ISSN 1343-9499

TOHOKU
MATHEMATICAL
PUBLICATIONS

Number 35

Su Buqing Memorial Lectures

No. 1

On the occasion of the Centennial of the Faculty of Science

Edited by

Reiko Miyaoka

September 2011

This series aims to publish material related to the activities of the Mathematical Institute of Tohoku University. This may include:

1. Theses submitted to the Institute by grantees of the degree of Doctor of Science.
2. Proceedings of symposia as well as lecture notes of the Institute.

A primary advantage of the series lies in quick and timely publication. Consequently, some of the material published here may very likely appear elsewhere in final form.

Tohoku Mathematical Publications

Mathematical Institute
Tohoku University
Sendai 980-8578, Japan

ISSN 1343-9499

TOHOKU
MATHEMATICAL
PUBLICATIONS

Number 35

Su Buqing Memorial Lectures

No. 1

On the occasion of the Centennial of the Faculty of Science

Edited by

Reiko Miyaoka

September 2011

Su Buqing Memorial Lectures

No. 1

held at

Tohoku University and Fudan University

2008–2011

Edited by

Reiko Miyaoka

Preface

COOPERATION AGREEMENT BETWEEN
MATHEMATICAL INSTITUTE, GRADUATE SCHOOL OF SCIENCE,
TOHOKU UNIVERSITY, JAPAN
AND
SCHOOL OF MATHEMATICAL SCIENCES,
FUDAN UNIVERSITY, CHINA

was signed by Professor Akihiko Yukie, Chair of Mathematical Institute of Tohoku University, and Professor Quanshui Wu, Dean of School of Mathematics of Fudan University, in November 2009. The background and details of activities are given in the articles by Professor Motoko Kotani and Professor Emeritus Katsuei Kenmotsu, and in the record of the activities in this volume, which features lectures held during 2008-2010 under this agreement.

We are grateful to all the lecturers for their splendid lectures and their contributions to this issue. We also thank all those, especially doctoral students, who took the notes, or did proof-reading. We sincerely thank Ms. Junko Bannai for her effective work in the preparation of the manuscript.

The Faculty of Science of Tohoku University celebrates its centennial in September 2011. On this occasion, the publication of these lecture notes in memory of Professor Su Buqing (蘇　歩青), a great mathematician and a leader of modern China, Doctor of Science, Tohoku Imperial University in 1931, is one of the most suitable and honorable events.

We intend to continue our activities for the development of mathematics on both sides through this exchange program.

September 2011

Reiko Miyaoka

Foreword

Tohoku University has been proud of maintaining "Research First" and "Open Door Policy" as the school principles since its foundation as the third imperial university following University Tokyo and Kyoto University. It has been known as the first university of Japan which admitted female students against the government in 1913. From its early stage, Tohoku University also opened its doors to international students, with two graduating in 1911. The great author Lu Xun (魯迅), a father of modern Chinese philosophy, studied at Tohoku University.

Department of Mathematics contributed to the history. The mathematician Chen Jiangong (陳 建功) enrolled in Tohoku University in 1920, and in 1929 became the first international student to obtain a doctorate in Japan. Su Buqing (蘇 歩青), who entered Department of Mathematics in 1924, not only continued to the graduate school upon graduation, but also began teaching as a lecturer at Tohoku University. He obtained a doctorate in sciences in 1931. Both contributed to establish modern mathematics in China. As of May 2010, there are around 1,500 international students enrolled at Tohoku University from 80 countries with 744 from China.

In 2007, on behalf of Graduate School of Science, Professor Reiko Miyaoka and I paid a visit to Fudan University together with Vice President Osamu Hashimoto of international relationship. It was to reconfirm our long history of friendship and discuss about a possible action to extend our research exchange and collaboration into a more intense and clear form. On that occasion, we visited the School of Mathematical Science to discuss about our academic programs and agreed to enter a new stage of relationship for the benefit not only to professors but to our students and made a contract. It says that we shall in turn invite professors of one of the two institutes to stay at the other's to give a series of lectures for the graduate students every year and publish the Su Buqing Lecture Note based on the lectures.

Since then, Professor Jiaxing Hong, Professor Quanshui Wu, Professor Yuaolong Xin and Professor Reiko Miyaoka have put their valuable effort to establish the new flow of exchange and accomplished the agreement.

The present lecture note is an outcome of our five years activities. We hope it will contribute to promote our collaboration on both research and education.

June 13, 2011

Motoko Kotani
Former Chair
Mathematical Institute
Graduate School of Science
Tohoku University

A BRIEF HISTORY OF THE MATHEMATICS EXCHANGES BETWEEN TOHOKU AND FUDAN UNIVERSITIES

Katsuei KENMOTSU

The exchange of mathematicians between Tohoku and Fudan Universities began when Prof. Su Buqing came to Sendai on December 6, 1955, to present a talk at the colloquium of the Mathematical Institute at Tohoku University. On the occasion of the visit, he wrote the following poem:

念載重来一夢中
髪糸凋尽冷秋風
仙台故舊幾人在
獨對向山斜日紅

1955 年 12 月 6 日来訪
母校萬感倶集書此記念
蘇　歩青

Coming here again over twenty years as in my dream.
Feeling autumn chilly as less and less my hair.
A few old friends in Sendai are still alive ?
Sitting alone and facing mountains to see red sunset.

Visiting Alma Mater, so many reflections and thoughts come out and write down to commemorate.

Su Buqing, December 6, 1955 (translated by Yuan-long Xin)

In fact, Prof. Su graduated from Tohoku Imperial University, which is now called Tohoku Universty, in 1927; he received the degree of Doctor of Science at the same University in March, 1931. One month later, he returned to China with a wife, Yoneko Matsumoto, and took a position at Zhejiang University. After that, he worked in China and became an outstanding leader within the mathematics community.

In August of 1981, when the International Symposium of Differential Geometry and Differential Equations was held at Fudan University in honor of Prof. Su's 80th birthday, six Japanese mathematicians were invited to give talks. Among them, the youngest was Seiki Nishikawa, who is now a leading figure at the Mathematical Institute of Tohoku University.

In April of 1983, Prof. Su, who was then the Honorable president of Fudan University, visited Japan again with Professors H. Hu and Y. Wang. He gave an invited talk at the annual meeting of the Mathematical Society of Japan held in Hiroshima, and he also delivered an address at the reception in honor of Prof. Su held at Sendai city hall.

In September of 1991, there was an international symposium on differential geometry in honor of Prof. Su's 90th birthday, at which seven Japanese mathematicians, including myself, were invited. At the reception of the symposium, it was an honor for me to give an address as the representative of the foreign participants. At that time, I met many young Chinese graduate students. Later, some of them came to Tohoku University as post doctoral fellows, invited by Japanese foundations.

As the scientific exchanges between Tohoku and Fudan Universities were extended to other faculties and increased in activity since 1992, both universities agreed to sign a treaty for cultural exchange. On April, 2001, the president of Tohoku University, Prof. Abe, and I visited Fudan University to sign this document.

In the field of mathematics, to promote increased, systematic exchange, both departments of mathematics participated in the exchange program in 2008; it now works very well.

I would like to thank my younger friends Yuan-long Xin of Fudan University and Seiki Nishikawa of Tohoku University for their help in preparing these notes.

May 12, 2011
Sendai, Japan

Record of the activities 2007–2011

1. Courtesy visit to the President of Fudan University, Wang Shenghong 王生洪 復旦大学学長 (6 Dec. 2007)

   Visitors: Osamu Hahimoto (Dean of the faculty of Science of Tohoku University)

   Motoko Kotani, Reiko Miyaoka: Professors (Mathematics)

   Masahiro Yamaguchi: Professor (Physics)

2. Meeting: Professor Jiaxing Hong, Dean of School of Mathematical Sciences, Professors Yuanlong Xin, M. Kotani and R. Miyaoka (7 Dec. 2007)

3. A visit to Fudan University (9–12 Sept. 2010) by R. Miyaoka (meeting with Professor Jiaxing Hong, and Professor Quanshui Wu, Dean of School of Mathematical Sciences.)

Exchange intensive lectures:

Lecturer	Title	Dates and Place
Jixiang Fu (Diff. Geom.)	Balanced Metric	23–25 Jul. 2008 Tohoku
Takashi Shioya (Diff. Geom.)	Geometric Analysis on Alexandrov spaces	3,5 Dec. 2008 Fudan
Yuanlong Xin (Diff. Geom.)	Topics in Minimal Submanifolds	7–9 Oct. 2009 Tohoku
Masayoshi Takeda (Probability)	Some formulae on additive functionals of symmetric Markov Processes	2–3 Nov. 2009 Fudan
Jiangang Ying (Probability)	Introduction to Markov Processes	24–25 Nov. 2010 Tohoku
Masanori Ishida (Alg. Geom.)	On the construction of toric varieties	23,25 Nov. 2010 Fudan

Events held in Tohoku University

Colloquium:

1. J. Fu, *Balanced metrics*, (22 July 2008)
2. Y.L. Xin, *Minimal subanifolds of high codimension*, (5 Oct. 2009)
3. J.G. Ying, *From Douglas integral to Feller measure*, (22 Nov. 2009)

Geometry seminar:

Y.L. Xin, *Curvature estimates for submanifolds with prescribed Gauss image and mean curvatrure*, (6 Oct. 2009)

Probability seminar:

Y. Yijun, *Complex Structure on Noncommutative 2-Tori*, (22 Nov. 2010)

Mini-workshop:

1. (with Prof. Fu), Speakers: A. Takatsu (DC), K. Funano (DC), Q.M. Cheng (Professor at Saga University), (25 July 2008)
2. (with Prof. Xin), Speakers (all DC): S. Kikuta, T. Omori, M. Tanaka, A. Takatsu, M. Watanabe, (7 Oct. 2009)

Events held in Fudan University

Colloquium:

1. M. Kotani, *A mathematical challenge to a new phase of materials science*, (6 Dec. 2007)
2. R. Miyaoka, *Hypersurface geometry and its applications*, (6 Dec. 2007)
3. T. Shioya, *Collapsing three-manifolds with a lower curvature bound*, (2 Dec. 2008)
4. H. Chihara, *Geometric analysis of Schroedinger maps*, (2 Dec.2008)
5. M. Takeda, *Dirichlet forms L^2 theory in symmetric Markov processes*, (1 Dec. 2009)
6. T. Yamazaki, *Higher dimensional class field theory of a product of curves*, (1 Dec. 2009)
7. M. Ishida, *Counting lattice points by the Riemann-Roch theorem*, (24 Nov. 2010)
8. M. Ishikawa, *Milnor fibrations and contact structures in low-dimensional topology*, (24 Nov. 2010)
9. R. Miyaoka, *Recent Development of hypersurface geometry*, (10 Sept. 2010)

CONTENTS

BALANCED METRICS

Jixiang FU

0 Introduction

A smooth Calabi-Yau threefold is a complex three dimensional manifold with finite fundamental group and trivial canonical bundle. If a Calabi-Yau threefold is Kähler, then according to Yau's celebrated solution [34] to the Calabi conjecture, there exists a unique Ricci-flat Kähler metric in each Kähler class of the threefold. Such metrics, known as Calabi-Yau metrics, are the bedrocks of geometric studies of Calabi-Yau threefolds and the superstring theory.

On the other hand, by the Clemens-Friedman construction, a large class of non-Kähler Calabi-Yau threefolds is obtained from Kähler Calabi-Yaus by blowing down rational curves and then smoothing the singularities. For example, the connected sum $\#_k(S^3 \times S^3)$ of k copies of $S^3 \times S^3$ for any $k \geq 2$ can be given a complex structure in this way. Based on this construction, Reid speculated that any two projective Calabi-Yau threefolds can be connected by a sequence of deformations, contractions and smoothing through non-Kähler Calabi-Yau threefolds. This speculation demonstrates the potential role of non-Kähler complex manifolds.

Therefore, it is important to construct canonical metrics on such non-Kähler complex manifolds. First, one should choose in general a good hermitian metric which is weaker than Kähler. One proposal is the balanced metric, which is a hermitian metric with the property that the $(n-1)$ power of its hermitian form is d-closed, where n is the complex dimension of the manifold. (Such a metric is also called a semi-Kähler metric in older references.)

In 2008, J. Li, S.-T. Yau and Fu [14] constructed balanced metrics on above mentioned non-Kähler Calabi-Yau threefolds, so especially on $\#_{k\geq 2}(S^3 \times S^3)$. Combining this result with some discussions in [7], they also proved that there exists no hermitian metric on $\#_{k\geq 2}(S^3 \times S^3)$ such that its hermitian form is $\partial\bar{\partial}$-closed.

On the other hand, such results are also needful in superstring theory. A candidate internal space in heterotic string is a compact three-dimensional hermitian manifold with trivial canonical bundle. The $N = 1$ supersymmetry requires that the given hermitian metric is a conformal balanced metric. More precisely, the holonomy of the hermitian metric with respect to the spin connection is $SU(3)$. So the concept of balanced metric plays an important role in superstring theory.

In these lectures, we focus on balanced metrics. In first two sections, we give the definition of a balanced metric and its basic properties. We prove that the existence of balanced metrics is preserved under a proper holomorphic submersion. We also give some examples of hermitian manifolds which admit no balanced metrics. In sections 3 and 4, we describe two classes of balanced manifolds: one is constructed by Calabi, and the other is observed by Goldstein and Prokushkin. From section 5 to 7, we discuss small deformations of a balanced manifold. In section 5, we give an example which shows that the existence of balanced metrics is not preserved under small deformations. In section 6, we give a sufficient condition, i.e., the $\partial\bar{\partial}$-lemma condition, under which small deformations of a balanced manifold are also balanced. Then, in section 7, we weaken the $\partial\bar{\partial}$-lemma condition. We especially obtain the result that any small deformation of the twistor space over a compact self-dual four manifold admits a balanced metric. This section is recently added. In section 8, some existence theorems on balanced metrics are presented. In last two sections, we describe the results of Fu-Li-Yau's paper on the existence of balanced metrics on non-Kähler Calabi-Yau threefolds and of Fu-Yau's paper on solutions to the Strominger system on a class of non-Kähler threefolds.

Acknowledgment. These lectures were given when the author was visiting Department of Mathematics, Tohoku University in July, 2008. The author would like to thank their warm hospitality. These lectures were revised by Shin Kikuta whom the author would also like to thank.

The author would like to thank Professor J. Li and S.-T. Yau for useful discussions on complex geometry and geometric analysis. He would also like to thank Professor C.-H. Gu, H.-S. Hu, Y.-L. Xin and J.-X. Hong for their encouragement and support.

This work is partially supported by NSFC 10771037, 10831008 and 11025103.

1 The definition

Let (X, J) be a complex n-dimensional manifold. Let g be a hermitian metric on X, i.e., a riemannian metric with the property that at each point $x \in X$

$$g(V, W) = g(JV, JW) \quad \text{for all } V, W \in T_xX.$$

Associated with it is the hermitian form

$$\omega(V,W) = g(JV,W).$$

Note that in literature, ω is called the Kähler form of g. Then the standard complex hermitian metric is

$$h = g - i\omega.$$

Definition 1. *A hermitian metric g on a complex n-dimensional manifold is said to be* balanced *if its hermitian form ω satisfies*

$$d(\omega^{n-1}) = 0. \tag{1.1}$$

A complex manifold is called a balanced manifold *if it admits a balanced metric. Otherwise, it is called a non-balanced manifold.*

Remark 2. (1) *A hermitian metric g is called* Kähler *if its hermitian form ω satisfies $d\omega = 0$. Hence a Kähler metric is automatically balanced.*
(2) *When $n = 2$, the conditions of being balanced and Kähler are equivalent.*

In this section, we will use local coordinates to understand the balanced condition (1.1). Let $(z_1,\ldots,z_n)$ be a local complex coordinate system on X. Suppose the complex hermitian metric is given in this coordinate by

$$h = \sum h_{i\bar{j}} dz_i \otimes d\bar{z}_j, \quad h_{i\bar{j}} = h(\frac{\partial}{\partial z_i}, \frac{\partial}{\partial z_j}).$$

Then the hermitian form ω can be written as

$$\omega = i \sum h_{i\bar{j}} dz_i \wedge d\bar{z}_j.$$

Let $(h^{\bar{l}k})_{n\times n}$ be the inverse of the matrix $(h_{k\bar{l}})_{n\times n}$. Direct calculations show

$$\begin{aligned}\omega^{n-1} &= i^{n-1}(n-1)!\det(h_{i\bar{j}}) \\ &\quad \cdot \sum h^{\bar{l}k} dz_1 \wedge d\bar{z}_1 \wedge \cdots \wedge \widehat{dz_k} \wedge d\bar{z}_k \wedge \cdots \wedge dz_l \wedge \widehat{d\bar{z}_l} \wedge \cdots \wedge dz_n \wedge d\bar{z}_n\end{aligned}$$

and

$$\begin{aligned}\partial\omega^{n-1} &= i^{n-1}(n-1)!\det(h_{i\bar{j}}) \\ &\quad \cdot \sum (\frac{\partial h_{i\bar{j}}}{\partial z_k} - \frac{\partial h_{k\bar{j}}}{\partial z_i}) h^{\bar{j}i} h^{\bar{l}k} dz_1 \wedge d\bar{z}_1 \wedge \cdots \wedge dz_l \wedge \widehat{d\bar{z}_l} \wedge \cdots \wedge dz_n \wedge d\bar{z}_n.\end{aligned}$$

Since ω is real, $d\omega^{n-1} = 0$ if and only if $\partial\omega^{n-1} = 0$. So the balanced condition (1.1) is equivalent to

$$\sum_{i,j} \Big(\frac{\partial h_{i\bar{j}}}{\partial z_k} - \frac{\partial h_{k\bar{j}}}{\partial z_i}\Big) h^{\bar{j}i} = 0 \quad \text{for any} \quad 1 \le k \le n. \tag{1.2}$$

On the other hand, the hermitian connection $\bigtriangledown$ of the complex hermitian metric h can be defined by

$$\bigtriangledown\frac{\partial}{\partial z_j}=\sum_k \theta_{jk}\otimes\frac{\partial}{\partial z_k},\quad \theta_{jk}=\sum_l \partial h_{j\bar{l}}\cdot h^{\bar{l}k}.$$

Hence, the torsion $T^{\bigtriangledown}\in\Gamma(\Lambda^2T^*X\otimes TX)$ of the connection $\bigtriangledown$ has no $(1,1)$-component and its $(2,0)$ and $(0,2)$-components are the complex conjugates of one another. Therefore, we only need to consider its $(2,0)$-component:

$$\begin{aligned}
T^{\bigtriangledown}_{\frac{\partial}{\partial z_j},\frac{\partial}{\partial z_k}} &= \bigtriangledown_{\frac{\partial}{\partial z_j}}\frac{\partial}{\partial z_k}-\bigtriangledown_{\frac{\partial}{\partial z_k}}\frac{\partial}{\partial z_j}-[\frac{\partial}{\partial z_j},\frac{\partial}{\partial z_k}]\\
&=\sum_l \theta_{kl}(\frac{\partial}{\partial z_j})\frac{\partial}{\partial z_l}-\sum_l \theta_{jl}(\frac{\partial}{\partial z_k})\frac{\partial}{\partial z_l}\\
&=\sum_{i,l}\Big(\frac{\partial h_{k\bar{i}}}{\partial z_j}-\frac{\partial h_{j\bar{i}}}{\partial z_k}\Big)h^{\bar{i}l}\frac{\partial}{\partial z_l}.
\end{aligned}$$

If we denote

$$T^l_{jk}=\sum_i\Big(\frac{\partial h_{k\bar{i}}}{\partial z_j}-\frac{\partial h_{j\bar{i}}}{\partial z_k}\Big)h^{\bar{i}l}, \tag{1.3}$$

then the $(2,0)$-component of the torsion $T^{\bigtriangledown}$ is

$$T=\sum T^l_{jk}dz_j\wedge dz_k\otimes\frac{\partial}{\partial z_l}.$$

A contraction of T gives a $(1,0)$-form

$$\tau=\sum\tau_j dz_j,\quad \tau_j=\sum_k T^k_{jk}.$$

In [27], τ is called *the torsion $(1,0)$-form* of the complex hermitian metric h. Combining (1.2) with (1.3) gives

Proposition 3. *A hermitian metric on a complex n-dimensional manifold is balanced if and only if $\tau=0$.*

2 Basic properties and examples of non-balanced complex manifolds

Proposition 4. [27] *Suppose X is a compact complex n-dimensional manifold with a balanced metric. Then, every compact complex subvariety of dimension $n-1$ in X represents a non-zero class in $H_{2n-2}(X,\mathbb{R})$.*

Proof. Let ω be the hermitian form of a balanced metric on X. Let S be such a subvariety. Then

$$\int_S \omega^{n-1} = (n-1)!\mathrm{vol}(S) \neq 0.$$

On the other hand, if $[S] = 0 \in H_{2n-2}(X, \mathbb{R})$, then the Poincaré duality implies that there exists an exact 2-form $\eta_S = d\zeta_S$ such that

$$\int_S \omega^{n-1} = \int_X \omega^{n-1} \wedge d\zeta_S = \int_X d(\omega^{n-1} \wedge \zeta_S) - \int_X d\omega^{n-1} \wedge \zeta_S = 0.$$

It's a contradiction. □

Example 5. *Calabi and Eckmann* [9] *constructed a complex structure on* $S^{2p+1} \times S^{2q+1}$ *such that*

$$\pi : S^{2p+1} \times S^{2q+1} \to \mathbb{C}P^p \times \mathbb{C}P^q$$

is holomorphic. Here $S^{2p+1} \to \mathbb{C}P^p$ *and* $S^{2q+1} \to \mathbb{C}P^q$ *are the hopf fibrations. Hence, if we take the compact complex hypersurface* $S = \mathbb{C}P^{p-1} \times \mathbb{C}P^q$ *in* $\mathbb{C}P^p \times \mathbb{C}P^q$*, then* $\pi^{-1}(S)$ *is a compact complex hypersurface in the complex manifold* $S^{2p+1} \times S^{2q+1}$*. As* $H_{2(p+q)}(S^{2p+1} \times S^{2q+1}, \mathbb{R}) = 0$*, we see that the manifold possesses no balanced metrics.*

Proposition 6. [27] *Let X and Y be complex manifolds.*
(1) *If X and Y are balanced, then the product $X \times Y$ is also balanced.*
(2) *If X is balanced and if $\pi : X \to Y$ is a proper holomorphic submersion, then Y is also balanced.*

Proof. Let $\dim_{\mathbb{C}} X = n$ and $\dim_{\mathbb{C}} Y = m$.

(1) Let ω_X and ω_Y be, respective, hermitian forms of balanced metrics on X and Y. Then

$$\omega = \omega_X + \omega_Y$$

is a hermitian form of the product metric on $X \times Y$ and

$$\omega^{n+m-1} = \begin{pmatrix} n-1 \\ n+m-1 \end{pmatrix} \omega_X^{n-1} \wedge \omega_Y^m + \begin{pmatrix} m-1 \\ n+m-1 \end{pmatrix} \omega_X^n \wedge \omega_Y^{m-1}.$$

Obviously, we have $d(\omega^{n+m-1}) = 0$ and so the product metric is a balanced metric on $X \times Y$.

(2) Let ω_X be the hermitian form of a balanced metric on X. Let $\Omega_X = \omega_X^{n-1}$. Since π is a proper holomorphic submersion,

$$\Omega_Y = \pi_* \Omega_X$$

is an $(m-1,m-1)$-form on Y. Here π_* is the push-forward operator. In fact, for any 2-form ϕ on Y with compact support, $\pi^*\phi$ is also with compact support in X. Then $\pi_*\Omega_X$ is defined by

$$\int_Y \pi_*\Omega_X \wedge \phi = \int_X \Omega_X \wedge \pi^*\phi = \int_Y \Big(\int_{f^{-1}(y)} \Omega_X\Big) \wedge \phi.$$

For Ω_Y, we have the following facts:
(i) $d\Omega_Y = 0$, since $d\pi_* = \pi_* d$;
(ii) Ω_Y is a strictly positive $(m-1,m-1)$-form.

Actually, $\Omega_X = \omega_X^{n-1}$ is a strictly positive $(n-1,n-1)$-form; that is, with respect to any $\mathbb{C}$-basis $\theta_1,\cdots,\theta_n$ of the holomorphic cotangent space $T_x^{*1,0}X$, we have

$$\Omega_X(x) = i^{n-1}\sum a_{j\bar{k}}\theta_1\wedge\bar{\theta}_1\wedge\cdots\wedge\widehat{\theta_j}\wedge\bar{\theta}_j\wedge\cdots\wedge\theta_k\wedge\widehat{\bar{\theta}_k}\wedge\cdots\wedge\theta_n\wedge\bar{\theta}_n, \tag{2.1}$$

such that $(a_{j\bar{k}})_{n\times n}$ is a hermitian positive definite matrix. Now fix $y\in Y$ and fix a $\mathbb{C}$-basis $\theta_1,\ldots,\theta_m$ for $T_y^{*1,0}Y$. At each point $x\in F=f^{-1}(y)$, we can lift the forms $\theta_1,\ldots,\theta_m$ and supply some suitable forms to obtain a basis $\theta_1,\ldots,\theta_n$ of $T_x^{*1,0}X$. Since π is the holomorphic submersion, we can do this so that

$$\nu = i^{n-m}\theta_{m+1}\wedge\bar{\theta}_{m+1}\wedge\cdots\wedge\theta_n\wedge\bar{\theta}_n,$$

when restricted to F, is any given smooth volume element. So if we write $\Omega_X(x)$ for any $x\in f^{-1}(y)$ as in (2.1), then Ω_Y at y can be written as

$$\Omega_Y(y) = i^{m-1}\sum \tilde{a}_{j\bar{k}}(y)\theta_1\wedge\bar{\theta}_1\wedge\cdots\wedge\widehat{\theta_j}\wedge\bar{\theta}_j\wedge\cdots\wedge\theta_k\wedge\widehat{\bar{\theta}_k}\wedge\cdots\wedge\theta_m\wedge\bar{\theta}_m,$$

where

$$\tilde{a}_{j\bar{k}}(y) = \int_{f^{-1}(y)} a_{j\bar{k}}(x)\nu$$

for each j,k. Note that since $(a_{j\bar{k}})_{n\times n}$ is hermitian positive definite, $(\tilde{a}_{j\bar{k}})_{m\times m}$ is also hermitian positive definite. Hence, Ω_Y is a strictly positive $(m-1,m-1)$-form on Y.

We need the following

Claim. There exists a unique strictly positive $(1,1)$-form ω_Y on Y such that $\omega_Y^{m-1} = \Omega_Y$.

Proof. We assume that ω_Y is a strictly positive $(1,1)$-form on Y. Locally we write

$$\omega_Y = i\sum h_{i\bar{j}}\theta_i\wedge\bar{\theta}_j.$$

If $\omega_Y^{m-1} = \Omega_Y$ holds, then

$$(m-1)!\det(h_{i\bar{j}})h^{\bar{l}k} = \tilde{a}_{k\bar{l}}.$$

We can solve $h_{k\bar{l}}$ from the above equality

$$h_{k\bar{l}} = \frac{\big(\det\tilde{a}_{i\bar{j}}\big)^{\frac{1}{m-1}}}{((m-1)!)^{\frac{m}{m-1}}}\tilde{a}^{\bar{l}k}. \tag{2.2}$$

Hence we define $h_{k\bar{l}}$ by (2.2) to get ω_Y. Then such an ω_Y is unique, strictly positive and satisfies $\omega_Y^{m-1} = \Omega_Y$. □

Of course, ω_Y determines a unique hermitian metric g on Y by the formula $g(V, W) = \omega(JV, W)$ for all vector fields V and W. Since $d\omega_Y^{m-1} = d\Omega_Y = 0$, g is a balanced metric on Y. □

Remark 7. (1) *A good reference on positive forms is* [29].
(2) *From the proof of the above proposition, we know that, to find a balanced metric on a complex n-dimensional manifold, we only need to find a d-closed and strictly positive* $(n-1, n-1)$*-form.*

Example 8. *If M is a balanced manifold, then according to proposition 6, $M \times T^{2k}$ is also a balanced manifold for all k-dimensional complex torus T^{2k}.*

On the other hand, if M is a compact non-balanced complex manifold, for example, if M is a compact non-Kähler complex surface, then the complex manifold $M \times T^{2k}$ is also not balanced for any $k \geq 1$.

3 Examples: Calabi's construction

Let $\mathbb{O} \cong \mathbb{R}^8$ denote the set of Cayley numbers. We fix a basis $\{I_1, \ldots, I_7\}$ such that
(1) $I_i \cdot I_j = \delta_{ij}$ with respect to the standard inner product on $\mathbb{R}^8$;
(2) The table of the cross product $I_j \times I_k$ is the following

(3.1)

$\times$	I_1	I_2	I_3	I_4	I_5	I_6	I_7
I_1	0	I_3	$-I_2$	I_5	$-I_4$	I_7	$-I_6$
I_2	$-I_3$	0	I_1	I_6	$-I_7$	$-I_4$	I_5
I_3	I_2	$-I_1$	0	$-I_7$	$-I_6$	I_5	I_4
I_4	$-I_5$	$-I_6$	I_7	0	I_1	I_2	$-I_3$
I_5	I_4	I_7	I_6	$-I_1$	0	$-I_3$	$-I_2$
I_6	$-I_7$	I_4	$-I_5$	$-I_2$	I_3	0	I_1
I_7	I_6	$-I_5$	$-I_4$	I_3	I_2	$-I_1$	0

Then via this basis, we have the isomorphism $\mathbb{R}^7 \cong \mathrm{Im}(\mathbb{O})$.

Next let $X^6 \hookrightarrow \mathbb{R}^7$ be a smooth oriented hypersurface. There is a natural almost complex structure $J : TX \to TX$ induced by the Cayley multiplication as follows. Let N be a unit normal vector field to X. Since for any $V \in T_xX$ at $x \in X$,

$$N \cdot (N \times V) = 0,$$

So in the following, a strictly positive $(1,1)$-form is also called a hermitian metric.

we have $N \times V \in T_xX$. Then we define $J : T_xX \to T_xX$ by

$$J(V) = N \times V.$$

One can check that

$$J^2(V) = N \times (N \times V) = -(N \times N) \times V + (N \cdot V)N - (N \cdot N)V = -V.$$

So J is an almost complex structure on X.

Theorem 9. [8] *J is integrable if and only if*

$$B(JV, JW) = -B(V, W)$$

for all pairs of tangent vectors V and W. Here B is the second fundamental form of X.

Calabi then constructed compact complex manifolds as follows. Let C be a compact riemann surface carrying 3 holomorphic differentials ϕ_1, ϕ_2, ϕ_3 with the following properties:

(1) linear independent;

(2) $\phi_1^2 + \phi_2^2 + \phi_3^2 = 0$;

(3) $i\phi_1 \wedge \bar{\phi}_1 + i\phi_2 \wedge \bar{\phi}_2 + i\phi_3 \wedge \bar{\phi}_3 > 0$.

We lift ϕ_1, ϕ_2, ϕ_3 to the universal covering $\tilde{C} \to C$ and denote them by the same symbols. Moreover, we fix a point $p' \in \tilde{C}$ and set

$$x^j(p) = \mathrm{Re} \int_{p'}^{p} \phi_j, \quad j = 1, 2, 3$$

for any point $p \in \tilde{C}$. Then, we obtain a conformal minimal immersion

$$\psi = (x^1, x^2, x^3) : \tilde{C} \to \mathbb{R}^3.$$

This mapping is regular, since the differentials ϕ_j satisfy (3); furthermore, by the Weierstrass representation, property (2) is equivalent to the statement that ψ is minimal and conformal; finally, it follows from property (1) that $\tilde{C}$ is not mapped into a plane.

Now we consider the hypersurface of the type

$$(\psi, id) : \tilde{C} \times \mathbb{R}^4 \to \mathbb{R}^3 \times \mathbb{R}^4 = \mathrm{Im}(\mathbb{O}),$$

where we regard $\mathbb{R}^3 = \mathrm{span}_{\mathbb{R}}\{I_1, I_2, I_3\}$ and $\mathbb{R}^4 = \mathrm{span}_{\mathbb{R}}\{I_4, I_5, I_6, I_7\}$. Since $\psi : \tilde{C} \to \mathbb{R}^3$ is minimal, $\tilde{C} \times \mathbb{R}^4$ is a complex manifold by theorem 9. If $g : \tilde{C} \to \tilde{C}$ denotes a covering transformation, then $\psi(gp) = \psi(p) + t_g$ for some vector $t_g \in \mathbb{R}^3$ because ϕ_j are invariant by g. It follows that the complex structure on $\tilde{C} \times \mathbb{R}^4$ is invariant by the covering transformations of C and so descends to $C \times \mathbb{R}^4$. On the other hand, for $\mathbb{R}^4$, we can further divide by a lattice Λ of translation of $\mathbb{R}^4$, and thereby produce a compact complex manifold $X_\Lambda = C \times T^4$.

Let $p' \in \tilde{C}$ and $q' \in \mathbb{R}^4$. The unit normal vector field N to $\tilde{C} \times \mathbb{R}^4$ at (p', q') is just the unit normal vector field to $\tilde{C}$ at p'. Thus, we can let the unit normal vector field be

$$N(p') = N(p', q') = \sum_{j=1}^{3} a_j(p') I_j, \quad \sum_{j=1}^{3} a_j^2 = 1.$$

If we let

$$e_1(p') = \sum_{j=1}^{3} b_j(p') I_j, \quad \sum_{j=1}^{3} b_j^2 = 1$$

be a unit vector field to $\tilde{C}$ at p', then from table (3.1), we have

$$e_2(p') \triangleq N(p') \times e_1(p') \in \mathbb{R}^3.$$

We also have

$$\begin{aligned} e_2(p') \cdot e_2(p') =& \big(N(p') \times e_1(p')\big) \cdot \big(N(p') \times e_1(p')\big) \\ =& \big((N(p') \times e_1(p')) \times N(p')\big) \cdot e_1(p') \\ =& e_1(p') \cdot e_1(p') = 1 \end{aligned}$$

and

$$e_2(p') \cdot e_1(p') = (N(p') \times e_1(p')) \cdot e_1(p') = 0.$$

Therefore, the complex structure on X_Λ, when restricted to $C \times \{q\}$ for any $q \in T^4$, is the rotation by $-90°$ on $C \times \{q\}$. Since ψ is conformal, this induced complex structure on $C \times \{q\}$ is the same as the given complex structure on the riemann surface. Hence, we get the following picture:

$$\begin{array}{ccc} T^4 & \longrightarrow & X_\Lambda \\ & & \Big\downarrow {\scriptstyle \pi:\ \text{hol.}} \\ & & C \end{array}$$

On the other hand, by table (3.1) again,

$$I_j \times I_\alpha \in \mathbb{R}^4, \quad \text{for all} \quad j = 1, 2, 3; \quad \alpha = 4, 5, 6, 7.$$

Then we have $N \times I_\alpha \in \mathbb{R}^4$ for all $\alpha = 4, 5, 6, 7$. So for each $p \in C$, $\{p\} \times T^4$ is a complex submanifold of X_Λ. However, the unit vector field $N(p)$ depends on p. Therefore, we can view X_Λ as a family of complex tori, parameterized by the riemann surface.

Proposition 10. [8] *The complex manifolds X_Λ are non-Kähler.*

Proposition 11. [21] *The complex manifolds X_Λ are balanced.*

Proof. We use the method in [16]. We can write down the natural metric on X_Λ. Define a 2-form on $\tilde{C} \times \mathbb{R}^4$ by

$$\omega(V, W) = N \cdot (V \times W)$$

for any $V, W \in T_x(\tilde{C} \times \mathbb{R}^4)$ at any point $x \in \tilde{C} \times \mathbb{R}^4$. Then, clearly we have

$$\omega(V, W) = -\omega(W, V);$$

Using the formula

$$N \cdot (V \times W) = (N \times V) \cdot W,$$

we also have

$$\omega(JV, JW) = \omega(V, W);$$
$$\omega(V, JV) = (N \times V) \cdot (N \times V) > 0, \quad \text{if} \quad V \neq 0.$$

Thus, ω is a strictly positive $(1,1)$-form on $\tilde{C} \times \mathbb{R}^4$. Since ω is invariant under the covering transformations of X_Λ, it descends to X_Λ and defines a hermitian metric on X_Λ, which is still denoted by ω.

Next, we should check that ω is a balanced metric. Let the unit normal vector field to be

$$N = \sum_{j=1}^{3} a_j I_j, \qquad \sum_{j=1}^{3} a_j^2 = 1,$$

where a_j, for $j = 1, 2, 3$, are functions on C. Let (x_4, x_5, x_6, x_7) be the standard local coordinates on T^4. Note that ω can be written as

$$\omega = \pi^* \omega_C + \varphi_0,$$

where ω_C is a Kähler metric on C as ψ is conformal, and

$$\begin{aligned}\varphi_0 =& a_1 dx_4 \wedge dx_5 + a_2 dx_4 \wedge dx_6 - a_3 dx_4 \wedge dx_7 \\ &- a_3 dx_5 \wedge dx_6 - a_2 dx_5 \wedge dx_7 + a_1 dx_6 \wedge dx_7.\end{aligned}$$

A direct calculation shows

$$\varphi_0^2 = 2 dx_4 \wedge dx_5 \wedge dx_6 \wedge dx_7.$$

Therefore,

$$d(\omega^2) = d(2\pi^* \omega_C \wedge \varphi_0 + \varphi_0^2) = 2\pi^*(d\omega_C) \wedge \varphi_0 + 2\pi^* \omega_C \wedge d\varphi_0 = 0,$$

since ω_C is the Kähler metric on C, and all functions a_j for $j = 1, 2, 3$ are only defined on C. □

Remark 12. *Gray* [21] *used the different method to prove the above proposition. Actually he showed*

$$d^*\omega(V) = -\sum_{i=0}^{6}\langle\bar{\nabla}_{E_i}(\times)(V,N),E_i\rangle.$$

Here d^* *is the coderivative on* X_Λ *with respect to* ω *and* $\{E_0,\ldots,E_6\}$ *is any frame field. As the vector cross product* $\times$ *is parallel,* $d^*\omega = 0$. *Then from the identities* $d^*\omega = -*d*\omega = -*\frac{d\omega^{n-1}}{(n-1)!}$, ω *is a balanced metric.*

Remark 13. *As a differential manifold,* X_Λ *is just the product manifold* $C\times T^4$. *Define* $\pi_1 : X_\Lambda \to T^4$ *to be the natural projection. For any smooth function* u *on* T^4, *we define*

$$\omega_u = (e^{-u}\circ\pi_1)\cdot\pi^*\omega_C + (e^u\circ\pi_1)\cdot\phi_0.$$

Then, ω_u *is also a balanced metric on* X_Λ.

4 Examples: Goldstein-Prokushkin's observation

We deal with this section as in [17].

Theorem 14. [20] *Let* ω_1 *and* ω_2 *be two closed* 2*-forms on a complex* m*-dimensional manifold* M *such that the following two conditions hold*:
(1) $\omega_1 + i\omega_2$ *has no component in* $\Lambda^{0,2}T^*M$;
(2) $\frac{\omega_1}{2\pi}$ *and* $\frac{\omega_2}{2\pi}$ *represent integral cohomology classes.*
Then there is a complex $(m+1)$*-dimensional manifold* X *and a holomorphic fibration* $\pi : X\to M$ *such that all fibers are* T^2.

Proof. We choose a good cover $\mathcal{U} = \{U_\alpha\}$ on M, i.e., all nonempty finite intersections $U_{\alpha_0\cdots\alpha_p} \triangleq U_{\alpha_0}\cap\cdots\cap U_{\alpha_p}$ are diffeomorphic to $\mathbb{R}^{2m}$. Since $d\omega_1 = 0$ and U_α is diffeomorphic to $\mathbb{R}^{2m}$, by the Poincaré lemma, there exists a smoothly real 1-form $\xi_{1\alpha}$ on U_α such that

$$\omega_1|_{U_\alpha} = d\xi_{1\alpha}.$$

If $U_{\alpha\beta}\neq\emptyset$, then

$$d(\xi_{1\alpha}-\xi_{1\beta})|_{U_{\alpha\beta}} = 0.$$

Also, since $U_{\alpha\beta}$ is diffeomorphic to $\mathbb{R}^{2m}$, by the Poincaré lemma again, there exists a smooth real function $f_{\alpha\beta}$ on $U_{\alpha\beta}$ such that

$$(\xi_{1\alpha}-\xi_{1\beta})|_{U_{\alpha\beta}} = df_{\alpha\beta}.$$

If $U_{\alpha\beta\gamma}\neq\emptyset$, then we have

$$d(f_{\alpha\beta}+f_{\beta\gamma}+f_{\gamma\alpha})|_{U_{\alpha\beta\gamma}} = (\xi_{1\alpha}-\xi_{1\beta}+\xi_{1\beta}-\xi_{1\gamma}+\xi_{1\gamma}-\xi_{1\alpha})|_{U_{\alpha\beta\gamma}} = 0.$$

Hence, $(f_{\alpha\beta} + f_{\beta\gamma} + f_{\gamma\alpha})|_{U_{\alpha\beta\gamma}}$ is a real constant, which we denote by $c_{\alpha\beta\gamma}$. If we use the language of the Cěch cohomology, we have

$$\{c_{\alpha\beta\gamma}\} \in C^2(\mathcal{U}, \mathbb{R}).$$

It can be easily checked that

$$(\delta c)_{\alpha\beta\gamma\eta} = c_{\beta\gamma\eta} - c_{\alpha\gamma\eta} + c_{\alpha\beta\eta} - c_{\alpha\beta\gamma} = 0,$$

i.e., $[c] \in \breve{H}^2(M, \mathbb{R})$.

Since $\frac{\omega_1}{2\pi}$ represents an integral class, $[\frac{c}{2\pi}] \in \breve{H}^2(M, \mathbb{Z})$, i.e., $\frac{c_{\alpha\beta\gamma}}{2\pi} \in \mathbb{Z}$. If we let

$$g_{\alpha\beta} = e^{if_{\alpha\beta}},$$

then on $U_{\alpha\beta\gamma}$,

$$g_{\alpha\beta}g_{\beta\gamma}g_{\gamma\alpha} = \exp\{i(f_{\alpha\beta} + f_{\beta\gamma} + f_{\gamma\alpha})\} = \exp\{ic_{\alpha\beta\gamma}\} = 1.$$

Hence, $\{g_{\alpha\beta}\}$ defines a line bundle $\pi_1 : L_1 \to M$. A localization of L_1

$$\varphi_\alpha : L_1|_{U_\alpha} \cong U_\alpha \times \mathbb{C}$$

satisfies

$$\varphi_\alpha \varphi_\beta^{-1}(x, z_\beta) = (x, g_{\alpha\beta}(x)z_\alpha).$$

This localization defines local sections

$$s_\alpha(x) = \varphi_\alpha^{-1}(x, 1_\alpha)$$

satisfying

$$s_\beta(x) = g_{\alpha\beta}s_\alpha(x).$$

We define a hermitian metric h on the line bundle L_1 as

$$h(s_\alpha, s_\alpha) = 1.$$

It is well-defined since

$$h(s_\beta, s_\beta) = h(g_{\alpha\beta}s_\alpha, g_{\alpha\beta}s_\alpha) = |g_{\alpha\beta}|^2 h(s_\alpha, s_\alpha) = 1.$$

The hermitian connection on L_1 can be defined by

$$\nabla s_\alpha = i\xi_{1\alpha} \otimes s_\alpha.$$

This definition is reasonable, since

$$\begin{aligned}\nabla s_\beta = \nabla(g_{\alpha\beta}s_\alpha) =& dg_{\alpha\beta}\otimes s_\alpha + g_{\alpha\beta}\nabla s_\alpha\\ =& dg_{\alpha\beta}\otimes s_\alpha + ig_{\alpha\beta}\xi_{1\alpha}\otimes s_\alpha\\ =& ig_{\alpha\beta}(\xi_{1\alpha} - ig_{\alpha\beta}^{-1}dg_{\alpha\beta})\otimes s_\alpha\\ =& ig_{\alpha\beta}(\xi_{1\alpha}+df_{\alpha\beta})\otimes s_\alpha\\ =& i\xi_{1\beta}\otimes s_\beta.\end{aligned}$$

We then compute

$$\nabla^2 s_\alpha = \nabla(i\xi_{1\alpha}\otimes s_\alpha) = id\xi_{1\alpha}\otimes s_\alpha - \xi_{1\alpha}\wedge\xi_{1\alpha}\otimes s_\alpha = i\omega_1\otimes s_\alpha.$$

Thus, the curvature of the connection is $i\omega_1$ and the first Chern class of L_1 is $[\frac{\omega_1}{2\pi}]$.

Let

$$\mathbb{S}_1 = \{v\in L_1 | h(v,v) = 1\}.$$

Then, $\pi_1 : \mathbb{S}_1 \to M$ is a S^1-bundle over M. The section s_α defines a local coordinate x_α on its fibers; that is, points on fibers can be described by $e^{ix_\alpha}s_\alpha$. For any point $p\in U_{\alpha\beta}$, a point on the fiber $\pi^{-1}(p)$ can be described by $e^{ix_\alpha}s_\alpha$ and $e^{ix_\beta}s_\beta$. Then from

$$e^{ix_\alpha}s_\alpha = e^{ix_\beta}s_\beta = e^{ix_\beta}g_{\alpha\beta}s_\alpha = e^{i(x_\beta+f_{\alpha\beta})}s_\alpha,$$

we get a relation between fiber's local coordinates x_α and x_β: For some $k\in\mathbb{Z}$,

$$x_\alpha = x_\beta + f_{\alpha\beta} + 2k\pi.$$

Therefore,

$$\xi_{1\alpha} + dx_\alpha = \xi_{1\alpha} + dx_\beta + df_{\alpha\beta} = \xi_{1\beta} + dx_\beta.$$

This equality means that $\xi_{1\alpha}+dx_\alpha$ is a globally defined real 1-form on $\mathbb{S}_1$, which is denoted by $dx+\xi_1$.

Similarly, we can use ω_2 to construct another line bundle L_2 such that its first Chern class is $[\frac{\omega_2}{2\pi}]$. A natural hermitian metric on L_2 also defines another circle bundle $\mathbb{S}_2$. Moreover, if we write locally $\omega_2|_{U_\alpha} = d\xi_{2\alpha}$, we can choose a local coordinate y_α on its fibers such that $dy_\alpha + \xi_{2\alpha}$ is a globally defined real 1-form on $\mathbb{S}_2$, which is denoted by $dy+\xi_2$.

Now we define the direct sum $X = \mathbb{S}_1\oplus\mathbb{S}_2$. There is a natural map $\pi : X\to M$. We need to define a complex structure on X. Let $(z_1,\dots,z_m)$ be a holomorphic local coordinate system on M. Let

$$\theta = (dx+\xi_1) + i(dy+\xi_2).$$

If we define $\{\pi^* dz_1, \pi^* dz_2, \dots, \pi^* dz_m, \theta\}$ as a basis of the linear space of $(1,0)$-forms on X, then it defines an almost complex structure J on X. Since $d(\pi^* dz_j) = 0$ for all $j = 1, \dots, m$, and $d\theta = \pi^*(\omega_1 + i\omega_2)$ has no $(0,2)$-component, J is integrable. Thus X is a complex manifold such that $\pi : X \to M$ is a holomorphic T^2 bundle. □

Remark 15. (1) *Calabi-Eckmann's example in section 2 can be gotten in this way: take* $M = \mathbb{C}P^p \times \mathbb{C}P^q$, $[\frac{\omega_1}{2\pi}] = 1 \in H^2(\mathbb{C}P^p, \mathbb{Z}) \cong \mathbb{Z}$ *and* $[\frac{\omega_2}{2\pi}] = 1 \in H^2(\mathbb{C}P^q, \mathbb{Z}) \cong \mathbb{Z}$.

(2) *If* $M = T^4$ *and* ω_1 *is trivial, we can choose* $[\frac{\omega_2}{2\pi}] \in H^2(T^4, \mathbb{Z})$ *such that the manifold* X *is just the Iwasawa manifold.*

If g_M is a hermitian metric on M, then

$$g = \pi^* g_M + (dx + \xi_1)^2 + (dy + \xi_2)^2$$

is a hermitian metric on X. If we denote the hermitian form of the metric g_M by ω_M, then the hermitian form of the metric g is

$$\omega = \pi^* \omega_M + i\theta \wedge \bar{\theta}.$$

Proposition 16. [20] *Let* $(Y, \omega_Y, \mathcal{V}_Y)$ *be a smooth Calabi-Yau 2-fold with a nonvanishing holomorphic 2-form* $\mathcal{V}_Y$ *such tat* $\| \mathcal{V}_Y \|_{\omega_Y} = 1$. *Let* ω_1 *and* ω_2 *be d-closed anti-self dual real* $(1,1)$*-forms on* Y *such that* $[\frac{\omega_1}{2\pi}], [\frac{\omega_2}{2\pi}] \in H^2(Y, \mathbb{Z})$. *Then, there is a hermitian 3-fold* X *such that* $\pi : X \to Y$ *is a holomorphic* T^2*-fibration over* Y *and the following holds:*

(1) X *admits a nowhere vanishing holomorphic 3-form*

$$\mathcal{V} = \pi^* \mathcal{V}_Y \wedge \theta.$$

(2) *If either* ω_1 *or* ω_2 *represents a non-trivial cohomological class, then* X *is non-Kähler.*

(3) *The natural metric* $\omega_0 = \pi^* \omega_Y + i\theta \wedge \bar{\theta}$ *is a balanced metric on* X.

(4) *Furthermore, for any smooth function* u *on* Y, *the hermitian metric*

$$\omega_u = \pi^*(e^u \omega_Y) + i\theta \wedge \bar{\theta}$$

is conformal balanced. More precisely, $(\omega_u, \mathcal{V})$ *satisfies the equation*

$$d(||\mathcal{V}||_{\omega_u} \omega_u^2) = 0.$$

Proof. Note that θ is not a holomorphic 1-form. Actually, since $d\theta = \pi^*(\omega_1 + i\omega_2)$ is a $(1,1)$-form, then

$$\partial\theta = 0, \quad \bar{\partial}\theta = \pi^*(\omega_1 + i\omega_2). \tag{4.1}$$

On the other hand, since ω_1 and ω_2 are $\bar{\partial}$-closed, according to the $\bar{\partial}$-Poincaré lemma, there exist $(1,0)$-forms ζ_1 and ζ_2 on U_α such that

$$\omega_1|_{U_\alpha} = \bar{\partial}\zeta_1 \quad \text{and} \quad \omega_2|_{U_\alpha} = \bar{\partial}\zeta_2.$$

If we define

$$\theta_0 = \theta - \pi^*(\zeta_1 + i\zeta_2),$$

then $\bar{\partial}\theta_0 = 0$. So θ_0 is a holomorphic 1-form on $\pi^{-1}(U_\alpha)$.

(1) Although θ is not a holomorphic 1-form on X, $\mathcal{V} = \pi^*\mathcal{V}_Y \wedge \theta$ is a holomorphic 3-form on X since we can write $\mathcal{V}$ locally as

$$\mathcal{V} = \pi^*\mathcal{V}_Y \wedge (\theta - \pi^*(\zeta_1 + i\zeta_2)).$$

(2) Goldestein and Prokushkin proved that

$$\begin{aligned} h^{1,0}(X) =& h^{1,0}(Y) \\ h^{0,1}(X) =& h^{0,1}(Y) + 1. \end{aligned}$$

Here 1 stands for the dimension of the extra subspace of $H^{0,1}_{\bar{\partial}}(X, \mathbb{C})$ which is spanned by $[\bar{\theta}]$ since $\bar{\partial}\bar{\theta} = 0$ by (4.1) and $[\bar{\theta}]$ is non-trivial by assumptions on ω_1 and ω_2. So X is non-Kähler.

(3) We compute

$$d\omega_0 = id\theta \wedge \bar{\theta} - i\theta \wedge d\bar{\theta} = i\pi^*(\omega_1 + i\omega_2) \wedge \bar{\theta} - i\theta \wedge \pi^*(\omega_1 - i\omega_2).$$

Since ω_1 and ω_2 are anti-self-dual, $\omega_1 \wedge \omega_Y = 0$ and $\omega_2 \wedge \omega_Y = 0$. Therefore

$$d(\omega_0^2) = 2d\omega_0 \wedge \omega_0 = 0.$$

(4) As $||\mathcal{V}_Y||_{\omega_Y} = 1$, $||\mathcal{V}||_{\omega_0} = 1$. Then we have

$$||\mathcal{V}||^2_{\omega_u} = \frac{||\mathcal{V}||^2_{\omega_u}}{||\mathcal{V}||^2_{\omega_0}} = \frac{\omega_0^3}{\omega_u^3} = \frac{1}{e^{2u}}$$

and

$$||\mathcal{V}||_{\omega_u}\omega_u^2 = \omega_0^2 + \pi^*((e^u - 1)\omega_Y^2).$$

Since u is the function on Y, we have

$$d(||\mathcal{V}||_{\omega_u}\omega_u^2) = d(\omega_0^2) + \pi^* d(e^u - 1) \wedge \pi^*\omega_Y^2 = 0.$$

□

5 Small deformations–a counterexample

Let $\{X_t | t \in \triangle(\epsilon)\}$ be a complex analytic family of compact complex manifolds. Here $\triangle(\epsilon) = \{t \in \mathbb{C} | |t| < \epsilon\}$. We will only consider the small deformation; that is, we can assume, if necessary, ϵ will be small enough. Kodaira and Spencer proved the following result.

Theorem 17. [24] *Any small deformation of a compact Kähler manifold is also Kähler.*

Thus, it is reasonable to ask whether the property of being balanced is stable under small deformations. The answer is no, as shown by Alessandrini and Bassanelli. They observed that there exist no balanced metrics on the small deformation of the Iwasawa manifold which was built up by Nakamura.

Let us first describe the Iwasawa manifold. Let

$$G = \left\{ \begin{pmatrix} 1 & z_2 & z_3 \\ 0 & 1 & z_1 \\ 0 & 0 & 1 \end{pmatrix}; z_i \in \mathbb{C} \right\} \cong \mathbb{C}^3$$

and let

$$\Gamma = \left\{ \begin{pmatrix} 1 & \omega_2 & \omega_3 \\ 0 & 1 & \omega_1 \\ 0 & 0 & 1 \end{pmatrix}; \omega_i \in \mathbb{Z} \oplus \sqrt{-1}\mathbb{Z} \right\}.$$

We define an action of Γ on G by

$$\begin{pmatrix} 1 & z_2 & z_3 \\ 0 & 1 & z_1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & \omega_2 & \omega_3 \\ 0 & 1 & \omega_1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & z_2 + \omega_2 & z_3 + z_2\omega_1 + \omega_3 \\ 0 & 1 & z_1 + \omega_1 \\ 0 & 0 & 1 \end{pmatrix},$$

i.e.,

$$\begin{cases} z_1' = z_1 + \omega_1 \\ z_2' = z_2 + \omega_2 \\ z_3' = z_3 + \omega_1 z_2 + \omega_3. \end{cases}$$

We let $X = G/\Gamma$. Then X is a complex manifold and is a holomorphic T^2-bundle over T^4. X is called the Iwasawa manifold.

Lemma 18. *The Iwasawa manifold X is a balanced manifold.*

Proof. On $G \cong \mathbb{C}^3$, we define $\phi_1 = dz_1$, $\phi_2 = dz_2$, and $\phi_3 = dz_3 - z_1 dz_2$. One can check

$$\begin{aligned} \phi_1' &= dz_1' = dz_1 = \phi_1; \\ \phi_2' &= dz_2' = dz_2 = \phi_2; \\ \phi_3' &= dz_3' - z_1' dz_2' = d(z_3 + \omega_1 z_2 + \omega_3) - (z_1 + \omega_1) d(z_2 + \omega_2) \\ &= dz_3 - z_1 dz_2 = \phi_3. \end{aligned}$$

Hence, ϕ_1, ϕ_2 and ϕ_3 descend to forms on X, which are still denoted by ϕ_1, ϕ_2 and ϕ_3 respectively. Clearly $\{\phi_1, \phi_2, \phi_3\}$ is a basis of the space of holomorphic $(1,0)$-forms on X. Let

$$\omega = i(\phi_1 \wedge \bar{\phi}_1 + \phi_2 \wedge \bar{\phi}_2 + \phi_3 \wedge \bar{\phi}_3).$$

Then, ω is a hermitian metric on X. On the other hand, since

$$d\phi_3 = -dz_1 \wedge dz_2 = -\phi_1 \wedge \phi_2,$$

and

$$d\omega = i(d\phi_3 \wedge \bar{\phi}_3 - \phi_3 \wedge d\bar{\phi}_3) = i(-\phi_1 \wedge \phi_2 \wedge \bar{\phi}_3 + \phi_3 \wedge \bar{\phi}_1 \wedge \bar{\phi}_2),$$

we have

$$d\omega^2 = 2d\omega \wedge \omega = 0.$$

□

Now we should prove that X is not a Kähler manifold. A compact Kähler manifold satisfies the following $\partial\bar{\partial}$-lemma.

Definition 19. *A complex manifold X satisfies* the $\partial\bar{\partial}$-lemma *if for every differential form α on X such that $\partial\alpha = \bar{\partial}\alpha = 0$ and such that $\alpha = d\gamma$ for some differential form γ on X, there is some differential form β such that $\alpha = i\partial\bar{\partial}\beta$.*

Remark 20. *There are several equivalent versions of the $\partial\bar{\partial}$-lemma, see* [10].

Lemma 21. *The Iwasawa manifold does not satisfy the $\partial\bar{\partial}$-lemma. Thus, it is non-Kähler.*

Proof. Let $\alpha = d\phi_3$. Then $\alpha = -\phi_1 \wedge \phi_2$ is a $(2,0)$-form and $\partial\alpha = \bar{\partial}\alpha = 0$. Clearly α can not be written as $\alpha = i\partial\bar{\partial}\beta$ for any β. So the $\partial\bar{\partial}$-lemma does not hold on X and X is non-Kähler. □

Next let us recall Nakamura's construction. From the definition of the Iwasawa manifold, we define a smooth family of diffeomorphisms $\Phi_t : \mathbb{C}^3 \to \mathbb{C}^3$ as follows:

$$\begin{cases} \zeta_1 = z_1 + t\bar{z}_2 \\ \zeta_2 = z_2 \\ \zeta_3 = z_3, \end{cases} \qquad \text{for} \quad t \in \triangle(1).$$

For any $(\omega_1, \omega_2, \omega_3) \in \Gamma$, we denote

$$(\tilde{\omega}_1(t), \tilde{\omega}_2(t), \tilde{\omega}_3(t)) = \Phi_t(\omega_1, \omega_2, \omega_3),$$

i.e.,

$$\begin{cases} \tilde{\omega}_1(t) = \omega_1 + t\bar{\omega}_2 \\ \tilde{\omega}_2(t) = \omega_2 \\ \tilde{\omega}_3(t) = \omega_3. \end{cases}$$

Let $\Gamma_t = \{(\tilde{\omega}_1(t), \tilde{\omega}_2(t), \tilde{\omega}_3(t)) | (\omega_1, \omega_2, \omega_3) \in \Gamma\}$. We define an action of Γ_t on $\mathbb{C}^3$ as

$$\begin{cases} \zeta_1' = \zeta_1 + \tilde{\omega}_1(t) \\ \zeta_2' = \zeta_2 + \tilde{\omega}_2(t) \\ \zeta_3' = \zeta_3 + \tilde{\omega}_3(t) + (\tilde{\omega}_1(t) - t\bar{\tilde{\omega}}_2(t))\zeta_2, \end{cases}$$

and then let $X_t = \mathbb{C}^3/\Gamma_t$. We can check that

$$\begin{cases} \zeta_1' = z_1 + t\bar{z}_2 + \omega_1 + t\bar{\omega}_2 = z_1' + t\bar{z}_2' \\ \zeta_2' = z_2 + \omega_2 = z_2' \\ \zeta_3' = z_3 + \omega_1 z_2 + \omega_3 = z_3'. \end{cases}$$

That is, the following diagram is commutative:

$$\begin{array}{ccc} \mathbb{C}^3 & \xrightarrow{\Phi_t} & \mathbb{C}^3 \\ {\scriptstyle \Gamma-\text{action}}\downarrow & & \downarrow{\scriptstyle \Gamma_t-\text{action}} \\ \mathbb{C}^3 & \xrightarrow{\Phi_t} & \mathbb{C}^3. \end{array}$$

Therefore, the diffeomorphisms Φ_t descend to a smooth family of diffeomorphisms $\phi_t : X \to X_t$

$$\begin{array}{ccc} \mathbb{C}^3 & \xrightarrow{\Phi_t} & \mathbb{C}^3 \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi_t} \\ X = \mathbb{C}^3/\Gamma & \xrightarrow{\phi_t} & X_t = \mathbb{C}^3/\Gamma_t. \end{array}$$

Thus, we get a complex analytic family X_t such that X_0 is the Iwasawa manifold.

Proposition 22. [1] *The above small deformation X_t of the Iwasawa manifold is not balanced when $t \neq 0$.*

Proof. Since $d\zeta_1' = d\zeta_1$ and $d\zeta_2' = d\zeta_2$, $d\zeta_1$ and $d\zeta_2$ descend to two forms on X_t, which are denoted respectively by $\phi_{1,t}$ and $\phi_{2,t}$. Then we let

$$\phi_3 = d\zeta_3 + (t\bar{\zeta}_2 - \zeta_1)d\zeta_2.$$

By the definition, we have

$$\begin{aligned} \phi_3' &= d\zeta_3' + (t\bar{\zeta}_2' - \zeta_1')d\zeta_2' \\ &= d\zeta_3 + (t\bar{\zeta}_2 - \zeta_1)d\zeta_2 = \phi_3, \end{aligned}$$

i.e., ϕ_3 defines a $(1,0)$-form $\phi_{3,t}$ on X_t. We compute

$$d\phi_{3,t} = -t\phi_{2,t} \wedge \bar{\phi}_{2,t} - \phi_{1,t} \wedge \phi_{2,t}.$$

If there exists a balanced metric ω_t on X_t, then for $t \neq 0$,

$$\int_{X_t} d\phi_{3,t} \wedge \omega_t^2 = -t \int_{X_t} \phi_{2,t} \wedge \bar{\phi}_{2,t} \wedge \omega_t^2 \neq 0.$$

On the other hand,

$$\int_{X_t} d\phi_{3,t} \wedge \omega_t^2 = \int_{X_t} d(\phi_{3,t} \wedge \omega_t^2) + \int_{X_t} \phi_{3,t} \wedge d\omega_t^2 = 0.$$

It's a contradiction. □

6 Small deformations–a sufficient condition

In this section, we will give a sufficient condition under which small deformations of a balanced manifold are also balanced.

Let X_t be a complex analytic family of compact complex manifolds. We denote by $\Lambda_t^{p,q}$ the space of (p,q)-forms on X_t. We choose a smooth family of hermitian metrics ω_t on X_t and consider the Kodaira-Spencer operators $E_t : \Lambda_t^{p,q} \to \Lambda_t^{p,q}$ as follows:

$$E_t = \partial_t \bar{\partial}_t \bar{\partial}_t^* \partial_t^* + \bar{\partial}_t^* \partial_t^* \partial_t \bar{\partial}_t + \bar{\partial}_t^* \partial_t \partial_t^* \bar{\partial}_t + \partial_t^* \bar{\partial}_t \bar{\partial}_t^* \partial_t + \bar{\partial}_t^* \bar{\partial}_t + \partial_t^* \partial_t.$$

Lemma 23. [24] *The differential operators E_t are self-adjoint and strongly elliptic of order 4. A form $\varphi \in \Lambda_t^{p,q}$ satisfies the equation $E_t\varphi = 0$ if and only if*

$$\partial_t \varphi = \bar{\partial}_t \varphi = 0 \quad \textit{and} \quad \bar{\partial}_t^* \partial_t^* \varphi = 0.$$

Then, we let

$$F_t^{p,q} = \{\varphi \in \Lambda_t^{p,q} | E_t \varphi = 0\}$$

be the kernel of E_t and let $F_t : \Lambda_t^{p,q} \to F_t^{p,q}$ be the orthogonal projection with respect to the hermitian metric ω_t. Let G_t be Green's operator associated to E_t so that for each $\varphi \in \Lambda_t^{p,q}$,

$$\varphi = E_t G_t \varphi + F_t \varphi.$$

Theorem 24. [32, 33] *If X_0 satisfies the $\partial\bar{\partial}$-lemma, then for small enough t, X_t also satisfies the $\partial\bar{\partial}$-lemma and*

$$h_t^{p,q} = \dim F_t^{p,q} = h_0^{p,q}.$$

Theorem 25. [33] *If X_0 satisfies the $\partial\bar{\partial}$-lemma and admits a balanced metric, then X_t admits a balanced metric for small enough t.*

Proof. Let ω be a balanced metric on X_0. Pick a smooth family of hermitian metrics ω_t on X_t so that $\omega_0 = \omega$. For example, we can pick ω_t as follows. Let $x_t : X_t \to X_0$ be a family of diffeomorphisms smoothly depending on t and $x_0 = id$. Then $\Omega_t = x_t^* \omega_0^{n-1}$ is a real $(2n-2)$-form on X_t. We decompose Ω_t as

$$\Omega_t = \Omega_t^{n,n-2} + \Omega_t^{n-1,n-1} + \Omega_t^{n-2,n}.$$

Since $\Omega_0 = \omega_0^{n-1}$ is strictly positive, by continuity, we see that for small t, $\Omega_t^{n-1,n-1}$ are also strictly positive definite. Hence, there exists a unique hermitian metric ω_t on X_t such that $\omega_t^{n-1} = \Omega_t^{n-1,n-1}$. Clearly ω_t depends smoothly on t.

By theorem 24, $\dim h_t^{p,q} = h_0^{p,q}$ is constant for small t. Then, according to Kodaira and Spencer's upper semi-continuity theorem, F_t and G_t are differential in t.

Now we let

$$\Phi_t = \partial_t \bar{\partial}_t \bar{\partial}_t^* \partial_t^* G_t \omega_t^{n-1} + F_t \omega_t^{n-1},$$

and let

$$\tilde{\Omega}_t = \frac{1}{2}(\Phi_t + \bar{\Phi}_t).$$

Then for small t, $\tilde{\Omega}_t$ are real $(n-1, n-1)$-forms depending smoothly on t and $d\tilde{\Omega}_t = 0$. So we only need to prove $\tilde{\Omega}_t$ is strictly positive. If we can prove that $\tilde{\Omega}_0 = \omega_0^{n-1}$, then $\tilde{\Omega}_0$ is strictly positive, and by continuity again, $\tilde{\Omega}_t$ is also positive for small t. Therefore, we can find a balanced metric $\tilde{\omega}_t$ on X_t such that $\tilde{\omega}_t^{n-1} = \tilde{\Omega}_t$.

So we only need to prove $\tilde{\Omega}_0 = \omega_0^{n-1}$. From the definition of E_0, we can write formally

$$\begin{aligned} \omega_0^{n-1} =& E_0 G_0 \omega_0^{n-1} + F_0 \omega_0^{n-1} \\ =& \partial_0 \bar{\partial}_0 \bar{\partial}_0^* \partial_0^* G_0 \omega_0^{n-1} + \partial_0^* \varphi + \bar{\partial}_0^* \psi + F_0 \omega_0^{n-1} \end{aligned}$$

for some forms φ and ψ. Hence,

$$\partial_0^* \varphi + \bar{\partial}_0^* \psi = \omega_0^{n-1} - \partial_0 \bar{\partial}_0 \bar{\partial}_0^* \partial_0^* G_0 \omega_0^{n-1} - F_0 \omega_0^{n-1}.$$

Since the right hand side of the above equality is ∂_0-closed and $\bar{\partial}_0$-closed,

$$\int_{X_0} \langle \partial_0^* \varphi + \bar{\partial}_0^* \psi, \partial_0^* \varphi + \bar{\partial}_0^* \psi \rangle_{\omega_0} = 0.$$

Thus, $\partial_0^* \varphi + \bar{\partial}_0^* \psi = 0$ and

$$\omega_0^{n-1} = \partial_0 \bar{\partial}_0 \bar{\partial}_0^* \partial_0^* G_0 \omega_0^{n-1} + F_0 \omega_0^{n-1}.$$

As ω_0^{n-1} is real, by the definition of $\tilde{\Omega}_0$, we have $\tilde{\Omega}_0 = \omega_0^{n-1}$. □

7 Small deformations of twistor spaces

This section is newly added. We will weaken the $\partial\bar{\partial}$-lemma condition such that the existence of balanced metrics is still preserved under small deformations. First we give

Definition 26. *A compact complex n-dimensional manifold satisfies the* $(n-1,n)$-th weak $\partial\bar{\partial}$-lemma *if for its any real $(n-1,n-1)$-form φ such that $\bar{\partial}\varphi$ is a ∂-exact form, there exists an $(n-2,n-1)$-form ψ such that*

$$\bar{\partial}\varphi = i\partial\bar{\partial}\psi. \tag{7.1}$$

Then we have

Theorem 27. [18] *Let X_t be a complex analytic family of compact complex manifolds. Suppose X_t satisfies the $(n-1,n)$-th weak $\partial\bar{\partial}$-lemma for small $t \neq 0$. If X_0 admits a balanced metric, then X_t also admits a balanced metric for small enough t.*

Certainly the condition of the $(n-1,n)$-th weak $\partial\bar{\partial}$-lemma is weaker than the $\partial\bar{\partial}$-lemma. According to theorem 24, if X satisfies the $\partial\bar{\partial}$-lemma, then X_t also satisfies the $\partial\bar{\partial}$-lemma for small t. So theorem 27 is stronger than theorem 25.

We should check that the small deformation of the Iwasawa manifold constructed in section 5 does not satisfy the $(n-1,n)$-th weak $\partial\bar{\partial}$-lemma. Actually, if we use the notations of section 5, we find that on X_t for $t \neq 0$,

$$\bar{\partial}_t\big(\phi_{1,t}\wedge\bar{\phi}_{1,t}\wedge\phi_{3,t}\wedge\bar{\phi}_{3,t}\big) = t\partial_t\big(\phi_{3,t}\wedge\bar{\phi}_{1,t}\wedge\bar{\phi}_{2,t}\wedge\bar{\phi}_{3,t}\big) \tag{7.2}$$

can not be written as a $\partial_t\bar{\partial}_t$-exact form.

An application of theorem 27 is

Corollary 28. [18] *Let X be a compact complex manifold with a balanced metric. If $H^{2,0}(X,\mathbb{C}) = 0$, then any small deformation of X admits a balanced metric.*

Proof. Since the function $h^{p,q}(t) = \dim H^{p,q}(X_t,\mathbb{C})$ is upper semicontinuous in t, $h^{2,0}(t) \leq h^{2,0}(0) = 0$ for small t. So $H^{2,0}(X_t,\mathbb{C}) = 0$, and by Serre duality, $H^{n-2,n}(X_t,\mathbb{C}) = 0$. Now let φ_t be a real $(n-1,n-1)$-form on X_t such that $\bar{\partial}_t\varphi_t$ is a ∂_t-exact form, i.e., there exists an $(n-2,n)$-form η_t such that $\bar{\partial}_t\varphi_t = \partial_t\eta_t$. Since $\bar{\partial}_t\eta_t = 0$ and $H^{n-2,n}(X_t,\mathbb{C}) = 0$, there exists an $(n-2,n-1)$-form ψ_t such that $\eta_t = i\bar{\partial}_t\psi_t$. Hence $\bar{\partial}_t\varphi_t = i\partial_t\bar{\partial}_t\psi_t$. Therefore X_t satisfies the $(n-1,n)$-th weak $\partial\bar{\partial}$-lemma and the corollary follows. □

We can use above corollary to twistor spaces. Given an oriented riemannian 4-manifold M, there is associated a 6-manifold Z which is called the twistor space of M. This is an S^2 bundle over M whose fiber over a point is the set of all unit anti-self-dual 2-forms.

The twistor space Z has a canonical almost complex structure. M. F. Atiyah, N. Hitchin and I. Singer [5] proved that the canonical almost complex structure on the twistor space Z over M is integrable if and only if M is self-dual. Hitchin [22] then showed that the only compact twistor spaces which are Kähler are those associated to S^4 and $\mathbb{C}P^2$. On the other hand, the natural metric on the twistor space over a compact self-dual four manifold is balanced (c.f. [19]). Moreover, M. Eastwood and M. Singer [11] observed that for any twistor space Z, $H^{2,0}(Z, \mathbb{C}) = 0$. Thus, corollary 28 implies

Corollary 29. [18] *Any small deformation of the twistor space over a compact self-dual four manifold admits a balanced metric.*

8 Some existence theorems

Michelsohn characterized the notion of balanced metrics in terms of positive currents; more precisely, she proved

Theorem 30. [27] *A compact complex manifold is balanced if and only if there exists no non-zero positive current T of degree $(1,1)$, such that T is the $(1,1)$-component of a boundary.*

Michelsohn then used this theorem to the following situation. Let $f : X \to C$ be a holomorphic map from a compact complex manifold onto a complex curve with irreducible fibers. f is called *essential* if the pull back $f^*\omega_C$ of a Kähler matric ω_C on C is not a $(1,1)$-boundary. Clearly f being essential is a necessary condition of X being balanced.

Theorem 31. [27] *If a compact complex manifold admits an essential holomorphic map onto a complex curve such that its smooth fibers are balanced, then X is a balanced manifold.*

Note that this theorem gives another proof of existence of balanced metrics on the complex manifolds X_Λ constructed in section 3, see [27].

The next important result on the existence of balanced metrics was given by Alessandrini and Bassanelli. Using theorem 30, they proved that the class of compact balanced manifolds is invariant under modification.

Theorem 32. [2, 3] *Let $\bar{X}$ and X be compact complex manifolds. Let $f : \bar{X} \to X$ be a modification. Then $\bar{X}$ is balanced if and only if X is balanced.*

A modification $f : \bar{X} \to X$ is a holomorphic map such that, for a suitable analytic subset Y in X, $E \triangleq f^{-1}(Y)$ is a hypersurface and $f|_{\bar{X}-E} : \bar{X} - E \to X - Y$ is a biholomorphism. Thus a compact manifold of Fujiki class C is balanced since it is bimeromorphic to a compact Kähler manifold.

Alessandrini and Bassanelli also proved

Theorem 33. [4] *If X is a compact complex manifold of dimension $n \geq 3$ and is Kähler outside an irreducible curve, then X is balanced.*

9 Examples: non-Kähler Calabi-Yau threefolds

Definition 34. *A smooth Calabi-Yau threefold is a compact complex three dimensional manifold with finite fundamental group and trivial canonical bundle.*

Thus, a smooth Calabi-Yau threefold may be *Kähler* or *non-Kähler*. If a Calabi-Yau threefold is Kähler, then according to Yau's solution to the Calabi conjecture, there exists a unique Ricci-flat Kähler metric in each Kähler class of the threefold. Such metrics are known as Calabi-Yau metrics.

On the other hand, a large class of non-Kähler Calabi-Yau threefolds can be obtained by the conifold transition from Kähler Calabi-Yau threefolds. The conifold transition was developed by Clemens and Friedman.

Let Y be a smooth Kähler Calabi-Yau threefold that contains a collection of mutually disjoint $(-1,-1)$-curves $E_1, \ldots, E_l$; these are smooth, isomorphic to $\mathbb{C}P^1$ and have normal bundles isomorphic to

$$\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{C}P^1.$$

By contracting all E_i, we obtain a singular Calabi-Yau threefold (a conifold) X_0 with l ordinary double points $p_1, \ldots, p_l$. Then there exists a holomorphic map $\varphi : Y \to X_0$ such that

$$\varphi|_{Y-(\cup_{i=1}^l E_i)} : Y - (\cup_{i=1}^l E_i) \to X_0 - \{p_1, \ldots, p_l\}$$

is a biholomorphism. We have the following facts:

Theorem 35. (1) [13] *There is an infinitesimal smoothing of X_0 if and only if the fundamental classes $[E_i]$ in $H_2(Y, \mathbb{Z})$ satisfy a relation*

$$\Sigma_i n_i [E_i] = 0 \quad \text{such that } n_i \neq 0 \text{ for every } i. \tag{9.1}$$

(2) [23, 31] *The infinitesimal smoothing can always be realized by a real smoothing, i.e., X_0 can be smoothed to a complex analytic family X_t.*

(3) [13] $\pi_1(X_t) = \pi_1(Y)$ *and the canonical bundle of X_t is also trivial.*

So under the condition (9.1), the geometric transition:

$$Y \to X_0 \dashrightarrow X_t$$

can always be carried out and is called a conifold transition. Moreover, according to definition 34, X_t are also Calabi-Yau threefolds. In general, X_t are non-Kähler. For example, Friedman observed that $\#_k(S^3 \times S^3)$ for any $k \geq 2$ can be given a complex structure in this way (Friedman: $k \geq 103$ and Lu-Tian [26]: $2 \leq k \leq 102$). Since the Hodge number $h^{1,1}$ of these manifolds is zero, they can not be Kähler. However, we have

Theorem 36. [14] *Let Y be a smooth Kähler Calabi-Yau threefold and let $Y \to X_0 \dashrightarrow X_t$ be a conifold transition. Then for sufficiently small t, X_t admits a smooth balanced metric.*

Corollary 37. [14] *The complex structure on $\#_k(S^3 \times S^3)$ for any $k \geq 2$ constructed from the conifold transition admits a balanced metric.*

Combining this corollary with the discussions in [7], we also have

Corollary 38. [7, 14] *There exists no $\partial\bar{\partial}$-closed hermitian metric on the complex structure on $\#_k(S^3 \times S^3)$ for any $k \geq 2$ constructed from the conifold transition.*

The conifold transition is important in string theory. P. S. Green and T. Hübsch proved that the conifold transition provides a connection between all known Kähler Calabi-Yau compactifications in string theory. This result supports the Reid conjecture that Kähler Calabi-Yau threefolds may have a universal moduli space even though they are of different homotopy types. Furthermore, Hübsch thought that threefolds $\#_{k\geq 2}(S^3 \times S^3)$ should be the universal covering spaces of *all* Kähler Calabi-Yau threefolds in some sense.

10 Superstring theory

In heterotic string theory, the internal space X is a compact complex three-dimensional manifold with trivial canonical bundle, i.e., with a non-vanishing holomorphic three-form $\mathcal{V}$. It also involves a holomorphic vector bundle V over X. Let ω be a hermitian metric on X and H a hermitian metric on V. In 1986, Strominger [30] proposed a system for (ω, H):

$$d(\| \mathcal{V} \|_\omega \, \omega^2) = 0;$$

$$F_H^{2,0} = F_H^{0,2} = 0, \quad F_H \wedge \omega^2 = 0;$$

$$\sqrt{-1}\partial\bar{\partial}\,\omega = \frac{\alpha'}{4}\left(\mathrm{tr}\big(R_\omega \wedge R_\omega\big) - \mathrm{tr}\big(F_H \wedge F_H\big)\right).$$

The first equation says that the metric ω is a conformal balanced metric. So the concept of balanced metric plays an important role in superstring theory. The second equation is the hermitian-Yang-Mills equation. The existence of its solution is, by the Li-Yau theorem which is the non-Kähler version of the Donaldson-Uhlenbeck-Yau theorem, equivalent to that V is stable with respective to the conformal balanced metric ω. The third equation is called the anomaly equation. Following Strominger, we take the curvature R in the third equation to be defined by the hermitian connection. Thus the term $tr(R \wedge R)$ is always a $(2,2)$-form.

When V is the holomorphic tangent bundle TX and ω is Kähler, (ω, H) is a solution to the Strominger system if and only if $\omega = H$ is a Calabi-Yau metric. So this system should be viewed as a generalization of the Calabi conjecture for the case of non-Kähler Calabi-Yau threefolds with balanced metrics.

The existence of smooth solutions of the Strominger system has been studied since 2004. Using the perturbation method, J. Li and S.-T. Yau [25] constructed irreducible smooth solutions to a class of Kähler Calabi-Yau threefolds on some $U(4)$ and $U(5)$ principle bundles. Shortly after, Fu and Yau constructed solutions to this system on a class of non-Kähler Calabi-Yau threefolds. Their solutions were orbifolded by M. Becker, L.-S. Tseng and Yau [6] to give many more solutions. Fu, Tseng and Yau [15] also presented explicit solutions on T^2-bundles over the Eguchi-Hanson space. We note further that nilmanifold solutions with different connections have been discussed recently in [12]. Now more solutions have been obtained.

Fu and Yau constructed their solutions on a class of torus bundles X over a $K3$-surface Y twisted by two d-closed anti-self-dual real $(1,1)$-forms ω_1 and ω_2, which have been mentioned in section 4. On such manifolds, we have showed that the natural metric

$$\omega_0 = \pi^*\omega_Y + i\theta \wedge \bar{\theta}$$

is a balanced metric. There also exists a non-vanishing holomorphic three form

$$\mathcal{V} = \pi^*(\mathcal{V}_Y) \wedge \theta,$$

where $\mathcal{V}_Y$ is a non-vanishing holomorphic two form on the $K3$ surface Y.

Moreover, one can define a hermitian metric on X:

$$\omega_u = \pi^*(e^u \omega_Y) + \sqrt{-1}\, \theta \wedge \bar{\theta}.$$

Here u is an arbitrary function on Y. This metric need not be a balanced metric. The key point is that for any function u, the metric ω_u still satisfies the first equation of the Strominger system, see proposition 16.

Next we consider the second equation. Take a stable vector bundle E over the $K3$ surface with respect to the metric ω_Y. By the Donaldson-Uhlenbeck-Yau theorem, there exists a hermitian-Yang-Mills metric H on E, i.e. its hermitian curvature F_H satisfies

$$F_H \wedge \omega_Y = 0.$$

Thus, $\pi^* F_H \wedge \omega_u^2$ is also zero. This means that $\pi^* H$ is a hermitian-Yang-Mills metric on $V = \pi^* E \to X$ with respect to any conformal balanced metric ω_u. So given a stable vector bundle E over the $K3$ surface Y, the second equation for the vector bundle $V = \pi^* E$ can always be solved for any metric ω_u.

Therefore, we only need to consider the third equation. Certainly the term $\mathrm{tr} F_H \wedge F_H$ is a $(2,2)$-form defined on the $K3$ surface. For the metric ω_u, by explicit calculation, we found that the terms $\mathrm{tr}(R_{\omega_u} \wedge R_{\omega_u})$ and $\sqrt{-1}\partial\bar{\partial}\omega_u$ are also defined on the surface. Thus we reduced the third equation to the following Monge-Ampere equation defined on the $K3$ surface:

$$\triangle(e^u - \frac{\alpha'}{2} f e^{-u}) + 4\alpha' \frac{\det u_{i\bar{j}}}{\det g_{i\bar{j}}} + \mu = 0,$$

where f and μ are two functions on the $K3$ surface satisfying $f \geq 0$ and $\int_Y \mu \omega_Y^2 = 0$. The last compatibility condition is equivalent to the condition

$$\alpha'(24 - c_2(E)) + Q(\omega_1/2\pi) + Q(\omega_2/2\pi) = 0. \tag{10.1}$$

Here 24 stands for the second Chern number of the $K3$ surface and $Q(\omega_i/2\pi)$, for $i = 1, 2$, denotes the intersection number of $\omega_i/2\pi$. The above equation can be solved by the continuity method. The estimate of the volume form is very complicated. Summarizing the above discussion we have

Theorem 39. [17] *Let Y be a $K3$ surface with a Calabi-Yau metric ω_Y. Let ω_1 and ω_2 be d-closed anti-self-dual real $(1,1)$-forms on Y such that $[\omega_1/2\pi], [\omega_2/2\pi] \in H^2(Y, \mathbb{Z})$. Let X be the T^2-bundle over Y twisted by ω_1 and ω_2. Let E be a stable bundle over Y with gauge group $SU(r)$. Suppose ω_1, ω_2 and $c_2(E)$ satisfy the topological constraint (10.1). Then there exist a smooth function u on Y and a hermitian-Yang-Mills metric H on E such that (ω_u, H) is a solution of the Strominger system.*

References

[1] L. Alessandrini and G. Bassanelli, Small deformations of a class of compact non-Kähler manifolds, Proc. Amer. Math. Soc. 109(1990), 1059–1062.

[2] L. Alessandrini and G. Bassanelli, Metric properties of manifolds bimeromorphic to compact Kähler spaces, J. Differential Geom. 37(1993), 95–121.

[3] L. Alessandrini and G. Bassanelli, Modifications of compact balanced manifolds, C. R. Acad. Sci. Paris Sér. I Math. 320(1995), 1517–1522.

[4] L. Alessandrini and G. Bassanelli, Wedge product of positive currents and balanced manifolds, Tohoku Math. J. 60(2008), 123–134.

[5] M. F. Atiyah, N. Hitchin and I. Singer, Self-duality in four-dimensional Riemannian geometry, Proc. Roy. Soc. London Ser. A 362(1978), 425–461.

[6] M. Becker, L.-S. Tseng and S.-T. Yau, New heterotic non-Kähler geometries, Adv. Theor. Math. Phys. 13(2009), 1815–1845.

[7] Y. Bozhkov, The specific Hermitian geometry of certain three-folds, Riv. Mat. Univ. Parma 4(1995), 61–68.

[8] E. Calabi, Construction and properties of some 6-dimensional almost complex manifolds, Trans. Amer. Math. Soc. 87(1958), 407–438.

[9] E. Calabi and B. Eckmann, A class of compact complex manifolds which are not alegbraic, Ann. of Math. 58(1953), 494–500.

[10] P. Deligne, P. Griffiths, J. Morgan and D. Sullivan, Real homotopy theory of Kähler manifolds, Invent. Math. 29(1975), 245–274.

[11] M. Eastwood and M. Singer, The Frölicher spectral sequence on a twistor space, J. Differential Geom. 38(1993), 653–669.

[12] M. Fernández, S. Ivanov, L. Ugarte and R. Villacampa, Non-Kaehler heterotic string compactifications with non-zero fluxes and constant dilaton, Comm. Math. Phys. 288(2009), 677–697.

[13] R. Friedman, On threefolds with trivial canonical bundle, Complex geometry and Lie theory (Sundance, UT, 1989), 103–134, Proc. Sympos. Pure Math. 53, Amer. Math. Soc., Providence, RI, 1991.

[14] J. Fu, J. Li and S.-T. Yau, Balanced metrics on non-Kähler Calabi-Yau threefolds, A new version of arXiv:0809.4748v1[math.DG].

[15] J. Fu, L.-S. Tseng and S.-T. Yau, Local heterotic torsional models, Comm. Math. Phys. 289(2009), 1151–1169.

[16] J. Fu, Z. Wang and D. Wu, Form-type Calabi-Yau equations on Kähler manifolds of nonnegative orthogonal bisectional curvature, arXiv:1010.2022v2 [math.DG].

[17] J. Fu and S.-T. Yau, The theory of superstring with flux on non-Kähler manifolds and the complex Monge-Ampère equation, J. Differential Geom. 78(2008), 369–428.

[18] J. Fu and S.-T. Yau, A note on small deformations of balanced manifolds, arXiv:11050026v2[math.DG].

[19] P. Gauduchon, Structures de Weyl et théorèmes d'annulation sur une variété conforme autoduale, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 18(1991), 563–629.

[20] E. Goldstein and S. Prokushkin, Geometric model for complex non-Kähler manifolds with $SU(3)$ structure, Comm. Math. Phys. 251(2004), 65–78.

[21] A. Gray, Some examples of almost Hermitian manifolds, Illinois J. Math. 10(1966), 353–366.

[22] N. Hitchin, Kählerian twistor spaces, Proc. London Math. Soc. 43(1981), 133–150.

[23] Y. Kawamata, Unobstructed deformations, A remark on a paper of Z. Ran: "Deformations of manifolds with torsion or negative canonical bundle", J. Algebraic Geom. 1(1992), 183–190.

[24] K. Kodaira and D. C. Spencer, On deformations of complex analytic structures, III, Stability theorems for complex structures, Ann. of Math. (2) 71(1960), 43–76.

[25] J. Li and S.-T. Yau, The existence of supersymmetric string theorey with torsion, J. Differential Geom. 70(2005), 143–181.

[26] P. Lu and G. Tian, The complex structures on connected sums of $S^3 \times S^3$, Manifolds and geometry (Pisa, 1993), 284–293, Sympos. Math. XXXVI, Cambridge Univ. Press, Cambridge, 1996.

[27] M. L. Michelsohn, On the existence of special metrics in complex geometry, Acta Math. 149(1982), 261–295.

[28] J. Morrow and K. Kodaira, Complex Manifolds, Holt, Rinehart and Winston, Inc., New York-Montreal, Que.-London, 1971.

[29] Y.-T. Siu, Analyticity of sets associated to Lelong numbers and the extension of closed positive currents, Invent. Math. 27(1974), 53–156.

[30] A. Strominger, Superstings with Torsion, Nuclear Phys. B 274(1986), 253–284.

[31] G. Tian, Smoothing 3-folds with trivial canonical bundle and ordinary double points, Essays on mirror manifolds, 458–479, Int. Press, Hong Kong, 1992.

[32] C. Voisin, Hodge theory and complex algecbraic geometry, I, Translated from the French original by Leila Schnels. Translated from the French by Leila Schneps, Reprint of the 2002 English ed., Cambridge Studies in Advanced Mathematics 76. Cambridge Univ. Press, Cambridge, 2007.

[33] C.-C. Wu, On the geometry of superstrings with torsion, Thesis (Ph.D.)-Harvard Univ. 2006, 60 pp.

[34] S.-T. Yau, On the Ricci curvature of a compact Kähler manifold and the complex Monge-Ampère equation, I, Comm. Pure Appl. Math. 31(1978), 339–411.

Institute of Mathematics
Fudan University
Shanghai 200433
China

E-mail address: majxfu@fudan.edu.cn

GEOMETRIC ANALYSIS ON ALEXANDROV SPACES

Takashi SHIOYA

1 Introduction

An Alexandrov space is a metric space with a lower curvature bound. This is a natural geometric object and also is important in connection with the study of the topology of Riemannian manifolds. In fact, Alexandrov spaces have been used in G. Perelman's proof of the geometrization conjecture.

In this article, we explain some ideas in the study of geometric analysis on Alexandrov spaces of curvature bounded from below. This article is written in a friendly but rough style, so the author suggests that the reader does not try to understand every details of proofs and, instead, just tries to grasp their ideas. We refer [S2] for the formal and more complete article on the same subject.

The organization of this article is as follows. In §2, we review the basics for Alexandrov spaces, which was established by Burago-Gromov-Perelman [BGP]. In §3, we prove that the set of singular points in an Alexandrov space is of measure zero, and in §4, we see the differentiable structure and the Riemannian metric studied by Otsu-Shioya [OS], Perelman [P] and Kuwae-Machigashira-Shioya [KMS]. In §5, we discuss the infinitesimal Bishop-Gromov condition, which is useful to prove a Poincaré inequality as seen in §6. We see, in §7, that the infinitesimal Bishop-Gromov condition implies the Laplacian comparison for the distance function, which is applied to prove a splitting theorem in §8.

2 Basics

We begin with the following theorem for a Riemannian manifold. Denote by $M^2(\kappa)$ a complete simply connected two-dimensional space form of constant curvature κ.

Theorem 2.1 (Triangle Comparison Theorem (Alexandrov-Toponogov)). *Let M be a complete Riemannian manifold and κ a real number. Then the following* (1) *and* (2) *are equivalent to each other.*

(1) *The sectional curvature K_M of M satisfies $K_M \geq \kappa$ on M.*

(2) *For any minimal geodesic triangle $\triangle pqr$ in M and any point s on the edge qr there exists a triangle $\triangle \tilde{p}\tilde{q}\tilde{r}$ in $M^2(\kappa)$ such that $d(p,q) = d(\tilde{p},\tilde{q})$, $d(q,r) = d(\tilde{q},\tilde{r})$, $d(r,p) = d(\tilde{r},\tilde{p})$, and such that, if $\tilde{s}$ is a point on the edge $\tilde{q}\tilde{r}$ with $d(q,s) = d(\tilde{q},\tilde{s})$, then we have*

$$d(p,s) \geq d(\tilde{p},\tilde{s})$$

(see Figure 1).

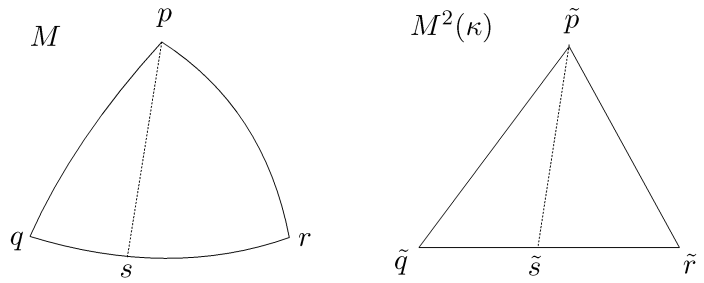

Figure 1: Triangle comparison condition

We call the condition (2) in the theorem the *triangle comparison condition.* For a given triangle $\triangle pqr$ in M, we call $\triangle \tilde{p}\tilde{q}\tilde{r}$ in $M^2(\kappa)$ as in (2) a *comparison triangle* of $\triangle pqr$.

Definition 2.2. A complete metric space X is called an *Alexandrov space of curvature* $\geq \kappa$ if the following (i), (ii), and (iii) are satisfied.

(i) For any two points $x, y \in X$, there exists a length-minimizing curve γ joining x and y such that $L(\gamma) = d(x,y)$, where $L(\gamma)$ is the length of γ. Such a γ is called a *minimal geodesic.*

(ii) X satisfies the triangle comparison condition for κ.

(iii) $\dim_H X < +\infty$, where $\dim_H X$ is the Hausdorff dimension of X.

It is known that the Hausdorff dimension of an Alexandrov space coincides with its covering dimension, which is always an integer. The *dimension* of an Alexandrov space means the Hausdorff dimension. By the finiteness of the dimension, an Alexandrov space is always a proper metric space, i.e., any bounded subset of it is relatively compact.

Examples-Propositions (cf. [BGP]).

(1) *Complete Riemannian manifolds of sectional curvature $\geq \kappa$ are Alexandrov spaces of curvature $\geq \kappa$ by Theorem 2.1.*

(2) *The boundary of a convex body in* $\mathbf{R}^n$ *is an Alexandrov space of nonnegative curvature.*

(3) *The quotient of an Alexandrov space by an isometric group action is an Alexandrov space.*

(4) *Let* $\{X_i\}$ *be a sequence of* n*-dimensional Alexandrov spaces of curvature* $\geq \kappa$*. If* X_i *converges to a proper metric space* X *in the (pointed) Gromov-Hausdorff topology, then the limit* X *is an Alexandrov space of curvature* $\geq \kappa$ *and dimension* $\leq n$*.*

(5) *The Euclidean cone over an Alexandrov space of curvature* ≥ 1 *is an Alexandrov space of curvature* ≥ 0*.*

The *Euclidean cone over a metric space* X is defined to be the cone $[0,+\infty)\times X/(0\times X)$ with the metric

$$d((s,u),(t,v)) := \sqrt{s^2+t^2-2st\cos d(u,v)}, \qquad (s,u),(t,v)\in[0,+\infty)\times X.$$

The point in the cone corresponding to $0\times X$ is called the *vertex* of the cone. Note that X is isometric to an $(n-1)$-dimensional unit sphere S^{n-1} in $\mathbf{R}^n$ with Riemannian distance if and only if its Euclidean cone is isometric to $\mathbf{R}^n$, where $n\geq 1$. A 0-dimensional unit sphere consists of two points with distance π.

Let X be an Alexandrov space of curvature $\geq \kappa$. For a given triangle $\triangle pqr$ in X, we denote by $\tilde{\angle} pqr$ the angle $\angle\tilde{p}\tilde{q}\tilde{r}$ between $\tilde{q}\tilde{p}$ and $\tilde{q}\tilde{r}$ for a comparison triangle $\triangle\tilde{p}\tilde{q}\tilde{r}$ in $M^2(\kappa)$. By the triangle comparison condition, we have

$$\tilde{\angle} pqr \leq \tilde{\angle} pqs,$$

where $s\in qr$ and $\tilde{s}\in\tilde{q}\tilde{r}$ are such that $d(q,s)=d(\tilde{q},\tilde{s})$, as in the triangle comparison condition. This implies that, for any two minimal geodesics $\sigma:[0,l]\to X$ and $\tau:[0,l']\to X$ with $\sigma(0)=\tau(0)=p$, $\tilde{\angle}\sigma(s)\,p\,\tau(t)$ is monotone non-increasing in $s,t>0$.

Definition 2.3. The *angle between* σ *and* τ is defined by

$$\angle(\sigma,\tau)=\angle qpr := \lim_{s,t\to 0+}\tilde{\angle}\sigma(s)\,p\,\tau(t),$$

where $q:=\sigma(l)$, $r:=\tau(l')$.

Clearly, $\angle(\cdot,\cdot)$ is symmetric, i.e., $\angle(\sigma,\tau)=\angle(\tau,\sigma)$. We have the triangle inequality $\angle(\sigma,\gamma)\leq\angle(\sigma,\tau)+\angle(\tau,\gamma)$ for any minimal geodesics σ, τ, and γ emanating from a point p (see [BBI, Theorem 3.6.34]). It is easy to see that $\angle(\sigma,\tau)=0$ if and only if $\sigma\subset\tau$ or

$\tau \subset \sigma$, where σ and τ indicate their images in X. The relation $\angle(\sigma, \tau) = 0$ between two minimal geodesics σ and τ from p is an equivalence relation and the quotient

$$\Sigma'_p := \{ \text{ minimal geodesics from } p \,\}/(\angle = 0)$$

becomes a metric space, which is not necessarily complete.

Definition 2.4. We denote the completion of Σ'_p by Σ_p and call it the *space of directions at* p. The *tangent cone at* p, say K_p, is defined to be the Euclidean cone over Σ_p, i.e., for $(s, u), (t, v) \in K_p = [\,0, +\infty\,) \times \Sigma_p/(0 \times \Sigma_p)$,

$$d((s, u), (t, v)) := \sqrt{s^2 + t^2 - 2st \cos \angle(u, v)}.$$

We denote the vertex of K_p by o_p.

The space of directions and the tangent cone are generalizations of the unit tangent sphere and the tangent space respectively.

Denoting by λX, $\lambda > 0$, the space X with the metric multiplied by λ times, we have the following.

Theorem 2.5 ([BGP]). *For any point $p \in X$, as $\lambda \to +\infty$, the scaled pointed space $(\lambda X, p)$ converges to the tangent cone (K_p, o_p) at p in the sense of Gromov-Hausdorff convergence.*

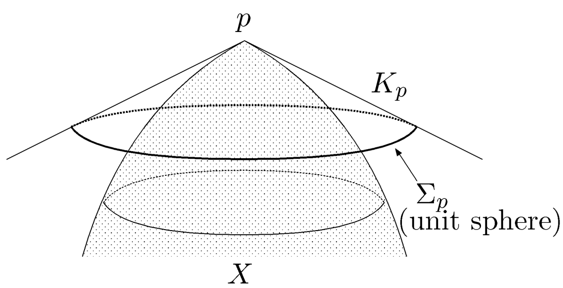

Figure 2: The space of directions and the tangent cone

Theorem 2.5 together with some discussion leads us to the following. We omit the details.

Corollary 2.6 ([BGP]). (1) *The tangent cone K_p at any point $p \in X$ is an Alexandrov space of nonnegative curvature and* $\dim K_p = \dim X$.

(2) *The space of directions Σ_p at any point $p \in X$ is a compact Alexandrov space of curvature ≥ 1 and* $\dim \Sigma_p = \dim X - 1$.

3 Singularity

Let X be an n-dimensional Alexandrov space of curvature $\geq \kappa$. A point $p \in X$ is said to be *singular* if the space of directions Σ_p is isometric to an $(n-1)$-dimensional unit sphere S^{n-1} in $\mathbf{R}^n$ (or equivalently K_p is isometric to $\mathbf{R}^n$). Denote by S_X the set of singular points in X. It is a theorem that $\dim_H S_X \leq n-1$ [BGP, OS]. The purpose of this section is to prove:

$$\mathcal{H}^n(S_X) = 0,$$

which is weaker than $\dim_H S_X \leq n-1$.

Definition 3.1 (Radial expansion map)**.** For $p \in X$ and $0 < t \leq 1$, we define a subset $W_{p,t} \subset X$ and a map $\Phi_{p,t} : W_{p,t} \to X$ as follows. We first set $\Phi_{p,t}(p) := p \in W_{p,t}$. A point x $(\neq p)$ belongs to $W_{p,t}$ if and only if there exists $y \in X$ such that $x \in py$ and $r_p(x) : r_p(y) = t : 1$, where py is a minimal geodesic from p to y. Since a geodesic does not branch on an Alexandrov space, for a given point $x \in W_{p,t}$ such a point y is unique and we set $\Phi_{p,t}(x) := y$ (see Figure 3). The triangle comparison condition implies the local Lipschitz continuity of the map $\Phi_{p,t} : W_{p,t} \to X$. We call $\Phi_{p,t}$ the *radial expansion map.*

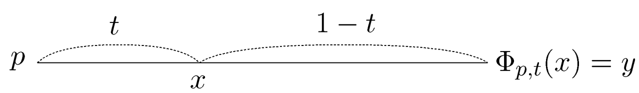

Figure 3: The radial expansion map

Lemma 3.2. $\Phi_{p,t}|_{W_{p,t}\cap B(p,R)} : W_{p,t} \cap B(p,R) \to X$ *is a Lipschitz map with Lipschitz constant* $\leq 1 + \theta_{\kappa,R}(1-t)$*, where* $\theta_{\kappa,R}$ *is some function such that* $\theta_{\kappa,R}(\delta) \to 0$ *as* $\delta \to 0$.

Proof. With the notations as in Figure 4, the triangle comparison condition implies

$$d(x_1, x_2) \geq d(\tilde{x}_1, \tilde{x}_2) \geq (1 - \theta_{\kappa,R}(1-t))\, d(\tilde{y}_1, \tilde{y}_2),$$

which completes the proof. □

Definition 3.3 (Cut locus)**.** The *cut-locus of* $p \in X$ is defined by

$$\mathrm{Cut}_p := \{\ x \in X \mid px \text{ does not extend beyond } x \\ \text{as a minimal geodesic.}\ \}$$

It is easy to see that $\bigcup_{0<t<1} W_{p,t} = X \setminus \mathrm{Cut}_p$ and $W_{p,t_1} \subset W_{p,t_2}$ for any t_1, t_2 with $t_1 < t_2$.

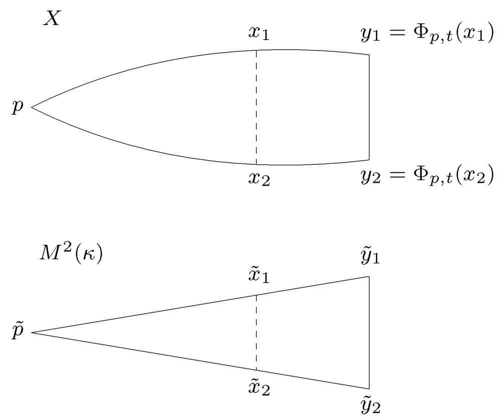

Figure 4:

Proposition 3.4 ([OS]). *For any point $p \in X$ we have*

$$\mathcal{H}^n(\mathrm{Cut}_p) = 0.$$

Proof. For any $R > 0$ and $0 < t < 1$,

$$\Phi_{p,t}(W_{p,t} \cap B(p, tR)) \supset B(p, R).$$

By Lemma 3.2,

$$\mathcal{H}^n(W_{p,t} \cap B(p, tR)) \geq (1 + \theta_{\kappa,R}(1-t))^n \mathcal{H}^n(B(p, R)).$$

As $t \to 1$, we have

$$\mathcal{H}^n(B(p, R) \setminus \mathrm{Cut}_p) \geq \mathcal{H}^n(B(p, R)),$$

so that $\mathcal{H}^n(B(p, R) \cap \mathrm{Cut}_p) = 0$ for all $R > 0$. This completes the proof. □

We use the following theorem.

Theorem 3.5 (Milka). *If an nonnegatively curved Alexandrov space contains a straight line, then it is isometric to the product $Y \times \mathbf{R}$ for some Alexandrov space Y.*

Proof of "$\mathcal{H}^n(S_X) = 0$". Since X is proper, there is a dense countable subset $\{p_i\}_{i=1}^\infty$ of X. We take any point $x \in X \setminus \bigcup_{i=1}^\infty \mathrm{Cut}_{p_i}$. For each i, the minimal geodesic $p_i x$ extends beyond x, so that K_x has a straight line γ_i corresponding to $p_i x$. It follows that $\bigcup_i \gamma_i \subset K_x$ is dense. By Theorem 3.5, K_x is isometric to $\mathbf{R}^n$, so that x is non-singular. Therefore we have $S_X \subset \bigcup_{i=1}^\infty \mathrm{Cut}_{p_i}$, which together with Proposition 3.4 completes the proof. □

4 Differentiable and Riemannian structure

Let X be an n-dimensional Alexandrov space and set, for $\delta > 0$,

$$X_\delta := \{\, x \in X \mid d_{GH}(\Sigma_x, S^{n-1}) < \delta \,\},$$

where d_{GH} denotes the Gromov-Hausdorff distance function. The purpose in this section is to show the idea of the proof of the following theorem.

Theorem 4.1 ([OS, BGP, KMS]). (1) *There exists a positive real number $\delta(n)$ depending only on the dimension n of X such that $X_{\delta(n)}$ is a C^∞-manifold.*

(2) *There exists a unique continuous Riemannian metric g on $X_{\delta(n)} \setminus S_X$ such that the distance function induced from g coincides with the original one.*

Definition 4.2. A finite sequence of points $\{p_i\}_{i=\pm 1,\dots,\pm n}$ in X is called a *δ-strainer* at a point $x \in X$, $\delta > 0$, if we have

$$\tilde{\angle} p_i x p_j > \begin{cases} \frac{\pi}{2} - \delta & \text{for } i \neq \pm j, \\ \pi - \delta & \text{for } i = -j. \end{cases}$$

If there is a δ-stainer at a point $x \in X$, we say that x is *δ-strained*. The map $\varphi : X \to \mathbf{R}^n$ defined by

$$\varphi(y) := (d(p_1, y), \dots, d(p_n, y)) \in \mathbf{R}^n, \quad y \in X,$$

is called the *strain map* associated with the δ-strainer $\{p_i\}_{i=\pm 1,\dots,\pm n}$.

For example, for an orthonormal basis $\{e_i\}_{i=1}^n$ on a Euclidean space $\mathbf{R}^n$, by setting $p_{\pm i} := \pm e_i$ for $i = 1, 2, \dots, n$, we have a δ-strainer $\{p_i\}_{i=\pm 1,\dots,\pm n}$ at the origin of $\mathbf{R}^n$ for any $\delta > 0$.

If a point $x \in X$ is δ-strained with strainer $\{p_i\}_{i=\pm 1,\dots,\pm n}$, then there is a neighborhood U of x such that all points in U is δ-strained with strainer $\{p_i\}_{i=\pm 1,\dots,\pm n}$.

Lemma 4.3. *If $x \in X_\delta$, then x has a $\theta(\delta)$-strainer, where $\theta(\delta) \to 0$ as $\delta \to 0$.*

Proof. As $\lambda \to +\infty$, the scaled space $(\lambda X, x)$ converges to the tangent cone (K_x, o) in the sense of Gromov-Hausdorff convergence. It follows from $x \in X_\delta$ that (K_x, o) is close to $(\mathbf{R}^n, o)$. Since $\mathbf{R}^n$ has a strainer at the origin o, we have the lemma. □

We omit the proof of the following.

Theorem 4.4 ([BGP]). *There exists $\delta(n) > 0$ such that if $\{p_i\}_{i=\pm 1,\dots,\pm n}$ is a $\theta(\delta(n))$-strainer at a point $x \in X_{\delta(n)}$, then there exists a neighborhood U of x such that the strain map $\varphi : U \to \varphi(U) \subset \mathbf{R}^n$ is a bi-Lipschitz homeomorphism and the image $\varphi(U)$ is an open subset of $\mathbf{R}^n$.*

Outline of Proof of Theorem 4.1.

Step 1. Let $x \in X$ be a non-singular point. Then there is a strainer $\{p_i\}$ at x such that $p_i \notin \mathrm{Cut}_x$ for all $i = 1, \dots, n$. By Theorem 4.4, the strain map $\varphi : U \to \varphi(U) \subset \mathbf{R}^n$ is a bi-Lipschitz homeomorphism. A Riemannian metric at x is defined by

$$g_{ij}(x) := \cos \angle p_i x p_j, \qquad i, j = 1, \dots, n.$$

We remark that (U, φ) depends on x and that g_{ij} is not necessarily unique if $p_i \in \mathrm{Cut}_x$ for some i.

We take another strainer $\{q_i\}$ at x such that $q_i \notin \mathrm{Cut}_x$. Let $\psi : V \to \psi(V) \subset \mathbf{R}^n$ be the strain map associated with $\{q_i\}$.

Lemma 4.5. *$\psi \circ \varphi^{-1}$ is differentiable at x.*

For the proof of this lemma, the following is essential.

Lemma 4.6 (First variation formula)**.** *If px is unique, then*

$$d(p, y) = d(p, x) - d(x, y) \cos \angle pxy + o(d(x, y)),$$

for all $y \in X$, where $o(\cdot)$ is uniform for y.

We omit the proof of the first variation formula.

The first variation formula implies that, if x is non-singular and if px is unique, then the distance function $r = d(p, \cdot)$ is differentiable at x, which proves Lemma 4.5.

Step 2. For $\epsilon > 0$ small enough and for $y \in X$, we define

$$\tilde{\varphi}_i(y) := \frac{1}{\mathcal{H}^n(B(p_i, \epsilon))} \int_{B(p_i, \epsilon)} d(y, z) \, d\mathcal{H}^n(z),$$
$$\tilde{\varphi}(y) := (\tilde{\varphi}_1(y), \dots, \tilde{\varphi}_n(y)).$$

By using $\mathcal{H}^n(\mathrm{Cut}_y) = 0$, $y \in X$, we show that the map $\tilde{\varphi}$ is differentiable at any non-singular point. $(U, \tilde{\varphi})$ is a chart better than (U, φ).

Step 3. Deform $\tilde{\psi} \circ \tilde{\varphi}$ to a C^∞ diffeomorphism, say F, between $\tilde{\varphi}(U \cap V)$ and $\tilde{\psi}(U \cap V)$ and then consider the gluing $\tilde{\varphi}(U)$ and $\tilde{\psi}(V)$ by F, which is a C^∞ manifold bi-Lipschitz homeomorphic to $U \cup V$ (see Figure 5). This bi-Lipschitz homeomorphism is C^1 diffeomorphic on $X \setminus S_X$. If the whole space $X_{\delta(n)}$ is covered by U and V, then a C^∞ structure on $X_{\delta(n)} = U \cup V$ is induced from the C^∞ structure of the gluing. If you have more than three charts covering $X_{\delta(n)}$, then we repeat this procedure to obtain a C^∞ structure on $X_{\delta(n)}$. The Riemannian metric as defined in Step 1 induces a metric on $X_{\delta(n)}$. □

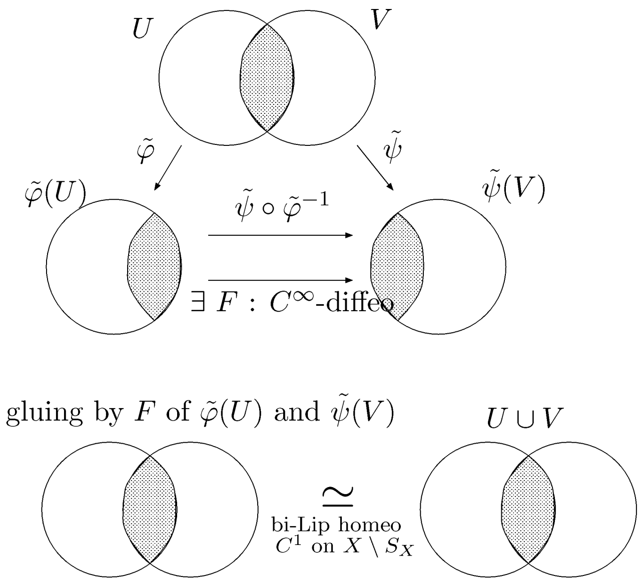

Figure 5:

5 Infinitesimal Bishop-Gromov condition

We want to consider the condition of a lower bound of Ricci curvature for an Alexandrov space. However, the Ricci curvature tensor cannot be defined on an Alexandrov space because of the low regularity of the Riemannian metric. Instead, we consider a volume comparison condition as follows.

We first define a function

$$s_K(r) := \begin{cases} \sin(\sqrt{K}r)/\sqrt{K} & \text{if } K > 0, \\ r & \text{if } K = 0, \\ \sinh(\sqrt{|K|}r)/\sqrt{|K|} & \text{if } K < 0. \end{cases}$$

This is a solution to the Jacobi equation ${s_K}''(r) + K s_K(r) = 0$ with $s_K(0) = 0$ and ${s_K}'(0) = 1$.

The following is the volume comparison condition.

Definition 5.1 (Infinitesimal Bishop-Gromov condition)**.** Let K and N be two real numbers with $N \geq 1$ and μ a positive Radon measure with full support in X. μ satisfies the *infinitesimal Bishop-Gromov condition*, $BG(K, N)$, if we have

$$\int_{W_{p,t}} f \circ \Phi_{p,t}(x)\, d\mu(x) \geq \int_X f(y)\, \frac{t\, s_K(tr(y))^{N-1}}{s_K(r(y))^{N-1}}\, d\mu(y).$$

for any point $p \in X$, any number $0 < t < 1$, and any Borel measurable function $f : X \to [\,0, +\infty\,)$ with compact support, where $\Phi_{p,t} : W_{p,t} \to X$ is the radial expansion map (see Definition 3.1).

We are going to prove:

Theorem 5.2 ([KS3])**.** *If X is an n-dimensional Alexandrov space of curvature $\geq \kappa$, then the n-dimensional Hausdorff measure $\mathcal{H}^n$ satisfies $BG(\kappa, n)$.*

Assume that $\mathcal{H}^n(\partial W_{p,t}) = 0$ and $X = X_{\delta(n)}$ for simplicity. (The proof in the general case is much more technical.) By the area formula,

$$\int_{W_{p,t}} f \circ \Phi_{p,t}(x)\, d\mathcal{H}^n(x) \geq \int_X \frac{f(y)}{|\det d\Phi_{p,t}|}\, d\mathcal{H}^n(y).$$

So, it suffices to prove the following:

Lemma 5.3. *For almost all $x \in X$, $y := \Phi_{p,t}(x)$, we have*

$$|\det d\Phi_{p,t}(x)| \leq \frac{s_\kappa(r(y))^{n-1}}{t\, s_\kappa(r(x))^{n-1}}.$$

Proof. (a) We prove that $|d\Phi_{p,t}(v)| = |v|/t$ for any $v \in K_x$ tangent to the ray from p, where $|v| := d(v, o_x)$. This is obvious by the definition of the radial expansion map (see

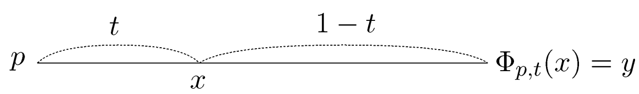

Figure 6:

Figure 6).

(b) We prove that $|d\Phi_{p,t}(v)| \leq \frac{s_\kappa(r(y))}{s_\kappa(r(x))}|v|$ for any $v \in K_x$ orthogonal to the ray from p. To prove it, we take any point $z \in \partial B(p, r(x))$ and set $y := \Phi_{p,t}(x)$, $w := \Phi_{p,t}(z)$. Then we have

$$\frac{d(y,w)}{d(x,z)} \leq \frac{d(\tilde{y},\tilde{w})}{d(\tilde{x},\tilde{z})} \xrightarrow{z \to x} \frac{s_\kappa(r(x))}{s_\kappa(r(y))},$$

which implies (b).

Combining (a) and (b) completes the proof of the lemma. $\square$

6 Poincaré inequality

The following theorem is essential for the existence of heat kernel, the Rellich compactness, and convergence of spectra under a Gromov-Hausdorff convergence $X_i \to X$.

Theorem 6.1 ([KMS]). *Let $0 < r \leq R$ and $x \in X$. For any Lipschitz function $u : B(x,3r) \to \mathbf{R}$ we have*

$$\int_{B(x,r)} |u - \bar{u}_{B(x,r)}|^2 \, d\mathcal{H}^n \leq \frac{Cr^{n+2}}{\mathcal{H}^n(B(x,r))} \int_{B(x,3r)} |\nabla u|^2 d\mathcal{H}^n,$$

where $C = C(n, \kappa, R)$ is some constant depending on n, κ, and R and we set

$$\bar{u}_{B(x,r)} := \frac{1}{\mathcal{H}^n(B(x,r))} \int_{B(x,r)} u \, d\mathcal{H}^n.$$

Proof. For any $y \in B(x,r)$,

$$\begin{aligned} |u(y) - \bar{u}_{B(x,r)}| &\leq \frac{1}{\mathcal{H}^n(B(x,r))} \int_{B(x,r)} |u(y) - u(z)| \, d\mathcal{H}^n(z) \\ &\leq \frac{2r}{\mathcal{H}^n(B(x,r))} \int_{B(x,r)} \int_0^1 |\nabla u(\Phi_{y,t}^{-1}(z))| \, dt d\mathcal{H}^n(z). \end{aligned}$$

Since $t \mapsto \Phi_{y,t}^{-1}(z)$ is a minimal geodesic from y to z, we have

$$\begin{aligned} |u(y) - \bar{u}_{B(x,r)}|^2 &\leq \frac{4r^2}{\mathcal{H}^n(B(x,r))} \int_0^1 \int_{B(y,2r)} |\nabla u(\Phi_{y,t}^{-1}(z))|^2 \, d\mathcal{H}^n(z) dt \\ &\leq \frac{4r^2}{\mathcal{H}^n(B(x,r))} \int_0^1 \int_{B(y,2tr)} |\nabla u(z)|^2 \, \frac{s_\kappa(r(z)/t)^{n-1}}{t\, s_\kappa(r(z))^{n-1}} d\mathcal{H}^n(z) dt \\ &\leq \frac{4r^2}{\mathcal{H}^n(B(x,r))} \int_0^1 \int_{B(y,2tr)} |\nabla u(z)|^2 \, \frac{C}{t^n} \, d\mathcal{H}^n(z) dt \\ &\leq \frac{4Cr^2}{\mathcal{H}^n(B(x,r))} \int_0^1 \int_{B(x,3r)} \frac{|\nabla u(z)|^2}{t^n} \, I(z,t) \, d\mathcal{H}^n(z) dt, \end{aligned}$$

where $I(z,t) := 1$ if $d(z,y) \leq 2tr$; $I(z,t) := 0$ otherwise,

$$\begin{aligned} &= \frac{4Cr^2}{\mathcal{H}^n(B(x,r))} \int_{B(x,3r)} |\nabla u(z)|^2 \int_{d(z,y)/2r}^1 \frac{1}{t^n} \, dt \, d\mathcal{H}^n(z) \\ &\leq \frac{2^{n+1}Cr^{n+1}}{(n-1)\mathcal{H}^n(B(x,r))} \int_{B(x,3r)} \frac{|\nabla u(z)|^2}{d(z,y)^{n-1}} d\mathcal{H}^n(z). \end{aligned}$$

Therefore,

$$\begin{aligned} &\int_{B(x,r)} |u(y) - \bar{u}_{B(x,r)}|^2 \, d\mathcal{H}^n(y) \\ &\leq \frac{2^{n+1}Cr^{n+1}}{(n-1)\mathcal{H}^n(B(x,r))} \int_{B(x,3r)} |\nabla u(z)|^2 \int_{B(x,r)} \frac{1}{d(z,y)^{n-1}} d\mathcal{H}^n(y) d\mathcal{H}^n(z). \end{aligned}$$

Since there is a 1-Lipschitz map from $M^n(\kappa)$ to X keeping the distance from z, we have

$$\begin{aligned}\int_{B(x,r)} \frac{1}{d(z,y)^{n-1}} d\mathcal{H}^n(y) &\le \int_{B(z,2r)} \frac{1}{d(z,y)^{n-1}} d\mathcal{H}^n(y)\\ &\le \int_{B(o,2r)\subset M^n(\kappa)} \frac{1}{d(o,\tilde{y})^{n-1}}\, d\,\mathrm{vol}(\tilde{y}) \le C'r.\end{aligned}$$

We thus obtain

$$\int_{B(x,r)} |u - \bar{u}_{B(x,r)}|^2 \, d\mathcal{H}^n \le \frac{2^{n+1}CC'r^{n+2}}{(n-1)\mathcal{H}^n(B(x,r))} \int_{B(x,3r)} |\nabla u|^2 \, d\mathcal{H}^n.$$

□

The discussion above relies only on $BG(K,N)$ and the Bishop volume comparison and works for a more general setting (cf. [KS1]).

We omit the proof of the following two lemmas.

Lemma 6.2 ([KMS]). *For any $x, y \in X$, we have*

$$d(x,y) = \sup\{\, u(y) - u(x) \mid u \in W^{1,2}_{\mathrm{loc}}(X) \cap C(X),\ |\nabla u| \le 1 \ a.e.\ \}.$$

Lemma 6.3 ([KMS]). *Let $\Omega \subset X$ be a relatively compact open subset, let $\delta > 0$, and let $S_\delta := X \setminus X_\delta$. Then, there exist two sequences of numbers $\{r_k\}_{k=1}^\infty$ and $\{s_k\}_{k=1}^\infty$ with $0 < r_k < s_k \to 0$ and a sequence of Lipschitz functions $\{\psi_k : \Omega \to [0,1]\}_{k=1}^\infty$ such that for any k we have*

(1) $\psi_k = 0$ *on* $\Omega \cap B(S_\delta, r_k)$,

(2) $\psi_k = 1$ *on* $\Omega \setminus B(S_\delta, s_k)$,

(3) $\int_\Omega |\nabla \psi_k|^2 \, d\mathcal{H}^n \to 0$ *as* $k \to \infty$.

The Poincaré inequality (Theorem 6.1) and Lemmas 6.2, 6.3 together imply the following:

- There exists a continuous heat kernel on X by using the method of Grigor'yan, Saloff-Coste, Sturm, etc.
- We obtain a continuous Brownian motion, which does not hit S_X almost surely by using the general theory of Dirichlet form.

7 Laplacian comparison

In this section we prove:

Theorem 7.1 (Laplacian comparison theorem [KS4]). *Let $K, N \in \mathbf{R}$ with $N > 1$. If μ satisfies $BG(K,N)$, then*

$$(*)\qquad \int_X \langle \nabla r, \nabla f\rangle \, d\mu \geq \int_X \{-(N-1)\cot_K \circ r\} f \, d\mu$$

for any $p \in X$ and for any Lipschitz function $f : X \to [0, +\infty)$ with compact support in $X \setminus \{p\}$, where $r(x) := d(p,x)$ and $\cot_K := {s_K}'/s_K$.

Note that $(*)$ is a weak form of $\Delta r - \langle \nabla \log \varphi, \nabla r\rangle \geq -(N-1)\cot_\kappa \circ r$.

Proof. Recalling the definition of the radial expansion map, we see

$$t\, r(\Phi_{p,t}(x)) = r(x).$$

Differentiating this by t at $t = 1$ yields

$$r(x) + \langle \nabla r, \frac{d}{dt}\Phi_{p,t}(x)|_{t=1}\rangle = 0,$$

and hence

$$\frac{d}{dt}\Phi_{p,t}(x)|_{t=1} = -r(x)\nabla r(x).$$

Setting $\Omega := \operatorname{supp} f$ we have

$$\begin{aligned}
&\int_\Omega (\langle \nabla r, \nabla(rf)\rangle - f)\, d\mu = \int_\Omega \langle r\nabla r, \nabla f\rangle \, d\mu \\
&= -\int_\Omega \langle \frac{d}{dt}\Phi_{p,t}(x)|_{t=1}, \nabla f\rangle \, d\mu(x) = -\int_\Omega \frac{d}{dt} f \circ \Phi_{p,t}(x)|_{t=1} \, d\mu(x) \\
&= \lim_{t\to 1-0} \left[\int_\Omega \frac{f\circ \Phi_{p,t}(x)}{1-t} \, d\mu(x) - \int_\Omega \frac{f(x)}{1-t} \, d\mu(x)\right] \\
&\geq \liminf_{t\to 1-0} \Big[\int_\Omega \frac{t\, s_K(r(x))^{N-1} f(x)}{(1-t)s_K(r(x)/t)^{N-1}} \, d\mu(x) - \int_\Omega \frac{f(x)}{1-t} \, d\mu(x)\Big] \quad (\text{by } BG(K,N)) \\
&\geq \int_\Omega \liminf_{t\to 1-0} \frac{t\, s_K(r(x))^{N-1} - s_K(r(x)/t)^{N-1}}{(1-t)s_K(r(x)/t)^{N-1}} f(x) d\mu(x) \\
&= \int_\Omega \frac{d}{dt}\left(\frac{s_K(r(x)/t)^{N-1}}{t}\right)\Big|_{t=1} \frac{f(x)}{s_K(r(x))^{N-1}} \, d\mu(x) \\
&= \int_\Omega \{-1 - (N-1)r(x)\cot_K(r(x))\} f(x) \, d\mu(x).
\end{aligned}$$

Therefore,

$$\int_\Omega \langle \nabla r, \nabla(rf)\rangle \, d\mu \geq \int_\Omega \{-(N-1)\cot_K \circ r\}\, rf \, d\mu.$$

Replace rf by f. □

8 Splitting theorem

In this section, we prove:

Theorem 8.1 (Splitting theorem [KS4]). *If a positive Radon measure μ satisfies $BG(0,N)$ for some $N > 1$ and if X contains a straight line, then X is homeomorphic to $Y \times \mathbf{R}$ for some metric space Y.*

The idea of the proof is based on a splitting theorem by Cheeger-Gromoll.

Definition 8.2. A *ray* in X is a curve $\gamma : [\,0, +\infty\,) \to X$ such that $d(\gamma(s), \gamma(t)) = |s - t|$ for all $s, t \geq 0$. The *Busemann function* for a ray $\gamma : [\,0, +\infty\,) \to X$ is defined by

$$b_\gamma(x) := \lim_{t \to +\infty} (t - d(x, \gamma(t))), \quad x \in X.$$

Proposition 8.3. *Let $\gamma : [\,0, +\infty\,) \to X$ be a ray. If a positive Radon measure μ satisfies $BG(0,N)$, then*

$$\int_X \langle \nabla b_\gamma, \nabla f \rangle \, d\mu \leq 0$$

for any $f : X \to [\,0, +\infty\,)$ Lipschitz with compact support, i.e., b_γ is subharmonic w.r.t. μ.

Proof. Take a sequence $t_i \to +\infty$. Since $-\nabla d(\gamma(t_i), \cdot) \to \nabla b_\gamma$ a.e. as $i \to \infty$, we have

$$\begin{aligned}
&\int_X \langle \nabla b_\gamma, \nabla f \rangle \, d\mu = -\lim_{i \to \infty} \int_X \langle \nabla d(\gamma(t_i), \cdot), \nabla f \rangle \, d\mu \\
&\leq (N-1) \lim_{i \to \infty} \int_{\operatorname{supp} f} \frac{f(x)}{d(\gamma(t_i), x)} \, d\mu(x) = 0.
\end{aligned}$$

□

Outline of Proof of Theorem 8.1. Let $b := b_{\gamma|_{[0,+\infty)}} + b_{\gamma|_{(-\infty,0]}}$. By Proposition 8.3, b is subharmonic w.r.t. μ. It is easy to see that $b \leq 0$ and $b \circ \gamma = 0$. We have the maximum principle, which is quite nontrivial. Therefore, we have $b = 0$ on X, which implies that X is covered by disjoint straight lines bi-asymptotic to γ, so that X is homeomorphic to $b_\gamma^{-1}(t) \times \mathbf{R}$. □

Assume for an n-dimensional Alexandrov space X,

- S_X is closed,
- $X^* = X \setminus S_X$ is an (incomplete) C^∞ Riemannian manifold,
- $V : X \to \mathbf{R} \in C(X) \cap C^\infty(X^*)$.

The *Bakry-Emery Ricci curvature* on X^* is defined by

$$\mathrm{Ric}_{N,V} := \begin{cases} \mathrm{Ric} & \text{if } N = n, \\ \mathrm{Ric} + \mathrm{Hess}\, V - \frac{dV \otimes dV}{N-n} & \text{if } n < N < +\infty, \\ \mathrm{Ric} + \mathrm{Hess}\, V & \text{if } N = +\infty. \end{cases}$$

Corollary 8.4. *Let $n \le N \le +\infty$. Assume that $\sup_X V < +\infty$ if $N = +\infty$. If $\mathrm{Ric}_{N,V} \ge 0$ on X^* and if X contains a straight line, then X is isometric to $Y \times \mathbf{R}$ for some Alexandrov space Y.*

Borzellino-Zhu proved the corollary in the case of complete Riemannian orbifolds and $N = n$. Lichnerowicz proved it in the case of complete Riemannian manifolds.

References

[BBI] D. Burago, Yu. Burago and S. Ivanov, A course in metric geometry, Graduate Studies in Mathematics 33, Amer. Math. Soc., Providence, RI, 2001.

[BGP] Yu. Burago, M. Gromov and G. Perel'man, A. D. Aleksandrov spaces with curvatures bounded below, Uspekhi Mat. Nauk 47(1992), no. 2(284), 3–51, 222, translation in Russian Math. Surveys 47 (1992), no. 2, 1–58.

[KMS] K. Kuwae, Y. Machigashira and T. Shioya, Sobolev spaces, Laplacian, and heat kernel on Alexandrov spaces, Math. Z. 238(2001),269–316.

[KS1] K. Kuwae and T. Shioya, Sobolev and Dirichlet spaces over maps between metric spaces, J. Reine Angew. Math. 555(2003), 39–75.

[KS2] K. Kuwae and T. Shioya, Laplacian comparison on Alexandrov spaces, preprint.

[KS3] K. Kuwae and T. Shioya, Infinitesimal Bishop-Gromov condition for Alexandrov spaces, Probabilistic Approach to Geometry, 293–302, Adv. Stud. Pure Math. 57, Math. Soc. Japan, Tokyo, 2010.

[KS4] K. Kuwae and T. Shioya, A topological splitting theorem for weighted Alexandrov spaces, preprint.

[OS] Y. Otsu and T. Shioya, The Riemannian structure of Alexandrov spaces, J. Differential Geom. 39(1994), 629–658.

[P] G. Perelman, DC-structure on Alexandrov space, preprint.

[S1] T. Shioya, Convergence of Alexandrov spaces and spectrum of Laplacian, J. Math. Soc. Japan 53(2001), 1–15.

[S2] T. Shioya, Geometric Analysis on Alexandrov spaces, (Japanese) Sugaku 61(2009), 1–20. (an English translation will be appeared in Sugaku Expositions)

TOPICS IN MINIMAL SUBMANIFOLDS

Y. L. XIN

(Institute of Mathematics, Fudan University, Shanghai)

CHAPTER I

Certain Techniques in Hypersurfaces

We outline basic notions on minimal submanifolds in this chapter. For more detail account we refer to consult author's book [X2].

1.1 Basic Notions and Formulas

* The second fundamental form

Let N be a Riemannian manifold of dimension m, M be an n-dimensional Riemannian manifold. We assume that $m = n + k$, $k > 0$. Let $M \to N$ be an isometric immersion which means that the natural induced Riemannian metric on M from the ambient space N coincides with the original one on M. The number k is called codimension of M in N. If $k = 1$, the submanifold M is called a hypersurface in N.

For each $p \in M$ the tangent space T_pN can be decomposed to a direct sum of T_pM and its orthogonal complement N_pM in T_pN. Such a decomposition is differentiable. So that we have an orthogonal decomposition of the tangent bundle TN along M

$$TN|_M = TM \oplus NM.$$

Let $(\cdots)^T$ and $(\cdots)^N$ denote the orthogonal projections into the tangent bundle TM and the normal bundle NM respectively.

Let $\bar{\nabla}$ be the Levi-Civita connection on N. As vector bundles TM, NM over M, they carry the induced metrics as their fiber metrics.

Definition 1.1.1 For $V, W \in \Gamma(TM)$, $\nu \in \Gamma(NM)$, the induced connections on TM and NM are defined by

$$\nabla_V W \overset{def.}{=} (\bar{\nabla}_V W)^T,$$

$$\nabla_V \nu \overset{def.}{=} (\bar{\nabla}_V \nu)^N.$$

Proposition 1.1.2 ∇ *is just the Levi - Civita connection on* M.

As done in the above proposition, the induced connection ∇ on the normal bundle also preserves the inner product.

Consider

$$B_{VW} \overset{def.}{=} \left(\bar{\nabla}_V W\right)^N = \bar{\nabla}_V W - \nabla_V W$$

for $V, W \in \Gamma(TM)$. B is a symmetric bilinear form on TM with values in NM. We call B to be second fundamental form of M in N.

For $\nu \in \Gamma(NM)$ we define the shape operator $A^\nu : TM \to TM$ by

$$A^\nu(V) = -(\bar{\nabla}_V \nu)^T.$$

It is easy to check that A^ν is a symmetric operator on the tangent space at each point, moreover, it satisfies the Weingarten equations:

$$\langle B_{XY}, \nu \rangle = \langle A^\nu(X), Y \rangle. \tag{1.1.1}$$

Definition 1.1.3 If $B \equiv 0$, then M is called a totally geodesic submanifold in N.

From the definition of the second fundamental form, we see that M is a totally geodesic submanifold, if and only if any geodesic in M is also a geodesic in the ambient manifold N.

Taking the trace of B gives the mean curvature vector H of M in N and

$$H \overset{def.}{=} \frac{1}{n}\mathrm{trace}(B) = \frac{1}{n}\sum_{i=1}^{n} B_{e_i e_i},$$

where $\{e_i\}$ is a local orthonormal frame field of M. The mean curvature vector is a cross-section of the normal bundle.

Remark The definition of the mean curvature in some references is different from one here by a constant factor which is equal to the dimension of the submanifold.

Definition 1.1.4 If $H \equiv 0$, then M is a minimal submanifold in N.

Definition 1.1.5 If H is a parallel cross-section on the normal bundle, then M is defined to be a submanifold with parallel mean curvature.

From the definitions one immediately sees that a totally geodesic submanifold M in N is necessarily a minimal submanifold and any minimal submanifold is a manifold with parallel mean curvature.

Note the special case that M is a hypersurface in N. Fix a unit normal vector field ν locally. Then the second fundamental form is determined by

$$A \stackrel{def.}{=} A^{\nu}.$$

This is symmetric on tangent space at each point. Its eigenvalues $k_1, \cdots, k_n$ are called the principal curvatures. The product of all principal curvatures is called the Gauss - Kronecker curvature. It is easy to see that the mean curvature is the mean value of all principal curvatures. In this case there is a notion of constant mean curvature hypersurfaces instead of manifolds with parallel mean curvature.

We can define the curvature tensors $R_{XY}Z$ and $R_{XY}\mu$, corresponding to the connections in the tangent bundle and the normal bundle respectively:

$$R_{XY}Z = -\nabla_X \nabla_Y Z + \nabla_Y \nabla_X Z + \nabla_{[X,Y]}Z,$$

$$R_{XY}\mu = -\nabla_X \nabla_Y \mu + \nabla_Y \nabla_X \mu + \nabla_{[X,Y]}\mu,$$

where X, Y, Z are tangent vector fields, μ is a normal vector field. Those are related to the curvature tensor $\bar{R}$ of the ambient manifold N and the second fundamental form B.

Proposition 1.1.6 (Gauss equation)

$$\langle R_{XY}Z, W\rangle = \left\langle \bar{R}_{XY}Z, W\right\rangle - \langle B_{XW}, B_{YZ}\rangle + \langle B_{XZ}, B_{YW}\rangle, \qquad (1.1.2)$$

where X, Y, Z, W are tangent vector fields in M, their images under the isometric immersion are tangent vector fields in N. For the simplicity we use the same notations.

Remark From the Gauss equation we obtain the famous Theorem Egiregium of Gauss: Let M be a surface in $\mathbb{R}^3$. Then the sectional curvature of M is equal to the Gauss - Kronecker curvature of M.

Proposition 1.1.7 (Codazzi equations)

$$(\nabla_X B)_{YZ} - (\nabla_Y B)_{XZ} = -(\bar{R}_{XY}Z)^N \tag{1.1.3}$$

Proposition 1.1.8 (Ricci equations)

$$\langle R_{XY}\mu, \nu\rangle = \langle \bar{R}_{XY}\mu, \nu\rangle + \langle B_{Xe_i}, \mu\rangle \langle B_{Ye_i}, \nu\rangle - \langle B_{Xe_i}, \nu\rangle \langle B_{Ye_i}, \mu\rangle\,, \tag{1.1.4}$$

where $\{e_i\}$ is a local orthonormal frame field, μ, ν are normal vector fields in M. Here and in the sequel we use the summation convention.

The equations of Gauss, Codazzi and Ricci are fundamental equations for the local theory of the immersed submanifolds. It is possible to state a generalization of the fundamental theorem of local surface theory in $\mathbb{R}^3$. We refer the readers to the book [Spi] (vol. IV, pp 64-74).

$*$ **The First Variational Formula**

The notion of totally geodesic submanifolds is a higher dimensional generalization of geodesics. But, those are very few in general situation. Note that geodesics are critical points of the arc length functional.

A minimal submanifold is defined to be one with vanishing mean curvature. This definition seems to have no relation with the "minimal" terminology. In fact, Lagrange found minimal surfaces in his investigation of the calculus of variations. Now, we generalize Lagrange's study to more general setting. Consider the space $\mathcal{I}(M, N)$ of all immersions from M into N. Then the volume $\mathrm{vol}(f(M))$ is a functional on the space. The critical points of the volume functional are minimal submanifolds by the following first variational formula. Thus, the notion of minimal submanifolds is an adequate generalization of that of geodesics

Proposition 1.1.9 *Let M be a compact Riemannian manifold, $f : M \to N$ an isometric immersion with mean curvature vector H. Let f_t, $|t| < \varepsilon$, $f_0 = f$, be a smooth family of immersions satisfying $f_t|_{\partial M} = f|_{\partial M}$. Denote $V = \left.\frac{\partial f_t}{\partial t}\right|_{t=0}$ to be the variational vector field along f. Then*

$$\left.\frac{d}{d\,t}\mathrm{vol}(f_t M)\right|_{t=o} = -\int_M \langle n\,H, V\rangle\, d\,\mathrm{vol}. \tag{1.1.5}$$

Remark 1.1.10 The first variational formula (1.1.5) shows that the $-n\,H$ represents the gradient of the volume functional. The equation $H = 0$ is the Euler - Lagrange equation for the functional.

If we restrict the variation above to be normal, namely V is normal to M everywhere and $V^T = 0$, then the formula remains valid without the boundary condition.

If M is not compact, then the formula can be used for compactly supported variations.

$*$ Minimal Submanifolds in Euclidean Space

The study of minimal surfaces in $\mathbb{R}^3$ is an interesting subject since Lagrange's time. Up to now the subject still attracts many mathematicians. The present section starts with its interesting feature on the coordinate functions. Then, we derive the equation for minimal graphs of codimension one in $\mathbb{R}^{n+1}$.

Let M be a Riemannian manifold of dimension m. Consider the Laplace operator $\Delta : C^\infty(M) \to C^\infty(M)$. For $f \in C^\infty(M)$ choose a local orthonormal frame field $\{e_1, \cdots, e_m\}$ in M. Then

$$\Delta f = e_i e_i(f) - (\nabla_{e_i} e_i)\, f. \tag{1.1.6}$$

Around each point p, there are local coordinates $(x^1, \cdots, x^m)$, where the Riemannian metric on M can be written as $ds^2 = g_{ij} dx^i dx^j$. If we denote $(g^{ij}) = (g_{ij})^{-1}$ and $g = \det(g_{ij})$, then

$$\Delta f = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i} \left(\sqrt{g}\, g^{ij} \frac{\partial f}{\partial x^j} \right). \tag{1.1.7}$$

In general, for any differential form with values in a vector bundle we can define exterior differential operator d and codifferential operator δ and the Hodge - Laplace operator $d\delta + \delta d$. The minus sign of the Hodge -Laplace operator acting on a smooth function f, a cross-section of the trivial bundle $M \times \mathbb{R}$, is just the ordinary Laplace operator

$$\Delta f = -\delta d\, f. \tag{1.1.8}$$

We omit the verification of the equivalence of those three definitions, which is left to the readers as an exercise.

Any $f \in C^\infty(M)$ satisfying $\Delta f = 0$ is called a harmonic function. We have the Hopf maximum principle for harmonic functions: any harmonic function on a Riemannian manifold has to be a constant, if it attains the local maximum in an interior point.

Now let us study the minimal submanifolds in Euclidean space.

Proposition 1.1.11 *Let $\psi : M \to \mathbb{R}^n$ be an isometric immersion with the mean curvature vector H, then*

$$\Delta\psi = m\, H, \tag{1.1.9}$$

where $\Delta\psi = (\Delta\psi^1, \cdots, \Delta\psi^n)$.

Corollary 1.1.12 *An isometric immersion $\psi : M \to \mathbb{R}^{n+1}$ is a minimal immersion if and only if each component of ψ is a harmonic function on M.*

From Corollary 1.1.12 and the Hopf maximum principle we have immediately: There is no compact minimal submanifold in Euclidean space.

From the first variational formula (1.1.5) we know that $H = 0$ is the Euler-Lagrangian equations for the volume functional of immersed submanifolds in an ambient manifold. What is the equations look like? Let us see the simplest situation.

In $\mathbb{R}^{n+1}$ a minimal graph M is defined by

$$x^{n+1} = f(x^1, \cdots, x^n).$$

We denote $f_i = \frac{\partial f}{\partial x^i}$. The induced metric on M is

$$ds^2 = g_{ij}\, dx^i\, dx^j,$$

where

$$g_{ij} = \delta_{ij} + f_i f_j.$$

Denote $v = \sqrt{1 + \sum_i f_i^2}$. We have $g^{ij} = \delta_{ij} - \frac{1}{v^2} f_i\, f_j$. The unit normal vector to M is

$$\nu = \frac{1}{v}(f_1, \cdots, f_n, -1).$$

It is obvious that

$$\bar{\nabla}_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \frac{\partial}{\partial x^i}\left(0, \cdots, 0, 1, 0, \cdots, 0, \frac{\partial f}{\partial x^j}\right) = (0, \cdots, f_{ij})$$

and

$$\left\langle B_{\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}}, \nu\right\rangle = \left\langle \bar{\nabla}_{\frac{\partial}{\partial xi}} \frac{\partial}{\partial x^j}, \nu\right\rangle = -\frac{1}{v} f_{ij}.$$

From $H = 0$ it follows that $g^{ij} f_{ij} = 0$. Thus, we obtain the minimal hypersurface equation

$$(1 + \sum_i f_i^2)\, f_{jj} - f_i f_j f_{ij} = 0, \tag{1.1.10}$$

which is equivalent to

$$\frac{\partial}{\partial x^i}\left(\frac{1}{v}\frac{\partial f}{\partial x^i}\right) = 0.$$

When $n = 2$ (1.1.10) reduces to

$$(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy} = 0, \tag{1.1.11}$$

where we denote $x = x^1,\ y = x^2$.

It is a nonlinear elliptic PDE.

On a minimal submanifold in $\mathbb{R}^n$ there is another important equation. In fact, we have

PROPOSITION 1.3.5 *Let M be an oriented hypersurface with constant mean curvature in $\mathbb{R}^{n+1}$ and with second fundamental form B. Let ν be the unit normal vector to M. Then for any fixed vector $a \in \mathbb{R}^{n+1}$,*

$$\Delta \langle a, \nu\rangle + |B|^2 \langle a, \nu\rangle = 0. \tag{1.1.12}$$

When M is a graph defined by $x^{n+1} = f(x^1, \cdots, x^n)$ in $\mathbb{R}^{n+1}$. Put $a = (0, \cdots, 0, -1)$ and $\nu = \frac{1}{v}(f_1, \cdots, f_n, -1)$. Then $\langle a, \nu\rangle = \frac{1}{v} \overset{def.}{=} w$ and we have

$$\Delta w + |B|^2 w = 0. \tag{1.1.13}$$

$*$ **Examples**

The minimal surface equation (1.1.11) is a nonlinear partial deferential equation. It is hard to solve. Besides the linear functions, what are its solutions? As early as 1776 J. L. Meunier obtained two nonlinear solutions to the equation firstly. Their graphs are catenoid and helicoid.

The catenoid is defined by

$$z = \cosh^{-1}\sqrt{x^2 + y^2}, \tag{1.1.14}$$

Take a catenary in Y-Z coordinate plane. Letting it rotating about Z-axis gives the catenoid.

The helicoid is defined by

$$z = \tan^{-1}\frac{x}{y}. \tag{1.1.15}$$

Let a line in $X-$axis screw about $Z-$axis. The resulting surface is a helicoid

We now give some examples of minimal submanifolds in the sphere.

Let $\psi : M \to S^n \subset \mathbb{R}^{n+1}$ and $\psi' : M' \to S^{n'} \subset \mathbb{R}^{n'+1}$ be minimal immersions. For any constants c and c'

$$c\psi \oplus c'\psi' : M \times M' \to \mathbb{R}^{n+n'+2}$$

is also an isometric immersion of the product manifold $M \times M'$ to $\mathbb{R}^{n+n'+2}$. If we choose c and c' with $c^2 + {c'}^2 = 1$, then the image of $M \times M'$ under $c\psi \oplus c'\psi'$ lies in the sphere $S^{n+n'+1}$. We know that the induced metric on M under $c\psi$ is $c^2\, ds^2$, where ds^2 is the original metric on M.

If c and c' also satisfy

$$\frac{m}{c^2} = \frac{m'}{{c'}^2},$$

we obtain a minimal immersion $c\psi \oplus c'\psi' : M \times M' \to S^{n+n'+1}$. In particular, $M = S^n$ and $M' = S^{n'}$ we have the Clifford minimal hypersurface

$$S^n\left(\sqrt{\frac{n}{n+n'}}\right) \times S^{n'}\left(\sqrt{\frac{n'}{n+n'}}\right) \to S^{n+n'+1}. \tag{1.1.16}$$

We have

$$S^2(\sqrt{3}) \to S^4,$$

which can be realized by the map

$$\begin{aligned}\Psi(x,y,z) = \Big(&\frac{1}{\sqrt{3}}xy, \frac{1}{\sqrt{3}}xz, \frac{1}{\sqrt{3}}yz, \\ &\frac{1}{2\sqrt{3}}(x^2-y^2), \frac{1}{6}(x^2+y^2-2z^2)\Big),\end{aligned} \tag{1.1.17}$$

where $x^2 + y^2 + z^2 = 3$. It is called the Veronese surface which is an imbedding of the real projective plane of curvature $\frac{1}{3}$ into S^4.

The Clifford minimal hypersurface and the Veronese surface are important minimal submanifolds in the sphere.

∗ Bochner-SimonsType Formula and Rigidity Theorems

A geometric invariant described by nonlinear equations possesses rigidity properties in many cases. The squared norm of the second fundamental form is in the case. This phenomenon was revealed by J. Simons [Si]. He firstly applied the useful Bochner technique to minimal submanifold theory.

Let $M \to N$ be a minimal immersion with the second fundamental form B, which can be viewed as a cross-section of the vector bundle $\mathrm{Hom}(\odot^2 TM, NM)$ over M. A connection on $\mathrm{Hom}(\odot^2 TM, NM)$ can be induced from those of TM and NM naturally. There is the trace-Laplace operator ∇^2 acting on any cross-section of a Riemannian vector bundle E. We know that if the base manifold is compact, then ∇^2 is a semi-negative and self-adjoint differential operator with respect to the global inner product on $\Gamma(E)$ (see [X3], p8). To compute $\nabla^2 B$ we introduce some relevant cross-sections in this bundle.

Definition 1.1.13

$$\tilde{\mathcal{B}} \overset{def.}{=} B \circ B^t \circ B,$$

where B^t is the conjugate map of B.

Definition 1.1.14

$$\underline{\mathcal{B}}_{X\,Y} \overset{def.}{=} \sum_{j=1}^{p} \left(B_{A^{\nu_j}\,A^{\nu_j}(X)\,Y} + B_{X\,A^{\nu_j}\,A^{\nu_j}(Y)} - 2\,B_{A^{\nu_j}(X)\,A^{\nu_j}(Y)}\right), \tag{1.1.18}$$

where ν_j are basis vectors of normal space and p is the codimension. It is obvious that $\underline{\mathcal{B}}_{X\,Y}$ is symmetric in X and Y, which is a cross-section of the bundle $\mathrm{Hom}(\odot^2 TM, NM)$.

Lemma 1.1.15

$$\langle \underline{\mathcal{B}}_{X\,Y}, \nu\rangle = \sum_{j=1}^{p} \langle ad\,A^{\nu_j}\,ad\,A^{\nu_j}A^{\nu}(X),\,Y\rangle\,,$$

where ν is a normal vector and $(adA)B = [A,B]$.

Definition 1.1.16.

$$\tilde{\mathcal{R}}_{X\,Y} \overset{def.}{=} \sum_{j=1}^{n} \left[(\bar{\nabla}_X \bar{R})_{Y\,e_j} e_j + (\bar{\nabla}_{e_j}\bar{R})_{X\,e_j} Y\right]^N, \tag{1.1.19}$$

where n is the dimension of M and $\{e_j\}$ is a local orthonormal frame field of M.

Lemma 1.1.17 *$\tilde{\mathcal{R}}_{X\,Y}$ is independent of the choice of $\{e_j\}$ and is symmetric: $\tilde{\mathcal{R}}_{X\,Y} = \tilde{\mathcal{R}}_{Y\,X}$.*

Definition 1.1.18

$$\begin{aligned}\underline{\mathcal{R}}_{X\,Y} = \sum_{j=1}^{n} \Big[&2\,\bar{R}_{Y\,e_j}(B_{X\,e_j}) + 2\,\bar{R}_{X\,e_j}(B_{Y\,e_j}) - B_{X\,(\bar{R}_{Y\,e_j}e_j)^T} \\ &- B_{Y\,(\bar{R}_{X\,e_j}e_j)^T} + \bar{R}_{B_{X\,Y}\,e_j}e_j - 2\,B_{e_j\,(\bar{R}_{X\,e_j}Y)^T}\Big]^N.\end{aligned} \tag{1.1.20}$$

It is a cross-section of $\mathrm{Hom}(\odot^2 TM, NM)$ obviously. It is easily seen that the former 5 terms in (1.1.20) are symmetric in X and Y. As for the last term of (1.1.20), it is also symmetric in X and Y, since

$$B_{e_j\,(\bar{R}_{X\,e_j}Y)^T} = B_{e_j\,e_k}\left\langle \bar{R}_{X\,e_j}Y, e_k\right\rangle$$

and the symmetric properties of B and R.

THEOREM 1.1.19 ([SI]) *Let M be a minimal submanifold in N with the second fundamental form B. Then*

$$\nabla^2 B = -\tilde{\mathcal{B}} - \underline{\mathcal{B}} + \tilde{\mathcal{R}} + \underline{\mathcal{R}} \tag{1.1.21}$$

If the ambient manifold $\bar{M}$ is local symmetric, namely, $\bar{\nabla}\bar{R} \equiv 0$, then $\tilde{\mathcal{R}} \equiv 0$. In particular, if $\bar{M}$ has constant sectional curvature c, besides $\tilde{\mathcal{R}} = 0$ we have $\underline{\mathcal{R}} = ncB$. In fact, since $\bar{R}_{X\,Y}Z = c\left(\langle X, Z\rangle Y - \langle Y, Z\rangle X\right)$, we have

$$\bar{R}_{Y\,e_j}(B_{X\,e_j}) = 0, \qquad \left(\bar{R}_{Y\,e_j}e_j\right)^T = c\,(1-n)Y,$$

$$\bar{R}_{B_{X\,Y}\,e_j}e_j = -c\,nB_{X\,Y},\ \left(\bar{R}_{X\,e_j}Y\right)^T = c\left(\langle X, Y\rangle e_j - \langle Y, e_j\rangle X\right),$$

$$B_{e_j\,(\bar{R}_{X\,e_j}Y)^T} = -c\,B_{X\,Y}.$$

Thus,

$$\underline{\mathcal{R}}_{X\,Y} = nc\,B_{X\,Y}.$$

In summary we have

THEOREM 1.1.20 ([SI]) *Let $\bar{M}$ be a Riemannian manifold with constant sectional curvature c and M a minimal submanifold in N with the second fundamental form B. Then*

$$\nabla^2 B = -\tilde{\mathcal{B}} - \underline{\mathcal{B}} + nc\,B. \tag{1.1.22}$$

If M has codimension one, from the definition we know that $\underline{\mathcal{B}} = 0$, moreover,

$$\begin{aligned}
\left\langle \tilde{\mathcal{B}}, B\right\rangle &= \left\langle B^t \circ B, B^t \circ B\right\rangle \\
&= \left\langle B^t \circ B_{e_i e_j}, e_k \odot e_l\right\rangle \left\langle B^t \circ B_{e_i e_j}, e_k \odot e_l\right\rangle \\
&= \left\langle B_{e_i e_j}, B_{e_k e_l}\right\rangle \left\langle B_{e_i e_j}, B_{e_k e_l}\right\rangle \\
&= |B|^4
\end{aligned}$$

In this case we have

$$\left\langle \nabla^2 B, B \right\rangle = -|B|^4 + nc\,|B|^2. \tag{1.1.23}$$

THEOREM 1.1.21 ([Si]) *Let $M \to S^{n+1}$ be a compact oriented minimal hypersurface in the unit sphere with the second fundamental form B. If $|B|^2 < n$, then $B \equiv 0$, namely M is a totally geodesic hypersurface in S^{n+1}.*

Now we study the cases of higher codimension.

LEMMA 1.1.22

$$\left\langle \tilde{\mathcal{B}} + \underline{\mathcal{B}}, B \right\rangle \le \left(2 - \frac{1}{p}\right) |B|^4, \tag{1.1.24}$$

where p is the codimension.

We now have the following theorem due to J. Simons [Si]

THEOREM 1.1.23 *Let $M \to S^{n+p}$ be a compact minimal submanifold in the unit sphere. If*

$$|B|^2 < \frac{n}{2 - \frac{1}{p}}, \tag{1.1.25}$$

then $|B|^2 = 0$, namely M is a totally geodesic submanifold in S^{n+p}.

This theorem tells us that the squared norm of the second fundamental form of a compact minimal submanifold in the sphere can not take every value. It omits the values in the interval $\left(0, \frac{n}{2-\frac{1}{p}}\right)$. It seems that $|B|^2$ is an extrinsic invariant. In fact, by the Gauss equation (1.1.2) its scalar curvature

$$s = n(n-1) - |B|^2 \le n(n-1).$$

Thus, the scalar curvature omits the values of the interval

$$\left(n(n-1) - \frac{n}{2 - \frac{1}{p}},\ n(n-1) \right).$$

Therefore, this is an intrinsic rigidity theorem.

Chern-do Carmo-Kobayashi [C-doC-K] (see also [L] for codimension one case) studied minimal submanifolds in the sphere satisfying

$$|B|^2 = \frac{n}{2 - \frac{1}{p}}.$$

Thus the second fundamental form can be determined in a suitable frame field, so did the connection form with respect to an adapted frame field. If the submanifolds are compact, they are either the Clifford minimal hypersurface or the Veronese surface.

Remark Consider a minimal hypersurface M in $S^n \subset \mathbb{R}^{n+1}$. Let ν be a unit normal vector field of M in S^n, a a fixed vector in $\mathbb{R}^{n+1}$. A similar calculation leads the same equation as (1.1.12)

$$\Delta \langle a, \nu \rangle = - \langle a, \nu \rangle |B|^2, \tag{1.1.26}$$

where B is the second fundamental form of M in S^n. Integrating this formula and using Stokes' theorem gives the Simons' extrinsic rigidity theorem as follows:

Suppose M is a compact minimal hypersurface in S^n, whose normal vector makes a positive inner product with a fixed vector in $\mathbb{R}^{n+1}$. Then M has to be a totally geodesic submanifold in S^n.

$*$ The Second Variational Formula

We know from the first variational formula (1.1.5) that minimal immersion $f : M \to N$ is a critical point on the immersion space from M into N. It is natural to ask if f is a local minimum of the volume functional, namely, for any smooth variations $f_t : M \to N$, and $t > 0$ small enough whether

$$\mathrm{vol}(f) \le \mathrm{vol}(f_t)$$

holds true. To answer the problem we need to derive the second variational formula. First of all let us consider the relevant geometric invariants.

1. For cross-sections on a vector bundle we can define the trace - Laplace operator ∇^2. For the immersion $f : M \to N$ we have normal bundle NM, where there define an induced connection $\nabla = (\overline{\nabla})^N$ on the normal bundle. Hence,

$$\nabla^2 : \Gamma(NM) \to \Gamma(NM).$$

We assume that M has boundary $\partial M \ne \emptyset$, $-\nabla^2$ is self adjoint, semi-positive operator on

$$\mathcal{N}_0 = \{\nu \in \Gamma(NM);\ \nu|_{\partial M} \equiv 0\}.$$

It is also an elliptic operator.

2. The second invariant is defined by curvature of the ambient manifold. Let $\bar{R}$ be the Riemannian curvature tensor on N. Define $\bar{\mathcal{R}} \in Hom(N_xM, N_xM)$ as follows:

$$\bar{\mathcal{R}}(\nu) = \{\bar{R}_{\nu e_i}(e_i)\}^N, \tag{1.1.27}$$

where $\nu \in N_xM$ and $\{e_i\}$ is a local orthonormal frame field near $x \in M$. It is a symmetric operator owing to the properties of the curvature tensor.

Remark If the codimension of M in N is one,

$$\bar{\mathcal{R}}(\nu) = \left(\bar{R}_{\nu e_i}(e_i)\right)^N = -\overline{\mathrm{Ric}}(\nu,\nu)\nu.$$

3. The last invariant involves the second fundamental form B of M in N. Recall that $B \in \Gamma(Hom(S^2TM, NM))$. Its adjoint operator is $A = B^t \in Hom(NM, S^2TM)$. We define $\mathcal{B} \in \Gamma(Hom(NM, NM))$ by

$$\mathcal{B} = B \circ B^t. \tag{1.1.28}$$

By definition it follows that

$$\langle \mathcal{B}(\nu), \mu\rangle = \left\langle B_{e_i e_j}, \nu\right\rangle \left\langle B_{e_i e_j}, \mu\right\rangle.$$

Hence, $\mathcal{B}$ is symmetric and semi-positive. Now, we can prove the following second variational formula.

Theorem 1.1.24 *Let $f : M \to N$ be a compact minimal immersion, $\nu \in \mathcal{N}_0$ be a normal vector field which vanishes on ∂M. Assume that $f_t : M \to N$ is a smooth one-parameter family of immersions, such that for $|t| < \varepsilon$*

$$\begin{cases} f_0 = f, \\ \frac{\partial f_t}{\partial t}|_{t=0} = \nu, \\ f_t|_{\partial M} = f|_{\partial M}, \text{ for each } \quad t. \end{cases}$$

Then

$$\frac{d^2}{dt^2} vol(f_t M)\bigg|_{t=0} = \int_M \left\langle -\nabla^2\nu + \bar{\mathcal{R}}(\nu) - \mathcal{B}(\nu),\ \nu\right\rangle * 1. \tag{1.1.29}$$

Remark the second variational formula is also valid for non-compact M, provided that ν has compact support.

The second variational formula (1.1.29) indicates that it is useful to study the elliptic differential operator of second order defined on $\mathcal{N}_0$

$$\mathcal{S} = -\nabla^2 + \bar{\mathcal{R}} - \mathcal{B}.$$

This is so-called Jacobi operator. We thus can define a symmetric bilinear form on $\mathcal{N}_0$

$$I(\mu,\nu) = \int_M \langle \mathcal{S}(\mu), \nu\rangle * 1.$$

The general self-adjoint elliptic operator theory tells us that the eigenvalues of $\mathcal{S}$ are

$$\lambda_1 < \lambda_2 < \cdots \to \infty$$

and for each i the corresponding eigenspace $E_{\lambda_i} \subset \mathcal{N}_0$ is finite.

Definition 1.1.25 Let $M \to N$ be a minimal immersion. If for any $\mu \in \mathcal{N}_0$

$$I(\mu, \mu) \geq 0,$$

then M is called stable minimal submanifold.

We now study the codimension 1 case, moreover the normal bundle is assumed to be trivial. The second variational formula then becomes

$$I(\mu, \mu) = \int_M \left(|\nabla \phi|^2 - (\langle \overline{\mathrm{Ric}}\, \nu, \nu \rangle + |B|^2)\phi^2\right) * 1. \tag{1.1.30}$$

When the ambient manifold is Euclidean space $\mathbb{R}^{n+1}$,

$$I(\mu, \mu) = \int_M \left(|\nabla \phi|^2 - |B|^2 \phi^2\right) * 1 \tag{1.1.31}$$

Proposition 1.1.26 *The minimal graph M in $\mathbb{R}^{n+1}$, which is defined by $x^{n+1} = f(x^1, \cdots, x^n)$, is stable.*

In fact, the minimal graphs are area-minimizing.

1.2 Curvature Estimates for Minimal Hypersurfaces

E. Heinze [H] in 1952 considered the minimal graph defined over a disc $D_R \subset \mathbb{R}^2$ and gave curvature estimates. The classical Bernstein theorem can be obtained by letting $R \to +\infty$ in his curvature estimates. In this section we describe some important curvature estimates for minimal hypersurfaces.

$*$ **S-S-Y Curvature Estimates**

Now, we introduce curvature estimates for stable minimal hypersurfaces, due to R. Schoen, L. Simon and S. T. Yau [S-S-Y].

Let $M \to N$ be a minimal hypersurface with the second fundamental form B. In the last section we derive a fundamental equation (1.1.21)

$$\nabla^2 B = -\tilde{\mathcal{B}} - \underline{\mathcal{B}} + \tilde{\mathcal{R}} + \underline{\mathcal{R}}.$$

In the case of codimension one we already showed that

$$\underline{\mathcal{B}} = 0, \quad \left\langle \tilde{\mathcal{B}}, B \right\rangle = |B|^4.$$

In addition we assume that the ambient manifolds is Euclidean space. We have

$$\tilde{\mathcal{R}} = \underline{\mathcal{R}} = 0.$$

We thus have

$$\left\langle \nabla^2 B, B \right\rangle \geq -|B|^4. \tag{1.2.1}$$

It follows that

$$\Delta |B|^2 \geq 2\,|\nabla B|^2 - 2\,|B|^4, \tag{1.2.2}$$

and moreover,

$$|B|\Delta|B| + |\nabla|B||^2 \geq |\nabla B|^2 - |B|^4. \tag{1.2.3}$$

To estimate $|\nabla B|^2$ in terms of $|\nabla|B||^2$ set $B_{e_i\, e_j} = h_{ij}\nu$ with the normal vector field ν. So $|B|^2 = \sum_{i,j} h_{ij}^2$. Then

$$|\nabla B|^2 = \langle \nabla_{e_i} B, \nabla_{e_i} B \rangle = \sum_{i,j,k} h_{ijk}^2,$$

$$|\nabla |B||^2 = \left\langle \nabla_{e_k} \sqrt{\sum_{i,j} h_{ij}^2}, \nabla_{e_k} \sqrt{\sum_{i,j} h_{ij}^2} \right\rangle = \frac{1}{\sum_{i,j} h_{ij}^2} \sum_k (\sum_{i,j} h_{ij} h_{ijk})^2,$$

$$\begin{aligned}
|\nabla B|^2 - |\nabla |B||^2 &= \sum_{i,j,k} h_{ijk}^2 - \frac{1}{\sum_{s,t} h_{st}^2} \sum_k (\sum_{i,j} h_{ij} h_{ijk})^2 \\
&= \frac{1}{|B|^2} \left[\sum_{s,t} h_{st}^2 \sum_{i,j,k} h_{ijk}^2 - \sum_k (\sum_{i,j} h_{ij} h_{ijk})^2 \right] \\
&= \frac{1}{2|B|^2} \sum_{i,j,s,t,k} (h_{ij} h_{stk} - h_{st} h_{ijk})^2. \qquad (1.2.4)
\end{aligned}$$

For any $p \in M$, we can choose a local frame field $\{e_1, \cdots, e_n\}$ around p so that $h_{ij} = \lambda_i \delta_{ij}$ at p. Then we have

$$\begin{aligned}
\sum_{i,j,s,t,k} (h_{ij} h_{stk} - h_{st} h_{ijk})^2 &= \sum_{i,s,t,k} (h_{ii} h_{stk} - h_{st} h_{iik})^2 \\
&\qquad + \sum_{s,t} h_{st}^2 \sum_{i \neq j,k} h_{ijk}^2 \\
&\geq \sum_{i,k,s \neq t} h_{ii}^2 h_{stk}^2 + \sum_{s,t} h_{st}^2 \sum_{i \neq j,k} h_{ijk}^2 \\
&= 2 \sum_{s,t} h_{st}^2 \sum_{i \neq j,k} h_{ijk}^2 \\
&= 2\,|B|^2 \sum_{i \neq j,k} h_{ijk}^2.
\end{aligned}$$

Substituting it into (1.2.4) gives

$$\begin{aligned}
|\nabla B|^2 - |\nabla |B||^2 &\geq \sum_{i\neq j,k} h_{ijk}^2 \\
&\geq \sum_{i\neq j} h_{iji}^2 + \sum_{i\neq j} h_{ijj}^2 \\
&= 2\sum_{i\neq j} h_{iji}^2. \qquad (1.2.5)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
|\nabla |B||^2 &= \frac{1}{|B|^2}\sum_k \Big(\sum_{i,j} h_{ij}h_{ijk}\Big)^2 \\
&= \frac{1}{|B|^2}\sum_k \Big(\sum_i h_{ii}h_{iik}\Big)^2 \\
&\leq \sum_{i,k} h_{iik}^2 \\
&= \sum_{i\neq k} h_{iik}^2 + \sum h_{iii}^2 \\
&= \sum_{i\neq k} h_{iik}^2 + \sum_i \Big(\sum_{j\neq i} h_{jji}\Big)^2 \\
&\leq \sum_{i\neq k} h_{iik}^2 + (n-1)\sum_{j\neq i} h_{jji}^2 = n\sum_{i\neq j} h_{iij}^2.
\end{aligned}$$

Substituting it into (1.2.5) gives

$$|\nabla B|^2 - |\nabla |B||^2 \geq \frac{2}{n}|\nabla |B||^2. \qquad (1.2.6)$$

(1.2.3) and (1.2.6) yield

$$|B|\Delta |B| + |B|^4 \geq \frac{2}{n}|\nabla |B||^2 \qquad (1.2.7)$$

For oriented stable minimal hypersurfaces we have the stability inequality (see (1.1.31))

$$\int_M |\nabla \phi|^2 * 1 \geq \int_M |B|^2\phi^2 * 1, \qquad (1.2.8)$$

where supp ϕ is compact. Replacing ϕ by $|B|^{1+q}\phi$ in (1.2.8) for $q \geq 0$ gives

$$\int_M |B|^{4+2q}\phi^2 * 1 \leq \int_M \Big[(1+q)^2|B|^{2q}|\nabla|B||^2\phi^2 + |B|^{2(1+q)}|\nabla\phi|^2 + 2(1+q)\phi|B|^{1+2q}(\nabla\phi)\cdot(\nabla|B|)\Big] * 1. \tag{1.2.9}$$

Multiplying $\phi^2|B|^{2q}$ with both sides of (1.2.7) and integrating by parts, we have

$$\frac{2}{n}\int_M \phi^2|B|^{2q}|\nabla|B||^2 * 1 \leq -(1+2q)\int_M \phi^2|B|^{2q}|\nabla|B||^2 * 1 + \int_M |B|^{2q+4}\phi^2 * 1 - 2\int_M \phi|B|^{2q+1}(\nabla\phi)\cdot(\nabla|B|) * 1. \tag{1.2.10}$$

Adding up both sides of (1.2.9) and (1.2.10) yields

$$\left[\frac{2}{n} - q^2\right]\int_M \phi^2|B|^{2q}|\nabla|B||^2 * 1 \leq \int_M |B|^{2q+2}|\nabla\phi|^2 * 1 + 2q\int_M \phi|B|^{2q+1}(\nabla\phi)\cdot(\nabla|B|) * 1. \tag{1.2.11}$$

Since

$$2q\phi|B|^{2q+1}(\nabla\phi)\cdot(\nabla|B|) \leq 2q\phi|B|^{2q+1}|\nabla\phi||\nabla|B|| \leq \varepsilon q^2\phi^2|B|^{2q}|\nabla|B||^2 + \varepsilon^{-1}|B|^{2q+2}|\nabla\phi|^2,$$

(1.2.11) becomes

$$\left[\frac{2}{n} - (1+\varepsilon)q^2\right]\int_M \phi^2|B|^{2q}|\nabla|B||^2 * 1 \leq \int_M (1+\varepsilon^{-1})|B|^{2q+2}|\nabla\phi|^2 * 1. \tag{1.2.12}$$

By using (1.2.9) and (1.2.12) we are able to prove the following result.

THEOREM 1.2.1 ([S-S-Y]) *For any* $p \in \left[4, 4+\sqrt{\frac{8}{n}}\right)$ *and any non-negative function* ϕ *with compact support*

$$\int_M |B|^p\phi^p * 1 \leq \beta\int_M |\nabla\phi|^p * 1, \tag{1.2.13}$$

where β *is a constant depending only on* n, p.

PROOF. Set $q = \frac{p-4}{2}$. Then

$$q \geq 0, \quad q^2 = \frac{(p-4)^2}{4} < \frac{2}{n}.$$

Choose ε sufficiently small, such that

$$\frac{2}{n} - (1+\varepsilon)q^2 > 0.$$

Thus, (1.2.12) becomes

$$\int_M \phi^2 |B|^{2q} |\nabla |B||^2 * 1 \\ \leq \beta_1 \int_M |B|^{2q+2} |\nabla \phi|^2 * 1, \tag{1.2.14}$$

where β_1 is a constant depending only on n and p.

On the other hand, (1.2.9) shows that

$$\begin{aligned} \int_M |B|^p \phi^2 * 1 \leq & \int_M \Big[(1+q)^2 |B|^{2q} |\nabla |B||^2 \phi^2 \\ & + 2(1+q)(|B|^q \phi \nabla |B|) \cdot (|B|^{\frac{p}{2}-1} \nabla \phi) \Big] * 1 + \int_M |B|^{p-2} |\nabla \phi|^2 * 1 \\ \leq & \int_M \Big[(1+q)^2 |B|^{2q} |\nabla |B||^2 \phi^2 + (1+q) |B|^{2q} \phi^2 |\nabla |B||^2 \\ & + (1+q) |B|^{p-2} |\nabla \phi|^2 \Big] * 1 + \int_M |B|^{p-2} |\nabla \phi|^2 * 1. \end{aligned} \tag{1.2.15}$$

Replacing ϕ by $\phi^{\frac{p}{2}}$ in (1.2.14) and (1.2.15) yields

$$\int_M |B|^{2q} |\nabla |B||^2 \phi^p * 1 \\ \leq \beta_1' \int_M |B|^{p-2} \phi^{p-2} |\nabla \phi|^2 * 1, \tag{1.2.16}$$

and

$$\begin{aligned} \int_M |B|^p \phi^p * 1 \leq \int_M & \Big[(1+q)^2 |B|^{2q} |\nabla |B||^2 \phi^p + (1+q) |B|^{2q} |\nabla |B||^2 \phi^p \\ & + (2+q) |B|^{p-2} \phi^{p-2} |\nabla \phi|^2 \Big] * 1. \end{aligned} \tag{1.2.17}$$

By using Young's inequality, namely for any positive real number α, a, b, p, q with $\frac{1}{p}+\frac{1}{q}=1$

$$\frac{\alpha^p a^p}{p}+\frac{\alpha^{-q}\, b^q}{q}\geq ab,$$

we have

$$|B|^{p-2}\phi^{p-2}|\nabla\phi|^2\leq\varepsilon|B|^p\phi^p+\beta_2|\nabla\phi|^p, \tag{1.2.18}$$

where β_2 are dependent only on ε and p. Applying (1.2.16), (1.2.18) to (1.2.17) gives

$$(1-\beta_3\varepsilon)\int_M |B|^p\phi^p * 1\leq\beta_4\int_M|\nabla\phi|^p * 1, \tag{1.2.19}$$

where β_3 is dependent only on n, p and β_4 is dependent only on n, p, and ε. To obtain our aim (1.2.13) it suffices to choose $\varepsilon<\frac{1}{\beta_3}$.

Let $B_R(x)\subset\mathbb{R}^{m+n}$ be a ball of radius R and centered at $x\in M$. Its restriction on M is denoted by $D_R(x)=B_R(x)\cap M$.

(1.2.13) enables us to have the following Bernstein type theorem.

Theorem 1.2.2 ([S-S-Y]) *Let M be a complete stable minimal hypersurface in $\mathbb{R}^{n+1}$. If for $p\in\left[4, 4+\sqrt{\frac{8}{n}}\right)$*

$$\lim_{R\to\infty} R^{-p}vol(D_R)=0,$$

then M is a totaly geodesic hypersurface.

Proof. By (1.2.13) we have an estimate

$$\int_M|B|^p\phi^p * 1\leq\beta\int_M|\nabla\phi|^p * 1.$$

Choose a cut-off function ϕ to be

$$\phi=\begin{cases}1, \text{ in } D_R;\\ 0, \text{ outside of } D_{2R}\end{cases}$$

with $\phi\geq 0$, and $|\nabla\phi|\leq\frac{C}{R}$ almost everywhere. We thus have

$$\int_{D_R}|B|^p * 1\leq\int_{D_{2R}}|B|^p\phi^p * 1\leq\beta\,\frac{C^p}{R^p}\int_{D_{2R}} *1=AR^{-p}\mathrm{vol}D_R,$$

where A is a constant depending only on n and p. Letting $R\to\infty$ gives $|B|\equiv 0$ and M is totally geodesic.

Since minimal graphs are stable and have Euclidean volume growth in the sense that

$$\mathrm{vol}(D_R) \le AR^n,$$

where A depending on n. Then, we obtain the Bernstein results for dimension up to 5.

Corollary 1.2.3 ([S-S-Y]) *Let M be an entire minimal graph in $\mathbb{R}^{n+1}$, then M has to be an affine linear space, provided $n \le 5$.*

There is a weak version of the Bernstein type theorem. It was J. Moser [M] who proved that the entire solution f to the minimal surface equation is affine linear, provided $|\nabla f|$ is uniformly bounded. There is no dimension limitation. Ecker-Huisken improved Moser's result by an interesting curvature estimates.

$*$ **E-H Curvature Estimates**

Let M be an entire graph in $\mathbb{R}^{n+1}$ defined by $f : \mathbb{R}^n \to \mathbb{R}$. As shown in the last section we have $v = \sqrt{1 + |\nabla f|^2}$. From (1.1.13) we obtain

$$\Delta v = v\,|B|^2 + \frac{2}{v}|\nabla v|^2. \tag{1.2.20}$$

From (1.2.7) and (1.2.20) we obtain for any real q and s

$$\begin{aligned}\Delta\left(v^q|B|^s\right) \ge & q\,(q+1-\varepsilon^{-1}s)\,v^{q-2}|B|^s|\nabla v|^2 \\ & + s\left(s - \frac{n-2}{n} - \varepsilon\, q\right) v^q|B|^{s-2}|\nabla|B||^2 + (q-s)\; v^q|B|^{s+2}.\end{aligned}$$

Choose s sufficiently large, we have

$$\Delta(v^s|B|^s) \ge 0,$$

$$\Delta(v^s|B|^{s-1}) \ge v^s|B|^{s+1}. \tag{1.2.21}$$

We have the mean value inequality for any subharmonic function on minimal submanifold M in $\mathbb{R}^{n+p}$ which gives

$$\begin{aligned} v^s|B|^s(o) \le & \frac{C}{R^n}\int_{D_R} v^s|B|^s * 1 \\ & \le \frac{C\,\mathrm{vol}(D_R)^{\frac{1}{2}}}{R^n}\left(\int_{D_R} v^{2s}|B|^{2s} * 1\right)^{\frac{1}{2}}, \end{aligned} \tag{1.2.22}$$

where we assume that $o \in M \subset \mathbb{R}^{n+p}$ is any fixed point, C is a constant depending only on n.

Multiplying by $v^s|B|^{s-1}\phi^{2s}$, where ϕ is any smooth function with compact support, in (1.2.21), then integrating by parts and using the Cauchy inequality, we have

$$\begin{aligned}\int_M v^{2s}|B|^{2s}\phi^{2s} * 1 \le \int_M & v^s|B|^{s-1}\phi^{2s}\Delta(v^s|B|^{s-1}) * 1 \\ &= -\int_M \big\langle \nabla(v^s|B|^{s-1}\phi^{2s}), \nabla(v^s|B|^{s-1}\big\rangle * 1 \\ &= -\int_M |\nabla(v^s|B|^{s-1})|^2\phi^{2s} \\ &\quad - 2s\int_M \big\langle \phi^{s-1}|B|^{s-1}v^s\nabla\phi, \phi^s\nabla(v^s|B|^{s-1})\big\rangle \\ &\le C_1(s)\int_M v^{2s}|B|^{2s-2}\phi^{2s-2}|\nabla\phi|^2 * 1. \end{aligned} \tag{1.2.23}$$

By using Young's inequality

$$ab \le \frac{\alpha^p a^p}{p} + \frac{\alpha^{-q} b^q}{q}$$

for any real numbers p, q, α, a, b with $\frac{1}{p} + \frac{1}{q} = 1$, (1.2.23) becomes

$$\int_M v^{2s}|B|^{2s}\phi^{2s} * 1 \le C_2(s)\int_M v^{2s}|\nabla\phi|^{2s} * 1. \tag{1.2.24}$$

Choosing ϕ as the standard cut-off function, we obtain

$$\begin{aligned}\int_{D_R} v^{2s}|B|^{2s} * 1 &\le C_2(s)\, R^{-2s}\int_{D_{2R}} v^{2s} * 1 \\ &\le C_2(s)\, R^{-2s}\mathrm{vol}(D_{2R})\sup_{D_{2R}} v^{2s}, \end{aligned} \tag{1.2.25}$$

Noting that the minimal graph has Euclidean volume growth , then (1.2.22) and (1.2.25) gives an estimate

$$|B|(o) \le C(n)R^{-1}\sup_{D_R} v. \tag{1.2.26}$$

This estimate yields the following result:

Theorem 1.2.4 ([E-H]) *An entire smooth solution f of the minimal surface equation satisfying*

$$|\nabla f| = o\left(\sqrt{|x|^2 + |f(x)|^2}\right) \tag{1.2.27}$$

is a linear function and its graph is an affine subspace.

Corollary 1.2.5 ([M]) *An entire smooth solution f of the minimal surface equation with $|\nabla f| \leq C$ for any constant $C > 0$ is a linear function and its graph is an affine subspace.*

CHAPTER II

Geometry of Grassmannian Manifolds

In order to prove Bernstein type theorems in higher codimension we study the Gauss maps whose images in our case lie in Grassmannian manifolds. It is natural to study their geometric properties, which is interesting in its own right.

2.1 Riemannian Metric on $\mathbf{G}_{n,m}$

Let $\mathbb{R}^{m+n}$ be an $(m+n)-$dimensional Euclidean space. The set of all oriented $n-$subspaces (called $n-$planes) constitutes the Grassmannian manifold $\mathbf{G}_{n,m}$, which is the irreducible symmetric space.

Let P and Q be two points in $\mathbf{G}_{n,m}$. The angles between P and Q are defined by the critical values of the angle θ between a nonzero vector x in P and its orthogonal projection x^* in Q as x runs through P. Assume that $e_1, \cdots, e_n$ are orthonormal vectors which span P, and $f_1, \cdots, f_n$ for Q. For a nonzero vector

$$x = \sum_\alpha x_\alpha e_\alpha,$$

its orthogonal projection in Q is

$$x^* = \sum_\alpha x^*_\alpha f_\alpha.$$

Thus, for any y in Q we have

$$\langle x - x^*, y \rangle = 0.$$

Assume that

$$a_{\alpha\beta} = \langle e_\alpha, f_\beta \rangle .$$

We then have

$$x^*_\beta = \sum_\alpha a_{\alpha\beta} x_\alpha,$$

and

$$\cos\theta = \frac{\left\langle \sum_\alpha x_\alpha e_\alpha, \sum_\beta x^*_\beta f_\beta \right\rangle}{\sqrt{\sum_\alpha x^2_\alpha}\sqrt{\sum_\alpha {x^*_\alpha}^2}} = \frac{\sqrt{\sum_{\alpha,\,\beta} A_{\alpha\beta} x_\alpha x_\beta}}{\sqrt{\sum_\alpha x^2_\alpha}},$$

where $A_{\alpha\beta} = \sum_\gamma a_{\alpha\gamma} a_{\beta\gamma}$ is symmetric in α and β. It follows that the angles θ_α between P and Q (K. Jordan, 1875) are

$$\theta_\alpha = \cos^{-1}(\lambda_a), \qquad 0 \le \theta_\alpha \le \frac{\pi}{2},$$

where λ^2_α are the eigenvalues of the symmetric matrix $(A_{\alpha\beta})$. it is independent of the choices of $\{e_i\}$ in P and $\{f_i\}$ in Q.

The distance between P and Q are defined by

$$d(P,Q) = \sqrt{\sum \theta^2_\alpha}. \tag{2.1}$$

The canonical Riemannian metric on $\mathbf{G}_{n,m}$ can be defined this way. Let $\{e_\alpha, e_{n+i}\}$ be a local orthonormal frame field in $\mathbb{R}^{m+n}$, where $i, j, \cdots = 1, \cdots, m;\quad \alpha, \beta, \cdots = 1, \cdots, n;\quad a, b, \cdots = 1, \cdots, m+n$ (say, $n \le m$). Let $\{\omega_\alpha, \omega_{n+i}\}$ be its dual frame field so that the Euclidean metric is

$$g = \sum_\alpha \omega^2_\alpha + \sum_i \omega^2_{n+i}.$$

The Levi-Civita connection forms ω_{ab} of $\mathbb{R}^{m+n}$ are uniquely determined by the structure equations

$$\begin{aligned} d\omega_a &= \omega_{ab} \wedge \omega_b, \\ \omega_{ab} + \omega_{ba} &= 0. \end{aligned} \tag{2.2}$$

The Riemannian metric on $\mathbf{G}_{n,m}$ can be written as

$$ds^2 = \sum_{\alpha,\,i} \omega^2_{\alpha\, n+i} \tag{2.3}$$

From (2.2) and (2.3) it is easily seen that the curvature tensor of $\mathbf{G}_{n,m}$ is

$$\begin{aligned} \mathrm{R}_{\alpha i\,\beta j\,\gamma k\,\delta l} =& \delta_{\alpha\beta}\delta_{\gamma\delta}\delta_{ik}\delta_{jl} + \delta_{\alpha\gamma}\delta_{\beta\delta}\delta_{ij}\delta_{kl} \\ &- \delta_{\alpha\beta}\delta_{\gamma\delta}\delta_{il}\delta_{kj} - \delta_{\alpha\delta}\delta_{\beta\gamma}\delta_{ij}\delta_{kl} \end{aligned} \tag{2.4}$$

in a local orthonormal frame field $\{e_{\alpha\, n+i}\}$, which is dual to $\{\omega_{\alpha\, n+i}\}$

The canonical Riemannian metric of $\mathbf{G}_{n,m}$ can also be expressed by matrix calculation in a local coordinates. Let us introduce now. Let P_0 be an oriented n-plane in $\mathbb{R}^{m+n}$. We represent it by n vectors e_α, which are complemented by m vectors e_{n+i}, such that $\{e_\alpha, e_{n+i}\}$ form an orthonormal base of $\mathbb{R}^{m+n}$. Then we can span the n-planes P in a neighborhood $\mathbb{U}$ of P_0 by n vectors f_α:

$$f_\alpha = e_\alpha + z_{\alpha i} e_{n+i},$$

where $(z_{\alpha i})$ are the local coordinates of P in $\mathbb{U}$. The metric (2.3) on $\mathbf{G}_{n,m}$ in those local coordinates can be described as

$$ds^2 = tr((I_n + ZZ^T)^{-1} dZ (I_m + Z^T Z)^{-1} dZ^T) \tag{2.5}$$

where $Z = (z_{\alpha i})$ is an $(n \times m)$-matrix and I_n (res. I_m) denotes the $(n \times n)$-identity (res. $m \times m$) matrix.

2.2 Geodesic Convex Sets on $\mathbf{G}_{n,m}$

We know the usual convex geodesic ball $B_R(x_0)$ from a fixed point x_0 in a Riemannian manifold N. When the sectional curvature of N is bounded above by κ, then $R \le \frac{\pi}{2\sqrt{\kappa}}$. For Grassmannian manifolds $\mathbf{G}_{n,m}$ we know that $\kappa = 2$. It is interesting to find the larger one.

Let N be a Riemannian manifold with curvature tensor $R(\cdot, \cdot)$. Let γ be a geodesic issuing from x_0 with $\gamma(0) = x_0$ and $\gamma(t) = x$, where t is the arc length parameter. Define a self-adjoint map along the geodesic γ

$$R_{\dot\gamma} : w \to R(\dot\gamma, w)\dot\gamma.$$

Let v be a unit eigenvector of $R_{\dot\gamma(0)}$ with eigenvalue μ and $\langle v, \dot\gamma(0)\rangle = 0$. Let $v(t)$ be the vector field obtained by parallel translation of v along γ. In the case of N being a locally symmetric space with nonnegative sectional curvature, $v(t)$ is an eigenvector of $R_{\dot\gamma(t)}$ with eigenvalue $\mu \ge 0$, namely

$$R(\dot\gamma(t), v(t))\dot\gamma = \mu v(t).$$

Thus,

$$J(t) = \begin{cases} \frac{1}{\sqrt{\mu}} \sin(\sqrt{\mu} t) v(t), & \text{when} \quad \mu > 0 \\ t\, v(t), & \text{when} \quad \mu = 0 \end{cases}$$

is a Jacobi field along $\gamma(t)$ with $J(0) = 0$. On the other hand, the Hessian of the distance function r from x_0 can be computed by those Jacobi fields. Now, we assume γ is a geodesic without a conjugate point up to distance r from x_0. For orthonormal vectors $X, Y \in T_{\gamma(r)}S_r(x_0)$ there exist unique Jacobi fields J_1 and J_2 such that

$$J_1(0) = J_2(0) = 0, \; J_1(r) = X, \; J_2(r) = Y,$$

since there is no conjugate point of x_0 along γ. We then have

$$\mathrm{Hess}(r)(X, Y) = \langle \nabla_{\dot\gamma} J_1, J_2 \rangle$$

Assume that μ_i and $v_i(t)$ are eigenvalues and orthonormal eigenvectors of $R_{\dot\gamma(t)}$. Then

$$J_i(t) = \frac{1}{\sqrt{\mu_i}} \sin(\sqrt{\mu_i} t) v_i(t)$$

are $n-1$ orthogonal Jacobi fields, where $\mu_i > 0$.

$$\mathrm{Hess}(r)(J_i, J_j) = \frac{1}{\sqrt{\mu_j}} \sin(\sqrt{\mu_j} r) \cos(\sqrt{\mu_i} r) \delta_{ij}$$

and

$$\mathrm{Hess}(r)(v_i(r), v_j(r)) = \sqrt{\mu_i} \cot(\sqrt{\mu_i} r) \delta_{ij} \tag{2.6}$$

(In the case $\mu_i = 0, \mathrm{Hess}(r)(v_i(r), v_i(r)) = \frac{1}{r}$).

On the other hand

$$\mathrm{Ric}(\dot\gamma, \dot\gamma) = \sum_i \langle R(\dot\gamma, v_i)\dot\gamma, v_i \rangle = \sum_i \mu_i.$$

Let us now compute those eigenvalues μ_i for $\mathbf{G}_{n,m}$.

Let $\dot\gamma = x_{\alpha i} e_{\alpha\, i}$ and $v = v_{\alpha i} e_{\alpha\, i}$. Then from (2.4)

$$\langle \mathrm{R}(\dot\gamma, e_{\beta j})\dot\gamma, v \rangle = x_{\beta i} x_{\alpha i} v_{\alpha j} + x_{\alpha j} x_{\alpha l} v_{\beta l} - 2\, x_{\beta l} x_{\alpha j} v_{\alpha l}. \tag{2.7}$$

By an action of $SO(m) \times SO(n)$

$$x_{\alpha i} = \lambda_\alpha \delta_{\alpha i},$$

where $\sum_\alpha \lambda_\alpha^2 = 1$. Then, we have

$$\begin{aligned}
\mathrm{R}_{\dot\gamma} v &= (\lambda_\alpha\lambda_\beta\delta_{\alpha i}\delta_{\beta i}v_{\alpha j} + \lambda_\alpha^2\delta_{\alpha j}\delta_{\alpha l}v_{\beta l} - 2\,\lambda_\alpha\lambda_\beta\delta_{\alpha j}\delta_{\beta l}v_{\alpha l})e_{\beta j}\\
&= (\lambda_\beta^2 v_{\beta j} + \lambda_\alpha^2\delta_{\alpha j}v_{\beta\alpha} - 2\,\lambda_\alpha\lambda_\beta\delta_{\alpha j}v_{\alpha\beta})e_{\beta j}\\
&= \begin{cases} (\lambda_\beta^2 v_{\beta\alpha} + \lambda_\alpha^2 v_{\beta\alpha} - 2\,\lambda_\alpha\lambda_\beta v_{\alpha\beta})e_{\beta\alpha}, \\ \qquad\qquad \text{when} \quad j = \alpha = 1, \cdots, n; \\ \lambda_\beta^2 v_{\beta s}e_{\beta s}, \qquad \text{when} \qquad j = s = n+1, \cdots, m. \end{cases}
\end{aligned}$$

For any $n \times m$ matrix V, there is an orthogonal decomposition

$$V = (V_1, 0) + (V_2, 0) + (0, V_3),$$

where V_1 is an $n \times n$ symmetric matrix, V_2 is an $n \times n$ skew-symmetric matrix and V_3 is an $n \times (m-n)$ matrix. For $v_{\beta\alpha} = v_{\alpha\beta}$

$$\mathrm{R}_{\dot\gamma} v_{\beta\alpha}e_{\beta\alpha} = (\lambda_\alpha - \lambda_\beta)^2 v_{\beta\alpha}e_{\beta\alpha}.$$

For $v_{\beta\alpha} = -v_{\alpha\beta}$

$$\mathrm{R}_{\dot\gamma} v_{\beta\alpha}e_{\beta\alpha} = (\lambda_\alpha + \lambda_\beta)^2 v_{\beta\alpha}e_{\beta\alpha}.$$

In summary, $\mathrm{R}_{\dot\gamma}$ has eigenvalues:

$$\begin{array}{ll}
\lambda_1^2 & \text{with multiplicity } m-n \\
\vdots & \vdots \\
\lambda_n^2 & \text{with multiplicity } m-n \\
(\lambda_\alpha + \lambda_\beta)^2 & \text{with multiplicity} \quad 1 \\
(\lambda_\alpha - \lambda_\beta)^2 & \text{with multiplicity} \quad 1 \\
0 & \text{with multiplicity} \quad n-1
\end{array}$$

for each pair α and β with $\alpha \neq \beta$. From (2.6) it follows that the eigenvalues of the Hessian of the distance function r from a fixed point at the direction $X = (x_{\alpha i}) = (\lambda_\alpha\delta_{\alpha i})$ are the same as the ones at $X_1 = (|\lambda_\alpha|\delta_{\alpha i})$. They are as follows.

$$\begin{array}{ll}
\lambda_1 \cot(\lambda_1 r) & \text{with multiplicity } m-n \\
\vdots & \vdots \\
\lambda_n \cot(\lambda_n r) & \text{with multiplicity } m-n \\
(\lambda_\alpha + \lambda_\beta)\cot(\lambda_\alpha + \lambda_\beta)r & \text{with multiplicity} \quad 1 \\
(\lambda_\alpha - \lambda_\beta)\cot(\lambda_\alpha - \lambda_\beta)r & \text{with multiplicity} \quad 1 \\
\dfrac{1}{r} & \text{with multiplicity} \quad n-1
\end{array} \tag{2.8}$$

where $\lambda_\alpha > 0$ without loss of generality.

Let $(x_{\alpha i}) = (\lambda_\alpha \delta_{\alpha i})$ be a unit tangent vector at P_0. The geodesic γ from P_0 at the direction $(x_{\alpha i})$ in $\mathbb{U}$ is (see [W])

$$(z_{\alpha i}(t)) = \begin{pmatrix} \tan(\lambda_1 t) & & 0 & \\ & \ddots & & 0 \\ 0 & & \tan(\lambda_n t) & \end{pmatrix} \tag{2.9}$$

where t is the arc length parameter and $0 \le t < \frac{\pi}{2|\lambda_\alpha|}$ with $|\lambda_\alpha| = \max(|\lambda_1|, \cdots, |\lambda_n|)$.

Now, let us define an open set $B_{JX}(P_0)$ in $\mathbb{U} \subset \mathbf{G}_{n,m}$. In $\mathbb{U}$ we have the normal coordinates around P_0, and then the normal polar coordinates around P_0. Define $B_{JX}(P_0)$ in normal polar coordinates around P_0 as follows:

$$B_{JX}(P_0) = \left\{ (X,t); X = (\lambda_\alpha \delta_{\alpha i}), 0 \le t < t_X = \frac{\pi}{2(|\lambda_{\alpha'}| + |\lambda_{\beta'}|)} \right\}, \tag{2.10}$$

where $\lambda_{\alpha'}$ and $\lambda_{\beta'}$ are two eigenvalues with largest absolute values. From (2.9) we see that $B_{JX}(P_0)$ lies inside the cut locus of P_0. We also know from (2.8) that the square of the distance function r^2 from P_0 is a strictly convex smooth function in $B_{JX}(P_0)$.

Remark The above definition of $B_{JX}(P_0)$ is for the case of $m \ge n > 1$. If $n = 1$, $\mathbf{G}_{1,m}$ is the usual sphere S^m and the defined set is the open hemisphere as usual.

We verify that $B_{JX}(P_0)$ is a geodesic convex set below. Let $P = ((\lambda_\alpha \delta_{\alpha i}), t)$ and $Q = ((\lambda'_\alpha \delta_{\alpha i}), t')$ be two points in $B_{JX}(P_0)$. Then the local expression of P in $\mathbb{U}$ is the $n \times m$ matrix $(\tan(\lambda_\alpha t)\delta_{\alpha i})$, similarly that of Q is $(\tan(\lambda'_\alpha t')\delta_{\alpha i})$. Consider a curve Γ between P and Q defined by

$$(\tan(\lambda_\alpha t(1-h) + \lambda'_\alpha t' h)\delta_{\alpha i})$$

in $\mathbb{U}$, there $0 \le h \le 1$ is the parameter for Γ. We can prove that Γ is a geodesic.

Let P' be the middle point defined by

$$\left(\tan\left(\frac{\lambda_\alpha t + \lambda'_\alpha t'}{2}\right)\delta_{\alpha i}\right)$$

in $\mathbb{U}$. We can prove that geodesic Γ has the following properties:

(1) $\Gamma \subset B_{JX}(P_0)$;
(2) $\Gamma \subset B_{JX}(P')$.

These two properties of the geodesic Γ mean that any minimal geodesic γ from P to Q lies in $B_{JX}(P_0)$ and has length less than $2t_X$, where X is the unit tangent vector of γ at P. On the other hand, we already computed all Jacobi fields along any geodesic in $\mathbf{G}_{n,m}$, which means that any geodesic from P of length $< 2t_X$ has no conjugate points and that the squared distance function from P remains strictly convex along this geodesic for length $< t_X$. Thus, we conclude that $B_{JX}(P_0)$ shares all the properties of the usual convex geodesic ball. In summary, we have

Theorem 2.1 ([J-X] *In $B_{JX}(P_0)$ the square of the distance function from its center P_0 is a smooth strictly convex function. Furthermore, $B_{JX}(P_0)$ is a convex set, namely any two points in $B_{JX}(P_0)$ can be joined in $B_{JX}(P_0)$ by a unique geodesic arc. This arc is the shortest connection between its end points and thus in particular does not contain a pair of conjugate points.*

Remark In the Grassmannian manifold there is the usual convex geodesic ball $B_R(P_0)$ of radius

$$R \le \begin{cases} \frac{\pi}{2\sqrt{2}} & \text{when} \quad \min(m,n) > 1; \\ \frac{\pi}{2} & \text{when} \quad \min(m,n) = 1. \end{cases}$$

From (2.10) it is seen that $B_R(P_0) \subset B_{JX}(P_0)$.

Any point in a Grassmannian manifold can be described by an $n-$vector. The inner product of two $n-$vectors is also related to its distance in $\mathbf{G}_{n,m}$. We study this relation in $B_{JX}(P_0)$.

Let $P(t)$ be any $n-$plane in $\mathbb{U}$ of P_0 which is spanned by

$$f_\alpha = e_\alpha + z_{\alpha i} e_{n+i},$$

where $z_{\alpha i}$ is defined by (2.9). Let

$$\tilde{f}_1 = \cos(\lambda_1 t) f_1, \ \cdots, \ \tilde{f}_n = \cos(\lambda_n t) f_n.$$

Since $|f_\alpha| = \frac{1}{\cos(\lambda_\alpha t)}$, the vectors $\tilde{f}_1, \cdots, \tilde{f}_n$ are orthonormal.

Therefore, we can define the inner product $\langle P_0, P\rangle$ of $n-$planes $P_0 = e_1 \wedge \cdots \wedge e_n$ and $P = \tilde{f}_1 \wedge \cdots \wedge \tilde{f}_n$ by

$$\langle P_0, P\rangle = \det\left(\left\langle e_\alpha, \tilde{f}_\beta \right\rangle\right).$$

It follows that

$$\langle P_0, P(t)\rangle = \det \begin{pmatrix} \cos(\lambda_1 t) & & & 0 \\ & \cos(\lambda_2 t) & & \\ & & \ddots & \\ 0 & & & \cos(\lambda_n t) \end{pmatrix} = \prod_{\alpha=1}^{n} \cos(\lambda_\alpha t).$$

By an elementary method we prove that

Theorem 2.2 ([FC])

$$max\{\langle P_0, P\rangle\,;\ P \in \partial B_{\frac{\pi}{2\sqrt{2}}}(P_0)\} = \cos^{-s}\left(\frac{\pi}{2\sqrt{2s}}\right),\ s = \min(m,n). \tag{2.11}$$

Theorem 2.3 ([J-X])

$$max\{\langle P_0, P\rangle\,;\ P \in \partial B_{JX}(P_0)\} = \frac{1}{2}. \tag{2.12}$$

2.3 Convex functions on $\mathbf{G}_{n,m}$

We view now the Grassmannian manifolds as submanifolds in Euclidean space via Plücker imbedding.

Fix $P_0 \in \mathbf{G}_{n,m}$ in the sequel, which is expressed by a unit $n-$vector $\varepsilon_1 \wedge \cdots \wedge \varepsilon_n$. For any $P \in \mathbf{G}_{n,m}$, expressed by an $n-$vector $e_1 \wedge \cdots \wedge e_n$, we define an important function on $\mathbf{G}_{n,m}$

$$w \overset{def.}{=} \langle P, P_0\rangle = \langle e_1 \wedge \cdots \wedge e_n, \varepsilon_1 \wedge \cdots \wedge \varepsilon_n\rangle = \det W,$$

where $W = (\langle e_i, \varepsilon_j\rangle)$.

Denote

$$\mathbb{U} = \{P \in \mathbf{G}_{n,m} : w(P) > 0\}.$$

Let $\{\varepsilon_{n+\alpha}\}$ be m vectors such that $\{\varepsilon_i, \varepsilon_{n+\alpha}\}$ form an orthornormal basis of $\mathbb{R}^{m+n}$. Then we can span arbitrary $P \in \mathbb{U}$ by n vectors f_i:

$$f_i = \varepsilon_i + z_{i\alpha}\varepsilon_{n+\alpha},$$

where $Z = (z_{i\alpha})$ are the local coordinates of P in $\mathbb{U}$. Here and in the sequel we use the summation convention and agree the range of indices:

$$1 \le i, j \le n; \qquad 1 \le \alpha, \beta \le m.$$

The Jordan angles between P and P_0 are defined by

$$\theta_\alpha = \arccos(\lambda_\alpha),$$

where $\lambda_\alpha \ge 0$ and λ_α^2 are the eigenvalues of the symmetric matrix $W^T W$. On $\mathbb{U}$ we can define

$$v = w^{-1}.$$

Then it is easily seen that

$$v(P) = \left[\det(I_n + ZZ^T)\right]^{\frac{1}{2}} = \prod_{\alpha=1}^{m} \sec\theta_\alpha.$$

Let $E_{i\alpha}$ be the matrix with 1 in the intersection of row i and column α and 0 otherwise. Denote $g_{i\alpha,j\beta} = \langle E_{i\alpha}, E_{j\beta}\rangle$ and let $\big(g^{i\alpha,j\beta}\big)$ be the inverse matrix of $\big(g_{i\alpha,j\beta}\big)$. Then,

$$(1+\lambda_i^2)^{\frac{1}{2}}(1+\lambda_\alpha^2)^{\frac{1}{2}} E_{i\alpha}$$

form an orthonormal basis of $T_P\mathbf{G}_{n,m}$, where $\lambda_\alpha = \tan\theta_\alpha$. Denote its dual basis in $T_P^*\mathbf{G}_{n,m}$ by $\omega_{i\alpha}$.

A lengthy computation yields

$$\begin{aligned}\mathrm{Hess}(v)_P = &\sum_{m+1\le i\le n,\alpha} v\,\omega_{i\alpha}^2 + \sum_\alpha (1+\lambda_\alpha^2) v\,\omega_{\alpha\alpha}^2 + v^{-1}\,dv\otimes dv\\ &+ \sum_{\alpha<\beta}\Big[(1+\lambda_\alpha\lambda_\beta)v\Big(\frac{\sqrt{2}}{2}(\omega_{\alpha\beta}+\omega_{\beta\alpha})\Big)^2\\ &+ (1-\lambda_\alpha\lambda_\beta)v\Big(\frac{\sqrt{2}}{2}(\omega_{\alpha\beta}-\omega_{\beta\alpha})\Big)^2\Big].\end{aligned} \tag{2.13}$$

For any real number a let $\mathbb{V}_a = \{P \in \mathbf{G}_{n,m},\quad v(P) < a\}$. From Theorem 2.2 we know that

$$\mathbb{V}_2 \subset B_{JX} \qquad \text{and} \qquad \overline{\mathbb{V}}_2 \cap \overline{B}_{JX} \ne \emptyset$$

$\mathrm{Hess}(v)_P$ is positive definite if and only if $\theta_\alpha + \theta_\beta < \frac{\pi}{2}$ for arbitrary $\alpha \ne \beta$, i.e., $P \in B_{JX}(P_0)$.

From (2.13) it is easy to get an estimate

$$\mathrm{Hess}(v) \ge v(2-v)g + v^{-1} dv \otimes dv \qquad \text{on } \overline{\mathbb{V}}_2.$$

For later applications the above estimate is not accurate enough. Using the radial compensation technique the estimate could be refined.

THEOREM 2.3 ([X-Y2])
v is a convex function on $B_{JX}(P_0) \subset \mathbb{U} \subset \mathbf{G}_{n,m}$, and

$$Hess(v) \geq v(2-v)g + \Big(\frac{v-1}{pv(v^{\frac{2}{p}}-1)} + \frac{p+1}{pv}\Big) dv \otimes dv \qquad (2.14)$$

on $\overline{\mathbb{V}}_2$, where g is the metric tensor on $\mathbf{G}_{n,m}$ and $p = min(n,m)$.

REMARK For any $a \leq 2$, the sub-level set $\mathbb{V}_a$ is a convex set in $\mathbf{G}_{n,m}$.

REMARK The sectional curvature varies in $[0,2]$ under the canonical Riemannian metric on $\mathbf{G}_{n,m}$. By the standard Hessian comparison theorem we have

$$\mathrm{Hess}(\rho) \geq \sqrt{2}\cot(\sqrt{2}\rho)(g - d\rho \otimes d\rho), \qquad (2.15)$$

where ρ is the distance function from a fixed point in $\mathbf{G}_{n,m}$.

CHAPTER III

Bernstein Type Theorems for Higher Codimension

3.1 Harmonic Gauss Maps

In this section some basic notions on harmonic Gauss maps will be described. Here we only introduce some related notions and formulas. For more detail, please consult author's book [X3].

Let (M, g) and (N, h) be Riemannian manifolds with metric tensors g and h, respectively. Harmonic maps are described as critical points of the following energy functional

$$E(f) = \frac{1}{2}\int_M e(f) * 1,$$

where $e(f)$ stands for the energy density. The Euler-Lagrange equation of the energy functional is

$$\tau(f) = 0,$$

where $\tau(f)$ is the tension field. In local coordinates

$$\begin{aligned} e(f) &= g^{ij}\frac{\partial f^\beta}{\partial x^i}\frac{\partial f^\gamma}{\partial x^j}h_{\beta\gamma}, \\ \tau(f) &= (\Delta_M f^\alpha + g^{ij}\Gamma^\alpha_{\beta\gamma}\frac{\partial f^\beta}{\partial x^i}\frac{\partial f^\gamma}{\partial x^j})\frac{\partial}{\partial y^\alpha}, \end{aligned}$$

where $\Gamma^\alpha_{\beta\gamma}$ denotes the Christoffel symbols of the target manifold N.

A Riemannian manifold M is said to be simple, if it can be described by coordinates x on $\mathbb{R}^n$ with a metric

$$ds^2 = g_{ij}dx^i dx^j,$$

for which there exist positive numbers λ and μ such that

$$\lambda|\xi|^2 \le g_{ij}\xi^i\xi^j \le \mu|\xi|^2 \tag{3.1}$$

for all x and ξ in $\mathbb{R}^n$. In other words, M is topologically $\mathbb{R}^n$ with a metric for which the associated Laplace operator is uniformly elliptic on $\mathbb{R}^n$.

Hildebrandt-Jost-Widman derived Hölder estimates for harmonic maps with values in Riemannian manifolds with an upper bound for sectional curvature and by a scaling argument then concluded a Liouville type theorem for harmonic maps under certain assumptions. The Hölder estimates needed a bound on the radius of the image, and examples show that [H-J-W] had achieved the optimal bound in the general framework for that paper. Precisely, they proved that

THEOREM 3.1 ([H-J-W]) *Let f be a harmonic map from a simple or compact Riemannian manifold M into a complete Riemannian manifold N, the sectional curvature of which is bounded above by a constant $\kappa \geq 0$. Denote by $B_R(Q)$ a geodesic ball in N with radius $R < \frac{\pi}{2\sqrt{\kappa}}$ which does not meet the cut locus of its center Q. Assume also that the range $f(M)$ of the map f is contained in $B_R(Q)$. Then f is a constant map.*

REMARK In the case where $B_R(Q)$ is replaced by $B_{JX}(P_0)$, which is constructed in the last section, the iteration technique in [H-J-W] is still applicable and the result remains true.

Let $M \to \mathbb{R}^{m+n}$ be an n-dimensional oriented submanifold in Euclidean space. Choose an orthonormal frame field $\{e_1, ..., e_{m+n}\}$ in $\mathbb{R}^{m+n}$ such that the $e'_\alpha s$ are tangent to M. Let $\{\omega_1, ..., \omega_{m+n}\}$ be its coframe field. By the structure equations of $\mathbb{R}^{m+n}$ along M

$$\omega_{n+i\,\alpha} = h_{i\alpha\beta}\omega_\beta,$$

where the $h_{i\alpha\beta}$, the coefficients of the second fundamental form of M in $\mathbb{R}^{m+n}$, are symmetric in α and β. Let 0 be the origin of $\mathbb{R}^{m+n}$. Let $SO(m+n)$ be the manifold consisting of all the orthonormal frames $(0; e_\alpha, e_{n+i})$. Let $P = \{(x; e_1, ..., e_n); x \in M, e_\alpha \in T_xM\}$ be the principal bundle of orthonormal tangent frames over $M, Q = \{(x; e_{n+1}, ..., e_{m+n}); x \in M, e_{n+i} \in N_xM\}$ be the principal bundle of orthonormal normal frames over M, then $\bar{\pi} : P \oplus Q \to M$ is the projection with fiber $SO(m) \times SO(n)$, $i : P \oplus Q \hookrightarrow SO(m+n)$ is the natural inclusion.

The generalized Gauss map $\gamma : M \to \mathbf{G}_{n,m}$ is defined by

$$\gamma(x) = T_xM \in \mathbf{G}_{n,m}$$

via the parallel translation in $\mathbb{R}^{m+n}$ for $\forall x \in M$. Thus, the following commutative diagram holds

$$\begin{array}{ccc} P \oplus Q & \xrightarrow{\ i\ } & SO(m+n) \\ {\scriptstyle \bar{\pi}}\big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ M & \xrightarrow{\ \gamma\ } & \mathbf{G}_{n,m} \end{array}$$

The energy density of the Gauss map (see [X3] Chap.3, §3.1) is

$$e(\gamma) = \frac{1}{2}\langle \gamma_* e_i, \gamma_* e_i \rangle = \frac{1}{2}|B|^2.$$

E. Ruh and J. Vilms discovered the relation between the property of the submanifold and the harmonicity of its Gauss map in [R-V] (see §3.1.5 in [X3] for its simplified proof).

Theorem 3.2 *Let M be a submanifold in $\mathbb{R}^{m+n}$. Then the mean curvature vector of M is parallel if and only if its Gauss map is a harmonic map.*

3.2 Hildebrandt-Jost-Widman's Theorem

The geometric meaning of the condition in the Hildebrandt-Jost-Widman's theorem is that the image under the Gauss map lies in a closed subset of an open geodesic ball of the radius $\frac{\sqrt{2}}{4}\pi$. From Theorem 2.2, Theorem 3.1 and Theorem 3.2, we immediately obtain the result as follows. It is a generalization of Moser's Theorem (see Corollary 1.2.5) to higher codimension.

Theorem 3.3 ([H-J-W]) *Let $z^\alpha = f^\alpha(x), \alpha = 1, \cdots, m,\ x = (x^1, \cdots, x^n) \in \mathbb{R}^n$ be the C^2 solution to the system of minimal surface equations. Let there exist β, where*

$$\beta < \cos^{-s}\left(\frac{\pi}{2\sqrt{s\,K}}\right), \quad K = \begin{cases} 1 & if \quad s = 1 \\ 2 & if \quad s \geq 2 \end{cases}, \quad s = \min(m, n) \qquad (3.2)$$

such that for any $x \in \mathbb{R}^n$,

$$\Delta_f = \left[det\left(\delta_{ij} + \sum_\alpha \frac{\partial f^\alpha}{\partial x^i}\frac{\partial f^\alpha}{\partial x^j} \right) \right]^{\frac{1}{2}} \leq \beta, \qquad (3.3)$$

then $f^1, \cdots, f^m$ are affine linear functions on $\mathbb{R}^n$, whose graph is an affine $n-$plane in $\mathbb{R}^{m+n}$.

3.3 Jost-Xin's Theorem

From Theorem 2.3, Theorem 3.1 and Theorem 3.2, we obtain an improved result.

THEOREM 3.4 ([J-X]) *Let* $z^\alpha = f^\alpha(x^1, \cdots, x^n)$, $\alpha = 1, \cdots, m$, *be smooth functions defined everywhere in* $\mathbb{R}^n$. *Suppose their graph* $M = (x, f(x))$ *is a submanifold with parallel mean curvature in* $\mathbb{R}^{n+m}$. *Suppose that there exists a number* β_0

$$\beta_0 < \begin{cases} 2 & \text{when } m \geq 2, \\ \infty & \text{when } m = 1; \end{cases} \tag{3.4}$$

such that

$$\Delta_f = \left[det \left(\delta_{ij} + \sum_\alpha \frac{\partial f^\alpha}{\partial x^i} \frac{\partial f^\alpha}{\partial x^j} \right) \right]^{\frac{1}{2}} \leq \beta_0. \tag{3.5}$$

Then $f^1, \cdots, f^m$ *has to be affine linear representing an affine* n*-plane.*

3.4 Curvature Estimates for Higher Codimension

It is natural to study the situations:

(1) the image under the Gauss map lies in an open geodesic ball of radius $\frac{\sqrt{2}}{4}\pi$ in $\mathbf{G}_{n,m}$ in Theorem 3.3.

(2) the image under the Gauss map lies in open set $\mathbb{V}_2 \subset B_{JX}$, namely, when β_0 in the condition (3.4) and (3.5) of the Theorem 3.4 approach to 2.

Theorem 3.3 and Theorem 3.4 use Theorem 3.1, the Liouville type theorem for harmonic maps. The derivation of the Theorem need closed conditions and it is fail to deal with the above questions. Therefore, we are interested to generalize the curvature estimates to higher codimension.

$*$ Simons-Bochner Type Formula in Higher Codimension

From (1.1.22) we have

$$\nabla^2 B = -\tilde{\mathcal{B}} - \underline{\mathcal{B}}. \tag{3.6}$$

In the case when $m \geq 2$, there is a refined estimate

$$\left\langle \tilde{\mathcal{B}} + \underline{\mathcal{B}}, B \right\rangle \leq \frac{3}{2}|B|^4.$$

Substituting it into (3.6) gives

$$\langle \nabla^2 B, B \rangle \geq -\frac{3}{2}|B|^4.$$

It follows that

$$\Delta|B|^2 \geq -3|B|^4 + 2|\nabla B|^2. \tag{3.7}$$

Schoen-Yau's formula can also be generalized to higher codimension, namely (in [X1] we derived this type formula in more general situation)

$$|\nabla B|^2 \geq \big(1 + \frac{2}{n}\big)\big|\nabla|B|\big|^2.$$

At last, we have

$$\Delta|B|^2 \geq 2\left(1 + \frac{2}{n}\right)\big|\nabla|B|\big|^2 - 3|B|^4. \tag{3.8}$$

$*$ **Schoen-Simon-Yau's Estimates in Higher Codimension**

We consider smooth functions on an open geodesic ball $B_{\frac{\sqrt{2}}{4}\pi}(P_0) \subset \mathbf{G}_{n,m}$ of radius $\frac{\sqrt{2}}{4}\pi$ and centered at P_0. Those are useful for our curvature estimates later. Let

$$u = \cos(\sqrt{2}\rho),$$

where ρ is the distance function from P_0 in $\mathbf{G}_{n,m}$. We have

$$u' = -\sqrt{2}\sin(\sqrt{2}\rho),$$

$$u'' = -2\cos(\sqrt{2}\rho).$$

Then, from (2.15)

$$\begin{aligned}\text{Hess}(u) &= u'\text{Hess}(\rho) + u'' d\rho \otimes d\rho \\ &\leq -2\cos(\sqrt{2}\rho)(g - d\rho \otimes d\rho) - 2\cos(\sqrt{2}\rho)d\rho \otimes d\rho = -2ug.\end{aligned}$$

The composition function $h_1 = u \circ \gamma$ of u with the Gauss map γ defines a function on M. Using the composition formula, we have

$$\begin{aligned} \Delta h_1 &= \text{Hess}(u)(\gamma_* e_i, \gamma_* e_i) + du(\tau(\gamma)) \\ &\le -2|B|^2 h_1, \end{aligned} \tag{3.9}$$

where $\tau(\gamma)$ is the tension field of the Gauss map, which is zero, provided M has parallel mean curvature by the Ruh-Vilms theorem mentioned above.

For a stable minimal hypersurface there is the stability inequality, which is one of main ingredient for Schoen-Simon-Yau's curvature esimates for stable minimal hypersurfaces. For minimal submanifolds with the Gauss image restriction we have stronger inequality as shown in (3.10) below. With the aid of (3.9), we immediately have the following lemma.

LEMMA 3.5 *Let M be an n-dimensional minimal submanifold of $\mathbb{R}^{n+m}$ (M needs not be complete), if the Gauss image of M is contained in an open geodesic ball of radius $\frac{\sqrt{2}}{4}\pi$ in $\mathbf{G}_{n,m}$, then we have*

$$\int_M |\nabla\phi|^2 * 1 \ge 2\int_M |B|^2\phi^2 * 1 \tag{3.10}$$

for any function ϕ with compact support $D \subset M$.

(3.8) and (3.10) enable us to carry out Schoen-Simon-Yau type estimates. Let r be a function on M with $|\nabla r| \le 1$. For any $R \in [0, R_0]$, where $R_0 = \sup_M r$, suppose

$$M_R = \{x \in M, \quad r \le R\}$$

is compact.

THEOREM 3.6 ([X-Y1]) *Let M be an n-dimensional minimal submanifolds of $\mathbb{R}^{n+m}$. If the Gauss image of M_R is contained in an open geodesic ball of radius $\frac{\sqrt{2}}{4}\pi$ in $\mathbf{G}_{n,m}$, then we have the L^p-estimate*

$$\big\||B|\big\|_{L^p(M_{\theta R})} \le C(n,p)(1-\theta)^{-1}R^{-1}\,Vol(M_R)^{\frac{1}{p}} \tag{3.11}$$

for arbitrary $\theta \in (0,1)$ and

$$p \in \left[4, 4 + \frac{2}{3} + \frac{4}{3}\sqrt{1+\frac{6}{n}}\right).$$

Based on the estimate (2.14) we can define auxiliary functions on $\mathbb{V}_2 = \{P \in \mathbf{G}_{n,m}, \quad v(P) < 2\}$. Let

$$h = v^{-k}(2-v)^k$$

define a positive function on $\mathbb{V}_2$, where $k = \frac{3}{4} + \frac{1}{2s}$ and $s = \min(m,n)$. From (2.14) we have (see (4.4) in [X-Y2])

$$\text{Hess}(h) \le -\left(\frac{3}{2} + \frac{1}{s}\right) h\ g, \tag{3.12}$$

where g is the metric tensor on $\mathbf{G}_{n,m}$.

We assume that the image of M under the Gauss map is contained in $\mathbb{V}_2 \subset \mathbf{G}_{m,n}$. Thus, we have the function $\tilde{h} = h \circ \gamma$ defined on M. We denote h for $\tilde{h}$ in the sequel for simplicity. From (3.12) and the composition formula we have

$$\Delta h \le -\left(\frac{3}{2} + \frac{1}{s}\right) |B|^2 h.$$

We then have strong stability inequality as follows.

LEMMA 3.7 *Let M be an n-dimensional minimal submanifold of $\mathbb{R}^{n+m}$ (M needs not be complete), if the Gauss image of M is contained in $\{P \in \mathbb{U} \subset \mathbf{G}_{n,m} : v(P) < 2\}$, then we have*

$$\int_M |\nabla \phi|^2 * 1 \ge (\frac{3}{2} + \frac{1}{s}) \int_M |B|^2 \phi^2 * 1 \tag{3.13}$$

for any function ϕ with compact support $D \subset M$.

(3.8) and (3.13) enables us to carry out the Schoen-Simon-Yau type estimates.

THEOREM 3.8 ([X-Y2]) *Let M be an n-dimensional minimal submanifolds of $\mathbb{R}^{n+m}$. If the Gauss image of M_R is contained in $\{P \in \mathbb{U} \subset \mathbf{G}_{n,m} : v(P) < 2\}$, then we have the estimate*

$$\||B|\|_{L^p(M_{\theta R})} \le C(n,,p)(1-\theta)^{-1} R^{-1} Vol(M_R)^{\frac{1}{p}} \tag{3.14}$$

for arbitrary $\theta \in (0,1)$ and

$$p \in \left[4, 4 + \frac{4}{3s} + \frac{2}{3}\sqrt{(3+\frac{2}{s})(\frac{6}{n}+\frac{2}{s})}\right).$$

$*$ Ecker-Huisken's Estimates in Higher Codimension

We consider smooth function h_2 on an open geodesic ball $B_{\frac{\sqrt{2}}{4}\pi}(P_0) \subset \mathbf{G}_{n,m}$ of radius $\frac{\sqrt{2}}{4}\pi$ and centered at P_0 defined by

$$h_2 = \sec^2(\sqrt{2}\rho),$$

where ρ is the distance function from P_0 in $\mathbf{G}_{n,m}$. We then have

$$\begin{aligned} Hess(h_2) &= h_2' Hess(\rho) + h_2'' d\rho \otimes d\rho \\ &\geq 4h_2\, g + \frac{3}{2} h_2^{-1} dh_2 \otimes dh_2 \end{aligned} \tag{3.15}$$

We assume that the image under the Gauss map of a minimal submanifold $M \to \mathbb{R}^{m+n}$ is contained in an open geodesic ball of $B_{\frac{\sqrt{2}}{4}\pi}(P_0) \subset \mathbf{G}_{n,m}$ of radius $\frac{\sqrt{2}}{4}\pi$ and centered at P_0. Then we have a function $\tilde{h}_2 = h_2 \circ \gamma$ on M. We denote h_2 for $\tilde{h}_2$ in the sequel for simplicity. From (3.15) and the composition formula we have

$$\Delta h_2 \geq 4h_2 |B|^2 + \frac{3}{2} h_2^{-1} |\nabla h_2|^2. \tag{3.16}$$

(3.8) and (3.16) we can carry out Ecker-Huisken's estimates and obtain the following result.

THEOREM 3.9 ([X-Y1]) *Let $x \in M$, $R > 0$ such that the image of $D_R(x)$ under the Gauss map lies in an open geodesic ball of radius $\frac{\sqrt{2}}{4}\pi$ in $\mathbf{G}_{n,m}$. Then, there exists $C_1 = C_1(n)$, such that*

$$|B|^{2p}(x) \leq C(n,p) R^{-(n+2p)} (\sup_{D_R(x)} h_2)^p Vol(D_R(x)),$$

for arbitrary $p \geq C_1$.

In the case of the image of M under the Gauss map is contained in $\mathbb{V}_2 \subset \mathbf{G}_{m,n}$. Let $h_3 = h^{-2}$. We then have

$$\Delta h_3 \geq 3\; h_3 |B|^2 + \Big(\frac{3}{2} + \frac{1}{3s}\Big) h_3^{-1} |\nabla h_3|^2. \tag{3.17}$$

From (3.8) and (3.17) we have the result as follows.

THEOREM 3.10 ([X-Y2]) *Let $x \in M$, $R > 0$ such that the image of $D_R(x)$ under the Gauss map lies in $\{P \in \mathbb{U} \subset \mathbf{G}_{n,m} : v(P) < 2\}$. Then, there exists $C_1 = C_1(n)$, such that*

$$|B|^{2p}(x) \leq C(n,p) R^{-(n+2p)} (\sup_{D_R(x)} h_3)^p Vol(D_R(x)), \tag{3.18}$$

for arbitrary $p \geq C_1$

$*$ **Geometric Conclusions**

Let $P_0 \in \mathbf{G}_{n,m}$ be a fixed point which is described by $P_0 = \varepsilon_1 \wedge \cdots \wedge \varepsilon_n$, where $\varepsilon_1, \cdots, \varepsilon_n$ are orthonormal vectors in $\mathbb{R}^{m+n}$. Choose complementary orthonormal vectors $\varepsilon_{n+1}, \cdots, \varepsilon_{n+m}$, such that $\{\varepsilon_1, \cdots, \varepsilon_n, \varepsilon_{n+1}, \cdots, \varepsilon_{n+m}\}$ is an orhtonormal base in $\mathbb{R}^{m+n}$.

Let $p : \mathbb{R}^{m+n} \to \mathbb{R}^n$ be the natural projection defined by

$$p(x^1, \cdots, x^n; x^{n+1}, \cdots, x^{m+n}) = (x^1, \cdots, x^n),$$

which induces a map from M to $\mathbb{R}^n$. It is a smooth map from a complete manifold to $\mathbb{R}^n$. It is not difficult to see that p increases the distance with respect to a homothetic change of the induced metric on M, provided $w-$function on M has a positive lower bound. In our first cosideration $w \geq w_0 = \left(\cos \frac{\sqrt{2}}{4}\pi\right)^s$ and the second case $w \geq w_0 = \frac{1}{2}$.

PROPOSITION 3.11 *Let M be a complete submanifold in $\mathbb{R}^{m+n}$. If the $w-$function is bounded below by a positive constant w_0. Then M is an entire graph with Euclidean volume growth. In particular, if the Gauss image of M is contained in a geodesic ball of radius $\frac{\sqrt{2}}{4}\pi$ or the Gauss image of M is contained in $\{P \in \mathbb{U} \subset \mathbf{G}_{n,m} : v(P) < 2\}$, then M is an entire graph with Euclidean volume growth.*

From Proposition 3.11 and Theorem 3.6, Theorem 3.8, Theorem 3.9 and Theorem 3.10 we obtain Bernstein type theorems.

THEOREM 3.12 ([X-Y1]) *Let M be a complete minimal n-dimensional submanifold in $\mathbb{R}^{n+m}$ with $n \leq 6$ and $m \geq 2$. If the Gauss image of M is contained in an open geodesic ball of $\mathbf{G}_{n,m}$ centered at P_0 and of radius $\frac{\sqrt{2}}{4}\pi$, then M has to be an affine linear subspace.*

THEOREM 3.13 ([X-Y2]) *Let $M = (x, f(x))$ be an n-dimensional entire minimal graph given by m functions $f^\alpha(x^1, \cdots, x^n)$ with $n \leq 5$ and $m \geq 2$. If*

$$\Delta_f = \left[det\left(\delta_{ij} + \sum_\alpha \frac{\partial f^\alpha}{\partial x^i}\frac{\partial f^\alpha}{\partial x^j}\right)\right]^{\frac{1}{2}} < 2,$$

then f^α has to be affine linear functions representing an affine n-plane.

For larger dimension n, we need growth conditions (3.19) and (3.20) below when the Gauss images approach to the boundary of the open geodesic convex sets.

THEOREM 3.14 ([X-Y1]) *Let M be a complete minimal n-dimensional submanifold in $\mathbb{R}^{n+m}$. If the Gauss image of M is contained in an open* *geodesic ball of* $\mathbf{G}_{n,m}$ *centered at* P_0 *and of radius* $\frac{\sqrt{2}}{4}\pi$, *and* $\big(\frac{\sqrt{2}}{4}\pi - \rho \circ \gamma\big)^{-1}$ *has growth*

$$\Big(\frac{\sqrt{2}}{4}\pi - \rho \circ \gamma\Big)^{-1} = o(R), \tag{3.19}$$

where ρ *denotes the distance on* $\mathbf{G}_{n,m}$ *from* P_0 *and* R *is the Euclidean distance from any point in* M. *Then* M *has to be an affine linear subspace.*

THEOREM 3.15 ([X-Y2]) *Let* $M = (x, f(x))$ *be an* n*-dimensional entire minimal graph given by* m *functions* $f^\alpha(x^1, \cdots, x^n)$ *with* $m \geq 2$. *If*

$$\Delta_f = \left[det \left(\delta_{ij} + \sum_\alpha \frac{\partial f^\alpha}{\partial x^i} \frac{\partial f^\alpha}{\partial x^j} \right) \right]^{\frac{1}{2}} < 2,$$

and

$$(2 - \Delta_f)^{-1} = o(R^{\frac{4}{3}}), \tag{3.20}$$

where $R^2 = |x|^2 + |f|^2$. *Then* f^α *has to be affine linear functions and hence* M *has to be an affine linear subspace.*

REFERENCES

[FC] D. Fischer-Colbrie, *Some rigidity theorems for minimal submanifolds of the sphere*, Acta math. **145** (1980), 29-46.

[C-doC-K] S. S. Chern, M. do Carmo and S. Kobayashi, *Minimal submanifolds of a sphere with second fundamental form of constant length*, Functional analysis and related topics ed. F. Brower, Springer-Verlag (1970), 59-75.

[H-J-W] S. Hildebrandt, J. Jost, J and K. O. Widman, *Harmonic mappings and minimal submanifolds*, Invent.math. **62** (1980), 269-298.

[J-X] J. Jost and Y. L. Xin, *Bernstein type theorems for higher codimension*, Calc. Var. PDE **9** (1999), 277-296.

[L] H.B.Lawson, *Local rigidity theorems for minimal hypersurfaces*, Ann. Math. **89(2)** (1969), 187-197.

[M] J. Moser, *On Harnack's theorem for elliptic differential equations*, Comm. Pure Appl. Math. **14** (1961), 577-591.

[R-V] E. A. Ruh and J. Vilms, *The tension field of the Gauss map*, Trans. A. M. S. **149** (1970), 569-573.

[Si] J. Simons, *Minimal varieties in Riemannian manifolds*, Ann. Math. **88** (1968), 62-105.

[Spi] M. Spivak, *A comprehensive introduction to differential geometry*, Publish or Perish Inc, 1979.

[W] Yung-Chow Wong, *Differential geometry of Grassmann manifolds*, Proc. N.A.S. **57** (1967), 589-594.

[X1] Y.L. Xin, *Curvature estimates for submanifolds with prescribed Gauss image and mean curvature*, Calc. Var. and PDE (to appear).

[X2] Yuanlong Xin, *Minimal submanifolds and related topics*, World Scientific, Singapore, 2003.

[X3] Yuanlong Xin, *Geometry of harmonic maps*, Birkhäuser PNLDE 23, 1996.

[X-Y1] Y. L. Xin and Ling Yang, *Curvature estimates for minimal submanifolds of higher codimension*, Chin. Ann. Math. **30B(4)** (2009), 379-396.

[X-Y2] Y. L. Xin and Ling Yang, *Convex functions on Grassmannian manifolds and Lawson-Osserman problem*, Adv. Math. **219** (2008), 1298-1326.

[E-H] K. Ecker andG. Huisken, *A Bernstein result for minimal graphs of controlled growth*, J.Differ. Geom. **31**, (1990) 397–400.

[S-S-Y] R. Schoen, L. Simon and S. T. Yau,*Curvature Estimate for Minimal Hypersurfaces*, Acta Math. **134** (1975) 275–288.

(added by the editor)

SOME FORMULAE ON ADDITIVE FUNCTIONALS OF SYMMETRIC MARKOV PROCESSES

Masayoshi TAKEDA

1 Introduction

A. Beurling and J. Deny [4], [5] initiated the theory of Dirichlet forms. Using potential theory of Dirichlet forms, M. Fukushima [26] succeeded in the construction of symmetric Hunt processes associated with Dirichlet forms. Since then, the theory of Dirichlet forms has been developed by many persons as a useful tool for studying symmetric Markov processes (see e.g. [8], [28], [42], [54]). The theory of Dirichlet forms is an L^2-theory, and which is a reason why the theory is suitable for treating singular Markov processes. On the other hand, the theory of Markov processes is, in a sense, an L^1-theory. To bridge this gap, we have studied the L^p-independence of growth bounds of Markov semigroups, more generally, of generalized Feynman-Kac (Schrödinger) semigroups ([60], [63], [69]). The L^p-independence enables us to control L^∞-properties of the symmetric Markov process; in fact, we can state, in terms of the bottom of L^2-spectrum, a necessary and sufficient condition for the integrability of Feynman-Kac functionals ([61]) and for the stability of Gaussian both side estimates of Schrödinger heat kernels ([62]). For the proof of the L^p-independence, we apply arguments in the Donsker-Varadhan large deviation theory ([22], [23]). In particular, the identification of the rate function with its Dirichlet form is crucial. The main objective of this note is to derive asymptotic properties of additive functionals by applying the L^p-independence to time-changed processes.

The theory of random time-changes of Markov processes is an fundamental tool for studying positive continuous additive functionals (PCAF's). We realize in [15] that the random time-change theory is still finding wide application in boundary theory of Markov processes. If a Markov process is symmetric, then the time-changed process is also symmetric and its generating Dirichlet form is completely identified ([15], [28]). By employing the time-change theory in Dirichlet forms we can give simple proofs of following two formulas of M. Kac and extend them.

The first formula is in [36]: let $(B_t, \mathbb{P}_x^W)$ be the Brownian motion on the 3-dimensional Euclidean space $\mathbb{R}^3$. Let K be a compact set of $\mathbb{R}^3$ with smooth boundary (so-called *Kac regularity*). Then

$$\lim_{\beta\to\infty}\frac{1}{\beta}\log \mathbb{P}_x^W\left(\int_0^\infty 1_K(B_t)dt > \beta\right) = -\frac{1}{\lambda_2}. \tag{1.1}$$

Here $1/\lambda_2$ is the maximum eigenvalue of the operator G defined by

$$Gf(x) = \frac{1}{\pi}\int_K \frac{f(y)}{|x-y|}dy, \quad f \in L^2(K;dx). \tag{1.2}$$

This formula is regarded as one on the lifetime of a time-changed process. More precisely, let $(Y_t, \check{\mathbb{P}}_x)$ be the time-changed process by the PCAF $\int_0^t 1_K(B_s)ds$. Noting that the lifetime $\check{\zeta}$ of Y_t equals $\int_0^\infty 1_K(B_t)dt$, we see that the probability in the equation (1.1) equals $\check{\mathbb{P}}_x(\check{\zeta} > \beta)$. The operator G is the Green operator of the time-changed process Y_t and λ_2 the L^2-principal eigenvalue of the generator of Y_t. Hence the asymptotic of $\check{\mathbb{P}}_x(\check{\zeta} > \beta)$ is controlled by an L^2-quantity λ_2, and thus we can think that the L^p-independence for the time-changed process is behind the first formula; indeed, we extend, in Section 4, the first formula by employing the L^p-independence for the time-changed process.

The second formula is in [37]: for a positive function V in $L^1(\mathbb{R}^3)$, define

$$\Gamma(V) = \lim_{t\to\infty}\frac{1}{t}\int_{\mathbb{R}^3}\left(1 - \mathbb{E}_x^W\left(e^{-\int_0^t V(B_s)ds}\right)\right)dx. \tag{1.3}$$

This is a probabilistic representation of the *scattering length* (M. Kac [37]). He proved that for a compact set K of $\mathbb{R}^3$ with smooth boundary,

$$\Gamma(\alpha 1_K) \longrightarrow \mathrm{Cap}(K) \quad \text{as } \alpha \to \infty, \tag{1.4}$$

where Cap is the Newtonian capacity. This formula is also regarded as one on the lifetime $\check{\zeta}$ of the time-changed process Y_t above. In fact, we can show that

$$\Gamma(\alpha 1_K) = \alpha\int_K \mathbb{E}_x^W\left(\exp\left(-\alpha\int_0^\infty 1_K(B_t)dt\right)\right)dx,$$

and the right hand side is written by $\alpha\int_K \check{\mathbb{E}}_x(\exp(-\alpha\check{\zeta}))dx$. As an extension of the second formula (1.4), M. Kac conjectured in [37] that for any positive L^1-function V with compact support, the limit

$$\gamma_V := \lim_{\alpha\to\infty}\Gamma(\alpha V)$$

equals the capacity of the support of V. M. Taylor [75] probabilistically verified the conjecture, and H. Tamura [74] proved it analytically. Y. Takahashi [57] gave a new probabilistic representation of $\Gamma(V)$ for more general symmetric Markov processes and proved that if V is a positive continuous function with compact support, then the limit γ_V exists and depends only on the set $\{x : V(x) > 0\}$. In Section 6, we extend their results and give another simple proof of the conjecture of Kac by using the time-change argument in the Dirichlet space theory.

In other sections, we will treat the topics relevant to the L^p-independence and the random time-change. In Section 5, we treat the gaugeability, i.e. the integrability of Feynman-Kac functionals, as an application of the L^p-independence. As applications of the gaugeability, we consider the penalization problem in Section 7, 8 and 9 and the stability of heat kernels in Section 10. In the large deviation principle and the Feynman-Kac penalization, the ergodic property of symmetric Markov processes plays an important role. In Section 11, we summarize the ergodic theory in Dirichlet spaces.

2 Donsker-Varadhan type large deviation principle

The Donsker-Varadhan large deviation theory for the occupation time distributions of Markov processes is considerably tractable in symmetric situations. M. Donsker and S.R.S. Varadhan introduced the so-called *I-function* as the rate function in their large deviation principle. While the evaluation of the I-function is generally hard, it becomes easier for symmetric Markov processes; the I-function has been identified with the Dirichlet form (Donsker-Varadhan [22]). Moreover, we can derive the Donsker-Varadhan type large deviation principle for a general, not necessarily conservative symmetric Markov processes by invoking an original idea in Donsker-Varadhan [21], where the one-dimensional Brownian motion was treated.

Let X be a locally compact separable metric space and m a positive Radon measure on X with full support. Let $\mathbf{M} = (\Omega, \mathcal{F}, \mathcal{F}_t, X_t, \mathbb{P}_x, \zeta)$ be an m-symmetric Markov process. Ω is specifically taken to be the space of all right continuous functions from $[0,\infty]$ into the one point comactification $X_\Delta = X \cup \{\Delta\}$ of X possessing the left limits such that $\omega(t) = \Delta$ for any $t \geq \zeta(\omega) = \inf\{s \geq 0 : w(s) = \Delta\}$ and $\omega(\infty) = \Delta$. The random variable ζ is called the lifetime which can be finite and X_t is defined by $X_t(\omega) = \omega(t)$ for $\omega \in \Omega,\ t \geq 0$. $\{\mathcal{F}_t\}$ is the minimal (augmented) admissible filtration.

Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be the Dirichlet form on $L^2(X;m)$ generated by $\mathbf{M}$:

$$(2.1)\qquad \begin{cases} \mathcal{D}(\mathcal{E}) = \left\{ u \in L^2(X;m) : \lim_{t\to 0} \dfrac{1}{t}(u - p_t u, u)_m < \infty \right\} \\ \mathcal{E}(u,v) = \lim_{t\to 0} \dfrac{1}{t}(u - p_t u, v)_m. \end{cases}$$

We assume that the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is *regular*, that is, $\mathcal{D}(\mathcal{E}) \cap C_0(X)$ is dense in $\mathcal{D}(\mathcal{E})$ with respect to $\mathcal{E}_1$-norm and dense in $C_0(X)$ with respect to the uniform norm. Here

$C_0(X)$ is the space of continuous functions on X with compact support and $\mathcal{E}_1(u,u) = \mathcal{E}(u,u)+(u,u)_m$. We define the *extended Dirichlet space* $\mathcal{D}_e(\mathcal{E})$ by the family of measurable function u on X such that $|u| < \infty$ m-a.e. and there exists an $\mathcal{E}$-Cauchy sequence $\{u_n\}$ of functions in $\mathcal{D}(\mathcal{E})$ such that $\lim_{n\to\infty} u_n = u$ m-a.e.

We define the *(1-)capacity* Cap associated with the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ as follows: for any open set $O \subset X$,

$$\mathrm{Cap}(O) = \inf\{\mathcal{E}_1(u,u) : u \in \mathcal{D}(\mathcal{E}), u \geq 1,\ m\text{-a.e. on } O\}$$

and for any Borel set $A \subset X$,

$$\mathrm{Cap}(A) = \inf\{\mathrm{Cap}(O) : O \text{ is open, } O \supset A\}.$$

We define the 0-order capacity $\mathrm{Cap}_{(0)}$ by replacing $\mathcal{E}_1$ and $\mathcal{D}(\mathcal{E})$ with $\mathcal{E}$ and $\mathcal{D}_e(\mathcal{E})$. Let A be a subset of X. A statement depending on $x \in A$ is said to hold q.e. on A if there exists a set $N \subset A$ of zero capacity such that the statement is true for every $x \in A \setminus N$. "q.e." is an abbreviation of "quasi-everywhere". A real valued function u defined q.e. on X is said to be *quasi continuous* if for any $\epsilon > 0$ there exists an open set $G \subset X$ such that $\mathrm{Cap}(G) < \epsilon$ and $u|_{X\setminus G}$ is finite and continuous. Here, $u|_{X\setminus G}$ denotes the restriction of u to $X \setminus G$. Each function u in $\mathcal{D}(\mathcal{E})$ admits a quasi-continuous version $\tilde{u}$, that is, $u = \tilde{u}$ m-a.e. In the sequel, we always assume that every function $u \in \mathcal{D}(\mathcal{E})$ is represented by its quasi-continuous version.

We denote by $\{p_t\}_{t\geq 0}$ and $\{R_\alpha\}_{\alpha>0}$ the semigroup and the resolvent of $\mathbf{M}$, $p_t f(x) = \mathbb{E}_x(f(X_t))$ and $R_\alpha f(x) = \mathbb{E}_x(\int_0^\infty e^{-\alpha t} f(X_t)dt)$. We now make following assumptions:

I. (**Irreducibility**) If a Borel set A is p_t-invariant, i.e., $p_t(1_A f)(x) = 1_A p_t f(x)$ m-a.e. for any $f \in L^2(X;m) \cap \mathcal{B}_b(X)$ and $t > 0$, then A satisfies either $m(A) = 0$ or $m(X \setminus A) = 0$. Here $\mathcal{B}_b(X)$ is the space of bounded Borel functions on X.

II. (**Strong Feller Property**) $R_1(\mathcal{B}_b(X)) \subset C_b(X)$, where $C_b(X)$ is the space of bounded continuous functions.

III. (**Tightness Property**) For any $\epsilon > 0$, there exists a compact set K such that $\sup_{x\in X} R_1 1_{K^c}(x) \leq \epsilon$. Here 1_{K^c} is the indicator function of the complement of K.

Remark 2.1. (i) It follows from the assumption II that the resolvent kernel $R_1(x, dy)$ is absolutely continuous with respect to m and so is the transitions probability $p_t(x, dy)$.
(ii) If $m(X) < \infty$ and $\|R_1\|_{1,\infty} < \infty$, then $\|R_1 1_{K^c}\|_\infty \leq \|R_1\|_{1,\infty} m(K^c)$ and the assumption III is fulfilled. Here $\|R_1\|_{1,\infty}$ is the operator norm from $L^1(X;m)$ to $L^\infty(X;m)$.
(iii) If $R_1 1 \in C_\infty(X)$, then the assumption III is fulfilled. Here $C_\infty(X)$ is the set of continuous functions vanishing at infinity. If $C_\infty(X)$ is invariant under R_1, $R_1(C_\infty(X)) \subset C_\infty(X)$,

then $R_1 1 \in C_\infty(X)$ is equivalent to the assumption III. If the Markov process $\mathbf{M}$ is conservative, then the assumption III is equivalent to that for any $\epsilon > 0$, there exists a compact set K such that $\inf_{x\in X} R_1 1_K(x) \geq 1-\epsilon$, which implies a strong ergodicity.
(iv) Let $\mathcal{P}$ be the set of probability measures on X equipped with the weak topology. We define the subset $\mathcal{P}_M$ of $\mathcal{P}$ by

$$\mathcal{P}_M = \left\{ u^2 \cdot m : u \in \mathcal{D}(\mathcal{E}),\ \int_X u^2 dm = 1,\ \mathcal{E}(u,u) \leq M \right\}, \quad M > 0.$$

By combining the assumption III with the inequality (5.14) below, we see that $\mathcal{P}_M$ is tight; indeed, for any compact set $K \subset X$ and any $u^2 \cdot m \in \mathcal{P}_M$

$$\int_{K^c} u^2 dm \leq \|R_1 1_{K^c}\|_\infty \cdot \left(\mathcal{E}(u,u) + \int_X u^2 dm \right) \leq (M+1)\|R_1 1_{K^c}\|_\infty. \tag{2.2}$$

([55]).

We define the function $I_\mathcal{E}$ on $\mathcal{P}$ by

$$I_\mathcal{E}(\mu) = \begin{cases} \mathcal{E}(\sqrt{f},\sqrt{f}) & \text{if } \mu = f\cdot m,\ \sqrt{f} \in \mathcal{D}(\mathcal{E}) \\ \infty & \text{otherwise.} \end{cases}$$

For $\omega \in \Omega$ with $\zeta(\omega) > t$, we define the normalized occupation time distribution $L_t(\omega) \in \mathcal{P}$ by

$$L_t(\omega)(A) = \frac{1}{t}\int_0^t 1_A(X_s(\omega))ds, \quad A \in \mathcal{B}(X).$$

We then have a version of Donsker-Varadhan type large deviation principle.

Theorem 2.2. (i) ([58]) *For any open set G of $\mathcal{P}$*

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\nu(L_t \in G, t < \zeta) \geq -\inf_{\mu\in G} I_\mathcal{E}(\mu) \quad \textit{for all } \nu \in \mathcal{P}.$$

(ii) *For any closed set K of $\mathcal{P}$*

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{x\in X} \mathbb{P}_x(L_t \in K, t < \zeta) \leq -\inf_{\mu\in K} I_\mathcal{E}(\mu).$$

We would like to make comments on the theorem above. Denote by A the generator of the Markov process $\mathbf{M}$ and set

$$\mathcal{D}^+(A) = \left\{ R_\alpha f :\ \alpha > 0,\ f \in L^2(X;m) \cap C_b^+(X) \text{ and } f \not\equiv 0 \right\},$$

where $C_b^+(X)$ denotes the set of non-negative bounded continuous functions. We see that any function in $\mathcal{D}^+(A)$ is strictly positive by the assumption I. We define the multiplicative functional L_t^ϕ, $\phi = R_\alpha f \in \mathcal{D}^+(A)$, by

$$L_t^\phi = \frac{\phi(X_t)}{\phi(X_0)} \exp\left(-\int_0^t \frac{A\phi}{\phi}(X_s)ds \right) 1_{\{t<\zeta\}}, \quad A\phi = \alpha R_\alpha f - f. \tag{2.3}$$

Let $\mathbf{M}^\phi = (\Omega, X_t, \mathbb{P}^\phi_x, \zeta)$ be the transformed process of $\mathbf{M}$ by L^ϕ_t:

$$\mathbb{P}^\phi_x, (F; t < \zeta) = \mathbb{E}_x(L^\phi_t 1_F; t < \zeta), \quad F \in \mathcal{F}_t.$$

For the proof of Theorem 2.2 (i), the symmetry of the Markov process $\mathbf{M}$ is crucial; indeed, even if $\mathbf{M}$ is explosive or of killing inside, $\mathbf{M}^\phi$ turns out to be an ergodic Markov process with invariant measure $\phi^2 m$. For the sketch of the proof of this fact, see Section 11. If K is a compact subset of $\mathcal{P}$, then Theorem 2.2 (ii) holds without the assumption III. The assumption III is crucial to strengthen the statement to any closed set. Moreover, the assumption III is crucial to show the uniform upper bound in initial points. In fact, the uniform upper bound is not valid for the Ornstein-Uhlenbeck process, while the locally uniform upper bound is valid ([23],[78]). Note that the Ornstein-Uhlenbeck process does not satisfy the assumption III, because $\lim_{x\to\infty} R_1 1_{K^c}(x) = 1$. As stated in Remark 2.1 (iii), the assumption III implies a strong ergodicity for conservative Markov processes.

For one-dimensional diffusion processes, see Example 2.2 below.

Let us define the function I on $\mathcal{P}$ by

$$I(\mu) = -\inf_{\substack{u\in\mathcal{D}^+(A)\\ \epsilon>0}} \int_X \frac{Au}{u+\epsilon} d\mu. \tag{2.4}$$

The function I is a version of the so-called Donsker-Varadhan I-function introduced in [22]. Note that $u = R_\alpha f \in \mathcal{D}^+(A)$ is not alway a function uniformly lower-bounded by a positive constant even if f is so, because the Markov process $\mathbf{M}$ is allowed to be explosive. Thus we add positive constant ϵ to make the function $Au/(u+\epsilon)$ a bounded continuous function. This is necessary for the proof of Theorem 2.2 (ii), because the weak topology is embedded in the space $\mathcal{P}$. In fact, we first prove the upper bound with respect to I and then identify I with $I_\mathcal{E}$:

Proposition 2.3.

$$I(\mu) = I_\mathcal{E}(\mu), \qquad \mu \in \mathcal{P}.$$

To treat symmetric Markov processes with general state space, Jain and Krylov [35] introduce another modification of I-function. We see from Proposition 2.3 that the function $I_\mathcal{E}$ is lower semi-continuous with respect to the weak topology on $\mathcal{P}$.

Corollary 2.4. (**Extended variational formula for Dirichlet forms**) *For* $f \in \mathcal{D}(\mathcal{E})$

$$\mathcal{E}(f,f) = \sup_{\substack{u\in\mathcal{D}^+(A)\\ \epsilon>0}} \int_X \frac{-Au}{u+\epsilon} f^2 dm. \tag{2.5}$$

Let λ_2 be the bottom of spectrum:

$$\lambda_2 = \inf \left\{ \mathcal{E}(f,f) : f \in \mathcal{D}(\mathcal{E}),\ \int_X f^2 dm = 1 \right\}. \tag{2.6}$$

Using Corollary 2.4, we have for any $u \in \mathcal{D}^+(A)$ and $\epsilon > 0$,

$$\lambda_2 \geq \inf_{x \in X} \frac{-Au}{u + \epsilon}(x). \tag{2.7}$$

For a generalization of Corollary 2.4, see [53]. For an extensive application of the variational formula for Dirichlet forms, see [12].

S.R.S. Varadhan [77] gave an abstract formulation for the large deviation principle. Theorem 2.2 is slightly different from the lower estimate and the upper estimate in his formulation; since the Markov process is not supposed to be conservative, we can not regard Theorem 2.2 as the large deviation principle from the invariant measure. By this reason, we consider the normalized probability measure $\widetilde{\mathbb{Q}}_{x,t}$ on $\mathcal{P}$ defined by

$$\widetilde{\mathbb{Q}}_{x,t}(B) = \frac{\mathbb{P}_x(L_t \in B, t < \zeta)}{\mathbb{P}_x(t < \zeta)}, \quad B \in \mathcal{B}(\mathcal{P}) \tag{2.8}$$

The family of probability measures $\{\widetilde{\mathbb{Q}}_{x,t}\}_{t>0}$ then satisfies the large deviation principle with the rate function $J(\nu) := I_{\mathcal{E}}(\nu) - \lambda_2,\ \nu \in \mathcal{P}$, as $t \to \infty$ in Varadhan's formulation, where λ_2 is the bottom of the spectrum of the L^2-generator A for $\mathcal{E}$ defined by (2.6). In other words, $\{\widetilde{\mathbb{Q}}_{x,t}\}_{t>0}$ obeys the *full* large deviation principle with the *good* rate function $J(\nu)$. In addition, we shall see that the ground state ϕ_0 of the operator A exists and $\phi_0^2 \cdot m$ is a unique probability measure for which $J(\nu) = 0$. On account of these facts, we shall reinterpret Theorem 2.2 as a large deviation principle from the ground state.

A function ϕ_0 on X is called a *ground state* of the L^2-generator A for $\mathcal{E}$ if $\phi_0 \in \mathcal{F},\ \|\phi_0\|_2 = 1$ and $\mathcal{E}(\phi_0, \phi_0) = \lambda_2$.

Lemma 2.5. ([70]) *Assume that* $\mathbf{M}$ *satisfies I∼III. Then there exists a ground state* ϕ_0 *of* A *uniquely up to a sign.* ϕ_0 *can be taken to be strictly positive on* X.

Proof. In our proof of the existence of the minimizer in the right hand side of (2.6), the identification of the I-function with the Dirichlet form (Proposition 2.3) plays a crucial role. In fact, let $\{u_n\}_{n=1}^{\infty} \subset \mathcal{D}(\mathcal{E})$ be a minimizing sequence, that is, $\|u_n\|_2 = 1$ and $\lambda_2 = \lim_{n\to\infty} \mathcal{E}(u_n, u_n)$.

We see from (2.2) that for any $\epsilon > 0$ there exists a compact set K such that

$$\sup_n \int_{K^c} u_n^2 \cdot dm \leq \|R_1 I_{K^c}\|_\infty \cdot \left(\sup_n \mathcal{E}(u_n, u_n) + 1 \right) < \epsilon,$$

that is, the subset $\{u_n^2 \cdot m\}$ of $\mathcal{P}$ is tight. Hence there exists a subsequence $\{u_{n_k}^2 \cdot m\}$ such that $u_{n_k}^2 m$ converges weakly to a probability measure ν. It follows from Proposition 2.3 that the function $I_{\mathcal{E}}$ is lower semi-continuous with respect to the weak topology,

$$I_{\mathcal{E}}(\nu) \leq \liminf_{k\to\infty} I_{\mathcal{E}}(u_{n_k}^2 \cdot m) = \liminf_{k\to\infty} \mathcal{E}(u_{n_k}, u_{n_k}) < \infty.$$

Therefore we see that ν is expressed as $\nu = \phi_0^2 \cdot m,\ \phi_0 \in \mathcal{D}(\mathcal{E}),\ \phi_0 \geq 0$. ϕ_0 is just a ground state of A.

It follows from the inequality $\|\phi_0 + \epsilon g\|_{\mathcal{E}}^2 \geq \lambda_2 \|\phi_0 + \epsilon g\|_2^2$ holding for any $g \in \mathcal{D}(\mathcal{E})$ and any $\epsilon > 0$ that $\mathcal{E}(\phi_0, g) = \lambda_2(\phi_0, g)$. Hence $\alpha R_{\alpha-\lambda_2}\phi_0 = \phi_0,\ \alpha > \lambda_2$, which implies that ϕ_0 is strictly positive by the irreducibility.

To prove the uniqueness of the ground state, we introduce a closed symmetric form $(\mathcal{E}^{\phi_0}, \mathcal{D}(\mathcal{E}^{\phi_0}))$ on $L^2(X; \phi_0^2 m)$ by

$$\begin{cases} \mathcal{E}^{\phi_0}(u,v) = \mathcal{E}(u\phi_0, v\phi_0) - \lambda_2(u\phi_0, v\phi_0) \\ \mathcal{D}(\mathcal{E}^{\phi_0}) = \{u \in L^2(X; \phi_0^2 \cdot m) : u\phi_0 \in \mathcal{D}(\mathcal{E})\}. \end{cases} \tag{2.9}$$

Since $1 \in \mathcal{D}(\mathcal{E}^{\phi_0})$, $\mathcal{E}^{\phi_0}(1,1) = 0$ and the associated resolvent $R_\alpha^{\phi_0}$ satisfies $R_\alpha^{\phi_0} f = \phi_0^{-1} R_{\alpha-\lambda_2}(f\phi_0),\ \alpha > \lambda_2$, we see from Lemma 11.4 that $(\mathcal{E}^{\phi_0}, \mathcal{D}(\mathcal{E}^{\phi_0}))$ is an irreducible recurrent Dirichlet form so that f is constant whenever $f \in \mathcal{D}(\mathcal{E}^{\phi_0}),\ \mathcal{E}^{\phi_0}(f,f) = 0$. Let ψ_0 be another ground state of A. Then $\psi_0 = f\phi_0$ with $f = \psi_0/\phi_0 \in \mathcal{D}(\mathcal{E}^{\phi_0}),\ \mathcal{E}^{\phi_0}(f,f) = \mathcal{E}(\psi_0, \psi_0) - \lambda_2 = 0$, which yields that f is constant and $\psi_0 = \pm\phi_0$. □

In a usual proof of the existence of the minimizer, the $\mathcal{E}_1$-weak compactness of $\{u_n\}_{n=1}^\infty$ in $\mathcal{D}(\mathcal{E})$ and the lower semi-continuity of $\mathcal{E}$ are used (e.g. [40]). We would like to emphasis that we use the tightness of $\{u_n^2 \cdot m\}_{n=1}^\infty \subset \mathcal{P}$ and the lower semi-continuity of the function $I_{\mathcal{E}}$ with respect to the weak topology. We see from the proof of Lemma 2.5 that the level set $\{\nu \in \mathcal{P} : I_{\mathcal{E}}(\nu) \leq \ell\}$ is a compact subset of $\mathcal{P}$. Hence we have the next lemma.

Lemma 2.6. *The function J satisfies*:
(i) $0 \leq J(\nu) \leq \infty$.
(ii) *J is lower semi-continuous.*
(iii) *For each $l < \infty$, the set $\{\nu \in \mathcal{P} : J(\nu) \leq l\}$ is compact.*
(iv) *$J(\phi_0^2 \cdot m) = 0$ and $J(\nu) > 0$ for $\nu \neq \phi_0^2 \cdot m$.*

Lemma 2.6 states that the function $J(\nu) = I_{\mathcal{E}}(\nu) - \lambda_2,\ \nu \in \mathcal{P}$, enjoys the properties as a good rate function in the large deviation principle. We note that the identity

$$J(\nu) = I_{\mathcal{E}^{\phi_0}}(\nu), \quad \nu \in \mathcal{P}, \tag{2.10}$$

holds true, where $I_{\mathcal{E}^{\phi_0}}$ is defined in terms of the Dirichlet form (2.9) by

$$I_{\mathcal{E}^{\phi_0}}(\nu) = \begin{cases} \mathcal{E}^{\phi_0}(\sqrt{f}, \sqrt{f}) & \text{if } \nu = f\phi_0^2 \cdot m,\ \sqrt{f} \in \mathcal{D}(\mathcal{E}^{\phi_0}) \\ \infty & \text{otherwise.} \end{cases} \tag{2.11}$$

By Theorem 2.2 we obtain the next large deviation principle:

Theorem 2.7. ([70]) *Let* $\{\widetilde{\mathbb{Q}}_{x,t}\}_{t>0}$ *be a family of probability measures defined by* (2.8). *Then the sequence* $\{\widetilde{\mathbb{Q}}_{x,t}\}_{t>0}$ *obeys the large deviation principle with rate function* J:
(i) *For each open set* $G \subset \mathcal{P}$

$$\liminf_{t\to\infty} \frac{1}{t} \log \widetilde{\mathbb{Q}}_{x,t}(G) \geq - \inf_{\nu \in G} J(\nu).$$

(ii) *For each closed set* $K \subset \mathcal{P}$

$$\limsup_{t\to\infty} \frac{1}{t} \log \widetilde{\mathbb{Q}}_{x,t}(K) \leq - \inf_{\nu \in K} J(\nu).$$

Corollary 2.8. *The measure* $\widetilde{\mathbb{Q}}_{x,t}$ *converges weakly to* $\delta_{\phi_0^2\cdot m}$ *as* $t \to \infty$.

Proof. If a closed set K does not contain $\phi_0^2 \cdot m$, then $\inf_{x\in K} J(x) > 0$ by Lemma 2.6 (iv). Hence Theorem 2.7 (ii) says that $\lim_{t\to\infty} \widetilde{\mathbb{Q}}_{x,t}(K) = 0$ and $\lim_{t\to\infty} \widetilde{\mathbb{Q}}_{x,t}(K^c) = 1$. For a positive constant δ and a bounded continuous function f on $\mathcal{P}$, define the closed set $K \subset \mathcal{P}$ by $K = \{\nu \in \mathcal{P} : |f(\nu) - f(\phi_0^2 \cdot m)| \geq \delta\}$. Then we have

$$\begin{aligned}
&\left| \int_{\mathcal{P}} f(\nu)\widetilde{\mathbb{Q}}_{x,t}(d\nu) - f(\phi_0^2 \cdot m) \right| \leq \int_{\mathcal{P}} |f(\nu) - f(\phi_0^2 \cdot m)|\widetilde{\mathbb{Q}}_{x,t}(d\nu) \\
&= \int_K |f(\nu) - f(\phi_0^2 \cdot m)|\widetilde{\mathbb{Q}}_{x,t}(d\nu) + \int_{K^c} |f(\nu) - f(\phi_0^2 \cdot m)|\widetilde{\mathbb{Q}}_{x,t}(d\nu) \\
&\leq 2\|f\|_\infty \widetilde{\mathbb{Q}}_{x,t}(K) + \delta\widetilde{\mathbb{Q}}_{x,t}(K^c) \longrightarrow \delta
\end{aligned}$$

as $t \to \infty$. Since δ is arbitrary, the proof of the corollary is complete. □

On account of Corollary 2.8, we can regard Theorem 2.7 as a genuine large deviation principle from the ground state.

Setting $G = K = \mathcal{P}$ in Theorem 2.2, we have

Corollary 2.9.

$$\begin{aligned}
\lim_{t\to\infty} \frac{1}{t} \log \sup_{x\in X} \mathbb{P}_x(t < \zeta) &= \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_x(t < \zeta) \\
&= -\inf \left\{ \mathcal{E}(u,u) : u \in \mathcal{D}(\mathcal{E}),\ \int_X u^2 dm = 1 \right\}.
\end{aligned} \tag{2.12}$$

Let us denote by $\|p_t\|_{p,p}$ the operator norm of p_t from $L^p(X;m)$ to $L^p(X;m)$ and put

$$-\lambda_p = \lim_{t\to\infty} \frac{1}{t} \log \|p_t\|_{p,p}, \quad 1 \leq p \leq \infty.$$

$-\lambda_p$ is the long time exponential growth bound of the semigroup $\{p_t\}$. Note that $\sup_{x\in X}\mathbb{P}_x(t<\zeta)=\|p_t\|_{\infty,\infty}$ and the right hand side of (2.12) is equal to $-\lambda_2$ by the spectral theorem. We then see from Corollary 2.9 that

$$\lambda_\infty=\lambda_2. \tag{2.13}$$

It follows from the symmetry of p_t that $\|p_t\|_{2,2}\le\|p_t\|_{\infty,\infty}$. Hence Riesz-Thorin interpolation theorem tells us that

$$\|p_t\|_{2,2}\le\|p_t\|_{p,p}\le\|p_t\|_{\infty,\infty},\quad 1\le p\le\infty.$$

Theorem 2.10. ([60]) *Under the assumptions I $\sim$ III, λ_p $(1\le p\le\infty)$ is independent of p.*

Example 2.1. Let us consider the symmetric bilinear form

$$\mathcal{E}(u,v)=\frac{1}{2}\sum_{i,j=1}^{d}\int_{\mathbb{R}^d}a_{ij}(x)\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}dx,\qquad u,v\in C_0^\infty(\mathbb{R}^d),$$

where $(a_{ij}(x))$ is a symmetric matrix satisfying

$$\lambda(2+|x|)^2\log(2+|x|)^\beta|\xi|^2\le\sum_{i,j=1}^{d}a_{ij}(x)\xi_i\xi_j\le\Lambda(2+|x|)^2\log(2+|x|)^\beta|\xi|^2$$

for some positive constant λ,Λ. Let $\mathcal{D}(\mathcal{E})$ be the closure of $C_0^\infty(\mathbb{R}^d)$. Then, $(\mathcal{E},\mathcal{D}(\mathcal{E}))$ becomes a strongly local Dirichlet form on $L^2(\mathbb{R}^d)$. Denote by $\mathbf{M}=(\Omega,X_t,\mathbb{P}_x,\zeta)$ the associated diffusion process on $\mathbb{R}^d$.

Let us define a metric ρ (so-called *intrinsic metric*) on $\mathbb{R}^d$ as follows.

$$\rho(x,y)=\sup\left\{u(x)-u(y):u\in\mathcal{D}_{loc}(\mathcal{E})\cap C(\mathbb{R}^d),\ \sum_{i,j=1}^{d}\int_{\mathbb{R}^d}a_{ij}(x)\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}\le 1\text{ a.e.}\right\}. \tag{2.14}$$

Then, we can show that if $\beta\le 2$, $(\mathbb{R}^d,\rho)$ is a complete metric space and the induced topology is equivalent to the usual one. Let $B_\rho(r)=\{x\in\mathbb{R}^d;\rho(0,x)<r\}$. Then

$$m(B_\rho(r))\simeq\begin{cases}e^{e^r} & \beta=2\\ e^{r^{\frac{2}{2-\beta}}} & \beta<2,\end{cases} \tag{2.15}$$

where $g(r)\simeq f(r)$ means $0<\liminf_{x\to\infty}\frac{f(x)}{g(x)}\le\limsup_{x\to\infty}\frac{f(x)}{g(x)}<\infty$. We see from Note 6.6 in [18] that if $a_{ij}(x)$ are smooth and $\beta>1$, the function $R_1 1$ belongs to $C_\infty(\mathbb{R}^d)$, and thus obtain

$$\lim_{t\to\infty}\frac{1}{t}\log\mathbb{P}_x(t<\zeta)=-\inf\left\{\mathcal{E}(u,u);u\in\mathcal{D}(\mathcal{E}),\int_{\mathbb{R}^d}u^2dm=1\right\}\quad\text{for all }x\in\mathbb{R}^d. \tag{2.16}$$

For $\beta \leq 1$, the diffusion process $\mathbf{M}$ is conservative. Thus, for $\beta < 0$, the both side of (2.16) are equal to zero. If $0 \leq \beta \leq 1$, then $\mathrm{Sp}(H) \subset [\frac{1}{8}\lambda d^2, \infty)$ (Theorem 1.5.14 in [19]). Here H is the self-adjoint operator associated to the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ and $\mathrm{Sp}(H)$ is the set of spectrum of H. Hence, the relation (2.16) does not hold for $0 \leq \beta \leq 1$ because the left hand side is equal to zero by the conservativeness of $\mathbf{M}$.

The L^p-independence of Markov semigroups can be extended to that of generalized Feynman-Kac semigroups by extending Theorem 2.2 to symmetric Markov processes with Feynman-Kac functional([20], [63], [68], [69], [76]). The L^p-independence of Feynman-Kac semigroups implies the existence of logarithmic moment generating functions of additive functionals. This is a prerequisite condition when we apply Gärtner–Ellis theorem to the proof of large deviations for additive functionals. See [71], [72].

The probabilistic interpretation of λ_∞ is known:

Theorem 2.11. ([51])

$$\sup_{x \in X} \mathbb{E}_x(\exp(\lambda\zeta)) < \infty \quad \textit{if and noly if} \quad \lambda < \lambda_\infty.$$

Therefore, we obtain

Corollary 2.12. *Assume I $\sim$ III. Then*

$$\sup_{x \in X} \mathbb{E}_x(\exp(\lambda\zeta)) < \infty \quad \textit{if and only if} \quad \lambda < \lambda_2.$$

Let $K \subset X$ be a compact set and D is the complement of K, $D := X \setminus K$. Let X^D be the part process on D:

$$X^D = \begin{cases} X_t & t < \tau_D \\ \Delta & t \geq \tau_D, \quad \tau_D = \inf\{t > 0 : X_t \notin D\}. \end{cases}$$

We suppose that X^D satisfies I$\sim$III for any K. Define the Dirichlet form $(\mathcal{E}^D, \mathcal{D}(\mathcal{E}^D))$ on D by

$$\begin{cases} \mathcal{E}^D = \mathcal{E} \\ \mathcal{D}(\mathcal{E}^D) = \{u \in \mathcal{D}(\mathcal{E}) : u = 0 \text{ q.e. } K\}. \end{cases} \tag{2.17}$$

Let λ^D be the principal eigenvalue of the spectrum of $(\mathcal{E}^D, \mathcal{D}(\mathcal{E}^D))$ and ϕ^D the ground state (Lemma 2.1). It follows from (2.2) that

$$1 = \int_D (\phi^D)^2 dm \leq \|R_1 1_D\|_\infty (\lambda^D + 1).$$

Since $\|R_1 1_D\|_\infty \to 0$ as $K \uparrow X$ by III, we see that under I $\sim$ III

$$\lambda_D \uparrow \infty \quad \text{as} \quad K \uparrow X. \tag{2.18}$$

If $\mathbf{M}$ is conservative, the lifetime of X^D equals the first hitting time σ_K of K and

$$\sup_{x\in D} \mathbb{E}_x(\exp(\lambda\sigma_K)) < \infty \Longleftrightarrow \lambda < \lambda_D.$$

Hence we can conclude that for any $\lambda > 0$ there exists a compact set K such that $\sup_{x\in X} \mathbb{E}_x(\exp(\lambda\sigma_K)) < \infty$. In other words, $\mathbf{M}$ have the *uniform hyper-exponential recurrence* in the sense of [78].

Example 2.2. Let us consider a one-dimensional diffusion process $\mathbf{M} = (X_t, P_x, \zeta)$ on an open interval $I = (r_1, r_2)$ such that $P_x(X_{\zeta-} = r_1 \text{ or } r_2,\ \zeta < \infty) = P_x(\zeta < \infty),\ x \in I$, and $P_a(\sigma_b < \infty) > 0$ for any $a, b \in I$. The diffusion $\mathbf{M}$ is symmetric with respect to its canonical measure m and it satisfies I and II. The boundary point r_i of I is classified into four classes: *regular boundary, exit boundary, entrance boundary and natural boundary* ([34, Chapter 5]):

(a) If r_2 is a regular or exit boundary, then $\lim_{x\to r_2} R_1 1(x) = 0$.

(b) If r_2 is an entrance boundary, then $\lim_{r\to r_2} \sup_{x\in(r_1,r_2)} R_1 1_{(r,r_2)}(x) = 0$.

(c) If r_2 is a natural boundary, then $\lim_{x\to r_2} R_1 1_{(r,r_2)}(x) = 1$ and thus $\sup_{x\in(r_1,r_2)} R_1 1_{(r,r_2)}(x) = 1$.

Therefore, III is satisfied if and only if no natural boundaries are present. As a corollary of (2.18), we see that if r_2 is an entrance boundary, then for any $\lambda > 0$ there exists $r_1 < r < r_2$ such that

$$\sup_{x>r} \mathbb{E}_x(\exp(\lambda\sigma_r)) < \infty.$$

Here σ_r is the first hitting time of $\{r\}$. The statement above implies the uniqueness of quasi-stationary distribution ([11]).

3 Random Time Change

In this section we treat time-changed processes of the Brownian motion by PCAF's associated with Kato measures. We will see that PCAF's associated with Kato measures are a suitable class in the theory of random time-change.

Let $\mathbf{D}$ be the classical Dirichlet form:

$$\mathbf{D}(u, v) = \int_{\mathbb{R}^d} \nabla u \cdot \nabla v dx \quad \text{for } u, v \in H^1(\mathbb{R}^d), \tag{3.1}$$

where $H^1(\mathbb{R}^d)$ is the Sobolev space of order 1. Let us denote by $(B_t, \mathbb{P}_x^W)$ the d-dimensional Brownian motion on $\mathbb{R}^d$.

A positive Radon measure μ on $\mathbb{R}^d$ is said to be in the *Kato class* K_d if

$$\lim_{\alpha\downarrow 0}\sup_{x\in\mathbb{R}^d}\int_{|x-y|<\alpha}\frac{\mu(dy)}{|x-y|^{d-2}}=0,\quad d\geq 3$$

$$\lim_{\alpha\downarrow 0}\sup_{x\in\mathbb{R}^d}\int_{|x-y|<\alpha}(\log|x-y|^{-1})\mu(dy)=0,\quad d=2$$

$$\sup_{x\in\mathbb{R}^d}\int_{|x-y|\leq 1}\mu(dy)<\infty,\quad d=1.$$

For $d\geq 3$ we introduce a subclass K_d^∞ of K_d by following [80]:

$$K_d^\infty=\left\{\mu\in K_d:\lim_{R\to\infty}\left[\sup_{x\in\mathbb{R}^d}\int_{|y|\geq R}\frac{d\mu(y)}{|x-y|^{d-2}}\right]=0\right\}.$$

For $\mu\in K_d$, there exists a unique (up to equivalence) positive continuous additive functional A_t^μ which is in *Revuz correspondence* with μ: for any γ-excessive function h ($\gamma\geq 0$) and any positive Borel function f

$$\lim_{t\to 0}\frac{1}{t}\int_{\mathbb{R}^d}\mathbb{E}_x\left(\int_0^t f(B_s)dA_s^\mu\right)h(x)dx=\int_{\mathbb{R}^d}f(x)h(x)d\mu \tag{3.2}$$

(Cf. [28, Theorem 5.1.3]). If μ is absolutely continuous with respect to the Lebesgue measure, say $\mu=V(x)dx$, then A_t^μ is nothing but $\int_0^t V(B_s)ds$.

Let A_t^μ be the continuous additive functional associated with $\mu\in K_d$. Let $\{\tau_t\}_{t\geq 0}$ be the right continuous inverse of A_t^μ:

$$\tau_t=\inf\{s>0:A_s^\mu>t\}.$$

The time-changed process Y_t^μ of B_t with respect to A_t^μ is defined by

$$Y_t^\mu=B_{\tau_t}.$$

Then, Y_t^μ is a μ-symmetric Markov process on a finely closed set $F=\{x\in\mathbb{R}^d:\mathbb{P}_x^W(\tau_0=0)=1\}$ with lifetime $\zeta=A_\infty^\mu$ (Theorem 6.2.1 in [28] and Theorem 65.9 in [52]). We assume that the set F equals the topological support of μ:

$$F=\operatorname{supp}[\mu]. \tag{3.3}$$

Set

$$H_F u(x)=\mathbb{E}_x^W(u(B_{\sigma_F});\sigma_F<\infty),$$

where $\sigma_F = \inf\{t > 0; B_t \in F\}$. Then, the Dirichlet form $(\check{\mathcal{E}}, \mathcal{D}(\check{\mathcal{E}}))$ on $L^2(F;\mu)$ generated by the time-changed process Y_t^μ is identified:

$$\left\{\begin{array}{l}\mathcal{D}(\check{\mathcal{E}}) = \{\varphi \in L^2(F;\mu) : \varphi = u \ \mu\text{-a.e. on } F \text{ for some } u \in H_e^1(\mathbb{R}^d)\} \\ \check{\mathcal{E}}(\varphi,\varphi) = \frac{1}{2}\mathbf{D}(H_F u, H_F u), \ \varphi \in \mathcal{D}(\check{\mathcal{E}}).\end{array}\right. \tag{3.4}$$

Here $H_e^1(\mathbb{R}^d)$ stands for the extended Dirichlet space of $(\frac{1}{2}\mathbf{D}, H^1(\mathbb{R}^d))$ (Theorem 6.2.1 in [28]).

From now on, we assume that $d \geq 3$ and denote by $R(x,y)$ the Green function of the Brownian motion. For $\mu \in K_d^\infty$,

$$\int_{\mathbb{R}^d} R(x,y)d\mu(y) \in C_\infty(\mathbb{R}^d). \tag{3.5}$$

Indeed, denote $R(x,y) \wedge n$ by $R^n(x,y)$. Then there exists a sequence of positive numbers α_n such that $R^n(x,y) = R(x,y)$ for $|x-y| \geq \alpha_n$, $\alpha_n \to 0$ as $n \to \infty$. On account of $\mu \in K_d$, we have

$$\begin{aligned} & \sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d} R(x,y)1_{\{|y|<R\}}d\mu(y) - \int_{\mathbb{R}^d} R^n(x,y)1_{\{|y|<R\}}d\mu(y)\right| \\ \leq \ & 2\sup_{x\in\mathbb{R}^d}\int_{|x-y|<\alpha_n} R(x,y)1_{\{|y|<R\}}d\mu(y) \to 0 \quad n \to \infty. \end{aligned}$$

Since $\int_{\mathbb{R}^d} R^n(x,y)1_{\{|y|<R\}}d\mu(y) \in C_\infty(\mathbb{R}^d)$, it holds that

$$\int_{\mathbb{R}^d} R(x,y)1_{\{|y|<R\}}d\mu(y) \in C_\infty(\mathbb{R}^d).$$

Hence, (3.5) follows from the definition of K_d^∞.

Let $\{R_\alpha^\mu(x,dy)\}_{\alpha\geq 0}$ be the resolvent kernel of Y_t^μ. Note that

$$\mathbb{E}_x^W\left(\int_0^\infty f(Y_t)dt\right) = \mathbb{E}_x^W\left(\int_0^\infty f(B_t)dA_t^\mu\right),$$

and thus

$$R_0^\mu(x,dy) = R(x,y)\mu(dy). \tag{3.6}$$

Since $R(x,y) > 0$ for any (x,y),

$$R_0^\mu 1_A(x) = \int_F R(x,y)1_A(y)d\mu(y) > 0$$

for any $A \in \mathcal{B}(F)$ with $\mu(A) > 0$, which implies the irreducibility of Y_t^μ. Hence we have

Lemma 3.1. *The time-changed process Y_t^μ satisfies $I \sim III$.*

Lemma 3.1 tells that the Kato class is a suitable class in the random time-change theory.

4 A Theorem of M. Kac on total occupation time

Proposition 4.1. If $\mu \in K_d^\infty$ satisfies (3.3), then

$$\lim_{\beta\to\infty}\frac{1}{\beta}\log\mathbb{P}_x^W(A_\infty^\mu > \beta) = -\inf\left\{\check{\mathcal{E}}(u,u) : u \in \mathcal{D}(\check{\mathcal{E}}),\ \int_F u^2 d\mu = 1\right\}. \tag{4.1}$$

Proof. As mentioned above, the time-changed process Y_t^μ satisfies Assumption I $\sim$ III. Since A_∞^μ is the lifetime of Y_t^μ, Corollary 2.9 tells us that the equation (4.1) holds for any $x \in F$. Since $A_{\sigma_F}^\mu = 0$ and thus

$$A_\infty^\mu = A_{\sigma_F}^\mu + A_\infty^\mu(\theta_{\sigma_F}) = A_\infty^\mu(\theta_{\sigma_F}),\quad \mathbb{P}_x^W\text{-a.s. on } \sigma_F < \infty,$$

we obtain

$$\begin{aligned}\mathbb{P}_x^W(A_\infty^\mu > \beta) &= \mathbb{P}_x^W(A_\infty^\mu > \beta; \sigma_F < \infty)\\ &= \mathbb{E}_x^W(\mathbb{P}_{X_{\sigma_F}}(A_\infty^\mu > \beta); \sigma_F < \infty)\\ &= \mathbb{P}_\nu^W(A_\infty^\mu > \beta),\end{aligned}$$

by the strong Markov property. Here, ν is the positive measure on F defined by $\nu(B) = \mathbb{P}_x(X_{\sigma_F} \in B; \sigma_F < \infty)$, $B \in \mathcal{B}(F)$. Therefore, (4.1) holds for any $x \in \mathbb{R}^d$. □

Lemma 4.2. *It holds that*

$$\begin{aligned}&\inf\left\{\check{\mathcal{E}}(u,u) : u \in \mathcal{D}(\check{\mathcal{E}}),\ \int_F u^2 d\mu = 1\right\}\\ =\ &\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : u \in C_0^\infty(\mathbb{R}^d),\ \int_{\mathbb{R}^d} u^2 d\mu = 1\right\}.\end{aligned} \tag{4.2}$$

Proof. On account of the regularity of $(\check{\mathcal{E}}, \mathcal{D}(\check{\mathcal{E}}))$ (Theorem 6.2.1 (iii) in [28]), the left hand side of (4.2) equals to

$$\inf\left\{\frac{1}{2}\mathbf{D}(H_F u, H_F u) : u \in C_0^\infty(\mathbb{R}^d),\ \int_{\mathbb{R}^d} u^2 d\mu = 1\right\},$$

and the above is equal to the right hand side of (4.2) because

$$\mathbf{D}(H_F u, H_F u) \le \mathbf{D}(u,u)$$

by the Dirichlet principle (Theorem 4.3.2 in [28]). □

For $\mu \in K_d$ with $\|R\mu\|_\infty < \infty$, denote by L^μ the generator of the time-changed process Y^μ and by λ_2^μ the bottom of the spectrum of L^μ. Since

$$\frac{L^\mu R\mu}{R\mu + \epsilon} = \frac{L^\mu R_0^\mu 1}{R\mu + \epsilon} = \frac{1}{R\mu + \epsilon},$$

we see from (2.7) that $\lambda_2^\mu \ge 1/\|R\mu\|_\infty$. Hence we have we have

Corollary 4.3. ([55]) *For $\mu \in K_d$,*

$$\int_{\mathbb{R}^d} u^2 d\mu \leq \frac{\|R\mu\|_\infty}{2} \mathbf{D}(u,u) \quad \text{for } u \in H^1(\mathbb{R}^d).$$

Combining Proposition 4.1 with Lemma 4.2, we now obtain

Theorem 4.4. ([59]) *It holds that for $\mu \in K_d^\infty$,*

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{P}_x^W(A_\infty^\mu > \beta) = -\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : u \in C_0^\infty(\mathbb{R}^d),\ \int_{\mathbb{R}^d} u^2 d\mu = 1\right\}.$$

Remark 4.5. Considering the absorbing Brownian motion, we can extend Theorem 4.4 as follows: for a Green bounded domain $D \subset \mathbb{R}^2 (d = 1, 2)$ and any domain ($d \geq 3$),

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{P}_x^W(A_{\tau_D}^\mu > \beta) = -\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : u \in C_0^\infty(D),\ \int_D u^2 d\mu = 1\right\}.$$

Here $\tau_D = \inf\{t > 0 : B_t \notin D\}$.

Example 4.1. Let $d = 3$. Let $\mu(dx) = 1_{B(0,1)}(x)dx$. Then, by Theorem 4.4

$$\begin{aligned} &\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{P}_x^W\left(\int_0^\infty 1_{B(0,1)}(B_t)dt > \beta\right) \\ = \quad &-\inf\left\{\frac{1}{2}\mathbf{D}(u,u); u \in C_0^\infty(\mathbb{R}^3),\ \int_{\{|x|<1\}} u^2 dx = 1\right\}. \end{aligned}$$

The right hand side equals $\frac{\pi^2}{8}$ (e.g. [28, Exercise 6.4.10]). $\frac{\pi^2}{8}$ is also the principal eigenvalue of the Dirichlet Laplacian on the interval $(-1, 1)$. Hence, for any $x \in \mathbb{R}^3$ and $y \in (-1, 1)$

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log \mathbb{P}_x^W\left(\int_0^\infty 1_{B(0,1)}(B_t)dt > \beta\right) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_y^W(t < \tau_{(-1,1)}).$$

Here, $\mathbb{P}_y^W$ means the one-dimensional Brownian motion. When both x and y are the origin of $\mathbb{R}^3$ and $\mathbb{R}$ respectively, the equation above is a corollary of Ciesielski-Taylor theorem ([48]): $\tau_{(-1,1)}$ with respect to the one-dimensional Wiener measure $\mathbb{P}_o^W$ has the same distribution as $\int_0^\infty 1_{B(0,1)}(B_t)dt$ with respect to the three-dimensional Wiener measure $\mathbb{P}_o^W$.

5 A necessary and sufficient condition for gaugeability

In this section, we consider the integrability of Feynman-Kac functionals, so-called, gaugeability as an application of Corollary 2.12.

Theorem 5.1. ([60]) *Suppose that* $\mu \in K_d^\infty$ *satisfies* (3.3). *Then*

$$\sup_{x\in D} \mathbb{E}_x^W \left(\exp(A^\mu_{\tau_D})\right) < \infty \tag{5.1}$$

if and only if

$$\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : u \in C_0^\infty(D),\ \int_D u^2 d\mu = 1\right\} > 1. \tag{5.2}$$

Proof. The time changed process Y_t of the part process B_t^D by $A^\mu_{\tau_D \wedge t}$ satisfies Assumption I $\sim$ III. Note $A^\mu_{\tau_D}$ is the lifetime of Y_t. Denote by $\check{\mathcal{E}}^D$ the Dirichlet form generated by Y_t. Then, Corollary 2.12 tells us that the equation (5.1) holds if and only if

$$\inf\left\{\check{\mathcal{E}}^D(u,u) : u \in \mathcal{D}(\check{\mathcal{E}}^D),\ \int_F u^2 d\mu = 1\right\} > 1.$$

By Lemma 4.2, the left hand side above is equal to the left hand side of (5.2). $\square$

Example 5.1. Let $\mu \in K_d^\infty$. For any compact set $K \subset D$, define

$$\pi(K,D) = \begin{cases} \frac{\mu(K)}{\mathrm{Cap}(K,D)} & \text{for } \ \mathrm{Cap}(K,D) > 0, \\ 0 & \text{for } \ \mathrm{Cap}(K,D) = 0. \end{cases}$$

Here, $\mathrm{Cap}(K,D) = \inf\{\mathbf{D}(u,u); u \geq 1 \text{ on } K,\ u \in C_0^\infty(D)\}$. It is known in Theorem 2.5.2/1 in [43] that

$$\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : u \in C_0^\infty(D),\ \int_D u^2 d\mu = 1\right\} > \begin{cases} 0 & \text{if } \sup_{K\subset D} \pi(K,D) < \infty \\ 1 & \text{if } \sup_{K\subset D} \pi(K,D) < \frac{1}{8}. \end{cases}$$

Let $d = 3$. Let H be a 2-dimensional hyperplane in $\mathbb{R}^3$ and M a Borel subset of H with regular boundary. Let μ be the positive measure defined by $\mu(B) = m(M \cap B)$, where m is the 2-dimensional Lebesgue measure. Then, it is known in [43, p.139] that

$$\sup_{F\subset\mathbb{R}^3} \pi(F,\mathbb{R}^3) \leq \frac{\pi^{1/2}}{8} m(M)^{1/2}.$$

As a result, if $m(M) < \frac{1}{\pi}$, then $\mathbb{E}_x^W(e^{A^\mu_\infty}) < \infty$.

For $d \geq 3$ and a closed set F

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}_x^W \left(\exp\left(\int_0^\infty 1_F(B_t)dt\right)\right) < \infty$$

if

$$\sup_{K\subset\mathbb{R}^d} \frac{|F \cap K|}{\mathrm{Cap}(K)} < \frac{1}{8}. \tag{5.3}$$

Here $|\cdot|$ means the Lebesgue measure on $\mathbb{R}^d$ and $\mathrm{Cap}(K) = \mathrm{Cap}(K, \mathbb{R}^d)$. For a closed set F denote by B_F the ball with the same volume as F:

$$B_F = B(0, r_F), \qquad r_F = \frac{\left(|F|\Gamma(\frac{d}{2}+1)\right)^{1/d}}{\sqrt{\pi}}.$$

Since by [43, 2.2.3, 2.2.4]

$$\begin{aligned}\frac{|F\cap K|}{\mathrm{Cap}(K)} &\leq \frac{|F\cap K|}{\mathrm{Cap}(F\cap K)} \leq \frac{|F\cap K|}{\mathrm{Cap}(B_{F\cap K})} \\ &= \frac{|B_{F\cap K}|}{\mathrm{Cap}(B_{F\cap K})} \leq \frac{|B_F|}{\mathrm{Cap}(B_F)} = \frac{{r_F}^2}{d(d-2)},\end{aligned}$$

the equation (5.3) holds if

$$|F| < \frac{\left(\frac{1}{8}d(d-2)\pi\right)^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}.$$

Example 5.2. Let (M, g) be a spherically symmetric Riemannian manifold with a pole o and consider the Brownian motion $(\mathbb{P}_x, X_t)$ on M. The Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ generated by the Brownian motion is as follows:

$$\begin{cases} \mathcal{E}(u,u) = \frac{1}{2}\int_M (\nabla u, \nabla v) dv_g, \quad u,\ v \in \mathcal{D}(\mathcal{E}) \\ \mathcal{D}(\mathcal{E}) = \text{the clousre of } C_0^\infty \text{ with respect to } \mathcal{E} + (\ ,\ )_{v_g}, \end{cases}$$

where v_g is the Riemannian volume.

Let $B_r = \{x \in M : \rho(o,x) < r\}$ and ∂B_r its boundary. Let σ_r be the surface measure of ∂B_r and $S(r)$ the area of ∂B_r, $S(r) = \sigma_r(\partial B_r)$. The measure σ_r belongs to $\mathcal{K}_\infty(G)$ (We can define $\mathcal{K}_\infty(G)$ by the same way as $\mathcal{K}_d^\infty$. Suppose that M is hyperbolic, i.e.,

$$\int_1^\infty \frac{dr}{S(r)} < \infty.$$

(see [31]). On account of the Dirichlet principle, we see that for $R > 0$

$$\begin{aligned} &\inf\left\{\frac{1}{2}\int_M (\nabla v, \nabla v) dv_g : v \in \mathcal{F},\ \int_{\partial B_R} v^2 d\sigma = 1\right\} \\ = &\inf\left\{\frac{1}{2}\int_M (\nabla v, \nabla v) dv_g : v = H_{\partial B_R} f(x),\ \int_{\partial B_R} f^2 d\sigma = 1\right\}. \end{aligned}$$

Here $H_{\partial B_R} f(x) = \mathbb{E}_x(f(X_{\sigma_{\partial B_R}}); \sigma_{\partial B_R} < \infty)$, $\sigma_{\partial B_R} = \inf\{t > 0 : X_t \in \partial B_R\}$. By the spherical symmetry, the infimum is attained by the function $v(x)$:

$$v(x) = c \cdot \mathbb{P}_x(\sigma_{\partial B_R} < \infty),$$

where $c = 1/\sqrt{S(R)}$. Since the Green function $R(o,x)$ is written as

$$R(o,x) = 2\int_{d(o,x)}^\infty \frac{dr}{S(r)}$$

([31, Example 4.1]), we see that

$$v(x) = \begin{cases} \frac{1}{\sqrt{S(R)}\int_R^\infty \frac{dr}{S(r)}} \int_{d(o,x)}^\infty \frac{dr}{S(r)} & d(o,x) > R \\ \frac{1}{\sqrt{S(R)}} & d(o,x) \leq R, \end{cases} \tag{5.4}$$

and thus

$$\frac{1}{2}\int_M (\nabla v, \nabla v) dv_g = \frac{1}{2S(R)\int_R^\infty \frac{dr}{S(r)}}. \tag{5.5}$$

Therefore, we can conclude that

$$2S(R)\int_R^\infty \frac{1}{S(r)}dr < 1 \iff \sup_{x\in M} \mathbb{E}_x\left(e^{\ell_R(\infty)}\right) < \infty, \tag{5.6}$$

where $\ell_R(t)$ the PCAF corresponding to σ_R. For $M = \mathbb{R}^d$ ($d \geq 3$), $S(r) = \omega_d r^{d-1}$ (ω_d: the area of the unit sphere in $\mathbb{R}^d$), and we see that the measure σ_R is gaugeable if and only if $\frac{d-2}{2} > R$.

If M is 2-dimensional hyperbolic space H^2, then $S(r) = \omega_2 \sinh r$ and

$$2S(R)\int_R^\infty \frac{1}{S(r)}dr = (e^R - e^{-R})\log\left(\frac{e^R+1}{e^R-1}\right).$$

Put

$$f(r) = (e^r - e^{-r})\log\left(\frac{e^r+1}{e^r-1}\right), \quad r > 0.$$

Then $f(r)$ is strictly increasing, $\lim_{r\to 0} f(r) = 0$, and $\lim_{r\to\infty} f(r) = 2$. Hence the equation $f(r) = 1$ has a unique root $r_0(\approx 0.22767)$, and if $R < r_0$, then σ_R is gaugeable.

Let us consider 3-dimensional hyperbolic space H^3. Then $S(r) = \omega_3 \sinh^2 r$ and

$$2S(R)\int_R^\infty \frac{1}{S(r)}dr = \frac{e^{2R}-1}{e^{2R}} < 1. \tag{5.7}$$

Hence, σ_R is gaugeable for all $R > 0$, and from which σ_R is expected to be gaugeable for all $R > 0$ in case $d \geq 4$. In fact,

$$\begin{aligned} 2S(R)\int_R^\infty \frac{1}{S(r)}dr &= 2(e^R - e^{-R})^{d-1}\int_R^\infty \frac{1}{(e^r - e^{-r})^{d-1}}dr \\ &\leq 2(e^R - e^{-R})^{d-1}\int_R^\infty \frac{1}{(e^r - e^{-R})^{d-1}}dr \\ &< \frac{2}{d-1} < 1. \end{aligned}$$

The left hand side of (5.7) equals to $\mathrm{Cap}(\partial B_R)/S(R)$. Hence we can also say that the measure σ_R is gaugeable if and only if R satisfies

$$\mathrm{Cap}(\partial B_R) > S(R).$$

As a typical example of jump processes we deal with symmetric α-stable process, the Markov process generated by $(-\Delta)^{\alpha/2}$. Let $\mathbf{M}^\alpha = (\Omega, \mathcal{F}, \mathcal{F}_t, \theta_t, \mathbb{P}_x, X_t)$ be the symmetric α-stable process on $\mathbb{R}^d$ with $0 < \alpha < 2$. Here $\{\mathcal{F}_t\}_{t\geq 0}$ is the minimal (augmented) admissible filtration and θ_t, $t \geq 0$, is the shift operators satisfying $X_s(\theta_t) = X_{s+t}$ identically for $s, t \geq 0$. When $\alpha = 2$, $\mathbf{M}^\alpha$ is the Brownian motion. Let $p(t, x, y)$ be the transition density function of $\mathbf{M}^\alpha$ and $R_\beta(x, y)$, $\beta \geq 0$, be its β-Green function,

$$R_\beta(x, y) = \int_0^\infty e^{-\beta t} p(t, x, y) dt.$$

If the process $\mathbf{M}^\alpha$ is transient, that is, $0 < \alpha < d$, then 0-Green function $R_0(x, y)$ is given by

$$R_0(x, y) = \int_0^\infty p(t, x, y) dt = C(d, \alpha)|x - y|^{\alpha - d}, \tag{5.8}$$

where $C(d, \alpha) = 2^{-\alpha}\pi^{-d/2}\Gamma(\frac{d-\alpha}{2})\Gamma(\frac{\alpha}{2})^{-1}$ and Γ is the Gamma function. R_0 is called *Riesz kernel* of index α. We write $R(x, y)$ for $R_0(x, y)$ simply. For a positive measure μ, the β-potential of μ is defined by

$$R_\beta\mu(x) = \int_{\mathbb{R}^d} R_\beta(x, y)\mu(dy).$$

We write $R\mu$ for $R_0\mu$. We Let P_t be the semigroup of $\mathbf{M}^\alpha$,

$$P_t f(x) = \int_{\mathbb{R}^d} p(t, x, y) f(y) dy = \mathbb{E}_x(f(X_t)).$$

Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be the Dirichlet form generated by $\mathbf{M}^\alpha$: for $0 < \alpha < 2$

$$\begin{cases} \mathcal{E}(u, v) = \mathcal{A}(d, \alpha) \displaystyle\iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \frac{(u(x) - u(y))(v(x) - v(y))}{|x - y|^{d+\alpha}} dx dy \\ \mathcal{D}(\mathcal{E}) = \left\{ u \in L^2(\mathbb{R}^d) : \displaystyle\iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \frac{(u(x) - u(y))^2}{|x - y|^{d+\alpha}} dx dy < \infty \right\}, \end{cases} \tag{5.9}$$

where $\Delta = \{(x, x) : x \in \mathbb{R}^d\}$ and

$$\mathcal{A}(d, \alpha) = \frac{\alpha 2^{d-1}\Gamma(\frac{\alpha+d}{2})}{\pi^{d/2}\Gamma(1 - \frac{\alpha}{2})}$$

([28, Example 1.4.1]).

Let $\mathcal{D}_e(\mathcal{E})$ denote the extended Dirichlet space ([28, Section 1.5]). If $\alpha < d$, that is, the process $\mathbf{M}^\alpha$ is transient, then $\mathcal{D}_e(\mathcal{E})$ is a Hilbert space with inner product $\mathcal{E}$ ([28, Theorem 1.5.3]).

We now define classes of measures which play an important role in this paper.

Definition 5.2. (I) A positive Radon measure μ on $\mathbb{R}^d$ is said to be in the *Kato class* ($\mu \in K_{d,\alpha}$ in notation), if

$$\lim_{\beta\to\infty} \sup_{x\in\mathbb{R}^d} R_\beta\mu(x) = 0. \tag{5.10}$$

(II) A measure μ is said to be β-*Green-tight* ($\mu \in K^\infty_{d,\alpha}(\beta)$ in notation), if μ is in $K_{d,\alpha}$ and satisfies

$$\lim_{A\to\infty} \sup_{x\in\mathbb{R}^d} \int_{|y|>A} R_\beta(x,y)\mu(dy) = 0. \tag{5.11}$$

For the Browian motion, the definition (5.10) is equivalent to that defined in Section 3 ([41]). We see from the resolvent equation that for $\beta > 0$

$$K^\infty_{d,\alpha}(\beta) = K^\infty_{d,\alpha}(1).$$

When $d > \alpha$, that is, $\mathbf{M}^\alpha$ is transient, we write $K^\infty_{d,\alpha}$ for $K^\infty_{d,\alpha}(0)$. For $\mu \in K_{d,\alpha}$, define a symmetric bilinear form $\mathcal{E}^\mu$ by

$$\mathcal{E}^\mu(u,u) = \mathcal{E}(u,u) - \int_{\mathbb{R}^d} u^2 d\mu, \quad u \in \mathcal{D}(\mathcal{E}). \tag{5.12}$$

Since $\mu \in K^\infty_{d,\alpha}$ charges no set of zero capacity by [1, Theorem 3.3], the form $\mathcal{E}^\mu$ is well defined. We see from [1, Theorem 4.1] that $(\mathcal{E}^\mu, \mathcal{D}(\mathcal{E}))$ becomes a lower semi-bounded closed symmetric form. Denote by $\mathcal{H}^\mu$ the self-adjoint operator generated by $(\mathcal{E}^\mu, \mathcal{D}(\mathcal{E}))$: $\mathcal{E}^\mu(u,v) = (\mathcal{H}^\mu u, v)$. Let P^μ_t be the L^2-semigroup generated by $\mathcal{H}^\mu$: $P^\mu_t = \exp(-t\mathcal{H}^\mu)$. We see from [1, Theorem 6.3(iv)] that P^μ_t admits a symmetric integral kernel $p^\mu(t,x,y)$ which is jointly continuous function on $(0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d$.

For $\mu \in K^\infty_{d,\alpha}$, let A^μ_t be a PCAF which is in the Revuz correspondence to μ. By the Feynman-Kac formula, the semigroup P^μ_t is written as

$$P^\mu_t f(x) = \mathbb{E}_x(\exp(A^\mu_t) f(X_t)). \tag{5.13}$$

Theorem 5.3. ([55]) *Let $\mu \in K_{d,\alpha}$. Then*

$$\int_{\mathbb{R}^d} u^2(x)\mu(dx) \le \|R_\beta\mu\|_\infty \mathcal{E}_\beta(u,u), \quad u \in \mathcal{D}(\mathcal{E}), \tag{5.14}$$

where $\mathcal{E}_\beta(u,u) = \mathcal{E}(u,u) + \beta\int_{\mathbb{R}^d} u^2 dx$.

Theorem 5.4. ([71, Theorem 3.4], [64, Theorem 2.7]) *If $\mu \in K^\infty_{d,\alpha}(1)$, then the embedding of $\mathcal{D}(\mathcal{E})$ into $L^2(\mu)$ is compact. If $d > \alpha$ and $\mu \in K^\infty_{d,\alpha}$, then the embedding of $\mathcal{D}_e(\mathcal{E})$ into $L^2(\mu)$ is compact.*

Example 5.3. Let σ_r be the surface measure of ∂B_r, where ∂B_r is the sphere with radius $r > 0$ at 0. Since the symmetric α-stable process hits the sphere ∂B_r if $1 < \alpha \leq 2$, the surface measure σ_r is smooth. Denote by $\ell_r(t)$ the additive functional corresponding to σ_r. The surface measure σ_r is then gaugeable if and only if

$$\inf \left\{ \mathcal{E}^{(\alpha)}(u,u) : \int_{\{|x|=r\}} u^2 d\sigma_r = 1 \right\} > 1.$$

Since the measure σ_r is spherically symmetric, the infimum is attained by the function

$$u(x) = c\,\mathbb{P}_x(\sigma_{\partial B_r} < \infty), \ x \in \mathbb{R}^d,$$

where $c = 1/\sqrt{\sigma(\partial B_r)}$. Let $\mathrm{Cap}^{(\alpha)}(\cdot)$ be the 0-order capacity with respect to the symmetric α-stable process. Then the infimum above becomes

$$\frac{\mathrm{Cap}^{(\alpha)}(\partial B_r)}{\sigma_r(\partial B_r)},$$

because

$$\mathcal{E}^{(\alpha)}(\mathbb{P}_\cdot(\sigma_{\partial B_r} < \infty), \mathbb{P}_\cdot(\sigma_{\partial B_r} < \infty)) = \mathrm{Cap}^{(\alpha)}(\partial B_r).$$

It is known that

$$\mathrm{Cap}^{(\alpha)}(\partial B_r) = \frac{2\pi^{(d+1)/2}\Gamma\left(\dfrac{\alpha+d}{2} - 1\right)\Gamma\left(\dfrac{\alpha}{2}\right)}{\Gamma\left(\dfrac{d}{2}\right)\Gamma\left(\dfrac{\alpha-1}{2}\right)\Gamma\left(\dfrac{d-\alpha}{2}\right)} r^{d-\alpha}. \tag{5.15}$$

Therefore, since $\sigma_r(\partial B_r) = 2\pi^{d/2}\Gamma(d/2)^{-1}r^{d-1}$ for $r > 0$, the surface measure σ_r is gaugeable if and only if

$$\left\{ \frac{\sqrt{\pi}\,\Gamma\left(\dfrac{d+\alpha}{2} - 1\right)\Gamma\left(\dfrac{\alpha}{2}\right)}{\Gamma\left(\dfrac{\alpha-1}{2}\right)\Gamma\left(\dfrac{d-\alpha}{2}\right)} \right\}^{\frac{1}{\alpha-1}} > r.$$

Let μ_r be the equilibrium measure of ∂B_r. Since the rotation invariance of the set ∂B_r, we see that, $\mu_r = A\sigma_r$ for some constant $A > 0$. Then by the definition of the equilibrium measure, we have

$$\mu_r(\partial B_r) = \mathrm{Cap}(\partial B_r).$$

So, it follows that $A = \mathrm{Cap}(\partial B_r)/\sigma_r(\partial B_r)$, hence,

$$\mu_r = \frac{\mathrm{Cap}(\partial B_r)}{\sigma_r(\partial B_r)}\sigma_r.$$

Therefore

$$\begin{aligned}\sup_{x\in\mathbb{R}^d}\mathbb{E}_x(\ell_r(\infty)) &= \sup_{x\in\mathbb{R}^d} R\sigma_r(x)\\ &= \sup_{x\in\partial B_r} R\sigma_r(x) \qquad \text{(by the maximum principle)}\\ &= \frac{\sigma_r(\partial B_r)}{\mathrm{Cap}(\partial B_r)}\sup_{x\in\partial B_r} R\mu_r(x)\\ &= \frac{\sigma_r(\partial B_r)}{\mathrm{Cap}(\partial B_r)}.\end{aligned}$$

Thus

$$\sup_{x\in\mathbb{R}^d}\mathbb{E}_x(\ell_r(\infty)) < 1 \iff \sup_{x\in\mathbb{R}^d}\mathbb{E}_x(\exp(\ell_r(\infty))) < \infty.$$

The implication ($\Longrightarrow$) follows generally from Khas'minskii's lemma.

Example 5.4. Let $\mu = \mu^+ - \mu^- \in K_{d,\alpha}^{\infty}$ with $\mu^+ \not\equiv 0$ and $\mu^- \not\equiv 0$. Consider the Schrödinger type operator

$$\mathcal{L}^\theta = \frac{1}{2}(-\Delta)^{\alpha/2} + \theta\mu, \qquad \theta \in \mathbb{R}^1.$$

Then it follows from Theorem 3.1 that the operator $\mathcal{L}^\theta$ $(\theta > 0)$ is subcritical if and only if

$$\begin{aligned}\lambda(\theta\mu) &= \inf\Big\{\mathcal{E}^{(\alpha)}(u,u) + \theta\int_{\mathbb{R}^d} u^2(x)\mu^+(dx) :\ u \in \mathcal{D}(\mathcal{E}^{(\alpha)}),\\ &\qquad\qquad \theta\int_{\mathbb{R}^d} u^2(x)\mu^-(dx) = 1\Big\} > 1.\end{aligned} \tag{5.16}$$

Let $\mathbb{R}^d = F + F^c$ be the Hahn decomposition: $\mu(F) = \mu^+(\mathbb{R}^d)$, $\mu^-(F^c) = -\mu(\mathbb{R}^d)$. Take $R > 0$ so large that $\mu^-(F^c \cap B_R) > 0$. Let $A = F^c \cap B_R$ and take a sequence of non-negative functions f_n in $C_0^\infty(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d}(1_A(x) - f_n(x))^2|\mu|(dx) \longrightarrow 0 \quad \text{as } n \to \infty.$$

It then holds that

$$\lim_{n\to\infty}\int_{\mathbb{R}^d} f_n^2(x)\mu^-(dx) = \mu^-(A) > 0, \quad \lim_{n\to\infty}\int_{\mathbb{R}^d} f_n^2(x)\mu^+(dx) = \mu^+(A) = 0,$$

and consequently, there exists a function $f \in C_0^\infty(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} f^2(x)\mu^-(dx) = 1, \quad \int_{\mathbb{R}^d} f^2(x)\mu^+(dx) < 1. \tag{5.17}$$

Put

$$F(\theta) = \inf\left\{\mathcal{E}^{(\alpha)}(u,u) + \theta\int_{\mathbb{R}^d} u^2(x)\mu^+(dx) :\ u \in \mathcal{D}(\mathcal{E}^{(\alpha)}),\ \int_{\mathbb{R}^d} u^2(x)\mu^-(dx) = 1\right\}.$$

Then $F(\theta)$, $\theta \geq 0$, is a concave function with $F(0) > 0$ by the definition. Moreover, $F(\theta)$ is dominated by the function $G(\theta) := \mathcal{E}^{(\alpha)}(f,f) + \theta\int_{\mathbb{R}^d} f^2(x)\mu^+(dx)$, where f is a function satisfying (5.17). On account of these properties of F and Theorem 5.4, we see that there exists a unique $\theta^+ > 0$ such that $F(\theta^+) = \theta^+$. By the same argument, there exists a unique $\theta^- < 0$ such that $F(\theta^-) = \theta^-$. Noting that the right hand side of (5.16) is equal to $F(\theta)/\theta$, we see that the operator $\mathcal{L}^\theta$ is subcritical for $\theta^- < \theta < \theta^+$. The operators $\mathcal{L}^\theta$ is *critical* for $\theta = \theta^\pm$ and is *supercritical* for $\theta < \theta^-$ or $\theta > \theta^+$ in the sense of [44]: for $\theta = \theta^\pm$, the equation $\mathcal{L}^\theta u = 0$ has a strictly positive continuous solution h which satisfies

$$k^{-1}R(0,x) \leq h(x) \leq kR(0,x) \quad x \in B_r^c,$$

where $k > 1$ (see (4.19) in [71]).

6 Scattering length and capacity

In Section 6, we extend their results in H. Tamura [74], Y. Takahashi [57], and give another simple proof of the conjecture of Kac by the time-change argument for Dirichlet forms. We have:

Theorem 6.1. ([65]) *Let μ be a finite smooth measure with fine support $\widetilde{Y}$. Then*

$$\lim_{\alpha\to\infty} \Gamma(\alpha\mu) = \mathrm{Cap}(\widetilde{Y}). \tag{6.1}$$

Here Cap *is the capacity.*

Let

$$\Gamma(\mu) = \int_X \mathbb{E}_x\left(e^{-A_\zeta}\right)\mu(dx). \tag{6.2}$$

If $\mathbf{M}$ is conservative, then $\Gamma(\mu)$ equals

$$\Gamma(\mu) = \lim_{t\to\infty}\frac{1}{t}\int_X \left(1 - \mathbb{E}_x\left(e^{-A_t^\mu}\right)\right) m(dx). \tag{6.3}$$

Indeed, let $\mathbf{M}^A = (\mathbb{P}_x^A, X_t)$ be the subprocess by $e^{-A_t^\mu}$. We then see from [25, Theorem 2.22] that the Revuz measure of A_t^μ with respect to the subprocess $\mathbf{M}^A$ is also μ and from [52, (62.13)] that

$$\mathbb{E}_x\left(\int_0^t e^{-A_s^\mu} dA_s^\mu\right) = \mathbb{E}_x^A\left(A_t^\mu\right).$$

Hence by Theorem 5.1.3 (iii),

$$\begin{aligned}\left\langle m, 1-\mathbb{E}_x\left(e^{-A_t^\mu}\right)\right\rangle &= \left\langle m, \mathbb{E}_x\left(\int_0^t e^{-A_s^\mu} dA_s^\mu\right)\right\rangle = \langle m, \mathbb{E}_x^A\left(A_t^\mu\right)\rangle \\ &= \int_0^t \langle \mu, p_s^A 1\rangle ds.\end{aligned}$$

Since $p_t^A 1(x)$ converges to $\mathbb{E}_x(e^{-A_\infty^\mu})$ as $t \to \infty$, we have the equation (6.2).

Lemma 6.2.

$$\Gamma(\alpha\mu) = \alpha \int_X \mathbb{E}_x\left(e^{-\alpha A_\infty^\mu}\right)\mu(dx) \uparrow \operatorname{Cap}_{(0)}(\widetilde{Y}), \quad \alpha \uparrow \infty. \tag{6.4}$$

Proof. First note that the lifetime $\check{\zeta}$ of $\check{\mathbf{M}}$ is A_ζ. For $x \in \widetilde{Y}$

$$\begin{aligned}\mathbb{E}_x\left(e^{-\alpha A_\infty}\right) &= \check{\mathbb{E}}_x\left(e^{-\alpha\check{\zeta}}\right) = 1 - \alpha\,\check{\mathbb{E}}_x\left(\int_0^{\check{\zeta}} e^{-\alpha t} dt\right) \\ &= 1 - \alpha \check{R}_\alpha 1(x),\end{aligned}$$

where 1 is the identity function on $\widetilde{Y}$, $1 = 1_{\widetilde{Y}}(x)$. Hence the left hand side of (6.4) equals $\alpha(1, 1-\alpha\check{R}_\alpha 1)_\mu$.

Noting that the function $1_{\widetilde{Y}}$ is in $L^2(\widetilde{Y};\mu)$ by the finiteness of μ, we know that if $1_{\widetilde{Y}} \in \mathcal{D}(\check{\mathcal{E}})$, then $\alpha(1, 1-\alpha\check{R}_\alpha 1)_\mu$ is non-decreasingly convergent to $\check{\mathcal{E}}(1,1)$ as $\alpha \uparrow \infty$. We see from [28] that

$$\check{\mathcal{E}}(1,1) = \mathcal{E}(H_{\widetilde{Y}}1, H_{\widetilde{Y}}1), \quad H_{\widetilde{Y}}1(x) = \mathbb{E}_x(1_{\widetilde{Y}}(X_{\sigma_{\widetilde{Y}}}); \sigma_{\widetilde{Y}} < \infty),$$

where $\sigma_{\widetilde{Y}} = \inf\{t > 0 : X_t \in \widetilde{Y}\}$. We thus have

$$\Gamma(\alpha\mu) = \alpha(1, 1-\alpha\check{R}_\alpha 1)_\mu \uparrow \check{\mathcal{E}}(1,1) = \mathcal{E}(H_{\widetilde{Y}}1, H_{\widetilde{Y}}1) \tag{6.5}$$

as $\alpha \uparrow \infty$. Since $\widetilde{Y}$ is a nearly Borel, finely closed set, $\mathbb{P}_x(X_{\sigma_{\widetilde{Y}}} \in \widetilde{Y}) = 1$ and thus

$$H_{\widetilde{Y}}1(x) = \mathbb{P}_x(\sigma_{\widetilde{Y}} < \infty).$$

Therefore, the right hand side of (6.5) equals $\operatorname{Cap}_{(0)}(\widetilde{Y})$ by [28, Theorem 4.3.3].

If $1_{\widetilde{Y}} \notin \mathcal{D}(\check{\mathcal{E}})$, then $\lim_{\alpha\to\infty} \alpha(1, 1-\alpha\check{R}_\alpha 1)_\mu \uparrow \infty$ as $\alpha \uparrow \infty$ and $\operatorname{Cap}_{(0}(\widetilde{Y}) = \infty$. The proof of the lemma is complete. □

P. He [33] extends Theorem 6.1 to Markov processes under weak duality.

7 Feynman-Kac penalization for symmetric α-stable processes

In Sections 7, 8 and 9, we consider Feynman-Kac penalization problem as an application of Theorem 5.1. In [49], [50], B. Roynette, P. Vallois and M. Yor have studied limit theorems for Wiener processes normalized by some weight processes. In [79], K. Yano, Y. Yano and M. Yor studied the limit theorems for the one-dimensional symmetric stable process normalized by non-negative functions of the local times or by negative (killing) Feynman-Kac functionals. They call the limit theorems for Markov processes normalized by Feynman-Kac functionals the *Feynman-Kac penalizations*. Our aim is to extend their results on Feynman-Kac penalizations to positive Feynman-Kac functionals of multi-dimensional symmetric α-stable processes.

Let $\mathbf{M}^\alpha = (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}_x, X_t)$ be the symmetric α-stable process on $\mathbb{R}^d$ with $0 < \alpha < 2$ and $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ the Dirichlet form of $\mathbf{M}^\alpha$ defined in (5.9). Let μ be a positive Radon measure in the class $K_{d,\alpha}^\infty$ of Green-tight Kato measures (Definition 5.2). We denote by A_t^μ the positive continuous additive functional (PCAF in abbreviation) in the Revuz correspondence to μ. We define the family $\{\mathbb{Q}_{x,t}^\mu\}$ of normalized probability measures by

$$\mathbb{Q}_{x,t}^\mu(B) = \frac{1}{Z_t^\mu(x)} \int_B \exp(A_t^\mu(\omega))\mathbb{P}_x(d\omega), \quad B \in \mathcal{F}_t,$$

where $Z_t^\mu(x) = \mathbb{E}_x(\exp(A_t^\mu))$. Our interest is the limit of $\mathbb{Q}_{x,t}^\mu$ as $t \to \infty$, mainly in transient cases, $d > \alpha$. They in [79] treated negative Feynman-Kac functionals in the case of the one-dimensional recurrent stable process, $\alpha > 1$. In this case, the decay rate of $Z_t^\mu(x)$ is important, while in our cases the growth order is.

We define

$$\lambda(\theta) = \inf\left\{\mathcal{E}_\theta(u,u) : \int_{\mathbb{R}^d} u^2 d\mu = 1\right\}, \quad 0 \le \theta < \infty, \tag{7.1}$$

where $\mathcal{E}_\theta(u,u) = \mathcal{E}(u,u) + \theta \int_{\mathbb{R}^d} u^2 dx$. We see from [28, Theorem 6.2.1] and [59, Lemma 3.1] that the time-changed process by A_t^μ is symmetric with respect to μ and $\lambda(0)$ equals the bottom of the spectrum of the time-changed process. We now classify the set $K_{d,\alpha}^\infty$ in terms of $\lambda(0)$:

(i) $\lambda(0) < 1$

In this case, there exist a positive constant $\theta_0 > 0$ and a positive continuous function h in the Dirichlet space $\mathcal{D}(\mathcal{E})$ such that

$$1 = \lambda(\theta_0) = \mathcal{E}_{\theta_0}(h,h)$$

(Lemma 8.1, Theorem 5.4). We define the multiplicative functional (MF in abbreviation) L_t^h by

$$L_t^h = e^{-\theta_0 t}\frac{h(X_t)}{h(X_0)}e^{A_t^\mu}. \tag{7.2}$$

(ii) $\lambda(0) = 1$

In this case, there exists a positive continuous function h in the extended Dirichlet space $\mathcal{D}_e(\mathcal{E})$ such that

$$1 = \lambda(0) = \mathcal{E}(h, h)$$

([71, Theorem 3.4]). We define

$$L_t^h = \frac{h(X_t)}{h(X_0)}e^{A_t^\mu}. \tag{7.3}$$

(iii) $\lambda(0) > 1$

In this case, the measure μ is *gaugeable*, that is,

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}_x\left(e^{A_\infty^\mu}\right) < \infty$$

(see Theorem 9.4 below). We put $h(x) = \mathbb{E}_x(e^{A_\infty^\mu})$ and define

$$L_t^h = \frac{h(X_t)}{h(X_0)}e^{A_t^\mu}. \tag{7.4}$$

The cases **(i)**, **(ii)**, and **(iii)** are corresponding to the *supercriticality*, *criticality*, and *subcriticality* of the operator, $-(-\Delta)^{\alpha/2} + \mu$, respectively (Theorem 9.4). We will see that L_t^h is a martingale MF for each case, i.e., $\mathbb{E}_x(L_t^h) = 1$. Let $\mathbf{M}^h = (\Omega, \mathbb{P}_x^h, X_t)$ be the transformed process of $\mathbf{M}^\alpha$ by L_t^h:

$$\mathbb{P}_x^h(B) = \int_B L_t^h(\omega)\mathbb{P}_x(d\omega), \quad B \in \mathcal{F}_t.$$

We then see from [14, Theorem 2.6] and Proposition 8.3 below that if $\lambda(0) \le 1$, then $\mathbf{M}^h$ is an h^2dx-symmetric Harris recurrent Markov process.

To state the main result of this paper, we need to introduce a subclass K_S^∞ of $K_{d,\alpha}^\infty$; a measure $\mu \in K_{d,\alpha}^\infty$ is said to be in K_S^∞ if

$$\sup_{x\in\mathbb{R}^d}\left(|x|^{d-\alpha}\int_{\mathbb{R}^d}\frac{d\mu(y)}{|x-y|^{d-\alpha}}\right) < \infty. \tag{7.5}$$

This class is relevant to the notion of *special* PCAF's which was introduced by J. Neveu ([45]); we will show in Lemma 9.5 that if a measure μ belongs to K_S^∞, then $\int_0^t (1/h(X_s))dA_s^\mu$ is a *special* PCAF of $\mathbf{M}^h$. This fact is crucial for the proof of the main theorem below. In fact, a key to the proof lies in the application of the Chacon-Ornstein type ergodic theorem for special PCAF's of Harris recurrent Markov processes ([9, Theorem 3.18]).

We then have the next main theorem.

Theorem 7.1. ([66]) (i) *If $\lambda(0) \neq 1$, then*

$$\mathbb{Q}_{x,t}^\mu \overset{t\to\infty}{\longrightarrow} \mathbb{P}_x^h \quad \text{along } (\mathcal{F}_t), \tag{7.6}$$

that is, for any $s \geq 0$ and any bounded $\mathcal{F}_s$-measurable function Z,

$$\lim_{t\to\infty} \frac{\mathbb{E}_x\left(Z\exp(A_t^\mu)\right)}{\mathbb{E}_x\left(\exp(A_t^\mu)\right)} = \mathbb{E}_x^h(Z).$$

(ii) *If $\lambda(0) = 1$ and $\mu \in K_S^\infty$, then* (7.6) *holds.*

8 Construction of ground states

For $d \leq \alpha$ (resp. $d > \alpha$), let μ be a non-trivial measure in $K_{d,\alpha}^\infty(1)$ (resp. $K_{d,\alpha}^\infty$). Define

$$\lambda(\theta) = \inf\left\{\mathcal{E}_\theta(u,u) : \int_{\mathbb{R}^d} u^2 d\mu = 1\right\}, \quad \theta \geq 0. \tag{8.1}$$

Lemma 8.1. *The function $\lambda(\theta)$ is increasing and concave. Moreover, it satisfies $\lim_{\theta\to\infty}\lambda(\theta) = \infty$.*

Proof. It follows from the definition of $\lambda(\theta)$ that it is increasing. For $\theta_1, \theta_2 \geq 0$, $0 \leq t \leq 1$

$$\begin{aligned}\lambda(t\theta_1 + (1-t)\theta_2) &= \inf\left\{\mathcal{E}_{t\theta_1+(1-t)\theta_2}(u,u) : \int_{\mathbb{R}^d} u^2 d\mu = 1\right\} \\ &\geq t\inf\left\{\mathcal{E}_{\theta_1}(u,u) : \int_{\mathbb{R}^d} u^2 d\mu = 1\right\} + (1-t)\inf\left\{\mathcal{E}_{\theta_2}(u,u) : \int_{\mathbb{R}^d} u^2 d\mu = 1\right\} \\ &= t\lambda(\theta_1) + (1-t)\lambda(\theta_2).\end{aligned}$$

We see from Theorem 5.3 that for $u \in \mathcal{D}(\mathcal{E})$ with $\int_{\mathbb{R}^d} u^2 d\mu = 1$, $\mathcal{E}_\theta(u,u) \geq 1/\|R_\theta\mu\|_\infty$. Hence we have

$$\lambda(\theta) \geq \frac{1}{\|R_\theta\mu\|_\infty}. \tag{8.2}$$

By the definition of the Kato class, the right hand side of (8.2) tends to infinity as $\theta \to \infty$. □

Lemma 8.2. *If $d \leq \alpha$, then $\lambda(0) = 0$.*

Proof. Note that for $u \in \mathcal{D}(\mathcal{E})$

$$\lambda(0) \int_{\mathbb{R}^d} u^2 d\mu \leq \mathcal{E}(u,u).$$

Since $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is recurrent, there exists a sequence $\{u_n\} \subset \mathcal{D}(\mathcal{E})$ such that $u_n \uparrow 1$ q.e. and $\mathcal{E}(u_n, u_n) \to 0$ ([28, Theorem 1.6.3]). Hence if $\lambda(0) > 0$, then $\mu = 0$, which is contradictory. □

We see from Theorem 5.4 and Lemma 8.2 that if $d \leq \alpha$, then there exist $\theta_0 > 0$ and $h \in \mathcal{D}(\mathcal{E})$ such that

$$\lambda(\theta_0) = \inf\left\{\mathcal{E}_{\theta_0}(h,h) : \int_{\mathbb{R}^d} h^2 d\mu = 1\right\} = 1.$$

We can assume that h is a strictly positive continuous function (e.g. Section 4 in [71]).

Let $M_t^{[h]}$ be the martingale part of the Fukushima decomposition ([28, Theorem 5.2.2]):

$$h(X_t) - h(X_0) = M_t^{[h]} + N_t^{[h]}. \tag{8.3}$$

Define a martingale by

$$M_t = \int_0^t \frac{1}{h(X_{s-})} dM_s^h$$

and denote by L_t^h the unique solution of the Doléans-Dade equation:

$$Z_t = 1 + \int_0^t Z_{s-} dM_s. \tag{8.4}$$

Then we see from the Doléans-Dade formula that L_t^h is expressed by

$$\begin{aligned} L_t^h &= \exp\left(M_t - \frac{1}{2}\langle M^c\rangle_t\right) \prod_{0<s\leq t} (1+\Delta M_s)\exp(-\Delta M_s) \\ &= \exp\left(M_t - \frac{1}{2}\langle M^c\rangle_t\right) \prod_{0<s\leq t} \frac{h(X_s)}{h(X_{s-})} \exp\left(1 - \frac{h(X_s)}{h(X_{s-})}\right). \end{aligned}$$

Here M_t^c is the continuous part of M_t and $\Delta M_s = M_s - M_{s-}$. By Itô's formula applied to the semi-martingale $h(X_t)$ with the function $\log x$, we see that L_t^h has the following expression:

$$L_t^h = e^{-\theta_0 t} \frac{h(X_t)}{h(X_0)} \exp(A_t^\mu). \tag{8.5}$$

Let $d > \alpha$ and suppose that $\theta_0 = 0$, that is,

$$\lambda(0) = \inf\left\{\mathcal{E}(u,u) : \int_{\mathbb{R}^d} u^2 d\mu = 1\right\} = 1.$$

We then see from [71, Theorem 3.4] that there exists a function $h \in \mathcal{D}_e(\mathcal{E})$ such that $\mathcal{E}(h,h) = 1$. We can also assume that h is a strictly positive continuous function and satisfies

$$\frac{c}{|x|^{d-\alpha}} \leq h(x) \leq \frac{C}{|x|^{d-\alpha}}, \quad |x| > 1 \tag{8.6}$$

(see (4.19) in [71]). We define the MF L_t^h by

$$L_t^h = \frac{h(X_t)}{h(X_0)} \exp(A_t^\mu). \tag{8.7}$$

We denote by $\mathbf{M}^h = (\Omega, \mathbb{P}_x^h, X_t)$ the transformed process of $\mathbf{M}^\alpha$ by L_t^h,

$$\mathbb{P}_x^h(d\omega) = L_t^h(\omega) \cdot \mathbb{P}_x(d\omega).$$

Proposition 8.3. The transformed process $\mathbf{M}^h = (\mathbb{P}_x^h, X_t)$ is Harris recurrent, that is, for a non-negative function f with $m(\{x : f(x) > 0\}) > 0$,

$$\int_0^\infty f(X_t)dt = \infty \quad \mathbb{P}_x^h\text{-a.s.}, \tag{8.8}$$

where m is the Lebesgue measure.

Proof. Set $A = \{x : f(x) > 0\}$. Since $\mathbf{M}^h$ is an h^2dx-symmetric recurrent Markov process,

$$\mathbb{P}_x(\sigma_A \circ \theta_n < \infty, \ \forall n \geq 0) = 1 \ \text{ for q.e. } x \in \mathbb{R}^d \tag{8.9}$$

by [28, Theorem 4.7.1](iii). Moreover, since the Markov process $\mathbf{M}^h$ has the transition density function

$$e^{-\theta_0 t} \cdot \frac{p^\mu(t,x,y)}{h(x)h(y)}$$

with respect to h^2dx, (8.9) holds for all $x \in \mathbb{R}^d$ by [28, Exercise 4.7.1]. Using the strong Feller property and the proof of [47, Chapter X, Proposition (3.11)], we see from (8.9) that $\mathbf{M}^h$ is Harris recurrent. □

We see from [71, Theorem 4.15] : If $\theta_0 > 0$, then $h \in L^2(\mathbb{R}^d)$ and $\mathbf{M}^h$ is positive recurrent. If $\theta_0 = 0$ and $\alpha < d \leq 2\alpha$, then $h \notin L^2(\mathbb{R}^d)$ $\mathbf{M}^h$ is null recurrent. If $\theta_0 = 0$ and $d > 2\alpha$, then $h \in L^2(\mathbb{R}^d)$ $\mathbf{M}^h$ is positive recurrent.

9 Penalization problems

In this section, we prove Theorem 7.1.

(1°) **Recurrent case** ($d \leq \alpha$)

Theorem 9.1. ([66]) *Assume that $d \le \alpha$. Then there exist $\theta_0 > 0$ and $h \in \mathcal{D}(\mathcal{E})$ such that $\lambda(\theta_0) = 1$ and $\mathcal{E}_{\theta_0}(h,h) = 1$. Moreover, for each $x \in \mathbb{R}^d$*

$$e^{-\theta_0 t}\mathbb{E}_x\left(e^{A_t^\mu}\right) \longrightarrow h(x)\int_{\mathbb{R}^d} h(x)dx \quad as \quad t \longrightarrow \infty.. \tag{9.1}$$

Proof. The first assertion follows from Theorem 5.4 and Lemma 8.2. Note that

$$e^{-\theta_0 t}\mathbb{E}_x\left(e^{A_t^\mu}\right) = h(x)\mathbb{E}_x^h\left(\frac{1}{h(X_t)}\right)$$

Then by [64, Corollary 4.7] the right hand side converges to $h(x)\int_{\mathbb{R}^d} h(x)dx$. □

Here we would like to make a remark that [64, Corollary 4.7] is proved by Corollary 11.6 in Section 11.

Theorem 9.1 implies (7.6). Indeed,

$$\begin{aligned}
\frac{\mathbb{E}_x\left(\exp(A_t^\mu)|\mathcal{F}_s\right)}{\mathbb{E}_x\left(\exp(A_t^\mu)\right)} &= \frac{e^{-\theta_0 t}\mathbb{E}_x\left(\exp(A_t^\mu)|\mathcal{F}_s\right)}{e^{-\theta_0 t}\mathbb{E}_x\left(\exp(A_t^\mu)\right)} \\
&= \frac{e^{-\theta_0 s}\exp(A_s^\mu)e^{-\theta_0(t-s)}\mathbb{E}_{X_s}\left(\exp(A_{t-s}^\mu)\right)}{e^{-\theta_0 t}\mathbb{E}_x\left(\exp(A_t^\mu)\right)} \\
\longrightarrow \quad & \frac{e^{-\theta_0 s}\exp(A_s^\mu)h(X_s)\int_{\mathbb{R}^d} h(x)dx}{h(x)\int_{\mathbb{R}^d} h(x)dx} = L_s^h \quad as \quad t \longrightarrow \infty.
\end{aligned}$$

We showed in [14, Theorem 2.6 (b)] that the transformed process $\mathbf{M}^h$ is recurrent. We see from this fact that L_t^h is martingale, $\mathbb{E}(L_t^h) = 1$. Therefore Scheff's lemma leads us to Theorem 7.1 (i) (e.g. [49]).

(2°) **Transient case** ($d > \alpha$)

If $\lambda(0) < 1$, there exist $\theta_0 > 0$ and $h \in \mathcal{D}(\mathcal{E})$ such that $\lambda(\theta_0) = 1$ and $\mathcal{E}_{\theta_0}(h,h) = 1$. Then we can show the equation (9.1) in the same way as above. If $\lambda(0) > 1$, then A_t^μ is gaugeable (see Theorem 9.4 below), that is,

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}_x\left(e^{A_\infty^\mu}\right) < \infty,$$

and thus

$$\lim_{t\to\infty} \mathbb{E}_x\left(e^{A_t^\mu}\right) = \mathbb{E}_x\left(e^{A_\infty^\mu}\right).$$

Hence for any $s \ge 0$ and any $\mathcal{F}_s$-measurable bounded function Z

$$\frac{\mathbb{E}_x\left(Ze^{A_t^\mu}\right)}{\mathbb{E}_x\left(e^{A_t^\mu}\right)} = \frac{\mathbb{E}_x\left(Ze^{A_s^\mu}\mathbb{E}_{X_s}\left(e^{A_{t-s}^\mu}\right)\right)}{\mathbb{E}_x\left(e^{A_t^\mu}\right)}$$

$$\longrightarrow \frac{\mathbb{E}_x\left(Ze^{A_s^\mu}\mathbb{E}_{X_s}\left(e^{A_\infty^\mu}\right)\right)}{\mathbb{E}_x\left(e^{A_\infty^\mu}\right)} = \frac{1}{h(x)}\mathbb{E}_x\left(Ze^{A_s^\mu}h(X_s)\right) = \mathbb{E}_x^h(Z)$$

as $t \to \infty$.

In the remainder of this section, we consider the case when $\lambda(0) = 1$. It is known that a measure $\mu \in K_{d,\alpha}^\infty$ is Green-bounded,

$$\sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{d\mu(y)}{|x-y|^{d-\alpha}} < \infty. \tag{9.2}$$

To consider the penalization problem for μ with $\lambda(0) = 1$, we need to impose a condition on μ.

Definition 9.2. (I) A measure $\mu \in K_{d,\alpha}$ is said to be *special* if

$$\sup_{x\in\mathbb{R}^d}\left(|x|^{d-\alpha}\int_{\mathbb{R}^d}\frac{d\mu(y)}{|x-y|^{d-\alpha}}\right) < \infty. \tag{9.3}$$

We denote by K_S^∞ the set of special measures.
(II) A PCAF A_t is said to be *special* with respect to $\mathbf{M}^h$, if for any positive Borel function g with $\int_{\mathbb{R}^d} g dx > 0$

$$\sup_{x\in\mathbb{R}^d}\mathbb{E}_x^h\left(\int_0^\infty \exp\left(-\int_0^t g(X_s)ds\right)dA_t\right) < \infty.$$

A Kato measure with compact support belongs to K_S^∞. The set K_S^∞ is contained in $K_{d,\alpha}^\infty(\beta)$,

$$K_S^\infty \subset K_{d,\alpha}^\infty. \tag{9.4}$$

Indeed, since for any $R > 0$

$$M(\mu) := \sup_{x\in\mathbb{R}^d}\left(|x|^{d-\alpha}\int_{\mathbb{R}^d}\frac{d\mu(y)}{|x-y|^{d-\alpha}}\right) \geq R^{d-\alpha}\sup_{x\in B(R)^c}\int_{\mathbb{R}^d}\frac{d\mu(y)}{|x-y|^{d-\alpha}},$$

we have

$$\begin{aligned}\sup_{x\in\mathbb{R}^d}\int_{B(R)^c}\frac{d\mu(y)}{|x-y|^{d-\alpha}} &= \sup_{x\in B(R)^c}\int_{B(R)^c}\frac{d\mu(y)}{|x-y|^{d-\alpha}} \\ &\leq \frac{M(\mu)}{R^{d-\alpha}} \longrightarrow 0, \quad R\to\infty.\end{aligned}$$

Lemma 9.3. *Let B_t be a PCAF. Then*

$$\mathbb{E}_x\left(\int_0^\infty e^{(A_t^\mu - B_t)}dA_t^\mu\right) = h(x)\mathbb{E}_x^h\left(\int_0^\infty e^{-B_t}\frac{dA_t^\mu}{h(X_t)}\right).$$

Proof. We have

$$\begin{aligned} h(x)\mathbb{E}_x^h\left(\int_0^s e^{-B_t}\frac{dA_t^\mu}{h(X_t)}\right) &= \mathbb{E}_x\left(e^{A_s^\mu}h(X_s)\int_0^s e^{-B_t}\frac{dA_t^\mu}{h(X_t)}\right) \\ &= \mathbb{E}_x\left(\int_0^s e^{A_s^\mu}h(X_s)e^{-B_t}\frac{dA_t^\mu}{h(X_t)}\right). \end{aligned}$$

Put $Y_t = e^{A_s^\mu}h(X_s)e^{-B_t}/h(X_t)$. Then since Y_t is a right continuous process, its optional projection is equal to $\mathbb{E}_x(Y_t|\mathcal{F}_t)$ (e.g. [48, Theorem 7.10]). Hence the right hand side equals

$$\mathbb{E}_x\left(\int_0^s \mathbb{E}_x\left(Y_t|\mathcal{F}_t\right)dA_t^\mu\right) = \mathbb{E}_x\left(\int_0^s e^{A_t^\mu}e^{-B_t}\frac{1}{h(X_t)}\mathbb{E}_{X_t}\left(e^{A_{s-t}^\mu}h(X_{s-t})\right)dA_t^\mu\right).$$

Since $\mathbb{E}_{X_t}\left(e^{A_{s-t}^\mu}h(X_{s-t})\right) = h(X_t)$, the right hand side equals

$$\mathbb{E}_x\left(\int_0^s e^{A_t^\mu - B_t}dA_t^\mu\right).$$

Hence the proof is completed by letting $s \to \infty$. □

The next theorem is an extension of Theorem 5.1 to the symmetric α-stable process. For more general Markov processes, see [13], [61].

Theorem 9.4. ([73]) *Suppose $d > \alpha$. For $\mu = \mu^+ - \mu^- \in K_{d,\alpha}^\infty - K_{d,\alpha}^\infty$, let $A_t^\mu = A_t^{\mu^+} - A_t^{\mu^+}$. Then the following conditions are equivalent*:

(i) $\sup_{x\in\mathbb{R}^d} \mathbb{E}_x(e^{A_\infty^\mu}) < \infty$.

(ii) *There exists the Green function $R^\mu(x,y) < \infty$ $(x \neq y)$ of the operator $-\frac{1}{2}(-\Delta)^{\alpha/2} + \mu$ such that*

$$\mathbb{E}_x\left(\int_0^\infty e^{A_t^\mu}f(X_t)dt\right) = \int_{\mathbb{R}^d} R^\mu(x,y)f(y)dy.$$

(iii) $\inf\left\{\mathcal{E}(u,u) + \int_{\mathbb{R}^d} u^2 d\mu^- : \int_{\mathbb{R}^d} u^2 d\mu^+ = 1\right\} > 1.$

We see from (4.19) in [71] that if one of the statements in Theorem 9.4 holds, then $R^\mu(x,y)$ satisfies

$$R(x,y) \leq R^\mu(x,y) \leq CR(x,y). \tag{9.5}$$

Lemma 9.5. *If $\mu \in K_S^\infty$, then $\int_0^t \frac{dA_s^\mu}{h(X_s)}$ is special with respect to $\mathbf{M}^h$.*

Proof. We may assume that g is a bounded positive Borel function with compact support. Note that by Lemma 9.3

$$\begin{aligned}
& \mathbb{E}_x^h\left(\int_0^\infty \exp\left(-\int_0^t g(X_s)ds\right)\frac{dA_t^\mu}{h(X_t)}\right) \\
= & \frac{1}{h(x)}\mathbb{E}_x\left(\int_0^\infty \exp\left(A_t^\mu - \int_0^t g(X_s)ds\right)dA_t^\mu\right) \\
= & \frac{1}{h(x)}R^{\mu - g\cdot dx}\mu(x).
\end{aligned}$$

If the measure μ satisfies $\lambda(0)=1$, then $\mu - g\cdot dx \in K_{d,\alpha}^\infty - K_{d,\alpha}^\infty$ satisfies Theorem 9.4 (iii), and $R^{\mu-g\cdot dx}(x,y)$ is equivalent with $R(x,y)$ by (9.5). Therefore the inequality (8.6) implies that (9.3) is equivalent to that $\sup_{x\in\mathbb{R}^d}\left\{(1/h(x))R^{\mu-g\cdot dx}\mu(x)\right\}<\infty$. □

We note that by Lemma 9.3

$$\mathbb{E}_x\left(e^{A_t^\mu}\right) = 1 + \mathbb{E}_x\left(\int_0^t e^{A_s^\mu}dA_s^\mu\right) = 1 + h(x)\mathbb{E}_x^h\left(\int_0^t \frac{dA_s^\mu}{h(X_s)}\right).$$

Thus for a finite positive measure ν,

$$\mathbb{E}_\nu\left(e^{A_t^\mu}\right) = \nu(\mathbb{R}^d) + \langle \nu, h\rangle \mathbb{E}_{\nu^h}^h\left(\int_0^t \frac{dA_s^\mu}{h(X_s)}\right) \tag{9.6}$$

where $\nu^h = h\cdot\nu/\langle\nu,h\rangle$. For a positive smooth function k with compact support, put

$$\psi(t) = \mathbb{E}_x^h\left(\int_0^t k(X_s)ds\right).$$

Then $\lim_{t\to\infty}\psi(t)=\infty$ by the Harris recurrence of $\mathbf{M}^h$. Moreover,

$$\lim_{t\to\infty}\frac{\psi(t+s)}{\psi(t)} = 1. \tag{9.7}$$

Indeed,

$$\begin{aligned}
\psi(t+s) &= \mathbb{E}_x^h\left(\int_0^t k(X_u)du\right) + \mathbb{E}_x^h\left(\mathbb{E}_{X_t}^h\left(\int_0^s k(X_u)du\right)\right) \\
&\le \psi(t) + \|k\|_\infty s,
\end{aligned}$$

and

$$1 \le \frac{\psi(t+s)}{\psi(t)} \le 1 + \frac{\|k\|_\infty s}{\psi(t)}.$$

We see from [24, Lemma 4.4] that the Revuz measure of A_t^μ is $h^2\mu$ as a PCAF of $\mathbf{M}^h$. Since by (9.6)

$$\frac{1}{\psi(t)}\mathbb{E}_\nu\left(e^{A_t^\mu}\right) = \frac{\nu(\mathbb{R}^d)}{\psi(t)} + \langle\nu,h\rangle\frac{\mathbb{E}_{\nu^h}^h\left(\int_0^t (1/h(X_s))dA_s^\mu\right)}{\mathbb{E}_x^h\left(\int_0^t k(X_s)ds\right)}$$

and $\int_0^t (1/h(X_s))dA_s^\mu$ and $\int_0^t k(X_s)ds$ are special with respect to $\mathbb{M}^h$, we see from Chacon-Ornstein type ergodic theorem in [9, Theorem 3.18] that

$$\frac{1}{\psi(t)}\mathbb{E}_\nu\left(e^{A_t^\mu}\right) \longrightarrow \langle \nu, h\rangle \cdot \frac{\langle \mu, h\rangle}{\int_{\mathbb{R}^d} kh^2 dx} \tag{9.8}$$

as $t \to \infty$. Note that $\langle \mu, h\rangle < \infty$ by (8.6) and (9.2).

For a bounded $\mathcal{F}_s$-measurable function Z, define a positive finite measure ν by

$$\nu(B) = \mathbb{E}_x\left(Ze^{A_s^\mu}; X_s \in B\right), \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Then by the Markov property,

$$\mathbb{E}_x\left(Ze^{A_t^\mu}\right) = \mathbb{E}_\nu\left(e^{A_{t-s}^\mu}\right).$$

Therefore

$$\begin{aligned}\lim_{t\to\infty}\frac{\mathbb{E}_x\left(Ze^{A_t^\mu}\right)}{\mathbb{E}_x\left(e^{A_t^\mu}\right)} &= \lim_{t\to\infty}\frac{\mathbb{E}_x\left(Ze^{A_t^\mu}\right)/\psi(t)}{\mathbb{E}_x\left(e^{A_t^\mu}\right)/\psi(t)} \\ &= \lim_{t\to\infty}\frac{(\psi(t-s)/\psi(t))\mathbb{E}_\nu\left(e^{A_{t-s}^\mu}\right)/\psi(t-s)}{\mathbb{E}_x\left(e^{A_t^\mu}\right)/\psi(t)}.\end{aligned}$$

By (9.7) and (9.8), the right hand side equals

$$\frac{(\langle \nu, h\rangle\langle \mu, h\rangle)/\int_{\mathbb{R}^d} kh^2dx}{(h(x)\langle \mu, h\rangle)/\int_{\mathbb{R}^d} kh^2dx} = \frac{\langle \nu, h\rangle}{h(x)} = \frac{1}{h(x)}\mathbb{E}_x\left(Ze^{A_s^\mu}h(X_s)\right) = \mathbb{E}_x^h(Z). \tag{9.9}$$

Remark 9.1. We suppose that $d > \alpha$ and $\lambda(0) = 1$. If $d > 2\alpha$, then $h \in L^2(\mathbb{R}^d)$ on account of (8.6). Hence $\mathbf{M}^h$ is an ergodic process with the invariant probability measure h^2dx, and thus for a smooth function k with compact support,

$$\frac{\psi(t)}{t} = \frac{1}{t}\mathbb{E}_x^h\left(\int_0^t k(X_s)ds\right) \longrightarrow \int_{\mathbb{R}^d} kh^2dx.$$

Hence we see that for $\mu \in K_S^\infty$

$$\lim_{t\to\infty}\frac{1}{t}\mathbb{E}_x\left(e^{A_t^\mu}\right) = h(x)\langle \mu, h\rangle. \tag{9.10}$$

10 Stability of heat kernels

Let $\mathbf{M}^\alpha = (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}_x, X_t)$ be the symmetric α-stable process on $\mathbb{R}^d$ with $\alpha < d$. In this section, we consider the stability of heat kernels of Schrödinger operators as another application of Theorem 5.1. We denote the potential of a positive measure μ by

$$R\mu(x) = \int_{\mathbb{R}^d} R(x,y)d\mu(y),$$

where R is the Riesz kernel defined in (5.8). Next lemma is shown in ([28, Example 2.2.1]); however, we give another proof by using 0-order version of the equality (2.2).

Lemma 10.1. ([28, Example 2.2.1]) *If* $\mu \in K_{d,\alpha}^{\infty}$ *satisfies* $\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} R(x,y)d\mu(x)d\mu(y) < \infty$, *then* $R\mu$ *belongs to* $\mathcal{D}_e(\mathcal{E})$.

Proof. First note that for $\mu \in K_{d,\alpha}^{\infty}$

$$\int_{\mathbb{R}^d} u^2 d\mu \leq \|R\mu\|_\infty \mathcal{E}(u,u), \quad u \in \mathcal{D}_e(\mathcal{E}) \tag{10.1}$$

(cf. [55]). Let K be a compact set of $\mathbb{R}^d$. Then by applying (10.1) to $\mu_K(\cdot) = \mu(K \cap \cdot)$, we have

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi d\mu_K &\leq (\mu(K))^{1/2}\left(\int_{\mathbb{R}^d} \varphi^2 d\mu_K\right)^{1/2} \\ &\leq (\mu(K))^{1/2}\|R\mu_K\|_\infty^{1/2}\mathcal{E}(\varphi,\varphi)^{1/2}.\end{aligned}$$

Hence the measure μ_K is of finite energy integral, and thus

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi d\mu_K &\leq \mathcal{E}(R\mu_K, R\mu_K)^{1/2}\mathcal{E}(\varphi,\varphi)^{1/2} \\ &\leq \left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} R(x,y)d\mu_K(x)d\mu_K(y)\right)^{1/2}\mathcal{E}(\varphi,\varphi)^{1/2}.\end{aligned}$$

By letting K increase to $\mathbb{R}^d$, we find that μ is of finite energy integral, and thus $R\mu$ is in $\mathcal{D}_e(\mathcal{E})$. □

Lemma 10.1 says that μ is of finite energy integral. Assume that

$$\lambda(\mu) := \inf\left\{\mathcal{E}(u,u) : \int_{\mathbb{R}^d} u^2 d\mu = 1\right\} > 1 \tag{10.2}$$

and set

$$h(x) = \mathbb{E}_x(e^{A_\infty^\mu}). \tag{10.3}$$

Then by Theorem 9.4, $1 \leq h(x) \leq M(= \sup_x \mathbb{E}_x(e^{A_\infty^\mu})) < \infty$.

Lemma 10.2. *Assume that* $\lambda(\mu) > 1$. *Then it holds that*

$$h(x) = R(h\mu)(x) + 1.$$

Proof. Define $M_t := \mathbb{E}_x(\exp(A_\infty^\mu)|\mathcal{F}_t)$. Then by the Markov property

$$\begin{aligned}h(X_t) &= \mathbb{E}_{X_t}(\exp(A_\infty^\mu)) = \mathbb{E}_x(\exp(A_\infty^\mu(\theta_t))|\mathcal{F}_t) \\ &= \mathbb{E}_x(\exp(A_\infty^\mu - A_t^\mu)|\mathcal{F}_t) = \exp(-A_t^\mu)M_t,\end{aligned}$$

and thus

$$\begin{aligned}\mathbb{E}_x\left(\int_0^t h(X_s)dA_s^\mu\right) &= \mathbb{E}_x\left(\int_0^t \exp(-A_s^\mu)M_s dA_s^\mu\right)\\ &= \mathbb{E}_x(M_0) - \mathbb{E}_x(\exp(-A_t^\mu)M_t) + \mathbb{E}_x\left(\int_0^t \exp(-A_s^\mu)dM_s\right) \qquad (10.4)\\ &= h(x) - \mathbb{E}_x(h(X_t)).\end{aligned}$$

Noting that

$$\lim_{t\to\infty} h(X_t) = \lim_{t\to\infty} \exp(-A_t^\mu)M_t = \exp(-A_\infty^\mu)\exp(A_\infty^\mu) = 1,$$

we have the lemma by letting t to ∞ in (10.4). □

Suppose that $\mu \in K_{d,\alpha}^\infty$ satisfies $\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} R(x,y)d\mu(x)d\mu(y) < \infty$. Then by Lemma 10.1, $R(h\mu)$ belongs to $\mathcal{D}_e(\mathcal{E})$. Thus by Fukushima decomposition

$$R(h\mu)(X_t) - R(h\mu)(X_0) = M_t^{[R(h\mu)]} + N_t^{[R(h\mu)]}.$$

We put $M_t^{[h]} = M_t^{[R(h\mu)]}$ and define the multiplicative functional L_t^h by the unique solution of (8.4). We then see from [17] that the Dirichlet form generated by the transformed process $\mathbf{M}^h$ by L_t^h is identified:

Theorem 10.3. ([17]) *The transformed process $\mathbf{M}^h$ is an h^2dx-symmetric and its Dirichlet form $(\mathcal{E}^h, \mathcal{D}(\mathcal{E}^h)$ on $L^2(\mathbb{R}^d, h^2dx)$ is*

$$\begin{cases} \mathcal{E}^h(u,v) = K \displaystyle\iint_{\mathbb{R}^d\times\mathbb{R}^d\setminus d} \frac{(u(x)-u(y))^2 h(x)h(y)}{|x-y|^{d+\alpha}}dxdy,\\ \mathcal{D}(\mathcal{E}^h) = \mathcal{D}(\mathcal{E}^{(\alpha)}).\end{cases}$$

In [17, Section 3], they treat a class of functions $h = e^u$ such that u is a bounded function in $\mathcal{D}_e(\mathcal{E})$. Since $\log(1+R(h\mu)) \in \mathcal{D}_e(\mathcal{E})$, the function h in (10.3) belongs to the class treated in [17] and Theorem 10.3 can be applied.

We denote by $p_t^\mu(x,y)$ the heat kernel of $(-\Delta)^{\alpha/2} - \mu$:

$$\mathbb{E}_x\left(e^{A_t^\mu} f(X_t)\right) = \int_{\mathbb{R}^d} p_t^\mu(x,y)f(y)dy.$$

The notation $p_t^\mu(x,y) \simeq \dfrac{1}{t^{d/\alpha}} \wedge \dfrac{t}{|x-y|}$ means that there exist positive constants c, C such that

$$c\left(\frac{1}{t^{d/\alpha}} \wedge \frac{t}{|x-y|}\right) \le p_t^\mu(x,y) \le C\left(\frac{1}{t^{d/\alpha}} \wedge \frac{t}{|x-y|}\right).$$

Theorem 10.4. *Let $\mu \in K_\infty^{d,\alpha}$ with $\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|^{\alpha-d} d\mu(x)d\mu(y) < \infty$. Then*

$$\lambda(\mu) > 1 \Longleftrightarrow p_t^\mu(x,y) \simeq \frac{1}{t^{d/\alpha}} \wedge \frac{t}{|x-y|}. \tag{10.5}$$

Proof. ($\Longleftarrow$) By the assumption, $p^\mu(t,x,y) \le \frac{c}{t^{d/\alpha}}$, and thus

$$R^\mu(x,y) \le \int_0^1 p^\mu(t,x,y)dt + c\int_1^\infty \frac{1}{t^{d/\alpha}} dt < \infty.$$

Hence Theorem 9.4 says that $\lambda(\mu) > 1$.

($\Longrightarrow$) Denote $m(dx) = h^2(x)dx$. Then the Dirichlet form $\mathcal{E}^h$ is written as

$$\mathcal{E}^h(u,v) = \iint_{\mathbb{R}^d\times\mathbb{R}^d\setminus d} \frac{(u(x)-u(y))^2 c(x,y)}{|x-y|^{d+\alpha}} m(dx)m(dy).$$

Here $0 < c \le c(x,y) = 1/(h(x)h(y)) \le C < \infty$. Let $p_t^h(x,y)$ be the heat kernel of $\mathbf{M}^h$, $\mathbb{E}_x^h(f(X_t)) = \int_{\mathbb{R}^d} p_t^h(x,y)f(y)h^2(y)dy$. We then see from [16] that the heat kernel $p_t^h(x,y)$ satisfies

$$p_t^h(x,y) \simeq \frac{1}{t^{d/\alpha}} \wedge \frac{t}{|x-y|}.$$

Since $p_t^\mu(x,y) = h(x)p_t^h(x,y)h(y)$ by the definition of $\mathbf{M}^h$, the proof is completed. □

The heat kernel of $\mathbf{M}^\alpha$ satisfies the right hand side of (10.5). Theorem 10.4 says that the measure μ satisfying the conditions in Theorem 10.4 is so small that it does not cause an essential change of the heat kernel. We can consider the same probelem for diffusion processes.

Let M be a complete, non-compact Riemannian manifold. Let $d(x,y)$ be the geodesic distance and m the Riemannian volume. Let $p(t,x,y)$ the heat kernel associated with half the Laplace-Beltrami operator $\frac{1}{2}\Delta$. Assume that $p(t,x,y)$ satisfies global Gaussian lower and upper bounds: for every $x,y \in M$ and $t > 0$,

$$\frac{C_1 \exp\left(-c_1\frac{d^2(x,y)}{t}\right)}{m(B(x,\sqrt{t}))} \le p(t,x,y) \le \frac{C_2 \exp\left(-c_2\frac{d^2(x,y)}{t}\right)}{m(B(x,\sqrt{t}))}, \tag{10.6}$$

where C_1, c_1, C_2, and c_2 are positive constants and $B(x,r)$ is the geodesic ball of radius r centered at the point $x \in M$. Following [32], we say that the heat kernel $p(t,x,y)$ satisfies the *Li-Yau estimate*, if it has the estimate (10.6). For a measure μ in a certain class, let $p^\mu(t,x,y)$ be the heat kernel associated with the Schrödinger operator, $\frac{1}{2}\Delta + \mu$. We establish, in the same manner as above, a necessary and sufficient condition on the potential μ for the heat kernel $p^\mu(t,x,y)$ also to satisfy the Li-Yau estimate ([62]): a positive Radon measure μ on M is said to be in the class $\mathcal{S}_\infty$, if for any $\epsilon > 0$ there exists a compact set $K \subset M$ and $\delta > 0$ such that

$$\sup_{(x,z)\in M\times M\setminus\Delta} \int_{K^c} \frac{G(x,y)G(y,z)}{G(x,z)}\mu(dy) \leq \epsilon$$

and for all measurable sets $B \subset K$ with $\mu(B) < \delta$,

$$\sup_{(x,z)\in M\times M\setminus\Delta} \int_{B} \frac{G(x,y)G(y,z)}{G(x,z)}\mu(dy) \leq \epsilon.$$

When μ is absolutely continuous with respect to m, the class $\mathcal{S}_\infty$ was essentially introduced by P. Pinchover, M. Murata, and the function in $\mathcal{S}_\infty$ is said to be *G-small at infinity*. The above definition is due to Z.-Q. Chen [13]. A Kato measure with compact support belongs to the class $\mathcal{S}_\infty$. For $\mu \in \mathcal{S}_\infty$, denote by $p^\mu(t,x,y)$ the heat kernel of the Schrödinger operator $\frac{1}{2}\Delta + \mu$.

Theorem 10.5. ([62]) *Suppose $\mu \in \mathcal{S}_\infty$. Then $p^\mu(t,x,y)$ satisfies the Li–Yau estimate if and only if $\lambda(\mu) > 1$.*

11 Ergodic properties of Dirichlet forms

Every symmetric Markov process generated a regular Dirichlet space can be transformed to an ergodic one. This fact is crucial for the proof of the lower bound in the large deviation principle (Theorem 2.2 (i)). In addition, an operator theoretical ergodic theorem (Theorem 9.1) is crucial in Feynman-Kac penalization problem. In this section we summarize the ergodic theory for symmetric Markov processes.

Take $\phi = R_\alpha g \in \mathcal{D}^+(A)$ and fix it hereafter. Let $\mathbf{M}^\phi = (\Omega, X_t, \mathbb{P}^\phi_x, \zeta)$ the transformed process of $\mathbf{M}$ by L^ϕ_t in (2.3). We then see from Lemma 6.3.2 in [28] and Theorem 62.5 in [52] that $\mathbf{M}^\phi$ is a $\phi^2 m$-symmetric Markov process on X. Denote by $(\mathcal{E}^\phi, \mathcal{D}(\mathcal{E}^\phi))$ the Dirichlet form on $L^2(X;\phi^2 m)$ associated with $\mathbf{M}^\phi$. It is known that the Dirichlet form $\mathcal{E}$ has the Beurling-Deny decomposition: for $u \in \mathcal{D}(\mathcal{E})$

$$\mathcal{E}(u,u) = \frac{1}{2}\int_X d\mu^c_{\langle u\rangle} + \iint_{X\times X-d} (u(x)-u(y))^2 J(dx,dy) + \int_X u^2 dk.$$

Theorem 11.1. ([14]) *The Dirichlet space $\mathcal{D}(\mathcal{E})$ is included in $\mathcal{D}(\mathcal{E}^\phi)$ and for $u \in \mathcal{D}(\mathcal{E})$*

$$\mathcal{E}^\phi(u,u) = \frac{1}{2}\int_X \phi^2 d\mu^c_{\langle u\rangle} + \iint_{X\times X-d} (u(x)-u(y))^2 \phi(x)\phi(y) J(dx,dy). \tag{11.1}$$

Moreover, the identity function 1 belongs to $\mathcal{D}(\mathcal{E}^\phi)$ and $\mathcal{E}^\phi(1,1) = 0$.

Theorem 11.2. *The transformed process $\mathbf{M}^\phi$ is conservative, $p_t^\phi 1(x) = 1$ for any $t > 0$ and any $x \in X$, where p_t^ϕ is the semigroup of $\mathbf{M}^\phi$. Moreover, $\mathbf{M}^\phi$ is ergodic in the sense that if $\Lambda \in \mathcal{F}$ is θ_t-invariant, $(\theta_t)^{-1}(\Lambda) = \Lambda$, then $\mathbb{P}^\phi_{\phi^2 m}(\Lambda) = 0$ or $\mathbb{P}^\phi_{\phi^2 m}(\Omega \setminus \Lambda) = 0$.*

For the proof of Theorem 11.2, we need the next general theorem taken from [30]. In the sequel, we assume that m is finite, $m(X) < \infty$.

Theorem 11.3. ([30]) *Suppose $m(X) < \infty$. Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be the Dirichlet form associated with an m-symmetric right process $(\Omega, \mathcal{F}, \mathcal{F}_t, X_t, \mathbb{P}_x, \theta_t)$ on X and suppose that*

$$1 \in \mathcal{D}(\mathcal{E}) \quad \text{and} \quad \mathcal{E}(1,1) = 0.$$

Then, the following statements are equivalent each other.
(i) $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ *is irreducible.*
(ii) *If a function $u \in \mathcal{D}(\mathcal{E})$ satisfies $\mathcal{E}(u,u) = 0$, then u is constant m-a.e.*
(iii) *If a function $u \in L^2(X;m)$ satisfies $p_t u = u$ m-a.e. for any $t > 0$, then u is constant m-a.e.*
(iv) *$(\Omega, \mathbb{P}_m, \mathcal{F}, \theta_t)$ is ergodic, i.e., if $\Lambda \in \mathcal{F}$ is θ_t-invariant, then $\mathbb{P}_m(\Lambda) = 0$ or $\mathbb{P}_m(\Omega \setminus \Lambda) = 0$.*

Proof. ((i)$\Longrightarrow$ (ii)) Let u be a function in $\mathcal{F}$ with $\mathcal{E}(u,u) = 0$. For $\lambda \in \mathbb{R}$, let $u_\lambda = (u - \lambda)_+$. Since $\mathcal{E}(u_\lambda, u_\lambda) \le \mathcal{E}(u - \lambda, u - \lambda) = 0$,

$$\mathcal{E}(u_\lambda, v) = 0 \qquad \text{for } \forall v \in \mathcal{F},$$

and so $Au_\lambda = 0$ ($p_t u_\lambda = u_\lambda$). Set $B_\lambda = \{x \in X : u_\lambda(x) = 0\}$. Noting that

$$p_t(1_{B_\lambda^c} u_\lambda) = p_t(u_\lambda) = u_\lambda = 0, m\text{-a.e. on } B_\lambda,$$

we have for any n

$$p_t(1_{B_\lambda^c} 1_{\{u_\lambda \ge \frac{1}{n}\}}) = 0 \quad \text{on } B_\lambda.$$

Thus, $p_t(1_{B_\lambda^c}) = 0$ m-a.e. on B_λ ($1_{B_\lambda} p_t(1_{B_\lambda^c}) = 0$). By the symmetry, $1_{B_\lambda^c} p_t(1_{B_\lambda}) = 0$. Therefore, for $f \in L^2(X;m)$

$$\begin{aligned} p_t(1_{B_\lambda} f) &= 1_{B_\lambda} p_t(1_{B_\lambda} f) + 1_{B_\lambda^c} p_t(1_{B_\lambda} f) \\ &= 1_{B_\lambda} p_t(1_{B_\lambda} f) \\ &= 1_{B_\lambda} p_t(1_{B_\lambda} f) + 1_{B_\lambda} p_t(1_{B_\lambda^c} f) \\ &= 1_{B_\lambda} p_t(f), \end{aligned}$$

and $m(B_\lambda) = 0$ or $m(B_\lambda^c) = 0$ by the assumption. Let $\lambda_0 = \sup\{\lambda : m(B_\lambda) = 0\}$. Then, for any $\lambda > \lambda_0$, $m(B_\lambda) \ne 0$, which implies $m(B_\lambda^c) = 0$. Hence $m(\{u > \lambda_0\}) = 0$. On the other hand, for any $\lambda < \lambda_0$, $m(B_\lambda) = 0$ and so $m(\{u < \lambda_0\}) = 0$. Therefore, we can conclude that $u = \lambda_0$ m-a.e.

((ii)$\Longrightarrow$ (iii)) and ((iii)$\Longrightarrow$ (i)) are trivial.

((iii)$\Longrightarrow$ (iv)) Let $F \in L^2(\mathbb{P}_m)$ with $F = F \circ \theta_t$ $\mathbb{P}_m$-a.e. for $\forall t > 0$. Put $f(x) = \mathbb{E}_x(F)$. Note that

$$\begin{aligned}\mathbb{E}_m(F \circ \theta_t) &= \mathbb{E}_m(\mathbb{E}_{X_t}(F)) = \mathbb{E}_m(f(X_t)) \\ &= \mathbb{E}_m(p_t f(X_0)).\end{aligned}$$

Hence for any bounded Borel function g on X,

$$\begin{aligned}0 &= \mathbb{E}_m((F \circ \theta_t - F)g(X_0)) = \mathbb{E}_m((p_t f(X_0) - f(X_0))g(X_0)) \\ &= (p_t f - f, g)_m,\end{aligned}$$

and thus $p_t f = f$ m-a.e. and $f = k$ (constant) m-a.e. by assumption. Therefore

$$k = f(X_n) = \mathbb{E}_{X_n}(F) = \mathbb{E}_m(F \circ \theta_n | \mathcal{F}_n) = \mathbb{E}_m(F|\mathcal{F}_n) \longrightarrow F, \ \mathbb{P}_m\text{-a.e. } n \longrightarrow \infty.$$

((iv)$\Longrightarrow$(iii)) Let f be an L^2-function such that $f = p_t f$. We set

$$\Omega_f = \left\{\omega \in \Omega : \int_0^T |f(X_t)| dt < \infty \ \text{ for } \forall T \in (0, \infty)\right\}$$

and define

$$F_T(\omega) = \begin{cases} \frac{1}{T}\int_0^T f(X_t(\omega))dt & \text{if } \omega \in \Omega_f \\ 0 & \text{if } \omega \notin \Omega_f. \end{cases}$$

Then, Ω_f^c is θ_t-invariant and thus $\mathbb{P}_m(\Omega_f^c) = 0$. Since for any $\varphi \in L^2(X; m)$

$$\begin{aligned}(f, \varphi)_m &= \left(\frac{1}{T}\int_0^T p_t f dt, \varphi\right)_m = (\mathbb{E}_x(F_T), \varphi)_m \\ &= \mathbb{E}_m(F_T \varphi(X_0)).\end{aligned}$$

Since

$$F_T = \frac{1}{T}\int_0^T f(X_0 \circ \theta_t)dt \to \frac{\mathbb{E}_m(f(X_0))}{m(X)} = \frac{\int_X f dm}{m(X)}$$

by the ergodic Theorem,

$$\begin{aligned}(f, \varphi)_m &= \lim_{T\to\infty} \mathbb{E}_m(F_T \varphi(X_0)) \\ &= \mathbb{E}_m\left(\frac{\int_X f dm}{m(X)}\varphi(X_0)\right) = \frac{\int_X f dm}{m(X)}\int_X \varphi dm,\end{aligned}$$

and thus $f = \int_X f dm / m(X)$, m-a.e. $\square$

Theorem 11.3 is extended to the case when $m(X) = \infty$ ([15], [39]).

Lemma 11.4. *A $\mathbb{P}_m$-integrable bounded random variable Z is θ_t-invariant ($Z = Z \circ \theta_t$, $\mathbb{P}_m$-a.e., $t > 0$) if and only if $Z = g(X_0)$ $\mathbb{P}_m$-a.e. for some m-integrable bounded function g on X which is p_t-invariant, $p_t g = g$ m-a.e., $t > 0$.*

Proof. Since $\mathbb{E}_m(Z|\mathcal{F}_t) = \mathbb{E}_m(Z \circ \theta_t|\mathcal{F}_t) = g(X_t)$ P_m-a.e. with $g(x) = \mathbb{E}_x(Z)$ by the Markov property,

$$\begin{aligned}\mathbb{P}_m(|Z - g(X_0)| > \epsilon) &= \mathbb{P}_m(|Z \circ \theta_t - g(X_0 \circ \theta_t)| > \epsilon) = \mathbb{P}_m(|Z - g(X_t)| > \epsilon) \\ &= \mathbb{P}_m(|Z - \mathbb{E}_m(Z|\mathcal{F}_t)| > \epsilon) \longrightarrow 0 \quad \text{as} \quad t \to \infty.\end{aligned}$$

Hence $Z = g(X_0)$ $\mathbb{P}_m$-a.e. and so $g(X_t) = \mathbb{E}_m(g(X_0)|\mathcal{F}_t) = g(X_0)$ $\mathbb{P}_m$-a.e., and thus for any bounded function h on X

$$\begin{aligned}\mathbb{E}_m((g(X_t) - g(X_0))h(X_0)) &= \mathbb{E}_m((p_t g(X_0) - g(X_0))h(X_0)) \\ &= \int_X (p_t g(x) - g(x))h(x)dm = 0,\end{aligned}$$

which implies $p_t g = g$ m-a.e.

Conversely, for a p_t-invariant function g, $g(X_0)$ is θ_t-invariant because

$$\mathbb{E}_m((g(X_t) - g(X_0))^2) = \mathbb{E}_m((p_t g(X_0) - g(X_0))^2) = 0.$$

□

On account of the symmetry, we have the ergodic theorem due to Fukushima [27].

Theorem 11.5. *Assume $m(X) < \infty$. For any $f \in L^p(X;m)$, $p \geq 1$, there exists a $L^p(X;m)$-function g such that*

$$\lim_{t \to \infty} p_t f = g, \quad m\text{-a.e. and in } L^p(X;m).$$

Moreover, g is p_t-invariant.

Proof. Let $\mathcal{G}_t = \sigma\{X_s : s \geq t\}$. By the symmetry, $Y_t := \mathbb{E}_m(f(X_0)|\mathcal{G}_t) = p_t f(X_t)$ $\mathbb{P}_m$-a.e. and thus $p_{2t} f(X_0) = \mathbb{E}_m(Y_t|\mathcal{F}_0)$. Since Y_t is an inverse martingale, the limit $\lim_{t\to\infty} Y_t = Z$ exists $\mathbb{P}_m$-a.e. and in $L^p(\mathbb{P}_m)$. Hence $\lim_{t\to\infty} p_{2t} f(X_0) = \mathbb{E}_m(Z|\mathcal{F}_0)$ $\mathbb{P}_m$-a.e. and in $L^p(\mathbb{P}_m)$. The theorem follows with $g(x) = \mathbb{E}_x(Z)$. □

Corollary 11.6. *Assume $m(X) < \infty$. If $\mathbf{M}$ is irreducible, then for $f \in L^p(X;m)$, $p \geq 1$*

$$\lim_{t \to \infty} p_t f(x) = \frac{1}{m(X)} \int_X f dm, \quad m\text{-a.e. and in } L^p(X;m).$$

Moreover, if for the conjugate number of $q (:= p/(p-1))$ the transition probability density satisfies $p_t(x, \cdot) \in L^q(X;m)$ for any x, then the limit holds for any $x \in X$.

Proof. For $g \in L^p(X;m)$, put $c_g = \frac{1}{m(X)} \int_X g dm$. Let $f^n = (-n \vee f) \wedge n$, $n = 1, 2, \cdots$. Then by combing Lemma 11.4 with Theorem 11.5,

$$\lim_{t \to \infty} p_t f^n(x) = c_{f^n}, \quad m\text{-a.e. and in } L^p(X;m).$$

Since

$$\begin{aligned} \|p_t f - c_f\|_p &\leq \|p_t f - p_t f^n\|_p + \|p_t f^n - c_{f^n}\|_p + \|c_{f^n} - c_f\|_p \\ &\leq 2\|f - f^n\|_p + \|p_t f^n - c_{f^n}\|_p, \end{aligned}$$

the first part of corollary follows.

The second part follows because by Hölder's inequality

$$\begin{aligned} &|p_t f(x) - c_f| = \left| \int_X p_1(x,y) \left(\int_X p_{t-1}(y,z) f(z) dm(z) - c_f \right) dm(y) \right| \\ &\leq \left(\int_X p_1^q(x,y) dm(y) \right)^{1/q} \left(\int_X \left| \int_X p_{t-1}(y,z) f(z) dm(z) - c_f \right|^p dm(y) \right)^{1/p} \\ &\longrightarrow 0 \quad \text{as } t \to \infty. \end{aligned}$$

□

References

[1] S. Albeverio, P. Blanchard and Z. M. Ma, Feynman-Kac semigroups in terms of signed smooth measures, Random partial differential equations (Oberwolfach, 1989), 1–31, Internat. Ser. Numer. Math. 102, Birkhäuser, Basel, 1991.

[2] S. Albeverio, F. Gesztesy, R. Hoegh-Krohn and H. Holden, Solvable Models in Quantum Mechanics, Texts and Monogr. Phys., Springer-Verlag, New York, 1988.

[3] S. Albeverio and Z. M. Ma, Perturbation of Dirichlet forms - Lower boundedness, closability, and form cores, J. Funct. Anal. 99(1991), 332–356.

[4] A. Beurling and J. Deny, Espaces de Dirichlet I, Le cas élémentaire, Acta Math. 99(1958), 203–224.

[5] A. Beurling and J. Deny, Dirichlet spaces, Proc. Nat. Acad. Sci. U.S.A. 45(1959), 208–215.

[6] P. Blanchard and Z. M. Ma, Semigroup of Schrödinger operators with potentials given by Radon measures, Stochastic processes, physics and geometry(Ascona and Locarno, 1988), 160–195, World Sci. Publ., Teaneck, NJ, 1990.

[7] R. M. Blumenthal and R. K. Getoor, Markov Processes and Potential Theory, Pure and Applied Mathematics, Vol. 29 Academic Press, New York-London 1968.

[8] N. Bouleau and F. Hirsch, Dirichlet forms and analysis on Wiener space, de Gruyter Stud. Math. 14, Walter de Gruyter, Berlin, 1991.

[9] M. Brancovan, Fonctionnelles additives spéciales des processus récurrents au sens de Harris, Z. Wahrsch. Verw. Gebiete. 47(1979), 163–194.

[10] J. Brasche, P. Exner, Yu. Kuperin and P. Šeba, Schrödinger operators with singular interactions, J. Math. Anal. Appl. 184(1994), 112–139.

[11] P. Cattiaux, P. Collet, A. Lamberrt, S. Martinez, S. Meleard and J. San Martin, Quasi-stationary distributions and diffusion models in population dynamics, Ann. Probab. 37(2009), 1926–1969.

[12] M.-F. Chen, Speed of stability for birth-death processes, Front. Math. China 5(2010), 379–515.

[13] Z.-Q. Chen, Gaugeability and Conditional Gaugeability, Trans. Amer. Math. Soc. 354(2002), 4639–4679.

[14] Z.-Q. Chen, P. J. Fitzsimmons, M. Takeda, J. Ying, and T.-S. Zhang, Absolute continuity of symmetric Markov processes, Ann. Probab. 32(2004), 2067–2098.

[15] Z.-Q. Chen and M. Fukushima, Symmetric Markov Processes, Time Change and Boundary Theory, Book manuscript, 2009.

[16] Z.-Q. Chen and T. Kumagai, Heat kernel estimates for jump processes of mixed types on metric measure spaces, Probab. Theory Related Fields 140(2008), 277–317.

[17] Z.-Q. Chen and T.-S. Zhang, Girsanov and Feynman-Kac type transformations for symmetric Markov processes, Ann. Inst. H. Poincaré Probab. Statist. 38(2002), 475–505.

[18] E. B. Davies, L^1 properties of second order elliptic operators, Bull. London Math. Soc. 17(1985), 417–436.

[19] E. B. Davies, Heat Kernels and Spectral Theory, Cambridge Tracts in Math. 92, Cambridge Univ. Press, Cambridge, 1989.

[20] G. De Leva and D. Kim, K. Kuwae, L^p-independence of spectral bounds Feynman-Kac semigroups by continuous additive functionals, J. Funct. Anal. 259(2010), 690–730.

[21] M. D. Donsker and S. R. S. Varadhan, Asymptotic evaluation of certain Wiener integrals for large time, Functional integration and its applications (Proc. Internat. Conf., London, 1974),15–33. Clarendon Press, Oxford, 1975.

[22] M. D. Donsker and S. R. S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, I, Comm. Pure Appl. Math. 28(1975), 1–47.

[23] M. D. Donsker and S. R. S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, III, Comm. Pure Appl. Math. 29(1976), 389–461.

[24] P. J. Fitzsimmons, Absolute continuity of symmetric diffusions, Ann. Probab. 25(1997), 230–258.

[25] P. J. Fitzsimmons and R. K. Getoor, Revuz measures and time changes, Math. Z. 199(1988), 233–256.

[26] M. Fukushima, Dirichlet spaces and strong Markov processes, Trans. Amer. Math. Soc. 162(1971), 185–224.

[27] M. Fukushima, A note on irreducibility and ergodicity of symmetric Markov processes, Stochastic processes in quantum theory and statistical physics (Marseille, 1981), 200–207, Lecture Notes in Phys. 173, Springer, Berlin, 1982.

[28] M. Fukushima, Y. Oshima and M. Takeda, Dirichlet Forms and Symmetric Markov Processes, 2nd rev. and ext. ed., de Gruyter Studies in Math., 19. Walter de Gruyter & Co., Berlin, 2011.

[29] M. Fukushima and M. Takeda, A transformation of a symmetric Markov processes and the Donsker-Varadhan theory, Osaka J. Math. 21(1984), 311–326.

[30] M. Fukushima and M. Takeda, Markov Processes (in Japanese), Baifukan, Tokyo, 2008. (Chinese translation by P. He, ed. by J. Ying, Science Press, Beijing, to be published 2011.)

[31] A. Grigor'yan, Analytic and geometric background of recurrence and non-explosion of the Brownian motion on Riemannian manifolds, Bull. Amer. Math. Soc.(N.S.) 36(1999), 135–249.

[32] A. Grigor'yan, Heat kernels on weighted manifolds and applications, Contemp. Math. 398(2006), 93–191.

[33] P. He, A formula on scattering length of dual Markov processes, Proc. Amer. Math. Soc. 139(2011), 1871–1877.

[34] K. Itô, Essentials of stochastic processes, Transl. Math. Monogr. 231, American Mathematical Society, Providence, RI, 2006.

[35] N. Jain and N. Krylov, Large deviations for occupation times of Markov processes with L_2 semigroups, Ann. Probab. 36(2008), 1611–1641.

[36] M. Kac, On some connections between probability theory and differential equations, Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability 1950, 189–215, Univ. of California Press, Berkeley and Los Angeles, 1951

[37] M. Kac, Probabilistic methods in some problems of scattering theory, Rocky Mountain J. Math. 4(1974), 511–537.

[38] M. Kac and J.-M. Luttinger, Scattering length and capacity, Ann. Inst. Fourier (Grenoble) 25(1975), 317–321.

[39] N. Kajino, Equivalence of recurrence and Liouville property for symmetric Dirichlet forms, preprint, 2010.

[40] E. H. Lieb and M. Loss, Analysis, Grad. Stud. Math. 14, American Mathematical Society, Providence, RI, 2001.

[41] K. Kuwae and M. Takahashi, Kato class measures of symmetric Markov processes under heat kernel estimates, J. Funct. Anal. 250(2007), 86–113.

[42] Z. M. Ma and M. Röckner, Introduction to the theory of (nonsymmetric) Dirichlet forms, Universitext, Springer-Verlag, Berlin, 1992.

[43] V. G. Maz'ja, Sobolev Spaces, Springer Ser. Soviet Math., Springer-Verlag, Berlin, 1985

[44] M. Murata, Structure of positive solutions of $(-\Delta + V)u = 0$ in $\mathbb{R}^d$, Duke Math. J. 53(1986), 869–943.

[45] J. Neveu, Potentiel Markovien recurrent des chaînes de Harris, Ann. Inst. Fourier(Grenoble) 22(1972), 85–130.

[46] L. Rogers and D. Williams, Diffusions, Markov Processes, and Martingales, Vol. 2, Wiley Ser. in Probab. Math. Statist. Probab. Math. Statist., John Wiley & Sons, Inc., New York, 1987.

[47] D. Revuz and M. Yor, Continuous Martingales and Brownian Motion, 3rd ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 293.

[48] L. C. G. Rogers and D. Williams, Diffusions, Markov processes, and martingales Vol. 1, Foundations, Reprint of the 2nd (1994) edition. Cambridge Math. Lib., Cambridge Univ. Press, Cambridge, 2000.

[49] B. Roynette, P. Vallois and M. Yor, Some penalisations of the Wiener measure, Jpn. J. Math. 1(2006), 263–290.

[50] B. Roynette, P. Vallois and M. Yor, Limiting laws associated with Brownian motion perturbed by normalized exponential weights I, Studia Sci. Math. Hungar. 43(2006), 171–246.

[51] S. Sato, An inequality for the spectral radius of Markov processes, Kodai Math. J. 8(1985), 5–13.

[52] M. Sharpe, General theory of Markov processes, Pure Appl. Math. 133, Academic Press, Academic Press, Inc., Boston, MA, 1988.

[53] Y. Shiozawa and M. Takeda, Variational formula for Dirichlet forms and estimates of principal eigenvalues for symmetric α-stable processes, Potential Anal. 23(2005), 135–151.

[54] M. Silverstein, Symmetric Markov Processes, Lect. Notes in Math. 426, Springer-Verlag, Berlin-New York, 1974.

[55] P. Stollmann and J. Voigt, Perturbation of Dirichlet forms by measures, Potential Anal. 5(1996), 109–138.

[56] D. W. Stroock, The Kac approach to potential theory I, J. Math. Mech. 16(1967), 829–852.

[57] Y. Takahashi, An integral representation on the path space for scattering length, Osaka J. Math. 27(1990), 373–379.

[58] M. Takeda, On a Large deviation for symmetric Markov processes with finite lifetime, Stochastics and Stochastic Rep. 59(1996), 143–167.

[59] M. Takeda, Exponential decay of lifetimes and a theorem of Kac on total occupation times, Potential Anal. 11(1999), 235–247.

[60] M. Takeda, L^p-independence of the spectral radius of symmetric Markov semigroups, Stochastic Processes, Physics and Geometry: New Interplays II, new interplays(Leipzig, 1999), 613–623, CMS Conf. Proc. 29, Amer. Math. Soc., Providence, RI, 2000.

[61] M. Takeda, Conditional gaugeability and subcriticality of generalized Schrödinger operators, J. Funct. Anal. 191(2002), 343–376.

[62] M. Takeda, Gaussian bounds of heat kernels for Schrödinger operators on Riemannian manifolds, Bull. London Math. Soc. 39(2007), 85–94.

[63] M. Takeda, L^p-independence of spectral bounds of Schrödinger type semigroups, J. Funct. Anal. 252(2007), 550–565.

[64] M. Takeda, Large deviations for additive functionals of symmetric stable processes, J. Theoret. Probab. 21(2008), 336–355.

[65] M. Takeda, A formula on scattering length of positive smooth measures, Proc. Amer. Math. Soc. 138(2010), 1491–1494.

[66] M. Takeda, Feynman-Kac penalisations of symmetric stable processes, Electron. Comm. Probab. 15(2010), 32–43.

[67] M. Takeda, A large deviation principle for symmetric Markov processes with Feynman-Kac functional, to appear in J. Theoret. Probab.

[68] M. Takeda, L^p-independence of growth bounds of Feynman-Kac semigroups, Surveys in Stochastic Processes, to appear in Proceedings of the 33rd SPA, European Mathematical Society.

[69] M. Takeda and Y. Tawara, L^p-independence of spectral bounds of non-local Feynman-Kac semigroups, Forum Math. 21(2009), 1067–1080.

[70] M. Takeda and Y. Tawara, A Large deviation principle for symmetric Markov processes normalized by Feynman-Kac Functionals, preprint.

[71] M. Takeda and K. Tsuchida, Differentiability of spectral functions for symmetric α-stable processes, Trans. Amer. Math. Soc. 359(2007), 4031–4054.

[72] M. Takeda and K. Tsuchida, Large deviations for discontinuous additive functionals of symmetric stable processes, to appear in Math. Nachr.

[73] M. Takeda and T. Uemura, Subcriticality and gaugeability for symmetric α-stable processes, Forum Math. 16(2004), 505–517.

[74] H. Tamura, Semi-classical limit of scattering length, Lett. Math. Phys. 24(1992), 205–209.

[75] M. E. Taylor, Scattering length and perturbations of $-\Delta$ by positive potentials, J. Math. Anal. Appl. 53(1976), 291–312.

[76] Y. Tawara, L^p-independence of spectral bounds of Schrödinger-type operators with non-local potentials, J. Math. Soc. Japan 62(2010), 767–788.

[77] S. R. S. Varadhan, Asymptotic probabilities and differential equations, Comm. Pure Appl. Math. 19(1966), 261–286.

[78] L. Wu, Some notes on large deviations of Markov processes, Acta Math. Sin. (Engl. Ser.) 16(2000), 369–394.

[79] K. Yano, Y. Yano and M. Yor, Penalising symmetric stable Lévy paths, J. Math. Soc. Japan. 61(2009), 757–798.

[80] Z. Zhao, Subcriticality and gaugeability of the Schrödinger operator, Trans. Math. Soc. 334(1992), 75–96.

Mathematical Institute
Tohoku University
Aoba, Sendai, 980-8578
Japan

E-mail address: takeda@math.tohoku.ac.jp

INTRODUCTION TO MARKOV PROCESSES

Jiangang YING

1 Introduction

I have planned for years and am now trying hard to write a book on theory of Markov processes and symmetric Markov processes so that graduate students in this field can move to the frontier quickly. In my impression, Markov processes are very intuitive to understand and manipulate. However to make the theory rigorously, one needs to read a lot of materials and check numerous measurability details it involved. This is a kind of dirty work a fresh graduate student hates to do, but has to do. My purpose is to help young researchers who are interested in this field to reduce the fear when they face it.

We shall state some fundamental results in general theory of stochastic processes mainly developed by Strasbourg school of probability. Let $(\Omega, \mathscr{F}, \mathsf{P})$ be a complete probability space. A family of σ-algebra $(\mathscr{F}_t) = \{\mathscr{F}_t : t \geq 0\}$ is called a **filtration** if for any $0 \leq s < t$, $\mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{F}$. We say a filtration $(\mathscr{F}_t)$ satisfies the **usual condition** if each $\mathscr{F}_t$ contains all null sets in $\mathscr{F}$ and it is right continuous, namely, for each $t \geq 0$,

$$\mathscr{F}_t = \mathscr{F}_{t+} = \bigcap_{s>t} \mathscr{F}_s. \tag{1.1}$$

Let $(\mathscr{F}_t)$ be a filtration and $X = (X_t : t \geq 0)$ a real-valued stochastic process. X is $(\mathscr{F}_t)$-**adapted** (or adapted, if no confusion will be caused) if for each $t \geq 0$, X_t is $\mathscr{F}_t$-measurable. Moreover X is $(\mathscr{F}_t)$-**progressively measurable** if for each $t \geq 0$, the map $(s, \omega) \mapsto X_s(\omega)$ is measurable as a map from $([0,t] \times \Omega, \mathscr{B}[0,t] \times \mathscr{F}_t)$ to $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$, where $\mathscr{B}[0,t]$ and $\mathscr{B}(\mathbf{R})$ are Borel σ-algebra on $[0,t]$ and $\mathbf{R}$, respectively. If no confusion is caused, $(\mathscr{F}_t)$ in the front may be omitted. A subset $A \subset \mathbf{R} \times \Omega$ is progressively measurable if so is the process $(t, \omega) \mapsto 1_A(t, \omega)$. A process is called right continuous or continuous or left continuous if almost all sample path has such regularity.

Theorem 1.1 A right continuous and adapted process is progressively measurable.

The least σ-algebra on $\mathbf{R} \times \Omega$ such that all adapted right continuous real processes are measurable is denoted by $\mathscr{O}$, an optional σ-field. A process which is $\mathscr{O}$-measurable is called **optional**. Then the theorem above implies that an optional process is progressively measurable. A map $\tau : \Omega \mapsto [0,\infty]$ is called an $(\mathscr{F}_t)$-stopping time if for each $t \geq 0$, $\{\tau \leq t\} \in \mathscr{F}_t$. For a stochastic process X and a subset $A \subset \mathbf{R}$, define the hitting time of A as

$$T_A = \inf\{t > 0 : X_t \in A\}. \tag{1.2}$$

Theorem 1.2 If the filtration $(\mathscr{F}_t)$ satisfies the usual condition and X is progressively measurable, then for any Borel subset $A \subset \mathbf{R}$, T_A is a stopping time.

The following theorem is called the section theorem, which is fundamental.

Theorem 1.3 Let X be a bounded progressively measurable process. If for any bounded decreasing sequence of stopping times $\{T_n\}$,

$$\lim_n \mathsf{E}[X_{T_n}] = \mathsf{E}[X_{\lim_n T_n}], \tag{1.3}$$

then X is right continuous.

2 Right continuous Markov processes

In this section we shall first introduce the notion of right processes, which, essentially due to P.A. Meyer, makes classical potential theory operate almost naturally on it. Though, more or less, right processes are right continuous Markov processes with strong Markov property, it is a difficult task to give the definition clearly and concisely. Let $(E,\mathscr{E})$ be a topological space with its Borel σ-algebra. For any probability measure μ on E, $\mathscr{E}^\mu$ is the completion of $\mathscr{E}$ under μ and set

$$\mathscr{E}^* = \bigcap_\mu \mathscr{E}^\mu \tag{2.1}$$

where μ runs over all probability measures on E. A set in $\mathscr{E}^*$ is called a universally measurable subset of E. Any probability measure on $(E,\mathscr{E})$ may be uniquely extended on $\mathscr{E}^*$. The requirement for topology on E may vary, but in most cases, Radon space or Lusin space, which is a universally measurable subset or Borel subset of a compact metric space, respectively. One reason why we need to start from seemly so general topology is that in this way the class of right processes keeps stable under usual transformation in Markov processes such as killing transform, time change and drift transform.

Definition 2.1 Let E be a Radon space. A family of kernels $(P_t)_{t\geq 0}$ on $(E,\mathscr{E}^*)$ is called a transition semigroup if $P_tP_s = P_{t+s}$ for any $t,s \geq 0$ and $P_t(x,E) \leq 1$ for any $t \geq 0$ and $x \in E$. In addition, if $P_t(x,E) = 1$ for any $t \geq 0$ and $x \in E$, it is called a transition probability semigroup. A transition semigroup (P_t) is called a Borel semigroup if E is Lusin space and each P_t is a kernel on $(E,\mathscr{E})$, or maps a Borel measurable function to a Borel measurable function.

It is known that by joining a point Δ, called a cemetery point, to E, (P_t) may be extended into a transition probability semigroup on E_Δ.

Definition 2.2 A group of notations

$$X = (\Omega, \mathscr{G}, \mathscr{G}_t, X_t, \theta_t, \mathsf{P}^x)$$

is called a **right continuous Markov process** on state space E with transition semigroup (P_t) if the following conditions are satisfied.

(1) $(\Omega, \mathscr{G}, \mathscr{G}_t)$ is a filtered measurable space and (X_t) is an E_Δ-valued process $\mathscr{E}^*_\Delta$-adapted to $(\mathscr{G}_t)$, more precisely for any $t \geq 0$, X_t is a measurable mapping from $(\Omega, \mathscr{G}_t)$ to $(E_\Delta, \mathscr{E}^*_\Delta)$.

(2) $(\theta_t)_{t\geq 0}$ is a family of shift operators for X, i.e., $\theta_t : \Omega \to \Omega$ and, identically for any $t, s \geq 0$,

$$\theta_t \circ \theta_s = \theta_{t+s} \text{ and } X_t \circ \theta_s = X_{t+s}.$$

(3) For every $x \in E_\Delta$, P^x is a probability measure on $(\Omega, \mathscr{G})$ and

$$\mathsf{P}^x(X_0 = x) = 1.$$

Moreover $x \mapsto \mathsf{P}^x(H)$ is universally measurable for any $H \in \mathscr{G}$.

(4) For every $t, s \geq 0$, $f \in b\mathscr{E}^*$ and $x \in E$, it holds P^x-a.s.

$$\mathsf{P}^x(f(X_{t+s})|\mathscr{G}_s) = P_t f(X_s). \tag{2.2}$$

(5) For any $x \in E$ and P^x-almost all $\omega \in \Omega$, $t \mapsto X_t(\omega)$ is a right continuous process on $E_\Delta = E \cup \{\Delta\}$.

(6) Define $\zeta(\omega) := \inf\{t : X_t = \Delta\}$. Then $X_t(\omega) \in E$ for $t < \zeta(\omega)$, and $X_t(\omega) = \Delta$ for all $t \geq \zeta(\omega)$. Hence ζ is called the lifetime of X.

In this case we sometimes say that X or its semigroup (P_t) satisfies (HD1).

The word **for almost all** means "for any $x \in E$ and P^x-almost all", i.e., a measurable subset Ω_0 of Ω such that $\mathsf{P}^x(\Omega_0) = 1$ for all $x \in E$. Notice that the measurability in (3) is not so much restricted. Let $(\mathscr{F}^{0*}_t)$ (resp., $(\mathscr{F}^0_t)$) be the natural filtration of (X_t) generated by $\mathscr{E}^*$ (resp., $\mathscr{E}$), precisely,

$$\mathscr{F}^{0*}_t = \sigma\left(\bigcup_{s\leq t} X_s^{-1}(\mathscr{E}^*)\right), \ \mathscr{F}^0_t = \sigma\left(\bigcup_{s\leq t} X_s^{-1}(\mathscr{E})\right). \tag{2.3}$$

Clearly for any $t \geq 0$, $\mathscr{F}_t^{0*} \subset \mathscr{G}_t$ and $\mathscr{F}_\infty^{0*} \subset \mathscr{G}$. By monotone class theorem, $x \mapsto \mathsf{P}^x(H)$ is $\mathscr{E}^*$-measurable for any $H \in \mathscr{F}_\infty^{0*}$. Furthermore if (P_t) is Borel semigroup, $x \mapsto \mathsf{P}^x(H)$ is $\mathscr{E}$-measurable for any $H \in \mathscr{F}_\infty^0$.

Fix now such a process X on E. For any probability measure μ on $(E, \mathscr{E})$, define

$$\mathsf{P}^\mu(H) = \int_E \mathsf{P}^x(H)\mu(dx), \; H \in \mathscr{G}. \tag{2.4}$$

Denote by $\mathscr{G}^\mu$ the completion of $\mathscr{G}$ with respect to P^μ and $\mathscr{G}_t^\mu$ the σ-field generated by $\mathscr{G}_t$ and all P^μ-null sets in $\mathscr{G}^\mu$. We usually say that $\mathscr{G}_t^\mu$ is the **augmentation** of $\mathscr{G}_t$ in $(\Omega, \mathscr{G}, \mathsf{P}^\mu)$.

Exercise 2.1 Prove that the completion of $\mathscr{F}_\infty^0$ with respect to P^μ is equal to the completion of $\mathscr{F}_\infty^{0*}$ with respect to P^μ. The same conclusion holds for the augmentation of $(\mathscr{F}_t^0)$ and $(\mathscr{F}_t^{0*})$ in $(\Omega, \mathscr{F}, \mathsf{P}^\mu)$.

Set

$$\widetilde{\mathscr{G}} = \bigcap_\mu \mathscr{G}^\mu, \widetilde{\mathscr{G}_t} = \bigcap_\mu \mathscr{G}_t^\mu, \tag{2.5}$$

where μ runs over all probability measures on $(E, \mathscr{E})$. The filtration $(\widetilde{\mathscr{G}_t})$ is called the augmentation of $(\mathscr{G}_t)$. It is not hard to see that the process has Markov property with respect to $(\widetilde{\mathscr{G}_t})$ and actually for any probability measure μ on E, it holds P^μ-a.s. for $t, s \geq 0$, $f \in b\mathscr{E}^*$

$$\mathsf{P}^\mu(f(X_{t+s})|\mathscr{G}_s^\mu) = P_t f(X_s). \tag{2.6}$$

The procedure to get $(\widetilde{\mathscr{G}_t^\mu})$ and $(\widetilde{\mathscr{G}_t})$ is called **augmentation** of the filtration of X with respect to the laws (P^x). This is a 'dirty' work which has to be done for a Markov process. Therefore we may assume from the beginning that $\mathscr{G}$ and $(\mathscr{G}_t)$ are augmented. The augmentation of the natural filtration $(\mathscr{F}_t^0)$ is denoted by $(\mathscr{F}_t)$, which is also the augmentation of $(\mathscr{F}_t^{0*})$. After the augmentation, we have to check that we still have the necessary measurability such as

(1) For $B \in \mathscr{G}$, $x \mapsto \mathsf{P}^x(B)$ is universally measurable;

(2) X_t is measurable from $(\Omega, \mathscr{G}_t)$ to $(E, \mathscr{E}^*)$;

(3) θ_t is measurable on $(\Omega, \mathscr{G})$.

The good news about augmentation which we shall prove later is that $(\mathscr{G}_t)$ will satisfy the usual condition when a slight more condition is imposed, and then the hitting time of any optional set is then a stopping time.

For $\alpha > 0$, a $[0, \infty]$-valued measurable function f on $(E, \mathscr{E}^*)$ is α-supermidian if $e^{-\alpha t} P_t f \leq f$ for each $t > 0$ and α-excessive if, in addition, $e^{-\alpha t} P_t f \uparrow f$ as $t \downarrow 0$. Let $\mathbf{S}^\alpha$ be the set of all α-excessive functions.

Definition 2.3 Let E be a Radon space and (P_t) a transition semigroup on E. Assume that the collection

$$X = (\Omega, \mathscr{G}, \mathscr{G}_t, X_t, \theta_t, \mathsf{P}^x)$$

is a right continuous Markov process (HD1) on E with (P_t) as its semigroup. Then X is said to be a **right process** provided X satisfies (HD2), or precisely, for any α-excessive function f, $t \mapsto f(X_t)$ is right continuous almost surely. Moreover if E is Lusin space and (P_t) is a Borel semigroup, then X is called a Borel right process.

We shall now introduce the notion of potential which plays an essential role in general theory of Markov processes. To define α-potentials, some measurability needs to be clarified in advance. For a bounded continuous function f on E, $t \mapsto f(X_t)$ is right continuous and hence $(t,x) \mapsto \mathsf{E}^x[f(X_t)] = P_t f(x)$ jointly measurable on $(\mathbf{R}^+ \times E, \mathscr{B}(\mathbf{R}^+) \times \mathscr{E}^*)$. It is also true for bounded Borel measurable f by monotone class theorem. The following result makes us define resolvent of (P_t) legally by using Fubini theorem.

Exercise 2.2 For $f \in b\mathscr{E}^*$, $(t,x) \mapsto \mathsf{E}^x[f(X_t)]$ is measurable for the completion of $(\mathscr{B}(\mathbf{R}^+) \times \mathscr{E}^*)$ with respect to the product measure of any finite measure on $\mathbf{R}^+$ and a finite measure on $\mathscr{E}^*$.

For $\alpha > 0$ and $f \in b\mathscr{E}^*$, define the resolvent or α-potential of f

$$U^\alpha f(x) = \int_0^\infty e^{-\alpha t} P_t f(x) dt \tag{2.7}$$

$$= \mathsf{E}^x \int_0^\infty e^{-\alpha t} f(X_t) dt. \tag{2.8}$$

Then we have the well-known resolvent equation

$$U^\alpha = U^\gamma + (\gamma - \alpha) U^\alpha U^\gamma \tag{2.9}$$

for $\alpha, \gamma > 0$.

The first important property of right processes is strong Markov property. We now give two fundamental theorems for right processes. Note that we may always assume that Ω is the canonical space, i.e., the space of right continuous maps from $[0, \infty)$ to E. To state strong Markovian property, we assume that readers are familiar with the theory related to stopping times.

Theorem 2.4 Let X be a right process on E with transition semigroup (P_t). Then

(1) X has strong Markov property with respect to $(\mathscr{F}^0_{t+})$, i.e., for any $(\mathscr{F}^0_{t+})$-stopping time σ, $f \in b\mathscr{E}^*$, $t > 0$ and $x \in E$,

$$\mathsf{E}^x[f(X_{t+\sigma}) 1_{\{\sigma<\infty\}} | \mathscr{F}^0_{\sigma+}] = 1_{\{\sigma<\infty\}} \mathsf{E}^{X_\sigma}[f(X_t)], \tag{2.10}$$

P^x-a.s.;

(2) for any probability μ on E, $(\mathscr{F}_t^\mu)$ is right continuous, and then $(\mathscr{F}_t)$ is right continuous.

Proof. (1) Let f be a uniformly continuous bounded function on E. Assume that $\sigma < \infty$. Set

$$\sigma_n = \sum_{k\geq 1} \frac{k}{2^n} 1_{\{(k-1)2^{-n}\leq\sigma<k2^{-n}\}}$$

Then $\sigma_n \downarrow \sigma$ and σ_n is $(\mathscr{F}_t^0)$-stopping time. By the right continuity and simple Markov property of X, we have for any probability μ on E,

$$\begin{aligned}
\mathsf{E}^\mu \int_0^\infty \mathrm{e}^{-\alpha t} f(X_{t+\sigma})\mathrm{d}t &= \lim_n \mathsf{E}^\mu \int_0^\infty \mathrm{e}^{-\alpha t} f(X_{t+\sigma_n})\mathrm{d}t \\
&= \lim_n \sum_k \mathsf{E}^\mu \left(\int_0^\infty \mathrm{e}^{-\alpha t} f(X_{t+\frac{k}{2^n}})\mathrm{d}t; \sigma_n = \frac{k}{2^n} \right) \\
&= \lim_n \sum_k \mathsf{E}^\mu \left(\mathsf{E}^{X(\frac{k}{2^n})} \int_0^\infty \mathrm{e}^{-\alpha t} f(X_t)\mathrm{d}t; \sigma_n = \frac{k}{2^n} \right) \\
&= \lim_n \mathsf{E}^\mu U^\alpha f(X_{\sigma_n}) = \mathsf{E}^\mu U^\alpha f(X_\sigma) \\
&= \int_0^\infty \mathrm{e}^{-\alpha t} \mathsf{E}^\mu \mathsf{E}^{X(\sigma)} f(X_t)\mathrm{d}t.
\end{aligned}$$

This means that $t \mapsto \mathsf{E}^\mu f(X_{t+\sigma})$ and $t \mapsto \mathsf{E}^\mu[\mathsf{E}^{X(\sigma)} f(X_t)]$ have the same Laplace transform and it implies they are identical because they are both right continuous. Hence

$$\mathsf{E}^\mu f(X_{t+\sigma}) = \mathsf{E}^\mu[\mathsf{E}^{X(\sigma)} f(X_t)]$$

from which, the strong Markov property with respect to $(\mathscr{F}_{t+}^0)$ follows.

(2) Obviously (1) implies that X has simple Markov property with respect to $(\mathscr{F}_{t+}^0)$, i.e., for any bounded random variable Y on $(\Omega, \mathscr{F}_\infty^0)$ and a probability μ on E, it holds P^μ-a.s.

$$\mathsf{E}^\mu(Y\circ\theta_t|\mathscr{F}_{t+}^0) = \mathsf{E}^{X_t}(Y) = \mathsf{E}^\mu(Y\circ\theta_t|\mathscr{F}_t^0). \tag{2.11}$$

It is easy to verify that when $Y = f_1(X_1)\cdots f_n(X_{t_n})$,

$$\mathsf{E}^\mu(Y|\mathscr{F}_{t+}^0) = \mathsf{E}^\mu(Y|\mathscr{F}_t^0),$$

and actually it holds for any $Y \in b\mathscr{F}_\infty^0$ by monotone class theorem. Then for $A \in \mathscr{F}_{t+}^0$, we have P^μ-a.s. $1_A = \mathsf{E}^\mu(1_A|\mathscr{F}_t^0)$, and hence $A \in \mathscr{F}_t^\mu$. It implies that

$$\mathscr{F}_{t+}^0 \subset \mathscr{F}_t^\mu.$$

The conclusion follows from an assertion that the σ-field generated by $\mathscr{F}_{t+}^0$ and P^μ-null sets equals $\mathscr{F}_{t+}^\mu$, which is left to the readers as an exercise. □

By (2) in Theorem 2.4, we may always assume without loss of generality that the filtration $(\mathscr{G}_t)$ satisfies usual condition, i.e., it contains all null sets and right continuous. By augmentation, the strong Markov property may be stated as follows. For any probability μ and a non-negative function $f \in \mathscr{E}^*$, if σ is an $(\mathscr{F}_t^\mu)$-stopping time, then

$$\mathsf{E}^\mu(f(X_{t+\sigma})1_{\{\sigma<\infty\}}|\mathscr{F}_t^\mu) = P_t f(X_\sigma)1_{\{T<\infty\}}. \tag{2.12}$$

Exercise 2.3 Prove that

$$\mathsf{E}^\mu \int_T^\infty e^{-\alpha t} f(X_t)\mathrm{d}t = \mathsf{E}^\mu \left(e^{-\alpha T} U^\alpha f(X_T)\right). \tag{2.13}$$

For $\alpha \geq 0$, a non-negative measurable function f (may take $+\infty$) is called α-excessive and written as $f \in \mathbf{S}^\alpha$ if (1) $e^{-\alpha t} P_t f \leq f$ for any $t > 0$; (2) $\lim_{t\downarrow 0} e^{-\alpha t} P_t f = f$. Write simply $\mathbf{S}^0$ as $\mathbf{S}$. The following lemma lists some properties of excessive functions and is easy to verify.

Lemma 2.5 (1) $\mathbf{S}^\alpha$ is a cone.
(2) $\mathbf{S}^\alpha$ is stable under increasing limit.
(3) If $\alpha > \beta \geq 0$, $\mathbf{S}^\alpha \supset \mathbf{S}^\beta$ and $\mathbf{S}^\beta = \bigcap_{r>\beta} \mathbf{S}^r$.
(4) If $f, g \in \mathbf{S}^\alpha$, $f \wedge g \in \mathbf{S}^\alpha$.
(5) If $f \in \mathbf{S}^\alpha$ and μ is a probability measure on E satisfying $\mu(f) < \infty$, then the process $(e^{-\alpha t} f(X_t))$ is a super-martingale with respect to P^μ.

Actually the proof of (4) needs to use (HD2) on the process.

Lemma 2.6 (1) For $\alpha \geq 0$ and $\mathscr{E}^*$-measurable $f \geq 0$, $U^\alpha f \in \mathbf{S}^\alpha$.
(2) For $\alpha \geq 0$, $f \in \mathbf{S}^\alpha$ if and only if $\beta U^{\alpha+\beta} f \uparrow f$ as $\beta \uparrow +\infty$.
(3) For $\alpha > 0$ and $f \in \mathbf{S}^\alpha$, there exist $g_n \in b\mathscr{E}_+^*$ such that $U^\alpha g_n \uparrow f$ as $n \uparrow +\infty$.

Since a super-martingale which is the limit of a sequence of right continuous super-martingales is also right continuous, we shall state following weaker forms of (HD2). A negative function f on E is nearly Borel for X if for each probability μ on E there exist $f_1, f_2 \in \mathscr{E}$ with $f_1 \leq f \leq f_2$ such that two processes $(f_1(X_t))$ and $(f_2(X_t))$ are P^μ-indistinguishable. A measurable function f on $(E, \mathscr{E}^*)$ is called **optional** if $t \mapsto f(X_t)$ is indistinguishable from an $(\mathscr{F}_t)$-optional process, and **nearly optional** if for any probability measure μ on E, $t \mapsto f(X_t)$ is indistinguishable from an $(\mathscr{F}_t^\mu)$-optional process. A set $A \in \mathscr{E}^*$ is optional or nearly optional if so is 1_A. Let $\mathscr{E}^{no}$ be the set of nearly optional subsets of E which is a σ-algebra.

Exercise 2.4 Prove that f is nearly optional if f is $\mathscr{E}^{no}$-measurable.

The next theorem follows from the section theorem as stated in Theorem 1.3.

Theorem 2.7 Assume that (HD1) holds. If X is strong Markov and each α-excessive function is nearly optional, then (HD2) holds. Therefore if (P_t) is Borel and X has strong Markov property, then (HD2) holds and each α-excessive function is nearly Borel.

Proof. For any probability μ, an increasing sequence $\{T_n\}$ of stopping times with $T = \lim T_n$, and a non-negative bounded function $g \in \mathscr{E}^*$, we have by strong Markov property

$$\begin{aligned} \mathsf{E}^\mu \left(e^{-\alpha T_n} U^\alpha g(X_{T_n})\right) &= \mathsf{E}^\mu \int_{T_n}^\infty e^{-\alpha t} g(X_t) \mathrm{d}t \\ &\to \mathsf{E}^\mu \int_T^\infty e^{-\alpha t} g(X_t) \mathrm{d}t = \mathsf{E}^\mu \left(e^{-\alpha T} U^\alpha g(X_T)\right). \end{aligned}$$

Combining the assumption that $U^\alpha g$ is nearly optional, it follows that $t \mapsto U^\alpha g(X_t)$ is right continuous from Theorem 1.3. Finally by Lemma 2.6(3), $t \mapsto f(X_t)$ is right continuous for any α-excessive function f. □

Remark 2.8 Though Theorem 2.7 hints that (HD2) may be equivalent to strong Markov property, an example, when (P_t) is not Borel, is presented by Salisbury to show that a right continuous Markov process with strong Markov property may not be a right process.

Theorem 2.9 Assume (HD1) holds. Let $\mathbf{C}$ be a linear subspace of $C(E)$, closed under function multiplication, which generates $\mathscr{E}$. If, for any bounded $f \in \mathbf{C}$, the process $t \mapsto U^\alpha f(X_t)$ is right continuous, then (HD2) holds.

Proof. It suffices to show that $U^\alpha g$ is nearly optional for non-negative and bounded $g \in \mathscr{E}^*$. It is true by monotone class theorem for $g \in \mathscr{E}$ and it follows from the proof of Theorem 2.7 that $t \mapsto U^\alpha g(X_t)$ is right continuous. Let now $g \in \mathscr{E}^*$ be bounded. For any probability μ on E, there exist $g_1, g_2 \in \mathscr{E}$ such that $g_1 \leq g \leq g_2$ and $\mu U^\alpha(g_2 - g_1) = 0$. Then for any $t > 0$, $U^\alpha g_1(X_t) \leq U^\alpha g(X_t) \leq U^\alpha g_2(X_t)$ and

$$\mathsf{E}^\mu[U^\alpha(g_2 - g_1)(X_t)] = \mu P_t U^\alpha(g_2 - g_1) \leq e^{\alpha t} \mu U^\alpha(g_2 - g_1) = 0.$$

Therefore two processes $U^\alpha g_2(X_\cdot)$ and $U^\alpha g_1(X_\cdot)$ are P^μ-distinguishable, i.e., $U^\alpha g$ is nearly optional. □

Example 2.10 (α-subprocess) Let X be a right process on E with transition semigroup (P_t). For $\alpha > 0$, it is known that $P_t^\alpha = e^{-\alpha t} P_t$ is also a transition semigroup on E. Is it a transition semigroup of a right process? Sure it is. But how do we construct the right process? Introduce the killing operators (k_t) on Ω:

$$X_s \circ k_t = \begin{cases} X_s, & s < t, \\ \Delta, & s \geq t, \end{cases} \tag{2.14}$$

Intuitively k_t makes no change before time t but sends the path after t to cemetery. For $x \in E$, define probability $\mathbb{Q}^x$ on $(\Omega, \mathscr{F})$ by

$$\mathbb{Q}^x(Y) = \mathsf{E}^x \int_0^\infty Y \circ k_u \mathrm{d}(-e^{-\alpha u}) = \alpha \mathsf{E}^x \int_0^\infty Y \circ k_u e^{-\alpha u} \mathrm{d}u, \tag{2.15}$$

where Y is a bounded or non-negative random variable on Ω. Note that we use $\mathbb{Q}$ for both probability and expectation. Let

$$X^\alpha = (\Omega, \mathscr{F}, \mathscr{F}_t, X_t, \theta_t, \mathbb{Q}^x)$$

which is called α-subprocess of X. It is easy to check that X is a right process with transition semigroup (P_t^α). In fact,

$$\begin{aligned}\mathbb{Q}^x(f(X_t)) &= \alpha \mathsf{E}^x \int_0^\infty f(X_t) \circ k_u e^{-\alpha u} \mathrm{d}u \\ &= \alpha \mathsf{E}^x \int_t^\infty f(X_t) e^{-\alpha u} \mathrm{d}u \\ &= e^{-\alpha t} \mathsf{E}^x (f(X_t)) = P_t^\alpha f(x).\end{aligned}$$

The verification of (HD2) is left for those who are interested. ∎

Example 2.11 (Killing at leaving) Let X be a right process on E with transition semigroup (P_t). Intuitively for a subset B, killing X at leaving B shall give us a process which certainly inherits Markov process from X. Rigorously speaking, let $B \in \mathscr{E}^o$ and $T = T_B$ the hitting time of B. Define a map $\omega \mapsto k_T\omega$ on Ω by

$$k_T\omega(t) = \begin{cases} \omega(t), & t < T; \\ \Delta, & t \geq T. \end{cases} \tag{2.16}$$

Hence the new lifetime is $\zeta \wedge T$. ∎

3 Feller processes and Lévy processes

A question we must ask is when and how we will have a right Markov process. There are basically two ways: one is from Feller semigroup and the other is through transformation as the example in the last section shows. In this section we shall introduce Feller semigroup and prove that it may be realized as a Markov process much better than a right process.

Definition 3.1 Let E be a locally compact metrizable space with a countable base. A transition semigroup (P_t) on E is called a Feller semigroup if
(1) $P_t C_\infty(E) \subset C_\infty(E)$ for each $t > 0$;
(2) for each $f \in C_\infty(E)$,

$$\lim_{t \downarrow 0} \|P_t f - f\|_\infty = 0.$$

With other conditions, (2) above is equivalent to a weaker one: for each $f \in C_\infty(E)$ and $x \in E$, $P_t f(x) \to f(x)$ as $t \downarrow 0$. The proof is a good exercise. Since $C_\infty(E)$ is a Banach space and Feller semigroup (P_t) is a strongly continuous semigroup on $C_\infty(E)$, its infinitesimal generator determines (P_t) completely by Hille-Yosida theorem. The following theorem is actually a corollary of regularization theorem of super-martingales.

Theorem 3.2 Let (P_t) be a Feller semigroup on E. Then (P_t) has a realization which is a Borel right process, which is called a **Feller process**.

Proof. Add a point Δ to E such that E_Δ is compact and (P_t) is extended to a probability transition semigroup on E_Δ. Any function f on E may be always viewed as a function on E_Δ by defining $f(\Delta) = 0$. In this way

$$C_\infty(E) = \{f \in C(E_\Delta) : f(\Delta) = 0\}.$$

Let $X = (X_t, \mathsf{P}^x)$ be a realization of (P_t) on E_Δ. For any non-negative $f \in C_\infty(E)$ and $\alpha > 0$, the process $(e^{-\alpha t} U^\alpha f(X_t) : t \geq 0)$ is a super-martingale with respect to P^x for each $x \in E$. It follows that $t \mapsto U^\alpha f(X_t)$ has right and left limits P^x-almost surely. We may take a countable subset D of $\{U^\alpha f : \alpha > 0, f \in C_\infty(E_\Delta)\}$ separating points in E. Since D is countable, there exists $N_0 \subset \Omega$ such that $\mathsf{P}^x(N_0) = 0$ for all $x \in E$ and for any $g \in D$ and $\omega \notin N_0$, $t \mapsto g(X_t(\omega))$ has right and left limits. From the facts that D separates points in E and any function in D is continuous, it follows that for $\omega \notin N_0$, $t \mapsto X_t(\omega)$ has right and left limits. Let $Y = (Y_t)$ is the right limit process of X, namely

$$Y_t(\omega) = \lim_{s \downarrow t, s > t} X_s(\omega),\ t \geq 0, \omega \notin N_0.$$

It suffices to show that Y is a version of X. Fix $t \geq 0$ and $s > 0$. Take any non-negative functions $f, g \in C_\infty(E)$ and

$$\begin{aligned} \mathsf{E}^x(f(X_t)g(X_{s+t})) &= \mathsf{E}^x(f(X_t)P_s g(X_t)) \\ &= P_t(f P_s g)(x). \end{aligned}$$

As $s \downarrow 0$, $X_{s+t} \to Y_t$, $P_s g \to g$ and hence we have

$$\mathsf{E}^x(f(X_t)g(Y_t)) = P_t(fg)(x) = \mathsf{E}^x(f(X_t)g(X_t)).$$

It follows from the monotone convergence theorem that for any continuous function $h \geq 0$ on $E \times E$,

$$\mathsf{E}^x[h(X_t, Y_t)] = \mathsf{E}^x[h(X_t, X_t)]$$

and we have $X_t = Y_t$ a.s.

Hence $Y = (Y_t)$ is a right continuous realization of (P_t) and it is easy to see that Y is a Borel right process, due to Theorem 2.9 and the fact that $U^\alpha f$ is continuous for any $f \in C_\infty(E)$. □

An important example of Feller semigroup is the convolution semigroup on Euclidean space, whose right continuous realization is called a Lévy process.

Definition 3.3 A family of probability measures $\{\nu_t : t > 0\}$ on $\mathbf{R}^d$ is called a **convolution semigroup** if
(1) $\nu_t * \nu_s = \nu_{t+s}$ for any $t, s > 0$;
(2) $\nu_t \to \varepsilon_0$ weakly as $t \downarrow 0$ where ε_0 is the point mass at 0.

Let $\{\nu_t\}$ be a convolution semigroup on $\mathbf{R}^d$ and set $P_t(x, dy) = \nu_t(dy - x)$. Then (P_t) is a Feller semigroup on $\mathbf{R}^d$ and its right continuous realization is called a **Lévy process** on $\mathbf{R}^d$. Actually many well-known Markov processes such as Brownian motion, Poisson process, stable process, are Lévy processes. The law of a Lévy process is determined by its convolution semigroup, which is in turn determined by its so-called Lévy exponent.

Let $\widehat{\nu}_t$ denote the characteristic function of ν_t which is bounded and continuous on $\mathbf{R}^d$. There exists a complex-valued continuous function φ on $\mathbf{R}^d$ such that

$$\widehat{\nu}_t = \exp(-t\varphi), \tag{3.1}$$

and this φ determines $\{\nu_t\}$ uniquely by the uniqueness of Fourier transform and called the Lévy exponent of $\{\nu_t\}$. Obviously $\varphi(0) = 0$ and it is well-known that φ has the following representation: for $x \in \mathbf{R}^d$,

$$\varphi(x) = i(a, x) + \frac{1}{2}(Sx, x) + \int_{\mathbf{R}^d} \left(1 - e^{i(x,y)} + \frac{i(x, y)}{1 + |y|^2}\right) \pi(dy), \tag{3.2}$$

where $a \in \mathbf{R}^d$, S is a $d \times d$ non-negative definite symmetric matrix, and π is a Radon measure on $\mathbf{R}^d \setminus \{0\}$ having the following integrability

$$\int_{\mathbf{R}^d} \frac{|y|^2}{1 + |y|^2} \pi(dy) < \infty. \tag{3.3}$$

The matrix S and measure π are uniquely determined. But the vector a depends on the way we write (3.2). Conversely given a function φ as in (3.2), there must be a unique convolution semigroup $\{\nu_t\}$ on $\mathbf{R}^d$ such that (3.1) holds. This characterization is the famous **Lévy-Khinchin formula**, which tells us that every character about a Lévy process may be retrieved from its Lévy exponent.

It is easy to verify that Lévy exponent of Brownian motion is $\varphi(x) = \frac{1}{2}|x|^2$. When π is a finite measure and

$$\varphi(x) = \int_{\mathbf{R}^d} (1 - e^{i(x,y)}) \pi(dy), \tag{3.4}$$

the corresponding semigroup (resp., Lévy process) is called the compound Poisson semigroup (resp., compound Poisson process). In this case, let $\lambda = \pi(\mathbf{R}^d)$ and $\pi_0 = \lambda^{-1}\pi$. At each step, the process will stay freezing at a position x for an exponentially distributed time with parameter λ and then jump to somewhere according to distribution $\pi_0(\cdot - x)$.

For a Lévy process X on $\mathbf{R}^d$ with convolution semigroup $\{\nu_t\}$, the Lebesgue measure m is always an invariant measure for X, since it is easy to check that $\int_{\mathbf{R}^d} m(dx)\nu_t(A-x) = m(A)$ for any Borel subset A. X is called **symmetric** if

$$\nu_t(-A) = \nu_t(A) \tag{3.5}$$

for any Borel subset A of $\mathbf{R}^d$. It can be seen that in this case

$$m(dx)\nu_t(dy - x) = m(dy)\nu_t(dx - y). \tag{3.6}$$

Clearly X is symmetric if and only if its Lévy exponent φ is real, i.e.,

$$\varphi(x) = \frac{1}{2}(Sx, x) + \int_{\mathbf{R}^d} (1 - \cos(x, y))\pi(dy), \;\; x \in \mathbf{R}^d. \tag{3.7}$$

Theorem 3.4 If X is symmetric, then any Radon invariant measure of X is a multiple of Lebesgue measure if and only if its Lévy exponent φ has unique zero.

4 Fine topology and balayage

The Blumenthal 0-1 law is easy to prove but very important.

Theorem 4.1 (Blumenthal) For any $A \in \mathscr{F}_0$ and $x \in E$, $\mathsf{P}^x(A)$ is either zero or one.

Proof. For any probability μ on E, there exists $B \in \mathscr{F}_0^0$ such that $\mathsf{P}^\mu(A \triangle B) = 0$. By Markov property, $\mathsf{P}^\mu(\theta_0^{-1}A \triangle \theta_0^{-1}B) = 0$. Since $\theta_0^{-1}B = B$, $\mathsf{P}^\mu(\theta_0^{-1}A \triangle A) = 0$. Then by Markov property again, for $x \in E$,

$$\mathsf{P}^x(A) = \mathsf{P}^x(A \cap \theta_0^{-1}A) = \mathsf{E}^x[\mathsf{P}^{X_0}(A); A] = (\mathsf{P}^x(A))^2$$

and it follows that $\mathsf{P}^x(A) = 0$ or 1. $\square$

If Blumethal 0-1 law was only talking about a set in $\mathscr{F}_0^0$, it would mean nothing. Its importance is due to the fact that $\mathscr{F}_0$ is much richer than $\mathscr{F}_0^0$.

Before we go any further we should answer a question: for what kind of subset B of E, the hitting time T_B is a stopping time for the augmented filtration $(\mathscr{F}_t)$? Let's start from two basic results. Given a filtration $(\mathscr{M}_t)$ and a measurable space $(S, \mathscr{S})$, an S-valued stochastic process (Y_t) is $(\mathscr{M}_t)$-**progressively measurable** if for any $t \geq 0$, $(s, \omega) \mapsto Y_s(\omega)$ is $\mathscr{B}([0, t]) \times \mathscr{M}_t/\mathscr{S}$-measurable.

Exercise 4.1 If Y is $(\mathscr{M}_t)$-progressively measurable and $\varphi : S \to \mathbf{R}$ is Borel measurable, then so is $t \mapsto \varphi \circ Y_t$.

For a set $A \subset E$, the **hitting time** and **entrance time** for A of X are $T_A = \inf\{t > 0 : X_t \in A\}$ and $D_A = \inf\{t \geq 0 : X_t \in A\}$. It is easy to see that T_A is a terminal time, i.e., almost surely $T_A \circ \theta_t + t = T_A$ on $\{T_A > t\}$ for all $t > 0$.

Lemma 4.2 (1) A right continuous and adapted process is progressively measurable. Therefore (X_t) is progressively measurable. (2) If the filtration satisfies the usual condition, the hitting time of a real progressively measurable process for a Borel set is a stopping time.

For $f \in \mathbf{S}^\alpha$, $t \mapsto f(X_t)$ is right continuous and so f is optional. Let $\mathscr{E}^e$ denote the σ-algebra generated by all excessive functions. Then

$$\mathscr{E} \subset \mathscr{E}^e \subset \mathscr{E}^{no} \subset \mathscr{E}^*. \tag{4.1}$$

Theorem 4.3 If A is nearly optional, then the hitting time T_A is an $(\mathscr{F}_t)$-stopping time.

Proof. By the definition and lemma above, T_A is an $(\mathscr{F}_t^\mu)$-stopping time for any probability μ on E and hence an $(\mathscr{F}_t)$-stopping time. □

For any $A \in \mathscr{E}^{no}$, since $\{T_A = 0\} \in \mathscr{F}_0$, $\mathsf{P}^x(T_A = 0) = 0$ or 1 for each $x \in E$, by Blumenthal 0-1 law. If it equals 1, we say x is regular for A or otherwise x is irregular for A. Let A^r denote the set of regular points for A. A nearly optional subset G of E is called **finely open**, if for any $x \in G$, $\mathsf{P}^x(T_{G^c} = 0) = 0$ or equivalently x is irregular for G^c. Intuitively G is finely open if X, starting from any point in G, will not leave G immediately. It is routine to show that the set of finely open subsets in E is a topology, which we call the **fine topology** of X on E. Since X is right continuous, any point in an open subset G will not leave G immediately and hence an open set is finely open, namely, the fine topology is really **finer** than the original topology on E. The fine topology carries some intrinsic characteristics of the process and is usually hard to trace. The following theorem presents a lot of information on fine topology.

Theorem 4.4 (1) If f is nearly optional, then f is finely continuous if and only if $t \mapsto f(X_t)$ is right continuous. (2) If $f \in \mathbf{S}^\alpha$, f is finely continuous. (3) For $\alpha > 0$, the fine topology is generated by $\mathbf{S}^\alpha$.

Exercise 4.2 For $A \in \mathscr{E}^{no}$, A^r is finely closed and $A \cup A^r$ is the fine closure of A.

Theorem 4.5 For $A \in \mathscr{E}^{no}$, $X_{T_A} \in A \cup A^r$ on $\{T_A < \infty\}$ almost surely.

Proof. By definition of T_A, $\{X_{T_A} \notin A\} \subset \{T_A \circ \theta_{T_A} = 0\}$. Hence for any $x \in E$, using strong Markov property

$$\begin{aligned}
&\mathsf{P}^x(X_{T_A} \notin A \cup A^r, T_A < \infty) \\
&\quad = \mathsf{P}^x(X_{T_A} \notin A \cup A^r, T_A \circ \theta_{T_A} = 0, T_A < \infty)
\end{aligned}$$

$$= \mathsf{E}^x[\mathsf{P}^{X_{T_A}}(T_A = 0); X_{T_A} \notin A \cup A^r, T_A < \infty] = 0,$$

since $\mathsf{P}^{X_{T_A}}(T_A = 0) = 0$ for $X_{T_A} \notin A^r$. □

For an $(\mathscr{F}_t)$-stopping time T, define α-balayage kernel

$$P_T^\alpha(x, A) = \mathsf{E}^x\left[e^{-\alpha T} 1_A(X_T)\right], \ x \in E, A \in \mathscr{E}^*. \tag{4.2}$$

When $\alpha = 0$, this means $P_T(x, A) = \mathsf{P}^x(X_T \in A, T < \infty)$. If $T = T_A$, write P_T^α as P_A^α.

Lemma 4.6 For $g \in \mathscr{E}_+^*$,

$$P_T^\alpha U^\alpha g(x) = \mathsf{E}^x\left[\int_T^\infty e^{-\alpha t} g(X_t) dt\right]. \tag{4.3}$$

Proof. By strong Markov property,

$$\begin{aligned} P_T^\alpha U^\alpha g(x) &= \mathsf{E}^x\left[e^{-\alpha T} \mathsf{E}^{X_T}\left(\int_0^\infty e^{-\alpha t} g(X_t) dt\right)\right] \\ &= \mathsf{E}^x\left[\int_0^\infty e^{-\alpha(T+t)} f(X_{t+T}) dt\right] \\ &= \mathsf{E}^x\left[\int_T^\infty e^{-\alpha t} f(X_t) dt\right]. \end{aligned}$$

□

Lemma 4.7 (1) If $f \in \mathbf{S}^\alpha$, then $P_T^\alpha f \le f$. (2) If, in addition, T is a hitting time, then $P_T^\alpha(\mathbf{S}^\alpha) \subset \mathbf{S}^\alpha$.

Proof. (1) Assume that $f(x) < \infty$. Since $t \mapsto e^{-\alpha t} f(X_t)$ is a non-negative super-martingale, we then apply the Doob's sampling theorem to get the conclusion.

(2) By Markov property, we have

$$P_t^\alpha P_T^\alpha = P_{t+T\circ\theta_t}^\alpha.$$

For T is a terminal time, $T\circ\theta_t + t \ge T$ for all $t \ge 0$ a.s. Hence if $f = U^\alpha g$, it is obvious that $P_T^\alpha U^\alpha g$ is α-super-median. T is a hitting time so $T\circ\theta_t + t \downarrow T$ as $t \downarrow 0$ and then

$$P_T^\alpha(U^\alpha \mathscr{E}_+^*) \subset \mathbf{S}^\alpha.$$

Finally the conclusion follows from Lemma 2.6(3) and Lemma 2.5(3). □

Definition 4.8 Let $A \in \mathscr{E}^{no}$. It is **polar** if $\mathsf{P}^x(T_A < \infty) = 0$ for all $x \in E$, **thin** if $\mathsf{P}^x(T_A > 0) = 1$ for all $x \in E$ and **semi-polar** if A is contained in a countable union of thin sets. A universally measurable subset A is **potential zero** if $U(x, A) = 0$ for all $x \in E$. The definition may apply to any subset if it is contained in a set with the respective property.

Intuitively, almost surely, X never meets a polar set and amount of time in a set of potential zero has Lebesgue measure zero. Therefore a polar set is potential zero. The following theorem asserts that semipolar sets are somewhat between.

Theorem 4.9 If A is a semipolar set, then almost surely $\{t : X_t \in A\}$ is at most countable.

Proof. Assume that A is thin. Let $0 < a < 1$, and $B = \{x \in A : P^1_A 1(x) \leq a\}$. Set $T_1 = T_B$, $T_{n+1} = T_n + T_1 \circ \theta_{T_n}$. It is enough to show that $T_n \to \infty$ a.s. Since B is thin, $B^r = \emptyset$ and $X_{T_n} \in B$ for $T_n < \infty$ by Theorem 4.5. By strong Markov property

$$\begin{aligned}\mathsf{E}^x[e^{-T_{n+1}}] &= \mathsf{E}^x[e^{-T_n}(e^{-T_1})\circ\theta_{T_n}] \\ &= \mathsf{E}^x[e^{-T_n}\mathsf{E}^{X_{T_n}}(e^{-T_B})] \\ &\leq \mathsf{E}^x[e^{-T_n}\mathsf{E}^{X_{T_n}}(e^{-T_A})] \\ &\leq a\mathsf{E}^x[e^{-T_n}],\end{aligned}$$

and hence $\mathsf{E}^x[e^{-T_n}] \to 0$, i.e., $T_n \to \infty$ a.s. □

Hence it is evident that a semipolar set is potential zero.

Theorem 4.10 If A is nearly optional, then $A \setminus A^r$ is semipolar.

Proof. $P^1_A 1$ is 1-excessive and finely continuous. Let

$$A_n = \{x \in A : P^1_A 1(x) \leq 1 - 1/n\}.$$

Then $A \setminus A^r = \bigcup_n A_n$ and it suffices to verify that A_n is thin. For any $x \in E$, if $P^1_A 1(x) < 1$, then $P^1_{A_n} 1 < 1$ or $x \notin A^r_n$. If $P^1_A 1(x) = 1$, then x is in the finely open set $\{P^1_A 1(x) > 1 - 1/n\}$, which is disjoint with A_n, and hence $x \notin A^r_n$. This means that A_n is thin. □

Exercise 4.3 If f is α-super-median, define $\bar{f} = \lim_{t\downarrow 0} e^{-\alpha t} P_t f$. Show that $\bar{f} \in \mathbf{S}^\alpha$, $f \geq \bar{f}$ and $\{f > \bar{f}\}$ is potential zero.

Definition 4.11 X or (P_t) is called **transient** if U is proper, i.e., there exists a strictly positive $g \in \mathscr{E}^*$ such that $Ug < \infty$.

Since 0-potential of the semigroup $(e^{-\alpha t}P_t)$ is U^α which is proper when $\alpha > 0$, $(e^{-\alpha t}P_t)$ is always transient when $\alpha > 0$.

Lemma 4.12 If X is transient, then there exists strictly positive f such that $Uf \leq 1$.

Proof. Let g be as in the definition. Set

$$A_n = \{g \geq \frac{1}{n}, Ug \leq n\}$$

for $n \geq 1$. Then $A_n \uparrow E$. Clearly $1_{A_n} \leq ng$ and A_n is contained in a finely closed set $\{Ug \leq n\}$. By Theorem 4.5, $X_{T_{A_n}} \in \{Ug \leq n\}$ for $T_{A_n} < \infty$. Now

$$\begin{aligned} U1_{A_n}(x) &= \mathsf{E}^x \int_{T_{A_n}}^{\infty} 1_{A_n}(X_t)dt \\ &= P_{A_n} U1_{A_n}(x) \\ &\leq nP_{A_n} Ug(x) \\ &\leq n\mathsf{E}^x[Ug(X_{T_{A_n}}), T_{A_n} < \infty] \leq n^2. \end{aligned}$$

Then write $f = \sum_n 2^{-n} n^{-2} 1_{A_n}$ which is strictly positive and $Uf(x) \leq 1$. □

It is shown in the proof that there exists $A_n \in \mathscr{E}^*$ such that $A_n \uparrow E$ and each $U1_{A_n}$ is bounded.

Theorem 4.13 If X is transient and $f \in \mathbf{S}$, then there exist $g_n \in \mathscr{E}^*_+$ such that $Ug_n \uparrow f$ and both g_n and Ug_n are bounded for each n.

Proof. Take A_n as above. Set

$$h_n = n1_{A_n},\ f_n = (Uh_n) \wedge f,\ g_{n,k} = k(f_n - P_{1/k} f_n).$$

Then Uh_n is bounded, $Uh_n \uparrow +\infty$ and

$$\begin{aligned} \int_0^t P_s g_{n,k} \mathrm{d}s &= k\Big(\int_0^t P_s f_n \mathrm{d}s - \int_{\frac{1}{k}}^{t+\frac{1}{k}} P_s f_n \mathrm{d}s\Big) \\ &= k\Big(\int_0^{\frac{1}{k}} P_s f_n \mathrm{d}s - \int_t^{t+\frac{1}{k}} P_s f_n \mathrm{d}s\Big). \end{aligned}$$

Since $P_t f_n \leq P_t U h_n = \int_t^\infty P_s h_n ds \downarrow 0$ as $t \uparrow \infty$, $Ug_{n,k}$ increases with both n and k. Hence $Ug_{n,n} \uparrow f$ as $n \uparrow \infty$. □

5 Symmetric Borel right processes

Let X be a Borel right process on E with transition semigroup (P_t).

Definition 5.1 Let m be a σ-finite measure on $(E, \mathscr{E})$. The process X is called symmetric with respect to m if for any non-negative measurable functions f, g and $t \geq 0$,

$$\int g(x) P_t f(x) m(dx) = \int f(x) P_t g(x) m(dx).$$

If we write the inner product of $f, g \in L^2(E; m)$ as $\langle f, g \rangle$, this means

$$\langle P_t f, g \rangle = \langle f, P_t g \rangle.$$

It follows that m is an excessive measure, namely $mP_t \leq m$ for any $t > 0$.

Lemma 5.2 If ξ is an **excessive measure** and $p \geq 1$, then (P_t) may be extended to a contraction semigroup on $L^p(E;\xi)$.

Proof. For $f, g \in \mathscr{E}$ with $f = g$ ξ-a.e., since $\xi|P_t f - P_t g| \leq \xi P_t|f-g| \leq \xi|f-g| = 0$, $P_t f = P_t g$ ξ-a.e. Hence for any $f \in L^p(E;\xi)$ with $p \geq 1$, $P_t f$ does not depend on any particular version of f. By Hölder's inequality

$$\begin{aligned}|P_t f(x)| &= \left|\int_E P_t(x,dy)f(y)\right| \\ &\leq \int_E P_t(x,dy)|f(y)| \leq \left(\int_E P_t(x,dy)|f(y)|^p\right)^{1/p}.\end{aligned}$$

Hence we have

$$\|P_t f\|_{L^p}^p = \int_E |P_t f(x)|^p \xi(dx) \leq \int_E P_t(|f|^p)(x)\xi(dx) = \xi P_t(|f|^p) \leq \|f\|_{L^p}^p,$$

i.e., (P_t) is also a contraction semigroup on $L^p(E;\xi)$. □

Theorem 5.3 (P_t) is a strongly continuous contraction semigroup on $L^2(E;m)$.

Proof. Take $\alpha > 0$ and set

$$D = \{U^\alpha f: \ f \in b\mathscr{E} \cap L^1(E;m)\} \subset L^2(E;m).$$

Then D is dense in $L^2(m)$. Indeed, it suffices to show that if $g \in L^2(E;m)$ and $\langle g, U^\alpha f\rangle = 0$ for all $f \in b\mathscr{E} \cap L^1(E;m)$, then $g = 0$ a.e. By the resolvent equation, it follows that $\langle g, U^\beta f\rangle = 0$ for all $\beta > 0$. Choose $h = U^1 k$ where $k \in b\mathscr{E} \cap L^1(E;m)$ and strictly positive. Then for any bounded $f \in C(E)$, $t \mapsto f(X_t)h(X_t)$ is right continuous and hence $\beta U^{\beta+1}(fh) \longrightarrow fh$ a.e. as $\beta \to \infty$. However $\langle g, \beta U^{\beta+1}(fh)\rangle = 0$. Since h is 1-excessive,

$$\beta|g \cdot U^{\beta+1}(fh)| \leq \beta|g|U^{\beta+1}|fh| \leq \|f\|_\infty |g| \beta U^{\beta+1} h \leq \|f\|_\infty |g| h.$$

It follows from the dominated convergence theorem that

$$\langle g, f\rangle_{h\cdot m} = \langle g, fh\rangle = \lim_{\beta\to\infty} \langle g, \beta U^{\beta+1}(fh)\rangle = 0.$$

Since $h \cdot m$ is a finite measure, $C(E)$ is dense in $L^2(E; h\cdot m)$ and then $g = 0$ a.e. m.

We now prove the strong continuity. Fix $\alpha > 0$.Let $u = U^\alpha f \in D$. Then

$$u(x) = \int_0^t e^{-\alpha s} P_s f(x) ds + e^{-\alpha t} P_t u(x)$$

and, as $t \downarrow 0$,

$$\|P_t u - u\|_{L^2} \leq (1 - e^{-\alpha t})\|P_t u\|_{L^2} + \left\|e^{-\alpha t} P_t u - u\right\|_{L^2}$$

$$\leq (1 - e^{-\alpha t})\|u\|_{L^2} + t\|f\|_{L^2} \to 0.$$

For any $u \in L^2(E; m)$, take $u_n \in D$ such that $u_n \to u$ in L^2. Then

$$\begin{aligned}\|P_t u - u\|_{L^2} &\leq \|P_t u - P_t u_n\|_{L^2} + \|P_t u_n - u_n\|_{L^2} + \|u_n - u\|_{L^2} \\ &\leq 2\|u_n - u\|_{L^2} + \|P_t u_n - u_n\|_{L^2}.\end{aligned}$$

It follows that (P_t) is strongly continuous. □

Let $(L, D(L))$ be the infinitesimal generator of strongly continuous contraction semigroup (P_t) on $L^2(E; m)$. By Theorem 5.3, the bilinear form $(\mathscr{E}, \mathscr{F})$ defined by

(5.1) $$\mathscr{E}(f, g) = (\sqrt{-L}f, \sqrt{-L}g),\ \ \mathscr{F} = D(\sqrt{-L}),$$

is a symmetric form on $L^2(E; m)$ and it may be represented by its approximating form

(5.2) $$\begin{cases}\mathscr{E}(f, g) = \lim_{t\downarrow 0} \dfrac{1}{t}\langle f - P_t f, g\rangle, \\ \mathscr{F} = \left\{ f \in L^2(E; m) : \sup_t \dfrac{1}{t}\langle f - P_t f, f\rangle < \infty \right\}.\end{cases}$$

Recall that we usually write

$$\mathscr{E}^{(t)}(f, g) = \frac{1}{t}\langle f - P_t f, g\rangle,\ \ \mathscr{E}^{[\beta]}(f, g) = \beta\langle f - \beta U^\beta f, g\rangle.$$

Theorem 5.4 $(\mathscr{E}, \mathscr{F})$ is a Dirichlet form on $L^2(E; m)$.

Proof. It suffices to prove that $(\mathscr{E}, \mathscr{F})$ is Markovian. By symmetry, $m(dx)P_t(x, dy)$ is symmetric and then for $f \in \mathscr{E}^*$,

$$\begin{aligned}\mathscr{E}^{(t)}(f, f) &= \frac{1}{t}\langle f, f - P_t f\rangle \\ &= \frac{1}{t}\int_E f(x)\left(f(x) - (P_t f)(x)\right) m(\mathrm{d}x) \\ &= \frac{1}{t}\int_E f(x)\left(f(x) - \int_E f(y)P_t(x, dy)\right) m(dx) \\ &= \frac{1}{t}\int_E f(x)\left(\int_E (f(x) - f(y))P_t(x, dy)\right) m(\mathrm{d}x) + \frac{1}{t}\int_E f^2\left(1 - P_t 1\right)\mathrm{d}m \\ &= \frac{1}{t}\int_{E\times E} f(x)\left(f(x) - f(y)\right) P_t(x, \mathrm{d}y) m(\mathrm{d}x) + \frac{1}{t}\int_E f^2\left(1 - P_t 1\right)\mathrm{d}m \\ &= \frac{1}{2t}\int_{E\times E} \left(f(x) - f(y)\right)^2 P_t(x, \mathrm{d}y) m(\mathrm{d}x) + \frac{1}{t}\int_E f^2\left(1 - P_t 1\right)\mathrm{d}m\end{aligned}$$

If g is a normal contraction of f, it is then obvious that

$$\mathscr{E}^{(t)}(g, g) \leq \mathscr{E}^{(t)}(f, f)$$

and hence $f \in \mathscr{F}$ implies that $g \in \mathscr{F}$ and $\mathscr{E}(g, g) \leq \mathscr{E}(f, f)$. □

6 Irreducibility and uniqueness of symmetrizing measure

When the process X is symmetric with respect to m, m is called a **symmetrizing measure** of X. The existence and uniqueness of symmetrizing measures of X are always interesting to explore. In this section we shall introduce the notion of fine irreducibility and prove that it implies the uniqueness. The process X is called *finely irreducible* if $\mathsf{P}^x(T_D < \infty) > 0$ for any $x \in E$ and any non-empty finely open subset D, where T_D is the hitting time of D. Intuitively the fine irreducibility means that any point can reach any non-empty finely open set, while the usual irreducibility means that any point can reach any non-empty open set. Certainly the fine irreducibility is stronger than the usual irreducibility. Since the fine irreducibility is hard to be characterized, we shall give a few equivalent statements which may be useful in some circumstances.

Lemma 6.1 The following statements are equivalent.
(1) X is finely irreducible.
(2) $U^\alpha 1_D$ is positive everywhere on E for any non-empty finely open set D.
(3) $U^\alpha 1_A$ is either identically zero or positive everywhere on E for any Borel set A or, in other words, $\{U^\alpha(x,\cdot) : x \in E\}$ are all mutually absolutely continuous.
(4) All non-trivial excessive measures are mutually absolutely continuous.

Proof. The equivalence of (1) and (2) is easy. We shall prove that they are equivalent to (3). We may assume $\alpha = 0$. Suppose (1) is true. If $U1_A$ is not identically zero, then there exists $\delta > 0$ such that $D := \{U1_A > \delta\}$ is non-empty. Since $U1_A$ is excessive and thus finely continuous, D is finely open and the fine closure of D is contained in $\{U1_A \geq \delta\}$. Then by Lemma 4.7,

$$U1_A(x) \geq P_D U1_A(x) = \mathsf{E}^x\left(U1_A(X_{T_D})\right) \geq \delta \mathsf{P}^x(T_D < \infty) > 0.$$

Conversely suppose (3) is true. Then for any finely open set D, by the right continuity of X, $U1_D(x) > 0$ for any $x \in D$. Therefore $U1_D$ is positive everywhere on E.

Let ξ be an excessive measure. Since $\alpha\xi U^\alpha \leq \xi$, $\xi(A) = 0$ implies that $\xi U^\alpha(A) = 0$. However ξ is non-trivial. Thus it follows from (3) that $U^\alpha 1_A \equiv 0$, i.e., A is potential zero. Conversely if A is potential zero, then $\xi(A) = 0$ for any excessive measure ξ. Therefore (3) implies (4).

Assume (4) holds. Since $U(x,\cdot)$ is excessive for all x and hence they are equivalent. This implies (3). □

Theorem 6.2 Assume that X is finely irreducible. Then the symmetrizing measure of X is unique up to a constant. More precisely if both μ and ν are non-trivial symmetrizing measures of X, then $\nu = c\mu$ with a positive constant c.

Proof. First of all there exists a measurable set H such that both $\mu(H)$ and $\nu(H)$ are positive and finite, because μ and ν are equivalent by Lemma 6.1. This is actually true when both measures are σ-finite and one is absolutely continuous with respect to another. Indeed, assume that $\nu \ll \mu$. Since ν is non-trivial and σ-finite, we may find a measurable set B such that $0 < \nu(B) < \infty$. Then $\mu(B) > 0$. Since μ is σ-finite, there exist $A_n \uparrow E$ such that $0 < \mu(A_n) < \infty$. Then $\nu(A_n \cap B) \uparrow \nu(B)$ and $\mu(A_n \cap B) \uparrow \mu(B)$. Hence there exists some n such that $\nu(A_n \cap B) > 0$. Take $H = A_n \cap B$, which makes both $\mu(H)$ and $\nu(H)$ positive and finite.

Set $c = \nu(H)/\mu(H)$. We may assume that $c = 1$ without loss of generality. Let $m = \mu + \nu$. Then there is $f_1, f_2 \geq 0$ such $\mu = f_1 \cdot m$ and $\nu = f_2 \cdot m$. Let $A = \{f_1 > f_2\}$, $B = \{f_1 = f_2\}$ and $C = \{f_1 < f_2\}$.

We shall show that $\nu = \mu$. Otherwise $\mu(A) > 0$ or $\nu(C) > 0$. We assume that $\mu(A) > 0$ without loss of generality. Since μ is σ-finite, there is $A_n \in \mathscr{B}(E)$ such that $A_n \subseteq A$, $\mu(A_n) < \infty$ and $A_n \uparrow A$. Let $D = B \cup C$. For any integer n and $\alpha > 0$,

$$(U^\alpha 1_{A_n}, 1_D)_\mu \leq (U^\alpha 1_{A_n}, 1_D)_\nu = (U^\alpha 1_D, 1_{A_n})_\nu \leq (U^\alpha 1_D, 1_{A_n})_\mu.$$

Since $(U^\alpha 1_{A_n}, 1_D)_\mu = (U^\alpha 1_D, 1_{A_n})_\mu$, it follows that $(U^\alpha 1_D, 1_{A_n})_\nu = (U^\alpha 1_D, 1_{A_n})_\mu$. Thus we have

$$(U^\alpha 1_D, (1 - \frac{f_2}{f_1})1_{A_n})_\mu = (U^\alpha 1_D, 1_{A_n})_\mu - (U^\alpha 1_D, 1_{A_n})_\nu = 0.$$

Since $1 - \frac{f_2}{f_1} > 0$ on A, let n go to infinity and by the monotone convergence theorem we get that $(U^\alpha 1_D, 1_A)_\mu = 0$. The fine irreducibility of X implies that $U^\alpha 1_D = 0$ identically or D is of potential zero. Therefore

$$\mu(D) = \nu(D) = 0.$$

Consequently,

$$0 = \mu(H) - \nu(H) = \int_{H \cap A} (1 - \frac{f_2}{f_1}) d\mu$$

which leads to that $\mu(H \cap A) = 0$ and also $\mu(H) = 0$. The contradiction implies that $\nu = \mu$. □

The following example shows that the usual irreducibility is not enough to guarantee the uniqueness of symmetrizing measure, while the fine irreducibility might be too strong.

Example 6.3 Let

$$J = \frac{1}{4}(\delta_1 + \delta_{-1} + \delta_{\sqrt{2}} + \delta_{-\sqrt{2}})$$

defined on $\mathbb{R}$ and $\pi = \{\pi_t\}_{t>0}$ the corresponding convolution semigroup; i.e., $\widehat{\pi}_t(x) = e^{-t\phi(x)}$ with

$$\phi(x) = \int (1 - \cos xy) J(dy) = \frac{1}{2}(1 - \cos x) + \frac{1}{2}(1 - \cos \sqrt{2}x).$$

Let X be the corresponding Lévy process. Then X is symmetric with respect to the Lebesgue measure. Let $N = \{n + m\sqrt{2} \ : \ n, m \text{ are integers}\}$ and $\mu = \sum_{x \in N} \delta_x$. Then μ is σ-finite and also a symmetrizing measure. It is easy to see that any point x can reach any point in $x + N$ and cannot reach outside of $x + N$. Since $x + N$ is dense in $\mathbb{R}$, any point can reach any non-empty open set, namely X is irreducible. However any compound Poisson process will stay at the starting point for a positive period of time, i.e., any singleton is finely open. Hence X is not finely irreducible.

Another interesting example is also a compound Poisson process X, where the Lévy measure J is a probability measure on $\mathbb{R}$ with a continuous even density. In this case, we can show that X has a unique symmetrizing measure, the Lebesgue measure, but X is still irreducible, while not finely irreducible. Actually any single point can not reach any other point.

It is known that the fine topology is determined by the process and hard to identify usually. Hence it is hard to verify sometimes the fine irreducibility defined in the theorem.

Definition 6.4 We say X is **LSC** or **strong Feller**, if $U^\alpha 1_B$ is lower-semi-continuous or continuous, respectively, for any Borel subset B of E.

Lemma 6.5 If X is LSC or strong Feller, the fine irreducibility is equivalent to the usual one.

Proof. It suffices to prove that $\mathsf{P}^x(T_D < \infty) > 0$ for any $x \in E$ and non-empty open subset $D \subset E$. In fact, take $A \in \mathscr{B}(E)$ with $U1_A \neq 0$ identically. There is $b > 0$ such that $G = \{U1_A > b\} \neq \emptyset$ and is open due to the property LSC. Again by Lemma 4.7, for any $x \in E$,

$$U1_A(x) \geq P_G U1_A(x) = \mathsf{P}^x\left(U1_A(X_{T_G}), T_G < \infty\right).$$

But $X_{T_G} \in \bar{G}$ on $\{T_G < \infty\}$ by Theorem 4.5 and then $U1_A(X_{T_G}) \geq b$ on $\{T_G < \infty\}$. Hence by the irreducibility, we have

$$U1_A(x) \geq b\mathsf{P}^x\left(T_G < \infty\right) > 0.$$

□

Question If X is a Lévy process, what conditions imposed on its Lévy exponent guarantee that X is irreducible or finely irreducible?

References

[1] O. D. Kellogg, Foundations of potential theory, Springer-Verlag, Berlin-New York, 1967.

[2] P.-A. Meyer, Processus de Markov, Lecture Notes in Math. 26, Springer-Verlag, Berlin-New York, 1967.

[3] R. M. Blumenthal and R. K. Getoor, Markov processes and potential theory, Pure Appl. Math. 29 Academic Press, New York-London, 1968.

[4] R. K. Getoor, Markov processes: Ray processes and right processes, Lecture Notes in Math. 440, Springer-Verlag, Berlin-New York, 1975.

[5] C. Berg and G. Forst, Potential theory on locally compact abelian groups, Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd 87, Springer-Verlag, New York-Heidelberg, 1975.

[6] S. C. Port and C. J. Stone, Brownian motion and classical potential theory, Probability and Mathematical Statistics, Academic Press, New York-London, 1978.

[7] M. Fukushima, Dirichlet forms and Markov processes, North-Holland Math. Library 23, North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1980.

[8] K. L. Chung, From Markov processes to Brownian motion, Grundlehren Math. Wiss. 249, Springer-Verlag, New York-Berlin, 1982.

[9] J. L. Doob, Classical potential theory and its probabilistic counterpart, Grundlehren Math. Wiss. 262, Springer-Verlag, New York, 1984.

[10] J. Bliedtner and W. Hansen, Potential Theory, An analytic and probabilistic approach to balayage, Universitext, Springer-Verlag, Berlin, 1986.

[11] M. Sharpe, General theory of Markov processes, Pure Appl. Math. 133, Academic Press, Inc., Boston, MA, 1988.

[12] R. K. Getoor, Excessive measures, Probab. Appl., Birkhäuser Boston, Inc., Boston, MA, 1990.

[13] C. Dellacherie and P.-A. Meyer, Probabilités et Potentiel, Chapitres I à IV, Actualites Sci. Indust. 1372, Hermann, Paris, 1975; Chapitres V à VIII, Actualites Sci. Indust. 1385, Hermann, Paris, 1980; Chapitres IX à XI, Publ. Inst. Math. Univ. Strasbourg XVIII, Hermann, Paris, 1983; Chapitres XIIà XVI, Publ. Inst. Math. Univ. Strasbourg XIX, Hermann, Paris, 1987; Erratum, Lecture Notes in Math. 1321, Springer, Berlin, 1988.

[14] Z.-M. Ma and M. Röckner, Introduction to the theory of Dirichlet forms, Universitext, Springer-Verlag, Berlin, 1992.

[15] M. Fukushima, Y. Oshima and M. Takeda, Dirichlet forms and symmetric Markov processes, de Gruyter Stud. Math. 19, Walter de Gruyter & Co., Berlin, 1994.

ON THE CONSTRUCTION OF TORIC VARIETIES

Masanori ISHIDA

Introduction

A normal algebraic variety X is called a toric variety if it contains an algebraic torus T as an open subvariety and the group action of T on itself is extended regularly to X. It is known that the toric variety X is described by a fan of a real space of dimension $r = \dim X$ with a lattice $N \simeq \mathbf{Z}^r$. A fan is a collection of cones in the real space, and each cone gives a finitely generated subsemigroup of the dual $\mathbf{Z}$-module $M = N^*$. The coordinate ring of the corresponding affine open set of the toric variety is the semigroup ring of this semigroup. One of the important properties of a toric variety is that it is compact if and only if the associated fan is complete, i.e., it is finite and the support is the whole real space. In this note, we will introduce a variation of the construction of the toric variety from a fan. Namely, we define a topological space based on a topological commutative semigroup from a fan.

In Oda's book [O], the affine toric variety over $\mathbf{C}$ associated to a cone σ is given as the set $\mathrm{Hom}(M \cap \sigma^\vee, \mathbf{C})$ of semigroup homomorphisms $x : M \cap \sigma^\vee \to \mathbf{C}$ with $x(0) = 1$, where $\mathbf{C}$ is considered as a multiplicative semigroup (see [O, Proposition 1.2]). In Section 3, we will replace $\mathbf{C}$ by a topological commutative semigroup A with some condition. Considering the case when A is the closed unit disk in $\mathbf{C}$, we give a new explanation of the completeness condition of a fan.

Section 4 is on a different topic, which is an application of the theory of toric varieties. We consider the algebraic surfaces defined by tetrahedra without lattice points on the boundaries except vertices. Normal and nonsingular complete models of such an algebraic surface are described by embedding in the toric variety associated to the tetrahedron. By applying the Riemann-Roch theorem to the nonsingular model, we give a formula of the geometric genus of the surface which is equal to the number of the lattice points in the interior of the tetrahedron.

Acknowledgment This note is based on my talk at Fudan University in November 2010 as a course of the exchange program of Tohoku and Fudan Universities. I would like to express my sincere thanks to Professors Quanshui Wu, Jiaxing Hong and Meng Chen for giving me a chance to stay at Fudan University.

1 Cones and fans

Let r be a non-negative integer and N a free $\mathbf{Z}$-module of rank r. We consider cones and fans in the real space $N_{\mathbf{R}} = N \otimes_{\mathbf{Z}} \mathbf{R}$, which contains N as a lattice. A non-empty subset C of $N_{\mathbf{R}}$ is said to be a cone if $u \in C$ and $a \in \mathbf{R}_0$ imply $au \in C$, where $\mathbf{R}_0 = \{c \in \mathbf{R} \; ; \; c \geq 0\}$. A cone C is convex if and only if $x, y \in C$ and $a, b \in \mathbf{R}_0$ imply $ax + by \in C$. For any subset $S \subset N_{\mathbf{R}}$, the minimal convex cone containing S is

$$C = \{a_1 u_1 + \cdots + a_l u_l \; ; \; a_1, \ldots, a_l \in \mathbf{R}_0, u_1, \ldots, u_l \in S\} \; ,$$

and we say that C is generated by S. A rational polyhedral cone is a cone generated by a finite subset of N. We assume that *a cone denoted by a Greek letter*, for example σ, is always a strongly convex rational polyhedral cone, where *strongly convex* means $\sigma \cap (-\sigma) = \{0\}$.

The dual $\mathbf{Z}$-module of N is denoted by M, which is a lattice of $M_{\mathbf{R}} = M \otimes_{\mathbf{Z}} \mathbf{R}$. The natural pairing $\langle \; , \; \rangle \; ; \; M \times N \to \mathbf{Z}$ is extended to the bilinear map

$$\langle \; , \; \rangle : M_{\mathbf{R}} \times N_{\mathbf{R}} \longrightarrow \mathbf{R} \; .$$

For an element $x \in M_{\mathbf{R}}$, we denote $(x = 0) = \{u \in N_{\mathbf{R}} \; ; \; \langle x, u \rangle = 0\}$ and $(x \geq 0) = \{u \in N_{\mathbf{R}} \; ; \; \langle x, u \rangle \geq 0\}$. If $x \neq 0$, then $(x = 0)$ is a hyperplane and $(x \geq 0)$ is a closed half space of $N_{\mathbf{R}}$. We use similar notation for $u \in N_{\mathbf{R}}$. Namely, $(u = 0)$ is a hyperplane and $(u \geq 0)$ is a closed half space of $M_{\mathbf{R}}$ if $u \neq 0$. Although it is not so easy, we can show that a subset $C \subset N_{\mathbf{R}}$ is a rational polyhedral cone if and only if

$$C = (x_1 \geq 0) \cap \cdots \cap (x_s \geq 0)$$

for a finite subset $\{x_1, \ldots, x_s\}$ of the lattice M. A subset ρ of a cone σ is said to be a *face* and we denote $\rho \prec \sigma$ if there exists $x \in M_{\mathbf{R}}$ with $\sigma \subset (x \geq 0)$ and $\sigma \cap (x = 0) = \rho$.

For a cone σ in $N_{\mathbf{R}}$, we define

$$\sigma^{\vee} = \{x \in M_{\mathbf{R}} \; ; \; \langle x, u \rangle \geq 0 \text{ for all } u \in \sigma\} = \{x \in M_{\mathbf{R}} \; ; \; \sigma \subset (x \geq 0)\} \; ,$$

which is a polyhedral cone of dimension r and we call it the dual cone of σ. The linear subspace $\sigma^{\perp} = \{x \in M_{\mathbf{R}} \; ; \; \langle x, u \rangle = 0 \text{ for all } u \in \sigma\}$ is contained in $\sigma^{\vee}$. In particular, $\sigma^{\vee}$ is not strongly convex if $\dim \sigma < r$. There exists a natural bijection

$$(\text{the set of faces of } \sigma) \longrightarrow (\text{the set of faces of } \sigma^{\vee})$$

defined by $\rho \mapsto \sigma^{\vee} \cap \rho^{\perp}$, where $\dim(\sigma^{\vee} \cap \rho^{\perp}) = r - \dim \rho$.

A nonempty set X of strongly convex rational polyhedral cones is said to be a *fan* if it satisfies the following conditions.

(1) $\sigma \in X$ and $\rho \prec \sigma$ imply $\rho \in X$.
(2) If $\sigma, \tau \in X$, then $\sigma \cap \tau$ is a common face of σ and τ.

Although it is common to denote a fan by a Greek capital, it is my favor to use a Roman letter so that it can be considered as a kind of scheme. Namely, for each $\sigma \in X$, the set $F(\sigma)$ of faces of σ is an affine piece and X is the union of them. The condition (2) is the separation condition. The union $\bigcup_{\sigma \in X} \sigma$ of cones in X is called the support of X and denoted by $|X|$. We say X is complete if it is finite and $|X| = N_{\mathbf{R}}$.

For a cone σ in $N_{\mathbf{R}}$, we denote by $H(\sigma)$ the linear subspace of $N_{\mathbf{R}}$ generated by σ. Clearly, $H(\sigma)$ is equal to $\sigma + (-\sigma) = \{u - v \ ; \ u, v \in \sigma\}$. The relative interior $\mathrm{rel.\,int}\,\sigma$ of σ is the set of interior points of σ as a subset of $H(\sigma)$. Note that $\mathrm{rel.\,int}\,\sigma$ is always nonempty and any cone π is the disjoint union of $\mathrm{rel.\,int}\,\sigma$ for $\sigma \in F(\pi)$. Namely, a point u in π belongs to $\mathrm{rel.\,int}\,\sigma$ for the unique minimal face σ of π which contains u. If σ is an element of a fan X, then σ is a face of $\tau \in X$ if and only if $\tau \cap \mathrm{rel.\,int}\,\sigma \neq \emptyset$.

For a cone σ, we define also $N[\sigma] = N/(N \cap H(\sigma))$ and $M[\sigma] = M \cap \sigma^{\perp}$, which are mutually dual free $\mathbf{Z}$-modules. If σ is a face of another cone τ, we denote by $\tau[\sigma]$ the image of τ in $N[\sigma]_{\mathbf{R}} = N_{\mathbf{R}}/H(\sigma)$, which is a strongly convex rational polyhedral cone in the real space $N[\sigma]_{\mathbf{R}}$.

When σ is an element of a fan X, we denote $X(\sigma\prec) = \{\tau \in X \ ; \ \sigma \prec \tau\}$ and

$$X[\sigma] = \{\tau[\sigma] \ ; \ \tau \in X(\sigma\prec)\} \ .$$

We can see that $X[\sigma]$ is a fan of $N[\sigma]_{\mathbf{R}}$. For each $\tau \in X(\sigma\prec)$, the semigroup $M[\sigma] \cap \tau[\sigma]^{\vee}$ is equal to $M \cap \tau^{\vee} \cap \sigma^{\perp}$. Here recall that $\tau^{\vee} \cap \sigma^{\perp}$ is a face of $\tau^{\vee}$.

Proposition 1.1. *Let X be a finite fan of $N_{\mathbf{R}}$, σ an element of X. If σ intersects the interior of $|X|$, then $X[\sigma]$ is a complete fan of $N[\sigma]_{\mathbf{R}}$.*

Proof. Let $\phi_{\mathbf{R}} : N_{\mathbf{R}} \to N[\sigma]_{\mathbf{R}}$ be the natural surjection. We set

$$|X(\sigma\prec)| = \bigcup_{\tau \in X(\sigma\prec)} \tau \quad \text{and} \quad E := \bigcup_{\tau \in X \setminus X(\sigma\prec)} \tau$$

Since E is a closed subset disjoint from the relative interior of σ, we know that σ intersects the interior of $|X(\sigma\prec)|$. Since $\phi_{\mathbf{R}}$ is an open map, $|X[\sigma]| = \phi_{\mathbf{R}}(|X(\sigma\prec)|)$ contains the origin of $N[\sigma]_{\mathbf{R}}$ as an interior point. Since $|X[\sigma]|$ is a cone, we have $|X[\sigma]| = N[\sigma]_{\mathbf{R}}$ and $X[\sigma]$ is complete. *q.e.d.*

2 Fan spaces over a semigroup

Let A be a topological commutative semigroup with unity. Namely, A is a commutative multiplicative semigroup with $1 \in A$, and is a topological space such that the map $A \times A \to$

A defined by $(a, b) \mapsto ab$ is continuous. Also, A might be an additive group with $0 \in A$. We will further assume the following conditions.

(a) The set $A^\times$ of invertible elements in A is open or closed in A and the map $A^\times \to A^\times$ defined by $z \mapsto z^{-1}$ is continuous.

(b) There exists an element $0 \in A$ such that $ab = 0$ if and only if $a = 0$ or $b = 0$, and the one point set $\{0\}$ is closed in A.

We mainly consider the following examples of A, where the element 0 in (b) is ∞ for the additive semigroups (6) and (7).

(1) The complex plane $\mathbf{C}$ as a multiplicative semigroup.

(2) The closed unit disk $D = \{z \in \mathbf{C}\ ;\ |z| \leq 1\}$ as a multiplicative semigroup.

(3) The real line $\mathbf{R}$ as a multiplicative semigroup.

(4) $\mathbf{R}_0 = \{a \in \mathbf{R}\ ;\ a \geq 0\}$ as a multiplicative semigroup.

(5) $[0, 1]$ as a multiplicative semigroup.

(6) $\overline{\mathbf{R}} = \mathbf{R} \cup \{\infty\}$ as an additive semigroup where $a + \infty = \infty$ for every $a \in \overline{\mathbf{R}}$. The subsets $(a, \infty) \cup \{\infty\}$ for $a \in \mathbf{R}$ are fundamental open neighborhoods of ∞.

(7) $\overline{\mathbf{R}}_0 = \mathbf{R}_0 \cup \{\infty\}$ as an additive semigroup. This is the one-point compactification of $\mathbf{R}_0$.

Note that there are surjective homomorphisms $\mathbf{C} \to \mathbf{R}_0$ and $D \to [0, 1]$ defined by $z \to |z|$, and isomorphisms $\mathbf{R}_0 \to \overline{\mathbf{R}}$ and $[0, 1] \to \overline{\mathbf{R}}_0$ defined by $a \mapsto -\log a$, while there is the inclusion map $\mathbf{R} \hookrightarrow \mathbf{C}$.

For a cone σ in $N_{\mathbf{R}}$, we consider the additive semigroup $M \cap \sigma^\vee$ and denote by $F(\sigma)_A$ the set $\mathrm{Hom}_{\mathrm{sgr}}(M \cap \sigma^\vee, A)$ of semigroup homomorphisms such that 0 is mapped to the unity. Let $\{m_1, \ldots, m_l\}$ be a set of generators of the semigroup $M \cap \sigma^\vee$. Then the map $F(\sigma)_A \to A^l$ defined by $\alpha \mapsto (\alpha(m_1), \ldots, \alpha(m_l))$ is an injection. We consider the topology of $F(\sigma)_A$ induced from the product topology of A^l.

Lemma 2.1. *The topology on $F(\sigma)_A$ does not depend on the choice of the finite set of generators $\{m_1, \ldots, m_l\}$.*

Proof. Since any subset of $M \cap \sigma^\vee$ containing a set of generators is also a set of generators, we should show that $\{m_1, \ldots, m_l\}$ and $\{m_1, \ldots, m_l, m_{l+1}, \ldots, m_s\}$ define the same topology on $F(\sigma)_A$ for any $m_{l+1}, \ldots, m_s \in M \cap \sigma^\vee$. If we use induction, it suffices to show the case $s = l + 1$. Since $\{m_1, \ldots, m_l\}$ is a set of generators, we have an expression $m_{l+1} = c_1 m_1 + \cdots + c_l m_l$ with $c_1, \ldots, c_l \geq 0$.

The induced topologies are the weakest topologies on $F(\sigma)_A$ such that the maps $\alpha \mapsto \alpha(m_i)$ are continuous for $i = 1, \ldots, l$ and $i = 1, \ldots, l + 1$, respectively. Since the map $A^l \to A$ defined by $(x_1, \ldots, x_l) \mapsto x_1^{c_1} \cdots x_l^{c_l}$ is continuous, these conditions are equivalent, and hence they define a same topology. *q.e.d.*

Let ρ be a face of σ. Then, since $M\cap\sigma^\vee \subset M\cap\rho^\vee$, there exists a map $\phi_{\sigma/\rho} : F(\rho)_A \to F(\sigma)_A$ defined by the restriction $x \mapsto x|M\cap\sigma^\vee$. In the proof of the following lemma, we use the fact that, if we take a point z in the relative interior of $\sigma^\vee \cap \rho^\perp$, the face $\rho \prec \sigma$ is defined by z and we have $\rho = \sigma \cap (-z \geq 0)$ and

$$\rho^\vee = (\sigma \cap (-z \geq 0))^\vee = \sigma^\vee + \mathbf{R}_0(-z)\ .$$

Lemma 2.2. *This map $\phi_{\sigma/\rho}$ is injective. An element $y \in F(\sigma)_A$ is in the image if and only if $y(m) \in A$ is invertible for every $m \in M \cap \sigma^\vee \cap \rho^\perp$.*

Proof. Take an element m_0 in $M\cap \mathrm{rel.\,int}(\sigma^\vee\cap\rho^\perp)$. Then, for any element $m \in M\cap\rho^\vee$, there exists a positive integer a and $m' \in M \cap \sigma^\vee$ such that $m = m' - am_0$. Hence for $x \in F(\rho)_A$, we have $x(m') = x(m)x(m_0)^a$. Since $M \cap \rho^\perp$ is a group and m_0 is in it, we have $-m_0 \in M \cap \rho^\vee$ and $x(m_0)$ has the inverse $x(-m_0)$. Hence $x(m) = x(m')x(m_0)^{-a}$. Since $m', m_0 \in M \cap \sigma^\vee$, this means that x is determined by its image in $F(\sigma)_A$, i.e., this map is injective. Since $m_0 \in M \cap \mathrm{rel.\,int}(\sigma^\vee \cap \rho^\perp)$, the homomorphism y can be extended to $M \cap \rho^\vee$ if and only if $y(m_0)$ is invertible. For $m \in M \cap \sigma^\vee \cap \rho^\perp$, $y(m)$ is invertible if so is $y(m_0)$, since there exists a positive integer a with $am_0 - m \in M \cap \sigma^\vee \cap \rho^\perp$ and $y(m_0)^a = y(m)y(am_0 - m)$. Hence we get the second assertion. *q.e.d.*

For a cone σ in $N_\mathbf{R}$ and its face ρ, we regard $F(\rho)_A$ as a subset of $F(\sigma)_A$ by the injection $\phi_{\sigma/\rho}$.

Lemma 2.3. *Let ρ be a face of σ. then the topology of $F(\rho)_A$ is equal to the topology induced from $F(\sigma)_A$. If $A^\times$ is open (resp. closed) in A, then $F(\rho)_A$ is open (resp. closed) in $F(\sigma)_A$.*

Proof. Let $\{m_1, \ldots, m_l\}$ be a set of generators of $M \cap \sigma^\vee$. We may assume that $m_1, \ldots, m_s$ are in $M \cap \sigma^\vee \cap \rho^\perp$ and others are not for some $0 \leq s \leq l$. By Lemma 2.2, the topology of $F(\rho)_A$ is the weakest topology such that maps $x \mapsto x(m_i)$ and $x \mapsto x(-m_i)$ for $i = 1, \ldots, s$ and $x \mapsto x(m_i)$ for $i = s+1, \ldots, l$ are continuous. Since $x(-m_i) = x(m_i)^{-1}$ for $i = 1, \ldots, s$, this is equivalent to the condition that $x \mapsto x(m_i)$ for $i = 1, \ldots, l$ are continuous by the condition (a) of A. Hence it is the induced topology of $F(\sigma)_A$. Since $F(\rho)_A$ is equal to $\{x \in M \cap \sigma^\vee\ ;\ x(m_i) \in A^\times \text{ for } i = 1, \ldots, s\}$ by Lemma 2.2, we have $F(\rho)_A = F(\sigma)_A \cap ((A^\times)^s \times A^{l-s})$. Hence, it is open if $A^\times$ is open, and closed if $A^\times$ is so. *q.e.d.*

3 Compact toric varieties

Let M and N be mutually dual free $\mathbf{Z}$-modules of rank $r \geq 0$ as before. We define the morphism

$$\mu_N : T_N \longrightarrow N_\mathbf{R}$$

from $T_N = N \otimes \mathbf{C}^\times$ to the real space $N_\mathbf{R} = N \otimes \mathbf{R}$ by $\mu_N = 1_N \otimes (-\log|\cdot|)$. Namely, if we consider $T_N = (\mathbf{C}^\times)^r$ and $N_\mathbf{R} = \mathbf{R}^r$ by fixing a basis of N, then

$$\mu_N(t_1, \ldots, t_r) = (-\log|t_1|, \ldots, -\log|t_r|) .$$

This is a continuous surjective homomorphism of groups. We can write also $T_N = \mathrm{Hom}_{\mathrm{gr}}(M, \mathbf{C}^\times)$ and $N_\mathbf{R} = \mathrm{Hom}_{\mathrm{gr}}(M, \mathbf{R})$. In this description, μ_N is the morphism mapping $\alpha \in \mathrm{Hom}_{\mathrm{gr}}(M, \mathbf{C}^\times)$ to $-\log|\alpha| \in \mathrm{Hom}_{\mathrm{gr}}(M, \mathbf{R})$.

Lemma 3.1. *Let σ be a cone of $N_\mathbf{R}$. For x in T_N, $\mu_N(x)$ is a point of σ if and only if $|\mathbf{e}(m)(x)| \leq 1$ for every $m \in M \cap \sigma^\vee$.*

Proof. Since $\mu_N(x) \in N_\mathbf{R} = \mathrm{Hom}_{\mathrm{gr}}(M, \mathbf{R})$ for $x \in T_N$, we have

$$-\log|\mathbf{e}(m)(x)| = \langle m, \mu_N(x) \rangle .$$

From this equality, $|\mathbf{e}(m)(x)| \leq 1$ is equivalent to $\langle m, \mu_N(x) \rangle \geq 0$.

If we replace the inequality $|\mathbf{e}(m)(x)| \leq 1$ in the lemma by $\langle m, \mu_N(x) \rangle \geq 0$, the second condition is $\langle m, \mu_N(x) \rangle \geq 0$ for every $m \in M \cap \sigma^\vee$. Since $\sigma^\vee$ is generated by $M \cap \sigma^\vee$ and $\sigma^{\vee\vee} = \sigma$, this is equivalent to the first condition $\mu_N(x) \in \sigma$. *q.e.d.*

For the closed unit disk D in the complex plane, we get the following lemma.

Lemma 3.2. *$F(\sigma)_D$ is compact for every cone σ.*

Proof. Let $\{m_1, \ldots, m_s\}$ be a set of generators of the semigroup $M \cap \sigma^\vee$. Any element $m \in M \cap \sigma^\vee$ can be expressed as

$$m = c_1 m_1 + \cdots + c_s m_s$$

by non-negative integers $c_1, \ldots, c_s$. Then, we have

$$\alpha(m) = \alpha(m_1)^{c_1} \cdots \alpha(m_s)^{c_s}$$

for any $\alpha \in F(\sigma)_\mathbf{C}$. Hence, α is in $F(\sigma)_D$ if and only if the s complex numbers $\alpha(m_1), \ldots, \alpha(m_s)$ are of the absolute values at most one. On the other hand, since the coordinate ring $\mathbf{C}[M \cap \sigma^\vee]$ of $F(\sigma)_\mathbf{C}$ is generated by $\{\mathbf{e}(m_1), \ldots, \mathbf{e}(m_s)\}$, we can regard $F(\sigma)_\mathbf{C}$ an affine algebraic variety defined in $\mathbf{C}^s$ with this set of coordinate functions. Then, since

$$F(\sigma)_D = F(\sigma)_\mathbf{C} \cap D^s ,$$

$F(\sigma)_\mathbf{C} \subset \mathbf{C}^s$ is closed and D^s is a bounded closed set, $F(\sigma)_D$ is also a bounded closed subset, and hence is compact. *q.e.d.*

Let X be a fan. We introduce an equivalence relation in $\coprod_{\sigma\in X} F(\sigma)_A/\sim$ as follows. For $x \in F(\sigma)_A$ and $y \in F(\tau)_A$, we define $x \sim y$ if there exists $z \in F(\eta)_A$ for $\eta = \sigma \cap \tau$ such that $x = z|M \cap \sigma^\vee$ and $y = z|M \cap \tau^\vee$. It is easy to see that this is an equivalence relation, and we denote by X_A the quotient space $\coprod_{\sigma\in X} F(\sigma)_A/\sim$.

If $A^\times$ is open in A, then the natural map $F(\sigma)_A \to X_A$ is an open immersion while it is a closed immersion if $A^\times$ is closed.

Let ρ be a face of $\sigma \in X$. We define

$$V_\sigma(\rho)_A = \{x \in F(\sigma)_A \,;\, x(m) = 0 \text{ for } m \in M \cap (\sigma^\vee \setminus \rho^\perp)\} .$$

We denote by $V(\rho)_A$ the union of $V_\sigma(\rho)$ for $\sigma \in X(\rho\prec)$ in X_A. Since

$$M[\rho] \cap \sigma[\rho]^\vee = M \cap \sigma^\vee \cap \rho^\perp \subset M \cap \sigma^\vee ,$$

each element $x \in V_\sigma(\rho)_A$ has an image in $F(\sigma[\rho])_A$ as the restriction.

For each $x \in X_A$, we can find the minimal $\rho \in X$ with $x \in V(\rho)_A$ as follows. Let $\sigma \in X$ be an element with $x \in F(\sigma)_A$. Since $\{m \in M \cap \sigma^\vee \,;\, x(m) \neq 0\}$ is a subsemigroup, there exists a maximal face $\sigma^\vee \cap \rho^\perp$ of $\sigma^\vee$ such that $x(m) \neq 0$ for an element $m \in M \cap \text{rel.}\operatorname{int}(\sigma^\vee \cap \rho^\perp)$. Then $x(m') \neq 0$ for every $m' \in M \cap \sigma^\vee \cap \rho^\perp$ by condition (b) since there exists $a > 0$ such that $am - m' \in M \cap \sigma^\vee \cap \rho^\perp$ and $x(m')x(am - m') = x(m)^a \neq 0$. By the maximality of $\sigma^\vee \cap \rho^\perp$, $x(m) = 0$ for $m \in M \cap (\sigma^\vee \setminus \rho^\perp)$. Namely, x is in $V(\rho)_A$ and this ρ is minimal.

Lemma 3.3. *Let X be a fan of $N_{\mathbf{R}}$. For any $\rho \in X$, the subspace $V(\rho)_A \subset X_A$ is homeomorphic to $X[\rho]_A$.*

Proof. The map from $V_\sigma(\rho)_A$ to $F(\sigma[\rho])_A$ is injective since every $x \in V_\sigma(\rho)_A$ is zero outside $M \cap \sigma^\vee \cap \rho^\perp$. Since every element y in $F(\sigma[\rho])_A$ is extended by 0 to an element of $V_\sigma(\rho)_A$, it is surjective. This is a homeomorphism since both topologies are induced from those of the products of A.

Since $V(\rho)_A$ is the union of $V_\sigma(\rho)_A$ for $\sigma \in X(\rho\prec)$ and the compatibility of these homomorphisms with the restriction maps are clear, we get a homeomorphism $V(\rho)_A \to X[\rho]_A$. *q.e.d.*

By this homeomorphism, we regard $X[\rho]_A$ as a subset of X_A. If $\rho \prec \sigma$ and $\{m_1, \ldots, m_l\}$ is a set of generators of $M \cap \sigma^\vee$ such that $m_1, \ldots, m_s \in M \cap \sigma^\vee \cap \rho^\perp$ and $m_{s+1}, \ldots, m_l \in M \cap (\sigma^\vee \setminus \rho^\perp)$, then $x \in F(\sigma)_{\mathbf{A}}$ is in $V_\sigma(\rho)_A$ if and only if $x(m_i) = 0$ for $i = s+1, \ldots, l$. Hence, by condition (b) of A, the one point $\{0\}$ is closed in A. Therefore, $V_\sigma(\rho)_A$ is closed in $F(\sigma)_{\mathbf{A}}$ for every $\sigma \in X(\rho\prec)$, and $X[\rho]_A$ is a closed subset of X_A.

Theorem 3.4. *Let X be a finite fan of $N_{\mathbf{R}}$. The toric variety $X_{\mathbf{C}}$ is compact if and only if the fan X is complete.*

Proof. We assume that X is complete. We prove that $X_{\mathbf{C}}$ is compact by induction on r. If $r = 0$, then $X_{\mathbf{C}}$ is one point and the assertion is clear. Assume $r > 0$. Since $F(\sigma)_D$ is compact for each $\sigma \in X$ by Lemma 3.2, their finite union X_D is also compact. Hence, it suffices to show that $X_{\mathbf{C}} = X_D$.

Let x be an element of $X_{\mathbf{C}}$. We take $\sigma \in X$ with $x \in F(\sigma)_{\mathbf{C}}$. If we think x as the homomorphism $M \cap \sigma^\vee \to \mathbf{C}$, we find an element $\rho \in F(\sigma)$ such that $x(m) \neq 0$ for $m \in M \cap \sigma^\vee \cap \rho^\perp$ and $x(m) = 0$ otherwise. Then x is contained in $X[\rho]_{\mathbf{C}}$. Here note that $X[\rho]$ is a complete fan of $N[\rho]_{\mathbf{R}}$ by Proposition 1.1. If $\rho \neq \{0\}$, then we have $X[\rho]_{\mathbf{C}} = X[\rho]_D \subset X_D$ by the assumption of the induction. Assume $\rho = \{0\}$. Then x is an element of T_N. Since X is complete, $\mu_N(x)$ is contained in some $\sigma \in X$. Then $x \in F(\sigma)_D \subset X_D$ by Lemma 3.1.

Now assume that X is not complete. Then, since $N_{\mathbf{R}} \setminus |X|$ is a nonempty open set, there exists a rational point x in it. Let γ be the one-dimensional cone generated by x. Since γ intersects the cones of X only at the origin of $N_{\mathbf{R}}$, $X' := X \cup \{\gamma\}$ is also a fan of $N_{\mathbf{R}}$. If $X_{\mathbf{C}}$ is compact, the Hausdorff property of $X'_{\mathbf{C}}$ implies that $X_{\mathbf{C}}$ is a closed subset of the toric variety $X'_{\mathbf{C}}$. This is a contradiction, since

$$F(\gamma)_{\mathbf{C}} \cap X_{\mathbf{C}} = T_N$$

and T_N is not closed in $F(\gamma)_{\mathbf{C}} \simeq \mathbf{C} \times (\mathbf{C}^\times)^{r-1}$. Hence, $X_{\mathbf{C}}$ is not compact if X is not complete. *q.e.d.*

Remark 3.5. Let $\mathbf{R}_0^{(+)}$ be $\mathbf{R}_0$ as an additive semigroup, which does not satisfy the condition (b). We consider the space $X_{\mathbf{R}_0^{(+)}}$ for a fan X. Since $\sigma = F(\sigma)_{\mathbf{R}_0^{(+)}}$ for $\sigma \in X$, $X_{\mathbf{R}_0^{(+)}}$ is the union of cones in X. Although there is natural continuous bijection from $X_{\mathbf{R}_0^{(+)}}$ to the support $|X|$, this is not a homeomorphism if X is infinite. We see that $X_{\mathbf{R}_0^{(+)}}$ is homeomorphic to N_R if and only if X is complete.

4 The algebraic surface defined by a tetrahedron

Let M be a free $\mathbf{Z}$-module of finite rank. We regard M as a lattice of the real space $M_{\mathbf{R}} = M \otimes_{\mathbf{Z}} \mathbf{R}$. For a finite subset $W = \{v_0, \ldots, v_s\}$ of M which is not contained in any $(s-1)$-dimensional affine subspace of $M_{\mathbf{R}}$, we denote by Index(W) the index

$$[M \cap (\mathbf{R}(v_1 - v_0) + \cdots + \mathbf{R}(v_s - v_0)) : \mathbf{Z}(v_1 - v_0) + \cdots + \mathbf{Z}(v_s - v_0)] .$$

Let $V = \{m_0, \ldots, m_3\}$ be a set of points of $M = \mathbf{Z}^3$ which is not contained in any plane of $M_{\mathbf{R}} = \mathbf{R}^3$. We set

T	:	the tetrahedron spanned by V,
n	:	$\mathrm{Index}(V)$,
S_i	:	the triangle spanned by $V \setminus \{m_i\}$ for $i = 0, 1, 2, 3$,
b_i	:	$\mathrm{Index}(V \setminus \{m_i\})$ for $i = 0, 1, 2, 3$,
$E_{i,j}$	:	the edge with the ends $V \setminus \{m_i, m_j\}$ for $0 \le i < j \le 3$,
$l_{i,j}$	:	$\mathrm{Index}(V \setminus \{m_i, m_j\})$ for $0 \le i < j \le 3$.

The classical Pick's theorem implies that b_i is equal to

$$2I_i + B_i - 2 \ ,$$

where I_i (resp. B_i) is the number of lattice points in the interior (resp. on the boundary) of the triangle S_i. Later, we consider the case where each S_i has lattice points only at the vertices. Then $b_i = 1$ since $I_i = 0$ and $B_i = 3$. In this case, we also have $l_{i,j} = 1$ for all i, j since each $E_{i,j}$ has no lattice points in the interior.

For each index $i = 0, \ldots, 3$, there exists a unique primitive element v_i of $N = \mathrm{Hom}_{\mathbf{Z}}(M, \mathbf{Z})$ which is constant on the triangle S_i and satisfying $\langle m_i, v_i \rangle > \langle S_i, v_i \rangle$. In other words, v_i's are the inner normal primitive vectors of the faces of the tetrahedron T. In particular, any three of $\{v_0, \ldots, v_3\}$ are linearly independent. For linearly independent primitive elements $u, v \in N$, we denote by $n(u, v)$ the index $[N \cap (\mathbf{R}u + \mathbf{R}v) : \mathbf{Z}u + \mathbf{Z}v]$. Then there exists a unique integer $q = q(u, v)$ with $0 \le q < n$ and $(v + qu)/n \in N$. Actually, we can take a coordinate of $\mathbf{R}u + \mathbf{R}v$ such that $u = (1, 0)$ and $v = (a, n)$ for an integer a. Then there exists a unique $0 \le q < n$ such that $a + q$ is a multiple of n. We can check the equalities $n(u, v) = n(v, u)$ and $q(u, v)q(v, u) \equiv 1 \pmod{n}$. We set $n_{i,j} = n(v_i, v_j)$ and $q_{i,j} = q(v_i, v_j)$ for $0 \le i < j \le 3$. The equality $n = l_{i,j} n_{i,j}$ can be checked easily. In particular, $n_{i,j} = n$ if $l_{i,j} = 1$.

In this note, a sequence $A = [a_1, \ldots, a_s]$ of integers at least 2 is called a *string* of length $s =: \mathrm{len}(A)$. We denote by Str the set of strings, and by Str_s the set of strings of length s.

For a pair of integers (n, q) with $0 < q < n$ and $\mathrm{gcm}(n, q) = 1$, we denote by $\mathrm{str}(n, q)$ the string $[a_1, \ldots, a_s]$ obtained by the continued fractions:

$$\frac{n}{q} = a_1 - \cfrac{1}{a_2 - \cfrac{1}{a_3 - \cfrac{\ddots}{a_{s-1} - \cfrac{1}{a_s}}}}$$

If $A = \mathrm{str}(n, q)$, then the integers n and q are recovered from the string A by the continued fractions. We denote the n and the q for A by $n(A)$ and $q(A)$, or by $n(a_1, \ldots, a_s)$ and $q(a_1, \ldots, a_s)$, respectively. For the empty sequence $A = [\ ]$, we define $n(A) = 1$ and $q(A) = 0$.

If $A = [a_1, \ldots, a_s]$, then we have the following results (see [I]).

(1) If $s \geq 1$, then

$$(a_1 - 1)q(A) < n(A) \leq a_1 q(A) . \tag{4.1}$$

(2) If we define a $n \times n$ matrix $(a_{i,j})$ by

$$a_{i,j} = \begin{cases} a_i & \text{if } i = j \\ 1 & \text{if } i = j \pm 1 \\ 0 & \text{otherwise} , \end{cases}$$

then we have

$$n(A) = \det(a_{i,j}) . \tag{4.2}$$

(3) Let A^* be the reverse $[a_s, \ldots, a_1]$ of A, then

$$n(A) = n(A^*) . \tag{4.3}$$

(4) Let $q^*(A) := q(A^*)$, then

$$q(A)q^*(A) \equiv 1 \ (\bmod n(A)) . \tag{4.4}$$

For positive integers n, q with $0 < q < n$ and $\gcd\{n, q\} = 1$, we define the integer $\sigma(A)$ by

$$\sigma(A) = \sigma(n, q) := a_1 + \cdots + a_s - 3s + 1$$

and the rational number $\lambda(A)$ by

$$\lambda(A) = \lambda(n, q) := a_1 + \cdots + a_s - 3s + \frac{q + q^*}{n} .$$

Where q^* is the integer with $0 < q^* < n$ and $qq^* \equiv 1 \pmod n$. We also define $\sigma(1, 0) = \lambda(1, 0) = 0$. If $n > 1$, then we have the inequalities

$$\sigma(n, q) - 1 < \lambda(n, q) < \sigma(n, q) + 1 \tag{4.5}$$

since $0 < q, q^* < n$.

We call $\lambda(n, q)$ the *deviation* of the pair (n, q). The equalities

$$\lambda(n, q) = \lambda(n, q^*) = -\lambda(n, n - q) = -\lambda(n, n - q^*)$$

hold if $n > 1$ (cf.[I]).

Let $\zeta = e^{2\pi i/n}$, and let Z be the quotient of $\mathbf{C}^2$ by the cyclic group action generated by the automorphism $(x, y) \mapsto (\zeta^q x, \zeta y)$ of order n. Then Z has a cyclic quotient singularity of type (n, q) at the image of the origin. It is known that the exceptional divisor of the minimal resolution $\tilde{Z}$ of this normal surface singularity is a chain of nonsingular rational curves of self-intersection numbers $-a_1, \ldots, -a_s$. If X is a compact normal surface with a cyclic quotient singularity of this type and $\tilde{X}$ is its minimal resolution, then we can show that $c_1^2 + c_2$ of $\tilde{X}$ decreases from that of X by $\lambda(n, q) - 2 + 2/n$.

We define the genus $g(T)$ of the tetrahedron T as the number of the lattice points in the interior of T. By Khovanskii [K1], this number is equal to the geometric genus of the algebraic surface defined by the sum of four monomials corresponding to the four vertices of T. A nonsingular model of the surface is obtained as follows.

The dual cones of $\mathbf{R}_0(T - m_i)$ for $i = 0, \ldots, 3$ and their faces form a complete fan Δ of $N_{\mathbf{R}}$, where we write the fan by a Greek capital in this section. We consider the surface $X(f)$ defined by $f = 0$ for $f = \mathbf{e}(m_0) + \cdots + \mathbf{e}(m_3)$ in the compact toric variety $Z(\Delta)$ associated to Δ. Since the normal compact surface $X(f)$ may have cyclic quotient singularities, we desingularize them by subdividing the fan. The exceptional curves of these desingularizations are described by the continued fractions, and this description helps the calculation of the chern numbers c_1^2 and c_2 of the desingularization $\widetilde{X(f)}$.

Using the fact that $X(f)$ is an abelian covering of degree n of $\mathbf{P}^2(\mathbf{C})$ ramifying over a union of four lines, we have $c_1^2 + c_2$ of $X(f)$ as

$$e(X(f)) + K_{X(f)}^2 = 2n - 3\sum_{i=0}^{3} b_i + (\sum_{i=0}^{3} b_i)^2/n + \sum_{0 \le i < j \le 3} l_{i,j} \; .$$

The normal surface $X(f)$ has $l_{i,j}$ cyclic quotient singularities of type $(n_{i,j}, q_{i,j})$ for each pair (i, j) with $0 \le i < j \le 3$. Hence this value decreases by

$$\sum_{0 \le i < j \le 3} l_{i,j} \{\lambda(n_{i,j}, q_{i,j}) - 2 + 2/n_{i,j}\}$$

after the minimal resolution.

The irregularity of this surface is zero since it is birationally a quotient of a Fermat surface in $\mathbf{P}^3(\mathbf{C})$. Hence we can calculate the geometric genus by the Riemann-Roch formula $\chi = (c_1^2 + c_2)/12$ of the arithmetic genus of algebraic surfaces, which is also called Noether's formula. Namely we have $g(T) = (c_1^2 + c_2 - 12)/12$.

Theorem 4.1 ([I]). *The genus $g(T)$ of the tetrahedron T is given by the formula*

$$g(T) = \frac{2n - 3\sum_{i=0}^{3} b_i + \left(\sum_{i=0}^{3} b_i\right)^2 n^{-1} - \sum_{0\le i<j\le 3} l_{i,j}\left\{(\sigma(n_{i,j}, q_{i,j}) - 4 + \frac{q_{i,j}+q_{j,i}+2}{n_{i,j}}\right\} - 12}{12}$$

$$= \frac{2n - 3\sum_{i=0}^{3} b_i + \left(\sum_{i=0}^{3} b_i\right)^2 n^{-1} - \sum_{0\le i<j\le 3} l_{i,j}\left\{(\lambda(n_{i,j}, q_{i,j}) - 3 + \frac{2}{n_{i,j}}\right\} - 12}{12}.$$

In this note, we consider the special case

$$\partial T \cap M = \{m_0, m_1, m_2, m_3\},$$

where ∂T is the boundary of the tetrahedron T. Then, since $b_i = 1$ for all $i = 0, \dots, 3$ and $l_{i,j} = 1$, $n_{i,j} = n$ for all $0 \le i < j \le 3$ in the formula in Theorem 4.1, the number g of the lattice points in the interior of the tetrahedron is given by the genus formula

$$\text{(4.6)} \qquad 12g = 2n + \frac{4}{n} - \sum_{0\le i<j\le 3} \sigma(n, q_{i,j}) - \sum_{0\le i<j\le 3} \frac{q_{i,j}+q_{j,i}}{n}$$

$$\text{(4.7)} \qquad = 2n - 6 + \frac{4}{n} - \sum_{0\le i<j\le 3} \lambda(n, q_{i,j})$$

We will normalize such tetrahedron so that we can count up all tetrahedra up to isomorphisms. By a parallel translation of T, we may assume $m_0 = 0$. Since the triangle spanned by $\{0, m_1, m_2\}$ contains no other points of M, $\{m_1, m_2\}$ is a part of a $\mathbf{Z}$-basis of M. We take a $\mathbf{Z}$-coordinate of M such that $m_1 = (1,0,0)$, $m_2 = (0,1,0)$ and $m_3 = (p,q,n)$ with $0 \le p, q < n$. Since $\{m_2, m_3\}$ is also a part of $\mathbf{Z}$-basis, p and n are coprime. Similarly, q and n are coprime. Since the three points m_1, m_2, m_3 span another face of T, $\{m_2 - m_1, m_3 - m_1\} = \{(-1,1,0), (p-1,q,n)\}$ is also a part of $\mathbf{Z}$-basis. This implies that $p+q-1$ and n are coprime.

The inner normal primitive vectors of T are

$$v_0 = (-n, -n, p+q-1), \quad v_1 = (n, 0, -p), \quad v_2 = (0, n, -q), \quad v_3 = (0,0,1).$$

We can calculate the $q_{i,j}$'s modulo n as follows.

$$\begin{array}{lllllll}
q_{0,1} & = & (p+q-1)^* p & & q_{1,0} & = & p^*(p+q-1) \\
q_{0,2} & = & (p+q-1)^* q & & q_{2,0} & = & q^*(p+q-1) \\
q_{0,3} & = & -(p+q-1)^* & & q_{3,0} & = & -(p+q-1) \\
q_{1,2} & = & -p^* q & & q_{2,1} & = & -q^* p \\
q_{1,3} & = & p^* & & q_{3,1} & = & p \\
q_{2,3} & = & q^* & & q_{3,2} & = & q
\end{array}$$

For example, the equation $q(v_0, v_1) = q_{0,1} = (p+q-1)^*p$ can be checked as follow. There exists an integer d with $(p+q-1)^*(p+q-1) = dn+1$. Hence

$$\begin{aligned} & v_1 + (p+q-1)^*pv_0 \\ = \ & (n, 0, -p) + (-(p+q-1)^*pn, -(p+q-1)^*pn, (dn+1)p) \\ = \ & ((1-(p+q-1)^*p)n, -(p+q-1)^*pn, dnp) \\ = \ & n(1-(p+q-1)^*p, -(p+q-1)^*p, dp) \ . \end{aligned}$$

By the above data, it is easy to check the following equalities.

(1.1.1) $q_{i,j}q_{j,i} \equiv 1 \pmod{n}$ for all $0 \leq i < j \leq 3$.

(1.1.2) $\sum_{j \neq i} q_{i,j} \equiv 1 \pmod{n}$ for all $i = 0, \dots, 3$.

(1.1.3) $q_{i,j}q_{j,k}q_{k,i} \equiv -1 \pmod{n}$ for all triple (i, j, k) of distinct elements of $\{0, 1, 2, 3\}$.

Note that this set of equalities is invariant under the action of the symmetric group of the index set $\{0, 1, 2, 3\}$ which is induced by the renumberings of the vertices of T. Since the induced action of the symmetric group on the set of pairs (i, j) is transitive, we can replace any $q_{i,j}$ by $q_{0,1}$ or other by a renumbering of the vertices.

For any triple (p, q, r) of integers prime to n with $0 < p, q, r < n$ and $p+q+r \equiv 1$ (mod n), we have a tetrahedron with $q_{3,1} = p$, $q_{3,2} = q$ and $q_{3,0} = r$. Furthermore, the triple is unique for a tetrahedron if we impose the following conditions.

(1) $p \leq q \leq r$.

(2) For $i = 0, 1, 2$, let $p' \leq q' \leq r'$ be the ordering of $\{q_{i,j} \ ; \ j \neq i\}$. Then $p < p'$ or $p = p'$ and $q < q'$.

As we comment later, there are more than 80 million such tetrahedra with $g \geq 1$ for the range $2 \leq n \leq 2500$ up to isomorphisms.

A simplex P of dimension three in $M_{\mathbf{R}}$ is said to be *terminal* if $P \cap M$ is equal to the set of the vertices of P.

The following theorem, which is called the Terminal Lemma, is due to White [W].

Theorem 4.2. *If P is a terminal simplex, then there exist integers n, p with $0 \leq p < n$ and $\gcd\{n, p\} = 1$ such that the set of vertices of P is*

$$\{(0, 0, 0), (1, 0, 0), (0, 1, 0), (p, 1, n)\}$$

for an affine $\mathbf{Z}$-coordinate of M.

Note that two vertices of P has the second coordinates 0 and other two has the coordinates 1. Hence the second coordinate of an interior point is in the open interval $(0, 1)$. In particular, it is not a lattice point.

It is clear that the simplex T is terminal if $q = q_{3,2} = 1$. If there exists (i, j) with $q_{i,j} = 1$, then T is terminal since we may assume $q_{3,2} = 1$ by a permutation of indices.

Hence, by this theorem, there exists a lattice point in the interior of T if and only if none of $q_{i,j}$'s is 1.

We are interested in the range of genera which actually appear for a fixed n. By the equality $\lambda(n,q) = -\lambda(n, n-q)$, we know the average of the deviation $\lambda(n,q)$ for a fixed n is zero. The genus formula indicates that, although the average of g is around $2n-6+4/n$, it has a high value if most of $\lambda(n, q_{i,j})$'s are negative and has a low value if they are positive. We see easily the maximal value of $\lambda(n,q)$ is $n-3+2/n$ for $q=1$ while the minimal value is $-n+3-2/n$ for $q=n-1$. Table 1 is a list of A with relatively high deviations.

	A	d	q^*	σ	λ
(1)	$[d]$	n	1	$n-2$	$n-3+2/n$
(2)	$[2,d]$	$(n+1)/2$	2	$(n-5)/2$	$n/2-3+5/2n$
(3)	$[3,d]$	$(n+1)/3$	3	$(n-5)/3$	$n/3-7/3+10/3n$
(4)	$[2,2,d]$	$(n+2)/3$	3	$(n-10)/3$	$n/3-11/3+10/3n$
(5)	$[4,d]$	$(n+1)/4$	4	$(n-3)/4$	$n/4-3/2+17/4n$
(6)	$[2,2,2,d]$	$(n+3)/4$	4	$(n-17)/4$	$n/4-9/2+17/4n$
(7)	$[2,d,2]$	$(n+4)/4$	$(n+2)/2$	$(n-12)/4$	$n/4-3+2/n$
(8)	$[2,3,d]$	$(n+2)/5$	5	$(n-13)/5$	$n/5-3+26/5n$
(9)	$[5,d]$	$(n+1)/5$	5	$(n+1)/5$	$n/5-3/5+26/5n$
(10)	$[3,2,d]$	$(n+3)/5$	5	$(n-12)/5$	$n/5-3+26/5n$
(11)	$[6,d]$	$(n+1)/6$	6	$(n+7)/6$	$n/6+1/3+37/6n$
(12)	$[3,d,2]$	$(n+5)/6$	$(n+3)/2$	$(n-13)/6$	$n/6-7/3+13/6n$
(13)	$[2,2,d,2]$	$(n+7)/6$	$(n+3)/2$	$(n-23)/6$	$n/6-11/3+13/6n$
(14)	$[2,4,d]$	$(n+2)/7$	7	$(n-12)/7$	$n/7-15/7+50/7n$
(15)	$[4,d,2]$	$(n+6)/8$	$(n+4)/2$	$(n-10)/8$	$n/8-3/2+5/2n$
(16)	$[3,d,3]$	$(n+6)/9$	$(n+3)/3$	$(n-12)/9$	$n/9-5/3+2/n$

Table 1: Relatively high deviations

Although the description of the sequence by continued fractions is explicit, the evaluation of the value $\lambda(n,q)$ is difficult in general. In order to give a lower bound of the genus, we need several lemmas. Inside the brackets of $n()$ and $q()$, the elements of Str and integers greater than one are concatenated by commas. For example, if $A = [a_1, \ldots, a_s]$ and $B = [b_1, \ldots, b_t]$, then $n(A,B) = n(a_1, \ldots, a_s, b_1, \ldots, b_t)$ and $n(A,3) = n(a_1, \ldots, a_s, 3)$.

When $s > 0$, we have the following equalities:

(4.8) $$q(a_1, a_2, \ldots, a_{s-1}, a_s) = n(a_2, \ldots, a_{s-1}, a_s)$$

(4.9) $$q^*(a_1, a_2, \ldots, a_{s-1}, a_s) = n(a_1, a_2, \ldots, a_{s-1}) \ ,$$

while $n() = 1$ and $q() = q^*() = 0$ for the empty sequence.

Lemma 4.3. *Let A, B in* Str*. Then we have the equality*

$$n(A, B) = n(A)n(B) - q^*(A)q(B) .$$

Proof. Set $A = [a_1, \dots, a_s]$ and $B = [b_1, \dots, b_t]$. We prove the lemma by induction on $\mathrm{len}(A)$. If $\mathrm{len}(A) = 0$, then we get the equality by $n() = 1$ and $q() = 0$.

If $\mathrm{len}(A) = 1$, then we have

$$n(a_1, B) = a_1 n(B) - q(B) = n(a_1)n(B) - q^*(a_1)q(B) . \tag{4.10}$$

If $\mathrm{len}(A) = 2$, then

$$\begin{aligned} & n(a_1, a_2, b_1, \dots, b_t) \\ = \;& a_1 n(a_2, b_1, \dots, b_t) - n(b_1, \dots, b_t) \\ = \;& a_1(a_2 n(b_1, \dots, b_t) - n(b_2, \dots, b_t)) - n(b_1, \dots, b_t) \\ = \;& (a_1 a_2 - 1)n(b_1, \dots, b_t) - a_1 n(b_2, \dots, b_t) \\ = \;& n(A, B) = n(A)n(B) - q^*(A)q(B) . \end{aligned}$$

Assume $\mathrm{len}(A) \geq 3$. Then, by using the induction assumption, we have

$$\begin{aligned} & n(a_1, \dots, a_s, b_1, \dots, b_t) \\ = \;& a_1 n(a_2, \dots, a_s, b_1, \dots, b_t) - n(a_3, \dots, a_s, b_1, \dots, b_t) \\ = \;& a_1(n(a_2, \dots, a_s)n(b_1, \dots, b_t)) - a_1(n(a_2, \dots, a_{s-1})n(b_2, \dots, b_t)) \\ - \;& n(a_3, \dots, a_s)n(b_1, \dots, b_t) + n(a_3, \dots, a_{s-1})n(b_2, \dots, b_t) \\ = \;& (a_1 n(a_2, \dots, a_s) - n(a_3, \dots, a_s)n(b_1, \dots, b_t) \\ - \;& (a_1 n(a_2, \dots, a_{s-1}) - n(a_3, \dots, a_{s-1}))n(b_2, \dots, b_t) \\ = \;& n(A)n(B) - q^*(A)q(B) . \end{aligned}$$

q.e.d.

Let $A_1, \dots, A_6$ be elements of Str with $n(A_1) = \cdots = n(A_6) = n > 1$. We assume that $m(A_1), \dots, m(A_6) \geq 2$ and the equality

$$12g = 2n - 6 + \frac{4}{n} - (\lambda(A_1) + \cdots + \lambda(A_6)) \tag{4.11}$$

is satisfied for a nonnegative integer g. Here we list some lemmas which are obtained in this situation.

Lemma 4.4. *If either $q(A_6)$ or $q^*(A_6)$ is equal to $n - 1$ and $n \geq 5$, then we have*

$$g \geq \frac{n-3}{24} . \tag{4.12}$$

Lemma 4.5. *If $n \geq 5$, either $q(A_6)$ or $q^*(A_6)$ is equal to $n-2$ and at most one of $A_1, \ldots, A_5$ satisfies $m(A_i) = 2$, then*

$$g \geq \frac{n-8}{18} \; . \tag{4.13}$$

Lemma 4.6. *If $n \geq 5$, $m(A_1) = 2$, $m(A_2) = 3$, $m(A_3), m(A_4), m(A_5) \geq 2$ and either $q(A_6)$ or $q^*(A_6)$ is equal to $n-6$, then*

$$g \geq \frac{n-11}{36} \; . \tag{4.14}$$

Lemma 4.7. *If $n \geq 5$, $q(A_6) = n-4$ and at most two of $m(A_1), \ldots, m(A_5)$ are 2, then*

$$g \geq \frac{n-22}{48} \tag{4.15}$$

By using the these lemmas and by further case-by-case calculation, we can show the following lower bound of the genus of a non-terminal tetrahedron with respect to the volume.

Theorem 4.8. *Assume that none of $q_{i,j}$'s are 1. Then the number g of lattice points in the interior of the tetrahedron satisfies the inequality*

$$\frac{n-23}{48} \leq g \; . \tag{4.16}$$

However, we may have the following much better bounds.

Conjecture 4.9. *Assume that none of $q_{i,j}$'s are 1. Then the number g of lattice points in the interior of the tetrahedron satisfies the inequalities*

$$\frac{n-8}{12} \leq g \leq \frac{n-1}{3} \; . \tag{4.17}$$

We can check this conjecture for $n \leq 2500$ by a direct calculation using the genus formula by a computer. The C program for the calculation is uploaded at:

`http://www.math.tohoku.ac.jp/%7Eishida/tetra2010wcom.c`

By this calculation, we know that there are 82,996,779 tetrahedra up to isomorphisms with at least one lattice point in the interior for the range $2 \leq n \leq 2500$, and all of them satisfy the inequalities of the conjecture. Among them, 1474 tetrahedra have the genus of the lower bound $\lceil (n-8)/12 \rceil$ and 1673 have that of the upper bound $\lfloor (n-1)/3 \rfloor$. For example, for $n = 2477$, there exist only one tetrahedron with the lower bound $g = 206$ and only one tetrahedron with the upper bound $g = 825$ among 255,647 tetrahedra of this volume.

References

[I] M. Ishida, The algebraic surface defined by a sum of four monomials, Comment. Math. Univ. St. Pauli. 40(1991), 39–51.

[K1] A. G. Khovanskii, Newton polyhedra and toroidal varieties, Functional Analysis and Applications 11(1977), 56–64, 96.

[K2] A. G. Khovanskii, Newton polyhedra and the genus of complete intersections, Functional Analysis and Applications 12(1978), 51–61.

[O] T. Oda, Convex Bodies and Algebraic Geometry, An Introduction to the Theory of Toric Varieties, Ergeb. Math. Grenzgeb.(3) 15, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1988.
`http://www.math.uni-bielefeld.de/~rehmann/DML/dml_links_title_C.html`

[W] G. K. White, Lattice tetrahedra, Canad. J. Math. 16(1964), 389–396.

Mathematical Institute
Graduate School of Science
Tohoku University
Sendai 980-8578
Japan

E-mail address: ishida@math.tohoku.ac.jp

Tohoku Mathematical Publications

No.1 Hitoshi Furuhata: *Isometric pluriharmonic immersions of Kähler manifolds into semi-Euclidean spaces*, 1995.

No.2 Tomokuni Takahashi: *Certain algebraic surfaces of general type with irregularity one and their canonical mappings*, 1996.

No.3 Takeshi Ikeda: *Coset constructions of conformal blocks*, 1996.

No.4 Masami Fujimori: *Integral and rational points on algebraic curves of certain types and their Jacobian varieties over number fields*, 1997.

No.5 Hisatoshi Ikai: *Some prehomogeneous representations defined by cubic forms*, 1997.

No.6 Setsuro Fujiié: *Solutions ramifiées des problèmes de Cauchy caractéristiques et fonctions hypergéométriques à deux variables*, 1997.

No.7 Miho Tanigaki: *Saturation of the approximation by spectral decompositions associated with the Schrödinger operator*, 1998.

No.8 Y. Nishiura, I. Takagi and E. Yanagida: *Proceedings of the International Conference on Asymptotics in Nonlinear Diffusive Systems — towards the Understanding of Singularities in Dissipative Structures —*, 1998.

No.9 Hideaki Izumi: *Non-commutative L^p-spaces constructed by the complex interpolation method*, 1998.

No.10 Youngho Jang: *Non-Archimedean quantum mechanics*, 1998.

No.11 Kazuhiro Horihata: *The evolution of harmonic maps*, 1999.

No.12 Tatsuya Tate: *Asymptotic behavior of eigenfunctions and eigenvalues for ergodic and periodic systems*, 1999.

No.13 Kazuya Matsumi: *Arithmetic of three-dimensional complete regular local rings of positive characteristics*, 1999.

No.14 Tetsuya Taniguchi: *Non-isotropic harmonic tori in complex projective spaces and configurations of points on Riemann surfaces*, 1999.

No.15 Taishi Shimoda: *Hypoellipticity of second order differential operators with sign-changing principal symbols*, 2000.

No.16 Tatsuo Konno: *On the infinitesimal isometries of fiber bundles*, 2000.

No.17 Takeshi Yamazaki: *Model-theoretic studies on subsystems of second order arithmetic*, 2000.

No.18 Daishi Watabe: *Dirichlet problem at infinity for harmonic maps*, 2000.

No.19 Tetsuya Kikuchi: *Studies on commuting difference systems arising from solvable lattice models*, 2000.

No.20 Seiki Nishikawa: *Proceedings of the Fifth Pacific Rim Geometry Conference*, 2001.

No.21 Mizuho Ishizaka: *Monodromies of hyperelliptic families of genus three curves*, 2001.

No.22 Keisuke Ueno: *Constructions of harmonic maps between Hadamard manifolds*, 2001.

No.23 Hiroshi Sato: *Studies on toric Fano varieties*, 2002.

No.24 Hiroyuki Kamada: *Self-dual Kähler metrics of neutral signature on complex surfaces*, 2002.

No.25 Reika Fukuizumi: *Stability and instability of standing waves for nonlinear Schrödinger equations*, 2003.

No.26 Tôru Nakajima: *Stability and singularities of harmonic maps into spheres*, 2003.

No.27 Yasutsugu Fujita: *Torsion of elliptic curves over number fields*, 2003.

No.28 Shin-ichi Ohta: *Harmonic maps and totally geodesic maps between metric spaces*, 2004.

No.29 Satoshi Ishiwata: *Geometric and analytic properties in the behavior of random walks on nilpotent covering graphs*, 2004.

No.30 Soondug Kim: *Computing in the Jacobian of a* C_{34} *curve*, 2004.

No.31 Mitsuko Onodera: *Study of rigidity problems for* $C_{2\pi}$*-manifolds*, 2006.

No.32 Shuji Yoshikawa: *Global solutions for shape memory alloy systems*, 2007.

No.33 Wakako Obata: *Homogeneous Kähler Einstein manifolds of nonpositive curvature operator*, 2007.

No.34 Keita Yokoyama: *Standard and Non-standard Analysis in Second Order Arithmetic*, 2009.

No.35 Reiko Miyaoka: *Su Buqing Memorial Lectures No. 1*, 2011.

Tohoku Mathematical Publications

Mathematical Institute
Tohoku University
Sendai 980-8578, Japan