we first need a suitable mathematical *language* to talk about the discrepancy or closeness of two shapes. Therefore the first thing to do is to introduce such a measure of closeness between two geometries.

Let $\mathcal{G}$ and $\mathcal{G}'$ be two spatial geometries. Let $\{\lambda_m\}_{m=0}^{\infty}$ and $\{\lambda'_n\}_{n=0}^{\infty}$ be the spectra for $\mathcal{G}$ and for $\mathcal{G}'$, respectively. Then the measure of closeness between $\mathcal{G}$ and $\mathcal{G}'$ can be introduced by comparing $\{\lambda_m\}_{m=0}^{\infty}$ and $\{\lambda'_n\}_{n=0}^{\infty}$ [5, 7]:

$$d_N(\mathcal{G}, \mathcal{G}') := \sum_{n=1}^{N} \mathcal{F}\left(\frac{\lambda'_n}{\lambda_n}\right) \quad , \tag{1}$$

where $\mathcal{F}(x)$ $(x > 0)$ is any continuous function which satisfies

$$\begin{aligned} \mathcal{F}(1) &= 0 \ , \\ \mathcal{F}(y) &> \mathcal{F}(x) \ \text{ for } \ y > x \geq 1 \ , \\ \mathcal{F}(1/x) &= \mathcal{F}(x) \ . \end{aligned} \tag{2}$$

For practical applications, it is preferable to impose further assumptions on $\mathcal{F}$ that it is smooth at $x = 1$ and $\mathcal{F}''(1) > 0$ (See Eq.(11) below). On dimensional grounds, the lower (higher) spectrum reflects the larger (smaller) scale properties of the geometry. Comparing only the low-lying spectra up to the N-th spectrum λ_N, thus, corresponds to comparing $\mathcal{G}$ and $\mathcal{G}'$ in a coarse-grained manner, down to the scale $O(\lambda_N^{-1/2})$, neglecting the difference in the smaller scale behavior. Thus the measure of closeness $d_N(\mathcal{G}, \mathcal{G}')$ introduced in this way naturally takes into account the scale-dependent property of observed geometry (*Issue (b)* in Sec.1).

We also note that the elliptic operator (e.g. Δ) is in principle interpreted as the observational apparatus for measuring geometry. Thus the choice of the operator decides the aspects of geometry to be compared in defining the closeness between the spaces. In this manner, our measure of closeness automatically takes care of the apparatus-dependent nature of observed geometry, too (*Issue (c)* in Sec.1).

It is convenient to set $\mathcal{F}$ to be $\mathcal{F}_1(x) = \frac{1}{2}\ln\frac{1}{2}(\sqrt{x} + 1/\sqrt{x})$. Then Eq.(1) becomes [3, 4]

$$d_N(\mathcal{G}, \mathcal{G}') = \frac{1}{2}\sum_{n=1}^{N} \ln\frac{1}{2}\left(\sqrt{\frac{\lambda'_n}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda'_n}}\right) \ . \tag{3}$$

[2] Note that (Σ, h) is a space, and not a spacetime.

We mainly focus on this form hereafter.

It turns out that $d_N(\mathcal{G},\mathcal{G}')$ in Eq.(3) does not satisfy the triangle inequality [3] so that one further step is required to justify regarding it "distance". It is achieved as follows.

Let us consider the collection of all D-dimensional compact Riemannian manifolds, *Riem*, equipped with d_N. In the present setting, it is natural to identify the so-called isospectral manifolds [8]. From the viewpoint of spectral scheme, there is no surprise in the isospectral manifolds. It is obvious that geometry cannot be perfectly specified by just a single kind of observational apparatus (Δ)[5, 4]. Now we introduce a space

$$\mathcal{S}_N = (Riem, d_N)/_\sim \ , \tag{4}$$

where $/_\sim$ indicates the identification of the isospectral manifolds[3] . Now one can prove that the space $\mathcal{S}_N$ forms a metrizable space, so that d_N is essentially regarded as a distance [5]. Therefore we are entitled to call d_N the *spectral distance* of order N; $\mathcal{S}_N$ may be called the *spectral space* of order N, or the space of spaces.

We note that $\mathcal{S}_N$ provides us with the concept of "closeness" between spaces as a 'function' of the observational scale (N) and the observational apparatus (Δ). In this manner, the idea of the "scale-dependent topology" [9, 10], which is a fascinating, but difficult-to-formulate concept, is automatically materialized in the spectral scheme.

With its several desirable properties, the space of spaces $\mathcal{S}_N$ may serve as a basic arena for spacetime physics; it is particularly suitable for analyzing situations when more than two spaces are involved.

With the help of the space of spaces $\mathcal{S}_N$, it is now possible to picture the relation between the real Universe and its model as follows: Each point in $\mathcal{S}_N$ represents a certain geometry viewed with the apparatus Δ by the scale-limit λ_N. A set of models is represented by the set of points distributed throughout $\mathcal{S}_N$. The optimal model $\mathcal{G}'$ for reality $\mathcal{G}$ is viewed as the point among them closest to the point representing $\mathcal{G}$ [7]. Thus, the real geometry $\mathcal{G}$ and its model geometry $\mathcal{G}'$ are represented by two points that are very close to each other in $\mathcal{S}_N$. What we should do now is, then, to study the time evolution of the spectral distance $d_N(\mathcal{G},\mathcal{G}')$ between them. To do so, we need the evolution equations for the spectra of the universe since $d_N(\mathcal{G},\mathcal{G}')$ is defined solely in terms of the spectra.

In this way, the spectral scheme is completed by supplying the spectral evolution equations, which will be explained in the next section.

3 Spectral distance and its time evolution

Being the spectral scheme at hand, we are now free to discuss the discrepancy between two arbitrary universes.

Let us investigate the time evolution of the spectral distance between two universes $\mathcal{G}$ and $\mathcal{G}'$. We need to choose the time-slicing for each of two spacetimes. First it is notable that the metrizable structure of $\mathcal{S}_N$ induces a metrizable structure on $Slice(\mathcal{M},g)$, the space of all spatial slices of a given spacetime $(\mathcal{M},g)$ [7]. In other words, we can talk about the distance between two different slicings of the same spacetime, by means of the metrizable structure of $\mathcal{S}_N$. Now suppose we want to compare two spacetimes $(\mathcal{M},g)$ and $(\mathcal{M}',g')$. By a similar procedure as above, it is possible to select an optimal time-slicing for each spacetime, such that $(\mathcal{M},g)$ and $(\mathcal{M}',g')$ look as similar as possible (as a time sequence of spatial geometries). We assume that this procedure has been already completed so that a suitable time t for each spacetime has been selected (and identified with each other).

Now, by taking the time derivative of Eq.(3), we get

$$\dot{d}_N(\mathcal{G},\mathcal{G}') = \frac{1}{4}\sum_{n=1}^{N} \frac{\frac{\lambda_n'}{\lambda_n}-1}{\frac{\lambda_n'}{\lambda_n}+1}\left(\ln\frac{\lambda_n'}{\lambda_n}\right)^{\cdot} \ . \tag{5}$$

In order to proceed further, thus, the evolution equations for the spectra are needed. They can be derived by the following consideration: The spatial geometry (Σ,h) is evolved in time by evolving the initial value (h_{ab},K_{ab}) according to the Einstein equations in the Hamiltonian form. (Here h_{ab} and K_{ab}

[3] More rigorously, $\mathcal{S}_N$ is defined as the completion of the right-hand side of Eq.(4) by means of the latter's metrizable-space structure. [5]

are the spatial metric and the extrinsic curvature, respectively.) Then the spectra of (Σ, h) also evolve in time accordingly. What we need to do is essentially to study the response of the spectra to the change of the spatial metric h_{ab}. In this manner the evolution equations for the spectra can be derived, which has been achieved in Ref.[6]. The spectral evolution equations so obtained are regarded as the spectral representation of the Einstein theory.

Just like the original Einstein equations, the spectral evolution equations are complicated system of equations. They are in the form of hierarchy equations as is expected from the fact that the spectra are not local quantities but they reflect the global information of spatial geometry. This new kind of representation of the Einstein theory should have several advantages for analyzing a certain type of problems; further investigations on the properties of the spectral evolution equations are awaited.

Before going further, we here introduce convenient notations that are very similar to those in quantum theory: Let $\mathcal{A}$ and $\mathcal{A}_{ab}$ be any function and any symmetric tensor field, respectively, on a spatial geometry (Σ, h). Then, we define $\langle \mathcal{A} \rangle_{mn} := \int_\Sigma f_m \, \mathcal{A} \, f_n \, \sqrt{h}$ and $\langle \mathcal{A}_{ab} \rangle_{mn} := \frac{1}{\sqrt{\lambda_m \lambda_n}} \int_\Sigma \partial^a f_m \, \mathcal{A}_{ab} \, \partial^b f_n \, \sqrt{h}$. Furthermore, $\langle \mathcal{A} \rangle_n := \langle \mathcal{A} \rangle_{nn}$ and $\langle \mathcal{A}_{ab} \rangle_n := \langle \mathcal{A}_{ab} \rangle_{nn}$.

Here we only show a typical equation among the system of spectral evolution equations (see Ref.[6] for more details):

$$(\ln \lambda_n)^{\cdot} = -2H_n \ . \tag{6}$$

Here

$$H_n := H + \iota_n + \alpha_n \ , \tag{7}$$

with the quantities

$$\begin{aligned} H: &= \frac{1}{D-1}\frac{\dot{V}}{V} \ , \\ \iota_n: &= \frac{1}{D-1}\langle K - K_{av} \rangle_n - \frac{D-3}{4(D-1)}\frac{1}{\lambda_n}\langle \Delta K \rangle_n \ , \\ \alpha_n: &= \langle \epsilon_{ab} \rangle_n \ . \end{aligned} \tag{8}$$

(D= the spacetime dimension; $K = K_a{}^a$; $\sum'$ indicates that the zero-mode is not included in the summation; Q_{av} denotes the spatial average of a scalar quantity Q.) Here H corresponds to the Hubble parameter; ι_n and α_n can be interpreted as the inhomogeneity and the anisotropy of the geometry at the scale $\lambda_n^{-1/2}$. Thus, H_n is interpreted as the effective Hubble parameter at the scale $\lambda_n^{-1/2}$.

4 Discrepancy equations for two nearby geometries

Spectral scheme in principle allows the analysis of the discrepancy of two arbitrary geometries. Here, however, we focus on the simplest case when two universes are very close to each other in $\mathcal{S}_N$. We only consider the case when the discrepancy between two geometries $\mathcal{G}$ and $\mathcal{G}'$ is captured by a small difference in their spatial metrics,

$$\gamma_{ab} := h'_{ab} - h_{ab} \ , \tag{9}$$

where h_{ab} and h'_{ab} are the spatial metrics for $\mathcal{G}$ and $\mathcal{G}'$, respectively. This case simulates the relation of the real Universe with its model.

The discrepancy γ_{ab} in their metrics induces the discrepancy of their n-th spectra [6],

$$\delta \ln \lambda_n := \frac{\lambda'_n - \lambda_n}{\lambda_n} = -\langle \overline{\gamma}_{ab} \rangle_n - \frac{1}{2}\langle \gamma \rangle_n \ , \tag{10}$$

where $\gamma := h^{ab}\gamma_{ab}$ and $\overline{\gamma}_{ab} := \gamma_{ab} - \frac{1}{2}\gamma h_{ab}$. Then the leading term in Eq.(1) becomes [6]

$$d_N(\mathcal{G}, \mathcal{G}') = \frac{1}{2}\mathcal{F}''(1) \sum_{n=1}^{N} \left(\langle \overline{\gamma}_{ab} \rangle_n + \frac{1}{2}\langle \gamma \rangle_n \right)^2 \ . \tag{11}$$

We note that the leading term of $d_N(\mathcal{G}, \mathcal{G}')$ is not affected by the choice of the function $\mathcal{F}(x)$ except for the unimportant numerical factor. Whenever two geometries are very close to each other in $\mathcal{S}_N$, the

leading term of the spectral distance $d_N(\mathcal{G},\mathcal{G}')$ is universally given by Eq.(11), *irrespective either of the detailed form of the spectral distance or of the gravity theory.*

It is convenient to introduce a vector-like notation $\vec{\gamma}$ with its n-th component being

$$\gamma_n := \langle \overline{\gamma}_{ab} \rangle_n + \frac{1}{2}\langle \gamma \rangle_n \ .$$

Then Eq. (11) is expressed in a neat form as

$$d_N(\mathcal{G},\mathcal{G}') = \frac{1}{16}\vec{\gamma}\cdot\vec{\gamma} \ , \tag{12}$$

where a standard Euclidean inner-product is implied and we have set $\mathcal{F}''(1) = 1/8$ corresponding to Eq.(3).

Taking the time-derivative of Eq.(12), we get

$$\dot{d}_N(\mathcal{G},\mathcal{G}') = \frac{1}{4}\vec{\gamma}\cdot\delta\vec{H} \ , \tag{13}$$

where $\delta\vec{H}$ is a vector-like quantity whose n-th component is $\delta H_n := H_n' - H_n$, namely the difference of the effective Hubble parameter between $\mathcal{G}'$ and $\mathcal{G}$ at the scale $\lambda_n^{-1/2}$. [4]

Equation (13) shows that *the relative directions of $\vec{\gamma}$ and $\delta\vec{H}$ (in R^N) determine the sign of $\dot{d}_N(\mathcal{G},\mathcal{G}')$.*

Taking the time-derivative once more, Eq.(13) yields

$$\ddot{d}_N(\mathcal{G},\mathcal{G}') = \frac{1}{2}\delta\vec{H}\cdot\delta\vec{H} + \frac{1}{4}\vec{\gamma}\cdot\delta\dot{\vec{H}} \ . \tag{14}$$

We find out that *$\delta\vec{H}$, the discrepancy of the effective Hubble parameters up to order N, always induces a repulsive effect and that the sign and the magnitude of the quantity $\vec{\gamma}\cdot\delta\dot{\vec{H}}$ are crucial for the sign of $\ddot{d}_N(\mathcal{G},\mathcal{G}')$.* Equation (14) can be modified further. After some manipulations, Eq.(14) reduces to [6]

$$\begin{aligned}\ddot{d}_N(\mathcal{G},\mathcal{G}') \simeq \quad & \frac{1}{2}\delta\vec{H}\cdot\delta\vec{H} - \frac{1}{4}\Big(\sum_{n=1}^{N}\gamma_n\Big)\delta\{(1+q)H^2\} \\ & -\frac{1}{4}\vec{\gamma}\cdot\delta\left\{2H\vec{\iota} + (D-1)H\vec{\alpha} + \vec{c} + \frac{(D-3)}{2(D-2)}\frac{1}{\alpha}\vec{\mathcal{M}}\right\} \ .\end{aligned} \tag{15}$$

Here, $q := -(1+\frac{\dot{H}}{H^2})$ is the deceleration parameter; $\vec{\iota}$ and $\vec{\alpha}$ are the vectors whose n-th components are ι_n and α_n, respectively; $\vec{c}$ is the vector whose n-th component is $\langle r_{ab}\rangle_n$ with $r_{ab} := R_{ab} - \frac{1}{D-1}Rh_{ab}$; $\alpha := \frac{c^3}{16\pi G}$; $\vec{\mathcal{M}}$ is the vector whose n-th component is $\mathcal{M}_n$ defined by

$$\mathcal{M}_n := \frac{1}{D-1}\{\langle \rho - \rho_{\mathrm{av}}\rangle_n + \frac{D-1}{D-3}\langle p - p_{\mathrm{av}}\rangle_n\} - \frac{D-3}{4(D-1)}\frac{1}{\lambda_n}\{\langle \Delta\rho\rangle_n + \frac{D-1}{D-3}\langle \Delta p\rangle_n\} \ ,$$

which measures the inhomogeneity of the matter density (ρ) and its pressure (p) at the scale of $\lambda_n^{-1/2}$.

For two nearby geometries $\mathcal{G}$ and $\mathcal{G}'$ given in the form of initial data, we can formulate the criteria for them to form a pair of "similar" histories: Let τ be the typical time-scale we are interested in. Then, the criteria become

(C1) $d_N(\mathcal{G},\mathcal{G}')$ is small.

(C2) $\tau\dot{d}_N(\mathcal{G},\mathcal{G}')$ is negative, or at least, non-negative and small.

(C3) $\tau^2\ddot{d}_N(\mathcal{G},\mathcal{G}')$ is negative, or at least, non-negative and small.

[4] For definiteness, we define δQ as Q for $\mathcal{G}'$ (the second entry of $d_N(\cdot,\cdot)$) minus Q for $\mathcal{G}$ (the first entry of $d_N(\cdot,\cdot)$). Whole of the theory is so constructed as invariant under the exchange of the roles of $\mathcal{G}$ and $\mathcal{G}'$, as is inferred from Eqs.(1) and (2). When appropriate, we write $\delta\{\ \cdot\ \}$ (rather than $\delta(\ \cdot\ )$) to avoid a confusion with the δ-function.

It is suggestive to imagine that $\mathcal{G}$ and $\mathcal{G}'$ represent the real Universe and its model. It is regarded as a great progress achieved by the spectral scheme that we can now talk about the relation of two different cosmological histories like in the above criteria, though of course these criteria are not written in the form directly connected to our daily cosmological observations.

We can do more explicit analysis by setting definite models. Let $\mathcal{G}$ be a regular T^3 and $\mathcal{G}'$ be a T^3-space obtained by perturbing $\mathcal{G}$. We consider only the terms linear in γ_{ab} (See Eq.(9)). In this case we can make use of the standard linear perturbation theory for structure formation and directly estimate the time evolution of $d_N(\mathcal{G}, \mathcal{G}')$, without resorting to the discrepancy equations Eqs.(13) and (14).

We get the estimation [11],

$$d_N(\mathcal{G}, \mathcal{G}') \sim \quad A\xi(t)\, t^{-\frac{2(1-\nu)}{1+\nu}} + Bk_N{}^2\, \eta(t)\, t^{-\frac{2(1-3\nu)}{3(1+\nu)}} \quad . \tag{16}$$

Here, $\nu = p/\rho$; A and B are appropriate positive numerical factors that make the expression dimension-free; $\xi(t) := 1 - Ct^{\frac{1+3\nu}{1+\nu}}$ and $\eta(t) := 1 - C't^{-\frac{2(1+3\nu)}{3(1+\nu)}}$ with C and C' being appropriate positive numerical factors that make $\xi(t)$ and $\eta(t)$ are positive for $t > \exists t_0$. We note that $\xi(t) \to 0$ and $\eta(t) \to 1$ as $t \to \infty$.

This is just the consideration of the linear regime, but it is notable that the spectral scheme enables us to reach explicit estimations such as Eq.(16).

5 Discussion

In this paper we analyzed the issue of the relation of the real Universe and its model in the framework of the spectral scheme. Though the analysis we have just sketched is still on a primitive stage, it is already clear that the spectral scheme provides us with a very powerful tool for dealing with problems that require quantitative description of geometrical information, or problems in which more than two geometries are involved. Indeed we have derived several important properties regarding the relation of the reality and its model which would not have been available without the spectral scheme. Here we realize the importance of preparing a suitable language first before talking about an entangled situation.

The problem we have just analyzed turns to a deeper issue than the original one when we put it in a wider context, as the issue of the relation among the triad (*Reality*, *Model*, *Dynamics*), which arises quite universally in various fields of science. Putting emphasis on the procedure of truncating information, it also can be regarded as the issue of classifying the objects into types, which may be our typical strategy of perceiving the complexity of reality. It may be concisely stated as the *issue of types*. The analysis presented here can be the first step for discussing the issue of types in natural sciences.

* * *

The author thanks the Ministry of Education, the Government of Japan and Inamori Foundation, Japan for financial support. A great portion of this work was completed when the author was a member of Institute of Cosmology, Tufts University. On this occasion, he would like to express his heartfelt thanks to the members of the institute for providing him with comfortable research environment.

References

[1] For instance, G.F.R. Ellis, in *"Proceedings of the Tenth International Conference on General Relativity and Gravitation"*, edited by B. Bertotti, F. De Felice and A. Pascolini (Reidel, Dordrecht, 1984). See also, A. Krasiński, *Inhomogeneous Cosmological Models* (Cambridge University Press, Cambridge, 1997), Chapter 8, and the references therein.

[2] M. Kac, Am. Math. Mon. **73**(4) (1966), 1.

[3] M. Seriu, Physical Review **D53** (1996) 6902.

[4] M. Seriu, in *"Proceedings of the 8th Workshop on General Relativity and Gravitation"* (K.Oohara et. al. (eds.), Niigata University, 1999).

[5] M. Seriu, Communications in Mathematical Physics **209** (2000) 393.

[6] M. Seriu, Physical Review **D62** (2000) #023516.

[7] M. Seriu, General Relativity and Gravitation **32** (2000) 1473.

[8] See e.g., I. Chavel, *Eigenvalues in Riemannian Geometry* (Academic Press, Orland, 1984).

[9] M. Visser, Physical Review **D41** (1990) 1116.

[10] M. Seriu, Physics Letters **B319** (1993) 74.

[11] M. Seriu, "Evolution of the discrepancy between the real Universe and its model", submitted for publication.

Screw Instability in Black-Hole Magnetospheres

Tomoya Matsuoka [1], Akira Tomimatsu [2]
Department of Physics,Nagoya University,Chikusa-ku,Nagoya 464-8602,Japan
and
Masaaki Takahashi [3]
Department of Physics and Astronomy,
Aichi University of Education Kariya,Aichi 448-8542,Japan

Abstract

We study the screw-instability for a force-free magnetic field around a Kerr black-hole. When the magnetic field lines rotate slowly and connect the event horizon to the astrophysical load at a far distant region, the criterion of the instability is found to be essentially determined by the ratio of angular velocities of the magnetic field and the hole. Then, by comparing the efficiency with the Blandford-Znajek mechanism, we discuss two different states of the magnetic activity in black-hole magnetospheres.

1 Introduction

As a central engine of active galactic nuclei (AGNs), a black hole magnetosphere involving axisymmetric magnetic field and accreting plasma has been extensively investigated. The spirally accreting plasma can generate the toroidal component of the magnetic field and the rotation of the field lines, so that we can expect that a persistent energy release of the rotational energy of the black hole by the Blandford-Znajek mechanism [1, 2]. Further, we would be able to expect plasma flares induced by some kind of plasma instabilities caused by the growing field. In fact, there is an observational suggestion that some flare-like events cause time variations of the radiation from AGNs [5, 6, 7]. Here, as an interesting example of current driven instabilities, we will investigate the screw instability. If the pitch of the field lines is smaller than a limit given by the Kruskal-Shafranov criterion (K-S criterion), which is the condition for the screw instability, the magnetic configuration becomes unstable against long-wave mode perturbations, and then the plasma column quickly twists into a helical shape. This causes a sudden release of magnetic free-energy into plasma's kinetic energy.

Recently, Gruzinov [3] has discussed the screw instability as a trigger of flares from a black hole magnetospheres; and Li [4] has shown that the Blandford-Znajek mechanism can only work efficiency close to the hole, because the long field line becomes unstable for the screw instability. Of cause, K-S criterion is applied to their works. However the criterion is formulated without any effect of the strong gravity and the field line rotation. The formulation for linear perturbations of the black hole magnetosphere is needed to study the instability.

Our main purpose in this report is to derive criterions of black-hole magnetospheres. In §2, by using the energy principle method [8, 9], we consider linear perturbations of the force-free magnetic field around

[1] t_matsu@allegro.phys.nagoya-u.ac.jp
[2] atomi@allegro.phys.nagoya-u.ac.jp
[3] takahasi@phyas.aichi-edu.ac.jp

Kerr black hole. The considering instability would be happened in a global scale, so that the effects of the field line rotation could be important rather than the gravity. Then, in §3, we extend K-S criterion to the slow rotating magnetic field at a far distance region (using flat spacetime approximation), and we derive an additional restriction on the magnetic field configuration for the screw-unstable. Finally, in §4, we discuss implications of the additional criterion for magnetic field lines threading the horizon, by using the fact that the current streaming is conserved along the field lines in the force-free magnetic field. Throughout this paper, we use the geometric units with $G = c = 1$.

2 Instability Condition for Rotating Force-Free field

In order to discuss the energy principle method in Kerr spacetime, we present the basic equations of a force-free electromagnetic field with the 3+1 formalism [2]. The metric of spacetime, in terms of Boyer-Lindquist time t and any arbitrary spatial coordinates x^j, has the form

$$ds^2 = -\alpha^2 dt^2 + h_{jk}(dx^j + \beta^j dt)(dx^k + \beta^k dt), \tag{1}$$

where α and β^j denote the lapse function and the shift function. The angular velocity of the dragging of inertial frames is denoted by $\omega \equiv -\beta^\phi$. The Maxwell equations and the force-free condition for zero-angular-momentum observers (ZAMOs) are follows

$$(\partial_t + L_\beta)\, \boldsymbol{B} = -\nabla \times \alpha \boldsymbol{E}, \tag{2}$$

$$(\partial_t + L_\beta)\, \boldsymbol{E} = \nabla \times \alpha \boldsymbol{B} - \alpha \boldsymbol{j}, \tag{3}$$

$$\rho \boldsymbol{E} + \boldsymbol{j} \times \boldsymbol{B} = 0, \tag{4}$$

where L_β denotes the Lie derivative along $\boldsymbol{\beta}$, $\boldsymbol{B}$ and $\boldsymbol{E}$ are the magnetic and electric fields, ρ and $\boldsymbol{j}$ are the electric charge and the current densities multiplied by 4π. Here, we use the terminology "force-free" means as the magnetically dominated limit of magnetohydrodynamics (MHD). We introduce the plasma velocity $\boldsymbol{v}$ defined by

$$\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B} = 0. \tag{5}$$

From these equations, we obtain

$$(\partial_t + L_\beta)\, \boldsymbol{B} = \nabla \times (\boldsymbol{v} \times \alpha \boldsymbol{B}), \tag{6}$$

$$\partial_t (\boldsymbol{B} \times \boldsymbol{E}) + \boldsymbol{B} \times L_\beta \boldsymbol{E} + (L_\beta \boldsymbol{B}) \times \boldsymbol{E} + \alpha (\nabla \cdot \boldsymbol{E})\, \boldsymbol{E} = \boldsymbol{B} \times (\nabla \times \alpha \boldsymbol{B}) + \boldsymbol{E} \times (\nabla \times \alpha \boldsymbol{E}). \tag{7}$$

A equilibrium magnetic field satisfy the stream equation [10]:

$$\frac{\Omega_F - \omega}{\alpha} (\nabla \Psi)^2 \frac{d\Omega_F}{d\Psi} + \nabla \cdot \left\{ \frac{\alpha}{\varpi^2} \left[1 - \left(\frac{\Omega_F - \omega}{\alpha} \varpi \right)^2 \right] \nabla \Psi \right\} + \frac{16\pi^2}{\alpha \varpi^2} I \frac{dI}{d\Psi} = 0, \tag{8}$$

where Ψ denotes the magnetic flux, and $\varpi^2 = h_{\phi\phi}$. Ω_F and I denote the angular velocity of the field lines and the current streaming into the hole:

$$\Omega_F = \alpha v^{\hat{\phi}}/\varpi + \omega, \tag{9}$$

$$2I = -\alpha \varpi B^{\hat{\phi}}. \tag{10}$$

In the force-free magnetic field, Ω_F and I is conserved along the field line.

To consider linear perturbations of (5)-(7) for a given equilibrium magnetic field, we denote small deviations of the electric, magnetic, and velocity fields from the equilibrium states by $\boldsymbol{E}_1, \boldsymbol{B}_1$ and $\boldsymbol{v}_1$. Further, we set the t- and ϕ- dependence to the form of $\exp(im\phi - i\sigma t)$, then we can introduce the small displacement $\boldsymbol{\xi}$ by $\boldsymbol{v}_1 = -i\sigma\boldsymbol{\xi}$. From perturbed equations of (5) and (6), we obtain $\boldsymbol{B}_1$ and $\boldsymbol{E}_1$:

$$\boldsymbol{B}_1 = \eta\,(\nabla \times \alpha\boldsymbol{\zeta})\,,\ \boldsymbol{E}_1 = \eta\frac{\Omega_F - \omega}{\alpha}\,[(\nabla \times \alpha\boldsymbol{\zeta}) \times \boldsymbol{e}_\phi] + i\sigma\boldsymbol{\zeta}, \tag{11}$$

where $\eta \equiv \sigma/(\sigma - m\Omega_F)$. For the simplification, we assume that $\Omega_F(\Psi)$ is constant all over the field lines. Substitution these expressions to perturbed equation of (7) and making inner products with the complex conjugate of the displacement $\boldsymbol{\xi}$, we can write as the following form

$$(\sigma a_i\,\xi^i + b_i\,\xi^i)(\sigma\overline{a_j}\,\overline{\xi^j} + \overline{b_j}\,\overline{\xi^j}) = G_{ij}\,\xi^i\,\overline{\xi^j}, \tag{12}$$

where a_i, b_i, G_{ij} are derivative operators including $\boldsymbol{B}_0$, which we determined the explicit form. To apply the energy principle method, we define an energy integral W as the integral of the right-hand side of equation (12):

$$W = \int G_{ij}\,\xi^i\,\overline{\xi^j}d^3r, \tag{13}$$

where this integral correspond to the potential energy associated with the small displacement in the case of no-rotating field ($\Omega_F = 0$). If $W < 0$, then σ is purely imaginary and the magnetic field is unstable. the instability problem for a given magnetosphere can replace the problem whether a displacement achieving $W < 0$ exist, or not. Thus, we only minimize W with a trial function ξ^j by using the variational principle. Now, we are ready to investigate the plasma instability around a Kerr black hole.

3 Slowly Rotating Limit

Let us now derive the criterion of the screw instability for a slowly rotating magnetic field. The screw instability would occur far from the black hole, because the long-wave mode dominates in the instability. So, it seems that the effect of strong gravity do not appear in extended criterion. Then, we can use the flat spacetime approximation $\alpha = 1, \omega = 0$ for a considering magnetic field. However, in spite of the flat spacetime approximation, the effect of the black hole is included in the streaming currents along the magnetic field, which can be specified at the event horizon.

We assume that an equilibrium force-free magnetic field is cylindrical within a plasma confined radius ϖ_0, and is denoted by $\boldsymbol{B}_0 = (0, B_{\hat{\phi}}(\varpi), B_{\hat{z}})$ in the cylindrical coordinates (ϖ, ϕ, z), and $|B_{\hat{z}}| << |B_{\hat{\phi}}|$. In order to evaluate the displacement $\boldsymbol{\xi}$, we add the assumption for z- dependence as $\boldsymbol{\xi} \propto \exp(im\phi + ikz - i\sigma t)$, where k is an arbitrary positive number. By using a variational principle, the minimization of W with respect to ξ can be carried out algebraically. W_{min} is denoted by the minimized W and has the following form

$$W \propto \int_0^{\varpi_0} d\varpi\,(f|\xi'|^2 + g|\xi|^2)\,, \tag{14}$$

where f, g are real function of $B_{\hat{\phi}}, B_{\hat{z}}$, and Ω_F [4].

[4]The explicit form of f, g is

First of all, for non-rotating magnetic field lines ($\Omega_F = 0$), we confirm that K-S criterion [11] is derived from this energy integral (14). In this case with $m = -1, k\varpi \ll 1$, W_{min} becomes simply

$$W_{min} \propto \int_0^{\varpi_0} \left[k^2\varpi(k\varpi B_{\dot{z}} - B_{\dot{\phi}})(3k\varpi B_{\dot{z}} + B_{\dot{\phi}})\right] d\varpi, \tag{15}$$

where $\xi = 1$ for $\varpi < \varpi_0$, and $\xi = 0$ for $\varpi > \varpi_0$ are chosen. Then, by $W < 0$, we can derive K-S criterion

$$B_{\dot{\phi}} > k\varpi B_{\dot{z}}. \tag{16}$$

Next, we consider the validity of the Kruskal-Shafranov criterion (16). As an example for the rotation dominant on the origin of the toroidal magnetic field, we consider a case that the relationship between $B_{\dot{\phi}}$ and $B_{\dot{z}}$ can be written as

$$B_{\dot{\phi}} \sim \varpi\Omega_F B_{\dot{z}}, \tag{17}$$

and define a non-dimensional parameter $\epsilon \equiv \varpi\Omega_F B_{\dot{z}}/B_{\dot{\phi}}$. Then, from the integrand of equation (14), under the approximation $\epsilon = O(1)$, $k\varpi \ll 1$, $m = O(1)$, we obtain

$$f|\xi'|^2 + g|\xi|^2 \sim \frac{\epsilon^2(B_{\dot{z}})^2\varpi}{m^2+\epsilon^2}\left[\left(a' + \frac{a}{\varpi}\right)^2 + \left(b' + \frac{b}{\varpi}\right)^2\right] > 0, \tag{18}$$

where we set $\xi = a + bi$. We can see that any magnetic field satisfying equation (17) are stable against the screw instability, even if the field configuration satisfy the K-S criterion. Thus, a new criterion for the screw instability, which would be concerned the field line rotation, should be needed.

Let us derive the additional criterion. To examine the effect of the field line rotation, we set the magnetic field configuration to become unstable when $\Omega_F = 0$

$$\begin{cases} B^{\dot{\phi}} > k\varpi B^{\dot{z}}, & (0 \le \varpi < \varpi_0), \\ B^{\dot{\phi}} = k\varpi B^{\dot{z}}, & (\varpi = \varpi_0), \end{cases} \tag{19}$$

$f = \frac{1}{C}\left[(mB_{\dot{\phi}} + k\varpi B_{\dot{z}})^2 + (\varpi\Omega_F)^2 f_1 + (\varpi\Omega_F)^4 f_2\right]$,

$f_1 = -(m^2 + k^2\varpi^2)B_{\dot{z}}^2 + x^2\left[B_{\dot{\phi}}^2 + B_{\dot{z}}^2\right]$,

$f_2 = -x^2 B_{\dot{z}}^2$,

$g = \frac{1}{\varpi}\left[\frac{C\varpi - 1}{C\varpi}(mB^{\dot{\phi}} + k\varpi B^{\dot{z}})^2 + \frac{D}{(C\varpi)^2}\right]$,

$D = 2k^2\varpi^2(k^2\varpi^2 B_{\dot{z}}^2 - m^2 B_{\dot{\phi}}^2) + (\varpi\Omega_F)^2 g_1 + (\varpi\Omega_F)^4 g_2 + (\varpi\Omega_F)^6 g_3$,

$g_1 = 2k^2\varpi^2(B_{\dot{z}}^2 - B_{\dot{\phi}}^2) + (m^2 + k^2\varpi^2)E + F$,

$g_2 = x^2 G + mx\varpi B_{\dot{\phi}}(k\varpi B_{\dot{\phi}} - mB_{\dot{z}})(x^2)'$,

$g_3 = x^4\left[(1 - k^2\varpi^2)B_{\dot{z}}^2 - m^2 B_{\dot{\phi}}^2\right]$,

$E = -x^2\left[B_{\dot{\phi}}^2 + \varpi^2(\frac{1}{\varpi}B_{\dot{z}}^2)'\right] - 2k\varpi x B_{\dot{\phi}}(2mB_{\dot{\phi}} + k\varpi B_{\dot{z}})$

$- m\varpi\left[xB_{\dot{\phi}}(k\varpi B_{\dot{\phi}} - mB_{\dot{z}})\right]' + (m^2 + k^2\varpi^2)(1 - k^2\varpi^2)B_{\dot{z}}^2$,

$F = (k^2\varpi^2 B_{\dot{\phi}}^2 - m^2 B_{\dot{z}}^2)\frac{1}{\varpi}(\varpi^2 x^2)' + 2k^2\varpi^2 mx B_{\dot{\phi}}(k\varpi B_{\dot{\phi}} - mB_{\dot{z}})$,

$G = E - m^2(m^2 + k^2\varpi^2)B_{\dot{\phi}}^2 + 2mxB_{\dot{\phi}}(k\varpi B_{\dot{\phi}} - mB_{\dot{z}})$.

and it is assumed that $\varepsilon \ll 1$, $k\varpi \ll 1$, $m = O(1)$. Then, equation (14) reduces to

$$W = W_k + W_p, \tag{20}$$

$$W_k \sim \int_0^{\varpi_0} d\varpi \left\{ \left[(B_{\hat{\phi}} - k\varpi B_{\hat{z}})^2 + \varepsilon^2 B_{\hat{z}}^2 \right] \varpi |\xi'|^2 + \varepsilon^2 B_{\hat{z}}^2 \left(|\xi|^2 \right)' \right\}, \tag{21}$$

$$W_p \sim \int_0^{\varpi_0} d\varpi \left[(k^2\varpi^2 + \varepsilon^2)(B_{\hat{\phi}} - k\varpi B_{\hat{z}})^2 + 2k^2\varpi^2(k^2\varpi^2 B_{\hat{z}}^2 - B_{\hat{\phi}}^2) + \varepsilon^2 (B_{\hat{z}})^2 \right] \frac{|\xi|^2}{\varpi}. \tag{22}$$

To obtain a negative value of W, the negative term in W must dominate to the other positive terms. Further, we take a trial function ξ as

$$\xi = \begin{cases} \varpi/\varpi_1, & (0 \le \varpi < \varpi_1), \\ 1, & (\varpi_1 < \varpi < \varpi_2), \\ (\varpi_0 - \varpi)/(\varpi_0 - \varpi_2), & (\varpi_2 < \varpi < \varpi_0), \end{cases} \tag{23}$$

and assume that $\varpi_1/\varpi_0 \ll 1, \Delta\varpi_2/\varpi_0 \ll 1$, where $\Delta\varpi_2 = \varpi_0 - \varpi_2$. We expect that the value of W would be minimized by this ξ. By estimation of W_k in both side regions $0 \le \varpi < \varpi_1$ and $\varpi_2 < \varpi < \varpi_0$, we found

$$W_k \sim \left(2\frac{\Omega_F}{k} B_{\hat{z}} \right)^2 \frac{\varpi_0}{\Delta\varpi_2}. \tag{24}$$

Thus, $W < 0$ is possible if the magnetic field satisfy

$$\left(2\frac{\Omega_F}{k} B_{\hat{z}} \right)^2 \ll k^2 \varpi_0^2 B_{\hat{\phi}}^2. \tag{25}$$

This is the additional criterion for screw instability. By using K-S criterion (16), equation (25) is rewritten as

$$\varpi_0 \Omega_F \ll \left(\frac{B_{\hat{\phi}}}{B_{\hat{z}}} \right)^3. \tag{26}$$

The magnetic field is unstable if both K-S criterion and this additional criterion are satisfied. If $\varpi_0\Omega_F$ is growing as the same order with $(B_{\hat{\phi}}/B_{\hat{z}})^3$, the field become stable against the screw instability, even if the field satisfy Kruskal-Shafranov criterion.

Finally, we would implicate of the additional criterion for black hole magnetospheres. In the force-free magnetic field, because the current streaming is conserved along the magnetic field lines, the toroidal component of the field lines induced by the current is determined by the boundary condition. Here, we assume that the field lines thread the event horizon. Then, by the requirement that the electromagnetic power F^2 is finite at the horizon, we obtain the ratio of the toroidal to the poloidal component of the field [1]:

$$\frac{B^{\hat{\phi}}}{B^{\hat{z}}} \sim \varpi(\omega_H - \Omega_F). \tag{27}$$

Substituting the configuration (27) into K-S criterion (16) and the additional criterion (26), we obtain two criterions to become unstable:

$$L(\omega_H - \Omega_F) > 2\pi, \tag{28}$$

$$(\varpi_0\omega_H)^3 \gg \varpi_0\Omega_F. \tag{29}$$

Equation (28) is also derived by Li [4].

Another criterion (29) is relationship of the two angular velocities of the hole and the field lines. Then, we can discuss the efficiency of the Blandford-Znajek mechanism, which also depend on the two angular velocities. the power of the mechanism is [1]

$$P \propto \Omega_F(\omega_H - \Omega_F). \tag{30}$$

From this expression, we can see when the power takes its maximum ($\Omega_F = \Omega_H/2$) the additional criterion (29) is violated, and the magnetosphere can't be unstable. On the other hands, in the unstable magnetosphere, Ω_F is so small that the Blandford-Znajek mechanism doesn't work efficiency. The flares induced by the screw instability would be dominant on the rotational-energy extraction process from the hole.

4 conclusion

We have formulated linear perturbations of a given force-free magnetic field in Kerr space-time, and found the additional criterion under slowly rotating approximation. It has been founded that there are two distinct states in black-hole magnetospheres; (1) the flaring state, in which twisted magnetic field lines become unstable. The power of the Blandford-Znajek mechanism has a low efficiency, and the flare would extract quasi-periodically the hole's rotational energy. (2) the persistent state, in which the field is stable. The Blandford-Znajek mechanism would work dominantly.

References

[1] R.D. Brandford and R.L. Znajek, MNRAS, **179** (1977) 433.

[2] K.S. Thorn, R.H. Price and D.A. Macdonald, *Black Holes: The Membrane Paradigm* (Yale University Press, New Haven, 1986).

[3] A. Gruzinov, preprint (astro-ph/9908101).

[4] L.-X. Li, ApJ, **531** (2000), L111.

[5] F. Fiore and E. Massaro, Accretion Disks and Magnetic Fields in Astrophysics. eds. G. Belvedere *et al* (Kluwer Academic Publishers, 1989) 267.

[6] F.Hamann, T.A. Barlow and V. Junkkarinen ApJ **478** (1997) 87.

[7] N. Hayashida, at al ApJ **504** (1998) L71.

[8] B.B. Kadomtsev Rev.Plasma.Phys., **2** (1966) 153.

[9] N.G. Kampen and B.U. Felderhof, *Theoretical Methods in Plasma Physics* (North-Holland Press, Amsterdam, 1967).

[10] V.P. Frolov and I.D. Novikov *Black Hole Physics* (Kluwer Academic Press, Netherlands, 1998).

[11] G. Bateman *MHD instabilities* (MIT press, Cambridge, 1978).

Variability in GRBs and Internal Shocks

Ryo Takagi[1]

Department of Physics, Tokyo Institute of Technology, Oh-Okayama, Tokyo 152-0033, Japan

Shiho Kobayashi[2]

Department of Earth and Space Science, Osaka University, Toyonaka, Osaka 560, Japan

Abstract

Gamma-ray bursts (GRBs) are believed to originate from internal shocks which arise in an irregular relativistic wind. In this scenario, the observed peaks are in almost one-to-one correlation with the activity of the emitting source. We analyze the fluence distribution of the BATSE gamma-ray bursts, and find that the dispersion is smaller for the burst group with many peaks. This is consistent with the internal shocks model.

1 Introduction

A widely accepted mechanism for producing a cosmological gamma-ray burst (GRB) is the deceleration of relativistically expanding shells (Piran 1999). The kinetic energy of the shells is converted into internal energy by relativistic shocks. These shocks can be due to a collision with the ambient medium (external shocks) or shocks inside the shell itself due to nonuniform velocity (internal shocks). Electrons are heated by the shocks, and the internal energy is radiated via synchrotron emission.

Most bursts have a highly variable temporal profile with variability time scale significantly shorter than the overall duration. In the external shocks scenario this variability is due to irregularity in the surrounding material, but the efficiency is extremely law (Fenimore, Madras & Nayakshin 1996; Sari & Piran 1997). Thus, GRBs are believed to be produced by the other alternative: internal shocks. The inner engine itself should be variable in this scenario, the observed temporal profile follows very closely the operation of the source (Kobayashi, Piran & Sari 1997).

In this paper, we study the variavility in the temporal profiles. In particular, we analyze the relation between the number of the peaks and the fluence to find some signature of the internal shocks. In section 2 we describe our data set, the method used for the analysis and the the resulting fluence distribution. In section 3 the results are compared with the internal shocks model. We in section 4 give conclusions.

2 The Observed Fluence Distribution

2.1 DATA

From the BATSE bursts, we select 938 bursts with the peak flux satisfying 0.4 ph cm^{-2} s^{-1} $< F_{1024} <$ 10 ph cm^{-2} s^{-1} and the duration $T_{90} > 3$ s (long burst). The BATSE trigger efficiency is better than 90 % above 0.4 ph cm^{-2} s^{-1}. On the other hand, to exclude the bright events occurring very close to us, we require the upper limit. Figure 1 shows the number counts distribution $N(> F_{1024})$. The distribution is incompatible with a homogeneous Euclidean distribution (dashed line), the paucity of weak bursts is explained by the cosmological expansion effect. We use only the cosmological bursts with $F_{1024} < 10$ ph cm^{-2} s^{-1}.

[1]E-mail:rtakagi@th.phys.titech.ac.jp
[2]E-mail:shiho@vega.ess.sci.osaka-u.ac.jp

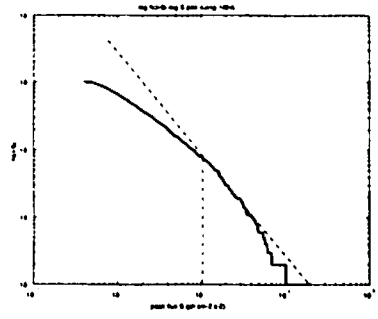


Figure 1: $N(> F_{1024})$ vs F_{1024}. The dash line corresponds to the homogeneous Euclidean distribution.

2.2 Peak-Finding Algorithm

We use almost the same algorithm proposed by Li and Fenimore (1996) to find the peaks in the temporal profile. The procedure is as follows: (1)First, we fit the background using a linear function to the pre and the post burst regions. (2)During the burst, every count bin that has more counts than the neighboring bins (both sides) is a candidate peak with counts C_p at time t_p. (3)We then search on right side of each candidate peak C_p (at t_p) for counts C_r (at t_r) that has more counts than C_p and search between t_p and t_r for counts C_1 (at t_1) that has minimum counts between t_p and t_r. And we search on left side of each candidate peak C_p (at t_p) for counts C_l (at t_l) that has more counts than C_p and search between t_p and t_l for counts C_2 (at t_2) that has minimum counts between t_p and t_l. If $C_p - C_1 \geq N_{var}\sqrt{C_p}$ and $C_p - C_2 \geq N_{var}\sqrt{C_p}$, C_p becomes a "true" peak.

2.3 the Number of Peaks

We classify the bursts by the number of the peaks into eight groups (table 1). Figure 2 shows the fluence distribution in each group, these are reasonably fit by the log-normal distributions. Figure 3 shows the relation between the number of peaks and the standard deviation in which we can see a clear correlation.

peaks	1	2	3	4	5~ 6	7~ 9	10~ 14	15~ 75
events	328	184	110	71	81	57	50	55

Table 1: the number of the peaks

3 Theoretical Model

Internal shocks arise in a relativistic wind with a non uniform Lorentz factor and convert a portion of the kinetic energy to a radiation. The irregular wind can be represent by a succession of relativistic shells in a similar manner as in Kobayashi and Sari (2000).

We consider a wind consisting of N equal mass shells with random Lorentz factors. The random Lorentz factor of each shell is assumed to be distributed uniformly in logarithmic space between $\log \gamma_{min}$ and $\log \gamma_{max}$. A collision of two shells is the elementary process in our model. A rapid shell catch up a slower and the two merge to form a single one. Using conservation of energy and momentum, the Lorentz factor of the merger and the internal energy produced by the collision can be estimated. After a fraction ϵ_e of the internal energy is given to electrons and emitted isotropically in the comoving frame, the shells spread to transform the remaining internal energy back to the kinetic energy. A simplified description of this process is to assume that the shells reflect each other with a smaller relativistic velocity, after the collision. The difference of the kinetic energy gives the radiated internal energy. We follow the evolution

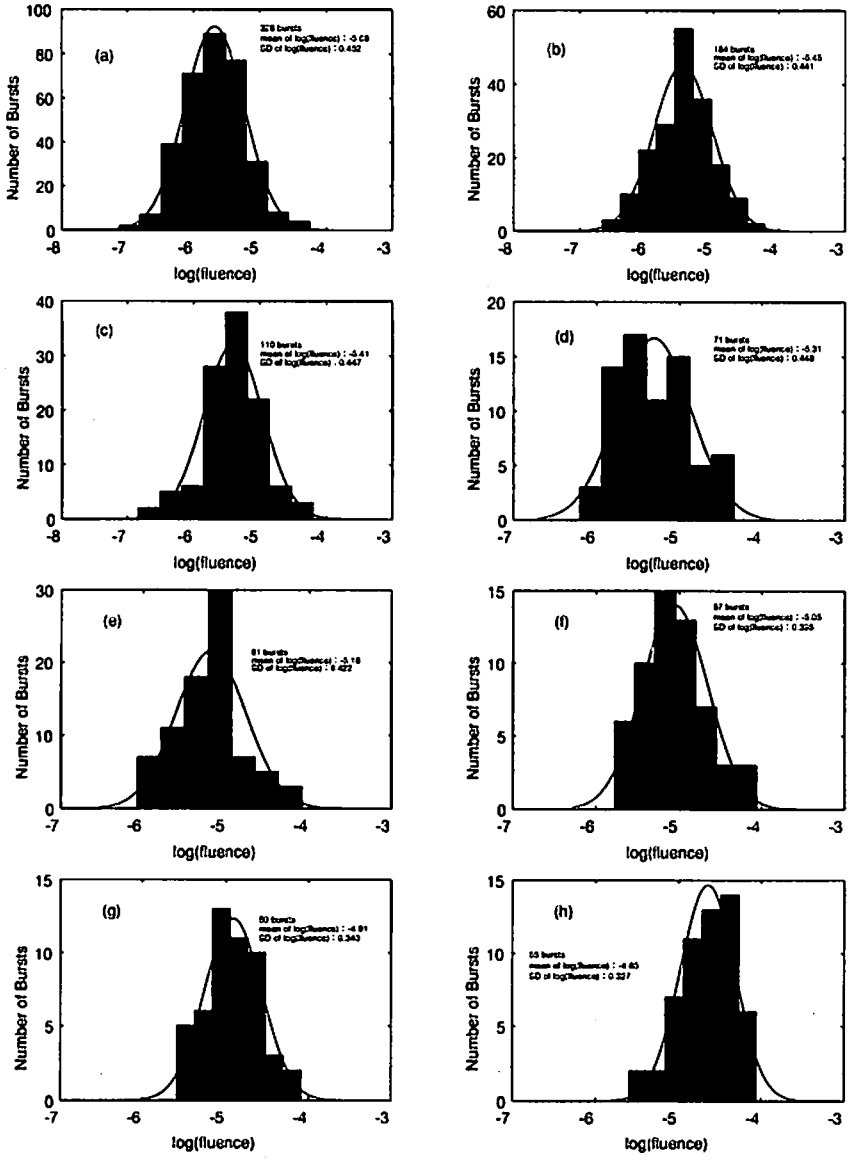


Figure 2: Fluence distribution. The number of peaks: (a)1, (b)2, (c)3, (d)4, (e)5~6, (f)7~9, (g)10~14 and (h)15~75. The solid line represents the best fit by the log-normal distribution.

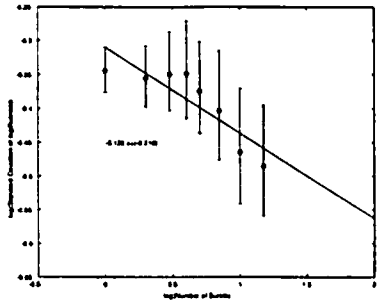


Figure 3: The standard deviation is plotted as a function of the number of the shells.

of shells until there are no more collisions, i.e. until the shells are ordered with increasing value of the Lorentz factors.

The conversion efficiency from the kinetic energy of the shells to the radiation depends on the model parameters $\{\gamma_{max}/\gamma_{min}, N, \epsilon_e\}$ and on the specific realization: the set of random Lorentz factors assigned to each shell. For fixed model parameters, the distribution of the efficiency follows the log-normal distribution (see figure 4). In this case, the mean value and the standard deviation are $\langle \log(\text{efficiency}) \rangle = 1.53$ and the $\sigma = 0.20$ respectively. Figure 5 shows the dispersion of the distribution as a function of N. The results are insensitive to ϵ_e, and the dispersion decreases as $N^{-0.7}$. This behavior can be explained as follows, the number of the main peaks in the temporal profile is almost the same as the number of the shells. If the released energy from each two shell collision E_i forms a normal distribution, the distribution of the total conversion efficiency $\propto (E_1 + E_2 + ... + E_N)/N$ for many realizations should have a dispersion $\propto 1/\sqrt{N}$ in the limit of a large N.

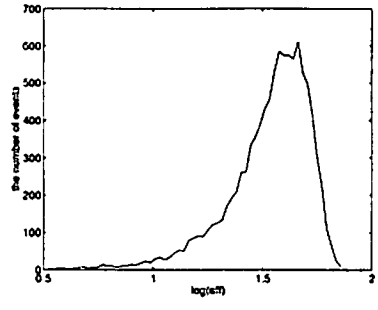


Figure 4: Efficiency distribution for 10^4 realizations: $N = 10, \epsilon_e = 1, \gamma_{min} = 10$ and $\gamma_{max} = 10^3$.

4 Conclusions

We have studies the fluence distributions of the BATSE bursts. We found that the distributions are described by the log-normal distributions, and that the standard deviation is smaller for the group with many peaks. The similar behavior was seen in the numerical simulation of the internal shocks model also. Since the differences of the beaming factors and the distances to the sources can affect the fluence, these effects should be considered to draw strong conclusions (Takagi and Kobayashi 2000).

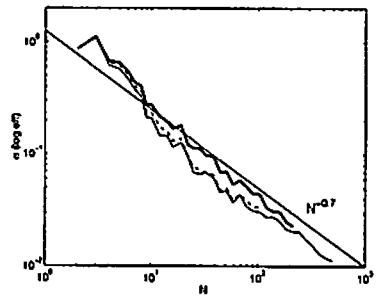


Figure 5: σ vs N. $\epsilon_e = 0.1$ (thick solid), 0.5 (dashed) and 1 (thin solid). $\gamma_{min} = 10$ and $\gamma_{max} = 10^3$.

References

[1] E.E. Fenimore, C.D. Madras and S. Nayakshin, Astrophys. J. **437** (1996) 998.

[2] S. Kobayashi and R. Sari, Submitted to Astrophys. J..

[3] S. Kobayashi, T. Piran, and R. Sari, Astrophys. J. **490** (1997) 92.

[4] H. Li and E.E. Fenimore, Astrophys. J. Lett. **469**(1996)L115.

[5] T. Piran, Physics Reports **314** (1999) 575.

[6] R. Sari and T. Piran, Astrophys. J. **485** (1997) 485.

[7] R. Takagi and S. Kobayashi, (2000) in preparation.

Neutrino Pair Annihilation above a Kerr Black Hole

Katsuaki Asano,[1]

Department of Earth and Space Science, Graduate School of Science, Osaka University, Toyonaka, 560-0043, Japan

Takeshi Fukuyama[2]

Department of Physics, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

Abstract

Using idealized models of the accretion disk, we investigate the relativistic effects on the energy deposition rate via neutrino pair annihilation near the rotation axis of a Kerr black hole. Neutrinos are emitted from the accretion disk. The bending of neutrino trajectories and the redshift due to the disk rotation and gravitation are taken into consideration. The Kerr parameter, a, affects not only behavior of the neutrinos but also the inner radius of the accretion disk. When the deposition energy is mainly contributed by the neutrinos coming from the central part, the redshift effect becomes dominant as a becomes large, and the energy deposition rate is reduced compared with that neglecting the relativistic effects. On the other hand, for a small a, the bending effect becomes dominant and makes the energy increase by factor of 2 compared with that which neglects the relativistic effects. For the disk with a temperature gradient, the energy deposition rate for a small inner radius of the accretion disk is smaller than that estimated by neglecting the relativistic effects. The relativistic effects, especially for large a, play a negative role in avoiding the baryon contamination problem in gamma-ray bursts.

1 Introduction

The light curves and polarization of gamma-ray burst afterglows indicate anisotropic central engines. These observations have been considered as evidence that gamma-ray bursts are highly beamed. One of the most probable candidates for the central engine of gamma-ray bursts is the accretion disk around a black hole. This system may be formed by the merging of two neutron stars, the merging of a black hole and a neutron star, or a failed supernova.

In these systems the hot accretion disk emits neutrinos and antineutrinos. The neutrino-antineutrino annihilation into electrons and positrons (hereafter neutrino pair annihilation) is a possible energy source of gamma-ray bursts. In order to get the gamma-ray bursts from the relativistic fireball, the baryon density in the fireball must be extremely small. Above the accretion disk, the baryon density has the lowest value near the rotation axis. The neutrino pair annihilation in this region has the possibility of making a clean fireball and of solving the baryon contamination problem. It is important to obtain the energy-density profile along the rotation axis in order to discuss the baryon contamination problem and the beaming effect.

The energy deposition rate (EDR) due to neutrino pair annihilation and the gravitational effects on it have been calculated. In a recent study, Asano & Fukuyama (2000, hereafter AF) calculated the EDR due to neutrinos emitted from the accretion disk. They incorporated the gravitational effects due to the central Schwarzschild black hole and concluded that the gravitational effects do not substantially change the EDR. However, AF did not consider the effects of the rotation and temperature gradient of the accretion disk. The EDR is proportional to $T_{\rm eff}^9$, where $T_{\rm eff}$ is the effective temperature of the accretion disk. Therefore, in order to discuss the EDR, we have to include the temperature dependence on the radius of the accretion disk.

[1] E-mail:asano@vega.ess.sci.osaka-u.ac.jp
[2] E-mail:fukuyama@se.ritsumei.ac.jp

The relativistic effects consist of three factors: the gravitational redshift, the bending of neutrino trajectories, and the redshift due to the disk rotation. The EDR is enhanced by the effect of neutrino bending. However, the redshift due to the disk rotation and gravitation reduces the EDR. Thus, these effects complicate the estimate of the EDR. It is not obvious whether the bending effect becomes dominant and the EDR is increased or the redshift effect becomes dominant and the EDR is decreased. AF did not consider the Kerr black hole. The rotation of the black hole affects the behavior of neutrinos. In addition, the rotation of the black hole makes the inner radius of the accretion disk smaller. These effects may drastically change the EDR differently than in the case for the Schwarzschild black hole.

In order to discuss quantitatively the central engine of gamma-ray bursts, we need comprehensive studies of the formation, evolution, temperature dependence, opacity, and geometry of the accretion disk, the mechanism of the energy deposition, its efficiency, the environment around the fireball, the evolution of the fireball, and the efficiency of radiation in gamma rays, etc., most of which are affected by the general relativistic effects. We think that these problems will be solved finally by laborious simulations. However, no one has demonstrated the reality of the central engine at the present stage and can say anything affirmative about it. In these situations we discuss qualitatively the problem limiting the relativistic effects on neutrino pair annihilation, which have not been treated seriously in numerical studies. We do not deal with a specific model of gamma-ray bursts but derive general conclusions on the EDR of gamma-ray bursts. Our aim in this work is not to estimate quantitatively the energy of gamma-ray bursts but to study semianalytically the relativistic effects on the EDR above the accretion disk around a rotating black hole. For the above purpose, we consider the most idealized situations and compare the results under the same conditions, except for the presence or absence of gravitation. We calculate the EDR along the rotation axis where the baryon contamination is minimum. The rotation and temperature gradient of the accretion disk are taken into consideration. We assume the simplest models of temperature gradient.

This paper is organized as follows: In section two we formulate the algorithm of the EDR calculation along the rotation axis of a Kerr black hole. The results caused by neutrinos emitted from the disk with $T_{\rm eff} \propto R^{-1}$ and $R^{-0.5}$ are given in section three. The last section is devoted to conclusions.

2 Formulation

The formulation developed in this section is obtained by modifying AF's work for the Schwarzschild black hole. In the present case we consider the following situations: Around a rotating black hole a hot thin accretion disk is formed and emits a vast number of neutrinos. Because we do not consider a specific model of the central engine of gamma-ray bursts, we consider the most idealized situations irrelevant to detailed models of the disk formation and so on. The inner (outer) edge of this accretion disk is denoted by $R_{\rm in}$ ($R_{\rm out}$). The geometrically thin disk is assumed to be sufficiently opaque to neutrinos over the whole region. We assume that the self-gravitational effects of the accretion disk are negligible and that neutrinos are gravitationally affected only by the central black hole. Of course, these idealizations may be far from the case of the realistic accretion disk. However, we consider this simple model to be sufficient for qualitatively studying the gravitational effects on the EDR. The metric around the rotating black hole is given by

$$\begin{aligned} ds^2 = g_{ij}dx^i dx^j = \left(1-\frac{r_g r}{\Sigma}\right)c^2dt^2 - \frac{\Sigma}{\Delta}dr^2 \\ - \left(\frac{r_g r}{\Sigma}a^2\sin^2\theta + r^2 + a^2\right)\sin^2\theta d\varphi^2 \\ - \Sigma d\theta^2 + 2\frac{r_g r}{\Sigma}a\sin^2\theta c dt d\varphi. \end{aligned} \tag{1}$$

Here $\Sigma \equiv r^2 + a^2\cos^2\theta$, $\Delta \equiv r^2 + a^2 - r_g r$ and $r_g = 2GM/c^2$ is the Schwarzschild radius. Also a is the Kerr parameter and takes a value from 0 to $r_g/2$. In this field the eikonal for a massless particle is written as

$$\psi = -\omega_0 t + L\varphi + \psi_r(r) + \psi_\theta(\theta), \tag{2}$$

where ω_0 and L are constants and connected with the energy and the angular momentum around the

rotation axis of the black hole, respectively. The values $\psi_r(r)$ and $\psi_\theta(\theta)$ satisfy the equations

$$\left(\frac{\partial\psi_r(r)}{\partial r}\right)^2 = \frac{1}{\Delta^2}\left[\frac{\omega_0}{c}(r^2+a^2) - aL\right]^2 - \frac{J^2}{\Delta}, \tag{3}$$

$$\left(\frac{\psi_\theta(\theta)}{\partial\theta}\right)^2 = J^2 - \left[\frac{a\omega_0}{c}\sin\theta - \frac{L}{\sin\theta}\right]^2, \tag{4}$$

where J is a constant (the separation parameter) and connected with the total angular momentum. From equation (2), we obtain the momentum of a neutrino by $p_i = \hbar\partial\psi/\partial x^i$.

As was mentioned in Introduction, we are most interested in the EDR via neutrino pair annihilation near the rotation axis, $\theta = 0$. The region of $\theta = 0$ allows analytical treatment because of its symmetric property. In the absence of gravitation, AF indicated that the θ-dependence of the EDR is weak for small values of θ. Although the θ-dependence of the gravitational effects may not necessarily be small, we calculate the EDR at $\theta = 0$ and take it as a rough standard in the vicinity of the rotation axis.

It should be remarked that $L = 0$ is required for neutrinos to reach to $\theta = 0$. The eikonal equation alone is not sufficient to obtain φ_ν, which is the angle from which meridian the particle is appoaching the rotation axis. Here and henceforth the subscript ν and $\bar{\nu}$ denote the quantities related to neutrinos and antineutrinos, respectively. Taking this φ_ν into consideration, the inner product of the momenta, p_ν and $p_{\bar{\nu}}$, at $\theta = 0$ becomes

$$(p_\nu \cdot p_{\bar{\nu}}) = \frac{\varepsilon_\nu\varepsilon_{\bar{\nu}}}{c^2}\left(1 - \sin\theta_\nu\sin\theta_{\bar{\nu}}\cos(\varphi_\nu - \varphi_{\bar{\nu}}) - \cos\theta_\nu\cos\theta_{\bar{\nu}}\right). \tag{5}$$

Here the proper energy of neutrinos has been written as

$$\varepsilon_\nu = \hbar\omega_{0\nu}\sqrt{\frac{r^2+a^2}{\Delta}} = \varepsilon_{0\nu}\sqrt{\frac{r^2+a^2}{\Delta}}, \tag{6}$$

where $\varepsilon_{0\nu}$ is the energy observed at infinity. The factor $\sqrt{(r^2+a^2)/\Delta}$ denotes the redshift. The value θ_ν is defined by

$$\sin\theta_\nu = \frac{\rho_\nu\sqrt{\Delta}}{r^2+a^2}, \tag{7}$$

with

$$\rho_\nu \equiv \frac{cJ_\nu}{\omega_{0\nu}}. \tag{8}$$

A neutrino is emitted from the disk at $(r,\theta) = (R,\pi/2)$, and it arrives at a point $(r,0)$. Then φ_ν can take values from 0 to 2π at $(r,\theta) = (r,0)$ from the symmetric property. On the other hand, a neutrino coming from $R_{\rm in}$ forms $\theta_{\rm m}$, the minimum value of θ_ν, and one coming from $R_{\rm out}$ forms $\theta_{\rm M}$, the maximum value of θ_ν. These values are derived from equation (7) and ρ_ν that is obtained by solving the trajectory equation ($\partial\psi/\partial J^2$ =const.),

$$\int_0^{\pi/2}\frac{d\theta}{\sqrt{1-(a/\rho_\nu)^2\sin^2\theta}} = \int_C \frac{dr'}{r'\sqrt{(r'/\rho_\nu)^2\left(1+(a/r')^2\right)^2 - (1 - r_g/r' + (a/r')^2)}}. \tag{9}$$

Here, in the case in which a neutrino passes through the nearest distance, r_0, before it arrives at $\theta = 0$, the integration for r' is performed from r_0 to R and r. When r' varies monotonically, the integration is performed from the smaller to the larger of r and R. The values r_0 and ρ_ν are related by

$$\rho_\nu = \frac{r_0(1+(a/r_0)^2)}{\sqrt{1 - r_g/r_0 + (a/r_0)^2}}. \tag{10}$$

As was derived in AF, the EDR via neutrino pair annihilation for a distant observer is expressed as

$$\frac{dE_0(r)}{dtdV} = \frac{21\pi^4}{4}\zeta(5)\frac{KG_{\rm F}^2}{h^6c^5}k^9T_{\rm eff}^9(3r_g)F(r), \tag{11}$$

where

$$F(r) = \frac{1}{T_{\rm eff}^9(3r_g)}\left(\frac{r^2+a^2}{\Delta}\right)^4\int_{\theta_{\rm m}}^{\theta_{\rm M}} d\theta_\nu \sin\theta_\nu \int_{\theta_{\rm m}}^{\theta_{\rm M}} d\theta_p \sin\theta_p \int_0^{2\pi} d\varphi_\nu \int_0^{2\pi} d\varphi_p$$
$$\times T_0^5(R_\nu)T_0^4(R_p)\left[1-\sin\theta_\nu\sin\theta_p\cos(\varphi_\nu-\varphi_p)-\cos\theta_\nu\cos\theta_p\right]^2. \quad (12)$$

Here the following notations and remarks are in order. By the temperature at $R = 3r_g$, we have represented the effective temperature of the disk $T_{\rm eff}$ that is observed in the comoving frame. The value T_0 is the temperature of the disk observed at infinity. The value $dtdV$ denotes $\sqrt{-g}d^4x$. The dimensionless parameter K is 0.12 for ν_e and 0.027 for ν_μ and ν_τ. The Fermi constant $G_{\rm F}^2$ is $5.29\times10^{-44}{\rm cm}^2\,{\rm MeV}^{-2}$. The values k and h are the Boltzmann and the Planck constants, respectively. The factor of the redshift in equation (12) and the equations to obtain $\theta_{\rm m}$ and $\theta_{\rm M}$ are different from those for the Schwarzschild metric discussed in AF. Furthermore, although it was assumed in AF that T_0 does not depend on R, we have here assumed that T_0 depends on R. From an angle θ_ν at a point $(r,0)$, we can trace back a trajectory of a neutrino to the emitted position R_ν on the disk. Thus, R_ν and therefore $T_0(R_\nu)$ depend on θ_ν, and T_0 is involved in the integrand θ_ν. This θ_ν-dependence of T_0 is obtained by the numerical calculation of equation (9).

In the locally nonrotating frame, a neutrino with $L = 0$ moves normally to the direction of the disk motion. In that case, the temperature suffers the redshift due to the rotation of the disk and gravitation. Namely,

$$T_0(R) = \frac{T_{\rm eff}(R)}{\gamma}\sqrt{g_{tt}-\frac{g_{t\varphi}^2}{g_{\varphi\varphi}}} = \frac{T_{\rm eff}(R)}{\gamma}\sqrt{\frac{R^2+a^2-r_gR}{R^2+a^2+(r_g/R)a^2}}, \quad (13)$$

where $\gamma \equiv 1/\sqrt{1-v^2/c^2}$ and

$$\frac{v^2}{c^2} = \frac{r_gR(R^2+a^2-a\sqrt{2r_gR})^2}{2(R^2+a\sqrt{r_gR/2})^2(R^2+a^2-r_gR)}. \quad (14)$$

We ignore the rate of deposition energy reabsorved by the central black hole. As was estimated in AF, the effect of reabsorption is not so large.

3 Results

In the following discussion we calculate the EDR incorporating the disk rotation and the gravitation of the central black hole. As a matter of convenience, AF assumed that the temperature of the accretion disk is relativistically isothermal, namely, T_0 is constant. In that case, we can estimate equation (11) by calculating the only two neutrino trajectories coming from $R_{\rm in}$ and $R_{\rm out}$. The value T_0 goes out of the integral in equation (12). However, it may be more natural to take the effective temperature $T_{\rm eff}(R)$ as a base since the temperature (and other thermodynamic quantities) is well defined in the comoving coordinates.

Integration of equation (11) over the volume $dV = (r^2+a^2\cos^2\theta)\sin\theta dr d\theta d\varphi$ gives the energy deposition per unit world time for a distant observer. However, we estimate the EDR within the infinitesimal angle $d\theta$ along the rotation axis over r = 1.5-10r_g. We do not consider the EDR at $r < 1.5r_g$ since in this region the baryon contamination occurs severely and the reabsorption rate of the deposited energy to the black hole is large. Then the EDR is proportional to the dimensionless integral of $G(r) \equiv F(r)(r^2+a^2)/r_g^2$ over $\hat{r} \equiv r/r_g$;

$$\frac{dE_0}{dt} \simeq 4.41\times10^{48}\left(\frac{d\theta}{10^\circ}\right)^2\left(\frac{kT_{\rm eff}(3r_g)}{10{\rm MeV}}\right)^9\left(\frac{r_g}{10{\rm km}}\right)^3\int_{1.5}^{10} G(\hat{r})d\hat{r} \quad {\rm ergs\ s^{-1}}, \quad (15)$$

for ν_e.

We discuss the cases for the various dimensionless quantity $a_* \equiv a/0.5r_g$. We adopt the innermost stable orbit as the inner edge of the disk, $R_{\rm in}$. As a_* takes a larger value, $R_{\rm in}$ becomes smaller. $R_{\rm in}$ are 3, 1.16, and $0.73r_g$ for $a_* = 0$, 0.9, and 0.99, respectively. For $a_* = 0.99$, $R_{\rm in}$ is inside of the ergosphere. The value $R_{\rm out}$ is fixed at $10r_g$ in every case.

When $R_{\rm in} = 3r_g$ and the relativistic effects are neglected, electron neutrinos, which are emitted from the isothermal disk, deposit over $r = 1.5\text{-}10r_g$ the energy,

$$\frac{dE_0}{dt} \simeq 2.23 \times 10^{51} \left(\frac{d\theta}{10^\circ}\right)^2 \left(\frac{kT_{\rm eff}(3r_g)}{10{\rm MeV}}\right)^9 \left(\frac{r_g}{10{\rm km}}\right)^3 \quad {\rm ergs\ s^{-1}}. \tag{16}$$

The EDR via neutrino pair annihilation is strongly dependent on the disk temperature $dE_0/dt \propto T_{\rm eff}^9$. Thus, the isothermal disk is too crude even for the zero-order approximation. In this subsection we discuss the pair annihilation of neutrinos emitted from the disk with temperature gradient. Here, as before, we do not consider the detailed structure and the formation process of the disk. We have no definite theory concerning the temperature gradient of the accretion disk as the source of gamma-ray bursts. Here we adopt $T_{\rm eff}(R) = T_{\rm eff}(3r_g) \cdot 3r_g/R \propto R^{-1}$ as the simplest and acceptable model. In this model the temperature profiles for different a_* remain common outside of $R = 3r_g$. Actual accretion disks may not be opaque and may not be hot in the central region. Here, however, we consider the idealized disk model mentioned in section 2. This disk model has the highest temperature at $R_{\rm in}$, where we need the most detailed survey of the gravitational effects. The efficiency of the energy deposition is very high because the innermost temperature is highest.

As $R_{\rm in}$ decreases, the neutrino luminosity becomes very large because of neutrinos coming from the hotter region inside of $R = 3r_g$ and enhances the EDR drastically. For instance, in the case ignoring the gravitation and disk rotation, the EDR in the region $r = 1.5\text{-}10r_g$ and $\theta \leq \pi/4$ is

$$\frac{dE_0}{dt} = 5.76 \times 10^{50} \left(\frac{kT_{\rm eff}(3r_g)}{10{\rm MeV}}\right)^9 \left(\frac{r_g}{10{\rm km}}\right)^3 \quad {\rm ergs\ s^{-1}\ for}\ R_{\rm in} = 3r_g, \tag{17}$$

$$= 1.05 \times 10^{53} \left(\frac{kT_{\rm eff}(3r_g)}{10{\rm MeV}}\right)^9 \left(\frac{r_g}{10{\rm km}}\right)^3 \quad {\rm ergs\ s^{-1}\ for}\ R_{\rm in} = 1.16r_g, \tag{18}$$

where we have adopted $R_{\rm out} = 10r_g$ as before. These values have been calculated using the formulation in AF, taking the temperature gradient into consideration. For $R_{\rm in} = 3r_g$, in comparison with the isothermal case, the EDR are reduced for the same $T_{\rm eff}(3r_g)$. For $R_{\rm in} = 1.16r_g$, on the other hand, the luminosity is about 0.7 times that of the isothermal disk with $R_{\rm in} = 3r_g$. However, the EDR becomes about 2 times larger, and the efficiency is improved compared to the isothermal disk for the same $T_{\rm eff}(3r_g)$. We have listed in Table 1 the EDR along the rotation axis over $r = 1.5\text{-}10r_g$ for various $R_{\rm in}$ in the case ignoring the relativistic effects. Values of 1.45 and 0.59 for $R_{\rm in}$ correspond to the innermost stable orbit for $a_* = 0.8$ and 0.999, respectively. Table 1 shows that the EDR increases rapidly because of the decrease of $R_{\rm in}$. The neutrino luminosity emitted from the disk with $R_{\rm in} = 0.73r_g$ becomes only 19 times that of the case for $R_{\rm in} = 3r_g$, whereas the EDR of the former case increases by factor 1400 compared with the latter one.

$R_{\rm in}/r_g$	3	1.45	1.16	0.73	0.59
EDR	0.011	0.70	2.1	15	31
$\mathcal{L}$	0.09	0.41	0.65	1.7	2.5

Table 1. EDR for the disk with $T_{\rm eff} \propto R^{-1}$ ignoring the relativistic effects. Values along the rotation axis over $r = 1.5\text{-}10r_g$. The given values are normalized by the value of equation (16). The neutrino luminosity $\mathcal{L}$ are also listed. Each $\mathcal{L}$ is normalized by the luminosity in the case of equation (16).

In the following discussion we estimate the EDR taking into consideration the gravitation of the central black hole and the disk rotation. The relativistic effects are estimated by comparing it to the EDR for the same $R_{\rm in}$ without the relativistic effects. We adopt $R_{\rm out} = 10r_g$ as before. We plot $G(r)$ in Figure 1 for the Schwarzschild black hole. In the case of $T_{\rm eff} \propto R^{-1}$, the EDR decreases rapidly with r. The EDR increases at small r because of the gravitational lensing. Since the neutrino energy decreases because of the gravitational redshift, the EDR is less efficient than that without the relativistic effects at r larger than $r = 6r_g$. The relativistic effects make the point of the maximum rate approach to the origin. As is shown in Table 2, the relativistic effects for $a_* = 0$ enhance the EDR by factor 1.7.

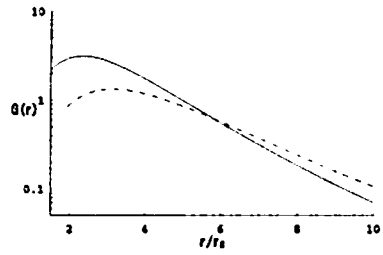


Figure 1. Plot of $G(r)$ (solid line) against r/r_g for the case of $a_* = 0$ and $T_{\rm eff} \propto R^{-1}$. The dashed line corresponds to the case for the same $R_{\rm in}$ ignoring the relativistic effects.

a_*	0	0.8	0.9	0.99	0.999
$R_{\rm in}/r_g$	3	1.45	1.16	0.73	0.59
EDR	1.7 (0.019)	1.1 (0.79)	0.73 (1.5)	0.18 (2.6)	0.087 (2.7)
$\mathcal{L}$	0.62 (0.056)	0.43 (0.18)	0.36 (0.24)	0.24 (0.40)	0.19 (0.47)

Table 2. EDR for the disk with $T_{\rm eff} \propto R^{-1}$ including the relativistic effects. Values along the rotation axis over $r = 1.5$-$10r_g$. The neutrino luminosities $\mathcal{L}$ observed at infinity are also listed. Each value is divided by that with the same $R_{\rm in}$ ignoring the relativistic effects. The values normalized by the value in the case of equation (16) are also listed as parenthetic values.

For $a_* = 0.8$, it is shown from Figure 2 that the relativistic effects enhance the EDR only in the very neighborhood of the black hole and reduce it at places distant because of the redshift. These relativistic effects are disadvantageous in avoiding the baryon contamination by the energy deposition at distant places from the central object. The peak of the EDR lies in $r < 1.5r_g$. Table 2 indicates that the EDR for $a_* = 0.8$ over $r = 1.5$-$10r_g$ is not changed very much compared to the case ignoring the relativistic effects.

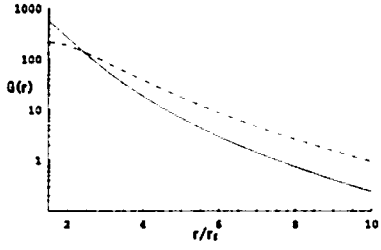


Figure 2. Same as Fig. 1 but for $a_* = 0.8$.

As a_* takes a larger value, the EDR itself increases. For $a_* \geq 0.9$, however, the energy reduction by the relativistic effects becomes larger as a_* increases (see Table 2). In Figure 3 $G(r)$ for $a_* = 0.99$ is much smaller than the $G(r)$ ignoring the relativistic effects in all regions, especially at distant places. These results are disadvantageous in avoiding the baryon contamination problem in gamma-ray bursts. In these cases, the EDR is largely owing to the neutrinos emitted from the innermost hot region of the disk. These neutrinos suffer the large redshift due to the rapid disk rotation and gravitation and cannot effectively deposit energy at $r > R$. Figure 3 clearly shows that the EDR substantially increases when $R_{\rm in}$ decreases in the case neglecting the relativistic effects. On the other hand, if we plot G(r) for $a_* = 0.999$ in Figure 3, taking into consideration the relativistic effects, $G(r)$ overlaps almost completely with those for $a_* = 0.99$ with the relativistic effects (solid line).

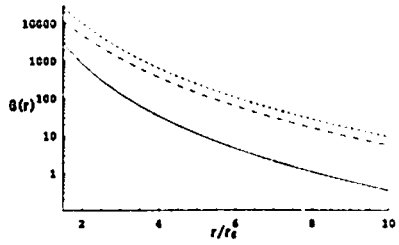


Figure 3. Same as Fig. 1 but for $a_* = 0.99$. The dashed line and dotted line correspond to the cases of $R_{in} = 0.73$ ($a_* = 0.99$) and $0.59r_g$($a_* = 0.999$) ignoring the relativistic effects, respectively.

The neutrino luminosity observed at infinity for $a_* = 0.99$, as an example, is reduced to 0.24 times the neutrino luminosity ignoring the relativistic effects. Thus, the redshift seriously affects the luminosity, which is shown in Table 2. However, the results of Table 2 cannot be explained only by the change of the luminosity. When a_* goes from 0.99 to 0.999, the luminosity increases by a factor of 1.2, but the EDR remains almost constant, whereas if we neglect the relativistic effects, the luminosity and the EDR increase, respectively, by factors of 1.5 and 2 for the same change in R_{in}. The different pattern of the EDR growth between these two cases indicates that the relativistic effects on the reaction rate of neutrino pair annihilation are also important. Even if a_* changes, the EDR at distant places is practically unchanged for $a_* > 0.8$. Most of the energy growth due to the increase of a_* occurs near the center.

As mentioned above, when a_* goes beyond 0.99, the EDR itself remains almost constant. For $T_{eff} \propto R^{-1}$, even if a_* approaches 1 beyond 0.99, the deposited energy does not increase because of the redshift due to the disk rotation and gravitation, unlike the case neglecting the gravitational effects. The EDR thus obtained is fairly small compared with that in the case where R_{in} approaches to $0.5r_g$ and the relativistic effects are neglected. Namely, as a_* approaches 1, the rough maximum value of the EDR in this model is about 2.7 times equation (16). The fraction of the EDR against the neutrino luminosity increases as a_* becomes large. For $a_* = 0.99$, if we take account of the relativistic effects, the neutrino luminosity observed at infinity reduces by a factor of 0.4 compared to the case of equation (16). Therefore, the efficiency of the EDR becomes 6.5 times that for the case of equation (16).

We have calculated the EDR for the disk, $T_{eff} \propto R^{-1}$, on behalf of various disk models. The calculations for $T_{eff} \propto R^{-1}$ show the following things: If the deposited energy is mainly contributed by the neutrinos coming from the central region, the bending effect that increases the EDR is dominant for large R_{in}, and the redshift effect that decreases the EDR is dominant for small R_{in}. Namely, as R_{in} approaches the horizon, the redshift effect overwhelms the bending effect. If we assume a model with a flatter temperature gradient than $T_{eff} \propto R^{-1}$, the contribution of the neutrinos coming from central region is relatively diminished compared to the case of $T_{eff} \propto R^{-1}$. Therefore the increase of the EDR due to the decrease of R_{in} is not so prominent as in the cases of Tables 1 and 2, and, correspondingly, the relativistic effects are not so manifest as those shown in Table 2. However, the qualitative result that the redshift effect becomes dominant according to the decrease of R_{in} (or according to the growth of a_*) is unchanged. We have listed the relativistic effects for the disk with $T_{eff}(R) = T_{eff}(3r_g) \cdot \sqrt{3r_g/R} \propto R^{-0.5}$ in Table 3. This table strongly supports our consideration mentioned above. Since the redshift effect becomes dominant as a_* increases, the enhancement factor of the EDR due to the relativistic effects has the maximum value in the Schwarzshild case. The enhancement factor in Table 3 is 1.7, which is the same as those in the former disk model. From these arguments it is conjectured that the enhancement factor due to the relativistic effects probably does not exceed 2 in the disk models with simple temperature dependences. Also in Table 3, for $a_* > 0.99$, the EDR itself remains almost constant as in the other two cases.

a_*	0	0.8	0.9	0.99	0.999
$R_{\rm in}/r_g$	3	1.45	1.16	0.73	0.59
EDR	1.7 (0.10)	1.4 (0.40)	1.2 (0.47)	0.81 (0.54)	0.68 (0.55)
$\mathcal{L}$	0.69 (0.17)	0.67 (0.23)	0.66 (0.24)	0.64 (0.33)	0.63 (0.35)

Table 3. EDR for the disk with $T_{\rm eff} \propto R^{-0.5}$ including the relativistic effects. The way of normalization of values is the same in Table 2.

We are not able to investigate all possible models with different temperature distributions. In order to study the qualitative properties of relativistic effects we have calculated only the simple models mentioned above. However, if the opacity and the density in the central region are low, the EDR is not dominated by the neutrinos coming from that region. The redshift effect is not so prominent as that of our cases. In that case, the role of $R_{\rm in}$ is replaced by the radius at which $T_{\rm eff}$ has the highest value. As this radius moves outward, the bending effect gradually becomes dominant, and the enhancement factor due to the relativistic effects may approach the 1.7 value of the $a_* = 0$ case.

4 Conclusions

In this article we have investigated the relativistic effects on the EDR via neutrino pair annihilation near the rotation axis of a Kerr black hole with a thin accretion disk. The energy deposited over $r = 1.5$-$10r_g$ is our concern in this paper. The bending of neutrino trajectories and the redshift due to the disk rotation and gravitation have been taken into consideration. The Kerr parameter, a, affects not only behavior of the neutrinos but also the inner radius of the accretion disk.

If the temperature of the accretion disk behaves as $T_{\rm eff} \propto R^{-1}$, the increase of a enhances the EDR drastically. However, most of the pair annihilation occurs near the central black hole. This is disadvantageous in avoiding the baryon contamination problem in gamma-ray bursts. When a_* goes beyond 0.99, the EDR remains almost constant, unlike the case without the relativistic effects, and is much smaller than the EDR in the case where $R_{\rm in}$ approaches to $0.5r_g$ and the relativistic effects are neglected. This is because the neutrinos, emitted from the innermost hot spot of the disk, experience the large redshift. This is the case also for the disk with the flatter temperature gradient. Therefore, these results indicate that the EDR hardly increases, regardless of the models of the disk temperature, even if we consider the disk with small $R_{\rm in}$.

From the above arguments the qualitative properties of the relativistic effects on neutrino pair annihilation are summarized as follows: When the deposition energy is contributed mainly by the neutrinos coming from the central part, the redshift effect becomes dominant as this active part approaches the horizon and the EDR is reduced compared with that neglecting the relativistic effects. On the other hand, as the part emitting the dominant neutrinos goes away from the horizon, the bending effect becomes dominant, making the EDR increase by factor of 2 compared to the EDR neglecting the relativistic effects.

These qualitative conclusions may be helpful for numerical simulations of gamma-ray burst sources that neglect the relativistic effects in the propagation of neutrinos. In order to discuss quantitatively the central engine of gamma-ray bursts, one needs numerical simulations, and several groups are working in this field. The gravitational effects on neutrino pair annihilation may change quantitatively the results of such simulations. As indicated by our calculations, the estimated power of a gamma ray burst may be changed by a factor of less than 2 if the temperature gradient is relativelly small and the neutrinos emitted close to the inner adge of the disk do not dominate. However, if a_* is close to 1 and the EDR is strongly dominated by the neutrinos coming from the central region, one cannot give a reliable result without taking the redshift effect into consideration.

References

[1] K. Asano and T. Fukuyama, Astrophysical Journal **531** (2000) 949.

[2] K. Asano and T. Fukuyama, to appear in Astrophysical Journal **546** (2001) (astro-ph/0009453).

A Generalized Equilibrium Condition in the Electro-Gravi Dynamics in 1 + 1 Dimensions

Tadayuki Ohta,[1]

Department of Physics, Miyagi University of Education, Aoba-Aramaki, Sendai 980, Japan

Robert Mann,[2]

Department of Physics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Abstract

In 1+1 dimensional electro-gravidynamics we find an exact solution to the motion of two charged particles from which a new generalized equilibrium condition is derived. This condition includes not only the Majumdar-Papapetrou condition for static balance but also a uniform motion with a fixed momentum in the center of inertia ststem.

1 Introduction

The static balance is a long-standing problem in gravitational physics. Originally motivated by attempts to find exact solutions of the ($N \geq 2$)-body system in general relativity [1], one seeks an equilibrium solution in which gravitational attraction is balanced with another repulsive force, typically electromagnetism. The first such solution was found in Einstein-Maxwell theory by Majumdar [2] and Papapetrou [3] (MJ) for $N = 2$ and was later generalized to N bodies on a line [4]. The MJ static balance condition is derived under the restriction that there is a functional relation between g_{00} and the electrostatic potential, and is given by

$$e_i = \pm\sqrt{4\pi G}\, m_i \qquad (i = 1, 2)\,. \tag{1}$$

This condition is considerably more stringent than the corresponding non-relativistic Newtonian condition

$$G m_1 m_2 - \frac{e_1 e_2}{4\pi} = 0\,. \tag{2}$$

The reason why these conditions differ have long intrigued theorists. There is no proof that (1) is a necessary condition for static balance, although it is sufficient. The question on this problem was first raised by Barker and O'Connell [5]. Although it has been conjectured [6] that an exact solution under the condition (2) should exist in general relativity, in the (2nd) post-Newtonian approximation (2) is incompatible with the static balance condition [7], and a test particle analysis [8] suggested that (2) is neither necessary nor sufficient. This suggests a wider range of possibilities for realizing equilibrium in general relativity that do not exist non-relativisitically, and several numerical studies [9, 10] have been carried out for finding a separation dependent balance condition by requiring the vanishing of strut singularity on the line connecting two particles. Until now no one has yet found – in any relativistic theory of gravity – equilibrium states in which $\sqrt{4\pi G m_i} > e_i$ for both bodies.

In this paper we present a new equilibrium solution to the static balance problem for which the condition (1) does not hold. Our exact solution is obtained for lineal gravity minimally coupled to electromagnetism, and allows for the possibility that the masses of the particles are both larger than their charges. This is the first example of its type, and our full solution is the first non-perturbative relativistic curved-spacetime treatment of this problem, providing new avenues for the study of lineal self-gravitating systems. Indeed, one-dimensional self-gravitating N-body systems have been quite fruitful in yielding insight into many problems in gravity [11]: they admit a considerably simpler level of computational and analytic analysis that can be applied to star systems (small $N \geq 2$) and galactic evolution (large N), whilst avoiding a number of difficulties inherent in three dimensions, including singularities, evaporation, and energy dissipation in the form of gravitational radiation.

[1] E-mail:t-ool@staff.miyakyo-u.ac.jp
[2] E-mail:@yukawa.kyoto-u.ac.jp

2 1+1 dimensional dilaton theory with charged particles

Our solution is derived in the context of the canonical theory for a charged N-body relativistic self-gravitating lineal system. We couple N charged point masses to Jackiw-Teitelboim lineal gravity [12], which in the absence of matter equates the scalar curvature to a cosmological constant

$$R - \Lambda = 0 \,. \tag{3}$$

As a model theory of quantum gravity [13] this model has been of considerable interest; our modification to include charged particles yields a generally covariant self-gravitating system with non-zero curvature outside the point sources. We do not include collisional terms, so that the bodies pass through each other.

We take the action to be

$$\begin{aligned} I &= \int d^2x \left[\frac{\sqrt{-g}}{2\kappa} g^{\mu\nu} \left\{ \Psi R_{\mu\nu} + \frac{1}{2} \nabla_\mu \Psi \nabla_\nu \Psi + g_{\mu\nu} \Lambda - \frac{\kappa}{2} F_\mu{}^\alpha F_{\nu\alpha} \right\} \right. \\ &\quad \left. - \sum_a \int d\tau_a \left\{ m_a \sqrt{\left(- \frac{dz_{a\mu}}{d\tau_a} \frac{dz_a^\mu}{d\tau_a} \right)} - e_a \frac{dz_a^\mu}{d\tau_a} A_\mu(x) \right\} \delta^{(2)}(x - z_a(\tau_a)) \right] , \end{aligned} \tag{4}$$

where Ψ is the dilaton field, which must be included since the Einstein action is a topological invariant in 2 spacetime dimensions. Here $g_{\mu\nu}$ and g are the metric and its determinant, R is the Ricci scalar, $\kappa = 8\pi G/c^4$, and the electromagnetic field $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, with τ_a the proper time of a-th particle whose mass is m_a and charge is e_a. Variation of the action (4) with respect to the metric, dilaton field, vector potential and particle coordinates yields the field equations

$$R - \Lambda = \kappa T^\mu{}_\mu, \; \frac{dz_a^\alpha}{d\tau_a} \nabla_\alpha \left\{ \frac{dz_a^\nu}{d\tau_a} \right\} = \frac{e_a}{m_a} \frac{dz_a^\alpha}{d\tau_a} F^\nu{}_\alpha(z_a) \,, \tag{5}$$

$$\partial_\nu \left(\sqrt{-g} F^{\mu\nu} \right) = \sum_a e_a \int d\tau_a \frac{dz_a^\mu}{d\tau_a} \delta^2(x - z_a(\tau_a)) \,, \tag{6}$$

$$\begin{aligned} &\frac{1}{2} \nabla_\mu \Psi \nabla_\nu \Psi - g_{\mu\nu} \left(\frac{1}{4} \nabla^\lambda \Psi \nabla_\lambda \Psi - \nabla^2 \Psi \right) - \nabla_\mu \nabla_\nu \Psi \\ &\qquad = \kappa T_{\mu\nu} + \frac{\Lambda}{2} g_{\mu\nu} \,, \end{aligned} \tag{7}$$

where the stress-energy is due to the electromagnetic field and the point masses

$$\begin{aligned} T_{\mu\nu} &= \sum_{a=1}^{N} m_a \int \frac{d\tau_a}{\sqrt{-g}} g_{\mu\sigma} g_{\nu\rho} \frac{dz_a^\sigma}{d\tau_a} \frac{dz_a^\rho}{d\tau_a} \delta^{(2)}(x - z_a(\tau_a)) \\ &\quad + \left\{ F_{\mu\alpha} F_\nu{}^\alpha - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right\} \end{aligned}$$

and is conserved. The set of the equations (5) and (6) takes the same form as that in 3+1 dimensional Einstein-Maxwell theory and is a closed system of $N+2$ equations whose solution yields the single metric and electromagnetic degrees of freedom and the N degrees of freedom of the point masses; it reduces to (3) if all the masses m_a and the charges e_a vanish. Note that the evolution of the charged point-masses governs that of the dilaton field via (7). The left-hand side of (7) is divergenceless (consistent with the conservation of $T_{\mu\nu}$), yielding only one independent equation to determine the single degree of freedom of the dilaton.

The canonical reduction of this theory was reported in the last year's Workshop [14] and is described in [15] in detail.
The essential point is as follows:

After all redundant variables (both electromagnetic and gravitational) are eliminated by using the constraints, the (ADM-like) Hamiltonian is defined in terms of dilaton field under an appropriate coordinate condition and boundary condition as

$$H = \int dx\mathcal{H} = -\frac{1}{\kappa}\int dx\Delta\Psi \tag{8}$$

where $\Psi(x, z_a(t), p_a(t))$ is determined by solving the constraints

$$\Delta\Psi - \frac{1}{4}(\Psi')^2 + \kappa^2\pi^2 + \frac{1}{2}(\kappa E^2 - \Lambda) + \kappa\sum_a \sqrt{p_a^2 + m_a^2}\,\delta\,(x - z_a) = 0\,, \tag{9}$$

$$2\pi' + \sum_a p_a\delta\,(x - z_a) = 0\,. \tag{10}$$

Here π is the conjugate momentum to g_{11} and E is the electric field given by

$$E = \frac{1}{2}\partial_x\sum_a e_a|x - z_a(t)|\,. \tag{11}$$

For the system of two particles, we can derive the determining equation of the Hamiltonian

$$\tanh(\frac{\kappa\mathcal{J}}{8}|r|) = \frac{\mathcal{J}(B_1 + B_2)}{\mathcal{J}^2 + B_1B_2}\,, \tag{12}$$

where $\mathcal{J}^2 = \left(\sqrt{H^2 + \frac{8\Lambda_e}{\kappa^2}} - 2\epsilon\tilde{p}\right)^2 - \frac{8e_1e_2}{\kappa} - \frac{8\Lambda_e}{\kappa^2}, B_{1,2} = H - 2\sqrt{p^2 + m_{1,2}^2}$, $\tilde{p}_i = p_i\mathrm{sgn}(z_1 - z_2)$, and $r \equiv z_1 - z_2$. The parameter $\epsilon = \pm 1$ is a constant of integration associated with the metric degree of freedom and changes sign under time reversal. We have written $\Lambda_e \equiv \Lambda - \frac{\kappa}{4}(\sum_a e_a)^2$, which is an effective cosmological constant in the spacetime. When $\Lambda_e = 0$,the Hamiltonian H reduces to that of two charged point particles in the non-relativistic limit [16], containing that in ref. [11] when $e_a = 0$. For a given $\Lambda_e \geq -(\kappa H)^2/8$, the equation (12) describes the surface in (r, p, H) space of all allowed phase-space trajectories. A given trajectory in the (r, p) plane is uniquely determined by setting $H = H_0$ in (12), since H is a constant of the motion (a fact easily verified by differentiation of (12) with respect to t).

Of course the canonical equations of motion obtained by differetiating the determining equation in terms of p and r are proved to be identical with the original geodesic equation. For the case of equal mass, the exact solution of particles' trajectories and the metric are rather complicated. Their explicit expressions in terms of the hyperbolic and /or trigonometric functions are given in Ref.[15]. A number of different types of motion are possible, depending upon a combination of four factors: gravitational attraction, the electric force between charges, the effect of the cosmological constant and relativistic effects.

3 Hamiltonian and the equilibrium condition

Our solution provides the first example of a new equilibrium solution to the static balance problem. From the relation (12) we compute the balance condition $\partial H/\partial r = 0$ with $\Lambda_e = 0$. This yields the relations

$$\mathcal{J}(B_1 + B_2) = \mathcal{J}^2 + B_1B_2 \; = 0 \tag{13}$$

which in turn imply that $B_1 = -B_2$ and $\mathcal{J}^2 = B_1^2$. These simplify to the condition

$$\frac{\kappa}{2}(\sqrt{p^2 + m_1^2} - \epsilon\tilde{p})(\sqrt{p^2 + m_2^2} - \epsilon\tilde{p}) - e_1e_2 = 0 \tag{14}$$

which we refer to as the force-balance condition. When this condition holds, the Hamiltonian becomes

$$H = \sqrt{p^2 + m_1^2} + \sqrt{p^2 + m_2^2}\,, \tag{15}$$

being the free Hamiltonian of the ststem. Only for $e_1e_2 > 0$ is the value of the momentum fixed

$$p = p_c = \pm \frac{|\left(\frac{\kappa}{2}\right)^2 m_1^2 m_2^2 - e_1^2 e_2^2|}{\sqrt{2\kappa e_1 e_2}\sqrt{(\frac{\kappa}{2}m_1^2 + e_1e_2)(\frac{\kappa}{2}m_2^2 + e_1e_2)}}\,, \tag{16}$$

in which case the two particles move with constant velocity. Eq. (14) can also be inferred by perturbatively solving for H in terms of r from (12). When the particles are initially at rest ($p_c = 0$), the condition (14) becomes

$$\frac{\kappa}{2}m_1m_2 - e_1e_2 = 0\,. \tag{17}$$

which is the static balance condition, identical to the non-relativistic condition (2) (also valid in (1+1) dimensions).

The condition (2) applies to both static and uniform motion, whereas the relativistic condition (17) represents only a static balance condition. The condition of force-balance (14) in general depends on the momentum and can be satisfied for some fixed momentum p_c whilst maintaining for both particles $\sqrt{4\pi G}m_i > e_i$ (or alternatively $\sqrt{4\pi G}m_i < e_i$). This is a qualitatively new feature and suggests that analogous equilibrium states might also exist in (3+1) dimensional general relativity.

4 Speculation in 3+1 dimensions

In (3+1) dimensions, one of the typical approach to the balance condition has been to impose the vanishing of strut singularity with $p = 0$. On the other hand, the Hamiltonian method is applied in (3+1) dimensions only on a perturbative level. Here we can show one speculation; let's take a Newtonian theory of gravitating charged particles, corrected to include the Darwin potential:

$$H = m_1 + m_2 + \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} - \frac{Gm_1m_2}{r} + \frac{e_1e_2}{4\pi r}\left\{1 - \frac{(p_1\cdot p_2)}{2m_1m_2} - \frac{(p_1\cdot r)(p_2\cdot r)}{2m_1m_2r^2}\right\}\,. \tag{18}$$

The balance condition $\partial H/\partial r = 0$ leads to

$$p_c = \pm\sqrt{\frac{Gm_1m_2 - \frac{e_1e_2}{4\pi}}{\frac{e_1e_2}{4\pi m_1m_2}}} \tag{19}$$

which shows that a uniform motion is realized for $0 < e_1e_2/4\pi < Gm_1m_2$.

It is an interesting open question as to whether or not this feature will survive in a full relativistic theory.

References

[1] *Exact solutions of Einstein's field equations* D. Kramer, H. Stephani, M. Maccullum and E. Herlt (Cambridge University Press, 1980).

[2] S.D. Majumdar, Phys. Rev. **72** (1947) 390.

[3] A. Papapetrou, Proc. Roy. Irish. Acad. **A51** (1947) 191.

[4] R. Gautreau, R.B. Hoffmann and A. Armenti jr., Nuovo Cimento **7B** (1972) 71.

[5] B.M. Barker and R.F. O'Connell, Phys. Letters **61A** (1977) 297; J. Math. Phys. **18** (1977) 18.

[6] W.B. Bonnor, Phys. Letters **83A** (1981) 414.

[7] T. Ohta and T. Kimura, Phys. Letters **63A** (1977) 193; **90A** (1982) 389; Progr. Theor. Phys. **68** (1982) 1175.

[8] W.B. Bonnor, Class. Quant. Grav. **10** (1993) 2077.

[9] G.P. Perry and F.I. Cooperstock, Class. Quant. Grav. **14** (1997) 1329.

[10] N. Bretón, V.S. Manko and J.A. Sánchez, Class. Quant. Grav. **15** (1998) 3071.

[11] See B.N. Miller and P. Youngkins, Phys. Rev. Lett. **81** (1998) 4794; K.R. Yawn and B.N. Miller, Phys. Rev. Lett. **79** (1997) 3561 and references therein.

[12] R. Jackiw, Nucl. Phys. B **252** (1985) 343; C. Teitelboim, Phys. Lett. B **126** (1983) 41.

[13] S.K. Kim et. al., Mod. Phys. Lett. **A10** (1995) 3009; A. Achucarro, Phys. Rev. Lett. **70** (1993) 1037; D. Cangemi and R. Jackiw, Phys. Rev. Lett. **69** (1992) 233;

[14] T. Ohta and R.B. Mann *Proceeding of the Ninth Workshop on General Relativity and Gravitation (Hiroshia, 1999)*, 247.

[15] R.B. Mann, D. Robbins, T. Ohta and M. Trott, Nucl. Phys. B **590**(2000) 367.

[16] T. Ohta and R.B. Mann, Class. Quant. Grav. **13** (1996) 2585.

Gamma Ray Bursts and Internal Shocks

Shiho Kobayashi [1]

Department of Earth and Space Science, Graduate School of Science, Osaka University, Toyonaka, 560-0043, Japan

Abstract

Gamma-ray bursts are believed to originate from internal shocks which arise in an irregular relativistic wind. The process has been thought to be inefficient, converting only a few percent of the kinetic energy into gamma-rays. We define ultra efficient internal shocks as those in which the fraction of emitted energy is larger than the fraction of energy given to the radiating electrons at each collision. We show that such a scenario is possible and even plausible. In our model, colliding shells which do not emit all their internal energy are reflected from each other and it causes subsequent collisions, allowing more energy to be emitted. As an example, we obtain about 60% overall efficiency even if the fraction of energy going to electrons is $\epsilon_e = 0.1$ provided that the shells' Lorentz factor varies between 10 and 10^4. The numerical temporal profile reflects well the activity of the source ejecting the shells, though numerous collisions take place in this model.

1 Introduction

A widely accepted mechanism for producing a cosmological gamma-ray burst (GRB) is the deceleration of relativistically expanding shells. The kinetic energy of the shell is converted into internal energy by relativistic shocks. These shocks can be due to collision with the ambient medium (external shocks) or shocks inside the shell itself due to nonuniform velocity (internal shocks). Electrons are heated by the shocks, and the internal energy is radiated via synchrotron and IC emission, with broken power law spectra [7].

Most bursts have a highly variable temporal profile with variability time scale significantly shorter than the overall duration. In the external shocks scenario this variability is due to irregularity in the surrounding material, but the efficiency is extremely law [2, 6]. Thus, GRBs are believed to be produced by the other alternative: internal shocks. The inner engine itself should be variable in this scenario, the observed temporal profile follows very closely the operation of the source [3, KPS97 hereafter]. Using a simple model, we have estimated that the hydrodynamic efficiency of this process (transforming kinetic energy to internal energy) is about 10% .

Kumar (1999) argued that the conversion efficiency from the bulk motion to gamma-ray is only 1%. His argument is based on three points: (I) The hydrodynamic efficiency is, as mentioned above, typically 10%. (II) It is only the electrons that are radiating. Even in equipartition among protons, magnetic field and electrons, the latter electrons only have a third of the internal energy. (III) The amount of the radiated energy within the gamma-ray band is about a third of the total radiated one. Combining these three factors give the low efficiency of $\frac{1}{10} \times \frac{1}{3} \times \frac{1}{3} = 1\%$.

Such a low efficiency results in sever energy demands on the source. Moreover, it has problems explaining the energy ratio between the GRB and its afterglow. According to the internal-external shock model, the remaining kinetic energy, which was not converted to radiation by internal shocks, is radiated during the afterglow stage. External shock does not suffer from problem (I), and the energy released in the afterglow should be considerably higher than that in the GRBs. However, it seems to be that the energy during the afterglow is only a tenth of that during the GRB, rather than ten times larger [5]. Even though the observational constraints are not very good since most of the energy is released at very early radiative stages where the afterglow was not observed, a factor of ten more energy in the afterglow seems to be excluded.

[1] E-mail:shiho@vega.ess.sci.osaka-u.ac.jp

A possible solution to these problems is to assume large angular fluctuation in the shells [5]. This model has clear predictions in the form of afterglow variability whose amplitude decay in time. It also predicts that the afterglow may be sometimes more energetic than the GRBs.

In this paper, we suggest a simple alternative solution, which overcomes the problems suggested by Kumar. We show that if the distribution of the Lorentz factor is not uniform, but instead its logarithm is distributed uniformly, then the typical ratio of Lorentz factors between neighboring shells is considerably larger. Then, the hydrodynamic efficiency can be close to hundred percents, even for a reasonable spread of Lorentz factors (Similar calculation was recently done by Beloborodov 2000). The main point of this paper is the possibility of "ultra-efficient" internal shocks. We define ultra-efficient internal shocks as a scenario in which the emitted fraction of kinetic energy is larger than ϵ_e, the fraction of internal energy that is going into electrons (and then radiated) at each collision. We will show that such a scenario is possible and even reasonable.

2 "Ultra Efficient" Internal Shocks

Internal shocks could occur within a variable relativistic wind produced by a highly variable source. We represent the irregular wind by a succession of relativistic shells with a random distribution of Lorentz factors in a similar manner as in KPS97. Beloborodov (2000) has shown that the internal shocks can convert most of the kinetic energy to internal energy if the fluctuation of the initial Lorentz factors $A^2 = (\langle\gamma^2\rangle - \langle\gamma\rangle^2)/\langle\gamma\rangle^2$ is large. Though it is maximally 1/3 if the initial Lorentz factors take random values between γ_{min} and γ_{max}, it is not limited if the distribution is uniform in logarithmic space between $\log\gamma_{min}$ and $\log\gamma_{max}$.

The efficiency can be estimated by an equation similar to equation 19 in KPS97. The most efficient case is that the masses of the shells are taken equal, and that the efficiency is given by a simple form:

$$\langle\epsilon\rangle \sim 1 - a^{1/2}\log a/(a-1) \tag{1}$$

where $a = \gamma_{max}/\gamma_{min}$. This analytic estimate fits the result of our numerical simulation, and has the asymptotic form $\langle\epsilon\rangle = A^2/2$ which Beloborodov estimated for a small fluctuation $A^2 = -1 + (a+1)\log a/2(a-1)$. Our expression generalizes Beloborodov's formula and gives reasonable estimate also for large fluctuations.

Though we have seen that a large hydrodynamic efficiency is possible if the fluctuation of the initial Lorentz factors is large, it is not reasonable that all the internal energy is emitted after each collision, since electrons do not have most of the internal energy. Defining ϵ_e as the fraction of energy given to electrons we expect $\epsilon_e < 1$. Even at equipartition with protons $\epsilon_e = 1/2$. Under these circumstances the total emitted energy in Beloborodov model is still limited by ϵ_e.

However, if $\epsilon_e < 1$ the merger produced by a collision is expected to stay hot after the emission. As a result, the merger will spread to transform the remaining internal energy back to the kinetic energy [4]. A simplified description of this process is to assume that the two shells reflect with a smaller relative velocity. The difference of the kinetic energy before and after the collision is the emitted internal energy. The reflecting shell will collide into the other neighbor shell. Since in this way a large amount of collisions are caused, the overall efficiency from the kinetic energy to the radiation could be larger than ϵ_e. This is the key ingredient of our model.

3 Two Shell Collision

A collision of two shells is the elementary process in our model. A rapid shell (denoted by the subscript r) catches up a slower one (s) and the two merge to form a single one (m). Using conservation of energy and momentum, the Lorentz factor of the merged shell γ_m and the internal energy E_{int} produced by the collision are given by

$$\gamma_m \sim \sqrt{(m_r\gamma_r + m_s\gamma_s)/(m_r/\gamma_r + m_s/\gamma_s)}\ ,\ E_{int} = m_r(\gamma_r - \gamma_m) + m_s(\gamma_s - \gamma_m). \tag{2}$$

After a fraction ϵ_e of the internal energy is emitted isotropically in local frame of merged shell (center of mass frame), the shells will spread to transform the remaining internal energy back to kinetic energy.

The Lorentz factor of the reflected shells in center of mass frame are given by

$$\bar{\Gamma}_r = (M^2 + m_r^2 - m_s^2)/2m_rM, \quad \bar{\Gamma}_s = (M^2 + m_s^2 - m_r^2)/2m_sM. \tag{3}$$

where $M = (m_r\gamma_r + m_s\gamma_s - \epsilon_e E_{int}/c^2)/\gamma_m$. The Lorentz factors in laboratory frame are

$$\Gamma_r = \bar{\Gamma}_r\gamma_m - \sqrt{(\bar{\Gamma}_r^2 - 1)(\gamma_m^2 - 1)}, \quad \Gamma_s = \bar{\Gamma}_s\gamma_m + \sqrt{(\bar{\Gamma}_s^2 - 1)(\gamma_m^2 - 1)}. \tag{4}$$

The shells are once compressed by shocks, but these spread when reflecting. For simplicity we assume that the width of the shells, l_i, does not change before and after the collision.

4 Multiple Shell Collision

We consider a wind consisting of N shells. Each shell is characterized by the four variables: γ_i, m_i, l_i and R_i. We assume that the initial Lorentz factor of each shell is distributed uniformly in logarithmic space between $\log\gamma_{min}$ and $\log\gamma_{max}$. The masses are assumed to be correlated with the Lorentz factors as $m \propto \gamma^\eta$. For $\eta = -1$ and 0 the shells have equal energy and equal mass respectively, and for $\eta = 1$ the shells have equal density under an assumption of the equal shell width. In order to reduce the number of parameters describing the wind, we assume a constant value l for the initial widths and the initial separations between the shells. Then, the initial position of the shells are $R_i = 2(i-1)l$.

The evolution of the system in time is basically same as in KPS97, but equation 4 is used to calculate the Lorentz factor of the reflecting shells for the next time step. We follow the evolution of shells until there are no more collisions, i.e. until the shells are ordered with increasing value of the Lorentz factors.

The conversion efficiency from the kinetic energy of the shells to the radiation can be calculated by using the initial and final kinetic energy as $\langle\epsilon\rangle = 1 - \Sigma m_i^{(f)}\gamma_i^{(f)}/\Sigma m_i^{(i)}\gamma_i^{(i)}$ where the superscripts (f) and (i) represent the initial and final values respectively. It depends on the model parameters $\{\gamma_{max}/\gamma_{min}, N, \eta, \epsilon_e\}$ and on the specific realization: the set of random Lorentz factors assigned to each shell. For each choice of the parameters of the model, we have evaluated the efficiency for 100 realizations. The mean efficiency and its standard deviation are listed in Table 1 for the equal mass case ($\eta = 0$).

Table.1

N	γ_{min}	γ_{max}	efficiency [%] ($\epsilon_e = 0.1$)	efficiency [%] ($\epsilon_e = 0.5$)
30	10^2	10^3	9.2 ± 2.3	16.2 ± 3.0
30	10	10^4	40.0 ± 9.2	67.5 ± 9.3
10^2	10^2	10^3	15.1 ± 1.5	17.7 ± 1.5
10^2	10	10^4	62.9 ± 4.8	72.4 ± 4.3

The efficiency approaches an asymptotic value as the Lorentz factor ratio $\gamma_{max}/\gamma_{min}$ and the number of the shells N increase. The asymptotic value depends on η and ϵ_e. For a fixed ϵ_e, the equal mass case ($\eta = 0$) is the most efficient.

The efficiency is plotted as a function of ϵ_e for the equal mass case in Figure 1. It is interesting that the efficiency can be larger than ϵ_e since the energy released in GRB can be much larger than that in the afterglow. For instance, consider the case of $\eta = 0$, $N = 100$, $\epsilon_e = 0.1$, $\gamma_{min} = 10$ and $\gamma_{max} = 10^4$, the internal shocks can convert about sixty percent of the kinetic energy to the radiation. The reminder is converted to the thermal energy by external shock. However, only the fraction $\epsilon_e = 0.1$ of that is emitted as the afterglow, the efficiency by the external shock is only four percent.

Figure 2a shows the resulting temporal structure in the equal mass case. It is a superposition of pulses from the elementary two shell collisions. Though numerous collisions take place during the evolution, the number of peaks in the profile is of order of N. Since all peaks have widths of the same order of magnitude, the different amplitude of the peaks originates mainly from the difference in the internal energy produced by the collisions. In Figure 2b, we plot the initial Lorentz factor as a function of time when shells were emitted by the source. We evaluate E_{int} for all pairs of the shells if inner shell is faster. We assign E_{int} to the ejection time of the inner shell (see Figure 2c). It resembles well the temporal profile in 2a.

Neighboring shells collide on a time scale of $2\gamma^2 l/c$. The matter is moving toward the observer, the resulting observer time scale is l/c. On the other hand, the difference in observer time due to the location

of the given shell within the wind is of order Nl/c. Then, we observe the pulses arising from the collisions according to their positions inside the wind. In Figure 2d we plot the shell's index against the time when a radiation from the shell is observed. Although the light curve in figure 2a is the superposition of all pulses, in figure 2d we plot only the pulses which are higher than a tenth of the highest. The number is $\sim$ 200 and that of pulses higher than 1/100 of the highest is $\sim$ 500 in this typical simulation. We can see a clear correlation between the index and the time, the temporal profile reflects the activity of the source.

The deviation from the correlation, e.g. some groups of circles in a line from left top to right bottom is due to that a rapid inner shell collides into the outer neighbor, and the boosted shell further collides into the outer. Since the masses of the shells are equal, the Lorentz factors of the inner and the outer shells are just switched at each collision if the radiation is negligible. Then, if an initial Lorentz factor of a shell is peculiarly high, the index of the shell which has the high Lorentz factor propagates outward. In this sequence with the radiation loss, the pulse from the collisions dumps quickly, then, the overall correlation is not destroyed.

5 Conclusions

We have shown that the conversion efficiency from kinetic energy of relativistic shells to radiation can be close to hundred percent if the source produces shells with comparable masses with different Lorentz factor, especially when the logarithms of the Lorentz factors are distributed uniformly. It had been assume in our previous work that the Lorentz factors themselves were distributed uniformly, and the efficiency was less than forty percent. With this distribution, the most efficient case is that the source produces shells with comparable energy, instead of comparable masses (KPS97).

However, this high efficiency is achieved assuming that all the internal energy is emitted at each collision, this is not reasonable since only electrons radiative effectively, and these do not have all the internal energy. After the electrons radiate, large amount of the internal energy remains in protons. Using a hydrodynamic simulation, it has been shown that the hot merger produced by a collision spreads to transform the remaining internal energy back to the kinetic one. A simplified description of this process is to assume that the shells reflect each other with a smaller relative velocity, after the collision. The difference of the kinetic energy gives the radiated internal energy.

Since the reflecting shell collide into the outer neighbor shell, the index of the shell which has a high Lorentz factor propagates outward until the high value decays by the radiation loss. The shell itself which has initially a high Lorentz factor might not go through many collisions, but its high "kinetic energy" does. Therefore, the internal shock process is very efficient even if the fraction of internal energy emitted at each collision is small. Previously, the efficiency in the case of $\epsilon_e < 1$ had been estimated as smaller by a factor of ϵ_e than that in the corresponding fully radiative case. Our ultra efficient internal shocks scenario shows this to be a significant underestimate. Though the efficiency that we have estimated is bolometric efficiency, the efficiency from the kinetic energy of the shells to gamma-ray radiation is also high if the faction of the energy radiated in the BATSE band is not very small.

Numerous collisions happened in our ultra-efficient internal shocks model, it made the peak width wider than in the previous internal shocks model. However, the number of main peaks is still almost the same as the number of shells that the source emitted. There is a strong correlation between the time at which we observe a pulse and the emission time of the corresponding shell from the source. This correlation persists even for a small ϵ_e case where larger number of collision happen. The temporal structure reproduces the activity of the source.

We have shown that the efficiency of the internal shock process is not limited by ϵ_e, while that of the external shock is so. If a fraction ζ of kinetic energy of an explosion is converted to the radiation by the internal shocks, all the remaining one is converted to the thermal by external shock in the afterglow stage, and a fraction ϵ_e of the thermal is emitted. Then, we can roughly estimate the ratio between the energy released in afterglow and that in GRB as $\sim \epsilon_e(1-\zeta)/\zeta$. Assuming $\epsilon_e = 0.1$, the ratio is 1/10 for $\zeta = 0.5$ and decrease as ζ increase. If the efficiency of internal shocks is indeed very large, the luminosity of GRB and the afterglow are expected to be anticorrelated.

References

[1] A. Beloborodov, (2000) astro-ph/0004360.

[2] E.E. Fenimore, C.D. Madras and S. Nayakshin Astrophysical Journal 437 (1996) 998.

[3] S. Kobayashi, T. Piran and R. Sari, Astrophysical Journal 490 (1997) 92.

[4] P. Kumar, Astrophysical Journal Letter 523 (1999) L113.

[5] P. Kumar and T. Piran, Astrophysical Journal 535 (2000) 152.

[6] R. Sari and T. Piran, Astrophysical Journal 485 (1997) 270.

[7] R. Sari, T. Piran and R. Narayan, Astrophysical Journal Letter 497 (1998) L17.

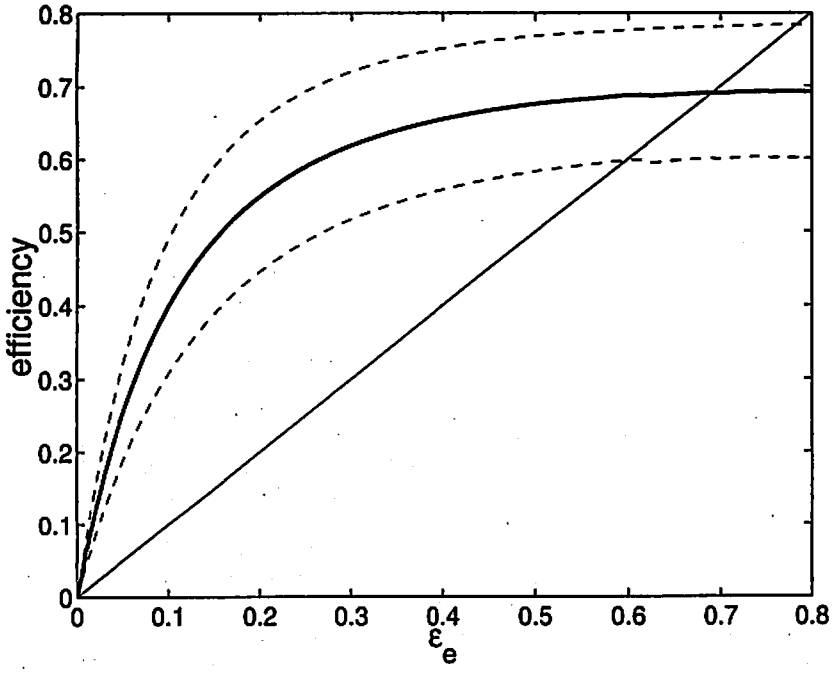


Figure 1: Efficiency vs ϵ_e with 1σ error bars of 100 random simulations. $N = 30$, $\eta = 0$, $\gamma_{min} = 10$ and $\gamma_{max} = 10^4$.

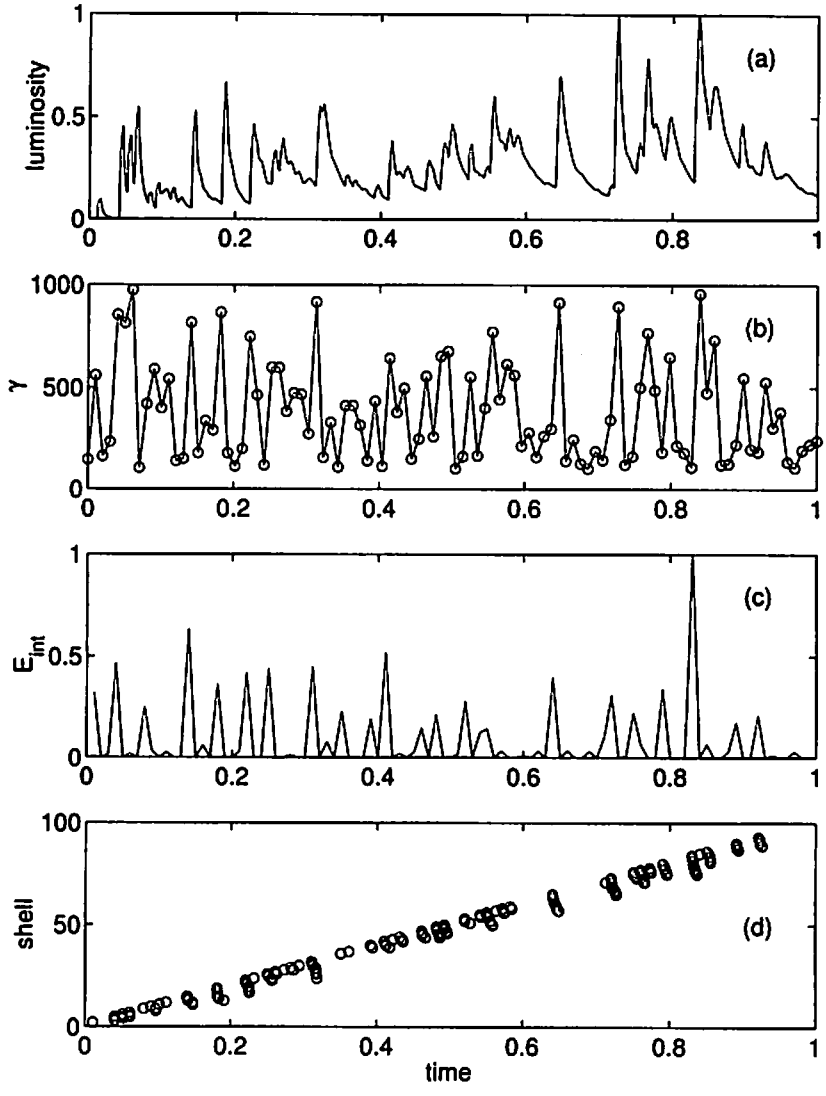


Figure 2: Temporal structure. (a) Numerical temporal structure. (b) Lorentz factor of a shell at the ejection vs ejection time (c) E_{int} between neighbor boring shells vs ejection time (d) Index of a shell vs observed time of the radiation produced in the shell. $N = 100, \eta = 0, \gamma_{min} = 10^2$ and $\gamma_{max} = 10^3$ and $\epsilon_e = 0.1$.

Multiple Sub-Jet Model of Gamma Ray Bursts and Possible Origin of X-ray Precursors and Post-cursors

Takashi Nakamura[1]

Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

Abstract

We assume that internal shocks of Gamma Ray Bursts (GRBs) consist of multiple sub-jets with a collimation half-angle $\sim$ several $\times \gamma_i^{-1}$, where γ_i is the Lorenz factor of each sub-jet. If by chance a sub-jet is first emitted off-axis from the line of sight, the observed peak energy can be in the X-ray region. Next if by chance a subsequent sub-jet is emitted along the line of sight, then the peak energy will be in the gamma ray region and the gamma ray may arrive after the X-ray precursor from the former sub-jet depending on parameters. This model predicts a new class of GRBs with extremely weak and short gamma ray emission but X-ray precursor and/or post-cursor as well as an afterglow.

1 Introduction

Recently evidences for collimation of GRBs in the afterglow came out for GRB990123 [1, 2, 3, 4] (see also [5] for non-jet interpretation) and GRB990510 [6, 7]. The rapid decline rate of GRB980519 [8] also suggests the collimation of GRBs. From the fact that the power index of the decline rate of the afterglow changed at $1 \sim 2$ days for GRB990123 and GRB990510, the collimation half-angle $\Delta\theta_a$ of the afterglow is estimated as ~ 0.1 [1, 6, 9], where $_a$ stands for afterglow. The unusually slowly declining X-ray afterglow of GRB980425 ($\propto T^{-0.2}$) can also be explained [2] if the afterglow is beamed and the angle between the jet and the line of sight is $\sim 30°$ [10].

In an internal shock model [11], which is one of the promising models of GRBs, the relativistic beaming half-angle ($= \gamma_i^{-1} < 0.01$) with γ_i being the initial gamma factor of the internal shock is smaller than the suggested collimation half-angle of the afterglow ($\Delta\theta_a \sim 0.1$). This means that it is possible to assume that the collimation half-angle of the internal shock is much smaller than the collimation half-angle of the afterglow. We here assume that the beam of internal shocks consist of such sub-jets with the collimation half-angle $\Delta\theta_\gamma < \Delta\theta_a$. This kind of a model has already been considered to interpret no correlation between the isotropic luminosity of gamma-rays and that of the afterglow [12] as well as the diversity of the isotropic luminosity for GRBs with known red-shifts.

Here I like to point out that such a sub-jet model has a potential ability to interpret X-ray precursors, which were first observed by Ginga satellite [13]. The evidence for X-ray precursors have been confirmed by GRANAT/WATCH [14] catalogue with 6 X-ray precursor events in 95 GRBs. This suggests that about $\sim$ 6% of GRBs have precursors. Recently Beppo/SAX found the X-ray precursor and prompt X-ray emission from GRB980519 [15]. For GRB980519 an analysis of the evolution of the X-ray spectrum in terms of a single power-law spectrum shows that the photon index evolved from -2.0 to -1.1 to -2.4. The onset of burst has such a soft spectrum that the 2-27 keV emission appears to precede the $>$ 107 keV emission by about 70s. One more important fact for GRB980519 is that the X-ray post-cursor is seen after $\sim$ 100s from the gamma ray emission.

Paczynski [16] has once suggested that precursors of GRBs might be the evidence for the association of GRBs with supernovae, that is, precursors might suggest baryon contaminated environment of GRBs, while the gamma ray might come from the clean fire ball. Here I like to interpret precursors in the framework of the standard fire ball model which is so successful so far [11].

[1] E-mail:takashi@yukawa.kyoto-u.ac.jp

[2] In this Letter I use T to express the time at the detector. For simplicity I neglect the cosmological time dilatation to avoid confusions.

In this article, I assume that internal shocks of Gamma Ray Bursts (GRBs) consist of multiple sub-jets with the collimation half-angle $\sim$ several $\times \gamma_i^{-1}$, where $\gamma_i \simeq 100 \sim 1000$ is the Lorenz factor of each sub-jet. If by chance a sub-jet is first emitted off-axis from the line of sight, the observed peak energy can be in the X-ray region. Next if by chance a subsequent sub-jet is emitted along the line of sight, then the peak energy will be in the gamma ray region and depending on parameters, the gamma ray can arrive after the X-ray precursor from the former sub-jet. This emission pattern agrees with that of GRBs with precursors. The event rate of precursors ($\sim 6\%$) is compatible with this model since precursors occur by chance. If the time order of the former and the latter sub-jets is reversed by chance, gamma-ray comes first and X-ray comes later. This emission pattern is compatible with the post-cursor. In next section we will argue a simple model in this picture.

2 Simple Model

Let us consider a simple model of GRBs in the framework of an internal shock picture. First at $t = 0$ in the laboratory frame a slow proton jet with Lorenz factor $\gamma_s > 100$ with the collimation half-angle $\Delta\theta_{\gamma_s}$ starts from the central engine toward the direction θ_v where θ_v is the angle between the direction to the detector and the axis of the sub-jet. This slow jet continues for Δt_s. At $t = \Delta t_s + t_r$, a rapid proton jet with Lorenz factor $\gamma_r > \gamma_s$ with the collimation half-angle $\Delta\theta_{\gamma_r}$ starts from the central engine toward the same direction θ_v as the slow jet. For simplicity we consider only the case of $\Delta\theta_{\gamma_s} = \Delta\theta_{\gamma_r} \equiv \Delta\theta_\gamma$. Then the rapid jet catches up the slow jet at $t_c = 2t_r\gamma_s^2 S^2/(S^2-1) + \Delta t_s$ where $S \equiv \gamma_r/\gamma_s > 1$. The internal shock breaks out at $t_b = t_c + 2\Delta t_s \gamma_s^2 S'^2/(S'^2-1)$, where $S' \equiv \gamma_{sh}/\gamma_s > 1$ with γ_{sh} being the gamma factor of the internal shock in the laboratory frame. Let us consider the case of $\theta_v > \gamma_s^{-1}$. Then the relation of the observed time T to t is given by

$$T = (1 - \beta\cos\theta_v)t \sim t\theta_v^2/2, \tag{1}$$

where we set $T = 0$ as the time when the photon starts from the central engine at $t = 0$ arrives at the detector. The observed frequency ν at the detector is related to ν' in the shock frame as

$$\nu = \frac{\nu'}{\gamma(1-\beta\cos\theta_v)} \sim 11.1\nu'(\frac{\gamma_{sh}}{200})^{-1}(\frac{\theta_v}{0.03})^{-2}. \tag{2}$$

Since in the internal shock model the peak energy in the comoving frame should be in the soft X-ray band(say $\sim$keV) , the observed peak energy from the off-axis emission can be in X-ray band (say $\sim$10keV) for appropriate values of γ_{sh} and θ_v. For $t_r \sim \Delta t_s$, the arrival time of this off-axis emission ($T_c(\theta_v)$) is given by

$$T_c(\theta_v) = 9t_r(\frac{\gamma_s}{100})^2(\frac{\theta_v}{0.03})^2\frac{S^2}{S^2-1}, \tag{3}$$

while the duration from the catch up time to the shock break out time $\Delta T(\theta_v)$ is given by

$$\Delta T(\theta_v) = 9\Delta t_s(\frac{\gamma_s}{100})^2(\frac{\theta_v}{0.03})^2\frac{S'^2}{S'^2-1} \tag{4}$$

Now let us assume at $t = t_0$, the central engine starts to emit another pair of the same slow and rapid sub-jets toward the line of site ($\theta_v = 0$). Then the catch up time (t_c^0) and the break out time (t_b^0) become $t_c^0 = t_0 + t_c$ and $t_b^0 = t_0 + t_b$, respectively, while the relation of the observed time T to t becomes

$$T(\theta_v = 0) = t_0 + \frac{t - t_0}{2\gamma_r^2}. \tag{5}$$

Then

$$T_c(\theta_v = 0) = t_0 + t_r\frac{1}{S^2-1}, \tag{6}$$

while the duration from the catch up time to the shock break out time $\Delta T(\theta_v = 0)$ is given by

$$\Delta T(\theta_v = 0) = \Delta t_s\frac{1}{S'^2-1}. \tag{7}$$

In this case the observed frequency ν at the detector is related to ν' in the shock frame as

$$\nu = 400\nu'(\frac{\gamma_{sh}}{200}). \tag{8}$$

This means that the observed emission will be in the gamma ray band if ν' is in the soft X-ray band (~keV).

Now the condition for the occurrence of the precursor ($T_c(\theta_v = 0) > T_c(\theta_v)$) is rewritten as

$$\frac{t_0}{t_r} > (9(\frac{\gamma_s}{100})^2(\frac{\theta_v}{0.03})^2 - \frac{1}{S^2})\frac{S^2}{S^2-1}. \tag{9}$$

Let us consider a specific numerical example;
Case 1

$$\gamma_s = 100, \gamma_r = 200, \gamma_{sh} \sim 200, t_r = 6s,$$
$$\Delta t_s = 6s, \theta_v = 0.03. \tag{10}$$

Then

$$T_c(\theta_v) = 72s, \Delta T(\theta_v) = 72s,$$
$$T_c(\theta_v = 0) = t_0 + 2s, \Delta T(\theta_v = 0) = 2s. \tag{11}$$

If the peak energy in the shock frame is ~ 0.5keV, then the peak energy from 72s to 144s is ~ 5keV while the peak energy from t_0+2s to t_0+4s is ~ 200keV. Therefore if $t_0 > 70s$ the X-ray emission starts ($t_0 - 70s$) before the gamma ray emission. The duration of the X-ray precursor is ~ 70s while the duration of the gamma ray emission is ~ 2s. The duration of the X-ray precursor in this specific example happens to be comparable to those observed by Ginga for GRB900126 [13] and by Beppo/SAX for GRB980519 [15] while the duration of the gamma ray is somewhat smaller.

The above model is extremely simple but clearly demonstrates a possible origin of precursors. In Case 1 by choosing t_0 much smaller than 70s we may explain the post-cursor. For example if we choose t_0 = 20s then the gamma ray emission starts at 22s and ends at 24s while the X-ray post-cursor emission starts at 72s.

3 X-ray flux from the precursor and the post-cursor in a sub-jet model

In the previous section, we discussed only the peak energy and the temporal structure of precursors and post-cursors. In this section we estimate the X-ray flux from the precursor in the sub-jet model. The observed X-ray flux of the precursor and post-cursor of GRB980519 [15] is $\sim 5 \times 10^{-9}$erg s^{-1}cm^{-2} in 2-10 keV band. The average X-ray flux of the precursor of GRB900126 [13] is $\sim 2.5 \times 10^{-9}$erg s^{-1}cm^{-2}. We argue that these values are compatible with our model of precursors and post-cursors.

For $\theta_v = 0$ and $\gamma\Delta\theta_\gamma \gg 1$, we assume that the observed spectrum of the gamma ray from the sub-jet is given by

$$F_\nu \propto \frac{1}{(1 + (\frac{\nu}{\nu_0})^2)^{\beta_B/2}}, \tag{12}$$

where $\beta_B \sim 1.5$ and $\nu_0 \sim 150$keV are constants. The above spectrum is essentially the same as the Band spectrum [17] with $\alpha = -1$ and $\beta = -\beta_B - 1$, where α and β are the spectrum parameters in the Band spectrum. The peak frequency ν_{peak}, that is, the maximum of νF_ν, is $\sqrt{1/(\beta_B - 1)}\nu_0$.

Garnot, Piran & Sari [18] as well as Woods & Loeb [19] derived a general formula to compute the off-axis emission from beamed GRBs. Here we adopt their formulations and notations. Let us use a spherical coordinate system $\mathbf{r} = (r, \theta, \phi)$ where the coordinate are measured in the lab frame; let the $\theta = 0$ axis (z-axis) points to the detector and $\mathbf{r} = 0$ be the central engine. Let also D be the distance to

the source and $\alpha = r\sin\theta/D$ be the angle that a given ray makes with the normal to the detector. Then the observed flux is given by

$$F_\nu(T) = \frac{\nu D}{\gamma\beta}\int_0^{2\pi} d\phi \int_0^{\alpha_m} \alpha^2 d\alpha$$
$$\times \int_{\nu\gamma(1-\beta)}^{\nu\gamma(1+\beta)} \frac{d\nu'}{\nu'^2}\ \frac{j'_{\nu'}[\Omega', \mathrm{r}, T+\frac{r\mu}{c}]}{\{1-\mu^2\}^{3/2}}, \tag{13}$$
$$\mu = (1-\nu'/\gamma\nu)/\beta, \tag{14}$$
$$\gamma = 1/\sqrt{1-\beta^2}, \tag{15}$$

where α_m, T and $j'_{\nu'}$ are the maximum value of α, the arrival time of a photon at the detector and the rest frame emissivity measured in erg s^{-1} cm^{-3} Hz^{-1} sr^{-1}, respectively. Note here that $'$ means the physical quantity in the rest frame.

In the case of $\theta_v = 0$, we assume that $j'_{\nu'}$ is expressed as

$$j'_{\nu'} \propto H(\Delta\theta_\gamma - \theta) f(\nu'), \tag{16}$$

where $f(\nu')$ and $H(x)$ are a certain function and the Heaviside step function, respectively. Then $F_\nu(T)$ is given by [10]

$$F_\nu(T) = 2\pi A\nu \int_{\nu/2\gamma}^{\nu\gamma\Delta\theta_\gamma^2/2} f(\nu')\frac{d\nu'}{\nu'^2}, \tag{17}$$

where A is a factor depending on the density, the gamma factor, the strength of the magnetic fields and the location of the shock as well as the distance to the source. Since we assume that the spectrum has the form of Eq. (12) for $\gamma\Delta\theta_\gamma \gg 1$, $f(\nu')$ is given by

$$f(\nu') = \frac{1+(\beta_B+1)(\frac{2\gamma\nu'}{\nu_0})^2}{2\gamma(1+(\frac{2\gamma\nu'}{\nu_0})^2)^{\beta_B/2+1}}. \tag{18}$$

Neglecting the term proportional to $1/(\gamma\Delta\theta_\gamma)^2$ we have

$$F_\nu(T) = 2\pi A\frac{1}{(1+(\frac{\nu}{\nu_0})^2)^{\beta_B/2}} \tag{19}$$

Then the peak frequency is in the gamma ray region as $\nu_{peak} \sim \nu_0 \sim 150$keV and the peak flux is given by

$$(\nu F_\nu)_\gamma \sim 2\pi A\nu_0/2^{\beta_B/2}. \tag{20}$$

Now let us consider the off axis case ($\theta_v > \Delta\theta_\gamma$). In this case $j'_{\nu'}$ is expressed as

$$j'_{\nu'} \propto H(\cos\phi - (\frac{\cos\Delta\theta_\gamma - \cos\theta_v\cos\theta}{\sin\theta_v\sin\theta}))$$
$$\times H(\Delta\theta_\gamma - \mid\theta-\theta_v\mid) f(\nu'), \tag{21}$$

The flux is given by [10]

$$F_\nu(T) = 2A\nu \int_{\nu\gamma(1-\beta\cos(\theta_v+\Delta\theta_\gamma))}^{\nu\gamma(1-\beta\cos(\theta_v-\Delta\theta_\gamma))} \phi_v f(\nu')\frac{d\nu'}{\nu'^2}, \tag{22}$$

where $\cos\phi_v = (\cos\Delta\theta_\gamma - \cos\theta\cos\theta_v)/\sin\theta_v\sin\theta$.

Eq.(22) is evaluated as

$$F_\nu(T) = 8A(\frac{\Delta\theta_\gamma}{\theta_v})^2(\frac{1}{\gamma\theta_v})^2 2\gamma f(\nu\gamma\theta_v^2/2). \tag{23}$$

Then the peak frequency is in the X-ray region as $\nu_{peak} \sim \nu_0/(\gamma\theta_v)^2 \sim 4(\gamma/200)^{-2}(\theta_v/0.03)^{-2}$keV and the peak flux is given by

$$(\nu F_\nu)_X \sim 8A(\frac{\Delta\theta_\gamma}{\theta_v})^2(\frac{1}{\gamma\theta_v})^4\frac{2+\beta_B}{2^{\beta_B/2+1}} \tag{24}$$

The ratio of the X-ray flux observed from the off axis of the sub-jet to the gamma ray flux observed along the axis is given by

$$\frac{(\nu F_\nu)_X}{(\nu F_\nu)_\gamma} = 10^{-3}(1+\beta_B/2)(\frac{\Delta\theta_\gamma}{\theta_v})^2(\frac{\gamma}{200})^{-4}(\frac{\theta_v}{0.03})^{-4}. \quad (25)$$

For the gamma ray flux of $\sim 5 \times 10^{-6}$erg s^{-1}cm^{-2}, we have the X-ray flux of $\sim 5 \times 10^{-9}$erg s^{-1}cm^{-2}. This X-ray flux is similar to the observed ones in the precursors of GRB980519 [15] and GRB900126 [13]. Therefore our model is compatible with the X-ray flux of the precursor and the post-cursor.

4 Discussion

Let us try to interpret the X-ray photon index evolution of GRB980519 [15] stated in Introduction. In our model X-ray precursor and post-cursor come from the off-axis emission from sub-jets so that the spectrum is expressed by Eq. (23) which leads to the photon index $\beta = -\beta_B - 1 \sim -2.5$. While the main gamma ray emission comes from the on-axis emission from the sub-jet so that the spectrum is expressed by Eq. (19). In Eq.(19) the photon index in the X-ray band is -1. Therefore our model predicts that the photon index evolves from -2.5 to -1 to -2.5 which agree with the observations.

As we discussed in section 2, the duration of the emission from the sub-jet depends on the viewing angle. From Eqs. (4) and (7) , we have

$$\frac{\Delta T(\theta_v)}{\Delta T(\theta_v = 0)} = 36(\frac{\gamma_s}{100})^2(\frac{\theta_v}{0.03})^2(\frac{S'}{2})^2. \quad (26)$$

This means that the duration of the lower energy emission ($\theta_v \neq 0$) is longer than that of the higher energy emission($\theta_v = 0$). If the X-ray precursor or post-cursors occur by chance, our model in this *Letter* predicts that their duration are longer than the duration of the main gamma ray emission. This agrees with observations of precursors and post-cursors [13, 15].

However for most cases precursors and post-cursors do not occur. Then what does our model predict for usual GRBs? In our model, in general, there exists a sub-jet whose axis makes an angle $\theta_v \neq 0$ with the normal to the detector. This sub-jet mainly emits X-ray delaying from the main gamma ray emission from the sub-jet with $\theta_v \sim 0$. The duration of X-ray emission is longer than the main gamma ray emission so that our model predicts that prompt and delayed X-ray emission with longer duration should associate with GRBs. The emission patterns of GRB960720, 970111, 970228,970508, 971214, 980329 and 980425 qualitatively agree with this prediction [20] while other possible explanations may exist.

What is a new prediction from our model ? Suppose that by chance none of the sub-jets point to the line of sight. In this case gamma ray emission is extremely weak and short or absent while the X-ray precursors and/or post-cursors as well as X-ray, optical and radio afterglow exist. If such an event is found, that will be compatible with our sub-jet model discussed in this article.

I am very grateful Murakami, Kawai and Yoshida for useful comments. This work was supported in part by Grant-in-Aid of Scientific Research of the Ministry of Education, Culture, and Sports, No.11640274 and 09NP0801.

References

[1] Kulkarni, S.R. *et al.* 1999, astro-ph/9902272

[2] Fruchter, A.S. *et al.* 1999, Astrophys. J., 519 , L13

[3] Galama, T.J. *et al.* 1999, astro-ph/9903021 (Nature in press)

[4] Castro-Tirado, A.J. *et al.* 1999, Science , 283 , 2069

[5] Dai, Z. G. & Lu, T. 1999, Astrophys. J.,519 , L155

[6] Harrison, F. A. *et al.* 1999, Astrophys. J.,523 , L121

[7] Stanek, K. Z. *et al.* 1999, Astrophys. J., 522 , L39

[8] Halpern,J. P. , Kemp,J., Piran, T. & Bershady, M. A. 1999, , Astrophys. J., 517 , L105

[9] Rhoads, J.E. 1997, Astrophys. J., 478, L1

[10] Nakamura, T. 1999, Astrophys. J., 522 , L101

[11] Piran, T. 1998, Phys. Rep. 314, 575 and references therein

[12] Kumar, P. & Piran, T. 1999 astro-ph/9909014

[13] Murakami, T. *et al.* 1991, Nature, 350, 592.

[14] Sazonov, S. Y. *et al.* 1998, AAS , 129, 1.

[15] in't Zand, J. J. *et al.* 1999, Astrophys. J.,516 , L57

[16] Paczynski, B. 1998, Astrophys. J., 494 , L45

[17] Band, D. *et al.* 1993 , Astrophys. J., 413 , 281

[18] Garnot,J., Piran, T. & Sari, R. 1998, Astrophys. J., 513 , 679

[19] Woods, E. & Loeb, A. 1999, astro-ph/9903377

[20] Frontera, F. *et al.* 1999, astro-ph/9911228

Gravitational Waves from the Giant Flares of the Soft Gamma-ray Repeaters

Kunihito Ioka,[1]

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

Abstract

We present one possible mechanism for the giant flares of the Soft Gamma-Ray Repeaters (SGRs) within the framework of magnetar, i.e., superstrongly magnetized neutron star model, motivated by the positive period increase associated with the August 27 event from SGR 1900+14. We assume that the global reconfiguration of the internal magnetic field suddenly occurs. Then, the shape of the star is deformed due to the change in the internal hydromagnetic stresses, and this deformation causes the moment of inertia to increase and the total energy to decrease. With the internal field of $H \gtrsim 10^{16}$ G, we can explain the positive period increase $\Delta P_t/P_t \sim 10^{-4}$ as well as the energy $\gtrsim 10^{44}$ ergs of the giant flares. In this mechanism, since the oscillation of the neutron star will be excited, a pulsation of $\sim$ ms period in the burst profile and an emission of the gravitational waves (GWs) are expected. The energy of the GWs will be much larger than that of the gamma-ray, and the GWs may be detected by the planned interferometers such as LIGO, VIRGO and LCGT.

1 Introduction

Soft Gamma-ray Repeaters (SGRs) are persistent X-ray sources with X-ray luminosity $\sim 10^{35}$-10^{36} erg s^{-1} that emit brief (~ 0.1 s), intense ($\sim 10^{39}$-10^{42} erg s^{-1}) recurrent bursts of soft (~ 30 keV) gamma-ray (see [1, 2] for a review). It is considered that SGRs are better understood within the framework of *magnetar*, i.e., superstrongly magnetized neutron stars with $H \gtrsim 10^{14}$ G.

During the past ~ 20 years, two giant flares have been recorded from two of the SGRs: 1979 March 5 event from SGR 0526-66 and 1998 August 27 event from SGR 1900+14 (see [3] for references). These flares differ from the more common bursts in larger energies ($\gtrsim 10^{44}$ ergs), longer durations ($\gtrsim 100$ s) and harder initial spectra ($\sim$ MeV). Remarkably, during an 80 day interval which contains the August 27 event, the positive period increase of $\Delta P_t/P_t \sim 10^{-4}$ was observed [4]. At present, the exact cause for the positive period increase of SGR 1900+14 and the giant flares remains unknown [2, 5].

Here we would like to show that the global reconfiguration of the *internal* magnetic field of $H \gtrsim 10^{16}$ G can cause the giant flare of $\gtrsim 10^{44}$ ergs as well as the positive period increase of $\Delta P_t/P_t \sim 10^{-4}$.

2 Order-of-magnitude estimate

Let us consider a rotating star. There is a relation $J = \mathcal{I}\Omega$ between the angular momentum J, the angular velocity $\Omega = 2\pi/P_t$ and the moment of inertia $\mathcal{I}$ of the star. There are two ways to cause the spin-down: one way is to decrease the angular momentum [6] and the other is to increase the moment of inertia. Our mechanism is the latter one. If the angular momentum J is conserved, the required increase of the moment of inertia for the positive period increase of SGR 1900+14 is of order $\Delta\mathcal{I}/\mathcal{I} \sim \Delta P_t/P_t \sim 10^{-4}$.

Let us consider the internal magnetic field of $H \gtrsim 10^{16}$ G. This is somewhat larger than the external dipole fields $H \sim 10^{14}$ G deduced from the period and the period derivative. Such larger internal field $H \gtrsim 10^{16}$ G is plausible if the magnetic field is generated by the helical dynamo in a new born neutron star. The magnetic field is a source of non-hydrostatic stress in the interior of the star. Hence the internal magnetic field induces a deformation of the star, and the magnetic deformation dominates the rotational

[1] E-mail:iokakuni@tap.scphys.kyoto-u.ac.jp

one when $H \gtrsim 10^{14}(P_t/1\mathrm{s})^{-1}$ G. The induced difference in the moment of inertia of the star is given by order of

$$\frac{\Delta\mathcal{I}}{\mathcal{I}} \sim \delta \sim \frac{(H^2/8\pi)(4\pi R^3/3)}{GM^2/R} \sim 10^{-4}\left(\frac{H}{2\times10^{16}\mathrm{G}}\right)^2\left(\frac{R}{10^6\mathrm{cm}}\right)^4\left(\frac{M}{1.4M_\odot}\right)^{-2}, \tag{1}$$

where δ is the ratio of the magnetic energy $\mathcal{M} \sim (H^2/8\pi)(4\pi R^3/3) \sim 7\times10^{49}(H/2\times10^{16}\mathrm{G})^2$ergs to the gravitational energy $|\mathcal{W}| \sim GM^2/R \sim 5\times10^{53}$ergs, and M and R are the mass and the radius of the star respectively. The deformation of the star also causes the gravitational potential energy $\mathcal{W} \sim -GM^2/R$ to change by order of

$$\Delta\mathcal{W} \sim \delta^2|\mathcal{W}|/5 \sim 10^{45}\left(\frac{H}{2\times10^{16}\mathrm{G}}\right)^4\left(\frac{R}{10^6\mathrm{cm}}\right)^7\left(\frac{M}{1.4M_\odot}\right)^{-2} \quad \text{ergs.} \tag{2}$$

Here we consider globally coherent fields so that the deformation is elliptical like a rotating star. In this case, δ has the same meaning as the oblateness parameter ϵ in equation (10.11.2) and (10.11.7) of [7], and the energy shift is of order δ^2.[2]

Therefore, if the global rearrangement of the internal magnetic field $H \sim 10^{16}$ G occurs, the moment of inertia and the energy will change by order of $\Delta\mathcal{I}/\mathcal{I} \sim 10^{-4}$ and $\Delta\mathcal{W} \sim 10^{45}$ ergs respectively, which is comparable with the required values for the positive period increase and the giant flare of SGR 1900+14. Note that the energy source of this mechanism is not the magnetic energy but the gravitational energy. The spin-down rate before and after the giant flare is almost the same since the moment of inertia and the external magnetic fields are nearly constant.

However one may wonder why the spin-down occurs with a release of energy. In a rotating star case, such as in the starquake model for the ordinary glitches, the energy release is always associated with the spin-up. Therefore, we shall address the following problem: *does the spin-down occur with a release of energy in a magnetized star ?* To approach this problem, we consider the most idealized model. We prepare several equilibrium polytropes with different magnetic configurations assuming axisymmetry and no rotations. Then, we investigate the relation of the energy and the moment of inertia between these equilibria to find at least one example in which the moment of inertia increases with a release of energy.

3 The equilibrium of magnetic polytropes

The following is just an outline. See [8] for more details.

The basic equations governing the hydrostatic equilibrium of a perfect conducting fluid are Euler equation, $0 = -\nabla p + \rho\nabla\Phi + (1/4\pi)(\nabla\times\boldsymbol{H})\times\boldsymbol{H}$, Poisson equation, $\nabla^2\Phi = -4\pi G\rho$, and Maxwell equation, $\nabla\cdot\boldsymbol{H} = 0$, where p is the pressure, ρ is the density, Φ is the gravitational potential and $\boldsymbol{H}$ is the magnetic field. We assume the polytropic equations of state,

$$p = K\rho^{1+1/n}, \tag{3}$$

and we normalize the quantities as $\rho = \rho_c\Theta^n$, $r = a\xi$, $\boldsymbol{H} = (4\pi G\delta)^{1/2}\rho_c a\boldsymbol{h}$ and $\Phi = 4\pi G a^2\rho_c\phi$, where a is given by $a = \left[K(n+1)\rho_c^{-1+1/n}/4\pi G\right]^{1/2}$, Θ, ξ, $\boldsymbol{h}$ and ϕ are dimensionless quantities, n is the polytropic index, ρ_c is the density at the center, and K is a constant. Then, the basic equations are written as

$$0 = -\nabla\Theta + \nabla\phi + \delta\frac{(\nabla\times\boldsymbol{h})\times\boldsymbol{h}}{4\pi\Theta^n}, \quad \nabla^2\phi = -\Theta^n, \quad \nabla\cdot\boldsymbol{h} = 0, \tag{4}$$

where δ has nearly the same meaning as δ in equation (1), and ∇ is the divergence with respect to ξ.

We shall assume that the magnetic field is axisymmetric about the z-axis. Then we can express the magnetic field as a superposition of a poloidal and a toroidal field in terms of two scalar functions $P(\xi,\mu)$ and $T(\xi,\mu)$ in the manner,

$$\frac{\boldsymbol{h}}{(4\pi)^{1/2}} = \nabla\times(\varpi P\hat{\boldsymbol{e}}_\varphi) + \varpi T\hat{\boldsymbol{e}}_\varphi, \tag{5}$$

[2] If the magnetic fields were well tangled, the deformation would not be elliptical but radial so that the energy shift is of order δ. However we don't consider such a case.

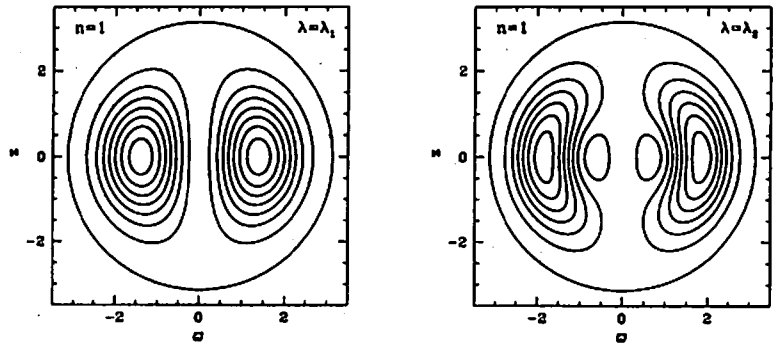


Figure 1: The projection of the lines of force on the meridional plane is shown for the case of $n = 1$. λ controls the configuration of the magnetic fields and takes the discrete values ($\lambda_2 > \lambda_1$). The lines of force for larger λ have the more complicated structure than that for smaller λ.

where $\hat{e}_\varphi$ is a unit vector along the φ-direction and (ϖ, φ, z) denotes the cylindrical coordinate, which is related to the spherical coordinate (ξ, θ, φ) as $\varpi = \xi \sin\theta = \xi(1-\mu^2)^{1/2}$ and $z = \xi\cos\theta = \xi\mu$. Since equation (4) requires $\nabla\times[(\nabla\times\boldsymbol{h})\times\boldsymbol{h}/\Theta^n] = 0$, it can be shown that

$$0 = \nabla\left[\Theta - \phi - \delta N_P(\varpi^2 P)\right], \tag{6}$$

$$\Delta_5 P = -\Theta^n N_P'(\varpi^2 P) - T N_T'(\varpi^2 P), \tag{7}$$

$$\varpi^2 T = N_T(\varpi^2 P), \tag{8}$$

where $N_T(x)$ and $N_P(x)$ are arbitrary functions which characterize the configuration of the magnetic fields, and $\Delta_5 = \partial^2/\partial\varpi^2 + (3/\varpi)\partial/\partial\varpi + \partial^2/\partial z^2$ is the five-dimensional axisymmetric Laplacian. Taking the divergence of equation (6), we obtain

$$\nabla^2\Theta = -\Theta^n + \delta\nabla^2 N_P(\varpi^2 P), \tag{9}$$

with equation (4).

We shall treat δ as a small perturbation parameter, and obtain the solutions for Θ, P and T to the second-order in δ,

$$\Theta = \Theta_0 + \delta\Theta_1 + \delta^2\Theta_2, \quad P = P_0 + \delta P_1, \quad T = T_0 + \delta T_1. \tag{10}$$

The basic equations (7), (8) and (9) can be solved order by order. It is obvious that depending upon the choice of the arbitrary functions $N_P(x)$ and $N_T(x)$, there can be many solutions for the functions P and T which are consistent with the demand for equilibrium. Here, for simplicity, we shall consider one parameter family of the magnetic fields which is characterized by

$$N_P(x) = -x, \quad N_T(x) = \lambda x, \tag{11}$$

where λ controls the configurations of the magnetic fields. We assume that the fields vanish at the boundary of the star. In Figure 1, the projection of the lines of force on the meridional plane is shown for the case of $n = 1$. The lines of force for larger λ have the more complicated structure than that for smaller λ, i.e., the number of the smaller structure like the inner loops for $\lambda = \lambda_2$ increases as λ increases.

Using the above solutions, we can calculate the total energy $\mathcal{E}$ and the kk-component of the moment of inertia tensor $\mathcal{I}_{kk}$ perturvatively in the following forms,

$$\mathcal{E}(\lambda) = \mathcal{E}_0 + \mathcal{M} + [\delta(\lambda)]^2\,\widehat{\mathcal{M}}_2(\lambda), \quad \mathcal{I}_{kk}(\lambda) = \mathcal{I}_0 + \delta(\lambda)\mathcal{I}_{kk;1}(\lambda), \tag{12}$$

where $\mathcal{E}_0$ is the zeroth-order total energy, $\mathcal{M}$ is the magnetic energy ($\mathcal{M}/\mathcal{E} \sim \delta$), $\delta^2\widehat{\mathcal{M}}_2$ is a correction due to the coupling of the first-order quantities, and $\mathcal{I}_0$ is the zeroth-order moment of inertia. Let us consider

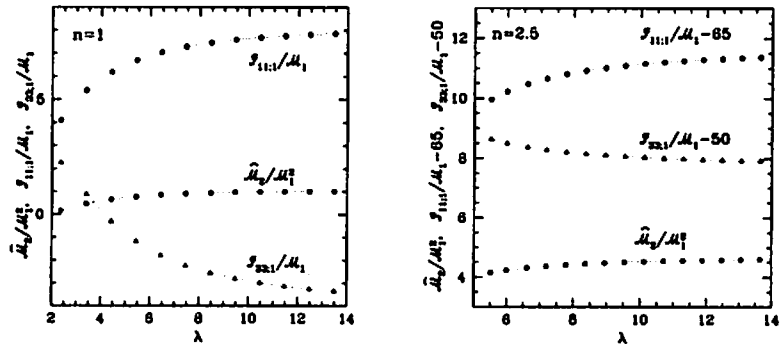


Figure 2: The quantities, $\widehat{\mathcal{M}}_2/\mathcal{M}_1^2$, $\mathcal{I}_{11;1}/\mathcal{M}_1$ and $\mathcal{I}_{33;1}/\mathcal{M}_1$, are plotted as a function of λ for the polytropic indices $n = 1$ and $n = 2.5$. λ controls the configuration of the magnetic fields and takes the discrete values. As λ decreases, $\widehat{\mathcal{M}}_2/\mathcal{M}_1^2$ decreases, $\mathcal{I}_{11;1}/\mathcal{M}_1$ decreases, and $\mathcal{I}_{33;1}/\mathcal{M}_1$ increases, which means that the energy decreases, the moment of inertia of 11-component decreases, and the moment of inertia of 33-component increases, respectively.

that the magnetic configuration changes from λ_i to λ_f. For simplicity, we assume that the magnetic energy $\mathcal{M}$ is nearly conserved in this transition. Then, the difference in the total energy and the moment of inertia between two configuration is given by

$$\frac{\Delta\mathcal{E}}{|\mathcal{E}_0|} = \delta_i^2 A\left[\frac{\widehat{\mathcal{M}}_2(\lambda_f)}{\mathcal{M}_1^2(\lambda_f)} - \frac{\widehat{\mathcal{M}}_2(\lambda_i)}{\mathcal{M}_1^2(\lambda_i)}\right], \quad \frac{\Delta\mathcal{I}_{kk}}{\mathcal{I}_0} = \delta_i B\left[\frac{\mathcal{I}_{kk;1}(\lambda_f)}{\mathcal{M}_1(\lambda_f)} - \frac{\mathcal{I}_{kk;1}(\lambda_i)}{\mathcal{M}_1(\lambda_i)}\right], \tag{13}$$

where A and B are positive quantities of order $O(1)$. In Figure 2, we plot $\widehat{\mathcal{M}}_2/\mathcal{M}_1^2$ and $\mathcal{I}_{kk;1}/\mathcal{M}_1$ as a function of λ for the cases of $n = 1$ and $n = 2.5$. λ takes discrete values. We can see that, if the magnetic configuration changes from the state of larger λ to that of smaller λ, $\widehat{\mathcal{M}}_2/\mathcal{M}_1^2$ decreases while $\mathcal{I}_{33;1}/\mathcal{M}_1$ increases. This means that the total energy decreases while the moment of inertia tensor along the magnetic axis increases from equation (13). Therefore, *the spin-down with a release of energy is possible in a magnetized star.* Note that the configuration of the polytrope is prolate rather than oblate since the moment of inertia tensor of 11-component, $\mathcal{I}_{11;1}$, is larger than that of 33-component, $\mathcal{I}_{33;1}$, from Figure 2.

4 Discussion

The above results can be understood as follows. The star is prolate at the first state. As the total energy decreases, the star becomes more spherical since the spherical configuration is the most stable. On the other hand, the moment of inertia increases since the radius in the equatorial plane increases. Hence the spin-down is a natural consequence of the internal field reconfiguration. In this way, with the internal magnetic field of $H \gtrsim 10^{16}$ G, we can explain the positive period increase $\Delta P_t/P_t \sim 10^{-4}$ as well as the energy $\gtrsim 10^{44}$ ergs of the giant flare of SGR 1900+14.

This model predicts a little bit larger energy than that of the observed gamma-ray energy. However, the released energy will be firstly converted into the neutron star oscillation, and almost all oscillational energy will be released as gravitational waves. We can make a rather crude estimate as follows. The luminosity of the gravitational waves can be estimated by the quadrupole formula,

$$L_{GW} = \frac{1}{5}\frac{G}{c^5}\left\langle \dddot{I}_{jk}\dddot{I}_{jk}\right\rangle \sim \frac{G}{c^5}\left(\frac{\Delta I}{\Delta T^3}\right)^2 \sim \delta^2\left(\frac{c^5}{G}\right)\left(\frac{GM}{c^2R}\right)^5, \tag{14}$$

where $\Delta I \sim \delta M R^2$ is the quadrupole of the neutron star, and $\Delta T \sim \sqrt{R^3/GM}$ is the free fall time. We estimate the luminosity of the electromagnetic waves by the dipole formula. Since there are relations,

$R^2H \sim$ const, $\mu \sim HR^3$, and $\Delta R \sim \delta R$, where μ is the magnetic dipole moment, we can obtain

$$L_{EM} = \frac{2}{3c^3}|\ddot{\mu}|^2 \sim \frac{1}{c^3}\left(\frac{\Delta\mu}{\Delta T^2}\right)^2 \sim \delta^2(cH^2R^2)\left(\frac{GM}{c^2R}\right)^2. \quad (15)$$

The ratio $L_{EM}/L_{GW} \sim \delta(c^2R/GM)$ is small for typical parameters. Since the above estimate is too crude, we need a more precise calculation for the energy conversion efficiency from the neutron star oscillation to the gamma-ray.

The difficulty in this model is the assumption of the sudden global reconfiguration of the internal magnetic field. There is no firm ground to assume that such a magnetic reconfiguration can occur inside the neutron star. In addition, the internal magnetic energy has to be nearly conserved in the magnetic reconfiguration ($|\Delta\mathcal{M}/\mathcal{M}| \lesssim 10^{-4}$). This is because the internal magnetic energy is so large ($\mathcal{M} \sim 10^{49}$ ergs) that a slight decay of the internal field leads to a large release of the energy. Anyway, it is an important fact that we can make two parameters (the energy and the moment of inertia) have the right values by using only one parameter (the decaying magnetic energy $\Delta\mathcal{M}$).

This model can be tested by the future observations [8], such as the observation with the fine time resolution ($\sim$ ms) to detect the neutron star oscillation and the observation of the gravitational waves by the planned interferometers (LIGO, VIRGO and LCGT). Such observations will be able to clarify whether the energy source is internal or external, and may open a window to the interior of the magnetar.

Acknowledgements

We would like to thank R. C. Duncan, J. I. Katz and P. M. Woods for useful comments. This work was supported in part by Grant-in-Aid for Scientific Research Fellowship (No.9627) of the Japanese Ministry of Education, Science, Sports and Culture.

References

[1] Hurley, K., 1999, astro-ph/9912061.

[2] Thompson, C., 2000, astro-ph/0010016.

[3] Feroci, M., et al., 2000, astro-ph/0010494.

[4] Woods, P. M., et al., 1999, ApJ, 524, L55.

[5] Ibrahim, A. I., et al., 2000, astro-ph/0007043.

[6] Thompson, C., et al., 1999, ApJ, 543, 340.

[7] Shapiro, S. L., Teukolsky, S. A., 1983, Black Holes, White Dwarfs and Neutron Stars (New York: Wiley).

[8] Ioka, K, 2000, astro-ph/0009327.

Origin of Gamma-Ray Bursts

— From the viewpoint of iron line observations —

Murakami Toshio and Yonetoku Daisuke,[1]

Institute of Space and Astronautical Science,
3-1-1, Yoshinodai, Sagamihara, Kanagawa 229-8510, Japan

Abstract

The recent observational results on Gamma-Ray Bursts (GRBs) in the X-ray, radio and optical bands, since the X-ray afterglow discovery with BeppoSAX in 1997, can be explained by the so-called fireball models. These observations clearly indicate that the distances to the GRBs are remote. The mean redshift is almost $z \sim 1$. Thus, these sources represent the largest explosions in the universe. However, the nature of the progenitor of GRBs is an unresolved issue. Merging of a binary neutron star system and the collapse of a massive star or a hypernova are the most promising models proposed. Clear observational evidence discriminating between these models is still lacking. I emphasize that the evidence can be gathered through measurements of the iron feature in the X-ray afterglow spectrum produced by the medium surrounding the GRB progenitor. ASCA, BeppoSAX and, most recently, Chandra have detected strong iron features in some of the GRBs, but curiously the lines were not common to all GRBs. The reported iron lines were strong, variable and somewhat broad. There are several proposed models but, in any case, the existence of the strong iron lines is very difficult to explain with merger models but is more consistent with a massive collapsar as a progenitor.

1 Introduction

Since the discovery of an X-ray afterglow with BeppoSAX in 1997[1], the distances to GRBs are known to be very remote. Most of them are at cosmological distances with a mean redshift (z) of nearly 1. The remotest case is $z = 4.5$ [2]. Thus, taking into account the probable beaming of GRBs, the required energy for a burst is estimated to be on the order of 10^{52} ergs [3]. Moreover, the Optical Transients (OTs), associated with GRBs, were discovered not so far from the center of their host galaxies but very close to the so-called star-forming region of the galaxies [5]. This does not mean that the GRBs are associated with the AGN activities of the galaxy. Some of the OT showed a very large extinction of R - K $\sim$ 4 mag., which might be expected from dense matter [4].

These observational results strongly favor the fireball model [6] ,which is produced by a hypernova or collapsar of a massive star as a progenitor of GRBs rather than the binary merger, which might not be associated with the star-forming region [7]. However, clear observational evidence discriminating between these various models is lacking. I think that this evidence can be obtained from iron line observations of the X-ray afterglow spectra. The iron line is thought to be the best evidence of exploded matter or circumstellar matter surrounding a GRB progenitor.

Although there were two early reports of iron line detections in 1997 with ASCA and BeppoSAX, these observations were statistically insignificant[1], [8]. Also, these observations were very much limited in the number of detections. In fact, there was only one burst each for the ASCA and the BeppoSAX satellite. Moreover, the statistical significance of the detections was not more than three sigma. Therefore, we require further observations to establish the existence of the iron lines.

Thus, further extensive efforts have been made with ASCA, Chandra and BeppoSAX to observe GRBs and test the existence of the iron line. I report on the present understanding of the iron line observations and the most preferred progenitor model derived from the iron line observations.

[1] E-mail:murakami@astro.isas.ac.jp

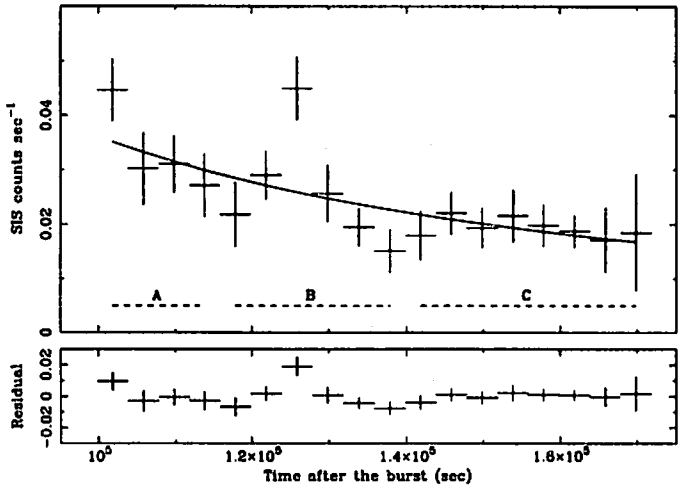


Figure 1: SIS light curve with time bins of 4000 s showing the time variation of the source brightness. A simple power-law decay model cannot fit the data; the solid line shows the "best-fit" model. Dashed lines labeled "A", "B" and "C" indicate the intervals for which spectral studies are done [8].

2 Iron line detection

The first iron line detections were reported independently from BeppoSAX and ASCA at the 2nd Rome GRB conference in 1998 as shown in figures 2 and 3 [1], [8]. Both iron lines were variable and strong. For the case of ASCA the iron line emission was detected only during the short interval B as shown in Figure 1. Thus the intensities of the detected lines was variable.

The lines were strong with equivalent widths of almost 1 to 3 keV and suggested some amount of line broadening. However, as mentioned above, neither detection was statistically significant. According to the F-test, the additional iron line parameters were significant at the 98.3% and 99.3% levels for ASCA and BeppoSAX respectively. Neither had a significance of more than three sigma. Moreover, although the redshift distance of the BeppoSAX observation was consistent with the redshift of the OT, the derived ASCA redshift of ~ 0.33 from the observed line energy of 5.04 keV, assuming an iron Kα line, did not match the OT distance of 0.9578 from the Keck observation of the host galaxy [8], [9]. Due to the low confidences and the discrepancy in the distance, the evidence for the existence of the iron line in the X-ray afterglow spectra is marginal. Thus, we are required to observe more GRBs with better S/N ratio.

ASCA has performed two more observations with better S/N than the first iron line detection of GRB 970828. However, we could not confirm the presence of an iron line, but on the contrary, could set extremely low upper limits of nearly 100 eV (EW) for GRB 990704 and GRB 990123, which has a known redshift distance of z=1.60 [11]. The ASCA iron feature detection from GRB 970828 is shown in figure 2 and the ASCA non-detection from GRB 990123 is shown in figure 4 together with the time interval of GRB 970828, where the iron line was detected.

In 1999, a new observation of GRB 991216 with Chandra resolved the question and clearly detected the iron feature, not only with the CCD detectors, but also with the grating spectrometer at the 4.7 sigma confidence level. Moreover, Chandra detected not only the redshifted iron line with a slight broadening but also the recombination edge of the fully ionized Hydrogen-like iron at 9.28 keV in the rest frame in the X-ray afterglow spectrum. The grating spectrum is shown in figure 5 [10]. The measured redshift is also consistent with that of the OT. Soon after the report from Chandra, BeppoSAX reported a much clearer iron line detection at the five sigma level from GRB 000214 as shown in figure 6 [12].

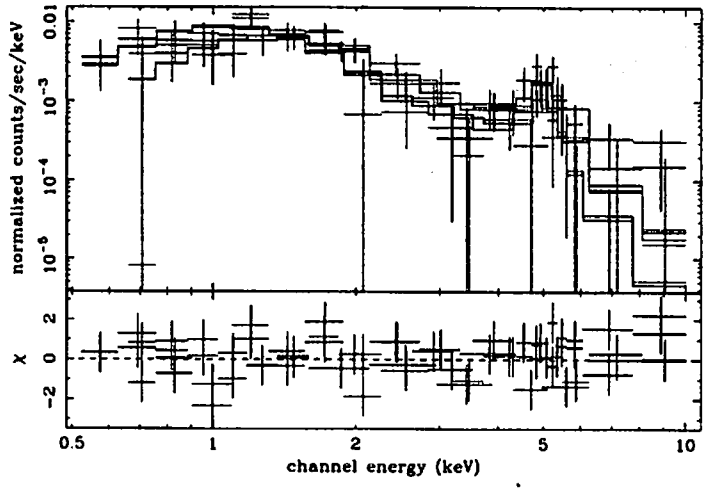


Figure 2: The X-ray spectrum for the period (B) shown in Fig. 1 fitted with the model of "absorbed power-law" ($\exp(-N_{\rm H}\sigma(E)) * E^{-\alpha}$). The line is somewhat broadened to 0.31 keV [8].

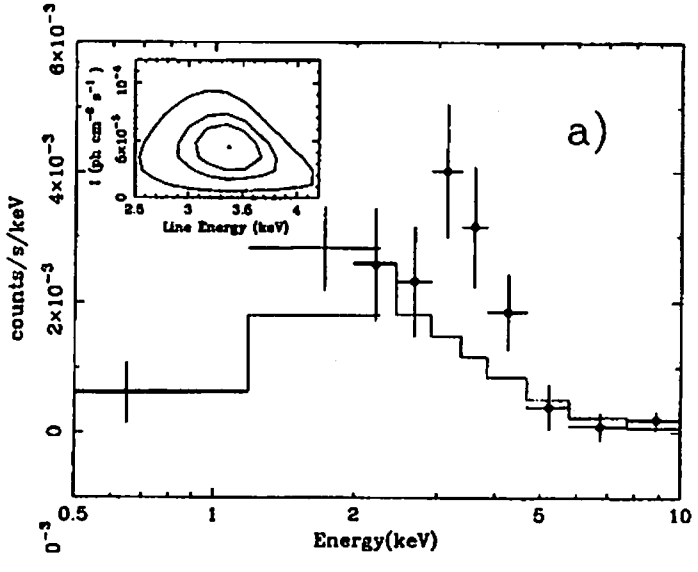


Figure 3: Spectrum of the afterglow of GRB 970508 with BeppoSAX. The continuous line represents the best-fit power-law model in the 0.1–10 keV band excluding the 3–4 keV interval. The inset shows contours of the line intensity vs. energy. Contours correspond to 68, 90 and 99% confidence levels [1].

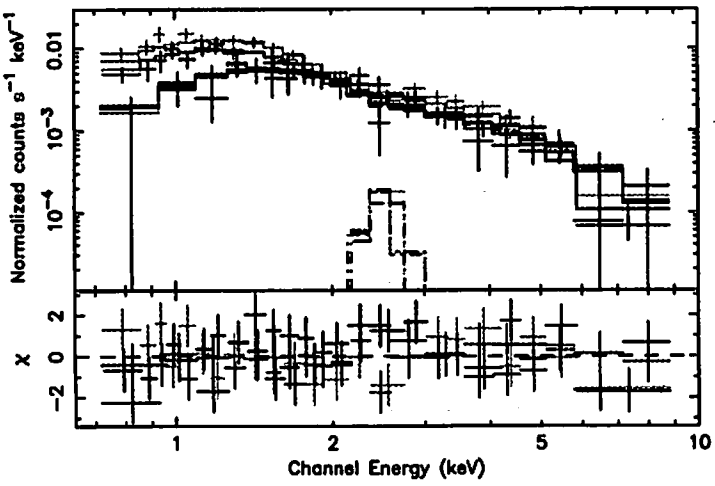


Figure 4: X-ray spectrum of GRB 990123 combined with the four ASCA detectors into one figure. The spectrum can be explained without any iron structure at $z = 1.60$. The 90 % upper limit for the 6.4 keV line intensity is 3.3×10^{-6} photons cm^{-2} s^{-1} or EW = 100 eV. The line inserted in this figure is the 90% upper limit [11].

Thus, it is now clear that the existence of the iron feature is established and also it is clear that all GRBs did not show the iron line structure. A very limited number of GRBs (less than 10 %) showed the iron line and the lines were variable.

3 Discussion

The observational results from ASCA. BeppoSAX and Chandra are now enough to confirm the existence of a strong iron feature in the X-ray afterglows of GRBs. However, the iron feature is not always associated with the GRBs. The line is very rare, appearing in 10% or less of GRBs, and two cases of iron line detection showed that the lines were only present during a short interval of 10^{4-5} s in duration [1], [8].

The iron feature really exists, so we must solve the discrepancy in distance between ASCA and Keck. The discrepancy between the ASCA distance of $z \sim 0.33$ and $z = 0.9578$ in the OT distance can be understood if the feature observed with ASCA was not the Kα line but the recombination edge of Hydrogen-like iron at 9.28 keV in the rest frame. It is very hard to imagine a strong recombination edge with no Kα line. In fact, the BeppoSAX observations detected the Kα lines and the Chandra observation indicates a strong recombination feature together with a strong Kα line [10]. However, the ASCA case can be realized in the non-equipartition condition of a low electron temperature but a high inonization degree of irons [13]. In this case, the recombination feature is strong but the iron Kα line is weak.

To explain the existence of the iron feature, two competing models of a progenitor of a fireball are proposed. One model is optically thin but assumes a large emission region of about 10^{16} cm in size because of a typical time scale of variability of 10^{4-5} s and a delay of the observation after the GRB explosion. The other model assumes a much smaller region with a thick target of almost 10^{17} cm^{-3} in density.

For the latter case, the recombination time is so short due to the high density that we require continuous illumination where the iron should be ionized to produce the intense iron feature. If we assume ordinary abundance of matter for the former scenario, the required mass of matter to produce

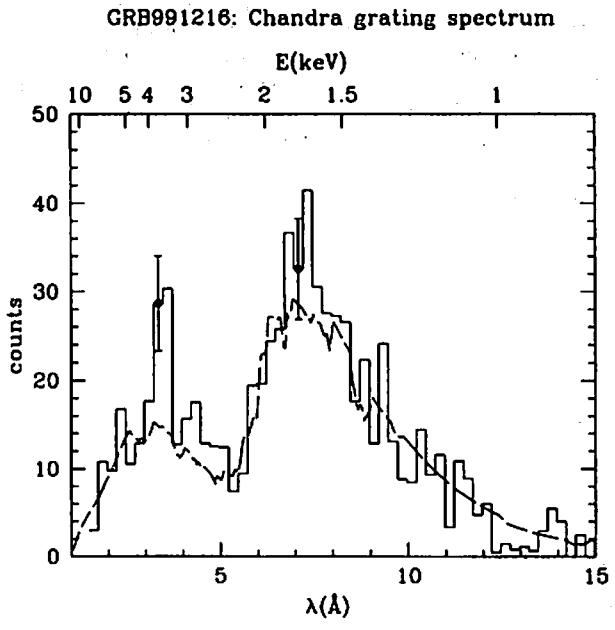


Figure 5: X-ray afterglow spectrum of GRB 991216 obtained with the Chandra high energy gratings. The grating spectrum has a bin size of 0.25 Å. The dashed line represents the best fit power-law on the 0th order ACIS-S spectrum. The peak at around 3.5 Å(E=3.5 keV) is detected with 4.7 sigma confidence [10].

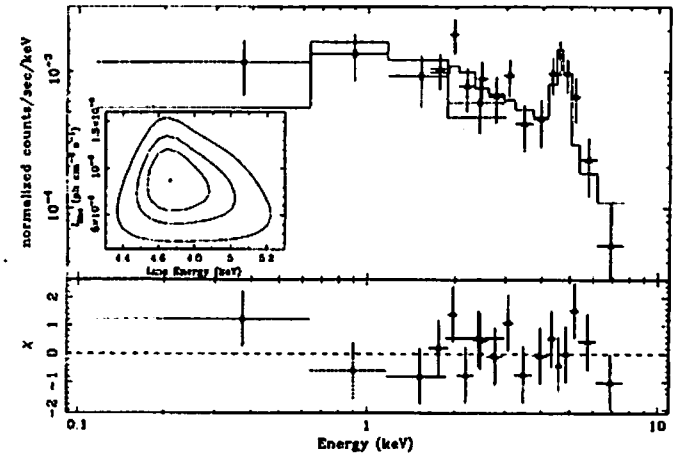


Figure 6: BeppoSAX spectrum of GRB 000214 X-ray afterglow. The LECS + MECS spectrum is fitted with an absorbed power-law plus a narrow Gaussian line. The insert shows the contour plot of the line intensity vs energy. Contours correspond to 68, 90 and 99% confidence levels [12].

the strong iron emission is huge, due mainly to the low emissivity of a thin ionized gas [14]. The required mass is nearly a solar mass. However, in the latter case, the density and the resultant emissivity are so high that the required mass is only 10^{-5} solar mass [15]. Which is better for explaining the observational facts? The short time variability favors the latter scenario but requires a continuous energy source like a magnetar in the center of the GRB progenitor to maintain the short recombination time [15].

In summary, the existence of the iron feature is now real and established and the iron feature's existence favors the massive collapsar type models to explain the strong iron line. It is difficult for the binary merger type models, which assume a location far from the star forming region, to produce the required mass of the observed iron.

I thank Damian Audley at ISAS for his critical reading.

References

[1] L. Piro, *et al.*, ApJL, **514**, (1999) L73.

[2] H. Pedersen, *et al.*, ESO press release 20/00, (2000).

[3] S. Kulkarni, *et al.*, Nature, **398**, (1999) 389.

[4] S. Klose, *et al.*, Astroph/0007201, (2000).

[5] J.S. Bloom, *et al.*, Astroph/0010176, (2000).

[6] M. Rees and P. Meszaros, MNRAS, **258** (1992) S41.

[7] T. Piran, Physics Report, **333**, (2000) 529.

[8] A. Yoshida, *et al.*, A&AS. **138**, (1999) 433.

[9] S.G. Djorgovski, *et al.*, ApJ. submitted, (2000)

[10] L. Piro, *et al.*, Science, **290** (2000) 953.

[11] D. Yonetoku, *et al.*, PASJ, **52**, (2000) 509.

[12] L.A. Antonelli, *et al.*, Astroph/0010221, (2000)

[13] K. Masai, ApJ, **437** (1994) 770.

[14] M. Vietri and L. Stella, Astroph/9910008 (1999).

[15] M.J. Rees and P. Meszaros, Astroph/0010258 (2000).

Ultra-High Energy Cosmic Rays and Superheavy Particles in the Universe

V.A. Kuzmin [1]

Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan
Institute for Nuclear Research of Russian Academy of Sciences, Moscow, Russia

Abstract

In this lecture we conjecture that the Ultra-High Energy Cosmic Rays (UHECR) with energies $E > E_{GZK}$, where $E_{GZK} \sim 5 \cdot 10^{19}$ eV is the Greisen-Zatsepin-Kuzmin cut-off energy of cosmic ray spectrum, may originate from decays of superheavy long-living X-particles populating the Universe thus providing a unique window into the very early epoch of the Universe, namely, that of preheating and reheating after inflation. These particles may constitute a considerable fraction of cold dark matter in the Universe. We argue that the unconventionally long lifetime of the superheavy particles, which should be in the range of $10^{10} - 10^{22}$ years, might require novel particle physics mechanisms of their decays, such as instantons. First of all I will describe a toy model illustrating the instanton scenario and then will describe the possible proper mechanism of creation of superheavy particles in the early Universe. It is my pleasure to emphasize that all I will talk about was done in collaboration with my friends V.A. Rubakov and I.I. Tkachev.

1 Introduction

1. Among the top-down mechanisms attempting to explain the observations [2] of ultra-high energy (UHE) cosmic rays beyond the Greisen-Zatsepin-Kuzmin cut-off [1] , decays of primordial heavy particles are most obvious possibility. It is clearly more conventional — at least at first sight — than scenarios invoking topological or non-topological defects [19], though the latter may be a viable alternative. Heavy particles with lifetime of the order of the age of the Universe or greater may constitute (a substantial fraction of) cold dark matter, so the two observed features of the Universe — ultra-high energy cosmic rays and dark matter — may be related to each other.

To get an idea of the range of properties of decaying particles (X-particles) that supposedly produce UHE cosmic rays, let us make the following simple observations.

First, assuming sizeable hadronic component (jets) among the decay products, the flux of protons or gammas of energy E on the Earth is estimated as

$$\frac{dF}{d\,lnE} = \frac{1}{4\pi}\frac{n_X}{\tau_X} R_{p,\gamma} N_j \frac{dN_{p,\gamma}(E)}{d\,lnE}, \qquad (1)$$

where N_j is the number of jets in a typical decay, $R_{p,\gamma}$ is the effective distance to X-particles, n_X is the number density of X-particles at the scale $R_{p,\gamma}$, τ_X is the X-particle lifetime and $dN/d\,lnE$ is the fragmentation function. In the following estimates we take $N_j \sim 1 - 10$ (we will soon see that large jet multiplicity may be favored in some decay scenarios) and $dN_{p,\gamma}/d\,lnE \sim$ (a few) $\cdot$ $(10 - 100)$ in the energy range of interest, $E >$ (a few) $\cdot\, 10^{10}$ GeV (the latter estimate comes from bold extrapolation of the fragmentation functions of ref.[20] to extremely high jet energies). For the effective distance we take $R \lesssim 100$ Mpc, with understanding that the actual value of R may be much smaller than 100 Mpc if X-particles are clumped. In fact, our conclusions will be fairly insensitive to the actual values of the above parameters.

The second relation is

$$m_X < n_X >= \Omega_X \rho_{crit}, \qquad (2)$$

[1] E-mails : kuzmin@ms2.inr.ac.ru, kuzmin@itp.phys.ethz.ch

where $< n_X >$ is the average number density of X-particles and $\Omega_X \lesssim 1$. If X-particles are clumped, the density n_X entering Eq.(2) may be several orders of magnitude larger than $< n_X >$. Again, this uncertainty will not affect our main conclusions, and we set $n_X \sim < n_X >$ in what follows. In order to produce cosmic rays of energies $E \gtrsim$ (a few) $\cdot 10^{11}$ GeV, the mass of X-particles is to be very large, $m_X \gtrsim 10^{13}$ GeV.

Let us now estimate the range of X-particle densities required. From Eq.(2) we find a bound for the present X-particle density-to-entropy ratio

$$n_X/s \lesssim 10^{-21}. \tag{3}$$

On the other hand, to produce the observed flux of UHE cosmic rays, the density of X-particles should not be too small. Keeping in mind that their lifetime, τ_X, cannot be much smaller than the age of the Universe,

$$\tau \gtrsim 10^{10}\ yr, \tag{4}$$

we obtain from Eq.(3)

$$n_X/s \gtrsim 10^{-33}. \tag{5}$$

Even though the window for n_X is very wide, the estimates (3) and (5) raise the issue of the production of X-particles in the early Universe.

Alternatively, Eqs.(3) and (5) may be used to place an upper bound on the lifetime of X-particles,

$$\tau_X \lesssim 10^{22}\ yr. \tag{6}$$

Again, the window for τ_X is wide, but the estimates (5) and (6) indicate another problem, namely, that of the particle physics mechanism responsible for long but finite lifetime of very heavy particles.

In the rest of this paper we propose possible scenarios for i) generating the abundance of X-particles in the range (3), (5) by processes in the early Universe, and ii) explaining the lifetime of X-particles in the range (4), (6).

Before coming to our main points, let us stress that these two problems inherent in theories with very heavy and almost stable particles were realized long ago (see, e.g., ref.[21] and references therein) in different contexts. In ref.[21] it was proposed that the problem i) may be solved by large entropy generation in the Universe after the heavy particles freeze out of thermal equilibrium, while their long lifetime may be due to very large dimension of operators responsible for the decay. We leave for the reader to judge how exotic are alternative possibilities that we discuss below.

2. If the temperature in the early Universe at some epoch exceeded the mass of X-particles and then decreased smoothly without large entropy generation, the freeze-out density of X-particles would greatly exceed the bound (3). A way out of this problem is provided by inflation and subsequent reheating. To explain the small abundance of X-particles, the reheating temperature T_r must be much smaller than m_X, so that X-particles were never at thermal equilibrium after inflation. In that case X-particles-to-entropy ratio is exponentially small,

$$n_X/s = \text{const} \cdot \exp(-2m_X/T_r), \tag{7}$$

where the constant depends on a number of factors (the coupling constant responsible for pair production of X-particles, the effective number of degrees of freedom, the ratios m_X/T_r and M_{Pl}/m_X, etc.) and is of order 10^{-3} with several orders of magnitude uncertainty. As the dominant suppression comes from the exponential factor, the reheating temperature can be estimated from (3), (5) fairly precisely,

$$T_r = \left(\frac{1}{20} - \frac{1}{35}\right) m_X, \tag{8}$$

and should be in the range $10^{11} - 10^{15}$ GeV, depending on m_X. Note that this range is realistic in many scenarios of inflation.

Hence, inflationary scenario can easily explain small value of the present density of X-particles in the space. Conversely, the determination of m_X from measurements of the upper end of cosmic ray spectrum would allow for rather precise estimate of the reheating temperature. Ultra-high energy cosmic rays may indeed serve as a window to reheating epoch in the early Universe.

3. Explaining long lifetime of X-particles is much harder. Conventional perturbative mechanisms cannot be responsible for cosmologically large τ_X (unless very high dimension operators are involved [21]), so one turns naturally to non-perturbative phenomena. A well known example is instantons that produce exponentially small effects in weakly coupled theories. If instantons are responsible for X-particle decays, the lifetime is roughly estimated as

$$\tau_X \sim m_X^{-1} \cdot \exp(4\pi/\alpha_X), \tag{9}$$

where α_X is the coupling constant of the relevant (spontaneously broken) gauge symmetry. From Eqs.(4), (6) we find that the coupling constant (at the scale m_X) is

$$\alpha_X = \frac{1}{10} - \frac{1}{12}. \tag{10}$$

Hence, we are lead to introduce additional non-Abelian gauge interactions with fairly large coupling constant at high energy scale.

To illustrate this possibility let us consider a toy model with $SU(2)_X$ gauge interactions added to the Standard Model. The $SU(2)_X$ gauge symmetry is assumed to be broken at sufficiently high energy scale. Some conventional quarks and leptons carry non-trivial $SU(2)_X$ quantum numbers (say, $SU(2)_X$ may be right-handed subgroup of a left-right symmetric theory, or it may be a horizontal group with generations forming $SU(2)_X$ triplets). Let there be two[2] left-handed $SU(2)_X$ fermionic doublets X and Y and four right-handed singlets, all of which are singlets under $SU(2)_L \times SU(3)_c$ of the Standard Model. After $SU(2)_X$ breaks down, all X- and Y-particles acquire large masses in a manner similar to the Standard Model. We further assume that X and Y carry different global quantum numbers, so there is no mixing between them.

Under these assumptions the lightest of X-particles and the lightest of Y-particles (we call them X and Y at certain risk of confusing notations) are perturbatively stable. However, $SU(2)_X$ instantons induce effective interactions violating global quantum numbers of X and Y. Say, if X is heavier than Y, then $SU(2)_X$-instantons lead to the decay

$$X \to Y + quarks + leptons \tag{11}$$

with the rate estimate given by Eq.(9). It is this type of processes that may be responsible for the production of ultra-high energy cosmic rays.

Let us point out a few features of this scenario which seem generic.

i) Decays induced by instantons typically lead to multiparticle final states. The number of quarks (jets) produced in the process (11) should be rather large, of order 10, and their distribution in energy should be fairly flat. In principle, the spectrum of cosmic rays within this scenario should be distinguishable from the spectrum predicted by other mechanisms (like two-jet decays of heavy particles born in the interactions of topological defects). Also, there are necessarily hard leptons among the decay products in the process (11).

ii) If Y-particles are indeed perturbatively stable, they are also stable against instanton-induced interactions (because of energy conservation and instanton selection rules). Then the dark matter in the Universe may consist predominantly of Y-particles, while the admixture of X-particles is small. Since the abundance of Y-particles is given by the formula similar to Eq.(7), and since the density of X-particles is bounded from above, Eq.(5), the mass splitting between X- and Y-particles should not be large,

$$m_X < 2m_Y. \tag{12}$$

Alternatively, the Higgs sector and its interactions with fermions may be organized in such a way that Y-particles are in fact perturbatively unstable (while X-particles remain perturbatively stable). In that case the heavy candidates for dark matter are X-particles, and the approximate degeneracy (12) need not hold.

iii) Because of the instanton selection rules, this scenario for the slow decay of heavy particles is rather restrictive. For example, X-particles cannot be colored (otherwise the $SU(2)_X$ instanton vertex would

[2] The $SU(2)_X$ anomaly prevents the number of $SU(2)_X$ doublets from being odd.

include at least three X-fields); if Y-particles are stable, they cannot be colored either. This fits nicely to the expectation that dark matter particles do not experience strong interactions. On the other hand, X-particles (and Y-particles, if stable) cannot be weak doublets for the same reason, so the heavy dark matter in this scenario does not have electroweak interactions, too.

To conclude, an explanation of ultra-high energy cosmic ray events beyond the GZK cut-off by decays of hypothetical heavy particles of cosmologically long lifetime is not unrealistic from both cosmological and particle physics points of view. Detailed study of the upper end of cosmic ray spectrum will provide insight into the decay mechanism involved, and allow for the determination of the mass of X-particles. As the latter has been argued to be related to the reheating temperature, the ultra-high energy cosmic rays may become a clue to the end-of-inflation epoch in the early Universe.

After this work was completed , we received a paper by Berezinsky, Kachelriess and Vilenkin [22] where the decays of heavy particles have been also considered as the origin of cosmic rays beyond the GZK cut-off. Their main point is that X-particles are expected to concentrate in the galactic halo, and in this way one easily avoids the constraints coming from the analysis of the cascade radiation [23]. Their proposal for long but finite lifetime of X-particles is that it is due to quantum gravity (wormhole) effects.

Now I would like to consider yet another problem concerning the population of the Universe with superheavy particles.

According to the Greisen-Zatsepin-Kuzmin [1] (GZK) bound, the Ultra High Energy (UHE) cosmic rays produced in any known candidate extra galactic source should have an exponential cut-off at energies $E \sim 5 \times 10^{10}$ GeV. On the other hand, the number of observed [2] cosmic rays events beyond the cut-off is growing and leads to a mounting paradox within standard frameworks of cosmological and particle physics models.

A wide variety of possible solutions were suggested. Resolution could be due to exotic particle which may be produced at cosmological distances were suitable conventional accelerators are found, be transmitted evading GZK bound, and yet which interact in the atmosphere like a hadron. A particle with correct properties was found in a class of supersymmetric theories [3]. Alternatively, high energy cosmic rays may have been produced locally. One possibility is connected to the events of destruction of (topological) defects [4], while another one to decays of primordial heavy long-living particles [5, 6]. The candidate particle must obviously obey constraints on the mass, density and lifetime.

In order to produce cosmic rays of energies $E \gtrsim 10^{11}$ GeV, the mass of X-particles has to be very large, $m_X \gtrsim 10^{13}$ GeV [5, 6]. The lifetime, τ_X, cannot be much smaller than the age of the Universe, $\tau \gtrsim 10^{10}$ yr. With this smallest value of the lifetime, the observed flux of UHE cosmic rays will be reproduced with rather low density of X-particles, $\Omega_X \sim 10^{-12}$, where $\Omega_X \equiv m_X n_X / \rho_{\rm crit}$, n_X is the number density of X-particles and $\rho_{\rm crit}$ is the critical density. On the other hand, X-particles must not overclose the Universe, $\Omega_X \lesssim 1$. With $\Omega_X \sim 1$, the X-particles may play the role of cold dark matter and the observed flux of UHE cosmic rays can be matched if $\tau_X \sim 10^{22}$ yr. The allowed windows are quite wide [5], but on exotic side, which may rise problems.

The problem of the particle physics mechanism responsible for a long but finite lifetime of very heavy particles can be solved in several ways. For example, otherwise conserved quantum number carried by the X-particle may be broken very weakly due to instanton transitions [5], or quantum gravity (wormhole) effects [6]. If instantons are responsible for X-particle decays, the lifetime is roughly estimated as $\tau_X \sim m_X^{-1} \cdot \exp(4\pi/\alpha_X)$, where α_X is the coupling constant of the relevant (spontaneously broken) gauge symmetry. Lifetime will fit the window if the coupling constant (at the scale m_X) is $\alpha_X \approx 0.1$ [5].

X-particles can be produced in the right amount by usual collision and decay processes if the reheating temperature after inflation never exceeded m_X, but the temperature should be in the range $10^{11} \lesssim T_r \lesssim 10^{15}$ GeV, depending on m_X, [5, 6]. This is a rather high value of reheating temperature and will lead to the gravitino problem in generic supersymmetric models [7].

In the present paper we investigate another process of X-particle creation, namely the direct production from vacuum fluctuations during inflation.

Any viable modern cosmological model invokes the hypothesis of inflation [8]. During inflation the Universe expands exponentially, which solves the horizon and flatness problems of the standard Big-Bang cosmology. Inflation is generally assumed to be driven by the special scalar field ϕ known as the *inflaton*. Fluctuations generated at inflationary stage can have strength and the power spectrum suitable for generation of the large scale structure. This fixes the range of parameters in the inflaton potential.

For example, the mass of the inflaton field has to be $m_\phi \sim 10^{13}$ GeV. During inflation, the inflaton field slowly rolls down towards the minimum of its potential. Inflation ends when the potential energy associated with the inflaton field became smaller than the kinetic energy. At that time all the energy of the Universe is contained entirely in the form of coherent oscillations of the inflaton field around the minimum of its potential. It is possible that a significant fraction of this energy is released to other Boson species after only a dozen or so inflaton field oscillations, in the regime of a broad parametric resonance [9]. This process was studied in details [10, 11]. Even particles with masses of order of magnitude larger than the inflaton mass can be produced quite abundantly. Applying these results to the case of our interest here, we find that stable very heavy particles, $m_X \gtrsim m_\phi$, generally will be produced in excess and will overclose the Universe.

However, if the parametric resonance is ineffective for some reason, and we estimate particle number density after inflation at the level of initial conditions used in Refs. [10], we find that Ω_X might prove to be of about the right magnitude. This level is saturated by the fundamental process of particle creation during inflation from *vacuum fluctuations* and it is the same process which generated primordial large scale density fluctuations. Parametric resonance for X particles is turned-off if X is either a fermion field or its coupling to inflaton is small, $g^2 \ll 10^4 (m_X/m_\phi)^4 (m_\phi/M_{\rm Pl})^2$ [10].

At some early epoch the metric of the Universe is conformally flat to a high accuracy, $ds^2 = a(\eta)^2(d\eta^2 - dx^2)$. We normalize the scale factor by the condition $a(0) = 1$ at the end of inflation. Number density of particles created in time varying cosmological background can be written as

$$n_X = \frac{1}{2\pi^2 a^3} \int |\beta_k|^2 k^2 dk \,, \tag{13}$$

where β_k are the Bogoliubov coefficients which relate "in" and "out" mode functions, and k is the comoving momentum. Massles conformally coupled particles (for scalars this means that $\xi = 1/6$ in the direct coupling to the curvature) are not created. For massive particles conformal invariance is broken. Therefore, for the power low (e.g., matter or radiation dominated) period of expansion of the Universe, one expects on dimensional grounds, $n_X \propto m_X^3/a^3$ at late times. Indeed, it was found in Ref. [12]

$$n_X \approx 5.3 \times 10^{-4} m_X^3 \, (m_X t)^{-3/2} \,, \tag{14}$$

for the radiation dominated Universe, and $n_X \propto m_X^3 \, (m_X t)^{-3q}$ for $a(t) \propto t^q$. Note that all particle creation occur in the region $mt = qm/H \lesssim 1$. When $mt \ll 1$, the number density of created particles remains on the constant level $n_X = m_X^3/24\pi^2$ independent of q [12]. At $qm/H \gg 1$ particle creation is negligible. Here H is the Hubble constant, $H \equiv \dot{a}/a$.

For the radiation dominated Universe one finds, $\Omega_X \sim (m_X^2/M_{\rm Pl}^2)\sqrt{m_X t_e}$, where t_e is time of equal densities of radiation and matter in $\Omega = 1$ Universe. This gives $\Omega_X \sim m_9^{5/2}$, where $m_9 \equiv m_X/10^9$ GeV. Stable particles with $m_X \gtrsim 10^9$ GeV will overclose the Universe even if they were created from the vacuum during regular Friedmann radiation dominated stage of the evolution. It is possible to separate vacuum creation from creation in collisions in plasma since X-particles may be effectively sterile.

This restriction can be overcomed if evolution of the Universe, as it is believed, was more complicated than simple radiation dominated expansion from singularity. Hubble constant may have never exceeded m_X, which is the case of inflation, $H(0) \approx m_\phi$. Moreover, compared to the case considered above, density of X-particles created during inflation is additionally diluted by late entropy release in reheating after inflation.

Particle creation from vacuum fluctuations during inflation (or in de Sitter space) was extensively studied [13, 14]. Characteristic quantity which is usually cited, the variance of the field $\langle X^2 \rangle$, is defined by an expression similar to Eq. (13). In the typical case $\alpha_k \approx -\beta_k$ the difference is given by the factor $2\sin^2(\omega_k \eta)/\omega_k$ in the integrand, where $\omega_k^2 = k^2 + a^2 m_X^2$. If $m_X \sim H(0) \approx m_\phi$, one has on dimensional grounds $n_X = C m_\phi^3/2\pi^2 a^3$ where the coefficient C is expected to be somewhat smaller than unity. Both Fermions and Bosons are produsedby this mechanism, exact numerical value of C being dependent on spin-statistics. In general, C is the function of the ratio $H(0)/m_X$, the function of self-coupling of X and the coupling ξ, depends on details of the transition between inflationary and matter (or radiation) dominated phases, etc. For example, for the scalar Bose-field, $\langle X^2 \rangle = 3H(0)^4/8\pi^2 m_X^2$ if $m_X \ll H(0)$. For massless self-interacting field $\langle X^2 \rangle \approx 0.132 H(0)^2/\sqrt{\lambda}$. C is expected to decrease exponentially when $m_X > m_\phi$.

Let us estimate the today's number density of X-particles. We consider massive inflaton, $V(\phi) = m_\phi^2\phi^2/2$. In this case inflation is followed by the matter domination stage. If there are light Bosons in a theory, $m_B \ll m_\phi$, even relatively weakly coupled to the inflaton, $g^2 \gtrsim 10^4 m_\phi^2/M_{\rm Pl}^2 \sim 10^{-8}$, this matter domination stage will not last long: inflaton will decay via parametric resonance and the radiation domination follows. This happens typically when the energy density in inflaton oscillations is redshifted by a factor $r \approx 10^{-6}$ compared to a value $m_\phi^2 M_{\rm Pl}^2$. Matter is still far from being in the thermal equilibrium, but it is still convenient to characterize this radiation dominated stage by an equivalent temperature, $T_* \sim r^{1/4}\sqrt{m_\phi M_{\rm Pl}}$. At this moment the ratio of energy density in X-particles to the total energy density retains its value reached at the end of inflation, $\rho_X/\rho_R \approx C\, m_\phi m_X/2\pi^2 M_{\rm Pl}^2$. Later on this ratio grows as $\propto T/T_*$ and reaches unity at $T = T_{\rm eq}$, where

$$T_{\rm eq} = \frac{Cr^{1/4}}{2\pi^2}\left(\frac{m_\phi}{M_{\rm Pl}}\right)^{3/2} m_X\,. \tag{15}$$

Using relation $T_{\rm eq} = 5.6\Omega_X h^2$ eV we find that $10^{-12} \lesssim \Omega_X \lesssim 1$ if

$$10^{-23} \lesssim Cr^{1/4} m_X/m_\phi \lesssim 10^{-11}\,. \tag{16}$$

For $m_X \sim$ (a few) $\cdot\, m_\phi$ this condition can be easily satisfied since the coefficient C is exponentially small. This condition may be satisfied even for $m_X \sim m_\phi$ since the coefficient $r^{1/4}$ (or equivalent reheating temperature) can be small too.

Our hypothesis has unique observational consequences. If UHE cosmic rays are indeed due to decay of superheavy particles which were produced from vacuum fluctuations during inflation, there has to be a new sharp cut-off in the cosmic ray spectrum at energy somewhat smaller m_X. Since the number density n_X depends exponentially upon m_X/m_ϕ, the position of this cut-off might be well predicted and has to be near $E_{\rm cut-off} < m_\phi \approx 10^{13}$ GeV, the very shape of the cosmic ray spectrum beyond the GZK cut-off being of quite generic form following from the QCD quark/gluon fragmentation. The Pierre Auger Project installation, the OWL project, the Telescope Array project might prove to be able to discover this fundamental phenomenon.

We conclude, observation of Ultra High Energy cosmic rays can probe the spectrum of elementary particles in its superheavy range and can be an additional opportunity (alongside with fluctuations in cosmic microwave background) to study the earliest epoch of the Universe evolution, starting from amplification of vacuum fluctuations during inflation through fine details of gravitational interaction and down to physics of reheating.

When our paper was at the very end of completion we became aware of the quite recent paper by Chung, Kolb and Riotto [18] where similar problems of superheavy dark matter creation were considered.

We are indebted to V. Berezinsky, G. Farrar, S. Khlebnikov, W. Ochs, G. Pivovarov, S. Sarkar, D. Semikoz, G. Sigl and L. Stodolsky for stimulating discussions. The work of V.K. was supported in part by Russian Foundation for Basic Research grant 95-02-04911a. The work of V.R. was supported in part by Russian Foundation for Basic Research grant 96-02-17449a and U.S. Civilian Research and Development Foundation for Independent States of FSU (CRDF) award RP1-187. V. A. Kuzmin and I. I. Tkachev thank Theory Division at CERN for hospitality where the major part of this work was done. I. I. T. was supported in part by the U.S. Department of Energy under Grant DE-FG02-91ER40681 (Task B) and by the National Science Foundation under Grant PHY-9501458.

References

[1] K. Greisen, Phys. Rev. Lett. **16**, 748 (1966); G. T. Zatsepin and V. A. Kuzmin, Pisma Zh. Eksp. Teor. Fiz. 4, 114 (1966).

[2] N. Hayashida et.al., Phys. Rev. Lett. **73**, 3491 (1994); D. J.Bird et.al., Astroph. J. **424**, 491 (1994); **441**, 144 (1995); T. A.Egorov et.al., in: *Proc. Tokyo Workshop on Techniques for the Study of Extremely High Energy Cosmic Rays*, ed. M.Nagano (ICRR, U. of Tokyo, 1993).

[3] G. R. Farrar, Phys. Rev. Lett. **76**, 4111 (1996); D. J. Chung, G. R. Farrar, and E. W. Kolb, astro-ph/9707036.

[4] C. T. Hill, Nucl. Phys. **B224**, 469 (1983); C. T. Hill, D. N. Schramm and T. P. Walker, Phys. Rev. **D36**, 1007 (1987); G. Sigl, D. N.Schramm and P. Bhattacharjee, Astropart. Phys. **2**, 401 (1994); V. Berezinsky, X. Martin and A. Vilenkin, Phys. Rev. **D56**, 2024 (1997); V. Berezinsky and A. Vilenkin, astro-ph/9704257.

[5] V. A. Kuzmin and V. A. Rubakov, astro-ph/9709187.

[6] V. Berezinsky, M. Kachelriess and A. Vilenkin, Phys. Rev. Lett. **79**, 4302 (1997).

[7] J. Ellis, J. E. Kim, and D. V. Nanopoulos, Phys. Lett. **B145**, 181 (1984).

[8] For a review and list of references, see A. D. Linde, *Particle Physics and Inflationary Cosmology*, (Harwood Academic, New York, 1990); E. W. Kolb and M. S. Turner, *The Early Universe*, (Addison-Wesley, Reading, Ma., 1990).

[9] L. A. Kofman, A. D. Linde and A. A. Starobinsky, Phys. Rev. Lett. **73**, 3195 (1994).

[10] S.Yu. Khlebnikov and I.I. Tkachev, Phys. Rev. Lett. **77**, 219 (1996); Phys. Lett. **B390**, 80 (1997); Phys. Rev. Lett. **79**, 1607 (1997); Phys. Rev. **D56**, 653 (1997).

[11] L. A. Kofman, A. D. Linde and A. A. Starobinsky, Phys. Rev. **D56**, 3258 (1997).

[12] S. G. Mamaev, V. M. Mostepanenko, and A. A. Starobinskii, ZhETF **70**, 1577 (1976) [Sov. Phys. JETP **43**, 823 (1976)].

[13] N. A. Chernikov and E. A. Tagirov, Ann. Inst. Henri Poinare **9A**, 109 (1968); E. A. Tagirov, Ann. Phys. **76**, 561 (1973); T. S. Bunch and P. C. W. Davies, Proc. R. Soc. **A360**, 117 (1978).

[14] A. D. Linde, Phys. Lett. **116B**, 335 (1982); A. A. Starobinsky, Phys. Lett. **117B**, 175 (1982); A. Vilenkin and L. H. Ford, Phys. Rev. **D26**, 1231 (1982); B. Allen, Phys. Rev. **D32**, 3136 (1985).

[15] A. A. Starobinsky and J. Yokoyama, Phys. Rev. **D50**, 6357 (1994).

[16] D. Lyth and D. Roberts, hep-ph/9609441.

[17] M. Boratav, for Pierre Auger Collaboration, *The Pierre Auger Observatory Project: An Overview.*, Proc. of 25th International Cosmic Ray Conference, Durban, v. 5, p. 205 (1997).

[18] D. J. H. Chung, E. W. Kolb and A. Riotto, hep-ph/9802238.

[19] C.T. Hill, D.N. Schramm and T.P. Walker, Phys.Rev. **D36**, 1007 (1987);
G.Sigl, D.N.Schramm and P.Bhattacharjee, Astropart. Phys. **2**, 401 (1994);
V. Berezinsky, X. Martin and A.Vilenkin, Phys.Rev. **D56**, 2024 (1997);
For a review see G.Sigl, *Topological Defect Models of Ultra-High Energy Cosmic Rays*, astro-ph/9611190.

[20] Yu. L. Dokshizer, V.A. Khoze, A.H. Mueller and S.I. Troyan, *Basics of Perturbative QCD* (Editions Frontiers, 1991).

[21] J. Ellis, J.L.Lopez and D.V. Nanopoulos, Phys. Lett. **B247**, 257 (1990);
J. Ellis, G.B. Gelmini, J.L. Lopez, D.V. Nanopoulos and S. Sarkar, Nucl.Phys. **B373**, 399 (1992).

[22] V. Berezinsky, M. Kachelriess, and A. Vilenkin, *Ultra-High Energy Cosmic Rays without GZK Cutoff*, 1997, unpublished.

[23] R. Protheroe and T. Stanev, Phys.Rev.Lett. **77**, 3708 (1996);
G. Sigl, S. Lee and P.Coppi, *Highest Energy Cosmic Rays, Grand Unified Theories and the Diffuse Gamma-Ray Background*, astro-ph/9604093 and references therein.

Magnetic Effects on the Moment of Inertia of a Relativistic Star

Kohkichi Konno[1]

Department of Physics, Faculty of Science,
Hiroshima University, Higashi-Hiroshima, 739-8526, Japan

Abstract

We consider principal moments of inertia of axisymmetric, magnetically deformed stars within general relativity. The general expression for the moment of inertia with respect to the symmetric axis is obtained. The numerical estimates are derived for several polytropic stellar models. We find that the values of the principal moments of inertia are modified by a factor 2 at most from Newtonian estimates.

1 Introduction

Various features of pulsars have been disclosed with the growing observational data. Most pulsars have stable pulse shapes, and spin down steadily with the typical time-scales of order several Myr (see, e.g., a standard text[1]). However, deviations from the linear spin-down trends are also observed on shorter time-scales in some pulsars. These deviations convey information about the surroundings, internal structure and dynamics of neutron stars. For example, the presence of glitches (see, e.g., Refs. [2, 3]) indicates the internal structure that two or more poorly coupled components coexist in the stars, while timing noise (see, e.g., Ref. [4]) is believed to be due to random movements of the fluid in the pulsars. Besides these short time-scale instabilities, spin-down features indicating the precession due to stellar deformation are also reported. Two anomalous X-ray pulsars (AXPs), 1E1048.1-5937 and 1E2259+586, are observed to indicate the irregular spin-down (see, e.g., Refs. [5, 6, 7, 8]). AXPs are a subclass of X-ray pulsars with period in a very narrow range (6-12s) and period derivatives in a range ($10^{-12} - 10^{-11}$ss^{-1}) (see, e.g., Ref. [9]), and those are most likely candidates of magnetars[10, 11, 12, 13] along with soft gamma repeaters (SGRs) (see, e.g., Refs. [14, 15]). As discussed by Melatos[16], the wobbles in their spin-down rates may be interpreted by an effect called radiative precession[17], which is related with an oscillating component of the electromagnetic torque. Another precessing pulsar PSR B1828-11 was reported very recently[18]. This gives the very clear observation of free precession of an isolated neutron star. These deformed precessing objects are analyzed by solving Euler equations of motion in the form $I_{ij}d\Omega^j/dt - \varepsilon_{ijk}I^{jl}\Omega^k\Omega_l = N_i$, where I_{ij} is the inertia tensor, Ω^i is the angular velocity and N_i is the torque acting on the object. When we take into account the electromagnetic torque by a rotating magnetic dipole[19, 20], the above-mentioned radiative precession can be found[16, 17] for magnetically deformed stars. Furthermore, since magnetically deformed rotating stars emit gravitational waves, we can consider the gravitational radiation reaction torque[21, 22]. The gravitational backreaction damps the wobble on a time-scale proportional to $[(I_z - I_x)^2/I_x]^{-1}$ for wobbling, axisymmetic rigid bodies, where I_i is the principal moments of inertia. Thus, the moments of inertia play a significant role in the analyses of pulsar precession. The estimates of the moments of inertia have been done within Newtonian gravity so far. However, neutron stars are fully general relativistic objects, and it is important to take into account the general relativistic effects. Therefore, we discuss the principal moments of inertia of magnetically deformed stars within general relativity.

In Newtonian gravity, once we have the the mass distribution of an object, $\rho(\vec{x})$, the moment of inertia with respect to any axis is calculated from $I = \int \rho(\vec{x})\chi^2 d^3x$, where χ denotes the length from the axis. In general relativity, only for axisymmetic objects, the moment of inertia with respect to the symmetric axis can be well defined. For this definition, we need slow rotation of the objects. The slow rotation ensures that the angular momentum J is linearly related to the angular velocity Ω, i.e., $J = I\Omega$. Here, I defines the general relativistic version of the moment of inertia. The moments of inertia of relativistic, spherically symmetric stars were discussed by Hartle[23] and Chandrasekhar & Millar[24].

[1] E-mail:konno@theo.phys.hiroshima-u.ac.jp

The approximate expression valid for various realistic equations of state was also derived by Ravenhall & Pethick[25].

In this paper, we discuss principal moments of inertia of magnetically deformed stars which are axisymmetric with respect to the magnetic axis. Deformation of relativistic magnetized stars was studied both in the numerical approach[26, 27] and in the analytic approach[28, 29]. Based on the latter, analytic approach developed by Konno et al.[28], we proceed with the discussions. In the previous work[28], the stellar magnetic field was regarded as corrections to non-rotating, spherical stars, and the magnetic deformation of stars was formulated using a perturbative approach. As mentioned above, we now take into account the slow rotation of the deformed stars on the symmetric axis to define the moment of inertia. The formulation of this configuration are given in §2. The general expression for the principal moment of inertia with respect to the symmetric axis is also derived. The numerical estimates are derived for several stellar models in §3. In §4, we discuss the other principal moments of inertia. Finally, we give conclusion in §5. Throughout this paper, we use the units in which $c = G = 1$.

2 Magnetically deformed rotating stars

We consider a slowly rotating star which is subject to quadrupole deformation due to the dipole magnetic field as described by Konno et al.[28], in which the magnetic field is treated as corrections to the background, spherical star, i.e., $B \sim O(\varepsilon_B)$. We now assume that the star rotates on the magnetic axis with a uniform angular velocity $\Omega \sim O(\varepsilon_\Omega)$, which is also regarded as a perturbation. In this paper, we take into account the rotational corrections up to first order in ε_Ω. The metric describing such a star is given by

$$\begin{aligned} ds^2 &= -e^{\nu(r)}\left[1 + 2\left(h_0(r) + h_2(r)P_2(\cos\theta)\right)\right]dt^2 \\ &\quad + e^{\lambda(r)}\left[1 + \frac{2e^{\lambda(r)}}{r}\left(m_0(r) + m_2(r)P_2(\cos\theta)\right)\right]dr^2 + r^2\left(1 + 2k_2(r)P_2\right)d\theta^2 \\ &\quad + r^2\sin^2\theta\left(1 + 2k_2(r)P_2(\cos\theta)\right)\left[d\phi - \left(W_1(r) - \frac{W_3(r)}{\sin\theta}\frac{dP_3(\cos\theta)}{d\theta}\right)dt\right]^2, \end{aligned} \tag{1}$$

where P_l is the Legendre's polynomial of order l, and the corrections h_0, h_2, m_0, m_2 and k_2 of second order in ε_B denote the deviation from the spherical shape. If we replace W_1 and W_3 with zero in Eq. (1), this expression reduces to the metric already investigated by Konno et al.[28]. The newly appeared functions W_1 and W_3 lead to the frame dragging due to the deformed rotating star. These functions include the corrections up to order $\varepsilon_\Omega\varepsilon_B^2$ in order to involve the effect of the deformation. The form of the terms including these functions in Eq. (1) corresponds to quadrupole deformation of the star. We can write down W_1 and W_3, showing the order explicitly, in the forms

$$W_1 = \omega + W_1^{(2)}, \quad W_3 = W_3^{(2)}, \tag{2}$$

where $\omega \sim O(\varepsilon_\Omega)$ and $(W_1^{(2)}, W_3^{(2)}) \sim O(\varepsilon_\Omega\varepsilon_B^2)$. These functions are very analogous to the rotational corrections up to third order in ε_Ω [30, 31].

The stress-energy tensor of the star which is composed of a perfect fluid and endowed with a dipole magnetic field has the form

$$T^\mu_{\ \nu} = (\rho + p)\,u^\mu u_\nu + p\delta^\mu_{\ \nu} + \frac{1}{4\pi}\left(F^{\mu\lambda}F_{\nu\lambda} - \frac{1}{4}F^{\sigma\lambda}F_{\sigma\lambda}\delta^\mu_{\ \nu}\right), \tag{3}$$

where the density ρ and pressure p are expanded as

$$\rho = \rho_0(r) + \left(\rho_{20}(r) + \rho_{22}(r)P_2\right), \tag{4}$$

$$p = p_0(r) + \left(p_{20}(r) + p_{22}(r)P_2\right). \tag{5}$$

The non-vanishing components of the 4-velocity u^μ are given by

$$u^\phi = \Omega u^t, \quad u^t = e^{-\frac{\nu}{2}}\left[1 - \left(h_0 + h_2P_2\right)\right]. \tag{6}$$

There is a difference from the non-rotating case also in the Faraday tensor $F_{\mu\nu}$, which is given by the 4-potential A_μ. In the current case, the 4-potential A_μ is written in the form[32, 33]

$$A_\mu = \left(a_{t0}(r) + a_{t2}(r)P_2, 0, 0, a_\phi(r)\sin\theta\frac{dP_1}{d\theta}\right), \tag{7}$$

where the t-component arises due to the electric field induced by the stellar rotation, and therefore a_{t0} and a_{t2} are of order $\varepsilon_\Omega\varepsilon_B$. The assumption of the perfectly conducting interior yields the following expressions[26, 27] inside the star

$$a_{t0} = \frac{2}{3}\Omega a_\phi + c, \quad a_{t2} = -\frac{2}{3}\Omega a_\phi, \tag{8}$$

where c is a constant fixed by the junction condition at the surface. The exterior solution for a_{t0} and a_{t2} can also be obtained in the analytic form[32, 33] by solving the Maxwell equations.

The differential equations which W_1 and W_3 obey can be obtained from the $(t\phi)$-component of the Einstein equations,

$$\frac{1}{r^4}\frac{d}{dr}\left[r^4 j\frac{d\overline{W}_1}{dr}\right] + \frac{4}{r}\frac{dj}{dr}\overline{W}_1 = -j\left(S_1 - \frac{S_2}{5}\right), \tag{9}$$

$$\frac{1}{r^4}\frac{d}{dr}\left[r^4 j\frac{dW_3}{dr}\right] + \left(\frac{4}{r}\frac{dj}{dr} - \frac{10}{r^2}e^{\frac{\lambda-\nu}{2}}\right)W_3 = j\frac{S_2}{5}, \tag{10}$$

where j and $\overline{W}_1$ are defined, respectively, as

$$j = e^{-\frac{\nu+\lambda}{2}}, \tag{11}$$

$$\overline{W}_1 = \Omega - W_1 = \varpi - W_1^{(2)} \quad (\varpi = \Omega - \omega). \tag{12}$$

The functions S_1 and S_2 appearing in the source terms in Eqs. (9) and (10), in general, include W_1 and W_3. However, if we consider S_1 and S_2 up to the order of our interest, these functions include only Ω and ω as will be shown in Eqs. (13) and (14). These differential equations are solved order by order. Firstly, the differential equation of order ε_Ω, in which both S_1 and S_2 vanish, is solved. Next, the differential equations of order $\varepsilon_\Omega\varepsilon_B^2$ are solved, in which case the source terms are given by

$$\begin{aligned} S_1 &= \varpi'\left(-h_0 - \frac{e^\lambda}{r}m_0\right)' - \frac{4}{r^2}e^\lambda(\nu'+\lambda')\varpi m_0 - 16\pi e^\lambda\varpi(\rho_{20}+p_{20}) \\ &\quad + \frac{16e^\lambda}{3r^4}a_\phi^2\omega + \frac{8}{3r^2}(a_\phi')^2\omega + \frac{8e^\lambda}{r^4}a_\phi a_{t2} - \frac{4}{r^2}a_\phi' a_{t0}', \end{aligned} \tag{13}$$

$$\begin{aligned} S_2 &= \varpi'\left(4k_2 - h_2 - \frac{e^\lambda}{r}m_2\right)' - \frac{4}{r^2}e^\lambda(\nu'+\lambda')\varpi m_2 - 16\pi e^\lambda\varpi(\rho_{22}+p_{22}) \\ &\quad + \frac{32e^\lambda}{3r^4}a_\phi^2\omega - \frac{8}{3r^2}(a_\phi')^2\omega + \frac{16e^\lambda}{r^4}a_\phi a_{t2} - \frac{4}{r^2}a_\phi' a_{t2}', \end{aligned} \tag{14}$$

where the prime denotes the differentiation with respect to r. These functions can be calculated by using the results for the density and pressure $(\rho_{20}, \rho_{22}, p_{20}, p_{22})$, the metric functions $(h_0, h_2, m_0, m_2, k_2)$ and the potential functions (a_ϕ, a_{t0}, a_{t2}), which were already calculated in Refs. [28, 33].

The differential equations (9) and (10) should be solved numerically by imposing boundary and junction conditions. In these conditions, W_1 and W_3 should be regular at the center of the star, vanish at infinity, and connected smoothly at the surface. Before deriving the numerical solution, it is worthwhile to investigate the behavior of W_1 and W_3 at large r in detail using Eqs. (9), (10), (13) and (14), because this inspection will give the expressions for the angular momentum and moment of inertia of the star.

At large r, we have

$$\lambda, \nu \to 0, \quad \text{i.e.,} \quad j \to 1. \tag{15}$$

From Eqs. (13) and (14), we derive the asymptotic behaviors

$$S_1 \propto \frac{1}{r^8}, \quad S_2 \propto \frac{1}{r^7}. \tag{16}$$

Using these expressions, from Eqs. (9) and (10), we obtain

$$W_1^{(2)} \propto \frac{1}{r^3}, \quad W_3^{(2)} \propto \frac{1}{r^5}, \tag{17}$$

since each function vanish at infinity. Hence, we can put

$$\overline{W}_1 = \Omega - \frac{2J}{r^3} - W_{1p}^{(2)}(r), \tag{18}$$

where J is the angular momentum of the deformed star, and $W_{1p}^{(2)} \sim O(1/r^6)$ is of order $\varepsilon_\Omega \varepsilon_B^2$. We now integrate Eq. (9) after multiplying both sides by r^4. Using the above expression (18), we can obtain the general expression for the principal moment of inertia

$$\begin{aligned} I_z &= \frac{J}{\Omega} \\ &= -\frac{2}{3}\int_0^R r^3 \frac{dj}{dr}\frac{\varpi}{\Omega}dr - \frac{R^4}{6\Omega}\left(\frac{dW_{1p}^{(2)}}{dr}\right)_R + \frac{2}{3}\int_0^R r^3 \frac{dj}{dr}\frac{W_1^{(2)}}{\Omega}dr - \frac{1}{6\Omega}\int_0^R r^4 j\left(S_1 - \frac{S_2}{5}\right)dr, \end{aligned} \tag{19}$$

where R is the radius of the background star. The moment of inertia of the background, spherical star, $I^{(0)}$, is given[23, 25] by

$$I^{(0)} = -\frac{2}{3}\int_0^R r^3 \frac{dj}{dr}\frac{\varpi}{\Omega}dr. \tag{20}$$

Therefore, we derive

$$I_z^{(2)} = -\frac{R^4}{6\Omega}\left(\frac{dW_{1p}^{(2)}}{dr}\right)_R + \frac{2}{3}\int_0^R r^3 \frac{dj}{dr}\frac{W_1^{(2)}}{\Omega}dr - \frac{1}{6\Omega}\int_0^R r^4 j\left(S_1 - \frac{S_2}{5}\right)dr, \tag{21}$$

where we have used the decomposition $I_z = I^{(0)} + I_z^{(2)}$. The magnetic modification of the principal moment of inertia, $I_z^{(2)}$, is the second order quantity in ε_B. The numerical estimates of $I_z^{(2)}$ for several stellar models will be given in the next section.

3 Numerical estimates of the moment of inertia

Using the boundary and junction conditions mentioned in the last section, we can derive the numerical solution for W_1 and W_3 for any stellar models. The numerical results for the magnetic correction of the principal moment of inertia, $I_z^{(2)}$, are also obtained by using Eq. (21). We now show the results for $I_z^{(2)}$ obtained for several polytropic stellar models. In these calculations, we adopted a different condition from that of the previous study[28], in which sequences with constant central density were investigated. In the current case, in order to clarify the correspondence between non-magnetized and magnetized stars having same mass, we use a condition that the total mass of the star does not change through the perturbative approach. This is also because the moment of inertia can be modified significantly by the mass shift rather than the magnetic deformation. In our formulation, this condition is accomplished by imposing a boundary condition that m_0 should vanish at infinity.

Fig. 1 displays the magnetic correction $I_z^{(2)}$ as a function of the general-relativistic factor M/R. The values are normalized by the typical value μ^2/M. We now discuss quadrupole deformation of a fluid body in the Newtonian limit as a simple example. If the deformed body has constant density, which corresponds to $n = 0$, then we can derive a result

$$I_z^{(2)} = \frac{10}{3}\frac{\mu^2}{M} \tag{22}$$

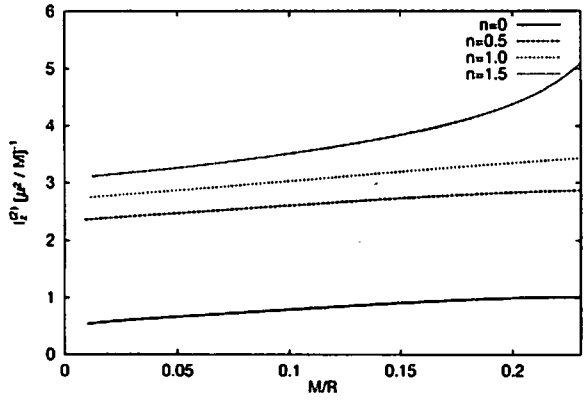


Figure 1: The magnetic correction of the principal moment of inertia, $I_z^{(2)}$, which is plotted against M/R. The values are normalized by μ^2/M. The polytropic index is denoted by n.

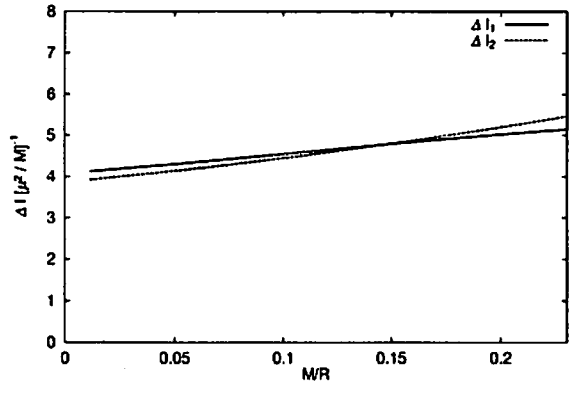


Figure 2: Comparison between ΔI_1 and ΔI_2 for $n = 1$. ΔI_1 and ΔI_2 are normalized by μ^2/M and plotted against M/R.

from the estimate of the ellipticity $25\mu^2/(2M^2R^2)$[34, 29]. As seen in Fig. 1, our numerical result for $n = 0$ is different from this simple estimate by a factor 6. This is because the perturbed star in our formulation does not have constant density by the added perturbation, even though we assume a background star with constant density. Owing to this reason, our numerical estimate of the magnetic correction $I_z^{(2)}$ take the different value from that of Eq. (22). Therefore, our result shown in Fig. 1 cannot simply be compared with the above estimate.

Concentrating on the general relativistic effects on the magnetic correction $I_z^{(2)}$, we can find from Fig. 1 that the value of $I_z^{(2)}$ for each stellar model increases with the general relativistic factor M/R. The increments are given by a factor 2 at most. However, the rates of increase may be neglected except for $n = 1.5$.

4 The other components of moments of inertia

Next, we discuss the other principal moments of inertia, i.e., $I_x^{(2)}$ and $I_y^{(2)}$. For this purpose, let us recall relations using the principal moments of inertia and the ellipticity which hold for an incompressible fluid within Newtonian gravity. This will be helpful in seeking the other principal moments of inertia. Firstly,

if the fluid body is subject to quadrupole deformation, then a relation between the principal moments of inertia exists

$$I_x^{(2)} = I_y^{(2)} = -\frac{1}{2} I_z^{(2)}. \tag{23}$$

Using this relation, the difference between the principal moments of inertia $I_z^{(2)}$ and $I_x^{(2)}$ is given by

$$\Delta I_1 = I_z^{(2)} - I_x^{(2)} = \frac{3}{2} I_z^{(2)}. \tag{24}$$

In the different way, this difference can be expressed also by using the ellipticity, which is defined by the difference between the equatorial radius and polar radius[24, 28, 29], in the form

$$\Delta I_2 = I^{(0)} \times (\text{ellipticity}). \tag{25}$$

Thus, the difference is derived by the two methods. There is no reason that the two expressions coincide, but empirically the agreement is good. Note that these expressions (24) and (25) coincide exactly in the current case.

If we assume that the relation (23) is applicable to general relativistic cases, it is seems that we should simply apply the relation (23) to the results in Fig. 1, and derive $I_x^{(2)}$ and $I_y^{(2)}$. However, in order to utilize this relation, the magnetic corrections must be purely quadrupole contributions. As seen in the metric (1), the magnetic deformation include monopole parts as well as quadrupole parts. Thereby, our results in Fig. 1 also include monopole parts. Hence, we have to subtract the monopole contributions from $I_z^{(2)}$, and derive $I_x^{(2)}$ and $I_y^{(2)}$. It is, nevertheless, not so easy to extract the monopole contributions accurately from those magnetically deformed stars.

However, there is no guarantee that all the results in Fig. 1 include the monopole contributions dominantly. If there is a case in which the monopole part can be neglected, then we can estimate $I_x^{(2)}$ and $I_y^{(2)}$ for its case by assuming that the relation (23) holds approximately. That case would also provide some extrapolation for the general relativistic effects on $I_x^{(2)}$ and $I_y^{(2)}$ in the other cases. In order to seek such a case, we now compare ΔI_1 derived by simply using the results in Fig. 1 with ΔI_2 calculated from the ellipticity, which was already estimated for several stellar models[29]. In a case that ΔI_1 is almost consistent with ΔI_2 about the values and tendency, it seems that the monopole part can be neglected, and we may use the relation (23). From the comparison, we find that ΔI_1 is almost consistent with ΔI_2 for $n = 1$. Fig. 2 displays the comparison between ΔI_1 and ΔI_2 for this stellar model. The two curves coincide within 10%. Hence, we can estimate $I_x^{(2)}$ and $I_y^{(2)}$ using Eq. (23) for this stellar model. Since the values of $I_x^{(2)}$ and $I_y^{(2)}$ are simply given by multiplying $I_z^{(2)}$ by a factor $(-1/2)$, the change of the absolute values of $I_x^{(2)}$ and $I_y^{(2)}$ due to the general relativistic effects is estimated to be specified by the same factor as in the case of $I_z^{(2)}$ (see Fig. 3). Consequently, we derive very similar result to that of $I_z^{(2)}$. However, $I_x^{(2)}$ and $I_y^{(2)}$ decrease with the general relativistic factor M/R, while $I_z^{(2)}$ increases. We expect that similar features exist also in the other stellar models.

5 Conclusions

Recent observations indicates precession of some pulsars. In order to understand the dynamics of these pulsars in detail, we have to know the moments of inertia, which play a crucial role in the Euler equations of motion. Motivated by this consideration, we have discussed the general relativistic effect on the moments of inertia of magnetically deformed stars. Developing the formulation by Konno et al.[28], we have considered the slow rotation of magnetically deformed stars on the symmetric axis in order to define the moment of inertia. The general expression for the magnetic correction, $I_z^{(2)}$, to the moment of inertia of spherically symmetric stars was obtained, and the numerical estimates for various polytropic stellar models were also obtained. The other components of the moment of inertia, $I_x^{(2)}$ and $I_y^{(2)}$, were discussed from the extension of relations in Newtonian gravity. From those results, we found that each principal moments of inertia is modified by a factor 2 at most owing to the general relativistic effect. Nevertheless, it seems that the general relativistic effect does not affect precessing motion noticeably in the case of free

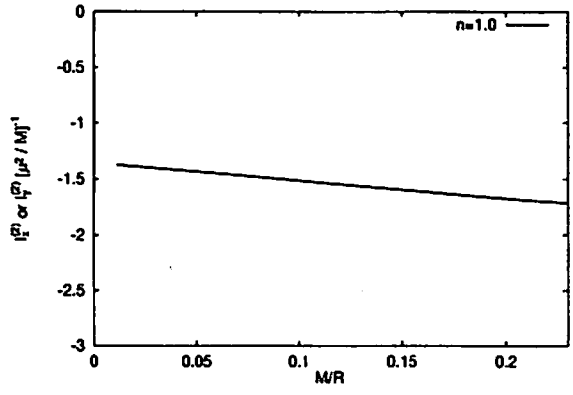


Figure 3: The other components of the principal moments of inertia, $I_x^{(2)}$ and $I_y^{(2)}$, which are derived for $n = 1$.

precession of pulsars. However, it is not clear whether or not the electromagnetic radiation reaction torque acting on the pulsars change in general relativity. If this torque is modified by the general relativistic effect, then the period of the wobbles would also be modified. Therefore, this point should be made clear in future investigation.

Acknowledgements

I would like to thank Y. Kojima for careful reading of the manuscript and for fruitful discussions and suggestions. I would also like to thank M. Kasai, H. Asada and K. Ioka for valuable and useful comments, and M. Hosonuma for useful discussions. This work was supported in part by a Grant-in-Aid for Scientific Research Fellowship of the Ministry of Education, Science, Sports and Culture of Japan (No.12001146).

References

[1] P. Mészáros, *High-Energy Radiation from Magnetized Neutron Stars* (The University of Chicago Press, USA, 1992).

[2] S.L. Shemar and A.G. Lyne, MNRAS **282**(1996) 677.

[3] N. Wang et al., MNRAS **317**(2000) 843.

[4] F. D'Alessandro, P.M. McCulloch, P.A. Hamilton and A.A. Deshpande, MNRAS **277**(1995) 1033.

[5] S. Mereghetti, ApJ **455**(1995) 598.

[6] A. Baykal and J. Swank, ApJ **460**(1996) 470.

[7] A. Baykal, J.H. Swank, T. Strohmayer and M.J. Stark, A&A **336**(1998) 173.

[8] T. Oosterbroek, A.N. Parmar, S. Mereghetti and G.L. Israel, A&A **334**(1998) 925.

[9] S. Mereghetti and L. Stela, ApJ **442**(1995) L17.

[10] R.C. Duncan and C. Thompson, ApJ **392**(1992) L9.

[11] C. Thompson and R.C. Duncan, ApJ **408**(1993) 194.

[12] C. Thompson and R.C. Duncan, MNRAS **275**(1995) 255.

[13] C. Thompson and R.C. Duncan, ApJ **473**(1996) 322.

[14] C. Kouveliotou et al., Nature **393**(1998) 235.

[15] C. Kouveliotou et al., ApJ **510**(1999) L115.

[16] A. Melatos, ApJ **519**(1999) L77.

[17] A. Melatos, MNRAS **313**(2000) 217.

[18] I.H. Stairs, A.G. Lyne and S.L. Shemar, 2000, Nature **406**(2000) 484.

[19] L. Davis and M. Goldstein, ApJ **159**(1970) L81.

[20] P. Goldreich, ApJ **160**(1970) L11.

[21] B. Bertotti and A.M. Anile, A&A **28**(1973) 429.

[22] C. Cutler and D.I. Jones, gr-qc/0008021

[23] J.B. Hartle, ApJ **150**(1967) 1005.

[24] S. Chandrasekhar and J.C. Miller, MNRAS **167**(1974) 63.

[25] D.G. Ravenhall and C.J. Pethick, ApJ **424**(1994) 846.

[26] S. Bonazzola, E. Gourgoulhon, M Salgado and J.A. Marck J.A., A&A **278**(1993) 421.

[27] M. Bocquet, S. Bonazzola, E. Gourgoulhon and J. Novak J., A&A **301**(1995) 757.

[28] K. Konno, T. Obata and Y. Kojima, A&A **352**(1999) 211.

[29] K. Konno, T. Obata and Y. Kojima, A&A **356**(2000) 234.

[30] H. Quintana, ApJ **207**(1976) 279.

[31] Y. Kojima and M. Hosonuma, ApJ **520**(1999) 788.

[32] A.G. Muslimov and A.I. Tsygan, SvA **30**(1986) 567.

[33] K. Konno and Y. Kojima, Prog. Theor. Phys. **104**(2000) in press.

[34] V.C.A. Ferraro, ApJ **119**(1954) 407.

Fireball generation by MHD jets

Masaaki Takahashi [1]

Department of Physics and Astronomy, Aichi University of Education, Kariya, Aichi 448-8542, Japan

Abstract

We discuss a model for the relativistic outflows associated with the gamma-ray burst remnants. Though the energy of the gamma-ray burst events is to be released as the outflow with high Lorentz factor, its origin at the central engine and its transport mechanism are unclear. When magnetic field lines thread a rotating black hole and an accreting plasma-torus surrounding the hole, and connect with the remnant as a remote astrophysical loads, the rotational energy of the black hole and torus are extracted by the outgoing plasma flow. Under the existence of the strong gravity, highly relativistic velocity can be obtained for tenuous plasma flows. Then, the initial magnetically dominated outflow can convert to the kinetic-energy dominated flow; the plasma's kinetic energy can be reconverted into gamma-ray by relativistic shocks.

1 Introduction

Sudden collapse of a compact object, or coalescence of a compact binary, can generate an relativistic wind/jet. Collapsing or coalescing compact objects may generate strong magnetic fields as high as $B_{\rm ini} \sim 10^{15}$ G, which is expected in young pulsars [2] or black hole–torus (disk) systems. In fact, at least three magnetars, neutron stars with fields $B \sim 10^{14}$–10^{15} G, have been observed (Their visible lifetime is only a few thousand years, because neutron stars slow down by emission of magnetic dipole radiation and then they no longer emit pulses).

In a black hole–torus system, though the black hole by itself cannot keep magnetic fields on it for a long time, the accreting torus (surrounding baryonic matters) can generate magnetic field lines threading the event horizon. The magnetic pressure of magnetic field on the black hole should be comparable to that of magnetic field on the inner part of the accretion disk [3]. The magnetic field on the disk should not be larger than the value from the equipartition argument $B_{\rm eqp}^2/8\pi \sim P_{\rm disk}$, where $B_{\rm eqp}$ is an upper limit to the magnetic field which can be supported by the accretion disk. Recent numerical calculations on the hyper-accreting black hole by Popham et al. [4] show $B_{\rm eqp} \sim 5 \times 10^{15}$ G.

The outflow from such magnetized objects is likely to carry a large magnetic energy. The magnetic energy would be transformed into kinetic energy of bulk relativistic outflow, and then the energy can be converted, with adequate efficiency, into gamma rays with a nonthermal spectrum by a relativistic shock formation. This process can be applied to cosmological gamma-ray burst (GRB) models as mentioned by Rees and Mészáros [1] (see Fig. 1). The strong disk field $B_{\rm eqp} \sim 5 \times 10^{15}$ G implies the accretion disk of the hyper-accreting black hole may support a magnetic field strong enough for GRBs, and the life time of the accreting torus is 0.1 – 150 s, which are not inconsistent with the GRB time scale; that is, during this timescale the strong magnetic field can be supported around the black hole.

It has been discussed how kinetic energy can be reconverted into gamma rays by relativistic shocks which form when the ejecta run into external matter (either ambient interstellar matter, or a non-relativistic outflow preceding the event itself). The acceptable models require bulk Lorentz factors Γ of order 1000 with the entrained baryonic mass below $10^{-4} M_\odot$. However, how the plasma is ejected from the central compact region (the size of the central engine $\sim 10^7$ cm) and how the flow is accelerated on the way to the shock front at radii in the range 10^{14} to 10^{16} cm are still uncertain.

The GRBs are clearly the signal of an energetic event lasting (typically) no more than a few seconds. The characteristic duration of bursts, which can extend to tens (or even hundreds) of seconds, is of course very long in comparison with the dynamical or orbital time scale for the instantaneous triggers,

[1] E-mail:takahasi@phyas.aichi-edu.ac.jp

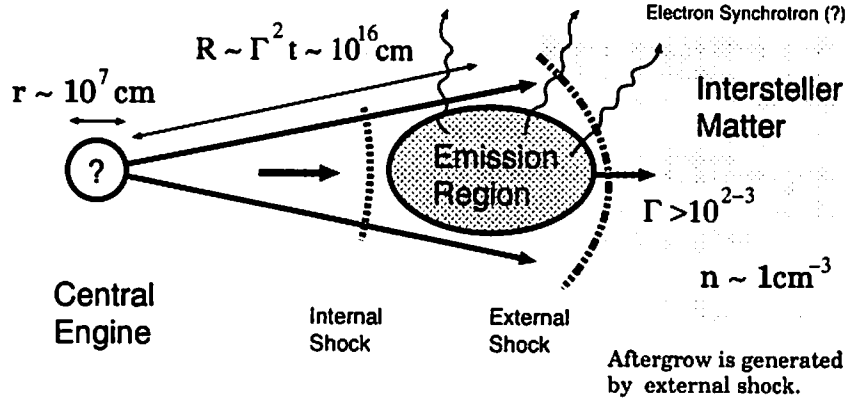


Figure 1: A schematic picture of gamma-ray burst remnant

which is measured in milliseconds. The burst event is approximated as an instantaneous fireball, or as an outflowing wind which is *steady* over the entire burst duration. In the former type of GRB models, the initial energy input is impulsive and the relativistic ejecta begins to decelerate when they have swept up external matter, leading to an external and a reverse shock that radiate a large fraction of the initial total energy (see, [1, 5, 6, 7]). The behavior resembles that of as ordinary supernova remnant expanding into as an undisturbed medium [8]. On the other hand, in the latter type of GRB models, the initial energy input continues and the resulting relativistic fireball wind produces a GRB that is due to internal shock in the wind itself. This situation is very similar to the relation between a pulsar wind and its nebula. There is a general consensus that the longer complex bursts must involve internal shocks, though simple shape pulses could arise from an external shock interaction (the latter would in effect be the precursor of the afterglow).

In this paper, we shall focus on the acceleration mechanism for the magnetized-wind internal shock models. We are proposing as a central engine of the GRBs is a black hole–torus (disk) systems. The rotational energy of the black hole and torus, which can be sufficient to power the most energetic GRBs with $\sim 10^{53}$ erg [9], would be extracted through the strong Poynting flux with the tenuous plasma flow [10, 11]. In sections 2 and 3, We will outline relativistic trans-fast magnetohydrodynamical (MHD) flows. The acceleration of the wind from a black hole magnetosphere is discussed in Section 4.

2 Basic Equations of Relativistic MHD Flow

In this section, we summarize the relativistic magnetohydrodynamical (MHD) flows. We assume a stationary and axisymmetric magnetosphere and ignore its self-gravity. We also require infinite conductivity for the plasma flow. The background metric is written by the Boyer-Lindquist coordinates with geometrical units ($c = G = 1$).

The basic equations of relativistic magnetized plasma flow are as follows:

(i) the particle number conservation

$$(nu^{\alpha})_{;\alpha} = 0 \; ; \tag{1}$$

(ii) the conservation of total energy and momentum

$$T^{\alpha\beta}_{\;\;;\beta} = 0 \, , \tag{2}$$

where the energy-momentum tensor is given by

$$T^{\alpha\beta} = n\mu u^\alpha u^\beta - Pg^{\alpha\beta} + \frac{1}{4\pi}\left(F^\alpha_\lambda F^{\lambda\beta} + \frac{1}{4}g^{\alpha\beta}F^2\right) , \tag{3}$$

n is the proper particle number density, μ is the relativistic specific enthalpy, P is the pressure and $F^2 = F^{\mu\nu}F_{\mu\nu}$;
(iii) MHD condition

$$u^\beta F_{\alpha\beta} = 0 . \tag{4}$$

The magnetic field and electric field seen by a distant observer are defined as

$$B_\alpha \equiv \frac{1}{2}\eta_{\alpha\beta\gamma\delta}k^\beta F^{\gamma\delta} \tag{5}$$

$$E_\alpha \equiv F_{\alpha\beta}k^\beta \tag{6}$$

where $k^\alpha = (1,0,0,0)$ is the time-like killing vector and $\eta_{\alpha\beta\gamma\delta} \equiv \sqrt{-g}\,\epsilon_{\alpha\beta\gamma\delta}$. Further, we introduce the angular velocity of the magnetosphere (see, [12])

$$\Omega_F \equiv -\frac{F_{tr}}{F_{\phi r}} = -\frac{F_{t\theta}}{F_{\phi\theta}} . \tag{7}$$

The energy and angular momentum of the fluid are

$$\mu u_t = \frac{M^2 E - eG_t}{M^2 - \alpha} , \tag{8}$$

$$\mu u_\phi = -\frac{M^2 L + eG_\phi}{M^2 - \alpha} , \tag{9}$$

where $G_t \equiv (g_{tt} + g_{t\phi}\Omega_F)$, $G_t \equiv (g_{t\phi} + g_{\phi\phi}\Omega_F)$ and $\alpha \equiv G_t + G_\phi\Omega_F$. $M^2 \equiv \alpha(u_p/u_{\rm AW})^2 = 4\pi\eta^2(\mu/n)$ is the Alfvén Mach number, where u_p is the poloidal velocity defined by $u_p^2 \equiv -u^A u_A$ $(A = r,\theta)$ and $u_{\rm AW}$ is the Alfvén wave speed. E and L are total (fluid plus magnetic) energy and angular momentum seen by a distant observer, and $e \equiv (E - \Omega_F L)$ is the total energy seen by a corotating observer with the magnetic field line. The particle number flux per magnetic flux tube η can be defined as

$$\eta = -\frac{nu^r}{B^r}G_t . \tag{10}$$

The poloidal equation of the magnetized flow (the relativistic Bernoulli equation) is written as [13, 14, 11]

$$(1 + u_p^2) = (E/\mu)^2[(\alpha - 2M^2)f^2 - k] , \tag{11}$$

where

$$f \equiv -\frac{G_\phi + G_t\tilde{L}}{\rho_w(M^2 - \alpha)} , \tag{12}$$

$$k \equiv -(g^{tt} - 2g^{t\phi}\tilde{L} + g^{\phi\phi}\tilde{L}^2) , \tag{13}$$

and $\tilde{L} \equiv L/E$. When we neglect dissipative effects and cooling in the plasma, the flow is adiabatic with a specific enthalpy of the form [15]

$$\mu = m_{\rm p} + \frac{\gamma}{\gamma - 1}\frac{P}{n} , \tag{14}$$

where we use an equation of state which is assumed to be a polytropic one, i.e. $P = K\rho^\gamma$, ρ is the rest mass density, and K is a measurement of the specific entropy of the flow, which is a constant in a shock-free flow. The toroidal component of the magnetic field is can be expressed to

$$B_\phi = -4\pi\eta E\rho_w f . \tag{15}$$

When $u_p = u_{\rm AW}$, the denominator in the function f vanishes. The physical plasma flow must satisfy the condition $\tilde{L} = (G_\phi/G_t)_{\rm A}$ at there. This point $r = r_{\rm A}$ is called the Alfvén point. Because the Alfvén points must locate between two light surfaces, the ratio of the total energy and the total angular momentum is restricted; the detail has been discussed by Takahashi et al. [11].

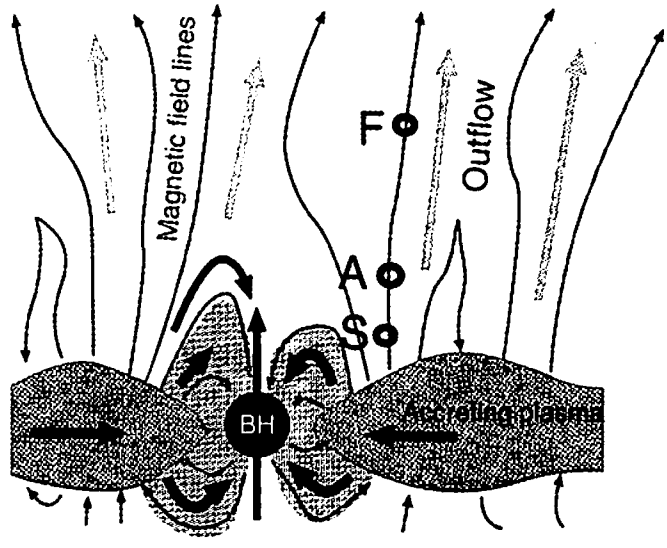


Figure 2: A schematic picture of black hole magnetosphere.

3 Magnetosonic Point Conditions

Here, to see the behavior of flow solutions, differentiating equation (11) with respect to r, we can get [11]

$$(\ln u_p)' = \frac{N}{(u_p^2 - u_{\rm AW}^2)^2(u_p^2 - u_{\rm FM}^2)(u_p^2 - u_{\rm SM}^2)} \tag{16}$$

where $N = N(r, u_p; \tilde{L}, \Omega_F, \mu_I; \Psi)$, and $u_{\rm FM}$ and $u_{\rm SM}$ are the fast and slow magnetosonic wave speed. There are additional singular points in the phase space [spanned by the pair (r, u_p)] where both the numerator and the denominator vanish also. These points $(r_{\rm cr}, u_{\rm cr})$ are the fast and slow magnetosonic points. These locations are obtained by $N_{\rm cr} = 0$ for a given parameter set of E, L, Ω_F, η, K and the considering field line geometry $\Psi = \Psi(r, \theta)$.

In order to construct a shock model with a highly accelerated wind/jet, we must search parameters for trans-fast magnetosonic flow solutions, which pass through the X-type fast magnetosonic point after passing the Alfvén point (and also slow magnetosonic point), and then become super-fast magnetosonic (see Fig. 2). The conditions at the Alfvén point and the fast and slow magnetosonic points restrict the five parameters of the flow which is ejected from a plasma source and streams along the magnetic field. The condition at the fast and slow magnetosonic point restrict the number flux per the magnetic flux-tube η and the total energy E [16]. The magnetic field geometry $\Psi = \Psi(r, \theta)$ and its angular velocity Ω_F should be determined at the boundary of plasma sources in the magnetosphere. Thus, at far distant region, we can expect the MHD shock formation as a GRB remnant.

The example of the restriction for η at the magnetosonic points are shown in Figures 3 and 4. We see that there are three type of branches for fast magnetosonic points; that is, (i) inner-fast point located inside the inner Alfvén point, (ii) middle-fast point located between the inner and outer Alfvén points and (iii) outer-fast point located inside the outer Alfvén point. In these figures, we assume the poloidal magnetic field as $B_p \propto (\Delta\Sigma)^{-1/2} r^{-\delta}$, where $\Delta \equiv r^2 + 2mr + a^2$, $\Sigma \equiv r^2 + a^2 \cos^2\theta$. The parameter δ is constant along a field line and is related to the divergence of a magnetic flux-tube (i.e., $\delta = 0$ means the split monopole magnetic field [10]).

When the rotational effect of the magnetosphere is not so effective ($\tilde{L} < \Omega_F$), the middle fast magnetosonic point appears on the r-u_p (or u^r) plane as seen in Fig. 3, while for $\tilde{L} \approx \Omega_F$ that point do not

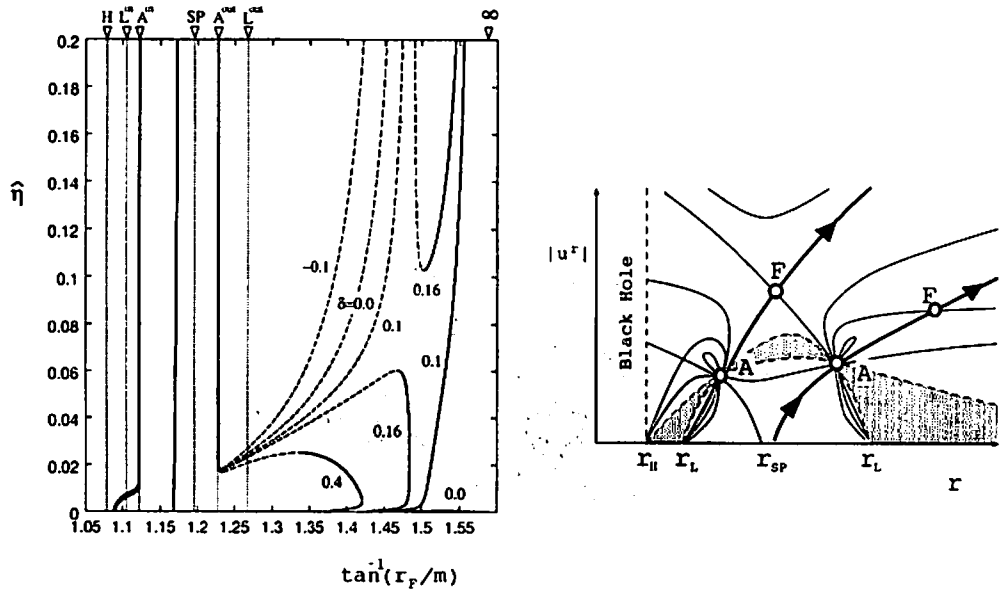


Figure 3: Left: Relations of the location of fast magnetosonic point r_F to the number-flux per magnetic flux-tube $\hat{\eta} \equiv \mu_{\rm cold}\eta/B_{\rm inj}$. $a = 0.5m$. Thick lines denote the X-type fast point, while the broken lines denote O-type (unphysical) point. Right: a schematic picture of outflow solutions. The hatched regions are forbidden regions of physical flow solutions.

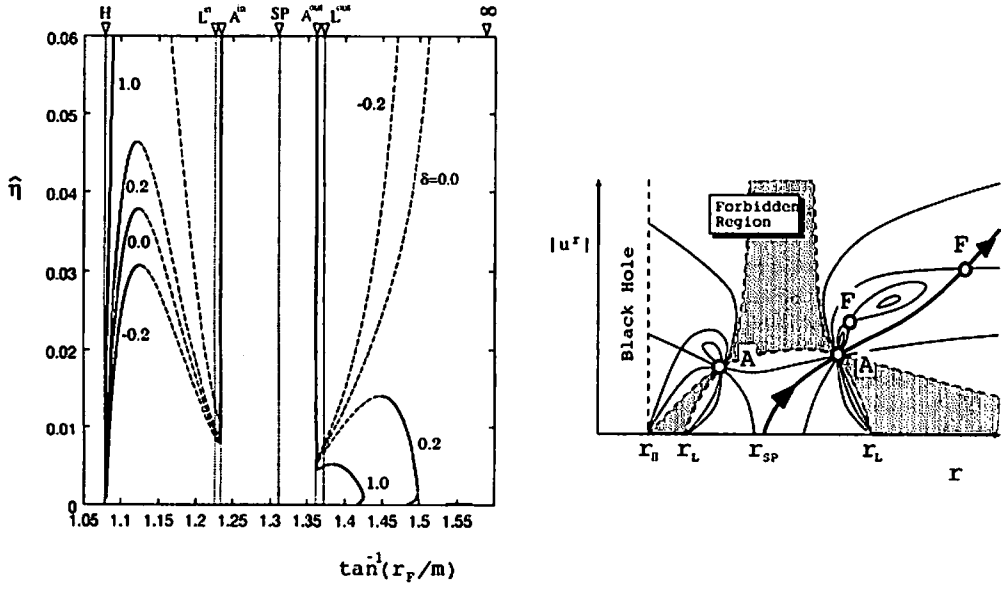


Figure 4: Left: Relations of r_F to $\hat{\eta}$. $a = -0.5m$. Right: The trans-fast magnetosonic solution must pass through the outer-fast magnetosonic point.

appear because of the appearance of the forbidden regions. In the former case, there are two type of solutions (see also [17] for accretions). One passes through the outer-fast magnetosonic point, and the other passes through the middle-fast point. In the fluid-dominated case, the outflow passing through the middle-fast point coincide to a hydrodynamical wind solution, while the solution passing through the outer-fast magnetosonic point disappears. By contrast with this, in the magnetically-dominated case, the middle-fast point disappears and the outer-fast magnetosonic point is only available.

4 Acceleration of Outflowing MHD flows

When the magnetic field is enoughly large, the trans-fast magnetosonic wind/jet, which passes through the outer-fast magnetosonic point, would be realized as seen in Fig. 5 and 6. Highly accelerating flow solutions are obtained for very small η-values ($\hat{\eta} \ll 1$), $\tilde{L}\Omega_F \sim 1$ and $\delta > 0$. In the magnetically-dominated case, we can obtain more accelerated outflows, which would give the large Lorentz factor $\Gamma \gtrsim 10^3$ for GRBs, at the far distant region (of course, the details of the acceleration depends on the field-aligned parameters restricted by the shock conditions and the magnetic field geometry). This situation is very similar to the pulsar wind model discussed by Takahashi [18] without gravity. However, under the strong gravity, the critical conditions are slightly modified; that is, the effects of gravity on the δ-dependence correspond to shift the δ-value down in comparison with non-gravity case. Here, we should note that the field-aligned component of the Lorentz force approximately becomes $(J \times B)_{||} \propto \delta B_\phi^2$; the negative δ means the deceleration of the flow (the gravity acts as negative δ).

For the hot plasma, the outer-fast magnetosonic point shifts toward the outer-Alfvén point, and it is easy to accelerate the plasma because of the gas-pressure. For the radial magnetic field geometry, the accelerating trans-fast magnetosonic is possible, while in the cold limit no accelerating solution is obtained. However, such a trans-fast magnetosonic solution is only available for smaller vales of $\hat{\eta}$. If the gas-pressure increase gradually (in secular time scale), a trans-fast magnetosonic solution which passes through the outer-fast magnetosonic point disappears. Such the hotter outflow remains sub-fast magnetosonic at a far distant region, which means that the shock formation is impossible along the flow.

5 Relativistic MHD Shock Formation

In a complete solution of the wind/jet which includes a shock, the flow must satisfy a set of conditions on either side of the discontinuity. The jump conditions for arbitrary shocks in a relativistic MHD flow are

$$[nu^a]n_a = 0\,, \quad \text{— the particle number conservation} \tag{17}$$

$$[T^{a\beta}]n_a = 0\,, \quad \text{— the energy momentum conservation} \tag{18}$$

$$[\mathrm{E}] \times \mathrm{n} = 0\,, \quad \text{— the continuity relations for the electric field} \tag{19}$$

$$[\mathrm{B}] \cdot \mathrm{n} = 0\,, \quad \text{— the continuity relations for the magnetic field} \tag{20}$$

where $n^a = (n_0, \mathrm{n})$ and the square brackets denote the difference between the values of a quantity on the two side of the shock. From these relations, we see that $\eta_1 = \eta_2$, $(\Omega_F)_1 = (\Omega_F)_2$, $E_1 = E_2$ and $L_1 = L_2$, but we also find $K_1 \neq K_2$, which means the entropy generation, where the subscripts "1" and "2" denote the pre-shock and the post-shock quantities, respectively. The above shock condition for relativistic MHD wind/jet has discussed by Appl and Camenzind [19]. All shock physical quantities are expressed in terms of the relativistic compression ratio. This compression ratio is the solution of a polynomial of seventh degree, which has to be solve simultaneously with an equation for the polytropic index for the shocked plasma. The upstream state of the wind/jet is then determined by the down-stream state of the shocked plasma as specified by the Lorentz factor, the ratio of the electromagnetic to the material energy content in the shocked plasma and the orientation of the magnetic field. If the states of the shocked plasma become clear, the field aligned five parameters of the wind/jet are restricted by the shock conditions. By combining the shock conditions with the constrains at the magnetosonic points, we would be able to specify the properties of the wind/jet generating the GRBs. Further, we will be able to clarify the structure of the central engine.

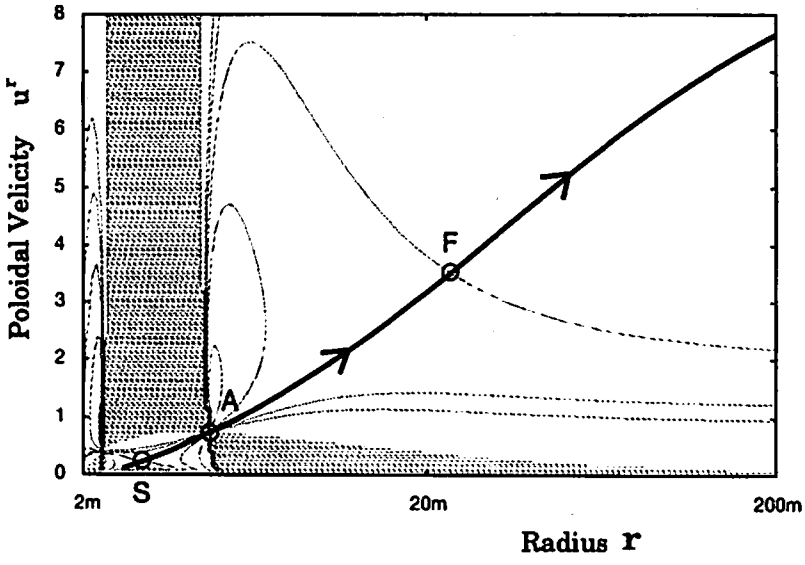


Figure 5: Solutions of trans-fast magnetosonic wind in the black hole magnetosphere. In a hot wind case in radial magnetic field geometry, the slow magnetosonic point (S) exists near the horizon, where $a = 0$. After passing the Alfvén point and the fast magnetosonic point, the wind can be highly accelerated.

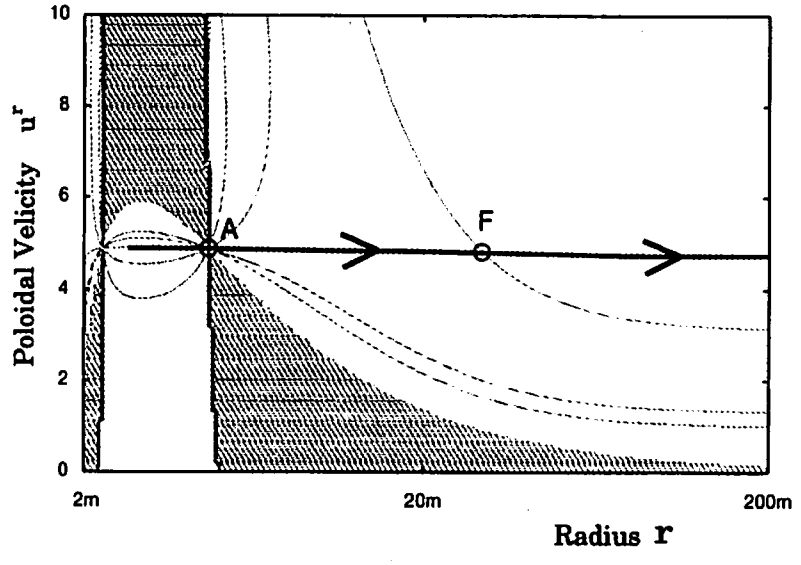


Figure 6: Solutions of trans-fast magnetosonic wind in the black hole magnetosphere. In a cold wind case in radial magnetic field geometry, the acceleration of fluid is inefficient. But an initially high velocity wind is possible for tenuous plasma without the slow magnetosonic point.

References

[1] M.J. Rees and P. Mészáros, MNRAS, **258** (1992) P41.

[2] V.V. Usov, Nature, **357** (1992) 472.

[3] P. Ghosh and M.A. Abramowicz, MNRAS, **292** (1997) 887.

[4] R. Popham, et al., ApJ, 518 (1999) 356.

[5] P. Mészáros and M.J. Rees, ApJ, **405** (1993) 278.

[6] J. Katz, ApJ, **422** (1994) 248.

[7] R. Sari, R. Narayan and T. Piran, ApJ, **473** (1996) 204.

[8] R.D. Blandford and C.F. McKee, Phys. Fluids, **19** (1976) 1130.

[9] H.K. Lee, R.A.M.J. Wijers and G.E. Brown, Physics Reports, **325** (2000) 83.

[10] R.D. Blandford and R.L. Znajek, MNRAS, **179** (1977), 433.

[11] M. Takahashi, S. Nitta, Y. Tatematsu and A. Tomimatsu, ApJ, **363** (1990), 206.

[12] J.D. Bekenstein and E. Oron, Phys. Rev. **D18** (1978) 1809.

[13] M. Camenzind, A&A, **156** (1986) 137.

[14] M. Camenzind, A&A, **162** (1986) 32.

[15] M. Camenzind, A&A, **184** (1987) 341.

[16] M. Takahashi, Il Nuovo Cimento, **115** (2000) 843.

[17] M. Takahashi, in *Proceedings of the 19th Texas Symposium on Relativistic Astrophysics and Cosmology*, edited by E. Aubourg, T. Montmerle, J. Paul and P. Peter (North-Holland, Amsterdam) 2000, CD-ROM 01/27.

[18] M. Takahashi, PASJ, **50** (1998) 271.

[19] S. Appl and M. Camenzind, A&A, **206** (1988), 258.

General Relativistic Simulations of Jet Formation in a Rapidly Rotating Black Holes

Shinji Koide,[1]

Faculty of Engineering, Toyama University, Gofuku 3190, Toyama 930-8555, Japan

David L. Meier,[2]

Jet Propulsion Laboratory, 4800 Oak Grove Dr. Pasadena, CA 91109, USA

Kazunari Shibata,[3]

Kwasan and Hida Observatories, Kyoto University, Yamashina, Kyoto, 607-8471, Japan

Takahiro Kudoh,[4]

National Astronomical Observatory, Mitaka, Tokyo 181-8588, Japan

Abstract

To investigate the formation mechanism of relativistic jets in active galactic nuclei and micro-quasars, we have developed a new general relativistic magnetohydrodynamic code in Kerr geometry. Here we report on the first numerical simulations of jet formation in a rapidly-rotating ($a = 0.95$) Kerr black hole magnetosphere. We study cases in which the Keplerian accretion disk is both co-rotating and counter-rotating with respect to the black hole rotation, and investigate the first ~ 50 light-crossing times. In the co-rotating disk case, our results are almost the same as those in Schwarzschild black hole cases: a gas pressure-driven jet is formed by a shock in the disk, and a weaker magnetically-driven jet is also generated outside the gas pressure-driven jet. On the other hand, in the counter-rotating disk case, a new powerful magnetically-driven jet is formed inside the gas pressure-driven jet. The newly found magnetically-driven jet in the latter case is accelerated by a strong magnetic field created by frame dragging in the ergosphere. Through this process, the magnetic field extracts the energy of the black hole rotation.

1 Introduction

Radio jets ejected from radio loud active galactic nuclei (AGNs) sometimes show proper motion with apparent velocity exceeding the speed of light c [1, 2]. The widely-accepted explanation for this phenomenon, called superluminal motion, is relativistic jet flow in a direction along the observer's line-of-sight with a Lorentz factor greater than 2 [3]. Such relativistic motion is

[1]E-mail:koidesin@ecs.toyama-u.ac.jp
[2]E-mail:dlm@cena.jpl.nasa.gov
[3]E-mail:shibata@kwasan.kyoto-u.ac.jp
[4]E-mail:kudoh.takahiro@nao.ac.jp

thought to originate from a region very close to the putative supermassive black hole which is thought to power each AGN [4, 5]. On the other hand, the great majority of AGNs are radio quiet and do not produce powerful relativistic radio jets [5]. These two classes of active objects (radio loud and quiet) are also found in the black hole candidates (BHCs) in our own Galaxy. Objects with superluminal jets, such as GRS 1915+105 and GRO J1655-40, belong to the radio loud class [6, 7]. Other objects such as Cyg X-1 and GS 1124-68 are relatively radio quiet and produce little or no jet.

What causes the difference between the two classes? Recent observations of the BHCs in our Galaxy suggest that the Galactic superluminal sources contain very rapidly rotating black holes (normalized angular momentum, $a \equiv J/[GM_{\rm BH}^2/c] = 0.9 - 0.95$, where G and $M_{\rm BH}$ are the gravitational constant and black hole mass, respectively), while the black holes in Cyg X-1 and GS 1124-68 are spinning much less rapidly ($a = 0.3 - 0.5$) [8]. A similar rapidly rotating black hole is also suggested in the AGN of the Seyfert 1 galaxy MCG-6-30-15 by the X-ray satellite ASCA [9]. According to recent (nonrelativistic) studies of magnetically-driven jets from accretion disks by Kudoh & Shibata [10, 11], the terminal velocity of the formed jet is comparable to the rotational velocity of the disk at the foot of the jet. Further nonrelativistic simulations of jet formation confirm these results [12, 13, 14, 15, 16], except for the extremely large magnetic field/high jet-power case [17, 18] in which very fast jets can be produced. The rotation velocity at the innermost stable orbit of the Schwarzschild black hole ($r = 3r_S$) is $0.5c$, where $r_S = 2GM_{\rm BH}/c^2$ is the Schwarzschild radius. In addition, it appears that the poloidal magnetic field strength in disks around non-rotating black holes may be not extremely strong if the magnetic field energy density is comparable with that of the radiation [19, 5]. Therefore, a jet produced by MHD acceleration from an accretion disk around a non-rotating black hole should be sub-relativistic and very weak. In fact, numerical simulations of jet formation in a Schwarzschild metric show only sub-relativistic jet flow [20], except for the case when the initial black hole corona is in hydrostatic equilibrium rather than free fall [21].

Several mechanisms for relativistic jet formation from rotating black holes have been proposed [22, 23]. However, up until now no one has performed a self-consistent numerical simulation of the dynamic process of jet formation in a rotating black hole magnetosphere, while there are a few simulations of plasma in the magnetosphere [24, 25]. To this end, we have developed a Kerr general relativistic magnetohydrodynamic (KGrMHD) code. In this paper we report briefly on what we believe are some of the first calculations of their kind — simulation of jet formation in a rotating black hole magnetosphere.

2 Numerical Method

We use a 3 + 1 formalism of the general relativistic conservation laws of particle number, momentum, and energy and Maxwell equations with infinite electric conductivity [26]. The Kerr metric, which describes the spacetime around a rotating black hole, is used in the calculation. When we use Boyer-Lindquist coordinates, $x^0 = ct$, $x^1 = r$, $x^2 = \theta$, and $x^3 = \phi$, the Kerr metric $g_{\mu\nu}$ is written as follows,

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -h_0^2(cdt)^2 + \sum_{i=1}^{3} h_i^2(dx^i)^2 - 2h_3\Omega_3 cdtdx^3. \qquad (1)$$

By modifying the lapse function in our Schwarzschild black hole code ($\alpha = \sqrt{1 - r_S/r}$) to be $\alpha = \sqrt{h_0^2 + \Omega_3^2}$, and adding some terms of Ω_3 to the time evolution equations, we were able to

develop a KGrMHD code relatively easily. (See Appendix C in [20] for more details on this procedure and the meaning of symbols used.)

We use the Zero Angular Momentum Observer (ZAMO) system for the 3-vector quantities, such as velocity v, magnetic field B, electric current density J, and so on. For scalars, we use the frame comoving with the fluid flow. The simulation is performed in the region $0.75r_S \leq r \leq 20r_S$, $0 \leq \theta \leq \pi/2$ with 210×70 mesh points, assuming axisymmetry with respect to the z-axis and mirror symmetry with respect to the plane $z = 0$. The assumptions of the symmetry tend to suppress the turbulence in the disk caused by Balbus-Hawley instability. A free boundary condition is employed at $r = 0.75r_S$ and $r = 20r_S$. In the simulations, we use simplified tortoise coordinates, $x = \log(r/r_H - 1)$, where r_H is the radius of the black hole horizon. To avoid numerical oscillations, we use a simplified TVD method [27, 28, 29, 20]. We checked the KGrMHD code by computing Kepler motion around a rotating black hole and comparing with analytic results [30].

3 Results

The simulations were performed for two cases in which the disk co-rotates and counter-rotates with respect to the black hole rotation. Figures 1a-c illustrate the time evolution of the counter-rotating disk case and Fig. 1d the final state of the co-rotating case. These figures show the rest mass density (gray-scale), velocity (vectors), and magnetic field (solid lines) in $0 \leq R \equiv r\sin\theta \leq 7r_S$, $0 \leq z \equiv r\cos\theta \leq 7r_S$. The black region at the origin shows the inside of the black hole horizon. The angular momentum parameter of the black hole is $a = 0.95$ and the radius is $r_H = 0.656r_S$. The initial state in the simulation consists of a hot corona and a cold accretion disk around the black hole (Fig. 1a). In the corona, plasma is assumed to be in nearly stationary infall, with the specific enthalpy $h/\rho c^2 = 1 + \Gamma p/[(\Gamma - 1)\rho c^2] = 1.3$, where ρ is the rest mass density, p is the pressure, and Γ is specific heat ratio and set $\Gamma = 5/3$. Far from the hole, it becomes the stationary transonic solution exactly. The accretion disk is located at $|\cot\theta| \leq 0.125$, $r \geq r_D = 3r_S$ and the initial velocity of the disk is assumed to be the velocity of a circular orbit around the Kerr black hole. The co-rotating disk is stable, but the counter-rotating disk is unstable in the region $R \leq 4.4r_S$. Except for the disk rotation direction, we use the same initial conditions in both cases. The mass density of the disk is 100 times that of the corona at the inner edge of the disk. The mass density profile is given by that of a hydrostatic equilibrium corona with a scale height of $r_c \sim 3r_S$. The disk is in pressure balance with the corona, and the magnetic field lines are perpendicular to the accretion disk. We use the azimuthal component of the vector potential A_ϕ of the Wald solution to set the magnetic field, which provides a uniform magnetic field far from the Kerr black hole [31]. Here the magnetic field strength far from the black hole is $0.3\sqrt{\rho_0 c^2}$, where ρ_0 is the initial corona density at $r = 3r_S$. However, we do not use the time component of the vector potential A_t from Wald solution; instead, we use the ideal MHD condition $\mathbf{E} + \mathbf{v} \times \mathbf{B} = \mathbf{0}$ to determine the electric field E. Here the Alfvén velocity and plasma beta value at the disk ($r = 3.5r_S$) are $v_A = 0.03c$ and $\beta \sim 3.4$, respectively.

Figure 1b shows the state at $t = 30\tau_S$, where τ_S is defined as $\tau_S \equiv r_S/c$. By this time the inner edge of the disk has rotated 0.75 cycles, *if* we assume the edge is at $R = 3r_S$.[5] Actually, the edge falls toward the black hole and rotates faster at $R = 2r_S$. The rapid infall produces a shock at $R = 3.1r_S$, and the high pressure behind it begins to produce the jet. This is the same pressure-driven jet formation process seen previously in the Schwarzschild case [20].

[5]To calculate inner disk rotation cycles, in this paper we always will assume that the inner edge is located at $R = 3r_S$, regardless of how far inward the edge actually has accreted.

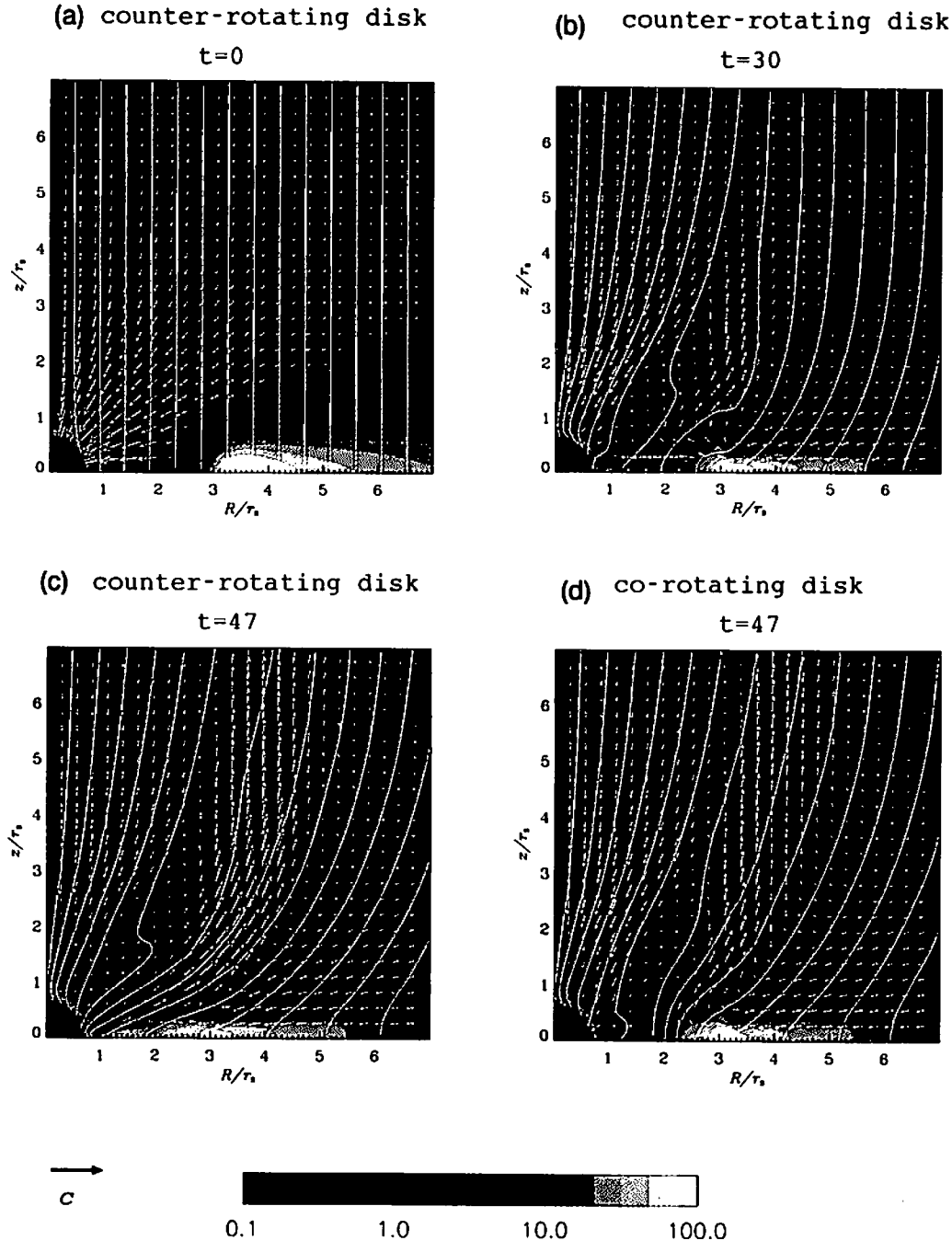


Figure 1: Time evolution of jet formation in the counter-rotating disk case and the final state of the co-rotating disk case. Gray-scale shows the logarithm of the proper mass density; vectors indicate velocity; solid lines show the poloidal magnetic field. The black fan-shaped region at the origin shows the horizon of the Kerr black hole ($a = 0.95$). The dashed line near the horizon is the inner boundary of the calculation region.

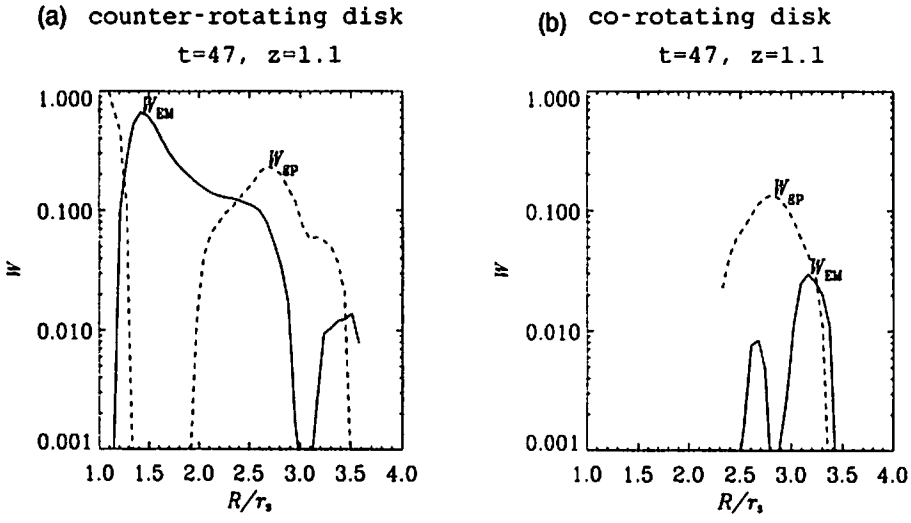


Figure 2: Power contribution to jet acceleration along the line, $z = 1.1r_S$ due to the gas pressure (W_{gp}) and the electromagnetic force (W_{EM}) for both the counter-rotating and co-rotating disk cases. The jet in the counter-rotating disk

Figure 1c shows the final state of the counter-rotating disk case at $t = 47r_S$ when the inner edge of the disk rotated 1.2 cycles. The accretion disk continues to fall rapidly toward the black hole, with the disk plasma entering the ergosphere and then crossing the horizon, as shown by the crowded magnetic field lines near $r = 0.75r_S$. The magnetic field lines become radial due to dragging by the disk infall near the black hole. The jet is ejected almost along the magnetic field lines. Its maximum total and poloidal velocities are the same, $v = v_p = 0.44c$ at $R = 3.2r_S$, $z = 1.6r_S$. The mass density plot (gray-scale) shows that the jet consists of two layers. One is an inner, low density, fast, magnetically-driven jet and the other is an outer, high density, slow, gas pressure-driven jet. The latter comes from the disk near the shock at $R = 3.1r_S$ and is, therefore, similar to the gas pressure-driven jet of Koide, Shibata, & Kudoh [21]. The former is new and has never been seen in the Schwarzschild black hole case. It comes from the disk near the ergosphere, and is accelerated as follows. As there is no stable orbit at $R \leq 4.4r_S$, the disk falls rapidly into the ergosphere. Inside the static limit, the velocity of frame dragging exceeds the speed of light ($c\Omega_3/\alpha > c$), causing the disk to rotate in the *same* direction of the black hole rotation (relative to the fixed Boyer-Lindquist frame), even though it was initially counter-rotating. The rapid, differential frame dragging enhances the azimuthal magnetic field, which then accelerates the flow upward and pinches it into a powerful collimated jet. The energetics of this phenomena are discussed in the last section.

Figure 1d shows a snapshot of the co-rotating disk case at $t = 47r_S$. The disk stops its infall near $R = 3r_S$ due to the centrifugal barrier with a shock at $r = 3.4r_S$. The high pressure behind the shock causes a gas pressure-driven jet with total and poloidal velocities of $v = v_p = 0.30c$ at $R = 3.4r_S$, $z = 2.4r_S$. A detailed analysis shows that a weak magnetically-driven jet is formed outside the gas pressure-driven jet with maximum total and poloidal velocities of $v = 0.42c$ and $v_p = 0.13c$, respectively. This two-layered shell structure is similar to that of Schwarzschild black hole case [21]. The centrifugal barrier makes the disk take much long time to reach the ergosphere, which causes the difference between the co-rotating and counter-rotating disk cases.

To confirm the jet acceleration mechanism, we estimate the power from the electromagnetic field, $W_{\rm EM} = \mathbf{v} \cdot (\mathbf{E} + \mathbf{J} \times \mathbf{B})$ and the gas pressure, $W_{\rm gp} = -\mathbf{v} \cdot \nabla p$ along the line, $z = 1.1 r_{\rm S}$ which crosses the jet foot (Fig. 2). At $t = 47\tau_{\rm S}$, the gas pressure is dominant in the co-rotating disk case (Fig. 2b). However, in the counter-rotating disk case, the electromagnetic power is dominant near the black hole even through the gas pressure power is the same as that of the co-rotating disk case (Fig. 2a). The magnetically-driven jet in this latter case is accelerated by the magnetic field anchored to the ergospheric disk. The frame dragging effect rapidly rotates the disk in the same direction as the black hole rotation, increasing the azimuthal component of the magnetic field and the magnetic tension which, in turn, accelerates the plasma by the magnetic pressure and centrifugal force, respectively. (A detailed analysis shows that both component of the magnetic forces are comparable.) This mechanism of jet production, therefore, is a kind of Penrose process that uses the magnetic field to extract rotational energy of the black hole and eject a collimated outflow from very near the horizon.

4 Discussion

We have presented general relativistic simulations of jet formation from both counter-rotating and co-rotating disks in a Kerr black hole magnetosphere. We have found that jets are formed in both cases. At the time when the simulations were stopped ($t = 47\tau_{\rm S}$ ($53\tau_{\rm S}$), after the inner edge of the disk had rotated 1.2 (1.4) cycles in the counter-rotating (co-rotating) disk case) the poloidal velocities of the jets were $v \sim 0.4c$ (counter-rotating), $\sim 0.3c$ (co-rotating), both sub-relativistic. In the co-rotating disk case, the jet has a two-layered structure: inner, gas pressure-driven jet and outer, magnetically-driven jet. On the other hand, in the counter-rotating case, a new magnetically-driven jet has been found inside the gas pressure-driven jet. The new jet is accelerated by the magnetic field induced by the frame dragging effect in the ergosphere. In this case, existence of a magnetically-driven jet is not clear outside the gas pressure-driven jet. A longer term simulation may show a three-layered structure including the outer, magnetically-driven jet.

Here, we discuss briefly the energy transport from the magnetic field enhancement due to the frame dragging effect of the rotating black hole to the newly found jet in the counter-rotating disk case. The detailed analysis of the transport (including spatial distribution and transport of energy and momentum) will be presented in our next paper. The general relativistic Maxwell equations yield the transport equation of the electromagnetic energy,

$$\frac{\partial e_\infty^{\rm EM}}{\partial t} = -\nabla \cdot \mathbf{Q}^{\rm EM} + \alpha W_{\rm EM} + W_{\rm fd},$$

where $e_\infty^{\rm EM} \equiv \alpha[B^2/2 + E^2/(2c^2)] + \Omega_3(\mathbf{E} \times \mathbf{B})_\phi/c$ is the density of the collected electromagnetic energy, called *energy-at-infinity* of the electromagnetic field, $\mathbf{Q}^{\rm EM}$ is the electromagnetic energy flux density, and $W_{\rm fd} \equiv -c\Omega_3[\rho_{\rm e} E_\phi + (\mathbf{J} \times \mathbf{B})_\phi]$ with the electric charge density, $\rho_{\rm e}$. The power density of the frame dragging effect to the magnetic field energy increase is expressed by $W_{\rm fd}$. The region where $W_{\rm fd}$ is significant is located almost exclusively in the ergosphere, and the value of $W_{\rm fd}$ is positive and huge ($W_{\rm fd} \sim 3\rho_0 c^2/\tau_{\rm S}$) in the region at $t = 47\tau_{\rm S}$. The integration of $W_{\rm fd}$ over the region ($L_{\rm fd}$) evaluates the magnetic field energy increase rate due to the frame dragging effect. The numerical result of the counter-rotating disk case at $t = 47\tau_{\rm S}$ provides $L_{\rm fd} = 3.6\rho_0 c^3 r_{\rm S}^2$. On the other hand, the integration of $\alpha W_{\rm EM}$ over the jet forming region ($L_{\rm EM}$) evaluates the power of the electromagnetic field to the jet acceleration. The numerical result gives $L_{\rm EM} = 0.72\rho_0 c^3 r_{\rm S}^2$. This shows that about 20% of the frame dragging power to the magnetic field is converted to the kinetic energy of the jet. The total kinetic energy of the jet is

$K_{\rm jet} = 0.77\rho_0 c^2 r_{\rm S}^3$. Therefore, the electromagnetic field induced by the frame dragging effect is capable of accelerating the jet plasma almost continuously.

Our simulation results suggest that the relativistic jet is ejected in the counter-rotating disk case easier (at least, more quickly) than the co-rotating disk case. (However, this is not confirmed yet because of the limitation of the simulation time.) The observed X-ray properties of the Galactic superluminal jet sources, GRO J1655-40 and GRS 1915+105, suggested that each contains a black hole spinning rapidly in the *same direction* as the accretion disk [32]. However, this does not mean that the relativistic jet is formed only with the co-rotating disk, that is, the plasma which radiates X-ray may be different from the source plasma which causes the relativistic jet. In fact, multiwavelength monitoring of the Galactic superluminal source GRS 1915+105 show that the onset of the jet ejection events are related to the sudden drop (with a number of sharp spikes) from a luminous state in the hard X-ray [33, 34]. This shows the disk becomes unstable when the jet is formed. For such non-steady disk, the method with the steady-state disk assumption used by Zhang *et al.* (1997) may not be valid to determined the disk rotation direction. The unstable period may be caused by the counter-rotating plasma. Such reversal of the disk rotation is not so surprising, for example, the recent model of the transition between the hard and soft states of Galactic black hole candidate, Cyg X-1 with the reversal [32]. (Of course, the situation of microquasars and Cyg X-1 may be completely different. Cyg X-1 disk reversal model is plausible because the companion star is a blue supergiant and the accretion onto the black hole appears to be fed by a stellar wind from that star. On the other hand, microquasars, GRS 1915+105 and GRO J1655-40 are low-mass binaries, thought to be accreting due to mass transfer through Roche lobe overflow. It is not clear whether a reversal can occur in such low-mass, Roche-lobe overflow binaries.) Furthermore, with respect to AGNs, the evidences for the counter-rotating disk or the rotation direction change have been discussed recently ([35, 36]; and references therein for further details regarding the observations). Detailed observations of the inner disk rotation in both Galactic superluminal sources and AGNs, especially during periods of jet ejection, will provide important clues to the jet formation mechanism in these objects.

S. K. thanks M. Koide for discussions and important comments for this study. We appreciate the support of the National Institute for Fusion Science in the use of their super-computers.

References

[1] Pearson, J. J. 1981, *et al.* Nature **290**, 365.

[2] Hughes, P. A. 1991, eds., *Beams and Jets in Astrophysics* (Cambridge University Press, New York).

[3] Rees, M. J. 1966, Nature **211**, 468.

[4] Linden-Bell, D. 1969, Nature **223**, 690.

[5] Rees, M. J. 1984, Ann. Rev. Astron. Ap. **22**, 471.

[6] Mirabel, I. F. & Rodriguez, L. F. 1994, Nature **371**, 46.

[7] Tingay, S. J. 1995, *et al.* Nature **374**, 141.

[8] Cui, W., Zhang, S. N., & Chen, W. 1998, ApJ **484**, 383.

[9] Iwasawa, K., *et al.* 1996, MNRAS, **282**, 1038.

[10] Kudoh, T. & Shibata, K. 1995, ApJ **452**, L41.

[11] Kudoh, T. & Shibata, K. 1997a, ApJ **474**, 362.

[12] Shibata, K. & Uchida, Y. 1986, PASJ **38**, 631.

[13] Kudoh, T. & Shibata, K. 1997b, ApJ **476**, 632.

[14] Ouyed, R., Pudritz, R. E., & Stone, J. M. 1997, Nature **385**, 409.

[15] Romanova, M. M., Ustyugova, G. V., Koldoba, A. V., Chechetkine, V. M., & Lovelace, R. V. E. 1997, ApJ **482**, 708.

[16] Ustyugova, G. V., Koldoba, A. V., Romanova, M. M., Chechetkine, V. M., & Lovelace, R. V. E. 1999, ApJ **516**, 221.

[17] Meier, D. L., Edgingon, S., Godon, P., Payne, D. G., & Lind, K. R. 1997, Nature **388**, 350.

[18] Meier, D. L. 1999, ApJ **522**, 753.

[19] Begelman, .M.C., Blandford, R.D., Rees, M.J. 1984, Review of Modern Physics, **56**, 255.

[20] Koide, S., Shibata, K., & Kudoh, T., 1999a, ApJ **522**, 727.

[21] Koide, S., Shibata, K., & Kudoh, T. 1998, ApJ **495**, L63.

[22] Blandford, R. D. & Znajek, R. 1977, MNRAS **179**, 433.

[23] Takahashi, M., Nitta, S., Tatematsu, Y., & Tomimatsu, A. 1990, ApJ **363**, 206.

[24] Yokosawa, M., Ishizuka, T., & Yabuki, Y. 1991, PASJ **43**, 427.

[25] Yokosawa, M. 1993, PASJ **45**, 207.

[26] Thorne, K. S., Price, R. H., & Macdonald, D. A. 1986, *Membrane Paradigm* (Yale University Press, New Haven and London).

[27] Davis, S. F. 1984, NASA Contractor Rep. 172373, ICASE Rep. No. 84-20.

[28] Koide, S., Nishikawa, K.-I., & Mutel, R. L. 1996, ApJ **463**, L71.

[29] Koide, S. 1997, ApJ **478**, 66.

[30] Shapiro, S. L. & Teukolsky, S. A. 1983, *Black Holes, White Dwarfs, and Neutron Stars*, (John Wiley & Sons Inc., New York)

[31] Wald, R. M. 1974, Phys. Rev. D **10**, 1680.

[32] Zhang, S. N., Cui, W., Chen, W. 1997, ApJ **482**, L155.

[33] Mirabel, I. F., Rodriguez, L. F., Chaty, S., Sauvage, M., Gerard, E., Duc, P.-A., Castro-Tirado, A., & Callanan, P 1996, ApJ **472**, L111.

[34] Mirabel, I. F. & Rodriguez, L. F. 1998, Nature **392**, 673.

[35] Lovelace, R. V. E. & Chou, T. 1996, ApJ **486**, L25.

[36] Kuznetsov, O. A., Lovelace, R. V. E., Romanova, M. M., & Chechetkin, V. M. 1999, ApJ **514**, 691.

Possible Direct Method to Determine the Radius of a Star

Takashi Nakamura [1] and Motoyuki Saijo [2]

[1] *Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan*

[2] *Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801-3080, U.S.A.*

Abstract

We computed the spectrum of gravitational waves from a dust disk star of radius R inspiraling into a Kerr black hole of mass M and specific angular momentum a. We found that when R is much larger than the wave length of the quasinormal mode, the spectrum has several peaks and the separation of peaks $\Delta\omega$ is proportional to R^{-1} irrespective of M and a. This suggests that the radius of the star in coalescing binary black hole - star systems may be determined directly from the observed spectrum of gravitational waves. This also suggests that the spectrum of the radiation may give us important information in gravitational wave astronomy as in optical astronomy.

A number of worldwide projects on the direct detection of gravitational waves using laser interferometers such as LIGO, VIRGO, GEO600, TAMA300, and LISA as well as several projects using resonant-type detectors, are in progress [2]. One of the most important sources of gravitational waves for these detectors is coalescing binary black hole (BH) - star systems. In the inspiral phase, the radius of the star has little effect on gravitational waves so that the star can be treated as a point particle. Comparing the waveform of the inspiral phase with a theoretical template using matched filtering techniques, we may determine each mass and spin of the star and BH, respectively [2]. While, in the final merging phase, the effect of the radius is important and some information on the radius of the star may be extracted from gravitational wave signals. However, this merger phase is much less well understood than the inspiral one.

In this Letter, as a model of this final merging phase, we computed the spectrum of gravitational waves from a dust disk star of radius R inspiraling into a Kerr BH of mass M and specific angular momentum a. We found that when R is much larger than the wave length of the quasinormal mode (QNM) of Kerr BH, the spectrum has several peaks and the separation of peaks $\Delta\omega$ is proportional to R^{-1} irrespective of M and a. This suggests that the radius of the star in coalescing binary BH – star systems may be determined directly from the observed spectrum of gravitational waves without any assumption to equation of state. [3]

First, let us consider gravitational waves from a particle plunging into a Kerr BH of mass M and specific angular momentum a. The Fourier component of the radial wave function of gravitational waves $X_{lm\omega}(r^*)$ obeys the Sasaki-Nakamura equation [Eq. (2.28) of Ref. [5]] given by

$$\left[\frac{d^2}{dr^{*2}} - F(r^*)\frac{d}{dr^*} - U(r^*)\right] X_{lm\omega}(r^*) = S_{lm\omega}(r^*), \tag{1}$$

where r^*, $S_{lm\omega}(r^*)$, $F(r^*)$ and $U(r^*)$ are the tortoise coordinate of Kerr BH, the source term from the test particle of mass μ [Eq. (2.29) of Ref. [5]] and two potential functions [Eqs. (2.12a) and (2.12b)

[1] Electronic address: takashi@yukawa.kyoto-u.ac.jp

[2] Electronic address: saijo@astro.physics.uiuc.edu

[3] Vallisneri [3] recently argued that the radius of the star may be determined if the frequency of gravitational waves at the tidal disruption is obtained observationally. However, his method needs some information about the equation of state [4] to know the frequency at the tidal disruption.

of Ref. [5]], respectively. To solve $X_{lm\omega}(r^*)$ using a Green function method, we need two independent homogeneous solutions whose boundary conditions are given by

$$X^{\rm in(0)}_{lm\omega}(r^*) = \begin{cases} e^{-ikr^*} & r^* \to -\infty, \\ A^{\rm in}_{lm\omega}e^{-i\omega r^*} + A^{\rm out}_{lm\omega}e^{i\omega r^*} & r^* \to \infty, \end{cases}$$

$$X^{\rm out(0)}_{lm\omega}(r^*) = \begin{cases} B^{\rm in}_{lm\omega}e^{-ikr^*} + B^{\rm out}_{lm\omega}e^{ikr^*} & r^* \to -\infty, \\ e^{i\omega r^*} & r^* \to \infty, \end{cases}$$

where $k = \omega - ma/[2(M + \sqrt{M^2 - a^2})]$. Then, the inhomogeneous solution to Eq. (1) becomes

$$X_{lm\omega}(r^*) = X^{\rm in(0)}_{lm\omega}\int_{r^*}^{\infty}\frac{S_{lm\omega}(r^*)}{W}X^{\rm out(0)}_{lm\omega}\,dr^* + X^{\rm out(0)}_{lm\omega}\int_{-\infty}^{r^*}\frac{S_{lm\omega}(r^*)}{W}X^{\rm in(0)}_{lm\omega}\,dr^*, \tag{2}$$

where W is the Wronskian,

$$W \equiv X^{\rm in(0)}_{lm\omega}\frac{dX^{\rm out(0)}_{lm\omega}}{dr^*} - X^{\rm out(0)}_{lm\omega}\frac{dX^{\rm in(0)}_{lm\omega}}{dr^*}. \tag{3}$$

The asymptotic behavior of the radial wave function $X_{lm\omega}(r^*)$ is given by

$$\begin{aligned} X_{lm\omega}(r^*) &= A_{lm\omega}e^{i\omega r^*}, \\ A_{lm\omega} &= \int_{-\infty}^{\infty}\frac{S_{lm\omega}(r^*)}{W}X^{\rm in(0)}_{lm\omega}\,dr^*. \end{aligned} \tag{4}$$

Then, the energy spectrum and waveform of gravitational waves [Eqs. (3.6) and (3.10) of Ref. [6]] are given by

$$\left(\frac{dE}{d\omega}\right)_{lm\omega} = 8\omega^2\left|\frac{A_{lm\omega}}{c_0}\right|^2 \qquad (-\infty < \omega < \infty), \tag{5}$$

$$h_+ - ih_\times = \frac{8}{r}\int_{-\infty}^{\infty}d\omega\, e^{i\omega(r^*-t)}\sum_{l,m}\left[\left(\frac{A_{lm\omega}}{c_0}\right){}_{-2}S^{a\omega}_{lm}(\theta)\frac{e^{im\varphi}}{\sqrt{2\pi}}\right], \tag{6}$$

where r is the coordinate radius from the center of BH, ${}_{-2}S^{a\omega}_{lm}$ is the spin -2 weighted spherical harmonics, c_0 is a constant given in Ref. [6].

Next, we consider gravitational waves from a dust disk [6, 7, 8]. We note that a geodesic in the equatorial plane in Kerr BH is characterized by the specific energy ($\tilde{E}$) and angular momentum ($\tilde{L}_z$). Let $t = T(r)$ and $\phi = \Phi(r)$ express an orbit of the geodesic for given $\tilde{E}$ and $\tilde{L}_z$. Then another geodesic for the same $\tilde{E}$ and $\tilde{L}_z$ is expressed by $t = T(r) + c_t$ and $\phi = \Phi(r) + c_\phi$ where c_t and c_ϕ are constants. Specifically the geodesic, whose location at $t = T(r_0)$ is $r = r_i$ and $\phi = \phi_i$, is expressed by $t = T(r) + T(r_0) - T(r_i)$ and $\phi = \Phi(r) + \phi_i - \Phi(r_i)$. It is possible to set a number of particles to form a disk of radius R whose center is $r = r_0$ and $\phi = \Phi(r_0)$ at $t = T(r_0)$.

Each particle in the disk emits the same gravitational waves as the particle in the center of the disk except that the phase in t and ϕ is different from the central particle [7]. Then the amplitude of the gravitational waves from the dust disk star is computed as

$$A^{\rm (disk)}_{lm\omega} = f_{m\omega}A^{\rm (particle)}_{lm\omega}, \tag{7}$$

$$f_{m\omega} = 2\frac{\mu}{S}\int_{r_0-R}^{r_0+R}dr\; r\frac{\sin(m\phi_0(r))}{m}e^{i[\omega[T(r)-T(r_0)]-m[\Phi(r)-\Phi(r_0)]]}, \tag{8}$$

$$\phi_0(r) = cos^{-1}\frac{r^2 + r_0^2 - R^2}{2rr_0}, \tag{9}$$

where $f_{m\omega}$ and $A^{\rm (particle)}_{lm\omega}$ are a form factor and the amplitude of gravitational waves at infinity for a single particle case, respectively. S is the normalization factor given by

$$S = 2\int_{r_0-R}^{r_0+R}dr\; r\phi_0(r). \tag{10}$$

In this Letter, we only show the result for $r_0 = 10M$, $a/M = 0.9$, $\tilde{E} = 1$ and $\tilde{L}_z = 2M$. In this case, the orbit rotates about π angle around the BH within the radius of $r = r_0$, and the velocity at $r = 3M$, the location where the particle might radiate the characteristic QNM frequency, is $v \sim 0.48c$, where c is the speed of light. As for the choice of the parameter $r_0 = 10M$, we can refer the result of the relativistic Roche limit [Eq. (5.1) of Ref. [4]] as

$$M < 4.64 M_\odot \left(\frac{1.4 M_\odot}{M_*}\right)^{1/2} \left(\frac{R}{10\ \mathrm{km}}\right)^{3/2} \left(\frac{6M}{r_c}\right)^{3/2}, \tag{11}$$

where M_* and r_c are the mass of the star and the Roche radius, respectively. For appropriate choice of parameters $r_0 = 10M$ corresponds to the Roche limit.

Figures 1(a) and 2(a) show the characteristics of gravitational waves from a test particle plunging into a Kerr BH of $a/M = 0.9$. Figure 1(a) shows the spectrum for $l = 2$ and Fig. 2(a) shows the waveform h_+ observed from the equatorial plane at infinity. From Fig. 1(a), we see that most of gravitational waves are radiated near the frequency of the QNM since the QNM frequency for $a/M = 0.9$ is Re $(M\omega_{\mathrm{QNM}}) = 0.7$ for $l = m = 2$, while the imaginary part of QNM [Im $(M\omega_{\mathrm{QNM}})$] in this case is 0.07 [Fig. 1a of [9] and Fig. 3(c) of [10]]. Figure 2(a) shows clearly that there exists a typical ringing tail at late time [Fig. 4 of Ref. [11] and Fig. 4 of Ref. [12]]. Figs. 1(b), 1(c), 2(b), and 2(c) show the spectra and the waveform from the circular dust disk star inspiraling into a Kerr BH with $a/M = 0.9$. The radius of the disk is set up as $R/M = 1.56$ for Figs. 1(b) and 2(b) while $R/M = 5.88$ for Figs. 1(c) and 2(c). For the case of Fig. 1(b), the peak of the spectrum is located at $M\omega \sim 0.5$ which is quite different from the QNM frequency Re $(M\omega_{\mathrm{QNM}}) \sim 0.7$ although the ringing tail can be seen in Fig. 2(b). When we turn to look at Fig. 1(c), there are several peaks in the spectrum and the ringing tail is very low or absent in the waveform of Fig. 2(c). This is a new finding. The reason for this behavior is as follows. The energy spectrum $(dE/d\omega)^{(\mathrm{disk})}_{lm\omega}$ of gravitational waves from the disk is expressed as

$$\left(\frac{dE}{d\omega}\right)^{(\mathrm{disk})}_{lm\omega} \propto |\, f_{m\omega}\,|^2 \left(\frac{dE}{d\omega}\right)^{(\mathrm{particle})}_{lm\omega}, \tag{12}$$

where $(dE/d\omega)^{(\mathrm{particle})}_{lm\omega}$ is the spectrum from the single test particle. The spectrum from the test particle has only one peak at the frequency ω_{QNM} [Fig. 1(a)] so that the square of the form factor $|\, f_{m\omega}\,|^2$ is responsible for this behavior. To confirm this, we show $|\, f_{m\omega}\,|^2$ for $R/M = 1.56$ [Fig. 3(a)] and $R/M = 5.88$ [Fig. 3(b)] and find that $|\, f_{m\omega}\,|^2$ is responsible for the behavior of the spectrum in Figs. 1(b) and 1(c). The existence of several peaks in Figs. 3(a) and 3(b) can be understood by the approximate estimation of $f_{m\omega}$ as

$$f_{m\omega} \propto \frac{\sin(\omega T'_{r=r_0} R)}{\omega}, \tag{13}$$

where $T' = dT(r)/dr$. The frequency where $f_{m\omega}$ takes zero is

$$\omega_n \sim \frac{(n+1)\pi}{T'_{r=r_0} R}, \qquad n = 0, 1, 2, \ldots, \tag{14}$$

which agrees quite well with the numerical results of Figs. 3(a) and 3(b). Equation (14) suggests that the separation of peaks of the spectrum $\Delta\omega$ may be in proportion to R^{-1}. In Table 1, we show $\Delta\omega$ for various R/M and we found that

$$R = C\frac{1}{\Delta\omega} \quad \text{for } R/M \gg 1.6, \tag{15}$$

where a constant C is close to 1 in the present case. In the physical unit, $\Delta\nu \equiv \Delta\omega/(2\pi)$ is given by

$$\Delta\nu = 5\ \mathrm{kHz} \left(\frac{R}{10\ \mathrm{km}}\right)^{-1} = 10\ \mathrm{Hz} \left(\frac{R}{5000\ \mathrm{km}}\right)^{-1}, \tag{16}$$

assuming that there are 4 peaks in the region of $\omega \in [0, \omega_{\mathrm{QNM}}]$ in the spectrum. Therefore, for neutron stars and white dwarfs, the frequency band is within some laser interferometers and resonant type detectors.

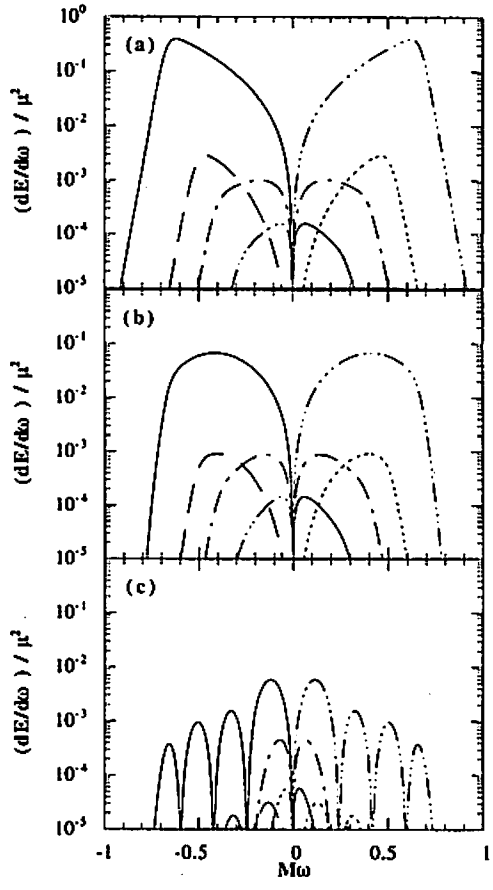


FIG. 1: Energy spectrum of gravitational waves from a dust disk star moving on an equatorial plane in Kerr spacetime whose radius is set up at $r_0 = 10M$ for the case of $a/M = 0.9$, $\tilde{L}_z/M = 2$ [(a) $R/M = 0$ (test particle), (b) $R/M = 1.56$, (c) $R/M = 5.88$]. We show only $l = 2$ mode. Solid, dashed, dash-dotted, dotted, and dash-three dotted lines denote the case of $m = -2, -1, 0, 1, 2$, respectively.

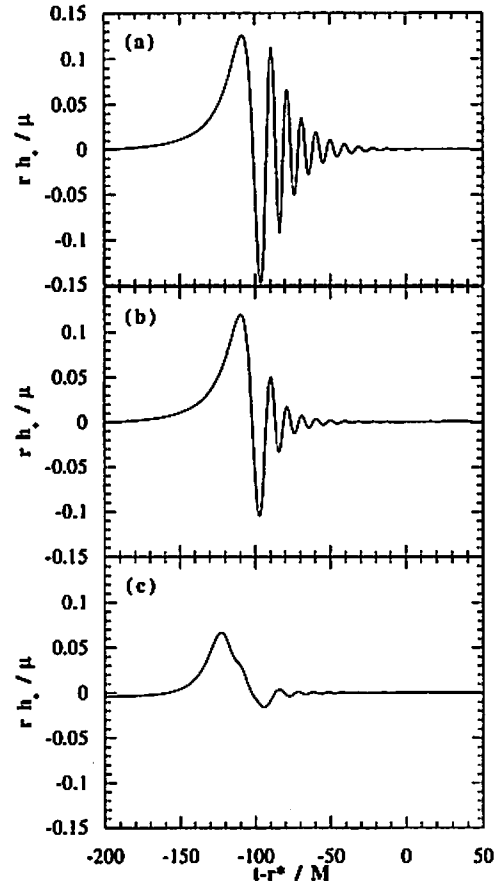


FIG. 2: Waveform of gravitational waves from a dust disk star moving on an equatorial plane in Kerr spacetime whose radius is set up at $r_0 = 10M$ for the case of $a/M = 0.9$, $\tilde{L}_z/M = 2$ [(a) $R/M = 0$ (test particle), (b) $R/M = 1.56$, (c) $R/M = 5.88$]. We set the angle of the observer at infinity as $\theta = \pi/2$, $\phi = 0$. We show only $l = m = 2$ mode to exclude m mode coupling effect.

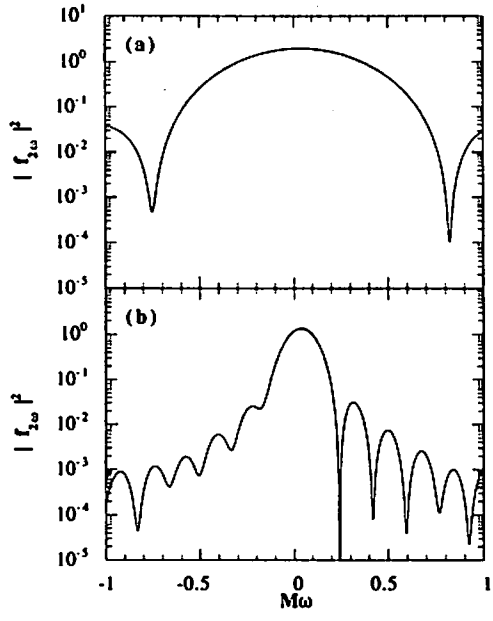


FIG. 3: Form factor of the dust disk star moving on an equatorial plane in Kerr spacetime whose radius is set up at $r_0 = 10M$ for the case of $a/M = 0.9$, $\tilde{L}_z/M = 2$ [(a) $R/M = 1.56$, (b) $R/M = 5.88$]. We show only $l = m = 2$ mode, because $m = 2$ is a dominant mode for above parameters in $l = 2$. It is clear that the form factor is responsible for the spectra in Fig. 1.

Equation (15) strongly suggests that, from the observed value of $\Delta\omega$ in the spectrum of the gravitational wave signals, we may determine the radius of the star if $R/M \gg 1.6$. From the physical point of view, the form factor $f_{m\omega}$ expresses the phase cancellation effect of gravitational waves [7], so that if $R < 1/\omega_{\rm QNM}$ gravitational waves from the dust disk star are essentially regarded as that from the test particle. [4] This suggests that our proposal is valid only if the inspiraling star is tidally disrupted by the BH finally. It is instructive to note here that in reality the form factor may bring us the information about the form of the source, *radius*.

Here, we showed only the results for a special set of parameter such as $r_0 = 10M$, $a/M = 0.9$, $\tilde{E} = 1$, and $\tilde{L}_z = 2M$. To check the robustness of our proposal, we also calculated the spectra from the dust disk star inspiraling into Kerr BH for the wide range of parameters ($\tilde{L}_z$: spiraling case, $0 \leq a/M \leq 0.9$) and found that there exist some peaks in the spectrum for $R \gg 1/\omega_{\rm QNM}$. In these cases we confirmed the relation $R \sim 1/\Delta\omega$. In this Letter, we did not take into account of the pressure and the self-gravity of the star so that it is urgent to confirm our proposal by full 3D numerical simulations including the determination of the value of the constant C in Eq. (15). If the relation of Eq. (15) is confirmed irrespective of the equation of state, Eq. (15) can be adopted as one of the direct methods to determine the radius of the star from gravitational waves. In any case, it is quite possible that the spectrum of the gravitational waves may give us important information in gravitational wave astronomy as in optical astronomy.

M. S. thanks Kei-ichi Maeda for a continuous encouragement. He thanks Masaru Shibata for discussion. He also thanks the visitor system of Yukawa Institute for Theoretical Physics and acknowledges Abhay Ashtekar, Lee Samuel Finn, Jorge Pullin, Hisaaki Shinkai for their kind hospitality at The Pennsylvania State University, where the part of this work was done. The numerical computations are mainly

[4] We also confirmed that this phase cancellation can also be seen in the 3D dust star assuming that each test particle moves the constant polar angle with no orbital angular momentum.

Table 1: Comparison with the characteristic length from the energy spectrum of gravitational waves to the radius of the disk. R denotes the coordinate radius.

R/M	$\Delta\omega$	$1/\Delta\omega$
3.09	0.335	2.99
3.83	0.270	3.70
4.54	0.233	4.29
5.22	0.200	5.00
5.88	0.175	5.71

performed by NEC-SX vector computer at Yukawa Institute for Theoretical Physics, Kyoto University, and FUJITSU-VX vector computer at Media Network Center, Waseda University. This work was supported in part by a JSPS Grant-in-Aid (No. 095689) and by Grant-in-Aid of Scientific Research of the Ministry of Education, Culture, and Sports, No.11640274 and 09NP0801.

References

[1] Original version of this paper is ;
M. Saijo and T. Nakamura, Phys. Rev. Lett. **85**, 2665 (2000).

[2] K. S. Thorne, in *Black Holes and Relativistic Stars*, edited by R. M. Wald (University of Chicago Press, Chicago, 1998), p. 41.

[3] M. Vallisneri, Phys. Rev. Lett. **84**, 3519 (2000).

[4] M. Shibata, Prog. Theor. Phys. **96**, 917 (1996).

[5] M. Sasaki and T. Nakamura, Prog. Theor. Phys. **67**, 1788 (1982).

[6] M. Saijo, H. Shinkai, and K. Maeda, Phys. Rev. **D56**, 785 (1997).

[7] T. Nakamura and M. Sasaki, Phys. Lett. **106B**, 69 (1981).

[8] M. P. Haugan, S. L. Shapiro and I. Wasserman, Astrophys. J. **257**, 283 (1982); S. L. Shapiro and I. Wasserman, Astrophys. J. **260**, 838 (1982); L. I. Petrich, S. L. Shapiro, and I. Wasserman, Astrophys. J. Suppl. **58**, 297 (1985).

[9] S. Detweiler, Astrophys. J. **239**, 292 (1980).

[10] E. W. Leaver, Proc. Roy. Soc. London Ser. A **402**, 285 (1985).

[11] Y. Kojima and T. Nakamura, Phys. Lett. **96A**, 335 (1983).

[12] Y. Kojima and T. Nakamura, Prog. Theor. Phys. **71**, 79 (1984).

Time evolution of purely axial displacements in isentropic neutron stars

Yasufumi Kojima[1]

Department of Physics, Hiroshima University, Higashi-Hiroshima 739-8526, Japan

Abstract

Evolution of purely axial displacement given as initial data is calculated on a slowly rotating relativistic star. The temporary behavior can be described by the classical r-mode oscillation with a single frequency approximately during a short-term period. The oscillation gradually becomes incoherent due to phase lag caused by the relativistic frame-dragging, which depends on the position. We estimate the time-scale, in which the time-evolution is approximated like the classical r-mode. The time-scale is smaller than that of the radiation reaction, so that the modification due to the frame-dragging may in part change the evolutionary picture based on the Newtonian r-mode instability.

1 Introduction

The axial oscillations in the rotating stars have been attractive, since the discovery of the r-mode instability [1, 2, 3]. The radiation reaction instability may play a significant role in the spin-down of neutron stars as well as being a strong detectable gravitational wave source. Most of the studies are however limited to idealized situation. Some realistic effects such as the magnetic field [4, 5] and the solid crust [6, 7] may significantly modify the previous estimate. In this paper, we do not consider such realistic ingredients, but the relativistic dragging effects only, since the previous estimates [4, 5, 6, 7, 8, 9] are based on the normal mode solution calculated in the Newtonian star with uniform density. The classical r-mode can be decoupled from other spherical harmonics. That is, e.g., the velocity perturbation is given by a single magnetic-type vector spherical harmonic. The gravitational radiation is calculated by time variation of the current multipole moments. Recently, it is showed that there is no such a purely axial mode in the relativistic star [10]. Instead, the normal mode solution should be a mixed state of the polar and axial functions.

One important question arises. Are the estimates based on the Newtonian r-mode still valid in the relativistic star? One method to answer it is to calculate the normal mode in the relativistic star and to compare the estimates using resultant frequencies and functions. We consider the axial oscillations from different approach in this paper. We numerically follow the time-evolution of the purely axial displacement given as the initial data. In general, the oscillation in the rotating star is mixed with the different spherical harmonics. In the slow rotation approximation, where the angular velocity Ω is a small expansion parameter, the coupling between the different spherical harmonic indices l and l' can formally be written as the power of Ω, $\sim \Omega^{|l-l'|}$. This is due to the addition rule of the angular momentum. The initial form is therefore preserved for small t, although it is difficult to address the valid domain of t beforehand. Our calculations are restricted to the lowest-order equations with respect to the rotational parameter. By the time-evolution within the approximation, we examine how the eigen-state in the Newtonian star breaks down as the system is relativistic. That is, the velocity perturbation is given in the similar form as in the Newtonian star and the subsequent evolution is numerically simulated. The initial data are not a pure state of a normal mode, rather mixed state of some modes. The mixture is small in the nearly Newtonian star. The approximation is expected to be not good in the relativistic star. It is of importance to examine when and how the approximation breaks down in the relativistic star. From this numerical test, we can find out the time-scale when the approximation breaks down. If the time-scale is long enough, e.g., the radiation reaction time-scale, then we may conclude that the estimate based on the approximation is good. Recently, Stergioulas and Font [11] performed numerical

[1] Email address:kojima@theo.phys.sci.hiroshima-u.ac.jp

simulation of the r-mode oscillation with 3-D general relativistic hydrodynamical code. They followed the time-evolution of fluids up to 20 rotations in fixed gravitational field. Our results will also give a useful test for the similar kinds of the general relativistic simulation.

In Sec.2, our formulation is briefly summarized. Details of the expressions as well as the derivation of the equations can be found in Ref.[12]. If the perturbation obeys the same equation of sate as in the background equilibrium state, the velocity function should be constrained. The function is used as the initial date, and the evolution is followed. The results are shown in Sec.3. Finally in Sec.4, discussion is given.

2 Formulation

A star is assumed to rotate with uniform angular velocity Ω. For the general oscillations with small amplitude, polar and axial parts are mixed in the lowest order of the rotation rate, and they form hybrid mode. However, the classical r-mode can be described by a single purely axial spherical harmonic. The dominant tangential 3-velocity on the sphere can be described by the axial vector spherical harmonic. The tetrad components of the velocity perturbation are

$$[\delta v^{\hat{\theta}}, \delta v^{\hat{\phi}}] = \Re\left(\delta v(t,r)\left[-\frac{1}{\sin\theta}\partial_\phi Y_{lm}(\theta,\phi), \partial_\theta Y_{lm}(\theta,\phi)\right]\right), \tag{1}$$

where the symbol $\Re$ means taking the real part, but is omitted from now on. Since we restrict our calculation to the lowest order, the radial equations governing the axial oscillations are given by

$$(\partial_t + im(\Omega - \chi))U_{lm} + 4\pi(p_0 + \rho_0)r^2 e^{-\nu}(\partial_t + im\Omega)\Phi_{lm} = 0, \tag{2}$$

$$\frac{j^2 r^2}{4}\left[\frac{1}{jr^4}(jr^4\Phi'_{lm})' - (v + 16\pi(p_0+\rho_0)e^{\lambda})\Phi_{lm}\right] = U_{lm}, \tag{3}$$

where a prime denotes a derivation with respect to r and

$$\chi = \frac{2}{l(l+1)}\varpi = \frac{2}{l(l+1)}(\Omega - \omega), \tag{4}$$

$$v = \frac{e^{\lambda}}{r^2}[l(l+1) - 2], \tag{5}$$

$$j = e^{-(\lambda+\nu)/2}. \tag{6}$$

The function Φ_{lm} is related to a component of the metric perturbations, and U_{lm} to the function in eq.(1) as

$$\delta v = \frac{U_{lm}}{(p_0+\rho_0)re^{-\nu}}. \tag{7}$$

Equation (2) can be integrated as

$$U_{lm}(t,r) = e^{-im(\Omega-\chi)t}U_{lm}(0,r) - 4\pi(p_0+\rho_0)r^2 e^{-\nu} \times\left[\Phi_{lm}(t,r) - e^{-im(\Omega-\chi)t}\left\{\Phi_{lm}(0,r) - im\chi\int_0^t e^{im(\Omega-\chi)s}\Phi_{lm}(s,r)ds\right\}\right]. \tag{8}$$

Neglecting the metric perturbations, i.e., in the Cowling approximation, we have

$$U_{lm}(t,r) = e^{-im(\Omega-\chi)t}U_{lm}(0,r). \tag{9}$$

The time-evolution of Φ_{lm} is determined by solving eq.(3) with the source term $U_{lm}(t,r)$, which is given by eq.(8). The boundary conditions should be given at the center and the stellar surface. There is a regularity condition at the center, i.e. $\Phi_{lm} \propto r^{l-1}$. The other condition comes from matching with an exterior solution. Outside the star, eq.(3) is reduced to

$$(r^4\Phi'_{lm})' - r^4 v e^{\lambda}\Phi_{lm} = 0. \tag{10}$$

The exterior solution, which is static and spatial regular at infinity, is given by the following power series

$$\Phi_{lm} = \left(\frac{M}{r}\right)^{l+2} \sum_{n=0}^{\infty} a_n \left(\frac{M}{r}\right)^n, \tag{11}$$

where M is a stellar mass, and the expansion coefficients a_n satisfy a certain recursion relation. The interior solution should be matched smoothly to the exterior one, so that Φ'_{lm}/Φ_{lm} should be continuous at the stellar surface $r = R$.

The initial data, $U_{lm}(0,r)$ and $\Phi_{lm}(0,r)$ seem to be arbitrary. They are however constrained by considering the higher order equations with respect to the rotation. This is because our perturbation scheme is degenerate one [12, 13], so that the lower radial functions are determined by solving the next order equations.

The pressure $\delta p_{l\pm1m}$ and density perturbations $\delta\rho_{l\pm1m}$, relevant to the axial functions with (l,m) are expressed by

$$4\pi\delta p_{l\pm1m} = 2\pi(p_0+\rho_0)(H_{0,l\pm1m} + 2T_\pm \Omega r^2 e^{-\nu}\Phi_{lm}) + 2S_\pm \varpi U_{lm}, \tag{12}$$

$$4\pi\delta\rho_{l\pm1m} = -4\pi\frac{\rho_0'}{\nu'}(H_{0,l\pm1m} + 2T_\pm\Omega r^2 e^{-\nu}\Phi_{lm}) - 4S_\pm\frac{e^{-\nu/2}}{\nu'}\left(e^{\nu/2}\varpi U_{lm}\right)' + 2T_\pm\frac{e^\nu}{\nu' r^2}(\varpi r^2 e^{-\nu})' U_{lm}, \tag{13}$$

where

$$S_\pm = \frac{2l+1\mp1}{2l+1\pm1}Q_\pm, \quad T_\pm = \pm\frac{2l+1\mp1}{2}Q_\pm, \tag{14}$$

$$Q_+ = \sqrt{\frac{(l+1)^2-m^2}{(2l+1)(2l+3)}}, \quad Q_- = \sqrt{\frac{l^2-m^2}{(2l-1)(2l+1)}}. \tag{15}$$

If the perturbations obey the same one-parameter equation of state as in the equilibrium, then we have

$$\delta p_{l\pm1m} = \frac{p_0'}{\rho_0'}\delta\rho_{l\pm1m}. \tag{16}$$

Substituting eqs.(12) and (13) into eq.(16), the terms with the metric functions, Φ_{lm} and $H_{0,l\pm1m}$ are cancelled out. In this way, we have the same relation as in the Cowling approximation. As shown in Ref. [14], these two equations are not simultaneously satisfied for a single function U_{lm} unless $m = \pm l$, in which one condition is trivial, $\delta p_{l-1m} = \delta\rho_{l-1m} = 0$ due to $Q_- = 0$. The other condition for $m = \pm l$ can be integrated with respect to r. We have

$$U_{lm} = (\rho_0 + p_0) r^{l+1} e^{-\nu} (\varpi e^{-\nu})^{(l-1)/2}. \tag{17}$$

In general, there is a time-dependent scale-factor in eq.(17). As the initial data the solution (17) is used with a suitable normalization constant, and the subsequent evolution is calculated by the coupled system, eqs.(3) and (8). We here neglected higher order rotational terms, so that the form of the velocity field is valid only for the early stage. The radial function will deviate from eq.(17) by the accumulated effects originated from higher order rotational terms like $\sim (\Omega t)^n$. When such a deviation becomes clear, higher order rotational terms in eqs.(12) and (13) should be included consistently.

3 Numerical results

3.1 sinusoidal oscillation

We use $n = 1$ polytropic equation of state to construct equilibrium stellar models. Three representative models with relativistic factor $M/R = 0.01$, 0.1 and 0.25 are adopted. We carry out the time-evolution of the perturbation on the equilibrium models, assuming that the initial function is given by (17) with $l = m = 2$. Our numerical calculation is limited to quadrupole only. The time-reversal symmetry is

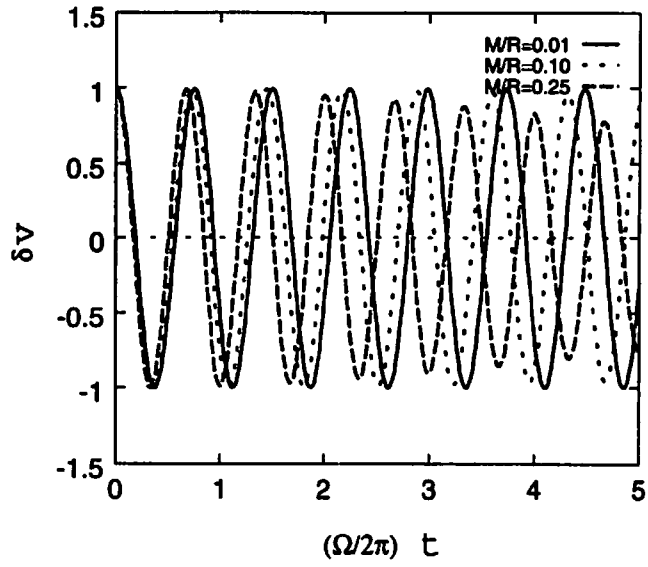


Figure 1: The amplitude of the velocity perturbation at the surface is shown as a function of time. The solid, dotted and dashed lines represent different stellar models as shown in the panel.

Table 1: Estimated frequencies for the time evolution of $l = m = 2$ purely axial displacement on stellar models with $n = 1$.

M/R	σ/Ω	σ_R/Ω	σ_0/Ω
0.01	1.34	1.34	1.36
0.10	1.40	1.37	1.52
0.25	1.51	1.46	1.80

checked as the accuracy of the numerical code. That is, when the function $\Phi_{lm}(t, r)$ is calculated from the initial data $\Phi_{lm}(0, r)$, then the calculation goes backward from t to 0, and the resulting function is compared with the original one. The accuracy of the numerical result is found to be good enough.

The time-evolution at the early phases is displayed in Fig.1. The amplitude of the velocity δv at surface is shown as a function of time. The sinusoidal time-dependence is clear for all the cases. The approximate frequency σ is numerically determined from each curve. The calculated values are listed in Table 1.

From the general argument [15, 16], the possible range of the frequency is limited as

$$\sigma_R \leq \sigma \leq \sigma_0, \tag{18}$$

where

$$\sigma_0 = \left(1 - \frac{2}{l(l+1)}\frac{\varpi_0}{\Omega}\right) m\Omega, \tag{19}$$

$$\sigma_R = \left(1 - \frac{2}{l(l+1)}\frac{\varpi_R}{\Omega}\right) m\Omega. \tag{20}$$

These values, σ_0 and σ_R respectively correspond to the classical r-mode frequency measured at the center and at the surface. Both are reduced to $(1 - 2/(l(l+1)))m\Omega$, when the frame-dragging is neglected, i.e.

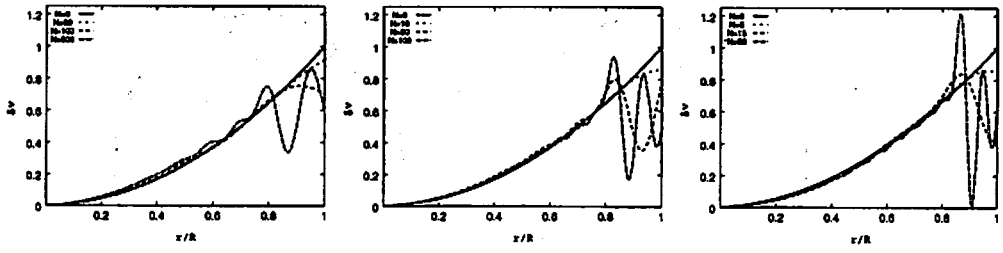


Figure 2: The velocity perturbation as a function of radius. The function after N-oscillation periods is compared with the initial one ($N = 0$). From left to right panels, relativistic factor is $M/R = 0.01$, 0.1 and 0.25.

$\varpi = \Omega$. The frequency for $l = m = 2$ is given by $(4/3)\Omega$. The frequency of the nearly Newtonian case $M/R = 0.01$ is close to the value. As the system becomes more relativistic, the frequency slightly shifts to higher value. The frequency is almost approximated as $\sigma \sim \sigma_R$, but not exactly the same. In addition to the sinusoidal oscillation, the gradual decrease of the amplitude can be seen for the relativistic case $M/R = 0.25$ in Fig.1. This indicates that the approximation as the form $\delta v_0(r)e^{-i\sigma t}$ is not good for the relativistic case. We will show how the function is by degrees modified in the time-scale longer than the oscillation period.

3.2 long-term behavior

In Figs.2-4, the velocity function δv at $t = (2\pi/\sigma)N$ is compared with the initial one ($N = 0$). The function δv after N-oscillation periods is shown as a function of radius. The long-term transition is clear in all the cases, although the time-scale depends on the relativistic factor M/R. The velocity field gradually deviates from the initial one, and some nodes in the radial direction appear after a lot of oscillations. In particular, the deviation is much larger at the outer region, whereas the function is the same in the inner region, $r/R < 0.8$. The duration, in which the velocity field can be approximated as the initial form, is estimated as $t \leq t_d = 2\pi/|\sigma_0 - \sigma_R| \propto \Delta\varpi^{-1}$. This definition of the time-scale is subtle, since the precise value significantly depends on the accuracy of the approximation. The value t_d should be regarded as a practical estimate. The radial function is not exactly the same even after one oscillation. But the difference can not just be seen. The accumulated change will become obvious around the time t_d. This also indicates that our approximation will break down after t_d, that is, the velocity field would not be expressed by the purely axial function. As the star becomes more relativistic, the dragging effect significantly depends on the position, so that the difference $\Delta\varpi$ of the dragging increases. Therefore, this effect considerably shortens the duration.

In Fig.3, we show the velocity field distribution, δv, on an equatorial plane by a contour map. From the symmetry, only δv_z component is relevant, and the others are zero. Four sectors due to $m = 2$ mode are by turns positive and negative in the azimuthal direction. The function after 50 oscillation periods in the model $M/R = 0.1$ is compared with the initial one. As the time goes, four sectors are slightly deformed, and the maximum points move in the radial direction. The complicated structure appears around the outer part, $r/R > 0.8$, whereas the initial structure is preserved in the inner part. The contour map at longer time-scale becomes more complicated and omitted here. These features can be inferred from the corresponding the radial function shown in Fig.2.

4 Discussion

In this paper, we have carried out the time-evolution of the purely axial displacement within the first-order slow-rotation formalism. The initial velocity perturbation is set up by solving an isentropic relation between the density and pressure perturbations. The initial data and the subsequent evolution would be

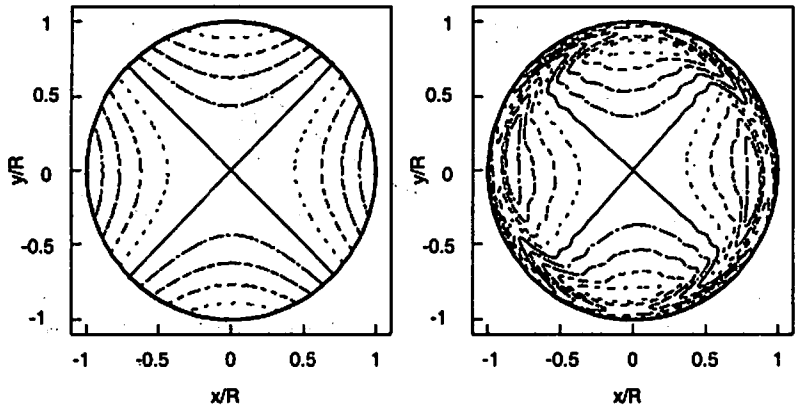


Figure 3: Contour map of δv_x on an equatorial plane for the model $M/R = 0.1$. Figures 5-a and 5-b correspond to the initial state and evolved state after 50-oscillation periods. The solids lines represent $\delta v_x = 0$, and divide four regions which are by turns positive and negative along the azimuthal direction.

different from those in the non-isentropic case. Indeed, there is no reason to restrict to $m = \pm l$. The short-term behavior of the evolution is well described by the sinusoidal oscillation with a certain frequency. The oscillation frequency is approximately determined by the classical r-mode frequency whose angular velocity is replaced by the value measured by the zero angular momentum observer at the surface. Note that the approximate frequency significantly depends on the initial data derived by the isentropic relation, that is, the oscillation is dominant near the surface. The functional behavior of the initial data singles out the oscillation frequency from a certain range of possible values. This consideration was lacking in the previous study of the first-order formalism [15, 16, 17].

In the long-term behavior of the evolution, however, the functional form gradually deviates from the initial one. This corresponds to no normal mode solution in such a restricted form except for the Newtonian star, as shown recently[10]. The initial data is not a pure state, but a mixed state of normal modes. The data are expected to be dominated by the counterpart of the Newtonian r-mode solution. We followed the time-evolution of the mixed state. The deviation due to the differential frame-dragging becomes obvious around t_d. This time-scale should be compared with others. For example, the second-order rotation changes the frequency as $\sigma(1 + \alpha(\Omega/\Omega_K)^2)$, where Ω_K is the maximum angular velocity of the star and α is some dimensionless value. The sinusoidal time-dependence is for the early phase approximated as $\exp\{-i\sigma t(1 + \alpha(\Omega/\Omega_K)^2)\} = \exp(-i\sigma t)\{1 - i\sigma t\alpha(\Omega/\Omega_K)^2 + \cdots\}$. The second order correction becomes obvious at $t > t_2 = 1/|\sigma\alpha(\Omega/\Omega_K)^2|$. In slowly rotating star, this time-scale t_2 is longer than the time-scale t_d associated with the frame-dragging effect. For the typical relativistic star, we have a crude estimate of the range $t_d < t_2$ as $\Omega < 0.3\Omega_K$. In other words, the frame-dragging effect can be neglected in rapidly rotating stars. This may be related to the fact in the recent numerical work that the normal calculation is easier in the rapidly rotating stars. The effects of gravitational radiation and the back reaction are relevant at a longer time-scale. Newtonian estimate indicates that it is of order $(\Omega/\Omega_K)^{2m+2}$ for $m \geq 2$. The time-scale associated with this effect is longer than t_d, and therefore the frame-dragging effect is in general important to consider the back reaction instability. Therefore, correct relativistic treatment may change the evolution of the radiation driven instability from the model based on the Newtonian r-mode solution.

References

[1] N. Andersson, ApJ **502**, 708, (1998).

[2] J. L. Friedman, and S. M. Morsink, ApJ **502**, 714 (1998).

[3] J. L. Friedman, and K. H. Lockitch, Prog. Theor. Phys. Suppl. **136**, 121 (1999).

[4] L. Rezzolla, F. K. Lamb, and S. L. Shapiro, ApJ **531**, L139 (2000).

[5] W. C. G. Ho, and D. Lai **astro-ph/9912296**.

[6] L. Bildsten, and G. Ushomirsky, ApJ **529**, L33 (2000).

[7] S. Yoshida, and U. Lee, **astro-ph/0006107**.

[8] L. Lindblom, B. J. Owen, and S. M. Morsink, Phys. Rev. Lett. **80**, 4843 (1998).

[9] B. J. Owen, L. Lindblom, C. Cutler, B. F. Schutz, A. Vecchio, and N. Andersson, Phys. Rev. D. **58**, 084020 (1998).

[10] K. H. Lockitch, N. Andersson, and J. L. Friedman (2000), preprint **gr-qc/0009019**.

[11] N. Stergioulas, and J. A. Font, preprint **gr-qc/0007086** .

[12] Y. Kojima, and M. Hosonuma, Phys. Rev. D. **62**, 044006 (2000).

[13] Y. Kojima, Prog. Theor. Phys. Suppl. **136**, 135 (1999).

[14] Y. Kojima, and M. Hosonuma, ApJ **520**, 788 (1999).

[15] Y. Kojima, Prog. Theor. Phys. Suppl. **128**, 251 (1997).

[16] Y. Kojima, Mon. Not. R. Astron. Soc. **293**, 49 (1998).

[17] H. R. Beyer, and K. D. Kokkotas, Mon. Not. R. Astron. Soc. **308**, 745 (1999).

R-mode oscillations of differentially and rapidly rotating Newtonian polytropic stars

Shigeyuki Karino[a] [1], Yoshiharu Eriguchi[a], and Shin'ichirou Yoshida[b]

[a] *Department of Earth Science and Astronomy, Graduate School of Arts and Sciences, University of Tokyo, Komaba, Meguro, Tokyo 153-8902, Japan*
[b] *SISSA, Via Beirut 2-4, 34013 Trieste, Italy*

Abstract

For the analysis of the r-mode oscillation of hot young neutron stars, it is necessary to consider the effect of *differential rotation*, because viscosity is not strong enough for differentially rotating young neutron stars to be lead to uniformly rotating configurations on a very short time scale after their birth. In this paper, we have developed a numerical scheme to solve r-mode oscillations of differentially rotating barotropic stars. This is the extended version of the method which was applied to compute r-mode oscillations of uniformly rotating stars. By using this new method, we have succeeded in obtaining eigenvalues and eigenfunctions of r-mode oscillations of differentially rotating stars. Our numerical results show that as the degree of differential rotation is increased, it becomes more difficult to solve r-mode oscillations for slightly deformed configurations from sphere compared to solving r-mode oscillations of considerably deformed stars. One reason for it seems that for slightly deformed stars corotation points appear near the surface region if the degree of differential rotation is strong enough. This explains that the perturbational approach of r-mode oscillations for *slowly rotating* stars in *general relativity* results in a singular eigenvalue problem.

1 Introduction

Recent intensive as well as extensive investigations of the r-mode instability have revealed its important role in neutron star physics ([1, 2]; for a review see e.g. [3]; for r-mode oscillations see [4, 5, 6]). At the early stage of those investigations, the r-mode oscillations were found to become significantly unstable for some range of the core temperature [7, 8, 9]. However, models used in those papers were too simplified to deduce a decisive conclusion for real neutron stars because one could only obtain r-mode oscillations of uniformly and slowly rotating Newtonian polytropic stars and apply the results to realistic neutron stars.

Real neutron stars are somewhat more complicated configurations. There are many factors to be included to get a definite answer about the effect of the r-mode instability on real neutron stars. For example, neutron stars are general relativistic objects, they are rapidly rotating at their birth, the physical states of the interior cannot be quantitatively approximated by polytropic relations, the surface of the neutron stars is not considered to be fluid and so on.

The rotation law is one of such factors. Although in almost all papers about the r-mode instabilities neutron stars are assumed to rotate uniformly, newly born neutron stars whose ages are less than 1 year should be considered to be rotating differentially. Since neutron stars are newly formed from collapse of massive stars or are recycled by accretion of matter onto old neutron stars on a dynamical time scale without enough time to redistribute the angular momentum due to viscosity, those neutron stars cannot rotate uniformly. Thus we need to study the effect of differential rotation of newly born or recycled neutron stars on the r-mode oscillations. Moreover, nonlinear growth of r-mode instabilities will lead to differential rotation within the star [10]. The presence of magnetic field may cause interaction between the magnetic field and the rotation and will establish differential rotation [11].

[1] E-mail:karino@valis.c.u-tokyo.ac.jp

In our present investigation, by extending the numerical method which we developed to compute r-mode oscillations of uniformly and rapidly rotating polytropes [12, 13], we have succeeded in showing that r-mode oscillations exist for differentially rotating polytropic stars (from a different point of view, the r-mode oscillations of configurations with *slightly* differential rotation are treated in [14, 15]). From our numerical results for differentially rotating stars, we have found that, if the degree of differential rotation is small, r-mode oscillations of the ordinary type can be solved. Here oscillations of the ordinary type mean oscillations whose eigenfrequencies distribute discretely. However, if the degree of differential rotation becomes large, no r-mode oscillations of the ordinary type seem to be allowed, because in the surface region of the star corotation points appear and the basic equation becomes singular there.

Since the purpose of this paper is to show the characteristic feature of the r-mode oscillation of differentially rotating stars, we only present a limited results of numerical computations, i.e. results for $N = 1$ polytropes and for the $m = 2$ modes, where N is the polytropic index and m is the azimuthal mode number of the perturbation.

2 Formulation of the problem

We will study the effect of differential rotation on the r-mode oscillations by solving the linearized equations of fluid motions of stars. The scheme employed in this paper is the extended version of that developed to analyze the r-mode oscillation of uniformly rotating Newtonian polytropes [12, 13]. Therefore, we will explain the scheme only briefly.

2.1 Assumptions for unperturbed stars

We construct axisymmetric equilibrium configurations of differentially rotating polytropic stars in Newtonian gravity by employing the SFNR method [16]. Here the polytropic relation is expressed by the following formula:

$$p = K\rho^{1+\frac{1}{N}}, \tag{1}$$

where p, ρ, K and N are the pressure, the density, the polytropic constant and the polytropic index, respectively.

As for the rotation law of differentially rotating unperturbed stars, the angular velocity, Ω, is assumed to be expressed as:

$$\Omega = \frac{\Omega_c A^2}{(R/R_{eq})^2 + A^2}, \tag{2}$$

where R is the distance from the rotation axis, R_{eq} is the equatorial radius of the star and Ω_c is the central angular velocity. The quantity A is a parameter which represents the degree of differential rotation. The rotation becomes more differentially as A becomes small. On the other hand, when we consider the limit of $A \to \infty$, the stellar rotation tends to uniform rotation. For instance, when $A = 1$, the central angular velocity is twice as large as the surface angular velocity on the equatorial plane.

2.2 Linearized basic equations for perturbed stars

The scheme to handle the r-mode oscillation of differentially rotating polytropes is the extended version of that to solve the perturbed fluid equations for uniformly rotating polytropes by the Newton-Raphson iteration scheme [12, 13].

Basic equations consist of 1) the linearized continuity equation, 2) the linearized compatibility equation between the $r-$component and the $\varphi-$ component of the equation of motion, 3) the linearized compatibility equation between the $r-$component and the $\theta-$component of the equation of motion, 4) the linearized $\varphi-$component of the equation of motion, and 5) the linearized gravitational potential equation. Here, the spherical coordinates (r, θ, φ) are used.

Concerning the perturbed quantities, we assume the harmonic expansion of the physical quantities as

$$\delta f(r, \theta, \varphi, t) = \sum_m \exp(-i\sigma t + im\varphi) f_m(r, \theta), \tag{3}$$

where f denotes a certain physical quantity, δ is the Euler perturbation of the corresponding quantity, σ is the frequency of the oscillation and m is the mode number.

The perturbed fluid equations are expressed as follows:

$$\rho_0\frac{\partial\delta u_r}{\partial r}+\Big(\frac{2\rho_0}{r}+\frac{\partial\rho_0}{\partial r}\Big)\delta u_r+\frac{\rho_0}{r}\frac{\partial\delta v_\theta}{\partial\theta}+\frac{1}{r}\Big(\rho_0\cot\theta+\frac{\partial\rho_0}{\partial\theta}\Big)\delta v_\theta+\frac{m\rho_0}{r\sin\theta}\delta w_\varphi=(\sigma-m\Omega)\delta\rho, \tag{4}$$

$$\begin{aligned}&\frac{r\sin^2\theta}{m}\Big(2\Omega+r\frac{\partial\Omega}{\partial r}\Big)\frac{\partial\delta u_r}{\partial r}+\frac{1}{m}\Big[m(m\Omega-\sigma)+\sin^2\theta\Big(2\Omega+4r\frac{\partial\Omega}{\partial r}+r^2\frac{\partial^2\Omega}{\partial r^2}\Big)\Big]\delta u_r\\&+\frac{r\sin^2\theta}{m}\Big(2\Omega\cot\theta+\frac{\partial\Omega}{\partial\theta}\Big)\frac{\partial\delta v_\theta}{\partial r}+\frac{r\sin^2\theta}{m}\Big[2\Omega\cot\theta+\frac{\partial\Omega}{\partial\theta}+2r\cot\theta\frac{\partial\Omega}{\partial r}+r\frac{\partial^2\Omega}{\partial r\partial\theta}\Big]\delta v_\theta\\&+\frac{r\sin\theta}{m}(m\Omega-\sigma)\frac{\partial\delta w_\varphi}{\partial r}+\frac{\sin\theta}{m}\Big[(m\Omega-\sigma)+2m\Omega+mr\frac{\partial\Omega}{\partial r}\Big]\delta w_\varphi=0,\end{aligned} \tag{5}$$

$$\begin{aligned}&(\sigma-m\Omega)\frac{\partial\delta u_r}{\partial\theta}-m\frac{\partial\Omega}{\partial\theta}+r(m\Omega-\sigma)\frac{\partial\delta v_\theta}{\partial r}+\Big[(m\Omega-\sigma)+mr\frac{\partial\Omega}{\partial r}\Big]\delta v_\theta\\&+2r\cos\theta\Omega\frac{\partial\delta w_\varphi}{\partial r}-2\sin\theta\Omega\frac{\partial\delta w_\varphi}{\partial\theta}+2\Big(r\cos\theta\frac{\partial\Omega}{\partial r}-\sin\theta\frac{\partial\Omega}{\partial\theta}\Big)\delta w_\varphi=0,\end{aligned} \tag{6}$$

$$\begin{aligned}&\Big(1+\frac{1}{N}\Big)\rho_0^{\frac{1}{N}-1}\frac{m}{r\sin\theta}\delta\rho++\frac{m}{r\sin\theta}\delta\phi+\sin\theta\Big(2\Omega+r\frac{\partial\Omega}{\partial r}\Big)\delta u_r\\&+\Big(2\Omega\cos\theta+\sin\theta\frac{\partial\Omega}{\partial\theta}\Big)\delta v_\theta+(m\Omega\sigma)\delta w_\varphi=0.\end{aligned} \tag{7}$$

Here, quantities with subscript $_0$ denote those of the equilibrium states, δu_r, δv_θ, δw_φ, $\delta\rho$ and δo are the Euler perturbations of the $r-$, $\theta-$ and $\varphi-$ components of the velocity, the density and the gravitational potential, respectively.

In these equations, we have made use of the following adiabatic perturbation between the pressure and the density:

$$\frac{\delta p}{p}=(1+\frac{1}{N})\frac{\delta\rho}{\rho}, \tag{8}$$

where δp is the Euler perturbation of the pressure.

2.3 Boundary conditions

Since we treat infinitesimal oscillations around equilibrium configurations, all the perturbed quantities must behave regularly throughout the whole space. In particular, the Euler change of the density and the Euler change of the gravitational potential must be regular function of the position.

As for the perturbed flow velocity, the boundary condition can be expressed as

$$\frac{\partial\rho_0}{\partial r}\Big(\delta u_r-\frac{1}{r_s}\frac{dr_s}{d\theta}\delta v_\theta\Big)+(m\Omega-\sigma)\delta\rho=0. \tag{9}$$

Here $r_s(\theta)$ denotes the surface of the equilibrium star. This condition comes from requirement that the fluid element on the surface is displaced to the new perturbed surface.

3 Numerical results

We can obtain eigenvalues and eigenfunctions by solving numerically the basic equations mentioned above by the Newton-Raphson iteration scheme, together with the boundary condition. In actual computations, we use 44 mesh points in $r-$direction and 11 mesh points in $\theta-$direction.

In Figure 1, the eigenvalues of equilibrium sequences with different values of the parameter A are plotted against the normalized central angular velocity for $m=2$ mode of $N=1$ polytropes. It is

clearly seen that the eigenvalues decrease as the central angular velocity increases. In this figure, for the sake of simplicity, the eigenvalues are normalized by using the central angular velocity. This choice of the normalization constant results in decreasing of the eigenfrequencies for the smaller value of A or for the larger degree of differential rotation. It should be noted that for smaller value of A the equilibrium sequences exist for larger values of the central angular velocity.

Figure 2 shows distributions of the perturbed velocity component in θ–direction on the equatorial plane of the $m = 2$ oscillation of $N = 1$ polytropes for three different values of A. The horizontal axis is the normalized stellar radius. The equilibrium stars are rotating rather rapidly and $r_p = 0.7$ where r_p is the axis ratio which is define by

$$r_p = r_{ax}/r_{eq}. \tag{10}$$

Here r_{ax} and r_{eq} are the stellar radius along the rotational axis and that on the equatorial plane, respectively.

Three curves in Figure 2 correspond to equilibrium models of rigid rotation, differential rotation with the parameter $A = 1.0$ and extremely differential rotation with $A = 0.6$, respectively. From this figure, we can see that if we take the differential rotation into consideration, amplitudes of oscillations becomes large only in the narrow layer at the surface of the star for larger degree of differential rotation. In other words, for differentially rotating models, if the degree of differential rotation is large, only the fluid near the surface region can oscillate appreciably but the bulk of the star almost remains at its original position. This behavior can be seen not only in the equatorial plane but also in all directions.

For differentially rotating stars, we need to pay attention to emergence of corotation points at some places in the stellar interior because appearance of corotation points corresponds to a singular behavior of the basic equations (see e.g. [17]). Here corotation points are defined as the positions where the following condition is satisfied:

$$\sigma \equiv \sigma - m\Omega = 0. \tag{11}$$

The condition of appearance of corotation points can be expressed by a simple inequality for differentially rotating stars. For the rotation law chosen in this paper (Eq. (2)), the quantity σ is rewritten as

$$\sigma = \Omega_c\left(\frac{\sigma}{\Omega_c} - \frac{mA^2}{(R/R_{eq})^2 + A^2}\right). \tag{12}$$

If this quantity σ becomes zero within the star, we cannot solve the eigenvalue problem because the matrix which governs the problem becomes singular as mentioned above. From Eq. (12), the quantity σ takes its minimum value on the rotational axis and its maximum value on the equatorial surface. Hence,

$$\Omega_c\left(\frac{\sigma}{\Omega_c} - m\right) \le \sigma - m\Omega \le \Omega_c\left(\frac{\sigma}{\Omega_c} - \frac{mA^2}{1 + A^2}\right). \tag{13}$$

For $m = 2$ modes of $N = 1$ polytropes, as seen from Fig. 1, the lower limit is always negative in the star. Thus the condition for which the quantity $(\sigma - m\Omega)$ becomes zero can be obtained by requiring that the upper limit of $(\sigma - m\Omega)$ must be positive inside the star. As a result, the condition that corotation points exist in the star can be written as

$$\frac{\sigma}{\Omega_c} \ge \frac{2A^2}{1 + A^2}, \tag{14}$$

where $m = 2$ is substituted because we only discuss $m = 2$ oscillation modes in this paper. In order to find out whether the corotation points appear in the star or not, we need to check this condition for oscillations of differentially rotating stars.

In Figure 3, values of eigenfrequencies which satisfy the *equality* of the above condition are plotted against the value of $\log A$. We will call this curve as a critical curve. If the eigenfrequency of a certain star is located above the critical curve, a corotation point appears in the star. On the other hand, if the eigenfrequency is below the critical curve, no corotation point appears. In this figure, eigenfrequencies of two equilibrium sequences, sequences with $r_p = 0.95$ and with $r_p = 0.70$, are also plotted. As seen from this figure, the eigenvalues locate well below the critical curve when the degree of differential rotation is small. When we decrease the value of A, i.e. increase the degree of differential rotation, however, the eigenfrequencies approach to the critical curve.

We show the ratio of the eigenfrequency to the critical value in Figure 4. From this figure, it is clear that eigenfrequencies are approaching toward the critical curve monotonically as the value of A is decreased. For nearly spherical configurations, i.e. the sequence with $r_p = 0.95$, the eigenfrequency of the model with $A = 0.615$ reaches 97% of the critical value, while the eigenfrequency of the sequence with $A = 0.585$ reaches 93% of the critical value for deformed configurations, i.e. the sequence with $r_p = 0.70$. For slowly rotating stars, sooner or later, the eigenvalues seem to cross the critical curve, though it is very difficult to get converged solutions of oscillations by using our numerical code. This is not related to the ability of the code because r-mode oscillations can be easily obtained for nearly spherical configurations with larger values of A.

4 Discussion and conclusions

4.1 R-mode oscillation of differentially rotating stars and the singular nature of the r-mode oscillation of the perturbational approach in GR

As discussed in Introduction, r-mode oscillations should be analyzed in the framework of general relativity. Although r-mode oscillations in general relativity have been investigated by several authors, there appear some features which do not appear in *uniformly rotating* Newtonian stars [18, 19, 20, 21, 22]. One of them is related to the strange nature of the eigenvalue problem if *barotropic* or *isentropic* configurations are treated. Some authors conclude the eigenvalue problem becomes singular [18, 19] and other authors claim that the system results in an overdetermined state [21, 22].

The strange nature of the eigenvalue problem for barotropic stars in general relativity arises from appearance of the following term in the coefficient of the highest order derivative of the master equation if the slow rotation approximation is employed:

$$\sigma \equiv \sigma - m(\Omega - \omega(r,\theta)), \tag{15}$$

where $\omega(r,\theta)$ is the dragging potential of the gravitational field of rotating general relativistic configurations.

If we introduce the following *effective* angular velocity

$$\Omega_{\rm eff} \equiv \Omega - \omega(r,\theta), \tag{16}$$

then we can write the quantity $\sigma_{\rm eff}$ as follows:

$$\sigma_{\rm eff} = \sigma - m\Omega_{\rm eff}. \tag{17}$$

Thus the peculiar nature of the eigenvalue problem in general relativistic slow rotation approximation is essentially the same phenomena of appearance of corotation points for differentially rotating stars which has been discussed in the previous section.

As our present numerical results show, if the corotation point is about to appear in the star, it becomes very difficult to solve ordinary type r-mode oscillations, i.e. r-mode oscillations with discrete eigenfrequencies and this seems consistent with the general relativistic analysis under the slow rotation approximation. Moreover, the eigenfunctions for such configurations behave as delta functions as seen from Fig. 2.

We may need to explain the reason why some authors [21, 22] could solve r-mode oscillations for isentropic stars in general relativity. As our numerical results show, weakly and mildly differentially rotating configurations do not suffer from appearance of corotation points. Thus for weakly and/or mildly relativistic configurations, there can be the ordinary type r-mode oscillations, although they do not have exactly the same characters as those of Newtonian configurations [21, 22].

From our "numerical experimental data" in this paper, there is a possibility of obtaining ordinary type r-mode oscillations for *rapidly rotating general relativistic* configurations, if one will be able to develop a scheme to handle linearized equations of the general relativistic fluid equations as well as the linearized Einstein equation for rapidly rotating stars.

4.2 Conclusions

We have succeeded in developing a new scheme which can be used to solve r-mode oscillations even for extremely young stars which rotate differentially. This is a necessary development in order to study the scenario that neutron stars have lost their angular momentum via gravitational waves just after their birth by the r-mode instability.

By analyzing the r-mode oscillations of differentially rotating polytropes, we have shown that the eigenvalue problem becomes singular for configurations with large degree of differential rotation.

References

[1] N. Andersson, Astrophys. J. **502** (1998) 708.

[2] J. L. Friedman and S. M. Morsink, Astrophys. J. **502**(1998) 714 .

[3] J. L. Friedman and K. H. Lockitch, Prog. Theor. Phys. Suppl. **136** (1999) 121.

[4] J. Papaloizou and J. E. Pringle, Mon. Not. R. Astr. Soc. **182**(1978) 423.

[5] J. Provost, G. Berthomieu and A. Rocca, Astron. Astrophys. **94** (1981) 126.

[6] H. Saio, Astrophys. J. **256** (1982) 717.

[7] L. Lindblom, B. J. Owen and S. M. Morsink, Phys. Rev. Lett. **80** (1998) 4843.

[8] B. J. Owen, L. Lindblom, C. Cutler, B. F. Schutz, A. Vecchio and N. Andersson, Phys. Rev. D**58** (1998) 084020.

[9] N. Andersson, K. Kokkotas and B. F. Schutz, Astrophys. J. **510** (1999) 846.

[10] L. Lindblom, J. E. Tohline and M. Vallisneri, preprint (astro-ph/0010653).

[11] L. Rezzola, F. K. Lamb and S. L. Shapiro, Astrophys. J. **531** (2000) L139.

[12] S'i. Yoshida, S. Karino, S. Yoshida and Y. Eriguchi, Mon. Not. R. Astr. Soc. **316** (2000) L1.

[13] S. Karino, S'i. Yoshida, S. Yoshida and Y. Eriguchi, Phys. Rev. D **62** (2000) 084012.

[14] C. L. Wolff, Astrophys. J. **502** (1998) 961.

[15] C. L. Wolff, Astrophys. J. **531** (2000) 591.

[16] Y. Eriguchi and E. Müller, Astron. Astrophys. **146** (1985) 260.

[17] J. R. Ipser and L. Lindblom, Astrophys. J. **355** (1990) 226.

[18] Y. Kojima, Mon. Not. R. Astr. Soc. **293** (1998) 49.

[19] Y. Kojima and M. Hosonuma, Astrophys. J. **520** (1999) 788.

[20] H. R. Beyer and K. D. Kokkotas, Mon. Not. R. Astr. Soc. **308** (1999) 745.

[21] K. H. Lockitch, preprint (gr-qc/9909029).

[22] K. H. Lockitch, N. Andersson and J. L. Friedman, Phys. Rev. D , (2000) in press: preprint (gr-qc/0008019).

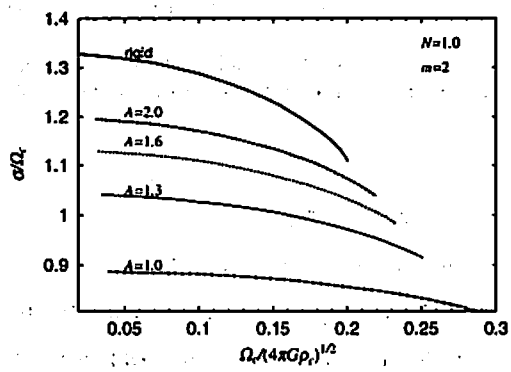


Figure 1: Eigenvalues of $m = 2$ mode oscillations for $N = 1$ polytropic equilibrium sequences with different values of A are plotted against the central angular velocity. The value of A is attached to each curve of an equilibrium sequence. The eigenfrequency and the central angular velocity are normalized by using Ω_c and $(4\pi G\rho_c)^{1/2}$, respectively. Here ρ_c is the central density.

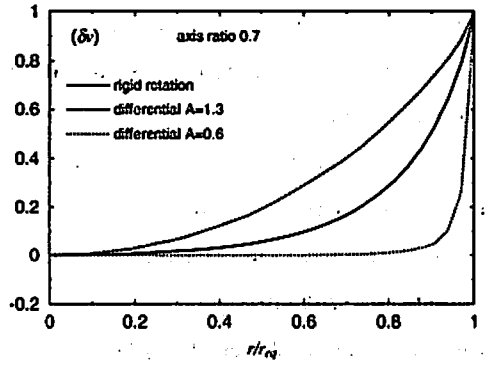


Figure 2: The θ−component of the perturbed velocity on the stellar equatorial plane is plotted against the distance from the center for $N = 1$ polytropes and the $m = 2$ mode. Three curves show distributions of the θ−component of the fluid velocity with different degree of differential rotation: the solid curve for a rigidly rotating configuration, dashed curve for a configuration with $A = 1.3$ and the dotted curve for a configuration with $A = 0.6$. The axis ratio for three configurations is chosen to be the same, i.e. $r_p = 0.7$. It should be noted that the large amplitude oscillations are confined to a narrow surface region for the highly differentially rotating model.

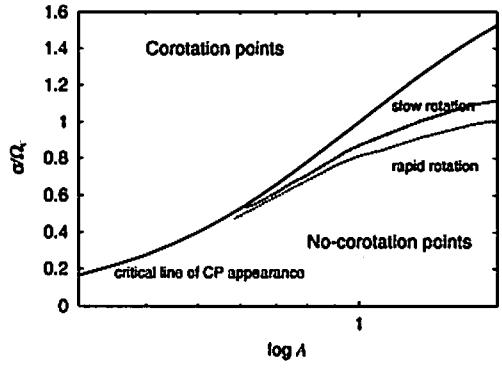


Figure 3: The critical curve for appearance of corotation points and eigenfrequencies of two equilibrium sequences are plotted against the value of $\log A$. Sold curve denotes the critical curve. Dashed and dotted curves show the eigenfrequencies for the equilibrium sequences with $r_p = 0.95$ and 0.70, respectively. The terminal points of eigenvalue curves correspond to models with $A = 0.615$ for the $r_p = 0.95$ sequence, and $A = 0.585$ for the $r_p = 0.70$ sequence, respectively.

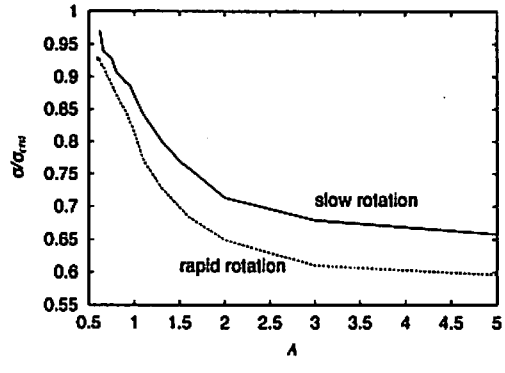


Figure 4: The ratio of the eigenfrequency to the critical value for the appearance of corotation points is plotted against the value of A. The maximum value of this ratio reaches 97% of the critical value for the appearance of corotation points along the $r_p = 0.95$ sequence.