Study of running bottom quark mass at 250GeV ILC

Department of Physics

Graduate School of Science

Tohoku University

Seidai Tairafune

February 12, 2021

Abstract

The Standard Model(SM) can not explain why each quark mass has different value each other and why the mass disparity of about 10^4 between measured quark masses. However, we can consider the energy dependence of quark mass, and these values change from current observed values at a higher energy scale. Furthermore, if some new particles such as SUSY contribute, the quark mass's energy dependence will deviate from the SM's expectation. Based on this idea, some models, such as GUTs, predict mass unification of 3rd generation particles (b, τ) at GUT scale ($\sim 10^{16}$GeV), and they are thought candidates which can approach problems of mass. Therefore, the verification of the b quark mass's energy dependence at a high energy scale provides a QCD theory test. Additionally, it can be a probe of future GUT scale physics.

The b mass at Z-pole ($\sim$ 91GeV) measured at LEP and SLD. It was in good agreement with SM, and we could not find any probe of new physics. As a next challenge, b quark mass measurement at a higher energy scale above Z-pole is expected to give a probe of new physics. This study simulated b quark pair production events at 250GeV ILC and estimated b mass measurement accuracy at the 250GeV energy scale.

The obtained accuracy of b quark mass measurement is

$$\Delta m_b(-,+) = 0.85(stat.) \pm 0.75(exp.) \pm 0.34(had.) \pm 0.07(theo.) \text{ GeV}$$
$$\Delta m_b(+,-) = 1.53(stat.) \pm 0.44(exp.) \pm 0.34(had.) \pm 0.07(theo.) \text{ GeV}.$$

If we combine them, it turns out that the precision of b quark mass measurement is 1GeV. Additionally, through this analysis, it turns out that the current Monte Carlo sample of $q\overline{q}$ process has some problems, and they affect the center value of the observable which this study uses and its statistical error, and it is updating now. Moreover, it is not a 250GeV experiment, but Giga-Z ILC can measure b quark mass at Z-pole at a better precision than LEP and SLD.

要旨

6 種類あるクォークの各質量値はどうしてバラバラな値を持つのか、またそこにはなぜ約 10^4 の質量格差があるのかということを標準模型は説明することができていない。しかしクォークの質量にはエネルギー依存性を考えることができて、高エネルギー領域でのクォーク質量は現在の測定値から変わる。さらにそこに SUSY などの新粒子が寄与してくるとクォーク質量のエネルギー依存性は標準模型の予測からずれ、最終的には第 3 世代粒子 (b, τ) が GUT スケール $(10^{16}\mathrm{GeV})$ で統一されることを予測するモデルもあり、上述の問題にアプローチできると考えられている。そのため b クォーク質量のエネルギー依存性の検証は QCD 理論の精密検証ができることに加え、GUT スケール物理の手がかりとなる可能性を持っている。

これまで LEP・SLD で既に測定された Z ボゾン質量領域 $(\sim 91\mathrm{GeV})$ での b クォーク質量の結果は標準模型によく従っており、そこに新物理の兆候は見られなかった。そこで Z-pole よりも高いエネルギースケールでの b クォーク質量測定で新物理の手がかりが得られる可能性がある。本研究は次に到達可能なエネルギースケールとして 250GeV ILC におけるボトムクォーク対生成事象をシミュレーションし、250GeV エネルギースケールにおけるボトムクォーク質量の測定精度を見積もる。

本研究では 250GeV での b クォークの測定精度は

$$\Delta m_b(-,+) = 0.85(stat.) \pm 0.75(exp.) \pm 0.34(had.) \pm 0.07(theo.)\ \mathrm{GeV}$$
$$\Delta m_b(+,-) = 1.53(stat.) \pm 0.44(exp.) \pm 0.34(had.) \pm 0.07(theo.)\ \mathrm{GeV}$$

となり、これらを合わせると 1GeV ほどの測定精度になることがわかった。また今回用いたモンテカルロサンプルには改善すべき点があることがわかり、この点は現在改良中である。さらにこれは 250GeV での実験ではないが、Giga-Z ILC では LEP・SLD を上回る精度で Z-pole での b クォーク質量を測定できることが今回の見積もりを通してわかった。

Acknowledgements

I would like to thank all the members of IFIC Valencia as a collaborative research institute and the members of the Experimental Particle Physics Laboratory of Tohoku University for conducting this research. Although I studied programming and the points of view of experiments in the undergraduate lectures, I came to the current Experimental Particle Physics Laboratory in a very immature state. I had many troubles since I became a graduate student because I had to acquire a massive amount of knowledge about physics and detectors and technical skills of software, but what I was most happy about was that the staff members of the laboratory went out with my questions and discussions and chats kindly without worrying about time. Until then, even if I had any questions in myself, I was atrophied and could not ask the teachers until getting off my chest so far, but I was able to relax and enjoy my study in the current environment. Professor Yamamoto gave the theme in the form of a collaborative study with Spanish researchers. Even after he was retired in March, he continued to pay attention to my progress and gave me various valuable comments. I was able to learn the points of view for study through him. Professor Sanuki commented on me from a bird's-eye view and an objective point of view to focus on points that I could not notice. Dr. Yonamine associated with my questions and discussions, especially on the technical and more detailed parts of the software, sometimes for a long time. I cannot thank you enough for giving me answers to software and physics analysis's fundamental questions politely and kindly. Thank you for joining in every working group meeting and working with me until the end of this study. In addition to the staff, I was also indebted to the seniors in the laboratory in various situations. In particular, I still remember that Mr. Sato kindly taught me how to tell what I was doing to others in an easy-to-understand manner. Like Mr. Sato, I am also a student of GPPU (Graduate Program on Physics for the Universe) at Tohoku University. He checked my slides for the GPPU interview or the progress report meeting and gave me objective advice. It allowed me to learn how to communicate things better, not just about study. In this way, everyone in the Experimental Particle Physics Laboratory at Tohoku University was very kind to me, and I was able to learn not only to study techniques but also necessary qualities as a researcher internally. Thank you again.

From the second year of my master's grade, I have been able to study with receiving a salary as an RA of the GP-PU program of Tohoku University. Not only receiving such financial support but asking researchers around the world questions and discussions, I once again realized that the discipline of physics connects people across national borders. I would like to thank all GP-PU staff for giving me such valuable opportunities.

This research was a collaboration study with IFIC Valencia. Professor Marcel Vos and Dr. Adrian Irles of the ATLAS / Future colliders group led by Professor Juan Faster shared the deeper aspects of this study. In particular, the code used for the analysis of this study was improved and used, which was shared by Dr. Adrian. I am really grateful to him for politely answering each question I asked about the basic part of the analysis and for working with me. Professor German Rodrigo greatly cooperated with my study in the form of answering my theoretical questions and making some plots. Originally, I was planning to stay in Valencia

around this summer to meet in person for study. However, due to the worldwide spread of the coronavirus, I began to proceed with analysis at remote meetings to the end. I actually started my research in March 2020, but I was able to achieve the result in less than a year, thanks solely to the staff of Tohoku University and IFIC Valencia. It is a miracle that everyone was able to stay healthy without being infected by the virus under this unprecedented situation.

I was trying my best to get used to a new environment for the first six months because the environment has changed dramatically from I was an undergraduate student. Meanwhile, in July 2019, I had symptoms of cholelithiasis and underwent surgery for cholecystectomy. It was a sudden thing, and the people around me were surprised. Because of such a situation, I could not start my study easily for the first year. But since I could finish this study thanks to my colleagues and my family and my friends. I am grateful to my family, who supported me in terms of health and the economy. Feeling my lack of power as I proceed during my study, I was sometimes wondering, "Is it really okay for me to go on the path of research?". But I learned from a teacher how I should live my life. He told me that be a researcher is just one method to live, but not the fundamental purpose of life. These words made me feel lighten up, and I was able to proceed with this study with a sense of responsibility, and I decided to go on to the Ph.D. program. I would like to express my gratitude to those who have beside me and supported many of my decisions so far.

謝辞

本研究を行なうにあたって東北大学 素粒子実験研究室のメンバーに加え共同研究機関としてIFIC Valencia のメンバーの皆様に大変お世話になりました。

私は学部時代にプログラミングや実験について学部の講義で触れてはいたものの、とても未熟な状態で現在の素粒子実験研究室に来るようになりました。身につけるべきことが物理や測定器の知識的にもソフトウェアの技術的にも膨大に増えてしまい大学院生になってからなかなか苦労していましたが、何より嬉しかったのは山本先生をはじめとする研究室のスタッフの方々が人格的で暖かく私の質問や議論ついには雑談まで時間も気にせず付き合ってくださったことでした。それまで自分の中で疑問を抱いてもなかなか自分の胸がスッキリするまで先生方と話し合うことができなかったのですが、心を楽にして楽しみながら研究を進められたという点で本当に救われた気持ちでした。山本先生にはスペインの研究者との共同研究という形でテーマをいただきました。3 月に退官されてからも続けて私の進捗状況に関心を払ってくださって様々に貴重なコメントをいただけましたし、それを通して研究をする際の着眼点を学ぶことができました。佐貫先生には俯瞰的・客観的な視点でコメントをくださり、なかなか自分では気づくことができなかった点に目が向くようにしてくださいました。また修論審査会の直前まで私の発表練習に付き合ってくださり、様々なアドバイスをしてくださりました。与那嶺先生は特にソフトウェアの技術的な部分や研究のより細かな部分に関して時には長時間にわたって私の質問や議論に付き合ってくださいました。ソフトウェアや物理解析のいろはから私の基本的な質問にも一つ一つ丁寧にそして優しく教えてくださって本当に感謝してもしきれません。working group でのミーティングにも毎回入ってくださり、最後まで私の研究に共にしてくださって本当にありがとうございました。またスタッフの方々だけではなく、研究室の先輩方にも様々な場面でお世話になりました。特に佐藤さんには自分が研究していることをいかにわかりやすく人に伝えるのか親身に教えてくださったことを今でも覚えています。私も佐藤さんと同じく東北大学の GP-PU(宇宙創成物理学国際共同大学院プログラム) の学生だったことから、GP-PU の interview や進捗報告会でのスライドを見てくださり客観的なアドバイスをいただけたことで私自身も研究に限らず物事をどうしたらよりよく人に伝えることができるかをよく学ぶことができました。このように東北大学 素粒子実験研究室の皆様はとても暖かく私に接してくださり研究の技術だけでなく研究者として内面的に必要な資質を学ぶことができました。改めてありがとうございました。

2 年目からは GP-PU の RA として給与をいただきながら研究させていただけるようになりました。このような経済的支援を受けながらさらに世界の研究者に質問や議論をさせていただく中で、改めて物理学という学問は国境を超えて人を繋ぐものであると実感しました。GP-PU の皆様にもこのような貴重な機会をいただけたことに感謝いたします。

また本研究はIFIC Valenica の皆様との共同研究でした。Juan Faster 教授率いる ATLAS/Future colliders グループの Marcel Vos 先生、Adrian Irles 先生には研究のより深い点で議論を共にしてくださいました。特に本研究の解析をする際のコードは Adrian 先生に共有していただいたものを改良して使わせていただきました。解析の基本的な部分に関して一つ一つ質問した私に丁寧に答えてくださり、共にしてくださって本当に感謝いたします。German Rodrigo 先生には理論的な質問に答えていただいたり、プロットを作成していただくという形で私の研究に大きく協力していただきました。本来今年の夏頃にバレンシアの方に滞在し直接会って研究することを予定していましたが、世界的なコロナウイルスの蔓延によりリモートでのミーティングで最後まで研究を進めるようになりました。実質的に研究を始めたのが 2020 年の 3 月からでしたが、1 年足ら

ずでこのように結果を出せたのは一重に東北大学と IFIC Valencia のスタッフの皆様のおかげです。この未曾有の事態の中で皆がウイルスに感染せず健康でいられたことはまさに奇跡と言っていいでしょう。

学部生の時から環境が大きく変わったことで最初の半年は新しい環境になれることに精一杯でした。そんな中で 2019 年の 7 月に私は胆石症の症状を起こしてしまい、胆嚢摘出の手術を受けました。あまりに突然のことで周囲を驚かせてしまいましたし、このような状況でなかなか最初の 1 年間は研究に着手することができなかったのですが、それでもここまで来れたのは私を精神的に支えてくれた友人たち、健康面・経済面に支えてくださった家族のおかげです。研究を進めていく中で自分の力不足を感じて「本当に自分は研究の道を進むべきだろうか」と悩むこともありましたが、1 人の先生から私は人生をどのように生きるべきかを学び、研究者という生き方があくまでも人生の根本の目的ではないことを教えていただきました。その言葉のおかげで心が軽くなり、程よい責任感と共にこの研究を進めることができましたし、博士課程への進学を決心するようにもなりました。ここに至るまで私の数多くの決心を見届け、共にし、支えてきてくださった方々にここに感謝の意を表します。

Introduction

Mass is one of the essential concepts in the construction of this universe. However, physicists have been attracted to many mysteries about its origin for several hundred years. In July 2012, the Higgs boson was discovered in the ATLAS experiment at LHC of CERN(European Organization for Nuclear Research). In the Standard Model of particle physics, the Higgs boson was predicted as the particle which gives the proper masses to other matter and forces particles, and there is no doubt that this discovery gives us significant development for the exploration of the origin of mass.

Thanks to the discovery of the Higgs boson, we can understand how particles obtain their mass, but there are unsolved points in each particle's mass value that each particle obtains from Higgs. In the Standard Model, there are six types (flavor) of quarks, and the smallest one is $2\mathrm{MeV}/c^2$, but whereas the biggest one is $173\mathrm{GeV}/c^2$ which is about 10^4 times. Although six quarks obtain each mass according to the same mechanism called the Higgs mechanism, why they have different mass values and why there is such a large disparity between them are unsolved problems of the Standard Model. However, we can consider the energy dependence for quark mass, and these values change at a higher energy scale. Furthermore, if some new particles contribute, the energy dependence of quark mass will deviate from the Standard Model's prediction. Based on this idea, some models predict that b quark and tau lepton (and may top quark also) are unified at the GUT scale, and such models are thought to approach mass problems. To approach these problems, we need to examine quark mass precisely and understand the physics behinds there. In this study, b quark, which is the second heavier quark, is focused on and estimates b quark mass measurement precision at 250GeV ILC by the simulation.

The constitution of this thesis is like this :

Chapter 1 : Theoretical Framework

This chapter reviews the outline of the theoretical framework of Quantum Field Theory and the Standard model. Especially, QCD(Quantum Chromodynamics) and renormalization theory are necessary base knowledge of this study, and they are reviewed here.

Chapter 2 : Study of running bottom quark mass

How to define quark mass is not trivial, unlike leptons. This chapter explains quark mass definitions and discusses the importance and usefulness of running quark mass study. Show the result at LEP(Large Electron-Positron Collider), which is run from the 1990s to the 2000s, and explain this study's motivation based on the previous study result.

Chapter 3 : Experimental apparatus; ILC and ILD detector

Since this study is a full-simulation study of the ILD detector, review the ILC and ILD detector's outline.

Chapter4 : Experimental framework for Simulation

Explain the tools which this study used and the flow of full-simulation.

Chapter5 : Estimation of b quark mass precision

Explain the procedures of event selection, and show the results of estimated b quark mass measurement.

Chapter6 : Discussion and Conclusion

Based on obtained results, discuss the viability of Beyond the Standard model and future bottom quark mass study scenarios at high energy.

はじめに

質量はこの宇宙を構築する上で重要な概念のうちの 1 つである。しかしながらその起源には多くの謎が残っており物理学者はその問題に魅了され今日に至るまで研究を進めてきた。そしてついに 2012 年 7 月に Higgs ボゾンが CERN(欧州原子核研究機構) にある LHC(Large Hadron Collider) にて行なわれた ATLAS 実験で発見された。素粒子の標準模型 (the Standard Model; SM) では Higgs ボゾンは他の物質及び力の粒子に固有の質量を与える粒子として予言されており、Higgs ボゾンの発見は質量に関する探究への大きな一歩となったことは疑いの余地はない。

Higgs ボゾンの発見により我々は素粒子がどのようにして自身の質量を獲得するのかということを理解することができたが、このようにして各粒子が Higgs ボゾンから獲得した質量値には未だ謎がある。実際 6 種類あるクォークには最も軽いもので約 $2\mathrm{MeV}/c^2$、また最も大きいもので約 $173\mathrm{GeV}/c^2$ と質量の獲得方法は共通していながらなぜ各質量値はバラバラなのか、またそれらの間にはなぜ約 10^4 という大きな質量格差があるのかは標準模型が説明できていない未解決問題の 1 つである。しかしクォークにはエネルギー依存性を考えることができ、高エネルギー領域でのクォーク質量の値は現在の測定値から変わる。またさらに超対称性粒子 (Super Symmetry; SUSY) などの新粒子が寄与してくるとそのエネルギー依存性は標準模型の予測からずれ、大統一理論 (Grand Unified Theory; GUT) などのいくつかのモデルでは b クォークと τ レプトン (場合によっては t クォークも) が GUT スケール 10^{16}GeV で統一されることを予言しているものもある。このようなモデルはバラバラな質量値の起源を説明できるモデルの候補と考えられている。この問題に実験的な側面からアプローチするためにはクォークの質量をより精密に調べ、その背景にある物理を検証・理解することが重要である。本研究では 2 番目に重いクォークであるボトムクォークに着目して、ILC を用いた 250GeV エネルギースケールにおける b クォーク質量の測定精度をシミュレーションにより見積もる。

本論文の構成は以下のようになっている：

Chapter 1 : 理論的枠組み

本章では量子場の理論 (Quantum Field Theory; QFT) と標準模型の理論的枠組みを概観する。特に繰り込み理論と QCD は本研究において必要な基礎知識になるのでその点に関しても本章で説明する。

Chapter 2 : running ボトムクォーク質量の研究

クォークはレプトンと違ってその質量の定義は簡単でない。本章ではクォークの質量の定義と running クォーク質量の研究の有用性を説明・議論する。1990 年代から 2000 年代にかけて行なわれた LEP(Large Electron Positron Collider) での先行研究の結果とそれを踏まえた本研究でのモチベーションを本章にて明示する。

Chapter 3 : 実験設備；ILC と ILD 検出器

本研究は ILD 検出器での質量測定をシミュレーションするので本章では ILC 及び ILD に関して概観する。

Chapter4 : シミュレーションのための実験的枠組み

本章ではシミュレーションと解析で用いるアルゴリズムや大まかなシミュレーションの流れを説明する。

Chapter5：b クォーク質量の測定精度の見積もり

前章にて説明した手順に則って事象選別/観測量の測定/最終的な結果を提示する。

Chapter6：議論と結論

本章では得られた結果をもとに 250GeV ILC における新物理の探索性能や今後のシナリオについて議論する。

Contents

List of Figures

List of Tables

Notation

- Mincowski metric : $g^{\mu\nu} = (+1, -1, -1, -1)$
- In this thesis, we use Natural unit system ： $c = 1,\ \hbar = 1$.

Terminology・Jargon

- Quantum Field Theory : QFT
- The Standard Model : the SM
- Beyond the Standard Model : BSM
- Quantum Electrodynamics : QED
- Vacuum Expectation Value : VEV
- Modified Minimal Subtraction scheme : $\overline{\text{MS}}$
- Quantum Chromodynamics : QCD
- Renormalization Group Equation : RGE
- Super Symmetry : SUSY
- Grand Unified Theory : GUT
- Minimal Super Symmetry Model : MSSM
- International Linear Collider : ILC
- International Large Detector : ILD
- Vertex Detector : VXD
- Particle Flow Algorithm : PFA

1 Theoretical Framework

1.1 Structure of the Standard model

This chapter explains the Standard Model(SM). In the SM, there are 17 type particles(See figure.1), and there are excellent structure and mechanism. The SM is based on Relativistic Quantum Field Theory(QFT), and particles appear under some beautiful symmetries. This chapter will show how these particles appear in theory and how they interact with each other to create this universe.

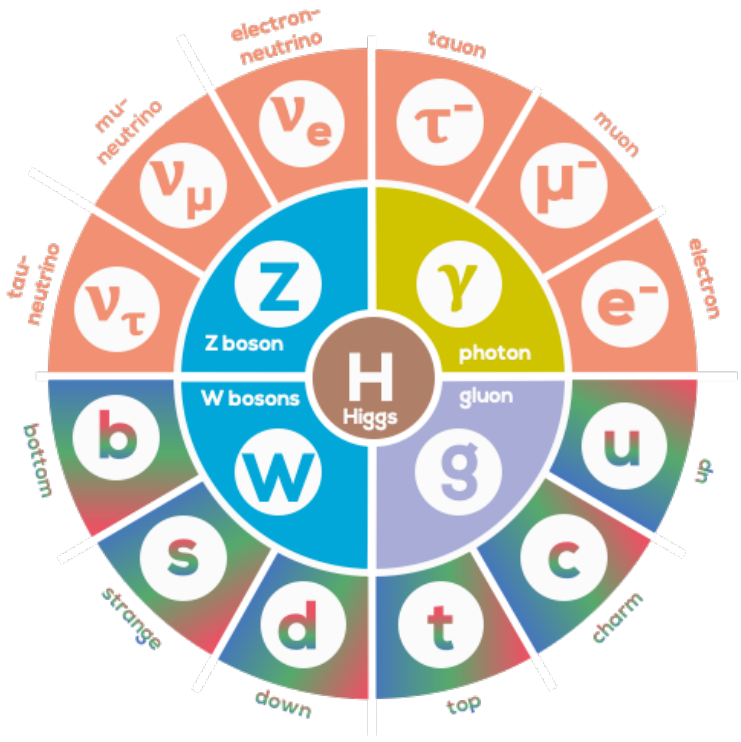


Figure 1: Particles of the Standard Model : There are six quarks, six leptons, four gauge bosons, and one Higgs boson. The SM is a successful theory that can explain many phenomena in the universe.

1.1.1 Categorization by representations of Lorentz Group

Lorentz group provides a fundamental category of particles in the SM. It is familiar that all SM particles appear as the objects of the Lorentz group's representation. At first, let us look at them[1].

Lorentz group is defined by Lorentz transformations, and let us consider the Lorentz group under proper-orthochronous transformations here[2]. Since both Lorentz boost and spatial rotation have 3 degrees of freedom each other, the Lorentz group has 6 degrees of freedom in total. Fortunately, we have already known the group with the degree of freedom 3; SU(2). Therefore, the Lorentz group is formulated by using the combination of 2 SU(2)s. SU(2) is one of the

[1]The first section of [1] is a very nice summary of this discussion.
[2]Lorentz transformation is included Lorentz boost, spatial rotation, Parity transformation, Time reversal. The former two transformations construct a subgroup, and it is defined as Lorentz Group here.

famous group in Lie groups[3] and its fundamental behavior is governed by Lie algebra :

$$[T^a, T^b] = \epsilon^{abc} T^c. \tag{1.1}$$

Here, ϵ^{abc} is a completely anti-symmetric tensor, and the index runs 1 to 3 (this number 3 comes from the degree of freedom of SU(2)). T^a generates SU(2) transformation elements by

$$U = \exp[i\theta^a T^a] \ \in \ \mathrm{SU}(2) \quad (\theta^a \in \mathbb{R}), \tag{1.2}$$

and in general, they are given as matrices. Such matrices T^a is called Generator of the Lie Group. Since we need 2 SU(2)s to construct the Lorentz group, there are six generators, and they correspond to three K_i for Lorentz boosts and three J_i for spatial rotations. Each of them satisfies the following SU(2) Lie algebra :

$$\begin{aligned}
[J_i, J_j] &= i\epsilon_{ijk} J_k \\
[K_i, K_j] &= -i\epsilon_{ijk} J_k \\
[J_i, K_j] &= i\epsilon_{ijk} K_k,
\end{aligned}$$

and they construct the Lorentz transformation

$$\Lambda = e^{i\boldsymbol{\theta}\cdot\boldsymbol{J} + i\boldsymbol{\phi}\cdot\boldsymbol{K}}.$$

We can check that J_is are closed itself, but K_i and J_i intertwine with each other, and it is an inconvenience. Then, let us define new generators which are consisted of mixing of J_i and J_i :

$$N_i^{\pm} \equiv \frac{1}{2}\left(J_i \pm iK_i\right). \tag{1.3}$$

It satisfies the following Lie algebras :

$$[N_i^+, N_j^+] = i\epsilon_{ijk} N_k^+ \tag{1.4}$$

$$[N_i^-, N_j^-] = i\epsilon_{ijk} N_k^- \tag{1.5}$$

$$[N_i^+, N_j^-] = 0. \tag{1.6}$$

N_i^+ and N_i^- are closed each other as SU(2) Lie algebra, and they are independent. This expression is useful when we think of the explicit formula of representation of the Lorentz group.

In principle, the size of the transformation matrix can be taken freely, and each selection of matrix size is called Representation of group[4]. In order to categorize the representation of SU(2), we can use a useful tool, the Casimir matrix. Casimir matrix is constructed by all generators of the considering representation, and in general, it takes the shape of $(Constant)\times\mathbf{1}$. This constant one-to-one corresponds to each representation of SU(2), so we can use it to categorize representations. Table.1 shows the correspondence between Casimir's constant(j in the following table) and the representation of SU(2).

[3] If the element in the group can be expressed by using continuous parameters, it is called Lie group. e.g.) Rotations in 2-dimensional plane

[4] If you want to know more detail, you should check texts of Lie Group. As a recommendation, [2] is a good introduction book.

j	matrix size	Generators number
0	1	3
1/2	2	3
1	3	3

Table 1: Representations of SU(2) : j takes an integer and half-integer number, and the matrix size becomes $2j+1$. Note that the number of generators does not change for each representation. Of course, we can think "higher" representations.

For the Lorentz group, since we should consider 2 SU(2) Lie algebras, we need to prepare two constants j_1, j_2 of Casimir matrices of each representation. As a result, we can consider the combinations(j_1, j_2), and $j = j_1 + j_2$(it is often called Spin) specifies the representation of Lorentz group(See table.2).

(j_1, j_2)	Spin j	matrix size	representation
(0 , 0)	0	1	Scalar rep.
(0 , 1/2)	1/2	2	Spinor rep. (Left Chiral)
(1/2 , 0)	1/2	2	Spinor rep. (Right Chiral)
(1/2 , 1/2)	1	3	Vector rep.

Table 2: Representations of Lorentz group : Note that $\left(\frac{1}{2}, 0\right)$ and $\left(0, \frac{1}{2}\right)$ are not equivalent each other.

The above three representations(Scalar, Spinor, Vector) are all representations in the SM. Why higher-representations do not appear in the SM is one of the mysteries, but for the moment, let us focus on these three representations.

By the way, before moving on to explanations of each representation, it should be explained about Field. As mentioned above, particles in the SM appear as objects of representations of the Lorentz group. Also, these objects are Operator Fields that create or annihilate particles. Such a field is called Quantum Field, and the picture of the particle appears as an excitation state of quantum fields. Since field means that it spreads on infinitely vast spacetime and has one or multiple values on each spacetime point, we can express it such as $\phi(x)$. Such fields that have spacetime dependence are objects of representations of the Lorentz group, and each field that introduces here corresponds to each particle in the SM.

Let us see each Lorentz transformation property of each object.

- Scalar representation (Spin-0 representation)
  Lorentz transformation property of object : $\phi(x) \to \phi(x)$

  Since $j = 0$ for scalar representation, the size of the transformation matrix is 1. It means that the generator should be 0, and the representation matrix is just identity $\mathbf{1}$. Thus, the Lorentz transformation of the scalar field does not change at all under spacetime transformations $x \to \Lambda x$. In the SM, only Higgs boson behaves on scalar representation.

- $\left(\frac{1}{2}, 0\right)$ Spinor representation (Left chirality)
  Lorentz transformation property of object : $\chi_a \to \left(e^{i\boldsymbol{\theta}\cdot\frac{\boldsymbol{\sigma}}{2}+\boldsymbol{\phi}\cdot\frac{\boldsymbol{\sigma}}{2}}\right)^b{}_a \chi_b$

  This is the case of $j_1 = \frac{1}{2}$ and $j_2 = 0$. So one side SU(2), which is generated by N_i^+, is a 2-dimensional representation, and another SU(2), which is generated by N_i^-, is a

1-dimensional representation(it is just identity). So each generator is decided :

$$N_i^+ = \frac{1}{2}\sigma_i \ , \ N_i^- = 0$$

It should be noted that $J_i = +iK_i$ from $N_i^- = 0$. Using it, N_i^+ is rewritten like this :

$$N_i^+ = iK_i = \frac{1}{2}\sigma_i.$$

According to them, we can decide the explicit expression of $\left(\frac{1}{2}, 0\right)$ representation :

$$\Lambda_{\left(\frac{1}{2},0\right)} = e^{i\boldsymbol{\theta}\cdot\frac{\boldsymbol{\sigma}}{2}+\boldsymbol{\phi}\cdot\frac{\boldsymbol{\sigma}}{2}} \tag{1.7}$$

Since the matrix size is 2, the object which is transformed by eq.(1.7) should have two components such as

$$\begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}. \tag{1.8}$$

Such object which is transformed by $\left(\frac{1}{2}, 0\right)$ representation is called left chiral spinor.

- $\left(0, \frac{1}{2}\right)$ Spinor representation (Right chirality)
  Lorentz transformation property of object : $\xi^{\dot{a}} \to \left(e^{i\boldsymbol{\theta}\cdot\frac{\boldsymbol{\sigma}}{2}-\boldsymbol{\phi}\cdot\frac{\boldsymbol{\sigma}}{2}}\right)^{\dot{a}}{}_{\dot{b}}\,\xi^{\dot{b}}$

  This is the case of $j_1 = 0$ and $j_2 = \frac{1}{2}$. So one side SU(2), which is generated by N_i^+, is a 1-dimensional representation(it is just identity), and another SU(2), which is generated by N_i^-, is a 2-dimensional representation. So each generator is decided :

$$N_i^+ = 0 \ , \ N_i^- = \frac{1}{2}\sigma_i$$

It should be noted that $J_i = -iK_i$ from $N_i^+ = 0$. Using it, N_i^- is rewritten like this :

$$N_i^- = -iK_i = \frac{1}{2}\sigma_i.$$

According to them, J_i and K_i are obtained as $J_i = \frac{\sigma_i}{2}, K_i = i\frac{\sigma_i}{2}$, and we can decide the explicit expression of $\left(0, \frac{1}{2}\right)$ representation :

$$\Lambda_{\left(0,\frac{1}{2}\right)} = e^{i\boldsymbol{\theta}\cdot\frac{\boldsymbol{\sigma}}{2}-\boldsymbol{\phi}\cdot\frac{\boldsymbol{\sigma}}{2}} \tag{1.9}$$

Since the matrix size is 2, the object which is transformed by eq.(1.9) should have two components such as

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}. \tag{1.10}$$

It should be noted that the representation eq.(1.9) are similar with eq.(1.7), but they are completely different from left chiral spinor. So the object eq.(1.10) has different Lorentz transformation property, and it's confusing that using same spinor index as left chiral spinor. In order to distinguish them, let us add the dot to spinor index like this : $\xi^{\dot{a}}$ [5]. Such object which is transformed by $\left(0, \frac{1}{2}\right)$ representation is called right chiral spinor, and eq.(1.8) and eq.(1.10) are collectively called Weyl spinors.

[5] Such notation is called Van der Waerden notation. You can check more detail by [1] or [2].

- $\left(\frac{1}{2},0\right)\oplus\left(0,\frac{1}{2}\right)$ representation

  We can combine two spinor representations. This summed representation has the size of representation matrix is 4. the object is like this :

$$\psi = \begin{pmatrix} \chi_{L\ a} \\ \xi_R^{\dot{a}} \end{pmatrix} \quad (a = 1, 2). \tag{1.11}$$

  χ_L is a left chiral spinor and ξ_R is a right chiral spinor. Such object is called Dirac spinor and its transformation property is give by[6]

$$\Lambda_{\left(\frac{1}{2},0\right)\oplus\left(0,\frac{1}{2}\right)} = \begin{pmatrix} \Lambda_{\left(\frac{1}{2},0\right)} & \\ & \Lambda_{\left(0,\frac{1}{2}\right)} \end{pmatrix}. \tag{1.12}$$

- Vector representation (Spin-1 representation)

  Lorentz transformation property of object : $A^\mu(x) \to \Lambda^\mu_\nu\ A^\nu(x)$

  In short, this representation is the same as the one of spacetime. So the object A^μ has a Lorentz index, which is the same as spacetime x^μ or four-momentum p^μ.

They are all representations of the Lorentz group, which appear in the Standard Model. Each particle(more precisely, fields) behaves according to each equation of motions, and each Lagrangian specifies them. So the next section introduces them.

1.1.2 Construction of Lagrangian

In the previous section, we saw that the Lorentz group is an important group to categorize elementary particles. The Lagrangian is constructed from these fields, and in order to construct relativistic theory, Lagrangian must keep the following requirements :

- Locality

  According to Causality, each term of Lagrangian must be consisted by information of only one spacetime point; e.g. $\phi(x)\phi(x)\phi(x)\phi(x)$.

- Lorentz Invariance

  In relativistic theory, Lagrangian or the equation of motion should not change under Lorentz transformations[7].

- Mass dimension is 4

  We believe that this world is 4-dimensional spacetime. Under this assumption, the action is defined as

$$S[\phi] = \int d^4x \mathcal{L}\left(\phi(x), \partial_\mu\phi(x)\right).$$

  Since the action is just a number, the mass dimension of the action should be 0. Additionally, the mass dimension of length is -1. Namely, the mass dimension of Lagrangian should be +4, and each term should be constructed so that it keeps. Since the field's mass dimension is decided by free Lagrangian, it may not be useful for free field theory. However, to the construction of interaction theory, which is more important, this requirement becomes a useful policy.

[6] It should be noted that it is not an irreducible representation.

[7] Here, we are thinking only proper-orthochronous subgroup; Lorentz boosts and spatial rotations.

- Renormalizability

  For interaction theory, mass and coupling constant, which appear in original Lagrangian (called bare Lagrangian), are not parameters that can measure physically. These mass and coupling constants(bare mass, bare coupling) include infinities[8] in them. However, of course, physical mass and coupling which can measure should be finite. To solve this problem, we need to eliminate these infinities from bare parameters, and this prescription is called Renormalization[9]. Therefore, physical mass and coupling, measured experimentally, is renormalized mass and renormalized coupling. In general, the theory in which UV divergence can be removed by finite counterterms is called the Renormalizable theory. Non-renormalizable theory needs infinite counterterms in its Lagrangian, and such theories are an inconvenience. So we want to construct renormalizable theory, and it is famous that renormalizable theory has a non-negative mass dimension for each coupling constant.

- Local Gauge Symmetry

  Gauge transformation corresponds to transformation at internal space of the field, but not spacetime[10]. As we can see in later sections, the interaction between matter field and force field is decided entirely by gauge symmetry. So, the construction of the Lagrangian, which keeps local gauge invariance, is an essential requirement. In the SM, important gauge symmetry is U(1), SU(2), and SU(3).

We should construct Lagrangian based on these requirements. For free field theory, each field's Lagrangian can be constructed like the following. Each equation of motion is derived easily by the action principle or Euler-Lagrange equation.

- Real Scalar Field $\phi(x)$
  $\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2$

- Complex Scalar Field $\Phi(x)$
  $\mathcal{L} = \partial_\mu\Phi^\dagger\partial^\mu\Phi - m^2\Phi^\dagger\Phi$

- Dirac Spinor Field $\psi(x)$
  $\mathcal{L} = \overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi$

- Massive Vector Field (Proca Field) $A^\mu(x)$
  $\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2A^\mu A_\mu \qquad (F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu)$

It should be noted that these Lagrangians do not include any interaction terms, but this free field picture is unrealistic because our world is constructed by interactions. So, we need to add interaction terms in Lagrangian, but how should we introduce them? Fortunately, we have a firm policy; local gauge symmetry and Gauge principle. Let us look at how interactions between matters are introduced into theory through this symmetry in the next section.

[8] There are some types of infinity. UV(Ultra-Violet) divergence appears at a high energy region. In contrast, IR(Infra-Red) divergence occurs at a low energy region. If particles with extremely little energy(called soft particles) are emitted, IR divergence appears. Collinear divergence occurs by parallel momentums between 2 particles.

[9] You can check more detail of renormalization at later section1.3

[10] Transformations of spacetime means Lorentz transformation here.

1.2 Electroweak Theory and Mass generation mechanism

Until previous sections, we have seen that particles in the SM appear as objects under the Lorentz group, and they are governed by each Lagrangian, which is introduced above. Such Lagrangians govern the dynamics of free particles. However, our actual world is constructed by various interactions between particles, and Lagrangian in the previous section can not explain these interactions. In order to incorporate forces into theory, local gauge symmetry is an essential symmetry. It is famous that this world is constructed by four fundamental forces; electromagnetic force, weak force, strong force, gravity. Former three interactions are introduced elegantly by corresponded local gauge symmetries. Interestingly, these three interactions can be expressed by the common theoretical structure of Gauge theory. This fact means that these three forces can be explained as "a unified force", and this Unified theory is one of the big dreams for all physicists. Electromagnetic and weak interactions were unified by S. Glashow, S. Weinberg, and A. Salam in the 1970s. This unified theory is called Electroweak(EW) Theory or Glashow-Weingerg-Salam(GWS) Theory, and one of the important things of EW theory is the mass generation mechanism of gauge bosons and fermions through spontaneous gauge symmetry breaking. It is called the Higgs mechanism. It solved the problem of how are massive weak bosons appeared and how does each fermion obtain proper mass by the introduction of a neutral Higgs boson. Let us see more details of them in this section.

1.2.1 Quantum Electrodynamics

As mentioned in the previous section, local gauge symmetry is an important element for constructing Lagrangian. The following example may be useful to understand it; Let us consider the transformation for the field $\phi(x)$ like this :

$$\phi(x) \to e^{i\theta}\phi(x). \tag{1.13}$$

θ is a real parameter. If we rewrite $\phi(x)$ as $\mathrm{Re}\phi(x) + i\mathrm{Im}\phi(x)$, we can interpret this transformation as the rotation on the complex plane of $\phi(x)$. Of course, these axes are not spacetime axes, and it can be interpreted that these axes span some "internal space" but not spacetime. Such transformation on the internal degree of freedom is called Gauge transformation. Especially since transformation eq.(1.13) belongs to U(1) group, it is called U(1) gauge transformation. These axes on the complex plane are taken by the observer freely, so physics must not be changed under this gauge transformation. Namely, the requirement which Lagrangian should keep the invariance under the gauge transformation makes sense, and it means imposing the internal symmetry to fields.

By the way, in the above case, the parameter θ is assumed "global" parameter, which means to operate the completely same rotation for all spacetime points. However, this transformation does not match with the sense of Causality of Special relativity. Causality claims that any information must not propagate between spacetime points by the speed which over the speed of light. Since the global transformation rotates the field for all spacetime points simultaneously, it looks unnatural. On the other hand, the local transformation, in which the degree of rotation

depends on each spacetime point, looks more natural. So let us consider the case of the phase parameter has spacetime dependence $\theta(x)$. This section's motivation is that the construction of theory keeps local U(1) gauge invariance. However, when we try to impose this local gauge symmetry to Lagrangian, the critical problem occurs. To see the fundamental point of it, let us consider the Lagrangian of Dirac Spinor field. The Lagrangian keeps the invariance under global U(1) transformation easily. On the other hand, in the case of local U(1) transformation, since the parameter $\theta(x)$ has spacetime dependence, the spacetime derivative in the first kinetic term acts on it not only Dirac field. As a result, Lagrangian does not keep the local U(1) symmetry because of appearing of an additional term $-(\partial_\mu\theta(x))\,\overline{\psi}\gamma^\mu\psi$.

To keep local U(1) gauge symmetry, we need to introduce the Gauge principle. The basic idea of Gauge principle like this; define Covariant derivative D_μ which includes new vector field $A_\mu(x)$ and replace spacetime derivative ∂_μ to D_μ. After that, specify the gauge transformation property of A_μ is determined so that it cancels appearing extra terms that break local gauge symmetry. Let us look at the practical example in the Dirac field case. At first, let us consider the local U(1) gauge transformation

$$U = e^{ig\theta(x)}. \tag{1.14}$$

Under this transformation, the covariant derivate is defined by

$$D_\mu \equiv \partial_\mu + igA_\mu(x), \tag{1.15}$$

and replace the derivative ∂_μ into this covariant derivative D_μ :

$$\overline{\psi}\,(iD_\mu - m)\,\psi \tag{1.16}$$

The field $A_\mu(x)$ is introduced to keep local gauge invariance, and its gauge transformation property is specified so that the above Lagrangian keeps local U(1) gauge invariance. To determine it, let us focus on the behavior of D_μ on gauge transformation. The gauge transformation property of covariant derivative $D_\mu\psi$ should be

$$D_\mu\psi \to UD_\mu\psi.$$

From this relation, the gauge transformation property of D_μ is decided :

$$D_\mu \to UD_\mu U^{-1}.$$

Furthermore, the transformation property of A_μ is also determined from this :

$$\begin{aligned} D_\mu \to D'_\mu = \partial_\mu + igA'_\mu(x) &= U\left(\partial_\mu + igA_\mu(x)\right)U^{-1} \\ &= U\partial_\mu U^{-1} + U\cdot U^{-1}\partial_\mu + igA_\mu(x) \\ &= \partial_\mu + ig\left(A_{\mu(x)} - \partial_\mu\theta(x)\right) \end{aligned}$$

$$A_\mu(x) \to A_\mu(x) - \partial_\mu\theta(x) \tag{1.17}$$

Therefore,

$$\mathcal{L} = \overline{\psi}\left(i\not{D} - m\right)\psi = \overline{\psi}\left(i\not{\partial} - m\right)\psi - g\overline{\psi}\not{A}\psi \tag{1.18}$$

is the Lagrangian, which keeps local U(1) gauge symmetry, and it is the theory that we want to derive here. The notable point is that appearing interaction terms between the Dirac field

and gauge field automatically by imposing local U(1) symmetry. The interaction term has the factor g, it specifies the strength of interaction, and it is called the Coupling constant. Then, we should add the kinetic term of the gauge field itself. It is accomplished by constructing the quantity, which is called strength of gauge field :

$$F_{\mu\nu} \equiv -\frac{i}{g}[D_\mu, D_\nu] \tag{1.19}$$

It can be verified that it is an anti-symmetric tensor :

$$\begin{aligned}
[D_\mu, D_\nu] &= [\partial_\mu + igA_\mu(x), \partial_\nu + igA_\nu(x)] \\
&= ig\left([\partial_\mu, A_\nu(x)] + [A_\mu(x), \partial_\nu]\right) \\
&= ig\left(\partial_\mu A_\nu(x) - A_\nu(x)\partial_\mu + A_\mu(x)\partial_\nu - \partial_\nu A_\mu(x)\right) \\
&= ig\left(\partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)\right) \\
&\rightarrow F_{\mu\nu} = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)
\end{aligned} \tag{1.20}$$

Since $F_{\mu\nu}$ is U(1) gauge invariance, the combination

$$-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{1.21}$$

is gauge and Lorentz invariant, and it corresponds to the kinetic term of gauge field A_μ. By applying the action principle, we can check that gauge field A_μ satisfies Maxwell equation $\partial_\mu F^{\mu\nu} = 0$. Additionally, it should be noted that the gauge field A_μ is prohibited from having its mass term because the mass term such as $m^2 A_\mu A^\mu$ does not gauge invariant. If we regard the U(1) gauge field A_μ as the field of the photon, the Lagrangian

$$\mathcal{L} = \overline{\psi}\left(i\not{D} - m\right)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{1.22}$$

explains the electromagnetic interaction between Dirac fields. Therefore the U(1) gauge theory eq.(1.22) is called Quantum Electrodynamics(QED). Actually, if we consider the infinitesimal gauge transformation of field $\psi \rightarrow U\psi \sim (1 + ig\theta)\,\psi$, the Noether current can be obtained as $j^\mu \equiv -g\overline{\psi}\gamma^\mu\psi$, and it satisfies the conservation $\partial_\mu j^\mu = 0$. According to them, the Noether charge is

$$Q \equiv -\int d^3x\, g\overline{\psi}\gamma^0\psi.$$

If we regard ψ as the electron field and rewrite $g = qe$(e is elementary charge, q is some constant), it can be interpreted as electric charge because $\overline{\psi}\gamma^0\psi = \psi^\dagger\psi$ is nothing but the density of electron[11] [12]. This section's important thing is the interaction between matter particles is derived by imposing local gauge symmetry. If we consider other gauge symmetries, we can incorporate new interactions. The next section explains the gauge theory, which explains weak nuclear force.

1.2.2 SU(2) gauge theory / Weak Theory

In the previous section, we have seen that local U(1) gauge symmetry derives QED. By the way, U(1) is the group that its elements are commutative, and it is called the Abelian gauge

[11] More accurately, $\psi^\dagger\psi$ means that the occupancy number operator.
[12] Note that the coupling constant g is dimensionless. It means that QED is renormalizable theory.

group. So QED is often called Abelian gauge theory. On the other hand, SU(2) is a Non-Abelian gauge group that its elements are not commutative with each other. In this section, let us see SU(2) non-Abelian gauge group.

The behavior of the Lie group is governed by Lie algebra. The Lie algebra of SU(2) is given by

$$[T^a, T^b] = i\epsilon_{abc}T^c \quad (a = 1, 2, 3) \tag{1.23}$$

and we can see that SU(2) is non-commutative for two elements, which are generated by generators T^a. Let us prepare 2 Dirac fields ψ_1 and ψ_2 in order to construct SU(2) gauge theory. If we introduce the Dirac doublet, which is the object under the fundamental representation of SU(2) like this :

$$\Psi \equiv \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{1.24}$$

the Langrangian of 2 Dirac fields ψ_1, ψ_2 is written as

$$\begin{aligned} \mathcal{L} &= \overline{\psi_1}\left(i\partial\!\!\!/ - m_1\right)\psi_1 + \overline{\psi_2}\left(i\partial\!\!\!/ - m_2\right)\psi_2 \\ &= \overline{\Psi}\left(i\partial\!\!\!/ - M\right)\Psi. \end{aligned} \tag{1.25}$$

Here, each mass of ψ_1 and ψ_2 is assumed not same, and M is defined as a matrix :

$$M \equiv \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}. \tag{1.26}$$

The Dirac doublet Ψ is transformed by SU(2) gauge transformation U as $\Psi \to U\psi$, and of course, the Lagrangian eq.(1.25) keeps global SU(2) transformation $U = e^{i\theta_a T^a}$ $(a = 1, 2, 3)$[13]. However, similar to the U(1) case, in the case of phase parameters θ_a are local, the SU(2) invariance is broken by appearing the extra terms from derivative of phase parameters. Additionally, since 2 Dirac fields do not necessarily have the same mass, its mass term also breaks gauge symmetry in general. Then, let us consider constructing the local SU(2) gauge invariant theory based on the idea of gauge principle again. However, the different point from U(1) case is that appeared extra terms in kinetic term are three but not one :

$$\overline{\Psi}i\partial\!\!\!/\Psi \to \overline{\Psi}i\partial\!\!\!/\Psi - (\partial_\mu\theta_a(x))\,\overline{\Psi}\gamma^\mu T^a\Psi.$$

Namely, we need to introduce three new gauge fields in this SU(2) case. Based on it, define the covariant derivative like this :

$$D_\mu \equiv \partial_\mu + igT^aW_\mu^a(x) \quad (a = 1, 2, 3). \tag{1.27}$$

So we should decide the gauge transformation of D_μ and W_μ^a so that the Lagrangian, which is replaced the spacetime derivative to the covariant derivative

$$\mathcal{L} = \overline{\Psi}iD\!\!\!/\Psi, \tag{1.28}$$

keeps local SU(2) gauge invariance. The gauge transformation property of D_μ is same as the U(1) case; $D_\mu \to UD_\mu U^{-1}$. The transformation property of W_μ^a is given by substitution

[13] Generators of fundamental representation of SU(2) is given as Pauli matrices $\frac{1}{2}\sigma^a$

eq.(1.27) to this relation :

$$\begin{aligned} D_\mu \to D'_\mu = \partial_\mu + igT^a W'_\mu a(x) &= U' \left(\partial_\mu + igT^a W^a_\mu(x)\right) U^{-1} \\ &= U\left(\partial_\mu U^{-1}\right) + \partial_\mu + igUT^a W^a_\mu(x) U^{-1} \end{aligned}$$

$$\therefore T^a W^a_\mu(x) \to UT^a W^a_\mu(x) U^{-1} + \frac{i}{g}\left(\partial_\mu U\right) U^{-1} \tag{1.29}$$

It should be noted that appeared gauge fields coupe with generators in its transformation property for non-abelian gauge theory. If we define the combination of gauge fields and generators as $T^a W^a_\mu(x) \equiv W_\mu(x)$, eq.(1.29) can be rewritten by

$$W_\mu(x) \to UW_\mu(x)U^{-1} + \frac{i}{g}\left(\partial_\mu U\right) U^{-1}. \tag{1.30}$$

Therefore,

$$\mathcal{L} = \overline{\Psi} i \not{D} \Psi = \overline{\Psi} i \not{\partial} \Psi - g \overline{\Psi} \not{W} \Psi \tag{1.31}$$

L is the Lagrangian, which keeps local SU(2) gauge invariance. The second term is the interaction term between the Dirac doublet and SU(2) gauge fields. The coefficient g specifies the strength of this interaction[14]. To introduce the kinetic term of SU(2) gauge fields, let us define the strength of non-abelian gauge fields. The definition itself is entirely the same :

$$W_{\mu\nu} = -\frac{i}{g}\left[D_\mu, D_\nu\right],$$

and you can calculate it like this[15] :

$$\begin{aligned} \left[D_\mu, D_\nu\right] &= \left[\partial_\mu + igT^a W^a_\mu(x), \partial_\nu + igT^b W^b_\nu(x)\right] \\ &= ig\left[\partial_\mu, T^a W^a_\nu(x)\right] + ig\left[T^a W^a_\mu(x), \partial_\nu\right] + (ig)^2\left[T^a W^a_\mu(x), T^b W^b_\nu(x)\right] \\ &= ig\left(\partial_\mu\left(T^a W^a_\nu(x)\right) - \partial_\nu\left(T^a W^a_\mu(x)\right)\right) + (ig)^2\left[T^a W^a_\mu(x), T^b W^b_\nu(x)\right] \\ &\to W_{\mu\nu} = \partial_\mu W_\nu(x) - \partial_\nu W_\mu(x) + ig[W_\mu(x), W_\nu(x)] \end{aligned} \tag{1.32}$$

If we compare it with eq.(1.20), you can see that the commutator of gauge field $W_\mu(x)$ is added in this non-abelian case, and it is an appearance of the non-commutativity of SU(2) gauge group. The appearance of the commutator of gauge fields means that the existence of the interaction between gauge fields themselves(it is called Self-coupling). The gauge transformation property of $W_{\mu\nu}$ is same as D_μ : $W_{\mu\nu} \to UW_{\mu\nu}U^{-1}$, but it is not gauge invariant unlike the case of U(1). Thus, $W_{\mu\nu}W^{\mu\nu}$ does not gauge invariant too, and we can not use this form as the kinetic term. But if we take the trace of it, it can be gauge invariant because of cyclicity of trace[16] :

$$\begin{aligned} \mathrm{Tr}\left(W_{\mu\nu}W^{\mu\nu}\right) \;\to\; \mathrm{Tr}\left(UW_{\mu\nu}W^{\mu\nu}U^{-1}\right) &= \mathrm{Tr}\left(U^{-1}UW_{\mu\nu}W^{\mu\nu}\right) \\ &= \mathrm{Tr}\left(W_{\mu\nu}W^{\mu\nu}\right). \end{aligned}$$

Therefore the kinetic term of SU(2) gauge field W_μ is given like this :

$$-\frac{1}{4}\mathrm{Tr}\left(W_{\mu\nu}W^{\mu\nu}\right). \tag{1.33}$$

[14] I used same character g as U(1) case, but it should be noted that it is different from the one of U(1) case.

[15] It should be noted that $W^a_\mu(x)$s are commutative with each other, but $W_\mu(x)$s are not commutative because $W_\mu(x)$s include generators T^a which are not commutative.

[16] If we use the normalization condition $\mathrm{Tr}\left(T^a T^b\right) = \frac{1}{2}\delta^{ab}$ for generators, it can be rewritten as $-\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu}$. Here, $W^a_{\mu\nu} \equiv \partial_\mu W^a_\nu(x) - \partial_\nu W^a_\mu(x)$

Finally, the Lagrangian, which has local SU(2) gauge symmetry, is

$$\mathcal{L} = \overline{\Psi} i \not{D} \Psi - \frac{1}{4} \mathrm{Tr}\left(W_{\mu\nu} W^{\mu\nu}\right). \tag{1.34}$$

SU(2) gauge theory is the theory of weak interaction, which is mediated by three gauge fields. One of the familiar phenomena by weak interaction is β decay. β decay is the action in which neutron is exchanged to proton or vice versa :

$$n \to p + e^- + \overline{\nu}_e$$
$$p \to n + e^+ + \nu_e$$

Since proton and neutron have the same mass approximately[17], if we consider the doublet[18]

$$\begin{pmatrix} p \\ n \end{pmatrix},$$

we can interpret that proton and neutron are exchanged by acting SU(2) fundamental representation. The force which occurs in the nuclear action, including β decay, is explained by SU(2) gauge group, and it is called weak force or weak interaction.

However, the theory of weak interaction is a little different from eq.(1.34). According to the experiment of ^{60}Co decay which was done by Wu in the 1950s, the parity violation of weak interaction was revealed. In eq.(1.34), there is only vector current interactions $\overline{\Psi}\gamma^\mu T^a \Psi W^a_\mu$, and the vector current keeps the invariance under P transformation[19] :

$$\overline{\Psi}\gamma^\mu T^a \Psi \ \to \ \mathrm{P}^\mu_\nu \, \overline{\Psi}\gamma^\nu T^a \Psi = \overline{\Psi}\gamma^\mu T^a \Psi.$$

So we need to add some interaction term which violates parity invariance to eq.(**??**) The pseudo-vector (or axial-vector) interaction such as $\overline{\Psi}\gamma^\mu\gamma_5 T^a \Psi$ flips the sign of interaction term under P transformation :

$$\overline{\Psi}\gamma^\mu\gamma_5 T^a \Psi \ \to \ \mathrm{P}^\mu_\nu \, \overline{\Psi}\gamma^\nu\gamma_5 T^a \Psi = -\Lambda^\mu_\nu \overline{\Psi}\gamma^\nu\gamma_5 T^a \Psi.$$

So the interaction term such as the combination of axial-vector currents and gauge fields can be used as additional terms of Lagrangian. Therefore, The Lagrangian

$$\begin{aligned} \mathcal{L} = & \ \overline{\Psi} i \not{\partial} \Psi \\ & - \frac{1}{4} \mathrm{Tr}\left(W_{\mu\nu} W^{\mu\nu}\right) \\ & - \frac{1}{2} g_W \overline{\Psi}\gamma^\mu \left(1 - \gamma_5\right) \Psi W^a_\mu \end{aligned} \tag{1.35}$$

gives the theory of weak interaction and such theory which includes both of vector and axial-vector currents is called V-A theory[20]. V-A theory tells us the important nature of weak interaction. To see it, let us consider the extraction of Left chiral spinor ψ_L and Right chiral spinor ψ_R from Dirac spinor through the projection operator

$$P_L \equiv \frac{\mathbf{1} - \gamma_5}{2} \quad \left(P_R \equiv \frac{\mathbf{1} + \gamma_5}{2}\right). \tag{1.36}$$

[17] Proton mass is about 93bMeV/c^2, neutron mass is about 940MeV/c^2.

[18] p and n in the doublet are fields which are called nucleon.

[19] It is often called Parity transformation too. It flips all spatial components, and the representation of P transformation on spinor space is $\mathrm{P}^\mu_\nu = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}$. It is the same form as γ^0 on chiral representation

[20] The factor is retaken so that it has $\frac{1}{2}$ for later discussion.

When swap this projection operator and γ matrices, the chirality of the projection operator is flipped because γ matrices and γ_5 are anti commutative each other :

$$\begin{aligned}\gamma^\mu P_{L,R} &= \frac{1}{2}\gamma^\mu\left(\mathbf{1} \mp \gamma_5\right) \\ &= \frac{1}{2}\left(\mathbf{1} \pm \gamma_5\right) = P_{R,L}\gamma^\mu\end{aligned}$$

Using it, the interaction term in eq.(1.35) can be rewritten as

$$\begin{aligned}-g_W\overline{\Psi}\gamma^\mu P_L\Psi W_\mu^a &= -g_W\overline{\Psi}\gamma^\mu P_L^2\Psi\, W_\mu^a \qquad \left(P_L^2 = P_L\right) \\ &= -g_W\overline{\Psi}\gamma^\mu P_L\cdot P_L\Psi\, W_\mu^a \\ &= -g_W\overline{\Psi}P_R\gamma^\mu\cdot P_L\Psi\, W_\mu^a \\ &= -g_W\Psi^\dagger\gamma^0 P_R\gamma^\mu P_L\Psi\, W_\mu^a \\ &= -g_W\Psi^\dagger P_L\gamma^0\gamma^\mu P_L\Psi W_\mu^a \\ &= -g_W\overline{\Psi}_L\gamma^\mu\Psi_L\, W_\mu^a \qquad \left(P_L^\dagger = P_L\right)\end{aligned}$$

.

According to this, weak interaction acts on only left chiral spinors. In other words, the object which can feel weak interaction is the left chiral Dirac doublet (it is called SU(2) doublet)

$$\Psi_L = \begin{pmatrix}\psi_{1L}\\ \psi_{2L}\end{pmatrix}, \tag{1.37}$$

and there is the charge that mediates weak interaction, similar to electric charge in U(1). The charge of ψ_{1L} and ψ_{2L} are decided by Noether theorem. Under the infinitesimal transformation of field $\Psi_L \to U\Psi_L \sim \left(\mathbf{1} + i\theta^a T^a\right)\Psi_L$, there are conserved currents $j^{\mu\,a} \equiv \overline{\Psi}\gamma^\mu T^a\Psi$. So there are three conserved quantities in SU(2) theory, and we can label each Dirac spinor ψ_{1L} and ψ_{2L} by using these quantities. To label, it is useful that choosing the diagonal element (it is called Cartan matrix), and in the case of SU(2), it is

$$T^3 = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

If we consider the third conserved component $j^{0\ 3} = \overline{\Psi}\gamma^0 T^3\Psi$, it becomes like this

$$\begin{aligned}j^{0\ 3} &= \Psi^\dagger\left(\gamma^0\right)^2 T^3\Psi \\ &= \begin{pmatrix}\psi_{1L}^\dagger & \psi_{2L}^\dagger\end{pmatrix}\begin{pmatrix}+\frac{1}{2} & 0\\ 0 & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}\psi_{1L}\\ \psi_{2L}\end{pmatrix} \\ &= +\frac{1}{2}\psi_{1L}^\dagger\psi_{1L} - \frac{1}{2}\psi_{2L}^\dagger\psi_{2L}.\end{aligned}$$

We can interpret that ψ_{1L} has a charge $+\frac{1}{2}$ and ψ_{2L} has a charge $-\frac{1}{2}$ each other. These values label each component of SU(2) doublet, and this charge $j^{0\ 3}$ is called Weak isospin. In QED, the particle with a non-zero electric charge can feel the electromagnetic force, but in the weak theory, the particle with non-zero weak isospin can feel weak interaction. On the other hand, the right chiral spinor can not feel weak interaction, and the parity violation is reflected in this nature. Right chiral spinors are incorporated as the form (called SU(2) singlet) of

$$\psi_{1R}\ ,\ \psi_{2R},$$

and they have just 0 as weak isospin. Do SU(2) gauge fields also has weak isospin? To understand this, we should consider the adjoint representation of SU(2)[21]. The diagonal element in the adjoint representation of SU(2) is

$$T^3 = \begin{pmatrix} +1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Similar to above, SU(2) fields can be labeled by eigenvalues of this matrix, and we can see that each three gauge fields have the isospin +1,0,-1.

This section shows that weak interaction is explained by SU(2) gauge theory. However, there is a critical issue here. According to various experiments, it is revealed that three particles that come from SU(2) fields (called Weak bosons) are massive. However, as described above, the mass term of gauge fields is prohibited by gauge symmetry. Additionally, of course, Dirac fermions also should have mass terms, but they also are prohibited by SU(2) gauge symmetry. It is a very critical problem, but there is a fantastic breakthrough. Let us see it in the next section.

1.2.3 Higgs Mechanism

Gauge symmetry is beautiful symmetry that derives interaction terms between matters naturally. However, it prohibits the existence of mass term of Dirac fermions and gauge fields, and it is a critical contradiction for SU(2) gauge theory. The numbers of the measured mass of weak bosons are like this :

Gauge bosons	mass
$W^{\pm}$	$80.379 \pm 0.012\text{GeV}/c^2$
Z^0	$91.1876 \pm 0.0021\text{GeV}/c^2$

Table 3: Mass numbers of weak bosons[5]

On the other hand, the photon is massless. So we need to devise the mechanism so that Dirac fermions and weak bosons can obtain their finite masses. In other words, we need to consider the mechanism which creates the difference between the photon and weak bosons. This section explains the Electroweak theory, which unifies electromagnetic and weak interactions, and how is the mass term's problem solved by Spontaneous symmetry breaking and the Higgs mechanism.

The basic scenario is like this: At a higher energy scale (at least $>$ 100GeV) such as the early universe, two gauge interactions, the electromagnetic and weak force, were unified. The universe becomes colder with time evolution, and the Higgs boson, a new scalar particle, has a non-zero vacuum expectation value at about 100GeV (this energy scale is called the Elecroweak scale). Because of it, Dirac fermions and weak bosons obtained each mass by the interaction with the Higgs boson, and the difference between electromagnetic force and weak force appears there. As a result, at the lower energy scale where we live now, SU(2) gauge symmetry is broken, and weak bosons have each mass. The following part of this section explains a more

[21] The representation which the generator size is the same as the degree of freedom. In this case, it is 3.

detailed formulation of this scenario.

First, let us construct the unified gauge theory of U(1) and SU(2) gauge group. The Lagrangian, which has local U(1) and SU(2) symmetries, is given by

$$\begin{aligned}\mathcal{L} = &\overline{\Psi} i\gamma_\mu \left(\partial_\mu - i\frac{g}{2}B^\mu Y - ig' W_\mu^a T^a\right)\Psi \\ &- \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \\ &- \frac{1}{4}\mathrm{Tr}W^{\mu\nu}W_{\mu\nu}.\end{aligned} \tag{1.38}$$

Here, the covariant derivative $\partial_\mu - i\frac{g}{2}B_\mu Y - ig'W_\mu^a T^a$ is defined by U(1) gauge field B^μ(coupling constant is g) and SU(2) gauge fields W_μ(coupling constant is g'). Y is the generator of U(1)[22]. We should think about how to incorporate mass terms of Dirac fermions and weak bosons without SU(2) symmetry breaking. To do this, let us introduce the doublet, which is constructed by two complex scalar fields

$$\Phi = \begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix}. \tag{1.39}$$

Each component ϕ_i satisfies the U(1) Lagrangian :

$$\mathcal{L} = \left(\left(\partial_\mu + i\frac{g}{2}B_\mu Y\right)\phi_i^\dagger\right)\left(\left(\partial^\mu - i\frac{g}{2}B^\mu Y\right)\phi_i\right) + \rho^2\phi_i^\dagger\phi_i - \lambda\left(\phi_i^\dagger\phi_i\right)^2,$$

and the scalar doublet Φ satisfies the Lagrangian :

$$\mathcal{L} = \left(\left(\partial_\mu + i\frac{g}{2}B_\mu Y + ig'W_\mu\right)\Phi^\dagger\right)\left(\left(\partial^\mu - i\frac{g}{2}B^\mu Y - ig'W^\mu\right)\Phi\right) + \rho^2\Phi^\dagger\Phi - \lambda\left(\Phi^\dagger\Phi\right)^2. \tag{1.40}$$

Customary, later 2 terms are called the Higgs potential :

$$V(\Phi) \equiv -\rho^2\Phi^\dagger\Phi + \lambda\left(\Phi^\dagger\Phi\right)^2. \tag{1.41}$$

In short, what we should consider here is this Lagrangian :

$$\begin{aligned}\mathcal{L} = &\overline{\Psi} i\not{D}\Psi \\ &- \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \\ &- \frac{1}{4}\mathrm{Tr}W^{\mu\nu}W_{\mu\nu} \\ &+ D_\mu\Phi^\dagger D^\mu\Phi + \rho^2\Phi^\dagger\Phi - \lambda\left(\Phi^\dagger\Phi\right)^2\end{aligned} \tag{1.42}$$

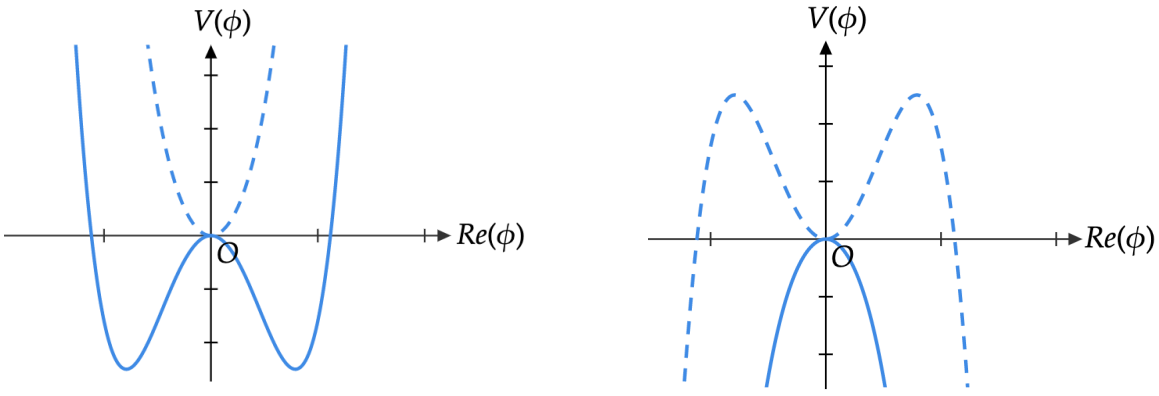
where $D_\mu \equiv \partial_\mu - i\frac{g}{2}B_\mu Y - ig'W_\mu$. This Lagrangian has two symmetries; the first one is the one which is generated by Y, and another one is generated by T^a. So the Lagrangian eq.(1.42) has the symmetry $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$. The index L indicates that SU(2) gauge fields couple to left chiral spinor only, and the index Y indicates that U(1) is generated by the generator Y. We should recognize it as the unified theory which is satisfied at high energy region. Then, how to generate mass terms from this theory? Higgs potential plays an important role in this. It can be rewritten as the function on the complex plane[23] :

$$\begin{aligned}V(\Phi) &= -\rho^2\phi_1^\dagger\phi_1 + \lambda\left(\phi_1^\dagger\phi_1\right)^2 - \rho^2\phi_2^\dagger\phi_2 + \lambda\left(\phi_2^\dagger\phi_2\right)^2 \\ &= V_1(\phi_1) + V_2(\phi_2).\end{aligned}$$

[22] In the previous section, the U(1) generator were not written explicitly. It is recognized as 2×2 identity.

[23] In this calculation, the cross term such as $\phi_1^\dagger\phi_1\phi_2^\dagger\phi_2$ can be eliminated by retaking the criterion of potential

The following figure.2 shows the plot of the function $V(\phi) = -\rho^2\phi^\dagger\phi + \lambda\left(\phi^\dagger\phi\right)^2$ on the complex plane.



(a) $\lambda > 0$ **(b)** $\lambda < 0$

Figure 2: Higgs potential : They are drawn at the 2-dimensional plane of $V(\phi)$ and Reϕ, but there is one more axis Imϕ which belongs to perpendicular direction to paper. The stable point is different according to the sign of ρ^2(dashed lines are $\rho^2 < 0$, solid lines are $\rho^2 > 0$). Additionally, the vacuum's stability is decided by whether λ is positive or not.

What we should focus on is the stable point of potential is changed by the sign of ρ^2. The stable point can be interpreted that the expectation value of the scalar field for the grand state(that is, vacuum), so this stable point is called the Vacuum Expectation Value(VEV). It should be noted that there is one more axis Imϕ which belongs to perpendicular direction to paper, so $V(\phi)$ should be a curved surface. Namely, for $\lambda > 0$[24], if the sign of ρ^2 is flipped negative to positive, the VEV of scalar field shifts to non-zero, and this vacuum state infinitely degenerates on the circle which the radius is $|\phi| = \sqrt{\frac{\rho^2}{2\lambda}}$. So, the VEV of the scalar field can be expressed by using the continuous phase parameter φ :

$$\phi_{min} = \sqrt{\frac{\rho^2}{2\lambda}} e^{i\varphi}.$$

At first, the vacuum places at the origin, but the vacuum choose one stable point from infinite ϕ_{min}s and "roll down" to there with the flipping of sign of ρ^2. Returning to the story, we are considering the scalar doublet model. So both Higgs potentials $V_1(\phi_1)$ and $V_2(\phi_2)$ should choose of one vacuum each other :

$$\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rightarrow \Phi_{min} = \begin{pmatrix} \phi_{1\ min} \\ \phi_{2\ min} \end{pmatrix}$$

Among infinity choices, let us take the following :

$$\begin{aligned} \Phi_{min} &= \begin{pmatrix} \phi_{1\ min} \\ \phi_{2\ min} \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \sqrt{\frac{\rho^2}{2\lambda}} \end{pmatrix} \equiv \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \quad \left(v \equiv \sqrt{\frac{\rho^2}{\lambda}} \right). \end{aligned} \tag{1.43}$$

[24] If $\lambda < 0$, the origin is not the perfect stable vacuum. Is the parameter λ positive really? It is still one of the mysteries which relates to the stability of vacuum.

We can see that new field $h(x)$ appears in the scalar doublet by expanding Φ around the VEV Φ_{min} (the detail of this calculation can be checked in Appendix B) :

$$\Phi = e^{i\theta_a \frac{\sigma_a}{2}} \begin{pmatrix} 0 \\ \frac{v+h(x)}{\sqrt{2}} \end{pmatrix} \tag{1.44}$$

Since Φ is transformed by local SU(2) transformation :

$$\Phi \to e^{ib_a(x) \frac{\sigma_a}{2}} \Phi,$$

if we choose the appropriate phase $b_a(x)$ so that it eliminates the phase factor of eq.(1.44)[25], we have only one physical field $h(x)$ which is called Higgs field[26] :

$$\Phi = \begin{pmatrix} 0 \\ \frac{v+h(x)}{\sqrt{2}} \end{pmatrix} \tag{1.45}$$

If we use the scalar doublet, which is imposed Unitary gauge eq.(1.45), the Lagrangian eq.(1.40) includes the coupling between Higgs field and gauge fields or self-coupling of Higgs field. Here, let us consider the kinetic term, which includes the VEV of the scalar doublet. The detail of the following calculation is shown in Appendix B.

$$eq.(1.40) \supset D_\mu \begin{pmatrix} 0 & \frac{v}{\sqrt{2}} \end{pmatrix} D^\mu \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = \frac{v^2}{8} \left| g'^2 \left((W_1^\mu + iW_2^\mu)^2 + (g'W_3^\mu - gB^\mu)^2 \right) \right| \tag{1.46}$$

After a long calculation, it outputs mass terms of weak bosons, which are the purpose of this section :

$$eq.(1.46) = \frac{v^2 g^2}{4} W_\mu^+ W_\mu^- + \frac{1}{8} v^2 \left(g^2 + g'^2 \right) Z_\mu Z^\mu + \frac{1}{8} v^2 \cdot 0 A_\mu A^\mu \tag{1.47}$$

Here, $W_\mu^\pm$ means $W^\pm$ bosons, Z_μ is Z boson, A_μ is photon. These fields are defined by

$$W_\mu^\pm \equiv \tfrac{1}{\sqrt{2}} (W_{1\mu} \mp W_{2\mu}) \tag{1.48}$$

$$Z_\mu \equiv \tfrac{1}{\sqrt{g^2+g'^2}} (g'W_{3\mu} - gB_\mu) = \sin\theta_W W_{3\mu} - \cos\theta_W B_\mu \tag{1.49}$$

$$A_\mu \equiv \tfrac{1}{\sqrt{g^2+g'^2}} (gW_{3\mu} + g'B_\mu) = \cos\theta_W W_{3\mu} + \sin\theta_W B_\mu \tag{1.50}$$

So mixed gauge fields correspond to each gauge boson of electromagnetic and weak forces, but not gauge fields which appeared in eq.(1.42) themselves. Photon and Z boson is appeared by mixing of $W_{3\mu}$ and B_μ, and the angle θ_W specifies the degree of mixing. θ_W is called Electroweak mixing angle, or Weinberg angle[27]. Let us organize the story. At a higher energy scale, $\mathrm{SU(2)_L \times U(1)_Y}$ Lagrangian eq.(1.42) is satisfied, and Higgs potential has a stable point at the origin. However, the universe becomes colder with time evolution, and the energy scale reaches to EW scale(about 100GeV). Then, the VEV of Higgs doublet shift to non-zero and mass terms of gauge bosons appear as we have seen at eq.(1.47)[28]. So in this phase, $\mathrm{SU(2)_L}$ is lost anymore, and only U(1) symmetry of electromagnetic force has remained. Therefore gauge symmetry $\mathrm{SU(2)_L \times U(1)_Y}$ is broken, and its smaller subgroup $\mathrm{U(1)_{QED}}$ is remained

[25] This choice is called Unitary gauge.

[26] Because of this reason, scalar doublet Φ is often called Higgs doublet too.

[27] Weinberg angle is measured precisely; $\sin^2\theta_W$=0.23122(15) [5].

[28] Obviously, shifting of VEV is the trigger of this story, and it is called Phase transition of vacuum. Why does such a phenomenon occur? It is one of the profound mysteries yet.

by at lower energy scale. This phenomenon is called Spontaneous (gauge) symmetry breaking. Additionally, based on spontaneous symmetry breaking, the mechanism which gives mass terms through Higgs doublet is called the Higgs mechanism.

Finally, let us see how Dirac fermions obtain these masses by the Higgs mechanism. We can imagine that the mass term of Dirac fields should be $\propto \overline{\Psi}\Psi$. It is a candidate because it is Lorentz invariant. By the way, it can be rewritten as the combination of left chiral and right chiral spinors :

$$\overline{\Psi}\Psi = \overline{\Psi}_L\Psi_R + \overline{\Psi}_R\Psi_L$$

To make the Lorentz invariance term, we need to couple a left chiral component and a right chiral component. However, since Ψ_L and Ψ_R have different transformation properties under $\mathrm{SU(2)_L}$, this combination does not SU(2) gauge invariant. To consistent both Lorentz invariance and gauge invariance, the combination $\overline{\Psi}_L\Phi\Psi_R$ is good because each gauge transformation property comes from $\overline{\Psi}_L$ and Φ are canceled, and Ψ_R does not change at all. Such interaction form is called Yukawa interaction. It should be noted that at this point, there are 2 Dirac fermions ψ_1 and ψ_2, and there are three SU(2) objects :

$$\Psi_L = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \ , \ \psi_{1R} \ , \ \psi_{2R}$$

Then, we can consider 2 combinations; Ψ_L and ψ_{1R}, Ψ_L and ψ_{2R}. At first, let us consider second one :

$$\mathcal{L} = -\lambda_2 \left(\overline{\Psi}_L\Phi\psi_{2R} + h.c.\right) \tag{1.51}$$

The calculation itself is simple :

$$\begin{aligned}
\mathcal{L} &= -\lambda_2 \left(\overline{\Psi}_L\Phi\psi_{2R} + h.c.\right) \\
&= -\frac{\lambda_1}{\sqrt{2}} \left(\overline{\psi}_{2L}\left(v+h\right)\psi_{2R} + \overline{\psi}_{2R}\left(v+h\right)\psi_{2L}\right) \\
&= -\frac{\lambda_2 v}{\sqrt{2}}\overline{\psi}_2\psi_2 - \frac{\lambda_2}{\sqrt{2}}h\overline{\psi}_2\psi_2
\end{aligned}$$

The first term is the mass term of Dirac fermion ψ_2. How about the mass term of ψ_1? We need ingenuity a little so that the upper component of Higgs doublet Φ to be non-zero, and Charge conjugated Higgs doublet is the answer. If we act charge conjugation on Higgs doublet,

$$\Phi = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \rightarrow \tilde{\Phi} \equiv \epsilon\Phi^* = \begin{pmatrix} & +1 \\ -1 & \end{pmatrix}\begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{v+h}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

In order to obtain the mass term of ψ_1, we need to introduce the following Lagrangian :

$$\mathcal{L} = -\lambda_1 \left(\overline{\Psi}_L\tilde{\Phi}\psi_{1R} + h.c.\right) \tag{1.52}$$

Similar to eq.(1.51), it outputs the mass term of ψ_1 :

$$\begin{aligned}
\mathcal{L} &= -\lambda_1 \left(\overline{\Psi}_L\tilde{\Phi}\psi_{1R} + h.c.\right) \\
&= -\frac{\lambda_1}{\sqrt{2}} \left(\left(v+h\right)\overline{\psi}_{1L}\psi_{1R} + \left(v+h\right)\overline{\psi}_{1R}\psi_{1L}\right) \\
&= -\frac{\lambda_1 v}{\sqrt{2}}\overline{\psi}_1\psi_1 - \frac{\lambda_1}{\sqrt{2}}h\overline{\psi}_1\psi_1
\end{aligned}$$

Therefore, mass terms of Dirac fermions are introduced by Yukawa interactions and the Higgs mechanism. In general, the fermion mass is give by

$$m_f = \frac{\lambda_f v}{\sqrt{2}}.$$

Customary, Yukawa coupling constant between Dirac fermion f and Higgs boson h is defined by

$$y_{ffh} \equiv \frac{\lambda_f}{\sqrt{2}} = \frac{m_f}{v} \tag{1.53}$$

Namely, fermion masses are decided by Yukawa couplings fundamentally.

In this section, we have seen that the unified theory of electroweak interaction and mass generation mechanism. The Lagrangian at EW scale is

$$\begin{aligned}\mathcal{L} = & \overline{\Psi} i \not{D} \Psi \\ & - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} \\ & - \frac{1}{4} \mathrm{Tr} W^{\mu\nu} W_{\mu\nu} \\ & + D_\mu \Phi^\dagger D^\mu \Phi + \rho^2 \Phi^\dagger \Phi - \lambda \left(\Phi^\dagger \Phi\right)^2 \\ & + \lambda_2 \overline{\Psi}_L \Phi \psi_{2R} + \lambda_1 \overline{\Psi}_L \tilde{\Phi} \psi_{1R} + h.c.\end{aligned} \tag{1.54}$$

In this paper, mass is the most important concept. However, we should know that the mass which appears in Lagrangian differs from the observed mass at experiments. The next section explains it.

1.3 Renormalization

1.3.1 Renormalization prescription

In the previous section, we have seen that Electroweak gauge theory as interaction theory. There is an obvious difference between the free field theory and the interaction field theory, whether the equation of motion can be solved exactly. In free field theory, the equation of motion(e.g. Klein-Gordon equation, Dirac equation) can be solved by expanding plane wave solutions exactly. Since such expanded solution has creation and annihilation operators of corresponded particles, and one particle state $|\boldsymbol{k}\rangle$ which has clear physical meaning (a particle which momentum is $\boldsymbol{k}$, energy is k^0, mass is $m^2 = \left(k^0\right)^2 - \boldsymbol{k}^2$) can be defined. On the other hand, in interaction field theory, we can not solve the equation of motion exactly, and creation and annihilation operators depend on time. One particle state can not have a clear physical meaning, such as free field theory, under such operators. So in the case of free theory, we can regard the parameter $m^2 = \left(k^0\right)^2 - \boldsymbol{k}^2$ in free Lagrangian as the mass of one particle state $\boldsymbol{k}$, but in the case of interaction theory, it is unclear that the mass parameter in Lagrangian can be regarded as the physical mass of one particle state. To understand this story, let us consider the following parable. Let us consider a particle, and it can move smoothly in vacuum space. However, if this particle is in some jelly box, it will looks "heavier" because of friction(interaction) between jelly and particle. Moreover, how to see the particle's motion will change according to the hardness of jelly(corresponds to the strength of interaction). Therefore,

how to see the "effective mass" of the particle will change in the free case(that is, mass in Lagrangian) or interaction case(mass we can observe actually). Based on this sense, the mass parameter in Lagrangian is not the same as the observed mass. For example, in the case of scalar ϕ^4, its Lagrangian is given as

$$\mathcal{L} = \frac{1}{2}\partial^\mu \phi_B \partial_\mu \phi_B - \frac{1}{2}m_B^2 \phi^2 - \frac{1}{4!}\lambda_B \phi^4 \tag{1.55}$$

and its mass appears in the second term, but it is not the same as the observed mass. Then, what are the relation between the mass in Lagrangian and the observed mass? The theory which answers this question is the Renormalization theory.

Before looking at the formulation of renormalization theory, let us see the policy of renormalization briefly. As described above, in general, the mass parameter m_B does not match with observed physical mass m_{phys}. Former mass is called bare mass, and later mass is called renormalized mass. In QFT, the n-point function is the most important object which connects to observables directly. So let us consider the 2-point function to look at how renormalization appears. It includes all possible diagrams, which has two external lines :

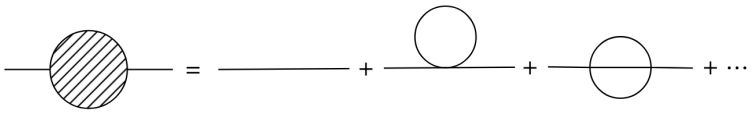


Figure 3: 2-point function diagrams : The first diagram is tree-level, and later diagrams appear as quantum corrections.

The first diagram is just a propagator, and later diagrams are contributed as Quantum corrections. The problem starts from the second loop diagram. According to Feynman rules, all these diagrams are translated to formulas, and the second diagram includes one momentum integration. This momentum integration has a divergence at a large momentum region called UV(Ultra Violet) divergence. It can be checked like this: a loop is translated to an integration measure $\sim \int d^4 l$ (l is the momentum on a loop), and it includes third-order momentum l^3 by the transformation to polar coordinates. Additionally, the loop in the second diagram is constructed by one propagator, and there is a negative second-order momentum l^{-2} per one propagator. So in the second diagram, the momentum integrand becomes $l^{3-2} = l$ approximately, and its momentum integration diverges at a large momentum region because the integrand is integrated as quadratic over its integration region $(-\infty, +\infty)$[29]. Such a type of divergence is called Quadratic divergence, and it is a critical problem because the 2-point function connects to observables.

How should we overcome this problem? The basic idea is like this: Observables that we can see at actual experiments must be finite quantities. However, other quantities that appear in the calculation only, but not observations, can be infinite. In the above case, the observable parameter is mass in the propagator, and it should be noted that this mass comes from the mass term of Lagrangian(so it is the bare mass). Namely, the bare mass which appears in the above

[29]It should be noted that all loop diagrams do not necessarily diverge. The third diagram in figure.3 does not diverge because it has three propagators for one loop. Its momentum order of integrand will be $l^{3-(2\times 3)} = l^{-3}$, it is safe. However, it should be noted that this is the case of 4-dimensional spacetime. Whether the diagram diverges or not depends on the dimension of spacetime.

diagrams is not observable, but this bare mass appears as the finite form in our side through "Renormalization prescription". According to this policy, renormalization needs to remove infinity in quantum correction, and it is accomplished by introducing counterterms that cancel divergences and remain finite quantities. In short, parameters in Lagrangian include divergence, but we cannot observe itself. We can observe "renormalized" finite quantities. It is the core idea of renormalization, and we need to construct theory so that it does not output infinity as observables. As mentioned in section 1.1.2, this is one of the essential requirements for the construction of Lagrangian. Then, let us look more concretely formulation of renormalization in the next section.

1.3.2 $\overline{\text{MS}}$ scheme

In the previous section, a more intuitive sense of renormalization is explained. To grasp the formulation of renormalization, let us consider the case of the scalar ϕ^4 model. As explained above, the second diagram in figure.3 has quadratic divergence. It can be seen from the result of the translation of diagram[30] :

$$2nd\ diagram \sim \lambda \int \frac{d^4l}{(2\pi)^4} \frac{i}{l^2 - m^2} \tag{1.56}$$

By transformation of the measure to the polar coordinates, it can be rewritten like this and it has divergence :

$$\sim \lambda \int^{\Lambda} \frac{dl\ l^3}{(2\pi)^4} \int d\Omega_4 \frac{1}{l^2} \sim \Lambda^2 \tag{1.57}$$

Here, the upper limit Λ is introduced, and it is taken the limit to infinity later. $d\Omega_4$ is the angular element. This integration is calculated, but there is a divergence with quadratic order under $\Lambda \to \infty$.

In order to deal with such divergence well mathematically, we apply the method which is called Regularization. There are various styles for regularization, but the core point of this prescription is to modify the integration with divergence so that it can be calculated finitely. For example, Pauli-Villars(PV) regularization is one way. In this method, the propagator is replaced to like the following :

$$\frac{1}{l^2 + m^2} \to \frac{1}{l^2 + m^2} \frac{\Lambda^2}{l^2 + \Lambda^2} \tag{1.58}$$

This method keeps Lorentz invariance but does not gauge invariance, and it is more complicated than other methods. On the other hand, we have one more way, which is easier than PV regularization and is gauge invariant. It is Dimensional regularization. Its basic idea is like this: since the momentum integration has divergence on 4-dimensional spacetime, replace it with the momentum integration on general d-dimensional spacetime so that the integration result to be finite. It should be reminded that such divergence can be determined by the relation between loop number and propagator number. So we can avoid this divergence by adjusting the dimension of spacetime. It may be eccentric, but good to overcome the problem of divergence. However, when we apply this method, we need to care about the mass dimension

[30] Note that each quantity ϕ, m, λ is bare quantity.

of Lagrangian. d-dimensional action is

$$S[\phi] = \int d^d x \left(\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \frac{1}{4!} \lambda \phi^4 \right), \tag{1.59}$$

and since action S is dimensionless, the mass dimension of scalar field ϕ is

$$[\phi] = \frac{d-2}{2}. \tag{1.60}$$

Of course, for $d = 4$, it is just 1, as we know. Namely, the coupling constant λ is dimensionless for 4-dimensional spacetime, but it also needs to be rewritten in this case. For d-dimensional spacetime, the mass dimension of λ is

$$[\lambda] = -d + 4, \tag{1.61}$$

and it is nonzero in general. However, dimensionless coupling tends to be more interesting than the nonzero dimensional one because observables such as scattering amplitude depend on coupling constants, and these quantities themselves should be dimensionless. Because of this reason, we should construct a dimensionless coupling constant even in the case of general dimensional spacetime. It is realized by replacing λ to like this :

$$\lambda \to \lambda \mu^{4-d} = \lambda \mu^{-[\lambda]}. \tag{1.62}$$

If we set the mass dimension of a new parameter μ to be $[\mu] = 1$, by regarding it as a new coupling constant, and it is a dimensionless coupling for general dimensional-spacetime.

Then, let us look at the concrete prescription of dimensional regularization. The method itself is straightforward; replace 4-dimension to general d-dimension. In the case of eq.(1.56), it is replaced like this :

$$2nd\ diagram \sim \lambda \mu^{4-d} \int \frac{d^d l}{(2\pi)^d} \frac{1}{l^2 + m^2}. \tag{1.63}$$

Under this replacement, let us calculate it by using some tricks. At first, let us transform this momentum integration to the polar coordinate's one :

$$\lambda \mu^{4-d} \int \frac{dl\ l^{d-1}}{(2\pi)^d} \int d\Omega_d \frac{1}{l^2 + m^2}. \tag{1.64}$$

Here, $\int d\Omega_d$ is the volume of d-dimensional unit sphere(See Appendix A). Using eq.(Appendix A.1) and eq.(Appendix A.2), the momentum integration eq.(1.64) can be evaluated as

$$\lambda \mu^{4-d} \frac{1}{(2\pi)^d} \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \frac{\Gamma\left(\frac{d}{2}\right) \Gamma\left(\frac{d}{2} - 1\right)}{2\Gamma(1)} = \lambda \mu^{4-d} \frac{\Gamma\left(1 - \frac{d}{2}\right)}{(4\pi)^{\frac{d}{2}}} \left(m^2\right)^{\frac{d}{2}-1}. \tag{1.65}$$

Let us define a new parameter $\epsilon \equiv 2 - \frac{d}{2}$. Note that taking the limit $\epsilon \to 0$ is equivalent to taking the limit $d \to 4$. Using this parameter, the gamma function in eq.(1.65) is rewritten as $\Gamma(\epsilon - 1)$ and it can be expanded as

$$\Gamma(\epsilon - 1) = -\frac{1}{\epsilon} + \gamma - 1 + O(\epsilon). \tag{1.66}$$

$\gamma = 0.5772\cdots$ is the Euler-Mascheroni constant. Expanding eq.(1.65) up to the order of $O(\epsilon)$, we obtain

$$\frac{\lambda}{(4\pi)^2} m^2 \left(\frac{4\pi\mu^2}{m^2} \right)^\epsilon \left(-\frac{1}{\epsilon} + \gamma - 1 + O(\epsilon) \right).$$

If ϵ is extremely small, we can use the following approximation :

$$\left(\frac{4\pi\mu^2}{m^2}\right)^{\epsilon} = \exp\left(\epsilon \ln\frac{4\pi\mu^2}{m^2}\right) \sim 1 + \epsilon \ln\frac{4\pi\mu^2}{m^2} + O\left(\epsilon^2\right),$$

and the result of evaluation of eq.(1.56) is

$$2nd\ diagram \sim \frac{\lambda}{(4\pi)^2} m^2 \left(-\frac{1}{\epsilon} + \gamma - \ln 4\pi - 1 + \ln\frac{m^2}{\mu^2}\right). \tag{1.67}$$

Thanks to regularization prescription, we could calculate it in a finite form. However, It should be noted that this result eq.(1.67) has the possibility of divergence. When we return to 4-dimensional spacetime, the first term will diverge. So we have not removed divergence yet, but it anymore can be removed easily. To do this, we need one more procedure; Renormalization. The basic idea is to add new terms so that they can absorb appeared divergence. In this 2-point function case, we should introduce the term proportional to ϕ^2. So let us add a term like this to Lagrangian :

$$\mathcal{L}_{counter} = -\frac{1}{2}\delta m^2 \phi_R^2. \tag{1.68}$$

ϕ_R is the renormalized scalar field and let us define it by the linear relation $\phi_B = \sqrt{Z}\phi_R$. We can recognize a factor $\sqrt{Z}$ as just a coefficient that connects bare field and renormalized field[31]. Here, δm^2 is the coefficient that has the mass dimension of 2, and it is adjusted so that it can cancel the divergence(See figure.4). Various methods exist in order to decide the explicit form of δm^2, but one of the familiar ways is called Minimal Subtraction(MS) scheme or Modified Minimal Subtraction($\overline{\text{MS}}$) scheme. In $\overline{\text{MS}}$ scheme, introduce the counter term so that it can absorb only divergent terms at a minimum[32] :

$$\mathcal{L}_{counter} = -\frac{1}{2}\delta m^2 \phi_R^2 \quad \left(\delta m^2 = \frac{\lambda}{(4\pi)^2} m^2 \left(\frac{1}{\epsilon} - \gamma + \ln 4\pi\right)\right). \tag{1.69}$$

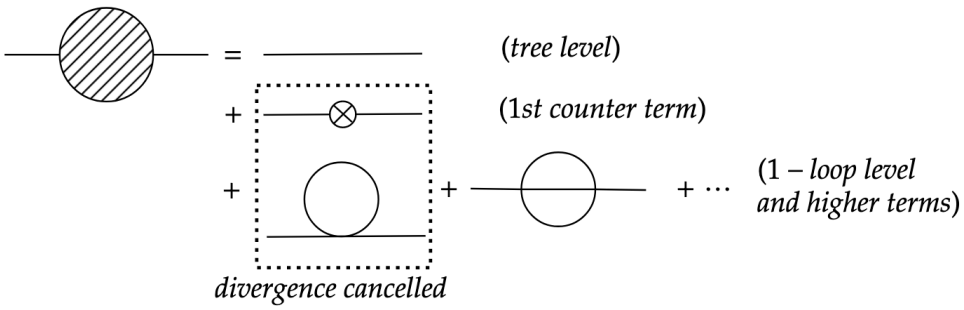


Figure 4: Renormalized 2-point function diagrams : By adding a new term(called Counterterm) to eq.(1.68), a new vertex and its diagrams are introduced. The cross vertex means the coefficient of counterterm.

It should be noted that the calculation result completely depends on how to define counter vertex, and this dependence is called Scheme dependence. We have three parameters in the

[31] If you care about the physical meaning of $\sqrt{Z}$, you should reference [4].

[32] In MS scheme, it subtracts only divergent term :

$$\delta m^2 = \frac{\lambda}{(4\pi)^2} m^2 \frac{1}{\epsilon}.$$

In $\overline{\text{MS}}$ scheme, it subtracts frequently appeared constants γ and $\ln 4\pi$ in addition to the above divergent term.

original Lagrangian; ϕ, m, λ. Since the counter vertex δm^2 corrects bare mass m_B to renormalized mass m_R, we need two more counterterms that correct the field and the coupling constant. Let us introduce them as[33] :

$$+\frac{1}{2}\delta Z \partial^\mu \phi_R \partial_\mu \phi_R \ , \ -\frac{1}{4!}\delta\lambda\phi_R^4 \tag{1.70}$$

δZ and $\delta\lambda$ is counter vertices. Renormalized Lagrangian $\mathcal{L}_R$ should be obtained by $\mathcal{L}_B + \mathcal{L}_{counter}$:

$$\begin{aligned}\mathcal{L}_R &= \frac{1}{2}\partial_\mu\phi_R\partial^\mu\phi_R - \frac{1}{2}m_R^2\phi_R^2 - \frac{1}{4!}\lambda_R\phi_R^4 \\ &= \frac{1}{2}\partial_\mu\phi_B\partial^\mu\phi_B - \frac{1}{2}m_B^2\phi_B^2 - \frac{1}{4!}\lambda_B\phi_B^4 \\ &+ \frac{1}{2}\delta Z\partial_\mu\phi_R\partial^\mu\phi_R - \frac{1}{2}\delta m^2\phi_R^2 - \frac{1}{4!}\delta\lambda\phi_R^4\end{aligned}$$

To satisfied this equation, each bare and renormalized parameter should satisfy the following relations :

$$Z = 1 + \delta Z \ , \ m_R^2 = Zm_B^2 - \delta m^2 \ , \ \lambda_R = Z^2\lambda_B - \delta\lambda. \tag{1.71}$$

So if we add this counter vertex, which includes divergence, to bare Lagrangian, the divergence which comes from bare theory cancel with the divergence of counter vertex, and only finite contribution remains. Remained finite quantity is what we can observe at actual experiments. It is the main story of renormalization. To absorb loop divergence, we need to introduce some counterterms, as we have seen. The theory which can be removed divergence by adding finite counterterms is called Renormalizable theory. On the other hand, the theory that needs infinite counterterms to remove divergence is called the Non-renormalizable theory. The non-renormalizable theory is inconvenient because we need infinite terms in Lagrangian. So the renormalizability of theory is one of the important requirements for the construction of Lagrangian. It is familiar that the theory which mass dimension of coupling constant is not negative is renormalizable.

Before moving to the next chapter, let us introduce one more important thing about renormalization; running parameters. As described above, we can choose $\delta Z, \delta m^2, \delta\lambda$ arbitrary under various renormalization scheme, and each condition to determine counter vertices is called Renormalization condition. For example, in the case of the coupling constant, the statement of renormalization condition is like this: renormalized 4-point vertex function(figure.5) should be the finite coupling constant(that is, renormalized coupling). It is a natural condition, and the formulation of this condition is given by

$$\Pi_4\left(p_1, \cdots, p_4\right)|_\mu = -i\lambda_R. \tag{1.72}$$

μ is called Renomalization point, and there are some choices. For example, fix external momentums p_i $(i = 1, \cdots, 4)$ by

$$\mu \ : \ p_1^\mu = p_2^\mu = -p_3^\mu = -p_4^\mu = (m, \mathbf{0}). \tag{1.73}$$

m is the (physical) mass of incoming or outgoing particles. It means that we are considering the physical system under the energy corresponds to m, and the 4-point vertex function should

[33] Here, I keep explanation in the only overview of renormalization theory so that you can grasp the sense of renormalization prescription. You can check more detail of renormalization theory in professional books. As references, [3], [4] are gentle books.

be finite as the renormalized coupling constant under this energy scale. It should be noted that $\Pi_4(p_1, \cdots, p_4)$ equals to renormalized coupling λ_R on the condition of eq.(1.73), but in general, it has different value at other energy regions. This thing can be interpreted that coupling constant λ_R depends on energy region(renormalization point), and it is no longer constant. This nature appears to mass too, and such phenomenon which renormalized parameters have energy dependence is called "running". How to run these mass and coupling constants is described by Renormalization Group Equation (RGE), and it can be seen in the later chapter together with running mass. In this paper, we are thinking about the running of quark mass. Quark is the particle that feels strong interaction. However, strong interaction has not been explained yet in this paper. Let us see about this new gauge interaction in the next section.

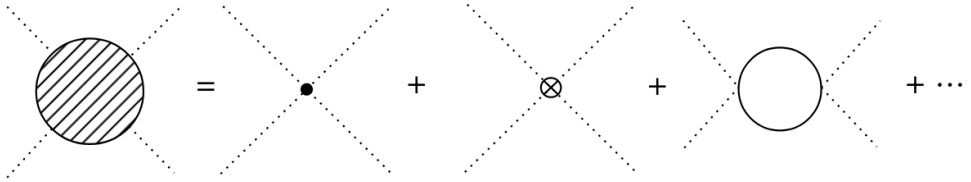


Figure 5: Renormalized 4-point vertex function : The cross vertex in the second diagram indicates that the counter vertex $\delta\lambda$. The third 1-loop diagram has logarithmic divergence, but it can be removed by introducing the counter vertex $\delta\lambda$.

1.4 Quantum Chromodynamics

1.4.1 SU(3) gauge theory

This section explains the third gauge interaction, strong interaction. It is explained by SU(3) gauge theory. SU(3) is also a non-abelian group, but it is more complicated than SU(2). Lie algebra of SU(3) gauge group is given by

$$[T^a, T^b] = i f_{abc} T^c \tag{1.74}$$

f_{abc} is the structure constant of SU(3), and they are given by

$$\begin{aligned} &f_{123} = 1 \\ &f_{147} = f_{165} = f_{246} = f_{257} = f_{345} = f_{376} = \frac{1}{2} \\ &f_{458} = f_{678} = \frac{\sqrt{3}}{2}. \end{aligned}$$

The basic idea of theory construction is almost the same as SU(2) gauge theory. Here, let us introduce the Dirac triplet, which is constructed by three Dirac fields :

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}, \tag{1.75}$$

and since the degree of freedom is 8 for SU(3), the covariant derivative is defined by

$$D_\mu = \partial_\mu + i g_s T^a G^a_\mu \quad (a = 1, \cdots, 8). \tag{1.76}$$

So the Lagrangian

$$\mathcal{L} = \overline{\Psi} i \not{D} \Psi = \overline{\Psi} i \not{\partial} \Psi - g_s \overline{\Psi} \not{G} \Psi \tag{1.77}$$

explains the strong interaction. The SU(3) gauge fields G_μ are called Gluon, and it is confirmed that they are massless particles. The mass terms of ψ_1, ψ_2, ψ_3 are provided by corresponded Yukawa interactions. The particle with an electric charge can feel an electromagnetic field, and the particle with weak isospin can feel weak interaction. From this analogy, strong interaction should have some charge to couple between gluon and Dirac fermions. For SU(3), there are two diagonal elements in the fundamental representation of SU(3) :

$$T^3 = \frac{1}{2}\begin{pmatrix} +1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ , \ T^8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} +1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

and we can construct eigenstates for them simultaneously. In this case, these eigenstates are SU(3) triplet components :

$$\begin{pmatrix} \psi_1 \\ 0 \\ 0 \end{pmatrix} \ , \ \begin{pmatrix} 0 \\ \psi_2 \\ 0 \end{pmatrix} \ , \ \begin{pmatrix} 0 \\ 0 \\ \psi_3 \end{pmatrix},$$

and they can be labeled by eigenvalues of $\left(T^3, T^8\right)$. The combination that we can consider is

$$\left(+\tfrac{1}{2} \ , \ +\tfrac{1}{2\sqrt{3}} \right) \ for \begin{pmatrix} \psi_1 \\ 0 \\ 0 \end{pmatrix} \ , \left(-\tfrac{1}{2} \ , \ +\tfrac{1}{2\sqrt{3}} \right) \ for \begin{pmatrix} 0 \\ \psi_2 \\ 0 \end{pmatrix} \ , \left(0 \ , \ -\tfrac{1}{\sqrt{3}} \right) \ for \begin{pmatrix} 0 \\ 0 \\ \psi_3 \end{pmatrix}.$$

These three components are classified by comparing to the three primary colors; Red ψ_R, Blue ψ_B, Green ψ_G, and the charge which label each component is called Color charge. When Red and Blue and Green are summed $R + G + B$, it becomes "White". In the case of QED, if the summed electric charge of the considering system is neutral, it does not feel electromagnetic force at all. Like this, in the case of SU(3) gauge theory, if the summed color charge of the considering system is White, it is neutral for the strong interaction. Such a state is called the Color singlet. As long as the particle has a not-white color charge, it continues to react under strong interaction, and when the summed color charge becomes white, the reaction is stopped. So we can only observe color singlet objects in nature[34]. What we should pay attention to is there is anti-version for the color charge; anti-Red, anti-Blue, anti-Green[35]. Anti-colors correspond to complementary colors, and we can construct color singlet by combine color and anti-color :

$$R\overline{R} \ or \ B\overline{B} \ or \ G\overline{G}.$$

Why do we need to introduce such a new interaction in addition to electromagnetic and weak interactions? As we know, matters are constructed by atoms. Moreover, each atom is constructed by electrons and protons, and neutrons(later two are called nucleons). Furthermore,

[34] Can we extract single color charge by our hand? The answer is No. It is because the strength of strong interaction becomes larger at lower energy(or long distance). This unique nature of strong interaction is explained in the next section.

[35] It is because the fundamental representation of SU(3)(it is called three representation too) is not equivalent with its $\overline{3}$ representation (conjugate representation of 3 representation).

these nucleons are constructed by three charged particles, which are called Quarks. Experiments revealed that each quark has a nonzero electric charge. However, why can these quarks cohesive and form proton or neutron even though they feel the repulsive electromagnetic force? So something new force which confines quarks is needed, and it is strong interaction. Namely, each quark has a color charge; red, blue, or green. Proton and neutron are formed by three quarks so that summed color charge to be a color singlet. We can confirm existence of color charge from the fact that the scattering cross-section $R \equiv \sigma(e^+e^- \to hadrons)/\sigma(e^+e^- \to \mu^+\mu^-)$ has a factor of 3. It is evidence that SU(3) triplet has 3 type color charges. Since quark feels strong interaction through the color charge, SU(3) gauge theory governs quarks called Quantum Chromodynamics(QCD)[36].

By the way, does the mediator of the strong interaction(that is, gluon) also have a color charge? Similar to the case of weak bosons, we can see it through the adjoint representation of SU(3) gauge group. Since the degree of freedom of SU(3) is 8, the size of adjoint representation is eight, and its object should have eight components. Each component indicates that different gluon state and they have eight color charges :

$$R\overline{G}\ ,\ G\overline{R}\ ,\ R\overline{B}\ ,\ B\overline{R}\ ,\ G\overline{B}\ ,\ B\overline{G}\ ,\ \frac{R\overline{R}-G\overline{G}}{\sqrt{2}}\ ,\ \frac{R\overline{R}+B\overline{B}+G\overline{G}}{\sqrt{6}}$$

It is called Color Octet. In this section, a new gauge interaction, strong interaction mediated by color charges, is introduced. The strong interaction is explained by SU(3) non-abelian gauge theory, called Quantum Chromodynamics (QCD). So at the point of non-abelian gauge theory, strong interaction and weak interaction are similar. However, there is a unique nature for strong interaction, and it is useful to understand this study's background physics. Let us see it in the next section.

1.4.2 Asymptotic freedom and Color confinement

As we have seen in the previous section, SU(3) gauge theory explains strong interaction. On the other hand, as described in section 1.3.2, coupling constants of interaction terms have energy dependence because of renormalization. In this section, The unique behavior of strong interaction through renormalization of strong gauge coupling constant is explained[37].

To see this, considering the vacuum polarization of QED is useful. Let us consider the 3-point vertex function(figure.6). For the considering process energy scale $Q \gg m_e$, the result which includes until the 1-loop correction is given like this :

$$\alpha_0 + \alpha_0\Pi^{(1)}\left(Q^2\right) = \alpha_0 - \frac{\alpha_0^2}{3\pi}\left(\Delta - \ln\left(\frac{Q^2}{m_e^2}\right)\right). \tag{1.78}$$

Here, Δ is a divergent quantity, α_0 is the fine structure constant at tree level. Based on it, we can calculate the 3-point vertex function easily :

$$\begin{aligned}\alpha_{eff}\left(Q^2\right) &= \alpha_0\left(1 + \Pi^{(1)} + \left(\Pi^{(1)}\right)^2 + \cdots\right) \\ &= \frac{\alpha_0}{1 - \Pi^{(1)}\left(Q^2\right)}\end{aligned} \tag{1.79}$$

[36] Chromo means color in Greek.
[37] The section 8.5 in [9] gives a good introduction of later discussion.

Since $\Pi^{(1)}$ includes infinity, this coupling eq.(1.79) is not physical coupling. However, if we choose the condition that effective coupling at lower energy limit $\left(Q^2 \to 0\right)$ which gives the measured coupling constant : $\alpha_{eff}(0) = \alpha_{mes} = 1/137$, we can eliminate this infinity :

$$\begin{aligned} \frac{1}{\alpha_{eff}\left(Q^2\right)} - \frac{1}{\alpha_{eff}(0)} &= \frac{1}{\alpha_0}\left(\Pi^{(1)}\left(0\right) - \Pi^{(1)}\left(Q^2\right)\right) \\ &= -\frac{1}{3\pi}\ln\left(\frac{Q^2}{m_e^2}\right) \\ \therefore\ \alpha_{eff}\left(Q^2\right) &= \frac{\alpha_{mes}}{1 - \frac{\alpha_{mes}}{3\pi}\ln\left(\frac{Q^2}{m_e^2}\right)} \end{aligned} \tag{1.80}$$

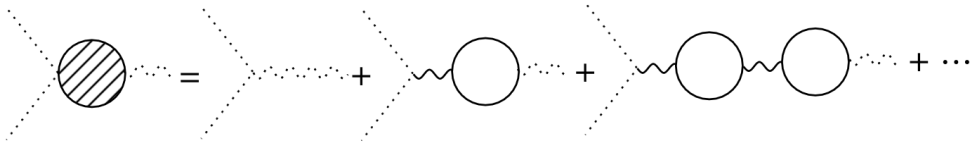


Figure 6: 3-point vertex function for QED

According to eq.(1.80), we can see that the coupling constant of electromagnetic interaction becomes larger at a higher energy scale, and it becomes infinity at the limit of $Q \to \infty$. It indicates that for QED, perturbative expansion bankrupts at enormous energy. It is a QED case, but we can look at the behavior of running of strong gauge coupling by the analogy of vacuum polarization. However, we need to take into account the self-coupling of gluons. The 3-point vertex function what we should consider is like this :

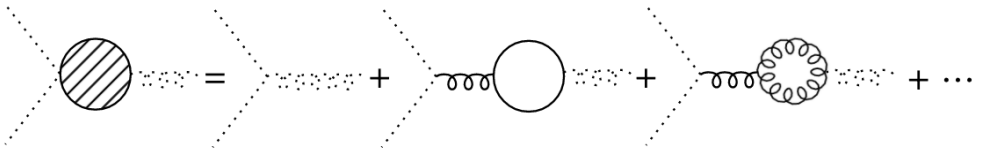


Figure 7: 3-point vertex function for QCD

Calculation is almost similar with QED case. The 1-loop level calculation is given like this :

$$\begin{aligned} &\alpha_{s0} + \alpha_{s0}\Pi^{(1)}\left(Q^2\right) = \alpha_{s0} + \frac{\alpha_{s0}^2}{4\pi}\beta_0\left(\Delta - \ln Q^2\right) \\ &\beta_0 = 11 - \frac{2}{3}n_f \end{aligned} \tag{1.81}$$

α_{s0} is gauge coupling of strong interaction at tree level, n_f is the number of active quark flavor with the considering energy scale Q. It should be noted that we can not use the condition that chooses the criterion point such as $\alpha_{eff}(0) = \alpha_{mes}$ in QED because we can only observe colorless particles and strong coupling can not be measured at lower energy region. Instead of the criterion at our energy scale, let us introduce some energy scale μ, which is called renormalized point or renormalized scale. Repeat same calculation with this scale μ,

$$\frac{1}{\alpha_s\left(Q^2\right)} - \frac{1}{\alpha_s(\mu^2)} = \frac{\beta_0}{4\pi}\ln Q^2$$

$$\therefore\ \alpha_s\left(Q^2\right) = \frac{\alpha_s\left(\mu^2\right)}{1 + \frac{\alpha_s(\mu^2)}{4\pi}\left(11 - \frac{2}{3}n_f\right)\ln Q^2} \tag{1.82}$$

The notable point is that the second term's sign in the denominator is opposite to eq.(refvacuumpol2). Since in the SM, the maximum number of $n_f = 6$, the sign becomes positive. As a result, strong coupling α_s is opposite to the QED case; strong coupling becomes smaller at higher energy, and it becomes larger at lower energy region. It means that QCD can not be applied perturbative expansion at lower energy region(at long distance). Because of this reason, quarks are confined within color singlet multiple particles. This phenomenon is called Color confinement. Oppositely, nature, which strong coupling becomes smaller at high energy, is called Asymptotic freedom. Because of this nature, single quark and single gluon can not appear, but they appear as multiple colorless particles. Such multiple color singlet particle is called Hadron, and the hadron, which is constructed by the combination of color and anti-color, is called Meson. On the other hand, the hadron, constructed by three color charged particles, is called Baryon. For example, proton and neutron are baryons. In high-energy experiments, single quark and single gluon themselves can not be measured because of color confinement, even if we want to examine quarks. However, we can observe hadrons that appear after quark and gluon decay, so quark physics should be explored through these observed hadrons that inherit quark information. Since this study's target is the quark, the review about the connection between Parton(that is, quark and gluon) and hadrons is essential. The next section explains how hadrons appear from quarks in high-energy experiments.

1.4.3 Hadronization and Jet

The previous section described that quarks and gluons do not appear lonely because of color confinement, but they appear as hadrons in nature. This section describes how we can observe hadrons that come from quark. Let us consider the case of the lepton collider, which collides leptons (typically, electron e^- and positron e^+) in the Centre-Mass(CM) frame. When e^- and e^+ collide and fly out quark pair $q\overline{q}$ by back to back like the following figure.**??**.

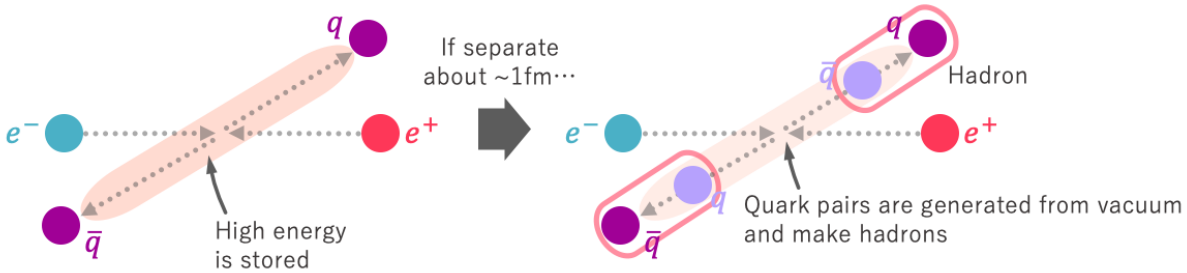


Figure 8: Conceptual figure of Hadronization

The potential energy that comes from strong interaction arises between quark pair[38]. Farther quarks separate, higher energy is stored between them. However, nature tends to go to the lower energy state and to make color singlet state. So if quarks separate about 1fm, it is more energetically stable to make hadrons by generating quark pairs from a vacuum, and this stage

[38]It looks like the spring's potential energy. In classical mechanics, the potential energy by spring is expressed by $V(x) = \frac{1}{2}kx^2$. The potential energy can be expressed by the combination of Coulomb potential and linear potential approximately: $V(r) \sim -\frac{1}{r} + r$

is called Hadronization. After hadronization, created hadrons are still energetically unstable yet, and they decay to more stable hadrons. These stable hadrons are which we can observe through detectors. In high-energy physics experiments, the generated quark pair has enormous energy (roughly half of collision CM energy). Therefore, appeared hadrons are biassed in the direction in which the original quark fly. As a result, if quarks are generated in the process, they appear as collimated sprays of stable hadrons, and it is called Jet. So we need to observe and study jets in order to explore the background physics of quark. In high-energy physics, the algorithms for the reconstruction of jets have been devised so far. However, unfortunately, we can not chase and specify which hadron belongs to which jet. More technical detail of it will be described in a later chapter.

1.5 Summary of the Standard model

We have seen the theoretical framework of QFT and the SM. This short section summarizes the essential things of the SM. At first, there are two types of important symmetries; Lorentz symmetry and Gauge symmetry. Since elementary particles appear with high energy, Lorentz symmetry is needed to construct relativistic theory. Local gauge symmetry is an important requirement to introduce interactions between matter particles. The gauge symmetry which governs the SM is

$$\mathrm{SU(3)_C} \times \mathrm{SU(2)_L} \times \mathrm{U(1)_Y}.$$

Under this gauge symmetry, Dirac fermions and three types of gauge fields are incorporated. However, they are massless because of gauge symmetry, and we need to add Higgs sector and Yukawa interactions. Based on them, the Lagrangian of the SM is given like this :

$$\begin{aligned}
\mathcal{L} = & \sum_{flavor} \overline{\Psi}_f i \not{D} \Psi_f \\
& - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} \\
& - \frac{1}{4} \mathrm{Tr} W^{\mu\nu} W_{\mu\nu} \\
& - \frac{1}{4} \mathrm{Tr} G^{\mu\nu} G_{\mu\nu} \\
& + |D_\mu \Phi|^2 - V(\Phi) \\
& + \sum_{flavor} \lambda_f \overline{\Psi}_{fL} \Phi \psi_{fR} + h.c.
\end{aligned} \tag{1.83}$$

$\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ is broken to $\mathrm{U(1)_{QED}}$ spontaneously at EW scale, and fermions and weak bosons obtain each mass.

We should know how many elementary particles exist through experiments. We have been confirmed six bosons experimentally; photon, $W^\pm$ bosons, Z boson, gluon, Higgs boson, as theory predicted. However, since only the Higgs boson is a scalar boson, it has spin 0, but other bosons have spin-1. Weak bosons are massive particles, but other gauge bosons are massless. For fermions, there are two classes; Quark and Lepton. The difference between them is whether they feel strong interaction or not. Lepton is a class of particles that do not have color charges. On the other hand, quarks have three color charges; red, blue, or green. Since

leptons and quarks satisfy Dirac Lagrangian, they have spin-1/2. We have discovered six quarks and six leptons, and they have the structure called Generation. Figure.9 is the summarized figure of the SM particles :

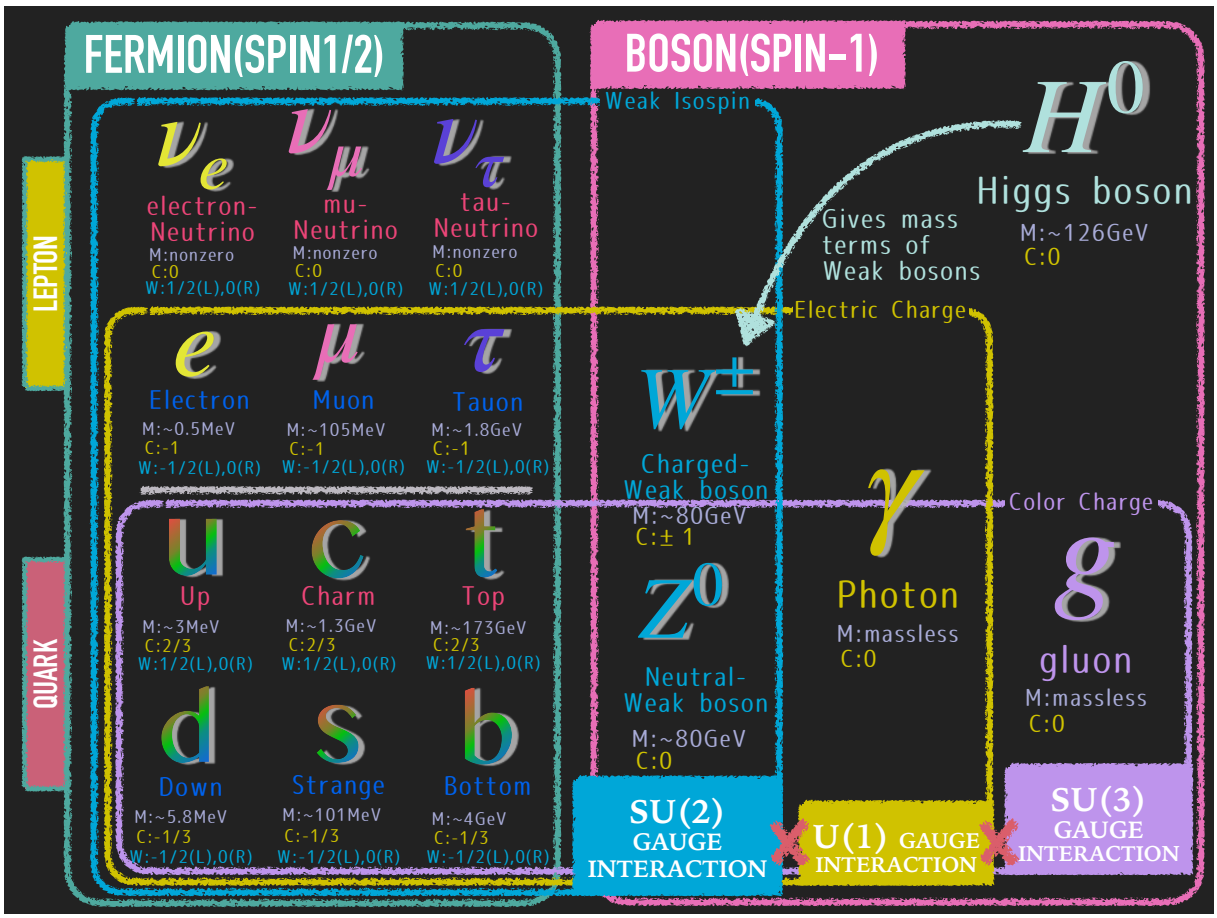


Figure 9: The SM particles and parameters : Each capital below particle names means parameters; M is mass, C is electric charge, W is weak isospin.

Is the SM the ultimate theory that can explain all phenomena completely? It is No. Indeed, the SM is a successful theory, and it can explain numerous experiments with fantastic precision. However, there are many fundamental problems that the SM can not explain. Why is the generation of quark and lepton 3? Why the representations of the Lorentz group which appear in the SM are only scalar, spinor, vector? How should we incorporate gravity into the SM? Humanity has just understood this universe. This study is a challenge to approach the mystery of Quark mass, which the SM can not solve.

2 Study of Running bottom quark mass

2.1 Quark mass definition

As we have seen in the chapter1, quarks and gluon are interacted by strong interaction with each other, and they have each color charge. So if the particle has some color, it feels strong interaction. However, if the particle has the color $\mathrm{R}+\mathrm{G}+\mathrm{B}$ or the pair of the color and its anti-color(e.g. $\mathrm{R}\overline{\mathrm{R}}$), it becomes "neutral" for strong interaction. Such a state is called the Color singlet. The particles which exist in nature are interacted and clustered by strong interaction, and when the clustered particle becomes a color singlet state, it no longer feels strong interaction. Because of this reason, particles that have color do not appear in nature, but they appear as colorless particles. Based on such a thing, how should we define the mass of "single" quark? Since a single quark can not exist in nature and many gluons appear around the quark so that the whole color becomes neutral, it is not trivial to define quark mass, unlike lepton mass. This section introduces how to define quark mass.

2.1.1 Pole mass

By convention, there are two typical definitions[39] in high-energy physics. At first, let us consider the quark propagator, which includes loop corrections. As we have seen in the previous chapter, we need to use the renormalization prescription to deal with loop divergence. When we consider loop corrections to the 2-point Green function, all loop effect is pushed into Self-energy $\Sigma(p)$, and quark propagator should be finite thanks to renormalization condition of $\Sigma(p)$. So renormalized quark propagator is given like this :

$$\Delta(p) = \frac{i}{\not{p} - M_q - \Sigma(p)} \tag{2.1}$$

M_q is regarded as the physical quark mass. So it can be used as a quark mass definition, and it is called Pole mass. It is one of the ways to define physical quark mass. Additionally, this definition is scheme and gauge independent. However, there is a problem; this renormalized propagator is sensitive to the non-perturbative effect. Perturbative QCD can not converge well at low energy region, and the theoretical uncertainties of Λ_{QCD} $\sim$300MeV that can not remove in principle appear in quark self-energy. Therefore, the pole mass receives such a non-perturbative effect, which cannot be avoided in principle.

2.1.2 $\overline{\mathrm{MS}}$ running mass

Let us move to another definition; running mass. We have already known about running coupling for renormalization theory[40]. The mass parameter in Lagrangian is called "bare" mass, which includes infinity, and physical mass(renormalized mass), which we can observe is obtained through the prescription of renormalization. This renormalized mass depends on the energy scale where we observe, and we call it Running mass. Of course, running mass appears through renormalization, and it depends on renormalization schemes. In this study, we focus

[39] You can check other definitions at [13].
[40] You can check it in section1.3.

on $\overline{\text{MS}}$ scheme.

The straightforward understanding of running mass is like this: the quark obtains its mass from the Yukawa interaction between the quark itself and Higgs boson. However, since a single quark can not exist in nature, many gluons appear around the quark. Higgs boson couples to the cluster of the quark and gluons(See Figure.10), and running mass is the mass that includes the energy of these gluons into quark mass. As described in section1.4.2, strong interaction between the quark and gluons has energy dependence. Namely, the number of gluons that contribute to the quark changes. As a result, quark mass also has energy dependence.

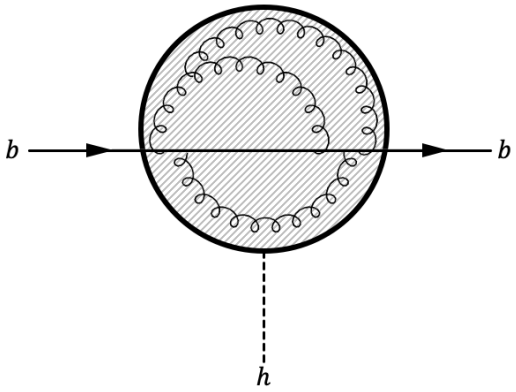


Figure 10: Intuitive understanding of running quark mass : A single quark cannot appear, and the gluon cloud appears around of quark so that the whole color be a color singlet. Higgs boson couples to this "effective" vertex, which includes the gluon cloud, and observed quark mass changes according to gluons' effect. Since each strong vertex between gluon and quark runs, quark mass also has energy dependence. In this sense, running quark mass will be more complicated than running of coupling constant.

Running quark mass(for all quark flavors) is governed by the Renormalized Group Equation(RGE) for QCD. RGE is derived under the requirement which considering physics should not change by taking different renormalization schemes. For example, let us consider the dimensionless observable[41] which depends on the process energy Q, the renormalization scheme μ, the strong coupling $\alpha_s(\mu)$ and the quark mass $m_q(\mu)$; $R\left(\frac{Q^2}{\mu^2}, \alpha_s(\mu), \frac{m_q(\mu)}{Q}\right)$. If we require that the observable R must not be changed by exchange of renormalization scheme $\mu \to \mu'$[42], R should be satisfied the following total derivative

$$\mu^2 \frac{d}{d\mu^2} R\left(\frac{Q^2}{\mu^2}, \alpha_s(\mu), \frac{m_q(\mu)}{Q}\right) = \left(\mu^2 \frac{\partial}{\partial \mu^2} + \mu^2 \frac{\partial \alpha_s(\mu)}{\partial \mu^2} \frac{\partial}{\partial \alpha_s(\mu)} + \mu^2 \frac{\partial m_q(\mu)}{\partial \mu^2} \frac{\partial}{\partial m_q(\mu)}\right) = 0.$$

Here, the factor of μ^2 is added in front of the equation following the convention. The second and third coefficients can be replaced with dimensionless functions like this :

$$\mu^2 \frac{\partial \alpha_s(\mu)}{\partial \mu^2} = \beta(\alpha_s(\mu)) \tag{2.2}$$

$$\mu^2 \frac{\partial m_q(\mu)}{\partial \mu^2} = -\gamma(\alpha_s(\mu)) m_q(\mu). \tag{2.3}$$

[41] What we can observe typically is cross-section or decay rate of the considering process, so thinking the dimensionless observable is reasonable.

[42] The semigroup of such transformations is called Renormalized Group. The semigroup is defined as the set which has associativity for a considering binary operation. So semigroup is not a group, but conventionally, it is called renormalized "group".

Based on these definitions, the above total derivative is rewritten by

$$\left(\mu^2\frac{\partial}{\partial\mu^2}+\beta(\alpha_s(\mu))\frac{\partial}{\partial\alpha_s(\mu)}+\gamma(\alpha_s(\mu))m_q(\mu)\frac{\partial}{\partial m_q(\mu)}\right)=0.$$

As we can see eq.(2.2) and eq.(2.3), these differential equations give the energy dependence of each parameter, and they are called Renormalized Group Equations(RGE). The RGE, which explains quark mass running, can be perturbatively expanded like this[35] [43] :

$$\mu^2\frac{dm_q(\mu)}{d\mu^2}=-\gamma(\alpha_s)m_q(\mu)=-\sum_{i=0}^{\infty}\gamma_i\left(\frac{\alpha_s(\mu)}{\pi}\right)^{i+1}m_q(\mu).$$

This γ is calculated up to the fourth-order of α_s, and they are given by

$$\begin{aligned}
\gamma_0&=1\\
\gamma_1&=\frac{1}{16}\left(\frac{202}{3}-\frac{20}{9}n_q\right)\\
\gamma_2&=\frac{1}{64}\left[1249-\left(\frac{2216}{27}+\frac{160}{3}\zeta(3)\right)n_q-\frac{140}{81}n_q^2\right]\\
\gamma_3&=\frac{1}{256}\Bigg[\frac{4603055}{162}+\frac{135680}{27}\zeta(3)-8800\zeta(5)\\
&-\left(\frac{91273}{27}+\frac{34192}{9}-880\zeta(4)-\frac{18400}{9}\zeta(5)\right)n_q\\
&+\left(\frac{5242}{243}+\frac{800}{9}\zeta(3)-\frac{160}{3}\zeta(4)\right)n_q^2-\left(\frac{332}{243}-\frac{64}{27}\zeta(3)\right)n_q^2\Bigg].
\end{aligned}$$

Here, n_q is the number of active quark flavors which these masses are $m_q<\mu$, and $\zeta(3)\sim 1.202057$, $\zeta(4)\sim 1.082323$, $\zeta(5)\sim 1.036928$ are Riemann Zeta functions. To solve this RGE of quark mass eq.(2.3), we should use RGE of strong coupling :

$$\mu^2\frac{\partial\alpha_s(\mu)}{\partial\mu^2}=\beta(\alpha_s(\mu))=-\sum_{i\geq 0}\beta_i\frac{\alpha_s^{i+2}(\mu)}{\pi^{i+1}}$$

$\beta(\alpha_s)$ is a beta function and it can be expanded by β_i similar to above $\gamma(\alpha_s)$. Up to the fourth-order of α_s, it is calculated by

$$\begin{aligned}
\beta_0&=\frac{1}{4}\left(11-\frac{2}{3}n_q\right)\\
\beta_1&=\frac{1}{16}\left(102-\frac{38}{3}n_q\right)\\
\beta_2&=\frac{1}{64}\left(\frac{2857}{2}-\frac{5033}{18}n_q+\frac{325}{54}n_q^2\right)\\
\beta_3&=\frac{1}{256}\Bigg[\frac{149753}{6}+3564\zeta(3)-\left(\frac{1078361}{162}+\frac{6508}{27}\zeta(3)\right)n_q\\
&+\left(\frac{50065}{162}+\frac{6472}{81}\zeta(3)\right)n_q^2+\frac{1093}{729}n_q^3\Bigg].
\end{aligned}$$

[43] In this paper and reference[35], $\overline{\text{MS}}$ scheme is used as renormalization scheme.

Based on these perturbative calculation of $\beta(\alpha_s)$ and $\gamma(\alpha_s)$, we can conclude the solution of eq.(2.3). The form itself of the solution can be derived easily :

$$m_q(Q) = \exp\left[-\int_{\alpha_s(\mu^2)}^{\alpha_s(Q^2)} d\alpha_s(\mu)\frac{\gamma(\alpha_s(\mu))}{\beta(\alpha_s(\mu))}\right] \times m_q(\mu) \tag{2.4}$$

The integration variable μ^2 is transformed to α_s by using eq.(2.2) here. The integrand $\gamma(\alpha_s(\mu))/\beta(\alpha_s(\mu))$ can be regarded as some polynomial with respect to α_s as we can check above, and it is positive[44]. Therefore, if we think μ is fixed on some energy scale, quark mass m_q at a higher energy scale Q will be smaller. It should be noted again that running mass depends on the process energy scale and the renormalization scheme. Running quark mass measurement by various energy scale processes becomes the verification of QCD theory. This point connects to the motivation of this study. Additionally, we can expect a useful discussion about Beyond the Standard Models(BSM) because it can be a probe of BSM if we find the deviation between measured quark mass evolution and the SM expectation. The next section explains its detail.

2.2 Mystery of Quark mass

In the chapter1, we have seen that six flavor quarks exist, and these masses are given by the Yukawa interaction between the quark and Higgs boson through EW symmetry breaking. Let us see each quark mass numbers which are measured experimentally so far :

Quark flavor	Up	Down	Strange	Charm	Bottom	Top
Mass	$2.16^{+0.49}_{-0.26}$MeV	$4.67^{+0.48}_{-0.17}$MeV	93^{+11}_{-5}MeV	1.27 ± 0.02GeV	$4.18^{+0.03}_{-0.02}$GeV	173.1 ± 0.9GeV

Table 4: Experimental Quark mass Numbers (2019 revised)[5] : u,d,s masses are masses at $\mu \sim 2$GeV. c,b masses are masses at each mass scale about 1 to 5GeV. t mass is given as its pole mass.

We can see that each quark mass is different, and there is a huge scale gap of $\sim 10^4$ between the lightest one(Up quark) and the heaviest one(Top quark). The SM can describe how each quark obtains its mass, but on the other hand, why does each quark mass differ and why there is such a huge disparity between obtained quark masses are mysteries that the SM can not explain.

However, as explained in the previous section, we can consider the energy dependence of quark mass by using the definition of running mass. So, each quark mass changes from the above-measured number at a higher energy scale.

Furthermore, if some new effects contribute to the quark, the energy dependence of quark mass will deviate from the SM's expectation. So each quark mass does not match at a higher energy scale in the SM, but if each mass is unified at a higher energy scale thanks to the effect of new physics, it can explain the above mass problems' origin.

Based on this idea, many new physics models have been proposed so far. One interesting idea is the theory that considers larger gauge group and unifies color symmetry $SU(3)_C$ and EW symmetry $SU(2)_L \times U(1)_Y$. Such a theory is called Grand Unified Theory(GUT). Some

[44] We can see it by substituting 6 to the number of active quarks n_q.

models include Super Symmetry(SUSY)[45], and predict mass unification of third-generation particles(for example, b quark and t quark, and τ lepton) at GUT scale($\sim 10^{16}$GeV)[6] [7]. Additionally, the GUT is expected to solve other critical mysteries; how is right-handed neutrino incorporated to describe the non-zero mass of neutrinos? is proton stable or not? etc... Because of such reasons, the GUT is an attractive future goal as new physics. However, any firm evidence of the existence of GUT has not been found.

Since $b-\tau$ mass unification at the GUT scale is predicted in some GUT models, the study of the energy dependence of b quark mass is a useful key to discuss the viability of GUT models. Since $b-\tau$ mass unification at the GUT scale is predicted in some GUT models, the study of the energy dependence of b quark mass is a useful key to discuss the viability of GUT models. According to RGE, quark mass runs to a higher energy scale, and if the measured behavior deviates from the SM expectation, it is a probe of new physics. Some GUT models are imposed Super Symmetry(SUSY), and new SUSY partner particles such as gaugino are predicted in these theories. If these new SUSY particles appear around b quark propagator at the higher energy scale in addition to gluon clouds(remember figure.10), the experimental result of running b quark mass deviates from the SM prediction. It is the usefulness and importance of running b mass study. Then, how is running b mass value decided at actual experiments? We have one precedent which is used at LEP(Large Electron-Positron collider) experiment. It provided the result of b mass at Z-pole energy scale, which is Z boson mass scale($\sim 91\text{GeV}/c^2$). How did they measure it? The next section shows it.

2.3 The measurement of running bottom mass

We have seen how to define quark mass and its importance and interesting point of running b mass study. Then, how to measure b quark mass at high energy is explained in this section. As a previous study, LEP experiments measured b quark mass at the Z-pole energy scale using the jet ratio as the observable. Additionally, the next challenge of b quark mass study based on LEP results is also discussed in this section.

2.3.1 The bottom mass measurement using jet

In the previous section, the importance and interesting point of running b quark mass study are explained. Here, how to measure the running b quark mass at actual experiments is explained. As described in section 1.4.2 and 1.4.3, single quark and gluon can not appear in nature, and they appear as the collimated spray of stable hadrons, which is called Jet. Interesting physics relate to QCD is appeared at Parton(quark and gluon) level[46], and jets inherit this information. So, we can verify QCD theory and explore new physics beyond QCD through observations of jets. In this study, we are interested in the mass of b quark, but of course, we can not observe it directly. We can use the following thing to measure it; gluon radiation from quark has quark mass sensitivity. The cross-section of gluon radiation from

[45] Super Symmetry is the symmetry between matter particles and force particles. Partner particles of SM fermions and bosons appear.

[46] It is because perturbative QCD is valid at a higher energy scale. Hadrons, which appear after hadronization from Parton, are objects at a low energy scale, and perturbative QCD can not predict its physics well.

quark is given as like this[11] :

$$\frac{1}{\sigma_0}\frac{d^2\sigma}{dx_1 dx_2} = C_F\frac{\alpha_s}{2\pi}\left[\frac{x_1^2+x_2^2}{(1-x_1)(1-x_2)} - \frac{4m_q^2}{s}\left(\frac{1}{1-x_1}+\frac{1}{1-x_2}\right)\right.$$
$$\left. - \frac{2m_q^2}{s}\left(\frac{1}{(1-x_1)^2}+\frac{1}{(1-x_2)^2}\right) - \frac{4m_q^4}{s^2}\left(\frac{1}{1-x_1}+\frac{1}{1-x_2}\right)^2\right] \tag{2.5}$$

According to this equation, we can see that heavier quark tends to emit fewer gluons than lighter quarks. This phenomenon looks like bremsstrahlung. For QED, charged particle emits photons according to the cross-section $\sigma \propto m^{-4}$. So, the lighter charged particle tends to emit more photons and lose more energy than the heavier particle case. So, gluon radiation can be interpreted as the QCD version of bremsstrahlung. Based on this phenomenon, the bottom quark mass can be decided experimentally by counting how often gluons are emitted from the bottom quark. So in this study, the signal process that we focus on is b quark pair production, including gluon radiation. In accelerator experiments, these b quarks and gluons occur hadronization. They decay to jets, and we can observe them as we have seen at 1.4.3. Therefore, we should count jet events instead of Parton(quarks and gluons). This study focuses on 3-jet events that originate from $e^+e^- \to b\bar{b}g$.

When we measure some quantity experimentally, it is important to define the observable, which has good sensitivity on the target quantity. How sensitive gluon radiation process to b quark mass? If we consider the total cross-section of $e^+e^- \to hadrons$(See figure.11), its b mass effect is $\sigma\left(e^+e^- \to hadrons\right) \sim m_q^2/s \sim 0.3\%$ at Z-pole, but it is too small and negligible. Instead of the inclusive quantity such as total cross-section, we can define a more exclusive quantity that extracts the ratio of 3-jet events(corresponds to the second diagram of figure.11[47]) from whole jet events. If we use such ratio as the observable, an additional parameter y_c that defines 3-jet event experimentally appears in the observable. Since it is set as $\ll 1$ typically, the b mass effect is enhanced by a factor of 10; $\sim m_q^2/(s \cdot y_c) \sim 3\%$. Therefore, such an exclusive quantity is appropriate to the observable for b mass measurement.

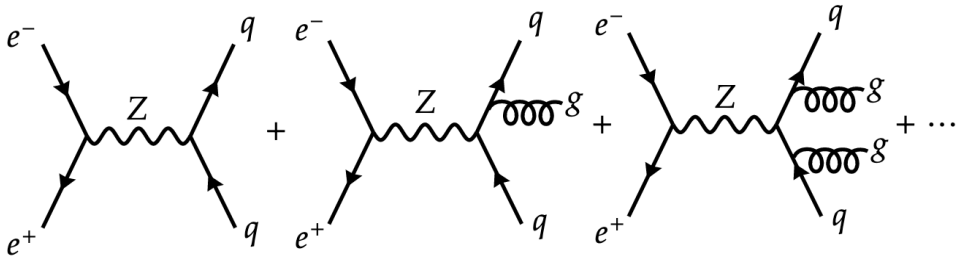


Figure 11: Contributions to $e^+e^- \to hadrons$: All diagrams drawn here are tree-level diagrams, but other processes also exist, and they include virtual gluon loops.

Actually, at LEP experiments[15], the following observable was defined, and b quark mass

[47]To be precise, a three-parton diagram that emits soft gluon is absorbed into a two-parton diagram to make IR and collinear safe observable. At NLO(Next-to-Leading Order), the decay width of 2-jet events(that is, two-parton) Γ_{2j} includes IR and collinear divergences. However, they are canceled completely with the divergence in the decay width of 3-jet events Γ_{3j} at LO(Leading Order)[8]. So if we divide Γ_{3j} into 2-jet region(has divergence by soft gluon radiation) and 3-jet region(safe), and combine this 2-jet region and Γ_{2j}(it also has divergences), redefined decay width Γ_{2j}(absorbed divergences) and Γ_{3j}(divergences are separated) become IR and collinear safe. The criterion that divides the 2-jet region and 3-jet ratio is y_c, which appears in the main text. Similarly, to make Γ_{3j} at NLO, we need to combine the hard part of three-parton processes and the soft part of four-parton processes.

at Z-pole was decided through its measurement[16][17] :

$$R_3^{bl} \equiv \frac{\Gamma_{3j}^b\left(y_c\right)/\Gamma^b}{\Gamma_{3j}^l\left(y_c\right)/\Gamma^l} = \frac{R_3^b}{R_3^l} \quad \left(R_3^q \equiv \frac{\Gamma_{3j}^q\left(y_c\right)}{\Gamma^q}\right). \tag{2.6}$$

We can see 4 quantities$\Gamma_{3j}^b\left(y_c\right)$, $\Gamma_{3j}^b\left(y_c\right)$, Γ^b, Γ^l in observable R_3^{bl}. $\Gamma_{3j}^b\left(y_c\right)$ is the decay rate of e^+e^- →3 b-jet events and Γ^b is the decay rate of all b-jet events(2-jet, 3-jet...). Indices b and l indicate that quark flavors originate each jet, and l means that light quark(here, we are considering u,d,s) cases. Why do we take into account c quark in light quarks? There are 2 reasons. First, since c quark mass is close to b quark mass, if we include c quark into the denominator Γ_{3j}^l/Γ^l, b mass sensitivity becomes low. Second, c quark identification is more difficult than b quark because c hadron flies shorter than b hadron. It is a technical reason, and it can be seen in later sections again. Because of such reasons, we did not include c quark here.

y_c means the size of reconstructed jets experimentally and defines 3-jet events as described. More detail of this parameter can be seen in chapter4. When we distinguish 3-jet events, we apply the criteria of y_c. Because of this reason, 3-jet decay rate Γ_{3j}^q depends on y_c.

Why do we need to take the fraction between 3-jet events and all jet events, take the fraction between b-jet events and light-jet events? There are theoretical and experimental advantages. The fraction $\Gamma_{3j}^q(y_c)/\Gamma$ cancels electroweak corrections(for example, electroweak loop correction at $Zq\overline{q}$ vertex) and extracts QCD mass effect. Whereas the fraction between b-jet events and light-jet events cancels corrections that come from hadronization and detector partially[48].

R_3^{bl} measurement can provide independent and more directly b quark mass measurement at the higher energy scale far from the threshold scale. It is an advantage of this method which is used R_3^{bl}. The next section shows the result obtained at LEP using the observable R_3^{bl} and discusses what the next challenge is based on it.

2.3.2 The result at LEP and importance of measurement at 250GeV

The running bottom quark mass at Z-pole(∼90GeV) was measured by using the observable R_3^{bl} at LEP and SLD. Figure.2.9 shows the results of LEP and SLD. This result is the bottom quark mass measurement at the highest-energy that has been done up to now. According to this result, it turns out that measured results are consistent with the current QCD theory. LEP measurement's pioneering point is that it can provide a direct measurement of b-mass on a high energy scale far from the threshold scale∼ 10GeV, independent of the b-mass at the threshold. However, there was no indication of new physics in these results. As the next challenge, progression to a higher energy scale above Z-pole can be expected to provide some probe of new physics.

[48] You can see this efficacy in estimation of systematic errors of chapter5.

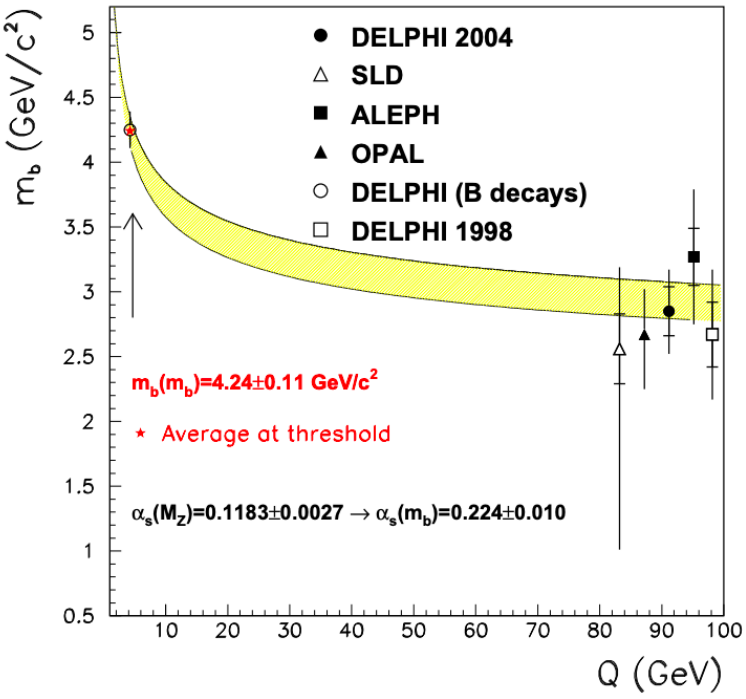


Figure 12: Result of b quark mass at LEP[18] : The horizontal axis indicates the energy scale of the considering physical process. The yellow band shows the theoretical uncertainties of RGE, which evolves $m_b\,(m_b)$ to Z-pole, and we can see that obtained results agreed with SM prediction. All five data around Z-pole were obtained at the same scale, but for display reasons, each data point is shifted to a different energy scale. The most left data point was the result which was obtained from the semi-leptonic B decays study at the DELPHI experiment.[19].

In this study, we are focusing on 250GeV ILC(International Linear Collider) as a next challenge. 250GeV is the next achievable energy scale, and we can make a strategic move for GUT scale physics by 250GeV b mass study. First, verification of SM(QCD theory) at 250GeV energy scale can be done. The b mass at 250GeV becomes an input parameter of BSM predictions of the b mass running at a higher energy scale above 250GeV. It will be a reference for discussion of the viability of GUT models. For such reasons, 250GeV b mass measurement is an important work for verifying SM and new theory building.

Some people may think like this; if we want to see some new physics effect on b mass running, we should go to a higher energy scale such as the TeV scale. However, unfortunately, it is not a clever way. It is because b quark mass can be regarded as massless at higher energy. For example, b mass(several GeV scale) can be regarded as massless under several TeV scale systems. So it will be more difficult to observe b mass effect at high energy in principle. We can see it from the explicit formula of R_3^{bl} as the function of b mass. NLO calculation of R_3^{bl} is given by like this[12] [14] :

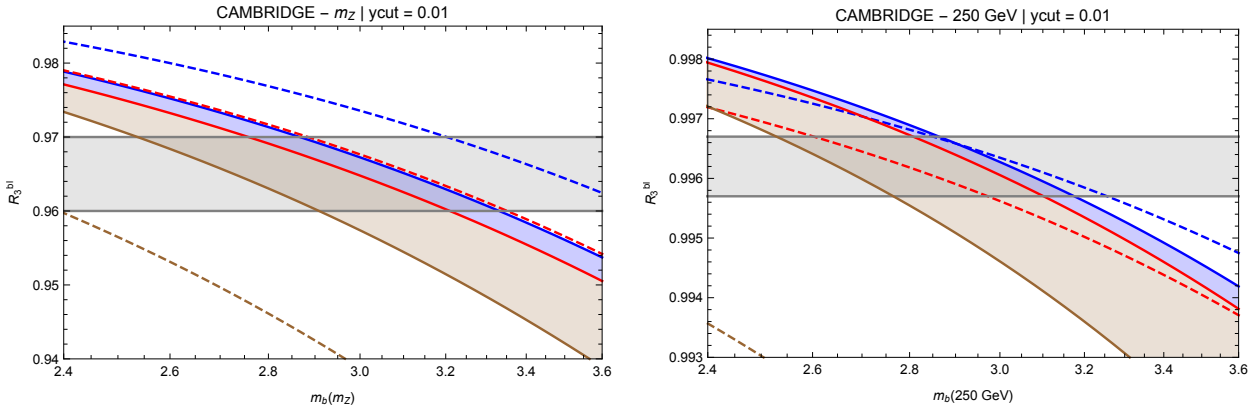
$$R_3^{bl} = 1 + \frac{\alpha_s(\mu)}{\pi} a_0\,(y_c) + \frac{m_b^2(\mu)}{s}\left\{ b_0 + \frac{\alpha_s(\mu)}{\pi}\left(b_1 + 2b_0\left(\frac{4}{3} - \log\frac{m_b^2(\mu)}{s} + \log\frac{\mu^2}{s} \right) \right) \right\}. \quad (2.7)$$

a_0 is a massless correction and it is very small($a_0(0.01) \sim 0.04$). b_0 and b_1 are LO and NLO mass corrections. Massive correction b_0 is negative[16], but it reflects that b quark is heavier than uds quarks. So, R_3^{bl} tends to close to one for smaller b quark mass. The sensitivity of b

mass on R_3^{bl} is calculated from this formula by simple diffferentiation for m_b :

$$\Delta R_3^{bl} \sim \frac{2\left(1-R_3^{bl}\right)}{m_b\left(\mu\right)}\Delta m_b(\mu). \tag{2.8}$$

The figure.13 shows that theoretical calculations of b mass sensitivity of observable R_3^{bl} at Z-pole and the next challenging energy scale 250GeV according to eq.(2.8). Let us estimate the necessary accuracy of R_3^{bl} measurement under the b mass precision of 0.4GeV for both Z-pole and 250GeV. At the Z-pole energy scale, if the center value of R_3^{bl} is assumed to be 0.965, and the center value of b mass is assumed to be 2.97GeV, the necessary accuracy of observable ΔR_3^{bl} should be ∼0.01. It means that if we want to determine b mass at Z-pole with the error of 0.4GeV, it corresponds to the measurement of R_3^{bl} at the precision of ∼1.0%. On the other hand, at 250GeV energy scale, if the center value of R_3^{bl} is assumed to be 0.996 and the center value of b mass is assumed to be 2.75GeV, ΔR_3^{bl} should be ∼0.001. It means that if we want to determine b mass at 250GeV with the same precision 0.4GeV, it corresponds to the measurement of R_3^{bl} at the precision of 0.1%, that is per mile level. Namely, if we want to see b quark mass at 250GeV, more accurate measurement of observable R_3^{bl} is needed.



(a) at $\sqrt{s}$ =Z-pole ($y_c = 0.01$) **(b)** at $\sqrt{s}$ =250GeV ($y_c = 0.01$)

Figure 13: Sensitivity of b quark mass on R_3^{bl} in the Cambridge jet-clustering algorithm : Brown lines are R_3^{bl} in terms of pole mass under $\mu = \sqrt{s}$. Red and blue lines are R_3^{bl} in terms of running mass by using the renormalized scale $\mu = \sqrt{s}/2$GeV for red and $\mu = 2\sqrt{s}$ for blue. Solid lines are NLO calculations, and dashed lines are LO calculations. It is reprinted from [14] with permission of the author.

The number of b quark mass of DELPHI Collaboration at LEP[18] was

$$m_b(M_Z) = 2.85 \pm 0.18(stat.) \pm 0.13(exp.) \pm 0.19(had.) \pm 0.12(theo.)\text{GeV}/c^2. \tag{2.9}$$

The first error means the statistical error, and the second and third ones are systematic errors which come from detector effect or hadronization model [49]. The fourth error comes from theoretical uncertainties. There are several sources; first source: renormalized scale μ (you can check it from spreads between each line of figure.13), second source: strong coupling $\alpha_s(\mu)$ (propagates to b mass through eq.(2.4)), third source: Mass ambiguity (Uncertainty which appears between mass definitions[50]).

[49] R_3^{bl} can cancel hadronization or detector uncertainties partially, but not completely. Remained uncertainties appear here.

[50] Of course, each mass definition, for example, pole mass and running mass should be equivalent. However, if we calculate running mass up to some finite order and compare it to pole mass, they are no longer equivalent, and there is a gap. This uncertainty reflects such a gap between different mass definitions. LO calculation depends on mass definition larger.

This study estimates the measurement precision of b mass at 250GeV ILC through its simulation. To take action, we need to know about the ILC. The next chapter explains the outline of ILC and its detector, the ILD detector.

3 Experimental Apparatus

As discussed in the previous section, this study focuses on the ILC and estimates b quark mass measurement precision at 250GeV. We need to grasp the overall picture of ILC experiments and ILD to do this study. This chapter explains the outline of the ILC accelerator and ILD detector.

3.1 ILC accelerator

3.1.1 Outline

ILC is a future linear collider. Why do we need a linear collider? Originally, observing the particles falling from the universe was the mainstream experiment method of elementary particle physics. However, more heavy new particles discovered by using accelerators, which can reproduce a high energy state. With the development of accelerator science, collision energy and luminosity have been improved so far. Luminosity is a parameter that means the accelerator's performance, and it depends on the beam size, density of particles in a bunch, and the frequency of collision. Furthermore, two types of accelerators have been developed in high-energy physics; Circular collider and Linear collider. For example, one of the famous circular colliders is LHC(Large Hadron Collider) in CERN. It collides protons by the center-mass energy of about 13TeV, and it is the highest record. So LHC collider is often called the energy frontier. One important character of the circular collider is that there is Synchrotron radiation. When the track of charged particle is bent by the magnetic field, it emits photons and loses energy. The energy of this radiation depends on the negative quartic of the mass of charged particle: $E \propto R^{-1} \cdot m^{-4}$ where (R is the radius of circular collider). So lighter charged particle emits more photons and loses energy, but heavier particle can be accelerated without large energy loss[51]. Based on such reason, circular colliders are suitable for heavier particle beams, for example, proton. If we use electron beams in a circular collider, we can accelerate it, but we can not expect the collision at higher energy such as LHC case[52].

LHC boasts the energy frontier, but it suffers from background contaminations. Since LHC uses proton, Parton(quarks and gluons) interact with each other when colliding them. As a result, obtained events become so complicated, and its analysis is laborious work. If we use elementary particles as beams, it provides more clean events. Instead of this advantage, we will suffer from synchrotron radiation, so how should we do...? The answer is taking $R \to \infty$, that is "line". Linear lepton collider free us from the above problems; radiation and a large number of background contaminations. Some linear colliders have been proposed so far, and the International Linear Collider(ILC) is the collider closest to realization. Figure.14 is the conceptual figure of ILC.

[51] It is similar with the physics of gluon radiation eq.(2.5).

[52] Super KEKB is also a circular collider in KEK, but it uses electron and positron beams. It is not the energy frontier, but it has the highest luminosity of world record.

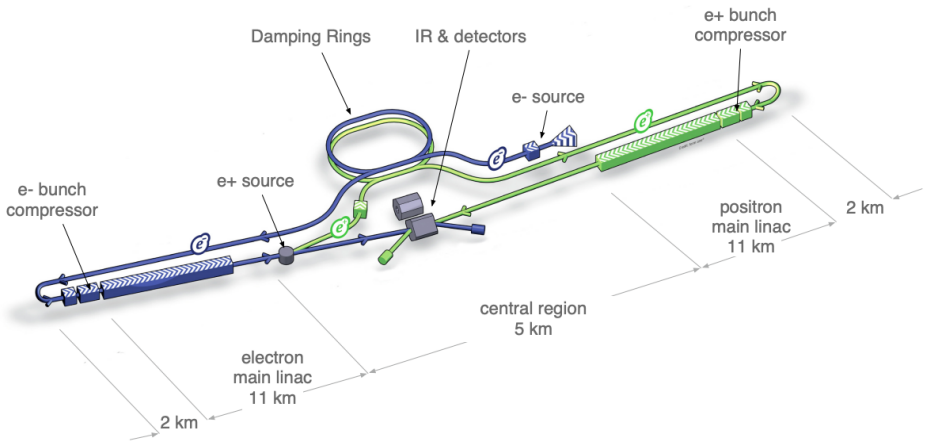


Figure 14: Conceptual figure of ILC accelerator [20]

e^- beam is generated by e^- source, and it is accelerated to 5GeV by normal and superconducting linac. Accelerated e^-s are transferred to e^+ source. In e^+ source, there are undulators. In this equipment, photons are emitted from e^- by bremsstrahlung radiation and create e^+ from photons by hitting titanium alloy. Both beams are accelerated and converged at dumping rings, and after that, they move on the main linuc. In the main linuc, SCRF(Super Conducting Radio Frequency) cavities, which have an average gradient of 31.5MV/m, are lined up, and beams are accelerated until goal energy. The first stage of ILC is designed as 250GeV ILC, which is called Higgs Factory. Moreover, ILC can be extended to 500GeV, and 1TeV. Some important parameters of ILC are like this[53] :

CM energy [GeV]	250(baseline)	500	1000(baseline)
Luminosity $[\times 10^{34} cm^{-2} s^{-1}]$	1.35	1.8	3.6
Electron Polarization [%]	80	80	80
Positron Polarization [%]	30	30	20

Table 5: Parameters of the ILC

It is the outline of ILC. ILC is not an energy frontier, but it can realize the precise measurement of particles such as the Higgs boson discovered at LHC in 2012. The self-coupling of Higgs boson and Yukawa couplings is an essential target for the ILC study program. Such precise measurements can be done at ILC, but not LHC. So ILC and LHC will be complementary experiments; LHC is good at exploring new particles by energy frontier, and ILC is good at the precise measurement of discovered particles. Of course, ILC also can be expected to discover SUSY partners or Dark Matters.

As described in the previous chapter, it is not easy to measure b quark mass at high energy, but if the observable eq.(2.6) can be measured at per-mile, we can expect a competitive result with LEP's one. It is a motivation to use ILC in this study. The next section introduces that beam polarization, which is a unique aspect of ILC. This configuration is used in this study, so let us see about it in the next section.

[53] You can check more detail in [20] [21].

3.1.2 e^-/e^+ beam and Polarization

ILC collides e^- and e^+ beams, but the beam structure of ILC is devised well to yield data efficiently. Electrons(or positrons) of 2.0×10^{10} per a bunch cross at the interaction point and 1312 bunches are lined up by the interval of 554nsec. Moreover, this train collides by the interval of 200msec(∼5Hz).

Beam polarization is one of the important and unique features of the ILC. Thanks to beam polarization, we can suppress background events and improve signal event creation. How does ILC make polarized beams? As explained in the previous section, e^- beam is generated at e^- source. In this e^- source, obtain polarized e^- beam by the photoelectric effect when irradiating polarized photon beam to GaAs photocathode. As the table.5 shows, 250GeV ILC aims to make 80% polarized e^- beam and 30% polarized e^+ beam.

Under such an environment, ILC can collect individual data sets by different polarization configurations. It is one of the advantages of ILC. Sharing total luminosity between different polarization data sets provides improved results rather than individual polarization configuration. Some combinations of shared luminosity for each CM energy has been proposed so far :

	fraction with sgn$(P(e^-), P(e^+)) =$			
	(-,+)	(+,-)	(-,-)	(+,+)
$\sqrt{s}$	[%]	[%]	[%]	[%]
250 GeV (2015)	67.5	22.5	5	5
250 GeV (update)	**45**	**45**	5	5
350 GeV	67.5	22.5	5	5
500 GeV	40	40	10	10

	int. luminosity with sgn$(P(e^-), P(e^+)) =$			
	(-,+)	(+,-)	(-,-)	(+,+)
$\sqrt{s}$	[fb^{-1}]	[fb^{-1}]	[fb^{-1}]	[fb^{-1}]
250 GeV (2015)	1350	450	100	100
250 GeV (update)	**900**	**900**	100	100
350 GeV	135	45	10	10
500 GeV	1600	1600	400	400

Figure 15: Proposed integrated luminosities of ILC for each scenario [21]

Here, $|P(e^-)|$ means that the polarization rate of e^- beam, and $|P(e^+)|$ is the polarization rate of e^+ beam. The polarization rate is defined as like this:

$$P(e^{\mp}) = \frac{f_R - f_L}{f_R + f_L}. \tag{3.1}$$

Here, f_L is the proportion of left handed components in a bunch, and f_R is the proportion of right handed components in a bunch. The scenarios in the table.15 called H20 scenario. There are 4 combinations of polarization configurations : $(-80\%, +30\%)$, $(+80\%, -30\%)$, $(-80\%, -30\%)$, $(+80\%, +30\%)$ and they are expressed by using signs + or - in table.15. For example, $P(e^-) = +80\%$ means that the electron beam which includes left components of 10% and right components of 90%. $P(e^-) = -80\%$ means vice versa; the electron beam which

includes left components of 90% and right components of 10%. The upper table shows the breakdown of each shared luminosity scenario, and the bottom table shows total luminosities for each polarization configuration. Sample files which this study used are prepared for each pure polarization; 100% left polarization or 100% right polarization with the luminosity of 250fb^{-1}. Here, 100% left polarization means that electrons are 100% left polarized and positrons are 100% right polarized. So, it is often written as $e^-_L e^+_R$. On the other hand, 100% right polarization means that electrons are 100% right polarized and positrons are 100% left polarized, and it is often written as $e^-_R e^+_L$. However, in actual scenarios, such pure polarizations can not be realized. So this study estimates the precision of b quark mass under the above H20 scenario by mixing obtained results from pure polarized samples and extrapolation 250fb^{-1} to 900fb^{-1}. This section explained the points of ILC, which relate to this study. However, the core point of what we want to do is a simulation of b running mass measurement at ILD(International Large Detector). So we need to know the ILD detector too. Let us see it in the next section.

3.2 ILD detector

3.2.1 Outline

So far, two detectors have been proposed for ILC; ILD and SiD. Why we need two detectors? It is because cross-checking of obtained results and they can be switched by push-pull method[20] and share one interaction point between them. ILD has been invented by the initiative of Asia and Europe. On the other hand, SiD has been invented by the initiative of north America. In this study, the ILD detector is used.

ILD detector is constructed by some sub-detectors and the figure.16 shows a conceptual figure of ILD. ILD is constructed by four main sub-detectors; Vertex Detector(VXD), Time Projection Chamber(TPC), Electromagnetic Calorimeter(ECAL) and Hadronic Calorimeter(HCAL). VXD measures the interaction point and finds the decay vertex of heavier particles that can fly longer. VXD plays an important role in quark flavor identification. TPC reconstructs the 3-dimensional track of charged particles. VXD, TPC, and Silicon Trackers[54] reconstruct tracks of charged particles, and they are collectively called Tracking system. ECAL measures the energy of charged particles through the shower of electrons and photons (it is called electromagnetic shower). On the other hand, HCAL measures neutral particles' energy through hadrons' showers (called hadron shower). These sub-detectors are surrounded by the superconducting magnet, which has the solenoidal magnetic field of about 3.5T belong to z-axis[55]. Since charged particles are bent by this magnetic field and its momentum proportional to the track's curvature, each charged particle's momentum can be measured at the Tracking system. One more detector called Return Yoke and Muon Tracker (it is often called Outer tracker because it is installed on the outermost) shields the magnetic field and detects muon, which has a long lifetime. It can be used to measure hadron showers, which overhung hadron shower.

This is the outline of ILD detector. Let us see more detail of them in later sections.

[54] They are equipped between TPC and other detectors.
[55] z-axis means the direction perpendicular to endcap plane of the detector.

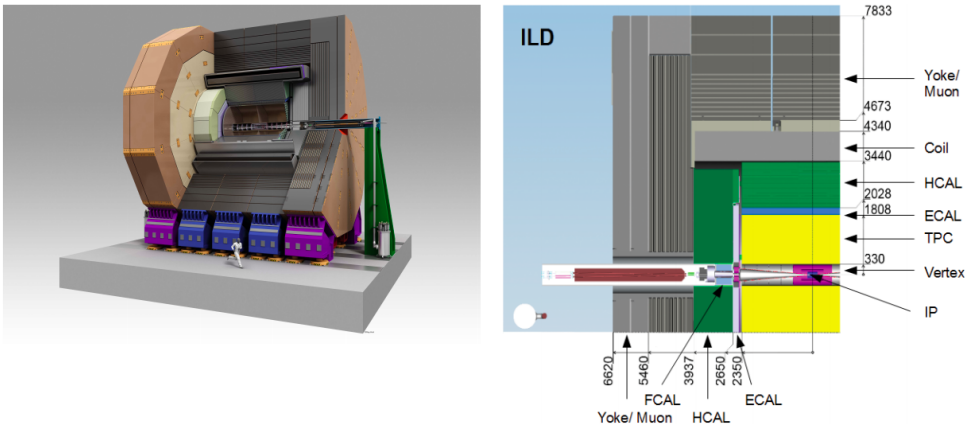


Figure 16: Conceptual figure of ILD detctor [20] : The innermost detector is VXD. TPC, ECAL, HCAL, Solenoidal coil, muon tracker, and return yoke are arranged in Baumkuchen shape to surround the interaction point.

3.2.2 Vertex Detector

The main mission of VXD is to identify b quark and c quark and distinguish them from uds quarks and gluon. As we can see eq.(2.6), the observable R_3^{bl} distinguishes b quark from uds quarks. So, identification of b quark is an important element for this b running mass measurement. We can identify and distinguish them by finding decay vertices of b or c quarks. Why can we do this? When b quark decays, the interaction that acts on here is weak interaction, and quark flavors are mixed under this force. The strength of quark flavor mixing is described by CKM(Cabbibo-Kobayashi-Masukawa) matrix[56][22] :

$$V_{\mathrm{CKM}} = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{tc}| & |V_{tb}| \end{pmatrix} \sim \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix}, \tag{3.2}$$

where $\lambda \sim 0.23$. In 250GeV ILC, top quarks do not appear, so let us focus on the quark mixing between udscb quarks. We can see that transitions of $b \to c$ and $b \to u$ are suppressed compare with other transitions. Figure.17 shows it by width and color denser of arrows. So b quark is difficult to decay, and it can be interpreted that b quark can survive longer than other lighter quarks. Because of such reason, b quark has a longer lifetime of 1.5 ps, and it can fly about several hundred ¯m to several mm from the interaction point(it is called Primary vertex). Afterward, when b decays to c, b quark leaves one more vertex(Secondary vertex). Furthermore, c quark leaves one more vertex (Tertiary vertex) when c will decay to s quark(See figure.18). As a result, the jet originated by b quark has three decay vertices, and the jet which comes from c quark has two decay vertices. On the other hand, the jet, which comes from uds light quarks and gluon has only one vertex. Therefore, if each vertex can be detected with sufficient resolution, we can identify quark flavors of detected jets.

[56] For qualitative explanation, I remained only characteristic contributions.

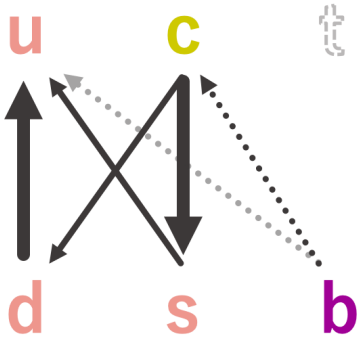


Figure 17: Strength of quark mixing : Direction of arrow means transition destination, and width(wider is stronger) and color denser (denser is stronger) of arrow means its strength. Here, t quark is ignored.

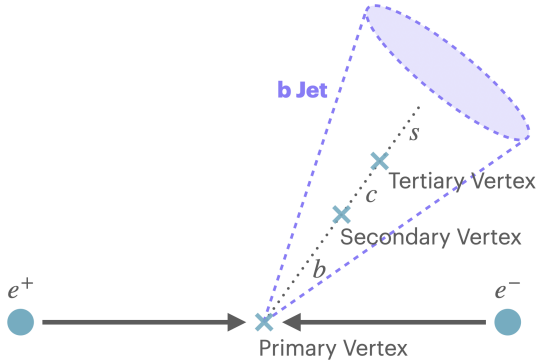


Figure 18: Conceptual figure of quark flavor ID

Such high-quality quark flavor identification, which is used vertex information, is one of the advantages of lepton colliders, such as ILC. The detector which is responsible for the detection of decay vertices is VXD. In order to accomplish sufficient quark flavor identification, the following conditions are required[10] :

- Point resolution less than 3μm
- Amount of material per a layer less than $0.15X_0$ in order to suppress multiple Coulomb scattering
- The radius of inner most layer is 1.6 to 6.0 cm
- Occupancy rate less than several % (If occupancy rate is high, we can not distinguish original signal from background contaminations.)
- Vertex resolution should be $\sigma \leq 5\mu m + \frac{10}{p\beta \sin^{3/2}\theta}\mu m$ where p is the momentum and β is velocity, θ is the polar angle of incident particle. First term means the resolution of detector itself(corresponds to the pixel size) and second term comes from multiple scattering.

These requirements are more strict than other experiments. Actually, following tables.6 and 7 shows the comparison between ILD and DELPHI, one of the detectors of LEP. Studies of Pixel VXD have been developed to satisfy such requirements so far, and there are some candidates; CMOS, DEPFET, and FPCCD. The VXD in ILD is constructed by three layers which each layer has a couple of layers, so there are six layers in total. Spatial resolution at each layer is summarized as the following table.6. These numbers are obtained under simulation and technical studies.

	radius(cm)	length(cm)	$\|\cos\theta\|$	$\sigma(\mu m)$
Layer 1	1.6	6.25	0.97	2.8
Layer 2	1.8	6.25	0.96	6
Layer 3	3.7	12.5	0.96	4
Layer 4	3.9	12.5	0.95	4
Layer 5	5.8	12.5	0.91	4
Layer 6	6.0	12.5	0.9	4

Table 6: Performance of VXD of ILD[10] : There are six layers and innermost two layers have half length of other layers to suppress the occupancy rate of beam background. Spatial resolution σ on $R\phi$ plane is based on CMOS.

Based on these spatial resolutions, the impact parameter resolution is provided as the function of the momentum of incident particles by figure.19. On the other hand, the performance of VXD in DELPHI detector is given as the table.7. According to table.6 and table.7, we can see that ILD superior to DELPHI for each element.

	radius(cm)	length(cm)	$\|\cos\theta\|$	$\sigma(\mu m)$
Layer 1	6.3	$\leq$14	$\leq$0.91	8.0
Layer 2	9	$\leq$14	$\leq$0.91	8.0
Layer 3	11	$\leq$14	$\leq$0.91	8.0

Table 7: Performance of VXD of DELPHI[11] : These numbers are based on 1994-5 version. We can see that VXD in ILD superior to DELPHI for each element.

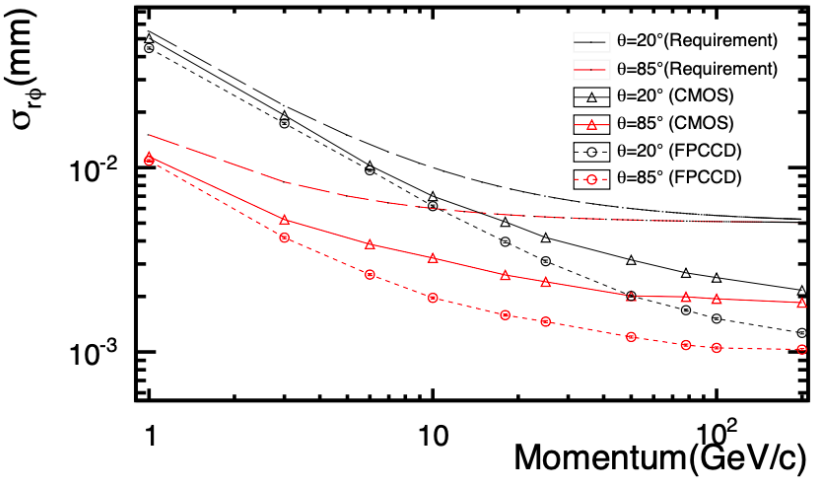


Figure 19: The resolution as the function of the incident particle's momentum[10] : Circle dots indicate the performance for CMOS, triangle dots are FPCCD. Black lines are each performance by $\theta = 20°$, red lines are the ones by $\theta = 85°$. Long dashed curves are performance goals with respect to each incident angle θ.

3.2.3 Time Projection Chamber

TPC is a detector that detects signals of charged particles and constructs the 3-dimensional track of them. The basic mechanism is like this: The gas mixture containing Ar as the main component(Ar-CF_4-isobutane) is filled inside of TPC[23]. When charged particles pass through it, the gas is ionized along the trajectory. Arisen electrons are drifted by an applied electric field, which belongs to the direction of the z-axis, and they are detected at the detector where the end cap of TPC. TPC can detect about 224 signal points per track, and we can obtain

2-dimensional track information on the $R-\phi$ plane from it. Additionally, we can know z-direction information from the drifted time of electrons. By summing them, 3-dimensional track information can be obtained. Obtained tracks have curvature because the magnetic field of 3.5T is applied to TPC by the superconducting coil. The momentum of charged particles is decided from obtained track information. It is one of the important roles of TPC. Moreover, we can make particle identification from dE/dx information of gas ionization. dE/dx is called Stopping power, which means the energy loss per unit length. It is given by like this [24] :

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ln\frac{2m_e c^2\beta^2\gamma^2 W_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2}\right]. \tag{3.3}$$

It is called the Bethe-Bloch equation. Here, the meaning of each quantity can be checked in the reference, but the first term means the energy loss by excitation of atoms in the matter, and the third term is called density effect, which gives the correction of several %. When the charged particle passes through in the matter, around matters are polarized, and its degree changes according to the charged particle's energy. As a result, the third term affects well for the charged particle, which has higher energy. This stopping power is unique for each material for momentum. So it is an index of particle identification.

3.2.4 Calorimeters

There are two calorimeters in ILD; ECAL and HCAL. Both detectors measure the energy of particles that occur in electromagnetic or hadronic showers. High jet energy resolution and good particle separation performance are required.

ECAL is designed by 30 layers sand-witch of absorber and sensor, but in order to accomplish the above requirements, tungsten is used as the absorber in ECAL. We can improve particle separation performance by using tungsten because it has a small Moliere radius. Moliere radius means the energy spread, which includes 90% of showers. When the Moliere radius is smaller, the shower size tends to be smaller, and showers' duplication is reduced. Sensor layers of ECAL have been developed so far, and there are some choices: Silicon sensor(it is called SiECAL), Scintillator sensor(ScECAL) and hybrid sensor of them[10] [25].

On the other hand, HCAL is a detector that detects neutral particles, and it needs a large volume because of its long hadron interaction length. Since electromagnetic shower also appears in hadron shower, the energy and particle separation resolution is lower than ECAL. HCAL is constructed by 48 layers sand-witch of absorber and sensor just as ECAL, but iron is used as absorber layers. Because iron has a smaller ratio of hadron interaction length and radiation length ($\lambda_I/X_0 = 9.5$) than other heavy metals, HCAL can also measure electromagnetic shower in hadronic showers. As candidates of sensor layer, there are two concepts[20]: One uses plastic scintillator tiles of $3\times3\text{cm}^2$ and readout with an analog system. Another one uses cell geometry of $1\times1\text{cm}^2$ with the gas-based detector and readout with a binary or semi-digital system.

3.2.5 Particle Flow Algorithm

As we have seen in the previous outline, the ILD detector is constructed by several sub-detectors. Created particles interact and drop energy to detectors, and dropped energy will be enhanced and output as electric signals. Each detector can measure energy, track, decay

point(vertex) by collecting these signals. So, it should be noted that what we can see is these signals, not but particles themselves. According to such obtained signals, tracks are made by combining them and identifying the corresponding particle to each track. Usually, this procedure is called Reconstruction. Namely, we reconstruct created particles by using detectors' information, and we analyze by using these reconstructed particles.

Charged particles are measured at both Tracker and ECAL. Particles that are interacted by strong interaction, such as hadrons, are measured at HCAL. Thus, each particle is measured at each detector, and some particles can be measured at some detectors. If the particle drop information to some detectors, we can reconstruct it by using information that comes from both detectors. However, in ILD experiments, each created particle is reconstructed at the detector, which is good at detecting this particle. For example, the photon is measured at ECAL only, so the photon is reconstructed at ECAL. Whereas charged particles such as the electron can be measured at ECAL and Tracker but, the momentum resolution of the tracking system has better precision than calorimeters up to a specific momentum range of ($\sim 100 - 150\mathrm{GeV}$). Therefore, the reconstruction of charged particles is done by the tracker's information, but not ECAL. Particle reconstruction at ILD is designed by dividing the roles between each sub-detector, and such method of particle reconstruction is called Particle Flow Algorithm(PFA)[26]. The following picture shows which type of particle is reconstructed at which detector.

In this chapter, we have seen that the outline of ILC and the role of each sub-detector in ILD. To simulate detector reaction, the data file[57] which reflects the geometric structure(VXD, ECAL, magnetic field of Solenoidal Coil etc...) of the ILD detector is used. The next section explains the experimental tools used in this study and how to simulate b running mass measurement at ILC.

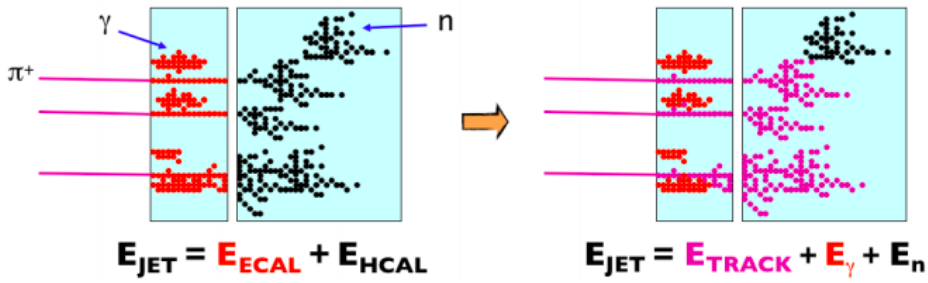


Figure 20: Conceptual figure of Particle Flow Algorithm [26] : The left figure indicates that the traditional reconstruction method. In this method, since charged hadrons such as π drop signals at both ECAL and HCAL, it is double-counting. The right figure shows that the particle reconstruction by PFA. In PFA, since each particle is reconstructed by only the detector, which is good at reconstructing it, we can avoid double-counting.

[57] We call it Gear file. They are made based on DBD(Detailed Baseline Document) of ILC.

4 Experimental Framework for Simulation

This chapter explains the experimental framework of detector simulation. As the first step, the simulation is started from event generation. This study used WHIZARD and PYTHIA for it. As the next step, the event generator's output is passed to the detector simulator, in this study, the Marlin processor. The simulation which realizes detector reaction for each cell and each sub-detector is called Full-Simulation. In this study, full-detector simulation of ILD for $e^+e^- \to q\overline{q}$ process is done, and the precision of b quark mass measurement at 250GeV is estimated. More detail of each procedure is explained in later sections of this chapter.

4.1 Event Generator and Hadronization model

4.1.1 WHIZARD

At first, we need to prepare the process of what we want to consider. The signal event and the background events of this study are shown in the next chapter, but as described in section 2.3.1, the process we want to focus on is $e^+e^- \to q\overline{q}$. So we should grasp the cross-section of this process, and the Event generator allows us to calculate it easily. This study used WHIZARD package[28]. WHIZARD calculates the cross-section up to a specified perturbation order, and it is used to generate the process information up to Parton level, which appears some quarks and gluons.[58] But the decay does not stop there, but they make Parton Shower and occur hadronization. Figure.21 shows the illustrative evolution of the $q\overline{q}$ process. Original partons have energy $O(\sqrt{s})$, but the energy scale becomes lower by Parton shower, which emits more gluons and quarks. Strong vertices α_s become bigger enough to be not able to apply perturbative QCD. After that, hadronization occurs, and hadrons appear. We can not longer use perturbative QCD to calculate Parton Shower and hadronization, but how is such information generated? The next section answers it.

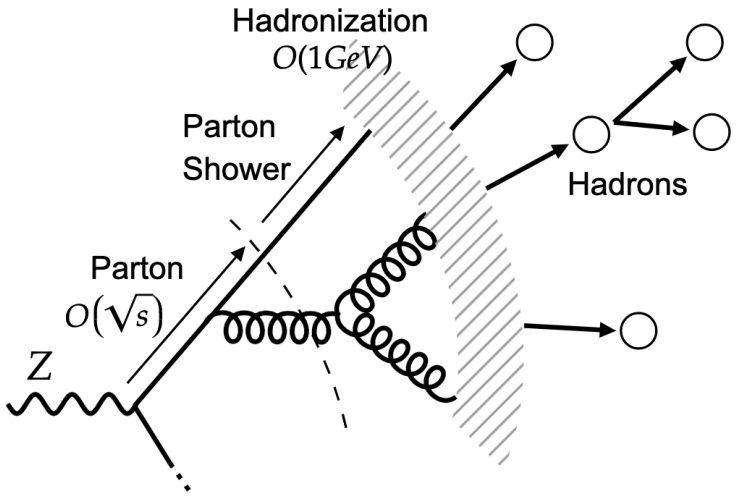


Figure 21: Evolution of hadronic decay : Strong coupling α_s becomes larger as the decay is progressed(energy scale goes to lower) and perturbative QCD can not be applied. To describe Parton Shower and hadronization, we need to rely on models.

[58] Parton level means the energy level which can be calculated strictly by perturbative QCD.

4.1.2 PYTHIA

Quark pair production for e^+e^- collision is already simulated in WHIZARD as described in the previous section. Later evolution from Parton Shower level can not be explained by perturbative theory, and we have no choice but to rely on models to describe hadronization and hadron decay. So we need one more tool to generate Parton Shower and hadronization according to models. Such models are called Fragmentation models, and in general, there are two styles: String Fragmentation or Cluster Fragmentation. The model which this study uses is PYTHIA [27] which implements the string fragmentation model. In PYTHIA, based on Parton, which is output by WHIZARD, Parton shower generation and hadronization are simulated. We could prepare a physical reaction of elementary particles. However, in the actual experiment, such particles interact with detectors and what we can see is its signal but not particles themselves. So we need to simulate interactions between generated particles and detectors under the Monte Carlo(MC) Simulation. The next section explains detector simulation.

4.2 Marlin Processor

4.2.1 About Marlin

Marlin processor is the package for detector simulation of the linear collider. When we run the detector simulation, we should mainly do the reconstruction of particles and jets and identify quark flavors. Each package for particle and jet reconstruction and flavor identification is already prepared. For example, there is the package called Pandora PFA for the reconstruction of particles[26]. In this study, we are using Fast Jet package[29] for jet reconstruction and LCFI Plus package[30] for quark flavor identification. Typically, we need to prepare one module for one operation (e.g. jet reconstruction), and it is called Processor in Marlin. Through each processor, detector reactions and the reconstruction of particles and jets are simulated. Marlin outputs the file, which includes Monte Carlo(MC) generated information and reconstructed information. We analyze it, and the result of what we want is yielded. Let us summarize them as the outline of simulation in the next section.

4.2.2 Flow of Detector simulation

The procedure of detector simulation in Marlin is like this :

1. Prepare input sample files, which are output by event generator (WHIZARD+PYTHIA)
   Prepare a gear file which provides detector geometry information
   (e.g. magnetic field of solenoidal coil etc...)
   The detector geometry based on ILD_o1_v05 is used in this study.

2. Create detector geometry.
   Initialize geometrical information of the detector

3. Reconstruct particles (run Pandora PFA)
   Reconstruct tracks of appeared particles based on signals from each sub-detector.

Yield important parameters from each reconstructed track(corresponds to a particle) such as momentum, energy etc...

4. Reconstruct jets (run Fast Jet)
   Reconstruct jets according to Jet-Clustering algorithm (in this study, CAMBRIDGE algorithm)
   It is explained in the next section.

5. Quark flavor ID (run LCFI Plus)
   Output likelihood of b and c quark according to Flavor-Tagging algorithm based on VXD signals.
   It is explained in the last section of this chapter.

There are two procedures that have not been explained yet; Jet-Clustering and Flavor-Tagging. They are essential concepts for this study because the observable eq.(2.6) is defined as the fraction of 3-jet ratio between b quark and uds quarks. Jet reconstruction and quark flavor identification are essential for this analysis. Then, let us see them in the rest sections of this chapter.

4.3 Jet-Clustering algorithm

4.3.1 Outline of Jet-Clustering algorithm

Detected particles are reconstructed by PFA(See 3.2.5). To reconstruct jets, we need to cluster these reconstructed particles in the shape of jets. However, in principle, at actual experiments, it is impossible to chase and specify which particle(hadron) belongs to which jet from obtained tracks. In other words, more simply, how to reconstruct jets is not trivial. Figure.22 may help your understanding of this point. If reconstructed tracks or particles are given just as black lines, some people may think it is a 2-jet event like Pattern A. On the other hand, other people may think it is a 3-jet event like Pattern B.

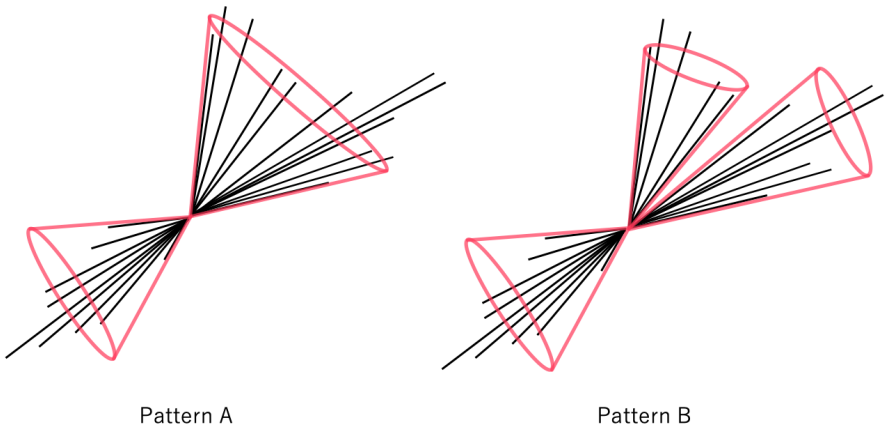


Figure 22: How are jets reconstructed? : Black lines mean reconstructed particles by PFA(sometimes they are called PFO(Particle Flow Object)). If PFOs are given like this, some people may reconstruct jets like the Pattern A(2-jet event). On the other hand, other people may reconstruct like the Pattern B(3-jet event). We need some objective method that does not depend on any observer for jet reconstruction.

The objective algorithm, called the Jet-Clustering algorithm, is used, and jets are reconstructed according to this algorithm. So far, two types of Jet-Clustering algorithm has been invented :

- Cone algorithm

  This algorithm sets "cones" and consists each jet by hadrons in it. This algorithm is famous among experimental physicists, but it includes infra-red and collinear(IRC) divergences.

- Cluster Sequence algorithm

  This algorithm is more natural and theory-friendly algorithm. The flowchart of this algorithm is like this :

  1. Define the distance measure y_{ij} between chosen 2 particles.
  2. Compare it with the clustering criterion (jet resolution parameter)y_c, which is decided in advance.
     $\rightarrow$ If $y_{ij} < y_c$, combine momentums of 2 particles, and regard them as a particle and return to the procedure1.
     $\rightarrow$ If $y_{ij} > y_c$, terminate algorithm.

  The parameter y_c specifies the size of the reconstructed jet, so this parameter defines the jet experimentally. y_c can be taken various values(typically, <1), we have many choices of the definition of the jet. In general, it is fixed so that theoretical uncertainty becomes smaller. Definitions of the distance y_{ij} and combination ways(called recombination scheme) of particle momentum in the procedure1 has been devised variously so far. There are some different algorithms in this style. The next sections explain them.

  It is often called an inclusive algorithm. There is an exclusive way, and its flowchart is like this :

  1. Define the distance measure y_{ij} between chosen 2 particles (Same as above).
  2. Set the jet number and execute the algorithm until the number of clusters matches it by adjusting the criterion y_c.

  Of course, these two modes should be equivalent, so there is no fundamental difference between them. In this study, the exclusive mode is used, and all events are forced to 2-jet events. This point is useful to cut background events. We will see it in the next chapter.

In this study, we use the Cluster Sequence algorithm with exclusive mode. Let us look at the famous algorithms of this style in later sections.

4.3.2 DURHAM algorithm

In DURHAM algorithm, the distance measure is defined by

$$y_{ij} = \frac{2\min\left(E_i^2, E_j^2\right)(1-\cos\theta)}{s^2} \tag{4.1}$$

Here, $\sqrt{s}$ is the CM energy. As a recombination scheme, we use $p_i, p_j \rightarrow p_i + p_j$ and regard two particles as a new particle which has the momentum $p_i + p_j$ (it is called E-scheme). According to this distance and this recombination scheme, this algorithm is executed as explained in 4.3.1.

4.3.3 kT algrorithm

In high energy experiments, photons cling to electron and positron beams. Since most of them have small energy, these photons are harmless. However, some photons have relatively large energy. If these photons interact and emit hadrons by the process of $\gamma\gamma \to q\overline{q} \to hadrons$, such jets are harmful because we want to focus on hadrons come from $e^+e^- \to q\overline{q}$. It is called the beam jet. In this case, since this beam jet event is a background event, such events should be eliminated. However, the mechanism to remove it is not equipped in DURHAM algorithm. In kT algorithm[29], the mechanism to remove such background jets. The distance measure is defined by

$$d_{ij} = \min\left(k_{\mathrm{T}i}^{2p}, k_{\mathrm{T}j}^{2p}\right)\frac{\Delta_{ij}}{R^2} \qquad (4.2)$$
$$\left(\Delta_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2\right)$$

Here, p, R are input parameters that we need to decide by our hand, and $k_{\mathrm{T}i}$ is the transverse momentum, $y_i \equiv \frac{1}{2}\ln\frac{E_i+p_{zi}}{E_i-p_{zi}}$ is rapidity, ϕ_i is azimuthal angle of i$^{\mathrm{th}}$ particle.

How should we eliminate beam jets? Since high-energy photons can make beam jets are boosted to the beam axis, beam jets emit near the beam pipe direction. How should we eliminate beam jets? Since high-energy photons can make beam jets are boosted to the beam axis, beam jets emit near the beam pipe direction. Therefore, beam jets tend to have low transverse momentum. In kT algorithm, there is one more additional distance measure d_{iB}. It is defined as the distance between i$^{\mathrm{th}}$ reconstructed particle and beam, and it is set by $d_{iB} = k_{\mathrm{T}i}^2$. Here, $k_{\mathrm{T}i}$ is the transverse momentum of i$^{\mathrm{th}}$ particle. kT algorithm is run by the following procedures :

1. Find the smallest distance :
   $\to$ If it is d_{ij}, combine them according to the recombination scheme and return to the procedure1 again.
   $\to$ If it is d_{iB}, remove this particle and regard it as a beam jet.

2. When d_{iB} or d_{ij} exceeds the criteria d_{cut}, rest clusters are non-beam jets what we want to yield.

3. Iterate them until there are not any left particles.

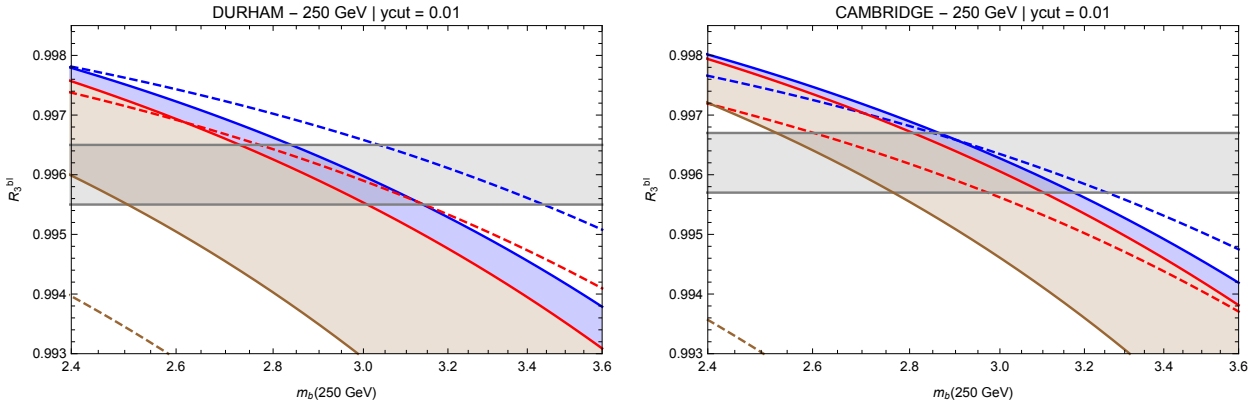
Thanks to using these two distances, we can discard beam jets. As described, we have two parameters that we can decide by hand; R and p. There are several algorithms for p, especially the algorithm which $p = 0$ is called Cambridge algorithm. This study used the generalized version for e^+e^- collider of this kT algorithm. The next section introduces it.

4.3.4 Generalized kT algorithm for ee collider

In this study, I used generalized kT algorithm for e^+e^- collider. In this algorithm, the distance is defined by

$$\begin{aligned} d_{ij} &= \min\left(E_i^{2p}, E_j^{2p}\right)\frac{1-\cos\theta_{ij}}{1-\cos R} \\ d_{iB} &= E_i^{2p}. \end{aligned} \qquad (4.3)$$

The case which set $p = 0$ corresponds to Cambridge algorithm. Is Cambridge algorithm is a better way? Figure.23 shows the comparison of the observable R_3^{bl} as the function of running b mass m_b between DURHAM and Cambridge algorithms.



(a) DURHAM algorithm at $\sqrt{s}$ =250GeV ($y_c = 0.01$) **(b)** CAMBRIDGE algorithm at $\sqrt{s}$ =250GeV ($y_c = 0.01$)

Figure 23: Theoretical calculations of R_3^{bl} as the function of m_b in DURHAM(left) and CAMBRIDGE(right) algorithms[14] : The usage guide is same as figure.13. Bands between each line means theoretical uncertainty.

According to these calculations, b quark mass seems to more stable(smaller uncertainty) in Cambridge algorithm than DURHAM's one. Because of such reason, Cambridge algorithm is used as Jet-Clustering algorithm in this study. We have one more parameter R what we should decide by hand. It is often called the jet radius means the maximum size of jets. It is introduced to eliminate beam jets[31]. It is chosen 1.25 as an optimized value of R for efficiency and radiative return background rejection maximization[14]. When the Jet-Clustering algorithm is run actually in Marlin processor, it is used in Fast Jet package[29]. The next section shows the result of the matching check of jet reconstruction performance of Cambridge algorithm.

4.3.5 Matching between MC Parton and reconstructed jet

Previous sections described how jets are reconstructed, and it is mentioned that this study uses Cambridge algorithm. This section discusses the performance of jet reconstruction in Cambridge algorithm. We compare the angle of a parton generated by MC and the angle of a reconstructed jet reconstructed by Cambridge algorithm. Of course, since parton occurs hadronization and decays to each jet, two angles which are generated Parton and reconstructed jet should be close even if they do not match completely. We can check the reconstruction performance of Jet-Clustering algorithm by such a matching check. At first, let us see two reconstructed jets are appeared back to back certainly. In the ILC, since the same energy e^- and e^+ collide at CM frame, generated quarks or jets should be appeared by back to back. Namely, two angles θ_1(first jet's angle) and θ_2(second jet's angle) should keep the relation of

$$\cos\theta_1 = -\cos\theta_2.$$

According to figure.24, we can see the negative linear correlation as the above relation expects.

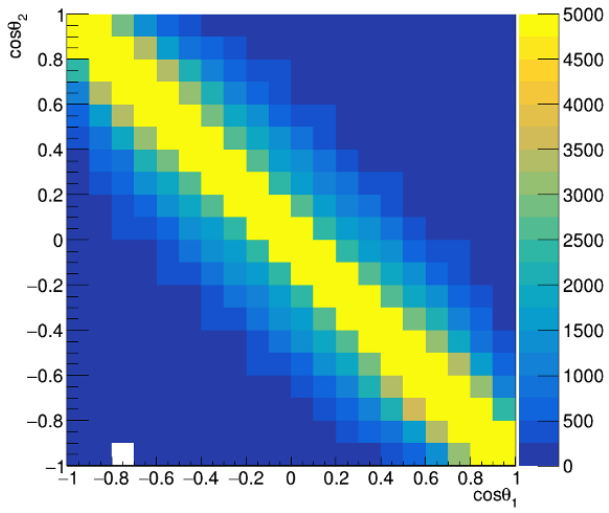
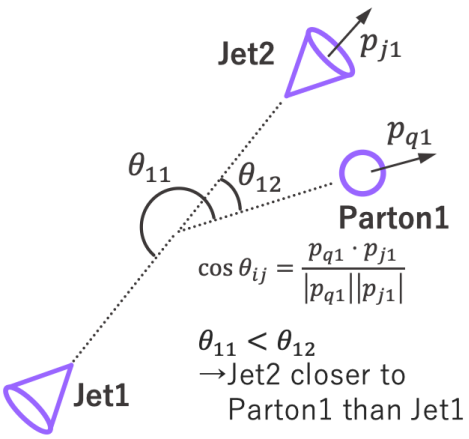
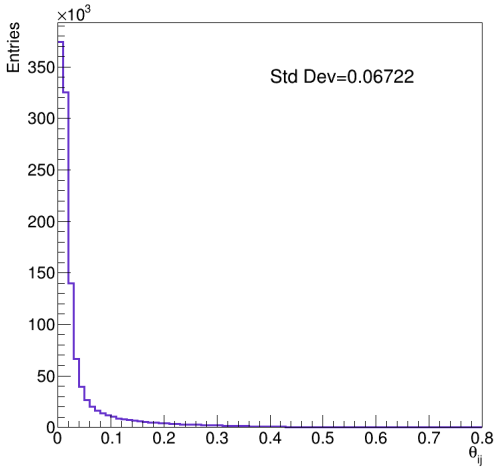


Figure 24: 2D distribution between angles of 2 reconstructed jets

Then, let us see how well these reconstructed jets match with generated parton. To check it, $\cos\theta_{ij}$, which θ_{ij} is the angle between ith generated Parton and jth reconstructed jet $(i, j = 1, 2)$, is used. For example, if $\cos\theta_{11}$ is smaller than $\cos\theta_{12}$, we can state that the first jet is closer to the first parton than the second jet. Additionally, if this $\cos\theta_{11}$ is sufficiently small, we can state that the first jet seems to come from the first parton(See figure.25(a)). We can grasp how well reconstructed jets match with generated parton by looking at this result. Since we have two jets and two parton, we have four combinations. Entries of the plot (figure.25(b)) includes all combinations. According to this result, the width of the distribution is about 0.07. So we can state that our jet reconstruction is reliable. So far, we have seen how to reconstruct jets. There is one more important technical element in this study; Flavor-Tagging. The next section explains it.



(a) Conceptual figure of matching



(b) Result of matching

Figure 25: Matching check : Checked how well are jets reconstructed by Jet-Clustering algorithm by taking the angle between MC generated Parton. According to (b), we can check that jets are reconstructed well by good performance.

4.3.6 Flavor-Tagging

As we can see in eq.(2.6), identification and distinction of b quark and uds quarks are one of the important work in this analysis. However, we can not distinguish them by using constituents of jets. How should we identify each flavor of reconstructed jets? The answer was already given in section 3.2.2. Since b quark has the Secondary vertex, which separates about several hundred ¯m to several mm from the interaction point(Primary vertex), if this vertex is found for the considering jet or reconstructed tracks through signals of VXD, we can state that it comes from b quark(See figure.18). In this study, Flavor-Tagging is executed with LCFI Plus package[30] in Marlin.

This chapter explained the experimental framework which is used in this study. The next chapter shows obtained results of the observable R_3^{bl} and the precision of b quark mass measurement at 250GeV under this experimental framework.

5 Estimation of b quark mass precision

5.1 Analysis Steps

This chapter describes how to analyze obtained information from detector simulation and the result. Each procedure or algorithm introduced in the previous section is executed in the ILC Soft package framework. This study used the ILC Soft version of 02-00-01. At first, let us review what we should do in the analysis. After the running of detector full-simulation by Marlin processor, Marlin outputs files by .root(or .slcio) extension. .root file can be used in ROOT[32] that is the software which is invented at CERN, and we can analyze simulated data from Marlin in ROOT. Let us look at the outline of analysis of this study :

1. Event Selection :
   Select signal events and reject background events by using various variables. The signal process is $e^+e^- \to q\overline{q}$. The process, which includes high energy photon radiation in $e^+e^- \to q\overline{q}$ decreases the energy of $q\overline{q}$ pair from 250GeV. Since this study focuses on b quark mass at 250GeV energy scale, such events should be cut. However, since the detector can not detect almost all radiation, the photon energy can not be used to cut it, and we need to think out some alternative way.

2. Selection by Flavor-tagging :
   Since we need to distinguish b quark from uds light quarks, as we can see at eq.(2.6), flavor identification of reconstructed jets is made here. To do this, Flavor-tagging, which is explained in section 4.4, is used.

3. Observable R_3^{bl} measurement :
   After the above selections, measure R_3^{bl} by counting each jet event's numbers. It is the parton level result that can compare with theoretical calculations, and it is obtained by corrections come from hadronization and detector effects to observed result. Since MC generates the parton level result in this study, corrections of hadronization and detector effects are estimated. We can estimate statistical error from statistics of $e^+e^- \to q\overline{q}$ sample.

4. Estimation systematic errors :
   Estimate systematic errors have two sources; hadronization model and detector effects(event selection, background rejection, quark flavor identification). They are estimated by Toy MC. In this study, these uncertainties are estimated as errors that appear on correction factors to the parton level.

5. Determination of b mass :
   After the estimation of errors on observable R_3^{bl}, estimate the precision of b quark mass measurement at 250GeV. The precision of b mass measurement is estimated from the precision of R_3^{bl} according to eq.(2.8).

This is the outline of this analysis to estimate the precision of b quark mass measurement at 250GeV. More detail of them is explained in later sections.

5.2 Event Selection

5.2.1 Signal and Background processes

As described in the chapter2, gluon radiation from the quark has quark mass sensitivity. So, the signal process that we should focus on is $e^+e^- \to q\overline{q} \to jets$ as drew the diagram below. We need to consider both flavors; b quark and u/d/s light quarks, because the observable R_3^{bl} is defined as the fraction of 3b-jet event ratio and 3uds-jet event ratio.

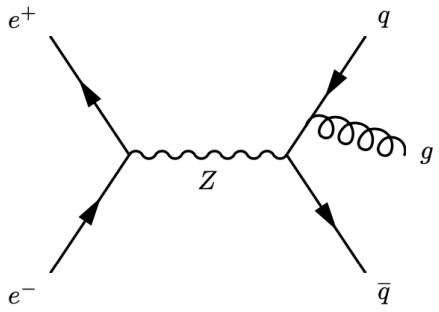


Figure 26: Signal event : q is b quark or uds light quarks. Created quark pair emits gluon, and these parton decay to the 3-jet event through hadronization.

Since the observable R_3^{bl} counts 3-jet events, how to define the 3-jet event experimentally is important for our analysis. In general, jets are reconstructed according to the Jet-Clustering algorithm, as explained in chapter4. Here, we are using it as exclusive mode; we force to cluster until the 2-jet event. Therefore, every event is forced to become a 2-jet event, and the criterion parameter y_c of clustering 3-jet event to 2-jet event has various values[59] The 3-jet event is defined as the event which the jet number is three on a chosen resolution parameter y_c. What is the "chosen y_c"? How should we choose y_c? We choose a y_c value so that the theoretical uncertainty of R_3^{bl} becomes small. Figure.27 shows the theoretical calculation of R_3^{bl} vs y_c under the CAMBRIDGE algorithm which we are using in this study.

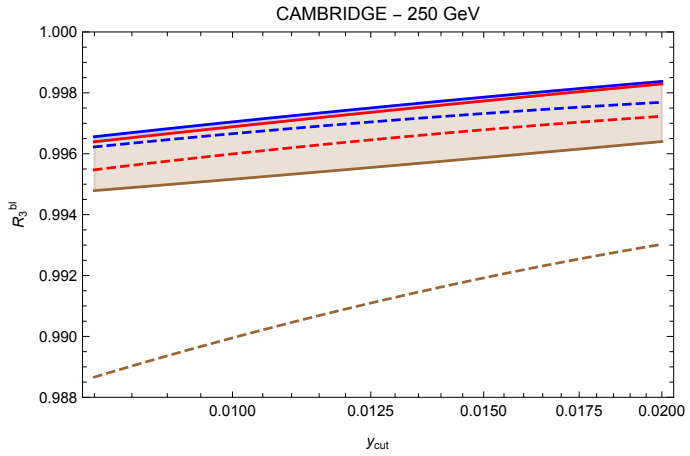


Figure 27: Theoretical prediction of R_3^{bl} as the function of y_c[14] : The usage guide is same as figure.13.

According to figure.27, we can see that R_3^{bl} tends to be stable(does not change largely) for the whole range. In this analysis, the y_c criteria is chosen as 0.01 and define the 3-jet event as

[59] Typically, y_c is in a range of 0 to 1.

the event which is the jet number is 3 on $y_c = 0.01$.

On the other hand, there are three types of background events which we should remove. The first one is the radiative return event. Before e^- and e^+ collide at high CM energy, they emit photons. Such radiation is called Initial State Radiation(ISR). In general, this ISR can take various energy sizes, and if ISR takes enormous energy, the "effective" CM energy we can use in the collision is decreased extremely. In this study, we want to execute the experiment at 250GeV, but higher energy ISR radiation, which returns the energy of $q\bar{q}$ pair to Z-pole can not be ignored(See table.9)[60]. If such radiation occurs, b quark mass effect on R_3^{bl} is no longer the one at 250GeV, but at Z-pole. Such a higher-ISR event is called Radiative return, and we should eliminate it. It should be noted that LEP did not go through the problem of radiative return because it runs at Z-pole. Such a situation appears at 250GeV ILC for the first time.

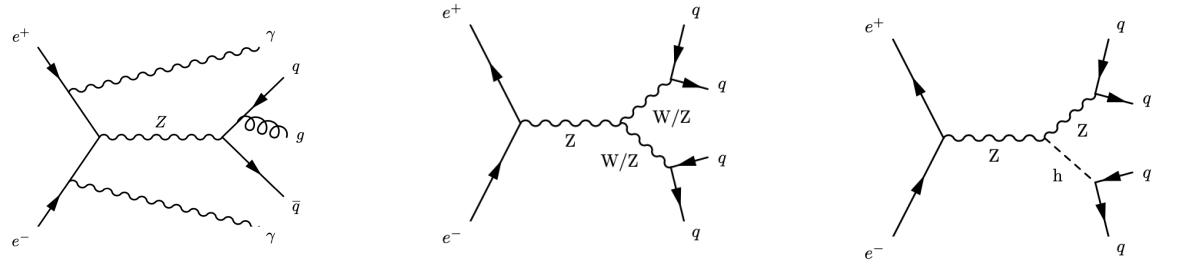
$E_\gamma < 50$GeV				
Polarization	$\sigma_{q\bar{q}}$[fb]	$\sigma_{b\bar{b}}$[fb]	$\sigma_{c\bar{c}}$[fb]	$\sigma_{l\bar{l}}[fb]$ $(l = uds)$
$e^-_L e^+_R$	34253.7	5970.9	8935.2	19347.6
$e^-_R e^+_L$	11007.6	1352.1	3735.1	5920.4

Table 8: Cross section of signal events at 250GeV : These numbers are calculated under the criteria of ISR energy of 50GeV. If the event that ISR energy larger than 50GeV, it is regarded as the radiative return event. This criterion is optimized to reject the radiative return and retain signal events as possible.

Polarization	$\sigma_{q\bar{q}\gamma}[fb]\,(E_\gamma > 50\text{GeV})$	$\sigma_{WW\to q_1\bar{q}_2 q_3\bar{q}_4}$[fb]	$\sigma_{ZZ\to q_1\bar{q}_2 q_3\bar{q}_4}$[fb]	$\sigma_{Zh\to q_1\bar{q}_2 q_3\bar{q}_4}[fb]$
$e^-_L e^+_R$	94895.3	14874.4	1402.1	346.0
$e^-_R e^+_L$	60265.3	136.4	605.0	222.0

Table 9: Cross section of background events at 250GeV : As we can see from the above numbers, radiative return that includes high energy ISR, which returns quark pair energy to Z-pole, can not be ignored compared with signal events. In addition to it, the other three channels also are not negligible.

In addition to radiative return, there is one more type of background events. They are full-hadronic decay mode of diboson events; $e^+e^- \to WW$ or ZZ or Zh as drew (b) and (c) in figure.28. Although they are 4-jet events, they can be misidentified. We can not distinguish them from signal events, so we need to remove them.



(a) Radiative return event

(b) Full-hadronic decay of $e^+e^- \to WW/ZZ$

(c) Full-hadronic decay of $e^+e^- \to Zh$

Figure 28: Background events : We are considering two types of background events; (a) is radiative return event. Because of missing energy by high energy ISR radiation, the CM energy of quark pair decreases and returns to Z-pole. (b) and (c) are full-hadronic decay modes of $e^+e^{\to}WW/ZZ/Zh$. They can be misidentified the number of jets when reconstructing jets.

[60]Since the Z-pole's CM energy matches the mass of the Z boson, it becomes a large contribution. It can be interpreted as the resonance.

5.2.2 Cut to remove Radiative return

One of the background events in this analysis is the radiative return event, as described in the previous section. It will be able to eliminate by imposing the cut of large ISR's energy. However, in principle, almost(more than 90%) of ISR is emitted at the nearby beam pipe[61]. So, ISR is invisible for our detector, and ISR energy and angles can not cut the high-energy ISR, which escaped at beam pipe. We should consider some alternative cut by something which is constructed by the reconstructed information from the detector. To cut invisible ISR, let us construct the energy of ISR by using angles of jets. If we consider the case of a photon is emitted like (a) of the figure.29, the energy and momentum conservations are given like this:

$$k^0 + p_1^0 + p_2^0 = 250\mathrm{GeV} \tag{5.1}$$

$$\boldsymbol{k} + \boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{0}. \tag{5.2}$$

Here, k^μ is the four-momentum of ISR, and p_1^μ and p_2^μ are four-momentums of jets. It should be reminded that all events are reconstructed as 2-jet events by the exclusive algorithm in this analysis. The angle between two jets ψ_{acol} is obtained from momentums of jets like this:

$$\sin\psi_{acol} = \frac{|\boldsymbol{p}_1 \times \boldsymbol{p}_2|}{|\boldsymbol{p}_1|\cdot|\boldsymbol{p}_2|} \tag{5.3}$$

The motivation is constructing the ISR's energy by using jet's information. By taking cross product of $\boldsymbol{p}_1$ and eq.(5.2), the following relation is obtained:

$$\begin{aligned}
&\boldsymbol{p}_1 \times \boldsymbol{k} = -\boldsymbol{p}_1 \times \boldsymbol{p}_2 \\
&\rightarrow |\boldsymbol{p}_1 \times \boldsymbol{k}| = |\boldsymbol{p}_1 \times \boldsymbol{p}_2| \\
&\rightarrow |\boldsymbol{p}_1|\cdot|\boldsymbol{k}|\sin\theta_1 = |\boldsymbol{p}_1|\cdot|\boldsymbol{p}_2|\sin\psi_{acol} \\
&\therefore |\boldsymbol{k}| = \frac{|\boldsymbol{p}_2|\cdot\sin\psi_{acol}}{\sin\theta_1}.
\end{aligned}$$

Here, θ_1 is the polar angle of the first jet. On the other hand, if we take the cross product of $\boldsymbol{p}_2$ and eq.(5.2), the following relation is obtained:

$$|\boldsymbol{k}| = \frac{|\boldsymbol{p}_1|\cdot\sin\psi_{acol}}{\sin\theta_2}.$$

Like this, the ISR's energy can be measured by using the momentum and angle of jets. However, the resolution of the jet's angle is better than momentum. If this ISR's energy could be expressed using the jet's angle only, it is better. Here, let us assume that each jet is massless, that is $p_1^0 = |\boldsymbol{p}_1|$, $p_2^0 = |\boldsymbol{p}_2|$. Since $k^0 = |\boldsymbol{k}|$, we can combine eq.(5.2) and above relation.

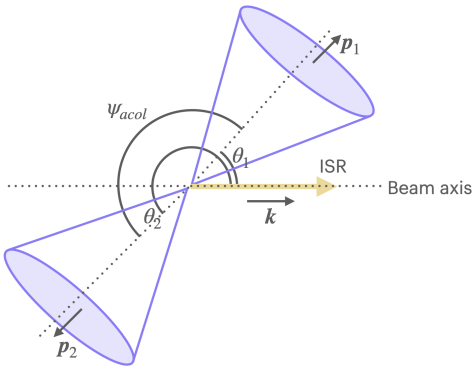
$$\begin{aligned}
k^0 &= \frac{\left(250\mathrm{GeV} - k^0 - p_1^0\right)\cdot\sin\psi_{acol}}{\sin\theta_2} \\
\rightarrow k^0 &= \frac{\left(250\mathrm{GeV} - p_1^0\right)\cdot\sin\psi_{acol}}{\sin\psi_{acol} + \sin\theta_2}
\end{aligned}$$

In this expression, the energy p_1^0 is still remained. However, we can remove it by the combination of this equation and the above relation between $|\boldsymbol{k}|$ and $|\boldsymbol{p}_1|$. As a result, the ISR's

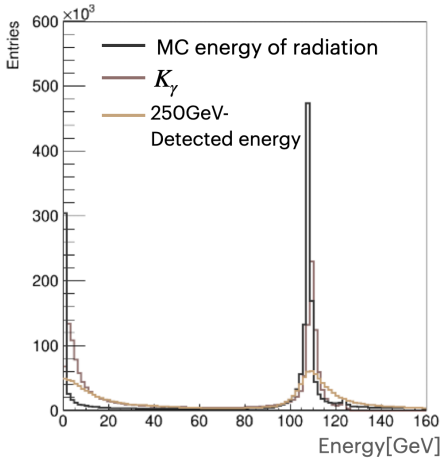
[61] It is because photons which are emitted from electron or positron beam are boosted to the direction of the beam axis.

energy is written by using the jet's angles only like this:

$$K_\gamma \equiv \frac{250\text{GeV} \cdot \sin\psi_{acol}}{\sin\psi_{acol} + \sin\theta_1 + \sin\theta_2}.$$



(a) The 2-jet event with ISR : All jet is reconstructed as 2-jet events in this analysis. ISR is emitted nearby beam pipe because it is emitted from accelerated beams.



(b) The distribution of K_γ : The black line is the MC generated ISR's energy. The darker brawn line shows the distribution of K_γ. The lighter brawn line is the invisible energy which is obtained from the difference of 250GeV and the total energy of reconstructed particles.

Figure 29: Cut invisible radiative return events : Almost of radiative return events escapes to the direction of the beam pipe. Cut them by constructing ISR's energy from jet's angles and imposing K_γ <50GeV.

(b) of the figure.29 shows distributions relate to ISR's energy. We can see that K_γ realizes that MC generated ISR's energy rather than using the invisible energy obtained from the difference of 250GeV and the total energy of reconstructed particles. The peak at ∼110GeV corresponds to Z-peak. Let us consider the case that the electron has the energy E_1 GeV and the positron has the energy E_2 GeV. When the positron(or electron) emits a photon, which has the energy X GeV[62], the electron and positron system is no longer symmetric(that is, electron and positron do not have the same energy). In the lab frame of the positron, the CM energy is given as $\sqrt{s} = 2\sqrt{E_1 E_2}$[63]. Since the electron's energy is 125GeV and the positron's energy is $125 - X$GeV, the CM energy becomes $\sqrt{s} = 2\sqrt{125 \cdot (125 - X)}$. When $X = 110$GeV, it corresponds to $\sqrt{s} = 90$GeV, that is, Z-pole. It matches the Z boson's mass, and the events that the CM energy returns to the Z-pole energy scale is a large contribution. Based on this plot, the condition $K_\gamma < 50$GeV is imposed in this study. Undetected high-energy ISR photons are cut up to about 97% by imposing this condition. Additionally, we need to care for detected high-energetic photons too. They are cut by using the invariant mass of all jets and PFO(that is, reconstructed particle) :

- Invariant mass of all jets should be greater than 130GeV
- Both jets should contain more than 5 neutral PFOs
- None of the jets should contain a neutral PFO with $E > 50$GeV in the very forward region $|\cos\theta| > 0.98$

[62] For the radiative return, the process, which has one photon, is dominant. The process, which has two processes, is rare.
[63] If high-energy ISR does not exist and the system of electron and positron is symmetric, the CM energy becomes $\sqrt{s} = 2E$ simply.

Former two conditions are imposed to guarantee that reconstructed jets be standard jets. The third condition is imposed to remove events, including high-energetic photons at the very forward region(that is, the beam axis direction). As we can see in the table.11, the radiative return is suppressed by about 99% by imposing these cuts.

5.2.3 Other backgrounds rejection

As explained in section 5.2, di-boson processes WW/ZZ/Zh also should be eliminated. Originally, they are 4-jet events. However, since all events are reconstructed as 2-jet events in this analysis, these jets' shapes tend to be wider. Based on this idea, the event shape variable called Thrust is used in this analysis. Thrust is defined like this:

$$T = \max_{|\boldsymbol{n}|} \frac{\sum_i^N |\boldsymbol{p}_i \cdot \boldsymbol{n}|}{\sum_i^N p_i}$$

Here, $\boldsymbol{n}$ is the unit vector of thrust axis, and $\max_{\boldsymbol{n}}$ means that choosing the direction which gives the maximum value of $\sum_i^N |\boldsymbol{p}_i \cdot \boldsymbol{n}| / \sum_i^N p_i$. As the figure.30, thrust is one of the indexes which means how sharp the jet shape is.

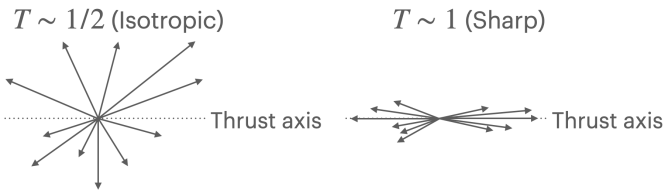


Figure 30: The conceptual figure of the variable Thrust

Distributions of thrust for signal and background events are shown in the figure.31 and the figure32. According to them, we can see that di-boson events have a lower value than signal events. In this analysis, the condition Thrust > 0.85 is imposed to cut ZZ/WW/Zh backgrounds. The table.11 shows the performance of this cut. As a remarkable point of this table, WW contamination remains relatively larger for the left polarization.

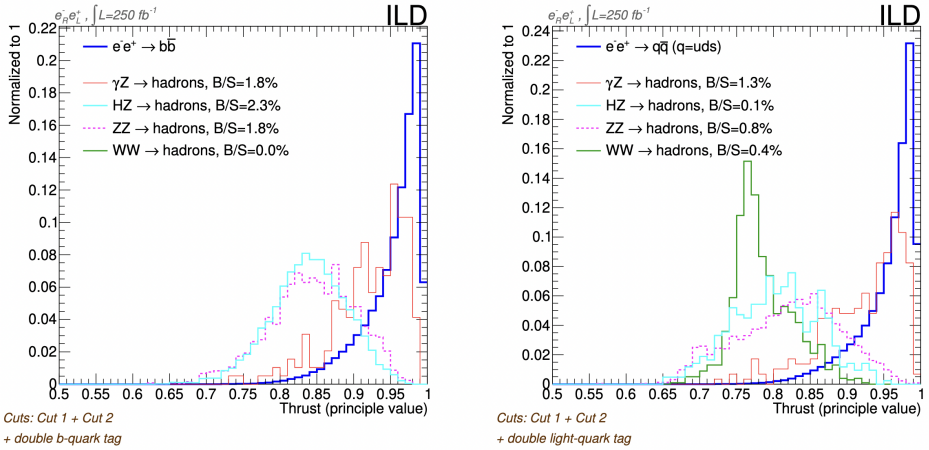


Figure 31: Thrust distributions of signal and backgrounds for the left polarization $e_L^- e_R^+$[14]

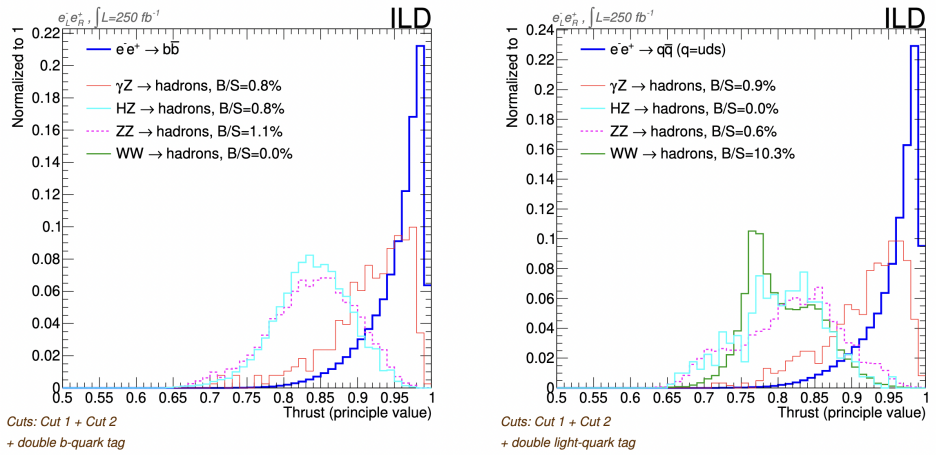


Figure 32: Thrust distributions of signal and backgrounds for the right polarization $e_R^- e_L^+$[14]

5.2.4 Selection by Flavor-Tagging

After the rejection of background events, we need to identify b and uds quarks each other, as explained in section4.3.6. The Flavor-Tagging algorithm is used, and it outputs the likelihood, which means how likely it looks like b or c quark. In general, the likelihood takes its value in the range of 0 to 1, and a high b-likelihood means it seems to be a b-ish event. So if we choose high b-likelihood events, we can extract clean b-like events. On the other hand, if we choose both low likelihoods for b and c, we can extract uds-like events. The figure.34 and the figure.33 show likelihood distributions. We can see that the distribution of c-like events has a relaxed tail. It is because c quark is difficult to identify compare with b quark identification. Since c hadrons have a shorter(about half) lifetime than b hadrons, it is not easy to detect its secondary vertex. Because of this reason, the performance of c-tagging tends to be worse than b-tagging[64].

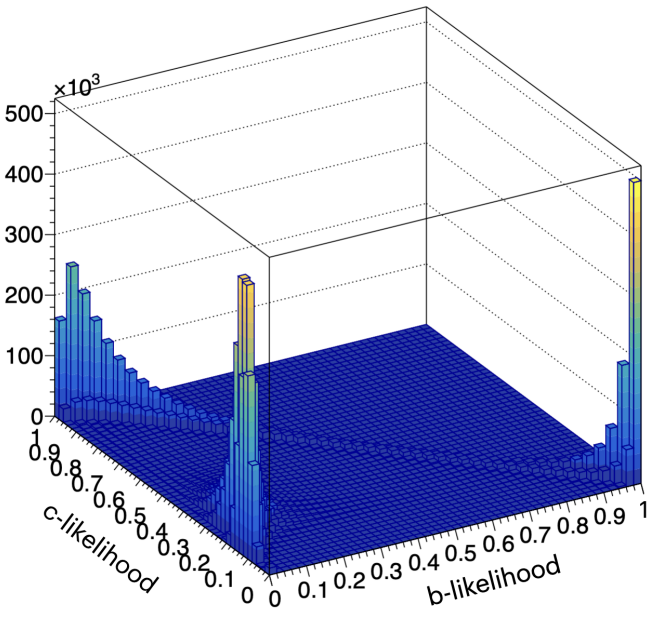
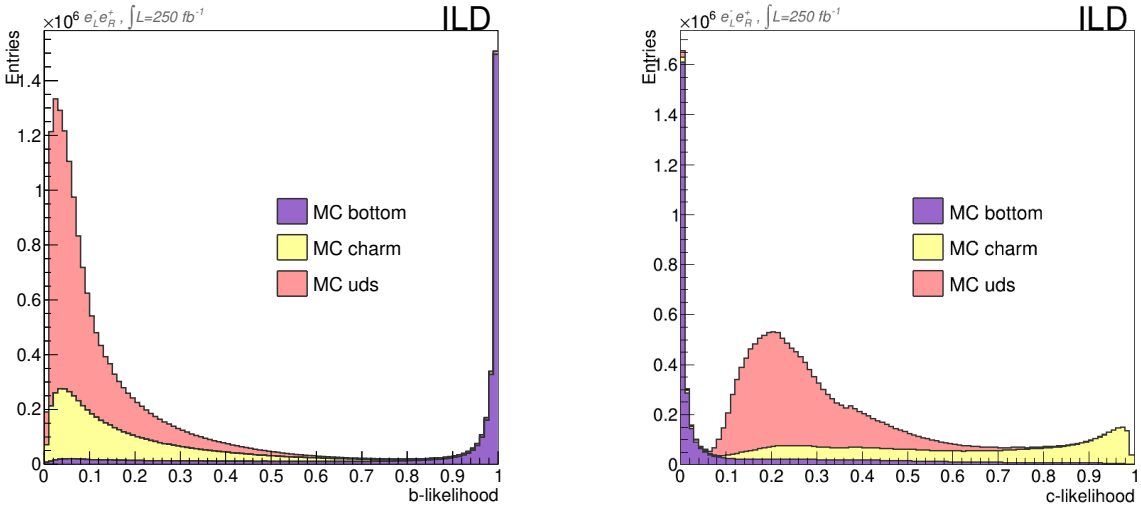


Figure 33: The 2D likelihood distribution

[64]It is a reason for removing c quark in the definition of R_3^{bl}.



(a) b-likelihood distribution

(b) c-likelihood distribution

Figure 34: The 1D likelihood distributions with respect to each MC truth flavor : We can see that $b\bar{b}$ sample concentrates around b-likelihood=1(higher b-likelihood is more b-ish.). Whereas, $b\bar{b}$ sample concentrates around c-likelihood=0(higher c-likelihood is more c-ish.).

Based on these distributions, we chose b-likelihood>0.8 as the cut for b-like events extraction, and b-likelihood<0.4 and c-likelihood<0.25 as the cut for uds-like events extraction. Each efficiency and each purity under these cuts are calculated as the following table.10. Here, the efficiency(for example, b quark case) is defined as

$$\epsilon_b \equiv \frac{N(b-tagged\ events\ in\ MC\ b)}{N(MC\ b\ events)}, \tag{5.4}$$

and the purity is defined as

$$\mathcal{P}_b \equiv \frac{N(b-tagged\ events\ in\ MC\ b)}{N(all\ b-tagged\ events)}. \tag{5.5}$$

The efficiency ϵ_b means how well MC b events are identified as b events by b-tagging. On the other hand, the purity $\mathcal{P}_b$ means that how many MC b events are in tagged b events. According to the numbers of the table.10, we can see that the performance of the flavor-tagging at ILD is superior to LEP's time.

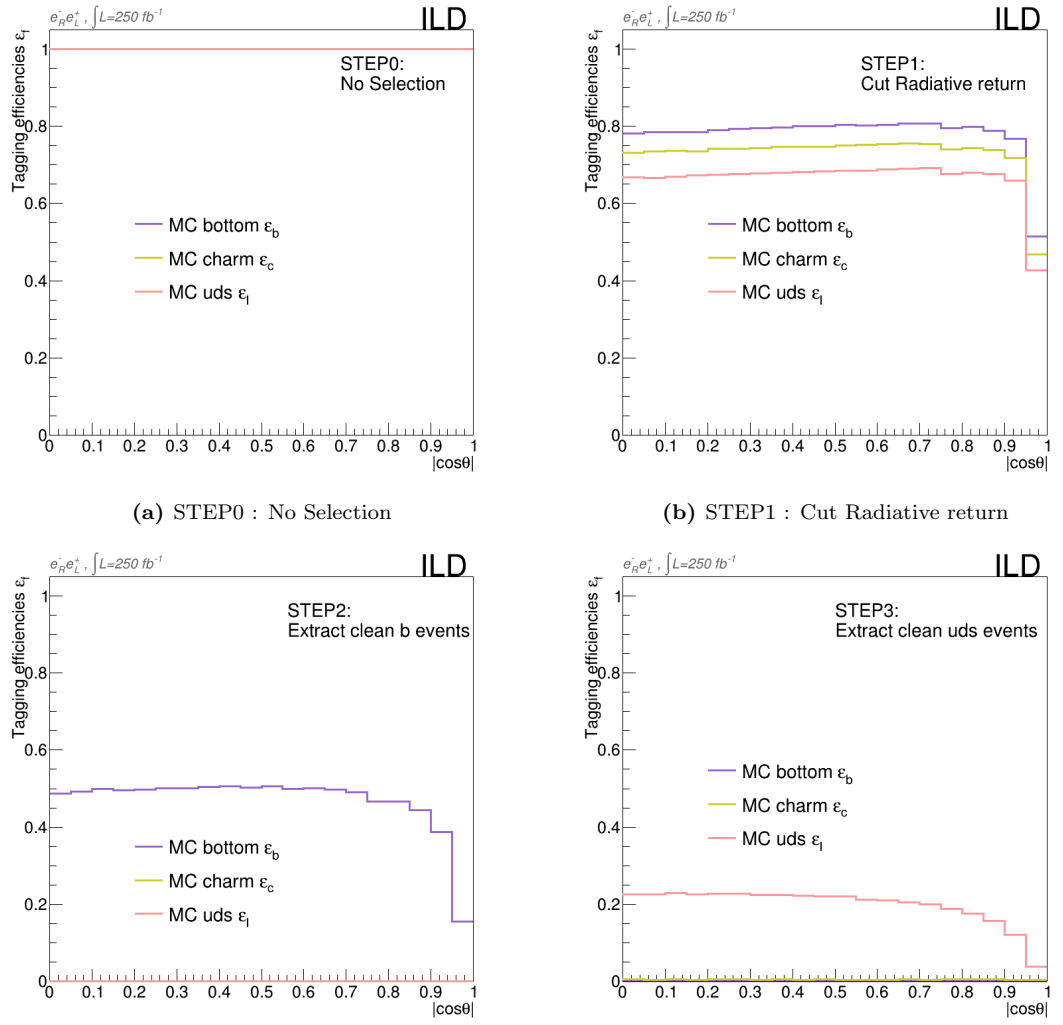
Experiment	b-tagging Efficiency[%]	b-tagging Purity[%]	uds-tagging Efficiency[%]	uds-tagging Purity[%]
ILD(this study)	80	98.7	58	96.1
DELPHI	47	86	51	82

Table 10: Flavor-Tagging efficiencies and purities : We can see that the performance of ILD is superior to LEP. It comes from the better performance of the Vertex detector of ILD, like the table.6 shows.

Figure.35 shows that the efficiency as the function of polar angle θ of reconstructed jets for each selection step. Each step means as below :

- STEP0 : Any selection conditions are not used but use only MC truth information of quark flavor and ISR energy. For MC truth level, we define the radiative return event as the event which has the ISR energy larger than 50GeV.

- STEP1 : Impose the radiative return cut conditions as explained in the previous section. Here, also MC information is used.
- STEP2 : Add b-like events extraction condition to STEP1; b-likelihood>0.8.
- STEP3 : Switch b to uds-like events extraction condition; b-likelihood<0.8 and c-likelihood<0.25.



(a) STEP0 : No Selection

(b) STEP1 : Cut Radiative return

(c) STEP2 : Extract clean b events

(d) STEP3 : Extract clean uds events

Figure 35: Polar angle distributions of Flavor-tagging efficiency as the function of polar angle : Each efficiency is calculated with MC truth information of quark flavor and ISR energy.

Until STEP1, since we do not have any flavor-tagged events, each efficiency's behaviors should be similar, and we can check it from the upper two plots. In STEP2, since b-like events extraction is imposed, the efficiency of $b\bar{b}$ sample should be higher, and others should be almost 0. Whereas, in STEP3, since we switch to uds-like events extraction, the efficiency of $l\bar{l}$ sample ($l = u, d, s$) should be larger, and others should be almost 0. According to the lower two plots, we can check them. Additionally, we have one more remarkable point. In STEP2 and 3, the efficiency becomes smaller in the super-front direction(about $|\cos\theta| > 0.8$). We should add it as a selection condition and cut the event at a super-front angle, which the efficiency becomes low.

5.3 Summary of Selection conditions

Summarize selection conditions in this analysis :

- For backgrounds rejection
  - $K_\gamma < 50$GeV
  - Invariant mass of all jets should be greater than 130GeV
  - Both jets should contain more than 5 neutral PFOs
  - None of the jets should contain a neutral PFO with $E > 50$GeV in the very forward region $|\cos\theta| > 0.98$
  - Thrust > 0.85
- For tagging selection
  - b-like events extraction : b-likelihood>0.8 for both jets
  - uds-like events extraction : b-likelihood<0.4 and c-likelihood<0.25 for both jets
  - Super-front angle events cut : $|\cos\theta| < 0.8$

Under such selections, we can confirm its performance as the signal efficiency and the B/S. Signal efficiency and B/S are defined for each flavor like this :

$$\varepsilon_b^{sig} = \frac{N(STEP2)}{N(STEP0)} \quad , \quad \varepsilon_l^{sig} = \frac{N(STEP3)}{N(STEP0)} \tag{5.6}$$

$$B/S_b = \frac{N(background\ at\ STEP2)}{N(signal\ at\ STEP2)} \quad , \quad B/S_l = \frac{N(background\ at\ STEP3)}{N(signal\ at\ STEP3)}. \tag{5.7}$$

STEP means that the selection step which as explained in the previous section. The signal efficiency ε_q^{sig} means that how many signal events are chosen by the event selection(bigger is better). On the other hand, the B/S means how many background events are remained compare with the selected signal events after imposing the event selection(smaller is better). It should be noted that MC truth information of quark flavor and ISR energy are used in these definitions. We can see that the biggest contamination of background events comes from the WW process for the left polarization. The next section explains measured result of the observable R_3^{bl} under this event selection.

	Signal efficiency[%]	B/S Radiative return[%]	B/S WW[%]	B/S ZZ[%]	B/S Zh[%]
Left polarization $e_L^- e_R^+$					
b quark	35.8	1.0	0.0	0.5	1.4
uds quark	15.6	1.1	2.1	0.2	0.0
Right polarization $e_R^- e_L^+$					
b quark	35.2	1.1	0.0	0.9	1.0
uds quark	15.9	1.0	0.0	0.3	0.1

Table 11: Signal efficiency and B/S of the event selection : Radiative return events are suppressed up to 99% by imposing the event selection. WW contamination for the left polarization is the largest contribution of background events.

5.4 Estimation of the b mass precision at 250GeV

5.4.1 The strategy of b mass determination through R_3^{bl}

Based on the selection condition which is explained in 5.3, the observable R_3^{bl} is measured. In this study, the MC samples of $e^+e^- \to q\overline{q}$ at 250GeV with $250fb^{-1}$ for both polarization configurations(100% left polarized and 100% right polarized)[65] are used in this analysis. In this analysis, R_3^{bl} is considered by 3 levels with respect to each decay evolution levels;

1. R_3^{bl} which includes up to Parton effect
2. R_3^{bl} which includes up to stable hadrons effect (after hadronization)
3. R_3^{bl} which includes up to detector effects (particle reconstruction, jet reconstruction, flavor-tagging)

The first two levels, parton level and stable hadrons level are calculated from MC generated samples. Detector reconstructed level is obtained from MC information after detector reconstruction. What we can see in actual experiments is the reconstructed level result. It is the parton level result that can compare with QCD theory directly, and there are corrections between parton and reconstructed levels like this :

$$R_3^{bl\ Par} = C^{had}C^{det}R_3^{bl\ Rec}. \tag{5.8}$$

$$C^{had} \equiv \frac{R_3^{bl\ Par}}{R_3^{bl\ Had}} \quad , \quad C^{det} \equiv \frac{R_3^{bl\ Had}}{R_3^{bl\ Rec}} \tag{5.9}$$

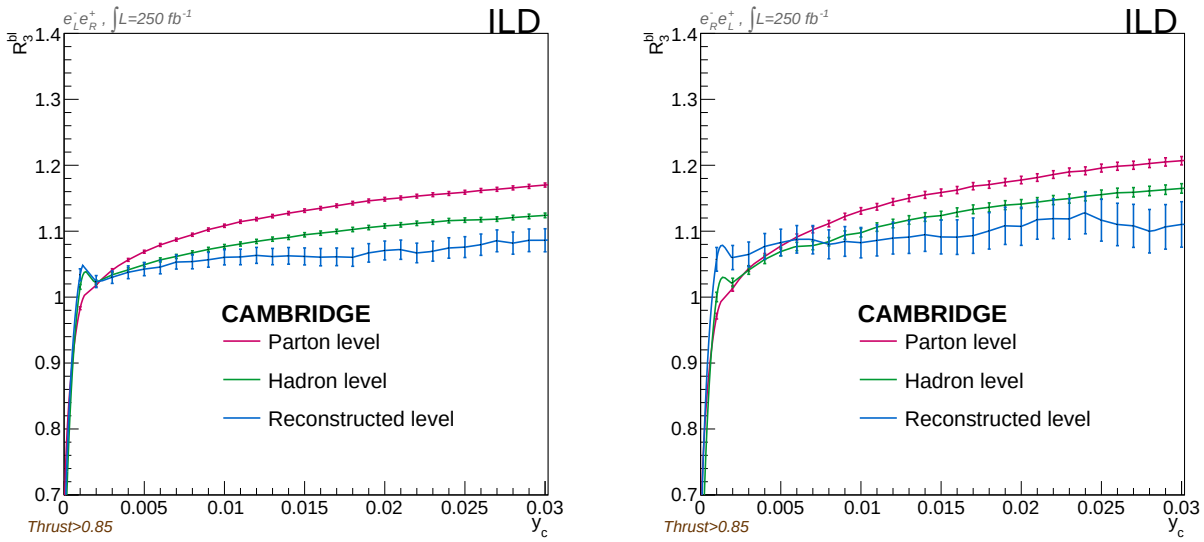
The two types of corrections fill the gaps of hadronization and detector contribution between parton level and reconstructed level. C^{had} is reflected hadronization model dependence(the correction between parton level and hadron level), C^{det} is reflected detector effects such as event selection(the correction between hadron level and reconstructed level).

Let us consider the detail of how to extract b quark mass from the measured reconstructed level result. When ILC runs, we will yield the reconstructed level result $R_3^{bl\ mes}$ which comes from raw data. In this study, MC simulated result is used instead of it. As described above, the reconstructed result is corrected to the parton level result by multiplying correction factors C^{had} and C^{det}. Since parton and stable hadrons levels are obtained as MC generated levels in this study, we can know the size of correction factors from MC simulation. In the actual experiment, we will quote same MC simulated(that is SM based) correction factors, and correct the measured result $R_3^{bl\ Rec}$ to the parton level result $R_3^{bl\ Par}$; $R_3^{bl\ Par} = C^{had}C^{det}R_3^{bl\ Rec}$. When we compare $R_3^{bl\ Par}$ to the theoretical calculation $R_3^{bl\ theo}$ which is provided by the SM, $R_3^{bl\ theo}$ should be calculated under the same perturbative order of the configuration in event generator(in this case, WHIZARD). Since $R_3^{bl\ theo}$ is obtained as the function of b quark mass just as figure.13, the b quark mass value is extracted by overlapping with $R_3^{bl\ Par}$ and $R_3^{bl\ theo}(m_b)$. When this measurement runs at the actual ILC experiment, $R_3^{bl\ Rec}$ may include new physics effects and extracted b quark mass may deviate from the SM expected mass. If we find the significant deviation there, it is a probe of new physics. It is the strategy of b quark mass determination using R_3^{bl}.

[65] You can see more detail of these configurations in section3.1.2

5.4.2 Measurement of Observable R_3^{bl}

The observable R_3^{bl} depends on the 3-jet event's number, as eq.(2.6) shows. As explained in section 4.3 and section5.2.1, the jet resolution parameter y_c specifies the size of the jet, and we can define the 3-jet event experimentally by choosing one y_c number. As mentioned in section 5.2.1, we are focusing on $y_c = 0.01$. So R_3^{bl} depends on y_c and we can check this behavior in figure.2.6. The left plot is the result for 100% left polarized configuration, and the right plot is the result for 100% right polarized configuration. There are three colored lines in figure.36. The red line is the result, which includes up to parton effect, the green line includes up to stable hadrons effect. These two results are obtained by clustering simulated particles(that is, parton or stable hadrons) from PYTHIA by the Cambridge algorithm. The remained blue line is the reconstructed level result after the event selection. It includes up to detector effects.



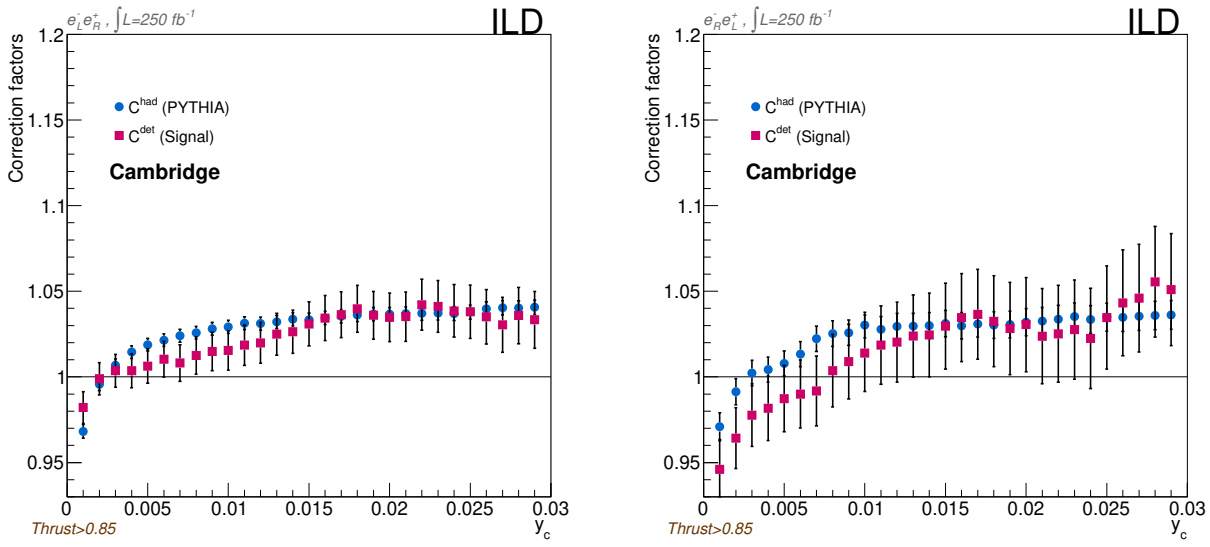
(a) Pure left polarization $e^-_L e^+_R$

(b) Pure right polarization $e^-_R e^+_L$

Figure 36: R_3^{bl} as the function of the jet resolution parametery_c : R_3^{bl} depends on y_c because y_c specifies the definition of 3-jet event and we count jet event numbers due to measure R_3^{bl}. Error bars reflect only each statistical error which is calculated by eq.(5.10).

It should be noted again that the parton level result(red) and the reconstructed level result(blue) are connected by correction factors that come from the hadronization model and detector effects as described eq.(5.8). Figure.37 shows that y_c dependence of correction factors which come from hadronization model(blue) and detector effects(red) for each polarization configuration[66]. We can check that each size of correction is close to 1. Namely, it turns out that the reconstructed level result can be regarded as the parton level result almost. Based on it, let us move to the consideration of the result.

[66] It should be reminded that we're focusing $y_c = 0.01$.



(a) for Left polarization

(b) for Right polarization

Figure 37: Correction factors: The left one gives each correction factor which comes from hadronization(red) and detector effects(blue) for only $q\bar{q}$ signal.

5.4.3 Important point of Monte Carlo $q\bar{q}$ sample

If we remind that b quark tends to emit less gluon than light quarks as explained in section2.3.1, R_3^{bl} should be smaller than 1. However, obtained results are larger than 1 for all levels, as the figure.36 shows. Why such a difference appear? When we examine our $q\bar{q}$ MC sample, it seems that there is a problem in handling the b quark mass of the current MC sample. For example, let us consider the case that the b quark mass is set to zero. Since b quark is massless just as uds quarks, its 3-jet ratio becomes similar level to uds, and R_3^{bl} becomes ~ 1. Our simulated R_3^{bl}s exceed 1 for all levels, so it seems that there is some problem in handling the b quark mass of the current MC sample.

Additionally, it turned out that the statistics of 3-jet events for b and uds were about ten times less than assumed numbers based on the theoretical(SM) calculation. The statistical error of R_3^{bl} is calculated by like this :

$$\frac{\Delta R_3^{bl}}{R_3^{bl}} = \frac{1}{\sqrt{N_b}} + \frac{1}{\sqrt{N_{3b}}} + \frac{1}{\sqrt{N_l}} + \frac{1}{\sqrt{N_{3l}}} \tag{5.10}$$

Namely, statistics of 3-jet events are essential for estimation of the statistical error on R_3^{bl}. However, because of the above problem, the statistical error estimated by the current MC sample will be unreliable, just as the center value of R_3^{bl}. Currently, the firm reason for this problem has not been found yet, and we are discussing with theoreticians the detail of QCD correction of gluon radiation.

As some papers[12] [8] [34] mentioned, the calculation result at LO(Leading Order) is affected larger by the definition of quark mass(e.g. pole mass or running mass) than NLO(Next-to-Leading Order) calculation. The current MC sample is output by LO calculation with massless quarks. Therefore, the MC sample should be calculated at NLO with massive quark mass(that is, normal b quark mass), and it is updating now by WHIZARD authors.

Then, which are results reliable, and what should we assume in this study? Center value

of R_3^{bl} and its statistical error are unreliable as long as the current MC sample is used, but correction factors C^{had} and C^{det} which are defined by these ratios are reliable regardless the current problem of MC sample. Here, we believe that the ratio can cancel the problem of R_3^{bl}s each other. Since systematic errors of detector and hadronization model appear on C^{had} and C^{det} themselves, there is no problem to conclude b quark mass precision at 250GeV ILC.

5.4.4 Estimation of Statistical error

As described in the previous section, the statistical error of R_3^{bl} is estimated using obtained statistics of each jet event. However, 3-jet event's numbers of the current MC sample are not simulated well[67]. Originally, obtained statistics after event selection are substituted to N_{3q} and N_q in eq.(5.10). However, unfortunately, the statistical error that is estimated by this way will not be reliable because of MC's problem. To estimate more reliable statistical error, we use assumed center values of R_3^{bl} and b quark mass and estimate the statistical error under assumed 3-jet event numbers. This study assumed that the 3-jet event's number is 30% of all jet event's number. Let us assume the central value of R_3^{bl} as 0.996 according to NLO theoretical calculation[14]. It corresponds to 2.75 of the center value of b quark mass. Since all jet event's number(or its cross-section) is simulated well in the current MC, the following numbers of the cross-section are used to estimate all jet event's numbers :

	100% Left polarization	100% Right polarization
σ_{bb}[fb]	5970	1352
σ_{ll}[fb]	19347	5920

Table 12: LO cross section for $e^+e^- \to q\bar{q}$ ($q = uds\ or\ b$) at 250GeV : They are obtained from the current MC, and they are the same as the table.8. It should be noted that statistics of 3-jet events are unreliable, but whole statistics of quark pair production are simulated well even if at LO. So, this study uses them as MC numbers. The numbers of 3-jet events are assumed 30% of them to obtain a more reliable result.

It should be noted that statistics of 3-jet events themselves are unreliable, as mentioned above, but numbers of the cross-section of quark pair production are simulated well, and we can use only them as MC numbers. In this study, let us estimate statistical error for actual ILC running program H20. Polarization configuration and integrated luminosity of H20 scenario are given by second row(250GeV(update)) in table.15. The flow of estimation is complicated a little, so let us see it step by step.

- Rewrite as the cross-sections for (-80%,+30%) and (+80%, -30%)

  At first, we need to translate the above cross-sections for $(P_{e^-}, P_{e^+}) = (-100\%, +100\%)$ and $(+100\%, -100\%)$ to ones for $(-80\%, +30\%)$ and $(+80\%, -30\%)$. In general, the cross-section of (P_{e^-}, P_{e^+}) is given like this [33][21] :

$$\begin{aligned}\sigma(P_{e^-}, P_{e^+}) = \frac{1}{4}\Big[&(1+P_{e^-})(1+P_{e^+})\sigma_{RR} + (1-P_{e^-})(1-P_{e^+})\sigma_{LL} \\ &+ (1-P_{e^-})(1+P_{e^+})\sigma_{LR} + (1+P_{e^-})(1-P_{e^+})\sigma_{RL}\Big].\end{aligned} \tag{5.11}$$

[67] In general, since statistics of 3-jet events N_{3b} and N_{3l} are smaller than ones of all jet events N_b and N_l, the statistical error eq.(5.10) is affected larger by statistics of 3-jet events than ones of all jet events.

Since $P_{e^\mp}$ is defined by eq.(3.1), $1+P_{e^\mp}$ extracts right components and $1-P_{e^\mp}$ extracts left components. Here, σ_{LR} means the cross-section for 100% left polarization$(-100\%, +100\%)$, and σ_{LL} means the cross-section for $(-100\%, -100\%)$. However, in the case of both electron and positron are left(or right) polarized, they can not create spin-1 particle by angular momentum conservation. So, σ_{LL} and σ_{RR} are zero completely in this study. As a result, the cross-section for (P_{e^-}, P_{e^+}) is

$$\sigma\left(P_{e^-}, P_{e^+}\right) = \frac{1}{4}\Bigg[\left(1-P_{e^-}\right)\left(1+P_{e^+}\right)\sigma_{LR} + \left(1+P_{e^-}\right)\left(1-P_{e^+}\right)\sigma_{RL}\Bigg]. \tag{5.12}$$

In the case of $|P_{e^-}|= 80\%$ and $|P_{e^+}|= 30\%$, the cross-section becomes like this :

$$\mathcal{O}\left(-,+\right) = 0.58\mathcal{O}_L + 0.035\mathcal{O}_R \tag{5.13}$$

$$\mathcal{O}\left(+,-\right) = 0.58\mathcal{O}_R + 0.035\mathcal{O}_L. \tag{5.14}$$

Here, $\mathcal{O}_L$ corresponds to each number of cross section for left polarization in table.12 and $\mathcal{O}_R$ is its right polarization version.

- Obtain the ideal statistics of all jet events

  After that, we can obtain statistics of $e^+e^- \to q\overline{q}$ process by multiplying the integrated luminosity $\mathcal{L} = 900\text{fb}^{-1}$.

- Obtain the expected statistics of all jet events

  However, we can not substitute them to eq.(5.10) directly because these numbers are not taken into account detector efficiency. So we should scale them to expected statistics of $q\overline{q}$ process by signal efficiency in table.11; 0.38 for b quark and 0.16 for uds quarks.

- Obtain the expected statistics of 3-jet events under the assumption of $R_3^{b(l)} = 0.3$

  Finally, expected statistics of 3-jet events can be estimated from them. We should introduce the assumption to avoid the current MC's unreliable statistics and obtain reliable statistics of 3-jet events. Let us assume that the numbers of 3-jet events are 0.3 times of ones all jet events :

  $$N_{3b} = 0.3 \times N_b \ , \ N_{3l} = 0.3 \times N_l$$

  It is equivalent with assuming that 3-jet ratios R_3^b and R_3^l are 0.3.

Under such procedures, statistics of all jet events N_q and statistics of 3-jet events N_{3q} are estimated. Table.13 shows the result of them. Here, we assume that errors of the jet event number N is given by Poisson distribution : $\delta N = \sqrt{N}$.

	$(-80\%, +30\%)$ polarization	$(+80\%, -30\%)$ polarization
all b-jet events	1,200,393	339,644
3b-jet events	360,118	101,893
all uds-jet events	1,645,698	591,947
3uds-jet events	495,692	178,297
$\Delta R_3^{bl}/R_3^{bl}$	0.0025	0.0045

Table 13: Results of R_3^{bl} for H20 scenario of 250GeV ILC : Each statistics is expected number which is taken into account the effect of detector efficiency.

For $(+,+)$ and $(-,-)$, statistics are several % of ones for $(-,+)$ and $(+,-)$. So, in this paper, $(+,+)$ and $(-,-)$ are neglected.

We should estimate each error(statistical error, systematic errors) independently, and the statistical error appear on $R_3^{bl\ Rec}$ in eq.(5.8). Other systematic error will appear on correction factors C^{had} and C^{det}. Namely, we can use it as the statistical error of parton level. Here, the statistical error on R_3^{bl} is translated to the statistical error on b quark mass through eq.(2.8) under the assumption of $R_3^{bl} = 0.996$ and $m_b(250\text{GeV}) = 2.75\text{GeV}$[68] :

$$\Delta m_b(stat.) \sim \frac{\Delta R_3^{bl}(stat.)}{2(1-R_3^{bl})} m_b \sim 0.85\text{GeV}\ for\ (-80\%, +30\%)$$
$$\sim 1.53\text{GeV}\ for\ (+80\%, -30\%)$$

It should be noted again that our $q\overline{q}$ sample has some problems on statistics of 3-jet events, but this statistical error is more reliable because it is estimated with the assumption which uses theoretical assumed 3-jet event ratios.

5.4.5 Estimation of Systematic errors

Let us estimate systematic errors on the observable in this section. It should be noted that what we should estimate at final is the systematic uncertainties on observable at the parton level $R_3^{bl\ Par}$. There are two sources of systematic errors :

- potential uncertainty in the Hadronization model
- potential uncertainty in Detector contributions
  - come from Flavor-Tagging performance
  - come from Signal selection performance
  - come from Backgrounds contamination

The hadronization error appear on C^{had}, and the detector error appear on C^{det}. $R_3^{bl\ Par}$ will inherit these uncertainties through C^{had} and C^{det} by eq.(5.8). So we should estimate them independently.

When we estimate the hadronization error, we need to consider the fluctuation of C^{had} by taking the variation of different models or different parameters. However, the raw data of 250GeV ILC is essential for tuning of models and parameters. So, this study estimates the uncertainty from the hadronization model by extrapolation of the LEP result. According to LEP result(Figure.2 in [18]), its uncertainties is given by $\Delta C^{had}/C^{had} \sim 0.2\%$[69]. If 250GeV ILC

[68] The total error of the parton level $R_3^{bl} = C^{had}C^{det}R_3^{bl\ Rec}$ is calculated by like this :

$$\Delta R_3^{bl} = \sqrt{\left(\frac{\partial R_3^{bl}}{\partial R_3^{bl\ Rec}}\right)^2 (\Delta R_3^{bl\ Rec})^2 + \left(\frac{\partial R_3^{bl}}{\partial C^{had}}\right)^2 (\Delta C^{had})^2 + \left(\frac{\partial R_3^{bl}}{\partial C^{det}}\right)^2 (\Delta C^{det})^2}.$$

The first term corresponds to the statistical error $\Delta R_3^{bl}(stat.) = C^{had}C^{det}\Delta R_3^{bl\ Rec}$. The second term and third term are systematic errors $\Delta R_3^{bl}(had.) = \Delta C^{had} \cdot C^{det} R_3^{bl\ Rec}$, $\Delta R_3^{bl}(exp.) = C^{had}\Delta C^{det} \cdot R_3^{bl\ Rec}$. The later two errors are estimated in the next section. It should be noted that correction factors are small(See figure.37). Additionally, other errors(that is, systematic errors) are neglected here. So, the statistical error of the parton level becomes like this :

$$\Delta R_3^{bl}(stat.) = C^{had} \cdot C^{det} \cdot \Delta R_3^{bl\ Rec} \sim \Delta R_3^{bl\ Rec}.$$

This $\Delta R_3^{bl\ Rec}$ is given by the table.13.

[69] The potentially associated error in the hadronization model is considering here. The effect of hadronization and its error on $R_3^{bl\ Par}$ will be appeared as C^{had} which means hadronization correction from measured $R_3^{bl\ Det}$. So we should look C^{had} to estimate the hadronization model uncertainty.

runs, we can expect that jets' misidentification is decreased because the jet's shape becomes sharper than LEP's time, thanks to the higher CM energy. Furthermore, the understanding of hadronization itself also will be developed from LEP's time. Because of such reasons, the hadronization model error can be expected to be smaller. So in this ILC case, let us assume that the uncertainty on the hadronization model is half of the LEP result; $\Delta C^{had}/C^{had} \sim 0.1\%$[70]. As mentioned above, each systematic error should be estimated independently. Therefore, the systematic error from the hadronization model on $R_3^{bl\ Par}$ is 0.1%. It means that the hadronization uncertainty on b quark mass is[71]

$$\Delta m_b(had.) \sim \frac{\Delta R_3^{bl}(had.)}{2(1-R_3^{bl})} m_b \sim \frac{\Delta C^{had} \cdot R_3^{bl}}{2(1-R_3^{bl})} m_b \sim 0.34\text{GeV}.$$

Here, m_b(250GeV)=2.75GeV and $R_3^{bl} = 0.996$ are used as assumptions just as above.

Let us move to detector uncertainties. They appear in C^{det}. Each element (Event selection, Flavor-Tagging) have uncertainties for its performance, and these uncertainties are propagated to the observable at reconstructed level R_3^{bl}. Our strategy is like this: mis-tagging, and backgrounds contamination contributes to observed jet event's number. So, consider $R_3^{bl\ Rec}$ which takes into account each contribution of mistagging(failure of Flavor-tagging), backgrounds contaminations(failure of background rejections), and estimate the uncertainty from each element on $R_3^{bl\ Rec}$. If we assume that other errors are negligible and statistics is infinity, the error of $R_3^{bl\ Rec}$ is almost the same as the error of C^{det}(See eq.(37)). Let us define the reconstructed level result R_3^q which includes effects of signal selection, and mistagging and background contaminations :

$$\begin{aligned} R_3^q(y_c) &= \frac{N_{3q-jet}(y_c)}{N_{all\ q-jet}} \\ &= \frac{\varepsilon^{sig}\left(\epsilon_q^2 \sigma_{q\overline{q}\to 3jets}(y_c) + \sum_{q'\neq q} \tilde{\epsilon}_{q'}^2 \sigma_{q'\overline{q}'\to 3jets}(y_c)\right) + \epsilon_{bkg}\sigma_{bkg\to 3jets}(y_c)}{\varepsilon^{sig}\left(\epsilon_q^2 \sigma_{q\overline{q}} + \sum_{q'\neq q} \tilde{\epsilon}_{q'}^2 \sigma_{q'\overline{q}'}\right) + \epsilon_{bkg}\sigma_{bkg}}. \end{aligned} \tag{5.15}$$

ε^{sig} is the selection efficiency without Flavor-Tagging, ϵ_q is the (single) q-tagging efficiency and $\tilde{\epsilon}_q$ is the mis-q tagging efficiency for $q\overline{q}$ sample. $\sigma_{q\overline{q}}$ is the number of all q-jet events and $\sigma_{q\overline{q}\to 3jets}(y_c)$ is the number of 3 q-jet events of $q\overline{q}$ sample. ϵ_{bkg} is the backgrounds rejection efficiency, and σ_{bkg} is the contamination to the number of all q-jet events and $\sigma_{bkg\to 3jets}$ is the contamination to the number of 3 q-jet events. Although the radiative return cut efficiency $\epsilon^{rad\ cut}$ is linear, why are (mis-)tagging efficiencies quadratic? It is because ϵ_q is the single q-tagging efficiency, which means tagging for a one-side jet. Since we have two jets by Jet-Clustering algorithm, ϵ_q should be considered for each jet. So tagging efficiencies should be

[70] This extrapolation is a rough estimation. However, it will no doubt that the hadronization error is improved in the actual experiment.

[71] Other errors(statistical error and detector error) are neglected here, and the center value of correction factors are close to 1(See the figure.37). So, the hadronization error of the parton level becomes like this :

$$\Delta R_3^{bl}(had.) = \Delta C^{had} \cdot C^{det} \cdot R_3^{bl\ Rec} \sim \Delta C^{had} \cdot R_3^{bl}.$$

contributed as quadratic.

The contribution of ϵ^{sig} means that how well the event selection can collect the events that should be identified as the signal. ϵ_q means that how well the flavor-tagging algorithm can identify quark flavor, and $\tilde{\epsilon_q}$ means that how well misidentified flavor events are contributed. ϵ_{bkg} means how many background events, which should be removed, are selected as signal events. Of course, they depend on performances of detector and selection algorithms, so they should be taken into account as a detector effect.

Then, let us estimate detector uncertainties on R_3^{bl}. To estimate them, this study operated 1.0×10^7 times Toy MC experiment. Its idea is like this :

1. Generate random numbers by Gaussian which its center value is the calculated efficiency, and the width is the assumed accuracy of flavor-tagging or selection.

2. Regard this fluctuation as potentially uncertainties on flavor-tagging or selection. Calculate R_3^b and R_3^l according to eq.(5.15) with each efficiency which includes random fluctuations. To estimate propagated uncertainties on R_3^q, the histogram of R_3^q is made, and its width of this distribution is used as the error on it.

3. Construct R_3^{bl} from calculated R_3^q and estimate propagated uncertainties on R_3^{bl} from histogram just as above.

Table.14 gives the results of detector uncertainties on C^{det}. As we have seen, the observable R_3^{bl} is the fraction between 3-jet ratios of b and uds quarks. Since uncertainties which common between jet's number and flavors are cancelled thanks to this definition, the detector error on C^{det} tends to be smaller. We can check it from the table.14.

	Uncertainties (assumed)	Uncertainties on C^{det} (100% left polarization)	Uncertainties on C^{det} (100% right polarization)
Flavor-Tagging	0.1-0.5%	0.07%	0.06%
Signal selection	0.5-1.0%	0.06%	0.06%
BKGs contamination	1.0%	0.20%	0.10%
Total	-	0.22%	0.13%

Table 14: Systematic errors which come from detector effects : There are three elements; Flavor-Tagging and signal selection, and backgrounds contamination.

The notable point is the effect that comes from di-boson events rejection is dominant. The table.11 shows that the WW contamination remains relatively larger after imposing the event selection. It is the reason for the above point. In this analysis, di-boson backgrounds are suppressed by the cut of thrust. Since 3-jet events are reconstructed as 2-jet events by the exclusive algorithm, such an event's shape tends to be wider, just like the di-boson event's case. So, some 3-jet events are also cut by imposing the cut of thrust. Because of this reason, the stronger condition of thrust that can eliminate more di-boson events should not be used.

As a result, the detector uncertainty is 0.23% for the left polarized sample and 0.21% for the right polarized sample. Therefore, the uncertainty which comes from detector effects on

$R_3^{bl\ Par}$ is 0.22% for the left polarization and 0.13% for the right polarization. Just as in the case of hadronization error, this detector uncertainty on $R_3^{bl\ Par}$ is translated to the error of b quark mass by eq.(2.8) :

$$\Delta m_b(exp.) \sim \frac{\Delta R_3^{bl}(exp.)}{2(1-R_3^{bl})} m_b \sim \frac{\Delta C^{det} \cdot R_3^{bl}}{2(1-R_3^{bl})} m_b \sim 0.75\text{GeV}\ for\ the\ left\ polarization$$
$$\sim 0.44\text{GeV}\ for\ the\ right\ polarization$$

5.4.6 Precision of bottom quark mass determination at 250GeV

Based on above two sections, the accuracy of $R_3^{bl\ Par}$ measurement is given by like this :

$$\frac{\Delta R_3^{bl\ Par}}{R_3^{bl\ Par}}(-,+) = 0.25(stat.) + 0.1(had.) + 0.22(exp.) \sim 0.34\%$$
$$\frac{\Delta R_3^{bl\ Par}}{R_3^{bl\ Par}}(+,-) = 0.45(stat.) + 0.1(had.) + 0.13(exp.) \sim 0.48\%.$$

Here, the detector error for the left polarization is combined to the result for $(-,+)$. It is because $e_L^- e_R^+$ is dominant(about 16 times larger) compare with $e_R^- e_L^+$ for $(-,+)$. Similraly, since $e_R^- e_L^+$ is dominant(about 17 times larger) compare with $e_L^- e_R^+$ for $(+,-)$, the detector error for the right polarization is combined to the result for $(+,-)$.

As we have seen, this result can be translated to the precision of b quark mass at 250GeV through eq.(2.8). The result of b mass measurement precision under the assumption that m_b=2.75GeV can be concluded that like the following :

$$\Delta m_b\,(-,+) = 0.85(stat.) \pm 0.75(exp.) \pm 0.34(had.) \pm 0.07(theo.)\text{GeV} \tag{5.16}$$

$$\Delta m_b\,(+,-) = 1.53(stat.) \pm 0.44(exp.) \pm 0.34(had.) \pm 0.07(theo.)\text{GeV}. \tag{5.17}$$

The 4th error, that is, theoretical error, is estimated from variations on b mass in the different renormalization scales, different mass definitions, and the strong gauge coupling.

These results eq.(5.16) and eq.(5.17) mean that b quark mass precision at 250GeV ILC is about 43% for $(-,+)$ and 59% for $(+,-)$. It does not compete with results at lower energy scales, but it is the result that paves the way for the higher energy scale, which has not to be seen at previous measurements for b quark mass study.

6 Discussion and Summary

6.1 Discussion of b mass study at 250GeV

Based on this result, can we discuss the viability of new physics? Unfortunately, finding the prediction of b quark mass under BSM models is not easy, and the material is not sufficient. As just illustrative examples in order to see deviation between SM and BSM, [35] and [36] predicted b mass evolution up to GUT scale under Minimal Super Symmetry Model(MSSM) or 2 Higgs Doublet Model(2HDM). Figure.38 shows the results of [35]. Each prediction is matched at Z-pole, but we can see that MSSM results deviate from SM prediction at a higher energy scale.

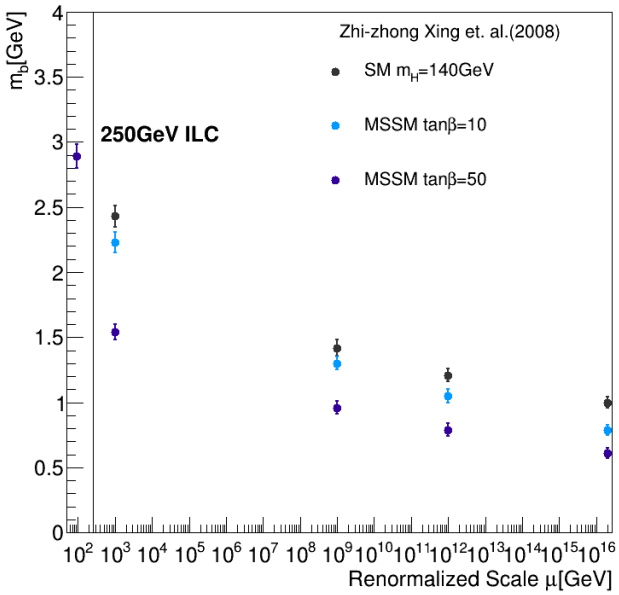


Figure 38: Theoretical prediction of b mass evolution from Z-pole to GUT scale under SM and MSSM : Gray is SM, blue and purple are MSSM predictions with respect to tan β.

The horizontal axis is the energy scale of considering processes, and the vertical axis is running b mass. The gray plot is SM prediction. Blue and purple ones are MSSM predictions with respect to each tan β. The parameter tan β means that the ratio of vacuum expectation values of SUSY Higgs, and larger tan β tends to give lighter SUSY particles. If tan β is larger, SUSY particles tend to be smaller.

We do not have a clear criterion to state a probe of new physics on b mass evolution. To evaluate the viability of BSM through the result of this study, let us assume that if we have a 10% deviation between SM and measured results on b quark mass at 250GeV, this is a probe of BSM. However, unfortunately, as we can see in eq.(5.16) and eq.(5.17), the statistical uncertainty is the largest contribution to b quark mass, and this result is concluded that does not have sufficient power to insist the viability of BSM. Can we improve it? Let us assume 10% as the goal precision which this measurement should aim[72]. To accomplish it, considerable scenarios to improve the precision of b quark mass in this measurement are discussed below.

[72]The accuracy of LEP experiment is about 22%.

- Scenario1 : Improvement at 250GeV

  We can expect that much statistics by the improvement of total luminosity at 250GeV ILC. In the previous chapter, statistical error is estimated under the integrated luminosity of 900fb^{-1}. When the total luminosity X is scaled to αX, improvement of the precision of R_3^{bl} is given by

$$\frac{\Delta R_3^{bl}}{R_3^{bl}}(\alpha X) = \frac{1}{\sqrt{\alpha}} \frac{\Delta R_3^{bl}}{R_3^{bl}}(X). \tag{6.1}$$

  It means that if we want to improve the precision of R_3^{bl} by a factor of $1/\sqrt{\alpha}$, the required total luminosity is αX.

  Currently, 250GeV ILC can extend its pulse rate 5 Hz to 10Hz, and thanks to this upgrade, the instantaneous luminosity at 250GeV is increased by 2 times. It needs the equivalent cryogenic capacity to 500GeV case. Therefore, if 500GeV ILC can run, the instantaneous luminosity at 250GeV ILC is 2 times increased by running with the pulse rate of 10Hz. Based on this upgrade and eq.(6.1), we can expect that the statistical uncertainty of R_3^{bl} is improved by $\sqrt{2}$ times; $\Delta R_3^{bl}(5Hz) \to \Delta R_3^{bl}/\sqrt{2}(10\text{Hz})$. According to eq.(2.8), the statistical error on b quark mass is also improved by $\sqrt{2}$ times. It means that $\Delta m_b(stat.)$=0.60GeV for $(-,+)$ and 1.08GeV for $(+,-)$. It corresponds to the precision of about % for $(-,+)$ and 44% for $(+,-)$. Certainly, we can confirm the improvement of b quark mass precision by luminosity upgrade at 250GeV ILC.

- Scenario2 : Upgrade of 250GeV to 500GeV/1TeV

  The accuracy on b quark mass measurement at 250GeV ILC is modest, but can we expect better results on 500GeV or 1TeV ILC? As we have seen so far, b mass can not be measured easily at a higher energy scale even if 250GeV ILC as the eq.(2.8) shows. We do not have corresponded results of 500GeV or 1TeV, but it is obvious that this factor becomes worse at these energy scales. If ILC is upgraded to 500GeV/1TeV, statistics and other elements are improved. However, since b quark mass sensitivity is quite mortal, such general advantages are no longer advantages on b mass measurement at high energy. So the b mass measurement at a higher energy scale at ILC also will be more difficult.

- Scenario3 : Using Giga-Z option

  We have seen that the significant improvement up to the sufficient precision($\sim$10GeV) which gives the probe of new physics can not be expected for this measurement at 250GeV. Is this measurement method of b quark mass at ILC stalemate? No, we can expect a hopeful result in Giga-Z ILC[39]. Giga-Z ILC is a running option of ILC, which creates a large number of Z bosons($\sim 10^9$) by setting the CM energy Z-pole. ILC can measure various important parameters, such as the Weinberg angle, with higher precision than previous measurements. Since the branching ratio of $Z \to hadronic$ is about 70%, let us assume that we can obtain 7.0×10^8 $Z \to hadronic$ events in Giga-Z ILC. It is 10^2 times statistics of LEP data. Based on obtained jet event numbers N_{3q}, N_q in this study, the statistical error of R_3^{bl} can be written by

$$\frac{\Delta R_3^{bl}}{R_3^{bl}} = \frac{1}{\sqrt{N_b}} + \frac{1}{\sqrt{N_{3b}}} + \frac{1}{\sqrt{N_l}} + \frac{1}{\sqrt{N_{3l}}}. \tag{6.2}$$

So, if we assume that each observable becomes 10^2 times larger, the statistical error is improved by 1/10 :

$$\left.\frac{\Delta R_3^{bl}}{R_3^{bl}}\right|_{\mathrm{LEP}} = \frac{1}{10}\left.\frac{\Delta R_3^{bl}}{R_3^{bl}}\right|_{\mathrm{Giga-Z\ ILC}}.$$

Additionally, since experimental tools such as Flavor-Tagging etc... has superior performance than LEP, we can expect to improve systematic errors. The detector uncertainties of ILC are estimated in table.14. However, it should be noted that we do not need to consider radiative return and di-boson events at Z-pole experiments. So let us use the detector uncertainty, which comes from Flavor-Tagging only here. It is 0.07 for left polarization and 0.06 for right polarization. If we assume that R_3^{bl}=0.96 and $m_b(M_Z)$=2.97GeV at Giga-Z ILC, detector uncertainty on b mass is

$$\Delta m_b(exp.) \sim \frac{\Delta R_3^{bl}(exp.)}{2(1-R_3^{bl})} m_b \sim \frac{\Delta C^{det}\cdot R_3^{bl}}{2(1-R_3^{bl})} m_b \sim 0.02\mathrm{GeV}$$

Under such improvements, the b quark mass result at Giga-Z ILC is like this :

$$\Delta m_b\left(M_Z\right) = \pm 0.02(stat.) \pm 0.02(exp.) \pm 0.09(had.) \pm 0.06(theo.). \tag{6.3}$$

If we compare it to actual LEP result eq.(2.9), theoretical and hadronization uncertainties are still dominant, but we can expect that $\Delta m_b\left(M_Z\right)/m_b\left(M_Z\right) \sim 3.7\%$ precision and it's superior to LEP result. So this measurement method at ILC Giga-Z can provide higher precision of running b mass at Z-pole. It will be a valuable development for quark mass study. Figure.39 shows results of b mass evolution, which includes estimations of this study.

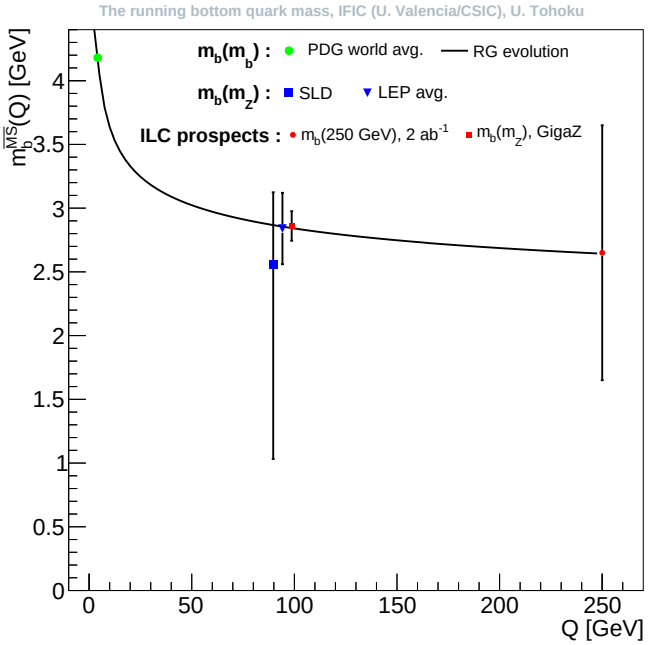


Figure 39: The evolution of b mass with the energy scale Q of processes(Preliminary)[14] : The black curve shows the evolution of b mass based on SM(RGE of QCD theory). The green marker is the reference data at m_b scale, which is the PDG world average. Blue markers are results that are obtained by measurements at LEP(averaged) and SLD. The red markers are estimations that are obtained in this study. The result at 250GeV is estimated by combining two results for each polarization configuration by Best Linear Unbiased Estimater.

In this plot, two results for different configurations at 250GeV are combined by Best Linear Unbiased Estimator, and the correlations among systematic errors are taken into account[14][73].

Based on the above discussions, a large improvement of 250GeV b quark mass up to 10% seems difficult. We expected to measure R_3^{bl} at per mile level by using ILC before starting this study. However, why could we not do this? It can be said that it is the appearance of the difficulty of b mass determination at a higher energy scale. However, this result provides verification of b mass evolution up to the energy scale that has not been seen at previous measurements.

By the way, as a higher energy scale than Z-pole, the b mass at around 125GeV, that is Higgs mass scale, can be known by measuring Yukawa coupling between Higgs and b quark. According to [38], it can be measured with the accuracy of about 1.3% at 250GeV ILC. Since b mass connects to Yukawa coupling as we have seen in eq.(1.53), the precision of b mass, which can be expected by Yukawa coupling measurement, is

$$\frac{\Delta g_{hbb}}{g_{hbb}} = \frac{\Delta m_b}{m_b} = 1.3\%.$$

So we can measure b mass at around Higgs mass scale with higher accuracy. Why did the accuracy suddenly deteriorate as soon as we proceed to 250 GeV from there? Its answer to this question is whether we care about the number of jets or not when measure b mass. In the case of Yukawa coupling g_{hbb} measurement, we use the observable $\sigma(e^+e^- \to Zb\bar{b})$ in addition to well-measured Higgs-Z coupling g_{hZZ} and Higgs total width Γ_h :

$$g_{hbb} = \sqrt{\frac{\sigma(e^+e^- \to Zb\bar{b})\Gamma_h}{Cg_{hZZ}^2}}$$

Here, $\sigma(e^+e^- \to Zb\bar{b})$ is the cross-section of $e^+e^- \to Zh$ ($h \to b\bar{b}$). C is a calculable factor which is decided from Feynman rule. 2 inputs g_{hZZ} and Γ_h are measured by the process of Higgs-Strahlung process $e^+e^- \to Zh$ and Vector boson fusion process $e^+e^- \to \nu_e\bar{\nu}_e h$. Whereas, $\sigma(e^+e^- \to Zb\bar{b})$ can be measured by counting b-jet events simply. When the cross-section $\sigma(e^+e^- \to Zb\bar{b})$ is measured, we observe jet process without caring about the number of jets. On the other hand, if we want to measure b mass at 250GeV, using the jet ratio is a way to obtain b mass more directly because there is no resonance process at 250GeV. Of course, we use jet processes for both cases, but the jet's number is needed to care for in this measurement. Extraction of a specified jet number(in this case, 3-jet event) depends on several factors such as how to define a 3-jet event, or how to configure gluon radiation from quarks in event generator etc. As a result, if we attempt to focus on some specified jet number, the measurement becomes more complicated than the case which does not need to care for jet number. Because of such reason, b mass at 250GeV can not be measured with high accuracy, unlike the Higgs mass-energy scale.

One of the remained tasks is that the point of the current Monte Carlo $q\bar{q}$ sample is simulated up to only LO with massless quarks in WHIZARD. We suspect that this point leads to an

[73] In this plot, the center value at 250GeV is taken as 2.65. It is different from the one(2.75) which is used in the above estimations. However, this study estimates the precision of b quark mass measurement at 250GeV, and it is no significant difference between whether 2.75 or 2.65.

unreliable number of R_3^{bl} which is larger than one and the less number of 3-jet event less than an assumed number, but its firm reason is discussing now. WHIZARD authors are updating gluon radiation in MC sample up to NLO with massive quarks now. So in the future, we will be able to do better b quark mass study.

6.2 Summary

The energy dependence of b quark mass can be expected to provide a probe of new physics, and b quark mass at up to Z-pole have been measured at LEP experiments. These results are consistent with the SM, but there are no indications of new physics from previous studies. The b quark mass study at a higher energy scale is a new challenge, and this study focuses on b mass determination at 250GeV ILC. The b mass study at 250GeV can be expected to provide a QCD verification and some probe to GUT scale new physics models. However, the b mass sensitivity becomes lower at a higher energy scale in principle, and the observable need to be measured at per mile level at 250GeV to competitive results to LEP.

In this study, the precision of b mass measurement at 250GeV ILC is obtained like this :

$$\Delta m_b\,(-,+) = 0.85(stat.) \pm 0.75(exp.) \pm 0.34(had.) \pm 0.07(theo.)\text{GeV}$$
$$\Delta m_b\,(+,-) = 1.53(stat.) \pm 0.44(exp.) \pm 0.34(had.) \pm 0.07(theo.)\text{GeV}.$$

We encountered two problems; the center value of R_3^{bl} is unreliable, and statistics of 3-jet events are less than the assumed numbers. It seems that they come from handling b quark mass and QCD corrections of MC. This $q\overline{q}$ sample is updating now by WHIZARD authors.

If this MC sample is updated, the central value of R_3^{bl} may be improved. Additionally, we can expect that about $1/\sqrt{2}$ times improvement of b mass precision at 250GeV by using some upgrade options of 250GeV ILC. Furthermore, if we return to Z-pole at ILC, Giga-Z ILC can be expected to publish better b mass results than LEP and SLD experiments.

Appendix A

Useful formula for renormalization theory

This part shows the proof of formulas that are useful for the calculation of renormalization.

- The volume of d-dimensional unit sphere

  It is needed to calculate the angular part in eq.(1.64). It can be calculated easily by using Gaussian integration :

$$\begin{aligned}
\pi^{\frac{d}{2}} = \left(\int dx e^{-x^2}\right)^d &= \int dx_1 \cdots dx_d \, \exp\left(-\sum_{i=1}^{d} x_i^2\right) \\
&= \int d^d x \; e^{-r^2} \qquad \left(\because r^2 \equiv \sum_{i=1}^{d} x_i^2\right) \\
&= \int d\Omega_d \int_0^\infty dr \; r^{d-1} e^{-r^2} \\
&= \int d\Omega_d \frac{1}{2} \int_0^\infty dr^2 \left(r^2\right)^{\frac{d}{2}-1} e^{-r^2}
\end{aligned}$$

  Using the Gamma function's definition, it can be concluded as

$$\int d\Omega_d = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)} \qquad \left(\Gamma(x) \equiv \int_0^\infty dt \; t^{x-1} e^{-t}\right). \tag{Appendix A.1}$$

- The radius part of loop integration

  To evaluate eq.(1.64), we need to calculate the radius part addition to the above angular part. The radius part of eq.(1.64) is given by

$$\int dl \; l^{d-1} \frac{1}{l^2 + m^2}.$$

  The following formula is useful to calculate it :

$$\int dl^2 \; \frac{\left(l^2\right)^{\frac{d}{2}-1}}{\left(l^2+m^2\right)^n} = \frac{\Gamma(\frac{d}{2})\Gamma(n-\frac{d}{2})}{\Gamma(n)} \left(m^2\right)^{\frac{d}{2}-n}. \tag{Appendix A.2}$$

  It is verified like this :

$$\because \int_0^\infty d\alpha \; e^{-A\alpha} = \frac{1}{A} \qquad \int_0^\alpha d\alpha \; \alpha^{n-1} e^{-A\alpha} = \frac{1}{A^n} \int_0^\infty dt \; t^{n-1} e^{-t} = \frac{\Gamma(n)}{A^n} \quad (t = A\alpha)$$

$$\therefore \; \frac{1}{\left(l^2+m^2\right)^n} = \frac{1}{\Gamma(n)} \int_0^\infty d\alpha \; \alpha^{n-1} e^{-\left(l^2+m^2\right)\alpha}$$

$$\begin{aligned}
\rightarrow \int dl^2 \frac{\left(l^2\right)^{\frac{d}{2}-1}}{\left(l^2+m^2\right)^n} &= \frac{1}{\Gamma(n)} \int dl^2 \; \left(l^2\right)^{\frac{d}{2}-1} \int_0^\infty d\alpha \; \alpha^{n-1} e^{-\left(l^2+m^2\right)\alpha} \\
&= \frac{1}{\Gamma(n)} \int_0^\infty d\alpha \; \alpha^{n-1} e^{-m^2\alpha} \int dl^2 (l^2)^{\frac{d}{2}-1} e^{-\alpha l^2} \\
&= \frac{\Gamma(\frac{d}{2})}{\Gamma(n)} \int_0^\infty d\alpha \; \alpha^{\left(n-\frac{d}{2}\right)-1} e^{-m^2\alpha} \\
&= \frac{\Gamma(\frac{d}{2})\Gamma(n-\frac{d}{2})}{\Gamma(n)} \left(m^2\right)^{\frac{d}{2}-n} \quad \blacksquare
\end{aligned}$$

Appendix B

Detailed calculation of mass generation in Higgs mechanism

- Expansion of Scalar doublet at around of the VEV

  At once we choose a vacuum such as eq.(1.43), the scalar doublet Φ can be expanded at around the VEV Φ_{min}. Since Φ is a doublet of complex scalar fields, the fluctuation around the minimum value Φ_{min} has both real and imaginary components :

$$\Phi = \Phi_{min} + \begin{pmatrix} \phi_{1r} + i\phi_{1c} \\ \phi_{2r} + i\phi_{2c} \end{pmatrix} = \begin{pmatrix} \phi_{1r} + i\phi_{1c} \\ \frac{v}{\sqrt{2}} + \phi_{2r} + i\phi_{2c} \end{pmatrix} \qquad \text{(Appendix B.1)}$$

  It can be rewritten as eq.(1.44). Its proof is shown below :

$$\begin{aligned} e^{i\theta_i \frac{\sigma_i}{2}} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} &\sim \left(\mathbf{1} + i\theta_i \frac{\sigma_i}{2}\right) \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} 1 + i\theta_3 \frac{1}{2} & \frac{1}{2}i\theta_1 + \frac{1}{2}\theta_2 \\ \frac{1}{2}i\theta_1 - \frac{1}{2}\theta_2 & 1 - i\theta_3 \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \end{aligned}$$

  If each real term is replaced to like this :

$$\begin{aligned} \phi_{1r} &\equiv \frac{\theta_2}{2\sqrt{2}}(v+h) \ , \ \phi_{1c} \equiv \frac{\theta_1}{2\sqrt{2}}(v+h) \\ \phi_{2r} &\equiv \frac{h}{\sqrt{2}} \ , \ \phi_{2c} \equiv \frac{\theta_3}{2\sqrt{2}}(v+h), \end{aligned}$$

  above equation is rewritten by

$$e^{i\theta_i \frac{\sigma_i}{2}} \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \phi_{1r} + i\phi_{1c} \\ \frac{v}{\sqrt{2}} + \phi_{2r} + i\phi_{2c} \end{pmatrix} = \Phi \quad \blacksquare$$

- Mass generation of gauge bosons

  The Lagrangian, which is invariant under $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ is given by eq.(1.42). After EW gauge symmetry breaking and selection of Unitary gauge, the kinetic term of Higgs Sector will be

$$D_\mu \Phi^\dagger D^\mu \Phi = D_\mu \begin{pmatrix} 0 & \frac{v+h(x)}{\sqrt{2}} \end{pmatrix} D^\mu \begin{pmatrix} 0 \\ \frac{v+h(x)}{\sqrt{2}} \end{pmatrix}$$

  Here, $h(x)$ is the Higgs field, which is remained a physical field under the unitary gauge, and v is the VEV of the Higgs field. Since the covariant derivative D_μ includes SU(2) and U(1) gauge fields, the above equation can be separated into two parts; the first one is the term which includes VEV v, and the second one is the interaction term between Higgs field and four EW gauge fields. To see the scenario of the mass generation of gauge

bosons, let us focus on the former one only. It can be calculated like this :

$$\begin{aligned}
&\left(\left(\partial_\mu + ig'\frac{\sigma_i}{2}W_{\mu i} + i\frac{1}{2}gB_\mu\right)\begin{pmatrix} 0 & \frac{v}{\sqrt{2}} \end{pmatrix}\right)\left(\left(\partial^\mu + ig'\frac{\sigma_i}{2}W^{\mu i} + i\frac{1}{2}gB^\mu\right)\begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}\right) \\
&= \frac{v^2}{8}\left|\left(g'\sigma_i W^{\mu i} + gB^\mu\right)\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right|^2 \\
&= \frac{v^2}{8}\left|\begin{pmatrix} g'W_3^\mu + gB^\mu & g'W_1^\mu - ig'W_2^\mu \\ g'W_1^\mu + ig'W_2^\mu & g'W_3^\mu + gB^\mu \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right|^2 \\
&= \frac{v^2}{8}\left|\begin{array}{c} g'W_1^\mu - ig'W_2^\mu \\ g'W_3^\mu + gB^\mu \end{array}\right|^2 \\
&= \frac{v^2}{8}\left((g')^2\left((W_1^\mu)^2 + (W_2^\mu)^2\right) + (g'W_3^\mu - gB^\mu)^2\right)
\end{aligned}$$

If we construt new fields :

$$W_\pm^\mu \equiv \frac{1}{\sqrt{2}}\left(W_1^\mu \mp iW_2^\mu\right), \qquad \text{(Appendix B.2)}$$

the first term of the above result is rewritten by

$$\left(\frac{g'v}{2}\right)^2 W_{+\mu}W_-^\mu. \qquad \text{(Appendix B.3)}$$

On the other hand, the second term of above result can be rewritten by

$$\begin{aligned}
(g'W_3^\mu - gB^\mu)^2 &= W_3^\mu\left((g')^2 W_3^\mu - gg'B^\mu\right) + B^\mu\left(-gg'W_3^\mu + g^2B_\mu\right) \\
&= \begin{pmatrix} W_3^\mu & B^\mu \end{pmatrix}\begin{pmatrix} (g')^2 & -gg' \\ -gg' & g^2 \end{pmatrix}\begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix}.
\end{aligned}$$

If we define

$$G \equiv \begin{pmatrix} (g')^2 & -gg' \\ -gg' & g^2 \end{pmatrix},$$

the matrix has the following eigenstates and eigenvalues :

$$\begin{cases} \lambda_1 = 0 \ : \ v_1 = \frac{1}{\sqrt{g^2+(g')^2}}\begin{pmatrix} g \\ g' \end{pmatrix} \\ \lambda_2 = g^2 + (g')^2 \ : \ v_2 = \frac{1}{\sqrt{g^2+(g')^2}}\begin{pmatrix} g' \\ -g \end{pmatrix} \end{cases}$$

So we can diagonalize G by using the matrix M which is constructed by these eigenstates :

$$\begin{aligned}
G_{diag} &= M^{-1}GM \quad \left(M \equiv \frac{1}{\sqrt{g^2 + (g')^2}}\begin{pmatrix} g & g' \\ g' & -g \end{pmatrix}\right) \\
&= \begin{pmatrix} 0 & 0 \\ 0 & g^2 + (g')^2 \end{pmatrix}
\end{aligned}$$

The second term of kinetic term will be

$$\frac{v^2}{8}(g'W_3^\mu - gB^\mu)^2 = \frac{v^2}{8}\begin{pmatrix} W_3^\mu & B^\mu \end{pmatrix} MG_{diag}M^{-1}\begin{pmatrix} W_3^\mu \\ B_\mu \end{pmatrix}.$$

If we define new fields from W_3^μ and B^μ as [74]

$$A^\mu \equiv \frac{1}{\sqrt{g^2+(g')^2}}\left(gW_3^\mu + g'B^\mu\right) \ , \ Z^\mu \equiv \frac{1}{\sqrt{g^2+(g')^2}}\left(g'W_3^\mu - gB^\mu\right), \quad \text{(Appendix B.4)}$$

it is rewritten by

$$\frac{v^2}{8}\left(g'W_3^\mu - gB^\mu\right)^2 = 0\cdot A_\mu A^\mu + \frac{v^2}{8}\left(g^2+(g')^2\right)Z_\mu Z^\mu \quad \text{(Appendix B.5)}$$

By combining eq.(Appendix B.3) and eq.(Appendix B.5), the result what we want finally is

$$\frac{v^2g^2}{4}W_{+\mu}W_-^\mu + \frac{v^2}{8}\left(g^2+(g')^2\right)Z_\mu Z^\mu + \frac{v^2}{8}\cdot 0\cdot A_\mu A^\mu. \quad \text{(Appendix B.6)}$$

If each coefficient is defined by $M_W^2 \equiv \frac{v^2g^2}{4}$, $M_Z^2 \equiv \frac{v^2}{4}\left(g^2+(g')^2\right)$, $M_A^2 \equiv \frac{v^2}{8}\cdot 0 = 0$, these coefficients can be regarded as masses of each gauge boson.

[74] If we define $\frac{g}{\sqrt{g^2+(g')^2}} \equiv \cos\theta_W$, $\frac{g'}{\sqrt{g^2+(g')^2}} \equiv \sin\theta_W$, we can see that A_μ and Z_μ appear through rotation by the angle of θ_W on the $W_3^\mu - B_\mu$ plain. This mixing angle θ_W is called Weinberg angle and it is one of the fundamental input parameters of the SM.

References

[1] ゲージ場の量子論 I, 久後 汰一郎, 培風館

[2] Physics from Symmetry second edition, Jakob Schwichtenberg, Springer

[3] Quantum Field Theory, Mark Srednicki, Cambridge university press

[4] 場の量子論 (II) ファインマン・グラフとくりこみを中心にして, 坂本 眞人, 裳華房

[5] Particle Data Group HP, https://pdg.lbl.gov/

[6] H. Baer, I. Gogoladze, A. Mustafayev, S. Raza, Q. Shafi, arXiv:1201.4412 [hep-ph] (2012)

[7] S. Chigusa, T. Moroi, arXiv:1604.02156 [hep-ph] (2016)

[8] G. Rodrigo, M. Bilenky, A. Santamaria, arXiv:hep-ph/9905276 (1999)

[9] Introduction to Elementary Particles, David Driffiths, Wiley-VCH

[10] ILC Technical Design Report volume.4, https://arxiv.org/pdf/1306.6329.pdf

[11] Costa.M.J, Determination of the b quark mass at the Mz scale with the DELPHI detector at LEP, CERN-DELPHI-THESIS-251 ; DELPHI-THESIS-251 (2003)

[12] G. Rodrigo, A. Santamaria, Phys. Rev. Lett. 79:193-196 (1997)

[13] Aida X. El-Khadra, M. Luke, Ann. Rev. Nucl. Part. Sci. 52 201-251 (2002).

[14] J. Fuster, A. Irles, S. Tairafune, G. Rodrigo, M. vos, H. Yamamoto, R. Yonamine, Prospects for the measurement of the bottom quark mass at the ILC, ILD note in preparation.

[15] R. Bailey, C. Benvenuti, S. Myers, D. Treille, C. R. Physique 3 1107-1120 (2002)

[16] G. Rodrigo, A. Santamaria, M. Bilenky, Phys. Rev. Lett. 79 193-196 (1997).

[17] G. Rodrigo, M. Bilenky, A. Santamaria, Nucl. Phys. B554 257-297 (1999).

[18] The DELPHI Collaboration, J. Abdallah, et. al., Eur. Phys. J.C46:569-583 (2006)

[19] M. Battaglia et al., Phys. Lett. B556 41-49 (2003)

[20] ILC Technical Design Report volume.1, arXiv:1306.6327 [physics.acc-ph] (2013)

[21] P. Bambade, et. al., arXiv:1903.01629 [hep-ex], (2019)

[22] Partcile Data Group HP, 12. CKM Quark-Mixing Matrix (2018)

[23] M. Kobayashi, et. al., Nuclear Instruments and Methods in Physics Research A 767 439-444 (2014)

[24] Particle Data Group HP, 33. Passage of Particles Through Matter (2019)

[25] A. Ishikawa, et. al., ILC の検出器, http://www.jahep.org/hepnews/2014/14-1-5-ILCDetector.pdf (2014)

[26] J. S. Marshell, M. A. Thomson, arXiv:1308.4537 [physics.ins-det] (2013)

[27] T. Sjöstrand, S. Mrenna, P. Skands, JHEP 026 (2006)

[28] W. Kilian, T. Ohl, J. Reuter, Eur.Phys.J.C71:1742 (2011)

[29] M. Cacciari, G. P. Salam, G. Soyez, arXiv:1111.6097 [hep-ph] (2011)

[30] T. Suehara, T. Tanabe, Nucl.Instrum.Meth. A808, 109-116 (2016)

[31] M. Boronat, J. Fuster, I. Garcia, P. Roloff, R. Simoniello, M. Vos, arXiv:1607.05039v2 [hep-ex] (2017)

[32] ROOT Reference Guide, https://root.cern/doc/master/

[33] K. Fujii, et. al., arXiv:1801.02840 [hep-ph] (2018)

[34] P. Abreu, et. al., Phys. Lett. B418 430-442 (1998)

[35] Zhi-zhong Xing, He Zhang, Shun Zhou, Phys. Rev. D77:113016 (2008)

[36] K. Bora, arXiv:1206.5909 [hep-ph](2013)

[37] Contributions to the European Particle Physics Strategy Update 2018-2020: https://indico.cern.ch/event/765096/contributions/, Document 77 (2018)

[38] T. Barklow, et. al., arXiv:1708.08912v5 [hep-ph] (2018)

[39] T. Barklow, et. al., arXiv:1506.07830 [hep-ex] (2015)