

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE  
**CERN** EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

**2013 European School of High-Energy Physics**


Parádfürdő, Hungary  
5 – 18 June 2013

**Proceedings**

Editors: M. Mulders  
G. Perez

ISBN 978-92-9083-420-5 (paperback)  
ISBN 978-92-9083-421-2 (PDF)  
ISSN 0531-4283  
DOI <http://dx.doi.org/10.5170/CERN-2015-004>  
Available online at <http://cds.cern.ch/>

Copyright © CERN, 2015

 Creative Commons Attribution 4.0

Knowledge transfer is an integral part of CERN's mission.

CERN publishes this report Open Access under the Creative Commons Attribution 4.0 license (<http://creativecommons.org/licenses/by/4.0/>) in order to permit its wide dissemination and use. The submission of a contribution to a CERN Yellow Report shall be deemed to constitute the contributor's agreement to this copyright and license statement. Contributors are requested to obtain any clearances that may be necessary for this purpose.

This report is indexed in: CERN Document Server (CDS), INSPIRE, Scopus.

This report should be cited as:

Proceedings of the 2013 European School of High-Energy Physics, Parádörd, Hungary, 5-18 June 2013, edited by M. Mulders and G. Perez, CERN-2015-004 (CERN, Geneva, 2015), <http://dx.doi.org/10.5170/CERN-2015-004>

A contribution in this report should be cited as:

[Author name(s)], in Proceedings of the 2013 European School of High-Energy Physics, Parádörd, Hungary, 5-18 June 2013, edited by M. Mulders and G. Perez, CERN-2015-004 (CERN, Geneva, 2015), pp. [first page]-[last page], <http://dx.doi.org/10.5170/CERN-2015-004>. [first page]

## **Abstract**

The European School of High-Energy Physics is intended to give young physicists an introduction to the theoretical aspects of recent advances in elementary particle physics. These proceedings contain lecture notes on the Standard Model of electroweak interactions, quantum chromodynamics, Higgs physics, physics beyond the Standard Model, flavour physics, and practical statistics for particle physicists.





## Preface

The twenty-first event in the series of the European School of High-Energy Physics took place in Parád-fürdő, Hungary, from 5 to 18 June 2013. It was organized jointly by CERN, Geneva, Switzerland, and JINR, Dubna, Russia, with support from several bodies in Hungary: Academy of Science; Scientific Research Fund; University of Debrecen; Wigner Research Centre for Physics. The local organization team was chaired by Professor Dezső Horváth. The other members of the local committee included: Csaba Hajdu, Gergő Hamar, János Karancsi, Ferenc Siklér, Balázs Ujvári, Viktor Veszprémi, and Anna Zsigmond.

A total of 114 students coming from 27 different countries attended the school, mainly from member states of CERN and/or JINR, but also a few from other regions. The participants were generally students in experimental High-Energy Physics in the final years of work towards their PhDs.

The School was hosted at the Erzsébet Park Hotel complex in Parád-fürdő, about 120 km from Budapest. According to the tradition of the school, the students shared twin rooms mixing participants of different nationalities.

A total of 31 lectures were complemented by daily discussion sessions led by six discussion leaders. The students displayed their own research work in the form of posters in an evening session in the first week, and the posters stayed on display until the end of the School. Each discussion group carried out a collaborative project, studying in detail the analysis from a published paper from one of the LHC experiments; a summary was presented by a student representative of each group in an evening session in the second week of the School. The full scientific programme was arranged in the on-site conference facilities.

Our thanks go to the local-organization team and, in particular, to Dezső Horváth and Csaba Hajdu for all their work and assistance in preparing the School, on both scientific and practical matters, and for their presence throughout the event. Our thanks also go to the efficient and friendly hotel management and staff who assisted the School organizers and the participants in many ways.

Very great thanks are due to the lecturers and discussion leaders for their active participation in the School and for making the scientific programme so stimulating. The students, who in turn manifested their good spirits during two intense weeks, undoubtedly appreciated listening to and discussing with the teaching staff of world renown.

We would like to express our appreciation to Professor Rolf Heuer, Director General of CERN, and Professor Victor Matveev, Director General of JINR, for their lectures on the scientific programmes of the two organizations and for discussing with the School participants.

In addition to the rich scientific programme, the participants enjoyed numerous sports and leisure activities in and around the Erzsébet Park Hotel complex. Particularly noteworthy were the very nice excursions to the nearby town of Eger, to Budapest with impressive views of the Danube in flood in addition to the normal sites of the capital city, and to the spectacular Aggtelek caves. Sports and leisure facilities, including the swimming pool, wellness centre and ten-pin bowling, provided an excellent environment for informal interactions between staff and students.

We would also like to thank Professor József Pálinkás, President of the Hungarian Academy of Sciences, and Professor Péter Lévai, Director General of the Wigner Research Centre for Physics in Budapest, for visiting the School and more generally for their interest and support. We are very grateful to Kate Ross and Tatyana Donskova for their untiring efforts in the lengthy preparations for and the day-to-day operation of the School. Their continuous care of the participants and their needs during the School was highly appreciated. We would also like to thank Hélène Haller for her help in the early stages of preparing the School.

The success of the School was to a large extent due to the students themselves. Their poster session was very well prepared and highly appreciated, their group projects were a great success, and throughout the School they participated actively during the lectures, in the discussion sessions and in the different activities and excursions.

Nick Ellis  
(On behalf of the Organizing Committee)









## PEOPLE IN THE PHOTOGRAPH

1	Nick Ellis	34	Ed Greening	67	Ferenc Sikler	100	Petr Balek
2	Csaba Hajdu	35	YanCJie Schnellbach	68	Mario Vormstein	101	Anna Usanova
3	Peter Levai	36	Francesca Ungaro	69	Guillaume Lefebvre	102	Martin Schultens
4	Andrej Arbuzov	37	Lucy Kogan	70	Rosemarie Aben	103	Veerle Heijne
5	Janos Karancsi	38	Lawrence Lee	71	Nikolay Anfimov	104	Ahmed Hasib
6	Dezso Horvath	39	Francesco Nuti	72	Hale Sert	105	Ahmed Bassalat
7	Tamas Biro	40	Priscilla Pani	73	Antonella Succurro	106	Alexander Tuna
8	Kate Ross	41	Claudia Giuliani	74	Andrey Belyaev	107	Radek Zlebcik
9	Bonnie Chow	42	Ehab Abbas	75	Christian Voss	108	Julia Fischer
10	Gabor Cynolter	43	Reza Goldouzian	76	Nils Ruthmann	109	Federico Scutti
11	Rebekka Hoenig	44	Predrag Cirkovic	77	Ivan Asin Cruz	110	Francesco Michelli
12	Viktor Veszpremi	45	Aurthorm Borodenko	78	Marco Tresch	111	Ivan Razumov
13	Martijn Mulders	46	Mark Turner	79	Marion Habiby Alaoui	112	Emrah Tiras
14	Balazs Ujvari	47	Vincent Andrieux	80	Eduard Boos	113	Emanuele Ripiccini
15	Zotan Trocsanyi	48	Maximilian Swiatlowski	81	Claudia Seitz	114	Elena Yatsenko
16	Marion Murumaa	49	Jessica Liebal	82	Livia Soffi	115	Monica Trovatelli
17	Johannes Stiller	50	Diedi Hu	83	Gagik Vardanyan	116	Lucia Perrini
18	Mykhailo Dalchenko	51	Emine Gurpinar	84	Zara Grout	117	Stefano Casasso
19	Kristin Heine	52	Tobias Verlage	85	Stefan Gadatsch	118	Jacob Howard
20	Julia Hofmann	53	Manuel Kambeitz	86	Paul Seyfert	119	Bing Li
21	Edmund Smith	54	Joern Mahlstedt	87	GoblirschCKolb Maximilian	120	Andrea De Simone
22	Geoffrey Herbert	55	Aurélien Demilly	88	Swagata Mukherjee	121	Giovanni Zurzolo
23	Jure Klucar	56	Bruno Galhardo	89	Ruth Sandbach	122	Charles Dietz
24	Agnieszka Dziurda	57	John Ellis	90	Jon Harrison	123	Spyridon Argyropoulos
25	Jessica Brinson	58	Tatyana Donskova	91	Thomas Neep	124	Antonio Salvucci
26	Elodie Tiouchichine	59	Bastian Kargoll	92	Michael Ughetto	125	Maria Zangoli
27	Bedrich Roskovec	60	Gergo Hamar	93	James Saxon	126	Krisztina Marton
28	Tova Holmes	61	Benjamin Nachman	94	Jonathan David Long	127	Anna Zsigmond
29	Gregor Hellwig	62	Florian Thibaud	95	Arseniy Rybnikov	128	Elmar Ritsch
30	Michal Meres	63	Christian Boser	96	Veronika Chobanova	129	Ilse Kratschmer
31	Katherine Pachal	64	Nika Valencic	97	Artem Ivanov	130	Joram Berger
32	Claudio Corti	65	Alexey Gladyshev	98	Monika Blanke	131	Karoly Uermoessy
33	Andrew Edmonds	66	Eduardo Navarro De Martino	99	Lucie Gauthier	132	Gyula Bencedi



# PHOTOGRAPHS (MONTAGE)





# PHOTOGRAPHS (MONTAGE)



# Contents

Preface	
<i>N. Ellis</i> .....	v
Photograph of participants .....	vi
Photographs (montage) .....	x
Quantum Field Theory and the Electroweak Standard Model	
<i>E. Boos</i> .....	1
QCD for collider experiments	
<i>Z. Trócsányi</i> .....	65
Higgs Physics	
<i>J. Ellis</i> .....	117
Beyond the Standard Model	
<i>C. Csáki and P. Tanedo</i> .....	169
Flavor Physics and CP Violation	
<i>Z. Ligeti</i> .....	269
Practical Statistics for Particle Physicists	
<i>H. B. Prosper</i> .....	301
Organizing Committee .....	329
Local Organizing Committee .....	329
List of Lecturers .....	329
List of Discussion Leaders .....	329
List of Students .....	330
List of Posters .....	331





# Quantum Field Theory and the Electroweak Standard Model

*E. Boos*

M. V. Lomonosov Moscow State University, Skobeltsyn Institute of Nuclear Physics (SINP MSU), Moscow 119991, Russia

## Abstract

The Standard Model is one of the main intellectual achievements for about the last 50 years, a result of many theoretical and experimental studies. In this lecture a brief introduction to the electroweak part of the Standard Model is given. Since the Standard Model is a quantum field theory, some aspects for understanding of quantization of abelian and non-abelian gauge theories are also briefly discussed. It is demonstrated how well the electroweak Standard Model works in describing a large variety of precise experimental measurements at lepton and hadron colliders.

## 1 Introduction

The Standard Model (SM) of strong and electroweak (EW) interactions is the basis for understanding of nature at extremely small distances. In high-energy physics usually the relativistic system of units is used in which the Planck constant  $\hbar$  and the speed of light  $c$  are equal to unity,  $\hbar = c = 1$ . Taking into account well-known values for  $\hbar = 1.055 \cdot 10^{27}$  erg s,  $c = 3 \cdot 10^{10}$  cm/s and the positron electric charge  $e = 1.6 \cdot 10^{-19}$  C and using the relation between the electronvolt and erg ( $1 \text{ eV} = e \cdot 1 \text{ V} = 1 \text{ V} \cdot 1.6 \cdot 10^{-19} \text{ C} = 1.6 \cdot 10^{-12}$  erg), one easily gets the following very useful relation between length and energy units:  $1/\text{GeV} = 2 \cdot 10^{-14}$  cm. Due to the Heisenberg uncertainty principle,  $\Delta x \Delta p \geq 1/2$ , the above relation allows us to understand which energies (momentum transfers) are needed approximately to probe certain distances:

$$100 \text{ GeV} \rightarrow 10^{-16} \text{ cm},$$

$$1 \text{ TeV} \rightarrow 10^{-17} \text{ cm},$$

$$10 \text{ TeV} \rightarrow 10^{-18} \text{ cm}.$$

Therefore, at the LHC one can study the structure of matter at distances of  $10^{-18}$ – $10^{-17}$  cm. For small distances of the order of  $10^{-16}$  cm or correspondingly 100 GeV energies the SM works very well, as follows from many studies and measurements.

The SM is a quantum field theory; it is based on a few principles and requirements:

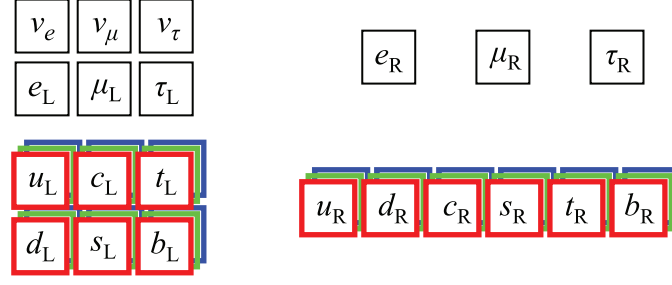
- gauge invariance with lowest dimension (dimension four) operators; SM gauge group:  $SU(3)_C \times SU(2)_L \times U(1)_Y$ ;
- correct electromagnetic neutral currents and correct charge currents with (V–A) structure as follows from four fermion Fermi interactions (1)

$$\frac{G_F}{\sqrt{2}} [\bar{\nu}_\mu \cdot \gamma_\alpha (1 - \gamma_5) \cdot \mu] [\bar{e} \cdot \gamma_\alpha (1 - \gamma_5) \cdot \nu_e] + \text{h.c.}; \quad (1)$$

- three generations without chiral anomalies;
- Higgs mechanism of spontaneous symmetry breaking.

Fermions are combined into three generations forming left doublets and right singlets with respect to the weak isospin (see Fig.1).

$$f_{L,R} = \frac{1}{2}(1 \mp \gamma_5)f,$$



**Fig. 1:** Fermion generation

$$I_f^{3L,3R} = \pm \frac{1}{2}, 0 : \begin{aligned} L_1 &= \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, e_{R1} = e_R^-, Q_1 = \begin{pmatrix} u \\ d \end{pmatrix}_L, u_{R1} = u_R, d_{R1} = d_R, \\ L_2 &= \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, e_{R2} = \mu_R^-, Q_2 = \begin{pmatrix} c \\ s \end{pmatrix}_L, u_{R1} = c_R, d_{R1} = s_R, \\ L_3 &= \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, e_{R3} = \tau_R^-, Q_3 = \begin{pmatrix} t \\ b \end{pmatrix}_L, u_{R1} = t_R, d_{R1} = b_R. \end{aligned}$$

The SM Lagrangian written in accord with the mentioned requirements looks very simple:

$$\begin{aligned} L &= -\frac{1}{4}W_{\mu\nu}^i(W^{\mu\nu})^i - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}G_{\mu\nu}^a(G^{\mu\nu})^a \\ &+ \sum_{f=\ell,q} \bar{\Psi}_L^f(iD_\mu^L \gamma^\mu) \Psi_L^\dagger + \sum_{f=\ell,q} \bar{\Psi}_R^f(iD_\mu^R \gamma^\mu) \Psi_R^\dagger + L_H, \\ L_H &= L_\Phi + L_{\text{Yukawa}}, \\ L_\Phi &= D_\mu \Phi^\dagger D^\mu \Phi - \mu^2 \Phi^\dagger \Phi - \lambda(\Phi^\dagger \Phi)^4, \\ L_{\text{Yukawa}} &= -\Gamma_d^{ij} \bar{Q}_L^i \Phi d_R^j + \text{h.c.} - \Gamma_u^{ij} \bar{Q}_L^i \Phi^C u_R^j + \text{h.c.} - \Gamma_e^{ij} \bar{L}_L^i \Phi e_R^j + \text{h.c.} \end{aligned}$$

The field strength tensors and covariant derivatives have very familiar forms:

$$\begin{aligned} W_{\mu\nu}^i &= \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g_2 \varepsilon^{ijk} W_\mu^j W_\nu^k, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \\ G_{\mu\nu}^a &= \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_S f^{abc} A_\mu^b A_\nu^c, \\ D_\mu^L &= \partial_\mu - ig_2 W_\mu^i \tau^i - ig_1 B_\mu \left( \frac{Y_L^f}{2} \right) - ig_S A_\mu^a t^a, \\ D_\mu^R &= \partial_\mu - ig_1 B_\mu \left( \frac{Y_R^f}{2} \right) - ig_S A_\mu^a t^a, \end{aligned}$$

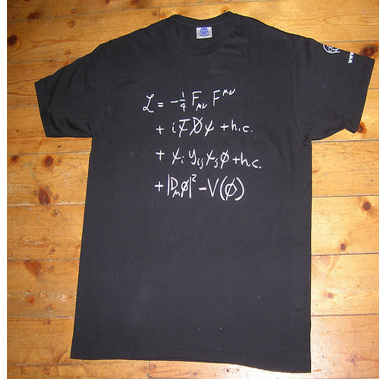
where  $i = 1, 2, 3$ ,  $a = 1, \dots, 8$ ;  $W_\mu^i$  are gauge fields for the weak isospin group,  $B_\mu$  are gauge fields for the weak hypercharge group and  $A_\mu^a$  are gluon gauge fields for the strong  $SU_C(3)$  colour group.

$$Y_f = 2Q_f - 2I_f^3 \Rightarrow Y_{L_i} = -1, Y_{e_{R_i}} = -2, Y_{Q_i} = \frac{1}{3}, Y_{u_{R_i}} = \frac{4}{3}, Y_{d_{R_i}} = -\frac{2}{3}.$$

The Lagrangian is so compact that its main part can be presented on the CERN T-shirt (see Fig. 2).

It is hard to imagine that such a simple Lagrangian allows one to describe basically all the phenomena of the microworld. But the SM Lagrangian, being expressed in terms of physics components, is not that simple, leading after quantization to many interaction vertices between particles or quanta of corresponding quantum fields.

This lecture is organized as follows. In the next section some aspects of quantum field theory are briefly discussed. After a motivation as to why do we need a quantum field theory, we consider scalar



**Fig. 2:** CERN T-shirt with SM Lagrangian

fields and introduce the Feynman propagator and functional integral approach as a quantization method. The functional integral given in the holomorphic representation allows us to clarify boundary conditions and show the connection between the Green functions and S-matrix elements. Feynman diagrams are introduced. The formalism is extended to the fermion and gauge fields stressing peculiarities in the quantization procedure and Feynman rule derivation. In the next section a construction of the EW SM Lagrangian is presented. We discuss experimental facts and theory principles based on which the EW part of the SM Lagrangian for fermion and gauge fields is constructed. We show explicitly which conditions on weak hypercharges allow us to get correctly electromagnetic and charge current (CC) interactions and predict additional neutral currents (NCs). We demonstrate how potentially dangerous chiral anomalies cancelled out. Then spontaneous symmetry breaking, the Goldstone theorem and the appearance of Nambu–Goldstone bosons are briefly discussed. The Brout–Englert–Higgs–Hagen–Guralnik–Kibble mechanism of spontaneous symmetry breaking is introduced leading to non-zero masses of the gauge fields and appearance of the Higgs boson. Very briefly we discuss in addition to the unitary gauge the covariant gauge, propagators of Goldstone bosons and ghosts. At the end of the section it is shown how the spontaneous symmetry breaking mechanism leads to non-zero masses for the fermions in the SM and how very naturally the Cabibbo–Kobayashi–Maskawa mixing matrix appears. In the next section we concentrate on some phenomenological aspects of the EW SM such as connections between the Fermi constant  $G_F$ , the Higgs vacuum expectation value  $v$ , consistency of low-energy measurements and W, Z mass measurements, W-, Z-boson decay widths and branching ratios, number of light neutrinos, two-fermion processes in  $e^+e^-$  collisions, tests of the gauge boson self-interactions, top-quark decays and the EW top production (single top). Briefly we discuss the EW SM beyond the leading order, renormalization and running coupling in quantum electrodynamics (QED), as a simplest example, running masses and running parameters in the SM, precision EW data and global parameter fits. Concluding remarks are given in the next section. The quantum chromodynamics (QCD) part of the SM and the phenomenology of the Higgs boson are not discussed in these lectures as they are addressed in other lectures of the School.

For a deeper understanding of the topics discussed, one can recommend a number of very good textbooks and reviews [1–9] and lectures given at previous schools and specialized reviews [10, 12–15], which have been used in preparation of this lecture.

## 2 Introductory words to quantum theory

In classical mechanics a system evolution follows from the principle of least action:

$$\delta S = \delta \int_{t_i}^{t_f} dt L(q(t), \dot{q}(t)) = 0; \quad \int_{t_i}^{t_f} \left[ \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta(\dot{q}) \right] = 0; \quad \delta(\dot{q}) = \frac{d}{dt} \delta q.$$

For an arbitrarily small variation  $\delta q$ , one gets the well-known Lagrange equation of motion

$$\frac{\partial L}{\partial q} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right). \quad (2)$$

For a non-relativistic system described by the Lagrangian  $L = \frac{m\dot{q}^2}{2} - V(q)$ , the second Newton law follows from Eq. (2),

$$m\ddot{q} = -\frac{\partial V}{\partial q} = F.$$

The Hamiltonian of the system is related to the Lagrangian in the following well-known way:

$$H(p, q) = p\dot{q} - L(q, \dot{q}),$$

where  $\dot{q}$  is a solution of the equation  $p = \frac{\partial L}{\partial \dot{q}}$ .

In quantum mechanics the coordinate and momentum are replaced by corresponding operators  $p, q \rightarrow \hat{p}, \hat{q}$  with postulated commutator relation  $[\hat{p}(0), \hat{q}(0)] = -i\hbar$ . In the Heisenberg picture the system evolution is described by the Heisenberg equation with time-dependent operators; for example, the equation for the coordinate operator has the following form:

$$\frac{\partial \hat{q}}{\partial t} = \frac{i}{\hbar} [\hat{H}, \hat{q}] \quad (3)$$

with a formal solution

$$\hat{q}(t) = e^{\frac{i}{\hbar}\hat{H}t} \hat{q}(0) e^{-\frac{i}{\hbar}\hat{H}t}.$$

This easily follows from the equalities

$$\frac{\partial \hat{q}}{\partial t} = \frac{i}{\hbar} \hat{H} e^{\frac{i}{\hbar}\hat{H}t} \hat{q}(0) e^{-\frac{i}{\hbar}\hat{H}t} + e^{\frac{i}{\hbar}\hat{H}t} \hat{q}(0) e^{-\frac{i}{\hbar}\hat{H}t} \hat{H} = \frac{i}{\hbar} \hat{H} \hat{q}(t) - \frac{i}{\hbar} \hat{q}(t) \hat{H} = \frac{i}{\hbar} [\hat{H}, \hat{q}].$$

For the coordinate and momentum operators, one can prove the following inequality:

$$\Delta q \cdot \Delta p \geq 1/2,$$

which is called the Heisenberg uncertainty principle.

Let us recall a simple proof of the Heisenberg uncertainty principle. The mid value of any operator  $\hat{A}$  and its dispersion are given by the relations

$$\langle \psi | \hat{A} | \psi \rangle = \bar{A}, \quad \langle \psi | (\hat{A} - \bar{A})^2 | \psi \rangle.$$

Let us take the following operator constructed from the momentum and coordinate operators ( $[\hat{p}, \hat{q}] = -i\hbar$ ) with arbitrary constant  $\gamma$ :

$$\hat{A} = \hat{p} + i\gamma\hat{q} - (\bar{p} + i\gamma\bar{q}).$$

Then the conjugated operator has the form

$$\hat{A}^\dagger = \hat{p} - i\gamma\hat{q} - (\bar{p} - i\gamma\bar{q}).$$

For any state  $|\psi\rangle$  we have

$$\langle \psi | \hat{A}^\dagger \hat{A} | \psi \rangle \geq 0$$

,

$$\begin{aligned} & \langle \psi | [(\hat{p} - \bar{p}) - i\gamma(\hat{q} - \bar{q})] [(\hat{p} - \bar{p}) + i\gamma(\hat{q} - \bar{q})] | \psi \rangle \\ & = (\Delta p)^2 + \gamma^2 \Delta q^2 - i\gamma(\hat{q}\hat{p} - \hat{p}\hat{q}) = (\Delta p)^2 + \gamma^2 \Delta q^2 + \gamma\hbar \geq 0. \end{aligned}$$

This is true for any value of  $\gamma$  and therefore the determinant is not positive:

$$\frac{\hbar^2}{4} - \Delta q^2 \Delta p^2 \leq 0.$$

Thus, we immediately arrive at the uncertainty principle:

$$\Delta q \Delta p \geq \frac{1}{2} \hbar.$$

Let us consider the simplest system, the harmonic oscillator, described by the Hamiltonian (we put  $\hbar = 1$ )

$$\hat{H} = \frac{1}{2}(\hat{p}^2 + \omega^2 \hat{q}^2),$$

and construct two operators  $\hat{a}$  and  $\hat{a}^\dagger$  which are called the annihilation and creation operators:

$$\begin{aligned} \hat{q} &= \frac{1}{\sqrt{2\omega}}(\hat{a} + \hat{a}^\dagger); & \hat{p} &= -i\sqrt{\frac{\omega}{2}}(\hat{a} - \hat{a}^\dagger). \\ \implies \hat{a} &= \sqrt{\frac{\omega}{2}}\hat{q} + i\frac{1}{\sqrt{2\omega}}\hat{p}; & \hat{a}^\dagger &= \sqrt{\frac{\omega}{2}}\hat{q} - i\frac{1}{\sqrt{2\omega}}\hat{p}. \end{aligned} \quad (4)$$

From the equation  $[\hat{p}, \hat{q}] = -i$ , one gets  $[\hat{a}, \hat{a}^\dagger] = 1$  and the Hamiltonian takes the form

$$\hat{H} = \frac{\omega}{2}(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}).$$

It is easy to show the following commutation relations with the Hamiltonian:

$$[\hat{H}, \hat{a}] = -\omega\hat{a} \quad \text{and} \quad [\hat{H}, \hat{a}^\dagger] = \omega\hat{a}^\dagger.$$

From (3),

$$\frac{d\hat{a}}{dt} = i[\hat{H}\hat{a}] = -i\omega\hat{a} \implies \hat{a}(t) = \hat{a}(0)e^{-i\omega t}; \quad \hat{a}^\dagger(t) = \hat{a}^\dagger(0)e^{i\omega t}.$$

Let us consider states with definite energy:

$$\hat{H}|E\rangle = E|E\rangle.$$

Then the state  $\hat{a}|E\rangle$  ( $\hat{a}^\dagger|E\rangle$ ) corresponds to the energy  $(E - \omega)$  ( $(E + \omega)$ ). Indeed,

$$\hat{H}\hat{a}|E\rangle = \hat{a}\hat{H}|E\rangle - \omega\hat{a}|E\rangle = (E - \omega)\hat{a}|E\rangle,$$

$$\hat{H}\hat{a}^\dagger|E\rangle = (E + \omega)\hat{a}^\dagger|E\rangle.$$

Let us construct states (Hilbert space of states) starting from the ‘vacuum’ state  $|0\rangle$ :

$$\hat{a}|0\rangle = 0.$$

What is the energy of the vacuum state? This is

$$\hat{H}|0\rangle = \frac{\omega}{2}(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})|0\rangle = \frac{\omega}{2}|0\rangle.$$

The state  $|n\rangle$  we introduce as  $|n\rangle = (\hat{a}^\dagger)^n|0\rangle$ ; its energy is given by

$$\hat{H}(\hat{a}^\dagger)^n|0\rangle = \omega(n + 1/2)|n\rangle.$$

As we know, such a construction is very successful in describing non-relativistic quantum phenomena (spectra of atoms, molecules, nuclei etc).

But there are well-known problems:

- In experiments we have not only particle creation and annihilation but also production of new particles and antiparticles.
- Relativity and causality might be in conflict with quantum principles.  
If in four-dimensional areas  $X$  and  $Y$  points are separated by the space-like interval  $(x - y)^2 < 0$ , events in the points  $x$  and  $y$  are causally independent. But in QM we have the uncertainty principle,  $\Delta p \Delta x \geq 1$ , and in some regions of the order of  $\Delta x \sim 1/m$  close to indistinct boundaries of the areas  $X$  and  $Y$  the causality might be violated.

Quantum field theory allows us to resolve both of these problems simultaneously.

To describe a particle and an antiparticle with mass  $m$ , momentum  $\vec{k}$  and energy  $\omega_k = k^0 = \sqrt{\vec{k}^2 + m^2}$ , let us consider two sets of creation and annihilation operators for each momentum point  $\vec{k}$   $\hat{a}, \hat{a}^\dagger$  and  $\hat{b}, \hat{b}^\dagger$ . Vacuum is defined by the requirements  $\hat{a}|0\rangle = \hat{b}|0\rangle = 0$ . The Hamiltonian for every momentum point  $\vec{k}$  obviously has the following form:

$$H_k = \frac{\omega_k}{2} (\hat{a}(\vec{k})\hat{a}^\dagger(\vec{k}) + \hat{a}^\dagger(\vec{k})\hat{a}(\vec{k}) + \hat{b}(\vec{k})\hat{b}^\dagger(\vec{k}) + \hat{b}^\dagger(\vec{k})\hat{b}(\vec{k})),$$

where  $\omega_k = k^0 = \sqrt{\vec{k}^2 + m^2}$ . Please note that the mass parameter  $m$  is the same for all oscillators with different  $\vec{k}$ .

One can use a different normalization of an integral measure in order to get the commutation relations:

$$\int \overline{dk} [\hat{a}(\vec{k})\hat{a}^\dagger(\vec{k})] = 1; \quad \int \overline{dk} [\hat{b}(\vec{k})\hat{b}^\dagger(\vec{k})] = 1.$$

We are using

$$\overline{dk} = \frac{d^3\vec{k}}{(2\pi)^3 2\omega_k} \Rightarrow [\hat{a}(\vec{k})\hat{a}^\dagger(\vec{k}')] = (2\pi)^2 2\omega_k \delta(\vec{k} - \vec{k}').$$

Total momentum and charge operators taken in so-called ‘normal ordering’ have the following form:

$$\begin{aligned} \hat{P}^\mu &= \int \overline{dk} k^\mu [\hat{a}^\dagger(\vec{k})\hat{a}(\vec{k}) + \hat{b}^\dagger(\vec{k})\hat{b}(\vec{k})], \\ \hat{Q} &= \int \overline{dk} [\hat{a}^\dagger(\vec{k})\hat{a}(\vec{k}) - \hat{b}^\dagger(\vec{k})\hat{b}(\vec{k})]. \end{aligned}$$

One can prove the following commutator relations, which clarify the meaning of the operators:

$$\begin{aligned} [\hat{P}^\mu, \hat{a}^\dagger(\vec{k})] &= k^\mu \hat{a}^\dagger(\vec{k}); \quad [\hat{P}^\mu, \hat{a}(\vec{k})] = -k^\mu \hat{a}(\vec{k}), \\ [\hat{Q}, \hat{a}^\dagger(\vec{k})] &= \hat{a}^\dagger(\vec{k}); \quad [\hat{Q}, \hat{b}^\dagger(\vec{k})] = -\hat{b}^\dagger(\vec{k}). \end{aligned}$$

Now one can construct the field operator:

$$\hat{\Phi}(x) = \int \overline{dk} [e^{-ikx} \hat{a}(\vec{k}) + e^{ikx} \hat{b}^\dagger(\vec{k})].$$

The momentum operator acts on the field operator leading to a coordinate translation:

$$e^{i\hat{P}_\mu y^\mu} \hat{\Phi}(x) e^{-i\hat{P}_\mu y^\mu} = \hat{\Phi}(x + y).$$

Indeed, one may prove this in a very simple way. Let us introduce an operator  $\hat{A}(\alpha)$  depending on some numerical parameter  $\alpha$ :

$$e^{i\alpha y^\mu \hat{P}_\mu} \hat{a}(k) e^{-i\alpha y^\mu \hat{P}_\mu} = \hat{A}(\alpha).$$

The operator obeys the following equation:

$$\frac{dA}{d\alpha} = iy^\mu \widehat{P}^\mu \widehat{A} - iy^\mu \widehat{P}^\mu \widehat{A} = -iy^\mu [\widehat{P}^\mu \widehat{A}].$$

One finds a solution in the form  $\widehat{A} = \widehat{a}(\vec{k})f(\alpha)$ , from which one can immediately see the needed translation relation:

$$\begin{aligned} \implies \widehat{a}(k) \frac{df}{d\alpha} &= -iy^\mu f(\alpha) k^\mu \widehat{a}(k) \implies \frac{df}{d\alpha} = -iy^\mu k^\mu f(\alpha) \implies \\ e^{i\alpha y^\mu \widehat{P}^\mu} \widehat{a}(k) e^{-i\alpha y^\mu \widehat{P}^\mu} &= \widehat{a}(k) e^{-ik^\mu y_\mu} \implies \widehat{\Phi}(x) = \widehat{\Phi}(x + y). \end{aligned}$$

It is very important to note that such a field operator obeys the Klein–Gordon equation:

$$[\square^2 + m^2] \widehat{\Phi}(x) = \int \overline{dk} \left[ e^{-ikx} (-k^2 + m^2) \widehat{a}(k) + e^{ikx} (-k^2 + m^2) \widehat{b}(k) \right] = 0$$

because of

$$-k^2 + m^2 = -k_0^2 + \vec{k}^2 + m^2 = -\omega_k^2 + \omega_k^2 = 0$$

What charge has the state created by the operator  $\widehat{\Phi}(x)$ ? Let us act on the vacuum state by the  $\widehat{\Phi}(x)$  operator:  $\widehat{\Phi}(x)|0\rangle$ . This state has the following charge:

$$\widehat{Q} \widehat{\Phi}(x)|0\rangle = -\widehat{\Phi}(x)|0\rangle.$$

In the same way, one gets

$$\widehat{Q} \widehat{\Phi}^\dagger(x)|0\rangle = \widehat{\Phi}^\dagger(x)|0\rangle.$$

This means that the field operator  $\widehat{\Phi}(x)$  acting on vacuum produces the state with the negative charge  $(-1)$  and the field operator  $\widehat{\Phi}^\dagger(x)$  produces the state with the positive charge  $(+1)$ .

Now let us consider two space–time points  $x_1$  and  $x_2$ :

$$\begin{array}{ccc} \bullet & & \bullet \\ & & \\ t_1, \vec{x}_1 & & t_2, \vec{x}_2 \end{array}$$

and two-point correlation functions—products of field operators between vacuum states. If  $t_1 < t_2$ , the operator  $\widehat{\Phi}(x_1)|0\rangle$  in the correlator

$$\langle 0 | \widehat{\Phi}^\dagger(x_2) \widehat{\Phi}(x_1) | 0 \rangle$$

creates the charge  $-1$  at  $t_1$ , and the operator  $\widehat{\Phi}^\dagger(x_2) \widehat{\Phi}(x_1)|0\rangle$  annihilates this charge at  $t_2$ . So, charge  $-1$  propagates from the point  $x_1$  to  $x_2$  and  $t_2 > t_1$ . If  $t_2 < t_1$ , the operator  $\widehat{\Phi}^\dagger(x_2)|0\rangle$  in the correlator

$$\langle 0 | \widehat{\Phi}(x_1) \widehat{\Phi}^\dagger(x_2) | 0 \rangle$$

creates the charge  $+1$  at  $t_2$ , and  $\widehat{\Phi}(x_1) \widehat{\Phi}^\dagger(x_2)|0\rangle$  annihilates this charge at  $t_1$ . So, charge  $+1$  propagates from the point  $x_2$  to  $x_1$  and  $t_2 < t_1$ .

Since both these actions do not change the vacuum, we should take both correlators into account to see the causal relation of events in points  $x_1$  and  $x_2$ :

$$\langle 0 | T \{ \widehat{\Phi}^\dagger(x_2) \widehat{\Phi}(x_1) \} | 0 \rangle$$

$$\begin{aligned}
&= \langle 0 | \widehat{\Phi}^\dagger(x_2) \widehat{\Phi}(x_1) | 0 \rangle \Theta(t_2 - t_1) + \langle 0 | \widehat{\Phi}(x_1) \widehat{\Phi}^\dagger(x_2) | 0 \rangle \Theta(t_1 - t_2) \\
&= i \int \frac{d^4 k}{(2\pi)^4} \frac{e^{-ik(x_2 - x_1)}}{k^2 - m^2 + i0} = D_c(x_2 - x_1).
\end{aligned}$$

The function

$$D_c(x) = i \int \frac{d^4 k}{(2\pi)^4} \frac{e^{-ikx}}{k^2 - m^2 + i0} \quad (5)$$

is called the Feynman propagator. Obviously, the Feynman propagator is a Green function of the Klein-Gordon equation. One can check that all the commutators between the field operators in the points  $x$  and  $y$  separated by the space-like interval  $(x - y)^2 < 0$  are equal to zero. So, the causality takes place. Also, one can construct multiparticle states by acting of the creation operators (operators have different quantum numbers corresponding to different kinds of particles) on the vacuum state  $|\vec{k}_1, \dots, \vec{k}_n\rangle = \prod_{i=1}^n \hat{a}^\dagger(\vec{k}_i) |0\rangle$ . The energy and momentum of the states are then obtained by acting of the operators

$$\begin{aligned}
\hat{P}^0 |\vec{k}_1, \dots, \vec{k}_n\rangle &= \hat{H} |\vec{k}_1, \dots, \vec{k}_n\rangle = \left( \sum_{i=1}^n k_i^0 \right) |\vec{k}_1, \dots, \vec{k}_n\rangle, \\
\hat{\vec{P}} |\vec{k}_1, \dots, \vec{k}_n\rangle &= \left( \sum_{i=1}^n \vec{k}_i \right) |\vec{k}_1, \dots, \vec{k}_n\rangle.
\end{aligned}$$

Of course, one can get the same results using the usual canonical quantization with the correspondence

$$\begin{aligned}
q(t), \dot{q}(t), L(q, \dot{q}), S &= \int dt L(q, \dot{q}) \iff \\
\varphi(x), \partial_\mu \varphi(x), L(\varphi, \partial_\mu \varphi), S &= \int d^4 x L(\varphi, \partial_\mu \varphi).
\end{aligned}$$

The field momentum is then

$$\pi(x) = \frac{\partial L}{\partial(\partial_0 \varphi(x))}; \quad \partial_0 \varphi(x) = \dot{\varphi}.$$

The Lagrangian of the complex scalar field has the form

$$\begin{aligned}
L &= \partial_\mu \varphi^\dagger \partial^\mu \varphi - m^2 \varphi^\dagger \varphi, \\
\pi(x) &= \dot{\varphi}^\dagger(x); \quad \pi^\dagger(x) = \dot{\varphi}(x).
\end{aligned}$$

In the same way, as in classical mechanics, the equation of motion comes from the principle of least action:

$$\frac{\partial L}{\partial \varphi} = \partial_\mu \frac{\partial L}{\partial \partial_\mu \varphi} \longrightarrow (\square - m^2)\varphi = 0.$$

The Feynman propagator  $D_c$  introduced above is a Green function of the equation of motion. For the field and momentum operators, one naturally assumes the equal-time commutation relation

$$[\hat{\pi}(\vec{x}, t), \hat{\varphi}(\vec{x}', t')] \Big|_{t=t'} = -i\delta(\vec{x} - \vec{x}').$$

The Lagrangian is invariant under a global phase shift:

$$\varphi(x) \rightarrow e^{i\alpha\varphi(x)}, \quad \alpha \equiv \text{Const.}$$

Such a combination, called a current, is conserved (this is a simple example of the first Noether theorem):

$$j^\mu(x) = i\partial^\mu \varphi^\dagger \varphi - i\varphi^\dagger \partial^\mu \varphi,$$



$$\partial_\mu j^\mu = i\Box\varphi^\dagger\varphi + i\partial^\mu\varphi^\dagger\partial_\mu\varphi - i\partial_\mu\varphi^\dagger\partial^\mu\varphi - i\varphi^\dagger\Box\varphi = 0.$$

Conservation of the current leads to the conservation of the charge:

$$Q = \int d^3\vec{x}j^0 = \int d\vec{x}(i\dot{\varphi}^\dagger\varphi - i\varphi^\dagger\dot{\varphi}),$$

$$\frac{dQ}{dt} = \int d\vec{x}\partial_0j^0 = \int d\vec{x}(\partial_i j^i) = \int_\Omega d\vec{n}\vec{j} = 0$$

for falling-off fields.

### 3 Functional integral in quantum mechanics

However, for our further consideration, the functional integral approach to quantum field theory is more useful. In particular, it allows us to quantize non-abelian gauge field theories, to clarify better boundary conditions and renormalization procedure and to get a reduction formula (connection between S-matrix elements and the Green functions).

Once more we begin with the quantum mechanics as a simple example.

$$L(q_i, \dot{q}_i) \rightarrow p_i = \frac{\partial L}{\partial \dot{q}_i}, \quad H(q_i, p_i) = \dot{q}_i p_i - L(q_i, \dot{q}_i)|_{\dot{q}=f(t)}$$

$$[\hat{q}^i(t), \hat{p}^j(t)] = i\hbar\delta^{ij}\hat{1}.$$

Let us consider a simple system described by the non-relativistic Hamiltonian

$$H(\hat{p}, \hat{q}) = \frac{\hat{p}^2}{2m} + V(\hat{q}).$$

In the Schrödinger picture, the evolution of a quantum system follows from the Schrödinger equation

$$i\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle.$$

The formal solution of the Schrödinger equation is

$$|\Psi(t)\rangle = e^{-i\hat{H}t}|\Psi(0)\rangle.$$

One can define states:

$$|q, t\rangle : \hat{q}|q, t\rangle = q|q, t\rangle.$$

Then the wave function coordinate representation is

$$\Psi(q, t) = \langle q|\Psi(t)\rangle = \left\langle q\left|e^{-i\hat{H}t}\right|\Psi(0)\right\rangle \text{ and } \langle q|q'\rangle = \delta(q - q').$$

If we introduce a complete set of states  $|q_0\rangle$  such that

$$\hat{1} = \int dq_0 |q_0\rangle\langle q_0|,$$

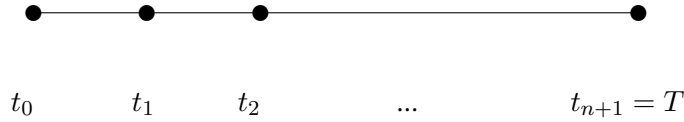
we can write

$$\Psi(q, t) = \int dq_0 \langle q|e^{-i\hat{H}t}|q_0\rangle\langle q_0|\Psi(0)\rangle = \int dq_0 K(q, q_0, t)\langle q_0|\Psi(0)\rangle,$$

where  $K$  is the so-called kernel of the Schrödinger equation.

Obviously,  $[\hat{H}, \hat{H}] = 0$  and therefore

$$e^{-i\hat{H}T} = e^{-i\hat{H}(t_{n+1}-t_n)} \cdot e^{-i\hat{H}(t_n-t_{n-1})} \dots e^{-i\hat{H}(t_1-t_0)}.$$



At each time moment  $t_i$  we can introduce a unity operator

$$\hat{1} = \int dq_i |q_i\rangle \langle q_i|.$$

Then, for the kernel  $K$ , we obtain the following formula:

$$K(q, q_0, T - t_0) = \int \lim_{n \rightarrow \inf} \prod_{i=1}^n dq_i |q_{n+1-i}\rangle \langle q_i | e^{-i\hat{H}\delta t} |q_{n-i}\rangle \dots \langle q_1 | e^{-i\hat{H}\delta t} |q_0\rangle,$$

$$e^{\varepsilon(\hat{A}+\hat{B})} = e^{\varepsilon\hat{A}} + e^{\varepsilon\hat{B}}(1 + o(\varepsilon^2)),$$

$$\langle q_{i+1} | e^{-i\hat{H}\delta t} |q_i\rangle = \langle q_{i+1} | e^{-i\frac{\hat{p}^2}{2m}\delta t} \cdot e^{-i\hat{V}(q)\delta t} |q_i\rangle.$$

For very small  $(\delta t)$ ,  $e^{-i\hat{V}(q)\delta t}$  could be factorized out, and therefore

$$\langle q_{i+1} | e^{-i\hat{H}\delta t} |q_i\rangle \approx e^{-i\hat{V}(q_i)\delta t} \cdot \langle q_{i+1} | e^{-i\frac{\hat{p}^2}{2m}\delta t} |q_i\rangle.$$

The last term can be expressed in the following way:

$$\langle q_{i+1} | e^{-i\frac{\hat{p}^2}{2m}\delta t} |q_i\rangle = \int \frac{dp}{2\pi} \langle q_{i+1} | e^{-i\frac{\hat{p}^2}{2m}\delta t} |p\rangle \langle p|q_i\rangle = \int \frac{dp}{2\pi} e^{-ip(q_{i+1}-q_i)\delta t} e^{-i\frac{p^2}{2m}\delta t},$$

where  $\langle q|p\rangle = e^{ipq}$ ,  $\langle p|q\rangle = e^{-ipq}$  and  $\hat{1} = \int \frac{dp}{2\pi} |p\rangle \langle p|$ .

One could make the following substitution of the integration variable:

$$p' = \left[ \frac{p}{\sqrt{2m}} - \sqrt{2m}(q_{i+1} - q_i) \frac{1}{2} \right] \sim e^{i\frac{m}{2}(q_{i+1}-q_i)^2} \frac{1}{\delta t}.$$

Then, for the kernel  $K$ , one gets

$$K(q, q_0; t) = N \int dq_i \prod_i e^{i \left[ \frac{m}{2} \left( \frac{q_{i+1}-q_i}{\delta t} \right)^2 - V(q_i) \right] \delta t} = N \int_{Dq} e^{i \int_0^t dt \left[ \frac{mv^2}{2} - V(q) \right]} = \int e^{iS} D(q).$$

For our consideration it is not needed, but if one takes the integral of  $dp$ , one gets the following representation for the functional integral measure:

$$D(q) = \lim_{n \rightarrow \inf, \delta t = \frac{t}{n} \rightarrow 0} \sqrt{\frac{m}{2\pi i \delta t}} \prod_{T=1}^n \left( \sqrt{\frac{m}{2\pi i \delta t}} dq_T \right).$$

We do not discuss mathematical aspects of how well such a construction is determined.

The formula for the kernel, being written as

$$K(q, q_0; t) = \int D(q) e^{iS} = \int D(q) e^{i \int_0^t dt L(q, \dot{q}, t)}, \quad (6)$$

can be generalized to the case of quantum field theory. For that we need to recall a few simple but important formulas for multidimensional Gaussian integrals.

#### 4 Gaussian integrals

Let us consider  $n$  coordinates  $y_1, \dots, y_n$  as a formal vector

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Obviously,  $y^T = (y_1 \dots y_n)$  with a formal definition of ‘scalar’ product

$$y^T \cdot x = \sum_i y_i x_i.$$

The well-known answer for the Gaussian integral has the form

$$Z = \int dx_1 \dots dx_n \cdot e^{-\frac{1}{2}x^T A x} = \frac{(2\pi)^{n/2}}{\sqrt{\det A}}, \quad (7)$$

where  $A$  is a positive-definite  $n \times n$  matrix.

*Problem.* Take the integral and obtain the above formula.

*Reminder:*  $\det A = \prod_{i=1}^n \lambda_i$ , where  $\lambda_i$  is the eigenvalue of the matrix  $A$ .

The integral (7) is an analogue of the integral in a field theory without external sources, as we shall see in the next part. An analogue of the functional integral with a source is as follows:

$$Z[J] = \int dx_1 \dots dx_n \cdot e^{-\frac{1}{2}x^T A x + J^T x}, \quad (8)$$

where  $J$  is some vector  $J^T = (J_1 \dots J_n)$ . If we make a substitution of integration variable

$$x' = x - (A^{-1})J,$$

the integral (8) will take the form

$$Z[J] = e^{\frac{1}{2}J^T A^{-1} J} \cdot \int dx'_1 \dots dx'_n \cdot e^{-\frac{1}{2}x'^T A x'}.$$

So, the integral  $Z[J]$  is given by

$$Z[J] = e^{\frac{1}{2}J^T A^{-1} J} \cdot Z = e^{\frac{1}{2}J^T A^{-1} J} \cdot \frac{(2\pi)^{n/2}}{\sqrt{\det A}}.$$

Generalization of the above formula for the case of complex variables of integration is straightforward.

Let  $z = x + iy$  and  $z^* = x - iy$ . We need to compute the integral

$$Z_C = \int \prod_{k=1}^n dz_k^* dz_k e^{-z^\dagger B z},$$

where  $B$  is a Hermitian matrix and  $z^\dagger = (z^*)^T$ . One can diagonalize the quadratic form  $z^{*\dagger} B z$  by applying a unitary transformation of variables  $z' = U \cdot z$  such that the matrix  $UBU^\dagger$  becomes diagonal:

$$UBU^\dagger = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

The integral  $Z_C$  then takes the form

$$Z_C = \int \prod_{k=1}^n dz_k'^* dz_k' e^{-\lambda_k |z_k'|^2} = \int \prod_{k=1}^n dx_k dy_k e^{-\lambda_k (x_k^2 + y_k^2)} = \frac{\pi^n}{\det B}. \quad (9)$$

If we add the external complex ‘source’  $J$ ,

$$Z_C[J] = \int \prod_{k=1}^n dz_k'^* dz_k' e^{-z^\dagger B z + J^\dagger z + z^\dagger J},$$

with the shift of variables of integration we get

$$Z_C[J] = e^{J^\dagger B^{-1} J} \cdot Z_C = e^{J^\dagger B^{-1} J} \cdot \frac{\pi^n}{\det B}. \quad (10)$$

In the field theory we have to consider interacting fields. So, we need to consider more complicated integrals involving source ‘interactions’:

$$Z_{\text{int}}[J] = \frac{1}{Z_{\text{int}}} \int \prod dx e^{-\frac{1}{2} x A x + J x - V(x)},$$

where  $Z_{\text{int}} = Z_{\text{int}}[0]$ .

If we expand the exponent  $e^{-V(x)}$  in ‘perturbation’ theory we can easily get the following form for the integral  $Z_{\text{int}}[J]$ :

$$Z_{\text{int}}[J] = \frac{1}{Z[0]} e^{-V[\frac{\partial}{\partial J}]} \cdot Z[J] = e^{-V[\frac{\partial}{\partial J}]} \cdot e^{\frac{1}{2} J A^{-1} J}.$$

## 5 Functional integral in quantum field theory

As we have seen already, a transition from mechanics to a field theory could be done by means of a formal correspondence between the coordinate and its derivative and the field  $\varphi(x)$  and its derivative:

$$q(t) \rightarrow \varphi(x); \quad \dot{q}(t) \rightarrow \partial_\mu \varphi(x).$$

With this analogy one can immediately write down the following formula for the evolution kernel in the case of quantum field theory:

$$Z[J] = \int D(\varphi) e^{i \int d^4 x L(\varphi, \partial_\mu \varphi) + i \int d^4 x J(x) \varphi(x)}, \quad (11)$$

where the measure  $D(\varphi) = \prod_x d\varphi(x)$  corresponds to the integration over all possible trajectories (field configurations).

Now all the formulas we derived for Gaussian integrals in the previous section can be applied here using the functional derivative instead of the usual one. For example,

$$\frac{\delta J(y)}{\delta J(x)} = \delta^{(4)}(x - y).$$

If we consider the Lagrangian for the free scalar field

$$L = \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - \frac{1}{2} m^2 \varphi^2 = -\frac{1}{2} \varphi (\square^2 + m^2) \varphi \equiv -\frac{1}{2} \varphi D_c^{-1} \varphi, \quad (12)$$

we can get for  $Z[J]$ ,

$$Z[J] = \exp \left( \frac{1}{2} \int d^4 x d^4 y J(x) D_c(x - y) J(y) \right), \quad (13)$$

where the normalization of the measure is taken such that

$$Z[0] = 1.$$

Here the function  $D_c$  is the Green function of the equation of motion:

$$D_c^{-1} \cdot D_c = 1,$$

which more accurately means that

$$i(\square^2 + m^2)_x D_c(x - y) = \delta^{(4)}(x - y).$$

In the momentum representation, by taking the Fourier transform of both sides of this equation, one gets

$$i(-p^2 + m^2) D(p) = 1.$$

The formal solution of the equation is

$$D(p) = \frac{i}{p^2 - m^2}.$$

But we need to fix how to deal with the pole. The only possible choice is to add  $+i\varepsilon$ . In this case the function  $D_c(p)$  has the familiar form of the Feynman propagator:

$$D_c(p) = \frac{i}{p^2 - m^2 + i\varepsilon}.$$

Indeed, such fixing of the denominator leads to the fact that in the expression for the functional integral

$$Z[J] = \int D(\varphi) \exp\left(i \int d^4x \left[ \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - \frac{1}{2} m^2 \varphi^2 + J\varphi \right]\right), \quad (14)$$

$m^2$  gets a shift  $m^2 \rightarrow m^2 - i\varepsilon$  and the term

$$\int D(\varphi) \exp\left(-\varepsilon \int dx \varphi^2(x)\right)$$

ensures the convergence of the integral.

On the other hand, as we have discussed already, such a form of the Feynman propagator leads to Feynman boundary conditions, namely if  $x^0 < y^0$  the particle propagates from  $\vec{x}$  to  $\vec{y}$ , and if  $x^0 > y^0$  the corresponding antiparticle propagates from  $\vec{x}$  to  $\vec{y}$ , as follows from the expression

$$D_c(x - y) = - \int \frac{d\vec{p}}{(2\pi)^3 2\omega_p} e^{i\vec{p}(\vec{x}-\vec{y})} [\Theta(t)e^{-i\omega_p t} + \Theta(-t)e^{i\omega_p t}].$$

So, the Feynman propagator is a Green function and, in other words, the inverse quadratic form in the action (12).

In the case of an interacting potential  $V(\varphi)$  we get from (13) the following general expression for the generating functional:

$$Z_V[J] = \exp\left(-i \int d^4x V\left(\frac{\delta}{i\delta J(x)}\right)\right) \cdot \exp\left(\frac{1}{2} \int dy dz J(y) D_c(y - z) J(z)\right). \quad (15)$$

As will be discussed, one can get the Green functions by taking the needed number of functional derivatives with respect to the source. However, this is not enough, since we do not know how exactly the functional integral, called the generalized functional integral, and the Green functions are related to the S-matrix elements needed to compute physics observables.

## 6 Functional integral in holomorphic representation

We start with the harmonic oscillator as we did before:

$$H(p, q) = \frac{\hat{p}^2}{2} + \frac{\omega^2 \hat{q}^2}{2}.$$

The creation and annihilation operators have the form (4) with the commutator

$$[\hat{a}, \hat{a}^\dagger] = 1. \quad (16)$$

The Hamiltonian of the system taken in a normal form (all creation operators are on the right-hand side) is

$$H = \omega \hat{a}^\dagger \hat{a}.$$

The commutator relation (16) has a very nice representation in terms of holomorphic functions, which are introduced by means of the following scalar product:

$$\langle f_1 | f_2 \rangle = \int (f_1(a^*))^* f_2(a^*) e^{-a^* a} \frac{da^* da}{2\pi i}.$$

With such a definition of the scalar product, the set of functions  $\Psi_n(a^*) = \frac{(a^*)^n}{\sqrt{n!}}$ ,  $n \geq 0$ , forms an orthonormal basis

$$\langle \Psi_n | \Psi_m \rangle = \frac{1}{\sqrt{n!m!}} \int a^n (a^*)^m e^{-a^* a} \frac{da^* da}{2\pi i} = \delta_{nm}. \quad (17)$$

One can easily prove that  $\sum_n |\Psi_n\rangle \langle \Psi_n| = 1$ .

*Problem.* Prove the relation (17).

The operators  $\hat{a}^\dagger$  and  $\hat{a}$  act according to the following rules:

$$\hat{a}^\dagger \cdot f(a^*) = a^* f(a^*), \quad \hat{a} f(a^*) = \frac{d}{da^*} f(a^*). \quad (18)$$

By direct substitution, one can prove the following relation:

$$\langle f_1 | \hat{a}^\dagger f_2 \rangle = \langle \hat{a} f_1 | f_2 \rangle, \quad (19)$$

which means that the operators  $\hat{a}^\dagger$  and  $\hat{a}$  are conjugate to each other.

Now we will show a few simple formulas for the holomorphic representation given above, which are useful for a construction of the S-matrix.

Let us take some operator  $\hat{A}$  with matrix element in our basis

$$A_{nm} = \langle \Psi_n | \hat{A} | \Psi_m \rangle. \quad (20)$$

The function

$$A(a^*, a) = \sum_{nm} A_{nm} \frac{(a^*)^n}{\sqrt{n!}} \frac{a^m}{\sqrt{m!}} = \sum_{nm} |n\rangle \langle m| \quad (21)$$

is called the kernel of the operator  $\hat{A}$ . The kernel of a product of two operators is given by the convolution of kernels:

$$A_1 A_2(a^* a) = \int A_1(a^* \alpha) A_2(\alpha^* a) e^{-\alpha^* \alpha} \frac{d\alpha^* d\alpha}{2\pi i}.$$

The operator  $\hat{A}$  can be decomposed into a formal series of normal ordered creation and annihilation operators:

$$\hat{A} = \sum_{nm} K_{nm} (\hat{a}^\dagger)^n (\hat{a})^m. \quad (22)$$

The following function is called the normal symbol of the operator  $\hat{A}$ :

$$K(a^*, a) = \sum_{nm} K_{nm} (a^*)^n a^m. \quad (23)$$

*Problem.* Prove the relation between the kernel and the normal symbol of the operator  $\hat{A}$ :

$$A(a^*, a) = e^{a^* a} K(a^*, a). \quad (24)$$

(Check the equality (24) for the particular case  $\hat{A} = \hat{a}^{\dagger n} \hat{a}^l$ .)

Now we use the relations (22), (23) and (24) to construct the functional integral in the holomorphic representation.

Let the Hamiltonian of some system be  $\hat{H}(\hat{a}^\dagger, a)$ . The evolution operator has the form

$$\hat{U} = e^{-i\hat{H} \cdot \Delta t}.$$

From (22) and (24), one can get the following formula for the kernel of the evolution operator:

$$U(a^*, a) = e^{[a^* a - i h(a^*, a)] \Delta t} \quad (25)$$

for a small time interval  $\Delta t$ .

In the case of a finite interval, we can split it into small pieces  $t'' - t' = N \cdot \Delta t$  and using our orthonormal bases (17) we get the following form for the normal symbol of the evolution operator, which is a convolution of products of the evolution operators:

$$U(a^*, a; t'', t') = \int \exp \left( [a^* \alpha_{N-1} - \alpha_{N-1}^* \alpha_{N-1} + \dots - \alpha_1^* \alpha_1 + \alpha_1^* \alpha_0] \right. \\ \left. - i \Delta t [h(a^*, \alpha_{N-1}) + \dots + h(\alpha^*, \alpha_0)] \right) \cdot \prod_{k=1}^{N-1} \frac{d\alpha_k^* d\alpha_k}{2\pi i}.$$

In the limit  $\Delta t \rightarrow 0$ ,  $N \rightarrow \infty$ ,  $\Delta N = t'' - t'$ , one gets

$$U(a^*, a; t'', t') = \int e^{a^* \alpha(t'')} \cdot \exp \left( \int_{t'}^{t''} [-\alpha^* \alpha - i h(\alpha^*, \alpha)] dt \right) \cdot \prod_t \frac{d\alpha^* d\alpha}{2\pi i}, \quad (26)$$

where the boundary conditions are  $\alpha^*(t'') = a^*$ ,  $\alpha(t') = a$ .

In our case for the harmonic oscillator  $h(a^*, a) = \omega a^* a$  the integral (26) can be easily computed. To do this, one should take a variation

$$\delta \left[ a^* \alpha(t'') + \int_{t'}^{t''} [-\alpha^* \alpha - i h(\alpha^*, \alpha)] dt \right] \\ = a^* \delta \alpha(t'') + \int_{t'}^{t''} dt [-\delta \alpha^* \dot{\alpha} - \alpha^* \delta \dot{\alpha} - i \omega \alpha^* \delta \alpha - i \omega \delta \alpha^* \alpha] \\ = a^* \delta \alpha(t'') - \alpha^*(t'') \delta \alpha(t'') + \int_{t'}^{t''} dt [-\delta \alpha^* (\dot{\alpha} + i \omega \alpha) + \delta \alpha (\dot{\alpha}^* - i \omega \alpha^*)].$$

The extremum condition gives us the answer. Note that the first two terms cancel each other because of the boundary condition  $\alpha^*(t'') = a^*$ . Extremum conditions can be simply solved:

$$a(t) = e^{-i\omega(t-t')} a(t'), \quad a^*(t) = e^{-i\omega(t''-t)} a^*(t'). \quad (27)$$

Then, for the normal symbol of the evolution operator, we get

$$U(a^*, a; t'', t') = \exp(a^* a e^{-i\omega(t-t')}), \quad (28)$$

substituting the solution (27) into the exponent in (26) where a non-zero contribution comes only from the first term in the exponent.

An important consequence of this is that for the operator  $\hat{A}$  with the kernel  $A(a^*, a)$  the kernel of the operator  $e^{i\hat{H}_0 t''} \hat{A} e^{-i\hat{H}_0 t'}$  is given by the convolution of corresponding kernels and is equal to

$$A(a^* e^{i\omega t''}, a e^{-i\omega t'}). \quad (29)$$

This relation demonstrates the power of the holomorphic representation in which the evolution is simply reduced to the substitution of arguments:

$$a \rightarrow a e^{-i\omega t}.$$

Now we come back to the field theory. The field operator is given by

$$\hat{\Phi}(x) = \int \frac{d\vec{k}}{(2\pi)^3 2k^0} [e^{-ikx} \hat{a} + e^{ikx} \hat{a}^\dagger(k)]$$

and the corresponding 4-momentum operator is

$$\hat{P}^\mu = \int \frac{d\vec{k}}{(2\pi)^3 2k^0} k^\mu \hat{a}^\dagger(k) \hat{a}(k).$$

Vacuum  $|0\rangle$  is the state  $\hat{a}(k)|0\rangle = 0$  and the one-particle state is  $|\vec{k}\rangle = \hat{a}^\dagger(k)|0\rangle$ . The operator  $\hat{P}^\mu$  acts on the one-particle state as

$$(\hat{P}^0 = \hat{H}) |\vec{k}\rangle = k^0 |\vec{k}\rangle, \quad \hat{P} |\vec{k}\rangle = \vec{k} |\vec{k}\rangle.$$

A multiparticle state is constructed as  $|\vec{k}_1, \dots, \vec{k}_n\rangle = \prod_{i=1}^n \hat{a}^\dagger(\vec{k}_i) |0\rangle$ . Obviously,

$$\begin{aligned} \hat{H} |\vec{k}_1, \dots, \vec{k}_n\rangle &= \left( \sum_{i=1}^n k_i^0 \right) |\vec{k}_1, \dots, \vec{k}_n\rangle, \\ \hat{P} |\vec{k}_1, \dots, \vec{k}_n\rangle &= \left( \sum_{i=1}^n \vec{k}_i \right) |\vec{k}_1, \dots, \vec{k}_n\rangle. \end{aligned}$$

If  $\hat{H}$  is the Hamiltonian of a system and  $\hat{H}^0$  is a free Hamiltonian, then the S-matrix is determined as the following limit of the evolution operator:

$$\hat{S} = \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} e^{i\hat{H}_0 t''} \hat{U}(t'', t') e^{-i\hat{H}_0 t'}. \quad (30)$$

With such a definition, it is clear that  $\hat{S} = 1$  if  $\hat{H} = \hat{H}^0$ .

From the formulas (26) and (29), we can obtain the following representation for the kernel of the S-matrix:

$$S(a^*, a) = \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} e^{\int \frac{d\vec{k}}{(2\pi)^3 2k^0} (\alpha^*(\vec{k}, t'') \alpha(\vec{k}, t') - \int_{t'}^{t''} [\alpha^*(\vec{k}, t) \dot{\alpha}(\vec{k}, t) + h(\alpha^*, \alpha)] dt)}. \prod_{t, \vec{k}} \frac{d\alpha^* d\alpha}{2\pi i}, \quad (31)$$



where  $\alpha^*(\vec{k}, t'') = a^*(\vec{k})e^{i\omega t''}$ ,  $\alpha(\vec{k}, t') = a(\vec{k})e^{-i\omega t'}$ .

Now we consider the system with an external source  $J(x)$ :

$$L = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 + J(x)\varphi.$$

The kernel of the interaction operator  $V = -J(x)\hat{\varphi}(x)$  is given by

$$V(a^*, a) = \int \frac{d\vec{k}}{(2\pi)^3 2k^0} \left[ \gamma(\vec{k}, t)a^*(\vec{k}) + \gamma^*(\vec{k}, t)a(\vec{k}) \right],$$

where  $\gamma(\vec{k}, t) = -\int J(\vec{x}, t)e^{-i\vec{k}\cdot\vec{x}}d\vec{x}$ . In order to take the integral (31) once more, we should solve the extremum conditions

$$\begin{aligned} \dot{\alpha}(\vec{k}, t) + i\omega(\vec{k})\alpha(\vec{k}, t) + i\gamma(\vec{k}, t) &= 0, \\ \dot{\alpha}^*(\vec{k}, t) - i\omega(\vec{k})\alpha^*(\vec{k}, t) - i\gamma^*(\vec{k}, t) &= 0 \end{aligned} \quad (32)$$

with the boundary conditions

$$\alpha^*(\vec{k}, t'') = a^*(\vec{k})e^{i\omega t''}, \quad \alpha(\vec{k}, t') = a(\vec{k})e^{-i\omega t'}. \quad (33)$$

The solution of these equations has the following form:

$$\begin{aligned} \alpha(\vec{k}, t) &= a(\vec{k})e^{-i\omega t} - ie^{-i\omega t} \int_{t'}^{t''} e^{i\omega s} \gamma(\vec{k}, s) ds, \\ \alpha^*(\vec{k}, t) &= a^*(\vec{k})e^{i\omega t} + ie^{i\omega t} \int_{t'}^{t''} e^{-i\omega s} \gamma^*(\vec{k}, s) ds. \end{aligned}$$

If one substitutes the obtained solutions into the integral (31), one will get the following formula for the kernel of the S-matrix:

$$\begin{aligned} S_J(a^*, a) &= \exp \left( \int \frac{d\vec{k}}{(2\pi)^3 2k^0} \left[ a^*(\vec{k})a(\vec{k}) \right. \right. \\ &+ \int dt \int d\vec{x} J(\vec{x}, t) \left( a^*(\vec{k})e^{i\omega t - i\vec{k}\cdot\vec{x}} + a(\vec{k})e^{-i\omega t + i\vec{k}\cdot\vec{x}} \right) / (2\omega) \\ &\left. \left. - \frac{1}{2} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} ds \int d\vec{x} d\vec{y} J(\vec{x}, t) J(\vec{x}, s) / (2\omega) e^{-i\omega|t-s|} e^{i\vec{k}\cdot\vec{x} - i\vec{k}\cdot\vec{y}} \right] \right). \end{aligned} \quad (34)$$

A transition from the kernel to the normal symbol for the S-matrix corresponds to omitting the first term in the exponent.

Now let us recall the solution of the free Klein–Gordon equation

$$\varphi_0(x) = \int \frac{d\vec{k}}{(2\pi)^3 2k^0} \left[ a(\vec{k})e^{i\omega t} + a^*(\vec{k})e^{-i\omega t} \right]$$

and the Green function (the propagator)

$$D_c(x) = i \int \frac{d\vec{k}}{(2\pi)^3 2\omega} e^{i\vec{k}\cdot\vec{x}} e^{-i\omega_k|t|} = i \int \frac{d^4k e^{-ikx}}{k^2 - m^2 + i0}.$$

In terms of  $\varphi_0(x)$  and  $D_c(x)$ , one can rewrite the normal symbol for the S-matrix in the following way:

$$S_J(\varphi_0) = \exp \left( i \int d^4x J(x)\varphi_0(x) + \frac{1}{2} \int d^4x d^4y J(x)D_c(x-y)J(y) \right). \quad (35)$$

Generalization of (35) to the case of the integration potential  $V(\varphi)$  is rather straightforward. We can explore the relation we already faced with

$$\exp\left(-i \int d^4x V(\varphi)\right) = \exp\left(-i \int V\left(\frac{\delta}{i\delta J}\right) d^4x\right) \exp\left(i \int \varphi J d^4y\right) \Big|_{J=0}. \quad (36)$$

Then the expression for the normal symbol of the S-matrix takes a simple and elegant form:

$$S_V(\varphi_0) = \exp\left(-i \int V\left(\frac{\delta}{i\delta J}\right) d^4x\right) \cdot \exp\left(i \int J(y)\varphi_0(y) d^4y + \frac{1}{2} \int d^4z d^4y J(y) D_c(y-z) J(z)\right) \Big|_{J=0}. \quad (37)$$

If we do not put the source  $J \rightarrow 0$  after taking functional derivatives, we have the normal symbol of the S-matrix in the presence of an external source:

$$S_V(\varphi_0, J) = \exp\left(-i \int V\left(\frac{\delta}{i\delta J}\right) d^4x\right) \cdot \exp\left(i \int J(y)\varphi_0(y) d^4y + \frac{1}{2} \int d^4z d^4y J(y) D_c(y-z) J(z)\right). \quad (38)$$

One can see that if we put  $\varphi_0 = 0$  we get the same formula we as already derived for the generating functional

$$Z_V(J) = \exp\left(-i \int V\left(\frac{\delta}{i\delta J}\right) d^4x\right) \cdot \exp\left(\frac{1}{2} \int d^4z d^4y J(y) D_c(y-z) J(z)\right). \quad (39)$$

This observation allows us to get a very important relation, which is called the Lehmann–Symanzik–Zimmermann reduction formula. To do this, we introduce in (38) some arbitrary external field  $\varphi(x)$  instead of  $\varphi_0(x)$ . Then, by direct computation of functional derivatives with respect to  $\varphi(x)$  and  $J(x)$ , one can check the following relation:

$$\int \prod_i dx_i \varphi_0(x_i) \left[ \frac{1}{i} \frac{\delta}{\delta \varphi_1} \dots \frac{\delta}{\delta \varphi_n} S_V(\varphi_0, J) \Big|_{\varphi, J=0} - \frac{1}{i} \frac{\delta}{\delta \tilde{J}_1(x_1)} \dots \frac{\delta}{\delta \tilde{J}_n(x_n)} Z(\tilde{J}) \Big|_{J=0} \right] = 0, \quad (40)$$

where  $\tilde{J}(x) = \int D_c(x-y) J(y)$ . If we keep in mind that the expression of the S-matrix

$$S_V(\varphi_0) = \sum_n \frac{1}{n!} \int dx_1 \dots dx_n \varphi_0(x_1) \dots \varphi_0(x_n) S_n(x_1 \dots x_n)$$

gives us the coefficient functions  $S_n(x_1, \dots, x_n)$  of S-matrix scattering elements, and on the other hand the expansion of the generating functional

$$Z_V(J) = \sum_n \frac{1}{n!} \int dx_1 \dots dx_n J(x_1) \dots J(x_n) \frac{\delta Z}{i\delta J(x_1) \dots i\delta J(x_n)}$$

gives us the Green functions

$$G_n(x_1 \dots x_n) = \frac{\delta Z}{i\delta J(x_1) \dots i\delta J(x_n)},$$

we observe a simple correspondence. The reduction formula (40) tells us how to compute S-matrix elements by computing corresponding Green functions:

1. one should compute the Green function;
2. multiply all legs to the inverse propagator or, in other words, apply the operator

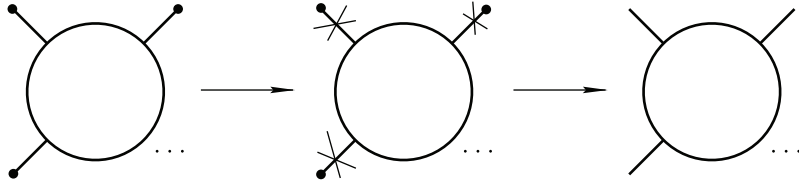
$$\prod_{i=1}^n (\square_i + m^2)$$

to each of the legs;

3. multiply the result by the product of the corresponding free fields:

$$\frac{1}{n} \prod_i \varphi_0(x_i).$$

Schematically, the procedure is shown in Fig. 3.



**Fig. 3:** Transition from the Green functions to the S-matrix elements

This rule is very general and could be applied for all types of fields, not only for the scalar fields.

## 7 Generating functional for Green functions and perturbation theory

Now we return to the generating functional written in the form

$$Z[J] = \int D(\varphi) \cdot \exp \left( i \int d^4x L(\varphi, \partial_\mu \varphi) + i \int d^4x J(x) \varphi(x) \right), \quad (41)$$

where  $L = \frac{1}{2} \partial_\mu \varphi \cdot \partial^\mu \varphi - \frac{1}{2} m^2 \varphi^2 - V(\varphi)$ . If we take the second derivative,

$$\left. \frac{\delta^{(2)} Z}{i \delta J(x_1) i \delta J(x_2)} \right|_{J=0} = \langle \varphi_1(x_1) \varphi_2(x_2) \rangle,$$

we obtain the Feynman propagator  $D_c(x_1 - x_2)$ , as can be easily seen from the form for  $Z[J]$ :

$$Z[J] = \exp \left( -i \int dx V \left( \frac{\delta}{\delta J} \right) \right) \cdot \exp \left( \frac{1}{2} \int dy dz J(y) D_c(y - z) J(z) \right). \quad (42)$$

We obtain the same function as we have obtained from the time-ordered product of field operators (see (5))

$$D_c(x_1 - x_2) = \langle 0 | T \{ \hat{\Phi}(x_1) \hat{\Phi}(x_2) \} | 0 \rangle = \langle \varphi_1(x_1) \varphi_2(x_2) \rangle.$$

This is always the case. Derivatives of the generating functional automatically give  $T$ -products of the corresponding field operators:

$$\frac{\delta^{(n)} Z}{i \delta J(x_1) \dots i \delta J(x_n)} \equiv \langle \varphi_1(x_1) \dots \varphi_n(x_n) \rangle = \langle 0 | T \{ \hat{\Phi}(x_1) \dots \hat{\Phi}(x_n) \} | 0 \rangle.$$

At this point let us consider as an example a theory with  $V(\varphi) = \frac{\lambda}{4!}\varphi^4$ . In such a theory if we take two derivatives and expand the exponent on  $\lambda$ , we get

$$\begin{aligned} \langle \varphi_1(x_1)\varphi_2(x_2) \rangle &= D_c(x_1 - x_2) + \lambda [D_c(x_1 - x_2)D_c^2(0) + \dots] \\ &\quad + \frac{1}{2} \left(\frac{\lambda}{4!}\right)^2 [72D_c^2(0)D_c^2(x_1 - x_2) + 24D_c^4(x_1 - x_2) + \dots] + \dots \end{aligned}$$

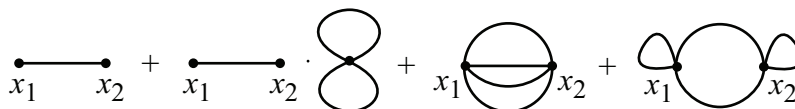
It is very useful to introduce the Feynman rules and Feynman diagrams. The closed line corresponds to the propagator

$$\frac{1}{i}D_c(x - y).$$

Each interaction vertex corresponds to

$$-i\lambda = -i \frac{d^4V(\varphi)}{d\varphi^4} \Big|_{\varphi=0}.$$

In terms of Feynman rules, the corrections to the two-point correlation function (43) are given by Feynman diagrams shown in Fig. 4.



**Fig. 4:** Illustration of (43)

One should add a symmetry factor, which corresponds to possible permutations of equivalent lines. If one takes the Fourier transformation, one can formulate the rules in momentum space, which are usually used in practical computations. In momentum space the integral

$$\int \frac{dp}{(2\pi)^4}$$

corresponds to each loop and in each vertex the momentum conservation law takes place.

We will be more specific and precise later in the formulation of Feynman rules for the case of the SM.

But now we need to consider a few more properties of the generating functional. As we have seen already from the two-point correlator the perturbative expansion contains disconnected diagrams, which are not really needed in computations. The way out of this problem is to consider the logarithmic function of the generating functional:

$$iW[J] = \ln Z[J], \quad Z[J] = e^{iW[J]}.$$

Then, for the functional derivatives, one gets

$$\frac{\delta W}{\delta J} = \frac{1}{Z} \frac{\delta Z}{i\delta J}, \quad \frac{\delta^2 W}{\delta J_1 \delta J_2} = i \frac{1}{Z} \frac{\delta^2 Z}{i\delta J_1 i\delta J_2} - \frac{1}{Z^2} \frac{\delta Z}{i\delta J_1} \frac{\delta Z}{i\delta J_2} \dots,$$

where additional terms exactly cancel out disconnected pieces in the Green functions. The property that the logarithmic function leads to connected diagrams is a particular example of a more general theorem in graph theory.

The functional

$$W[J] = \frac{1}{i} \ln Z[J] \tag{43}$$

is called the generating functional for connected Green functions.

The next important observation is related to the functional Legendre transformation:

$$\Gamma[\varphi_{\text{cl}}] = W[J] - \int d^4x J(x)\varphi_{\text{cl}}(x), \quad (44)$$

where  $\varphi_{\text{cl}} = \delta W/\delta J$  with a formal solution  $J = J(\varphi_{\text{cl}})$ . If we take functional derivatives of the functional  $\Gamma[\varphi_{\text{cl}}]$ , we get the so-called one-particle irreducible Green functions (Feynman diagrams corresponding to such functions cannot be split up into disconnected pieces by cutting only one internal line):

$$\frac{\delta \Gamma}{\delta \varphi_{\text{cl}}} = \frac{\delta W}{\delta J} \frac{\delta J}{\delta \varphi_{\text{cl}}} - J(x) - \varphi_{\text{cl}} \frac{\delta J}{\delta \varphi_{\text{cl}}}. \quad (45)$$

The functional  $\Gamma$  is called the effective action. Two terms in (45) are cancelled out and we get

$$\frac{\delta \Gamma}{\delta \varphi_{\text{cl}}(x)} = -J(x). \quad (46)$$

Now one can take functional derivatives from both sides of (46) and get the following relations:

$$\frac{\delta^2 \Gamma}{\delta \varphi_{\text{cl}}(x_1) \delta \varphi_{\text{cl}}(x_2)} = -\frac{\delta J(x_1)}{\delta \varphi_{\text{cl}}(x_2)} = -\left[ \frac{\delta W}{\delta J(x_1) \delta J(x_2)} \right]^{-1}. \quad (47)$$

If we introduce notation for the connected Green function:

$$G_n(x_1, \dots, x_n) = -i \frac{\delta W[J]}{i \delta J(x_1) \dots i \delta J(x_n)} \Big|_{J=0}$$

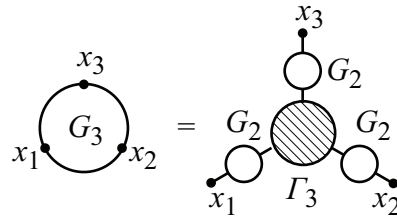
and for the one-particle irreducible Green function:

$$\Gamma_n(x_1, \dots, x_n) = -i \frac{\delta^{(n)} \Gamma[\varphi_{\text{cl}}]}{\delta \varphi_{\text{cl}}(x_1) \dots \delta \varphi_{\text{cl}}(x_n)} \Big|_{\varphi_{\text{cl}}=0},$$

then the formula (47) can be written in the following form:

$$\Gamma_2 = G_2^{-1},$$

which means that the irreducible two-point Green function is nothing but the inverse propagator. If we take more derivatives on both sides of (47) we obtain relations between connected and one-particle irreducible Green functions. For the three-point Green function, it is presented schematically in Fig. 5. If we restore the Planck constant  $\hbar$ , the generating functional has the form



**Fig. 5:** Connected three-point Green function is equal to one-particle irreducible three-point function convoluted with three propagators.

$$Z[J] = \int D(\varphi) \exp \left( \frac{i}{\hbar} S[\varphi] + i \int dx J(x)\varphi(x) \right),$$

where  $S[\varphi]$  is the action. In the quasiclassical limit ( $\hbar \rightarrow 0$ ), the functional integral is dominated by the stationary trajectory:

$$\left. \frac{\delta S[\varphi]}{\delta \varphi(x)} \right|_{\varphi=\varphi_{\text{cl}}} + J = 0$$

Therefore

$$Z[J] \sim \exp \left( \frac{i}{\hbar} S[\varphi_{\text{cl}}] + i \int dx J(x) \varphi_{\text{cl}}(x) \right)$$

and one can see from (43) that

$$W[J] = S[\varphi_{\text{cl}}] + \int dx J(x) \varphi_{\text{cl}}$$

and, by comparing with (44), we obtain

$$\Gamma[\varphi_{\text{cl}}] = S[\varphi_{\text{cl}}].$$

So, we can conclude that the irreducible Green functions are  $\Gamma_n(x_1, \dots, x_n)$ , the effective vertices of the theory:

$$\Gamma_n(x_1, \dots, x_n) = (-i) \frac{\delta^{(n)} S[\varphi_{\text{cl}}]}{\delta \varphi_{\text{cl}}(x_1) \dots \delta \varphi_{\text{cl}}(x_n)}. \quad (48)$$

We obtain this formula for the case of scalar fields as an example, but it remains true for any theory with corresponding obvious changes. Note that the formula (48) is very useful to get Feynman rules for complicated vertices in the interaction Lagrangian. For example, three- and four-gluon vertices are obtained with all needed symmetry properties.

Up to now we have considered basic ingredients of the quantum field theory for the case of scalar fields. However, there are many other fields and corresponding particles which have different spin properties. In the SM there are leptons and quarks, being fermions with the spin 1/2, and the boson fields with the spin 1. We begin our brief consideration with spin-1/2 fermion fields.

## 8 Fermion fields

Spin-1/2 particles with mass  $m$  are described by the four-component field  $\Psi$ . The Lagrangian for the field has the well-known form

$$L = \bar{\Psi} i \partial_\mu \gamma^\mu \Psi - m \bar{\Psi} \Psi. \quad (49)$$

The least-action principle leads to the famous Dirac equation of motion:

$$(i \gamma^\mu \partial_\mu - m) \Psi = 0, \quad (50)$$

where  $\gamma^\mu$  ( $\gamma^0, \gamma^1, \gamma^2, \gamma^3$ ) are the Dirac ( $4 \times 4$ ) matrices. The matrices  $\gamma_\mu$  obey the anticommutation relation

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}. \quad (51)$$

There are several representations for  $\gamma$ -matrices. In the SM chiral or Weyl spinors are of particular importance. Therefore, we use the Weyl representation of  $\gamma$ -matrices

$$\gamma_\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}. \quad (52)$$

where  $\sigma^0 = I$ ,  $\sigma^i = \tau^i$ ,  $\bar{\sigma}^0 = I$ ,  $\bar{\sigma}^i = -\tau^i$  and  $\tau^i$  are the ( $2 \times 2$ ) Pauli matrices.

In this representation the  $\gamma^5$ -matrix has the following form:

$$\gamma^5 = i \gamma^1 \gamma^2 \gamma^3 \gamma^4 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}. \quad (53)$$

Chiral spinors

$$\Psi_{L,R} = \frac{1 \mp \gamma_5}{2} \Psi \quad (54)$$

in this notation

$$\Psi_L = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ 0 \\ 0 \end{pmatrix}, \quad \Psi_R = \begin{pmatrix} 0 \\ 0 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}$$

are in fact two-component objects.

In momentum space the function  $\Psi(x)$  is decomposed into positive- and negative-energy parts

$$u_\lambda(p)e^{-ipx} \text{ and } v_\lambda(p)e^{ipx} \quad (55)$$

and the spinors  $u(p)$  and  $v(p)$  obey the Dirac equation in the following form:

$$\begin{aligned} (p_\mu \gamma^\mu - m) u_\lambda(p) &= 0, \\ (p_\mu \gamma^\mu + m) v_\lambda(p) &= 0. \end{aligned} \quad (56)$$

The concrete form of spinors is different in different parametrizations of  $\gamma$ -matrices, and in the Weyl representation the spinors are

$$u_\lambda = \begin{pmatrix} \sqrt{p^0 + \vec{p}\vec{\sigma}\xi_\lambda} \\ \sqrt{p^0 - \vec{p}\vec{\sigma}\xi_\lambda} \end{pmatrix}, \quad (57)$$

$$v_\lambda = \begin{pmatrix} \sqrt{p^0 + \vec{p}\vec{\sigma}\eta_\lambda} \\ -\sqrt{p^0 - \vec{p}\vec{\sigma}\eta_\lambda} \end{pmatrix}, \quad (58)$$

where  $\xi$  and  $\eta$  are two-component spinors determined by fixing some quantization axis. Left and right chiral spinors are then

$$\begin{aligned} u_{L,R} &= \frac{1 \mp \gamma_5}{2} u_\lambda, \\ v_{L,R} &= \frac{1 \mp \gamma_5}{2} v_\lambda. \end{aligned}$$

Normalization conditions and summation over indices are as follows:

$$\begin{aligned} \bar{u}_\lambda u_{\lambda'} &= 2m\delta_{\lambda\lambda'}, \quad \bar{v}_\lambda v_{\lambda'} = -2m\delta_{\lambda\lambda'}, \\ \sum_\lambda u_\lambda \bar{u}_\lambda &= p_\mu \gamma^\mu + m, \quad \sum_\lambda v_\lambda \bar{v}_\lambda = p^\mu \gamma_\mu - m. \end{aligned}$$

Quantization of the Dirac field is similar to the scalar case considered above with a very important difference. In order to have correct Fermi statistics and obey the Pauli principle, the commutation relations in the scalar case should be replaced by corresponding anticommutation relations:

$$\{\hat{\pi}_\alpha(t, \vec{x}), \Psi_\beta(t, \vec{x}')\} = -i\delta_{\alpha\beta}\delta(\vec{x} - \vec{x}'), \quad (59)$$

where  $\alpha = 1, 2, 3, 4$  and the field momentum is

$$\pi_\alpha(t, \vec{x}) = \frac{\partial L}{\partial \dot{\Psi}_\alpha} = i\Psi_\alpha^\dagger.$$

The fermionic field operator may be constructed with the help of spinors obeying the Dirac equation

$$\Psi(x) = \int \frac{d\vec{p}}{(2\pi)^3 p^0} \sum_{\lambda=1,2} \left[ \hat{b}_\lambda(p) u_\lambda(p) e^{-ipx} + \hat{d}_\lambda^\dagger(p) v_\lambda(p) e^{ipx} \right],$$

$$\bar{\Psi}(x) = \Psi^\dagger \gamma^0 = \int \frac{d\vec{p}}{(2\pi)^3 p^0} \sum_{\lambda=1,2} \left[ \hat{b}_\lambda^\dagger(p) \bar{u}_\lambda(p) e^{ipx} + \hat{d}_\lambda(p) \bar{v}_\lambda(p) e^{-ipx} \right].$$

It is easy to check that from the anticommutators (59) the creation and annihilation operators satisfy the anticommutation relations in the following form:

$$\left\{ \hat{b}_\lambda(\vec{p}), \hat{b}_{\lambda'}(\vec{p}') \right\} = \left\{ \hat{d}_\lambda(\vec{p}), \hat{d}_{\lambda'}(\vec{p}') \right\} = (2\pi)^3 2p^0 \delta(\vec{p} - \vec{p}') \delta_{\lambda\lambda'}. \quad (60)$$

Then one-particle and one-antiparticle states are obtained from the vacuum state  $|0\rangle$  by acting of the creation operators:

$$\hat{b}_\lambda^\dagger(\vec{p}) |0\rangle \text{ and } \hat{d}_\lambda^\dagger(\vec{p}) |0\rangle.$$

In the same way, by acting of creation operators for particle and antiparticle on the vacuum state, one gets two-, three-, . . . particle states. Because of zero anticommutators, for any creation operator one gets nicely the Pauli principle:

$$\left\{ \hat{b}_\lambda^\dagger(\vec{p}), \hat{b}_\lambda^\dagger(\vec{p}') \right\} = 0 \Rightarrow \hat{b}_\lambda^\dagger(\vec{p}), \hat{b}_\lambda^\dagger(\vec{p}') |X\rangle \equiv 0 \quad (61)$$

for any state  $|X\rangle$ . From the field operators one can get the Feynman propagator (T-ordered correlator) in a similar way as was done for the scalar case:

$$\langle 0| T (\bar{\Psi}(x_1) \Psi(x_2)) |0\rangle = \frac{-1}{i} \int \frac{dp}{(2\pi)^4} \frac{p_\mu \gamma^\mu + m}{p^2 - m^2 + i0}. \quad (62)$$

Considering the path-integral method for the fermion field, one can construct the holomorphic representation similar to the scalar case. However, now we have to deal with anticommuting numbers called Grassmann numbers, which form the Grassmann algebra:

$$\begin{aligned} (a_\alpha)^* &= a_\alpha^*, & (a_\alpha^*)^* &= a_\alpha, \\ \{a_\alpha a_\beta\} &= \{a_\alpha^* a_\beta^*\} = \{a_\alpha a_\beta^*\} = 0, \\ ca_\alpha &= a_\alpha c; & ca_\alpha^* &= a_\alpha^* c, \end{aligned}$$

where  $\alpha = 1, \dots, n$ ,  $c$  are the usual numbers.

A function of Grassmann variables has therefore the generic form

$$\begin{aligned} f(a, a^*) &= f_{00} + \sum_{\alpha_1} f_{\alpha_1|0} a_{\alpha_1} + \sum_{\alpha_1} f_{0|\alpha_1} a_{\alpha_1}^* \\ &+ \sum_{\alpha_1 \alpha_2} f_{\alpha_1 \alpha_2} a_{\alpha_1} a_{\alpha_2} + \dots + f_{1\dots n|n\dots 1} a_1 \dots a_n a_n^* \dots a_1^*. \end{aligned} \quad (63)$$

The expression (63) reminds us of the norm-ordering operator products we have used already. The operations of differentiation and integration are defined as follows:

$$\begin{aligned} \frac{\partial}{\partial a_\alpha} a_\beta &= \frac{\partial}{\partial a_\alpha^*} a_\beta^* = \delta_{\alpha\beta}, & \frac{\partial}{\partial a_\alpha} a_\beta^* &= \frac{\partial}{\partial a_\alpha^*} a_\beta = 0, \\ \frac{\partial}{\partial a_\alpha} f &= f_\alpha; & \frac{\partial}{\partial a_\alpha^*} f &= \bar{f}_\alpha, & \int da_\alpha f &= f_\alpha; & \int da_\alpha^* f &= \bar{f}_\alpha, \end{aligned}$$

where  $f_\alpha$  does not depend on  $a_\alpha$  and  $\bar{f}_\alpha$  does not depend on  $a_\alpha^*$ . One can check the anticommutation relation for differentials:

$$\{da_\alpha da_\beta\} = \{da_\alpha^* da_\beta^*\} = \{da_\alpha da_\beta^*\} = 0.$$

If we denote

$$da^* da = da_1^* \dots da_n^* da_n \dots da_1$$



and take into account that for the function defined in (63)

$$\int da^* da f(a, a^*) = f_{1\dots n|n\dots 1},$$

we can easily prove the following relation by expanding in series in  $a$  and  $a^*$ :

$$\int da^* da \exp \left( \sum a_\alpha^* A_{\alpha\beta} a_\beta + \sum \eta_\alpha^* a_\alpha + i \sum \eta_\alpha a_\alpha^* \right) = \det A \exp \left( \eta_\alpha^* A_{\alpha\beta}^{-1} \eta_\beta \right). \quad (64)$$

One can see from (64) that in contrast to the integration with the usual complex numbers (see (9)) the determinant appears in the numerator in the case of anticommuting Grassmann numbers (64).

Similar to the case of the scalar field, one can get a formula for the S-matrix normal symbol:

$$S_\eta(b^\dagger, d^\dagger, b, d) = \frac{1}{N} \exp \left( -i \int \bar{\eta} S_C \eta + i \int (\bar{\eta} \Psi_0 + \bar{\Psi}_0 \eta) \right),$$

where  $\bar{\Psi}_0 = \Psi_0^\dagger \gamma^0$ ,  $\Psi_0$  is a solution of the free Dirac equation

$$\Psi_0(x) = \int \frac{d\vec{k}}{(2\pi)^3 2k_0} \left[ \sum_{\lambda=1,2} b_\lambda(\vec{k}) u_\lambda(\vec{k}) e^{-ikx} + \sum_{\lambda=1,2} d_\lambda^*(\vec{k}) v_\lambda(\vec{k}) e^{ikx} \right],$$

$\eta(x)$  is the fermion source and

$$S_C(x) = \frac{i}{(2\pi)^4} \int dk e^{-ikx} \frac{\hat{k} + m}{k^2 - m^2 + i0}$$

is the Feynman propagator for the fermion field.

Taking into account the following relation for the functional measure:

$$\prod_{t,k,\lambda} db_\lambda^*(t, \vec{k}) db_\lambda(t, \vec{k}) dd_\lambda^*(t, \vec{k}) dd_\lambda(t, \vec{k}) = \text{Const.} \prod_{x,\alpha} d\bar{\Psi}_\alpha(x) d\Psi_\alpha(x),$$

one gets the answer for the generating functional for the fermion Green functions in compact form similar to the scalar case:

$$Z(\bar{\eta}, \eta) = N^{-1} \int \exp \left( i \int d^4x (L(x) + \bar{\eta} \Psi + \bar{\Psi} \eta) \right) \prod_x d\bar{\Psi}(x) d\Psi(x).$$

## 9 Quantization of theories with the gauge fields

The gauge field was first introduced in QED when the Maxwell equations were rewritten in terms of the 4-vector potential  $A_\mu(x)$ .

The equation in terms of the field  $A_\mu$ ,

$$\partial_\mu F^{\mu\nu} = 0,$$

where  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ , is invariant under the local  $U(1)$  transformation:

$$A'_\mu = A_\mu + \partial_\mu \alpha(x).$$

It is easy to check the  $U(1)$  invariance of the QED Lagrangian

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\Psi} (i\hat{D} - m) \Psi, \quad \Psi \rightarrow e^{ie\alpha} \Psi, \quad A_\mu \rightarrow A_\mu + \partial_\mu \alpha, \quad (65)$$

where  $D_\mu = \partial_\mu - ieA_\mu$  is the covariant derivative.

In the SM we deal not only with the abelian  $U(1)$  group but also with non-abelian  $SU(N)$  groups,  $SU(2)$  for the EW and  $SU(3)$  for the strong forces.

We are not discussing here generic structure and properties of Lie groups but introduce briefly the  $SU(N)$  group.  $SU(N)$  is a group of unitary matrices  $U$  ( $U^\dagger U = 1$ ) with determinant equal to 1 ( $\det U = 1$ ). Elements of the group  $U(x)$  may depend on the space-time point  $x^\mu$ .

If we want the theory to be invariant under  $SU(N)$  transformation the covariant derivative in the Lagrangian

$$L = \bar{\Psi}(i\hat{D} - m)\Psi \quad (66)$$

should transform as

$$\begin{aligned} D_\mu \Psi &\rightarrow (D_\mu \Psi)^U = U D_\mu \Psi, \\ (\partial_\mu - igA_\mu^U) U \Psi &= U (\partial_\mu - igA_\mu) \Psi. \end{aligned}$$

From this, one gets the following transformation form for the potential  $A$ :

$$A_\mu^U = U A_\mu U^{-1} + \frac{i}{g} U \partial_\mu U^{-1}. \quad (67)$$

The kinetic term for the non-abelian  $A_\mu$  field is constructed as the gauge-invariant operator

$$L_A = -\frac{1}{2} \text{Tr}(F_{\mu\nu} F^{\mu\nu}). \quad (68)$$

In the SM all gauge fields are taken in the adjoint representation:

$$A_\mu(x) = A_\mu^a(x) t^a,$$

where  $t^a$  ( $a = 1, \dots, N-1$  for  $SU(N)$ ) are so-called generators of the group. As do all generators for the Lie group, the generator  $t^a$  obeys the following commutation relation:

$$[t^a, t^a] = f^{abc} t^c, \quad \text{Tr}(t^a) = 0.$$

One may choose normalization conditions as

$$\text{Tr}(t^a t^b) = \frac{1}{2} \delta^{ab}. \quad (69)$$

Then the Lagrangian for the gauge field takes the form

$$L = -\frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu}, \quad (70)$$

where the field strength tensor is

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c. \quad (71)$$

One can express the unitary matrix  $U(1)$  in the form  $U(x) = e^{ig\alpha^a(x)t^a}$ . Then the transformation for the gauge field  $A_\mu^a$  takes the form

$$A_\mu^a \rightarrow (A^\alpha)_\mu^a = A_\mu^a + \partial_\mu \alpha^a + g f^{abc} A_\mu^b \alpha^c = A_\mu^a + D_\mu^{ac} \alpha^c, \quad (72)$$

where the covariant derivative in components is

$$D_\mu^{ac} = \partial_\mu \delta^{ac} + g f^{abc} A_\mu^b.$$

Part of the SM describing strong interactions is based on the  $SU_C(3)$  group. It is called QCD with the Lagrangian

$$L_{\text{QCD}} = -\frac{1}{4}G_{\mu\nu}^a G^{\mu\nu a} + \bar{q}_i \left( i(D_\mu)_{ij} \gamma_{ij}^\mu - m\delta_{ij} \right) q_j, \quad (73)$$

where  $q = 1, 2, 3$ ,  $(D_\mu)_{ij} = \partial_\mu \delta_{ij} - ig(t^a)_{ij} A_\mu^a$ .

On the classical level abelian and non-abelian theories look very elegant. However, problems appear when the theories are quantized. The reason can be seen already in QED, where we know that in a theory described by the field  $A_\mu$  we have four components but only two of them are physical degrees of freedom corresponding to two polarizations of the physics photon. This problem is manifested in the fact that the quadratic form of the Lagrangian

$$A^\mu D_{\mu\nu}^{-1} A^\nu = A^\mu (\square g_{\mu\nu} - \partial_\mu \partial_\nu) A^\nu$$

or, in momentum space,

$$(k^2 g_{\mu\nu} - k_\mu k_\nu)$$

does not have an inverse form. As we have seen in cases of scalar and fermion fields the inverse of the differential operator, the Green function, is the propagator. So, here one cannot get the propagator in such a way.

The way out of this problem is a correct quantization procedure called the quantization of constrained systems. The reason that the functional integral

$$\int \prod_{\mu,x} dA_\mu(x) \exp \left\{ i \int dx \left( -\frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} \right) \right\} \quad (74)$$

does not give a reasonable result is that there are an infinite number of gauge configurations  $(A^\alpha)_\mu^a$ , which differ only by the gauge transformation, leading to identical physics results since the action is gauge invariant. So, one should perform the functional integration taking only one representative from such gauge configuration.

Without going into details, the final recipe is as follows:

$$\int DA \delta(F(A)) \det(\Delta_{gh}^F) e^{iS(A)}, \quad (75)$$

where  $DA = \prod_{\mu,x} dA_\mu(x)$  and the so-called Faddeev–Popov determinant  $\det(\Delta_{gh}^F)$  is introduced to ensure the gauge invariance of the functional measure. The  $\delta$ -function fixes the gauge condition of gauge choice  $F(A) = 0$ .

We recall briefly the main ideas of the method proposed by Faddeev and Popov. Let us introduce a functional integral that is equal to unity:

$$1 = \int D\alpha \cdot \delta(F(A^\alpha)) \det\left(\frac{\delta F}{\delta\alpha}\right), \quad (76)$$

where  $D\alpha = \prod_x d\alpha(x)$ . Substituting (76) into the integral (16) and performing the gauge transformation  $(A^\alpha)_\mu^a(x) \rightarrow A_\mu^a(x)$ , one gets

$$\int D\alpha \int DA e^{iS[A]} \cdot \delta(F(A)) \cdot \det\left(\frac{\delta F}{\delta\alpha}\right). \quad (77)$$

The factor  $\det\left(\frac{\delta F}{\delta\alpha}\right)$  does not depend on  $\alpha$  and therefore the integration over the gauge group is factorized out. This infinite factor is included into the normalization factor of the functional measure and therefore could be dropped.

The main idea of quantization and the Faddeev–Popov method could be illustrated in a very simple example with the usual integrals. Let us consider an integral

$$I = \int \int_{-\infty}^{\infty} dx_1 dx_2 e^{-x_1^2 - x_2^2 + 2x_1 x_2} = \int dx_1 dx_2 e^{-x_i A_{ij} x_j}, \quad (78)$$

where

$$A_{ij} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \text{ and } \det A = 0.$$

We can obviously see that by making a substitution of variables we get

$$I = c \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy e^{-x^2} \text{ when } x = x_1 - x_2, \quad y = x_1 + x_2.$$

The ‘action’ in the integral  $I$  is invariant under translation:

$$x_1 \rightarrow x_1 + a, \quad x_2 \rightarrow x_2 + a.$$

And, the integral  $\int_{-\infty}^{\infty} dy$  gives simply the infinite volume of the algebra corresponding to the translation group.

Now let us substitute  $\hat{1}$  into the integral  $I$ :

$$\int_{-\infty}^{\infty} d\omega \delta(F(x_i + \omega)) \det \left( \frac{\delta F}{\delta \omega} \right) = 1,$$

$$\int d\omega \int dx_1 dx_2 e^{-x_i A_{ij} x_j} \delta(F(x_i + \omega)) \det \left( \frac{\delta F}{\delta \omega} \right).$$

After the ‘gauge’ substitution  $x_i \rightarrow x_i - \omega$ , one can drop the infinite group integral  $\int d\omega$  and get the final integral in a form symmetric with respect to integration variables:

$$I_G = \int dx_1 \int dx_2 e^{-x_i A_{ij} x_j} \delta(F(x_i)) \left| \frac{\partial F}{\partial \omega} \right|.$$

The determinant  $\det \left( \frac{\delta F}{\delta \alpha} \right)$  is the Faddeev–Popov determinant and  $\Delta_{\text{ch}} = \frac{\delta F}{\delta \alpha}$  is the so-called ghost operator.

Obviously, in the abelian case the Lorentz gauge condition,  $F = \partial_\mu A_\mu$ , leads to the ghost operator  $\Delta_{\text{ch}} = \partial_\mu \cdot \partial^\mu$ . Since the operator does not depend on the field, it is cancelled in the connected Green functions and does not give any contribution.

However, this is not so in the non-abelian case where  $\Delta_{\text{ch}}$  is a non-trivial operator depending on the gauge field. Technically it is convenient to express  $\det(\Delta_{\text{ch}})$  also as a functional integral. As we know, a determinant in the numerator appears when one integrates over anticommuting fields:

$$\det(\Delta_{\text{ch}}) = \int \prod d\bar{c} dc e^{\int \bar{c} \Delta_{\text{ch}} c}, \quad (79)$$

where the anticommuting fields  $c$  are called Faddeev–Popov ghosts.

As an example, let us consider the gauge condition in the covariant form:

$$F(A) = \partial_\mu A^\mu - a(x) \quad (80)$$

with an arbitrary function  $a(x)$ . The functional integral (77) in this case takes the form

$$e^{iW[J]} = \int \prod_{\mu, x} dA_\mu^a(x) d\bar{c}^a(x) dc^a(x) \cdot \delta(\partial_\mu A^{\mu a} - a^a(x)).$$

$$\exp \left( i \int d^4x \left[ -\frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} + \bar{c}^a \left( \square c^a - f^{abc} \partial_\mu (A_\mu^c c^b) \right) + J_\mu^a A^{a\mu} \right] \right).$$

The Green function does not depend on  $a(x)$ , so one can integrate (9) with the exponent

$$\int da \exp \left( -\frac{a^2}{2\xi} \right) \delta(\partial_\mu A^\mu - a) = \exp \left( -\frac{(\partial_\mu A^\mu)^2}{2\xi} \right).$$

As the result, we get the following quadratic part of the action in (9):

$$S = \int dx \left[ -\frac{1}{4} (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) (\partial^\mu A^{a\nu} - \partial^\nu A^{a\mu}) - \frac{1}{2\xi} \partial_\mu A^{a\mu} \partial_\nu A^{a\nu} + \bar{c} \square c \right],$$

where the numerical parameter  $\xi$  is the gauge parameter. Now there is the inverse form, and we get the following propagators for non-abelian gauge and ghost fields, respectively:

$$D_{\mu\nu}^{ab}(k) = -i \frac{\delta^{ab}}{k^2 + i0} \left[ g_{\mu\nu} - (1 - \xi) \frac{k_\mu k_\nu}{k^2} \right], \quad (81)$$

$$D_{\text{ch}}^{ab}(k) = i \frac{\delta^{ab}}{k^2 + i0}. \quad (82)$$

There are a few famous choices of the parameter  $\xi$  used in concrete computations:

$\xi = 1$ —the 't Hooft–Feynman gauge,

$\xi = 0$ —the Landau gauge,

$\xi = 3$ —the Frautschi–Yenni gauge.

Of course, any computed physics observable such as cross-section or distribution does not depend on  $\xi$ .

As we see,

$$k^\mu D_{\mu\nu}^{ab} = -\xi k_\nu i \frac{\delta^{ab}}{k^2 + i0} = -\xi k_\nu D_{\text{ch}}^{ab}(k).$$

So, ghosts are acting in a way to cancel a dependence on  $\xi$  in physics quantities.

Now one can use our formula (48) in momentum space:

$$\Gamma_{\mu_1 \dots \mu_n}^{a_1 \dots a_n}(p_1 \dots p_n) \cdot (2\pi)^4 \delta(p_1 + \dots + p_n) = i \frac{\delta^{(n)} S}{\delta A_{\mu_1}^{a_1}(p_1) \dots \delta A_{\mu_n}^{a_n}(p_n)} \quad (83)$$

in order to get Feynman rules in momentum space for all the vertices in abelian and non-abelian theories. Note that by taking functional derivatives in (83) one gets vertex functions with all needed symmetries. In the formula (83),  $\mu$  and  $a$  indicate proper Lorentz and other group indices identifying the field.

As an example, let us consider the QED Lagrangian

$$L = \bar{\Psi} (iD_\mu \gamma^\mu - m) \Psi, \quad D_\mu = \partial_\mu + ieQA_\mu. \quad (84)$$

For the interaction vertex of the fermion field  $\Psi$  with the photon field  $A_\mu$ , we obtain




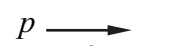



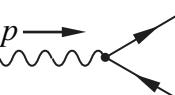
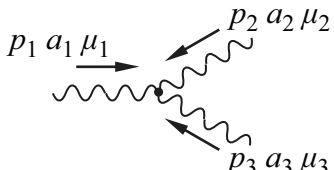
$$\begin{aligned} & \Gamma_\mu(p_1, p_2; p_3) (2\pi)^4 \delta(p_1 + p_2 + p_3) \\ &= i \frac{\delta^{(3)}}{\delta \bar{\Psi}(p_1) \delta \Psi(p_2) \delta A_\mu(p_3)} \cdot \int dx i \cdot (ieQ) \bar{\Psi}(x) \gamma^\mu A_\mu(x) \Psi(x) \\ &= -ieQ \frac{\delta^{(3)}}{\delta \bar{\Psi}(p_1) \delta \Psi(p_2) \delta A_\mu(p_3)} \cdot \\ & \cdot \int dx dq_1 dq_2 dq_3 \exp(-iq_1 x - iq_2 x - iq_3 x) \bar{\Psi}(q_1) \gamma^\mu \Psi(q_2) A_\mu(p_3) \end{aligned}$$

$$= -ieQ(2\pi)^4 \delta(p_1 + p_2 + p_3) \gamma_\mu$$

and so  $\Gamma_\mu(p_1, p_2; p_3) = -ieQ\gamma_\mu$ , where  $Q = -1$  for the fermion field (say, an electron) and  $+1$  for the antifermion field (say, a positron) etc.

In a similar way, one gets all other interaction vertices in the case of the SM Lagrangian, which will be considered later. Also, similar to the scalar field theory, the reduction formula allows us to compute S-matrix elements from the corresponding connected Green functions by cutting out propagators on all legs and multiplying by corresponding free-particle wave functions.

A well-known visual way for presenting and computing S-matrix elements is given by Feynman rules—lines for different types of propagators and external particles, and points for vertices:

	$\frac{-i}{k^2+i0} \delta^{ab} [g_{\mu\nu} - (1-\xi) \frac{k_\mu k_\nu}{k^2}]$ for massless gauge field
	$\frac{i}{k^2+i0} \delta^{ab}$ for the ghost field
	$u(p)$ for an incoming fermion
	$\bar{v}(p)$ for an incoming antifermion
	$u(p)$ for an outgoing fermion
	$\bar{v}(p)$ for an outgoing antifermion
	$i \frac{\hat{p} + m}{p^2 - m^2 + i0}$ for a fermion propagator
	$ig\gamma_\mu(t^a)$ for a fermion–gauge boson vertex
	$gf^{a_1 a_2 a_3} [g_{\mu_1 \mu_2} (p_1 - p_2)_{\mu_3} + g_{\mu_2 \mu_3} (p_2 - p_3)_{\mu_1} + g_{\mu_3 \mu_1} (p_3 - p_1)_{\mu_2}]$

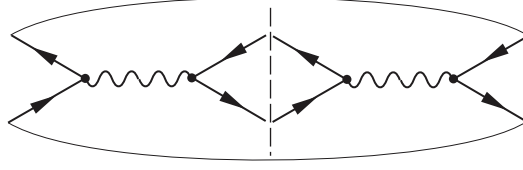
We do not derive here formulas for cross-sections and decay widths; they are given in many textbooks. We use notation of the Particle Data Group:

$$d\sigma_{ab} = \frac{|M|^2}{4\sqrt{(p_a p_b)^2 - m_a^2 m_b^2}} d\Phi_n,$$

where

$$d\Phi_n = (2\pi)^4 \delta(p_i - p_f) \cdot \frac{d^3 \vec{p}_1}{(2\pi)^3 2p_1^0} \cdots \frac{d^3 \vec{p}_n}{(2\pi)^3 2p_n^0},$$

$$d\Gamma = \frac{|M|^2}{2m_a} d\Phi_n.$$



**Fig. 6:**  $\mu^+\mu^-$  production in  $e^+e^-$  collisions

As we see, one needs to compute the matrix element squared in order to get a scattering cross-section and a decay width.

Also, one can formulate the Feynman rules for the matrix element squared directly.

In Fig. 6, the square diagram is shown for  $\mu^+\mu^-$ -pair production in  $e^+e^-$  collisions in QED. The crossed lines correspond to external particles summed over polarizations. Concrete expressions for the sums over polarizations for particles and antiparticles with different spins are as follows:

$$\begin{aligned}
 & 1 && \text{for scalar particles} \\
 & \left\{ \begin{aligned} \sum_{\lambda} u_{\lambda}(p) \times \bar{u}_{\lambda}(p) &= p_{\mu}\gamma^{\mu} + m \\ \sum_{\lambda} v_{\lambda}(p) \times \bar{v}_{\lambda}(p) &= -p_{\mu}\gamma^{\mu} + m \end{aligned} \right. && \text{for spin-1/2 Dirac particles} \\
 & \sum_{\lambda} e_{\lambda}^{\mu}(k)e_{\lambda}^{*\nu}(k) = g_{\mu\nu} - \frac{k_{\mu}k_{\nu}}{k^2} && \text{for massless gauge fields} \\
 & && \text{in the Landau gauge} \\
 & \sum_{\lambda} e_{\lambda}^{\mu}(k)e_{\lambda}^{*\nu}(k) = g_{\mu\nu} && \text{in the 't Hooft–Feynman gauge} \\
 & \sum_{\lambda} e_{\lambda}^{\mu}(k)e_{\lambda}^{*\nu}(k) = g_{\mu\nu} - \frac{k_{\mu}k_{\nu}}{M^2} && \text{for vector fields in the unitary gauge}
 \end{aligned}$$

The Feynman rules for propagators and vertices in the case of matrix elements squared are the same as for the case of amplitudes. Note that the sums over polarizations represent the spin-density matrix and coincide with numerators of the propagators of corresponding particles. Note also, in computations using the Feynman rules for matrix elements squared, that the ghosts should be added into initial and final lines together with corresponding gauge-boson lines. Each loop with crossed ghost lines should include extra factor  $(-1)$  with respect to the corresponding gauge-boson loop. This  $-1$  sign reflects the anticommuting property of the ghost fields.

## 10 Electroweak interactions in the SM

As we know from school textbooks, the weak interactions are responsible for decay of elementary particles. As we shall see, there are also scattering processes due to weak interactions, as were predicted by the SM and discovered in experiments. Studies of weak interactions started from decays, and have a long history, which we do not describe here. From various experimental studies it was realized that (1) electron and muon neutrinos are not the same, and the electron neutrino and antineutrino are different. There are processes  $\nu_{e n} \rightarrow e^{-}p$ ,  $\bar{\nu}_{e p} \rightarrow e^{+}n$ ,  $\nu_{\mu n} \rightarrow \mu^{-}p$ ,  $\bar{\nu}_{\mu p} \rightarrow \mu^{+}n$ , but there are no processes  $\bar{\nu}_{e n} \not\rightarrow e^{-}p$ ,  $\nu_{e p} \not\rightarrow e^{+}n$ ,  $\bar{\nu}_{\mu n} \not\rightarrow \mu^{-}p$ ,  $\nu_{\mu p} \not\rightarrow \mu^{+}n$ ;

(2) the decays  $\mu \rightarrow eX$  have never been observed;

(3) only left-handed leptons and right-handed antileptons participate in the process with  $|\Delta Q| = 1$  for leptons from the same generations;

(4) three generations have been observed.

$$\begin{array}{cc} \ell & \bar{\ell} \\ \longrightarrow p & \longrightarrow p \\ \longleftarrow \bar{J} & \longrightarrow \bar{J} \end{array}$$

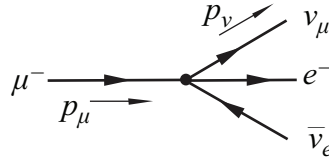
The observations lead to the assumption that the lepton interactions with  $|Q| = 1$  occur via charge current in so-called V–A form:

$$J_\ell \sim \bar{\ell} \gamma_\mu (1 - \gamma_5) \nu_\ell. \quad (85)$$

The corresponding four-fermion interaction Lagrangian for muon and electron currents is

$$L = \frac{G_F}{\sqrt{2}} \bar{\mu} \gamma_\sigma (1 - \gamma_5) \nu_\mu \bar{e} \gamma_\sigma (1 - \gamma_5) \nu_e + \text{h.c.}, \quad (86)$$

where the notation  $\mu, e, \nu$  stands for the corresponding fermion fields and  $G_F$  is the well-known Fermi constant with dimension  $[m]^{-2}$ . With the help of the Lagrangian (86), one can easily compute the decay width of  $\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$ , the Feynman diagram for which is shown in Fig. 7.



**Fig. 7:** Feynman diagram for  $\mu^-$  decay due to four fermion interaction

From dimensional analysis, one can say without any computations that  $\Gamma \sim G_F^2 \cdot m_\mu^5$ .

The formula for the width is

$$\Gamma_\mu = \frac{G_F^2 m_\mu^5}{192\pi^3} \cdot f \left( \frac{m_e^2}{m_\mu^2} \right), \quad (87)$$

where  $f = 1 + 0 \left( \frac{m_e^2}{m_\mu^2} \right)$ . Today's  $G_F$  is measured from the muon decay very precisely:

$$G_F = (1.166371 \pm 6 \cdot 10^{-6}) \cdot 10^{-5} \text{ GeV}^{-2}.$$

As we know, any fermion field  $\Psi(x)$  may be presented as

$$\Psi = \frac{1 - \gamma_5}{2} \Psi + \frac{1 + \gamma_5}{2} \Psi = \Psi_L + \Psi_R.$$

Therefore, the current (85) involves the left component of the fermion field only and has the form

$$J_\ell = \bar{\Psi}_L \gamma_\mu \Psi_L. \quad (88)$$

We want to construct a quantum field theory (the SM) which obeys a few requirements:

- correct electromagnetic neutral currents and V–A charge currents (Fermi);
- three generations without chiral anomalies;
- gauge-invariant dimension-four operators.

Because there are two leptons (charge lepton and corresponding neutrino) in each generation and the left components interact, a very natural assumption is to choose the gauge group for the EW part of the SM to be

$$SU_L(2) \otimes U_Y(1), \quad (89)$$



where  $SU_L(2)$  is called the weak isospin group (the weak isospin is an analogue of the usual isospin introduced by Heisenberg to describe the proton and the neutron) and  $U_Y(1)$  is the weak hypercharge group. The hypercharge group is needed because we need to have somehow the  $U(1)$  group in order to describe the electromagnetic interactions, as we already know. But if one includes simply the electromagnetic group  $U_{em}(1)$  instead of  $U_Y(1)$  the construction would not give us interactions correctly for all the fermions. The EW part  $SU_L(2) \otimes U_Y(1)$  and the  $SU_C(3)$  group for strong interactions are combined into the gauge group of the SM.

The fermion fields are taken to be in three generations, and in each generation the left components are combined into  $SU(2)$  doublets and the right components transform as  $SU(2)$  singlets:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \quad (90)$$

$e_R \quad \mu_R \quad \tau_R.$

Right-handed neutrinos are present in the original version of the SM.

For the quarks, a similar structure of representations is assumed but with singlet right-handed components for up and down types of quarks:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \quad \begin{pmatrix} c \\ s \end{pmatrix}_L \quad \begin{pmatrix} t \\ b \end{pmatrix}_L \quad (91)$$

$u_R, d_R \quad c_R, s_R \quad t_R, b_R.$

We begin with the construction of the gauge and fermion parts. Now the requirements of the gauge invariance and lowest possible dimension four of terms fix uniquely the form of the EW interaction Lagrangian. The strong interactions are described by the  $SU_C(3)$  gauge group:

$$\begin{aligned} L = & -\frac{1}{4}W_{\mu\nu}^i(W^{\mu\nu})^i - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}G_{\mu\nu}^a(G^{\mu\nu})^a \\ & + \sum_{f=\ell,q} \bar{\Psi}_L^f(iD_\mu^L\gamma^\mu)\Psi_L^\dagger + \sum_{f=\ell,q} \bar{\Psi}_R^f(iD_\mu^R\gamma^\mu)\Psi_R^\dagger, \end{aligned} \quad (92)$$

where the field strength tensors and covariant derivatives have forms very familiar to us:

$$\begin{aligned} W_{\mu\nu}^i &= \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g_2\varepsilon^{ijk}W_\mu^jW_\nu^k \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \\ G_{\mu\nu}^a &= \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_S f^{abc}A_\mu^bA_\nu^c, \\ D_\mu^L &= \partial_\mu - ig_2W_\mu^i\tau^i - ig_1B_\mu\left(\frac{Y_L^f}{2}\right) - ig_SA_\mu^at^a, \\ D_\mu^R &= \partial_\mu - ig_1B_\mu\left(\frac{Y_R^f}{2}\right) - ig_SA_\mu^at^a, \end{aligned} \quad (93)$$

$$\quad (94)$$

where  $i = 1, 2, 3$ ,  $a = 1, \dots, 8$ ;  $W_\mu^i$  are gauge fields for the weak isospin group,  $B_\mu$  are gauge fields for the weak hypercharge group and  $A_\mu$  are gluon gauge fields for the strong  $SU_C(3)$  colour group. The gauge fields are taken in the adjoint representations and the lepton and the quark fields are in the fundamental representation of  $SU_L(2)$  and  $SU_C(3)$  groups. The strongly interacting part of the SM related to the  $SU_C(3)$  colour group is called QCD, which is covered in a separate course of lectures.

Often when the SM is described the weak hypercharges  $Y_L^f$  and  $Y_R^f$  are chosen from the beginning such that the Gell-Mann–Nishijima formula is satisfied for each of the left and right chiral fermions:

$$\begin{aligned} Q_f &= (T_3^f)_L + \frac{Y_L^f}{2}, \\ Q_f &= (T_3^f)_R + \frac{Y_R^f}{2}, \end{aligned} \quad (95)$$

where  $(T_3)_{\nu_\ell} = 1/2$  and  $(T_3)_\ell = -1/2$  are projections of weak isospin operators,  $+1/2$  for the up-type and  $-1/2$  for the down-type fermions.

However, we do not know at this moment why the Gell-Mann–Nishijima formula should work in our case for the EW part. So, we do not assume from the beginning the Gell-Mann–Nishijima relations for weak hypercharges. Let us take weak hypercharges as free parameters for a moment, and try to fix them from two physics requirements:

1. correct electromagnetic interactions;
2. V–A weak charge currents.

Let us consider for simplicity only fermions from the first generation. We will consider the case of three generations later by introducing the quark mixing matrix.

From the covariant derivatives for the left and right chiral fields (94), one gets the following Lagrangian for leptons of the first generation:

$$L^\ell = -i^2 (\bar{\nu}_{e_L} \bar{e}_L) \gamma_\mu \begin{pmatrix} \frac{1}{2} g_2 W_\mu^3 + g_1 \frac{Y_L^\ell}{2} B_\mu & g_2 \frac{W_\mu^+}{\sqrt{2}} \\ g_2 \frac{W_\mu^-}{\sqrt{2}} & -\frac{1}{2} g_2 W_\mu^3 + g_1 \frac{Y_L^\ell}{2} B_\mu \end{pmatrix} \begin{pmatrix} \nu_{e_L} \\ e_L \end{pmatrix} \quad (96)$$

$$+ \bar{e}_R \gamma_\mu g_1 \frac{Y_R^\ell}{2} B_\mu e_R.$$

Here the relation following from the Pauli matrices is used:

$$\tau^i W^i = \frac{\sigma^i}{2} W^i = \frac{1}{2} \begin{pmatrix} W_\mu^3 & \sqrt{2} W_\mu^+ \\ \sqrt{2} W_\mu^- & -W_\mu^3 \end{pmatrix}, \quad (97)$$

where  $W_\mu^\pm = (W_\mu^1 \mp iW_\mu^2) / \sqrt{2}$ .

Products of non-diagonal elements give us the form of the charge current:

$$L_{CC}^\ell = \frac{g_2}{\sqrt{2}} \bar{\nu}_{e_L} \gamma_\mu W_\mu^+ e_L + \text{h.c.} = \frac{g_2}{2\sqrt{2}} \bar{\nu}_e \gamma_\mu (1 - \gamma_5) W_\mu^+ e + \text{h.c.} \quad (98)$$

The interaction Lagrangian (98) contains the lepton charge current with the needed V–A structure.

Products of diagonal elements in (96) lead to the neutral current interaction Lagrangian<sup>1</sup>:

$$L_{NC}^\ell = \bar{\nu}_{e_L} \gamma_\mu \left( \frac{1}{2} g_2 W_\mu^3 + g_1 \frac{Y_L^\ell}{2} B_\mu \right) \nu_{e_L} \\ + \bar{e}_L \gamma_\mu \left( -\frac{1}{2} g_2 W_\mu^3 + g_1 \frac{Y_L^\ell}{2} B_\mu \right) e_L \quad (99) \\ + \bar{e}_R \gamma_\mu g_1 \frac{Y_R^\ell}{2} B_\mu e_R.$$

Generically, the neutral component of the  $W$  field and the  $B$  field can mix with some mixing angle  $\theta_W$ :

$$W_\mu^3 = Z_\mu \cos \theta_W + A_\mu \sin \theta_W, \quad (100) \\ B_\mu = -Z_\mu \sin \theta_W + A_\mu \cos \theta_W.$$

The angle  $\theta_W$  is called the Weinberg mixing angle.

<sup>1</sup>Question for students: why can arbitrary hypercharge not exist in the case of non-abelian gauge symmetry?

One of these fields, say  $A$ , we try to identify with the photon—it should not interact with the neutrino and should have the well-known Dirac interaction with the electron field as needed in QED. These two physics requirements lead to three simple equations:

$$\frac{1}{2} \left( -\frac{g_2}{2} \sin \theta_W + \frac{g_1}{2} Y_L^\ell \cos \theta_W \right) + \frac{1}{2} \frac{g_1}{2} Y_R^\ell \cos \theta_W = Q_e e \quad (Q_e = -1), \quad (101)$$

$$\frac{1}{2} \left( \frac{g_2}{2} \sin \theta_W - \frac{g_1}{2} Y_L^\ell \cos \theta_W \right) + \frac{1}{2} \frac{g_1}{2} Y_R^\ell \cos \theta_W = 0, \quad (102)$$

$$\frac{g_2}{2} \sin \theta_W + \frac{g_1}{2} Y_L^\ell \cos \theta_W = 0. \quad (103)$$

The first equation (101) comes from the coefficient in front of the  $\gamma_\mu$  structure in the interaction of the electron with the  $A_\mu$  field, the second equation (102) follows from the coefficient in front of the  $\gamma_\mu \gamma_5$  structure and the third one (103) comes from the absence of the interaction of the neutrino field with  $A_\mu$ .

Therefore,

$$-\frac{g_2}{2} \sin \theta_W + \frac{g_1}{2} Y_L^\ell \cos \theta_W = \frac{g_1}{2} Y_R^\ell \cos \theta_W = Q_e e. \quad (104)$$

From the equations (104), we get

$$g_1 Y_L^\ell \cos \theta_W = -e,$$

$$g_2 \sin \theta_W = e$$

and

$$Y_R^\ell = 2Y_L^\ell. \quad (105)$$

As we can see, the hypercharges of the left and right chiral leptons are proportional but not fully fixed.

In the quark sector there are both left and right chiral components for up and down quarks:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \quad u_R, \quad d_R.$$

Then, from the Lagrangian (92), the interaction of the quarks with gauge fields is

$$\begin{aligned} (\bar{u}d)_L \gamma_\mu \left( \begin{array}{cc} \frac{1}{2} g_2 W_\mu^3 + g_1 \frac{Y_L^q}{2} B_\mu & g_2 \frac{W_\mu^+}{\sqrt{2}} \\ g_2 \frac{W_\mu^-}{\sqrt{2}} & -\frac{1}{2} g_2 W_\mu^3 + g_1 \frac{Y_L^q}{2} B_\mu \end{array} \right) \begin{pmatrix} u \\ d \end{pmatrix}_L \\ + \bar{u}_R \gamma_\mu g_1 \frac{Y_R^u}{2} B_\mu u_R + \bar{d}_R \gamma_\mu g_1 \frac{Y_R^d}{2} B_\mu d_R. \end{aligned} \quad (106)$$

The charge current has, as expected, the needed V–A form:

$$L_{CC}^q = \frac{g_2}{2\sqrt{2}} \bar{u} \gamma_\mu (1 - \gamma_5) W_\mu^+ d + \frac{g_2}{2\sqrt{2}} \bar{d} \gamma_\mu (1 - \gamma_5) W_\mu^- u. \quad (107)$$

In the same way as was done for the lepton case, one should require correct electromagnetic interactions for both  $u$  and  $d$  quarks. This means that one should have a QED electromagnetic Lagrangian with electric charges  $\frac{2}{3}$  for up-quark and  $-\frac{1}{3}$  for down-quark fields. Substituting  $W_\mu^3$  and  $B_\mu$  in terms of  $A_\mu$  and  $Z_\mu$  fields (99) into (106), we get the following equalities:

$$\begin{aligned} \frac{1}{2} g_2 \sin \theta_W + \frac{1}{2} g_1 Y_L^q \cos \theta_W &= \frac{1}{2} Y_R^u g_1 \cos \theta_W = \frac{2}{3} e, \\ -\frac{1}{2} g_2 \sin \theta_W + \frac{1}{2} g_1 Y_L^q \cos \theta_W &= \frac{1}{2} Y_R^d g_1 \cos \theta_W = -\frac{1}{3} e. \end{aligned} \quad (108)$$

From these four equations one gets the following four equalities:

$$\begin{cases} g_2 \sin \theta_W = e, \\ g_1 Y_L^q \cos \theta_W = \frac{1}{3}e, \\ Y_R^u = -2Y_R^d, \\ Y_R^u + Y_R^d = 2Y_L^q, \end{cases} \quad (109)$$

which are consistent with equalities we obtained from the lepton sector:

$$\begin{cases} g_2 \sin \theta_W = e, \\ g_1 Y_L^\ell \cos \theta_W = -e, \\ Y_R^\ell = 2Y_L^\ell. \end{cases} \quad (110)$$

Note that, as follows from (109) and (110),

$$Y_L^\ell = -3Y_L^q,$$

which means that there is only one independent hypercharge, say  $Y_L^\ell$ , and all the others may be expressed in terms of it.

Let us recall that up to now we did not assume any additional relations such as ( $Q = T_3 + Y/2$ ), which are usually assumed from the very beginning.

Now we can write the generic Lagrangian for neutral current interactions with introduced bosons  $A_\mu$  and  $Z_\mu$  in the following form:

$$L_{\text{NC}} = e \sum_f Q_f J_{f\mu}^{\text{em}} A^\mu + \frac{e}{4 \sin \theta_W \cos \theta_W} \cdot \sum_f J_{f\mu}^Z Z^\mu, \quad (111)$$

where  $J_{f\mu}^{\text{em}} = \bar{f} \gamma_\mu f$ ,  $Q_\nu = 0$ ,  $Q_e = -1$ ,  $Q_u = 2/3$ ,  $Q_d = -1/3$ ,

$$J_{f\mu}^Z = \bar{f} \gamma_\mu [v_f - a_f \gamma_5] f,$$

$$v_\nu = 1, \quad a_\nu = 1, \quad v_e = -1 + 4 \sin^2 \theta_W, \quad a_e = -1;$$

$$v_u = 1 - \frac{1}{3} \left( 4 + \frac{Y_R^u}{Y_L^q} \right) \sin^2 \theta_W, \quad a_u = 1 - \frac{1}{3} \left( 4 - \frac{Y_R^u}{Y_L^q} \right) \sin^2 \theta_W,$$

$$v_d = -1 + \frac{1}{3} \left( 2 - \frac{Y_R^d}{Y_L^q} \right) \sin^2 \theta_W, \quad a_d = -1 + \frac{1}{3} \left( 2 + \frac{Y_R^d}{Y_L^q} \right) \sin^2 \theta_W.$$

Since the structure of all three generations is the same, the equality (111) is the same for all leptons and quarks. Vector and axial-vector couplings  $v_f$  and  $a_f$  may be expressed for all the fermions in a compact common form via the fermion charge  $Q_f$  and a component of the fermion weak isospin  $T_3^f$ :

$$v_f = 2T_3^f - 4Q_f \sin^2 \theta_W, \quad a_f = 2T_3^f.$$

Note that the hypercharge parameters  $Y_L^\ell$  and  $Y_R^\ell$  are not present in (111) while in the quark sector one free parameter, which as we saw may be expressed in terms of  $Y_L^\ell$ , remains taking into account (109),

$$\frac{Y_R^u}{Y_L^q} + \frac{Y_R^d}{Y_L^q} = 2.$$

We have obtained the Lagrangian, the sum of (98), (107), (111) and the gauge boson kinetic terms from (92), which contains QED with massless fermions, an additional vector particle  $Z_\mu$  interacting with new neutral currents, two charge massless vector particles  $W_\mu^\pm$  interacting with V-A charge currents and self-interactions of the gauge particles.

Since our Lagrangian has a rather non-trivial chiral structure, an important question arises as to whether or not our construction is free of chiral anomalies, which is absolutely needed for a theory to be self-consistent.

## 11 Anomalies

Detailed discussion of anomalies is not a subject of our brief notes and can be found in a number of textbooks (see for example [6, 8]).

Generically, anomalies correspond to a situation in the field theory when some symmetry takes place at the level of a classical Lagrangian but it is violated at quantum loop level. For us an important anomaly is the chiral anomaly. In short, it means that, for example, the Lagrangian

$$L = \bar{\Psi}_L i\gamma^\mu (\partial_\mu - igA_\mu^a t^a) \Psi_L = \bar{\Psi} i\gamma^\mu (\partial_\mu - igA_\mu^a t^a) \frac{1 - \gamma_5}{2} \Psi \quad (112)$$

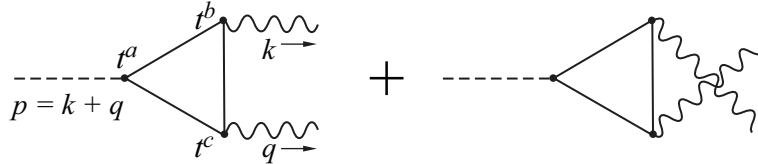
is invariant under the transformation

$$\Psi \rightarrow \exp\left(i\alpha^a t^a \frac{1 - \gamma_5}{2}\right) \Psi, \quad A_\mu \rightarrow A_\mu^a + \frac{1}{g} \partial_\mu \alpha^a + f^{abc} A_\mu^b \alpha^c. \quad (113)$$

This invariance according to the Noether theorem leads to conserving of the current:

$$j_\mu^a = \bar{\Psi} \gamma_\mu \frac{1 - \gamma_5}{2} t^a \Psi. \quad (114)$$

However, after quantization one finds that the current (114) cannot be conserved due to the triangle loop contributions shown in Fig. 8.



**Fig. 8:** Loop corrections

The diagrams in Fig. 8, after a convolution with the momentum  $p_\mu$  instead of being equal to zero, are proportional to the factor

$$\frac{g^2}{8\pi^2} \varepsilon^{\mu\nu\alpha\beta} k^\alpha q^\beta \cdot \text{Tr} \left[ t^a \{ t^b t^c \} \right]. \quad (115)$$

If the anomaly is not vanishing the theory loses its gauge invariance and therefore cannot be acceptable. (However, in cases of some currents external with respect to the theory, which have nothing to do with symmetries of the theory and the Noether theorem, anomalies or such currents may take place. This does not lead to any problems. Moreover, such type of anomalies may have very important physics consequences, as in the case of  $\pi^0$  decay to two photons.) In the SM there are simultaneously contributions from left and right chiral fermions which contribute to the anomaly with opposite signs. The anomaly is then proportional to the differences between traces of group generators coming from fermions with left and right chiralities:

$$\text{Anom} \sim \text{Tr} \left[ t^a \{ t^b t^c \} \right]_L - \text{Tr} \left[ t^a \{ t^b t^c \} \right]_R. \quad (116)$$

In the theories like QED or QCD there are no  $\gamma^5$  matrices involved in the Lagrangians. Therefore, the left and right chiral contributions exactly compensate each other, the anomaly is equal to zero and the theories perfectly make sense.

In the EW part of the SM left and right states couple to  $U_Y(1)$  gauge bosons with different hypercharges, and only the left components couple to the  $SU_2(2)$  gauge bosons. So, it is not obvious a priori that chiral anomalies vanish. In fact, zero anomalies is a request to the SM to be a reasonable quantum theory.

Since the generators of the  $SU_L(2)$  group are the matrices  $t^i = \sigma^i/2$ , then any combination of three  $SU_C(2)$  generators gives zero traces in (116) and therefore zero anomalies. The only potentially dangerous ones are

$$(SU_L(2))^2 \cdot U_Y(1) \text{ and } U_Y(1)^3$$

anomalies. For the first type we have to take into account  $\{t^i, t^j\} = \frac{1}{2}\delta^{ij}$  and then the only non-zero contribution is

$$\text{Anom} \sim \text{Tr} [Y\{t^i, t^j\}]_L = \frac{1}{2}\delta^{ij}\text{Tr}Y_L = \frac{1}{2}\delta^{ij} [N_C \cdot 2Y_L^q + 2Y_L^\ell]. \quad (117)$$

From the relations (108) and (109), as already was mentioned, we have the following relation between hypercharges:

$$Y_L^\ell = -3Y_L^q. \quad (118)$$

After substitution of (118) into (117), we get

$$\text{Anom} \sim \frac{1}{2}\delta^{ij}2Y_L^q(N_C - 3). \quad (119)$$

It is very interesting that the anomaly vanishes only for the number of colours equal to three, as it does in QCD. However, the value of the hypercharge  $Y_L^q$  is not fixed.

The second type  $((U_Y(2))^3)$  of anomaly for the fermions for each generation is proportional to

$$\begin{aligned} \text{Anom} &\sim \text{Tr}(Y_L^3) - \text{Tr}(Y_R^3) \\ &= N_C(Y_L^q)^3 \cdot 2 + (Y_L^\ell)^3 \cdot 2 - N_C [(Y_R^u)^3 + (Y_R^d)^3] - (Y_R^\ell)^3, \end{aligned} \quad (120)$$

where the factor (2) in the left-hand contribution comes from two (u and d) quarks and two (e and  $\nu_e$ ) leptons. Taking into account from (118) and (108) that  $Y_R^u + Y_R^d = 2Y_L^q$ , one gets from (120) the following:

$$\begin{aligned} \text{Anom} &\sim Y_L^\ell [2N_C(\frac{1}{3}Y_L^\ell + Y_R^u)^2 - 6(Y_L^\ell)^2] \\ &= Y_L^\ell \cdot 6(\frac{1}{3}Y_L^\ell + Y_R^u - Y_L^\ell)(\frac{1}{3}Y_L^\ell + Y_R^u + Y_L^\ell). \end{aligned} \quad (121)$$

In order to get zero for the anomaly,

$$Y_R^u = \frac{2}{3}Y_L^\ell \text{ or } Y_R^u = -\frac{4}{3}Y_L^\ell. \quad (122)$$

At this point one cannot prefer one of the relations in (122). This value will be finally fixed only after  $SU_L(2) \times U_Y(1)$  symmetry breaking.

So, we note once more that we have constructed a theory with the Lagrangian for massless fermions and gauge bosons, which gives us:

1. correct V-A charge currents;
2. correct electromagnetic interactions;
3. no chiral anomalies;

4. predictions of additional neutral currents observed experimentally.

However, obviously such a theory cannot describe nature correctly. We do not observe massless EW vector particles, except for the photon, and we do not observe massless fermions, except for, maybe, the neutrinos (or one of the neutrinos). Massive W and Z bosons, massive leptons and quarks are observed experimentally.

So, we have to introduce masses in the theory. But we cannot do it directly without violation of the basic principle of gauge invariance. Indeed, the mass term for the vector field  $m_V^2 V^\mu V_\mu$  is not invariant under the gauge transformation  $V_\mu \rightarrow V_\mu + \partial_\mu \alpha$ , and the mass term for fermions  $m_\Psi \bar{\Psi} \Psi$ , the Dirac mass, is equal to  $m (\bar{\Psi}_L \Psi_R + \bar{\Psi}_R \Psi_L)$ , and it is also not gauge invariant. Indeed, the left field  $\Psi_L$  is the doublet and the right field  $\Psi_R$  is the singlet with respect to the group  $SU_L(2)$ . How to make massive particles without violation of the basic principle of gauge invariance? There is a way to resolve this problem, which is related to spontaneous symmetry breaking, the Nambu–Goldstone theorem and the Brout–Englert–Higgs–Hagen–Guralnik–Kibble mechanism.

## 12 Spontaneous symmetry breaking and the Brout–Englert–Higgs–Hagen–Guralnik–Kibble mechanism

The situation when the Lagrangian is invariant under some symmetry while the spectrum of the system is not invariant is very common for spontaneous symmetry breaking (for example, Ginzburg–Landau theory). But a naive realization of ideas of spontaneous symmetry breaking leads to a problem manifested in the appearance of so-called Nambu–Goldstone bosons with zero masses.

To illustrate this, let us consider a very simple scalar model with the Lagrangian

$$L = \partial_\mu \varphi^\dagger \partial^\mu \varphi - \mu^2 \varphi^\dagger \varphi - \lambda (\varphi^\dagger \varphi)^2. \quad (123)$$

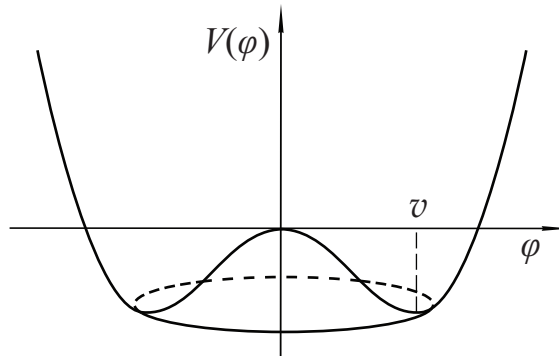
The Lagrangian (123) is invariant under the phase shift  $\varphi \rightarrow \varphi e^{i\omega}$  with  $\omega = \text{Const}$ . The case with  $\mu^2 > 0$  is trivial and not interesting for us. In the case  $\mu^2 = -|\mu^2| < 0$ , the potential shown in Fig.9

$$V(\varphi) = \mu^2 \varphi^\dagger \varphi + \lambda (\varphi^\dagger \varphi)^2 \quad (124)$$

has a non-trivial minimum:

$$\left. \frac{dV}{d\varphi^\dagger} \right|_{\varphi_0} = -|\mu^2| \varphi_0 + 2\lambda (\varphi_0^\dagger \varphi_0) \varphi_0 = 0 \Rightarrow |\varphi_0| = \sqrt{\frac{|\mu^2|}{2\lambda}} = \frac{v}{\sqrt{2}} > 0.$$

The system takes some concrete value for the vacuum solution, say  $\varphi_0 = +v/\sqrt{2}$ , which violates the phase-shift symmetry.



**Fig. 9:** Higgs potential

A complex scalar field can be parametrized by two real fields

$$\varphi = \frac{1}{\sqrt{2}}(v + h(x))e^{-i\xi(x)/v}. \quad (125)$$

In terms of the new fields  $h(x)$  and  $\xi(x)$ , the Lagrangian has the following form:

$$\begin{aligned} L = & \frac{1}{2}\partial_\mu h \partial^\mu h - \lambda v^2 h^2 - \lambda v h^3 - \lambda h^4/4 \\ & + \frac{1}{2}\partial_\mu \xi \partial^\mu \xi + \frac{2}{v}\partial_\mu \xi \partial^\mu \xi h + \frac{1}{v^2}\partial_\mu \xi \partial^\mu \xi h^2 + \lambda v^4/4. \end{aligned} \quad (126)$$

The Lagrangian (126) describes the system of a massive scalar field  $h$  with mass  $m_h^2 = 2\lambda v^2$  interacting with a massless scalar field  $\xi(x)$ . The field  $\xi(x)$  is called the Nambu–Goldstone boson field.

This is a particular case of the generic Goldstone theorem. If the theory is invariant under a global group with  $m$  generators but the vacuum is invariant under transformations generated only by  $\ell$ -generators, then in theory there exist  $m - \ell$  massless Nambu–Goldstone bosons.

Consider the system described by the Lagrangian

$$L = \frac{1}{2}\partial_\mu \varphi \partial^\mu \varphi - V(\varphi).$$

Let the Lagrangian be invariant under  $i = 1, \dots, m$  transformations:

$$\varphi \rightarrow \varphi' = \varphi + \delta\varphi, \quad \delta\varphi_i = i\delta\Theta_A t_{ij}^A \varphi_j.$$

The invariance of the potential means that

$$\delta V = (\partial V / \partial \varphi_i) \delta\varphi_i = i\delta\Theta^A (\partial V / \partial \varphi_i) t_{ij}^A \varphi_j = 0. \quad (127)$$

Let the potential have a minimum (vacuum) at some field value  $\varphi_i = \varphi_i^0$ :

$$(\partial V / \partial \varphi_i)(\varphi_i = \varphi_i^0) = 0. \quad (128)$$

We consider the case that the vacuum is invariant under transformations generated only by  $\ell$ -generators from all  $m$  generators corresponding to the symmetry, which means that

$$t_{ij}^A \varphi_j^0 = 0 \quad (129)$$

only for  $i = 1, \dots, \ell$  ( $\ell < m$ ).

The second derivative from the invariance condition (127) at the minimum leads to

$$\frac{\partial^2 V}{\partial \varphi_k \partial \varphi_i}(\varphi_i = \varphi_i^0) t_{ij}^A \varphi_j^0 + \frac{\partial V}{\partial \varphi_i}(\varphi_i = \varphi_i^0) t_{ij}^A = 0.$$

The second term here is equal to zero due to (128) and therefore

$$\frac{\partial^2 V}{\partial \varphi_k \partial \varphi_i}(\varphi_i = \varphi_i^0) t_{ij}^A \varphi_j^0 = 0.$$

For the first  $\ell$ , this equality takes place because of (129). However, for other fields with  $i = \ell + 1, \dots, m$ , the following equality has to be valid:

$$\frac{\partial^2 V}{\partial \varphi_k \partial \varphi_i}(\varphi_i = \varphi_i^0) = 0. \quad (130)$$



But the second derivative of the potential in (130) is nothing but the mass term for these  $i = \ell + 1, \dots, m$  fields. And, the equation (130) tells us that the masses of these fields are equal to zero. So, in such a situation when the vacuum of a system is not invariant under all the symmetry transformations of the Lagrangian there are  $m - \ell$  massless fields (Nambu–Goldstone bosons) corresponding to the number of generators violating the symmetry.

Let us return to the SM gauge group  $SU_L(2) \times U_Y(1)$  and add to the system one additional complex scalar field  $\Phi(x)$ , being an  $SU_L(2)$  doublet and a  $U_Y(1)$  singlet:

$$L_\Phi = D_\mu \Phi^\dagger D^\mu \Phi - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^4. \quad (131)$$

The Lagrangian  $L_\Phi$  is gauge invariant; the covariant derivative has the form

$$D_\mu = \partial_\mu - ig_2 W_\mu^i \tau^i - ig_1 \frac{Y_H}{2} B_\mu. \quad (132)$$

As in our previous example, let the mass parameter squared be negative,  $\mu^2 = -|\mu^2|$ , and therefore the field potential has a non-trivial minimum at  $\Phi = v/\sqrt{2}$ .

One can parametrize the complex field doublet  $\Phi(x)$  by four real fields in the following generic way:

$$\Phi(x) = \exp\left(-i \frac{\xi^i(x) t^i}{v}\right) \begin{pmatrix} 0 \\ (v+h)/\sqrt{2} \end{pmatrix}, \quad (133)$$

where four scalar fields  $\xi_1, \xi_2, \xi_3$  and  $h$  are introduced.

The Lagrangian (131) is invariant under an  $SU_L(2)$  transformation:

$$\Phi(x) \rightarrow \Phi'(x) = \exp(ig_2 \alpha^i t^i) \Phi(x), \quad (134)$$

where  $t^i = \sigma^i/2$  are the generators of the  $SU_L(2)$  gauge group. If we compare (133) and (134), we can choose a special gauge  $g_2 \alpha^i(x) = \xi^i(x)/v$  such that the unitary factor  $\exp(-i\xi^i(x)t^i/v)$  disappears from all the formulas. This gauge is called the unitary gauge. The Higgs field  $\Phi(x)$  in this gauge takes therefore the following form:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+h(x) \end{pmatrix}. \quad (135)$$

The field  $\Phi$  has a non-zero vacuum expectation value and as we know it leads to a violation of the symmetry of the system.

After such spontaneous symmetry breaking, the substitution of (135) into the Lagrangian (131) with the covariant derivative (132) yields the following Lagrangian in terms of fields  $W_\mu^\pm, A_\mu$  and  $Z_\mu$  introduced before:

$$L = \frac{1}{2}(\partial_\mu h)^2 - \frac{1}{2}(2\lambda v^2)h^2 - \lambda v h^3 - \frac{\lambda}{4}h^4 \quad (136)$$

$$+ M_W^2 W_\mu^+ W^{\mu-} (1+h/v)^2 + \frac{1}{2} M_Z^2 Z_\mu Z^\mu (1+h/v)^2,$$

where

$$M_h^2 = 2\lambda v^2 \quad (137)$$

is the mass of the scalar field  $h$  called the Higgs boson and

$$M_W = \frac{1}{2} g_2 v \quad \text{and} \quad M_Z = \frac{1}{2} (g_2 \cos \theta_W + g_1 Y_H \sin \theta_W) v \quad (138)$$

are masses of the vector fields  $W_\mu^\pm$  and  $Z_\mu$ . The field  $A_\mu$  is not present in the Lagrangian (136) and therefore remains massless only in the case where the corresponding coefficient in front of it is equal to zero:

$$-\frac{1}{2} g_2 \sin \theta_W + g_1 \frac{Y_H}{2} \cos \theta_W = 0.$$

The condition from the lepton sector (103)

$$g_2 \sin \theta_W = -g_1 Y_L^\ell \cos \theta_W$$

tells us that the field  $A$  has the correct electromagnetic interactions and has zero mass simultaneously if the charged lepton and Higgs field hypercharges have equal moduli and opposite signs:

$$Y_H = -Y_L^\ell.$$

One should note that if the vacuum is invariant under some group transformation, then the generator of the group gives zero acting on the vacuum. Indeed, the invariance of the vacuum means that

$$e^{iT_i \Theta_i} \Phi_{\text{vac}} = \Phi_{\text{vac}};$$

therefore, the generator  $T_i$  is given by

$$T_i \Phi_{\text{vac}} = 0.$$

In our case

$$\Phi_{\text{vac}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

and the generator of unbroken symmetry should have a generic form:

$$T \Phi_{\text{vac}} = \frac{1}{\sqrt{2}} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix} = 0, \quad \Rightarrow \quad T = \begin{pmatrix} a_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

taking into account the fact that the generator should be Hermitian. For the group  $SU_L(2) \times U_Y(1)$ , such a generator is

$$T_3 + \frac{1}{2} Y_H = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{1}{2} Y_H \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{only if } Y_H = +1.$$

This reflects the fact that the vacuum should be neutral, and the remaining group is naturally the unbroken electromagnetic group  $U_{\text{em}}(1)$ :

$$\begin{aligned} SU_L(2) \times U_Y(1) &\rightarrow U_{\text{em}}(1), \\ T_3 + \frac{1}{2} Y_H &= Q_H = 0, \quad Y_H = +1. \end{aligned} \quad (139)$$

Because  $Y_H = 1$ , one gets the following relation:

$$g_2 \sin \theta_W = g_1 \cos \theta_W. \quad (140)$$

If one substitutes (140) into the equality for  $M_Z$  (138), one gets the well-known relation between masses of W and Z bosons:

$$M_W = M_Z \cos \theta_W. \quad (141)$$

The value of the Higgs hypercharge  $Y_H = 1$  fixes the lepton hypercharge  $Y_L^\ell = -1$ . Now, from the connection between hypercharges (109, 110) and following from them

$$Y_L^\ell = -3Y_L^q,$$

all the values for hypercharges of leptons and quarks with left and right chiralities are fixed:

$$Y_R^\ell = -2, Y_L^q = Y_L^u = Y_L^d = 1/3, Y_R^u = 4/3, Y_R^d = -2/3.$$

This confirms the Gell-Mann–Nishijima relation

$$I_3 + \frac{Y}{2} = Q$$

for all three leptons and for all quarks with both chiralities. It should be so from the relation between  $SU_L(2)$  and  $U_Y(1)$  generators leading to an unbroken  $U_{\text{em}}(1)$  generator.

Coming back to the Lagrangian (136) and adding to it the kinetic terms and self-interactions of the gauge fields  $W_\mu^\pm$ ,  $A_\mu$  and  $Z_\mu$ , which come from the terms of the SM Lagrangian

$$-\frac{1}{4}W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu},$$

we get the Lagrangian describing the massive Higgs boson  $h$ , massive vector fields  $W_\mu^i$  and  $Z_\mu$  and massless field  $A_\mu$ . From the Goldstone theorem we expect  $4 - 1 = 3$  massless Nambu–Goldstone bosons. But they are not present in the Lagrangian. Three would-be Nambu–Goldstone bosons  $\xi_1$ ,  $\xi_2$  and  $\xi_3$  are ‘eaten’ by three longitudinal components of the fields  $W_\mu^-$ ,  $W_\mu^+$  and  $Z_\mu$ . One should stress that while the symmetry is spontaneously broken, the gauge symmetry of the Lagrangian itself remains unbroken.

This is the famous Brout–Englert–Higgs mechanism of spontaneous symmetry breaking (Nobel prize for 2012) confirmed by the discovery of the Higgs-like boson in ATLAS and CMS experiments at the LHC.

Now we consider the fermions, leptons and quarks, of the SM, and show how the mechanism of spontaneous symmetry breaking allows us to get massive fermions without violation of the gauge invariance.

As we know, in the SM the left fermions are the  $SU(2)$  doublets and the right fermions are the singlets. There are only two gauge-invariant dimension-four operators of the Yukawa-type preserving the SM gauge invariance:

$$\bar{Q}_L \Phi d_R \quad \text{and} \quad \bar{Q}_L \Phi^C u_R, \quad (142)$$

where

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$$

is the doublet of left fermions and

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \quad \text{and} \quad \Phi^C = i\sigma^2 \Phi^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} v + h \\ 0 \end{pmatrix}$$

are the Higgs and conjugated Higgs  $SU_L(2)$  doublet fields in the unitary gauge. Corresponding to (142), charge conjugated operators have the form

$$(\bar{Q}_L \Phi d_R)^\dagger = d_R^\dagger \Phi^\dagger (\bar{Q}_L)^\dagger = d_R^\dagger \gamma^0 \gamma^0 \Phi^\dagger \gamma^0 Q_L = \bar{d}_R \Phi^\dagger Q_L$$

and

$$(\bar{Q}_L \Phi^C u_R)^\dagger = \bar{u}_R (\Phi^C)^\dagger Q_L.$$

As one can easily see, the operators of (142) type lead after spontaneous symmetry breaking to the needed terms for the fermion masses. Indeed,

$$(\bar{u}_L \bar{d}_L) \begin{pmatrix} 0 \\ v \end{pmatrix} d_R + \bar{d}_R (0 \ v) \begin{pmatrix} u_L \\ d_L \end{pmatrix} = \bar{d}_L d_R + v \bar{d}_R d_L = v (\bar{d}_L d_R + \bar{d}_R d_L) = v \bar{d}d,$$

which is the Dirac mass term for the fermion. Similarly, the operator with the  $\Phi^C$  field leads to the correct Dirac mass term for the up-type fermions:

$$v \bar{u}u.$$

However, most general operators preserving the SM gauge invariance may include mixing of the fermion fields from various generations. The most general interaction Lagrangian including operators of (142) type has the following form:

$$L_{\text{Yukawa}} = -\Gamma_d^{ij} \bar{Q}_L^i \Phi d_R^j + \text{h.c.} - \Gamma_u^{ij} \bar{Q}_L^i \Phi^C u_R^j + \text{h.c.} - \Gamma_e^{ij} \bar{L}_L^i \Phi e_R^j + \text{h.c.}, \quad (143)$$

where there are no terms with a right-handed neutrino, and  $\Gamma_{u,d,e}$  are generically possible mixing coefficients. After spontaneous symmetry breaking, one can rewrite the Lagrangian (143) in the unitary gauge as follows:

$$L_{\text{Yukawa}} = - \left[ M_d^{ij} \bar{d}_L^i d_R^j + M_u^{ij} \bar{u}_L^i u_R^j + M_e^{ij} \bar{e}_L^i e_R^j + \text{h.c.} \right] \cdot \left( 1 + \frac{h}{v} \right), \quad (144)$$

where  $M^{ij} = \Gamma^{ij} v / \sqrt{2}$ .

The physics states are the states with definite mass. So, one should diagonalize the matrices in order to get the physics states for quarks and leptons. This can be done by unitary transformations for all left- and right-handed fermions:

$$\begin{aligned} d'_{Li} &= (U_L^d)_{ij} d_{Lj}; & d'_{Ri} &= (U_R^d)_{ij} d_{Rj}; & u'_{Li} &= (U_L^u)_{ij} u_{Lj}; & u'_{Ri} &= (U_R^u)_{ij} u_{Rj} \\ \ell'_L &= (U_L^\ell) \ell_L; & \ell'_R &= (U_R^\ell) \ell_R \\ U_L U_L^\dagger &= 1, & U_R U_R^\dagger &= 1, & U_L^\dagger U_L &= 1. \end{aligned}$$

The matrices  $U$  are chosen such that

$$\begin{aligned} (U_L^u)^\dagger M_u U_R^u &= \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix}; & (U_L^d)^\dagger M_d U_R^d &= \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} \\ (U_L^\ell)^\dagger M_\ell U_R^\ell &= \begin{pmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & m_\tau \end{pmatrix}. \end{aligned}$$

Therefore, the Yukawa-type Lagrangian (144) is

$$\begin{aligned} L_{\text{Yukawa}} &= - \left[ m_d^i \bar{d}_L^i d_R^i + m_d^{*i} \bar{d}_R^i d_L^i + \right. \\ &\quad \left. + m_u^i \bar{u}_L^i u_R^i + m_u^{*i} \bar{u}_R^i u_L^i + m_\ell^i \bar{\ell}_L^i \ell_R^i + m_\ell^{*i} \bar{\ell}_R^i \ell_L^i \right] \cdot \left( 1 + \frac{h}{v} \right). \end{aligned}$$

We consider only real mass parameters  $m \equiv m^*$ . So, the Yukawa Lagrangian after diagonalization of the mass matrices contains masses of fermions and their interactions with the Higgs boson:

$$\implies L_{\text{Yukawa}} = - \left[ m_d^i \bar{d}^i d^i + m_u^i \bar{u}^i u^i + m_\ell^i \bar{\ell}^i \ell^i \right] \cdot \left( 1 + \frac{h}{v} \right). \quad (145)$$

Now one can easily see what the fermion interactions with the gauge bosons look like in the basis of the fermion physics state with definite masses.

Neutral currents have the same structure (110) with respect to flavours as the mass terms. And they, after the unitary rotation  $\Psi' \rightarrow U\Psi$ , become diagonal simultaneously with the mass terms:

$$\bar{\Psi}' \hat{O}_N \Psi' \rightarrow \bar{\Psi} \hat{O} \Psi.$$

However, charge currents

$$J_C \sim \bar{u}_L \hat{O}_{\text{ch}} d_L + \text{h.c.}$$

contain fermions rotated by different unitary matrices for the up- and down-type fermions:

$$u' \rightarrow (U_L^u)u, \quad d' \rightarrow (U_L^d)d.$$

Therefore, after the rotation one gets for the charge current

$$J_C \sim (U_L^u)^\dagger U_L^d \bar{u}_L \hat{Q} d_L.$$

The unitary matrix

$$V_{\text{CKM}} = (U_L^u)^\dagger U_L^d$$

is called the Cabibbo–Kobayashi–Maskawa (CKM) mixing matrix:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (146)$$

Concrete values for the elements of the CKM matrix are not predicted in the SM. One can show that an arbitrary unitary matrix with  $N \times N$  complex elements may be parametrized by  $N(N-1)/2$  real angles and  $(N-1)(N-2)/2$  complex phases. So, the CKM  $3 \times 3$  matrix contains three real parameters and one complex phase. The presence of this phase leads to parity and charge (CP) violation, which in this sense is a prediction of the SM with three generations. In this lecture we do not discuss physics of the CKM matrix. The flavour physics is covered in a special lecture course at the School.

Now we have all parts of the EW part of the SM Lagrangian expressed in terms of physics fields in unitary gauge:

$$L_{\text{SM}} = L_{\text{Gauge}} + L_{\text{FG}} + L_{\text{H}}. \quad (147)$$

Here, as was mentioned, the self-interactions of the gauge fields  $W_\mu^\pm$ ,  $A_\mu$  and  $Z_\mu$  come from the terms of the SM Lagrangian  $-\frac{1}{4}W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu}$ :

$$\begin{aligned} L_{\text{Gauge}} = & -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} - \frac{1}{2}W_{\mu\nu}^+ W^{-\mu\nu} \\ & + e [W_{\mu\nu}^+ W^{-\mu} A^\nu + \text{h.c.} + W_\mu^+ W_\nu^- F^{\mu\nu}] \\ & + e \frac{c_W}{s_W} [W_{\mu\nu}^+ W^{-\mu} Z^\nu + \text{h.c.} + W_\mu^+ W_\nu^- Z^{\mu\nu}] \\ & - e^2 \frac{1}{4s_W^2} [(W_\mu^- W_\nu^+ - W_\nu^- W_\mu^+) W^{-\mu} W^{+\nu} + \text{h.c.}] \\ & - \frac{e^2}{4} (W_\mu^+ A_\nu - W_\nu^+ A_\mu) (W^{-\mu} A^\nu - W^{-\nu} A^\mu) \\ & - \frac{e^2}{4} \frac{c_W^2}{s_W^2} (W_\mu^+ Z_\nu - W_\nu^+ Z_\mu) (W^{-\mu} Z^\nu - W^{-\nu} Z^\mu) \\ & - \frac{e^2}{2} \frac{c_W}{s_W} (W_\mu^+ A_\nu - W_\nu^- A_\mu) (W^{+\mu} Z^\nu - W^{-\nu} Z^\mu) + \text{h.c.}, \end{aligned} \quad (148)$$

where  $c_W = \cos \theta_W$ ,  $s_W = \sin \theta_W$ ; the gauge for a photon field may be taken differently, for example  $(\partial_\mu A^\mu) = 0$ . The Lagrangian for the interactions of fermions with the gauge bosons is

$$L_{\text{FG}} = \sum_f \bar{f}(i\hat{D})f + L_{\text{NC}} + L_{\text{CC}}, \quad (149)$$

where  $L_{\text{NC}}$  and  $L_{\text{CC}}$  are given by (111), (99) and (107). The Lagrangian for the Higgs boson and its interactions with the gauge and fermion fields is

$$\begin{aligned} L_{\text{H}} = & \frac{1}{2}(\partial^\mu h)(\partial_\mu h) + \frac{M_h^2}{2}h^2 - \frac{M_h^2}{2v}h^3 - \frac{M_h^2}{8v^2}h^4 \\ & + (M_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2}M_Z^2 Z_\mu Z^\mu) \left(1 + \frac{h}{v}\right)^2 - \sum_f m_f \bar{f}f \left(1 + \frac{h}{v}\right). \end{aligned} \quad (150)$$

All Feynman rules following from the Lagrangian (147) can be obtained with the help of the formula (83). The kinetic parts of the Lagrangians  $L_{\text{Gauge}}$  (148) and  $L_{\text{FG}}$  (148) being taken together with the gauge and fermion mass terms from  $L_{\text{H}}$  (150) give the propagators for fermions, massless photons, massive  $W^\pm$  and  $Z$  bosons, and for the Higgs boson by inverting the corresponding quadratic forms, as we know already<sup>2</sup>.

The propagator for the massive gauge bosons requires special care. The propagators for massive vector fields follow directly from the Lagrangian by inverting the quadratic form:

$$V^\mu \left( \square g_{\mu\nu} - \partial_\mu \partial^\nu + g_{\mu\nu} M_V^2 \right) V^\nu. \quad (151)$$

Then the propagator for massive vector fields ( $V = W, Z$ ) has the following structure in the unitary gauge:

$$D_{\mu\nu}(p) = \frac{-i}{p^2 - M_V^2} \left[ g_{\mu\nu} - \frac{p_\mu p_\nu}{M_V^2} \right]. \quad (152)$$

However, the term  $p_\mu p_\nu / M_V^2$  has a bad ultraviolet behaviour. This leads to the problem of proving renormalizability of the SM. To resolve the problem, one can use another gauge in which the bad ultraviolet behaviour is absent. It is convenient to express the Higgs field as follows:

$$\Phi(x) = \begin{pmatrix} -i w_g^+ \\ (v + h + i z_g) / \sqrt{2} \end{pmatrix} \quad (153)$$

and  $\Phi^\dagger$  contains  $w_g^-$ , where the notation  $w_g^\pm$  and  $z_g$  for the Goldstone bosons is introduced. The covariant derivative, being expressed in terms of the fields  $W_\mu^\pm$ ,  $Z_\mu$  and  $A_\mu$ , and constants  $e$  and  $\sin \theta_W = s_W$ , takes the form

$$D_\mu \Phi = \begin{pmatrix} \partial_\mu - i \frac{e(1 - 2s_W^2)}{2s_W c_W} Z_\mu - i e A_\mu & -i \frac{e}{\sqrt{2} s_W} W_\mu^+ \\ -i \frac{e}{\sqrt{2} s_W} W_\mu^- & \partial_\mu + i \frac{e}{2s_W c_W} Z_\mu \end{pmatrix} \Phi \quad (154)$$

Simple algebraic manipulation leads to the following Lagrangian for the Higgs-gauge part of the SM:

$$\begin{aligned} L &= (D_\mu \Phi)^\dagger (D^\mu \Phi) - \lambda (\Phi \Phi^\dagger - v^2/2)^2 \\ &= \frac{1}{2} (\partial^\mu h) (\partial_\mu h) + M_W^2 W_\mu^+ W^{\mu-} (1 + h/v)^2 + \frac{1}{2} M_Z^2 Z_\mu Z^\mu (1 + h/v)^2 \\ &\quad - M_h^2 h^2 - \lambda v h^3 - \frac{\lambda}{4} h^4 \\ &\quad - M_W \partial_\mu w_g^+ W^{\mu-} - M_W \partial_\mu w_g^- W^{\mu+} - M_Z \partial_\mu z_g Z^\mu \\ &\quad + \partial_\mu w_g^+ \partial^\mu w_g^- + \frac{1}{2} \partial_\mu z_g \partial^\mu z_g \\ &\quad - \lambda h (h + 2v) (w_g^- w_g^+ + z_g/2) - \lambda (w_g^- w_g^+ + z_g/2)^2 \\ &\quad + \text{more cubic and quadratic terms involving } w_g^\pm \text{ and } z_g \text{ fields.} \end{aligned} \quad (155)$$

The first two lines in (155) are exactly the same as in the SM in the unitary gauge. The fourth line involves massless scalar fields, the Goldstone bosons  $w_g^\pm$  and  $z_g$ . There are many terms describing interactions of the Goldstone fields. But we would like to draw attention to the third line in (155) describing the kinetic mixing of the  $w_g^\pm$  and  $z_g$  fields with  $W^\pm$  and  $Z$  fields, respectively. Such a mixing should be removed from the Lagrangian. This can be achieved by choosing proper gauge conditions.

<sup>2</sup>The complete list of the Feynman rules for interaction vertices can be found in many books and listed explicitly, for example, in the model files for the SM used in computer codes like CompHEP, Grace, CalcHEP, MadGraph, Wizard, Sherpa etc.

Indeed, if we add to (155) the following gauge fixing terms:

$$L_{\text{GF}} = -\frac{1}{\xi} (\partial_\mu W_\mu^+ - \xi M_W w_g^+) (\partial_\nu W^{\mu-} - \xi M_W w_g^-) - \frac{1}{2\xi} (\partial_\mu Z^\mu - \xi M_Z z_g)^2, \quad (156)$$

the mixing terms are cancelled out. Then the quadratic part of the SM Lagrangian for  $W_\mu^\pm$  and  $Z_\mu$  fields which gives their propagators is

$$\begin{aligned} & -\frac{1}{4} Z_{\mu\nu} Z^{\mu\nu} + \frac{1}{2} M_Z^2 Z_\nu Z^\nu - \frac{1}{2\xi} (\partial_\mu Z^\mu)^2 \\ & -\frac{1}{2} W_{\mu\nu}^+ W^{-\mu\nu} + M_W^2 W_\nu^+ W^{-\nu} - \frac{1}{\xi} (\partial_\mu W^{+\mu}) (\partial_\nu W^{-\nu}), \end{aligned} \quad (157)$$

where terms with

$$Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu, \quad W_{\mu\nu}^\pm = \partial_\mu W_\nu^\pm - \partial_\nu W_\mu^\pm$$

come from the kinetic part of the SM Lagrangian, as in the unitary case. Inverting the quadratic form (157), one gets the propagator of the massive gauge field in the so-called  $R_\xi$  gauge:

$$D_{\mu\nu}^\xi = \frac{-i}{k^2 - M_V^2} \left[ g_{\mu\nu} - (1 - \xi) \frac{k_\mu k_\nu}{k^2 - \xi M_V^2} \right], \quad (158)$$

where  $M_V$  is  $M_W$  or  $M_Z$  and  $\xi$  is the gauge parameter.

The unitary gauge is restored by the formal limit  $\xi \rightarrow \infty$ . In the Landau gauge  $\xi = 0$ , we get the transverse structure  $(g_{\mu\nu} - k_\mu k_\nu / k^2)$ , while in the 't Hooft–Feynman gauge  $\xi = 1$  the propagator contains only the part with the  $g_{\mu\nu}$  tensor. However, one should stress that in both these gauges as well as in the generic  $R_\xi$  gauge one should take into account the appearance of Faddeev–Popov ghost fields. This is done with the help of the Faddeev–Popov method in the functional integral, which we have described already. Without going into details, the following ghost fields appear:  $c_W^\pm$ ,  $c_Z$  and  $c_A$ , corresponding to the gauge fixing terms (156) and the gauge fixing  $-(\partial_\mu A^\mu)^2 / 2\xi$  term for the photon field. In contrast to pure electrodynamics where the photon ghost fields do not interact and can be omitted, in the SM the photon ghost field  $c_A$  has non-trivial interactions with the ghost  $c_W^\pm$  and Goldstone  $w_g^\pm$  fields.

Propagators of all the ghost fields have the following form:

$$D^c = \frac{i}{p^2 - \xi M_V^2},$$

where  $M_V^2$  is equal to  $M_Z^2$  for  $c_Z$ ,  $M_W^2$  for  $c_W$  and 0 for  $c_A$  ghost fields. (The complete set of all Feynman rules for interaction vertices of Goldstone and ghost fields in the SM in the  $R_\xi$  gauge is rather long and can be found in the mentioned computer codes such as CompHEP.)

So, all the propagators for massive gauge, Goldstone and ghost fields have good ultraviolet behaviour, and therefore the SM is a renormalizable quantum field theory.

### 13 Phenomenology of the SM in lowest order

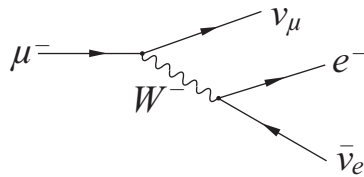
The Fermi constant  $G_F$  is measured with high precision from the muon lifetime:

$$G_F = 1.166\,378\,7(6) \times 10^{-5} \text{ GeV}^{-2}. \quad (159)$$

The decay is described in the SM by the Feynman diagram shown in Fig. 10.

Since the muon mass  $m_\mu \ll M_W$ , one can neglect the W-boson momentum in the propagator, and one immediately gets the following relation:

$$\frac{g_2^2}{8M_W^2} = \frac{G_F}{\sqrt{2}}. \quad (160)$$



**Fig. 10:**  $\mu \rightarrow e^- \bar{\nu}_e \nu_\mu$  decay diagram

As we have seen, the W-boson mass is obtained in the SM due to the Higgs mechanism and is proportional to the Higgs field vacuum expectation value  $v$ :

$$M_W^2 = \frac{1}{4} g_2^2 v^2. \quad (161)$$

From these two relations, we obtain

$$v = \frac{1}{\sqrt{\sqrt{2}G_F}} = 246.22 \text{ GeV}. \quad (162)$$

The Higgs field expectation value  $v$  is determined by the Fermi constant  $G_F$  introduced long before the Higgs mechanism appeared. At this point one can see the power of the gauge invariance principle;  $g_2$  is the same gauge coupling in the relations (160) and (161).

Now from (160) using the relation (140),  $g_1 c_W = g_2 s_W = e$  and keeping in mind  $M_W = M_Z c_W$ , one gets

$$M_W^2 \left(1 - \frac{M_W^2}{M_Z^2}\right) = \frac{\pi \alpha_{\text{em}}}{\sqrt{2}G_F} \equiv A_0^2, \quad (163)$$

where  $\alpha_{\text{em}} = e^2/4\pi$  is the usual electromagnetic fine structure constant. The low-energy

$$\alpha_{\text{em}} = (137.035999074(44))^{-1}$$

follows mainly from the electron anomalous magnetic measurements. One gets  $A_0$  very precisely from low-energy experimental results:

$$A_0 = 37.2804 \text{ GeV}. \quad (164)$$

On the other hand, one gets  $A_0$  from measured values of the masses of W and Z bosons:

$$\begin{aligned} M_W &= 80.385 \pm 0.015 \text{ GeV}, \\ M_Z &= 91.1876 \pm 0.0021 \text{ GeV} \end{aligned} \quad (165)$$

by substituting (165) into the left-hand side of (163):

$$A_0 = 37.95 \text{ GeV}. \quad (166)$$

The values (163) and (166) are rather close. The difference is about 1.5%. If one takes into account properly the higher-order corrections to the relation (163), the agreement between the two numbers will be improved further.

CC and NC interactions of the SM fermions, as has been shown in the previous section, have the following structure (see (111)):

$$\begin{aligned} L_{\text{CC}} &= \frac{g_2}{2\sqrt{2}} \sum_{ij} V_{ij} \bar{u}_i \gamma_\mu (1 - \gamma_5) d_j = \frac{e}{2\sqrt{2}s_W} \sum_{ij} V_{ij} \bar{u}_i \gamma_\mu (1 - \gamma_5) d_j, \\ L_{\text{NC}} &= e \sum_f Q_f \bar{f} \gamma_\mu f A^\mu + \frac{e}{4s_W c_W} \sum_f \bar{f} \gamma_\mu (v_f - a_f \gamma_5) f Z^\mu, \end{aligned} \quad (167)$$

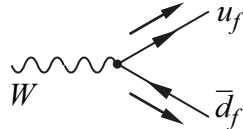


where  $V_{ij}$  are the CKM matrix elements,  $i, j = 1, 2, 3$  the number of the SM fermion generations and

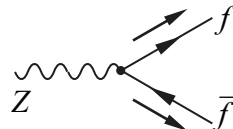
$$\begin{aligned} v_{u_i} &= 1 - \frac{8}{3}s_W^2, & a_{u_i} &= 1; & v_{d_i} &= -1 + \frac{4}{3}s_W^2, & a_{d_i} &= -1; \\ v_\ell &= -1 + 4s_W^2, & a_\ell &= -1; & v_\nu &= 1, & a_\nu &= 1 \end{aligned}$$

are the vector and axial-vector coupling constants.

The Feynman rules following from (166) allow us to get tree level formulas for the W- and Z-boson partial decay widths, as shown below at tree level:



$$\Gamma(W \rightarrow u_f \bar{d}_f) = |V_{ij}|^2 N_c \frac{\alpha}{12s_W^2} M_W, \quad (168)$$



$$\Gamma(Z \rightarrow f \bar{f}) = N_c \frac{\alpha M_Z}{12 \sin^2(2\theta_W)} [v_f^2 + a_f^2], \quad (169)$$

where the number of colours  $N_c = 3$  for quarks and  $N_c = 1$  for leptons.

The total W- and Z-boson widths are obtained by summing up all the partial widths (168) and (169). Since CCs for all SM fermions have the same V–A structure, one can very easily obtain branching fractions for W-decay modes:

$$\begin{aligned} \sum_q \text{Br}(W \rightarrow q\bar{q}) &= 2N_c \cdot \frac{1}{9} = \frac{2}{3}, \\ \sum_\ell \text{Br}(W \rightarrow \ell\nu) &= 3 \cdot \frac{1}{9} = \frac{1}{3}. \end{aligned} \quad (170)$$

The measured  $\text{Br}(W \rightarrow \ell\nu) = (10.80 \pm 0.09)\%$  is in reasonable agreement with the simple tree level result

$$\text{Br}(W \rightarrow \ell\nu) = \frac{1}{9} = (11.11)\%. \quad (171)$$

QCD corrections to the branching ratio  $\text{Br}(W \rightarrow q\bar{q})$  improve the agreement.

The decay width of the Z boson to neutrinos, the invisible decay mode, allows us to measure the number of light ( $m_\nu < M_Z/2$ ) neutrinos by comparing

$$\Gamma_{\text{inv}}^Z = \Gamma_{\text{tot}}^Z - \Gamma_{\text{had}}^Z - \Gamma_{\ell^+\ell^-}^Z \quad (172)$$

with the tree level formula obtained from (169):

$$\Gamma_{\text{inv}}^Z = \Gamma_{\nu\bar{\nu}}^Z = N_\nu \cdot \frac{\alpha M_Z}{12 \sin^2(2\theta_W)} (1 + 1). \quad (173)$$

Experimentally,  $\Gamma_{\text{tot}}$  is measured from the shape of the Z-boson resonance according to the well-known Breit–Wigner formula

$$\Gamma_{\text{tot}}^Z = 2.4952 \pm 0.0023 \text{ GeV}.$$

Decay widths to hadrons and charged leptons are measured directly in  $e^+e^-$  collisions (LEP1) to be

$$\Gamma_{\text{had}}^Z = 1744.4 \pm 2.0 \text{ MeV},$$

$$\Gamma_{\ell^+\ell^-}^Z = 83.984 \pm 0.086 \text{ MeV}.$$

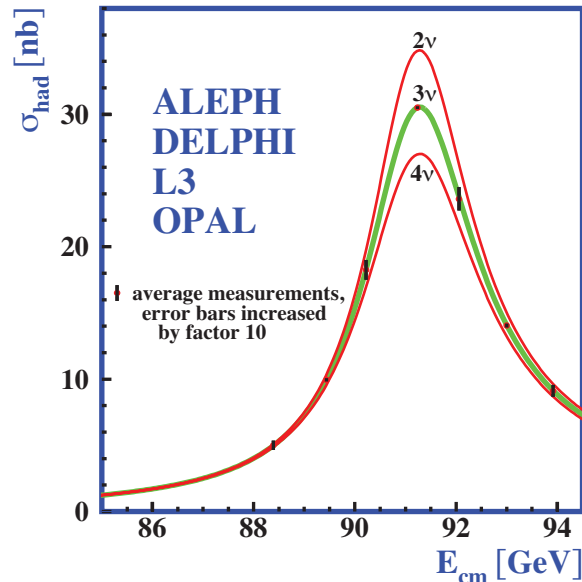
As a result,  $\Gamma_{\text{inv}}^Z$  obtained from (172) is

$$\Gamma_{\text{inv}}^Z = 0.4990 \pm 0.0015 \text{ GeV}.$$

This gives for  $N_\nu$

$$N_\nu = 2.984 \pm 0.008,$$

which is close to the number of known neutrinos. The test is an important confirmation of three generations of fermions assumed in the SM and observed in experiments as shown in Fig. 11.



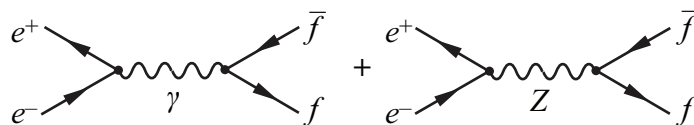
**Fig. 11:**  $N_\nu$  from LEP measurements

One can make this test by looking at the ratio  $\Gamma_{\text{inv}}^Z/\Gamma_{e^+e^-}^Z$ . In the SM, the ratio follows from (169) and (173):

$$\frac{\Gamma_{\text{inv}}^Z}{\Gamma_{e^+e^-}^Z} = \frac{2N_\nu}{1 + (1 - 4s_W^2)^2}. \quad (174)$$

The measured value ( $5.942 \pm 0.016$ ) is in agreement with 5.970 coming from the formula (174) for  $N_\nu = 3$  and  $s_W^2 = 0.2324$ .

An important part of information about the EW interactions and couplings of the SM fermions comes from  $e^+e^-$  annihilation to fermion–antifermion pairs. The differential cross-section computed from the diagrams shown in Fig. 12 has a well-known form neglecting fermion masses compared to the



**Fig. 12:**  $e^+e^-$  diagram

centre of mass energy  $\sqrt{s}$ :

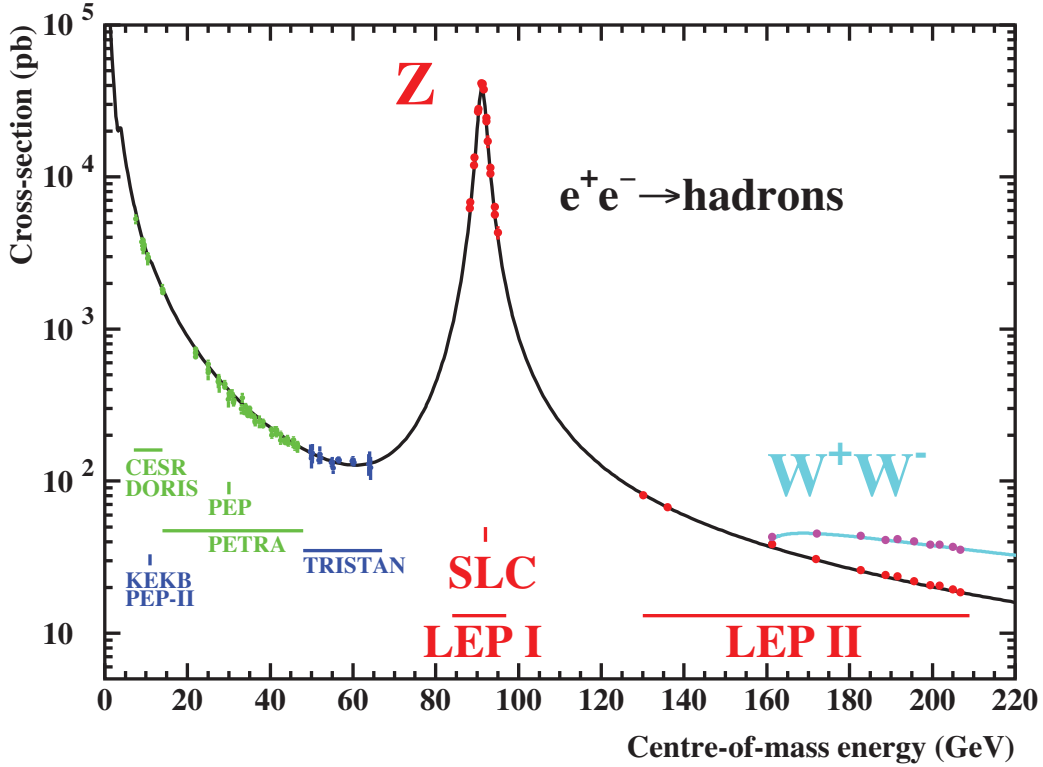
$$\begin{aligned} \frac{d\sigma}{d\cos\theta} &= \frac{2\pi\alpha^2}{4s} N_C \{ (1 + \cos^2\theta) \cdot \\ &\cdot [Q_f^2 - 2\chi_1 v_e v_f Q_f + \chi_2 (a_e^2 + v_e^2)(a_f^2 + v_f^2)] \\ &+ 2\cos\theta [-2\chi_1 a_e a_f Q_f + 4\chi_2 a_e a_f v_e v_f] \}, \end{aligned} \quad (175)$$

where

$$\chi_1 = \frac{1}{16s_W^2 c_W^2} \frac{s(s - M_Z^2)}{(s - M_Z^2)^2 + M_Z^2 \Gamma_Z^2},$$

$$\chi_2 = \frac{1}{256s_W^2 c_W^2} \frac{s^2}{(s - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}.$$

The cross-section obtained from the differential form (175) is in good agreement with the experimental data, as shown in Fig. 13. In the region far below the Z-boson pole, one can neglect the Z-boson exchange



**Fig. 13:** The cross section of hadron production in  $e^+e^-$  collisions showing a good agreement between the SM computation and various experiments in the energy range up to 220 GeV including the contribution of the Z-boson resonance.

diagram and restore the well-known QED formula

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2}{2s} Q_f^2 N_C (1 + \cos^2\theta), \quad \sigma = \frac{4\pi\alpha^2}{3} Q_f^2 N_C. \quad (176)$$

From the formula (175), one can get a number of asymmetries, which have been measured, in particular, at LEP1 and SLC. In the region close to the Z-boson pole the photon exchange part is small and can be neglected. Then the forward-backward asymmetry is

$$A_{FB} \equiv \frac{N_F - N_B}{N_F + N_B},$$

where

$$N_F = \int_0^1 d(\cos\theta) \frac{d\sigma}{d\cos\theta}, \quad N_B = \int_{-1}^0 d(\cos\theta) \frac{d\sigma}{d\cos\theta}.$$

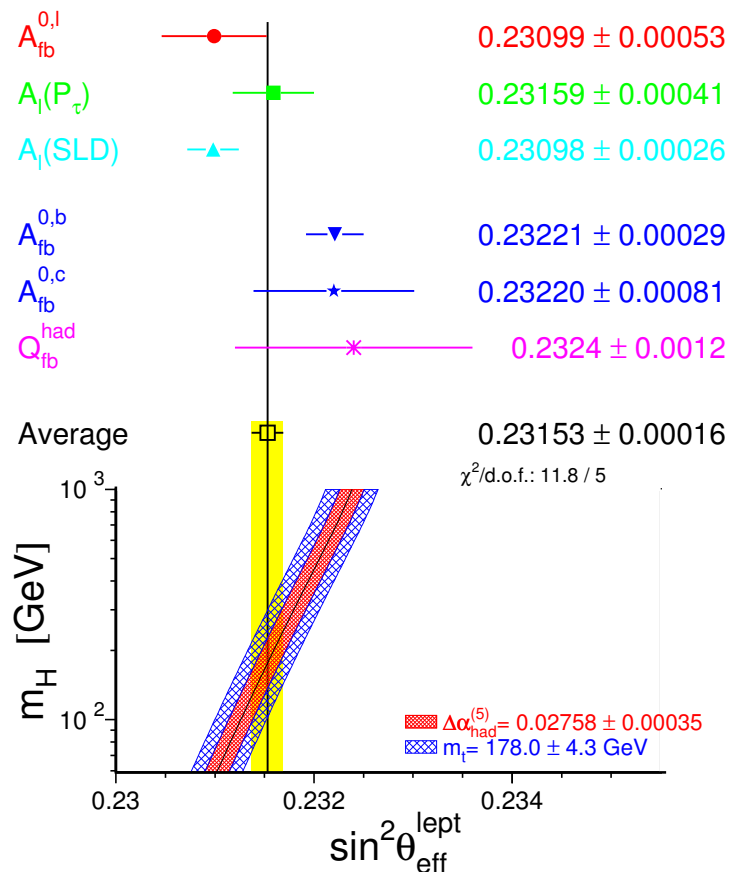
Simple integration of (175) gives the following result:

$$A_{FB} = \frac{3}{2} A_e \cdot A_f, \quad A_{e,f} = \frac{2a_{e,f} v_{e,f}}{a_{e,f}^2 + v_{e,f}^2}.$$

Measurements of decay widths being proportional to  $(a_f^2 + v_f^2)$  and asymmetries for different fermions  $f$  allow us to extract the coefficients  $a_f$  and  $v_f$ . Then one can get a precise value for the Weinberg mixing angle from the relation involving the lepton couplings:

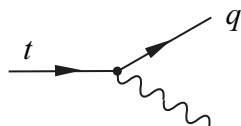
$$\sin^2 \theta_{\text{eff}}^{\text{lept}} = \frac{1}{4} \left( 1 - \frac{v_l}{a_l} \right).$$

Results of the measurements are shown in Fig.14 as obtained by the EW working group [16].



**Fig. 14:** Effective electroweak mixing angle from various observables [16]

In the SM there are no  $1 \rightarrow 2$  decays of fermions to the real Z boson due to absence of FCNCs (flavour-changing neutral currents). The top quark is heavy enough to decay to a W boson, as shown in Fig. 15. The decay mode to the b-quark is dominant in the top decays due to the CKM mixing matrix



**Fig. 15:** Top-quark decay

structure, where

$$V_{tb} \sim 1 \gg V_{ts}, V_{td}.$$

A direct tree level computation leads to

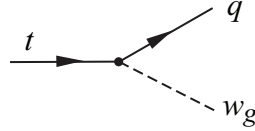
$$\Gamma_{\text{top}} = \frac{G_F M_t^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{M_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{M_t^2}\right), \quad (177)$$

where one neglects the b-quark mass:

$$\Gamma(t \rightarrow bW)_{\text{LO}} \simeq 1.53 \text{ GeV}, \quad \Gamma(t \rightarrow bW)_{\text{corr}} = 1.42 \text{ GeV}.$$

The top-quark lifetime  $\tau_{\text{top}} = 1/\Gamma_{\text{top}}$  is about  $5 \cdot 10^{-25}$  s, which is much smaller than the typical time of strong bound state formation  $\tau_{\text{QCD}} \sim 1/\Lambda_{\text{QCD}} \sim 3 \cdot 10^{-24}$  s. The top quark decays before hadronization. Therefore, there are no hadrons containing the top quark.

Since the top-quark mass is larger than the W-boson and b-quark masses, one can use the EW equivalence theorem to get the leading top width up to the term  $m_W^2/m_t^2$ . According to the EW equivalence theorem, amplitudes with external W and Z bosons are dominated by the longitudinal polarization of the bosons ( $e_L^{W,Z} \sim p^0/M_{W,Z}$ ). But the longitudinal component in the SM appears by ‘eating’ the Goldstone bosons  $w_g, z_g$ . So, one can compute simply the diagram

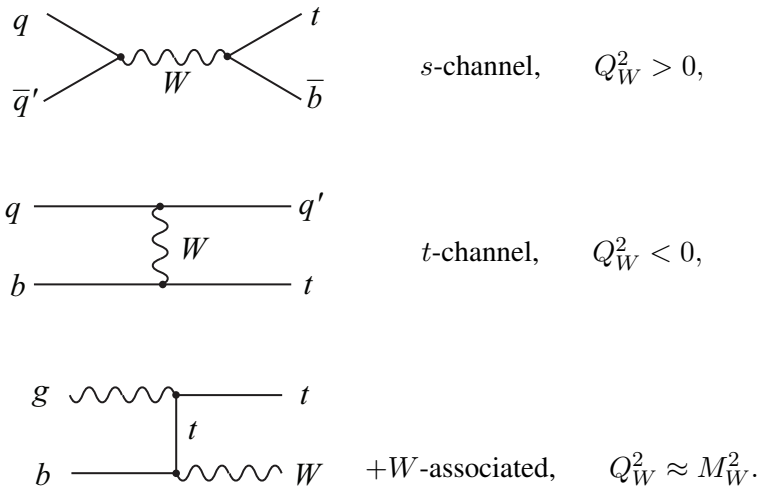


with the Yukawa vertex  $M_t/(v\sqrt{2})$ . Then one immediately obtains for the top width the following formula:

$$\Gamma = \frac{2}{32\pi} \left(\frac{M_t}{v}\right)^2 \cdot M_t = \frac{G_F M_t^3}{8\pi\sqrt{2}},$$

which is exactly equal to the first term in (177), as expected.

The EW single top quark production is another confirmation of the EW fermion structure of the SM. There are three mechanisms of single top production at hadron colliders differing by the typical virtuality ( $Q_W^2$ ) of the W boson involved:



quark  $t$ -channel and  $s$ -channel production mechanisms have been observed at the Tevatron, while  $t$ -channel and  $W$ -associated production was observed at the LHC. There are a number of important QCD

next-to-leading-order (NLO) and next-to-next-to-leading-order (NNLO) corrections which are needed to be taken into account in order to get SM predictions with needed accuracy to be compared to experimental results. Up to now a good agreement with SM computations was observed.

A well-known example demonstrating correctness of the Yang–Mills interactions of the gauge bosons is the gauge boson pair production. Triple gauge boson vertices  $WW\gamma$  and  $WWZ$  have been tested at LEP2 ( $e^+e^- \rightarrow W^+W^-$ ) and at the Tevatron ( $q\bar{q} \rightarrow W^+W^-$ ,  $q\bar{q}' \rightarrow W\gamma$ ,  $q\bar{q}' \rightarrow WZ$ ). The diagrams for the process  $e^+e^- \rightarrow W^+W^-$  form the so-called CC3 set of diagrams, as shown in Fig. 16.

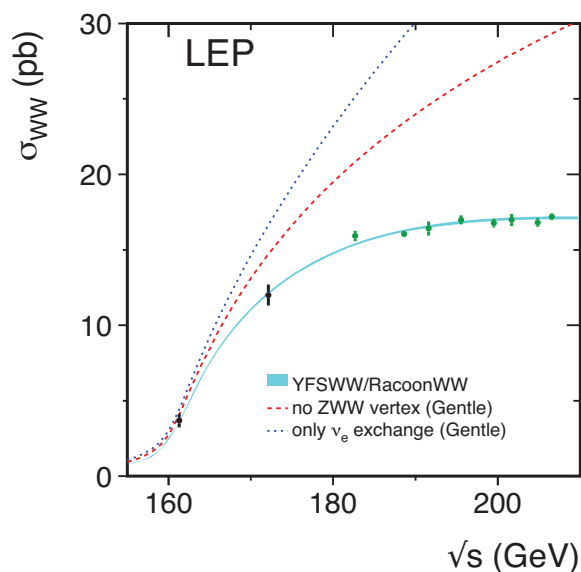


**Fig. 16:**  $e^+e^- \rightarrow W^+W^-$

The triple vertex of the Yang–Mills interaction:

$$\Gamma_{m_1 m_2 m_3}^{WW\gamma/Z}(p_1 p_2 p_3) = g_{\gamma,Z} [(p_1 - p_2)_{m_3} g_{m_1 m_2} + (p_3 - p_1)_{m_2} g_{m_1 m_3} + (p_2 - p_3)_{m_1} g_{m_2 m_3}],$$

where  $g_\gamma = e$ ,  $g_Z = g_2 c_W = e \frac{c_W}{s_W}$ , is confirmed perfectly experimentally, as shown in Fig. 17.



**Fig. 17:** Measurements of the  $W$ -pair production cross-section, compared to the theoretical predictions. For explanations, see [17].

The quartic gauge coupling  $WW\gamma\gamma$  has been tested recently at the Tevatron and the LHC in  $W$ -pair production in association with two protons or a proton and an antiproton. The quartic couplings  $WW\gamma Z$ ,  $WWZZ$  have not been tested yet. This is a challenging task for the LHC and will require a high-luminosity regime at a linear collider.

## 14 The electroweak SM beyond the leading order

All the above examples confirming the structure of the SM interactions are leading order processes. However, in many cases a high accuracy of experimental measurements requires the SM computations

beyond the leading order. In the SM, being a quantum field theory, computations of higher-order corrections face divergences of ultraviolet and infrared/collinear nature. In general, the ultraviolet (hard) divergences are treated with the help of the renormalization procedure while the infrared/collinear (soft) divergences are cancelled out due to the Kinoshita–Lee–Nauenberg theorem in summing virtual and real contributions to squared matrix elements.




The renormalization procedure is the usual way to deal with the ultraviolet divergences. We describe briefly only the main ideas of the procedure. In the SM the dimensions of all the coupling constants are equal to zero. This property has important consequences making the theory renormalizable. In simple words, the renormalizability means that all the UV divergences may be incorporated into a redefinition of a few constants such as coupling constants, masses and field normalization constants. In renormalizable theories only a few diagrams are UV divergent.

As an example, let us consider QED.

The divergency index of a diagram depends only on the number of external legs, and for QED can be expressed in a well-known form

$$w = 4 - L_\gamma - 3/2L_e,$$

where  $L_\gamma$  is the number of external photon lines and  $L_e$  is the number of external electron lines.

So, there are only three types of divergent diagrams with two external photon lines, the photon self-energy , with two external electron lines, the electron self-energy  and, with one photon and two electron external lines, the electron–photon vertex .

It is convenient to consider the renormalization procedure in the functional integral approach already familiar to us. The generating functional integral in QED in the covariant gauge is given by

$$\begin{aligned} Z[J, \eta, \bar{\eta}] &= \int D(\bar{\Psi}\Psi A) \exp \left( i \int d^4x \bar{\Psi}(i\rlap{\not{D}} - m)\Psi + ieA + J_\mu A^\mu \right. \\ &\quad \left. + \bar{\eta}\Psi + \bar{\Psi}\eta - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2\xi} \int d^4x (\partial_\mu A^\mu)^2 \right). \end{aligned} \quad (178)$$

The photon propagator is obtained from (178) by taking two functional derivatives on  $J$  and setting the sources  $J$ ,  $\eta$  and  $\bar{\eta}$  to zero:

$$\begin{aligned} iD_{\alpha\beta}(x_1, x_2) &= \int D(\bar{\Psi}\Psi A) A_\mu(x_1) A_\mu(x_2) \\ &\quad \cdot \exp \left( i \int d^4x \left[ -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\Psi}(i\rlap{\not{D}} - m)\Psi + e\bar{\Psi}A\Psi \right] \right). \end{aligned} \quad (179)$$



The Dyson–Schwinger equation for the photon propagator is obtained as a consequence of the invariance of the measure of the functional integral with respect to the shift  $A_\mu(x) \rightarrow A_\mu(x) + \varepsilon_\mu(x)$ .

The equation for the inverse propagator takes, after Fourier transformation, the form

$$D_{\alpha\beta}^{-1}(k) = (D_0)_{\alpha\beta}^{-1} + \Pi_{\alpha\beta} \quad (180)$$

or, graphically,

$$\left( \text{wavy line with loop} \right)^{-1} = \left( \text{wavy line} \right)^{-1} + \left( \text{circle with loop} \right),$$

where  denotes the dressed fermion propagator and  is the truncated one-particle irreducible vertex function  $\Gamma_\mu(p_1, p_2, k)$ .

At one-loop level the function  $\Pi_{\alpha\beta}(k)$  is given by the following Feynman integral:

$$(-ie)^2 \int \frac{d^4p}{(2\pi)^4} \text{Tr} \left[ \frac{\not{p} + m}{p^2 - m^2 + i0} \gamma_\alpha \frac{(\not{p} - \not{k}) + m}{(p - k)^2 - m^2 + i0} \gamma_\beta \right]. \quad (181)$$

The integral (181) is quadratically divergent from formal power counting. In order to deal with divergent integrals, we need to introduce some regularization. We use the dimensional regularization

$$d^4p \rightarrow d^D p (\mu^2)^{2-D/2}.$$

One can see that (181) gives zero, being convoluted with the external moments  $k_\alpha$  or  $k_\beta$ . Indeed,

$$k_\alpha \gamma^\alpha = (\not{p}) - (\not{p} - \not{k}) = (\not{p} - m) - [(\not{p} - \not{k}) - m]. \quad (182)$$

If we substitute (182) into (181), we get

$$(\mu^{2-D/2})(-ie)^2 \int \frac{d^D p}{(2\pi)^D} \text{Tr} \left[ \left( \frac{\not{p} + m}{p^2 - m^2 + i0} - \frac{(\not{p} - \not{k}) + m}{(p - k)^2 - m^2 + i0} \right) \gamma_\beta \right] = 0. \quad (183)$$

In fact, this result is valid to all perturbation orders due to the Ward identity

$$k^\mu \Gamma_\mu(p_1, p_2, k) = S^{-1}(p_1) - S^{-1}(p_2). \quad (184)$$

The identity (183) can be easily derived from  $U(1)$  gauge invariance of (178).

The property means that  $\Pi_{\alpha\beta}$  has the following form:

$$\Pi_{\alpha\beta}(k) = (g_{\alpha\beta} k^2 - k_\alpha k_\beta) \Pi(k^2). \quad (185)$$

Therefore, the dressed photon propagator can be written as

$$D_{\alpha\beta}(k) = -\frac{i}{k^2} \left[ \frac{1}{1 + \Pi_\gamma(k^2, \varepsilon, \mu^2)} \left( g_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2} \right) + \xi \frac{k_\alpha k_\beta}{k^2} \right]. \quad (186)$$

Of course, in the case where  $\Pi_\gamma = 0$ , we obtain the free photon propagator.

The factor  $(1 + \Pi_\gamma(k^2, \varepsilon, \mu^2))^{-1}$ , being taken at zero momentum, should be removed for the correct normalization of the kinetic term. It can be done by rescaling the field  $A_\mu(x)$  in the following way:

$$A_\mu(x) \rightarrow \frac{1}{\sqrt{Z_3}} A_\mu, \text{ where } Z_3^{(a)} = (1 + \Pi_\gamma(0, \varepsilon))^{-1}. \quad (187)$$

Direct computation of (181) with well-known Feynman techniques gives

$$\begin{aligned} \Pi_\gamma(k^2, \varepsilon, \mu^2) &= \frac{\alpha}{3\pi\varepsilon} + \Pi_{\text{finite}} \\ \text{and therefore } Z_3^{-1} &= 1 + \frac{\alpha}{3\pi\varepsilon}, \end{aligned} \quad (188)$$

where  $\varepsilon = (4 - D)/2$ .

Now let us consider the fermion propagator taking functional derivatives of (178) on the fermion sources  $\bar{\eta}$  and  $\eta$ .

Now the expression for the dressed inverse fermion propagator is

$$S^{-1}(p) = S_0^{-1}(p) - \Sigma(p). \quad (189)$$

The formula (188) is also the Dyson–Schwinger equation for the dressed fermion propagator, graphically presented as

$$\left( \text{---} \bigcirc \text{---} \right)^{-1} = \left( \text{---} \bullet \text{---} \right)^{-1} - \frac{k_\mu \text{---} \bigcirc \text{---} \Gamma}{p_1 \text{---} p}$$



The equation involves the same truncated vertex function  $\Gamma_\mu(p_1 p_2; k)$ . In the second order of perturbation theory,  $\Sigma^{(2)}(p)$  is as follows:

$$-i\Sigma^{(2)}(p) = (-ie)^2 \int \frac{d^D k}{(2\pi)^D} (\mu^2)^{2-D/2} \cdot \gamma^\mu D_{\mu\nu}(k) \frac{\not{p} - \not{k} + m}{(p-k)^2 - m^2} \gamma^\nu. \quad (190)$$

Direct computation gives the following answer:

$$\Sigma_2 = \frac{\alpha}{8\pi} (4m - \not{p}) \frac{2}{\varepsilon} + \Sigma_{\text{finite}}. \quad (191)$$

The generic structure of  $\Sigma(p)$  is as follows:

$$\Sigma(p) = \not{p} f_1(p^2) - m f_2(p^2).$$

Due to (189), this means that the fermion propagator has the form

$$S(p) = \frac{1}{\not{p}(1 - f_1(p^2)) - m(1 - f_2(p^2))} = -\frac{1}{1 - f_1(p^2)} \frac{1}{\not{p} - m \frac{1 - f_2(p^2)}{1 - f_1(p^2)}}. \quad (192)$$

Close to physics mass, one should have

$$m_{\text{phys}} = m \frac{1 - f_2(m_{\text{phys}}^2)}{1 - f_1(m_{\text{phys}}^2)}.$$

So, the fermion propagator has the following form close to physics mass:

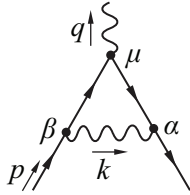
$$S(p) = \frac{Z_2(\varepsilon, \mu)}{\hat{p} - m_{\text{phys}}(\varepsilon, \mu)}, \quad (193)$$

where  $Z_2 = (1 - f_1)^{-1}$ ,  $m_{\text{phys}} = m \frac{Z_2}{Z_m}$  and  $Z_m = (1 - f_2)^{-1}$ . From the one-loop result (191), we obtain

$$Z_2^{-1} = 1 + \frac{\alpha}{4\pi\varepsilon} + O(\alpha), \quad (194)$$

$$Z_m^{-1} = 1 + \frac{\alpha}{\pi\varepsilon} + O(\alpha). \quad (195)$$

The remaining divergent QED diagram is the vertex function correction given by the integral



$$\begin{aligned} ie\Gamma_\mu^{(2)}(p, q) &= (-ie)^3 (\mu^2)^{2-D/2} \int \frac{d^D k}{(2\pi)^D} \\ &\cdot \gamma^\alpha(i) \frac{\not{p} - \not{q} - \not{k} + m}{(p-q-k)^2 - m^2 + i0} \\ &\cdot \gamma_\mu(i) \frac{\not{p} - \not{q} + m}{(p-q)^2 - m^2 + i0} \cdot \gamma^\beta \frac{-i}{k^2 + i0} g^{\alpha\beta} \end{aligned} \quad (196)$$

In order to compute the divergent part, one can simplify the problem and compute (196) in the limit  $q \rightarrow 0$ . The answer is

$$\Gamma_\mu^{(2)}(p, 0) = \gamma_\mu \left[ \frac{\alpha}{4\pi\varepsilon} + O(\alpha) \right]. \quad (197)$$

Therefore, the vertex function including one-loop correction may be written in the form

$$-ie\Gamma_\mu = -ieZ_1\gamma_\mu, \quad (198)$$

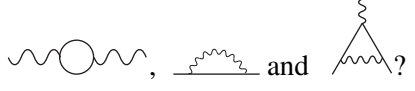


where

$$Z_1 = 1 - \frac{\alpha}{4\pi} \frac{1}{\varepsilon} + O(\alpha).$$

We can see from (194) and (198) that  $Z_1 = Z_2$  including the one-loop part. The equality  $Z_1 = Z_2$  takes place to all orders of perturbation theory due to the Ward identity

$$\Gamma_\mu(p, 0) = \partial_\mu S^{-1}(p), \quad (199)$$

as follows from (183) in the limit of the photon momentum  $k \rightarrow 0$ .

Why did we do all the above computations of divergent graphs ,  and ? Let us rewrite our initial (before renormalization) QED Lagrangian

$$L = -\frac{1}{4}F_{\mu\nu}^0 F^{0\mu\nu} + \bar{\Psi}_0(i\not{D}^0 - m_0)\Psi_0, \quad (200)$$

where  $F_{\mu\nu}^0 = \partial_\mu A_\nu^0 - \partial_\nu A_\mu^0$ ,  $D_\mu^0 = \partial_\mu - ie_0 A_\mu^0$  (we use the symbol (0) to stress that all the objects are not renormalized, or bare, as one usually says) in terms of physical fields and parameters labelled by the symbol ph and additional terms  $\Delta L$ :

$$L = -\frac{1}{4}F_{\mu\nu}^{\text{ph}} F^{\text{ph}\mu\nu} + \bar{\Psi}_{\text{ph}}(i\not{D}_{\text{ph}} - m_{\text{ph}})\Psi_{\text{ph}} + \Delta L, \quad (201)$$

where  $A_\mu^{\text{ph}} = Z_3^{-1/2} A_\mu^0$ ,  $\Psi_{\text{ph}} = Z_2^{-1/2} \Psi$ ,  $m_{\text{ph}} = (Z_2/Z_m)m_0$ ,  $e_0 = Z_1 Z_2^{-1} Z_3^{-1/2} (\mu)^{D/2-2} e_{\text{ph}}$ ,  $D_\mu^{\text{ph}} = \partial_\mu - ie_{\text{ph}} A_\mu^{\text{ph}}$  and

$$\begin{aligned} \Delta L = & -(Z_3 - 1) \frac{1}{4} F_{\mu\nu}^{\text{ph}} F^{\text{ph}\mu\nu} + (Z_2 - 1) \bar{\Psi}_{\text{ph}}(i\not{\partial})\Psi_{\text{ph}} \\ & + (Z_m - 1) m_{\text{ph}} \bar{\Psi}_{\text{ph}} \Psi_{\text{ph}} + (Z_1 - 1) e_{\text{ph}} \bar{\Psi}_{\text{ph}} (\not{A}_{\text{ph}}) \Psi_{\text{ph}}. \end{aligned} \quad (202)$$

The Lagrangian  $\Delta L$  contains so-called counter-terms. In the leading order we computed all the coefficients in front of the counter-terms. Now when one computes some effect using the Lagrangian (201) all UV divergences are cancelled out order by order in perturbation theory by contributions of the counter-terms.

Let us look in more detail at the relation for the coupling constant

$$e_0 = Z_1 Z_2^{-1} Z_3^{-1/2} (\mu)^{D/2-2} e_{\text{ph}}(\mu),$$

where  $(\mu)^{D/2-2}$  is the dimension of the charge. As we discussed above,  $Z_1 = Z_2$  due to the Ward identity, and

$$e_0 = Z_3^{-1/2} (\mu)^{D/2-2} e_{\text{ph}}(\mu). \quad (203)$$

Note that  $e_{\text{ph}}(\mu)$  is a function of the dimension regularization parameter  $\mu$ , while  $e_0$  does not depend on  $\mu$ . From (203), one gets the following equality for  $\alpha = \frac{e^2}{4\pi}$  ( $D = 4 - 2\varepsilon$ ):

$$\alpha_0 = Z_3^{-1} (\mu^2)^{-\varepsilon} \alpha_{\text{ph}}(\mu). \quad (204)$$

Taking the derivative  $\mu \frac{\partial}{\partial \mu}$  on both sides of (204), we get the following equation:

$$\mu \frac{\partial \alpha}{\partial \mu} = \frac{2\alpha^2}{3\pi} \equiv \beta(\alpha). \quad (205)$$

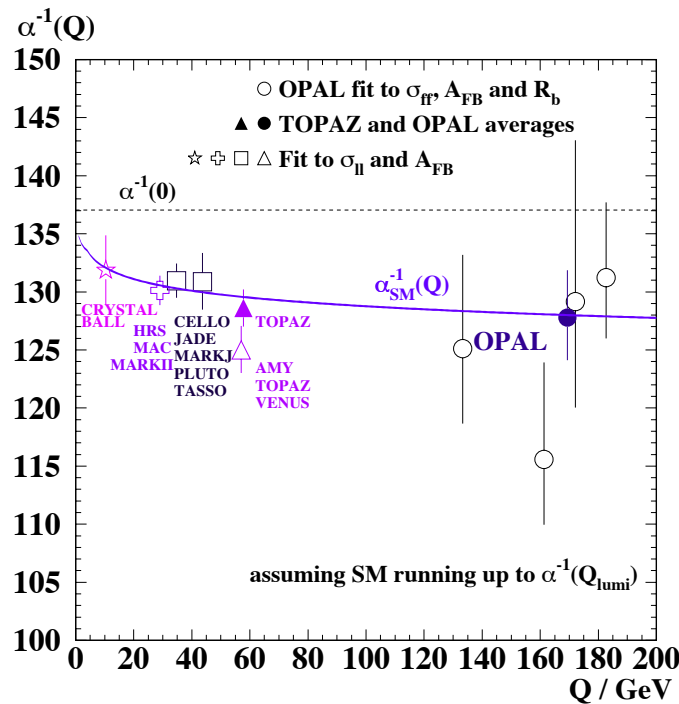
The equation (205) is a particular example of the renormalization group equation, which we do not discuss in this brief lecture course. The function on the right-hand side of (205) is called the  $\beta$ -function. So, at the one-loop level the  $\beta$ -function in QED is given by the following formula:

$$\beta(\alpha) = \frac{b_0}{\pi} \alpha^2, \quad b_0 = \frac{2}{3}. \quad (206)$$

One should stress that the coefficient  $b_0 = \frac{2}{3}$  in QED is positive. The equation (205) can be easily solved:

$$\alpha(\mu) = \frac{\alpha(\mu_1)}{1 - \frac{\alpha(\mu_1)}{3\pi} \ln(\mu/\mu_1)^2}. \quad (207)$$

This is the running coupling constant. If one measures the constant  $\alpha$  at some scale  $\mu_1$ , one gets values for the constant at other scales. The coupling constant  $\alpha = 1/137$  being measured at very small scale (small momentum transfer or large distance) in Thompson scattering increases with the scale growing and becomes  $\alpha(M_Z^2) \approx \frac{1}{129}$  at the  $Z$  mass. This fact was confirmed nicely in LEP experiments. Note that, in order to get  $1/129$ , one should take into account the contribution of all SM charged particles to the photon vacuum polarization function  $\Pi_\gamma$ . This means that the charged particle–antiparticle pairs screen the bare charge at small  $\mu^2$  or at large distances. The QED running coupling constant is illustrated in Fig.18 [18].



**Fig. 18:** Running electromagnetic coupling as a function of collision energy measured at various energies, in particular, at LEP by the OPAL collaboration [18].

Note that in QCD the  $\beta$ -function is negative, leading to an antiscreening effect;  $\alpha_S$  becomes smaller with increasing of the momentum scale (momentum transfer) or decreasing distances. This is a famous asymptotic freedom property in QCD. In QED the situation is the opposite. If the scale  $\mu$  increases to very large values the well-known Landau pole

$$\frac{b_0}{\pi} \ln(\mu/\mu_1)^2 = 1$$

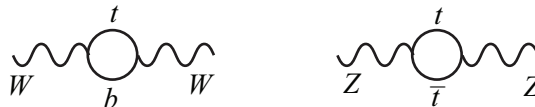
is approached where the perturbation picture in QED breaks down.

One should note that in QED all terms in the four-dimensional Lagrangian (gauge-invariant operators) have dimension four. As a result, the coupling constant in QED is dimensionless. This is crucial to have renormalizable theory, where all UV divergences are cancelled to all orders with the help of a finite number of counter-terms. This is also the case in the SM. All terms of the SM Lagrangian have dimension four and all the coupling constants are dimensionless. So, the SM is a renormalizable theory in the same manner as QED.

Naively, one may think that the EW higher-order corrections are not that important. The perturbation theory expansion parameters  $\alpha/\pi$  with  $\alpha_{\text{em}} \sim 1/129$  and  $\alpha_{\text{weak}} \sim 1/30$  are very small. However, the experimental accuracies are so high in various cases that even one-loop EW corrections might not be sufficient. Indeed, selected lists of measured parameters by LEP1, LEP2, SLD and Tevatron are given below:

$M_Z$	=	91.1875	±	0.0021	GeV	0.002%
$\Gamma_Z$	=	2.4952	±	0.0023	GeV	0.09%
$M_W$	=	80.385	±	0.015	GeV	0.02%
$M_{\text{top}}$	=	173.2	±	0.9	GeV	0.52%

The most important higher-order corrections come from resummation of the large logarithms,  $\log \frac{M_t^2}{m_e^2} \approx 24.2$ , as we have seen with running  $\alpha$  ( $1/137 \rightarrow 1/129$ ). The second class of large corrections comes from contributions of the order of  $M_{\text{top}}^2/M_W^2$ , which originate from EW Goldstone boson (or longitudinal W/Z polarization state) couplings to the quarks of the third generation. The later corrections lead to the shifts in W- and Z-boson masses coming from the diagram in Fig. 19.



**Fig. 19:** Loop corrections

Loop corrections lead to the fact that SM parameters, coupling constants and masses, are running parameters, and they are non-trivial functions of each other. A famous example is given in Fig.20 [17] showing the dependence of the W-boson mass as a function of the top-quark mass at different values of the Higgs boson mass.

One should recall that the top-quark mass has been determined indirectly from the analysis of loop corrections, it being

$$m_t = 178 \pm 8 \begin{matrix} +17 \\ -20 \end{matrix} \text{ GeV},$$

which is remarkably close to today's precise measured value  $173.2 \pm 0.9$  GeV.

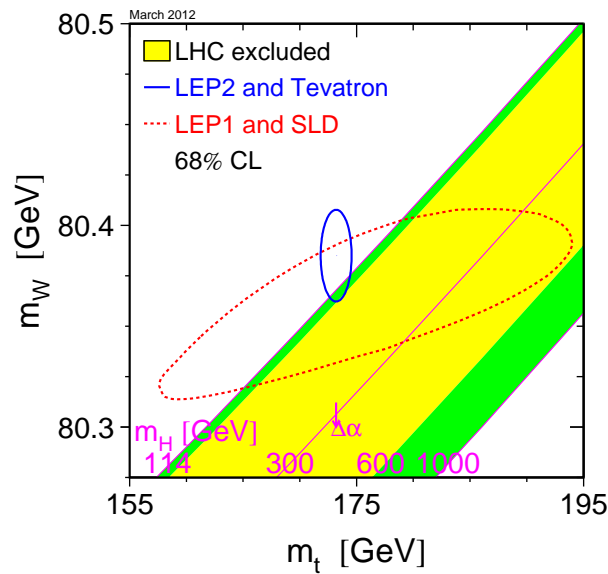
The low Higgs mass range was preferred by a similar analysis of the Tevatron and LEP data, as one can clearly see in Fig.20.

A summary of comparisons of the EW precision measurements at LEP1, LEP2, SLD and the Tevatron and a global parameter fit is given in the well-known plot shown in Fig.21 [16].

The only one discrepancy on the level of  $3\sigma$  is observed for  $b\bar{b}$ -pair forward-backward asymmetry.

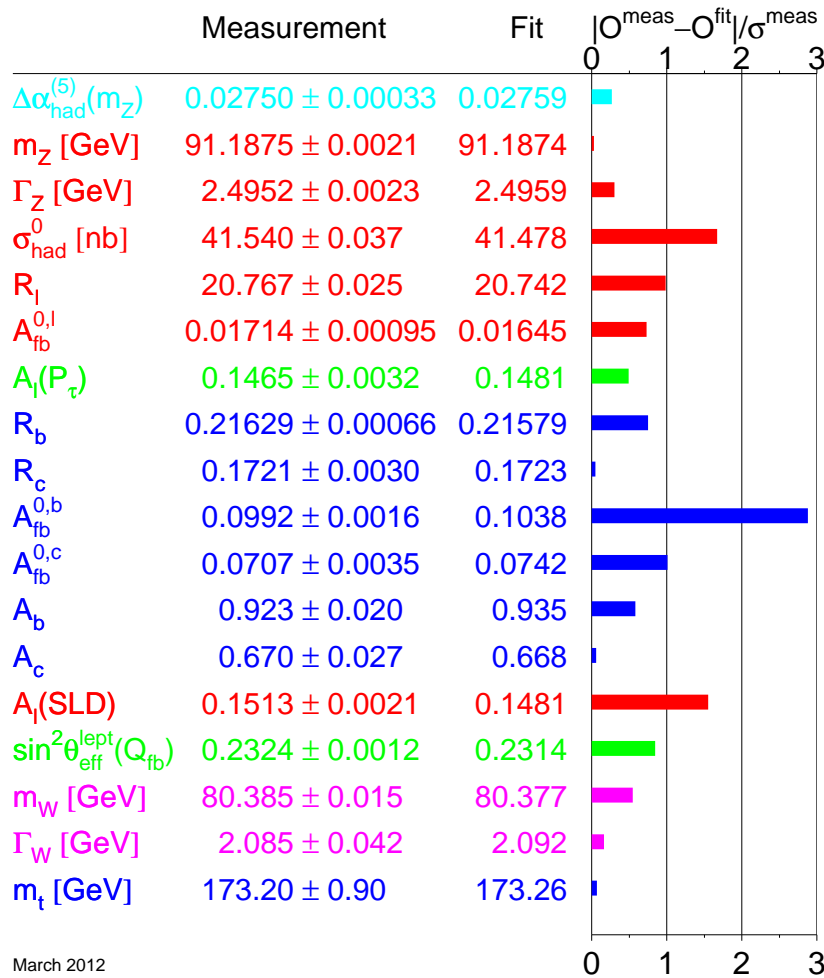
## 15 Concluding remarks

1. The SM is a renormalizable anomaly-free gauge quantum field theory with spontaneously broken EW symmetry. Remarkable agreement with many experimental measurements is observed.



**Fig. 20:** W boson mass as a function of the top quark mass at various masses of the Higgs boson (see details in [17]).

2. All SM leptons, quarks, gauge bosons and, very probably, the Higgs boson have been discovered.
3. The SM predicts the structure of all interactions: fermion-gauge, gauge self-interactions, Higgs-gauge, Higgs-fermion and Higgs self-interactions. As a result, the SM allows us to compute various cross-sections, distributions and decay rates taking into account higher-order corrections. However, not yet all of the interactions were tested experimentally. Parameters of the theory are not predicted by the theory itself but extracted from the measurements.
4. The EW SM has 17 parameters and QCD has one more parameter fixed from experiments:
  - the gauge-Higgs sector contains four parameters  $g_1, g_2, \mu^2, \lambda$  or in terms of best measured  $\alpha_{\text{em}}, G_F, M_Z, M_h$ ;
  - six quark masses, three lepton masses;
  - three mixing angles and one phase of the CKM matrix (more parameters come from the neutrino mixing matrix, which we do not consider here);
  - the QCD coupling constant  $\alpha_s$ .
5. There are facts which cannot be explained in the SM:
  - fermions have very much different masses ( $M_{\text{top}} = 173 \text{ GeV}$ ,  $M_e = 0.5 \text{ MeV}$ ) coming from the same mechanism;
  - dark matter exists in the Universe, and there are no dark-matter candidates in the SM;
  - the CKM phase as a source of CP violation in the SM is too small to explain particle-antiparticle asymmetry in the Universe;
  - neutrino masses, mixing and oscillations cannot be understood in the framework of the SM EW symmetry breaking mechanism;
  - there is some tension in explaining the muon anomalous magnetic moment.
6. As is well known, the simplest Higgs mechanism in the SM is not stable with respect to quantum corrections (naturalness problem). In the SM there is no symmetry which protects a strong



**Fig. 21:** Global fit of the EW precision measurements at LEP1, LEP2, SLD and the Tevatron by the SM computations including loop corrections [16] (the latest update version of the plot <http://lepewwg.web.cern.ch/LEPEWWG/plots/winter2012/>)

(quadratic) dependence of the Higgs mass on a possible new scale. Something is needed in addition to the SM to stabilize the mass parameter.

7. In addition, the SM does not give answers to many questions, such as:

- What is a generation? Why are there only three generations?
- How are quarks and leptons related to each other?; what is the nature of the quark–lepton analogy?
- What is responsible for gauge symmetries, why are charges quantized? Are there additional gauge symmetries?

- What is responsible for the formation of the Higgs potential?
- To which accuracy is the CPT (charge, parity, and time) symmetry exact?
- Why is gravity so weak compared to other interactions?

In our lecture we focused mainly on the EW part of the SM and aspects of the field theory needed clarifying, and we did not discuss QCD physics, Higgs boson physics, neutrino physics, flavour physics, problems of the SM models leading to BSM (beyond the Standard Model) scenarios and sequences for cosmology. These are the subjects of the lectures by Z. Trócsányi, J. Ellis, B. Gavelo, Z. Ligetti, C. Csaki and D. Gorbunov at the 2013 European School.

### Acknowledgements

I would like to thank the organizers and the participants of the School for creating a very warm and productive atmosphere. Many thanks to the organizers for their help and support. The work was partially supported by RFBR grant 12-02-93108 and NSh grant 3042.2014.2.

### References

- [1] J.D. Bjorken and S. Drell, *Relativistic Quantum Mechanics/Fields* (McGraw-Hill, New York, 1965), Vols. I, II.
- [2] E.S. Abers and B.W. Lee, *Phys. Rep.* **9** (1973) 1.
- [3] C. Itzykson and J. Zuber, *Introduction to Quantum Field Theory* (McGraw-Hill, New York, 1980).
- [4] F. Halzen and A.D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics* (Wiley, New York, 1984).
- [5] T.P. Cheng and L.F. Li, *Gauge Theory of Elementary Particle Physics* (Oxford University Press, New York, 1984).
- [6] L.D. Faddeev and A.A. Slavnov, *Gauge Fields, Introduction to Quantum Theory* (Addison-Wesley, Redwood, CA, 1991). Translation of: A.A. Slavnov and L.D. Faddeev, *Vvedenie v Kvantovuiu Teoriii Kalibrovochnykh Polei* (Nauka, Moscow, 1988), Izd. 3.
- [7] M.E. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, MA, 1995).
- [8] S. Weinberg, *The Quantum Theory of Fields* (Cambridge University Press, Cambridge, MA, 1996), Vols. I, II.
- [9] A. Zee, *Quantum Field Theory in a Nutshell* (Princeton University Press, Princeton, NJ, 2003).
- [10] W. Buchmüller and C. Lüdeling, Field theory and the Standard Model, European School of High-Energy Physics, Kitzbühel, Austria, 2005, Ed. R. Fleischer (CERN, Geneva, 2006), CERN-2006-014, p. 1 [arXiv:hep-ph/0609174].
- [11] A. Pich, The Standard Model of electroweak interactions, European School of High-Energy Physics, Aronsborg, Sweden, 2006, Ed. R. Fleischer (CERN, Geneva, 2007), CERN-2007-005, p. 1 [arXiv:hep-ph/0705.4264].
- [12] G. Altarelli, *The Standard Model of Electroweak Interactions*, Ed. H. Schopper, Springer Materials – The Landolt-Börnstein Database Groups I, V (Springer, Berlin Heidelberg, 2008), Vol. 21A, Chap. 3.
- [13] V. Rubakov, Field theory and the Standard Model, European School of High-Energy Physics, Herbeumont-sur-Semois, Belgium, 2008, Eds. N. Ellis and R. Fleischer (CERN, Geneva, 2009), CERN-2009-002, p. 1.
- [14] W. Hollik, Quantum field theory and the Standard Model, European School of High-Energy Physics, Bautzen, Germany, 2009, Eds. C. Grojean and M. Spiropulu (CERN, Geneva, 2010), CERN-2010-002, p. 1 [arXiv:hep-ph/1012.3883v1].

- [15] G. Altarelli, Collider physics within the Standard Model: a primer, CERN-PH-TH-2013-020, 2013, [arXiv:hep-ph/1303.2842].
- [16] S. Schael *et al.* [ALEPH and DELPHI and L3 and OPAL and SLD and LEP Electroweak Working Group and SLD Electroweak Group and SLD Heavy Flavour Group Collaborations], “Precision electroweak measurements on the  $Z$  resonance,” Phys. Rept. **427**, 257 (2006) [hep-ex/0509008].
- [17] S. Schael *et al.* [ALEPH and DELPHI and L3 and OPAL and LEP Electroweak Collaborations], “Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP,” Phys. Rept. **532** (2013) 119 [arXiv:1302.3415 [hep-ex]].
- [18] G. Abbiendi *et al.* [OPAL Collaboration], Eur. Phys. J. C **6** (1999) 1 [hep-ex/9808023].



## QCD for collider experiments

Z. Trócsányi\*

Department of Physics and MTA-DE Particle Physics Research Group, University of Debrecen, Debrecen, Hungary

### Abstract

These lectures are intended to provide the theoretical basis of describing high-energy particle collisions at a level appropriate to graduate students in experimental high energy physics. They are supposed to be familiar with quantum electrodynamics, the concept of Feynman rules, Feynman graphs and computation of the cross section in quantum field theory.

*When you measure what you are speaking about and express it in numbers, you know something about it, but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.*

Lord Kelvin

### 1 QCD as quantum field theory of the strong interaction

The Lagrangian of the quark-gluon field based on a non-abelian gauge symmetry was first proposed in Ref. [1] forty years ago. The paper discussed the advantages of the colour-octet picture. Since then an immense amount of research lead to a lot of interesting results and a deep understanding of the strong interaction based on this quantum field theoretical description of chromo-dynamics, QCD. Today we are convinced that QCD is the correct description of the strong interaction, yet we still lack a complete and satisfactory solution. In such a situation one may set two goals: (i) either an ambitious one: solve QCD, or (ii) a more pragmatic one: develop tools for modeling particle interactions in high energy collider experiments. In these lectures we go for the second one.

Our aim is to understand high-energy particle collisions quantitatively from first principles. Examples of such events recorded by the CMS experiment at the LHC are shown in Fig. 1. In these events kinematic characteristics of particles, such as energy and momentum, are collected. Analyzing many such events, we can produce distributions of kinematic variables, for instance, differential distribution of the inclusive jet cross section with respect to pseudorapidity,  $d\sigma/d\eta$ . There is a long way from the QCD Lagrangian to making predictions for such distributions, full of difficulties. I clearly do not expect students in high energy experimental physics to be able to solve those difficulties. Instead I would like to explain the consequences of solving the difficulties, because an incomplete understanding of these consequences can easily lead to false interpretation of correct measurements.

#### 1.1 The QCD Lagrangian

The quantum field theory (QFT) of the strong interactions is a part of the Standard Model (SM) of elementary particle interactions. The SM is based upon the principle of local gauge invariance. The underlying gauge group is

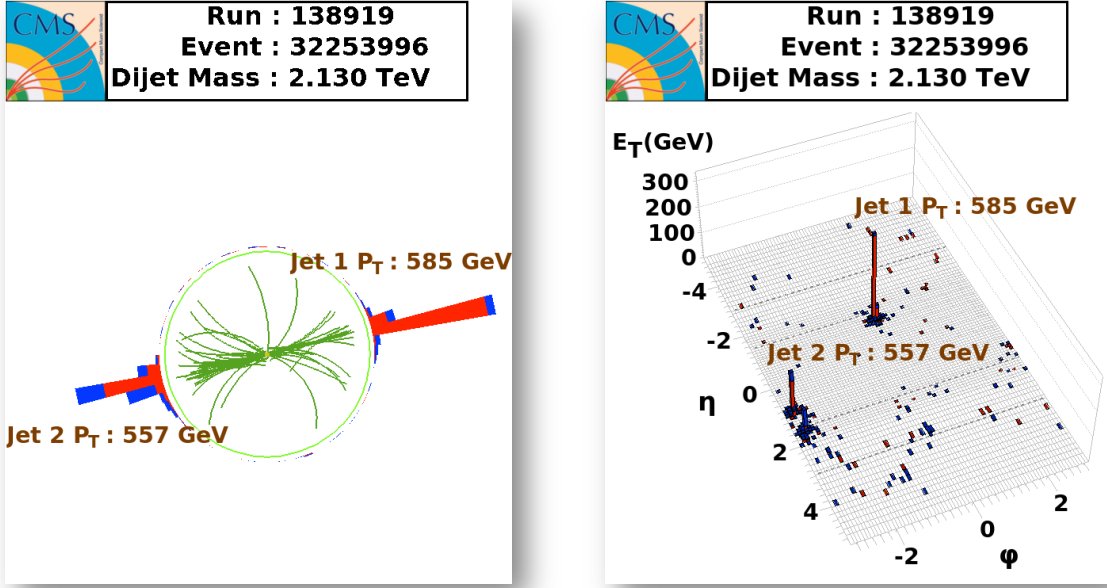
$$SU(3)_c \times SU(2)_L \times U(1)_Y,$$

where  $c$  stands for “colour”,  $L$  for “left” (or “weak isospin”) and  $Y$  for “hypercharge”. As we concentrate on QCD, which is based on  $SU(3)_c$  gauge symmetry, we can write the Lagrangian as

$$\mathcal{L}_{\text{QCD}} = \mathcal{L}_{\text{QCD}}^{(0)} + \mathcal{L}_{\text{sources}}, \quad (1)$$

---

\*Z.Trocsanyi@atomki.hu.



**Fig. 1:** A violent proton-proton collision resulting in two hard jets. Left: tracks and energy deposits in the calorimeters, right: energy deposits represented by towers in the pseudorapidity–azimuthal angle plane.

where

$$\mathcal{L}_{\text{QCD}}^{(0)} = \mathcal{L}_C + \mathcal{L}_{\text{GF}} + \mathcal{L}_G, \quad (2)$$

while the electroweak sector (based on the  $SU(2)_L \times U(1)_Y$  symmetry) act as sources. In Eq. (2)  $\mathcal{L}_C$  is the classical Lagrangian, while  $\mathcal{L}_{\text{GF}}$  is the gauge-fixing term. The last piece is the ghost Lagrangian, absent if we use physical gauges, which will be our choice.

To find the classical Lagrangian, one starts with the Lagrangian of free Dirac fields,

$$\mathcal{L}_q^{(0)}(q_f, m_f) = \sum_{k,l=1}^{N_c} \bar{q}_f^k (i\gamma_\mu \partial^\mu - m_f)_{kl} q_f^l, \quad (3)$$

where the  $\gamma_\mu$  matrices satisfy the Clifford algebra,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \quad \{\gamma^\mu, \gamma^5\} = 0. \quad (4)$$

The matter field content is dictated by the electroweak sector. The fermion fields are called *quark fields*:  $q_f^k$  with masses  $m_f$  and  $f = 1, \dots, n_f$ , where  $n_f$  is the number of different flavours. The quark fields also have an additional degree of freedom: colour, labelled by  $k$ , that can take  $N_c$  values,  $k = 1, \dots, N_c$ . The precise matter content is shown in Table 1.

If we apply a transformation  $q^k \rightarrow q'^k = U_{kl} q^l$ , with

$$U_{kl} = \exp \left\{ i \sum_{a=1}^{N_c^2-1} t^a \theta^a \right\}_{kl} \equiv \exp \{ i \mathbf{t} \cdot \boldsymbol{\theta} \}_{kl},$$

where  $\theta^a \in \mathbb{R}$ , then the Lagrangian of free Dirac fields remains invariant,  $\mathcal{L}_q^{(0)}(q) = \mathcal{L}_q^{(0)}(q')$ . The  $(t^a)_{kl}$  are  $N_c \times N_c$  matrices that constitute the fundamental representation of the generators  $T^a$  (called

**Table 1:** The six quark flavours as dictated by the electroweak sector. Their baryon number is  $B = 1/3$ . Each quark flavour comes in three colours, not shown. For light flavours (u, d and s) the mass values are not without controversy and still under investigation. For flavours c and b the mass values are running  $\overline{\text{MS}}$  quark masses at 2 MeV (see definition below), while for t it is the pole mass.

$f$	1	2	3	4	5	6
$q_f$	u	d	s	c	b	t
$m_f$	$\approx 3 \text{ MeV}$	$\approx 6 \text{ MeV}$	$\approx 100 \text{ MeV}$	1.2 GeV	4.2 GeV	172.6 GeV

colour-charge operators), which satisfy the Lie algebra:

$$\left[ T^a, T^b \right] = i f^{abc} T^c, \quad \text{with normalization } \text{Tr}(T^a T^b) = T_R \delta^{ab}. \quad (5)$$

For  $SU(3)$  the matrices  $t^a$  are the Gell-Mann matrices (see e.g., [2]).

Next we ask the question if we can make  $\mathcal{L}_q^{(0)}(q)$  invariant under local  $SU(N_c)$  transformations. The answer is yes, we can through the following steps:

1. Introduce  $A_\mu^a$  coloured vector field with the following transformation property under  $SU(N_c)$  transformations:

$$\mathbf{t} \cdot \mathbf{A}_\mu \longrightarrow \mathbf{t} \cdot \mathbf{A}'_\mu = U(x) \mathbf{t} \cdot \mathbf{A}_\mu U^{-1}(x) + \frac{i}{g_s} (\partial_\mu U(x)) U^{-1}(x),$$

where  $U(x) = \exp\{i\mathbf{t} \cdot \boldsymbol{\theta}(x)\}$ .

2. Replace  $\partial_\mu \delta_{kl}$  with  $D_\mu [A]_{kl} = \partial_\mu \delta_{kl} + ig_s (\mathbf{t} \cdot \mathbf{A}_\mu)_{kl}$ . This covariant derivative  $D_\mu [A]_{kl} q_l(x)$  transforms the same way as the quark field  $q_k(x)$ .
3. Introduce a kinetic term

$$\mathcal{L}_g(A) = -\frac{1}{4} F_{\mu\nu}^a [A] F^{a\mu\nu} [A],$$

with the non-abelian field strength  $F_{\mu\nu}^a$  given by

$$F_{\mu\nu}^a [A] = \partial_{[\mu} A_{\nu]}^a - g_s \underbrace{f^{abc} A_\mu^b A_\nu^c}_{\text{“}\mathbf{A}_\mu \times \mathbf{A}_\nu\text{”}}$$

so the Lagrangian contains cubic and quartic terms of the gauge field.

The constants  $f^{abc}$  are the structure constants of the Lie algebra. The structure constants are completely antisymmetric and are related to the adjoint representation of the generators  $F_{bc}^a$  by  $F_{bc}^a = -if^{abc}$ .

Thus we find that the gauge boson field, called *gluon field*, is a consequence of the local  $SU(N_c)$  (gauge) invariance. The classical Lagrangian of QCD is a sum of interacting Dirac Lagrangians for spin 1/2 fermion fields and a Lagrangian of a gauge field,

$$\mathcal{L}_C = \mathcal{L}_f + \mathcal{L}_g = \sum_{f=1}^{n_f} \mathcal{L}_q(q_f, m_f) + \mathcal{L}_g(A), \quad (6)$$

where

$$\mathcal{L}_q(q_f, m_f) = \sum_{k,l=1}^{N_c} \bar{q}_f^k (i\gamma_\mu D^\mu [A] - m_f)_{kl} q_f^l. \quad (7)$$

The gluon field is also coloured and self-interacting. In fact, these self-interactions are the sources of the main difference between QED and QCD. We shall see that as a result, QCD is a ‘perfect theory’ in the sense that it is asymptotically free. Furthermore, *among quantum field theories in  $d = 4$  dimensions only non-Abelian gauge theories are asymptotically free* (see discussion after Eq. (21)). It is also plausible that the self-interactions are the sources of *colour confinement*, i.e., the colour neutrality of hadrons, but we do not have a proof based on first principles.

It is clear that there is an unprecedented large number of degrees of freedom we have to sum over when computing a cross section:

1. spin and space-time as in any field theory, not exhibited above,
2. flavour, which also appears in electroweak theory, and colour, which is specific to QCD only.

As a result computations in QCD are rather cumbersome. During the last two decades a lot of effort was invested and great progress was made to find “simple” ways of computing QCD cross sections and to automate the computations.

---

**Exercise 1.1** Show that in QED the covariant derivative transforms the same way as the field itself, i.e., if  $f(x) \rightarrow U(x)f(x)$  then  $D_\mu f(x) \rightarrow U(x)D_\mu f(x)$ , where  $D_\mu = \partial_\mu + ie A_\mu$ .

**Exercise 1.2** Show that in QED

$$[D_\mu, D_\nu] = ie F_{\mu\nu},$$

where  $D_\mu = \partial_\mu + ie A_\mu$ .

**Exercise 1.3** Show that the generators of a special unitary group are traceless and hermitian.

**Exercise 1.4** The generators in the fundamental representation of  $SU(2)$  are the Pauli matrices divided by two:

$$t_f^a = \frac{\sigma^a}{2}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The adjoint representation of a group is defined as

$$(t_A^b)_{ac} = if^{abc}.$$

Compute the generators in the adjoint representation of  $SU(2)$ .

We define the constant  $T_R$  for a representation  $R$  by the condition

$$\text{Tr}[t_R^a t_R^b] = \delta^{ab} T_R.$$

Compute this constant from the explicit form of the fundamental ( $T_F$ ) and the adjoint ( $T_A$ ) representation.

The quadratic Casimir  $C_2(R)$  of a representation  $R$  is defined by

$$C_2(R)\mathbb{1} = \sum_a t_R^a t_R^a.$$

Compute the quadratic Casimir in the fundamental ( $C_F$ ) and the adjoint ( $C_A$ ) representation of  $SU(2)$  using the explicit form of the representation matrices.

**Exercise 1.5** Show that in  $SU(N)$  gauge theories

$$[D_\mu, D_\nu] = ig F_{\mu\nu}^a T^a \quad \text{with} \quad F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - gf^{abc} A_\mu^b A_\nu^c.$$

**Exercise 1.6** Show that  $F_{\mu\nu}^a$  transforms according to the adjoint representation of  $SU(N)$ :

$$F_{\mu\nu}^a \rightarrow F_{\mu\nu}^a - f^{abc} \theta^b F_{\mu\nu}^c.$$


---

## 1.2 Feynman rules

The Feynman rules can be derived from the action,

$$S = i \int d^4x (\mathcal{L}_f + \mathcal{L}_g) \equiv S_0 + S_I, \quad \text{where } S_0 = i \int d^4x \mathcal{L}_0, \quad \text{and } S_I = i \int d^4x \mathcal{L}_I.$$

In this decomposition  $\mathcal{L}_0$  contains the terms bilinear in the fields and  $\mathcal{L}_I$  does all other terms, called interactions. The gluon propagator  $\Delta_{g,\mu\nu}(p)$  is the inverse of the bilinear term in  $A_\mu$ . In momentum space we have the condition (we suppress colour indices as these terms are diagonal in colour space)

$$\Delta_{g,\mu\nu}(p) i [p^2 g^{\nu\rho} - p^\nu p^\rho] = \delta_\mu^\rho. \quad (8)$$

However,

$$[p^2 g^{\nu\rho} - p^\nu p^\rho] p_\rho = 0, \quad (9)$$

which means that the inverse does not exist, the matrix  $[p^2 g^{\nu\rho} - p^\nu p^\rho]$  is not invertible. We can exploit gauge invariance to rewrite the classical Lagrangian in a physically equivalent form (action remains the same) such that  $\Delta_{g,\mu\nu}$  exists, which is called gauge fixing. This amounts to imposing a constraint on  $A_\mu$  by adding a term to the Lagrangian with a Lagrange multiplier (like in classical mechanics). For example, the *covariant gauges* are defined by requiring  $\partial_\mu A^\mu(x) = 0$  for any  $x^\mu$ . Adding

$$\mathcal{L}_{\text{GF}} = -\frac{1}{2\lambda} (\partial_\mu A^\mu)^2, \quad \lambda \in \mathbb{R},$$

to  $\mathcal{L}$ , the action  $S$  remains the same. The bilinear term becomes in this case

$$i \left( p^2 g^{\nu\rho} - \left(1 - \frac{1}{\lambda}\right) p^\nu p^\rho \right),$$

with inverse

$$\Delta_{g,\mu\nu}(p) = -\frac{i}{p^2} \left[ g_{\mu\nu} - (1 - \lambda) \frac{p_\mu p_\nu}{p^2} \right].$$

Of course, physical results must be independent of  $\lambda$ . It is customary to choose  $\lambda = 1$  (called covariant Feynman gauge).

In covariant gauges unphysical degrees of freedom (longitudinal and time-like polarizations) also propagate, and these unphysical degrees of freedom are canceled by, so-called, ghost contributions. This can be avoided by choosing *axial (physical) gauges*, defined with an arbitrary, but fixed vector  $n^\mu$ , different from  $p^\mu$ :

$$\mathcal{L}_{\text{GF}} = -\frac{1}{2\lambda} (n^\mu A_\mu)^2,$$

which leads to

$$\Delta_{g,\mu\nu}(p, n) = -\frac{i}{p^2} \left( g_{\mu\nu} - \frac{p_\mu n_\nu + n_\mu p_\nu}{p \cdot n} + \frac{(n^2 + \lambda p^2) p_\mu p_\nu}{(p \cdot n)^2} \right).$$

Since  $p^2 = 0$ , we have:

$$\Delta_{g,\mu\nu}(p, n) p^\mu = 0, \quad \Delta_{g,\mu\nu}(p, n) n^\mu = 0.$$

Thus, only 2 degrees of freedom propagate (transverse ones in the  $n^\mu + p^\mu$  rest-frame). A usual choice is  $n^2 = 0$ ,  $\lambda = 0$ , called *light-cone gauge*. The price we pay by choosing this gauge instead of a covariant one is that the propagator looks more complicated and it diverges when  $p^\mu$  becomes parallel to  $n^\mu$ . In this gauge

$$\Delta_{g,\mu\nu}(p, n) = \frac{i}{p^2} d_{\mu\nu}(p, n)$$

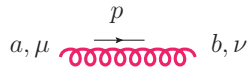
with

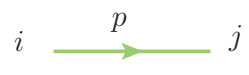
$$d_{\mu\nu}(p, n) = -g_{\mu\nu} + \frac{p_\mu n_\nu + n_\mu p_\nu}{p \cdot n} = \sum_{\lambda=1,2} \epsilon_\mu^{(\lambda)}(p) \epsilon_\nu^{(\lambda)}(p)^*, \quad (10)$$

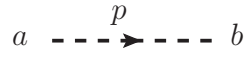
where  $\epsilon_\mu^{(\lambda)}(p)$  is the polarization vector of the gauge field (photon in QED, gluon in QCD).

### 1.3 Feynman rules for QCD

Propagators (Feynman's '+iε'-prescription is assumed, but not shown):

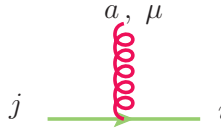
gluon propagator:  $\Delta_{g,\mu\nu}^{ab}(p) = \delta^{ab} \Delta_{g,\mu\nu}(p) \longrightarrow$  

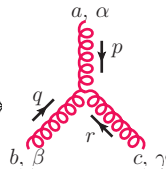
quark propagator:  $\Delta_q^{ij}(p) = \delta^{ij} i \frac{\not{p} + m}{p^2 - m^2} \longrightarrow$  

ghost propagator:  $\Delta^{ab}(p) = \delta^{ab} \frac{i}{p^2} \longrightarrow$  

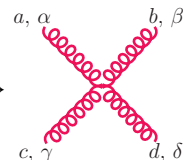
(not needed in physical gauges)

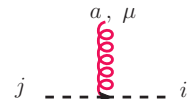
Vertices:

quark-gluon:  $\Gamma_{gq\bar{q}}^{\mu,a} = -ig_S (t^a)_{ij} \gamma^\mu \longrightarrow$  

three-gluon:  $\Gamma_{\alpha\beta\gamma}^{abc}(p, q, r) = -ig_S (F^a)_{bc} V_{\alpha\beta\gamma}(p, q, r) \longrightarrow$  

$$V_{\alpha\beta\gamma}(p, q, r) = (p - q)_\gamma g_{\alpha\beta} + (q - r)_\alpha g_{\beta\gamma} + (r - p)_\beta g_{\alpha\gamma}, \quad p^\alpha + q^\alpha + r^\alpha = 0$$

four-gluon:  $\Gamma_{\alpha\beta\gamma\delta}^{abcd} = -ig_S^2 \begin{bmatrix} +f^{xac} f^{xbd} (g_{\alpha\beta} g_{\gamma\delta} - g_{\alpha\delta} g_{\beta\gamma}) \\ +f^{xad} f^{xbc} (g_{\alpha\beta} g_{\gamma\delta} - g_{\alpha\gamma} g_{\beta\delta}) \\ +f^{xab} f^{xcd} (g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma}) \end{bmatrix} \longrightarrow$  

ghost-gluon:  $\Gamma_{g\eta\bar{\eta}}^{\mu,a} = -ig_S (F^a)_{ij} p^\mu \longrightarrow$   (not needed in physical gauges).

The four-gluon vertex differs from the rest of the Feynman rules in the sense that it is not in a factorized form of a colour and a tensor factor. This is an inconvenient feature because it prevents the separate summation over colour and Lorentz indices and complicates automation. We can however circumvent this problem by introducing a *fake field* with propagator

$$a \begin{array}{c} \gamma \\ \text{---} \\ \alpha \end{array} \text{---} \text{---} \text{---} \text{---} \begin{array}{c} \delta \\ \text{---} \\ \beta \end{array} b = \frac{i}{2} \delta^{ab} (g^{\alpha\beta} g^{\gamma\delta} - g^{\alpha\delta} g^{\beta\gamma}), \text{ that couples only to the gluon with vertex}$$

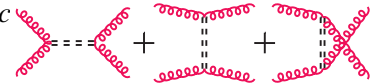
$$\begin{array}{c} a, \alpha \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ c, \gamma \end{array} \begin{array}{c} \xi \\ \text{---} \\ \zeta \end{array} x = i\sqrt{2}g_s f^{xac} g^{\alpha\xi} g^{\gamma\zeta}.$$

We can check that a single four-gluon vertex can be written as a sum of three graphs. This way the summations over colour and Lorentz indices factorize completely, which helps automation and makes possible for us to concentrate on the colour algebra independently of the rest of the Feynman rules.

Finally, we have to supply the following factors for incoming and outgoing particles:

- outgoing fermion:  $\bar{u}(p)$
- incoming fermion:  $u(p)$
- outgoing photon or gluon:  $\epsilon_\mu^{(\lambda)}(p)^*$
- outgoing antifermion:  $v(p)$
- incoming antifermion:  $\bar{v}(p)$
- incoming photon or gluon:  $\epsilon_\mu^{(\lambda)}(p)$ .

**Exercise 1.7** Show that the four-gluon vertex can be written as a sum of three graphs, with the help of the fake field such that in each graph the colour and Lorentz indices are fac



### 1.4 Basics of colour algebra

Examining the Feynman rules, we find that there are two essential changes as compared to QED. One is that there is an additional degree of freedom: colour. The second is that there are new kind of couplings: the self couplings of the gauge field. We now explore the effect of the first.

In order to see how to treat the colour degrees of freedom, we set to one all but the colour part of the Feynman rules and try first to develop an efficient technique to compute the coefficients involving the colour structure. This is possible because the colour degrees of freedom factorize from the other degrees of freedom completely. We use the following graphical representation for the colour charges in the fundamental representation:

$$\begin{array}{c} a \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ j \text{---} \text{---} i \end{array} = (t^a)_{ij}.$$

The normalization of these matrices is given by  $\text{Tr}(t^a t^b) = a \text{---} \text{---} \text{---} \text{---} b = T_R a \text{---} \text{---} b = T_R \delta^{ab}$ . The usual choice is  $T_R = \frac{1}{2}$ , but  $T_R = 1$  is also used often. We shall use both.

In the adjoint representation the colour charge  $T^a$  is represented by the matrix  $(F^a)_{bc}$  that is related to the structure constants by

$$(F^a)_{bc} = (F^b)_{ca} = (F^c)_{ab} = -i f_{abc} = \begin{array}{c} a \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ b \text{---} \text{---} c \end{array}$$

where  $F^a$  with  $a = 1, \dots, (N_c^2 - 1)$  are  $(N_c^2 - 1) \times (N_c^2 - 1)$  matrices which again satisfy the commutation relation (5). The graphical notation in the adjoint representation is not unique. For the matrix

$(F^a)_{bc}$  we assume an arrow pointing from index  $c$  to  $b$ , opposite to which we read the indices of  $(F^a)_{bc}$  (similarly as for the matrices  $t^a$ ). On the structure constants the indices are not distinguished, therefore arrows do not appear. However, these are completely antisymmetric in their indices, therefore, the ordering matters. By convention, in the graphical representation, the ordering of the indices is counter-clockwise. The representation matrices are invariant under  $SU(N)$  transformations.

The sums  $\sum_a t_{ij}^a t_{jk}^a$  and  $\text{Tr}(F^a F^b)$  have two free indices in the fundamental and adjoint representation, respectively. These are invariant under  $SU(N)$  transformations, therefore, must be proportional to the unit matrix,

$$\sum_{j,a} t_{ij}^a t_{jk}^a = C_F \delta_{ik}, \quad \text{Tr}(F^a F^b) = C_A \delta^{ab},$$

which is depicted graphically as

Here  $C_F$  and  $C_A$  are the eigenvalues of the quadratic Casimir operator in the fundamental and adjoint representation, respectively. In the familiar case of angular momentum operator algebra ( $SU(2)$ ), the quadratic Casimir operator is the square of the angular momentum with eigenvalues  $j(j + 1)$ . The fundamental representation is two dimensional, realized by the (half of the) Pauli matrices acting on two-component spinors, when  $j = 1/2$  and  $C_F = 1/2(1/2 + 1) = 3/4$ . In the adjoint representation  $j = 1$  and  $C_A = 2$ . Below we derive the corresponding values for general  $SU(N)$ .

The commutation relation (5) can be represented graphically by

Multiplying this commutator first with another colour charge operator with summing over the fermion index and then taking the trace over the fermion line (i.e., multiplying with  $\delta_{ik}$ ) we obtain the resolution of the three-gluon vertex as traces of products of colour charges:

$$= \text{Tr}(t^a t^b t^c) - \text{Tr}(t^c t^b t^a) = i T_R f^{abc}.$$

We now show some examples of how one can compute the colour algebra structure of a QCD amplitude, in particular we will also find an explicit value for  $C_F$  and  $C_A$ . Taking the trace of the identity in the fundamental and in the adjoint representation we obtain

$$= N_c, \quad = N_c^2 - 1,$$

respectively. Then, using the expressions for the fermion and gluon propagator corrections, we immediately find



$$\text{Diagram 1} = C_F N_c, \quad \text{Diagram 2} = C_A (N_c^2 - 1).$$

The generators are traceless,

$$\text{Diagram 3} = \text{Tr}(t^a) = 0, \quad \text{Diagram 4} = \text{Tr}(F^a) = 0.$$

We can now find the value of  $C_F$  as follows. On the one hand we know that

$$T_R \text{Diagram 2} = \text{Diagram 5} = \text{Diagram 6} = C_F N_c$$

while on the other, the left hand side is also equal to  $T_R (N_c^2 - 1)$ . Thus

$$C_F = T_R \frac{N_c^2 - 1}{N_c}.$$

Analogously one can find

$$C_A = 2 T_R N_c.$$

As the colour factors  $C_F$  and  $C_A$  depend on  $N_c$ , their measurement gives information on the number of colours. The experiments of the Large Electron Positron collider measured the values of the colour factors based on fits of theoretical predictions [3] to four-jet angular distributions that are sensitive to both  $C_F$  and  $C_A$ . The result of the simultaneous measurement of the colour factors and the strong coupling by the OPAL collaboration is shown in Fig. 2 [4]. The values corresponding to  $N_c = 3$  are marked with the star, just in the middle of the confidence-ellipses.

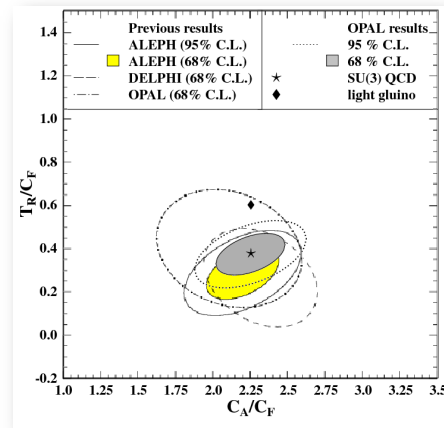
The expression  $\sum_a t_{ij}^a t_{kl}^a$  is invariant under  $SU(N)$  transformations, therefore has to be expressible as a linear combination of  $\delta_{il}\delta_{kj}$  and  $\delta_{ij}\delta_{kl}$  (the third combination of Kronecker  $\delta$ 's is not possible, the direction of arrows do not match). The two coefficients can be obtained by making contractions with  $\delta_{il}\delta_{jk}$  and  $\delta_{ij}\delta_{kl}$ . Thus we obtain the Fierz identity,

$$\sum_a t_{ij}^a t_{kl}^a = T_R \left( \delta_{il}\delta_{kj} - \frac{1}{N_c} \delta_{ij}\delta_{kl} \right),$$

or graphically:

$$\text{Diagram 7} = T_R \left( \text{Diagram 8} - \frac{1}{N_c} \text{Diagram 9} \right).$$

These graphical rules make the evaluation of colour algebra easy. Nevertheless, nowadays computer algebra codes make computation of colour sums an automated procedure. For instance, you may try `In[1]:= Import["http://www.feyncalc.org/install.m"]` in your *Mathematica* session to see one solution.



**Fig. 2:** Measurement of the colour factors by the LEP collaborations [4]

**Exercise 1.8** Consider the process  $q\bar{q} \rightarrow ggg$ . Compute the color structures that appear in the squared matrix element.

**Exercise 1.9** Try using the Fierz identity to obtain

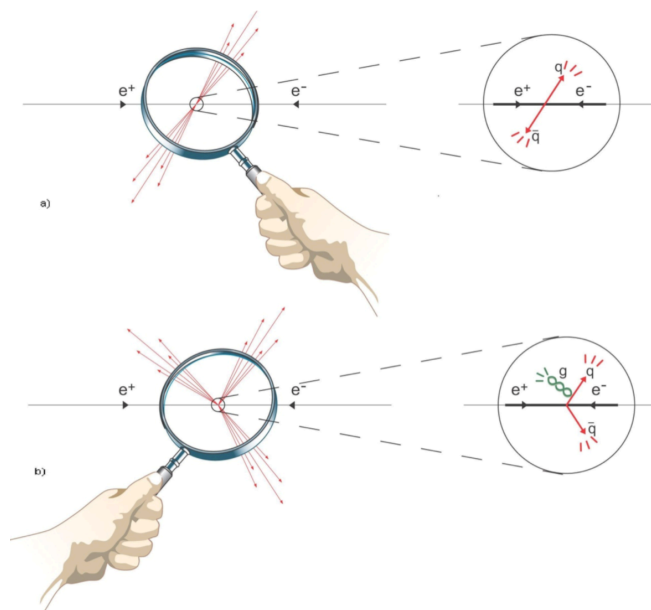
$$\text{Diagram} = -T_R C_F$$

**Exercise 1.10** Determine the color factors  $A, B, C$  in the following equations:

$$\text{Diagram 1} = A \text{Diagram 2}, \quad \text{Diagram 3} = B \text{Diagram 4}, \quad \text{Diagram 5} = C \text{Diagram 6}$$

### 1.5 Are we done?

We now have the Feynman rules with colour and the rest factorized, and we gained some insight how to perform the colour algebra. Thus it seems that we are in the position to compute the cross section of any process up to the desired accuracy in perturbation theory (PT), just as we can do in QED. So it may appear that conceptually we are done. Well, we are going to see big surprises!



**Fig. 3:** Illustration of the approximation of hadronic final states by partonic events in electron-positron annihilation: the sprays of hadrons (called jets) are assumed to originate from primary quarks and gluons, thus approximated by quarks and gluons as shown by the magnification

*The result of a low-order perturbative computation in QCD is an approximation to sufficiently inclusive hadronic cross section if (i) the total centre-of-mass energy  $Q$  of partons is much larger than the mass of quarks,  $Q \gg m_q$ , and (ii)  $Q$  is far from hadronic resonances and thresholds.*

The first conceptual challenge is due to a phenomenological observation. In QED, PT is applicable because the elementary excitations of the quantum fields, the electrons and photons, can be observed as stable, free particles. Thus asymptotic states are parts of the physical reality. On the contrary, free quarks and gluons (usually called simply partons) have never been observed in nature. This experimental fact can be reformulated saying that the probability of observing a final state with any *fixed number of on-shell partons* is zero. This negative result has been turned into the principle of ‘quark confinement’. Thus it is questionable whether a QFT of quark and gluon fields can describe the observed world of particles where in addition to leptons only hadrons have been found. In fact, a main research project at the LEP was to find an answer to this question in a well controlled quantitative manner. It turned out that the answer is positive if we make an assumption that we cannot prove from first principles:

We shall define precisely what ‘sufficiently inclusive’ means later. Predictions made on the basis of this assumption agree with measurements (e.g. made at LEP) within the expected accuracy of the prediction, which we are to define also later.

Based on this assumption, it makes sense to make predictions with quark and gluon asymptotic states. However, in QCD the complexity of the Feynman rules will make higher order computations prohibitive. Indeed, the largest effort in QCD computations during the past 20 years went into devising ever more efficient methods to decrease the algebraic complexity of the computations. This research is driven by the observation that the QCD Lagrangian is highly symmetric, which has to be reflected in the final results. Thus the complications somehow appear mainly because with our rules we artificially introduce complications at intermediate steps of the computations, which cancels to large extent in the final formulae. Learning about the symmetries of QCD is interesting and useful not only for technical purposes, so let us make an inventory of those.

### 1.6 Symmetries of the classical Lagrangian

The symmetries can be grouped into two large categories: exact symmetries and approximate ones. Space-time symmetries are exact. These consist of invariance against continuous transformations: translations and Lorentz-transformations (rotations and boosts). In addition  $\mathcal{L}_{\text{cl}}$  is invariant under scale transformation:

$$x^\mu \rightarrow \lambda x^\mu \quad A_\mu(x) \rightarrow \lambda^{-1} A_\mu(\lambda x) \quad q(x) \rightarrow \lambda^{-3/2} q(\lambda x),$$

and conformal transformations, which we do not detail here. The Lagrangian is also invariant under charge conjugation (C), parity (P) and time-reversal (T), in agreement with observed properties of strong interactions (C, P and T violating strong decays are not observed).

We already discussed exact symmetry in colour space: local gauge invariance. In addition to the classical Lagrangian of Eq. (6), there exists additional gauge invariant dimension-four operator, the  $\Theta$ -term:

$$\mathcal{L}_\Theta = \frac{\Theta g_s}{32\pi^2} \sum_a F_{\mu\nu}^a \tilde{F}^{a,\mu\nu}, \quad \text{with} \quad \tilde{F}^{a,\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}^a,$$

that violates P and T. As experimentally  $\Theta < 10^{-9}$ , we set  $\Theta = 0$  in perturbative QCD.

Another interesting feature of  $\mathcal{L}_{\text{cl}}$  is that it is almost supersymmetric. For one massless flavour

$$\mathcal{L}_{\text{cl}} = -\frac{1}{4} \sum_a F_{\mu\nu}^a F^{a,\mu\nu} + \bar{q} i \not{D} q,$$

which is very similar to the Lagrangian of  $N = 1$  supersymmetric gauge theory,

$$\mathcal{L}_{\text{cl}}^{\text{SUSY}} = -\frac{1}{4} \sum_a F_{\mu\nu}^a F^{a,\mu\nu} + \bar{\lambda} i \not{D} \lambda.$$

The only difference is that the quark  $q$  transforms under the fundamental, while the gluino  $\lambda$  under the adjoint representation of the gauge group.

An important approximate symmetry of the classical Lagrangian is related to the quark mass-matrix. Let us introduce the quark flavour triplet

$$\psi = \begin{pmatrix} u \\ d \\ s \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix},$$

with each component being a four-component Dirac spinor, and the combinations

$$P_\pm = \frac{1}{2} (1 \pm \gamma_5). \tag{11}$$

The latter are projections:

$$P_+ P_- = P_- P_+ = 0, \quad P_\pm^2 = P_\pm, \quad P_+ + P_- = 1.$$

It follows from Clifford-algebra that  $\gamma_\mu P_\pm = P_\mp \gamma_\mu$ . We define  $\psi_\pm = P_\pm \psi$ . Using  $\gamma_5^2 = 1$ , we find that  $\psi_\pm$  are eigenvectors of  $\gamma_5$  with  $\pm 1$  eigenvalues:

$$\gamma_5 \psi_\pm = \pm \psi_\pm .$$

From the definition of the Dirac adjoint,  $\bar{\psi} = \psi^\dagger \gamma_0$ , we obtain  $\bar{\psi}_\pm = \bar{\psi} P_\mp$ . Thus the quark sector of the Lagrangian can be rewritten in terms of the chiral fields  $\psi_\pm$ :

$$\begin{aligned} \mathcal{L}_{\text{cl}} &= \bar{\psi} i \gamma_\mu D^\mu \psi = \bar{\psi} (P_+ + P_-) i \gamma_\mu D^\mu (P_+ + P_-) \psi = \bar{\psi} P_+ i \gamma_\mu D^\mu P_- \psi + \bar{\psi} P_- i \gamma_\mu D^\mu P_+ \psi = \\ &= \bar{\psi}_- i \gamma_\mu D^\mu \psi_- + \bar{\psi}_+ i \gamma_\mu D^\mu \psi_+ = \mathcal{L}_- + \mathcal{L}_+ \equiv \mathcal{L}_L + \mathcal{L}_R . \end{aligned}$$

This decomposition would not work if the gluon field in the covariant derivative were not Lorentz-vector. In this chiral form the left- and right-handed fields decouple, so the Lagrangian is invariant under separate  $U(N_f)$  transformations for the left- and right-handed fields, i.e., under  $U_L(N_f) \times U_R(N_f)$ , hence it is called chiral symmetry. Indeed, if  $(g_L, g_R) \in U_L(N_f) \times U_R(N_f)$ , then under the transformation

$$\psi_L \rightarrow g_L \psi_L , \quad \bar{\psi}_L \rightarrow \bar{\psi}_L g_L^\dagger , \quad g_R = 1$$

$\mathcal{L}_L$  remains invariant. This symmetry is exact if the quarks are massless. The group elements can be parametrized using  $2N_f^2$  real numbers  $\{\alpha, \alpha_a, \beta, \beta_b\}$  ( $a, b = 1, \dots, N_f^2 - 1$ ),

$$\begin{aligned} (g_L, g_R) &= \left( \exp(i\alpha) \exp(i\beta) \exp\left(i \sum_a \alpha_a T^a\right) \exp\left(i \sum_b \beta_b T^b\right), \right. \\ &\quad \left. \exp(i\alpha) \exp(-i\beta) \exp\left(i \sum_a \alpha_a T^a\right) \exp\left(-i \sum_b \beta_b T^b\right) \right) \\ &\in U_V(1) \otimes SU_L(N_f) \otimes U_A(1) \otimes SU_R(N_f) , \end{aligned}$$

where the matrices  $T^a$  represent the generators of the group ( $N_f \times N_f$  matrices). The transformations  $(\exp(i \sum_a \alpha_a T^a), \exp(i \sum_a \alpha_a T^a))$ , acting as  $\psi \rightarrow \exp(i \sum_a \alpha_a T^a \mathbf{I}) \psi$ , form a vector subgroup  $SU_V(N_f)$ . The transformations  $(\exp(i \sum_b \beta_b T^b), \exp(-i \sum_b \beta_b T^b))$ , acting as  $\psi \rightarrow \exp(i \sum_b \beta_b T^b \gamma_5) \psi$  however, do not form an axial-vector subgroup because

$$[T^a \gamma_5, T^b \gamma_5] = i \sum_c f^{abc} T^c \mathbf{I} \quad (\gamma_5^2 = \mathbf{I} \neq \gamma_5) .$$

This chiral symmetry is not observed in the hadron spectrum. Therefore, we assume that vacuum has a non-zero VEV of the light-quark operator,

$$\langle 0 | \bar{q} q | 0 \rangle = \langle 0 | \bar{u} u + \bar{d} d | 0 \rangle \simeq (250 \text{ MeV})^3 ,$$

a chiral condensate that connect left- and right-handed fields,

$$\langle 0 | \bar{q} q | 0 \rangle = \langle 0 | \bar{q}_L q_R + \bar{q}_R q_L | 0 \rangle .$$

The condensate breaks chiral symmetry spontaneously to  $SU_V(N_f) \otimes U_V(1)$ . This remaining symmetry explains the existence of good quantum numbers of isospin and baryon number, as well as the appearance of  $N_f^2 - 1 = 8$  massless mesons, the Goldstone bosons. As non-zero quark masses violate the chiral symmetry, which is broken spontaneously, the Goldstone bosons are not exactly massless. Thus we have natural candidates for the Goldstone bosons: we can identify those with the pseudoscalar meson octet. In practice, we assume exact chiral symmetry and treat the quark masses as perturbation. This procedure leads us to chiral perturbation theory ( $\chi$ PT) [5], which is capable to predict the (ratios of)

masses of light quarks [6, 7], scattering properties of pions [8] and many more. Although,  $\chi$ PT is a non-renormalizable QFT, it can be made predictive order by order in PT if the measured values of sufficiently many observables are used to fix the couplings of interaction terms at the given order.

The QCD Lagrangian was written forty years ago. Since then many attempts were tried to solve it and mature fields emerged that aim at solving the theory in a limited range of physical phenomena. For instance,  $\chi$ PT is a PT that uses low-energy information (in the MeV range) to explain the world of hadrons and masses of light quarks. In the same energy range non-perturbative approaches, notably lattice QCD and sum rules have been developed for the same purpose. By now it is possible to explain the light hadron spectrum from first principles using lattice results [9]. The main goal at colliders, our focus in these lectures, is different. We shall prove that PT can give reliable predictions for scattering processes at high energies, which is the topic of jet physics.

We have seen that the classical QCD Lagrangian shows many interesting symmetry properties that can be utilized for (i) easing computations, (ii) checking results, (iii) hinting on solving QCD. We shall see that some of these symmetries are violated by quantum corrections, which leads to important physical consequences. In QCD an important example is scaling violations. Another example is the axial anomaly which provides strong constraints on possible QFT's, but it is discussed within the electroweak theory usually.

### 1.7 What is scaling?

Let us consider a dimensionless physical observable  $R$  that depends on a large energy scale  $R = R(Q^2)$ . Large means that  $Q$  is much bigger than any other dimensionful parameter, for instance, masses of quarks. Thus we assume that these other dimensionful parameters can be set zero.<sup>1</sup> Classically,  $\dim R = 0$  and, since  $Q$  is dimensionful, it follows that  $\frac{dR}{dQ} = 0$ . So  $\lim_{Q^2 \rightarrow \infty} R = \text{constant}$ , which is called *scaling*.

In these lectures we do not have room for a complete description of ultraviolet (UV) renormalization of QCD. We simply state that in a renormalized QFT  $R$  depends also on another scale, the renormalization scale  $\mu_R$ . So

$$R = \lim_{Q^2 \rightarrow \infty} R \left( \frac{Q^2}{\mu_R^2}, \alpha_s(\mu_R^2) \right) \neq \text{constant},$$

$R$  need not be a constant. This is called *scaling violation*. The first term in parenthesis is the only dimensionless combination of  $Q$  and  $\mu_R$ . However,  $\mu_R$  is arbitrary. If  $R$  depended on  $\mu_R$ , then its value could not be predicted. For simplicity from now on we drop the subscript "R" from  $\mu_R$ . As  $\mu$  is an arbitrary, un-physical parameter (the classical Lagrangian did not contain  $\mu$ ), we expect that measurable (physical) quantities cannot depend on it, which is expressed by the renormalization group equation (RGE):

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

We can simplify this equation a bit by introducing the new variable  $t$  and the function  $\beta(\alpha_s)$ ,

$$t = \ln \frac{Q^2}{\mu^2}, \quad \beta(\alpha_s) = \mu^2 \frac{\partial \alpha_s}{\partial \mu^2} \Big|_{\alpha_s \text{ fixed}}. \quad (12)$$

Then the RGE becomes

$$\left( -\frac{\partial}{\partial t} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} \right) R(e^t, \alpha_s) = 0. \quad (13)$$

---

<sup>1</sup>We shall study the validity of this assumption in the next subsection.

To present the solution of this partial differential equation, we introduce the *running coupling*  $\alpha_s(Q^2)$ , defined implicitly by

$$t = \int_{\alpha_s}^{\alpha_s(Q^2)} \frac{dx}{\beta(x)}, \quad \text{with } \alpha_s \equiv \alpha_s(\mu^2), \quad (14)$$

where  $\alpha_s \equiv \alpha_s(\mu^2)$  is an arbitrarily fixed number. The derivative of Eq. (14) with respect to the variable  $t$  gives

$$1 = \frac{1}{\beta(\alpha_s(Q^2))} \frac{\partial \alpha_s(Q^2)}{\partial t}, \quad \text{which implies } \beta(\alpha_s(Q^2)) = \frac{\partial \alpha_s(Q^2)}{\partial t}.$$

The derivative of Eq. (14) with respect to  $\alpha_s$  gives

$$0 = \frac{1}{\beta(\alpha_s(Q^2))} \frac{\partial \alpha_s(Q^2)}{\partial \alpha_s} - \frac{1}{\beta(\alpha_s)} \frac{\partial \alpha_s}{\partial \alpha_s},$$

from which it follows that

$$\frac{\partial \alpha_s(Q^2)}{\partial \alpha_s} = \frac{\beta(\alpha_s(Q^2))}{\beta(\alpha_s)}.$$

It is now easy to prove that the value of  $R$  for  $\mu^2 = Q^2$ ,  $R(1, \alpha_s(Q^2))$  solves Eq. (13):

$$-\frac{\partial}{\partial t} R(1, \alpha_s(Q^2)) = -\frac{\partial R}{\partial \alpha_s(Q^2)} \frac{\partial \alpha_s(Q^2)}{\partial t} = -\beta(\alpha_s(Q^2)) \frac{\partial R}{\partial \alpha_s(Q^2)}$$

and

$$\beta(\alpha_s) \frac{\partial}{\partial \alpha_s} R(1, \alpha_s(Q^2)) = \beta(\alpha_s) \frac{\partial \alpha_s(Q^2)}{\partial \alpha_s} \frac{\partial R}{\partial \alpha_s(Q^2)} = \beta(\alpha_s(Q^2)) \frac{\partial R}{\partial \alpha_s(Q^2)}.$$

It then follows that the scale-dependence in  $R$  enters only through  $\alpha_s(Q^2)$ , and we can predict the scale-dependence of  $R$  by solving Eq. (14), or equivalently,

$$\frac{\partial \alpha_s(Q^2)}{\partial t} = \beta(\alpha_s(Q^2)). \quad (15)$$

So far our analysis was non-perturbative. Assuming that PT is applicable, which we shall discuss at the end of this subsection, we may try to solve Eq. (15) in PT where the  $\beta$ -function has the formal expansion:

$$\beta(\alpha_s) = -\alpha_s \sum_{n=0}^{\infty} \beta_n \left( \frac{\alpha_s}{4\pi} \right)^{n+1}. \quad (16)$$

The first four coefficients are known from cumbersome computations [10]

$$\begin{aligned} \beta_0 &= \frac{11}{3} C_A - \frac{4}{3} T_R n_f > 0, & \beta_1 &= \frac{34}{3} C_A^2 - \frac{20}{3} C_A T_R n_f - 4 C_F T_R n_f, \\ \beta_2 &= \frac{2857}{2} - \frac{5033}{18} n_f + \frac{325}{54} n_f^2, & \beta_3 &= 29243 - 6946.3 n_f + 405.9 n_f^2 + 1.5 n_f^3. \end{aligned} \quad (17)$$

The first two coefficients in the expansion of the  $\beta$  function are independent of the renormalization scheme. The second two coefficients in Eq. (17) are valid in the  $\overline{\text{MS}}$  renormalization scheme.<sup>2</sup>

Another often used convention is

$$\beta(\alpha_s) = -b_0 \alpha_s^2 \left[ 1 + \sum_{n=1}^{\infty} b_n \alpha_s^n \right], \quad (18)$$

<sup>2</sup>As we have not gone through the renormalization procedure, we cannot define precisely what we mean by ‘renormalization scheme’. Various schemes differ by finite renormalization of the parameters and fields in the Lagrangian.

where  $b_0 = \frac{\beta_0}{4\pi}$  and  $b_0 b_1 = \frac{\beta_1}{(4\pi)^2}$ , thus  $b_1 = \frac{\beta_1}{4\pi\beta_0}$ .

If  $\alpha_s(Q^2)$  is small we can truncate the series. The solution at leading-order (LO) accuracy is

$$Q^2 \frac{\partial \alpha_s}{\partial Q^2} = \frac{\partial \alpha_s}{\partial t} = -b_0 \alpha_s^2 \Rightarrow - \left[ \frac{1}{\alpha_s(Q^2)} - \frac{1}{\alpha_s(\mu^2)} \right] = -b_0 t \Rightarrow \alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + b_0 t \alpha_s(\mu^2)},$$

which gives  $\alpha_s(Q^2)$  as a function of  $\alpha_s(\mu^2)$  if both are small;  $\alpha_s(\mu^2)$  is a number to be measured. We observe that:

$$\alpha_s(Q^2) \xrightarrow{Q^2 \rightarrow \infty} \frac{1}{b_0 t} \xrightarrow{Q^2 \rightarrow \infty} 0. \quad (20)$$

This behaviour is called *asymptotic freedom*. The sign of  $b_0$  (positive for QCD) plays a crucial role in establishing whether or not a theory is asymptotically free. If it is, then the use of PT is justified: the higher  $Q^2$ , the smaller the coupling. The coefficient  $b_0$  is easiest to compute in *background field gauge* [11] where only three graphs contribute, the quark and gluon loops:


(21)

and a similar ghost loop. The contribution of the quark loop is negative  $-\frac{4}{3}T_R n_f$ , while that of the gluon+ghost loop is positive  $\frac{11}{3}C_A$ . (We knew the colour factors immediately, only the coefficients have to be computed!) The net result is positive up to  $n_f < 17$  in QCD. In 2004 D.J. Gross, H.D. Politzer and F. Wilczek were awarded the Nobel prize for their discovery of asymptotic freedom in QCD [12, 13].

Clearly, it is the gluon self-interaction that makes QCD perfect in PT. In QED, in the absence of photon self-interaction,  $b_0 < 0$ , hence the coupling increases with energy, but remains perturbative up to the Planck scale ( $\simeq 10^{19}$  GeV) where we expect that any known physics breaks down.

Asymptotic freedom gives rationale to perturbative QCD, but we shall see that LO accuracy is not enough. The analysis is also simple at next-to-leading order (NLO):

$$[\alpha_s^2(1 + b_1 \alpha_s)]^{-1} \frac{\partial \alpha_s}{\partial t} = -b_0.$$

$\alpha_s(Q^2)$  is then given implicitly by the equation

$$\frac{1}{\alpha_s(Q^2)} - \frac{1}{\alpha_s(\mu^2)} + b_1 \ln \frac{\alpha_s(Q^2)}{\alpha_s(\mu^2)} - b_1 \ln \frac{1 + b_1 \alpha_s(Q^2)}{1 + b_1 \alpha_s(\mu^2)} = b_0 t,$$

which can be solved numerically.

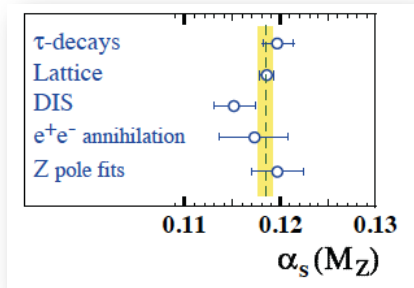
Using the formula for the sum of the geometric series,  $(1 + x)^{-1} = \sum_{j=0}^{\infty} (-x)^j$  and recalling Eq. (19), we find that the running coupling sums logarithms,

$$R(1, \alpha_s(Q^2)) = R_0 + R_1 \alpha_s(\mu^2) \sum_{j=0}^{\infty} [-\alpha_s(\mu^2) b_0 t]^j.$$

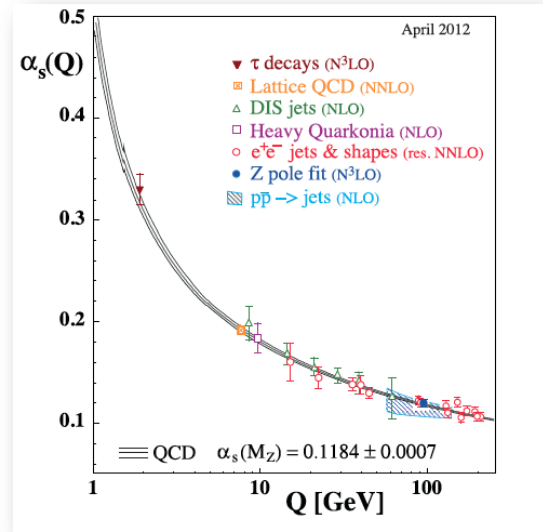
The NLO term  $R_2 \alpha_s^2$  gives logarithms with one power less in each term.

## 1.8 Measuring $\alpha_s(\mu^2)$

We know  $\alpha_s(Q^2)$  if  $\alpha_s(\mu^2)$  is known. We therefore, have to measure  $\alpha_s$  at some scale  $\mu$ . The perturbative solution of the renormalization group equation (RGE, Eq. (13)) is never unique. The difference between two solutions at  $\mathcal{O}(\alpha_s^n)$  is suppressed by  $\alpha_s$ , i.e. at  $\mathcal{O}(\alpha_s^{n+1})$ . Nevertheless, this difference can lead to significant difference in  $\alpha_s(Q^2)$  if  $\mu^2$  and  $Q^2$  are far from each other, which is important in present day precision measurements. Therefore, the scale  $\mu$  is chosen to be  $\mu = M_Z$  because  $M_Z = 91,2 \text{ GeV}$  is not far from the scales where  $\alpha_s(Q^2)$  is used in current experimental analyses. In Figs. 4 and 5 we show the present status of  $\alpha_s$  measurements from Ref. [14].



**Fig. 5:** Results of different measurements of the strong coupling run to  $\mu = M_Z$



**Fig. 4:** Results of measurements of the strong coupling at different scales. The theoretical prediction with four-loop running, fixed at  $\mu = M_Z$  is marked as ‘QCD’.

Another approach to solving the RGE is to introduce a reference scale  $\Lambda$  by

$$\ln \frac{Q^2}{\Lambda^2} = \int_{\alpha_s(Q^2)}^{\infty} \frac{dx}{\beta(x)}.$$

The scale  $\Lambda$  indicates where the coupling becomes strong. The following exercise is to explore the characteristics of this choice.

**Exercise 1.11** *The running of the strong coupling constant is given by Eq. (12). The perturbative expansion of the QCD beta function is given by Eq. (18) with  $b_0, b_1 \geq 0$ . Determine (i) the expression for the coupling constant in leading order ( $b_0 \neq 0, b_1 = 0$ ) and the corresponding scale  $\Lambda_0$  (see below) (ii) the expression for the coupling constant in next-to-leading order ( $b_0 \neq 0, b_1 \neq 0$ ) and the corresponding scale  $\Lambda_1$  (see below).*

*Hints:*

1. Solve the differential equation for  $\alpha(\mu)$ ; you’ll get an integration constant.
2. Express your result in the form

$$\alpha(\mu) = \frac{1}{K \ln\left(\frac{\mu^2}{\Lambda_0^2}\right)}$$



where  $K$  is a constant.

3. Solve the differential equation using  $b_1 \neq 0$

$$\int d\alpha \frac{1}{-b_0\alpha^2 - b_1\alpha^3} = \frac{b_0 + b_1\alpha \log(\alpha) - b_1\alpha \log(b_0 + b_1\alpha)}{b_0^2\alpha} + K$$

4. This time the solution cannot be solved for  $\alpha$  analytically. One can nevertheless find an approximate solution by expanding  $\alpha$  in  $\log \frac{\mu^2}{\Lambda_1^2}$ . The constant  $K$  is not equal to the one in the first part of this exercise.
5. Cast your equation for  $\alpha$  into the form

$$\alpha = \frac{1}{K \ln \frac{\mu^2}{\Lambda_1^2}} \frac{1}{1 + c_1 \frac{\ln(c_2 + b_0\alpha)}{\ln \frac{\mu^2}{\Lambda_1^2}}}$$

with a suitable choice of  $\Lambda_1$ .

6. Expand the right hand side of your equation in  $t = \frac{1}{\ln \frac{\mu^2}{\Lambda_1^2}}$  and keep only the first order term. Use the expansion

$$\frac{1}{1 + C_1 t \ln(C_3 \frac{1}{t} + C_2)} = 1 + t C_1 \ln\left(\frac{1}{t}\right) + O(t).$$

## 1.9 Quark masses and massless QCD

Quark masses  $m_q$  are *parameters* of  $\mathcal{L}_{QCD}$  like the gauge coupling, which need to be renormalized. In QED the electron mass is measured in the laboratories at  $\mu_R^2 = 0$  (classical limit). We cannot similarly isolate a quark at  $\mu_R^2 = 0$  (at low scale quarks are confined). Instead, we can perform a similar RGE analysis as with  $\alpha_s$ . For simplicity we assume one quark flavour with mass  $m$ , which is yet another dimensionful parameter, so the RGE becomes:

$$\left[ \mu^2 \frac{\partial}{\partial \mu^2} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} - \gamma_m(\alpha_s) m \frac{\partial}{\partial m} \right] R\left(\frac{Q^2}{\mu^2}, \alpha_s, \frac{m}{Q}\right) = 0, \quad (22)$$

where  $\gamma_m$  is called the mass anomalous dimension and the minus sign before  $\gamma_m$  is a convention. In PT we can write the mass anomalous dimension as

$$\gamma_m(\alpha_s) = c_0 \alpha_s (1 + c_1 \alpha_s + \mathcal{O}(\alpha_s^2)),$$

with known coefficient up to  $c_3$ . At NLO accuracy we need only  $c_0 = \frac{1}{\pi}$  and  $c_1 = \frac{303-10n_f}{72\pi}$ . As  $R$  is dimensionless, the dependence on the dimensionful parameters has to cancel

$$\left( Q^2 \frac{\partial}{\partial Q^2} + \mu^2 \frac{\partial}{\partial \mu^2} + m^2 \frac{\partial}{\partial m^2} \right) R\left(\frac{Q^2}{\mu^2}, \alpha_s, \frac{m}{Q}\right) = 0. \quad (23)$$

The difference of Eqs. (22) and (23) gives the dependence of  $R$  on  $Q$ :

$$\left[ Q^2 \frac{\partial}{\partial Q^2} - \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} + \left(\frac{1}{2} + \gamma(\alpha_s)\right) m \frac{\partial}{\partial m} \right] R\left(\frac{Q^2}{\mu^2}, \alpha_s, \frac{m}{Q}\right) = 0. \quad (24)$$

This equation is solved by introducing the running mass (in addition to the running coupling)  $m(Q^2)$  obeying

$$Q^2 \frac{\partial m}{\partial Q^2} = -\gamma_m(\alpha_s) m(Q^2), \quad \Rightarrow \quad \ln \frac{m(Q^2)}{m(\mu^2)} = - \int_{\mu^2}^{Q^2} \frac{dq^2}{q^2} \gamma_m(\alpha_s(q^2)). \quad (25)$$

Exponentiating, changing integration variable from  $q^2$  to  $\alpha_s$  and using the definition of the  $\beta$  function, we obtain

$$m(Q^2) = m(\mu^2) \exp \left[ - \int_{\alpha_s(\mu^2)}^{\alpha_s(Q^2)} d\alpha_s \frac{\gamma_m(\alpha_s)}{\beta(\alpha_s)} \right] \xrightarrow{Q^2 \rightarrow \infty} 0, \quad (26)$$

which means that asymptotically free QCD is a massless theory at asymptotically large energies. At LO in PT theory the solution of (26) is given by

$$-\frac{\gamma_m(\alpha_s)}{\beta(\alpha_s)} = \frac{c_0}{b_0 \alpha_s} \quad \Rightarrow \quad m(Q^2) = \bar{m} [\alpha_s(Q^2)]^{\frac{c_0}{b_0}},$$

where we introduced the abbreviation  $\bar{m} = m(\mu^2) [\alpha_s(\mu^2)]^{-\frac{c_0}{b_0}}$ . At NLO the solution becomes

$$m(Q^2) = \bar{m} [\alpha_s(Q^2)]^{\frac{c_0}{b_0}} \left( 1 + \frac{c_0}{b_0} (c_1 - b_1) (\alpha_s(Q^2) - \alpha_s(\mu^2)) + \mathcal{O}(\alpha_s^2) \right).$$

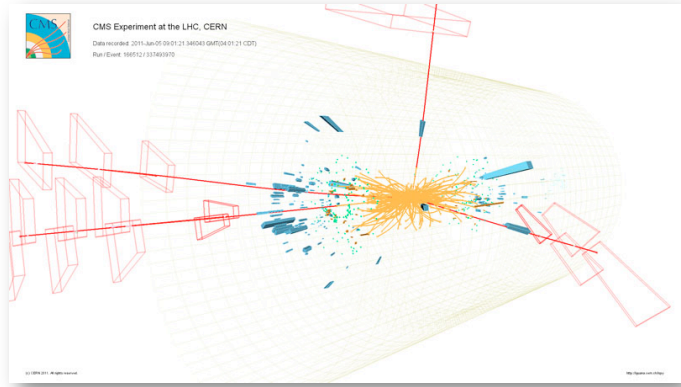
In terms of the running coupling and mass,  $R(1, \alpha_s(Q^2), \frac{m(Q^2)}{Q})$  is a solution of Eq. (24), proven similarly as  $R(1, \alpha_s(Q^2))$  being the solution of Eq. (13). Expanding around  $m(Q^2) = 0$ , we obtain

$$R\left(1, \alpha_s(Q^2), \frac{m(Q^2)}{Q}\right) = R\left(\frac{Q^2}{\mu^2}, \alpha_s, 0\right) + \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{m(Q^2)}{Q}\right)^n R^{(n)}\left(\frac{Q^2}{\mu^2}, \alpha_s, 0\right). \quad (27)$$

We see from Eq. (27) that derivative terms are suppressed by factors of  $1/Q^n$  at large  $Q^2$ . From the dependence of  $R$  on  $\frac{m(Q^2)}{Q}$  we can conclude that the effect of mass is suppressed at high  $Q^2$  by its physical dimension and also by its anomalous dimension, which justifies the assumption about negligible quark masses. The expansion in Eq. (27) has a deeper consequence. The dimensionless observable  $R$  may depend on  $\ln \frac{m(Q^2)}{Q}$  that can become large when  $Q^2$  is large. If we want to avoid such large logarithms, we should consider physical observables (that is physically measurable quantities) that have a finite zero-mass limit.

## 2 Predictions in perturbative QCD

In a typical collider experiment we collect collision events with something interesting in the final state. For instance, in searching for the Higgs boson, events with four hard muons such as in Fig. 6 are interesting. Counting the event rate of such events we obtain measured cross sections, which compare to theoretical predictions. Following our assumption about the use of low-order perturbative predictions in QCD, for such comparisons we need predictions for cross sections with partons. We start with the simplest possible case when partons appear only in the final state: electron-positron annihilation into hadrons (and possibly other particles).



**Fig. 6:** An event with four hard muons in the CMS detector

Let us consider a measurable quantity  $O$ , that has non-vanishing value for at least  $m$  partons in the final state. At LO accuracy the basic formula for the differential cross section in  $O$  is

$$\frac{d\sigma}{dO} = \mathcal{N} \int d\phi_m(p_1, \dots, p_m; Q) \frac{1}{S_{\{m\}}} |\mathcal{M}_m(p_1, \dots, p_m)|^2 \delta(O - O_m^{(m)}(p_1, \dots, p_m)), \quad (28)$$

where  $\mathcal{N}$  contains non-QCD factors (e.g., the flux factor),  $d\phi_m$  is the phase space of  $m$  particles,  $S_m$  is a symmetry factor,  $|\mathcal{M}_m(p_1, \dots, p_m)|^2$  is the squared matrix element (SME), and  $O_m^{(m)}$  is the value of  $O$  computed from the  $m$  final state momenta. The integration is usually done by Monte Carlo integration and the hard part of the computation is to obtain the SME. In these lectures we can compute hardly any SME explicitly. Fortunately, there are freely available computer programs [15–19] that can be used to check the formulae. Even more, these programs can often be used to obtain the cross sections at LO accuracy, too.


We now use Eq. (28) to make predictions for the cross section of electron-positron annihilation into hadrons.

## 2.1 $R$ ratio at lowest order

The leading-order (LO) perturbative contribution to the cross section  $\sigma(e^+e^- \rightarrow \text{hadrons})$  is  $e^+e^- \rightarrow q\bar{q}$ . The calculation is like in the case of  $e^+e^- \rightarrow \mu^+\mu^-$ , supplemented with colour and fractional electric charge of  $q_j$ . The colour diagram is a loop in the fundamental representation which corresponds to a factor  $N_c$  as we have seen in the previous chapter. While the annihilation into  $\mu^+\mu^-$  contains only one flavour in the final state, quarks can have three, four or five flavours depending on the centre-of-mass energy.<sup>3</sup> We have, therefore, to sum over all possible flavours which can appear. The ratio of the two cross sections is thus given by

$$R \equiv \frac{\sigma(e^+e^- \rightarrow q\bar{q})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} = \left( \sum_q e_q^2 \right) N_c, \quad (29)$$

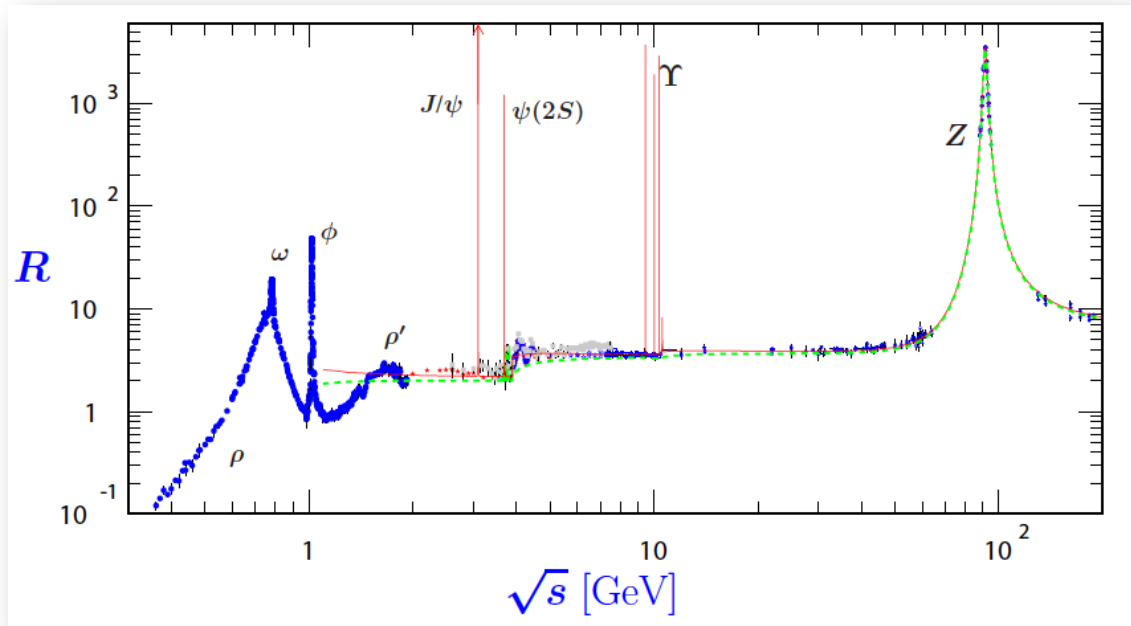
where  $e_u = e_c = \frac{2}{3}$  and  $e_d = e_s = e_b = -\frac{1}{3}$ . If we consider only the up, down, strange and charm quarks  $\sum_q e_q^2 = 2\frac{4}{9} + 2\frac{1}{9} = \frac{10}{9}$ . Considering also the bottom quark  $\sum_q e_q^2 = \frac{11}{9}$ . This step-wise increasing behavior of the  $R$ -ratio was observed (see Fig. 7), providing an experimental confirmation of the existence of 3 families of quarks and of the  $SU(N_c)$  gauge-symmetry of QCD with  $N_c = 3$ .

According to our basic assumption, pQCD cannot give predictions for the resonances in Fig. 7. However, there is one exception, the impressive  $Z$  peak. The LO prediction uses the cross section for the  $e^+e^- \rightarrow q\bar{q}$  process. A  $2 \rightarrow 2$  process has only a single free kinematic variable, the scattering angle  $\vartheta$ . In the full SM the differential cross section for electron-positron annihilation into a massless and colourless fermion pair  $f\bar{f}$  is  from the square of a single Feynman graph,  $\left| \dots \right|^2$ , as

$$\frac{d\sigma}{d\cos\vartheta} = \frac{\pi\alpha^2}{2s} \left\{ (1 + \cos^2\vartheta) \left[ e_f^2 + (A_e^2 + V_e^2)(A_f^2 + V_f^2) \frac{\kappa^2 s^2}{(s - M_Z^2)^2 + \Gamma_Z^2 M_Z^2} + \dots \right] \right\}, \quad (30)$$

where we neglected terms that vanish at centre-of-mass energy  $\sqrt{s} = M_Z$ , or after integration.  $e_f$ ,  $A_f$  and  $V_f$  denote the fractional charge, axial-vector and vector electroweak couplings of the fermions and  $\kappa = \sqrt{2}G_F M_Z^2 / (16\pi\alpha_{\text{em}}) \simeq 0.374$  is a number. Well below the  $Z$  peak the  $Z$  propagator becomes negligible and the total cross section is obtained by integrating over the scattering angle and we find the LO prediction  $\sigma_{\text{LO}}(s) = \sigma_0(s)e_f^2$ , where  $\sigma_0(s) = \frac{4\pi\alpha^2}{3s}$ . On the  $Z$  peak the same integration results in  $\sigma_{\text{LO}}(M_Z^2) = \sigma_0(M_Z^2) \left[ e_f^2 + (A_e^2 + V_e^2)(A_f^2 + V_f^2) \kappa^2 \frac{M_Z^2}{\Gamma_Z^2} \right]$ . Then we can make prediction for the hadronic  $R$  ratio at LO accuracy by simply counting the contributing final states and relating their total charge factors to that of the muon and find  $R_{\text{LO}} = 3 \sum_q e_q^2$  away from the  $Z$  peak and  $R_{Z,\text{LO}} = 3 \sum_q (A_q^2 + V_q^2) / (A_\mu^2 + V_\mu^2)$  on the  $Z$  peak. The factor three is due to the three colours of quarks. Considering five quark flavours, i.e.,  $m_b \ll s \ll m_t$ , we find  $R_{\text{LO}} = 11/3$  and  $R_{Z,\text{LO}} = 20.09$ . We have seen on Fig. 7 that  $11/3$  is fairly close to the measured value away from the  $Z$  peak. The measured

<sup>3</sup>The sixth flavour, the top is so heavy that it cannot contribute at CM energies attained in  $e^+e^-$  experiments so far.



**Fig. 7:** Experimental measurements of the  $R$ -ratio as a function of the total centre-of mass energy (taken from Ref. [14]).

value of  $R_Z$  at LEP is  $R_Z = 20.79 \pm 0.04$  [20]. The LO prediction works amazingly well. The 3.5% difference is mainly due to QCD radiation effects that we call NLO corrections. Our next goal is to understand the origin of those corrections.

**Exercise 2.1** Derive the result in Eq. (30) (at least below the  $Z$  peak, where you consider only photon intermediate state) and integrate it over  $\vartheta$ .

**Exercise 2.2** Use *Mathematica* and the *Package Tracer.m* (or *FORM*) to compute the following traces:

$$\begin{aligned} & \text{Tr} \left( \not{p}_2 \gamma^\nu (\not{p}_1 - \not{k}_1) \gamma^\mu \not{p}_1 \gamma_\mu (\not{p}_1 - \not{k}_1) \gamma_\nu \right) \\ & \text{Tr} \left( \gamma^{\mu_1} \gamma^{\mu_2} \gamma^{\mu_3} \gamma^{\mu_4} \gamma^{\mu_5} \gamma^{\mu_6} \gamma^{\mu_7} \gamma^{\mu_8} \gamma^{\mu_9} \gamma^{\mu_{10}} \gamma_{\mu_1} \gamma_{\mu_2} \gamma_{\mu_3} \gamma_{\mu_4} \gamma_{\mu_5} \gamma_{\mu_6} \gamma_{\mu_7} \gamma_{\mu_8} \gamma_{\mu_9} \gamma_{\mu_{10}} \right) \end{aligned}$$

## 2.2 Ultraviolet renormalization of QCD

The strong coupling is rather large as compared to the other couplings in the SM, and as a result, the QCD radiative corrections are also large. Therefore, it is always important to compute at least the NLO accuracy, but if possible, even higher order corrections.<sup>4</sup>

The computation of QCD radiative corrections is technically quite involved and a good organization of the calculations is very important. Thus, first we introduce some notation. The tensor product of the ket vectors  $|c_1, \dots, c_m\rangle \otimes |s_1, \dots, s_m\rangle$  denotes a basis vector in colour and helicity space,

<sup>4</sup>There is even a more severe reason that we shall discuss later.

$|\mathcal{A}_m(p_1, \dots, p_m)\rangle$  is a state vector of  $n = m - 2$  final-state particles in colour and helicity space. The amplitude for producing  $n$  final-state particles of colour  $(c_1, \dots, c_n)$ , spin  $(s_1, \dots, s_n)$ , momentum  $(p_1, \dots, p_n)$  is

$$\mathcal{A}_m^{c_1 \dots c_m, s_1 \dots s_m}(p_1 \dots p_m) \equiv \langle c_1 \dots c_m | \otimes \langle s_1 \dots s_m | \mathcal{A}_m(p_1, \dots, p_m) \rangle \quad (31)$$

( $m = n + 2$ ), so

$$\sum_{\text{colour}} \sum_{\text{helicity}} |\mathcal{A}_m^{c_1 \dots c_m, s_1 \dots s_m}(\{p_i\})|^2 = \langle \mathcal{A}_m(\{p_i\}) | \mathcal{A}_m(\{p_i\}) \rangle. \quad (32)$$

The loop expansion in terms of the bare coupling, i.e., the coupling that appears in the classical Lagrangian,  $g_s^{(0)} \equiv \sqrt{4\pi\alpha_s^{(0)}}$  is:

$$|\mathcal{A}_m\rangle = \left( \frac{\alpha_s^{(0)} \mu^{2\epsilon}}{4\pi} \right)^{\frac{q}{2}} \left[ |\mathcal{A}_m^{(0)}\rangle + \left( \frac{\alpha_s^{(0)} \mu^{2\epsilon}}{4\pi} \right) |\mathcal{A}_m^{(1)}\rangle + \mathcal{O}((\alpha_s^{(0)})^2) \right], \quad (33)$$

where  $q \in \mathbb{N}$ ,  $\mu$  is the dimensional regularization scale, introduced to keep  $\alpha_s^{(0)}$  dimensionless in  $d = 4 - 2\epsilon$  dimensions. The exponent  $\frac{q}{2}$  in the prefactor takes account of the power of  $\alpha_s$  at LO, the loop-expansion is an expansion in the strong coupling  $\alpha_s$ . For instance,  $q = 0$  for  $e^+e^- \rightarrow q\bar{q}$ , while  $q = 1$  for  $e^+e^- \rightarrow q\bar{q}g$ . The tree amplitude  $|\mathcal{A}_m^{(0)}\rangle$  is finite, while the one-loop correction  $|\mathcal{A}_m^{(1)}\rangle$  is divergent in  $d = 4$  dimensions, which is manifest in terms of  $1/\epsilon^2$  and  $1/\epsilon$  poles if dimensional regularization is used. These poles have both ultraviolet (UV) and infrared (IR) origin.

The UV poles can be removed by multiplicative redefinition of the fields and parameters in the Lagrangian, systematically order by order in PT. This is a hard task even at one loop, but presently known up to four loops [21] – a truly remarkable computation! It turns out that when computing scattering amplitudes in massless QCD at one-loop accuracy, the renormalization amounts to the simple substitution

$$\alpha_s^{(0)} \mu^{2\epsilon} \longrightarrow \alpha_s(\mu_R^2) \mu_R^{2\epsilon} S_\epsilon^{-1} \left[ 1 - \frac{\alpha_s(\mu_R^2)}{4\pi} \frac{\beta_0}{\epsilon} + \mathcal{O}(\alpha_s^2) \right], \quad (34)$$

with  $S_\epsilon = \frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}$ . Note that on the left of this substitution  $\mu$  is the dimensional regularization scale to keep  $\alpha_s^{(0)}$  dimensionless, while on the right  $\mu_R$  is the renormalization scale. We discussed in Sect. 1.8 when we extract  $\alpha_s$  from measurements, we have to define  $\mu_R$ . The dimensional regularization scale turns into the renormalization scale through the substitution (34).

Why does the substitution (34) work? Each Feynman graph consists of vertices with propagators connecting those and external lines. Moreover,

- each vertex receives a factor  $Z_g$  (or  $Z_g^2$  for quartic vertex) and factors of  $\sqrt{Z_i}$ , ( $i = q, A$ ) for each field connected to the vertex,
- each propagator of field  $i$  receives a factor of  $Z_i^{-1}$ ,
- each external leg of field  $i$  receives a factor of  $Z_i^{-\frac{1}{2}}$ .

Thus the renormalization field factors cancel from each graph and only the charge renormalization ( $Z_g$ ) is needed in practice! This can be seen as a consequence of the fact that in massless QCD the only free parameter besides the gauge-fixing parameter  $\lambda$  is  $\alpha_s$ . The scattering amplitudes are physical, and any physical quantity has to be independent of  $\lambda$ , so the only remaining parameter, which the amplitudes may depend on, is the coupling. The renormalization factor  $Z_g$  is most easily computed in background field gauge, defined by

$$\mathcal{L}_{\text{GF}} = -\frac{1}{2\lambda} \sum_a \left( \partial^\mu A_\mu^a + g f^{abc} A_\mu^b A^{\mu c} \right)^2, \quad (35)$$

where  $\mathcal{A}_\mu^b$  is a background field and  $\mathcal{A}_\mu^c$  describes the quantum fluctuations on this background. It can be shown [11] that in this gauge the field and coupling renormalization factors are related by the Ward identity  $Z_A^{-\frac{1}{2}} = Z_g$ , and  $Z_A$  can be computed from loop insertions into the propagator shown in (21).

The simple substitution rule (34) for the coupling leads to a simple shift in the amplitude. As

$$\left[ 1 - \frac{\alpha_s(\mu_R^2)}{4\pi} \frac{\beta_0}{\epsilon} \right]^{\frac{q}{2}} = 1 - \frac{q}{2} \frac{\alpha_s(\mu_R^2)}{4\pi} \frac{\beta_0}{\epsilon} + O(\alpha_s^2)$$

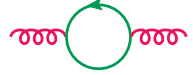
we obtain for the renormalized amplitudes  $|\mathcal{M}_m^{(i)}\rangle$  ( $i = 0, 1$ )

$$\begin{aligned} |\mathcal{M}_m^{(0)}\rangle &= \left( \frac{\alpha_s(\mu_R^2) \mu_R^{2\epsilon}}{4\pi} S_\epsilon^{-1} \right)^{\frac{q}{2}} |\mathcal{A}_m^{(0)}\rangle \quad q \in \mathbb{N}, \\ |\mathcal{M}_m^{(1)}\rangle &= \left( \frac{\alpha_s(\mu_R^2) \mu_R^{2\epsilon}}{4\pi} S_\epsilon^{-1} \right)^{\frac{q}{2}} \frac{\alpha_s(\mu_R^2)}{4\pi} S_\epsilon^{-1} \left( \mu_R^{2\epsilon} |\mathcal{A}_m^{(1)}\rangle - \frac{q}{2} \frac{\beta_0}{\epsilon} S_\epsilon |\mathcal{A}_m^{(0)}\rangle \right). \end{aligned} \quad (36)$$

The renormalized theory is UV finite, yet  $|\mathcal{M}_m^{(1)}\rangle$  is still infinite in  $d = 4$  dimensions, as it is divergent also in the infrared. After UV renormalization is achieved we can use dimensional regularization to regulate the amplitudes in the IR by continuing into  $d > 4$  ( $\epsilon < 0$ ). The integrals that are scaleless in  $d = 4$  have  $(q^2)^{-\epsilon}$  mass dimension in  $d = 4 - 2\epsilon$  dimensions. Therefore, in the massless limit all integrals can depend only on momentum invariants raised to a positive fractional power ( $\epsilon < 0$ ). We conclude that when all external invariants vanish, the continued integral must also vanish (“scaleless integrals vanish in dimensional regularization”).

For *IR-safe* observables these IR poles vanish and we can set  $d = 4$  at the end of the computations, and we obtain the UV finite, IR regularized SME that can be used to compute cross sections.

**Exercise 2.3** Compute the contribution to the beta function from the fermion loop:



1. Write down carefully the amplitude and compute the trace.
2. The following types of integrals occur:

$$I_2^\mu = \int \frac{d^d \ell}{(2\pi)^d} \frac{\ell^\mu}{\ell^2 (\ell - p)^2}, \quad I_2^{\mu\nu} = \int \frac{d^d \ell}{(2\pi)^d} \frac{\ell^\mu \ell^\nu}{\ell^2 (\ell - p)^2} \quad (37)$$

Express these as linear combination of

$$I_2(p) = \int \frac{d^d \ell}{(2\pi)^d} \frac{1}{\ell^2 (\ell - p)^2}. \quad (38)$$

3. Obtain  $I_2$  from

$$I_2(p, m) = \int \frac{d^d l}{(2\pi)^d} \frac{1}{[(l-p)^2 - m^2] l^2} = \frac{i}{(4\pi)^{2-\epsilon}} \Gamma(\epsilon) (p^2)^{-\epsilon} \int_0^1 dx \left( x \frac{m^2}{p^2} - x(1-x) - i\epsilon \right)^{-\epsilon} \quad (39)$$

and find the divergent pieces.

The contribution to  $\beta_0$  is the coefficient of the  $1/\epsilon$  pole without the coupling factor.

### 2.3 $R$ ratio at NLO accuracy

This is by far the simplest example of computing QCD radiative corrections. As we saw in Sect. 2.1 it requires the total hadronic cross section that depends only on a single kinematic invariant, the total centre-of-mass energy  $\sqrt{s}$ . As a result, the emerging integrals in this computation can be evaluated exactly. Nevertheless, the complete computation is still too lengthy, and we shall be able to present the main step and filling the details is left to the student.

There are two kinds of corrections that contribute at NLO accuracy. One is the *real correction*, with an additional gluon in the final state, so the SME is computed from Feynman graphs as

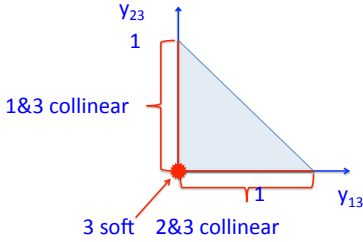
$\langle \mathcal{M}_3^{(0)} | \mathcal{M}_3^{(0)} \rangle = \left| \left\langle \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \end{array} \right. \right|^2$ , which gives an  $\mathcal{O}(\alpha_s)$  correction. The other kind of contribution

is the *virtual correction*, with an additional gluon providing a loop in the final state,

$$\langle \mathcal{M}_2^{(1)} | \mathcal{M}_2^{(0)} \rangle + \langle \mathcal{M}_2^{(0)} | \mathcal{M}_2^{(1)} \rangle = 2\text{Re} \left\langle \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \\ \text{---} \text{---} \text{---} \end{array} \right\rangle.$$

The real correction has three particles in the final state. The three-particle phase space has five independent variables: two energies and three angles. As we are looking for the total cross section, we integrate over the angles and use  $y_{ij} = (p_i + p_j)^2/s = 2p_i \cdot p_j/s$  scaled two-particle invariants to write both the phase space and the SME. Momentum conservation implies  $1 = (p_1 + p_2 + p_3)^2/s = y_{12} + y_{13} + y_{23}$ . The complete real contribution to the total cross section is

$$\sigma^R = \sigma_0 R_0 \int_0^1 dy_{13} \int_0^1 dy_{23} C_F \frac{\alpha_s}{2\pi} \left( \frac{y_{23}}{y_{13}} + \frac{y_{13}}{y_{23}} + \frac{2y_{12}}{y_{13}y_{23}} \right) \Theta(1 - y_{13} - y_{23}). \quad (40)$$



**Fig. 8:** Region of integration for real correction

This integral is divergent along the boundaries at  $y_{13} = 0$ ,  $y_{23} = 0$  as well as in the point  $y_{13}y_{23} = 0$ , so the singularities are in the IR parts of the phase space. As  $y_{i3}s = 2E_i E_3(1 - \cos \vartheta_{i3})$ , the divergence occurs either when  $E_3 \rightarrow 0$ , which is called *soft-gluon* singularity, or when  $\vartheta_{i3} \rightarrow 0$ , which is called *collinear* singularity (the gluon is collinear to either of the quarks). The region of integration with the singular places is shown in Fig. 8.

To make sense of the integral, we use dimensional regularization, which amounts to the computation of the phase space and the SME in  $d = 4 - 2\epsilon$  dimensions. The result is

$$\begin{aligned} \sigma^R(\epsilon) &= \sigma_0 R_0 H(\epsilon) \int_0^1 \frac{dy_{13}}{y_{13}^\epsilon} \int_0^1 \frac{dy_{23}}{y_{23}^\epsilon} \Theta(1 - y_{13} - y_{23}) \\ &\times C_F \frac{\alpha_s}{2\pi} \left[ (1 - \epsilon) \left( \frac{y_{23}}{y_{13}} + \frac{y_{13}}{y_{23}} \right) + \frac{2y_{12}}{y_{13}y_{23}} - 2\epsilon \right], \end{aligned} \quad (41)$$

where  $H(\epsilon) = 1 + \mathcal{O}(\epsilon)$  (the exact form of this function will turn out to be irrelevant). The integrals can be evaluated exactly, but actually the Laurent-expansion around  $\epsilon = 0$  is sufficient,

$$\sigma^R(\epsilon) = \sigma_0 R_0 H(\epsilon) C_F \frac{\alpha_s}{2\pi} \left[ \frac{2}{\epsilon^2} + \frac{3}{\epsilon} + \frac{19}{2} - \pi^2 + \mathcal{O}(\epsilon) \right]. \quad (42)$$

The computation of the virtual correction is even more cumbersome due to the loop integral. We present only the result:

$$\sigma^V(\epsilon) = \sigma_0 R_0 H(\epsilon) C_F \frac{\alpha_s}{2\pi} \left[ -\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + \pi^2 + \mathcal{O}(\epsilon) \right], \quad (43)$$

We now see that the *sum of the real and virtual contribution is finite in  $d = 4$* , so for the sum we can set  $\epsilon \rightarrow 0$  and find the famous  $\alpha_s/\pi \simeq 0.037$  correction:  $R = R_0(1 + \frac{\alpha_s}{\pi} + \mathcal{O}(\alpha_s^2))$ . The correction is the same for  $R_Z$ .

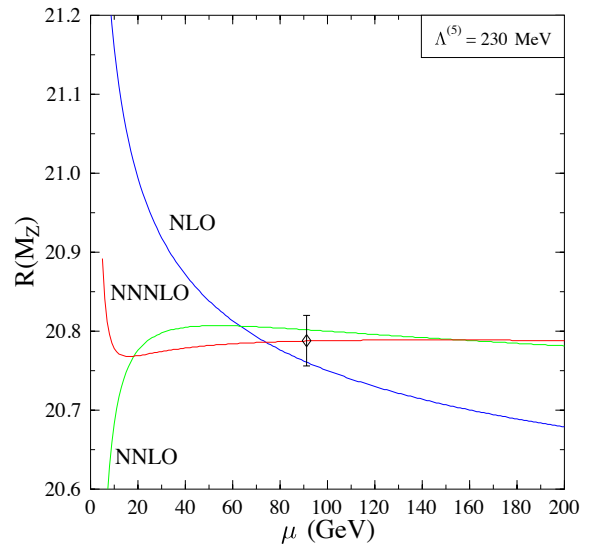
Actually there is a much easier way of computing the radiative corrections to the total cross section from the imaginary part of the hadronic vacuum polarization, using the optical theorem ( $\sigma \propto \text{Im}f(\gamma \rightarrow \gamma)$ ). The state of the art is  $R$  at  $\mathcal{O}(\alpha_s^4)$  [22]. The result of the computation at next-to-next-to-leading order (NNLO) accuracy,

$$\begin{aligned} \frac{R}{R_0} = 1 + c_1\alpha_s(\mu) + \left[ c_2 + c_1 b_0 \ln \frac{\mu}{Q^2} \right] \alpha_s(\mu)^2 \\ + \left[ c_3 + \left( 2c_2 b_0 + c_1 b_1 + c_1 b_0^2 \ln \frac{\mu^2}{Q^2} \right) \ln \frac{\mu^2}{Q^2} \right] \alpha_s(\mu)^3 + \mathcal{O}(\alpha_s^4) \end{aligned} \quad (44)$$

satisfies the RGE to order  $\alpha_s(\mu)^4$ . The coefficients  $c_1 = 1/\pi$ ,  $c_2 = 1.409/\pi^2$ ,  $c_3 = -12.85/\pi^3$  suggest that the perturbation series is convergent. Our more complicated way of computing  $R$  is instructive for our studies in the next section.

The predictions at the first three fixed orders in PT for  $R_Z$  are shown in Fig. 9. The  $R$  ratio at LO accuracy does not depend on the strong coupling, hence it is independent of the scale. The figure is meant to show the general pattern of QCD predictions which, with the exception of the  $R$  ratio, depend on the scale already at LO. The NLO curve shows the typical feature of LO predictions: it depends on the renormalization scale in a monotonically decreasing way. As this scale is unphysical, in principle, its value can be arbitrary. Thus the prediction at LO is in general only an order of magnitude indication of the cross section, but not a precision result. (In the case of the hadronic ratio the QCD corrections are actually quite small as compared to many other QCD cross sections and the precision is actually better than usual.) As a result, if we want to make reliable predictions in pQCD, the NLO accuracy (NNLO for  $R$ ) is absolutely necessary unless we have some way to fix the scale.

However, there is no theorem that tells us the proper scale choice. The usual practice is to set the scale at a characteristic physical scale of the process. A reasonable assumption that the strength of the QCD interaction for a process involving a momentum transfer  $Q$  is given by  $\alpha_s(Q)$ , so  $\mu = Q$  is the proper scale choice, to minimize logarithmic contributions  $\ln(\mu/Q)$  in higher-order terms. For instance, in case of electron-positron annihilation the total centre-of-mass energy is the usual choice, while for a jet cross section in proton-proton collisions the transverse momentum of the jet<sup>5</sup> is used. The application of this recipe appears clear as long as there is only one hard scale in the process. In the state of the art computations there are complex processes with several scales and it is not obvious which one to choose. For instance, in vector boson hadroproduction in association with  $m$  jets ( $m \leq 5$ ) [23], in addition to the transverse momentum of the vector boson  $E_{V,\perp} = \sqrt{p_{V,x}^2 + p_{V,y}^2}$  there are the transverse momenta of the jets. In this example, the choice  $\mu = E_{\perp}^V$  was found to result in a badly behaving perturbation series



**Fig. 9:** Dependence on the renormalization scale of the hadronic ratio on the  $Z$  pole

<sup>5</sup>We discuss jets in the next section.



with corrections driving the  $E_{\perp}$  distribution of the second hardest jet at NLO accuracy even unphysically negative for  $m = 3$  and  $E_{\perp} > 475$  GeV at the LHC. Choosing a dynamical scale, set event by event, appears a better choice. For instance, half the total transverse energy of the final-state particles (both QCD partons and leptons from the decay of the vector boson),  $\mu = \hat{H}_T/2$ , leads to much milder scale dependence and a similar shape of the distributions at LO and NLO accuracies.

There are suggestions on making educated guesses for the best scale. Among those are the principle of fastest apparent convergence (FAC), that of minimal sensitivity (PMS), or the BLM scale choice [24–26], beyond the scope of these lectures. The experience is that in hadron collisions there is no choice that works well for any process and it is best to choose a dynamical scale chosen by examining the process.

As there is no unique scale, the standard procedure is to choose a default scale  $\mu_0$ , related to the typical momentum transfer in the process, and to assign a theoretical uncertainty by varying the scale within a certain range around the default choice  $\mu_0$ . The usual range is between half and twice the default choice. However, this is again an indication only of the scale uncertainties and there is no mathematical theorem that states this procedure yields the true theoretical uncertainty due to neglected higher order terms. In order to have a measure on the effect of neglected higher orders, i.e., to understand the reliability of the assigned theoretical uncertainty one has to compute the NNLO corrections. The latter are very demanding computations both technically and numerically and predictions at NNLO accuracy for some fairly simple processes, with one or two final-state particles in the prediction at LO, constitute the state of the art of pQCD.

**Exercise 2.4** Show that the  $d$ -dimensional three-particle phase space for  $q \rightarrow p_1 + p_2 + p_3$  can be expressed in terms of the Lorentz-invariants  $s_{ij} = (p_i + p_j)^2$

$$d\phi_3 = (2\pi)^{3-2d} 2^{-1-d} (q^2)^{\frac{2-d}{2}} d^{d-2}\Omega d^{d-3}\Omega (s_{12}s_{13}s_{23})^{\frac{d-4}{2}} ds_{12} ds_{13} ds_{23} \delta(q^2 - s_{12} - s_{13} - s_{23}) .$$

where  $d^d\Omega$  is the measure of the hypersurface element in  $d$  dimensions,  $\int d^{d-3}\Omega = \Omega_d = 2\pi^{d/2}/\Gamma(d/2)$ .  
Hints:

1. The  $d$ -dimensional volume measure in spherical coordinates is recursively given by

$$d^{d+1}p = E dE d^d p_E = E dE E^{d-1} d^d\Omega, \quad d^d\Omega = (\sin\theta_1)^{d-1} d\theta_1 d^{d-1}\Omega .$$

2. Show that

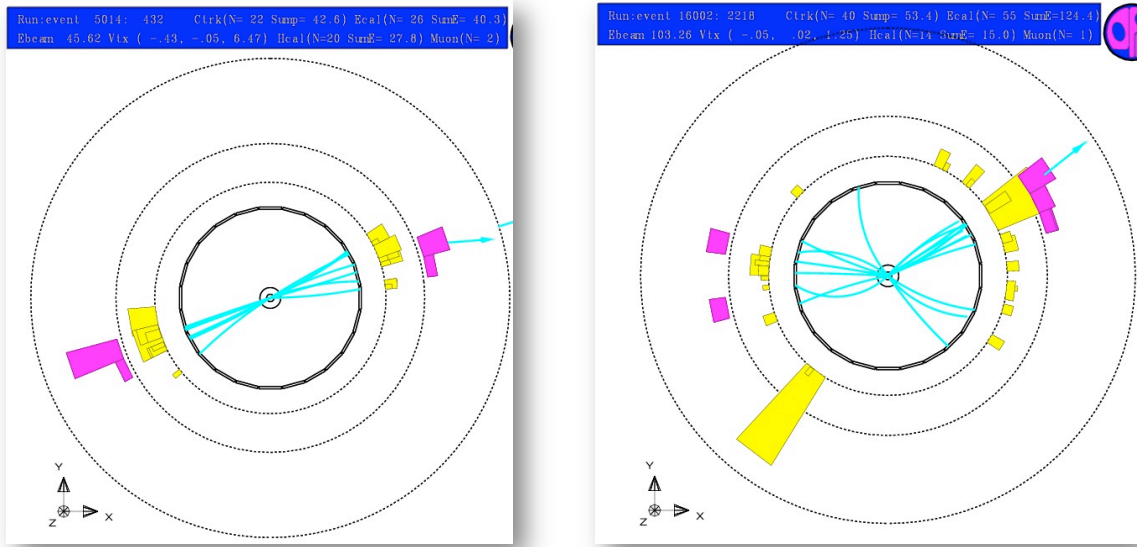
$$\sin^2\theta_1 = \frac{1}{4} \frac{s_{12} s_{13} s_{23}}{q^2 E_1^2 E_2^2} ,$$

where  $\theta_1$  is the angle between  $p_1$  and  $p_2$ .

**Exercise 2.5** Let  $y_{ij} = \frac{s_{ij}}{q^2}$ . Using the previous exercise, compute the real correction to the process  $e^+ e^- \rightarrow q\bar{q}$  given in Eq. (42). Hint: Transform the triangular integration region into the unit square and evaluate the  $B$  (Euler  $\beta$ ) functions.

### 3 Jet cross sections

In the first two sections we established our theoretical playground to make predictions for hadronic cross sections. Based on RGE analysis we showed that PT can only be fully consistent in an asymptotically free QFT, like QCD. We found that predictions can be made only for those quantities that remain finite in the limit of vanishing masses of light quarks. We computed the radiative corrections for such a quantity, the total hadronic cross section in electron-positron annihilation. We found that at intermediate steps of the computations there are singular contributions of two types: of UV and IR origin. The UV singularities can be removed by renormalization, and the remaining IR ones can be regularized in dimensional regularization where IR singularities appear as  $1/\epsilon$  poles. When adding all contributions, these poles cancel and we obtain the finite correction after setting  $\epsilon = 0$ . Our question in this section is whether there are more exclusive observables than the totally inclusive one for which this procedure can be applied.



**Fig. 10:** Two events observed in the OPAL detector

It is clear from experiments that typical final states have structures. For instance, Fig. 10 shows two events, one with two and the other with three sprays of hadrons, called *hadron jets*. If we count the relative number of events with two, three, four jets, an interesting pattern emerges:

$$\# \text{ of events with 2 jets} : \# \text{ of events with 3 jets} : \# \text{ of events with 4 jets} \simeq \mathcal{O}(\alpha_s^0) : \mathcal{O}(\alpha_s^1) : \mathcal{O}(\alpha_s^2).$$

Recalling our basic assumption and Fig. 3 we find that jets reflect the partonic structure of the events. We now use our pQCD formalism to describe these structures theoretically. For this purpose, we need a function of the final state momenta  $J_m(\{p_i\})$  that quantifies the structure of the final state in some ways (we give examples below). This function is called *jet function*.

Let us consider again the process  $e^+e^- \rightarrow \text{hadrons}$ . If we are not interested in the orientation of the final state events, we can average over the orientation and find that the SME  $|\mathcal{M}_2|^2$  has no dependence on the parton momenta. Then the two-particle phase space is  $d\phi_2 = dy_{12}\delta(1 - y_{12})$ , and contribution of the process  $e^+e^- \rightarrow q\bar{q}$  to the cross section is

$$\sigma_{\text{LO}} = |\mathcal{M}_2|^2 \int_0^1 dy_{12} \delta(1 - y_{12}) J_2(p_1, p_2), \quad (45)$$

which sets our normalization of  $|\mathcal{M}_2|^2$ . The two kinds of NLO corrections are

$$\begin{aligned} d\sigma^R &= |\mathcal{M}_2|^2 S_\epsilon \frac{dy_{13}}{y_{13}^\epsilon} \frac{dy_{23}}{y_{23}^\epsilon} C_F \frac{\alpha_s}{2\pi} \left[ (1-\epsilon) \left( \frac{y_{23}}{y_{13}} + \frac{y_{13}}{y_{23}} \right) + \frac{2y_{12}}{y_{13}y_{23}} - 2\epsilon \right] J_3(p_1, p_2, p_3), \\ d\sigma^V &= |\mathcal{M}_2|^2 S_\epsilon C_F \frac{\alpha_s}{2\pi} \left( \frac{\mu^2}{s} \right)^\epsilon \left[ -\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + \pi^2 + \mathcal{O}(\epsilon) \right] dy_{12} \delta(1-y_{12}) J_2(p_1, p_2). \end{aligned} \quad (46)$$

Contrary to the case of the total cross section, where  $J_m = 1$ , we cannot simply perform the integration analytically and combine the results, neither we can combine the integrands. The general method to deal with this problem is to regularize both with a properly chosen subtraction,

$$d\sigma_3^{\text{NLO}} = d\sigma^R J_3 - d\sigma^A J_2, \quad \text{and} \quad d\sigma_2^{\text{NLO}} = (d\sigma^V + d\sigma^A) J_2,$$

such that both terms are separately integrable in  $d = 4$  dimensions. This requires a special property of the jet function  $J_n$ , called IR safety, expressed analytically as

$$\lim_{y_{13}, y_{23} \rightarrow 0} J_3 = J_2. \quad (47)$$

Qualitatively IR safety means that the jet function is insensitive to an additional soft particle, or to a collinear splitting in the final state.

How can we construct such an approximate cross section? For this simple process we can follow the steps:

$$\begin{aligned} \frac{y_{23}}{y_{13}} + \frac{y_{13}}{y_{23}} + \frac{2y_{12}}{y_{13}y_{23}} &= \frac{y_{23}}{y_{13}} + \frac{1}{y_{13}} \frac{2y_{12}}{y_{13} + y_{23}} + (1 \leftrightarrow 2) \\ &= \frac{1}{y_{13}} \left[ y_{23} + \left( 2 \frac{\overbrace{y_{12} + y_{13} + y_{23}}^{=1}}{y_{13} + y_{23}} - 2 \right) \right] + (1 \leftrightarrow 2). \end{aligned} \quad (48)$$

Then introduce the new variable  $z_1 \equiv z_{1,2} = \frac{y_{12}}{y_{12} + y_{23}}$ , so that  $y_{13} + y_{23} = 1 - y_{12} = 1 - z_1(1 - y_{13})$  and

$$\frac{y_{23}}{y_{13}} - \frac{1}{y_{13}} = \frac{y_{23}(1 - y_{13}) - y_{12} - y_{23}}{y_{13}(y_{12} + y_{23})} = \frac{-y_{23}y_{13} - y_{12}}{y_{13}(y_{12} + y_{23})} = -\frac{y_{23}}{y_{12} + y_{23}} - \frac{z_1}{y_{13}},$$

and substitute these into Eq. (48):

$$\frac{y_{23}}{y_{13}} + \frac{y_{13}}{y_{23}} + \frac{2y_{12}}{y_{13}y_{23}} = \left[ \frac{1}{y_{13}} \left( \frac{2}{1 - z_1(1 - y_{13})} - 1 - z_1 \right) - \frac{y_{23}}{y_{12} + y_{23}} \right] + (1 \leftrightarrow 2). \quad (49)$$

The term  $y_{23}/y_{12} + y_{23}$  never becomes infinite, thus the approximate cross section

$$d\sigma^A = |\mathcal{M}_2|^2 S_\epsilon \frac{dy_{13}}{y_{13}^\epsilon} \frac{dy_{23}}{y_{23}^\epsilon} C_F (V_{13,2} + V_{23,1}), \quad \text{with} \quad (50)$$

$$V_{ij,k} = \frac{\alpha_s}{2\pi} \left[ \frac{1}{y_{ij}} \left( \frac{2}{1 - z_{i,k}(1 - y_{ij})} - 1 - z_{i,k} \right) - \epsilon(1 - z_{i,k}) \right] \quad (51)$$

is a proper subtraction term that regularizes the real contribution in all of its singular limits in  $d$  dimensions. Consequently, the difference  $d\sigma_3^{\text{NLO}} = d\sigma^R J_3 - d\sigma^A J_2$  can be integrated in any dimensions, in particular, we can set  $\epsilon = 0$  and integrate in  $d = 4$  numerically.

To obtain  $d\sigma_2^{\text{NLO}}$  we integrate the two terms separately. For  $V_{13,2}$  we change variables in the phase space to  $y_{13}$  and  $z_1$ , and find

$$d\sigma^A = |\mathcal{M}_2|^2 S_\epsilon \left( \frac{\mu^2}{s} \right)^\epsilon \int_0^1 dy_{13} \int_0^1 dz_1 y_{13}^{-\epsilon} (1 - y_{13})^{1-2\epsilon} C_F V_{13,2} + (1 \leftrightarrow 2). \quad (52)$$

We shall see that this factorization of the singular terms is universal. We can now perform the integration over the factorized one-particle phase space, independently of the jet function, and obtain the *integrated subtraction term* in the form  $d\sigma^A = |\mathcal{M}_2|^2 \mathbf{I}(\epsilon)$  with *insertion operator*

$$\mathbf{I}(\epsilon) = C_F \frac{\alpha_s}{2\pi} \frac{1}{\Gamma(1-\epsilon)} \left( \frac{4\pi\mu^2}{s} \right)^\epsilon \left[ \frac{2}{\epsilon^2} + \frac{3}{\epsilon} + 10 - \frac{\pi^2}{3} + \mathcal{O}(\epsilon) \right]. \quad (53)$$

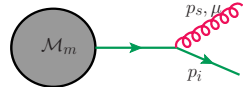
Comparing this integrated subtraction to Eq. (46), we see that the sum  $d\sigma_2^{\text{NLO}} = (d\sigma^V + d\sigma^A)J_2$  is finite if  $\epsilon = 0$ ,

$$\sigma_2^{\text{NLO}} = |\mathcal{M}_2|^2 C_F \frac{\alpha_s}{\pi} \left( 1 + \frac{\pi^2}{3} \right) \int_0^1 dy_{12} \delta(1-y_{12}) J_2(p_1, p_2) + \mathcal{O}(\epsilon), \quad (54)$$

and so can be integrated in  $d = 4$  dimensions.

### 3.1 Infrared safety

A natural question is if we can construct the approximate cross section universally, i.e., independently of the process and observable. Our presentation above suggests the affirmative answer. To understand how, we have to study the origin of the singular behaviour in the SME. This singularity arises from propagator factors that diverge



The diagram shows a circular vertex labeled  $\mathcal{M}_m$  on the left. A green arrow points from the vertex to the right, representing a propagator. This propagator then splits into two outgoing particles: a green arrow labeled  $p_i$  and a red wavy arrow labeled  $p_s, \mu$ .

$$\propto \frac{1}{(p_i + p_s)^2} = \frac{1}{2p_i \cdot p_s} = \frac{1}{2E_i E_s (1 - \cos\theta)} \simeq \frac{1}{E_i E_s \theta^2}$$

In the collinear limit,  $\theta \rightarrow 0$  and  $\mathcal{M}_{m+1} \simeq \mathcal{M}_m/\theta$  + less singular terms (a factor of  $\theta$  appears in the numerator factors). In the soft limit,  $E_s \rightarrow 0$  and  $\mathcal{M}_{m+1} \simeq \mathcal{M}_m/E_s$  + less singular terms. The gluon phase space is

$$\frac{d^3 p_s}{2E_s} = \frac{1}{2} E_s dE_s d\cos\theta d\phi \simeq \frac{1}{4} E_s dE_s d\theta^2 d\phi,$$

so in the cross section we find logarithmic singularities in both the soft and the collinear limits:  $\frac{dE_s}{E_s}$  or  $\frac{d\theta^2}{\theta^2}$ . These are the IR singular limits. In dimensional regularization the logarithmic singularities appear as poles:

$$\int dy_{is} y_{is}^{-1-\epsilon} = -\frac{1}{\epsilon}.$$

Thus, the singular behaviour arises at kinematically degenerate phase space configurations, which at the NLO accuracy means that one cannot distinguish the following configurations: (i) a single hard parton, (ii) the single parton splitting into two nearly collinear partons, (iii) the single parton emitting a soft gluon (on-shell gluon with very small energy). Then an answer to the question posed at the beginning of Sect. 3 is given by the Kinoshita-Lee-Nauenberg (KLN) theorem [27, 28]:

*In massless, renormalized field theory in four dimensions, transition rates are IR safe if summation over kinematically degenerate initial and final states is carried out.*

For the  $e^+e^- \rightarrow$  hadrons process, the initial state is free of IR singularities. Typical IR-safe quantities are (i) event shape variables and (ii) jet cross sections.

### 3.2 Event shapes

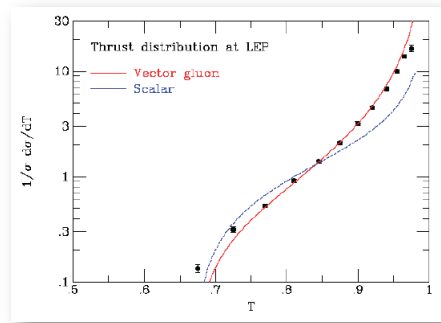
Thrust, thrust major/minor, C- and D-parameters, oblateness, sphericity, aplanarity, jet masses, jet-broadening, energy-energy correlation, differential jet rates are examples of event shape variables. The value of an event shape does not change if a final-state particle further splits into two collinear particles, or emits a soft gluon, hence it is (qualitatively) IR safe. As an example we consider the thrust  $T$ , which is defined by

$$T = \max_{\vec{n}} \frac{\sum_{i=1}^m |\vec{p}_i \cdot \vec{n}|}{\sum_{i=1}^m |\vec{p}_i|}, \quad (55)$$

where  $\vec{n}$  is a three-vector (the direction of the thrust axis) such that  $T$  is maximal. The particle three-momenta  $\vec{p}_i$  are defined in the  $e^+e^-$  centre-of-mass frame.  $T$  is an example of the jet function  $J_m$ . It is infrared safe because neither  $p_j \rightarrow 0$ , nor replacing  $p_i$  with  $zp_i + (1-z)p_i$  change  $T$ . At LO accuracy it is possible to perform the phase space integrations and

$$\frac{1}{\sigma} \frac{d\sigma}{dT} = C_F \frac{\alpha_s}{2\pi} \left[ \frac{2(3T^2 - 3T + 2)}{T(1-T)} \ln \left( \frac{2T-1}{1-T} \right) - 3(3T-2) \frac{2-T}{1-T} \right]. \quad (56)$$

We see that the perturbative prediction for the thrust distribution is singular at  $T = 1$ . In addition to the linear divergence in  $1 - T$  there is logarithmic divergence, too. The latter is characteristic to event shape distributions. In PT at  $n$ th order logarithms of  $1 - T$  in the form  $\alpha_s^n \ln^m(1/(1 - T))$ ,  $m \leq n + 1$  appear in the exponent of the cross-section. These spoil the convergence of the perturbation series and call for resummation if we want to make reliable prediction near the edge of the phase space, for large values of  $T$  where the best experimental statistics are available. Resummations of leading ( $m = n + 1$ ) and next-to-leading ( $m = n$ ) logarithms are available for many event shape variables, but the discussion of this technique is beyond the scope of these lectures.



**Fig. 11:** Distribution of thrust as measured at LEP compared to pQCD predictions obtained with vector and scalar gluon

---

**Exercise 3.1** Verify that  $T$  as defined in Eq. (55) is infrared and collinear safe. What is the range of values that  $T$  can take if (i) there are only two particles in the final state, or (ii)  $m \rightarrow \infty$  and all  $\vec{p}_i$  are distributed spherically?

---

### 3.3 Jet algorithms

Jets are sprays of energetic, on-shell, nearly collinear hadrons. The number of jets does not change if a final-state particle further splits into two collinear particles, or emits a soft gluon, hence it is again qualitatively IR safe. To quantify the jet-like structure of the final states jet algorithms have been invented. These have a long history with rather slow convergence. The reason is that the experimental and theoretical requirements posed to a jet algorithm are rather different. Experimentally we need cones that include almost all hadron tracks at cheap computational price. Theoretically the important requirements are IR safety, so that PT can be employed to make predictions and resummability, so that we can make predictions in those region of the phase space where most of the data appear.

For many years experimenters preferred cone jet algorithms (according to the ‘Snowmass accord’) [29]. These start from a cone seed (centre of the cone) in pseudorapidity ( $\eta$ ) and azimuthal-angle ( $\phi$ )

plane:  $(\eta_c, \phi_c)$ . We define a distance of a hadron track  $i$  from the seed by  $d_{ic} = \sqrt{(\eta_i - \eta_c)^2 + (\phi_i - \phi_c)^2}$ . A track belongs to the cone if  $d_{ic} < R$ , with a predefined value for  $R$  (usually 0.7). It turned out however, that (i) this is an IR unsafe definition and (ii) there is a problem how to treat overlapping cones, so the cone jet definition has been abandoned.

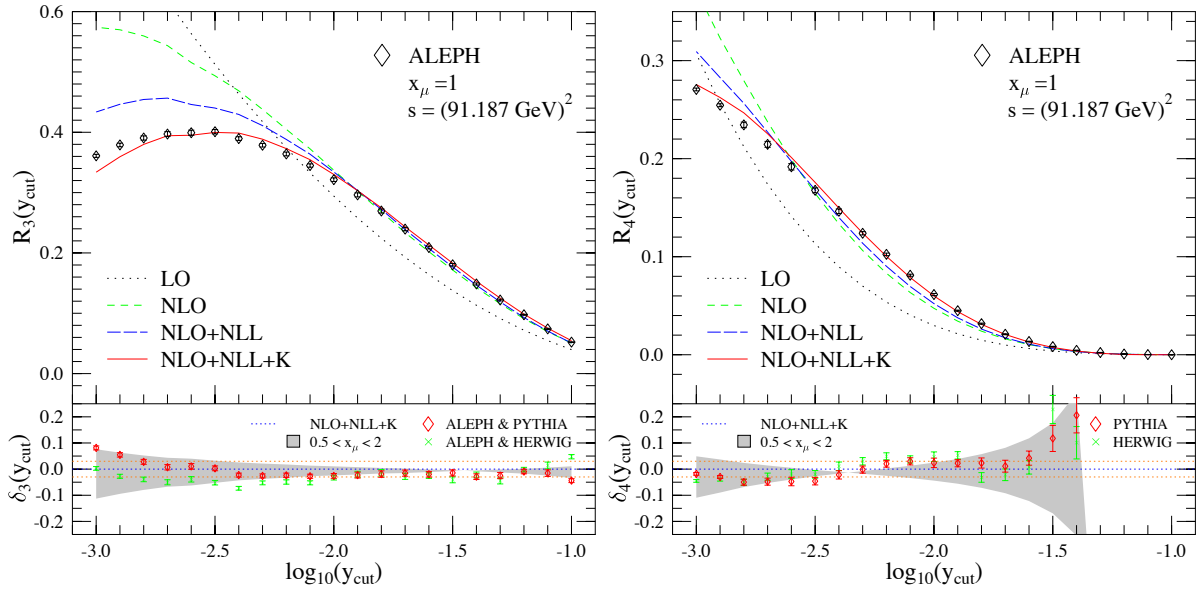
Theoreticians prefer iterative jet algorithms, consisting of the following steps. (i) First we define a *distance between two momenta* (of partons or hadron tracks) and a *rule to combine two momenta*,  $p_i$  and  $p_j$  into  $p_{(ij)}$ . (ii) Then we select a *value for resolution*  $d_{\text{cut}}$  and consider all pairs of momenta. If the minimum of  $\{d_{ij}\}$  is smaller than  $d_{\text{cut}}$ , then we combine the momenta  $p_i$  and  $p_j$  and start the algorithm again. If the minimum is larger than  $d_{\text{cut}}$ , then the remaining momenta (after the combinations) are considered the *momenta of the jets*, and the algorithm stops. The drawback of this algorithm is that it becomes very expensive computationally for many particles in the final state. This is not an issue in pQCD computations because according to our basic assumption there are only few partons, but a major problem for the final states in the detectors where hundreds of hadrons may appear in a single event.

At LEP theory won and the Durham (or  $k_{\perp}$ ) algorithm was used. It was invented so that resummation of large logarithms could be achieved [30]. The distance measure is  $d_{ij} = 2 \frac{\min(E_i^2, E_j^2)}{s} R_{ij}$ , where  $R_{ij} = 1 - \cos \theta_{ij}$  and the recombination scheme is simple addition of the four momenta,  $p_{(ij)}^{\mu} = p_i^{\mu} + p_j^{\mu}$ . The resolution parameter  $y_{\text{cut}} = d_{\text{cut}}/s$  can take values in  $[0, 1]$ . The pQCD prediction contains logarithmically enhanced terms of the form  $\alpha_s^n \ln^m(1/y_{\text{cut}})$ , at any order, which has to be resummed if we want to use small value of  $y_{\text{cut}}$ , where we find the bulk of the data (see Figure 12). Predictions are available with leading- ( $m = 2n$ ) and next-to-leading ( $2n - 1$ ) logarithms (LL and NLL) summed up to all orders [30].

Figure 12 shows the fixed order LO and NLO predictions, as well as predictions where NLO and NLL are matched.  $K$  denotes an improvement in the resummation accuracy. The curve at NLO accuracy gives a good description of the measure data by the ALEPH collaboration [31], but only for  $y_{\text{cut}} > 0.01$ . As  $\alpha_s \ln^2(100) = 2.5$ , for smaller values of  $y_{\text{cut}}$  resummation is indispensable. The resummed prediction however, is not expected to give a good description at large  $y_{\text{cut}}$  because in the resummation formula only the collinear approximation of the matrix element is used. Matching the two predictions gives a remarkably good description of the data over the whole phase space.

At hadron colliders the  $k_{\perp}$  algorithm needs modifications. First, instead of energy, the boost-invariant measure of hardness, transverse momentum is used to define the distance between tracks,  $d_{ij} = \min(p_{\perp,i}^2, p_{\perp,j}^2) \frac{R_{ij}^2}{R^2}$  where  $R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  (distance in rapidity–azimuthal-angle plane),  $R$  is a small positive real number, and we need a distance from the beam  $d_{iB} = p_{\perp,i}^{2n}$ , too. Also, the algorithm needs modification because either  $d_{ij}$  or  $d_{iB}$  can be the smallest distance. If a  $d_{ij}$  is the smallest value, then  $i$  and  $j$  are merged, while if the smallest is a  $d_{iB}$ , then momentum  $p_i$  becomes a jet momentum and is removed from the tracks to be clustered. We then call jet candidates with transverse momentum  $p_{\perp,i} > E_R$  resolved jets. The merging rule changes as well. In one merging scheme we add transverse momenta,  $p_{\perp,(ij)} = p_{\perp,i} + p_{\perp,j}$ , and we add rapidities  $y$  and azimuthal angles  $\phi$  weighted,  $y_{(ij)} = (w_i y_i + w_j y_j)/(w_i + w_j)$  and  $\phi_{(ij)} = (w_i \phi_i + w_j \phi_j)/(w_i + w_j)$ , where the weight can be  $w_i = p_{\perp,i}, p_{\perp,i}^2, E_{\perp,i}$ , or  $E_{\perp,i}^2$ . Such a merging is boost invariant along the beam. In another widely used scheme at hadron colliders one simply adds the four-momenta of the particles. The parameter  $R$  plays a similar role as  $d_{\text{cut}}$  in electron-positron annihilation or the cone radius  $R$  in the cone algorithms: the smaller  $R$ , the narrower the jet.

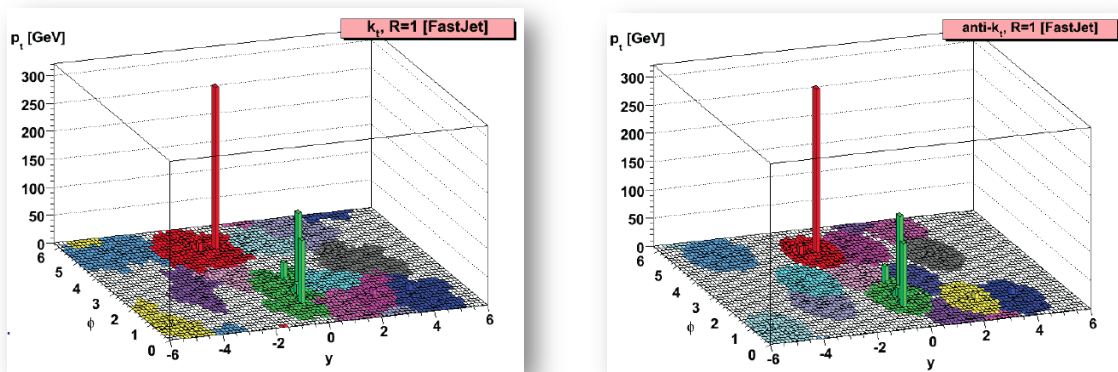
The iterative  $k_{\perp}$ -algorithm is infrared safe and resummation of large logarithmic contributions of the form  $\alpha_s^n \ln^{2n}$  and  $\alpha_s^n \ln^{2n-1}$  is possible, which is a clear advantage from the theoretical point of view. The logarithms are those of  $1/R$  and/or  $Q/E_R$ ,  $Q$  being the hard process scale. By taking  $E_R$  sufficiently large in hadron-hadron collisions, we avoid such leading contributions from initial-state showering and the underlying event, so these terms are determined by the time-like showering of final-state partons (when the virtuality of the parent parton is always positive). Particles within angular separation  $R$  tend



**Fig. 12:** Comparison of pQCD predictions to data for three- and four-jet lepton production [32]. The data points include corrections from hadrons to partons based on Monte Carlo simulations.

to combine and particles separated by larger distance than  $R$  from all other particles become jets. The algorithm assigns a clustering sequence to particles within jets, so we can look at jet substructure.

Nevertheless, at the TeVatron experiments the  $k_{\perp}$ -algorithm did not become a standard for several reasons. The jets have irregular, often weird shapes as seen on Fig. 13(a) because soft particles tend to cluster first (even arbitrary soft particles can form jets). As a result there is a non-linear dependence on soft particles, energy calibration and estimating acceptance corrections are more difficult. The underlying event correction depends on the area of the jet (in  $\eta - \phi$  plane). It was also very expensive computationally, so experimenters had a clear preference of cone algorithms.



**Fig. 13:** Jets in a proton-proton scattering event obtained with the (a)  $k_{\perp}$ , (b)  $\text{anti-}k_{\perp}$  clustering algorithms

The breakthrough occurred with Refs. [33, 34] where variants of the  $k_{\perp}$  algorithm and an improved, fast implementation was introduced. The distance formula was modified to  $d_{ij} = \min(p_{\perp,i}^{2n}, p_{\perp,j}^{2n}) \frac{R_{ij}^2}{R^2}$

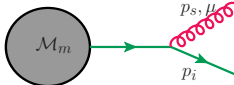
( $n = -1, 0, 1$ ). IR-safety is independent of  $n$ , as well as NLL resummation of large logarithms. It was found that with  $n = -1$  (called anti- $k_{\perp}$ -algorithm) particles close in angle cluster first, which results in regular cone-like shapes as seen on Fig. 13(b) without using stable cones. As a result it became the standard jet algorithm at the LHC experiments. Yet, one should keep in mind that there is no ‘perfect’ jet algorithm. For instance, the anti- $k_{\perp}$  one does not provide useful information on jet substructure. It is important to remember that *in pQCD theoretical prediction can be made only with IR-safe jet functions, but among those the goal of the study may help decide which algorithms to use.*

#### 4 Towards a general method for computing QCD radiative corrections

We have seen that (i) in pQCD the computation of radiative corrections at NLO accuracy is indispensable, (ii) the NLO corrections are of two kinds: real and virtual, that are separately divergent and contain different number of particles in the final state, (iii) these singularities cancel for IR-safe cross sections. To find the finite NLO corrections we have to develop a method for combining the real and virtual corrections. In order to be able to automate the NLO computations such a method has to be general, i.e., independent of the measurable quantity and the process. To devise such a general method, we need to study the origin of the singularities in a more precise way than we did in the previous section. We shall find factorization formulae of the SME’s that find many important applications in QCD, and so belong to the most important features of QCD.

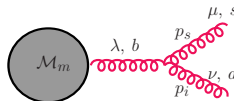
##### 4.1 Factorization of $|\mathcal{M}_n|^2$ in the soft limit

The soft limit is defined by  $p_s^\mu = \lambda q^\mu$ , with  $\lambda \in \mathbb{R}^+$  and  $\lambda \rightarrow 0$  for  $q^\mu$  fixed. In this limit the emission of the soft gluon from (internal) propagators is IR finite. If we consider the emission of a soft gluon off an external quark we find



$$\propto \mathcal{M}_m T_i^s g_s \bar{u}(p_i, s_i) \gamma^\mu \frac{\not{p}_i + \not{p}_s}{s_{is}} \xrightarrow{p_s \rightarrow 0} \mathcal{M}_m T_i^s g_s \frac{p_i^\mu}{p_i \cdot p_s} \bar{u}(p_i, s_i).$$

In taking the limit, we used the anti-commutation relation (4) to write  $\gamma^\mu \not{p}_i = -\not{p}_i \gamma^\mu + 2p_i^\mu$  and the Dirac equation of the massless bi-spinor,  $\bar{u}(p_i) \not{p}_i = 0$ . The factor  $\frac{p_i^\mu}{p_i \cdot p_s}$  is the ‘‘square root’’ of the eikonal factor  $S_{ik}(s) = \frac{2s_{ik}}{s_{is}s_{ks}}$ . In the same limit, we can derive after a bit more algebra the factorization formula for soft-gluon emission off a gluon line. The emission of a soft gluon off an external gluon (in light-cone gauge) is given by



$$\propto \mathcal{M}_m \varepsilon^\mu(p_s, n) \frac{1}{s_{is}} d^{\lambda\lambda'}(p_i + p_s, n) \Gamma_{\nu\mu\lambda'}^{asb}(-p_i, -p_s, p_i + p_s) \varepsilon^\nu(p_i, n),$$

where in the three-gluon vertex

$$\begin{aligned} V_{\nu\mu\lambda}(-p_i, -p_s, p_i + p_s) &= -(p_i + 2p_s)_\nu g_{\mu\lambda} + (2p_i + p_s)_\mu g_{\nu\lambda} - (p_i - p_s)_\lambda g_{\mu\nu} \\ &= 2p_{i\mu} g_{\lambda\nu} + [-(p_i + p_s)_\lambda g_{\mu\nu} - p_{i\nu} g_{\mu\lambda}] + [p_{s\mu} g_{\nu\lambda} + 2p_{s\lambda} g_{\mu\nu} - 2p_{s\nu} g_{\mu\lambda}] \\ &\xrightarrow{p_s \rightarrow 0} 2p_{i\mu} g_{\nu\lambda} - [(p_i + p_s)_\lambda g_{\mu\nu} + p_{i\nu} g_{\mu\lambda}]. \end{aligned}$$

We use  $d^{\lambda\lambda'}(p_i + p_s, n) (p_i + p_s)_\lambda = 0$  and  $\varepsilon^\nu(p_i, n) p_{i\nu} = 0$ , thus

$$\frac{1}{s_{is}} d^{\lambda\lambda'}(p_i + p_s, n) \Gamma_{\nu\mu\lambda'}^{asb}(-p_i, -p_s, p_i + p_s) \varepsilon^\nu(p_i, n) \xrightarrow{p_s \rightarrow 0} -T_b^s g_s \frac{p_{i\mu}}{p_i \cdot p_s} \underbrace{\left[ d^{\lambda\lambda'}(p_i, n) g_{\lambda'\nu} \varepsilon^\nu(p_i, n) \right]}_{-\varepsilon^\lambda(p_i, n)}.$$

These two results can be unified and formalized by

$$\hat{S}_s \langle c_s | \mathcal{M}_{m+1}(p_s, \dots) \rangle = g_s \varepsilon^\mu(p_s) \mathbf{J}_\mu(s) | \mathcal{M}_m(\dots) \rangle,$$



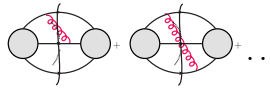
where  $c_s$  is the colour index of the soft gluon  $s$ ,  $\hat{S}_s$  is an operator which takes the soft limit and keeps the leading  $\frac{1}{\lambda}$  singular term, and the soft gluon current  $\mathbf{J}_\mu(s)$  is given by

$$\mathbf{J}_\mu(s) = \sum_{k=1}^m \mathbf{T}_k^s \frac{p_{k\mu}}{p_k \cdot p_s}.$$

The soft gluon can be emitted from any of the external legs, therefore the sum in the previous formula runs over all external partons. A soft quark leads to an integrable singularity because the fermion propagator is less singular than that of the gluon. Colour conservation implies that the current  $\mathbf{J}_\mu(s)$  is conserved,

$$p_s^\mu \mathbf{J}_\mu(s) |\mathcal{M}_m\rangle = \sum_{k=1}^m \mathbf{T}_k^s |\mathcal{M}_m\rangle =$$


Then the soft limit of the SME  $\langle \mathcal{M}_m^{(0)} | \mathcal{M}_m^{(0)} \rangle$  is as follows:

$$\begin{aligned} \hat{S}_s |\mathcal{M}_{m+1}(p_s, \dots)|^2 &= 4\pi\alpha_s \mu^{2\epsilon} \sum_{i=1}^m \sum_{k=1}^m \underbrace{\epsilon_\mu(s) \epsilon_\nu^*(s)}_{d_{\mu\nu}(p_s, n)} \frac{p_i^\mu p_k^\nu}{p_i \cdot p_s p_k \cdot p_s} \langle \mathcal{M}_m | \mathbf{T}_i \cdot \mathbf{T}_k | \mathcal{M}_m \rangle \\ &= -8\pi\alpha_s \mu^{2\epsilon} \sum_{i,k=1}^m \frac{1}{2} S_{ik}(s) \left| \mathcal{M}_{m(i,k)}^{(0)} \right|^2 + \text{gauge terms} = \end{aligned} \quad (57)$$


The gauge terms give zero contribution on on-shell matrix elements due to gauge invariance.

## 4.2 Factorization of $|\mathcal{M}_n|^2$ in the collinear limit

The collinear limit of momenta  $p_i$  and  $p_r$  is defined by Sudakov parametrization:

$$p_i^\mu = z_i p^\mu + k_{i\perp}^\mu - \frac{k_{i\perp}^2}{z_i} \frac{n^\mu}{2p \cdot n}, \quad p_r^\mu = z_r p^\mu + k_{r\perp}^\mu - \frac{k_{r\perp}^2}{z_r} \frac{n^\mu}{2p \cdot n}$$

where  $k_{i\perp}^\mu + k_{r\perp}^\mu = 0$  and  $z_i + z_r = 1$ . The momentum  $p^\mu$  is the collinear direction and

$$p^2 = p_i^2 = p_r^2 = n^2 = 0, \quad k_{i\perp} \cdot p = k_{r\perp} \cdot n = 0,$$

In the collinear limit  $k_{i\perp}^\mu, k_{r\perp}^\mu \rightarrow 0$  and  $s_{ir} = -\frac{k_{r\perp}^2}{z_i z_r}$ . We now state the following theorem

*In a physical gauge, the leading collinear singularities are due to the collinear splitting of an external parton.*

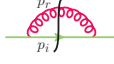
This means that we need to compute  in the collinear limit. There are three cases:

$$\begin{array}{lcl} f_{ir} & \rightarrow & f_i + f_r \\ q & \rightarrow & q + g \\ g & \rightarrow & q + \bar{q} \\ g & \rightarrow & g + g \end{array}$$

We compute explicitly the first case and leave the second and the third as exercise.

For the case of a quark splitting into a quark and a gluon we have

$$\begin{aligned} &= C_F g_s^2 \mu^{2\epsilon} \frac{\not{p}_i + \not{p}_r}{s_{ir}} \gamma^\mu \not{p}_i \gamma^\nu d_{\mu\nu}(p_r, n) \frac{\not{p}_i + \not{p}_r}{s_{ir}} \\ &= C_F 4\pi\alpha_s \mu^{2\epsilon} \frac{\not{p}_i + \not{p}_r}{s_{ir}} \left( -\gamma^\mu \not{p}_i \gamma_\mu + \frac{\not{p}_r \not{p}_i \not{n} + \not{n} \not{p}_i \not{p}_r}{p_r \cdot n} \right) \frac{\not{p}_i + \not{p}_r}{s_{ir}}. \end{aligned} \quad (58)$$



Using

$$-\gamma^\mu \not{p}_i \gamma_\mu = (d-2) \not{p}_i, \quad \not{p}_i \not{p}_i = p_i^2 \mathbb{1}, \quad \text{and} \quad \not{p}_i \not{p}_r \not{p}_i = s_{ir} \not{p}_i - p_i^2 \not{p}_r = s_{ir} \not{p}_i,$$

we find

$$\begin{aligned} (\not{p}_i + \not{p}_r) (-\gamma^\mu \not{p}_i \gamma_\mu) (\not{p}_i + \not{p}_r) &= (d-2) s_{ir} \not{p}_r, \\ \not{p}_r \not{p}_i \not{\eta} &= -\not{p}_i \not{p}_r \not{\eta} + s_{ir} \not{\eta} = \not{p}_i \not{\eta} \not{p}_r - 2 \not{p}_i p_r \cdot n + s_{ir} \not{\eta} = -\not{\eta} \not{p}_i \not{p}_r + 2 p_i \cdot n \not{p}_r - 2 \not{p}_i p_r \cdot n + s_{ir} \not{\eta}. \end{aligned}$$

Then

$$\begin{aligned} (\not{p}_i + \not{p}_r) (\not{p}_r \not{p}_i \not{\eta} + \not{\eta} \not{p}_i \not{p}_r) (\not{p}_i + \not{p}_r) &= \\ &= 2 (\not{p}_i + \not{p}_r) (p_i \cdot n \not{p}_r - p_r \cdot n \not{p}_i + p_i \cdot p_r \not{\eta}) (\not{p}_i + \not{p}_r) \\ &= 2 \left[ p_i \cdot n s_{ir} \not{p}_i - p_r \cdot n s_{ir} \not{p}_r + p_i \cdot p_r \left( 2 (p_i + p_r) \cdot n (\not{p}_i + \not{p}_r) - (p_i + p_r)^2 \not{\eta} \right) \right] \\ &= s_{ir} \left[ 4 p_i \cdot n \not{p}_i + 2 p_i \cdot n \not{p}_r + 2 p_r \cdot n \not{p}_i - s_{ir} \not{\eta} \right]. \end{aligned}$$

Substituting these results and then the Sudakov parametrization of the momenta into Eq. (58) we obtain

$$\begin{aligned} &\stackrel{p_i \parallel p_r}{\simeq} \frac{1}{s_{ir}} C_F 4\pi\alpha_s \mu^{2\epsilon} \left[ 2(1-\epsilon) z_r + 4 \frac{z_i^2}{z_r} + 4z_i + \mathcal{O}(k_\perp) \right] \not{p} \\ &= \frac{1}{s_{ir}} C_F 8\pi\alpha_s \mu^{2\epsilon} \left[ 2 \frac{z_i}{z_r} + (1-\epsilon) z_r \right] \not{p} = \frac{1}{s_{ir}} C_F 8\pi\alpha_s \mu^{2\epsilon} \left[ \frac{1+z_i^2}{1-z_i} - \epsilon(1-z_i) \right] \not{p} \end{aligned}$$

Similarly to the soft case we can define an operator  $\hat{C}_{ir}$  which takes the collinear limit and keeps the leading singular ( $\mathcal{O}(1/k_\perp^2)$ ) terms:

$$\hat{C}_{ir} \left| \mathcal{M}_{m+1}^{(0)} \right|^2 = 8\pi\alpha_s \mu^{2\epsilon} \frac{1}{s_{ir}} \left\langle \mathcal{M}_m^{(0)}(p, \dots) \left| \hat{P}_{qg}^{(0)}(z_i, z_r, k_\perp; \epsilon) \right| \mathcal{M}_m^{(0)}(p, \dots) \right\rangle. \quad (59)$$

The kernel  $\hat{P}_{qg}$ , called Altarelli-Parisi splitting function for the process  $q \rightarrow q + g$ , is diagonal in the spin-state of the parent (splitting) parton:

$$\langle s | \hat{P}_{qg} | s' \rangle = C_F \left[ 2 \frac{z_i}{z_r} + (1-\epsilon) z_r \right] \delta_{ss'}.$$

Similar calculations give the splitting kernels for the gluon splitting processes, which however, contain azimuthal correlations of the parent parton

$$\langle \mu | \hat{P}_{q\bar{q}}^{(0)}(z_i, z_r, k_\perp; \epsilon) | \nu \rangle = T_R \left[ -g^{\mu\nu} + 4z_i z_r \frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2} \right] \quad (60)$$

$$\langle \mu | \hat{P}_{gg}^{(0)}(z_i, z_r, k_\perp; \epsilon) | \nu \rangle = 2C_A \left[ -g^{\mu\nu} \left( \frac{z_i}{z_r} + \frac{z_r}{z_i} \right) - 2(1-\epsilon) z_i z_r \frac{k_\perp^\mu k_\perp^\nu}{k_\perp^2} \right]. \quad (61)$$

The soft and collinear limits overlap when the soft gluon is also collinear to its parent parton:

$$\hat{C}_{jr} \hat{S}_r \left| \mathcal{M}_{m+1}^{(0)}(p_r, \dots) \right|^2 = -8\pi\alpha_s \mu^{2\epsilon} \sum_{k \neq j} \frac{2z_j}{s_{jr} z_r} \left| \mathcal{M}_{m(j,k)}^{(0)}(\dots) \right|^2 = 8\pi\alpha_s \mu^{2\epsilon} \mathbf{T}_j^2 \frac{2}{s_{jr}} \frac{z_j}{z_r} \left| \mathcal{M}_m^{(0)} \right|^2.$$

The notation for the splitting kernels in these lectures is different from the usual notation in the literature. Usually,  $\hat{P}_{ij}^{(0)}(z, k_\perp; \epsilon)$  denotes the splitting kernel for the process  $f_i(p) \rightarrow f_j(zp) + f_k((1-z)p)$ , which does not lead to confusion for  $1 \rightarrow 2$  splittings because the momentum fraction of parton  $j$  determines that of parton  $k$  as their sum has to be one. For splittings involving more partons, it is more appropriate to introduce as many momentum fractions  $z_i$  as the number of offspring partons, with the constraint  $\sum_i z_i = 1$ , and use the flavour indices to denote the offspring partons in the order of the momentum fractions in the argument. For  $1 \rightarrow 2$  splittings this means the use of  $\hat{P}_{ir}^{(0)}(z_i, z_r, k_\perp; \epsilon)$  for the splitting process  $f_k(p) \rightarrow f_i(z_i p) + f_r(z_r p)$ . The flavour of the parent parton  $f_k$  is determined uniquely by the flavour summation rules,  $q + g = q$ ,  $q + \bar{q} = g + g = g$ . These flavour summation rules are unique also for  $1 \rightarrow 3$  splittings.

**Exercise 4.1** Compute the Altarelli-Parisi-splitting function  $\hat{P}_{qg}(z)$  for the process  $q \rightarrow qg$  from the collinear limit of the matrix element for the process  $e^+e^- \rightarrow q\bar{q}g$ :

$$|\mathcal{M}(e^+e^- \rightarrow q\bar{q}g)|^2 \propto \left( (1-\epsilon) \left( \frac{y_{23}}{y_{13}} + \frac{y_{13}}{y_{23}} \right) + 2 \left( \frac{y_{12}}{y_{13}y_{23}} - \epsilon \right) \right).$$

**Exercise 4.2** The Altarelli-Parisi splitting function  $\hat{P}_{q\bar{q}}(z)$  for the process  $g \rightarrow q\bar{q}$  is defined by the following collinear limit:

$$\begin{aligned} \langle \mathcal{M}_{n+1}^{(0)}(p_i, p_r, \dots) | \mathcal{M}_{n+1}^{(0)}(p_i, p_r, \dots) \rangle &\stackrel{p_i \parallel p_r}{\simeq} \frac{1}{s_{ir}} 8\pi\alpha_s \mu^{2\epsilon} \langle \mathcal{M}_n^{(0)}(p, \dots) | \hat{P}_{q\bar{q}}^{(0)}(z, k_\perp) | \mathcal{M}_n^{(0)}(p, \dots) \rangle \\ &= \frac{1}{s_{ir}} 8\pi\alpha_s \mu^{2\epsilon} \langle \mathcal{M}_n^{(0)}(p, \dots) | \mu \rangle \langle \mu | \hat{P}_{q\bar{q}}^{(0)}(z, k_\perp) | \nu \rangle \langle \nu | \mathcal{M}_n^{(0)}(p, \dots) \rangle \\ &= \frac{1}{s_{ir}} 8\pi\alpha_s \mu^{2\epsilon} \langle \mathcal{M}_n^{(0)}(p, \dots) | \mu \rangle \frac{d_{\mu\rho}}{s_{ir}} \Pi^{\rho\sigma} \frac{d_{\sigma\nu}}{s_{ir}} \langle \nu | \mathcal{M}_n^{(0)}(p, \dots) \rangle. \end{aligned}$$

Compute  $\langle \mu | \hat{P}_{q\bar{q}}^{(0)}(z, k_\perp) | \nu \rangle$  in leading order in  $k_\perp$ . Hint: In which sense does  $\Pi_{\mu\nu} = d_{\mu\rho} \Pi^{\rho\sigma} d_{\sigma\nu}$  hold?

**Exercise 4.3** Derive the flavour summation rules for  $1 \rightarrow 3$  splittings.

**Exercise 4.4** Compute the soft limit of Eq. (59) and the collinear limit of Eq. (57).

### 4.3 Regularization of real corrections by subtraction

The cross section at NLO accuracy is a sum of two terms, the LO prediction and the corrections at one order higher in the strong coupling,

$$\sigma_{\text{NLO}} = \sigma^{\text{LO}} + \sigma^{\text{NLO}},$$

where  $\sigma^{\text{LO}}$  is the integral of the fully differential Born cross section over the available phase space defined by the jet function, while  $\sigma^{\text{NLO}}$  is the sum of the real and virtual corrections:

$$\sigma^{\text{LO}} = \int_m d\sigma^B J_m(\{p\}_m), \quad \sigma^{\text{NLO}} = \int_{m+1} d\sigma^R J_m(\{p\}_{m+1}) + \int_m d\sigma^V J_m(\{p\}_m).$$

Both contributions to  $\sigma^{\text{NLO}}$  are divergent in four dimensions, but their sum is finite for IR-safe jet functions.

The factorization of the squared matrix elements in the soft and collinear limits allows for a process and observable independent method to regularize the real corrections in their singular limits. The essence of the method is to devise an approximate cross section  $d\sigma^A$  that matches the singular behaviour of the real cross section  $d\sigma^R$  in all kinematically degenerate regions of the phase space when one parton becomes soft or two partons become collinear. Then we subtract this approximate cross section from the real one and the difference can be integrated in four dimensions. Next, we integrate  $d\sigma^A$  over the phase space of the unresolved parton and we add it to  $d\sigma^V$ . The integrated subtraction term cancels the explicit poles in the virtual correction and the sum can also be integrated in four dimensions. The key for this procedure is a proper mapping of the  $(m + 1)$ -parton phase space to the  $m$ -parton one which respects the limits, thus the approximate cross section is defined with the  $m$ -parton jet function. This way we can rewrite the NLO correction as a sum of two finite terms,

$$\sigma^{\text{NLO}} = \int_{m+1} [d\sigma^R J_m(\{p\}_{m+1}) - d\sigma^A J_m(\{\tilde{p}\}_m)]_{\epsilon=0} + \int_m \left[ d\sigma^V + \int_1 d\sigma^A \right]_{\epsilon=0} J_m(\{p\}_m). \quad (62)$$

The definition of the approximate cross section is not unique and the best choice may depend on further requirements that we do not discuss here. We also skip the precise definition of the momenta  $\tilde{p}^\mu$  which are obtained by mapping the  $(m + 1)$ -particle phase space onto an  $m$ -particle phase space times a one-particle phase space. A widely used general subtraction scheme that can be used also for processes including massive partons with smooth massless limits is presented in Ref. [35], where these definitions are given explicitly. This method uses the factorization of the SME in the soft and collinear limits. The challenge posed by the overlapping singularity in the soft-collinear limit is solved by a smooth interpolation between these singular regions.

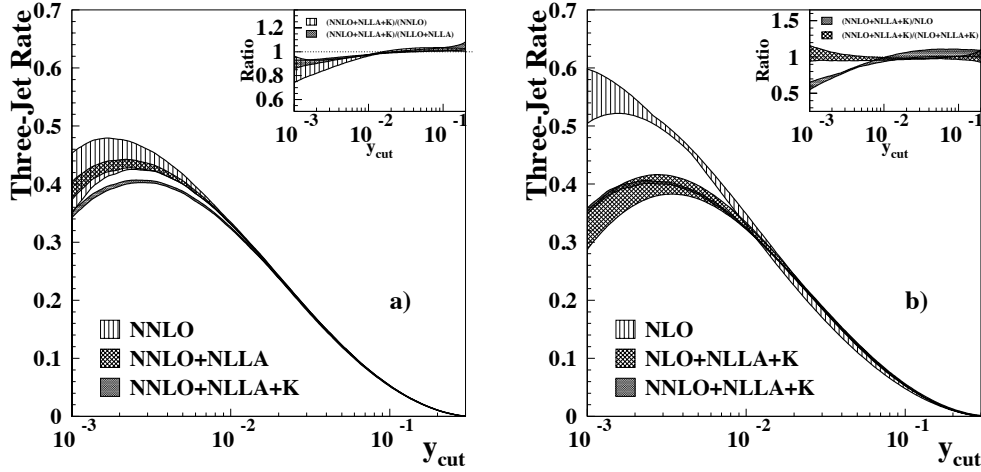
The factorization properties of Eqs. (57) and (59) play other very important roles in pQCD. The numerical implementation of the SME is in general a process prone to errors. Testing the factorization in the kinematically degenerate phase space regions serves a good check of the implementation. The computation is even more difficult for the virtual corrections. Similar factorization holds for those. The factorized form of the SME can be used in resumming logarithmically enhanced terms at all orders, or in devising a parton shower algorithm for modelling events (see Sect. 6.3). The splitting kernels that appear in the collinear factorization have a role in the evolution equations of the parton distribution functions (see Sect. 5.6).

The state of the art in making precision predictions assaults on the one hand the full automation of computations at NLO, and on the other the realm of next-to-next-to-leading order (NNLO) corrections. The automation of computing jet cross sections at NLO accuracy has been accomplished and several programs are available with the aim to facilitate automated solutions for computing jet cross sections at NLO accuracy:

- aMC@NLO (<http://amcatnlo.web.cern.ch>)
- BlackHat/Sherpa (<https://blackhat.hepforge.org>)
- FeynArts/FormCalc/LoopTools (<http://www.feynarts.de>)
- GoSam (<https://gosam.hepforge.org>)
- HELAC-NLO (<http://helac-phegas.web.cern.ch>)
- MadGolem (<http://www.thphys.uni-heidelberg.de/lopez/madgolem-corner.html>).

In the NNLO case the IR singularity structure is much more involved than in the case of NLO computations due to complicated overlapping singly- and doubly-unresolved configurations. Several subtraction methods have been proposed for the regularization of the IR divergences and there is intense research to find a general one that can be automated. To provide an impression about the importance of

NNLO corrections, we present QCD predictions at various accuracies for the three-jet rate computed with  $\alpha_s = 0.118$  and at a centre-of-mass energy of  $\sqrt{s} = 35$  GeV in Fig. 14. Figure 14(a) shows comparison of prediction at NLO with that at matched NLO and resummed next-to-leading logarithmic (denoted by NLLA in the figure) accuracy, while Fig. 14(b) presents comparison of prediction at NNLO with that at matched NNLO and resummed NLL accuracy. The inserts in both cases show the ratio between the matched and the unmatched predictions. For all calculations the uncertainty band reflects the uncertainty due to the variation of the renormalization scale around the default scale  $\mu = \sqrt{s}$  by factors of 2 in both directions.



**Fig. 14:** QCD predictions for the three-jet rate in electron-positron annihilation [36]

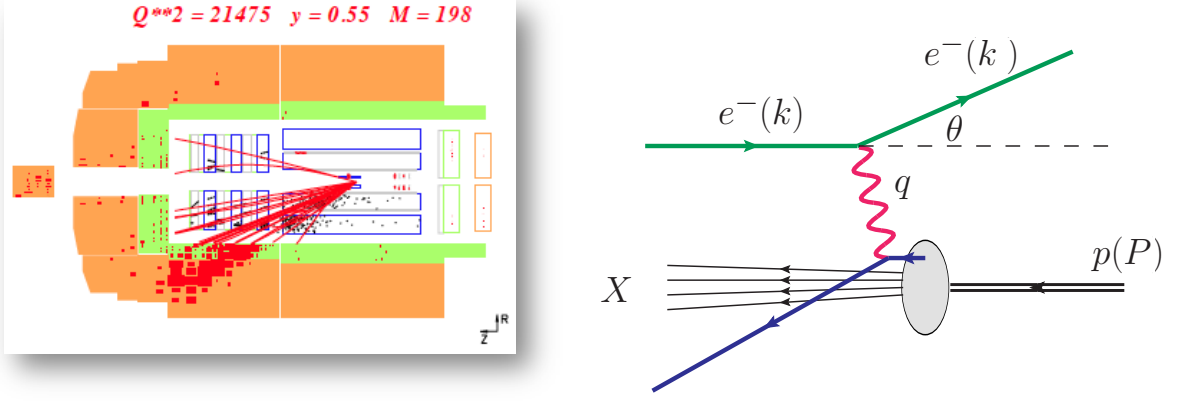
## 5 Deeply inelastic lepton-proton scattering

Perturbative QCD stems from the parton model that was developed to understand deeply inelastic lepton-hadron scattering (DIS). The purpose of those experiments was to study the structure of the proton by measuring the kinematics of the scattered lepton. In Fig. 15(a) we show a real event in the H1 experiment at the HERA collider. The value of  $Q^2$ , which is the modulus squared of the momentum transfer between the lepton and the proton is  $21475 \text{ GeV}^2 \gg 1 \text{ GeV}^2$ , signifying that the scattering is well in the deeply inelastic region. The parton model interpretation of the event is shown in Fig. 15(b): the lepton is scattered by an angle  $\theta$  due to the exchange of a virtual photon with one of the constituents of the proton (a parton). The measurement is inclusive from the point of view of hadrons ( $X$  means any number of hadrons that are not observed separately), thus the process can be described in pQCD.

The DIS kinematics is described by the following variables

$$\begin{aligned}
 \text{centre-of-mass energy}^2 &= s = (P + k)^2, \\
 \text{momentum transfer} &= q^\mu = k^\mu - k'^\mu, \\
 |\text{momentum transfer}|^2 &= Q^2 = -q^2 = 2MExy, \\
 \text{scaling variable} &= x = Q^2/(2P \cdot q), \\
 \text{energy loss} &= \nu = (P \cdot q)/M = E - E', \\
 \text{relative energy loss} &= y = (P \cdot q)/(P \cdot k) = 1 - E'/E, \\
 \text{recoil mass}^2 &= W^2 = (P + q)^2 = M^2 + \frac{1-x}{x}Q^2,
 \end{aligned}$$

where we set the more important ones for these lectures in red.



**Fig. 15:** Deeply inelastic lepton-proton scattering (a) in the H1 detector and (b) parton model interpretation of such an event

### 5.1 Parametrization of the target structure

The cross section for  $e(k) + p(P) \rightarrow e(k') + X$  reads

$$d\sigma = \sum_X \frac{1}{4ME} \int d\phi \frac{1}{4} \sum_{\text{spin}} |\mathcal{M}|^2. \quad (63)$$

We factorize the phase space and the SME into two parts, one for the lepton and one for the hadrons:

$$d\phi = \frac{d^3k'}{(2\pi)^3 2E'} d\phi_X, \quad \frac{1}{4} \sum_{\text{spin}} |\mathcal{M}|^2 = \frac{e^4}{Q^4} L^{\mu\nu} H_{\mu\nu}.$$

Then the hadron part of the cross section is the dimensionless Lorentz tensor  $W_{\mu\nu} = \frac{1}{8\pi} \sum_X \int d\phi_X H_{\mu\nu}$  (the factor of  $\frac{1}{8\pi}$  is included here by convention). As it depends on two momenta  $P^\mu$  and  $q^\mu$ , the most general gauge invariant combination of the Lorentz tensor can be written as

$$W_{\mu\nu}(P, q) = \left( -g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2} \right) W_1(x, Q^2) + \left( P_\mu - q_\mu \frac{P \cdot q}{q^2} \right) \left( P_\nu - q_\nu \frac{P \cdot q}{q^2} \right) \frac{W_2(x, Q^2)}{P \cdot q},$$

where the structure functions  $W_i(x, Q^2)$  are dimensionless functions of the scaling variable and the momentum transfer.

For the lepton part we express the kinematical relations  $E' = (1 - y)E$ ,  $\cos \vartheta = 1 - \frac{xyM}{(1-y)E}$  to change variables to scaling variable and relative energy loss:

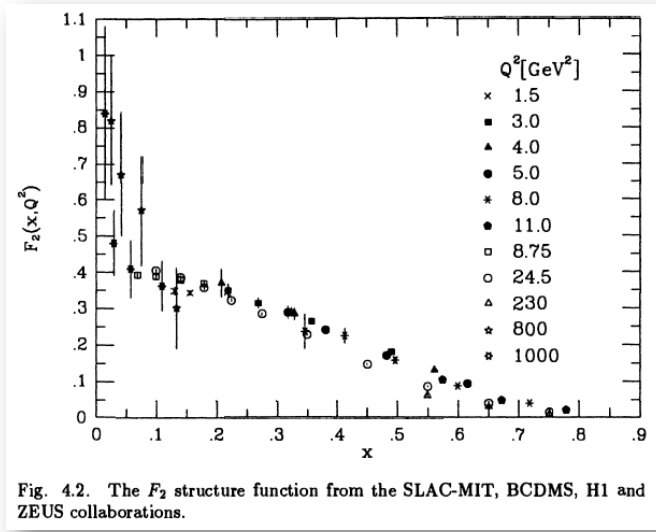
$$\frac{d^3k'}{(2\pi)^3 2E'} = \frac{d\varphi}{2\pi} \frac{E'}{8\pi^2} dE' d\cos \vartheta = \frac{d\varphi}{2\pi} \frac{yME}{8\pi^2} dy dx,$$

and compute the trace  $L^{\mu\nu} = \frac{1}{2} \text{Tr}[k \gamma^\mu k' \gamma^\nu] = k^\mu k'^\nu + k^\nu k'^\mu - g^{\mu\nu} k \cdot k'$ . Then the differential cross section in  $x$  and  $y$  is obtained from Eq. (63) as

$$\frac{d^2\sigma}{dx dy} = \frac{4\pi\alpha^2}{yQ^2} \left[ y^2 W_1(x, Q^2) + \left( \frac{1-y}{x} - xy \frac{M^2}{Q^2} \right) W_2(x, Q^2) \right],$$

which we rewrite in the scaling limit, defined by  $Q^2 \rightarrow \infty$  with  $x$  fixed, as

$$\frac{d^2\sigma}{dx dy} = \frac{4\pi\alpha^2}{yQ^2} \left[ (1 + (1-y)^2) F_1 + \frac{1-y}{x} (F_2 - 2xF_1) \right]. \quad (64)$$



**Fig. 16:** Measured value of the  $F_2$  structure function at several different values of  $Q^2$

section is

$$d\hat{\sigma} = \frac{1}{2\hat{s}} \int d\phi_2 \frac{1}{4} \sum_{\text{spin}} |\mathcal{M}|^2,$$

with  $\hat{s} = (p + k)^2$ . The SME is proportional to the product of the lepton tensor  $L^{\mu\nu}$  and a similar quark tensor  $Q_{\mu\nu} = \frac{1}{2} \text{Tr}[\not{q}\gamma^\mu \not{q}'\gamma^\nu] = q^\mu q'^\nu + q^\nu q'^\mu - g^{\mu\nu} q \cdot q'$ , i.e.,  $L^{\mu\nu} Q_{\mu\nu} = 2(\hat{s}^2 + \hat{u}^2)$ , where  $\hat{u} = (p - k')^2 = -2p \cdot k'$ . As  $y = P \cdot q / P \cdot k = 2p \cdot q / 2p \cdot k = (\hat{s} + \hat{u}) / \hat{s}$ , momentum conservation,  $p'_\mu = p_\mu + q_\mu$ , implies for the on-shell condition of the scattered quark  $0 = p'^2 = (p + q)^2 = 2p \cdot q + q^2 = \hat{s} + \hat{u} - Q^2$ . We have  $y = Q^2 / \hat{s}$  and  $\hat{u} = (y - 1)\hat{s}$ , so

$$\frac{1}{4} \sum_{\text{spin}} |\mathcal{M}|^2 = \frac{e_q^2 e^4}{Q^4} L^{\mu\nu} Q_{\mu\nu} = 2e_q^2 e^4 \frac{\hat{s}^2}{Q^4} (1 + (1 - y)^2).$$

Also  $Q^2 = 2p \cdot q = 2\xi P \cdot q$ , so  $p'^2 = Q^2(\xi/x - 1)$ . Then the two-particle phase space is

$$d\phi_2 = \frac{d^3 k'}{(2\pi)^3 2E_{k'}} \frac{d^4 p'}{(2\pi)^4} 2\pi \delta_+(p'^2) (2\pi)^4 \delta^4(k + p - k' - p') = \frac{d\varphi}{2\pi} \frac{E'}{4\pi} dE' d\cos\vartheta \frac{x}{Q^2} \delta(\xi - x),$$

or using  $E' = \frac{\sqrt{\hat{s}}}{2}(1 - y)$  and  $\cos\vartheta = 1 - \frac{2yx}{\xi(1 - y)}$ , we obtain  $d\phi_2 = \frac{d\varphi}{(4\pi)^2} \frac{y\hat{s}}{Q^2} dy dx \delta(\xi - x)$ . The differential cross section in  $x$  and  $y$

$$\frac{d^2 \hat{\sigma}}{dx dy} = \frac{4\pi\alpha^2}{Q^2} [1 + (1 - y)^2] \frac{1}{2} e_q^2 \delta(\xi - x). \quad (65)$$

Comparing Eqs. (64) and (65), we find the parton model predictions

$$F_1(x) \propto e_q^2 \delta(\xi - x), \quad F_2 - 2xF_1 = 0, \quad \text{called Callan-Gross relation.} \quad (66)$$

Thus  $F_2$  probes the quark constituent of the proton with  $\xi = x$ . However, this prediction for  $F_2$  cannot be correct because  $F_2(x)$  is not a  $\delta$  function as seen from Fig. 16, which leads us to formulate the naïve parton model in the following way:

The dimensionless functions  $F_1$  and  $F_2$  were first measured by the SLAC-MIT experiment [37]. The result of that measurement supplemented by some later ones is shown in Fig. 16. The interesting feature is that in the scaling limit  $F_2$  becomes independent of  $Q^2$ ,  $F_2(x, Q^2) \rightarrow F_2(x)$  (in fact, the independence starts at quite low values of  $Q^2$ ).

## 5.2 DIS in the parton model

Let us now describe the same scattering process by assuming the proton is a bunch of free flying quarks and the lepton exchanges a hard virtual photon with one of those quarks as shown in Fig. 15(b). The struck quark carries a momentum  $p^\mu$ , which is a fraction of the proton momentum,  $p^\mu = \xi P^\mu$ , so we consider the process  $e(k) + q(p) \rightarrow e(k') + q(p')$ . The corresponding cross

the virtual photon scatters incoherently off the constituents (partons) of the proton; the probability that a quark  $q$  carries momentum fraction of the proton between  $\xi$  and  $\xi + \delta\xi$  is  $f_q(\xi)d\xi$ .

**Exercise 5.1** Compute the contribution to the DIS cross section in Eq. (64) with the exchange of a transversely polarized photon. Hint: Use Eq. (10) for the numerator in the propagator of the transversely polarized photon and the Callan-Gross relation in Eq. (66). Can you identify the result with any of the terms in Eq. (64)? What is the source of the remainder?

### 5.3 Measuring the proton structure

With the assumptions of the naïve parton model the Callan-Gross relation predicts

$$F_2(x) = 2xF_1(x) = \sum_q \int_0^1 d\xi f_q(\xi) x e_q^2 \delta(x - \xi) = x \sum_q e_q^2 f_q(x). \quad (67)$$

Taking into account four flavours and simplifying the notation by using  $f_q(x) \equiv q(x)$ , we obtain a prediction for the structure function measured in scattering of charged-lepton off proton (neutral current interaction):

$$F_2^{\text{em}}(x) = x \left[ \frac{4}{9} (u(x) + \bar{u}(x) + c(x) + \bar{c}(x)) + \frac{1}{9} (d(x) + \bar{d}(x) + s(x) + \bar{s}(x)) \right].$$

Similarly, in charged current interactions the prediction is

$$F_2^{\nu}(x) = 2x [u(x) + \bar{d}(x) + c(x) + \bar{s}(x)] \quad (\text{with } W^-), \quad F_2^{\nu}(x) = 2x [d(x) + \bar{u}(x) + s(x) + \bar{c}(x)] \quad (\text{with } W^+).$$

Further information can be obtained if we use different targets. Assuming two flavours and isospin symmetry, the proton (with uud valence quarks) structure is

$$F_2^{\text{proton}}(x) = x \left[ \frac{4}{9} (u_p(x) + \bar{u}_p(x)) + \frac{1}{9} (d_p(x) + \bar{d}_p(x)) \right], \quad (68)$$

and that of the neutron (with udd valence quarks) is

$$F_2^{\text{neutron}}(x) = x \left[ \frac{4}{9} (u_n(x) + \bar{u}_n(x)) + \frac{1}{9} (d_n(x) + \bar{d}_n(x)) \right] = x \left[ \frac{1}{9} (u_p(x) + \bar{u}_p(x)) + \frac{4}{9} (d_p(x) + \bar{d}_p(x)) \right]. \quad (69)$$

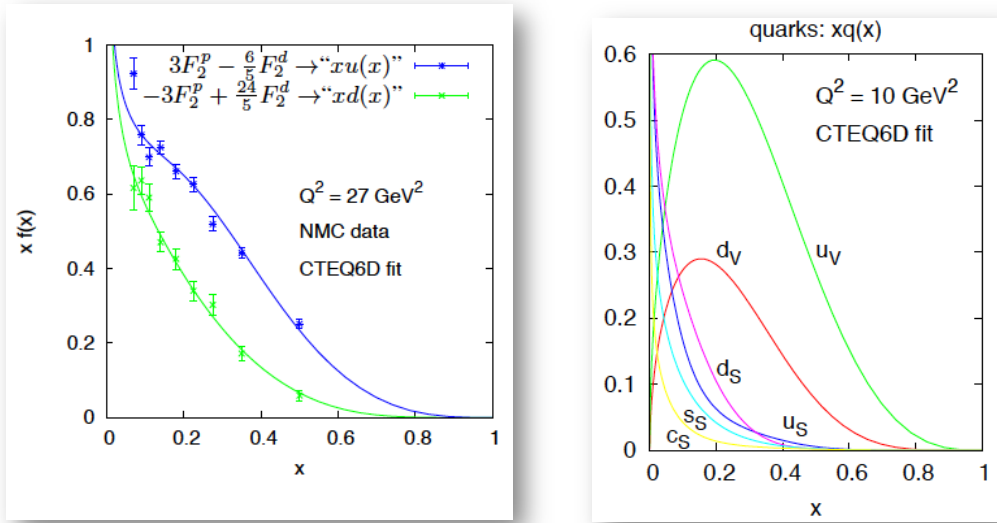
The measurements are supplemented by sum rules. For instance, as the proton consists of uud valence quarks, we have

$$\int_0^1 dx (u_p(x) - \bar{u}_p(x)) = 2, \quad \int_0^1 dx (d_p(x) - \bar{d}_p(x)) = 1, \quad \int_0^1 dx (s_p(x) - \bar{s}_p(x)) = 0.$$

The combination of the measurements and sum rules gives separate information on the quark distributions in the proton  $f_q(x)$ . The result of such measurements performed by the NMC collaboration [38] is shown in Fig. 17(a) together with a fit to the data by the CTEQ collaboration [39]. The parton distributions deduced from the fit are shown in Fig. 17(b).

We can infer the proton momentum from the measurements. The surprising result is that *quarks give only about half of the momentum of the proton*,  $\sum_q \int_0^1 dx x f_{q/p}(x) \simeq 0.5$ . By now we know that the other half is carried by gluons, but clearly the naïve parton model is not sufficient to interpret the gluon distribution in the proton. With our experience in pQCD we try to compute radiative corrections to the quark process to see if that helps to find the role of the gluon distribution.





**Fig. 17:** (a) Measurement of combination of  $F_2$  structure functions on proton and deuteron targets by the NMC collaboration fit by CTEQ6D PDF set, (b) CTEQ6D valence and sea quark distributions

**Exercise 5.2** It is not feasible to use a neutron target experimentally. Instead deuteron is used which is the bound state of a proton and a neutron. The corresponding structure function is  $F_2^{\text{deuteron}}(x) = \frac{1}{2}(F_2^{\text{proton}}(x) + F_2^{\text{neutron}}(x))$ , with  $F_2^{\text{proton}}$  and  $F_2^{\text{neutron}}$  given in Eqs. (68) and (69), respectively. Which combination of the structure function on proton and deuteron targets gives the  $u$ - and  $d$ -quark distributions?

#### 5.4 Improved parton model: pQCD

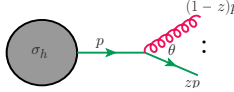
Using the relations  $dy = dQ^2/\hat{s}$  and  $\delta(\xi - x) = \frac{1}{\xi}\delta\left(1 - \frac{x}{\xi}\right)$ , we rewrite the differential cross section (65) in a more usual notation,

$$\frac{d^2\sigma}{dx dQ^2} = \int_0^1 \frac{d\xi}{\xi} \sum_i f_i(\xi) \frac{d^2\hat{\sigma}}{dx dQ^2} \left(\frac{x}{\xi}, Q^2\right), \quad (70)$$

which gives the cross section as a convolution of a long-distance component (the PDF) and a short-distance component (the hard scattering cross section). This form of the cross section is the main content of the *factorization theorem*, which we derived heuristically, but a rigorous proof, based on QFT exists.


The factorization formula (70) raises some questions. Knowing that the quarks do not give the total momentum of the proton, it is natural to include the contribution of gluons in Eq. (70). However, we do not yet know the corresponding hard scattering cross section. We also do not know how we can apply PT. Furthermore, the scaling was exact in the parton model. Is it so in QCD? There is a common answer to these questions: DIS in pQCD.

To develop pQCD for DIS, let us revisit the IR singularities once more. Let us denote the hard scattering cross section for some final state by  $\sigma_h$ . Then the cross section in the collinear approximation for the same final state with an extra gluon of relative transverse momentum  $k_\perp = E\theta$ , carrying momentum fraction  $(1 - z)$  is



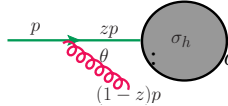
$$\sigma_{h+g} \simeq \sigma_h 2C_F \frac{\alpha_s}{\pi} \frac{dE}{E} \frac{d\theta}{\theta} = \sigma_h C_F \frac{\alpha_s}{\pi} \frac{dz}{1-z} \frac{dk_{\perp}^2}{k_{\perp}^2}.$$

Integrating over  $z$  up to one and over  $k_{\perp}$  we find soft and collinear divergence, respectively. In studying pQCD we found that these IR singularities in the final state cancel against IR divergences in the virtual correction for IR safe quantities:



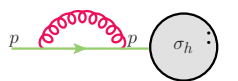
$$\sigma_{h+V} \simeq -\sigma_h C_F \frac{\alpha_s}{\pi} \frac{dz}{1-z} \frac{dk_{\perp}^2}{k_{\perp}^2}.$$

If there is a coloured parton in the initial state, then the splitting may occur before the hard scattering and the momentum of the parton that enters the hard process is reduced to  $zp^{\mu}$ , so



$$\sigma_{h+g}(p) \simeq \sigma_h(zp) 2C_F \frac{\alpha_s}{\pi} \frac{dE}{E} \frac{d\theta}{\theta} = \sigma_h C_F \frac{\alpha_s}{\pi} \frac{dz}{1-z} \frac{dk_{\perp}^2}{k_{\perp}^2}.$$

Integrating over  $z$  up to one and over  $k_{\perp}$  we again find soft and collinear divergence, respectively. The corresponding  $\epsilon$  poles multiply  $\sigma_h(zp)$ , while in the virtual correction the poles multiply  $\sigma_h(p)$ , irrespective whether the IR divergence is in the initial or final state:



$$\sigma_{h+V} \simeq -\sigma_h C_F \frac{\alpha_s}{\pi} \frac{dz}{1-z} \frac{dk_{\perp}^2}{k_{\perp}^2}.$$

The sum of the real and virtual corrections then contains an uncancelled singularity,

$$\sigma_{h+g} + \sigma_{h+V} \simeq C_F \frac{\alpha_s}{\pi} \underbrace{\int_{m_g^2}^{Q^2} \frac{dk_{\perp}^2}{k_{\perp}^2}}_{\text{infinite if } m_g=0} \underbrace{\int_0^1 \frac{dz}{1-z} [\sigma_h(zp) - \sigma_h(p)]}_{\text{finite}},$$

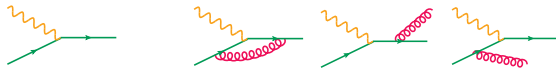
where we used a finite gluon mass to regulate the collinear divergence (instead of dimensional regularization) to make manifest that the collinear singularity remains, while the soft one (at  $z \rightarrow 1$ ) vanishes in the sum.

This uncancelled collinear singularity in the initial state is a general feature of pQCD computations with incoming coloured partons and its form is universal, so we can find its precise form studying the structure function at NLO accuracy. We know that in the parton model (QCD at LO) the prediction for hard scattering cross section  $\hat{F}_2$  is finite:

$$\hat{F}_{2,q}(x) = \left. \frac{d^2 \hat{\sigma}}{dx dQ^2} \right|_{F_2} = e_q^2 x \delta(1-x), \quad \hat{F}_{2,g}(x) = \left. \frac{d^2 \hat{\sigma}}{dx dQ^2} \right|_{F_2} = \sum_q e_q^2 x \cdot 0$$

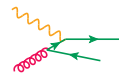
i.e., it is zero in the gluon channel because the virtual photon does not interact with the gluon directly. At one order higher in  $\alpha_s$  we find

$$\hat{F}_{2,q}(x) = \left. \frac{d^2 \hat{\sigma}}{dx dQ^2} \right|_{F_2} = e_q^2 x \left[ \delta(1-x) + \frac{\alpha_s}{4\pi} \left( P_{qq}(x) \ln \frac{Q^2}{m_q^2} + C_2^q(x) \right) \right], \quad (71)$$



and

$$\hat{F}_{2,g}(x) = \left. \frac{d^2 \hat{\sigma}}{dx dQ^2} \right|_{F_2} = \sum_q e_q^2 x \left[ 0 + \frac{\alpha_s}{4\pi} \left( P_{q\bar{q}}(x) \ln \frac{Q^2}{m_q^2} + C_2^g(x) \right) \right], \quad (72)$$



where  $P_{ij}(x)$  is the Altarelli-Parisi splitting function (regularized at  $x = 1$ ), obtained from the splitting kernel  $\hat{P}_{ij}$  in four dimensions by (i) averaging over the spin states of the splitting parton and (ii) adding the contribution from the loop graphs, while  $C_2(x)$  is the remaining finite term, called coefficient function. We see that at NLO the prediction for  $\hat{F}_2$  is finite in the UV and final state IR divergences cancel, but un-cancelled singularity remains in the initial state IR, regularized with a small mass here.

The hard scattering function is not measurable, only the structure function is physical:

$$F_{2,q}(x, Q^2) = x \sum_i e_{q_i}^2 \left[ f_{q_i}^{(0)}(x) + \frac{\alpha_s}{2\pi} \int_0^1 \frac{d\xi}{\xi} f_{q_i}^{(0)}(\xi) \left( P_{qg} \left( \frac{x}{\xi} \right) \ln \frac{Q^2}{m_g^2} + C_2^q \left( \frac{x}{\xi} \right) \right) \right].$$

However, this function appears divergent if the regulator is removed,  $m_g \rightarrow 0$ . While  $C_2(x)$  depends on the process under investigation, the divergence does not because it is multiplied with universal splitting functions.

**Exercise 5.3** Compute the coefficient  $C_2^g(x)$  in Eq. (72).

### 5.5 Factorization in DIS

If the remaining divergences are universal (and they are because do not depend on the hard scattering), we can absorb the singularity into the PDF's. For instance, defining

$$f_q(x, \mu_F) = f_q^{(0)}(x) + \frac{\alpha_s}{2\pi} \int_0^1 \frac{d\xi}{\xi} \left[ f_q^{(0)}(\xi) P_{qg} \left( \frac{x}{\xi} \right) \ln \frac{\mu_F^2}{m_g^2} + z_{qg} \left( \frac{x}{\xi} \right) \right], \quad (73)$$

the structure function becomes

$$F_{2,q}(x, Q^2) = x \sum_i e_{q_i}^2 \left[ f_i(x, \mu_F) + \frac{\alpha_s(\mu_R)}{2\pi} \int_0^1 \frac{d\xi}{\xi} f_i(\xi, \mu_F) \left( P_{qg} \left( \frac{x}{\xi} \right) \ln \frac{Q^2}{\mu_F^2} + (C_2^q - z_{qg}) \left( \frac{x}{\xi} \right) \right) \right]. \quad (74)$$

Defining the convolution in  $x$ -space,  $f \otimes_x g \equiv \int_0^1 \frac{d\xi}{\xi} f(\xi) g \left( \frac{x}{\xi} \right)$ , we see that the structure function is 'factorized' in the form of a convolution,

$$F_{2,q}(x, Q^2) = x \sum_i e_{q_i}^2 f_i(\mu_F) \otimes_x \hat{F}_{2,i}(\mu_R, t), \quad t = \ln \frac{Q^2}{\mu_F^2}.$$

The long distance physics is factored into the PDF's that depend on the *factorization scale*  $\mu_F$ . The short distance physics is factored into the hard scattering cross section that depends on both the factorization and the renormalization scales. Both scales are arbitrary, unphysical scales. The term  $z$  defines the *factorization scheme*. It is not unique, finite terms can be shifted between the short and long distance parts, but it is important that it must be chosen the same in all computations (the  $\overline{\text{MS}}$  scheme is the standard).

**Exercise 5.4** The regularization of the splitting functions at  $z = 1$  is achieved by the  $+$ -prescription defined by

$$\int_0^1 dx \frac{f(x)}{(1-x)_+} = \int_0^1 dx \frac{f(x) - f(1)}{1-x}$$

for any smooth test function  $f(x)$ . The contribution of the loop corrections has the same kinematics as the LO one, so it has to be proportional to  $\delta(1-x)$ . Thus the complete regularized splitting function has the form

$$P_{qg}(x) = C_F \left[ \frac{1+x^2}{(1-x)_+} + K\delta(1-x) \right].$$

We can obtain the parton distribution for quark in quark from Eq. (73) by the substitution  $f_q^{(0)}(x) \rightarrow \delta(1-x)$

$$f_q(x, \mu_F) = \delta(1-x) + \frac{\alpha_s}{2\pi} P_{qg}(x) \ln \frac{\mu_F^2}{m_g^2}.$$

Integration over  $x$  gives the number of quarks in a quark that has to be one, independently of  $\mu_F$ . Thus we have the condition  $\int_0^1 dx P_{qg}(x) = 0$ . Compute the regularized splitting function.

## 5.6 DGLAP equations

The short-distance component of the factorized structure function in Eq. (74) can be computed in pQCD. It depends on the renormalization scale, but recall that it has to satisfy the RGE.

We cannot compute the PDF's in PT, so it seems that this is the end of the story: pQCD appears non-predictive for processes with hadrons in the initial state. However, the arguments that lead to the RGE come to the rescue. While the right hand side of Eq. (74) depends on both renormalization and factorization scales, the measurable quantity  $F_2$  does not, which can be expressed by RGE. Of course, this statement has to be understood perturbatively, namely at any order in PT, the right hand side of the RGE is not exactly zero, but may contain terms that are higher order in PT. Only infinite order is expected to give exact independence of the scales. The RGE gives the missing piece of information needed to make the theory predictive.

To write the RGE, we introduce Mellin transforms defined by  $f(N) \equiv \int_0^1 dx x^{N-1} f(x)$ , which turns a convolution into a real product:

$$\begin{aligned} \int_0^1 dx x^{N-1} \int_0^1 \frac{d\xi}{\xi} f(\xi) g\left(\frac{x}{\xi}\right) &= \int_0^1 dx x^{N-1} \int_0^1 d\xi \int_0^1 dy \delta(x-y\xi) f(\xi) g(y) \\ &= \int_0^1 d\xi \int_0^1 dy (\xi y)^{N-1} f(\xi) g(y) = f(N)g(N). \end{aligned}$$

So  $F_{2,q}(N, Q^2) = x \sum_i e_{q_i}^2 f_i(N, \mu_F) \hat{F}_{2,i}(N, \mu_R, t)$  is independent of  $\mu_F$ , expressed as

$$\mu_F \frac{dF_2}{d\mu_F} = 0 \left( = \mathcal{O}(\alpha_s^{n+1}) \text{ in PT at } \mathcal{O}(\alpha_s^n) \right).$$

Let us explore the consequences of this RGE. For simplicity, let us assume one quark flavour,  $F_{2,q}(N, Q^2) = x e_{q_i}^2 f_q(N, \mu_F) \hat{F}_{2,i}(N, \mu_R, t)$ . Then the RGE reads

$$\hat{F}_{2,q}(N, t) \frac{df_q}{d\mu_F}(N, \mu_F) + f_q(N, \mu_F) \frac{d\hat{F}_{2,q}}{d\mu_F}(N, t) = 0.$$

Dividing with  $f_q \hat{F}_{2,q}$ , it turns into

$$\mu_F \frac{d \ln f_q}{d\mu_F}(N, \mu_F) = -\mu_F \frac{d \ln \hat{F}_{2,q}}{d\mu_F}(N, t) \equiv -\gamma_{qg}(N), \quad (75)$$

where  $\gamma_{qg}(N)$  is called the anomalous dimension because it acts as a factor  $\mu_F^{-\gamma_{qg}(N)}$  in the dimensionless function  $\ln f_q(N, \mu_F)$ . Taking the Mellin moment of Eq. (73) and then its derivative with respect to  $\mu_F$ , we obtain that the anomalous dimension is

$$\gamma_{qg}(N) = -\mu_F \frac{d \ln f_q}{d \mu_F}(N, \mu_F) = -\frac{\alpha_s(\mu_R)}{\pi} P_{qg}(N) + \mathcal{O}(\alpha_s^2), \quad (76)$$

i.e., it is the Mellin transform of the splitting function, which can be computed in PT. Equation (75) implies that *the scale dependence of the PDF can be predicted in PT*. This together with the universality of PDF's makes pQCD predictive: *we can measure the PDF's in one process at a certain scale and then use it in another process at another scale to make predictions*.

How shall we choose the renormalization and factorization scales? If we want to avoid large logarithms that spoil the convergence of the perturbative series, the scales should be chosen near the characteristic physical scale of the process  $Q$ , e.g.,  $\mu_R^2 = \mu_F^2 = Q^2$ . Then the RGE becomes

$$Q^2 \frac{d \ln f_q}{d Q^2}(N, Q^2) = -\frac{1}{2} \gamma_{qg}\left(N, \alpha_s(Q^2)\right), \quad (77)$$

which is the Mellin transform of

$$Q^2 \frac{d f_q}{d Q^2}(x, Q^2) = \frac{\alpha_s(Q^2)}{2\pi} P_{qg} \otimes_x f_q(Q^2). \quad (78)$$

Our discussion was highly simplified by considering only one quark flavour and neglecting the mixing of partons. If we make the full computation we obtain the gold-plated formula

$$Q^2 \frac{d f_{(ij)}}{d Q^2}(x, Q^2) = \frac{\alpha_s(Q^2)}{2\pi} \sum_i P_{ij} \otimes_x f_i(Q^2), \quad (79)$$

called DGLAP (for Dokshitzer [40], Gribov-Lipatov [41] and Altarelli-Parisi [42]) equation.

Let us now solve the (simplified) DGLAP equation in Mellin space, Eq. (77). It is a simple first order differential equation whose solution is

$$f_q(N, Q^2) = f_q(N, Q_0^2) \exp \left[ - \int_{t_0}^t dt \gamma_{qg}\left(N, \alpha_s(\Lambda^2 e^t)\right) \right].$$

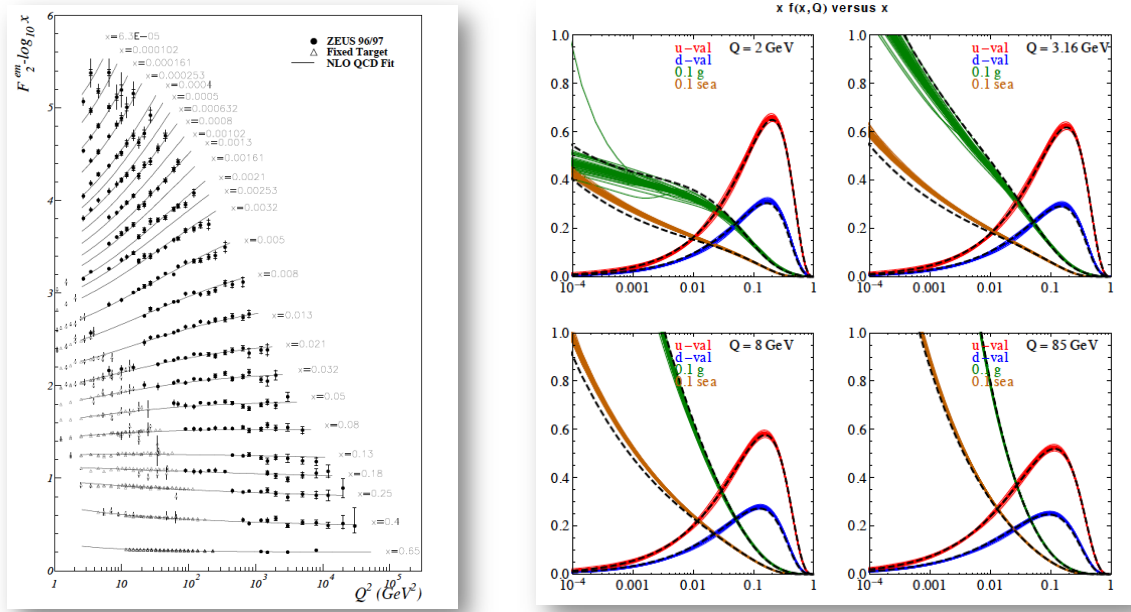
Let us recall the one-loop formula in Eq. (20),  $\alpha_s(Q^2) = \frac{1}{b_0 t}$ ,  $t = \ln \frac{Q^2}{\Lambda^2}$  and introduce the abbreviation  $d_{qg}(N) = -\frac{\gamma_{qg}(N)}{2\pi b_0} \leq 0$ . Then

$$f_q(N, Q^2) = f_q(N, Q_0^2) \exp \left[ d_{qg}(N) \int_{t_0}^t \frac{dt}{t} \right], \quad \text{or} \quad f_q(N, Q^2) = f_q(N, Q_0^2) \left( \frac{t}{t_0} \right)^{d_{qg}(N)}, \quad (80)$$

called *scaling violation*.

As  $\gamma_{qg}(1) = 0$ , the valence q-quark in the proton, given by the integral  $\int_0^1 dx f_q(x, Q^2)$ , is independent of  $Q^2$ . Higher moments vanish more rapidly, therefore, the average  $x$  decreases as  $Q^2$  increases. Thus we predict that  $f_q(x, Q^2)$  increases at small  $x$  and decreases at large  $x$ . This prediction is seen to be valid from the measurements shown in Fig. 18(a).

**Exercise 5.5** Compute the anomalous dimension  $\gamma_{qg}(x)$  using Eq. (76).



**Fig. 18:** (a) Measurement of  $F_2$  structure function at different  $Q^2$  as a function of  $x$ , (b) evolution of valence quark, sea quark and gluon distributions

## 6 Hadron collisions

While electron-positron annihilation and DIS played very important role in establishing pQCD for understanding high-energy scattering experiments, presently and in the mid-term future the experiments at the energy frontier can be found at the Large Hadron Collider (LHC). Thus we are most interested in the theoretical tools needed to understand high-energy proton-proton collisions.

### 6.1 Factorization theorem

Fortunately, the tools we have developed so far can be generalized straightforwardly to hadron collisions. The most general form of the factorization theorem includes convolution with two PDF's, one for each colliding parton, the hard scattering cross section, and possibly a convolution with a fragmentation function (FF) of a parton into an identified hadron in the final state. Thus, the differential cross section for a hypothetical process  $pp \rightarrow Z + \pi + X$  has the form

$$\begin{aligned}
 d\sigma_{pp \rightarrow Z + \pi + X}(s, x, \alpha_s, \mu_R, \mu_F) &= \sum_{i,j,k} \int_0^1 dx_1 f_{i/p}(x_1, \alpha_s, \mu_F) \int_0^1 dx_2 f_{j/p}(x_2, \alpha_s, \mu_F) \\
 &\times \int_x^1 \frac{dz}{z} d\hat{\sigma}_{ij \rightarrow Z+k+X}(\hat{s}, z, \alpha_s(\mu_R), \mu_R, \mu_F) D_{\pi/k}\left(\frac{x}{z}, \hat{s}\right) + \mathcal{O}\left(\frac{\Lambda}{Q}\right)^p.
 \end{aligned} \tag{81}$$

In Eq. (81)  $s$  is the total centre-of-mass energy squared,  $x/z$  is the longitudinal momentum fraction of the pion in the parton  $k$ ,  $\mu_R$  and  $\mu_F$  are the renormalization and factorization scales,  $f_{i/p}(x)$  is the PDF for parton  $i$  in the proton with momentum fraction  $x$ ,  $d\hat{\sigma}_{ij \rightarrow Z+k+X}(\hat{s})$  is the hard scattering cross section for the partonic process,  $D_{\pi/k}(x)$  is the FF for the process parton  $k \rightarrow \pi$ . The last term shows that contributions suppressed at high  $Q^2$  are neglected ( $p > 1$ ). Substituting the PDF's and FF's with  $\delta$  functions (in momentum and flavour) we obtain the cross section formulae in DIS and electron-positron annihilation.

The PDF's and FF's constitute the non-perturbative, long-distance components of the cross section that cannot be computed in pQCD, only extracted from measurements. Thus, it is a natural question whether or not the factorization theorem is predictive. The answer is a clear yes for the following reasons.

We can compute the hard scattering cross section in PT, which involves (i) renormalization of UV divergences (order by order in PT), (ii) cancellation of IR ones for IR safe observables using a subtraction method, (iii) absorbing initial state collinear divergences into renormalization of PDF's (and possibly uncanceled final state ones into that of the FF). The non-perturbative components are universal, so can be measured in one process and used to make prediction in another one. Furthermore, the evolution of these with  $Q^2$  can be predicted in PT (DGLAP equations), shown in Fig. 18(b).

In summary, we are prepared to make predictions for any high-energy scattering process. The theoretical framework for such predictions relies on pQCD and the factorization theorem. In PT we can compute the hard scattering cross section and the evolution of the PDF's. There are universal elements, such as the PDF's and FF's, as well as the subtraction method for computing radiative corrections.

### 6.2 Are we happy?

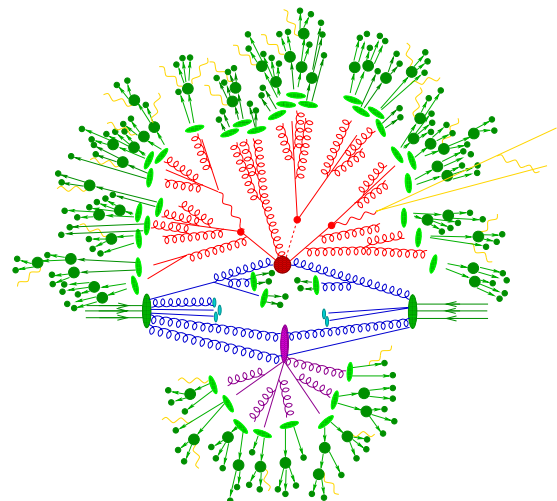
At this point theorists can make precision predictions for distributions of IR safe observables. The main bottleneck to make such predictions is the algebraic complexity of computing amplitudes and the analytic complexity of evaluating loop integrals. The state of the art considers the computation of NLO corrections a solved problem with automated implementations for processes up to about five partons in the final state (at tree level). The exact number depends on the process being considered because the numerical integrations become too expensive eventually. Nevertheless, all processes listed in the 'Les Houches wishlist (2011)' are known by now. Furthermore, there is also a computer code to compute seven-jet production in electron-positron annihilation [43].

For experimenters the situation is less satisfactory. While pQCD predictions are based on a solid theoretical ground, those lack important features. On the one hand pQCD gives predictions for final states with few partons, detectors detect hadrons. A tool that can simulate real events with hadrons at correct rates would be much more handy. To finish these lectures we look into modeling events in a qualitative way. A more detailed description can be found in Ref. [44].

### 6.3 Modelling events

Figure 19 shows our view of a proton-proton scattering event at high energy. The three parallel lines ending in discs from both sides represent the two incoming protons. At high energies these protons consist of (almost) free-flying partons, two of which (one from each) collide at high centre-of-mass energy and produce the hard scattering, with perturbatively computable cross section. This is where signs of new physics may appear. The hard scattering cross section is process dependent. We have discussed how it can be computed from first principles, which can be improved systematically by computing the radiative corrections.

Before collision the colliding partons may emit other partons collinear with the beam. These collinear emissions in the initial state give rise to divergences that can be factored into the renormalized parton distribution functions. After collision few energetic partons appear that may emit less energetic partons and each develops showers of partons. Emissions into



**Fig. 19:** Artistic view of a proton-proton scattering event at high energy (courtesy of F. Krauss)

almost the same direction as the original parton occur with enhanced probability (due to the collinear divergence) as well as emissions of soft gluon. This is represented in Figure 19 by red quark and gluon lines. Both factorization and parton showering can be described from first principles based upon known physics of QCD, and are universal, i.e., independent of the process and observable. We have seen how factorization works, but have not discussed how parton showers are modelled with shower Monte Carlo (SMC) programs [45,46]. We mentioned marginally how the large logarithms emerging in the final state splittings can be resummed, which gives improved prediction for the cross section (as seen in Fig. 14), but does not simulate events.

Parton showers still only give a description of events in terms of quarks and gluons, whereas detectors detect only hadrons. We do not know how to compute hadronization, the transition from quarks and gluon to hadrons, from first principles. Yet the idea of local parton-hadron duality (LPHD) provides some sort of theoretical understanding (see, e.g., Ref. [47]). It states that *after accounting for all gluon and quark production down to scales  $\simeq \Lambda_{\text{QCD}}$ , the transition from partons to hadrons is essentially local in phase space*. Thus the hadron directions and momenta will be closely related to that of the partons, and the hadron multiplicity will reflect the parton multiplicity, too. This is illustrated by the green lines with dots.

In addition to the energetic partons in the initial state, there are also low-energy ones that may collide, which is energy and process dependent. This low-energy physics is described in models of underlying event, which are also part of modern SMC's. The underlying event produces low-energy partons. Also at the end of the shower low-energy partons emerge. As QCD confines partons, these partons turn into hadrons before detection, a process called hadronization. We do not have a theory of hadronization based on first principles. Instead, SMC's include models that describe hadronization in a process independent way. These models contain parameters that are fixed experimentally.

## 7 Conclusions

In these lectures we discussed the theoretical basis of interpreting the results of high-energy collider experiments. We discussed how pQCD can be made predictive and also the main uncertainties in the predictions. We used the following key ingredients in this tour: (i) gauge invariance that allows us to write down the Lagrangian and which predicts many important features of the theory; (ii) renormalization that cancels ultraviolet divergences systematically order-by-order in perturbation theory and introduces a dimensionful scale into even the scaleless Lagrangian of massless QCD, leading to scaling violations of one-scale observables that would be scale independent in the classical theory; (iii) asymptotic freedom at high energies emerging from the quantum structure of the theory and the non-Abelian nature of the gauge group; (iv) need for infrared safety, emerging from asymptotic freedom, to ensure that the IR divergences, associated with unresolved parton emission, cancel between real and virtual contributions, allowing the perturbative calculation of jet cross sections, without a detailed understanding of the mechanism by which partons become jets; (v) factorization that makes possible to use perturbative QCD to calculate the interactions of hadrons, since all the non-perturbative physics gets factorized, into parton distribution functions; (vi) evolution and universality of PDF's that allows us to extract those measuring cross sections in one process, like DIS, and then used to predict the cross sections for any other process. Again, this factorization introduces a scale dependence into the parton model so that the structure functions of DIS, and other one-scale observables become scale dependent. These features make pQCD predictive, without forcing us to solve the theory at all possible scales: unknown or uncalculable high- and low-energy effects can be renormalized, factorized and cancelled away.

Of course, in four double lectures, it was impossible to give full treatment of any of the topics we encountered. For that I refer to any of the classic textbooks about QCD at colliders [48–50].



## Acknowledgements

My lectures grew partly out of research conducted during the past 20 years and supported in part by Hungarian Scientific Research Fund grants, most recently under grant OTKA K-101482, and partly out of lectures given at Universities of Debrecen and Zürich where I received useful feedback from students. I benefited greatly from the lectures of previous schools by G. Salam [2] and F. Maltoni (<http://physicschool.web.cern.ch/PhysicSchool/ESHEP/ESHEP2011/programme.html>). I am grateful to the organizers, discussion leaders and students of the 2013 European School of High-Energy Physics for providing a pleasant and stimulating atmosphere.

## References

- [1] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, *Phys.Lett.* **B47** (1973) 365–368.
- [2] G. P. Salam, *Perturbative QCD for the LHC, PoS ICHEP2010* (2010) 556, [[arXiv:1103.1318](https://arxiv.org/abs/1103.1318)].
- [3] Z. Nagy and Z. Trocsanyi, *Four jet angular distributions and color charge measurements: Leading order versus next-to-leading order*, *Phys.Rev.* **D57** (1998) 5793–5802, [[hep-ph/9712385](https://arxiv.org/abs/hep-ph/9712385)].
- [4] OPAL Collaboration, G. Abbiendi *et. al.*, *A Simultaneous measurement of the QCD color factors and the strong coupling*, *Eur.Phys.J.* **C20** (2001) 601–615, [[hep-ex/0101044](https://arxiv.org/abs/hep-ex/0101044)].
- [5] J. Gasser and H. Leutwyler, *Chiral Perturbation Theory to One Loop*, *Annals Phys.* **158** (1984) 142.
- [6] J. Gasser and H. Leutwyler, *Quark Masses*, *Phys.Rept.* **87** (1982) 77–169.
- [7] J. Gasser and H. Leutwyler, *Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark*, *Nucl.Phys.* **B250** (1985) 465.
- [8] G. Colangelo, J. Gasser, and H. Leutwyler,  *$\pi\pi$  scattering*, *Nucl.Phys.* **B603** (2001) 125–179, [[hep-ph/0103088](https://arxiv.org/abs/hep-ph/0103088)].
- [9] S. Durr, Z. Fodor, J. Frison, C. Hoelbling, R. Hoffmann, *et. al.*, *Ab-Initio Determination of Light Hadron Masses*, *Science* **322** (2008) 1224–1227, [[arXiv:0906.3599](https://arxiv.org/abs/0906.3599)].
- [10] T. van Ritbergen, J. Vermaseren, and S. Larin, *The Four loop beta function in quantum chromodynamics*, *Phys.Lett.* **B400** (1997) 379–384, [[hep-ph/9701390](https://arxiv.org/abs/hep-ph/9701390)].
- [11] L. Abbott, *The Background Field Method Beyond One Loop*, *Nucl.Phys.* **B185** (1981) 189.
- [12] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys.Rev.Lett.* **30** (1973) 1343–1346.
- [13] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys.Rev.Lett.* **30** (1973) 1346–1349.
- [14] Particle Data Group Collaboration, J. Beringer *et. al.*, *Review of Particle Physics (RPP)*, *Phys.Rev.* **D86** (2012) 010001.
- [15] M. Jamin and M. E. Lautenbacher, *TRACER: Version 1.1: A Mathematica package for gamma algebra in arbitrary dimensions*, *Comput.Phys.Commun.* **74** (1993) 265–288.
- [16] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, *JHEP* **1106** (2011) 128, [[arXiv:1106.0522](https://arxiv.org/abs/1106.0522)].
- [17] A. Pukhov, *CalcHEP 2.3: MSSM, structure functions, event generation, batchs, and generation of matrix elements for other packages*, [hep-ph/0412191](https://arxiv.org/abs/hep-ph/0412191).
- [18] A. Pukhov, E. Boos, M. Dubinin, V. Edneral, V. Ilyin, *et. al.*, *CompHEP: A Package for evaluation of Feynman diagrams and integration over multiparticle phase space*, [hep-ph/9908288](https://arxiv.org/abs/hep-ph/9908288).
- [19] T. Hahn, *Generating Feynman diagrams and amplitudes with FeynArts 3*, *Comput.Phys.Commun.* **140** (2001) 418–431, [[hep-ph/0012260](https://arxiv.org/abs/hep-ph/0012260)].
- [20] ALEPH, DELPHI, L3, OPAL, SLD Collaborations, LEP Electroweak Working Group, SLD

- Electroweak Group, SLD Heavy Flavour Group, S. Schael *et. al.*, *Precision electroweak measurements on the  $Z$  resonance*, *Phys.Rept.* **427** (2006) 257–454, [hep-ex/0509008].
- [21] K. Chetyrkin, *Four-loop renormalization of QCD: Full set of renormalization constants and anomalous dimensions*, *Nucl.Phys.* **B710** (2005) 499–510, [hep-ph/0405193].
- [22] P. Baikov, K. Chetyrkin, J. Kuhn, and J. Rittinger, *Complete  $\mathcal{O}(\alpha_s^4)$  QCD Corrections to Hadronic  $Z$ -Decays*, *Phys.Rev.Lett.* **108** (2012) 222003, [arXiv:1201.5804].
- [23] Z. Bern, L. Dixon, F. Febres Cordero, S. Hoeche, H. Ita, *et. al.*, *Next-to-Leading Order  $W + 5$ -Jet Production at the LHC*, *Phys.Rev.* **D88** (2013) 014025, [arXiv:1304.1253].
- [24] G. Grunberg, *Renormalization Group Improved Perturbative QCD*, *Phys.Lett.* **B95** (1980) 70.
- [25] P. M. Stevenson, *Optimized Perturbation Theory*, *Phys.Rev.* **D23** (1981) 2916.
- [26] S. J. Brodsky, G. P. Lepage, and P. B. Mackenzie, *On the Elimination of Scale Ambiguities in Perturbative Quantum Chromodynamics*, *Phys.Rev.* **D28** (1983) 228.
- [27] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J.Math.Phys.* **3** (1962) 650–677.
- [28] T. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Phys.Rev.* **133** (1964) B1549–B1562.
- [29] E. L. Berger, *Research directions for the decade. Proceedings, 1990 Summer Study on High-Energy Physics, Snowmass, USA, June 25 - July 13, 1990.*, .
- [30] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. Webber, *New clustering algorithm for multi-jet cross-sections in  $e^+e^-$  annihilation*, *Phys.Lett.* **B269** (1991) 432–438.
- [31] ALEPH Collaboration, R. Barate *et. al.*, *Studies of quantum chromodynamics with the ALEPH detector*, *Phys.Rept.* **294** (1998) 1–165.
- [32] Z. Nagy and Z. Trocsanyi, *Multijet rates in  $e^+e^-$  annihilation: Perturbation theory versus LEP data*, *Nucl.Phys.Proc.Suppl.* **74** (1999) 44–48, [hep-ph/9808364].
- [33] M. Cacciari and G. P. Salam, *Dispelling the  $N^3$  myth for the  $k_t$  jet-finder*, *Phys.Lett.* **B641** (2006) 57–61, [hep-ph/0512210].
- [34] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$  jet clustering algorithm*, *JHEP* **0804** (2008) 063, [arXiv:0802.1189].
- [35] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl.Phys.* **B627** (2002) 189–265, [hep-ph/0201036].
- [36] JADE Collaboration, J. Schieck, S. Bethke, S. Kluth, C. Pahl, and Z. Trocsanyi, *Measurement of the strong coupling  $\alpha_S$  from the three-jet rate in  $e^+e^-$  annihilation using JADE data*, *Eur.Phys.J.* **C73** (2013) 2332, [arXiv:1205.3714].
- [37] G. Miller, E. D. Bloom, G. Buschhorn, D. Coward, H. DeStaebler, *et. al.*, *Inelastic electron-Proton Scattering at Large Momentum Transfers*, *Phys.Rev.* **D5** (1972) 528.
- [38] New Muon Collaboration, M. Arneodo *et. al.*, *Measurement of the proton and deuteron structure functions,  $F_2(p)$  and  $F_2(d)$ , and of the ratio  $\sigma_L / \sigma_T$* , *Nucl.Phys.* **B483** (1997) 3–43, [hep-ph/9610231].
- [39] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, *et. al.*, *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **0207** (2002) 012, [hep-ph/0201195].
- [40] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and  $e^+e^-$  Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov.Phys.JETP* **46** (1977) 641–653.
- [41] V. Gribov and L. Lipatov, *Deep inelastic  $e p$  scattering in perturbation theory*, *Sov.J.Nucl.Phys.* **15** (1972) 438–450.
- [42] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl.Phys.* **B126** (1977) 298.

- [43] S. Becker, D. Goetz, C. Reuschle, C. Schwan, and S. Weinzierl, *NLO results for five, six and seven jets in electron-positron annihilation*, *Phys.Rev.Lett.* **108** (2012) 032005, [arXiv:1111.1733].
- [44] P. Skands, *Introduction to QCD*, arXiv:1207.2389.
- [45] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **0605** (2006) 026, [hep-ph/0603175].
- [46] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, *et. al.*, *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *JHEP* **0101** (2001) 010, [hep-ph/0011363].
- [47] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. I. Troyan, *Basics of Perturbative QCD*, Editions Frontieres, Gif-sur-Yvette, 1991.
- [48] R. Ellis, J. Stirling, and B. Webber, *QCD and Collider Physics*, Cambridge University Press, 1996.
- [49] G. G. Dissertori, I. Knowles, and M. Schmelling, *High-Energy Experiments and Theory*, Clarendon, Oxford, 2003.
- [50] **CTEQ** Collaboration, R. Brock *et. al.*, *Handbook of perturbative QCD: Version 1.0*, *Rev.Mod.Phys.* **67** (1995) 157–248.



## Higgs Physics

*J. Ellis*

Department of Physics, King's College London, London WC2R 2LS, United Kingdom;  
Theory Division, CERN, CH-1211 Geneva 23, Switzerland

### Abstract

These lectures review the background to Higgs physics, its current status following the discovery of a/the Higgs boson at the LHC, models of Higgs physics beyond the Standard Model and prospects for Higgs studies in future runs of the LHC and at possible future colliders.

KCL-PH-TH/2013-49, LCTS/2013-36, CERN-PH-TH/2013-315

## 1 Motivations and Context

### 1.1 To Higgs or not to Higgs?

The Standard Model describes all the visible matter in the Universe in terms of a limited number of fermionic constituents of matter: six quarks, three charged leptons and three light neutrinos. It also comprises three fundamental gauge interactions between these constituents, namely the electromagnetic, strong and weak forces, to which should be added gravitation. The Standard Model is in good agreement with all confirmed experimental results from particle accelerators. However, there was, until July 4th 2012, one crucial missing ingredient: the origin of particle masses. Without a mass for the electron there would be no atoms, as electrons would escape from nuclei at the speed of light, and the weak interactions would not live up to their name: they would be stronger than electromagnetism. So discovering the origin of particle masses has been a Big Deal.

In addition to elucidating the origin of particle masses and establishing whether a/the Higgs boson exists, there are many important open questions beyond the Standard Model. Why are there so many different types of matter particles? Why do weak interactions mix them the way they do? How/why do they discriminate between matter and antimatter, and might this difference explain the dominance of matter over antimatter in the Universe today? What is the nature of the invisible dark matter that, according to astrophysicists and cosmologists, dominates over the visible matter described by the Standard Model? Are the fundamental forces unified, and what is the full quantum theory of gravity?

The key to answering many of these other questions may be provided by finding the origin of particle masses. For example, decays of the Higgs boson may discriminate between matter and antimatter, and might lie at the origin of the cosmological matter-antimatter asymmetry. In many theoretical extensions of the Standard Model, the Higgs boson is accompanied by other new particles, the lightest of which might provide the astrophysical dark matter. A possible unified theory of all the elementary particle interactions might employ a symmetry-breaking mechanism analogous to that in the Standard Model, and the existence of any light 'elementary' scalar boson would pose a challenge for many quantum theories of gravity.

The equations describing the gauge interactions of the Standard Model do not discriminate between matter particles with the same quantum numbers, which differ only in their masses. The mass-generation mechanism must discriminate between these otherwise-identical particles of matter and between the different force-carrying vector bosons: it must break the symmetries between them. One way to achieve this might be to break the symmetry explicitly in the equations, but then the calculability of the theory would be lost. The alternative is to retain symmetric equations but break the symmetry in their solutions.

The issue then is whether to break the symmetry throughout space, or via boundary conditions. The latter is not possible in conventional three-dimensional space, since it has no boundaries. However, it would be possible in theories with additional dimensions of space: one could postulate different behaviours in the extra dimension(s) for different particle species. The discovery of a/the Higgs boson at the LHC [1, 2] has somewhat deflated interest in extra-dimensional models, unless their spectrum features a low-mass excitation that resembles closely the Higgs boson of the Standard Model.

This discovery seems to mark the latest success of a long-running theoretical strategy in particle physics: when in trouble, postulate one of more new particles. A partial list includes reconciling quantum mechanics and special relativity (antimatter), nuclear spectra (the neutron), the continuous spectrum in  $\beta$  decay (the neutrino), nucleon-nucleon interactions (the pion), the suppression of  $\mu \rightarrow e\gamma$  (the second neutrino), flavour SU(3) ( $\Omega^-$  and quarks), the suppression of flavour-changing neutral currents (charm), CP violation (the third generation), strong dynamics (the gluons), the discovery of the  $\tau$  lepton (the  $b$  and  $t$  quarks), weak interactions (the  $W^\pm$  and  $Z^0$ ), and their renormalizability (a/the Higgs boson)<sup>1</sup>.

The discovery of a/the Higgs boson marks the completion of the Holy Trinity of particle types seen in Table 1.2. It has been known for decades that the only type of field theory capable of making non-trivial over many magnitudes of energy is a renormalizable one. Also, it has been a theorem for some 40 years that such a theory could only contain (i) gauge vector bosons, (ii) spin-1/2 fermions, and (iii) scalar bosons. Specification of a renormalizable theory is completed by choosing (a) the gauge group, (b) the fermion representations, and (c) the scalar fields used to realize the desired symmetry-breaking pattern. We have long known that the answer (so far) to (a) is  $SU(3) \times SU(2) \times U(1)$  and that the fermion representations are triplets of SU(3) and singlets and doublets of SU(2). Finally we have an example of category (iii), a scalar boson, and there is strong evidence that it is responsible for (c) electroweak symmetry breaking.

The first of these lectures describes the long road towards the discovery of a/the Higgs boson, the second lecture describes the state of our knowledge after Run 1 of the LHC, and the third lecture outlines some of the prospects for future studies, including supersymmetric Higgs bosons and concepts for Higgs factories.

## 1.2 Summary of the Standard Model

Table 1.2 summarizes the particle content of the Standard Model [3–6]. The electromagnetic and weak interactions are described by an  $SU(2)_L \times U(1)_Y$  group [7], where the subscript  $L$  reminds us that the weak SU(2) group acts only the left-handed fermions, and  $Y$  is the hypercharge. The  $SU(2)_L \times U(1)_Y$  part of the Standard Model Lagrangian may be written as

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4} \mathbf{F}_{\mu\nu}^a \mathbf{F}^{a\mu\nu} \\ &+ i\bar{\psi} \not{D} \psi + h.c. \\ &+ \psi_i y_{ij} \psi_j \phi + h.c. \\ &+ |D_\mu \phi|^2 - V(\phi) , \end{aligned} \tag{1}$$

which is short enough to write on a T-shirt!

The first two lines of (1) have been confirmed in many different experiments with a high degree of accuracy. However, consistency with the precision electroweak measurements made at LEP and other accelerators agreed with the Standard Model if only if there was a relatively light Higgs boson weighing  $< 180$  GeV or so [8, 9]. However, until July 2012 there was no direct experimental evidence for the last two lines [1, 2]. One of the main objectives of the LHC was to discover whether they are right, need

<sup>1</sup>Outstanding examples include dark matter (the axion or a WIMP?) and the fine-tuning problem (supersymmetry?). Depressed advocates of supersymmetry should remember that it took 48 years to discover a/the Higgs boson, whereas at the time of writing four-dimensional supersymmetric gauge theories are ‘only’ 40 years old.

Gauge bosons	Scalar bosons
$\gamma, W^+, W^-, Z^0, g_{1\dots 8}$	$\phi$ (Higgs)
Spin-1/2 Fermions	
Quarks (each with 3 colour charges) Charges $+2/3$ : $\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}$	Leptons Charges $0$ : $\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}$ Charges $-1$ : $\begin{pmatrix} e^- \\ \mu^- \\ \tau^- \end{pmatrix}$

**Table 1:** The particle content of the Standard Model with a minimal Higgs sector

modification, or are simply wrong. The most important result of the first run of the LHC has been to find some evidence that these lines do indeed contain a grain of truth.

The first line in (1) is the kinetic term for the gauge sector of the electroweak theory, with  $a$  running over the total number of gauge fields: three associated with  $SU(2)_L$ , which we call  $B_\mu^1, B_\mu^2, B_\mu^3$ , and one associated with  $U(1)_Y$ , which we call  $A_\mu$ . Their field-strength tensors are

$$F_{\mu\nu}^a = \partial_\nu B_\mu^a - \partial_\mu B_\nu^a + g\varepsilon_{bca}B_\mu^b B_\nu^c \text{ for } a = 1, 2, 3; \quad (2)$$

$$f_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu. \quad (3)$$

In (2),  $g$  is the coupling constant of the weak-isospin group  $SU(2)_L$ , and the  $\varepsilon_{bca}$  are its structure constants. The last term in this equation stems from the non-Abelian nature of  $SU(2)$ . The gauge fields are massless in the absence of any scalar fields, but we will see later how specific linear combinations of the four electroweak gauge fields can acquire masses by spontaneous symmetry breaking induced by a scalar field.

The second line in Eq. (1) describes the interactions between the matter fields  $\psi$ , described by Dirac equations, and the gauge fields via covariant derivatives.

The third line is the Yukawa sector and incorporates the interactions between the matter fields and the scalar field  $\phi$  that is responsible for giving fermions their masses when electroweak symmetry breaking occurs.

The fourth and final line is the engine room of the scalar sector. The first piece is the kinetic term for  $\phi$  with the covariant derivative defined here to be

$$D_\mu = \partial_\mu + \frac{ig'}{2}A_\mu Y + \frac{ig}{2}\boldsymbol{\tau} \cdot \mathbf{B}_\mu, \quad (4)$$

where  $g'$  is the  $U(1)$  coupling constant, and  $Y$  and  $\boldsymbol{\tau} \equiv (\tau_1, \tau_2, \tau_3)$  are the Pauli matrices that generate, respectively,  $U(1)$  and  $SU(2)$ . The second piece of the final line of (1) is the effective potential  $V(\phi)$  constructed in such a way that its minimization gives rise to a non-zero v.e.v. for the scalar field, and hence spontaneous electroweak symmetry breaking.

### 1.3 The (NG)AEBHGHKMP Mechanism

This mechanism is often called the Higgs mechanism, but the history is quite complicated, with many antecedents to the Higgs papers, with antecedents in condensed-matter physics, e.g., in the theories of superfluidity and superconductivity [10]<sup>2</sup>. Spontaneous global symmetry breaking was introduced into four-dimensional, relativistic particle physics by Nambu [11], providing insight into the lightness of the pion and spontaneous chiral symmetry breaking. Subsequently, a simple field-theoretical model was formulated by Goldstone [12], and the appearance and number of massless scalar bosons appearing in such theories was specified by a theorem proved in [13]. In parallel, Nambu [14] and Anderson [15]

<sup>2</sup>The wheel has turned full circle with their importance for the construction of the LHC!

had shown how to interpret superconductivity in terms of a spontaneously-broken local U(1) symmetry. Moreover, Anderson had conjectured that this could occur also in the four-dimensional case, but Gilbert [16] argued in early 1964 that this would be impossible in a relativistic theory.

However, later in 1964 several papers successfully introduced spontaneously-broken local symmetry into particle physics. The initial paper by Englert and Brout [17] was followed a few weeks later by two papers written by Higgs, who did not know about their work at the time. The first Higgs paper [18] pointed out that Gilbert’s objection to a four-dimensional extension of Anderson’s approach could be circumvented, and the second proposed a specific four-dimensional model with a massive scalar boson [19]. The subsequent paper by Guralnik, Hagen and Kibble [20] developed the ideas proposed in these earlier papers. Also of note is a relatively-unknown 1965 paper by Migdal and Polyakov [21], which discusses the partial breaking of a local non-Abelian symmetry, ahead of the influential paper of Kibble [22].

Of all these authors, Higgs was the only one who mentioned explicitly the existence of a massive scalar boson (see equation (2b) of his second paper [19]), and he went on to write a third paper in 1966 [23] that discusses the properties of this ‘Higgs boson’ in surprising detail including, e.g., its decays into massive vector bosons.

### 1.4 Spontaneous Symmetry-Breaking in a U(1) Model

As an illustration of spontaneous symmetry breaking [24], consider first a single complex scalar field  $\phi = \phi_1 + \phi_2$  with the following effective potential:

$$V(\phi^* \phi) = \mu^2 (\phi^* \phi) + \lambda (\phi^* \phi)^2, \quad (5)$$

where  $\mu^2$  and  $\lambda > 0$  are real constants. This Lagrangian is clearly invariant under global U(1) phase transformations

$$\phi \rightarrow e^{i\alpha} \phi, \quad (6)$$

where  $\alpha$  is a phase (rotation) angle. If the parameter  $\mu^2$  in (5) is positive, there is a unique vacuum state with  $\langle \phi \rangle = 0$ . Perturbing around this vacuum reveals that, in this case,  $\phi_1$  and  $\phi_2$ , have the same mass. The symmetry of the original Lagrangian is explicit in this case.

Consider now the case  $\mu^2 < 0$ , corresponding to the ‘Mexican Hat’ potential illustrated in Fig.1. When we minimize the potential (5) we find a non-zero vacuum expectation value, or v.e.v., of the scalar field with:

$$|\phi|^2 = \phi_1^2 + \phi_2^2 = -\frac{\mu^2}{2\lambda}, \quad (7)$$

and the phase  $\alpha$  undetermined. Thus, when  $\mu^2 < 0$  there is a set of equivalent minima lying around a circle of radius  $\sqrt{-\mu^2/(2\lambda)}$ , and choosing one of them breaks the rotational symmetry spontaneously.

The U(1) symmetry is now implicit, since it relates the different equivalent vacua, corresponding to the appearance of spontaneous symmetry breaking. In order to see the particle content, we choose, without loss of generality, a particular ground state around which to perturb:

$$\phi_{1,\text{vac}} = \sqrt{-\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}, \quad \phi_{2,\text{vac}} = 0. \quad (8)$$

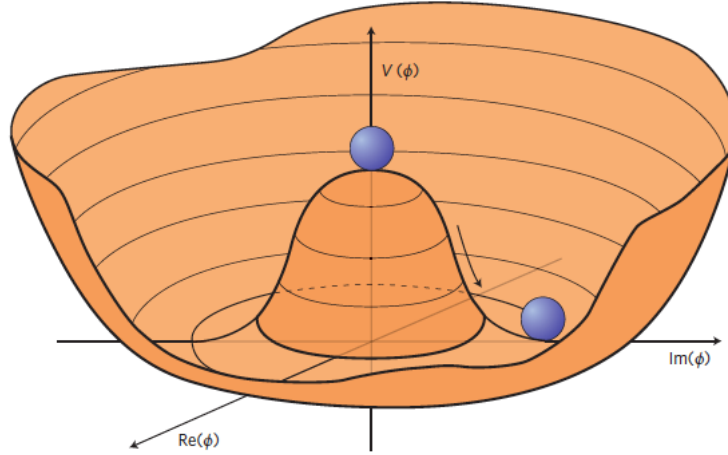
The perturbations may be parametrized by

$$\frac{\eta}{\sqrt{2}} \equiv \phi_1 - \frac{v}{\sqrt{2}}, \quad \frac{\xi}{\sqrt{2}} \equiv \phi_2, \quad (9)$$

so that  $\phi = (v + \eta + i\xi)/\sqrt{2}$ , where  $\eta$  and  $\xi$  are real fields. In terms of these, the effective potential becomes

$$\mathcal{L} = -\frac{\mu^2}{2}\eta^2 - \frac{\lambda}{2} \left[ (v + \eta)^2 + \xi^2 \right]^2 - \mu^2 v \eta - \frac{\mu^2}{2} \xi^2 - \frac{1}{2} \mu^2 v^2.$$





**Fig. 1:** An illustration of the Higgs potential (5) in the case that  $\mu^2 < 0$ , in which case the minimum is at  $|\phi|^2 = -\mu^2/(2\lambda)$ . Choosing any of the points at the bottom of the potential breaks spontaneously the rotational U(1) symmetry.

The scalar particle corresponding to  $\eta$  is massive with  $m_\eta^2 = -\mu^2 > 0$ , whereas the scalar particle corresponding to  $\xi$  is massless.

This particle is a prototype of a (Nambu-)Goldstone boson. It is massless because there is a direction in field space, corresponding to changing the phase, in which the potential energy does not change. Its appearance is a general feature of models with spontaneously-broken global symmetries, as proven in [13]. The total number of such massless particles corresponds in general to the number of field directions in which the potential is flat. Nambu introduced this idea into particle physics in order to describe the (relatively light) pion of QCD [11], which he identified as a (pseudo-)Goldstone boson of chiral symmetry that would have no mass if the up and down quarks were exactly massless. The simple field-theoretical model is due to Goldstone [12].

We now discuss how this spontaneous symmetry breaking of symmetry manifests itself in the presence of a U(1) gauge field [17, 19, 20]. In order to construct a theory that is invariant under local U(1) phase transformations, i.e.

$$\phi \rightarrow e^{i\alpha(x)} \phi, \quad (10)$$

we introduce a gauge field  $\mathcal{A}_\mu$  that transforms under U(1) as follows:

$$\mathcal{A}'_\mu \rightarrow \mathcal{A}_\mu + \frac{1}{q} \partial_\mu \alpha(x). \quad (11)$$

The space-time derivatives appearing in the kinetic term for the scalar field  $\phi$  are replaced by covariant derivatives

$$D_\mu = \partial_\mu + iq\mathcal{A}_\mu, \quad (12)$$

where  $q$  is the conserved charge. Including kinetic terms for both the scalar field and the  $\mathcal{A}_\mu$  field:  $(1/4) F^{\mu\nu} F_{\mu\nu}$  where  $F_{\mu\nu} \equiv \partial_\nu \mathcal{A}_\mu - \partial_\mu \mathcal{A}_\nu$ , which is invariant under the U(1) gauge transformation (11), we have the Lagrangian

$$\mathcal{L} = [(\partial_\mu - iq\mathcal{A}_\mu) \phi^*][(\partial^\mu + iq\mathcal{A}^\mu) \phi] - V(\phi^* \phi) - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}, \quad (13)$$

which we now analyze.

We minimize the potential  $V(\phi)$  as before, and write the Lagrangian in terms of the perturbations around the ground state (9):

$$\begin{aligned} \mathcal{L} = & \left\{ \frac{1}{2} [(\partial^\mu \eta)(\partial_\mu \eta) - \mu^2 \eta^2] + \frac{1}{2} (\partial^\mu \xi)(\partial_\mu \xi) - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} q^2 v^2 \mathcal{A}^\mu \mathcal{A}_\mu \right\} \\ & + v q^2 A^\mu \mathcal{A}_\mu \eta + \frac{q^2}{2} \mathcal{A}^\mu \mathcal{A}_\mu \eta^2 + q (\partial^\mu \xi) \mathcal{A}_\mu (v + \eta) - q (\partial^\mu \eta) \mathcal{A}_\mu \xi \\ & - \mu^2 v \eta - \frac{\mu^2}{2} \xi^2 - \frac{\lambda}{2} [(v + \eta) + \xi^2]^2 - \frac{\mu^2 v}{2}. \end{aligned} \quad (14)$$

As before, the first three terms describe a (real) scalar particle,  $\eta$ , with mass  $\sqrt{-\mu^2}$  and a massless Goldstone boson,  $\xi$ . The fourth term describes the free U(1) gauge field. However, whereas previously the Lagrangian (13) apparently described a massless gauge boson field, we now see in the spontaneously-broken phase (14) a term proportional to  $\mathcal{A}_\mu \mathcal{A}^\mu$ , corresponding to a mass for the gauge field:

$$m_{\mathcal{A}} = qv, \quad (15)$$

that is proportional to the vacuum expectation value of the Higgs field.

The other terms in (14) describe couplings between the fields  $A^\mu$ ,  $\eta$  and  $\xi$ , including a bilinear interaction coupling  $\propto A^\mu \partial_\mu \xi$ . The correct particle interpretation of (14) is obtained by diagonalizing the bilinear terms, which is easily done by using the gauge freedom of  $\mathcal{A}_\mu$  to replace

$$\mathcal{A}_\mu \rightarrow \mathcal{A}'_\mu = \mathcal{A}_\mu + \frac{1}{qv} \partial_\mu \xi, \quad (16)$$

and making the local phase transformation

$$\phi \rightarrow \phi' = e^{-i\xi(x)/v} \phi = \frac{v + \eta}{\sqrt{2}}. \quad (17)$$

Following this transformation, the field  $\xi$  disappears, and (14) takes the simple form

$$\mathcal{L} = \frac{1}{2} [(\partial^\mu \eta)(\partial_\mu \eta) - \mu^2 \eta^2] - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \frac{q^2 v^2}{2} \mathcal{A}^{\mu'} \mathcal{A}'_\mu + \dots \quad (18)$$

where the  $\dots$  represent trilinear and quadrilinear interactions.

The Goldstone boson  $\xi$  that appeared when the global U(1) symmetry was broken spontaneously by the choice of ground state when  $\mu^2 < 0$  has been absorbed (or ‘eaten’) by the gauge field  $\mathcal{A}_\mu$ , which thereby acquired a mass. Remember that, whereas a massless gauge boson has only two degrees of freedom (transverse polarization states), a massive gauge boson has a third (longitudinal) polarization state that is supplied by the Goldstone boson of the spontaneously-broken U(1) global symmetry. This is the Englert-Brout-Higgs mechanism.

In order for this mechanism to work, the magnitude of the v.e.v. of the scalar field must be fixed dynamically, which occurs in this model because the potential varies non-trivially in the radial ( $|\phi|$ ) direction. The mass term for the  $\eta$  field in (18) is a reflection of this variation in the potential. The appearance of such a massive scalar boson is an unavoidable signature of spontaneous symmetry breaking.

## 1.5 Spontaneous Symmetry Breaking in the Standard Model

As already mentioned, the gauge group of the Standard Model is  $SU(2)_L \times U(1)_Y$  [3, 4, 7], and the Lagrangian can be written in the form

$$\mathcal{L} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{leptons}}$$

$$\begin{aligned}\mathcal{L}_{\text{gauge}} &= -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} - \frac{1}{4}f_{\mu\nu} f^{\mu\nu} \\ \mathcal{L}_{\text{leptons}} &= \bar{\mathbf{R}} \left( \partial_\mu + i\frac{g'}{2}\mathcal{A}_\mu Y \right) \mathbf{R} + \bar{\mathbf{L}} i\gamma^\mu \left( \partial_\mu + i\frac{g'}{2}\mathcal{A}_\mu Y + i\frac{g}{2}\boldsymbol{\tau} \cdot \mathbf{B}_\mu \right) \mathbf{L},\end{aligned}\quad (19)$$

where the field-strength tensors,  $F_{\mu\nu}$  and  $f_{\mu\nu}$ , were defined in (2) and (3), respectively,  $g$  is the SU(2) coupling and  $g'$  is the U(1) hypercharge coupling. The symbol  $\mathbf{L}$  represents doublets of left-handed fermions and  $\mathbf{R}$  represents right-handed fermions. As written in (19), the theory contains four massless bosons ( $\mathcal{A}_\mu, B_\mu^1, B_\mu^2, B_\mu^3$ ).

We now introduce a scalar field that is a complex doublet of SU(2) [3,4]:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \quad (20)$$

and add to the Lagrangian

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi), \quad (21)$$

with an effective potential of similar form to (5):

$$V(\phi^\dagger \phi) = \mu^2 (\phi^\dagger \phi) + \lambda (\phi^\dagger \phi)^2, \quad (22)$$

with  $\mu^2 < 0$  and  $\lambda > 0$ . We also include Yukawa interactions between this scalar field and the matter fermions:

$$\mathcal{L}_{\text{Yukawa}} = -G_e \left[ \bar{\mathbf{R}} \phi^\dagger \mathbf{L} + \bar{\mathbf{L}} \phi \mathbf{R} \right], \quad (23)$$

which yield masses for the matter fermions, as we see later.

As in the U(1) case when  $\mu^2 < 0$ ,  $\langle \phi \rangle = 0$  is an unstable local maximum of the effective potential. The minimum has  $\langle \phi \rangle \neq 0$  with an arbitrary SU(2)  $\times$  U(1) orientation, leading to spontaneous symmetry breaking. Minimizing the effective potential as in the U(1) case, we obtain

$$\frac{\partial}{\partial (\phi^\dagger \phi)} V(\phi^\dagger \phi) = \mu^2 + 2\lambda \langle \phi \rangle_0 = \mu^2 + 2\lambda \left[ (\phi_{\text{vac}}^+)^2 + (\phi_{\text{vac}}^0)^2 \right] = 0. \quad (24)$$

Without loss of generality, we may set  $\phi_{\text{vac}}^+ = 0$  and take  $\phi_{\text{vac}}^0 = \sqrt{-\mu^2/(2\lambda)}$ . This choice breaks both the SU(2) and U(1) symmetries, but preserves invariance under a residual U(1) gauge symmetry that we may identify with electromagnetism. Since three of the four generators are broken spontaneously by the v.e.v., the spectrum of the global theory would contain three massless (Nambu-)Goldstone bosons.

To see how these are ‘eaten’ by three of the gauge bosons, we consider perturbations around the chosen vacuum, representing the scalar field as

$$\phi = \exp\left(\frac{i\xi \cdot \boldsymbol{\tau}}{2v}\right) \begin{pmatrix} 0 \\ (v + \eta)/\sqrt{2} \end{pmatrix}. \quad (25)$$

Just as in the U(1) case discussed in the previous section where we rotated away the Goldstone boson  $\xi$ , we are able in this case to make the following gauge transformation on the scalar  $\phi$  and the gauge and matter fields:

$$\phi \rightarrow \phi' = \exp\left(\frac{-i\xi \cdot \boldsymbol{\tau}}{2v}\right) \phi = \begin{pmatrix} 0 \\ (v + \eta)/\sqrt{2} \end{pmatrix}. \quad (26)$$

$$\boldsymbol{\tau} \cdot \mathbf{B}_\mu \rightarrow \boldsymbol{\tau} \cdot \mathbf{B}'_\mu \quad (27)$$

$$\mathbf{L} \rightarrow \mathbf{L}' = \exp\left(\frac{-i\xi \cdot \boldsymbol{\tau}}{2v}\right) \mathbf{L}, \quad (28)$$

where  $\tau$  is an SU(2) matrix, which leaves  $\mathcal{A}_\mu$  and  $\mathbf{R}$  invariant.

In this unitary gauge, and henceforward simplifying the notation:  $\phi' \rightarrow \phi$ , etc., so that

$$\phi = \begin{pmatrix} 0 \\ (v + \eta) / \sqrt{2} \end{pmatrix}, \quad (29)$$

and

$$\phi^\dagger \phi = \left( \frac{v + \eta}{\sqrt{2}} \right)^2, \quad (30)$$

we see that

$$V(\phi^\dagger \phi) = \mu^2 \left( \frac{v + \eta}{\sqrt{2}} \right)^2 + \lambda \left( \frac{v + \eta}{\sqrt{2}} \right)^4. \quad (31)$$

We also have

$$D_\mu \phi = \partial_\mu \phi + \frac{ig'}{2} \mathcal{A}_\mu Y \phi + \frac{ig}{2} \tau \cdot \mathbf{B}_\mu \phi, \quad (32)$$

which may be written in the form

$$D_\mu \phi = \begin{pmatrix} \frac{ig}{2} \left( \frac{v + \eta}{\sqrt{2}} \right) (B_\mu^1 - iB_\mu^2) \\ \frac{1}{\sqrt{2}} \partial_\mu \eta + \left( \frac{v + \eta}{\sqrt{2}} \right) \frac{i}{2} (ig' \mathcal{A}_\mu - igB_\mu^3) \end{pmatrix} \quad (33)$$

and hence

$$(D^\mu \phi)^\dagger (D_\mu \phi) = \frac{g^2}{8} (v + \eta)^2 |B_\mu^1 - iB_\mu^2|^2 + \frac{1}{2} (\partial_\mu \eta) (\partial^\mu \eta) + \frac{1}{8} (v + \eta)^2 (g' \mathcal{A}_\mu - gB_\mu^3)^2. \quad (34)$$

The final form of the scalar Lagrangian is therefore

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} &= \left\{ \frac{1}{2} (\partial_\mu \eta) (\partial^\mu \eta) - \frac{\mu^2}{2} \eta^2 + \frac{v^2}{8} [g^2 |B_\mu^1 - iB_\mu^2|^2 + (g' \mathcal{A}_\mu - gB_\mu^3)^2] \right\} \\ &+ \left\{ \frac{1}{8} (\eta^2 + 2v\eta) [g^2 |B_\mu^1 - iB_\mu^2|^2 + (g' \mathcal{A}_\mu - gB_\mu^3)^2] \right\} \\ &- \left\{ \frac{1}{4} \eta^4 - \lambda v \eta^3 - \frac{\mu^2}{2} \eta^2 - (\lambda v^3 + \mu^2 v) \eta - \left( \frac{\lambda v^4}{4} + \frac{\mu^2 v^2}{2} \right) \right\}, \quad (35) \end{aligned}$$

whose interpretation we now discuss.

The second term on the first line of (35) is a mass term for the  $\eta$  field: this is the Higgs boson, which appears in the same way as in the previous U(1) case. *A priori*, there is no theoretical prediction within the Standard Model for the Higgs mass

$$m_H = -2\mu^2, \quad (36)$$

since  $\mu$  is not determined by any of the known parameters of the Standard Model. The following terms on the first line of (35) are mass terms for the massive vector bosons, to which we return later. The second line includes interactions of the Higgs boson with these massive gauge bosons, and the last line describes self-interactions of the Higgs boson.

We define the charged gauge fields  $W_\mu^\pm$  as the combinations

$$W_\mu^\pm \equiv \frac{B_\mu^1 \mp iB_\mu^2}{\sqrt{2}}, \quad (37)$$

and identify the following neutral gauge boson mass eigenstates:

$$Z_\mu \equiv \frac{-g' \mathcal{A}_\mu + gB_\mu^3}{\sqrt{g^2 + g'^2}}, \quad (38)$$

$$A_\mu \equiv \frac{g\mathcal{A}_\mu + g'B_\mu^3}{\sqrt{g^2 + g'^2}}. \quad (39)$$

Substituting these expressions in the Lagrangian (35), we find

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} = & \left[ \frac{1}{2} (\partial^\mu \eta) (\partial_\mu \eta) - \frac{\mu^2}{2} \eta^2 \right] + \frac{v^2 g^2}{8} W^{+\mu} W_\mu^+ + \frac{v^2 g^2}{8} W^{-\mu} W_\mu^- + \frac{(g^2 + g'^2) v^2}{8} Z^\mu Z_\mu \\ & + \dots, \end{aligned} \quad (40)$$

and it is evident that the field  $\mathcal{A}_\mu$  is massless. This is due to the unbroken U(1) symmetry (i.e., the symmetry under  $e^{iQ\alpha(x)}$  rotations) that we identify with electromagnetism. On the other hand, the charged vector bosons  $W^\pm$  and the neutral vector boson  $Z^0$  have masses

$$m_W = \frac{gv}{2}, \quad m_Z = \frac{v}{2} \sqrt{g^2 + g'^2}. \quad (41)$$

We see that the couplings of the Higgs boson to the  $W^\pm$  and  $Z^0$  in (35) are related to these masses. They are related through

$$m_Z = m_W \sqrt{1 + g'^2/g^2}. \quad (42)$$

It is convenient to introduce the angle  $\theta_W$  to parametrize the mixing of the neutral gauge bosons, defined by

$$\tan \theta_W = \frac{g'}{g}, \quad (43)$$

so that

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}. \quad (44)$$

Eqs. (38) and (39) can then be written as

$$Z_\mu = -\sin \theta_W \mathcal{A}_\mu + \cos \theta_W B_\mu^3, \quad (45)$$

$$A_\mu = \cos \theta_W \mathcal{A}_\mu + \sin \theta_W B_\mu^3, \quad (46)$$

and the relation (42) between the masses of  $W^\pm$  and  $Z^0$  becomes

$$m_W = m_Z \cos \theta_W. \quad (47)$$

The ratio

$$\rho \equiv \frac{m_W^2}{m_Z^2 \cos^2 \theta_W} \quad (48)$$

is equal to unity at the tree level in the Standard Model. This is a direct consequence of the choice of isospin 1/2 for the Higgs field (20).

This choice also enables the Higgs field to give masses to the Standard Model fermions [3], as we now discuss. Looking at the fermion Lagrangian (23) in the unitary gauge, it becomes

$$\mathcal{L}_{\text{Yukawa}} = -G_e \left[ \bar{e}_R \phi^\dagger \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} + (\bar{\nu}_L \bar{e}_L) \phi e_R \right] = -G_e \frac{v + \eta}{\sqrt{2}} (\bar{e}_R e_L + \bar{e}_L e_R). \quad (49)$$

In terms of  $\bar{e} \equiv (\bar{e}_R, \bar{e}_L)$  and  $e \equiv (e_L, e_R)^T$ , (49) becomes

$$\mathcal{L}_{\text{Yukawa}} = -\frac{G_e v}{\sqrt{2}} \bar{e} e - \frac{G_e \eta}{\sqrt{2}} \bar{e} e, \quad (50)$$

and we see that the electron has a mass

$$m_e = G_e \frac{v}{\sqrt{2}}. \quad (51)$$

The same holds for all the Standard Model fermions, and their couplings to the Higgs boson are proportional to their masses. Thus, the Higgs boson prefers to decay into the heaviest fermions  $f$  that are kinematically accessible, i.e., have  $m_f < m_H/2$ .

At first sight, the Lagrangian (35) may look rather artificial. However, as described above, the spontaneous symmetry breaking mechanism for giving masses to vector bosons and fermions is in fact very generic. It suffices to have a scalar field  $\phi$  and choose the coefficient  $\mu^2$  of the quadratic term in its effective potential to be negative, and the mass-generation mechanism follows automatically. The original Lagrangian (21) is in fact very symmetric, and this symmetry is still present, though hidden, in (35).

Since the theory (35) still possesses symmetry, it is a renormalizable theory at the quantum level [25], which enables many detailed calculations to be compared with precise experimental measurements. In fact, not only is it *a* renormalizable theory, it is *the only* way to construct a renormalizable theory of interacting massive vector bosons [26–28]. In order to get some flavour why this is the case, consider  $W^+W^- \rightarrow W^+W^-$  scattering. At the tree level, the combination of  $\gamma$  and  $Z^0$  exchanges in the direct and crossed channels with the point-like quartic coupling, shown in the two upper rows of Fig. 2, yields a scattering amplitude that grows quadratically with energy:

$$\mathcal{M}_V = -g^2 \frac{E^2}{m_W^2} + \mathcal{O}(E^0). \quad (52)$$

This is a problem when one calculates loop diagrams, since the integral over the loop momenta is uncontrollably divergent. If one now includes a scalar with a coupling  $g_{HWW}$  to the vector bosons, the direct and crossed-channel scalar exchanges, shown in the bottom row of Fig. 2, yield an amplitude

$$\mathcal{M}_S = +g_{HWW}^2 \frac{E^2}{m_W^2} + \mathcal{O}(E^0). \quad (53)$$

Adding these contributions, we find

$$\mathcal{M} = \mathcal{M}_V + \mathcal{M}_S = \frac{m_H^2}{v^2} \left( 2 + \frac{m_H^2}{s - m_H^2} + \frac{m_H^2}{t - m_H^2} \right) + \dots, \quad (54)$$

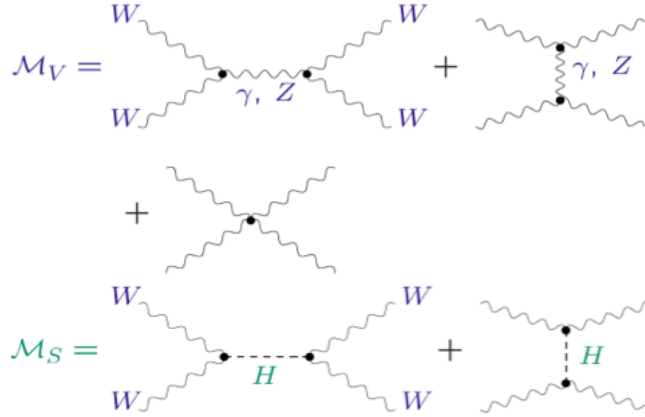
where the  $\dots$  represent terms that are subdominant at high energies, *iff* the  $HWW$  coupling coincides with the prediction from spontaneous symmetry breaking. The fact that the resultant amplitude  $\mathcal{M}$  is asymptotically constant for this particular choice of  $g_{HWW}$  ensures that the integration over the loop momentum is controllable and permits the theory to be renormalizable<sup>3</sup> - which it is if the other particle couplings also coincide with the spontaneously-broken gauge theory.

## 1.6 A Phenomenological Profile of the Higgs Boson

### 1.6.1 Before the LHC

For a decade after the original papers after spontaneously-broken gauge theories were formulated, and even for several years after they were proven to be renormalizable, rather few people took seriously Higgs' prediction of the boson that bears his name. Indeed, only a handful of papers discussed its possible experimental signatures before the paper that Mary Gaillard, Dimitri Nanopoulos and I wrote in 1975 [29] with the same title as this Section heading. In the last sentence of our paper we wrote that "*we do not want to encourage big experimental searches for the Higgs boson*". Fortunately, the experimental community did not pay attention to this caveat, and the ATLAS and CMS experiments announced the discovery of a candidate for a (the?) Higgs boson on July 4th 2012 [1, 2].

<sup>3</sup>A similar argument applies to fermion- $W$  scattering and the scalar-fermion coupling, which must also coincide with the prediction from spontaneous symmetry breaking.

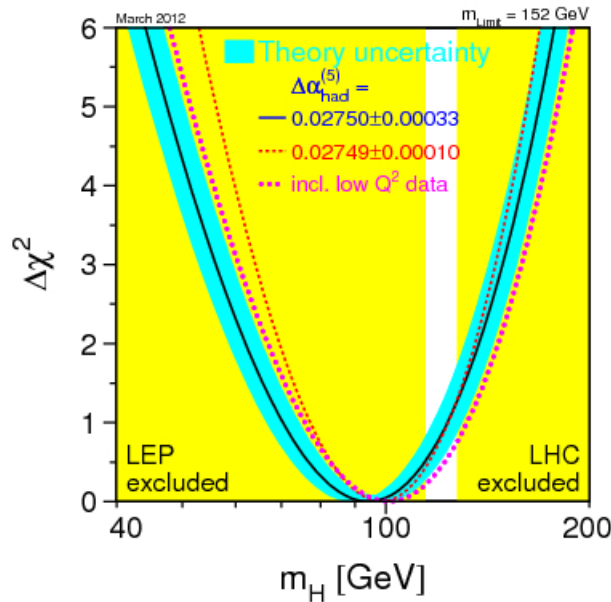


**Fig. 2:** Contributions to  $WW$  scattering from diagrams due to vector boson exchange and the four-point gauge interaction in the absence of a Higgs boson (upper two rows), and diagrams due to Higgs boson exchange (lowest row).

The search for the Higgs boson at LEP was advertized in the first survey of LEP physics made in 1976 [30] and featured strongly in the subsequent LEP experimental programme [31]. These direct searches for the Higgs boson resulted in the lower limit [32]

$$m_H > 114.4 \text{ GeV}, \tag{55}$$

shown as the left-hand of the two yellow excluded regions in Fig. 3. The search for the Higgs boson was also advertized at the first LHC physics workshop in 1984 and grew subsequently to become one of the major objectives of the LHC experimental programme.



**Fig. 3:** The status of the Higgs search in March 2012 [8]. The left-hand yellow-shaded region is the LEP exclusion, and the right-hand yellow-shaded region is the Tevatron exclusion at that time [33].

Although LEP found no direct evidence for the Higgs boson, the precision of electroweak measurements at LEP and elsewhere provided indirect indications on the Higgs mass through the sensitivity of electroweak observables to quantum loop corrections. For example, there is a one-loop correction  $\Delta r$

to the  $W^\pm$  and  $Z^0$  masses:

$$m_W^2 \sin^2 \theta_W = m_Z^2 \cos^2 \theta_W \sin^2 \theta_W = \frac{\pi\alpha}{\sqrt{2}G_F} (1 + \Delta r). \quad (56)$$

The one-loop correction  $\Delta r$  and other electroweak radiative corrections are in turn sensitive to the masses of heavy virtual particles, in particular  $m_t$  and  $m_H$ :

$$\propto \frac{3G_F}{8\pi^2\sqrt{2}}m_t^2, \frac{\sqrt{2}G_F}{16\pi^2} \left( \frac{11}{3} \ln \frac{m_H^2}{m_Z^2} + \dots \right) \quad (57)$$

in the limits of large  $m_t$  and  $m_H \gg m_{W,Z}$ . Note that the sensitivity to  $m_t$  is quadratic [34], whereas that to  $m_H$  is only logarithmic [35]. The large-mass divergences in (57) reflect the fact that the electroweak theory would become renormalizable if these particles were absent. In the case of the top quark, its absence would leave us with an incomplete fermion doublet, and the problems arising in the absence of the Higgs boson were discussed at the end of the previous Section.

First attempts to use precision measurements to constrain  $m_H$  were made before the discovery of the top quark [36], and already indicated that  $m_H = \mathcal{O}(m_W)$ . The discovery of the top quark with a mass consistent with predictions based on electroweak data and (56) enabled the prediction of  $m_H$  to be sharpened, as has the inclusion of QCD and higher-order electroweak effects, as illustrated by the blue band in Fig. 3.

A conservative estimate of the current estimate of  $m_H$  on the basis of precision electroweak data alone is [8, 9]

$$m_H = 100 \pm 30 \text{ GeV}, \quad (58)$$

which is quite compatible with the direct lower limit (55) and the exclusion by the Tevatron of  $m_H \sim 160$  to 170 GeV. Combining the LEP and Tevatron exclusions with the precision electroweak data led in mid-2011 to the prediction [37]:

$$m_H = 125 \pm 10 \text{ GeV}. \quad (59)$$

This has been an impressively successful prediction!

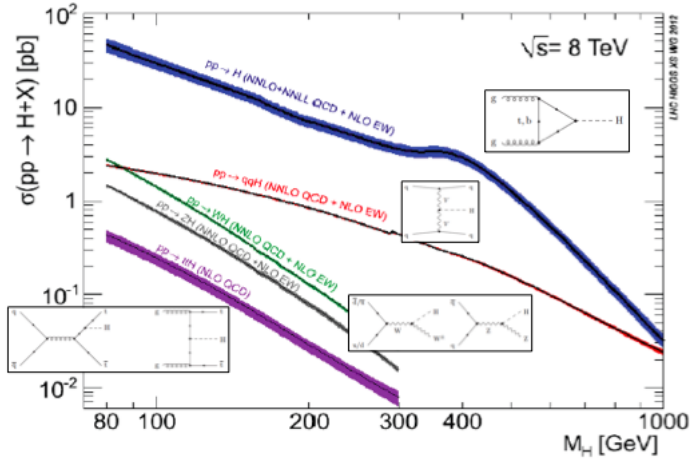
### 1.6.2 Higgs Production at the LHC

Several production modes are measurable at the LHC for a Higgs boson weighing  $\sim 125$  GeV, as displayed in Fig. 4 for Higgs production at the LHC at 8 TeV. The dominant mechanism is calculated to be gluon-gluon fusion:  $gg \rightarrow H$  via loops of heavy coloured particles [38], of which the most important in the Standard Model is the top quark. Since the leading-order contribution to the production amplitude is  $\mathcal{O}(\alpha_s)$ , making an accurate calculation is particularly challenging. However, the cross section has been calculated by different theoretical groups at next-to-next-to-leading order (NNLO), including also the most important logarithms at higher orders (NNLL), with good agreement between the calculations as seen in Fig. 5 [39]. The main remaining uncertainties in the calculation are associated with the choice of renormalization scale in the calculation, and in the gluon parton distribution function within the proton. The overall theoretical uncertainty is currently estimated to be about 10%, and a significant rise in the cross section is expected when the LHC reaches 13/14 TeV in 2015.

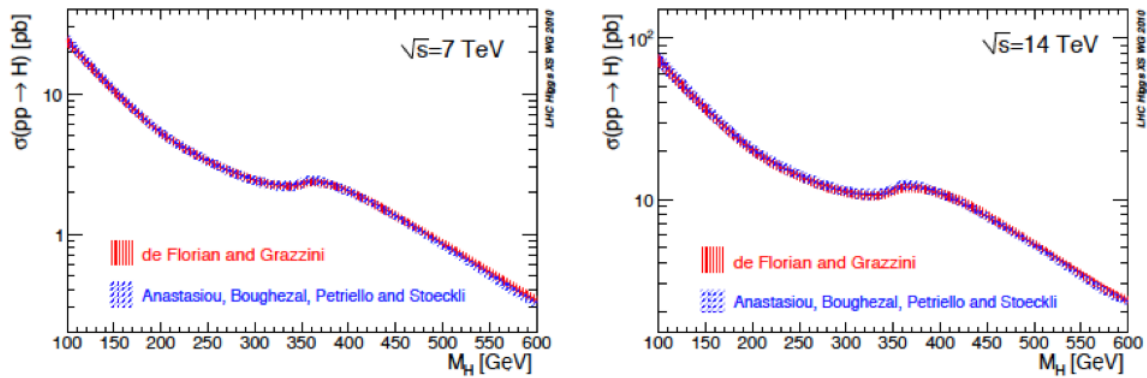
The second-largest contribution to Higgs production at the LHC is due to vector-boson fusion (VBF) [40]. This has now been calculated at NNLO in  $\alpha_s$  and including electroweak corrections at NLO, and the perturbation expansion is converging well. In this case, the uncertainties in quark parton distribution functions are relatively small, so this cross section is known more accurately, and it grows with energy more rapidly than gluon-gluon fusion [39].

The third-largest contribution is associated production of the Higgs with a massive vector boson  $V = W^\pm, Z^0$  [41]. This has also been calculated at NNLO in  $\alpha_s$  and including electroweak corrections

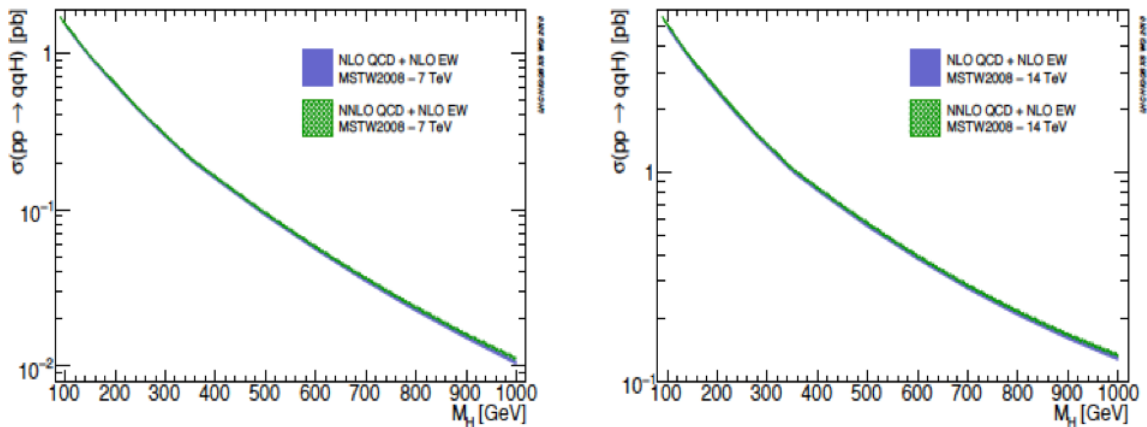




**Fig. 4:** The principal Higgs production cross sections at the LHC at 8 TeV. The insets depict the corresponding fundamental production subprocesses.

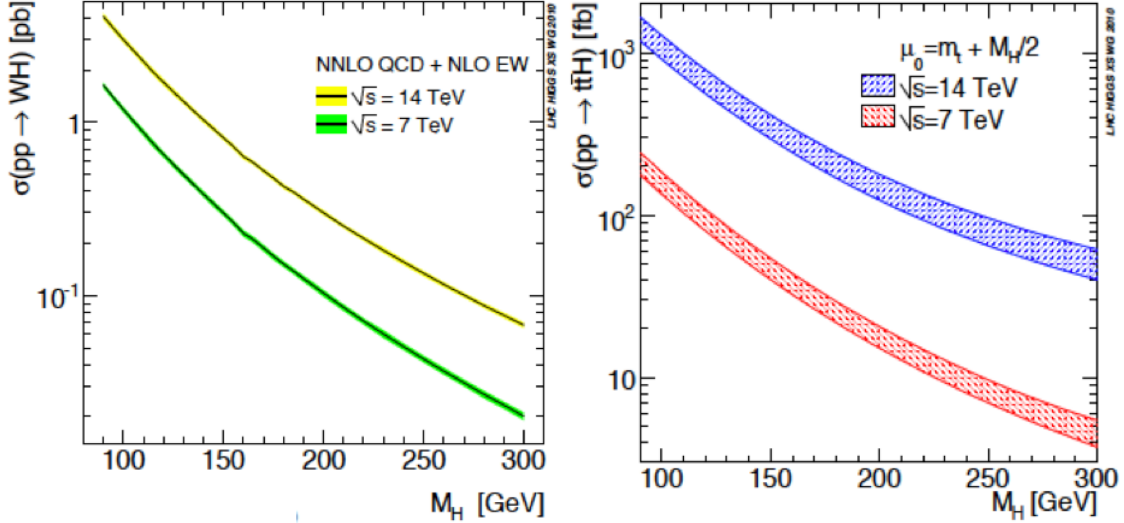


**Fig. 5:** The cross sections for  $gg \rightarrow H$  production at the LHC at 7 and 14 TeV, comparing two next-to-next-leading order (NNLO) calculations that also include leading higher-order logs (NNLL) [39].



**Fig. 6:** The cross sections for VBF production of the  $H$  at the LHC at 7 and 14 TeV, comparing calculations at NLO in the electroweak interactions and at NLO and NNLO in QCD [39].

at NLO, as seen in the left panel of Fig. 7. and the perturbation expansion again converges well. Here the rate of growth of the cross section is less rapid.



**Fig. 7:** The cross sections for production of the  $H$  in association with a massive vector boson  $V$  (left panel) and in association with  $\bar{t}t$  (right panel) at the LHC at 7 and 14 TeV [39].

The next contribution is from associated  $\bar{t}t + H$  production, which is currently known less accurately: it has been calculated at NLO in  $\alpha_s$ , so there are larger uncertainties in the perturbation expansion, and the choice of parton distributions is also an important uncertainty [39]. This process has the most rapid cross section increase with energy, as seen in the right panel of Fig. 7, offering interesting prospects for measurement at LHC 13/14.

Finally, interest has recently been attracted by  $H$  production in association with a single  $t$  or  $\bar{t}$  [42]. This has a relatively small cross section in the Standard Model, but it may be enhanced or suppressed significantly in models where the  $H\bar{t}t$  coupling differs from its Standard Model value [43, 44], as we discuss later.

LHC experimentalists are indeed fortunate that all of these mechanisms are potentially measurable at the LHC for  $m_H \sim 125$  GeV! This would not have been the case if the Higgs mass had been 400 GeV, say, in which case only  $gg \rightarrow H$  and VBF could have been measurable.

### 1.6.3 Higgs Decays [45]

The Higgs decay rate into a pair of fermions is given at the tree level by

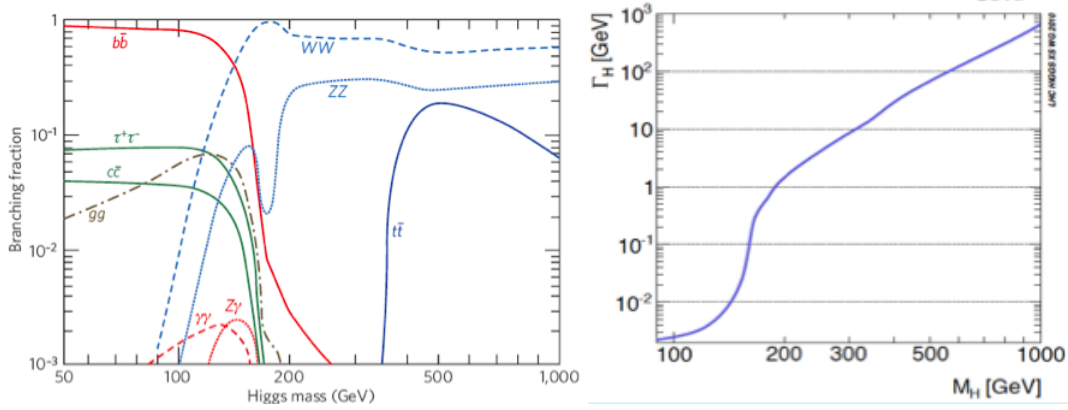
$$\Gamma(H \rightarrow \bar{f}f) = N_c \frac{G_F m_H}{4\pi\sqrt{2}} m_f^2, \quad (60)$$

where  $N_c = 3(1)$  for decays into quarks (leptons). Since the tree-level Higgs couplings to other particles are proportional to their masses (squared in the cases of massive vector bosons), the dominant Higgs decays are into the heaviest particles that are kinematically accessible. This implies that the dominant two-body decays of the Higgs weighing  $\sim 125$  GeV are expected to be into  $\bar{b}b$ ,  $\bar{c}c$  and  $\tau^+\tau^-$ , as seen in the left panel of Fig. 8. However, decays into  $\bar{b}b$  have yet to be confirmed, there is no evidence for  $\bar{c}c$ , and  $H \rightarrow \tau^+\tau^-$  decay has only recently been observed unambiguously [46].

The  $H$  decay rate into a pair of  $W^\pm$  bosons, one on- and one off-shell, is given by

$$\Gamma(H \rightarrow WW^*) = \frac{G_F m_H^3}{8\pi\sqrt{2}} F(r), \quad (61)$$

where  $F(r \equiv m_W/m_H)$  is a kinematic factor, and the corresponding decay into a pair of  $Z$  bosons is given by a similar formula with  $m_W \rightarrow m_Z$  and a symmetry factor of  $1/2$  [45]. The decays  $H \rightarrow$  virtual



**Fig. 8:** The most important decay branching fractions for the decays of a Standard Model Higgs boson (left panel), and the total decay rate  $\Gamma_H$  (right panel) [45].

$W^+W^-$  and  $Z^0Z^0$  are important for  $m_H \sim 125$  GeV, as seen in the left panel of Fig. 8, despite the fact that  $m_H < 2m_W$  and  $2m_Z$ .

Moreover, although the decays  $H \rightarrow gg$  and  $\gamma\gamma$  are absent at the tree level, they are generated by quantum loops, as discussed above in connection with  $gg \rightarrow H$  production. The dominant contributions to the  $H \rightarrow \gamma\gamma$  decay amplitude are due to massive charged particles [29], the most important in the Standard Model being the  $t$  quark and the  $W^\pm$  boson, whose contributions interfere destructively. At the one-loop level

$$\Gamma(H \rightarrow \gamma\gamma) = \frac{G_F \alpha^2 m_H^3}{128\pi^3 \sqrt{2}} \left| \sum_f N_c Q_f^2 A_{1/2}(r_f) + A_1(r_W) \right|^2, \quad (62)$$

where  $A_{1/2}$  and  $A_1$  are known functions of  $r_f \equiv m_f/m_H$  and  $r_f \equiv m_W/m_H$  that have opposite signs [45].

Decays into strongly-interacting final states have been evaluated at NNNLO in  $\alpha_s$ , while electroweak decays have been evaluated at NLO. The total Higgs decay rate in the Standard Model is expected to be  $\sim 4.2$  MeV for  $m_H \sim 125$  GeV, as seen in the right panel of Fig. 8 [45].

Once again, Nature has been kind in her choice of the Higgs mass, with half-a-dozen Higgs decays being observable at the LHC for  $m_H \sim 125$  GeV. If the Higgs mass had been 300 GeV, say, only the decays  $H \rightarrow W^+W^-$  and  $Z^0Z^0$  would have been measurable.

#### 1.6.4 From Discovery to Measurement

Following the dramatic announcements of the discoveries of the Higgs boson by the ATLAS and CMS Collaborations on July 4th, 2012 [1, 2], the emphasis is now on measurements and the information they provide about physics within and beyond the Standard Model.

The strengths of the signals observed in many channels are compatible with the Standard Model predictions [47]. However, the bulk of the evidence concerns production by gluon-gluon fusion. There have been several observations of VBF channels at the  $2\text{-}\sigma$  level, and the overall significance of the evidence for VBF is some  $3\sigma$ . So far the best evidence for production in association with massive vector bosons comes from the CDF and D0 experiments at the Fermilab Tevatron, and there are only upper limits on production in association with  $t\bar{t}$ .

The primary evidence for  $H$  decays is in final states involving vector bosons:  $ZZ^*$ ,  $\gamma\gamma$  and  $WW^*$ , whereas the direct evidence for decays into fermions is much weaker. Concerning leptons, although evidence for  $H \rightarrow \tau^+\tau^-$  decay is emerging [46], but there are only upper limits on the decay into

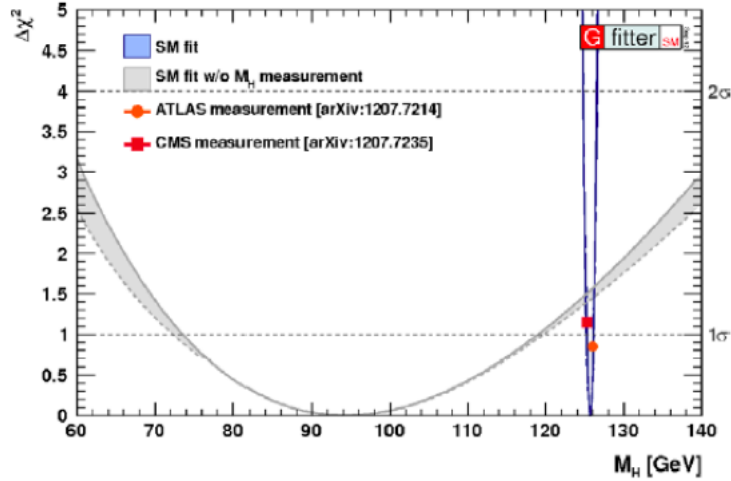
$\mu^+\mu^-$  [47] (which may be the only second-generation final state accessible at the LHC) <sup>4</sup>, and the prospects for measuring the  $He^+e^-$  coupling look very dim. Concerning quarks, evidence for  $H \rightarrow \bar{b}b$  decay is also emerging, but there is only indirect evidence for an  $Htt$  coupling via measurements of  $gg \rightarrow H$  production and  $H \rightarrow \gamma\gamma$  decay. In the future, more information could be provided by  $H \rightarrow Z^0\gamma$  decay and  $H$  production in association with a single  $t$  or  $\bar{t}$  [42–44], as well as  $H\bar{t}t$  production.

### 1.6.5 The Higgs Mass - Evidence for Physics beyond the Standard Model?

There are two ways to measure the Higgs mass accurately with the present data: using  $H \rightarrow ZZ^* \rightarrow 4\ell^\pm$  and  $\gamma\gamma$  decays. In the case of CMS, these two final states yield very similar masses, with  $m_{\gamma\gamma}$  slightly lower. In the case of ATLAS, there is some tension between the measurements in the two channels, with  $m_{\gamma\gamma}$  higher by  $\sim 2$  GeV, corresponding to a  $\sim 2\text{-}\sigma$  discrepancy. However, the CMS and ATLAS measurements [47] are quite consistent, and a naive global average is

$$m_H = 125.6 \pm 0.4 \text{ GeV} . \quad (63)$$

As seen in Fig. 9, this is consistent at the  $\Delta\chi^2 \sim 1.5$  level with the estimate of  $m_H$  provided by precision electroweak data [9]. A victory for the Standard Model at the quantum (loop) level!



**Fig. 9:** Comparison of the indirect estimate of the Higgs mass based on precision electroweak data with the direct measurement by ATLAS and CMS [9].

However, issues arise when we consider the effective Higgs potential. There are two important sources of renormalization of the quartic Higgs self-coupling  $\lambda$ : that due to the Higgs self-coupling itself:

$$\lambda(Q) = \frac{\lambda(v)}{1 - \frac{3}{4\pi^2} \lambda(v) \ln \frac{Q^2}{v^2}} + \dots , \quad (64)$$

where  $Q$  is some renormalization scale above the electroweak scale  $v$ , and that due to the  $H\bar{t}t$  coupling:

$$\lambda(Q) = \lambda(v) - \frac{3m_t^4}{4\pi^2 v^4} \ln \frac{Q^2}{v^2} + \dots , \quad (65)$$

where in each case the  $\dots$  represent subleading terms in the solution of the renormalization-group equation. We see in (64) that the self-renormalization tends to increase  $\lambda$  as  $Q$  increases, leading to an

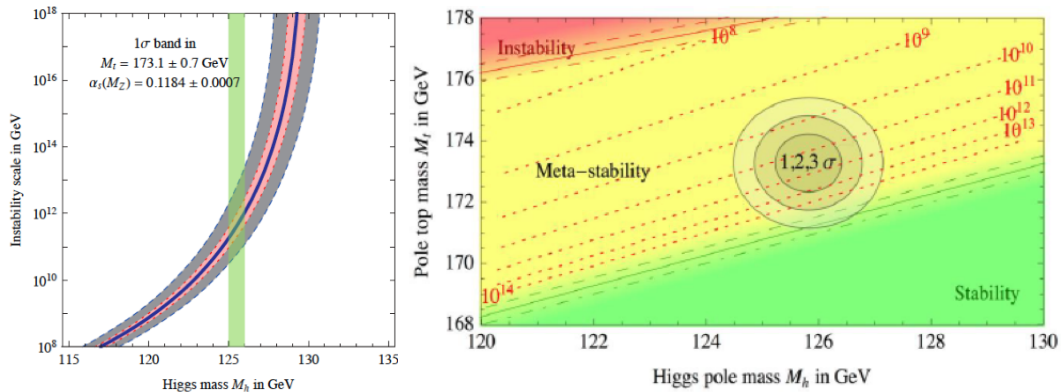
<sup>4</sup>It would also be interesting to search for the flavour-changing decays  $H \rightarrow \tau\mu$  and  $\tau e$ . These are highly suppressed in the Standard Model, but model-independent upper bounds from low-energy flavour-changing processes allow these decays to occur at rates similar to  $H \rightarrow \tau^+\tau^-$  [48].

apparent singularity (a so-called Landau pole). On the other hand, the renormalization by the  $t$  quark tends to reduce  $\lambda$  as  $Q$  increases, potentially driving it negative at some scale above  $v$ .

This would imply an instability in the electroweak vacuum if [49]

$$m_H < \left[ 129.4 + 1.4 \left( \frac{m_t - 173.1 \text{ GeV}}{0.7} \right) - 0.5 \left( \frac{\alpha_s(m_Z) - 0.1184}{0.0007} \right) \pm 1.0_{\text{TH}} \right] \text{ GeV}. \quad (66)$$

The measured values of  $m_t \sim 173 \text{ GeV}$  and  $m_H$  (63) would drive the quartic Higgs self-coupling negative at some scale  $\sim 10^{10}$  to  $10^{14} \text{ GeV}$ , as seen in the left panel of Fig. 10, if no physics beyond the Standard Model intervenes at some lower energy scale. (One example of possible new physics is supersymmetry, to which we return later.) However, the lifetime of the vacuum is estimated to be probably much longer than the age of the Universe, as seen in the right panel of Fig. 10, so it is not an immediate issue for the future of humanity, leading some people to suggest that this instability is not a problem. My own point of view is that such an instability would make it much more difficult to understand why the current vacuum energy (cosmological constant) is so close to zero in natural units. Why should our present vacuum energy be small if we are in a temporary state on the way to a state with vacuum energy much larger in magnitude than now (and negative)?



**Fig. 10:** Negative renormalization of the Higgs self-coupling by the top quark is calculated within the Standard Model to lead to an instability in the effective Higgs potential for field values  $\sim 10^{13}$  to  $10^{14} \text{ GeV}$  (left panel). The current estimates of  $m_t$  and  $m_H$  suggest that the current electroweak vacuum is in fact metastable (right panel), though a definite conclusion must wait a more accurate measurement of  $m_t$ , in particular. Figures taken from [49].

It should be emphasized, however, that the conclusion that the electroweak vacuum is unstable is not definite, even within the Standard Model. The stability or otherwise of the electroweak vacuum depends sensitively on  $m_t$  as well as  $m_H$  (and, to a lesser extent,  $\alpha_s$ ). In addition to the quoted experimental error in  $m_t$ , there is also a theoretical uncertainty associated with the way  $m_t$  is defined and introduced into experimental Monte Carlo programmes [50], which warrants more study.

### 1.6.6 The Higgs Discovery is a Big Deal

As already mentioned, without the Higgs boson (or something to replace it), there would be no atoms, because massless electrons would escape from nuclei at the speed of light without forming atoms, and the weak interactions would not be weak, everything would be radioactive, and life would be impossible. The discovery of a/the Higgs boson tells us how gauge symmetry is broken and whether there is such a thing as an elementary scalar field. It is likely to be the portal to new physics such as dark matter. The switch-on of the Higgs v.e.v. would have caused a phase transition in the Universe when it was about  $10^{-12}$  seconds old, and may have played a role then in generating the matter in the Universe via electroweak baryogenesis. A related inflaton might have made the Universe expand exponentially when it was about  $10^{-35}$  seconds old, and might contribute  $10^{60}$  too much to today's dark energy!

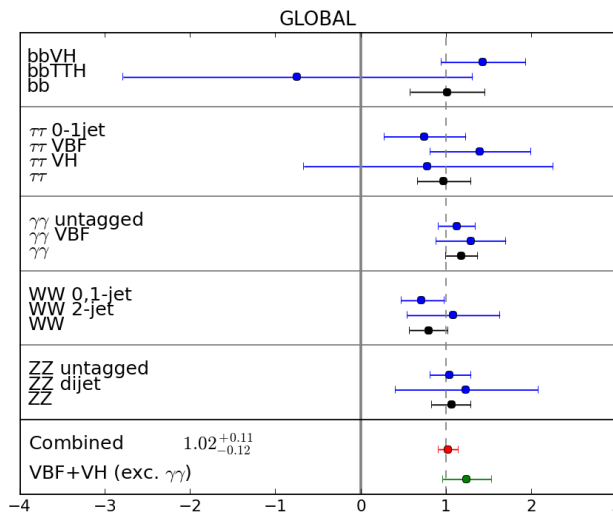
## 2 What we know now

### 2.1 The Particle Jigsaw Puzzle

Fig. 11 summarizes our state of knowledge concerning the  $H$  signal strengths in the various different final states:  $\bar{b}b, \tau^+\tau^-, \gamma\gamma, WW^*$  and  $ZZ^*$ , averaged over the results of ATLAS, CMS and the Tevatron experiments. The signal strengths in different production channels are presented (in blue), followed by the global combinations in each final state (in black). The combined mean signal strength in all channels (in red) is

$$\mu = 1.02^{+0.11}_{-0.12}, \tag{67}$$

and finally the combination of the signal strengths in the VBF and associated  $V + H$  channels is shown (in green). We see that there is no indication of any significant deviation from the Standard Model predictions: Peter Higgs should be smiling!



**Fig. 11:** A compilation of the Higgs signal strengths measured by the ATLAS, CDF, D0 and CMS Collaborations in the  $\bar{b}b, \tau^+\tau^-, \gamma\gamma, WW^*$  and  $ZZ^*$  final states. We display the combinations of the different channels for each final state, and also the combination of all these measurements, with the result for the VBF and VH channels (excluding the  $\gamma\gamma$  final state) shown separately in the bottom line. Figure taken from [51].

The current situation of particle physicists resembles that of someone who has spent more than 100 years putting together a jigsaw puzzle, and has finally (after 48 years) discovered what may be the last missing piece, hidden away in the back of the sofa and with the picture rubbed off. Have the LHC experiments really discovered the missing piece, or is it an impostor? Does it have the right shape to fit into the empty space in the puzzle, and does it have the right size?

The rest of this lecture is devoted to answering these questions as best we can on the basis of the present data.

## 2.2 Is it the Missing Piece?

### 2.2.1 Does it have Spin 0 or 2?

The question ‘does the newly-discovered  $H$  particle have the right shape to be the missing piece of the puzzle?’ can be parsed as the question ‘what is its spin?’ Since the  $H$  particle decays into pairs of photons - identical spin-1 bosons - it must have some spin  $\neq 1$ . The simplest possibilities are spin 0 and 2, the Standard Model being an example of the former case, a Kaluza-Klein graviton being an example

of the second case. A higher spin cannot be excluded *a priori*, but I am unaware of any model with spin  $> 2$ .

Several ways to diagnose the  $H$  spin have been proposed, including the characteristics of production in association with  $W^\pm$  or  $Z^0$  [52, 54], the angular distribution of  $\gamma\gamma$  decays [55, 56], and the kinematic correlations of leptons in  $WW^*$  and  $ZZ^*$  decays [57]. In general, a massive spin-2 particle has many possible couplings to Standard Model particles, and distinguishing the spin-0 and general spin-2 hypotheses is difficult. Here we consider the simplest spin-2 case (see [58] and references therein), in which it has minimal graviton-like couplings, as in simple models with extra dimensions:

$$\mathcal{L}_{int} = \sum_i \frac{c_i}{M_{eff}} G^{\mu\nu} T_{\mu\nu}^i, \quad (68)$$

where the sum is over Standard Model particle types  $i$  and the overall mass scale and the individual coefficients  $c_i$  are model-dependent.

For definiteness, we can consider warped compactifications of 5-dimensional theories [58], in which the metric takes the form:

$$ds^2 = w(z)^2 (\eta_{\mu\nu} dx^\mu dx^\nu - dz^2). \quad (69)$$

In such a scenario we expect identical coefficients for the couplings of a spin-2 particle  $X$  to the massless vector bosons  $g$  and  $\gamma$ :

$$c_g = c_\gamma = 1 / \int_{z_{UV}}^{z_{IR}} w(z) dz. \quad (70)$$

since their wave functions are uniform in the extra dimension. This implies the following simple relation between the decay rates of a spin-2 particle  $X$  into photons and gluons:

$$\Gamma(X \rightarrow gg) = 8\Gamma(X \rightarrow \gamma\gamma), \quad (71)$$

which is disfavoured by the data on the  $H(126)$  decay branching ratios, as seen in Fig. 12. The couplings of the other Standard Model particles are non-universal, reflecting their different wave functions in the extra dimension. In simple warped compactifications one expects

$$c_b \simeq c_t \gtrsim c_W \simeq c_Z = \mathcal{O}(35) \times (c_g = c_\gamma > c_{u,d}). \quad (72)$$

The experimental data on  $H(126)$  decays also disfavour the expected hierarchy (72) between  $c_W \simeq c_Z$  and  $c_g = c_\gamma$ , as also seen in Fig. 12<sup>5</sup>.

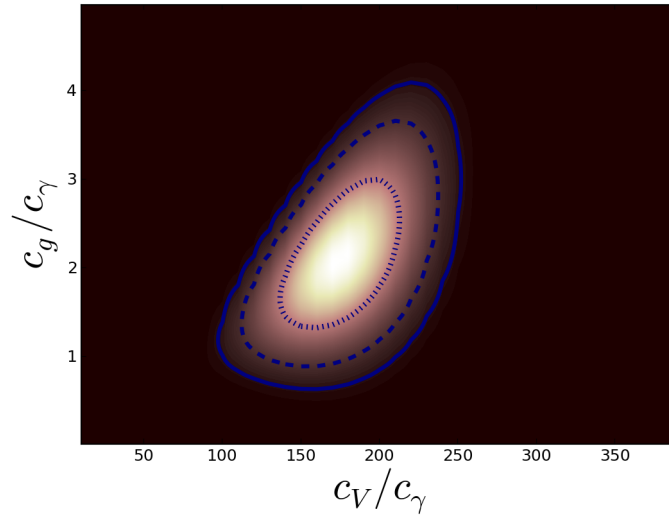
The difference between the spin-2 couplings (69) and spin-0 couplings also affect the kinematics of production in association with massive vector bosons  $W^\pm$  and  $Z^0$ . For example, the  $VX$  invariant mass distributions for spin 2 and a scalar  $0^+$  particle are very different, as seen in Fig. 13 [52], as also are the invariant mass distributions for a pseudoscalar  $0^-$  particle. The Tevatron experiments have studied the related transverse-mass distribution, as seen in Fig. 14, and have found that the spin-2 and  $0^-$  hypotheses are disfavoured at the 99% CL [53] - assuming that these experiments have indeed observed the same particle as discovered by ATLAS and CMS.

The differences in the couplings also lead to different energy dependences in the spin-2,  $0^+$  and  $0^-$  cases, as seen in the left panel of Fig. 15. If one accepts the Tevatron evidence for  $H(126)$  production, the ratio of the cross section to that at the LHC is also strong evidence against the spin-2 and  $0^-$  hypotheses, as seen in the right panel of Fig. 15.

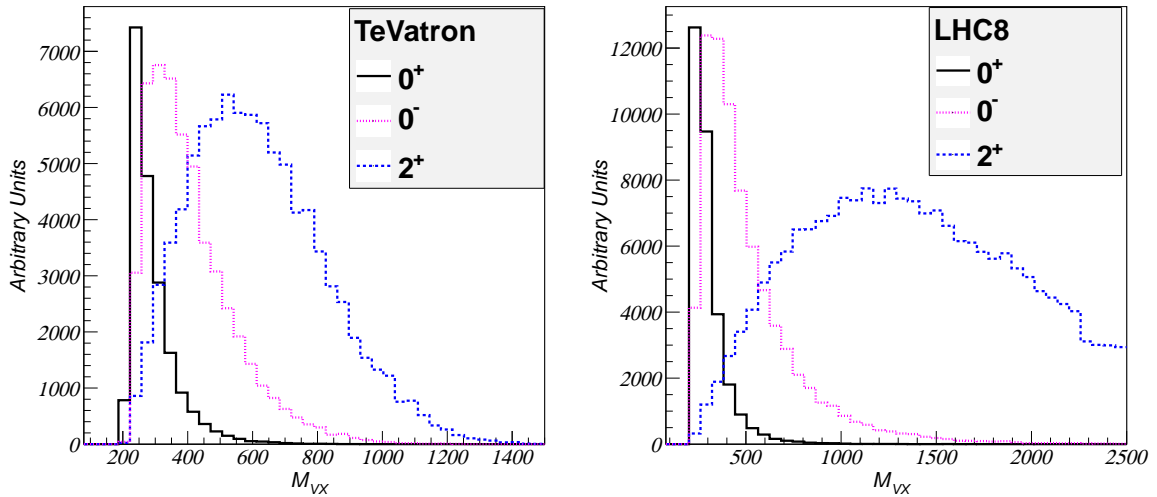
The polar angle distributions in the  $H(126)$  centre-of-mass frame under the spin-2 and  $0^-$  hypotheses are also expected to be easily distinguishable. In the case of  $gg \rightarrow H \rightarrow \gamma\gamma$  production, one

<sup>5</sup>Another problem for this scenario is that one expects, as in QCD, that the tensor boson should have a higher mass than the lightest Kaluza-Klein vector boson, which has not been seen.





**Fig. 12:** The correlation between the values of  $c_W/c_\gamma$  (horizontal axis) and  $c_g/c_\gamma$  (vertical axis) found in a global fit to the current experimental data under the spin- two hypothesis [58].

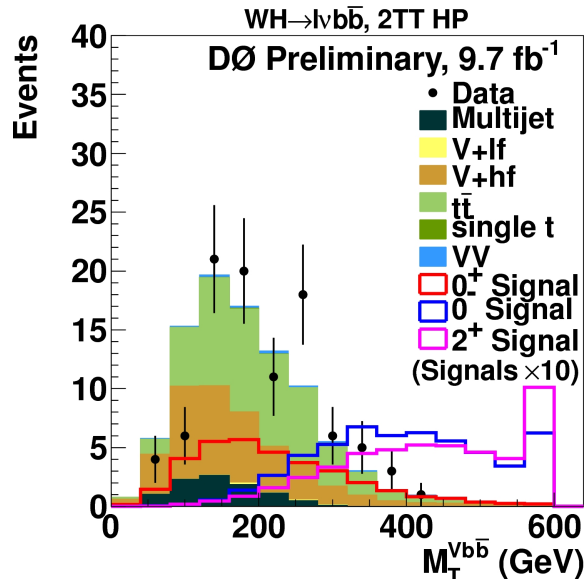


**Fig. 13:** The distributions in the  $Z + H$  invariant mass  $M_{ZH}$  for the  $0^+$  (solid black),  $0^-$  (pink dotted) and  $2^+$  (blue dashed) assignments for the  $H$  particle discovered by ATLAS [1] and CMS [2], calculated for the reaction  $\bar{p}p \rightarrow Z + H$  at the TeVatron (left) and for the reaction  $pp \rightarrow Z + H$  at the LHC at 8 TeV (right) [52].

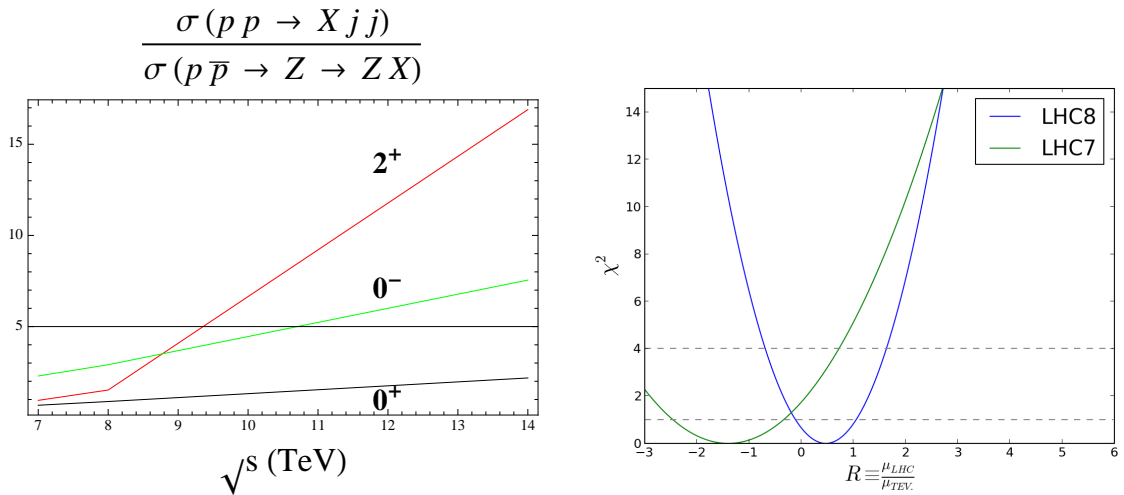
expects the initial state to be an incoherent superposition of parallel and antiparallel gluon spins along the proton-proton collision axis. This knowledge of the initial state enables the final-state  $\gamma\gamma$  polar-angle distribution to be calculated: it is expected to be non-uniform and peaked in the forward and backward directions, whereas the angular distribution would be isotropic in the spin-0 case. The ATLAS Collaboration has found that the spin-2 case is disfavoured at more than the 99% CL [59], and has extended the analysis to include an arbitrary admixture of  $\bar{q}q$  initial states (which would be suppressed in the warped compactification scenario discussed above).

The azimuthal and polar angle distributions of the charged leptons in  $H \rightarrow W^\pm W^{\mp*} \rightarrow \ell^+ \ell^- + \dots$  decays also provide significant power to distinguish between the spin-2 and -0 hypotheses [55, 57],



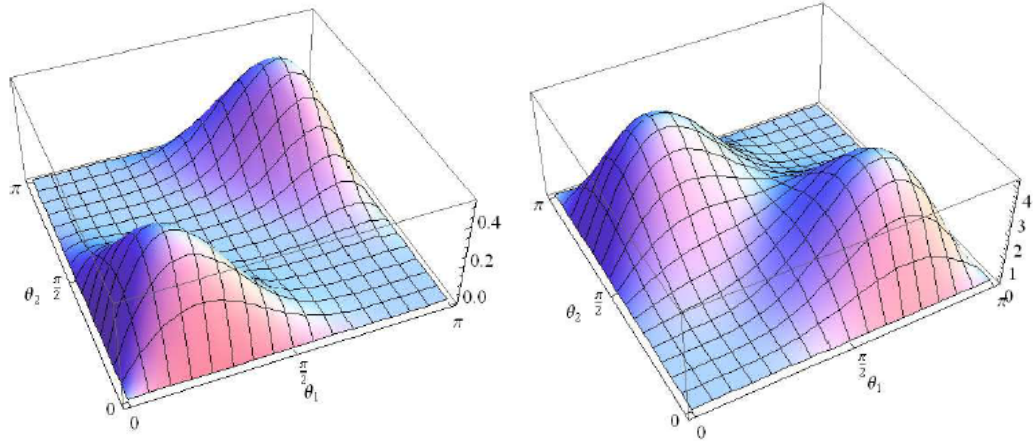


**Fig. 14:** The distribution in the  $Z + H$  transverse mass  $M_{ZH}$  measured by the D0 Collaboration compared with simulations for the  $0^+$  (red),  $0^-$  (blue) and  $2^+$  (mauve) hypotheses for the  $H$  particle [53].



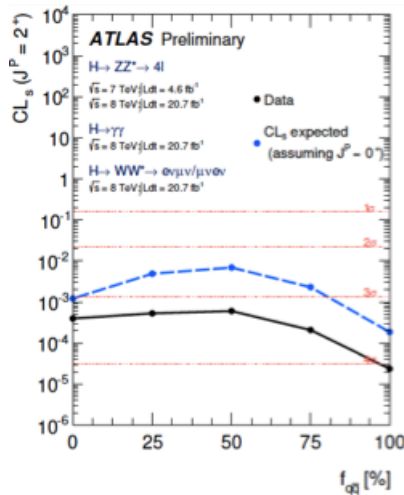
**Fig. 15:** Left - The energy dependence of the cross section for production of  $h$  in association with a  $Z$  boson under different hypotheses for the  $J^P$  of  $H$ :  $0^+$  (black),  $2^+$  (red) and  $0^-$  (green). Right - The likelihood for the ratio  $R_{\text{data}} = \mu_{\text{LHC8}} / \mu_{\text{TeVatron}}$  extracted from the experimental data at 8 TeV (blue) and 7 TeV (green). The spin-two expectations  $R_{\text{Spin } 2} = 5.4$  and  $6.7$  for 7 and 8 TeV, respectively, are excluded, and the  $0^-$  expectations  $R_{0^-} = 3.1$  and  $2.7$  for 7 and 8 TeV, respectively, are highly disfavoured, whereas the  $0^+$  expectation  $R = 1$  is quite consistent with the data [54].

as seen in Fig. 16, and also disfavour the interpretation of  $H(126)$  as a spin-2 particle with graviton-like couplings. Finally, the multiple kinematical observables in  $H \rightarrow Z^0 Z^{0*} \rightarrow 2\ell^+ 2\ell^-$  also provide many powerful ways to distinguish between different spin-parity assignments for the  $H(126)$  particle.



**Fig. 16:** Correlated distributions for the lepton polar angles in  $H \rightarrow W^\pm W^{\mp*} \rightarrow \ell^+ \ell^- + \dots$  decays for the  $0^+$  assignment (left panel) and for a graviton-like  $2^+$  particle (right panel) [55].

It is on the basis of a combination of  $\gamma\gamma$ ,  $W^\pm W^{\mp*}$  and  $Z^0 Z^{0*}$  measurements, the ATLAS Collaboration has been able to exclude the graviton-like spin-2 hypothesis at more than the 99.9% CL, as seen in Fig. 17 [59]. Peter Higgs can continue smiling!



**Fig. 17:** Combining measurements in the  $H \rightarrow \gamma\gamma, ZZ^* \rightarrow 4\ell^\pm$  and  $WW^* \rightarrow 2\ell^\pm \nu\bar{\nu}$  final states, the ATLAS Collaboration excludes the spin-2 hypothesis for  $H$  at more than the 99.9% CL for any combination of  $H$  production via  $gg$  and  $\bar{q}q$  collisions [59].

### 2.2.2 Is it Scalar or Pseudoscalar?

As has been discussed above, many of the analyses that discriminate between the spin-2 and  $0^+$  hypotheses can also be used to discriminate between  $0^-$  and  $0^+$ . For example, the  $V + H$  invariant mass distributions in associated production are different, and already disfavour  $0^-$  quite strongly. The angu-

lar distribution of the  $\gamma\gamma$  final state does not distinguish, but the angular and kinematic distributions in  $H \rightarrow W^\pm W^\mp \rightarrow \ell^+ \ell^- + \dots$  and  $H \rightarrow Z^0 Z^{0*} \rightarrow 2\ell^+ 2\ell^-$  do offer discrimination between  $0^+$  and  $0^-$  [57]. Thus the possibility of a pure  $0^-$  spin-parity assignment can also be excluded beyond the 99% CL [59].

On the other hand, in the presence of CP violation the  $H$  particle could decay as a mixture of scalar and pseudoscalar, and the fractions could be different in different final states. At the moment, the admixture of a substantial fraction of pseudoscalar final states in  $H \rightarrow W^\pm W^\mp$  and  $H \rightarrow Z^0 Z^{0*}$  decays cannot be excluded. It is important to extend probes of a possible  $0^-$  admixture to final states involving fermions, and measurements in  $\tau^+ \tau^-$  final states,  $\bar{t}t + H$  and single  $t + H$  production (see later) have been proposed [44]. But, for the time being, Peter Higgs can continue smiling!

### 2.2.3 Is it Elementary or Composite?

This question may be addressed by constructing a phenomenological Lagrangian  $\mathcal{L}$  with free parameters to describe the interactions of the ‘Higgs’ boson, and constraining the parameters using data on  $H$  production and decay. Motivated by the success of the Standard Model relation  $\rho \equiv m_W/m_Z \cos \theta_W = 1$ , it is usually assumed that this phenomenological Lagrangian possesses a custodial symmetry:  $SU(2) \times SU(2) \rightarrow SU(2)$ . In this case, one may parameterize the leading-order terms in  $\mathcal{L}$  as follows [60]:

$$\begin{aligned} \mathcal{L} = & \frac{v^2}{4} \text{Tr} D_\mu \Sigma^\dagger D^\mu \Sigma \left( 1 + 2a \frac{H}{v} + b \frac{H^2}{v^2} + \dots \right) \\ & - \bar{\psi}_L^i \Sigma \left( 1 + c \frac{H}{v} + \dots \right) \\ & + \frac{1}{2} (\partial_\mu H)^2 + \frac{1}{2} m_H^2 H^2 + d_3 \frac{1}{6} \left( \frac{3m_H^2}{v} \right) H^3 + d_4 \frac{1}{24} \left( \frac{3m_H^2}{v} \right) H^4 + \dots, \end{aligned} \quad (73)$$

where

$$\Sigma \equiv \exp \left( i \frac{\sigma^a \pi^a}{v} \right), \quad (74)$$

and the effective interaction with massless gauge bosons is written as

$$\mathcal{L}_\Delta = - \left( \frac{\alpha_s}{8\pi} c_g G_{a\mu\nu} G_a^{\mu\nu} + \frac{\alpha_{em}}{8\pi} c_\gamma F_{\mu\nu} F^{\mu\nu} \right) \left( \frac{H}{v} \right)^2. \quad (75)$$

The free coefficients  $a, b, c, d_3, d_4, c_g$  and  $c_\gamma$  are all normalized such that they are unity in the Standard Model: composite models may give observable deviations from these values. For example, in the composite model known as MCHM4 one has [51]

$$a = c = \sqrt{1 - \xi} : \xi \equiv \left( \frac{v}{f} \right)^2, \quad (76)$$

where  $f$  is an analogue of the pion decay constant in QCD. On the other hand, in the MCHM5 composite model, one has

$$a = \sqrt{1 - \xi}, \quad c = \frac{1 - 2\xi}{\sqrt{1 - \xi}}, \quad (77)$$

and in a pseudo-dilaton model one has

$$a = c = \frac{v}{V}, \quad (78)$$

where  $V$  is the dilaton v.e.v. that breaks scale invariance. One may also consider an ‘anti-dilaton’ scenario in which  $a = -c$ .

The signal strengths  $R$  in various channels relative to the Standard Model values are related in an obvious way to the parameters in (73). For example, for vector boson fusion and production in association with  $V = W, Z$  one has

$$R_{VBF} = R_{VH} = a^2, \quad (79)$$

and for production in association with  $\bar{t}t$  and the rates for decays into  $\bar{b}b$  and  $\tau^+\tau^-$  one has

$$R_{\bar{f}f} = c^2. \quad (80)$$

The corresponding ratio for the  $ggH$  coupling strength depends on the  $\bar{t}t$  coupling:

$$R_{gg} = c_g^2 = c^2 + \dots, \quad (81)$$

where the  $\dots$  represent possible contributions from particles beyond the Standard Model, and the ratio for the  $H\gamma\gamma$  coupling depends on both  $a$  and  $c$  as well as possible non-Standard Model contributions:

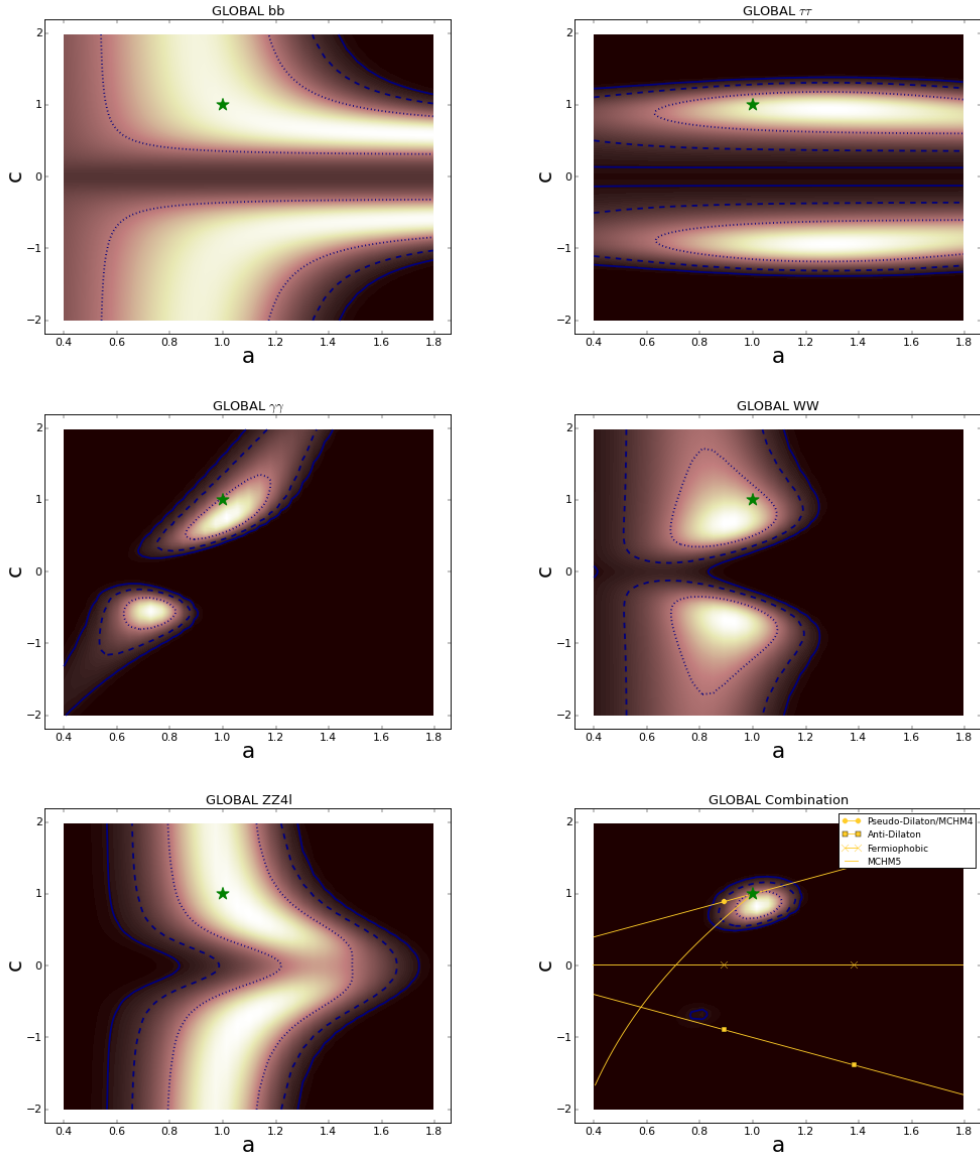
$$R_{\gamma\gamma} = c_\gamma^2 = \frac{\left(-\frac{8}{3}cF_t + aF_W\right)^2}{\left(-\frac{8}{3}F_t + F_W\right)^2} + \dots, \quad (82)$$

where  $F_{t,W}$  are form factors that depend on the ratios  $m_H/m_t$  and  $m_H/m_W$ , respectively. It is apparent from these expressions that only  $R_{\gamma\gamma}$  is sensitive to the relative sign of the  $H\bar{f}f$  and  $HVV$  couplings. The principal dependences of the signal strengths in various channels on the  $a$  and  $c$  parameters in (73) are summarized in Table 2.

channel	Production sensitive to		Decay sensitive to	
	$a$	$c$	$a$	$c$
$\gamma\gamma$	✓	✓	✓	✓
$\gamma\gamma$ VBF	✓	×	✓	✓
$WW$	✓	✓	✓	×
$WW + 2$ jets	✓	×	✓	×
$WW + 0,1$ jet	×	✓	✓	×
$\bar{b}b$	✓	×	×	✓
$ZZ$	✓	✓	✓	×
$\tau\tau$	✓	✓	×	✓
$\tau\tau$ VBF, VH	✓	×	×	✓

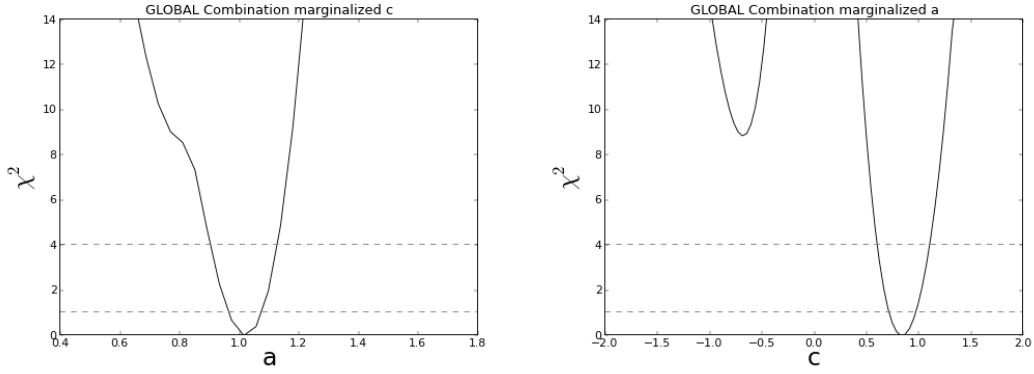
**Table 2:** The dominant dependences on the model parameters  $(a, c)$  (73) of the  $H$  signal strengths in various channels, from [61].

Fig. 18 shows how measurements of these various channels at the Tevatron collider and the LHC combine to constrain the parameters  $(a, c)$  [51]. We see in the top left panel that the data on  $\bar{b}b$  final states already disfavour leptophobic models in which the  $H$  particle has no couplings to fermions - here the Tevatron experiments play an important rôle. In the top right panel we see that data on the  $\tau^+\tau^-$  final state also disfavour leptophobic models. However, as expected on the basis of (80), these measurements by themselves offer no information about the sign of the fermion coupling coefficient  $c$ . The middle left panel shows the constraint imposed by the data on the  $\gamma\gamma$  final state. As seen in (82), this final state gives a constraint that is not symmetric between the signs of  $c$ , since there is interference between the virtual  $\bar{t}t$  and  $W^+W^-$  intermediate states that may be either constructive or destructive, depending on the sign of  $c$ . The middle right panel of Fig. 18 shows the constraint imposed by measurements of  $WW^*$  final states as well as  $H$  production via  $W^+W^-$  VBF and production in association with  $W^\pm$ , and the bottom left panel shows the corresponding constraint on the  $HZZ$  coupling. These measurements are highly consistent with custodial symmetry:  $SU(2) \times SU(2) \rightarrow SU(2)$ , as assumed in writing (73).



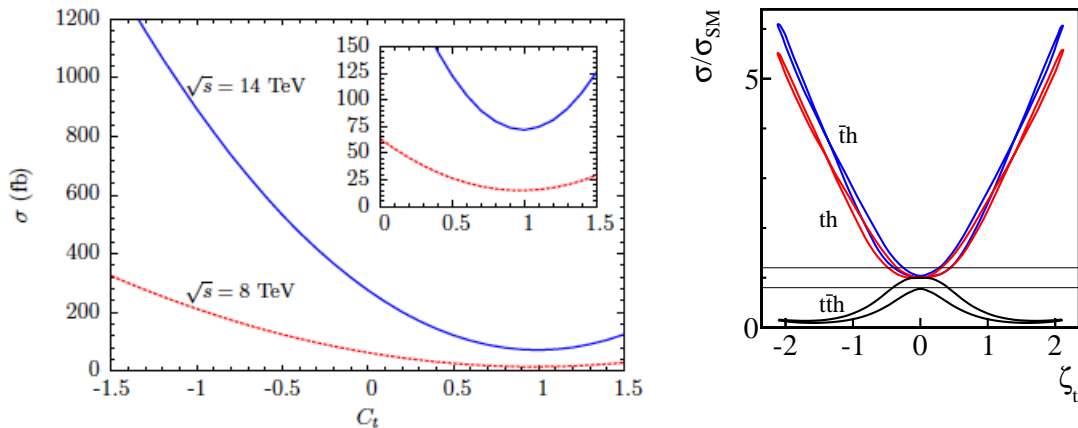
**Fig. 18:** The constraints in the  $(a, c)$  plane imposed by the measurements in Fig. 11 of the  $\bar{b}b$  final state (top left), of the  $\tau^+\tau^-$  final state (top right), of the  $\gamma\gamma$  final state (middle left), of the  $WW$  coupling (middle right) and in the  $ZZ$  coupling (bottom left). The combination of all these constraints is shown in the bottom right panel [51].

Finally, the bottom right panel of Fig. 18 displays the constraints in the  $(a, c)$  plane obtained in a global combination of these measurements [51]. We see that the positive sign of  $c$ , as expected in the Standard Model, is strongly favoured. This point is made explicitly in Fig. 19, where we see that the data favour  $a \sim 1$  and disfavour  $c < 0$  by  $\Delta\chi^2 \sim 9$ , i.e., 3 standard deviations. The continuous yellow lines the bottom right panel of Fig. 18 show the predictions of various composite alternatives to the Standard Model. As already mentioned, leptophilic models (represented by the horizontal line) are strongly disfavoured, as are ‘anti-dilaton’ models with  $a = -c$  (downwards-sloping line). The global analysis is compatible with the MCHM4 and dilaton models *iff* they are tuned to resemble the Standard Model, with  $\xi \sim 0, f \sim v$  in the MCHM4 (76) or  $V \sim v$  in the pseudo-dilaton (78) model (upwards-sloping line). Likewise, the MCHM5 model is compatible with the data only if  $\xi \sim 0$  in (77). Clearly, there is no evidence for any significant deviation from the Standard Model, and Peter Higgs may continue to smile!



**Fig. 19:** The one-dimensional likelihood functions for the boson coupling parameter  $a$  (left panel) and the fermion coupling parameter  $c$  (right panel), as obtained by marginalizing over the other parameter in the bottom right panel of Fig. 18 [51].

Before concluding this Section, it is interesting to discuss in more detail a Higgs production channel that could give direct information on the sign and magnitude of  $c$ , namely single  $t$  (or  $\bar{t}$ ) production in association with  $H$  [42, 43]. The two dominant amplitudes are due to Higgsstrahlung from an exchanged  $W$  boson and the final-state  $t$  quark. In the Standard Model with  $c > 0$ , these diagrams interfere destructively, as a precursor of the good high-energy behaviour expected in a spontaneously-broken gauge theory, whereas if  $c < 0$  the production cross section may be much larger. Even establishing an upper limit on single  $t$  (or  $\bar{t}$ ) production in association with  $H$  may be sufficient to determine the sign of  $c$ , independently of the  $\gamma\gamma$  measurement [43]. One may also consider the possibility of a CP-violating  $\bar{t}tH$  vertex  $\tilde{c}_t$  in addition to a conventional scalar vertex with coefficient  $c_t$  relative to the Standard Model value. The right panel of Fig. 20 shows the dependences of the  $\bar{t}tH$ ,  $tH$  and  $\bar{t}H$  cross sections on  $\zeta_t \equiv \arctan(\tilde{c}_t/c_t)$  [44] for choices of the  $c_t$  and  $\tilde{c}_t$  that are compatible with the constraints on the  $Hgg$  and  $H\gamma\gamma$  couplings shown in Fig. 23 [51]. We see that measurements of the  $\bar{t}tH$ ,  $tH$  and  $\bar{t}H$  cross sections could provide interesting information on the top- $H$  couplings.



**Fig. 20:** Left panel: The cross section for single  $t + H$  production as a function of the scalar top- $H$  coupling  $c_t$  normalized to its Standard Model value [43]. Right panel: The cross sections for  $\bar{t}t + H$ , single  $t$  and  $\bar{t} + H$  production relative to their Standard Model values for ranges of the scalar and pseudoscalar couplings ( $c_t, \tilde{c}_t$ ) compatible with current data on  $gg \rightarrow H$  production and  $H \rightarrow \gamma\gamma$  decay [51], as functions of  $\zeta_t \equiv \arctan(c_t/\tilde{c}_t)$  [44].

### 2.2.4 Are its Couplings Proportional to Particle Masses?

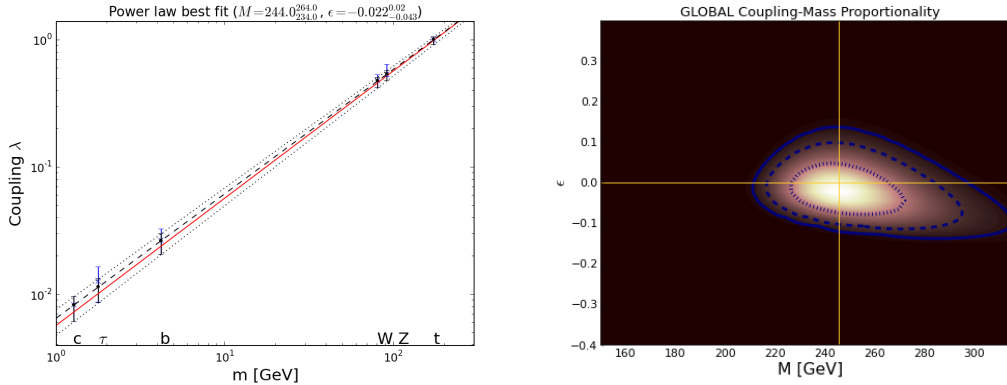
It is a key property of the Higgs boson of the Standard Model that its couplings to other particles should be proportional to their masses, and this is verified implicitly by the type of analysis reviewed in the previous Section. In order to verify it more explicitly, we may consider a parametrization of the  $H$  couplings to fermions  $\lambda_f$  and massive bosons  $g_V$  of the form [51, 61]

$$\lambda_f = \sqrt{2} \left( \frac{m_f}{M} \right)^{1+\epsilon}, \quad g_V = 2 \left( \frac{m_V^{2(1+\epsilon)}}{M^{1+2\epsilon}} \right). \quad (83)$$

In the Standard Model, one would expect the power  $\epsilon = 0$  and the scaling coefficient  $M = v = 246$  GeV. The results of a fit in terms of the two parameters  $(M, \epsilon)$  is shown in Fig. 21. It is represented in the left panel by the dashed line, with the one- $\sigma$  excursions shown as dotted lines. The solid red line is the prediction of the Standard Model, and the points with error bars are the predictions of the two-parameter fit. We see that these are completely compatible with the Standard Model predictions. In the right panel of Fig. 21 we see the 68 and 95% CL regions given by the fit in the  $(M, \epsilon)$  plane. Here the solid horizontal and vertical lines represent the Standard Model predictions  $\epsilon = 0$  and  $M = 246$  GeV. The data are quite close to the bull's eye! We display in the left panel of Fig. 22 the one-dimensional  $\chi^2$  function for  $\epsilon$ , marginalized over  $M$ , and in the right panel the one-dimensional  $\chi^2$  function for  $M$ , marginalized over  $\epsilon$ . The central values and the 68% CL ranges of  $M$  and  $\epsilon$  are:

$$M = 244_{-10}^{+20} \text{ GeV}, \quad \epsilon = -0.022_{-0.021}^{+0.042}. \quad (84)$$

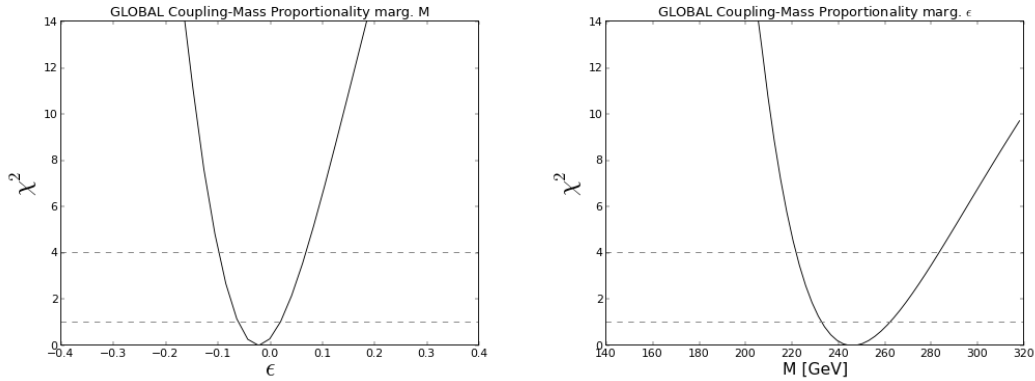
As we wrote in [61]: *"It walks and quacks like a Higgs boson."*



**Fig. 21:** The constraints on  $M$  and  $\epsilon$  (83) imposed by the measurements in Fig. 11. The left panel shows the strengths of the couplings to different fermion flavours and massive bosons predicted by this two-parameter  $(M, \epsilon)$  fit. The red line is the Standard Model prediction, the black dashed line is the best fit, and the dotted lines are the 68% CL ranges. For each particle species, the black error bar shows the range predicted by the global fit, and the blue error bar shows the range predicted for that coupling if its measurement is omitted from the global fit. The right panel displays the fit constraint in the  $(M, \epsilon)$  plane [51].

### 2.2.5 Are there Extra Contributions to its Loop Couplings?

The previous two Sections show that the tree-level  $H$  couplings are similar to those of a Standard Model Higgs boson. What can one say about its loop couplings to  $gg$  and  $\gamma\gamma$ ? Here we assume that its couplings to fermions and massive vector bosons are indeed Standard Model-like, so that  $a = c = 1$ , and investigate whether there is any evidence for other coloured (in the case of the  $gg$  coupling) or charged (in the case

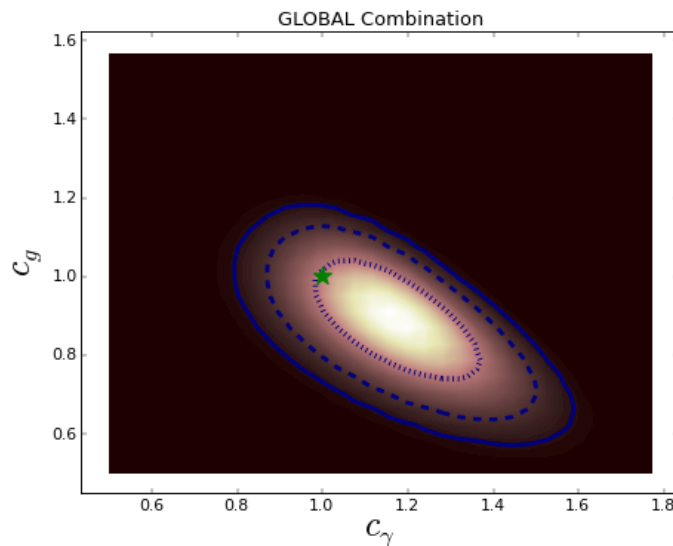


**Fig. 22:** The one-dimensional  $\chi^2$  functions for  $\epsilon$  (left panel) and  $M$  (right panel), as obtained by marginalizing over the other fit parameter [51].

of the  $\gamma\gamma$  coupling) particles contributing via triangular loop diagrams, so that  $c_g$  and/or  $c_\gamma \neq 1$  in (75). We see in Fig. 23 that the central value of  $c_\gamma > 1$  and the central value of  $c_g < 1$  [51]:

$$c_\gamma = 1.18 \pm 0.12, \quad c_g = 0.88 \pm 0.11. \quad (85)$$

However, the data are compatible with the Standard Model at the 68% CL, as seen by the location of the green star in Fig. 23. Thus, there is no good evidence for new particles circulating in loop diagrams. Fig. 24 displays the one-dimensional  $\chi^2$  functions for  $c_\gamma$  (left panel) and  $c_g$  (right panel), assuming, as above that  $a = c = 1$ , so that the the tree-level couplings to massive bosons and fermions have the Standard Model values.

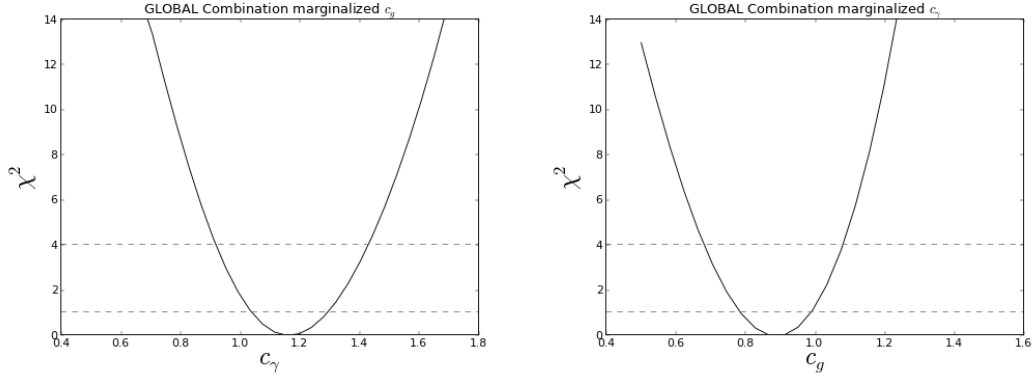


**Fig. 23:** The constraints in the  $(c_\gamma, c_g)$  plane imposed by the measurements in Fig. 11, assuming that  $a = c = 1$ , i.e., the Standard Model values for the tree-level couplings to massive bosons and fermions [51].

### 2.2.6 What is its Total Decay Rate?

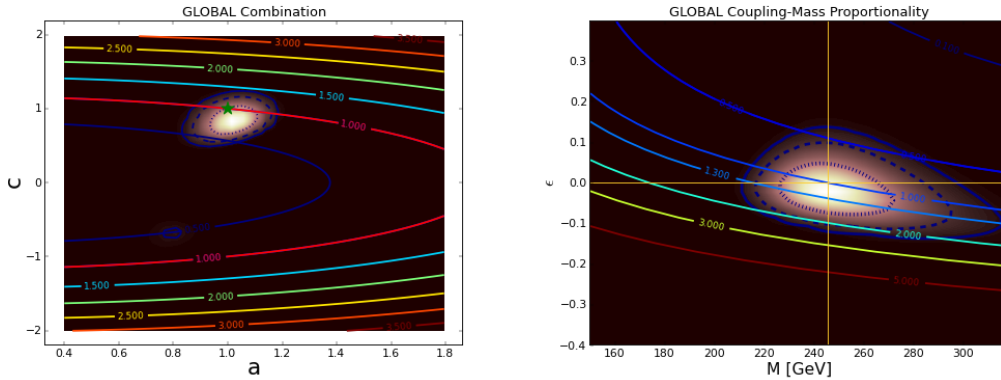
We now assume that the Higgs has no other decays beyond those in the Standard Model, and discuss the total Higgs decay rate in the two global fits discussed above, in terms of the parameters  $(a, c)$  and  $(M, \epsilon)$





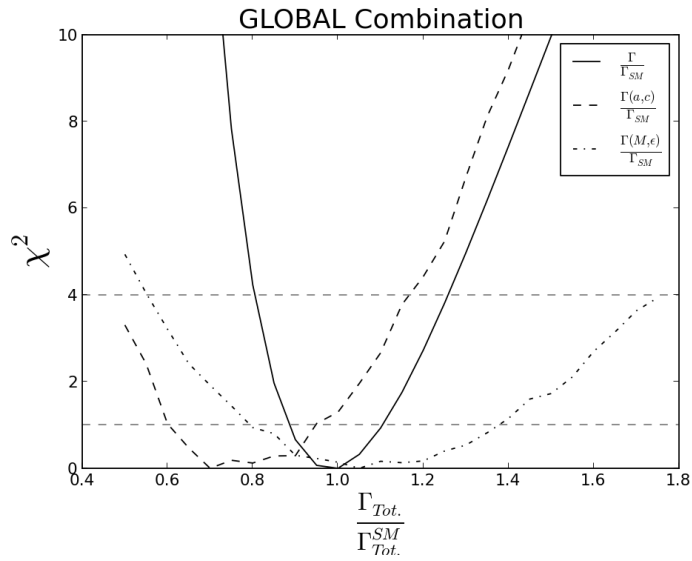
**Fig. 24:** The one-dimensional  $\chi^2$  functions for  $c_\gamma$  (left panel) and  $c_g$  (right panel), assuming that  $a = c = 1$ , so that the tree-level couplings to massive bosons and fermions have the Standard Model values [51].

and assuming no contributions from non-Standard-Model particles. The left panel of Fig. 25 displays contours of the Higgs decay width relative to the Standard Model prediction in the  $(a, c)$  plane shown in the bottom right panel of Fig. 18, and the right panel of Fig. 25 displays analogous contours in the  $(M, \epsilon)$  plane. We see that in each case the best fit has a total decay rate close to the Standard Model value. Fig. 26 displays the one-dimensional  $\chi^2$  function for the total Higgs decay width relative to its Standard Model value. The solid line is obtained assuming that  $a = c$  (or, equivalently, that  $\epsilon = 0$  but  $M$  is free), the dashed line is obtained by marginalizing over  $(a, c)$ , and the dot-dashed line is obtained by marginalizing over  $(M, \epsilon)$ . In all cases, we see that the total  $H$  decay width is compatible with the Standard Model prediction [51].



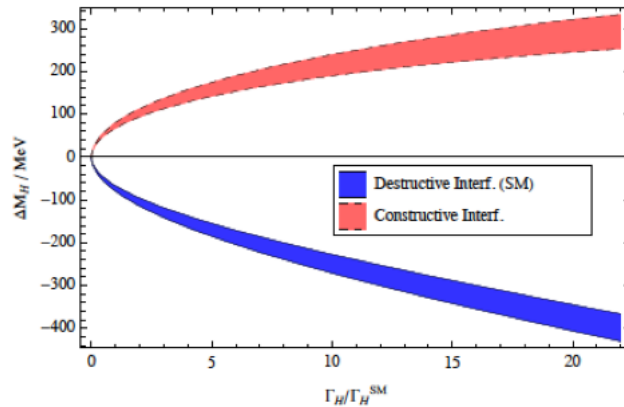
**Fig. 25:** Contours of the total Higgs decay rate relative to the Standard Model prediction in the  $(a, c)$  plane shown in the bottom right panel of Fig 18 (left) and the  $(M, \epsilon)$  plane shown in the right panel of Fig. 21 (right) [51].

In the absence of an assumption about  $H$  decays into non-Standard Model particles, it is difficult to obtain an accurate measurement of the total Higgs decay rate  $\Gamma_H$ . The CMS Collaboration has given a model-independent upper limit of 3 GeV, based on the width of the  $H \rightarrow \gamma\gamma$  signal peak that they observe, which is dominated by the experimental resolution. It has also been suggested [62] that one could establish an upper limit on  $\Gamma_H$  using measurements of  $ZZ$  final states mediated by off-shell  $H$  bosons. Using CMS data for  $m_{ZZ} \in (100, 800)$  GeV, it was estimated in [62] that  $\Gamma_H < 163$  MeV at the 95% CL, it was suggested that restricting to  $M_{ZZ} > 300$  GeV this bound could be improved to  $\Gamma_H < 88$  MeV, and it was suggested that the ultimate LHC sensitivity would be to  $\Gamma_H \sim 40$  MeV (see also [63]). Similar sensitivity may be obtained from an analysis of off-resonance  $W^+W^-$  final states [64].



**Fig. 26:** The one-dimensional  $\chi^2$  function for the total Higgs decay width relative to its value in the Standard Model,  $R \equiv \Gamma/\Gamma_{SM}$ , assuming decays into Standard Model particles alone and assuming  $a = c$  or equivalently  $\epsilon = 0$  (solid line), marginalizing over  $(a, c)$  (dashed line) and marginalizing over  $(M, \epsilon)$  (dot-dashed line) [51].

Another way to constrain or measure  $\Gamma_H$  may be via interference effects between the QCD and  $H$  contributions to the  $\gamma\gamma$  final state, which could shift the  $\gamma\gamma$  peak relative to its position in the  $ZZ^* \rightarrow 4\ell^\pm$  final state. (These are the only two observed  $H$  states where the invariant mass can be measured accurately.) This mass shift is sensitive to the sign and magnitude of the  $H\gamma\gamma$  coupling, by an amount that depends on the production kinematics. For  $\Gamma_H$  similar to the Standard Model value, the mass shift  $\sim 70$  MeV, as seen in Fig. 27 [65], so this is not a measurement for the faint-hearted! The published mass measurements have the problems that ATLAS and CMS find opposite signs for the  $\gamma\gamma$  and  $4\ell^\pm$  final states, though they are compatible within the experimental uncertainties, and their sensitivity is not very interesting.



**Fig. 27:** The shift  $\Delta M_H$  between measurements of the  $H$  mass in the  $\gamma\gamma$  and  $4\ell^\pm$  final states due to interference with QCD processes yielding  $\gamma\gamma$  final states, calculated in NLO QCD as a function of  $\Gamma_H$  relative to its value in the Standard Model [65].

### 2.2.7 *The Story so Far*

The discovery of the  $H$  particle has opened a new chapter in particle physics. Alternatives to the expectation that it is a scalar boson have been excluded with a high degree of confidence, and its couplings are consistent with those of a Standard Model Higgs boson. In particular, they exhibit the expected correlation with the masses of other particles. This is why we wrote in [51]: “*Beyond any reasonable doubt, it is a Higgs boson.*”<sup>6</sup>

Experiments are placing severe constraint on composite models, pushing upwards the possible scale of compositeness. On the other hand, an elementary scalar is a challenge for theorists, in particular because of the issue of quadratic divergences, that are symptomatic of extreme sensitivity to details of the ultraviolet completion. However, an elementary Higgs boson fits naturally within supersymmetry and, as we shall see in the next Section, simple supersymmetric models predict a Higgs mass below  $\sim 130$  GeV, as observed, and also predict that its couplings should be very similar to those of a Standard Model Higgs boson. As yet, there are no signs of supersymmetric particles, and we wait with interest to see what the LHC will reveal at 13/14 TeV, and in its high-luminosity incarnation.

In 1982 I was introduced to Mrs. Thatcher when she visited CERN, and she asked me what I did. I replied that “*My job as a theorist is to think of thing for the experiments to look for, but then we hope they find something different.*” Mrs. Thatcher liked things to be the way she wanted them to be, so she asked “*Wouldn’t it be better if they found what you predicted?*”. I responded along the lines that “*If they just find exactly what we predict, we would have no clues how to progress.*” In this spirit, let us all hope that the Higgs boson is not exactly that of the Standard Model, and that higher-energy LHC running will reveal other new physics beyond the Standard Model. In the next lecture I will discuss some of the prospects for these hopes.

## 3 What may the Future hold?

### 3.1 Theoretical Confusion

So far, though experiments at the LHC have discovered the Higgs boson, as yet they have found no direct hint of any new physics beyond the Standard Model such as supersymmetry or compositeness. The combination of these facts has caused a high mortality rate among theories, though not among theorists! As discussed previously, however, the fact that the measured values of  $m_H$  and  $m_t$  lie in a region where the electroweak vacuum would be unstable (OK, metastable) has led to suggestions that there should be new physics below  $10^{10}$  GeV to stabilize the vacuum. There have also been suggestions that the proximity of  $(m_H, m_t)$  to the stability boundary may be an indirect hint for some new physical principle<sup>7</sup>.

Supersymmetrists are among the most confused theorists. Motivated by the apparent fine-tuning of the electroweak scale and several phenomenological considerations such as dark matter, many of them had expected supersymmetry to appear during the first LHC run. In my view, this should be understood as the first run of the LHC at or close to its design energy of 14 TeV, so we should wait a while before jumping to conclusions. However, voices have been heard favouring very high-scale supersymmetry or split supersymmetry. The faint-hearted are asking whether we should modify or abandon the principle of naturalness. It is a reasonable question whether Nature needs to care about the naturalness of the electroweak scale, as long as she can find one set of parameters that includes the Standard Model. This is one possibility opened up by the string landscape, which comprises an exceedingly large number of possible vacua that all seem consistent with our current understanding. My own point of view is that supersymmetry anywhere would be better than nowhere, in terms of reducing the required amount of

<sup>6</sup>This phrase was quoted by the Royal Swedish Academy of Sciences in the Advanced Information it released about the award of the 2013 Nobel Physics Prize [66]. Ironically, this phrase had been removed from the published version of [51] at the request of the referee, who found the phrase “*unscientific*”.

<sup>7</sup>This is also sometimes linked to the fact that the cosmological constant (dark energy) lies close to the upper bound proposed by Weinberg [67].

fine-tuning. In any case, supersymmetry alone could not explain the hierarchy between the electroweak and gravitational scales: another mechanism would be needed to establish the hierarchy. New ideas are clearly needed!

In the absence of signatures of physics beyond the Standard Model at the LHC, there has been a tendency among some physicists to wonder whether the Standard Model is all there is, despite the persistence of a few loose ends such as the hierarchy, dark matter, the origin of matter, quantum gravity, etc. History gives many examples where such pessimism has turned out to be unwarranted: consider the examples of Albert Michelson (1894) “*The more important fundamental laws and facts of physical science have all been discovered*” or Lord Kelvin (1900) “*There is nothing new to be discovered in physics now. All that remains is more and more precise measurement*”, not to mention a Spanish Royal Commission, rejecting the proposal of Christopher Columbus to sail west (before 1492) “*So many centuries after the Creation, it is unlikely that anyone could find hitherto unknown lands of any value*”.

Perhaps we should rather follow the approach of Sherlock Holmes in the “Silver Blaze” story who, when asked by a policeman “*Is there any other point to which you would wish to draw my attention?*”, responded “*To the curious incident of the dog in the night-time.*” The policeman then remarked that “*The dog did nothing in the night-time*”, to which Holmes replied “*That was the curious incident.*” In our case, the “*curious incident*” is that no beyond the Standard Model dog has yet barked. Nevertheless, experiments have already provided theorists with many other clues: perhaps we need next to examine them more carefully, as well as planning ambitious future experiments. These are the themes of this Lecture.

## 3.2 Additional Topics in Higgs Studies

### 3.2.1 Higher-Dimensional Operators

A powerful way to probe indirectly possible physics beyond the Standard Model is to consider additional higher-dimensional operators that might be generated by new physics such as the exchanges of heavy particles, use data to constrain their coefficients, and thereby constrain the possibilities for physics beyond the Standard Model. In principle, this offers a way to constrain new physics in a coherent and effective way using a formalism that is consistent with all the established gauge and other symmetries. As an example, the CP-conserving dimension-6 operators in an effective Lagrangian involving just boson fields may be written in the form [68]:

$$\mathcal{L}_{CP+6} = \frac{\bar{c}_H}{2v^2} \partial^\mu [\Phi^\dagger \Phi] \partial_\mu [\Phi^\dagger \Phi] + \frac{\bar{c}_T}{2v^2} [\Phi^\dagger \overleftrightarrow{D}^\mu \Phi] [\Phi^\dagger \overleftrightarrow{D}_\mu \Phi] - \frac{\bar{c}_6 \lambda}{v^2} [H^\dagger H]^3 \quad (86)$$

$$+ \frac{ig \bar{c}_W}{m_W^2} [\Phi^\dagger T_{2k} \overleftrightarrow{D}^\mu \Phi] D^\nu W_{\mu\nu}^k + \frac{ig' \bar{c}_B}{2m_W^2} [\Phi^\dagger \overleftrightarrow{D}^\mu \Phi] \partial^\nu B_{\mu\nu} \quad (87)$$

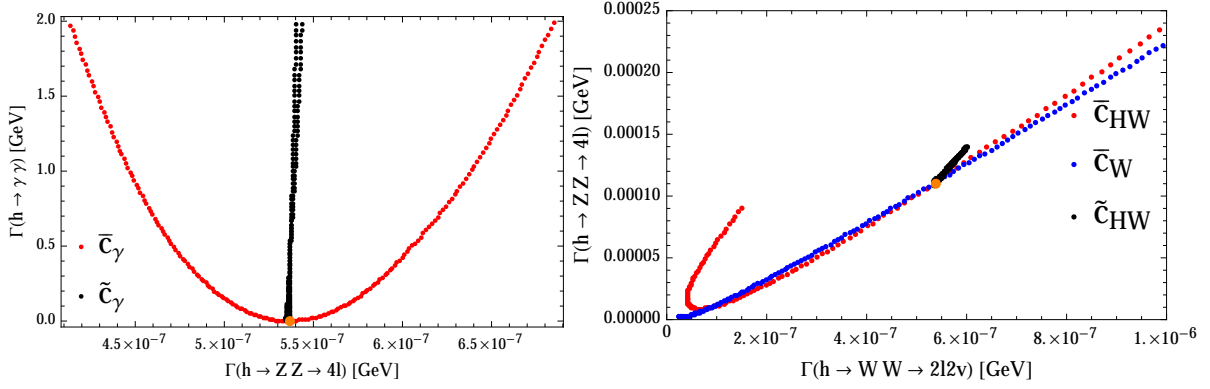
$$+ \frac{2ig \bar{c}_{HW}}{m_W^2} [D^\mu \Phi^\dagger T_{2k} D^\nu \Phi] W_{\mu\nu}^k + \frac{ig' \bar{c}_{HB}}{m_W^2} [D^\mu \Phi^\dagger D^\nu \Phi] B_{\mu\nu} \quad (88)$$

$$+ \frac{g'^2 \bar{c}_\gamma}{m_W^2} \Phi^\dagger \Phi B_{\mu\nu} B^{\mu\nu} + \frac{g_s^2 \bar{c}_g}{m_W^2} \Phi^\dagger \Phi G_{\mu\nu}^a G_a^{\mu\nu}. \quad (89)$$

The coefficients  $\bar{c}_H, \bar{c}_T, \bar{c}_W, \bar{c}_B, \bar{c}_{HW}, \bar{c}_{HB}, \bar{c}_\gamma$  and  $\bar{c}_g$  may then be constrained using precision electroweak data, measurements of Higgs production and decays, triple-gauge-boson couplings, etc. These constraints may then be compared with calculations in specific extensions of the Standard Model such as supersymmetry or composite models. One can also consider a CP-violating set of dimension-6 bosonic operators:

$$\mathcal{L}_{CP-6} = \frac{ig \tilde{c}_{HW}}{m_W^2} D^\mu \Phi^\dagger T_{2k} D^\nu \Phi \tilde{W}_{\mu\nu}^k + \frac{ig' \tilde{c}_{HB}}{m_W^2} D^\mu \Phi^\dagger D^\nu \Phi \tilde{B}_{\mu\nu} + \frac{g'^2 \tilde{c}_\gamma}{m_W^2} \Phi^\dagger \Phi B_{\mu\nu} \tilde{B}^{\mu\nu} \quad (90)$$

$$+ \frac{g_s^2 \tilde{c}_g}{m_W^2} \Phi^\dagger \Phi G_{\mu\nu}^a \tilde{G}_a^{\mu\nu} + \frac{g^3 \tilde{c}_{3W}}{m_W^2} \epsilon_{ijk} W_{\mu\nu}^i W_{\nu\rho}^j \tilde{W}^{\rho\mu k} + \frac{g_s^3 \tilde{c}_{3G}}{m_W^2} f_{abc} G_{\mu\nu}^a G_{\nu\rho}^b \tilde{G}^{\rho\mu c}, \quad (91)$$



**Fig. 28:** Effects of the dimension-6 operators  $\bar{c}_\gamma$ ,  $\tilde{c}_\gamma$ ,  $\bar{c}_{HW}$ ,  $\bar{c}_W$  and  $\tilde{c}_{HW}$  from (89, 90) on the  $H \rightarrow ZZ^*$  and  $\gamma\gamma$  partial widths (left panel) and the  $H \rightarrow WW^*$  and  $H \rightarrow ZZ^*$  partial widths (right panel). In each case, the Standard Model prediction is indicated by an orange dot [68].

where the dual field strength tensors are defined by

$$\tilde{B}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}B^{\rho\sigma}, \quad \tilde{W}_{\mu\nu}^k = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}W^{\rho\sigma k}, \quad \tilde{G}_{\mu\nu}^a = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}G^{\rho\sigma a}. \quad (92)$$

Two specific examples of possible effects on  $H$  decays due to higher-dimensional operators are shown in Fig. 28. In the left panel, we see the effects of the terms  $\propto \bar{c}_\gamma$  and  $\tilde{c}_\gamma$  on  $H \rightarrow ZZ^*$  and  $\gamma\gamma$  decays, and in the right panel we see the effects of the  $\propto \bar{c}_{HW}$ ,  $\bar{c}_W$  and  $\tilde{c}_{HW}$  on  $H \rightarrow ZZ^*$  and  $WW^*$  decays.

Some of the operators in (89) and (90) may also affect the production cross sections and kinematic distributions of the  $H$  boson. An example is provided by the double ratio of the cross sections for  $H$  production in association with a vector boson at 14 and 8 TeV:

$$\mathcal{R} \equiv \left( \frac{\sigma(\sqrt{S} = 14 \text{ TeV})}{\sigma(\sqrt{S} = 8 \text{ TeV})} \right)_{\tilde{c}_i} / \left( \frac{\sigma(\sqrt{S} = 14 \text{ TeV})}{\sigma(\sqrt{S} = 8 \text{ TeV})} \right)_{SM} \quad (93)$$

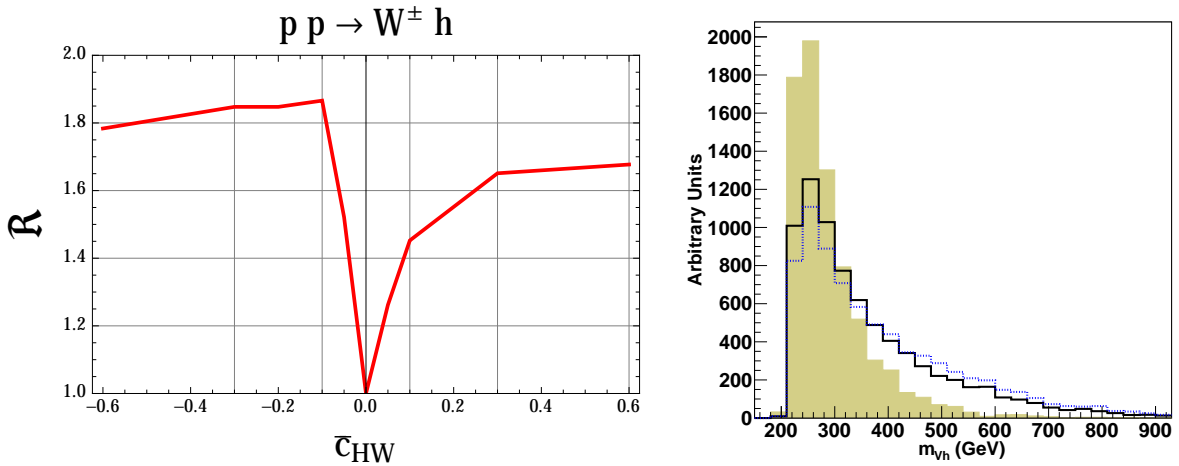
in the presence of an operator with coefficient  $\tilde{c}_i$ , as illustrated in the left panel of Fig. 29 for the case of  $\bar{c}_{HW}$ . The right panel of Fig. 29 illustrates the effects on the  $WH$  invariant mass distribution for the cases  $\bar{c}_{HW}$  (blue dotted histogram) and  $\bar{c}_W = 0.1$  (black histogram), the shaded histogram being the prediction of the Standard Model. We see that the double ratio (93) and the invariant mass distribution are interesting tools for constraining such operator coefficients, just as they provide discrimination between the  $0^+$ ,  $0^-$  and  $2^+$  hypotheses for the  $H$  spin.

### 3.2.2 A or The?

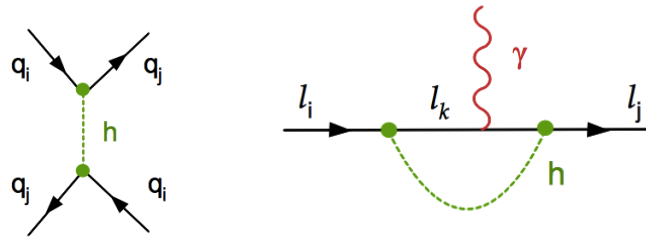
Now that the  $H$  particle has been established 'beyond any reasonable doubt' to be a Higgs boson, the questions arise whether it is *the* Higgs boson of the Standard Model, and whether there are any others. Possibilities proposed include models with an extra singlet field, models with a fermiophobic Higgs boson, and models with two Higgs doublets (2HDM) such as the minimal supersymmetric extension of the Standard Model (the MSSM). The ATLAS and CMS experiments have established upper limits on the couplings of possible massive  $H'$  boson, as we shall see later in connection with the MSSM.

Also on the agenda of the LHC experiments is to measure  $VV$  scattering and make a closure test, so as to check the Standard Model  $H$  cancellation discussed earlier. Does the Higgs boson discovered by ATLAS and CMS cure its high-energy behaviour so that the theory is indeed renormalizable?

Another strategy is to search for non-Standard Model decays, e.g., into invisible final states, or into pairs of light (pseudo)scalars  $aa$ , or into lepton-flavour-violating final states such as  $\mu\tau$  or  $e\tau$ , as discussed in the next Section.



**Fig. 29:** Left panel: the double ratio  $\mathcal{R}$  (93) of total cross sections at  $\sqrt{S} = 8$  TeV and 14 TeV for the associated production process  $pp \rightarrow W^\pm H \rightarrow \ell\nu b\bar{b}$  may provide a useful constraint on the dimension-6 operator coefficient  $\bar{c}_{HW}$ . Right panel: the invariant-mass distribution  $m_{VH}$  is displayed for the Standard Model (shaded histogram) and with additional couplings  $\bar{c}_{HW} = 0.1$  (blue-dotted histogram) and  $\bar{c}_W = 0.1$  (black histogram) [68].



**Fig. 30:** Left: Tree-level diagram contributing to a generic flavour-changing amplitude via  $H$  exchange. Right: One-loop  $H$  loop diagram contributing to anomalous magnetic moments and electric dipole moments of charged leptons ( $i = j$ ), or radiative LFV decay modes ( $i \neq j$ ) [48].

### 3.2.3 Flavour-Changing $H$ Decays

In the Standard Model one expects flavour-changing Higgs decays to occur only far below the sensitivity of present and prospective LHC measurements. However, this might not be true, e.g., in some composite Higgs models, so searches for flavour-violating  $H$  decays present an interesting opportunity to look for new physics beyond the Standard Model. Model-independent constraints on such decays are provided by measurements of flavour-violating processes at low energies, including the effects of four-fermion interactions generated by  $H$  exchange and  $H$  loop contributions to dipole moments, as illustrated in Fig. 30 [48].

Constraints from  $\Delta F = 2$  processes such as  $K - \bar{K}$ ,  $D - \bar{D}$ ,  $B - \bar{B}$  and  $B_s - \bar{B}_s$  mixing constrain flavour-changing  $H$  couplings so severely that quark-flavour-violating  $H$  decays are too suppressed to be detectable in the foreseeable future. However, the upper limits on lepton-flavour-changing  $H$  couplings are much weaker, and leave open the possibility that either  $H \rightarrow \tau\mu$  or  $H \rightarrow \tau e$  might have a branching ratio as large as  $\sim 10\%$ , comparable to the branching ratio for  $H \rightarrow \tau\tau$ . (The upper limit on  $\mu \rightarrow e\gamma$  forbids both branching ratios from being large simultaneously.) On the other hand, the constraints on the flavour-violating  $H\mu e$  coupling from anomalous  $\mu \rightarrow e$  conversion on nuclei and  $\mu \rightarrow e\gamma$  are much stronger, and require  $\text{BR}(H \rightarrow \mu e) \lesssim \mathcal{O}(10^{-9})$  [48].

### 3.2.4 Measuring the Triple-Higgs Coupling

If the Englert-Brout-Higgs field gives masses to elementary particles, and if this field is itself elementary, and if the Higgs boson is the particle associated with this field, what gives a mass to the Higgs boson? The answer within the Standard Model is the Englert-Brout-Higgs field itself, via the triple-Higgs coupling. Examining the effective Lagrangian (35), we see the following terms:

$$\mathcal{L}_{\text{Higgs}} \ni -\frac{\mu^2}{2}\eta^2 - \lambda v\eta^3 - \frac{1}{4}\eta^4, \quad (94)$$

where  $\eta$  denotes the quantum fluctuation in the Englert-Brout-Higgs field around its classical v.e.v., see (29).

The triple-Higgs coupling may be measured via  $H$  pair production [69], which should be within reach of the LHC with high luminosity<sup>8</sup>. The dominant mechanism for  $HH$  production is expected to be  $gg$  fusion:  $gg \rightarrow H^* \rightarrow HH$ , with an important background from  $t$  and  $b$  box diagrams for  $gg \rightarrow HH$ . Another strategy for measuring the triple-Higgs coupling is indirectly via its effects on the cross section for  $e^+e^- \rightarrow Z + H$  [70].

### 3.3 Supersymmetry

What else is there beyond the Higgs boson already discovered? Supersymmetry is my personal favourite candidate for physics beyond the Standard Model [5]. In my view, the discovery of a/the Higgs boson has strengthened the scientific case for supersymmetry. In addition to the traditional arguments that low-energy supersymmetry could resolve the fine-tuning (naturalness) aspect of the electroweak hierarchy problem, could provide the astrophysical dark matter, could facilitate grand unification and is essential (?) for string theory, we should remember that simple supersymmetric models stabilize the electroweak vacuum, predicted successfully the existence of a Higgs boson weighing  $< 130$  GeV [71], and also predict (successfully, so far) that Higgs couplings should be within a few % of their Standard Model values. *No wonder I wrote the word ‘supersymmetry’ in the largest possible font on one of my slides!*

Historically, the first motivation for supersymmetry at the TeV scale came from considerations of quantum (loop) corrections to the Higgs mass-squared,  $m_H^2$ , and thereby to the electroweak scale [5]. For example, a generic fermionic loop such as that in Fig. 31(a) yields a correction:

$$\Delta m_H^2 = -\frac{y_f^2}{8\pi^2}[2\Lambda^2 + 6m_f^2 \ln(\Lambda/m_f) + \dots], \quad (95)$$

where  $y_f$  is the Yukawa coupling:  $y_f H \bar{\psi} \psi$ , and  $\Lambda$  is an ultraviolet cutoff that represents the scale up to which the Standard Model remains valid, beyond which new physics appears. This contribution to the mass of the Higgs diverges quadratically with  $\Lambda$ . Hence if the Standard Model were to remain valid up to the Planck scale,  $M_P \simeq 10^{19}$  GeV, so that  $\Lambda = M_P$ , this correction would be  $\simeq 10^{34}$  times larger than the physical mass-squared of the Higgs, namely  $(10^2 \text{ GeV})^2$ . Moreover, the loop of a scalar field  $S$ , shown in Fig. 31(b), makes a similarly divergent contribution:

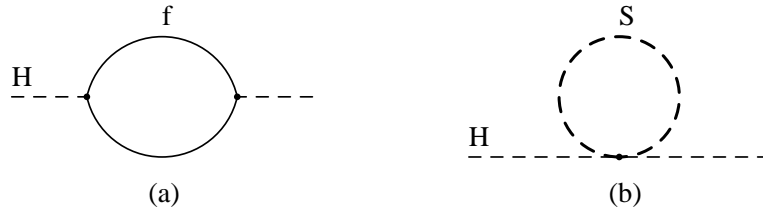
$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2}[\Lambda^2 - 2m_S^2 \ln(\Lambda/m_S) + \dots], \quad (96)$$

where  $\lambda_S$  is the quartic coupling of  $S$  to the Higgs boson.

Comparing (95) and (96), we see that the quadratically-divergent terms  $\propto \Lambda^2$  would cancel if, corresponding to every fermion  $f$  there is a scalar  $S$  with quartic coupling

$$\lambda_S = 2y_f^2. \quad (97)$$

<sup>8</sup>Measuring the quadruple-Higgs coupling would require measuring triple- $H$  production, which is likely to require a higher-energy collider such as the VHE-LHC described later.



**Fig. 31:** One-loop quantum corrections to the mass-squared of the Higgs boson due to (a) the loop of a generic fermion  $f$ , (b) a generic scalar  $S$ .

This is exactly the relationship imposed by supersymmetry! Therefore, there are no quadratic divergences in supersymmetric field theories, not just at the one-loop level discussed above, but also at the multi-loop level<sup>9</sup>. This means that if there is some dynamical mechanism that imposes a large hierarchy between different physical mass scales at the tree level, supersymmetry enables it to be maintained in a natural way.

A different motivation for supersymmetry is provided by the measured mass of the Higgs boson. As already remarked, the electroweak vacuum would not be stable in the absence of any new physics, since the (negative) renormalization by top quark loops would drive the quartic Higgs self-coupling negative at some scale  $\ll 10^{19}$  GeV, probably in the range  $10^{10}$  to  $10^{13}$  GeV. This could be averted if there were some new physics to counteract the negative renormalization by the top quark. In order to have the opposite sign to the top loop, this new physics should be bosonic, much like the stop squark [72]. But then one must consider all the quartic bosonic couplings permitted (enforced) by renormalizability, and ensure that none of them blow up or generate an instability, which requires fine-tuning to one part in  $10^3$  in the simplest case studied. However, this fine-tuning could be made more natural by postulating a new fermion, much like the Higgsino. Thus, one finishes up with a theory that looks very much like supersymmetry!

Within a supersymmetric theory, the renormalization due to the top quark could prove to be a blessing in disguise! After cancelling the quadratic divergences in (95, 96), one is left with residual logarithmic divergences that can be resummed using the renormalization-group equations (RGEs). Not knowing how supersymmetry is broken, one often assumes that this occurs far above the TeV scale, e.g., around the grand unification or Planck scale,  $M_{GUT}$  or  $M_P$ . In this case, the Higgs and other supersymmetry-breaking masses for scalars and gauginos are renormalized significantly by time the electroweak scale is reached. At leading order in the RGEs, which resum the leading one-loop logarithms, the renormalizations of the soft gaugino masses  $M_a$  coincide with the corresponding gauge couplings:

$$Q \frac{dM_a}{dQ} = \beta_a M_a, \quad (98)$$

where  $\beta_a$  is the one-loop renormalization coefficient including supersymmetric particles. Hence, to leading order

$$M_a(Q) = \frac{\alpha_a(Q)}{\alpha_{GUT}} m_{1/2} \quad (99)$$

if the gaugino masses are assumed to have a universal value  $m_{1/2}$  at the same large mass scale  $M_{GUT}$  as the gauge couplings  $\alpha_a$ . For this reason, one expects the gluino to be heavier than the wino:  $m_{\tilde{g}}/m_{\tilde{W}} \simeq \alpha_3/\alpha_2$  and the bino to be lighter again.

The gaugino masses contribute to the renormalizations of the soft supersymmetry-breaking scalar masses-squared  $m_i^2$  via the gauge couplings, and the scalar masses and the trilinear soft supersymmetry

<sup>9</sup>Moreover, many logarithmic corrections to couplings are also cancelled in a supersymmetric theory.



breaking parameters  $A_\lambda$  contribute via the Yukawa couplings:

$$\frac{Q dm_i^2}{dQ} = \frac{1}{16\pi^2} [-g_a^2 M_a^2 + \lambda^2(m_i^2 + A_\lambda^2)]. \quad (100)$$

For most of the scalar partners of Standard Model fermions, one has at leading order

$$m_i^2(Q) = m_i^2 + C_i m_{1/2}^2, \quad (101)$$

where the coefficients  $C_i$  depends on the gauge quantum numbers of the corresponding fermion. Since renormalization by the strong coupling is largest, one expects the squarks to be heavier than the sleptons. Specifically, if all the  $m_i$  and the  $M_a$  are each assumed to be universal at the GUT scale (a scenario known as the CMSSM), at the electroweak scale one finds:

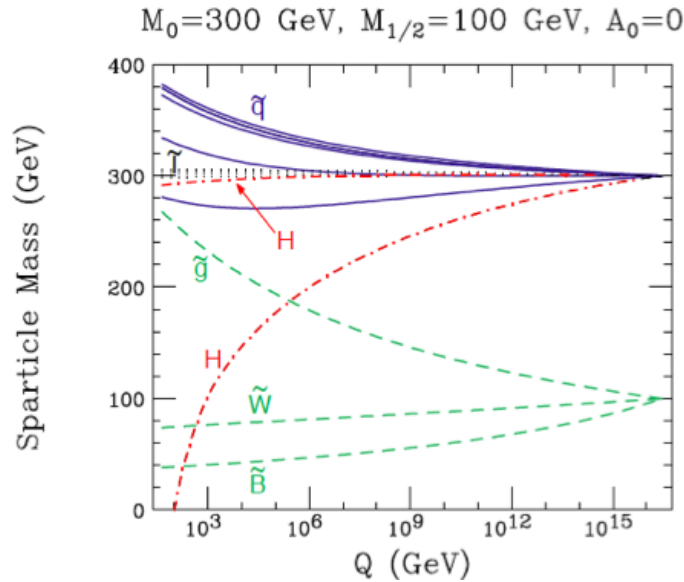
$$\text{Squarks : } m_{\tilde{q}}^2 \sim m_0^2 + 6m_{1/2}^2, \quad (102)$$

$$\text{Left - handed sleptons : } m_{\tilde{\ell}_L}^2 \sim m_0^2 + 0.5m_{1/2}^2, \quad (103)$$

$$\text{Right - handed sleptons : } m_{\tilde{\ell}_R}^2 \sim m_0^2 + 0.15m_{1/2}^2. \quad (104)$$

Typical results of calculations of these renormalization effects in the CMSSM are shown in Fig. 37.

Supersymmetry requires at least two Higgs doublets, one to give masses to charge-(+2/3) quarks,  $H_u$ , and the other to give masses to charge-(-1/3) quarks and charged leptons,  $H_d$ , and we denote the ratio of their v.e.v.s as  $\tan \beta$ . As we see in Fig. 32, renormalization by the top quark Yukawa coupling is important for one of the Higgs multiplets<sup>10</sup>, and may drive  $m_{H_u}^2$  negative at the electroweak scale. This may explain the negative sign of the quadratic term in the effective Standard Model potential, and would trigger electroweak symmetry breaking. If the top quark is heavy, it is possible for the electroweak scale to be generated naturally at a scale  $\sim 100$  GeV if  $m_t \sim 100$  GeV. For this reason, supersymmetry theorists actually suggested that the top quark should be heavy, before its discovery.



**Fig. 32:** Results of calculations of the renormalization of soft supersymmetry-breaking sparticle masses, assuming universal scalar and gaugino masses  $m_0, m_{1/2}$  at  $M_{GUT}$ . In general, strongly-interacting sparticles have larger physical masses at low scales, and the  $m_{H_u}^2$  is driven negative, triggering electroweak symmetry breaking.

<sup>10</sup>Renormalization by the other third-generation sfermions may also be important if  $\tan \beta$  is large.

The two complex Higgs doublets of the MSSM have eight degrees of freedom, of which three are used by the Higgs mechanism for electroweak breaking to give masses to the  $W^\pm$  bosons and to the  $Z^0$ , leaving five physical Higgs bosons in the physical spectrum. Of these, two ( $h, H$ ) are neutral Higgs bosons that are CP-even (scalar), one ( $A$ ) is neutral and CP-odd (pseudoscalar), and two are charged, the  $H^\pm$ . At tree level, the masses of the scalar supersymmetric Higgs bosons are:

$$m_{h,H}^2 = \frac{1}{2} \left( m_A^2 + m_Z^2 \mp \sqrt{(m_A^2 + m_Z^2)^2 - 4m_A^2 m_Z^2 \cos^2 2\beta} \right), \quad (105)$$

and the mass of the  $h$  is bounded from above by  $m_Z$ . This upper limit arises because the quartic Higgs coupling  $\lambda$  is fixed in the MSSM to be equal to the square of the electroweak gauge coupling (up to numerical factors), so that  $\lambda$  and hence  $m_{h^0}$  cannot be very large.

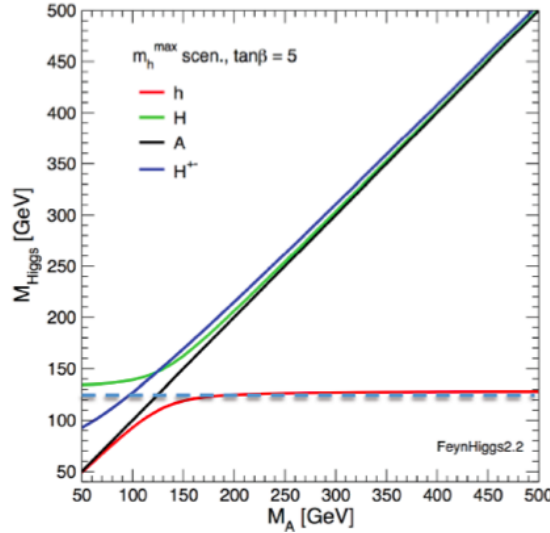
However, there are important radiative corrections to the above relations [71], the most important correction for  $m_h$  being the one-loop effect of the top quark and stop squark:

$$\Delta m_h^2 = \frac{3m_t^4}{4\pi^2 v^2} \ln \left( \frac{m_{\tilde{t}_1} m_{\tilde{t}_2}}{m_t^2} \right) + \dots, \quad (106)$$

where  $m_{\tilde{t}_{1,2}}$  are the physical masses of the stops. The correction  $\Delta m_h^2$  (106) depends quartically on the mass of the top, and after including this and higher-order corrections the mass of the lightest Higgs boson may be as large as [71, 73]:

$$m_h \lesssim 130 \text{ GeV}. \quad (107)$$

for stop masses of about a TeV, as seen in Fig. 33. The uncertainty in the calculation of  $m_h$  for given values of the supersymmetric model parameters is typically  $\sim 1.5$  GeV. As noted earlier, the range (107) is perfectly consistent with the mass measured by ATLAS and CMS, yet another attractive feature of supersymmetry.



**Fig. 33:** The masses of the supersymmetric Higgs bosons as functions of  $m_A$  for fixed values of the other MSSM parameters.

In general, the couplings of the supersymmetric Higgs bosons differ from those in the Standard Model.

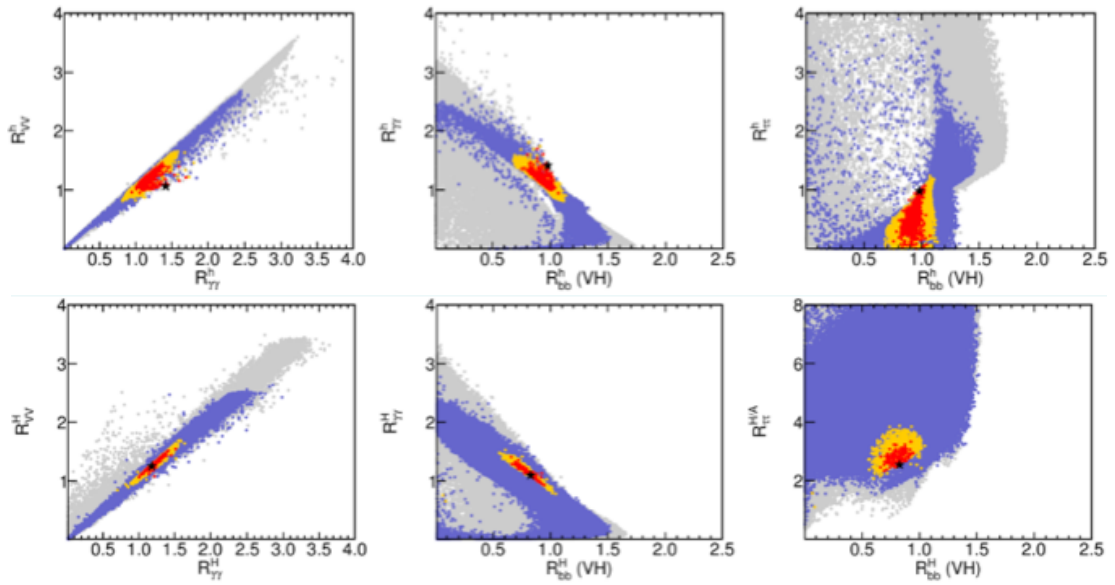
$$g_{hVV} = \sin(\beta - \alpha) g_{HVV}^{SM}, \quad (108)$$

$$g_{HVV} = \cos(\beta - \alpha) g_{HVV}^{SM}, \quad (109)$$

$$g_{hAZ} = \cos(\beta - \alpha) \frac{g'}{2 \cos \theta_W}, \quad (110)$$

$$g_{h\bar{b}b}, g_{h\tau^+\tau^-} = -\frac{\sin \alpha}{\cos \beta} g_{h\bar{b}b}^{SM}, g_{h\tau^+\tau^-}^{SM}. \quad (111)$$

If  $m_A \gg m_W$ , the masses of the other four Higgs bosons are very similar:  $m_H \sim m_A \sim m_{H^\pm}$ . However, there is a different and interesting possibility of  $m_A$  is small, namely that  $m_H \sim 125$  GeV, in which case the Higgs discovered at the LHC might actually be the second-lightest Higgs boson, and there might be a lighter one waiting to be discovered [74]<sup>11</sup>. Fig. 34 compares the predictions for various Higgs decays, relative to their Standard Model values, for fits in which the Higgs boson discovered is assumed to be the lightest one  $h$  (upper panel) and in which it is the heavier scalar  $H$  (lower panel). Overall, the quality of the conventional fit is better, but the unconventional fit may not yet be excluded. Experiments should continue the search for a lighter Higgs boson, remembering that it might have different couplings from those in the Standard Model.



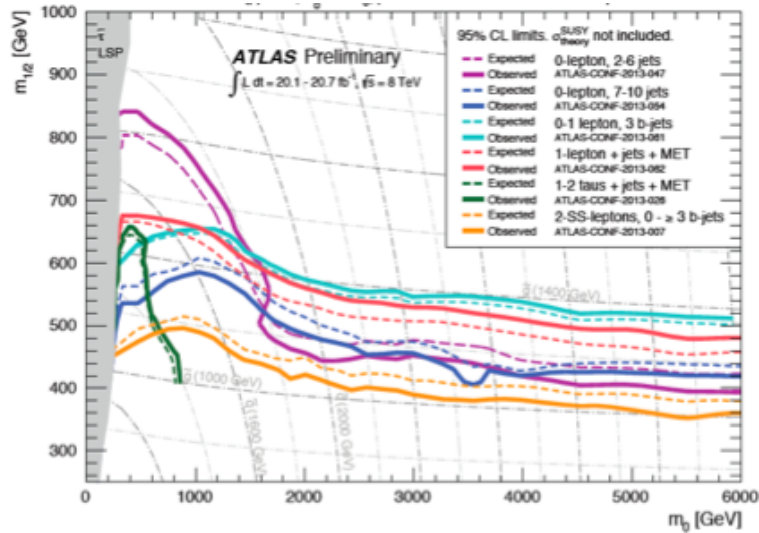
**Fig. 34:** Results of fits to Higgs data assuming (upper panel) that the Higgs boson discovered is the lightest supersymmetric Higgs  $h$  (upper panel) and in which it is the heavier scalar  $H$  (lower panel) [74].

### 3.4 Higgs and Supersymmetry

Let us now explore the implications for supersymmetry of the Higgs discovery, assuming that it is indeed the lighter scalar supersymmetric Higgs boson  $h$ . Important constraints on supersymmetric models are imposed by electroweak precision observables and flavour physics observables, the cosmological density [75] and astrophysical searches for cold dark matter [76], as well as LHC searches. In the following, we also take into account the experimental measurement of the anomalous magnetic moment of the muon,  $g_\mu - 2$ , which disagrees with theoretical calculations within the Standard Model by  $\sim 3$  standard deviations [77]. This discrepancy could be explained by supersymmetry at a relatively low mass scale, although this possibility is disfavoured in simple supersymmetric models by the LHC Higgs mass measurement and the absence (so far) of direct evidence for supersymmetric particles at the LHC. Fig. 35 displays the relevant constraints provided by various ATLAS searches for supersymmetry with

<sup>11</sup>There are more possibilities for a lighter Higgs boson in the next-to-minimal supersymmetric extension of the Standard Model (NMSSM).

the full Run 1 data set of  $\sim 20/\text{fb}$  of data at 8 TeV [78], using signatures with missing transverse energy (MET), jets, leptons and  $b$  quarks, interpreted within the CMSSM in which there are universal soft supersymmetry-breaking scalar masses  $m_0$ , gaugino masses  $m_{1/2}$  and trilinear parameters  $A_0 = -2m_0$  at the input GUT scale, assuming that  $\tan \beta = 30$ . We see that at small  $m_0$  the most important constraint is provided by searches for jets + MET, whereas searches for leptons,  $b$ -jets and MET are more important at large  $m_0$ .

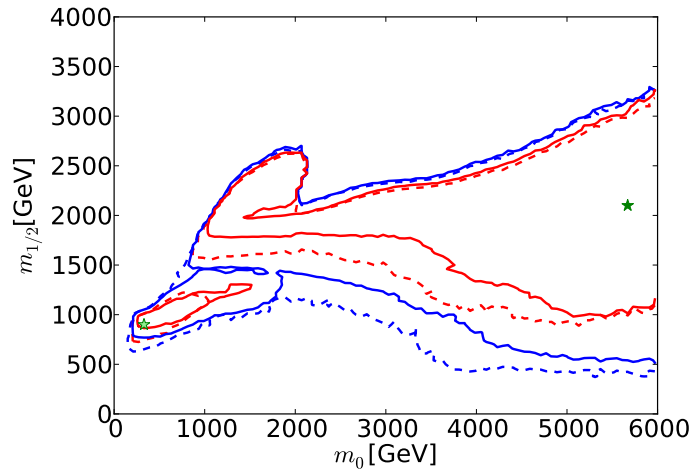


**Fig. 35:** Constraints on the universal soft supersymmetry-breaking scalar masses  $m_0$ , and gaugino masses  $m_{1/2}$  of the CMSSM from ATLAS searches for supersymmetry with the full Run 1 data set of  $\sim 20/\text{fb}$  of data at 8 TeV, using signatures with missing transverse energy (MET), jets, leptons and  $b$  quarks, assuming trilinear parameters  $A_0 = -2m_0$  at the input GUT scale and  $\tan \beta = 30$  [78].

In the following I present some results from a recent analysis [79] of these constraints made using the MasterCode framework [80], which incorporates a code for the electroweak observables based on [81], the flavour codes SuFla [82] and SuperIso 3.3 [83], SoftSUSY 3.3.9 [84] and FeynHiggs 2.10.0 [73] for spectrum calculations, and the MicrOMEGAS 3.2 [85] code for dark matter, which are interfaced using the SUSY Les Houches Accord [86]. We use the MasterCode framework to construct a global likelihood function ( $\chi^2$ ) that includes contributions from all the relevant observables.

Fig. 36 displays the regions of the  $(m_0, m_{1/2})$  plane allowed at the 95% CL (blue lines) and favoured at the 68% CL (red lines) after taking all these constraints into account [79]. The solid lines and filled star are obtained using the current 20/fb ATLAS constraints, and the dashed lines and open star are based on the previous constraints from 7/fb of LHC data at 7 TeV. The  $m_h$  constraint has the effect of favouring relatively large values of  $m_{1/2}$  beyond the reach of the direct LHC searches for supersymmetric particles, which have an impact only at low values of  $m_{1/2}$ . We note that the  $m_h$  constraint is relatively independent of  $m_0$ . Large values of  $m_{1/2}$  are excluded by the dark matter density constraint.

The one-dimensional  $\chi^2$  function for the gluino mass  $m_{\tilde{g}}$  resulting from this analysis of the CMSSM is shown in the upper left panel of Fig. 37 [79]. Again, the solid line is based on the current data set and the dotted line is based on the previous data set. We see that updating from the 7/fb 7-TeV data to the 20/fb 8-TeV data does not change the  $\chi^2$  function substantially. The current 95% CL lower limit on  $m_{\tilde{g}} \sim 1350$  GeV. A similar plot for the mass of a generic supersymmetric partner of a right-handed quark is shown in the upper right panel of Fig. 37. In this case, the 95% CL lower limit is  $m_{\tilde{q}_R} \sim 1650$  GeV. The lighter supersymmetric partner of the top quark may be significantly lighter, as shown in the lower left panel of Fig. 37, with a 95% CL lower limit  $m_{\tilde{t}_1} \sim 750$  GeV. Finally, the corresponding plot for the lighter supersymmetric partner of the  $\tau$  lepton is shown in the lower right



**Fig. 36:** The  $(m_0, m_{1/2})$  plane in the CMSSM after implementing the constraints from ATLAS MET searches, precision electroweak data, flavour physics,  $g_\mu - 2$ ,  $m_h$  and dark matter. The results of the current CMSSM fit are indicated by solid lines and a filled star, and a fit to previous data is indicated by dashed lines and open stars. The red lines denote  $\Delta\chi^2 = 2.30$  contours (corresponding approximately to the 68% CL), and the red lines denote  $\Delta\chi^2 = 5.99$  (95% CL) contours [79].

panel of Fig. 37. This is expected to be the next-to-lightest supersymmetric particle, after the dark matter particle  $\chi$ , and may have a mass as low as 330 GeV at the 95% CL.

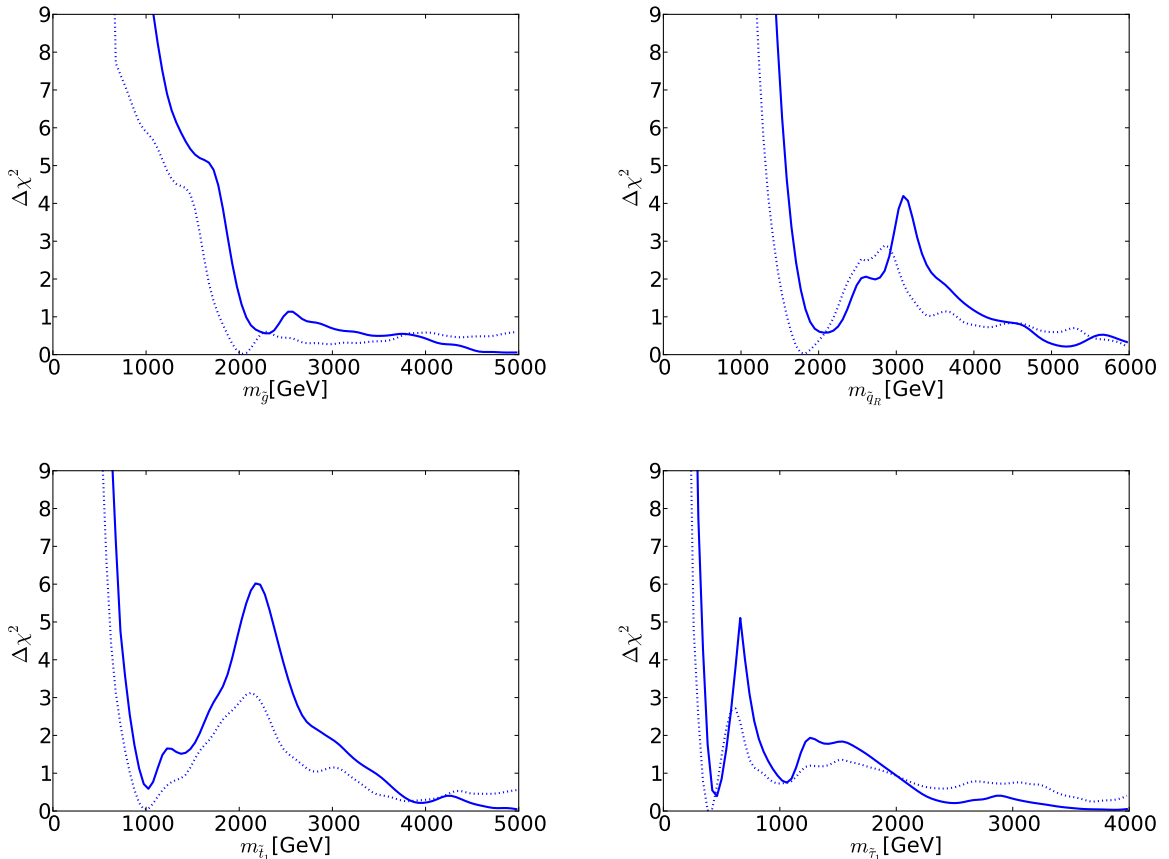
Ref. [87] provides estimates of the supersymmetry discovery reach of the LHC with 14 TeV, e.g., the  $(m_0, m_{1/2})$  plane displayed in Fig. 38. As seen there, the  $5\text{-}\sigma$  discovery reach for squarks and gluinos with 300/fb of luminosity should be to  $m_{\tilde{q}} \sim 3500$  GeV and  $m_{\tilde{g}_R} \sim 2000$  GeV in the CMSSM, and the discovery range with 3000/fb of luminosity would extend a few hundred GeV further. Thus, large parts of the CMSSM parameter space will be accessible in future runs of the LHC <sup>12</sup>. The priorities and prospects for future colliders will depend whether the LHC discovers supersymmetry during its runs at 14 TeV, but certainly more detailed studies of the Higgs boson will be on the agenda of the LHC and future accelerators, as discussed in the last Section of these lectures.

### 3.5 What Accelerator Next: a Higgs Factory?

One of the possible options for a future accelerator is a ‘Higgs Factory’ designed to study the Higgs boson in detail. In fact, we already have a Higgs Factory, namely the LHC, which has already produced millions of Higgs bosons, though only a small fraction of them have been observed. High-energy runs of the LHC will start in 2015, and luminosity upgrades are being planned. Accordingly, the capabilities of the LHC are being rethought from the Higgs Factory perspective. However, although the upgraded LHC will produce orders of magnitude more Higgs bosons, its capabilities are limited by theoretical uncertainties in the production cross section as well as by the backgrounds that render unobservable some interesting decay modes.

Both these shortcomings would be avoided at a lepton collider, and various options are being considered. The most mature concept is a linear  $e^+e^-$  collider, and two projects are being developed: the ILC that would operate initially at energies up to 500 GeV [88], and CLIC that could possibly operate at energies up to 3 TeV [89], where the Higgs production cross section would be larger.

<sup>12</sup>It should be emphasized that the likelihood estimates made here are specific to the models studied, as are the estimates of the physics reaches.



**Fig. 37:** The one-dimensional  $\chi^2$  likelihood functions in the CMSSM for  $m_{\tilde{g}}$  (upper left),  $m_{\tilde{q}_R}$  (upper right),  $m_{\tilde{t}_1}$  (lower left) and  $m_{\tilde{\tau}_1}$  (lower right). In each panel, the solid line is derived from a global analysis of the present data, and the dotted line is derived from an analysis of a previous data set, using current implementations of the constraints applied there [79].

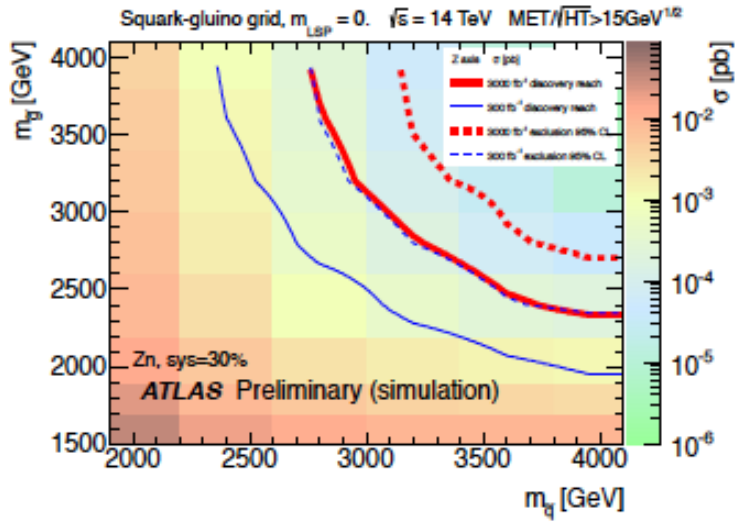
In the last couple of years, the alternative possibility of a circular  $e^+e^-$  collider has been revived [90]. Proposals include LEP3, a design for a high-luminosity  $\sqrt{s} = 240$  GeV  $e^+e^-$  collider that could be installed in the LHC tunnel, and TLEP, a  $\sqrt{s} = 350$  to 500 GeV collider that could be installed in a larger tunnel with a circumference of 80 to 100 km. Also under consideration is a  $\mu^+\mu^-$  collider, that would benefit from the (expected) larger coupling of the Higgs to the muon. Finally, there is the idea of a photon-photon collider, for example SAPPHiRE [91] that would exploit the recirculating linear accelerators proposed for the LHeC electron-proton collider.

Fig. 39 displays ATLAS estimates of the measurement uncertainties in Higgs signal strengths  $\mu$  (left) and ratios of partial decay widths (right) with integrated luminosities of 300/fb (green) and 3000/fb (blue) [87]<sup>13</sup>. We see good prospects for significant improvements with 300/fb relative to the current measurements, and for further improvements with 3000/fb that would enable several Higgs couplings to be measured with accuracies  $\lesssim 10\%$ .

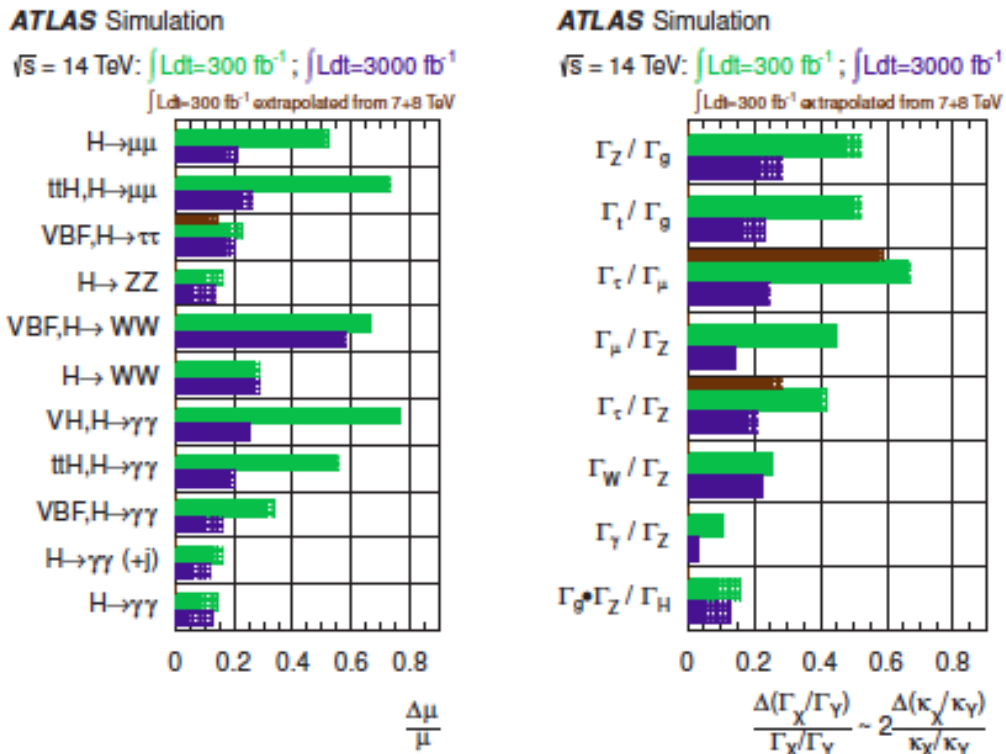
Fig. 40 displays estimates of the accuracies of measurements of the Higgs couplings to other particles that would be possible with the ILC [88], combining data from  $\sqrt{s} = 250, 500$  and 1000 GeV. This figure exhibits the prospective improvements in testing the linear dependence of the couplings on the other particle masses expected in the Standard Model.

<sup>13</sup>The possible improvements in  $\tau^+\tau^-$  measurements that could be provided by a more complete analysis are shown in brown.

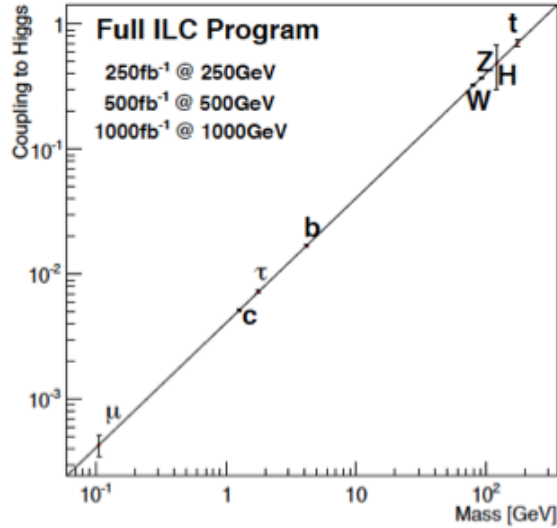




**Fig. 38:** The physics reach of the LHC in the  $(m_0, m_{1/2})$  plane provided by searches for squarks and gluinos assuming that the LSP mass is negligible. The different colours represent the production cross section at 14 TeV. The solid (dashed) lines display the  $5\text{-}\sigma$  discovery reach ( $95\%$  CL exclusion limit) with  $300/\text{fb}$  and  $3000/\text{fb}$  respectively [87].

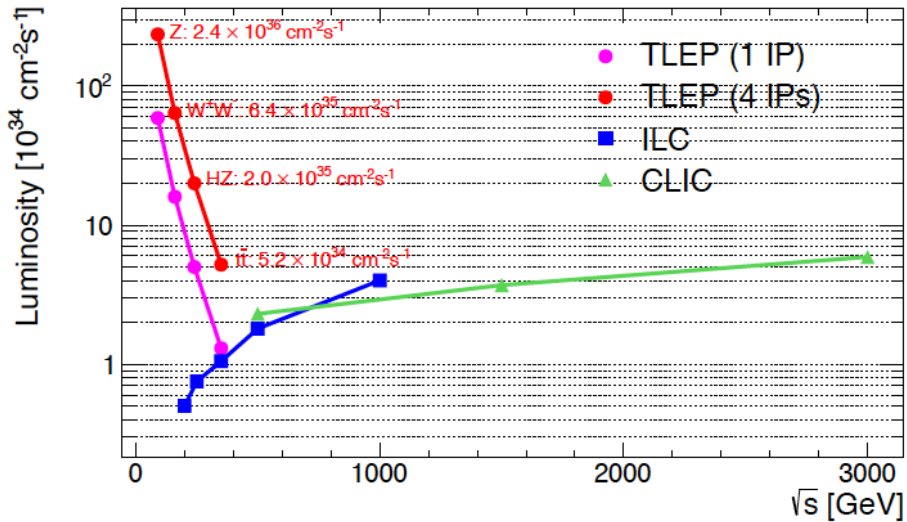


**Fig. 39:** Summary of ATLAS Higgs analysis sensitivities to signal strengths  $\mu$  (left) and ratios of partial decay widths (right) with integrated luminosities of  $300/\text{fb}$  (green) and  $3000/\text{fb}$  at  $\sqrt{s} = 14 \text{ TeV}$  for a Standard Model Higgs boson with a mass of  $125 \text{ GeV}$  [87].



**Fig. 40:** Summary of the possible ILC accuracies for measurements of the Higgs couplings to other particles that could be obtained by combining data at 250, 500 and 1000 GeV [88].

As compared to linear  $e^+e^-$  colliders, circular colliders possess the feature that the achievable luminosity increases at lower energies, assuming that a fixed amount of power can be supplied to the beams. This feature is illustrated in Fig. 41 for the cases of TLEP, CLIC and the ILC [92]. Another feature of a circular  $e^+e^-$  collider is that it can accommodate multiple interaction points (IPs), whereas a linear collider has only a single IP, possibly with multiple detectors operated alternately in push-pull mode.

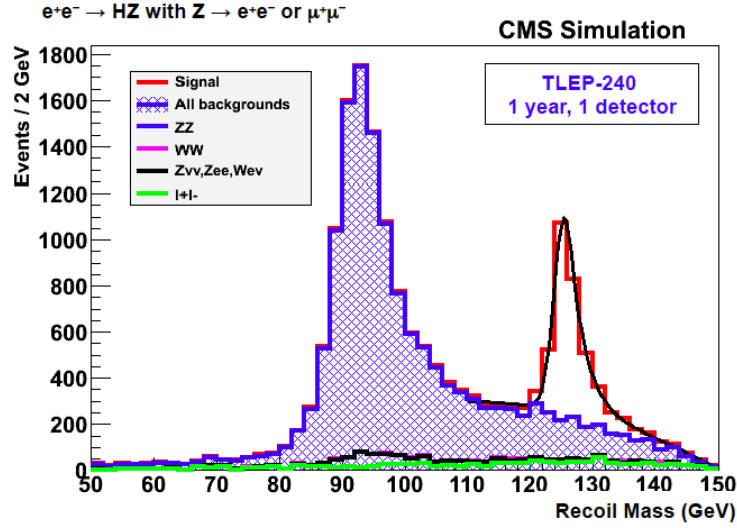


**Fig. 41:** Comparison of the luminosities estimated for TLEP with either one or four interaction points (IPs), CLIC and the ILC (which would have only a single IP) as functions of  $\sqrt{s}$  [92].

The experimental conditions at circular and linear  $e^+e^-$  colliders are similar, with the difference that the beam energies are spread by beamstrahlung in the linear case and by synchrotron radiation in the circular case, which yields a lower probability of large energy loss. Preliminary studies of possible Higgs measurements at TLEP have been made with simulations of the CMS detector that was designed for LHC physics [92]. Fig. 42 shows one example, that of the process  $e^+e^- \rightarrow H + Z$  followed by

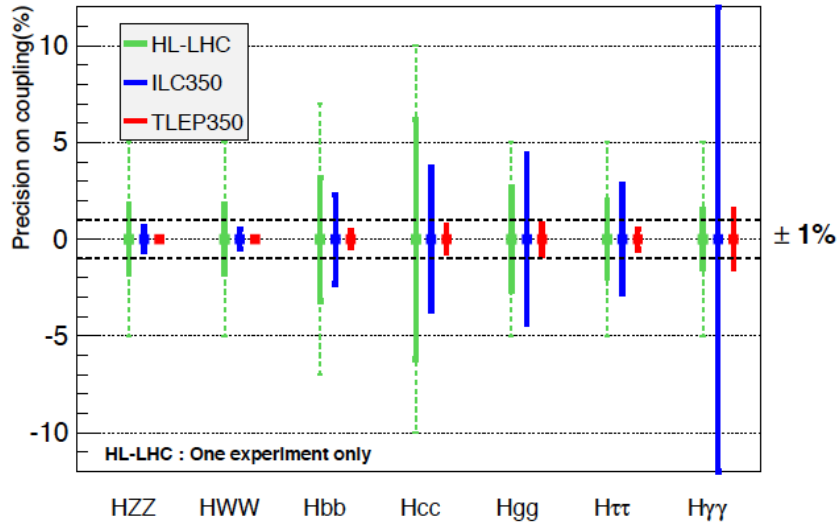


$H \rightarrow e^+e^-, \mu^+\mu^-$ : a detector specifically designed for  $e^+e^-$  collisions such as those developed for the ILC or CLIC would undoubtedly improve on these measurements.



**Fig. 42:** Example of a simulation of the process  $e^+e^- \rightarrow H + Z$  followed by  $H \rightarrow e^+e^-, \mu^+\mu^-$  at TLEP using a simulation of the CMS detector that was designed for LHC physics [92].

Fig. 43 shows a comparison of the estimated uncertainties in possible measurements of Higgs couplings with the high-luminosity upgrade of the LHC at  $\sqrt{s} = 14$  TeV (HL-LHC, green), the ILC (blue) and TLEP (red) operating at  $\sqrt{s} = 350$  GeV [92]. In the case of the HL-LHC, just one experiment is included, and the dashed lines neglect the possibility of improved theoretical calculations leading to reduced theoretical uncertainties. In each case, we see that TLEP could provide an accuracy considerably superior to that of the ILC, which is traceable directly to the higher statistics made available by the higher luminosity of TLEP visible in Fig. 41.



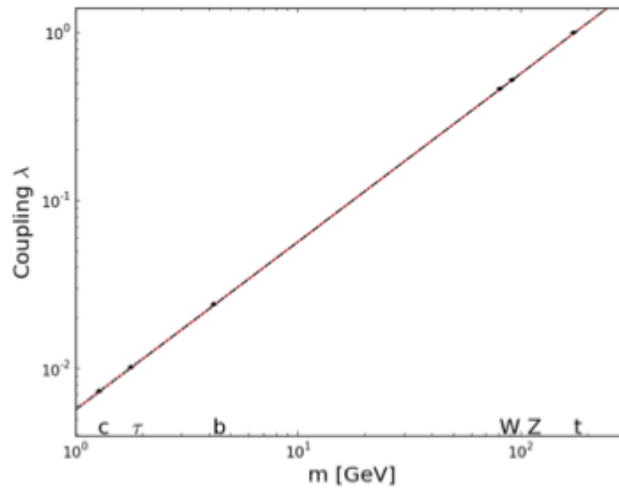
**Fig. 43:** Comparison of the estimated uncertainties of measurements of Higgs couplings with the high-luminosity upgrade of the LHC (HL-LHC, green), the ILC (blue) and TLEP (red) operating at  $\sqrt{s} = 350$  GeV [92].

Fig. 44 shows the result of a two-parameter  $(M, \epsilon)$  fit (83) to the TLEP coupling measurements

listed in Figs. 43, assuming the same central values as the Standard Model, which yields

$$M = 246.0 \pm 0.8 \text{ GeV}, \quad \epsilon = 0.0000_{-0.0010}^{+0.0015}, \quad (112)$$

offering the possibility of probing the Standard Model couplings at the few per-mille level. From this analysis we see that TLEP offers peerless capabilities for measuring Higgs properties. It is worth mentioning that TLEP also offers unique possibilities for other precision measurements of the Standard Model, e.g., at the  $Z^0$  peak and in  $W^+W^-$  production. However, the full exploitation of these accurate measurements will require a new generation of high-precision calculations within the Standard Model, posing a challenge to the theoretical community.



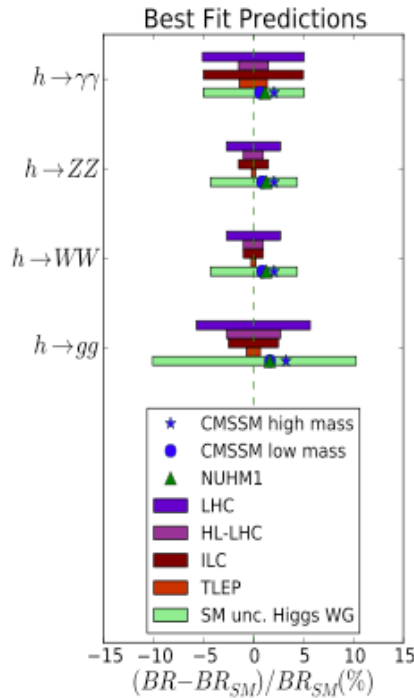
**Fig. 44:** The result of a two-parameter  $(M, \epsilon)$  fit (83) to the TLEP coupling measurements listed in Figs. 43, assuming the same central values as the Standard Model, to be compared with the left panel of Fig. 21.

This point is exemplified in Fig. 45. The horizontal bars represent the experimental accuracies with which various Higgs couplings can be measured at the LHC with 300/fb of luminosity, at the HL-LHC with 3000/fb of luminosity, at the ILC and at TLEP. Also shown are the deviations from the Standard Model predictions for various Higgs branching ratios calculated in typical supersymmetric fits within the models described in the previous Section. The good news is that TLEP would have sufficient precision to distinguish these models from the Standard Model. Unfortunately, there is some bad news as well. Fig. 45 that shows the current theoretical uncertainties in these branching ratios quoted by the LHC Higgs cross-section working group [39] dwarf the TLEP experimental uncertainties. More precise theoretical calculations will be sorely needed.

### 3.6 A Vision for the Future

It would be premature to decide on the top priority for a possible future large collider before we see at least some first results from the LHC at 13/14 TeV. The physics landscape will be completely different if supersymmetric particles or some other new physics is discovered at the TeV scale. In my opinion, in that case it would be a mistake to invest the world's particle physics resources in a collider incapable of studying the new TeV-scale physics.

Beyond any reasonable doubt, the ATLAS and CMS experiments at the LHC have discovered a Higgs boson, but it remains an open question whether there may be others. However, the existence of an (apparently) elementary scalar boson poses a big challenge for theoretical physics. Most of the responses to this challenge postulate some new physics at the TeV scale. There is some puzzlement that no such



**Fig. 45:** From top to bottom: uncertainties in the measurements of Higgs branching ratios that may be made at the LHC with 300/fb, the HL-LHC with 3000/fb, the ILC and TLEP, and finally the current theoretical uncertainties within the Standard Model. Also shown are the deviations from Standard Model predictions found in representative fits within supersymmetric models [92].

new physics turned up in the first LHC run at 7 and 8 TeV, but it is too soon for disappointment, still less despair. The LHC has broad possibilities for discovering new physics beyond the Standard Model when it restarts at 13/14 TeV. If it does discover new physics at the TeV scale, the top priority will be to study it, and beyond the LHC, a very-high-energy  $pp$  collider may offer the best prospects for long-term studies of this new physics. If the LHC does not discover more new physics, it would be natural to focus on studies of the Higgs boson that has already been discovered, in which case TLEP offers the best prospects, also for other high-precision physics.

The TLEP project is part of a vision for the future of particle physics that combines indirect exploration of possible new physics at the 10-TeV scale in  $e^+e^-$  collisions with direct exploration of this energy scale in very-high-energy  $pp$  collisions at  $\sqrt{s} \lesssim 100$  TeV in the same tunnel with a circumference of 80 to 100 km [93]. The communities interested in these complementary exploratory projects should work together to realize this vision, whose physics case will require a major effort to develop and convince those who control the global resources for scientific research.

**Acknowledgements**

I thank Tevong You for his collaboration on topics discussed in these lectures. I also thank fellow members of the MasterCode Collaboration, particularly Oliver Buchmueller, Sven Heinemeyer, Jad Marrouche, Keith Olive and Kees de Vries for many discussions. I also thank members of the TLEP Study Group, including Alain Blondel, Patrick Janot, Mike Koratzinos and Frank Zimmermann for many discussions. This work was supported in part by the London Centre for Terauniverse Studies (LCTS), using funding from the European Research Council via the Advanced Investigator Grant 267352.

## References

- [1] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716** (2012) 1 [arXiv:1207.7214 [hep-ex]].
- [2] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716** (2012) 30 [arXiv:1207.7235 [hep-ex]].
- [3] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).
- [4] A. Salam, in the *Proceedings of 8th Nobel Symposium*, Lerum, Sweden, 19-25 May 1968, pp 367-377.
- [5] M. Bustamante, L. Cieri and J. Ellis, CERN Yellow Report CERN-2010-001, 145-228 [arXiv:0911.4409 [hep-ph]].
- [6] E. Boos, Lectures at this School.
- [7] S. L. Glashow, Nucl. Phys. **22** (1961) 579.
- [8] LEP Electroweak Working Group, <http://lepewwg.web.cern.ch/LEPEWWG/>.
- [9] Gfitter Group <http://project-gfitter.web.cern.ch/project-gfitter/>.
- [10] J. Bardeen, L. N. Cooper and J. R. Schrieffer, Phys. Rev. **106** (1957) 162; V. L. Ginzburg and L. D. Landau, Zh. Eksp. Teor. Fiz. **20** (1950) 1064 and Phys. Rev. **108** (1957) 1175; L. N. Cooper, Phys. Rev. **104** (1956) 1189.
- [11] Y. Nambu, Phys. Rev. Lett. **4** (1960) 380.
- [12] J. Goldstone, Nuovo Cim. **19** (1961) 154;
- [13] J. Goldstone, A. Salam and S. Weinberg, Phys. Rev. **127** (1962) 965.
- [14] Y. Nambu, Phys. Rev. **117** (1960) 648.
- [15] P. W. Anderson, Phys. Rev. **130** (1963) 439;
- [16] W. Gilbert, Phys. Rev. Lett. **12** (1964) 713.
- [17] F. Englert and R. Brout, Phys. Rev. Lett. **13** (1964) 321.
- [18] P. W. Higgs, Phys. Lett. **12** (1964) 132.
- [19] P. W. Higgs, Phys. Rev. Lett. **13** (1964) 508.
- [20] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, Phys. Rev. Lett. **13** (1964) 585.
- [21] A. A. Migdal and A. M. Polyakov, Sov. Phys. JETP **24** (1967) 91 [Zh. Eksp. Teor. Fiz. **51** (1966) 135].
- [22] T. W. B. Kibble, Phys. Rev. **155** (1967) 1554.
- [23] P. W. Higgs, Phys. Rev. **145** (1966) 1156.
- [24] For some pre-LHC review of electroweak theory, see C. Quigg, Ann. Rev. Nucl. Part. Sci. **59** (2009) 505 [arXiv:0905.3187 [hep-ph]]; M. Bustamante, L. Cieri and J. Ellis, CERN Yellow Report CERN-2010-001, 145-228 [arXiv:0911.4409 [hep-ph]].
- [25] G. 't Hooft, Nucl. Phys. B **33** (1971) 173 and Nucl. Phys. B **35** (1971) 167; G. 't Hooft and M. J. G. Veltman, Nucl. Phys. B **44** (1972) 189.
- [26] J. M. Cornwall, D. N. Levin and G. Tiktopoulos, Phys. Rev. Lett. **30** (1973) 1268 [Erratum-ibid. **31** (1973) 572].
- [27] C. H. Llewellyn Smith, Phys. Lett. B **46** (1973) 233.
- [28] J. S. Bell, Nucl. Phys. B **60** (1973) 427.
- [29] J. R. Ellis, M. K. Gaillard and D. V. Nanopoulos, Nucl. Phys. B **106** (1976) 292.
- [30] J. Ellis and M. K. Gaillard, *Theoretical remarks*, in L. Camilleri *et al.*, *Physics with very high-energy  $e^+e^-$  colliding beams*, CERN report 76-18 (Nov. 1976), pp 21-94.
- [31] LEP Higgs Working Group, <http://lephiggs.web.cern.ch/LEPHIGGS/www/Welcome.html>.
- [32] R. Barate *et al.* [LEP Working Group for Higgs boson searches and ALEPH, DELPHI, L3 and OPAL Collaborations], Phys. Lett. B **565** (2003) 61 [hep-ex/0306033].
- [33] Tevatron New Phenomena and Higgs Working Group, <http://tevnpnphwg.fnal.gov/>.

- [34] M. J. G. Veltman, Nucl. Phys. B **123** (1977) 89.
- [35] M. J. G. Veltman, Acta Phys. Polon. B **8** (1977) 475.
- [36] J. R. Ellis and G. L. Fogli, Phys. Lett. B **249** (1990) 543, J. R. Ellis, G. L. Fogli and E. Lisi, Phys. Lett. B **274** (1992) 456 and Phys. Lett. B **318** (1993) 148.
- [37] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Ludwig, K. Moenig, M. Schott and J. Stelzer, Eur. Phys. J. C **72** (2012) 2003 [arXiv:1107.0975 [hep-ph]].
- [38] H. M. Georgi, S. L. Glashow, M. E. Machacek and D. V. Nanopoulos, Phys. Rev. Lett. **40** (1978) 692.
- [39] S. Dittmaier *et al.* [LHC Higgs Cross Section Working Group Collaboration], arXiv:1101.0593 [hep-ph]; arXiv:1201.3084 [hep-ph].
- [40] R. N. Cahn and S. Dawson, Phys. Lett. B **136** (1984) 196 [Erratum-ibid. B **138** (1984) 464].
- [41] S. L. Glashow, D. V. Nanopoulos and A. Yildiz, Phys. Rev. D **18** (1978) 1724.
- [42] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, Phys. Rev. D **64** (2001) 094023 [hep-ph/0106293].
- [43] S. Biswas, E. Gabrielli and B. Mele, JHEP **1301** (2013) 088 [arXiv:1211.0499 [hep-ph]]; S. Biswas, E. Gabrielli, F. Margaroli and B. Mele, JHEP **07** (2013) 073 [arXiv:1304.1822 [hep-ph]].
- [44] J. Ellis, D. S. Hwang, K. Sakurai and M. Takeuchi, KCL-PH-TH/2013-47, LCTS/2013-35, CERN-PH-TH/2013-312.
- [45] S. Heinemeyer *et al.* [LHC Higgs Cross Section Working Group Collaboration], arXiv:1307.1347 [hep-ph].
- [46] ATLAS Collaboration, <https://cds.cern.ch/record/1632191/files/ATLAS-CONF-2013-108>.  
CMS Collaboration, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/Hig13004TWikiUpd>.
- [47] ATLAS Collaboration, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HiggsPublicR>.  
CMS Collaboration, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsHI>.
- [48] G. Blankenburg, J. Ellis and G. Isidori, Phys. Lett. B **712** (2012) 386 [arXiv:1202.5704 [hep-ph]]; R. Harnik, J. Kopp and J. Zupan, JHEP **1303** (2013) 026 [arXiv:1209.1397 [hep-ph]].
- [49] G. Degrandi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori and A. Strumia, JHEP **1208** (2012) 098 [arXiv:1205.6497 [hep-ph]].
- [50] A. Juste, S. Mantry, A. Mitov, A. Penin, P. Skands, E. Varnes, M. Vos and S. Wimpenny, arXiv:1310.0799 [hep-ph].
- [51] J. Ellis and T. You, JHEP **1306** (2013) 103 [arXiv:1303.3879 [hep-ph]].
- [52] J. Ellis, D. S. Hwang, V. Sanz and T. You, JHEP **1211** (2012) 134 [arXiv:1208.6002 [hep-ph]].
- [53] D0 Collaboration, <http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/HIGGS/H138/>  
and <http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/HIGGS/H139/>.
- [54] J. Ellis, V. Sanz and T. You, Eur. Phys. J. C **73** (2013) 2507 [arXiv:1303.0208 [hep-ph]].
- [55] See, for example: J. Ellis and D. S. Hwang, JHEP **1209** (2012) 071 [arXiv:1202.6660 [hep-ph]].
- [56] J. Ellis, R. Fok, D. S. Hwang, V. Sanz and T. You, Eur. Phys. J. C **73** (2013) 2488 [arXiv:1210.5229 [hep-ph]].
- [57] S. Bolognesi, Y. Gao, A. V. Gritsan, K. Melnikov, M. Schulze, N. V. Tran and A. Whitbeck, Phys. Rev. D **86** (2012) 095031 [arXiv:1208.4018 [hep-ph]].
- [58] J. Ellis, V. Sanz and T. You, Phys. Lett. B **726** (2013) 244 [arXiv:1211.3068 [hep-ph]].
- [59] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **726** (2013) 120 [arXiv:1307.1432 [hep-ex]].
- [60] Here the analysis in [51] is used. For a partial list of previous references, see: D. Carmi, A. Falkowski, E. Kuflik and T. Volanski, arXiv:1202.3144 [hep-ph]; A. Azatov, R. Contino and J. Galloway, JHEP **1204** (2012) 127 [hep-ph/1202.3415]. J.R. Espinosa, C. Grojean, M. Muhlleitner and M. Trott, arXiv:1202.3697 [hep-ph]; P. P. Giardino, K. Kannike, M. Raidal and

- A. Strumia, arXiv:1203.4254 [hep-ph]; T. Li, X. Wan, Y. Wang and S. Zhu, arXiv:1203.5083 [hep-ph]; M. Rauch, arXiv:1203.6826 [hep-ph]; J. Ellis and T. You, JHEP **1206** (2012) 140, [arXiv:1204.0464 [hep-ph]]; A. Azatov, R. Contino, D. Del Re, J. Galloway, M. Grassi and S. Rahatlou, arXiv:1204.4817 [hep-ph]; M. Klute, R. Lafaye, T. Plehn, M. Rauch and D. Zerwas, arXiv:1205.2699 [hep-ph]; L. Wang and X.-F. Han, Phys. Rev. D **86** (2012) 095007, [arXiv:1206.1673 [hep-ph]]; D. Carmi, A. Falkowski, E. Kuflik and T. Volansky, arXiv:1206.4201 [hep-ph]; M. J. Dolan, C. Englert and M. Spannowsky, arXiv:1206.5001 [hep-ph]; J. Chang, K. Cheung, P. Tseng and T. Yuan, arXiv:1206.5853 [hep-ph]; S. Chang, C. A. Newby, N. Raj and C. Wanotayaroj, arXiv:1207.0493 [hep-ph]; I. Low, J. Lykken and G. Shaughnessy, arXiv:1207.1093 [hep-ph]; T. Corbett, O. J. P. Eboli, J. Gonzalez-Fraile and M. C. Gonzalez-Garcia, arXiv:1207.1344 [hep-ph]; P. P. Giardino, K. Kannike, M. Raidal and A. Strumia, arXiv:1207.1347 [hep-ph]; M. Montull and F. Riva, arXiv:1207.1716 [hep-ph]; J. R. Espinosa, C. Grojean, M. Muhlleitner and M. Trott, arXiv:1207.1717 [hep-ph]; D. Carmi, A. Falkowski, E. Kuflik, T. Volansky and J. Zupan, arXiv:1207.1718 [hep-ph]; S. Banerjee, S. Mukhopadhyay and B. Mukhopadhyaya, JHEP **10** (2012) 062, [arXiv:1207.3588 [hep-ph]]; F. Bonner, T. Ota, M. Rauch and W. Winter, arXiv:1207.4599 [hep-ph]; T. Plehn and M. Rauch, arXiv:1207.6108 [hep-ph]; A. Djouadi, arXiv:1208.3436 [hep-ph]; B. Batell, S. Gori and L. T. Wang, arXiv:1209.6832 [hep-ph]; G. Moreau, Phys. Rev. D **87** (2013) 015027, [arXiv:1210.3977 [hep-ph]]; G. Cacciapaglia, A. Deandrea, G. D. La Rochelle and J-B. Flament, arXiv:1210.8120 [hep-ph]; E. Masso and V. Sanz, arXiv:1211.1320 [hep-ph]; T. Corbett, O. J. P. Eboli, J. Gonzalez-Fraile and M. C. Gonzalez-Garcia, arXiv:1211.4580 [hep-ph]; R. Tito D’Agnolo, E. Kuflik and M. Zanetti, arXiv:1212.1165 [hep-ph]; A. Azatov and J. Galloway, arXiv:1212.1380 [hep-ph]; G. Bhattacharyya, D. Das and P.B. Pal, Phys. Rev. D **87** (2013) 011702, [arXiv:1212.4651 [hep-ph]]; D. Choudhury, R. Islam, A. Kundu and B. Mukhopadhyaya, arXiv:1212.4652 [hep-ph]; R. S. Gupta, M. Montull and F. Riva, arXiv:1212.5240 [hep-ph]; G. Belanger, B. Dumont, U. Ellwanger, J. F. Gunion and S. Kraml, arXiv:1212.5244 [hep-ph]; K. Cheung, J. S. Lee and P.-Y. Tseng, arXiv:1302.3794 [hep-ph].
- [61] J. Ellis and T. You, JHEP **1209** (2012) 123 [arXiv:1207.1693 [hep-ph]].
- [62] F. Caola and K. Melnikov, Phys. Rev. D **88** (2013) 054024 [arXiv:1307.4935 [hep-ph]].
- [63] J. M. Campbell, R. K. Ellis and C. Williams, arXiv:1311.3589 [hep-ph].
- [64] J. M. Campbell, R. K. Ellis and C. Williams, arXiv:1312.1628 [hep-ph].
- [65] L. J. Dixon and Y. Li, Phys. Rev. Lett. **111** (2013) 111802 [arXiv:1305.3854 [hep-ph]].
- [66] Class for Physics of the Royal Swedish Academy of Sciences, [http://www.nobelprize.org/nobel\\_prizes/physics/laureates/2013/advanced-physicsprize2013.pdf](http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/advanced-physicsprize2013.pdf).
- [67] S. Weinberg, Rev. Mod. Phys. **61** (1989) 1.
- [68] For an introduction to the extensive literature, see A. Alloul, B. Fuks and V. Sanz, arXiv:1310.5150 [hep-ph]; C. Y. Chen, S. Dawson and C. Zhang, arXiv:1311.3107 [hep-ph].
- [69] See, for example M. J. Dolan, C. Englert and M. Spannowsky, JHEP **1210** (2012) 112 [arXiv:1206.5001 [hep-ph]].
- [70] M. McCullough, arXiv:1312.3322 [hep-ph].
- [71] Y. Okada, M. Yamaguchi and T. Yanagida, Prog. Theor. Phys. **85** (1991) 1; J. R. Ellis, G. Ridolfi and F. Zwirner, Phys. Lett. B **257** (1991) 83; H. E. Haber and R. Hempfling, Phys. Rev. Lett. **66** (1991) 1815.
- [72] J. R. Ellis and D. Ross, Phys. Lett. B **506** (2001) 331 [hep-ph/0012067].
- [73] For the most recent calculations of supersymmetric Higgs masses, including leading and next-to-leading contributions beyond two loops, see: T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak and G. Weiglein, arXiv:1312.4937 [hep-ph].
- [74] P. Bechtle, S. Heinemeyer, O. Stal, T. Stefaniak, G. Weiglein and L. Zeune, Eur. Phys. J. C **73**

- (2013) 2354 [arXiv:1211.1955 [hep-ph]].
- [75] J. R. Ellis, J. S. Hagelin, D. V. Nanopoulos, K. A. Olive and M. Srednicki, Nucl. Phys. B **238** (1984) 453; H. Goldberg, Phys. Rev. Lett. **50** (1983) 1419 [Erratum-ibid. **103** (2009) 099905].
- [76] D. S. Akerib *et al.* [LUX Collaboration], arXiv:1310.8214 [astro-ph.CO].
- [77] G. Bennett *et al.* [The Muon g-2 Collaboration], Phys. Rev. D **73** (2006) 072003 [arXiv:hep-ex/0602035]; M. Benayoun, P. David, L. DelBuono and F. Jegerlehner, Eur. Phys. J. C **73** (2013) 2453 [arXiv:1210.7184 [hep-ph]]; T. Blum, A. Denig, I. Logashenko, E. de Rafael, B. L. Roberts, T. Teubner and G. Venanzoni, arXiv:1311.2198 [hep-ph].
- [78] ATLAS Collaboration,  
<http://cds.cern.ch/record/1547563/files/ATLAS-CONF-2013-047.pdf>.
- [79] O. Buchmueller, R. Cavanaugh, A. De Roeck, M. J. Dolan, J. R. Ellis, H. Flacher, S. Heinemeyer and G. Isidori *et al.*, arXiv:1312.5250 [hep-ph], see also O. Buchmueller, M. J. Dolan, J. Ellis, T. Hahn, S. Heinemeyer, W. Hollik, J. Marrouche and K. A. Olive *et al.*, arXiv:1312.5233 [hep-ph].
- [80] MasterCode Collaboration, <http://mastercode.web.cern.ch/mastercode/>.
- [81] S. Heinemeyer *et al.*, JHEP **0608** (2006) 052 [arXiv:hep-ph/0604147]; S. Heinemeyer, W. Hollik, A. M. Weber and G. Weiglein, JHEP **0804** (2008) 039 [arXiv:0710.2972 [hep-ph]].
- [82] G. Isidori and P. Paradisi, Phys. Lett. B **639** (2006) 499 [arXiv:hep-ph/0605012]; G. Isidori, F. Mezzadri, P. Paradisi and D. Temes, Phys. Rev. D **75** (2007) 115019 [arXiv:hep-ph/0703035], and references therein.
- [83] F. Mahmoudi, Comput. Phys. Commun. **178** (2008) 745 [arXiv:0710.2067 [hep-ph]]; Comput. Phys. Commun. **180** (2009) 1579 [arXiv:0808.3144 [hep-ph]]; D. Eriksson, F. Mahmoudi and O. Stal, JHEP **0811** (2008) 035 [arXiv:0808.3551 [hep-ph]].
- [84] B. C. Allanach, Comput. Phys. Commun. **143** (2002) 305 [arXiv:hep-ph/0104145].
- [85] G. Belanger, F. Boudjema, A. Pukhov and A. Semenov, Comput. Phys. Commun. **176** (2007) 367 [arXiv:hep-ph/0607059]; Comput. Phys. Commun. **149** (2002) 103 [arXiv:hep-ph/0112278]; Comput. Phys. Commun. **174** (2006) 577 [arXiv:hep-ph/0405253].
- [86] P. Skands *et al.*, JHEP **0407** (2004) 036 [arXiv:hep-ph/0311123]; B. Allanach *et al.*, Comput. Phys. Commun. **180** (2009) 8 [arXiv:0801.0045 [hep-ph]].
- [87] ATLAS Collaboration, arXiv:1307.7292 [hep-ex].
- [88] ILC TDR, H. Baer, T. Barklow, K. Fujii, Y. Gao, A. Hoang, S. Kanemura, J. List and H. E. Logan *et al.*, arXiv:1306.6352 [hep-ph].
- [89] CLIC CDR, eds. M. Aicheler, P. Burrows, M. Draper, T. Garvey, P. Lebrun, K. Peach, N. Phinney, H. Schmickler, D. Schulte and N. Toge, CERN-2012-007,  
<http://project-clic-cdr.web.cern.ch/project-CLIC-CDR/>.
- [90] A. Blondel and F. Zimmermann, arXiv:1112.2518; A. Blondel, M. Koratzinos, R. W. Assmann, A. Butterworth, P. Janot, J. M. Jimenez, C. Grojean and A. Milanese *et al.*, arXiv:1208.0504 [physics.acc-ph].
- [91] S. A. Bogacz, J. Ellis, L. Lusito, D. Schulte, T. Takahashi, M. Velasco, M. Zanetti and F. Zimmermann, arXiv:1208.2827 [physics.acc-ph].
- [92] M. Bicer, H. Duran Yildiz, I. Yildiz, G. Coignet, M. Delmastro, T. Alexopoulos, C. Grojean and S. Antusch *et al.*, arXiv:1308.6176 [hep-ex].
- [93] Future Circular Collider Study, <http://indico.cern.ch/conferenceDisplay.py?confId=28234>





## Beyond the Standard Model

C. Csáki<sup>a</sup> and P. Tanedo<sup>b</sup>

<sup>a</sup> Department of Physics, LEPP, Cornell University, Ithaca, NY 14853

<sup>b</sup> Department of Physics & Astronomy, University of California, Irvine, CA 92697

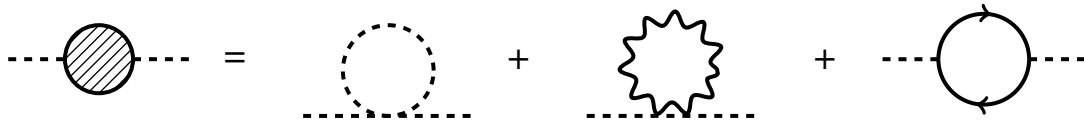
### Abstract

We introduce aspects of physics beyond the Standard Model focusing on supersymmetry, extra dimensions, and a composite Higgs as solutions to the Hierarchy problem. Lectures given at the 2013 European School of High Energy Physics, Parádörd, Hungary, 5 – 18 June 2013.

This document is based on lectures by C.C. on physics beyond the Standard Model at the 2013 European School of High-Energy Physics. We present a pedagogical introduction to supersymmetry, extra dimensions, and composite Higgs. We provide references to useful review literature and refer to those for more complete citations to original papers on these topics. We apologize for any omissions in our citations or choice of topics.

### 1 The Hierarchy Problem

At loop level, the Higgs mass receives corrections from self interactions, gauge loops, and fermion loops (especially the top quark). Diagrammatically,



These loops are quadratically divergent and go like  $\int d^4k (k^2 - m^2)^{-1} \sim \Lambda^2$  for some cutoff scale  $\Lambda$ . Explicitly,

$$\delta m_H^2 = \frac{\Lambda^2}{32\pi^2} \left[ 6\lambda + \frac{1}{4} (9g^2 + 3g'^2) - y_t^2 \right] \quad (1.1)$$

If  $\Lambda \gg 10$  TeV (for example,  $\Lambda \sim M_{\text{Pl}}$ ), then the quantum correction to the Higgs mass is much larger than the mass itself,  $\delta m_H^2 \gg m_H^2$ . This is the **Hierarchy problem**: the Higgs mass is quadratically sensitive to *any* mass scale of new physics. This problem is specific to elementary scalars.

Unlike scalars, the quantum corrections to fermion and gauge boson masses are proportional to the particle masses themselves. In this way, small fermion and gauge boson masses are technically natural: the loop corrections are suppressed by the smallness of the tree-level parameter. For fermions this is because of the appearance of a new chiral symmetry in the massless limit. For gauge bosons this is because gauge symmetry is restored in the massless limit. By dimensional analysis, the corrections to these mass parameters cannot be quadratically sensitive to the cutoff,  $\Lambda$ ,

$$\Delta m_e \sim m_e \ln \left( \frac{\Lambda}{m_e} \right) \quad (1.2)$$

$$\Delta M_W^2 \sim M_W^2 \ln \left( \frac{\Lambda}{m_e} \right). \quad (1.3)$$

The Hierarchy problem is independent of the renormalization scheme. It is sometimes argued that in dimensional regularization there are no quadratic divergences since the  $1/\epsilon$  poles correspond to logarithmic divergences. This is fallacious. The Hierarchy problem isn't about the cancellation of



**Fig. 1:** Heuristic two-loop contributions to the Higgs mass from heavy fermions,  $\Psi$ . Even though the  $\Psi$  do not directly couple to the Higgs, they reintroduce a quadratic sensitivity to the new scale.

divergences, it is about the separation of the electroweak and UV scales. *Any* new physics coupled to the Higgs will reintroduce the quadratic dependence on the scale at which the new physics appears. For example, suppose new physics enters at the scale  $m_S$  by a four-point interaction between the Higgs and an additional complex scalar,  $\Delta\mathcal{L} \supset \lambda_S |H|^2 |S|^2$ . The contribution to the Higgs mass from a loop of the  $S$  particle is

$$\delta m_H^2 = \frac{\lambda_S}{16\pi^2} \left[ \Lambda_{\text{UV}}^2 - 2m_S^2 \ln \left( \frac{\Lambda_{\text{UV}}}{m_S} \right) + (\text{finite}) \right]. \quad (1.4)$$

Suppose one chose to ignore the term quadratic in the loop regulator,  $\Lambda_{\text{UV}}^2$ —note that there’s no justification to do this—the logarithmically divergent piece (corresponding to the  $1/\epsilon$ ) and the finite pieces are proportional to the squared mass scale of the new physics,  $m_S^2$ . The regulator  $\Lambda_{\text{UV}}$  is not a physical scale, but  $m_S^2$  is the scale of new physics. The Higgs mass is quadratically sensitive to this scale, no matter how one chooses to regulate the loop.

This quadratic sensitivity is true even if these new states are *not* directly coupled to the Higgs but only interact with other Standard Model fields. For example, suppose there were a pair of heavy fermions  $\Psi$  which are charged under the Standard Model gauge group but don’t directly interact with the Higgs. One still expects two loop contributions to the Higgs mass from diagrams such as those in Fig. 1. These contributions are of the form

$$\delta m_H^2 \sim \left( \frac{g^2}{16\pi^2} \right)^2 \left[ a\Lambda_{\text{UV}}^2 + 48m_F^2 \ln \frac{\Lambda_{\text{UV}}}{m_F} + (\text{finite}) \right]. \quad (1.5)$$

This is indeed of the same form as (1.4). Note that in this case, the sensitivity to the new scale is softened by a loop factor.

The Higgs mass operator  $|H|^2$  is a relevant and thus grows in the infrared. From the Wilsonian perspective, the Hierarchy problem is the statement that it is difficult (finely tuned) to choose a renormalization group trajectory that flows to the correct Higgs mass. In summary, the Hierarchy problem is the issue that the Higgs mass  $m_H$  is sensitive to *any* high scale in the theory, even if it only indirectly couples to the Standard Model. Thus naïvely one would expect that  $m_H$  should be on the order of the scale of new physics. In the Wilsonian picture, the Higgs mass is a relevant operator and so its importance grows towards the IR. Indeed,  $m_H$  is the only relevant operator in the Standard Model.

The implication of the Hierarchy problem is that there should to be new physics at the TeV scale that eliminates the large loop contributions from above the TeV scale<sup>1</sup>. In these lectures we explore some of options for the physics beyond the SM that enforce naturalness. Before going into further detail, here is a brief overview of some of the possibilities for this to happen:

- **Supersymmetry:** relate the elementary scalar Higgs to fermions in such a way that the chiral symmetry protecting the fermion mass is extended to also protect the scalar mass.

<sup>1</sup>See [1] for a recent discussion of naturalness and fine-tuning in the post-Higgs era.

- **Gauge-Higgs unification:** relate the elementary scalar Higgs to an elementary gauge field so that gauge symmetry also protects the Higgs mass.
- **Technicolor, Higgsless:** there is no Higgs boson, just a dynamically generated condensate.
- **Composite Higgs, warped extra dimensions:** There is a Higgs, but it is not elementary. At the TeV scale the Higgs “dissolves”: it becomes sensitive to large form factors that suppresses corrections.
- **Pseudo-Goldstone Higgs:** The Higgs is a pseudo-Goldstone boson of a spontaneously broken symmetry. This gives some protection against quadratic divergences, usually removing the one-loop contribution. In practice one must still combine with additional mechanisms, such as in **little Higgs** models.
- **Large extra dimensions:** The fundamental Planck scale is actually  $\sim \text{TeV}$  and only appears much larger because gravity is diluted through its propagation in more directions.

## 2 Supersymmetry

Recall that under an infinitesimal transformation by an ‘ordinary’ internal symmetry, a quantum field  $\phi$  transforms as

$$\varphi_i \rightarrow (\mathbb{1}_{ij} + i\epsilon^a T_{ij}^a)\varphi_j, \quad (2.1)$$

where  $\epsilon^a$  is an infinitesimal parameter,  $T^a$  is the [bosonic] generator of the symmetry, and  $i, j$  label the representation of  $\phi$  with respect to this symmetry. These internal symmetries do not change the spin of  $\phi$ : bosons remain bosons and fermions remain fermions. **Supersymmetry** (SUSY) is a generalization of this ‘ordinary’ symmetry where generator is now fermionic. Thus a SUSY transformation changes fermions into bosons and vice versa.

**Further reading:** Wess and Bagger [?] is the canonical reference for the tools of supersymmetry. The text by Terning has a broad overview of SUSY and its modern applications in particle physics. Additional reviews include [?, ?, 2]. Key historical papers are collected in [3] and a more personal account is presented in [4]. More formal topics in SUSY that are beyond the scope of these lectures, but are key tools for model builders, can be found in [5–7].

### 2.1 The SUSY algebra

The ’60s were very successful for classifying hadrons based on Gell-Mann’s  $SU(3)$  internal symmetry. Physicists then tried to enlarge this group to  $SU(6)$  so that it would include

$$SU(3)_{\text{Gell-Mann}} \times SU(2)_{\text{spin}}, \quad (2.2)$$

but they were unable to construct a viable relativistic model. Later this was understood to be a result of the Coleman-Mandula ‘no go’ theorem which states that one cannot construct a consistent quantum field theory based on a nontrivial combination of internal symmetries with space-time symmetry [8]. The one exception came from Haag, Lopuszanski, and Sohnius: the only non-trivial combination of an internal and spacetime symmetry is to use a **graded Lie algebra** whose generators are fermionic [9]. Recall that fermionic objects obey anti-commutation relations rather than commutation relations. The main anti-commutation relation for SUSY is:

$$\{Q_\alpha^A, \bar{Q}_{\dot{\alpha}B}\} = 2P_\mu \sigma_{\alpha\dot{\alpha}}^\mu \delta_B^A, \quad (2.3)$$

where the  $Q$  and  $\bar{Q}$  are SUSY generators (supercharges) and  $P_\mu$  is the momentum operator. Here the  $\alpha$  and  $\dot{\alpha}$  are Lorentz indices while  $A, B$  index the number of supercharges. For completeness, the rest of the algebra is

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(M^{\mu\nu}\eta^{\nu\rho} + M^{\nu\rho}\eta^{\mu\sigma} - M^{\mu\rho}\eta^{\nu\sigma} - M^{\nu\sigma}\eta^{\mu\rho}) \quad (2.4)$$

$$[P^\mu, P^\nu] = 0 \quad (2.5)$$

$$[M^{\mu\nu}, P^\sigma] = i(P^\mu\eta^{\nu\sigma} - P^\nu\eta^{\mu\sigma}) \quad (2.6)$$

$$[Q_\alpha^A, M^{\mu\nu}] = (\sigma^{\mu\nu})_\alpha{}^\beta Q_\beta^A \quad (2.7)$$

$$[Q_\alpha^A, P^\mu] = 0 \quad (2.8)$$

$$\{Q_\alpha^A, Q_\beta^B\} = \epsilon_{\alpha\beta} Z^{AB}. \quad (2.9)$$

The  $Z^{AB}$  may appear for  $\mathcal{N} > 1$  and are known as **central charges**. By the Coleman-Mandula theorem, we know that internal symmetry generators commute with the Poincaré generators. For example, the Standard Model gauge group commutes with the momentum, rotation, and boost operators. This carries over to the SUSY algebra. For an internal symmetry generator  $T_a$ ,

$$[T_a, Q_\alpha] = 0. \quad (2.10)$$

This is true with one exception. The SUSY generators come equipped with their own internal symmetry, called  **$R$ -symmetry**. For  $\mathcal{N} = 1$  there exists an automorphism of the supersymmetry algebra,

$$Q_\alpha \rightarrow e^{it} Q_\alpha \quad \bar{Q}_{\dot{\alpha}} \rightarrow e^{-it} \bar{Q}_{\dot{\alpha}}, \quad (2.11)$$

for some transformation parameter  $t$ . This is a  $U(1)$  internal symmetry. Applying this symmetry preserves the SUSY algebra. If  $R$  is the generator of this  $U(1)$ , then its action on the SUSY operators is given by

$$Q_\alpha \rightarrow e^{-iRt} Q_\alpha e^{iRt}. \quad (2.12)$$

By comparing the transformation of  $Q$  under (2.12), we find the corresponding algebra,

$$[Q_\alpha, R] = Q_\alpha \quad [\bar{Q}_{\dot{\alpha}}, R] = -\bar{Q}_{\dot{\alpha}}. \quad (2.13)$$

Note that this means that different components of a SUSY multiplet have different  $R$  charge. For  $\mathcal{N} > 1$  the  $R$ -symmetry group enlarges to  $U(\mathcal{N})$ .

## 2.2 Properties of supersymmetric theories

Supersymmetric theories obey some key properties:

1. The number of fermionic degrees of freedom equals the number of bosonic degrees of freedom. To see this, first introduce an operator  $(-)^{N_F}$  such that,

$$(-)^{N_F} |q\rangle = \begin{cases} +|q\rangle & \text{boson} \\ -|q\rangle & \text{fermion} \end{cases} \quad (2.14)$$

where  $N_F$  is the fermion number operator. Note that

$$(-)^{N_F} Q_\alpha^A |q\rangle = -Q_\alpha^A (-)^{N_F} |q\rangle \quad (2.15)$$

so that  $(-)^{N_F}$  and the supercharges anticommute,  $\{(-)^{N_F}, Q_\alpha^A\} = 0$ . Next consider the operator in (2.3) weighted by  $(-)^{N_F}$ . When one sums over the states in a representation—which we write as a trace over the operator—one finds:

$$\text{Tr} \left[ (-)^{N_F} \{Q_\alpha^A, \bar{Q}_\beta^B\} \right] = \text{Tr} \left[ -Q_\alpha^A (-)^{N_F} \bar{Q}_\beta^B + (-)^{N_F} \bar{Q}_\beta^B Q_\alpha^A \right] = 0, \quad (2.16)$$

where in the last step we've used the cyclicity of the trace to convert the first term into the second term up to a minus sign. By (2.3) the left-hand side of this equation is simply  $\text{Tr} \left[ (-)^{N_F} 2\sigma_{\alpha\dot{\beta}}^{\mu} P_{\mu} \right]$ . Note that since Poincaré symmetry is assumed to be unbroken,  $P_{\mu}$  is identical for each state in a representation. Thus we are left with the conclusion that

$$\text{Tr}(-)^{N_F} = 0, \quad (2.17)$$

which implies that there is an equal number of fermions and bosons.

2. All states in a supersymmetry multiplet ('supermultiplet' or **superfield**) have the same mass. This follows from the equivalence of  $P_{\mu}$  acting on these states.
3. Energy for any state  $\Psi$  is positive semi-definite  $\langle \Psi | H | \Psi \rangle \geq 0$  and the energy for any vacuum with unbroken SUSY vanishes exactly,  $\langle 0 | H | 0 \rangle = 0$ .

### 2.3 Classification of supersymmetry representations

For the basic case of  $\mathcal{N} = 1$  SUSY there is a single supercharge  $Q$  and its conjugate  $\bar{Q}$ . The massless representations of this class of theories are separated into two cases:

- **(anti-)chiral superfield**: contains a complex scalar and a 2-component (Weyl) spinor.
- **vector superfield**: contains a 2-component (Weyl) spinor and a gauge field.

These are the *only*  $\mathcal{N} = 1$  representations that do not involve fields with spin greater than 1.

**Multiplets when there is more supersymmetry.** If there are more SUSY charges, e.g.  $\mathcal{N} = 2$ , then the smallest representation is the **hypermultiplet** which contains a 4-component (Dirac) fermion and two complex scalars. For supersymmetric extensions of the SM it is sufficient to focus only on the  $\mathcal{N} = 1$  case since this is the only case which admits the observed chiral fermions of the Standard Model.

One can compare the number of bosonic and fermionic degrees of freedom in these representations. In the chiral superfield, the complex scalar carries 2 degrees of freedom while the complex Weyl spinor carries 4 degrees of freedom. Recall, however, that fermions only have two helicity states so that in fact only 2 of these fermionic degrees of freedom propagate on-shell. Since one of the key points of using fields to describe physical particles is that we can describe off-shell propagation, we would like to also have supersymmetry hold off-shell. This requires adding two 'dummy' scalar degrees of freedom, which we package in a non-propagating 'auxiliary' complex field  $F$ :

Field	off-shell degrees of freedom	on-shell degrees of freedom
scalar, $\phi$	2	2
fermion, $\psi$	4	2
auxiliary, $F$	2	0

For the vector superfield the Weyl spinor has 4 (2) off-(on-)shell degrees of freedom while the massless gauge boson has 3 (2) off(on-)shell degrees of freedom after identifying gauge equivalent states. As in the chiral superfield, the number of on-shell degrees of freedom match automatically while the number of off-shell degrees of freedom require an additional non-propagating auxiliary field. In this case we introduce a real scalar,  $D$ :

Field	off-shell degrees of freedom	on-shell degrees of freedom
fermion, $\psi$	4	2
gauge boson, $A_{\mu}$	3	2
auxiliary, $D$	1	0

## 2.4 Superspace

The most convenient way to describe  $\mathcal{N} = 1$  supersymmetric field theories is to use the **superspace** formalism. Here we understand the supersymmetry transformation generated by  $Q$  and  $\bar{Q}$  to be a spacetime transformation in an additional fermionic dimension. To do this, we introduce Weyl spinor superspace coordinates  $\theta_\alpha$  and  $\bar{\theta}^{\dot{\alpha}}$ . Superfields are functions of  $x$ ,  $\theta$ , and  $\bar{\theta}$  and encode all of the off-shell degrees of freedom of a supermultiplet.

**Weyl spinors and van der Waerden notation.** We assume familiarity with two-component Weyl spinors. These are the natural language for fermions in four-dimensions. We use the van der Waerden notation with dotted and undotted indices to distinguish the indices of left- and right-chiral spinors. Readers unfamiliar with this notation may consult [?, ?]. The encyclopedic ‘two component bible’ is a useful reference for full details and as a template for doing calculations [10].

The SUSY algebra tells us that the effect of a SUSY transformation with infinitesimal parameters  $\epsilon$  and  $\bar{\epsilon}$  on a superspace coordinate  $(x, \theta, \bar{\theta})$  is

$$(x^\mu, \theta, \bar{\theta}) \rightarrow (x^\mu + i\theta\sigma^\mu\bar{\epsilon} - i\epsilon\sigma^\mu\bar{\theta}, \theta + \epsilon, \bar{\theta} + \bar{\epsilon}). \quad (2.18)$$

It is useful to define the superspace covariant derivatives,

$$D_\alpha = +\frac{\partial}{\partial\theta^\alpha} + i\sigma_{\alpha\dot{\alpha}}^\mu\bar{\theta}^{\dot{\alpha}}\partial_\mu \quad \bar{D}_{\dot{\alpha}} = -\frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}} - i\theta^\alpha\sigma_{\alpha\dot{\alpha}}^\mu\partial_\mu. \quad (2.19)$$

These are ‘covariant derivatives’ in that they anticommute with the SUSY generators<sup>2</sup>. They satisfy

$$\{D_\alpha, \bar{D}_{\dot{\beta}}\} = -2i(\sigma^\mu)_{\alpha\dot{\beta}}\partial_\mu \quad \text{and} \quad \{D_\alpha, D_\beta\} = \{\bar{D}_{\dot{\alpha}}, \bar{D}_{\dot{\beta}}\} = 0 \quad (2.20)$$

By expanding in the fermionic coordinates, a generic superfield  $F(x, \theta, \bar{\theta})$  can be written in terms of component fields of different spin that propagate on ordinary spacetime,

$$F(x, \theta, \bar{\theta}) = f(x) + \theta\psi(x) + \bar{\theta}\bar{\chi}(x) + \theta^2 M(x) + \bar{\theta}^2 N(x) + \theta\sigma^\mu\bar{\theta}v_\mu(x) + \theta^2\bar{\theta}\bar{\lambda}(x) + \bar{\theta}^2\theta\xi + \theta^2\bar{\theta}^2 D(x).$$

This expansion is exact because higher powers of  $\theta$  or  $\bar{\theta}$  vanish identically because an anticommuting number  $\theta_1$  satisfies  $(\theta_1)^2 = 0$ . As a sanity check, we are allowed quadratic terms in  $\theta$  since it is a Weyl spinor and  $\theta^2 = \theta^\alpha\theta_\alpha = \epsilon^{\alpha\beta}\theta_\beta\theta_\alpha = 2\theta_1\theta_2$ .

With modest effort, one can work out the transformation of each component of this general superfield by applying the transformation (2.18), expanding all fields in  $\theta$  and  $\bar{\theta}$ , and matching the coefficients of each term. Some of the terms require massaging by Fierz identities to get to the correct form. Fortunately, the general superfield above is a reducible representation: some of these fields do not transform into one another. We can restrict to irreducible representations by imposing one of the following conditions:

$$\text{chiral superfield} \quad \bar{D}_\alpha\Phi = 0 \quad (2.21)$$

$$\text{anti-chiral superfield} \quad D_{\dot{\alpha}}\bar{\Phi} = 0 \quad (2.22)$$

$$\text{vector (real) superfield} \quad V = V^\dagger \quad (2.23)$$

$$\text{linear superfield} \quad \bar{D}^2 L = D^2 L = 0 \quad (2.24)$$

The chiral and anti-chiral superfields carry Weyl fermions of left- and right-handed helicity respectively. It is convenient to write all anti-chiral superfields into chiral superfields, for example by swapping the right-handed electron chiral superfield with a left-handed positron superfield. The field content is identical, one is just swapping which is the ‘particle’ and which is the ‘anti-particle.’

<sup>2</sup>One may be used to thinking of covariant derivatives as coming from local symmetries with some gauge field. Here, however, we consider only *global* SUSY. Geometrically, the covariant derivative comes from the fact that even rigid superspace carries torsion [11].

**The linear superfield.** The defining condition for this superfield includes a constraint that the vector component is divergence free,  $\partial_\mu V^\mu = 0$ . It is thus a natural supersymmetrization of a conserved current. We will not consider linear superfields further in these lectures.

## 2.5 Supersymmetric Lagrangians for chiral superfields

One can check that because  $\bar{D}_{\dot{\alpha}}(x^\mu + i\theta\sigma^\mu\bar{\theta}) = 0$ , any function of  $y^\mu = x^\mu + i\theta\sigma^\mu\bar{\theta}$  is automatically a chiral superfield ( $\chi_{\text{SF}}$ ). Indeed, the most compact way of writing the components of a  $\chi_{\text{SF}}$  is

$$\Phi(y, \theta) = \varphi(y) + \sqrt{2}\theta\psi(y) + \theta^2 F(y). \quad (2.25)$$

Again, we point out that this expansion is exact since higher powers of the Weyl spinor  $\theta$  vanish by the antisymmetry of its components. Under a SUSY transformation with parameter  $\epsilon$ , the components of the  $\chi_{\text{SF}}$  each transform as

$$\delta_\epsilon \varphi(x) = \sqrt{2}\epsilon\psi(x) \quad (2.26)$$

$$\delta\psi(x) = i\sqrt{2}\sigma^\mu\bar{\epsilon}\partial_\mu\varphi(x) + \sqrt{2}\epsilon F(x) \quad (2.27)$$

$$\delta_\epsilon F(x) = i\sqrt{2}\bar{\epsilon}\bar{\sigma}^\mu\partial_\mu\psi(x). \quad (2.28)$$

Observe that the auxiliary field transforms into a total spacetime derivative. This is especially nice since a total derivative vanishes in the action and so the highest component of a  $\chi_{\text{SF}}$  is a candidate for a SUSY-invariant term in the Lagrangian. Thus we arrive at our first way of constructing supersymmetric Lagrangian terms: write the  $F$ -term of a chiral superfield.

To generate interesting interactions we don't want to write the  $F$ -terms of our fundamental fields—indeed, these are generally not even gauge invariant. Fortunately, one can check that a product of chiral superfields is itself a chiral superfield. Indeed, a general way of writing a supersymmetry Lagrangian term built out of chiral superfields is

$$\mathcal{L} = \int d^2\theta W(\Phi) + \text{h.c.}, \quad (2.29)$$

where  $W$  is a holomorphic function of chiral superfields called the **superpotential**. Note that the integral over  $d^2\theta$  is an ordinary fermionic integral that just picks out the highest component of  $W$ . Performing the fermionic integral gives Lagrangian terms

$$\mathcal{L} = -\frac{\partial^2 W(\varphi)}{\partial\Phi_i\partial\Phi_j}\psi_i\psi_j - \sum_i \left| \frac{\partial W(\varphi)}{\partial\Phi_i} \right|^2. \quad (2.30)$$

Observe that the superpotential is evaluated on the scalar components of the superfields,  $\Phi = \varphi$ . One can check that restricting to renormalizable terms in the Lagrangian limits the mass dimension of the superpotential to  $[W] \leq 3$ .

**Cancellation of quadratic divergences.** One can check from explicit calculations that the SUSY formalism ensures the existence of superpartner particles with just the right couplings to cancel quadratic divergences. A more elegant way to see this, however, is to note that the symmetries of superspace itself prevent this. While it is beyond the scope of these lectures, the superpotential is not renormalized perturbatively—see, e.g. [5, 12] for details. The holomorphy of  $W$  plays a key role in these arguments. The symmetries of the theory enforce the technical naturalness of parameters in  $W$ , including scalar masses.

Superpotential terms, however, do not include the usual kinetic terms for propagating fields. In fact, one can show that these terms appear in the  $\theta^2\bar{\theta}^2$  term of the combination

$$\Phi^\dagger\Phi \Big|_{\theta^2\bar{\theta}^2} = FF^* + \frac{1}{4}\varphi^*\partial^2\varphi + \frac{1}{4}\partial^2\varphi^*\varphi - \frac{1}{2}\partial_\mu\varphi^*\partial^\mu\varphi + \frac{i}{2}\partial_\mu\bar{\psi}\bar{\sigma}^\mu\psi - \frac{i}{2}\bar{\psi}\bar{\sigma}^\mu\partial_\mu\psi. \quad (2.31)$$

Two immediate observations are in order:

1. The complex scalar  $\varphi$  and Weyl fermion  $\psi$  each have their canonical kinetic term. The non-propagating field,  $F$ , does not have any derivative terms: its equation of motion is algebraic and can be solved explicitly. This is precisely what is meant that  $F$  is auxiliary.
2.  $\Phi^\dagger\Phi$  is not a chiral superfield. In fact, it's a real superfield and the  $\theta^2\bar{\theta}^2$  component is the auxiliary  $D$  field. Indeed, in the same way that the highest component of a  $\chi$ SF transforms into a total derivative, the highest component of a real superfield also transforms into a total derivative and is a candidate term for the Lagrangian.

We thus arrive at the second way to write supersymmetric Lagrangian terms: take the  $D$ -term of a real superfield. We may write this term as an integral over superspace,  $\int d^4\theta \Phi^\dagger\Phi$ , where  $d^4\theta = d^2\theta d^2\bar{\theta}$ .

More generally, we may write a generic real function  $K(\Phi, \Phi^\dagger)$  of chiral superfields,  $\Phi$  and  $\Phi^\dagger$ , whose  $D$  term is supersymmetric contribution to the Lagrangian. This is called the **Kähler potential**. The simplest Kähler potential built out of chiral superfields is precisely (2.31) and includes the necessary kinetic terms for the chiral superfield. One can check that restricting to renormalizable terms in the Lagrangian limits the mass dimension of the Kähler potential to  $[K] \leq 2$ . Combined with the condition that  $K$  is real and the observation that chiral superfields are typically not gauge invariant, this usually restricts the Kähler potential to take the canonical form,  $K = \Phi_i^\dagger\Phi_i$ .

The most general  $\mathcal{N} = 1$  supersymmetric Lagrangian for chiral superfields is thus

$$\mathcal{L} = \int d^4\theta K(\Phi, \Phi^\dagger) + \left( \int d^2\theta W(\Phi) + \text{h.c.} \right). \quad (2.32)$$

This expression is general, but renormalizability restricts the mass dimensions to be  $[K] \leq 2$  and  $[W] \leq 3$ . For theories with more supersymmetry, e.g.  $\mathcal{N} = 2$ , one must impose additional relations between  $K$  and  $W$ . Assuming a renormalizable supersymmetric theory of chiral superfields  $\Phi_i$ , we may plug in  $K = \Phi_i^\dagger\Phi_i$  and integrate out the auxiliary fields from (2.32). The result is

$$\mathcal{L} = \partial_\mu\varphi_i^*\partial^\mu\varphi_i + i\bar{\psi}_i\bar{\sigma}^\mu\partial_\mu\psi_i - \frac{\partial^2 W}{\partial\varphi_i\partial\varphi_j}\psi_i\psi_j - \sum_i \left| \frac{\partial W}{\partial\varphi_i} \right|^2. \quad (2.33)$$

Here the superpotential is assumed to be evaluated at its lowest component so that  $W[\Phi_i(y, \theta)] \rightarrow W[\varphi_i(x)]$ . Observe that dimension-2 terms in the superpotential link the mass terms of the Weyl fermion and the complex scalar. Further, dimension-3 terms in the superpotential connect Yukawa interactions to quartic scalar couplings.

## 2.6 Supersymmetric Lagrangians for vector superfields

Until now, however, we have only described supersymmetric theories of complex scalars and fermions packaged as chiral superfields. In order to include the interactions of gauge fields we must write down SUSY Lagrangians that include vector superfields.

Suppose a set of chiral superfields  $\Phi$  carry a  $U(1)$  charge such that  $\Phi(x) \rightarrow \exp(-i\Lambda)\Phi(x)$ . For an ordinary global symmetry this is an overall phase on each component of the chiral superfield. For a gauge symmetry, the transformation parameter is spacetime dependent,  $\Lambda = \Lambda(x)$ . Note, however, that this is now problematic because our definition of a chiral superfield,  $\bar{D}_\alpha\Phi = 0$ , contains a spacetime derivative. It would appear that the naïve gauge transformation is not consistent with the irreducible SUSY representations we've written because it does not preserve the chiral superfield condition.

This inconsistency is a relic of keeping  $\Lambda(x)$  a function of spacetime rather than a function of the full superspace. We noted above that a function of  $y^\mu = x^\mu + i\theta\sigma^\mu\bar{\theta}$  is a chiral superfield and, further, that a product of chiral superfields is also a chiral superfield. Thus a consistent way to include gauge transformations is to promote  $\Lambda(x)$  to a chiral superfield  $\Lambda(y)$  so that  $\exp(-i\Lambda(y))\Phi(y)$  is indeed chiral. In this way we see that supersymmetry has 'complexified' the gauge group.



Under this complexified gauge transformation, the canonical Kähler potential term that contains the kinetic terms transforms to

$$\Phi^\dagger \Phi \rightarrow \Phi^\dagger e^{-i(\Lambda - \Lambda^\dagger)} \Phi. \quad (2.34)$$

For gauge theories one must modify the Kähler potential to accommodate this factor. This is unsurprising since gauging an ordinary quantum field theory requires one to modify the kinetic terms by promoting derivatives to covariant derivatives which include the gauge field. To correctly gauge a symmetry, we introduce a vector (real) superfield (VSF)  $V$  which transforms according to

$$V \rightarrow V + i(\Lambda - \Lambda^\dagger) \quad (2.35)$$

and promote the Kähler potential to

$$K(\Phi, \Phi^\dagger) = \Phi^\dagger e^V \Phi. \quad (2.36)$$

A generic VSF has many components, but many can be eliminated by partially gauge fixing to the **Wess-Zumino** gauge where

$$V = -\theta^\mu \bar{\theta} V_\mu(x) + i\theta^2 \bar{\theta} \bar{\lambda}(x) - i\bar{\theta}^2 \theta \lambda(x) + \frac{1}{2} \theta^2 \bar{\theta}^2 D(x). \quad (2.37)$$

here  $V_\mu(x)$  is the gauge field of the local symmetry,  $\lambda(x)$  and  $\bar{\lambda}(x) = \lambda^\dagger(x)$  are **gauginos**, and  $D(x)$  is the auxiliary field needed to match off-shell fermionic and bosonic degrees of freedom. The two gauginos are the pair of two-component spinors that make up a Majorana four-component spinor. This gauge choice fixes the complex part of the ‘complexified’ gauge symmetry, leaving the ordinary spacetime (rather than superspace) gauge redundancy that we are familiar with in quantum field theory.

We have not yet written a kinetic term for the vector superfield. A useful first step is to construct the chiral superfield,

$$\mathcal{W}_\alpha = -\frac{1}{4} \bar{D}^{\dot{\alpha}} \bar{D}_{\dot{\alpha}} D_\alpha V \quad (2.38)$$

$$= -i\lambda_\alpha(y) + \theta_\beta \left[ \delta_\alpha^\beta D(y) - \frac{i}{2} (\sigma^\mu \bar{\sigma}^\nu)_\alpha^\beta F_{\mu\nu}(y) \right] + \theta^2 \sigma_{\alpha\dot{\alpha}}^\mu \partial_\mu \bar{\lambda}^{\dot{\alpha}}(y). \quad (2.39)$$

One can see that  $\mathcal{W}_\alpha$  is a chiral superfield because  $\bar{D}_{\dot{\beta}} \mathcal{W}_\alpha = 0$  from the antisymmetry of the components of  $\bar{D}$ , (2.20). Observe that unlike  $\Phi$ , the lowest component is a spin-1/2 field. Further,  $\mathcal{W}$  contains the usual gauge field strength. Indeed, one can write the supersymmetric Yang Mills kinetic terms for the vector superfield as

$$\mathcal{L}_{\text{SYM}} = \frac{1}{4} \mathcal{W}_\alpha \mathcal{W}^\alpha \Big|_\theta^2 + \text{h.c.} = \frac{1}{4} \int d^2\theta \mathcal{W}^2 + \text{h.c.} \quad (2.40)$$

One can check that this gives the usual kinetic terms for the gauge field and gauginos as well as an auxiliary term. For completeness, the general form of the field strength superfield for a non-Abelian supersymmetric gauge theory is

$$T^a \mathcal{W}_\alpha^a = -\frac{1}{4} \bar{D}^{\dot{a}} \bar{D}_{\dot{a}} e^{-T^a V^a} D_\alpha e^{T^a V^a}. \quad (2.41)$$

Under a non-Abelian gauge transformation the chiral and vector superfields transform as

$$\Phi \rightarrow e^{-g T^a \Lambda^a} \Phi \quad (2.42)$$

$$e^{T^a V^a} \rightarrow e^{T^a \Lambda^a} e^{T^a V^a} e^{T^a \Lambda^a}. \quad (2.43)$$

The final form of the renormalizable, gauge-invariant,  $\mathcal{N} = 1$  supersymmetric Lagrangian is

$$\mathcal{L} = \int d^4\theta \Phi_i^\dagger e^{gV} \Phi_i + \int d^2\theta \left( \frac{1}{4} \mathcal{W}_\alpha^a \mathcal{W}^{\alpha a} + \text{h.c.} \right) + \int d^2\theta (W(\Phi) + \text{h.c.}). \quad (2.44)$$

**Non-renormalization and the gauge kinetic term.** Although  $\mathcal{W}^2$  looks like it could be a superpotential term, it is important to treat it separately since it is the kinetic term for the gauge fields. Further the arguments that the superpotential is not renormalized in perturbation theory do *not* hold for the  $\mathcal{W}^2$  term. Indeed, the prefactor of  $\mathcal{W}^2$  can be identified with the [holomorphic] gauge coupling, which is only corrected perturbatively at one loop order. One way to see this is to note that for non-Abelian theories, the gauge kinetic term  $\mathcal{W}^2 d^2\theta + \text{h.c.}$  also includes a topological term,  $F\tilde{F}$ , which we know is related to anomalies. Another way to see this is the note that the simplest demonstration of non-renormalization of the superpotential makes use of holomorphy and the global symmetries of  $W$ : the vector (real) superfield from which  $\mathcal{W}_\alpha$  is built, however, is not holomorphic and its fields cannot carry have the  $U(1)$  global symmetries used in the proof.

## 2.7 Example: SUSY QED

As a simple example, consider the supersymmetric version of quantum electrodynamics, SQED. In ordinary QED we start with a Dirac spinor representing the electron and positron. Since we've seen above that a chiral superfield only contains a Weyl spinor, we require two chiral superfields,  $\Phi_\pm$ , which we may interpret to be the electron and positron superfields. Our only two inputs are the electromagnetic coupling  $e$  and the electron mass  $m$ . The latter suggests a superpotential

$$W(\Phi_+, \Phi_-) = m\Phi_+\Phi_-. \quad (2.45)$$

Writing out the resulting Lagrangian in components:

$$\begin{aligned} \mathcal{L}_{\text{SQED}} = & \left[ \frac{1}{2} D^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - i\lambda\sigma^\mu \partial_\mu \bar{\lambda} \right] \\ & + F_+^* F_+ + |D_\mu \varphi_+|^2 + i\bar{\psi}_+ D_\mu \bar{\sigma}^\mu \psi_+ \\ & + F_-^* F_- + |D_\mu \varphi_-|^2 + i\bar{\psi}_- D_\mu \bar{\sigma}^\mu \psi_- \\ & - \frac{ie}{\sqrt{2}} (\varphi_+ \bar{\psi}_+ \bar{\lambda} - \varphi_- \bar{\psi}_- \bar{\lambda}) + \text{h.c.} \\ & + \frac{e}{2} D (|\varphi_+|^2 - |\varphi_-|^2) \\ & + m (\varphi_+ F_- + \varphi_- F_+ - \psi_+ \psi_-) + \text{h.c.} \end{aligned} \quad (2.46)$$

We can write this out explicitly by solving for the auxiliary fields  $D, F_\pm$ . The equations of motion are

$$D = -\frac{e}{2} (|\varphi_+|^2 - |\varphi_-|^2) \quad F_\pm = -m\varphi_\mp^*. \quad (2.47)$$

Plugging this back into the Lagrangian gives

$$\begin{aligned} \mathcal{L}_{\text{SQED}} = & \sum_{i=\pm} (|D_\mu \varphi_i|^2 + i\bar{\psi}_i D_\mu \bar{\sigma}^\mu \psi_i) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - i\lambda\sigma^\mu \partial_\mu \bar{\lambda} \\ & - m^2 (|\phi_+|^2 + |\phi_-|^2) - m\psi_+\psi_- - m\bar{\psi}_+\bar{\psi}_- \\ & - \frac{e^2}{8} (|\varphi_+|^2 - |\varphi_-|^2)^2 - \left[ \frac{ie}{\sqrt{2}} (\varphi_+ \bar{\psi}_+ \bar{\lambda} - \varphi_- \bar{\psi}_- \bar{\lambda}) + \text{h.c.} \right]. \end{aligned} \quad (2.48)$$

The first line gives the kinetic terms for the electron  $\psi_-$ , positron  $\psi_+$ , selectron ( $\phi_-$ ), spositron ( $\phi_+$ ), photon  $A_\mu$ , and photino  $\lambda$ . The second line gives an equivalent mass to the chiral scalars and fermions. The last line gives vertices that come from the supersymmetrization of the kinetic terms: four-point scalar interactions from the  $D$  terms and a three-point Yukawa-like vertex with the 'chiral' scalars and photino. The relation between the gauge group and the four-point scalar interaction plays a central role in how the Higgs fits into SUSY, as we show below.

$\chi_{\text{SF}}$	$\text{SU}(3)_c$	$\text{SU}(2)_L$	$\text{U}(1)_Y$
$Q$	$\mathbf{3}$	$\mathbf{2}$	$1/6$
$\bar{U}$	$\bar{\mathbf{3}}$	$\mathbf{1}$	$-2/3$
$\bar{D}$	$\bar{\mathbf{3}}$	$\mathbf{1}$	$1/3$
$L$	$\mathbf{1}$	$\mathbf{2}$	$-1/2$
$\bar{E}$	$\mathbf{1}$	$\mathbf{1}$	$-1$
$H_d$	$\mathbf{1}$	$\mathbf{2}$	$1/2$
$H_u$	$\mathbf{1}$	$\mathbf{1}$	$-1/2$

Table 1: Matter content of the MSSM. Note that we have used  $\mathbf{2} = \bar{\mathbf{2}}$  for  $\text{SU}(2)_L$ .

## 2.8 The MSSM

We now focus on the minimal supersymmetric extension of the Standard Model, the MSSM. To go from the SM to the MSSM, it is sufficient to promote each SM chiral fermion into a chiral superfield and each SM gauge field into a vector superfield. Thus for each SM fermion there is a new propagating scalar sfermion (squarks or sleptons) and for each SM gauge field there is also a propagating gaugino, a fermion in the adjoint representation. As we showed above, off-shell SUSY also implies non-propagating auxiliary fields.

The matter ( $\chi_{\text{SF}}$ ) content of the MSSM is shown in Table 1. It is the same as the SM except that we require two Higgs doublet chiral superfields. This is necessary for the cancellation of the  $\text{SU}(2)_L^2 \times \text{U}(1)_Y$  and  $\text{SU}(2)_L$  Witten anomalies coming from the Higgs fermions, or Higgsinos. An additional hint that this is necessary comes from the observation that the superpotential is a holomorphic function of the chiral superfields while the Standard Model up-type Yukawa coupling requires the conjugate of the Higgs,  $\tilde{H} = i\sigma^2 H^*$ .

The most general renormalizable superpotential made with these fields can be split into two terms,  $W = W^{(\text{good})} + W^{(\text{bad})}$ ,

$$W^{(\text{good})} = y_u^{ij} Q^i H_u \bar{U}^j + y_d^{ij} Q^i H_d \bar{D}^j + y_e^{ij} L^i H_d \bar{E}^j + \mu H_u H_d \quad (2.49)$$

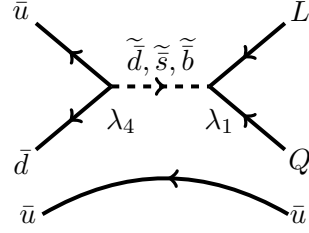
$$W^{(\text{bad})} = \lambda_1^{ijk} Q^i L^j \bar{D}^k + \lambda_2^{ijk} L^i L^j \bar{E}^k + \lambda_3^i L^i H_u + \lambda_4^{ijk} \bar{D}^i \bar{D}^j \bar{U}^k. \quad (2.50)$$

In  $W^{(\text{good})}$  one can straight forwardly identify the Standard Model Yukawa couplings which give the SM fermions their masses. Since these are packaged into the superpotential these terms also encode the additional scalar quartic interactions required by supersymmetry. The last term in  $W^{(\text{good})}$  is a supersymmetric Higgs mass known as the  $\mu$ -term. By supersymmetry this term also gives a mass to the Higgsinos, which we require since we do not observe any very light chiral fermions with the quantum numbers of a Higgs.

The  $W^{(\text{bad})}$  terms, on the other hand, are phenomenologically undesirable. These are renormalizable interactions which violate baryon ( $B$ ) and/or lepton ( $L$ ) number and are thus constrained to have very small coefficients. Compare this to the SM where  $B$  and  $L$  are accidental symmetries: all renormalizable interactions of SM fields allowed by the SM gauge group preserve  $B$  and  $L$ . Violation of these symmetries only occurs at the non-renormalizable level and are suppressed by what can be a very high scale, e.g.  $M_{\text{GUT}}$ .

We see that in the MSSM we must find ways to forbid, or otherwise strongly suppress, the terms in  $W^{(\text{bad})}$ . Otherwise one would be faced with dangerous rates for rare processes such as proton decay,  $p^+ \rightarrow e^+ \pi^0$  or  $\bar{\nu} \pi^+$  (or alternately with  $\pi$  replaced with  $K$ ) as shown in Fig. 2. Observe that this is a tree level process and all of the couplings are completely unsuppressed.

A simple way to forbid  $W^{(\text{bad})}$  is to impose **matter parity**, which is a  $\mathbb{Z}_2$  symmetry with assign-



**Fig. 2:** Proton decay mediated by squarks. Arrows indicate *helicity* and should not be confused with the ‘charge flow’ arrows of Dirac spinors [10]. Tildes indicate superpartners while bars are used to write right-chiral antiparticles into left-chiral fields in the conjugate representation.

ments:

Superfield	Matter parity
quark, lepton $\chi_{\text{SF}}$	$P_M = -1$
Higgs $\chi_{\text{SF}}$	$P_M = +1$
gauge $\nu_{\text{SF}}$	$P_M = +1$ .

Under these assignments, all terms in  $W^{(\text{good})}$  have  $P_M = +1$  while all terms in  $^{(\text{bad})}$  have  $P_M = -1$ . One can check that one may write matter parity in terms of baryon and lepton number as

$$P_M = (-)^{3(B-L)}. \quad (2.51)$$

A common variation of this is to impose the above constraint using *R-parity*,

$$P_R = (-)^{3(B-L)+2s}, \quad (2.52)$$

where  $s$  is the spin of the field. Conservation of matter parity implies conservation of *R-parity*. This is because the  $(-)^{2s}$  factor always cancels in any interaction term since Lorentz invariance requires that any such term has an even number of fermions. Observe that all SM fields have *R-parity*  $+1$  while all superpartner fields have *R-parity*  $-1$ . (This is similar to *T-parity* for Little Higgs models.) The diagrams associated with electroweak precision observables carry only SM external states. Since *R-parity* requires pair-production of superpartners, this means that electroweak precision corrections cannot occur at tree-level and must come from loop diagrams.

It is important to understand that *R-parity* (or matter parity) is an additional symmetry that we impose on top of supersymmetry. *R-parity* has some important consequences:

1. The lightest *R-parity* odd particle is stable. This is known as the **lightest supersymmetric particle** or LSP. If the LSP is an electrically neutral color singlet—as we shall assume—it is a candidate for WIMP-like DM.
2. Each superpartner (sparticle) other than the LSP will decay. At the end of any such sequence of decays one is left with an odd number (usually one) of LSPs.
3. In collider experiments, the initial state has  $P_R = +1$  so that only an even number of sparticles can be produced at a time (e.g. via pair production). At the end of the decay these end up as LSPs which manifest themselves as missing energy signals at colliders.

For most of this document we postulate that the MSSM has exact *R-parity* conservation—though this is something of an ad-hoc assumption.

## 2.9 Supersymmetry breaking

Any scalar partners to the SM leptons or quarks with exactly degenerate masses as their SM partner would have been discovered long ago. Thus, the next piece required to construct a realistic MSSM is a way to break supersymmetry and split the mass degeneracy between the SM particles and their superpartners. Since we want to keep the desirable ultraviolet behavior of supersymmetry, we assume that SUSY is a fundamental symmetry of nature which is spontaneously broken.

SUSY is unbroken when the supercharges annihilate the vacuum,  $Q|0\rangle = \bar{Q}|0\rangle = 0$ . The SUSY algebra,  $\{Q, \bar{Q}\} = 2\sigma^\mu P_\mu$  allows us to write the four-momentum operator as  $P^\mu = \frac{1}{4}\bar{\sigma}^\nu\{Q, \bar{Q}\}$  so that the Hamiltonian is

$$H = P^0 = \frac{1}{4} (Q_1\bar{Q}_1 + \bar{Q}_1Q_1 + Q_2\bar{Q}_2 + \bar{Q}_2Q_2). \quad (2.53)$$

Observing that this expression is positive semi-definite, we see that

$$\begin{aligned} \text{if SUSY is unbroken,} & \quad \langle 0|H|0\rangle = 0 \\ \text{if SUSY is broken,} & \quad \langle 0|H|0\rangle > 0 \end{aligned} .$$

The vacuum energy can be read from the scalar potential,

$$V[\phi] = V_F[\phi] + V_D[\phi] \quad (2.54)$$

$$V_F[\phi] = \sum_i \left| \frac{\partial W}{\partial \phi_i} \right|^2 = \sum_i |F_i|^2 \quad (2.55)$$

$$V_D[\phi] = \sum_a \frac{1}{2}g^2 \left| \sum_i \phi_i^\dagger T^a \phi_i \right|^2 = \sum_a \frac{1}{2}gD^a D^a. \quad (2.56)$$

We see that SUSY breaking corresponds to one of the auxiliary fields,  $F_i$  or  $D_i$ , picking up a vacuum expectation value (VEV). We refer to the case  $\langle F_i \rangle \neq 0$  as  $F$ -type SUSY breaking and the case  $\langle D \rangle \neq 0$  as  $D$ -type SUSY breaking.

When an ordinary global symmetry is spontaneously broken due to a field picking up a VEV there exists a massless boson in the spectrum of the theory known as the Goldstone boson. In the same way, when SUSY is broken spontaneously due to a auxiliary field picking up a VEV, there exists a massless *fermion* in the theory known as the **Goldstino**<sup>3</sup>. The spin of this field is inherited by the spin of the SUSY generators. Heuristically, the massless Goldstone modes correspond to acting on the VEV with the broken generators and promoting the transformation parameters to fields. Since the SUSY transformation parameter is fermionic, the Goldstone field must also be fermionic.

For example, if  $\langle F \rangle \neq 0$ , then the transformation of the fermion  $\psi$  under the broken (SUSY) generator is

$$\delta_\epsilon \psi = 2\epsilon \langle F \rangle. \quad (2.57)$$

SUSY acts as a shift in the fermion, analogously to the shift symmetry of a Goldstone boson under a spontaneously broken global internal symmetry. If there is more than one superfield with a non-zero  $F$  term, then

$$\delta_\epsilon \psi_i = 2\epsilon \langle F_i \rangle \quad (2.58)$$

$$\psi_{\text{Goldstone}} = \sum_i \frac{F_i}{\sqrt{\sum_i F_i^2}} \psi_i. \quad (2.59)$$

<sup>3</sup>This is somewhat unfortunate nomenclature. One would expect the massless mode coming from spontaneously broken SUSY to be called a Goldstone fermion whereas the ‘Goldstino’ should refer to the supersymmetric partner of a Goldstone boson coming from the spontaneous breaking of an ordinary symmetry.

Note that we have used the convention that, when there is no ambiguity,  $F$  refers to the SUSY breaking background value, dropping the brackets  $\langle \dots \rangle$  to avoid clutter. One can further generalize this to include a linear combination of gauginos when there is also  $D$ -term SUSY breaking.

When ordinary spontaneously broken internal symmetries are promoted to gauge symmetries, their Goldstone modes are ‘eaten’ and become the longitudinal polarization of the gauge fields. Similarly, gauging supersymmetry corresponds to writing a theory of supergravity. The gravitino then becomes massive by eating the Goldstino from spontaneous SUSY breaking.

## 2.10 Sum rule for broken SUSY

Even when it is spontaneously broken, SUSY is a strong constraint on the parameters of a theory. One of the most important constraints is the SUSY sum rule, which relates the traces of the mass matrices of particles of different spins.

First consider the mass terms for chiral fermions ( $\psi$ ) and gauginos ( $\lambda$ ):

$$i\sqrt{2}g (T^a)_j^i (\varphi_i \bar{\lambda}^a \bar{\psi}^j - \varphi^* \lambda \psi) - \frac{\partial^2 W}{\partial \varphi_i \partial \varphi_j} \psi_i \psi_j + \text{h.c.} \quad (2.60)$$

We may write this succinctly as a mass matrix,

$$(\psi_i \quad \lambda_a) \begin{pmatrix} F_{ij} & \sqrt{2}D_{bi} \\ \sqrt{2}D_{aj} & 0 \end{pmatrix} \begin{pmatrix} \psi_j \\ \lambda_b \end{pmatrix}, \quad (2.61)$$

where we use the shorthand notation

$$F_{ij} = \frac{\partial F_i}{\partial \varphi_j} = \frac{\partial^2 W}{\partial \varphi_i \partial \varphi_j} \quad D_{ai} = \frac{\partial D_a}{\partial \varphi_i} = g\varphi_i^* T^a. \quad (2.62)$$

Call this fermion mass matrix  $m^{(j=1/2)}$ . Next, the scalar mass matrix  $(m^2)_{ij}^{(j=0)}$  is obtained by the Hessian of the scalar potential,

$$\begin{pmatrix} \frac{\partial^2 V}{\partial \varphi_i \partial \varphi_j^*} & \frac{\partial^2 V}{\partial \varphi_i \partial \varphi_j} \\ \frac{\partial^2 V}{\partial \varphi_i^* \partial \varphi_j^*} & \frac{\partial^2 V}{\partial \varphi_i^* \partial \varphi_j} \end{pmatrix} = \begin{pmatrix} \bar{F}^{ij} F_{kj} + D_a^i D_{aj} + D_{aj}^i D_a & \bar{F}^{ijk} F_k + D_a^j D_a^j \\ F_{ijk} \bar{F}^k + D_{ai} D_{aj} & F_{ik} \bar{F}^{jk} + D_{ai} D_a^j + D_{ai}^j D_a \end{pmatrix}. \quad (2.63)$$

Finally, the gauge boson matrix comes from the kinetic terms

$$\sum_i g^2 |A_\mu^a T^{a\alpha}{}_\beta \phi_{i\alpha}|^2 = |A_\mu^a D_a^i|^2, \quad (2.64)$$

and may thus be written

$$(m^2)_{ab}^{(j=1)} = D_a^i D_{bi} + D_{ai} D_b^i. \quad (2.65)$$

The traces of the squared mass matrices are, respectively,

$$\text{Tr } m^{(j=1/2)} \left( m^{(j=1/2)} \right)^\dagger = F_{ij} \bar{F}^{ij} + 4|D_{ai}|^2 \quad (2.66)$$

$$\text{Tr } \left( m^{(j=0)} \right)^2 = 2F^{ij} \bar{F}_{ij} + 2D_a^i D_{ai} + 2D_a D_{ai}^i \quad (2.67)$$

$$\text{Tr } \left( m^{(j=1)} \right)^2 = 2D_{ai} D_a^i. \quad (2.68)$$

For convenience, we may define the **supertrace**, a sum of the squared mass matrices weighted by the number of states,

$$\text{STr} \left( m^{(j)} \right)^2 \equiv \sum_j \text{Tr} (2j+1)(-)^{2j} m^2 \quad (2.69)$$

$$= -2F\bar{F} - u|D_{ai}|^2 + 2F\bar{F} + 2D_a^i D_{ai} + 2D_a D_{ai}^i + 3 \cdot 2D_{ai} D_a^i \quad (2.70)$$

$$= 2D_a (D_a)_i^i \quad (2.71)$$

$$= 2D_{\tilde{a}} \sum_i q_i^{(\tilde{a})} \quad (2.72)$$

Note that  $\langle D_a \rangle \neq 0$  only for U(1) factors, so  $(D_a)_i^i = \sum q_i$ , the sum of all U(1) charges. We have written  $\tilde{a}$  to index only the U(1) factors of the gauge group. Note, however, that usually

$$\sum_i q_i^{(\tilde{a})} = 0 \quad (2.73)$$

due to anomaly cancellation. This leads to the very stringent constraint that

$$\text{STr} m^2 = 0. \quad (2.74)$$

Note that this is a tree-level result that assumes renormalizable interactions<sup>4</sup>.

## 2.11 Soft breaking and the MSSM

The sum rule (2.74) is a road block to SUSY model building. To see why, consider the scalar mass matrix (2.63) applied to squarks. In order to preserve  $SU(3)_c$ , the squarks should not obtain a VEV. This implies that the  $D$ -terms vanish,  $D_a^i = D_{\text{color}} = 0$ , for squarks. Thus further means that quarks only get their masses from the superpotential.

Similarly preserving  $U(1)_{\text{EM}}$  implies that the  $D$ -terms corresponding to the electrically charged  $SU(2)_L$  directions should also vanish:  $D_{\pm} = D_{1,2} = 0$ . This means that the only  $D$ -terms which are allowed to be non-trivial are  $D_3$  and  $D_Y$ , corresponding to the third generator of  $SU(2)_L$  and hypercharge. The scalar mass matrix for the up-type quarks is then

$$m_{2/3}^2 = \begin{pmatrix} m_{2/3} m_{2/3}^\dagger + \left( \frac{1}{2} g D_3 + \frac{1}{6} g' D_Y \right) \mathbb{1} & \Delta \\ \Delta^\dagger & m_{2/3} m_{2/3}^\dagger - \frac{2}{3} g' D_Y \mathbb{1} \end{pmatrix} \quad (2.75)$$

$$m_{1/3}^2 = \begin{pmatrix} m_{1/3} m_{1/3}^\dagger + \left( -\frac{1}{2} g D_3 + \frac{1}{6} g' D_Y \right) \mathbb{1} & \Delta' \\ \Delta'^\dagger & m_{1/3} m_{1/3}^\dagger + \frac{1}{3} g' D_Y \mathbb{1} \end{pmatrix}, \quad (2.76)$$

where the  $\Delta$  and  $\Delta'$  are the appropriate expressions from (2.63) and  $m_{2/3,1/3}$  correspond to the quadratic terms in the superpotential that contribute to the quark masses.

Charge conservation requires the sum of  $D$  terms to vanish, so that at least one  $D$  term is less than or equal to zero. For example, suppose that

$$\frac{1}{2} g D_3 + \frac{1}{6} g' D_Y \leq 0. \quad (2.77)$$

Let  $\beta$  be the direction in field space corresponding to the up quark. Then  $\beta$  is an eigenvector of the quark mass matrix  $m_{2/3}$  with eigenvalue  $m_u$ . Then (2.77) implies that

$$(\beta^\dagger \ 0) m_{2/3}^2 \begin{pmatrix} \beta \\ 0 \end{pmatrix} \leq m_u^2. \quad (2.78)$$

<sup>4</sup>Non-renormalizable terms in the Kähler potential, for example, modify how the superpotential terms contribute to the scalar potential since one has to rescale fields for them to be canonically normalized.

This implies that there exists a squark in the spectrum that has a tree-level mass less than the up quark. Such an object would have been discovered long ago and is ruled out.

More generally, the observation that there is at least one negative  $D$ -term combined with the form of the squark matrices (2.75) and (2.76) implies that there must exist a squark with mass less than or equal to either  $m_u$  or  $m_d$ . Thus even if SUSY is broken, it appears that any supersymmetric version of the Standard Model is doomed to be ruled out at tree level.

In order to get around this restriction, one typically breaks SUSY in a separate **supersymmetry breaking sector** (SUSY) that is not charged under the Standard Model gauge group. This SUSY sector still obeys a sum rule of the form (2.74) but the spectrum is no longer constrained by observed SM particles. In order for the SUSY sector to lend masses to the SM superpartners, one assumes the existence of a **messenger sector** which interacts with both the SM and the SUSY sectors. The messenger sector transmits the SUSY-breaking auxiliary field VEV to the SM sector and allows the SM superpartners to become massive without violating the sum rule (2.74). Note that this also allows a large degree of agnosticism about the details of the SUSY sector—as far as the phenomenology of the MSSM is concerned, we only need to know about the SUSY scale and the properties of the messenger sector.

There are two standard types of assumptions for the messenger sector depending on how one assumes it couples to the SM:

- **Gravity mediation:** here one assumes that the SM and SUSY breaking sectors only communicate gravitationally. The details of these interactions fall under the theory of local supersymmetry, or **supergravity** (SUGRA), but are typically not necessary for collider phenomenology.
- **Gauge mediation:** The messenger sector contains fields which are charged under the SM gauge group.

An alternative way around the SUSY sum rule is to construct a ‘single sector’ model based on strong coupling [13, 14]. These turn out to be dual to 5D models of SUSY breaking using tools that we introduce in Section 3.

Often we are only interested in the properties of the Standard Model particles and their superpartners. We can ‘integrate out’ the details of the messenger sector and parameterize SUSY breaking into non-renormalizable interactions. As an example, suppose that a superfield,  $X$ , breaks supersymmetry by picking up an  $F$ -term VEV:  $\langle X \rangle = \dots + \langle F \rangle \theta^2$ .  $X$  may also have a scalar VEV, but this does not break SUSY. We then parameterize the types of non-renormalizable couplings that are generated when we integrate out the messenger sector. We have four types of terms:

1. **Non-holomorphic scalar masses** are generated by higher order Kähler potential terms such as

$$\int d^4\theta \frac{X^\dagger X}{M^2} \Phi^\dagger \Phi = \left( \frac{F}{M} \right)^2 \varphi^* \varphi + (\text{SUSY preserving terms}). \quad (2.79)$$

$$\int d^4\theta \frac{X + X^\dagger}{M} \Phi^\dagger \Phi = \left( \frac{F^*}{M} \right) \int d^2\theta \Phi^\dagger \Phi + \text{h.c.} + (\text{SUSY preserving terms}). \quad (2.80)$$

We have written the SUSY-breaking part of (2.80) suggestively to appear as a non-holomorphic superpotential term. Since  $\Phi^\dagger$  only contains  $\bar{\theta}$ s and not  $\theta$ ,  $\int d^2\theta \Phi^\dagger \Phi = \varphi^* F_\varphi = \varphi^* W'[\varphi^*]$ . For renormalizable superpotentials, this can give an  $A$ -term of the form (2.82) or a  $b$ -term of the form (2.81) below; the latter with a slightly different scaling with  $F$ .

The mass scale  $M$  is required by dimensional analysis and is naturally the scale of the mediator sector that has been integrated out. For gravity mediation  $M \sim M_{\text{Pl}}$  while for gauge mediation  $M \sim M_{\text{mess}}$ , the mass of the messenger fields. Doing the Grassmann integral and picking the terms that depend on the SUSY breaking order parameter  $F$  gives a mass  $m^2 = (F/M)^2$  to the scalar  $\varphi$ . Note that  $F$  has dimension 2 so that this term has the correct mass dimension.



2. **Holomorphic scalar masses** are generated by a similar higher order Kähler potential term,

$$\int d^4\theta \frac{X^\dagger X}{M^2} \left[ \Phi^2 + (\Phi^\dagger)^2 \right] = \left( \frac{F}{M} \right)^2 (\varphi^2 + \varphi^{*2}) + (\text{SUSY preserving terms}). \quad (2.81)$$

These are often called  $b$ -terms. One may want to instead write these masses at lower order in  $F$  by writing a superpotential term  $W \supset X\Phi^2$ . This, however, is a renormalizable interaction that does not separate the SUSY sector from the visible sector—as one can see the mediator mass does not appear explicitly in such a term. Thus  $W \supset X\Phi^2$  is subject to the SUSY sum rule and is not the type of soft term we want for the MSSM.

3. **Holomorphic cubic scalar interactions** are generated from the superpotential,

$$\int d^2\theta \frac{X}{M} \Phi^3 + \text{h.c.} = \frac{F}{M} (\varphi^3 + \varphi^{*3}) + (\text{SUSY preserving terms}). \quad (2.82)$$

These are called  $A$ -terms and are the same order as the scalar mass.

4. **Gaugino masses** are generated from corrections to the gauge kinetic term,

$$\int d^2\theta \frac{X}{M} \mathcal{W}_\alpha \mathcal{W}^\alpha + \text{h.c.} = \frac{F}{M} \lambda\lambda + \text{h.c.} + (\text{SUSY preserving terms}). \quad (2.83)$$

This is a gaugino mass on the same order as the scalar mass and the  $A$ -term.

In principle one could also generate tadpole terms for visible sector fields, but we shall ignore this case and assume that all field are expanded about their minimum. These four types of terms are known as **soft supersymmetry breaking** terms. The key point is that these do not reintroduce any quadratic UV sensitivity in the masses of any scalars. This is clear since above the SUSY breaking mediation scale  $M$ , the theory is supersymmetric and these divergences cancel.

It is common to simply parameterize the soft breaking terms of the MSSM in the Lagrangian:

$$\mathcal{L}_{\text{soft}} = -\frac{1}{2} \left( M_3 \tilde{g}\tilde{g} + M_2 \tilde{W}\tilde{W} + M_1 \tilde{B}\tilde{B} \right) + \text{h.c.} \quad (2.84)$$

$$- \left( a_u \tilde{Q} H_u \tilde{u} + a_d \tilde{Q} H_d \tilde{d} + a_e \tilde{L} H_d \tilde{e} \right) + \text{h.c.} \quad (2.85)$$

$$- \tilde{Q}^\dagger m_Q^2 \tilde{Q} - \tilde{L}^\dagger m_L^2 \tilde{L} - \tilde{u}^\dagger m_u^2 \tilde{u} - \tilde{d}^\dagger m_d^2 \tilde{d} - \tilde{e}^\dagger m_e^2 \tilde{e} - m_{H_u}^2 H_u^* H_u - m_{H_d}^2 H_d^* H_d \quad (2.86)$$

$$- (b H_u H_d + \text{h.c.}). \quad (2.87)$$

This is simply a reparameterization of the types of soft terms described in (2.79 – 2.83), from which one can read off the scaling of each coefficient with respect to  $F/M$ .

Note that the trilinear soft terms,  $a_{u,d,e}$ , and the soft masses  $m_{Q,L,u,d,e}^2$  are  $3 \times 3$  matrices in flavor space. The trilinear terms are in a one-to-one correspondence with the Yukawa matrices except that they represent a coupling between three scalars. In general, the soft masses cause the squarks and sleptons to have different mass eigenstates than the SM fermions.

Phenomenologically, we assume that

$$M_{1,2,3}, a_{u,d,e} \sim m_{\text{SUSY}} \quad (2.88)$$

$$m_{Q,u,d,L,e,H_u,H_d}^2, b \sim m_{\text{SUSY}}^2, \quad (2.89)$$

where  $m_{\text{SUSY}}$  is between a few hundreds of GeV to a TeV. This is the range in which generic MSSM-like models provide a solution to the Hierarchy problem.

***R*-symmetry, gauginos, supersymmetry breaking.** Recall that when an *R*-symmetry exists, the different components of a superfield carry different *R* charges. Because the  $\mathcal{O}(\theta)$  component of  $\mathcal{W}^\alpha$ ,  $F_{\mu\nu}$ , is real, it cannot carry an *R* charge. This means that the lowest component, the gaugino  $\lambda$ , must have non-zero *R*-charge. Further, the gaugino mass term (2.83) breaks this symmetry. One will find that *R*-symmetry plays an important role in many non-perturbative results in SUSY. Two important results related to SUSY breaking and gaugino masses are [15, 16].

## 2.12 Electroweak symmetry breaking in the MSSM

The most important feature of the Standard Model is electroweak symmetry breaking. Recall that this is due to a tachyonic Higgs mass at the origin being balanced by a positive quartic coupling leading to a non-zero vacuum expectation value. In the MSSM we have two Higgs doublets,

$$H_u = \begin{pmatrix} H_u^+ \\ H_u^0 \end{pmatrix} \quad H_d = \begin{pmatrix} H_d^0 \\ H_d^- \end{pmatrix}. \quad (2.90)$$

We have already seen that supersymmetry relates the scalar quartic coupling to the other couplings of the theory. This then constrains the expected Higgs boson mass.

To preserve  $SU(3)_c$  and  $U(1)_{EM}$  we assume that no squarks or sleptons pick up VEVs. Then the quartic terms in the Higgs potential come from *D*-terms, (2.56):

$$\begin{aligned} V_D &= \frac{g^2}{4} \left( H_u^\dagger \sigma^a H_u + H_d^\dagger \sigma^a H_d \right) \left( H_u^\dagger \sigma^a H_u + H_d^\dagger \sigma^a H_d \right) + \frac{g'^2}{4} (|H_u|^2 - |H_d|^2)^2 \\ &= \frac{1}{2} g^2 |H_u^\dagger H_d|^2 + \frac{1}{8} (g^2 + g'^2) (|H_u|^2 - |H_d|^2)^2, \end{aligned} \quad (2.91)$$

where we have simplified the  $SU(2)_L$  terms using the relation  $\sigma_{ij}^a \sigma_{kl}^a = 2\delta_{il}\delta_{jk} - \delta_{ij}\delta_{kl}$ . We see immediately that the Higgs quartic  $\lambda$  coupling goes like the squared electroweak couplings,  $g^2$  and  $g'^2$ . This connection between the Higgs sector and the gauge parameters does not exist in the Standard Model

In addition to the *D*-term contribution, there is also the supersymmetric *F*-term contribution coming from the  $\mu$ -term in the superpotential. The quadratic contributions to the Higgs potential are,

$$V_F = |\mu|^2 |H_u|^2 + |\mu|^2 |H_d|^2 + \dots \quad (2.92)$$

We have dropped terms proportional to the Yukawa couplings since we assume the scalar partners of the SM fermions do not acquire VEVs. On top of this, there are the soft supersymmetry breaking terms. These include soft masses for each Higgs doublet and a ‘holomorphic’ *b*-term which is called  $B_\mu$  (or sometimes  $B\mu$ ),

$$V_{\text{soft}} = m_{H_u}^2 |H_u|^2 + m_{H_d}^2 |H_d|^2 + (B_\mu H_u \cdot H_d + \text{h.c.}). \quad (2.93)$$

Note that the contraction of  $H_u$  and  $H_d$  in the *D*-term (2.91) is different from that in the  $B_\mu$  term (2.95). Specifically,  $H_u \cdot H_d$  is contracted using the  $\epsilon_{ab}$  tensor and gives  $H_u^+ H_d^- - H_u^0 H_d^0$ . Further, the *D*-term couplings are real since they are part of a real superfield. The *F*-term couplings are made real because they are the modulus of a complex parameter. The couplings of the soft terms, on the other hand, carry arbitrary sign and phase.

Combining all of these factors, the full Higgs potential is

$$V_H = V_D + V_F + V_{\text{soft}} \quad (2.94)$$

$$\begin{aligned} &= \frac{1}{2} g^2 |H_u^\dagger H_d|^2 + \frac{1}{8} (g^2 + g'^2) (|H_u|^2 - |H_d|^2)^2 \\ &\quad + (|\mu|^2 + m_{H_u}^2) |H_u|^2 + (|\mu|^2 + m_{H_d}^2) |H_d|^2 + (B_\mu H_u \cdot H_d + \text{h.c.}). \end{aligned} \quad (2.95)$$

To simplify this, we can assume that the charged components of the doublets pick up no VEV and write everything in terms of only the neutral components (we address the validity of this assumption below):

$$V_H = \frac{1}{8}(g^2 + g'^2) (|H_u^0|^2 - |H_d^0|^2)^2 + \sum_{i=u,d} (|\mu|^2 + m_{H_i}^2) |H_i^0|^2 - 2B_\mu \text{Re}(H_u^0 H_d^0). \quad (2.96)$$

Observe that this potential has a direction in field space,  $|H_u^0|^2 = |H_d^0|^2$  where the  $D$ -term quartic vanishes. This is called a **D-flat** direction and requires caution. In order to break electroweak symmetry, we must destabilize the origin of field space with a tachyonic mass term to force a linear combination of the neutral Higgses to pick up a VEV. In the SM destabilization is balanced by the quartic coupling which forces the VEV to take a finite value. We see now in the MSSM that one has to take special care to make sure that the destabilized direction does *not* align with the  $D$ -flat direction or else the potential isn't bounded from below. In other words, we must impose a negative mass squared in one direction in the Higgs moduli space while making sure that there is a positive definite mass squared along the  $D$ -flat direction. This can be written as two conditions:

1. We require exactly one negative eigenvalue in the neutral Higgs mass matrix,

$$\begin{vmatrix} |\mu|^2 + m_{H_u}^2 & -B_\mu \\ -B_\mu & |\mu|^2 + m_{H_d}^2 \end{vmatrix} = (|\mu|^2 + m_{H_u}^2) (|\mu|^2 + m_{H_d}^2) - B_\mu^2 < 0. \quad (2.97)$$

2. The mass squared term is positive when  $|H_u^0| = |H_d^0|$ . For simplicity, suppose  $B_\mu$ ,  $\langle H_u^0 \rangle$ , and  $\langle H_d^0 \rangle$  are all real (see below). Then this imposes

$$(|\mu|^2 + m_{H_u}^2) + (|\mu|^2 + m_{H_d}^2) + 2B_\mu > 0. \quad (2.98)$$

The conditions (2.97) and (2.98) are the requirements for electroweak symmetry breaking in the MSSM.

Note that a natural choice for the soft masses,  $m_{H_u}^2 = m_{H_d}^2$ , does not obey the restrictions (2.97) and (2.98). One way to nevertheless enforce this relation is to impose it as a boundary condition at some high scale and allow the renormalization group flow to differentiate between them. This is actually quite reasonable, since the  $\beta$ -function for these soft masses include terms that go like the squared Yukawa coupling. The two soft masses flow differently due to the large difference in the top and bottom Yukawas. In fact, the up-type Higgs mass parameter shrinks in the IR and it is natural to assume

$$m_{H_u}^2 < m_{H_d}^2. \quad (2.99)$$

A convenient choice is  $m_{H_u}^2 < 0$  and  $m_{H_d}^2 > 0$ . In this way the MSSM naturally admits **radiative electroweak symmetry breaking** where the tachyonic direction at the origin is generated by quantum effects.

Since there are many parameters floating around, it use useful to summarize that the following all prefer electroweak symmetry breaking and no runaway directions:

- Relatively large  $B_\mu$
- Relatively small  $\mu$
- Negative  $m_{H_u}^2$ .

Be aware that these are only rough guidelines and are neither necessary nor sufficient.

It is standard to parameterize the VEVs of the two Higgses relative to the SM Higgs VEV by introducing an angle,  $\beta$ ,

$$\langle H_u^0 \rangle = \frac{v_u}{\sqrt{2}} = \frac{v}{\sqrt{2}} \sin \beta \quad \langle H_d^0 \rangle = \frac{v_d}{\sqrt{2}} = \frac{v}{\sqrt{2}} \cos \beta. \quad (2.100)$$

Minimizing the potential,  $\partial V/\partial H_u^0 = \partial V/\partial H_d^0 = 0$ , one obtains

$$\sin 2\beta = \frac{2B_\mu}{2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2} \quad (2.101)$$

$$\frac{M_Z^2}{2} = -|\mu|^2 + \frac{m_{H_d}^2 - m_{H_u}^2 \tan^2 \beta}{\tan^2 \beta - 1}. \quad (2.102)$$

The second relation is especially strange: it connects the supersymmetric  $\mu$  term to the soft-breaking masses, even though these come from totally different sectors of the theory. In other words, unlike the quartic and gauge couplings which are tied together by supersymmetry, these parameters have no reason to have any particular relation with each other. Further,  $M_Z^2$  is experimentally measured and much smaller than the typical expectation for either  $\mu$  or  $m_{H_{u,d}}^2$ , so it appears that there's some cancellation going on.

The Higgs sector of the MSSM contains the usual CP-even Higgs  $h$ , a heavier CP-even Higgs, the Goldstones of electroweak symmetry breaking, an additional pair of charged Higgses  $H^\pm$ , and a CP-odd Higgs  $A$ . With a little work, one can show that the CP-even Higgs masses are

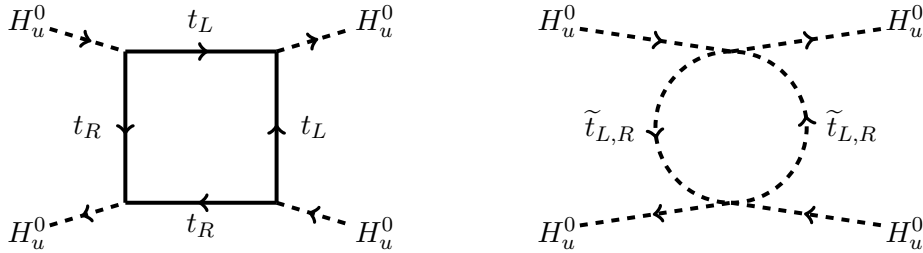
$$m_h^2 = \frac{1}{2} \left[ M_Z^2 + m_A^2 \pm \sqrt{(M_Z^2 + m_A^2)^2 - 4m_A^2 M_Z^2 \cos^2 2\beta} \right], \quad (2.103)$$

where  $m_A^2 = B_\mu/(\sin \beta \cos \beta)$ . One can further show that this is bounded from above,

$$m_h \leq M_Z |\cos 2\beta| \leq M_Z. \quad (2.104)$$

Of course, we now know that  $m_h \approx 125$  GeV. In fact, even before the LHC it was known from LEP that  $m_h \gtrsim 114$  GeV. While at first glance (2.104) appears to be ruled out experimentally, this is only a tree-level bound. What this is really saying is that one requires large corrections to the quartic self-coupling to pull up the Higgs mass from its tree level value. Due to the size of  $y_t$ , the main effect comes from top and stop loops.

To maximize the quartic coupling, we are pushed towards large values of  $\tan \beta$  since this would put most of the Higgs VEV in  $H_u$  and would make the light Higgs be primarily composed of  $H_u$ . Examining the  $H_u^4$  coupling at loop level, consider diagrams of the form:



Assuming negligible  $A$  terms, the result is

$$\lambda(m_t) = \lambda_{\text{SUSY}} + \frac{2N_c y_t^4}{16\pi^2} \ln \left( \frac{m_{\tilde{t}_1} m_{\tilde{t}_2}}{m_t^2} \right), \quad (2.105)$$

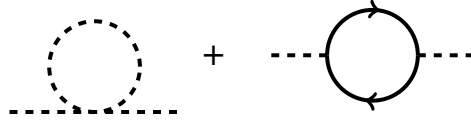
where  $\lambda_{\text{SUSY}}$  comes from the  $D$ -term potential and  $N_c$  is the number of colors. This equation tells us that in order to push the Higgs mass above the tree-level bound of  $M_Z$ , one must increase  $m_{\tilde{t}}$ . The correction is

$$\Delta m_h^2 = \frac{3}{4\pi^2} v^2 y_t^4 \sin^2 \beta \ln \left( \frac{m_{\tilde{t}_1} m_{\tilde{t}_2}}{m_t^2} \right). \quad (2.106)$$

For further details, we refer to the treatment in [17] or the encyclopedic reference [18].

### 2.13 The little hierarchy problem of the MSSM

It has been well known since LEP that in order to push  $m_h > 114$  GeV in the MSSM, one requires large stop masses,  $m_{\tilde{t}} \sim 1 - 1.4$  TeV. Pushing the stop mass this heavy comes at a cost, unfortunately. The stops contribute not only to the Higgs quartic—which we need to push the Higgs mass up—but also to the soft mass  $m_{H_u}^2$  from loops of the form



The larger one sets  $m_{\tilde{t}}$ , the larger the shift in  $m_{H_u}^2$ . Recall, however, the strange cancellation we noted in (2.102). This equation seems to want  $m_{H_u}^2 \sim M_Z^2/2$ . The loop corrections above contribute a shift of the form

$$\Delta m_{H_u}^2 = \frac{3y_t^2}{4\pi^2} m_{\tilde{t}}^2 \ln \left( \frac{\Lambda_{UV}}{m_{\tilde{t}}} \right). \quad (2.107)$$

For  $m_{\tilde{t}} = 1.2$  TeV and  $\Lambda_{UV} = 10^{16}$  this balancing act between  $m_{H_u}^2$  and  $M_Z^2/2$  requires a fine tuning of

$$\frac{M_Z^2/2}{\Delta m_{H_u}^2} \sim 0.1\%.$$

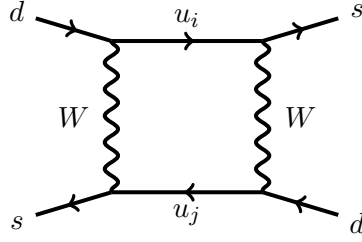
Physically what's happening is that the stop plays a key role in naturalness by canceling the sensitivity to the UV scale. By pushing the stop to be heavier to increase the Higgs quartic, one reintroduces quadratic sensitivity up to the scale of the stop mass. This is known as the **little hierarchy problem** of the MSSM.

### 2.14 SUSY breaking versus flavor

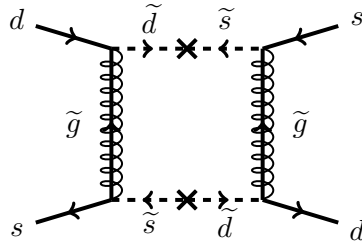
The soft breaking Lagrangian introduces many new masses, phases, and mixing angles on top of those found in the Standard Model for a total of 124 parameters [19]. Most of this huge parameter space, however, is already excluded from flavor and CP violating processes. Recall that in the SM, there are no tree-level flavor-changing neutral currents (FCNC) and loop-level contributions are suppressed by the GIM mechanism. Lepton number violation is similarly strongly suppressed. In the limit where the Yukawa couplings vanish,  $y \rightarrow 0$ , the Standard Model has a  $U(3)^5$  flavor symmetry where each of the five types of matter particles are equivalent. This flavor symmetry is presumably broken at some scale  $\Lambda_F$  in such a way that the only imprint of this UV physics at scales well below  $\Lambda_F$  are the Yukawa matrices. This flavor scale can be very large so that effects of this flavor breaking go like  $1/\Lambda_F$  and are plausibly very small.

In the MSSM, one must further check that the flavor breaking dynamics has already ‘frozen out’ at the SUSY breaking scale so that the only non-trivial flavor structure in the SUSY breaking parameters are the Yukawa matrices themselves. This means we would like the mediator scale  $M$  to be below the flavor scale,  $M \ll \Lambda_F$ . In gravity mediation, however,  $\Lambda_{med} = M_{Pl}$ , and we can no longer guarantee that the SUSY breaking mediators are insulated from flavor violating dynamics. This leads to strong constraints on the flavor structure of the MSSM soft parameters.

For example, consider one of the most carefully studied FCNC processes, kaon anti-kaon ( $K-\bar{K}$ ) mixing. The quark content of the mesons are  $K = d\bar{s}$  and  $\bar{K} = \bar{d}s$ . In the SM this process is mediated by diagrams such as



Each vertex picks up a factor of the CKM matrix. The GIM observation is the fact that the unitarity of the CKM matrix imposes an additional suppression. In the MSSM, on the other hand, the squark soft masses introduce an additional source of flavor violation so that the quark and squark mass matrices are misaligned. This manifests itself as flavor-changing mass insertions,  $\Delta m_{ds}^2$ , on squark propagators when written in terms of the Standard Model mass eigenstate combinations:



Note that rather than  $W$  bosons, this diagram is mediated by gluinos which carry much stronger coupling constants  $\alpha_3 \gg \alpha_2$ . Further, since there are no factors of  $V_{\text{CKM}}$ , there is no GIM suppression. The loop integral goes like  $d^4k/k^{10} \sim 1/m_{\text{SUSY}}^6$ . Thus we can estimate this contribution to kaon mixing to be

$$\mathcal{M}_{KK}^{\text{MSSM}} \sim \alpha_3^3 \left( \frac{\Delta m_{ds}^2}{m_{\text{SUSY}}^2} \right)^2 \frac{1}{m_{\text{SUSY}}^2}. \quad (2.108)$$

Comparing this to the experimental bound,

$$\frac{\Delta m_{ds}^2}{m_{\text{SUSY}}^2} \lesssim 4 \cdot 10^{-3} \left( \frac{m_{\text{SUSY}}}{500 \text{ GeV}} \right). \quad (2.109)$$

There are similar constraints on CP violating and lepton number violating processes (e.g. dipole moments and  $\mu \rightarrow e\gamma$ ). This is the **SUSY flavor problem**: a generic flavor structure for the MSSM soft parameters is phenomenologically ruled out. We are led to conclude that the off-diagonal flavor terms must be strongly suppressed to avoid experimental bounds.

One way to do this is to suppose an organizing principle in the SUSY breaking parameters, **soft-breaking universality**,

1. Soft breaking masses are all universal for all particles at some high scale. This means that  $m_Q^2 \propto \mathbb{1}$  in flavor space, and similarly for each MSSM matter multiplet.
2. If  $a$ -terms are not flavor-universal, then the Higgs VEV induces similar problematic mixings,

$$\mathcal{L}_a = a_{ij}^u Q_i \bar{U}_j H_u + a_{ij}^d Q_i \bar{D}_j H_d + a_{ij}^e L_i \bar{E}_j H_d. \quad (2.110)$$

To avoid this, assume that  $a_{ij}^I$  is proportional to the Yukawa matrix,

$$a_{ij}^I = A^I y_{ij}^I. \quad (2.111)$$

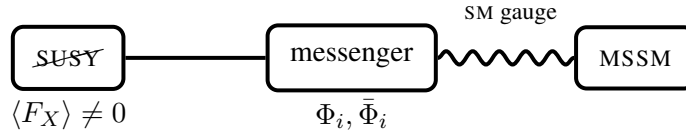
This way, the rotation that diagonalizes the SM fermions also diagonalizes their scalar partners.

3. To avoid CP violation, assume that all non-trivial phases beyond those in the Standard Model CKM matrix vanish.

These are phenomenological principles. Ultimately, one would like to explain why these properties should be true (or at least approximately so).

### 2.15 Gauge mediated SUSY breaking

One straightforward realization of soft-breaking universality is to have the messenger sector be flavor universal. A natural way to do this is **gauge mediation** since the SM gauge fields are blind to flavor [20–23]. See [24] for a review.



The main idea is that the SUSY breaking sector has some superfield (or collection of superfields)  $X$  which pick up  $F$ -term VEVs,  $\langle F_X \rangle \neq 0$ . This generates mass splittings in the messenger sector superfields,  $\Phi_i$  and  $\bar{\Phi}_i$ . These messengers obey the tree-level SUSY sum rules discussed above but are not problematic since all of the components can be made heavy. One then assumes that the messengers are charged under the SM gauge group so that the MSSM superfields will feel the effects of SUSY breaking through loops that include the messenger fields. Note that anomaly cancellation of the SM gauge group typically requires the messenger superfields to appear in vector-like pairs,  $\Phi$  and  $\bar{\Phi}$  with opposite SM quantum numbers.

The messenger fields generate non-renormalizable operators that connect the MSSM and the SUSY breaking sector without introducing any flavor dependence for the soft masses. Further, because the messenger scale is adjustable, one can always stay in regime where it is parametrically smaller than the flavor scale  $M \ll \Lambda_F$ . Recall the estimates in Section 2.11 for the size of the MSSM soft terms. For gauge mediation,  $M$  is the mass of the messenger sector fields  $\Phi_i$  and  $\bar{\Phi}_i$  and  $F$  is the SUSY breaking VEV,  $F_X$ . Below  $M$  we integrate out the messengers to generate the MSSM soft parameters.

The simplest realization of this is **minimal gauge mediation**. Here one assumes only one SUSY breaking field  $X$  and  $N_m$  mediators,  $\Phi_i$  and  $\bar{\Phi}_i$ , in the fundamental representation of an  $SU(5)$  GUT. The superpotential coupling between these sectors is

$$W = \bar{\Phi} X \Phi. \quad (2.112)$$

The contribution to the potential is

$$\left| \frac{\partial W}{\partial \Phi} \right|^2 = |\langle X \rangle|^2 |\varphi|^2 + |\langle X \rangle|^2 |\bar{\varphi}|^2 + \varphi \bar{\varphi} \langle F_X \rangle \quad (2.113)$$

The messenger masses are

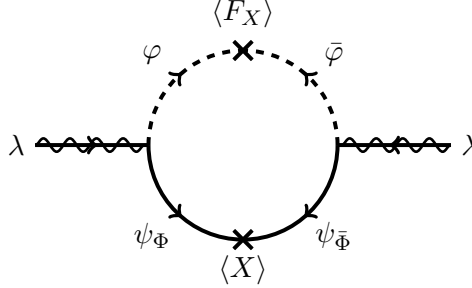
$$m_\psi = X \quad (2.114)$$

$$m_\varphi^2 = X^2 \pm F_X, \quad (2.115)$$

using the notation where the angle brackets  $\langle \dots \rangle$  are dropped when it is clear that we are referring to the VEV of a field. Observe that the messenger scale is set by the lowest component VEV of the SUSY

breaking parameter,  $M = X$ . In what follows we make the typical assumption that  $F/M^2 \ll 1$ . Note that these masses satisfy the SUSY sum rule.

Now let's consider the spectrum arising from this simple set up. The gauginos of the SM gauge group pick up a mass contribution from diagrams of the form



The  $\langle X \rangle$  insertion on the  $\psi_\Phi$  line is required to flip the gaugino helicity (recall that arrows on fermion indicate helicity). The  $F$  insertion on the  $\varphi$  line is required to connect to SUSY breaking so that this is indeed a mass contribution that is not accessible to the gauge boson. The  $F$  VEV is also required to flip from a  $\varphi$  to a  $\bar{\varphi}$  so that the scalar of the chiral superfield picks up a sense of chirality as well. Using powerful methods based on holomorphy [25, 26], the gaugino mass for the  $i^{\text{th}}$  gauge factor is

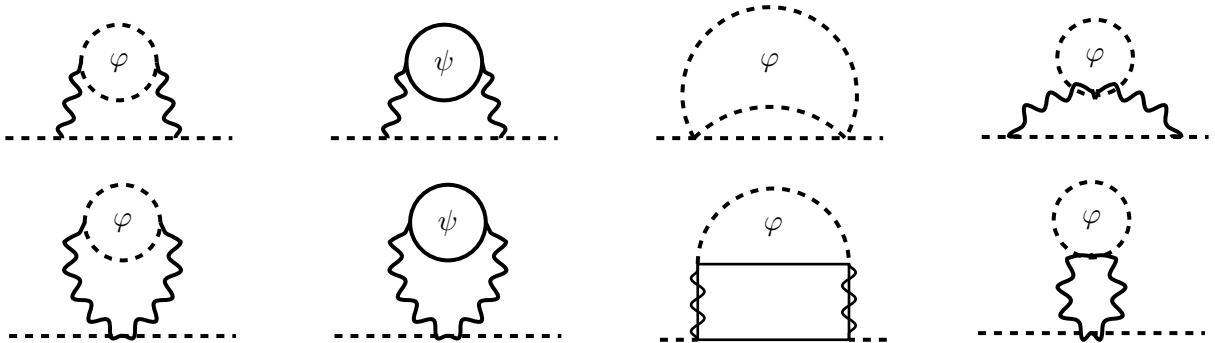
$$M_{\lambda_i} = \frac{FM}{M^2} \frac{g_i^2}{16\pi^2} N_m = \frac{\alpha_i}{4\pi} N_m \frac{F}{M}. \tag{2.116}$$

This expression—which one could have guessed from a back-of-the-envelope estimate—turns out to be exact to leading order in  $F/M^2$ . This is a reflection of the powerful renormalization theorems in supersymmetry, see e.g. [?]. One of the concrete predictions of minimal gauge mediation is the relation

$$M_{\lambda_1} : M_{\lambda_2} : M_{\lambda_3} = \alpha_1 : \alpha_2 : \alpha_3. \tag{2.117}$$

The heaviest superpartners are those which couple to the largest rank gauge group.

The scalar partners of the SM matter particles do not directly couple to the messengers. Thus the masses for the squarks and sleptons must be generated at two loop level. There are many diagrams that include loops of both the messenger scalar and fermion:



The loops either include a gauge boson or otherwise use the scalar quartic  $D$ -term interaction between messengers and sfermions. The result is that the soft scalar masses go like

$$m_{\text{soft}}^2 \sim \left( \frac{g^2}{16\pi^2} \right)^2 N_m \frac{F^2}{M^2} C_i, \tag{2.118}$$



where  $C_i$  is the relevant quadratic casimir. Observe that  $m_{\text{soft}}^2 \sim m_\lambda^2$  so that the sfermions which couple to the higher rank gauge factors pick up more mass. Including the various gauge charges and taking the limit  $\alpha_3 \gg \alpha_2 \gg \alpha_1$  gives a prediction for the sfermion spectrum in minimal gauge mediation,

$$m_q^2 : m_\ell^2 : m_E^2 = \frac{4}{3}\alpha_3^2 : \frac{3}{4}\alpha_2^2 : \frac{3}{5}\alpha_1^2. \quad (2.119)$$

Note that (2.117) and (2.119) are only predictions of minimal gauge mediation. A parameterization of the soft terms from a generic gauge mediation model is presented in [27,28] under the banner of **general gauge mediation**. Requiring that the superpartner masses are around the electroweak scale sets

$$\frac{F}{M} \sim 100 \text{ TeV}. \quad (2.120)$$

Note that since the messengers interact with the SM superfields only through gauge interactions, the holomorphic soft terms ( $A$  and  $B$  terms) are typically very small in gauge mediation.

One important phenomenological consequence of gauge mediation is that the lightest supersymmetric partner (LSP) is not one of the MSSM fields but rather the gravitino whose mass is [24],

$$m_{3/2} \sim \frac{F}{\sqrt{3}M_{\text{Pl}}} \sim \left( \frac{\sqrt{F}}{100 \text{ TeV}} \right)^2 2.4 \text{ eV}. \quad (2.121)$$

Thus the gravitino is much lighter than the electroweak scale, but is also similarly weakly coupled. The relevant couplings at low energies are not gravitational, but rather through the Goldstino component of the gravitino. This coupling is proportional to the SUSY breaking VEV  $F$ . Because of  $R$ -parity, any supersymmetric partner produced in the MSSM will eventually decay into the next-to-lightest superpartner (NLSP). This NLSP must eventually decay into the gravitino LSP since it is the only decay mode available. When  $\sqrt{F} \gtrsim 10^6 \text{ GeV}$ , the NLSP is so long lived that on collider scales it behaves effectively like the LSP. On the other hand, if  $\sqrt{F} \lesssim 10^6 \text{ GeV}$ , the NLSP decays within the detector. This gives a fairly unique signal with displaced photons and missing energy if the NLSP is the bino,  $\tilde{B}$ .

## 2.16 The $\mu$ - $B_\mu$ problem of gauge mediation

Let's return to an issue we addressed earlier when discussing electroweak symmetry breaking. We wrote two relations (2.101 – 2.102) satisfied at the minimum of the Higgs potential. We noted the  $\mu$ -problem associated with (2.102):  $\mu$  and  $m_{H_{u,d}}^2$  seem to come from different sectors of the theory but must conspire to be roughly the same scale. In principle, since  $\mu$  is a supersymmetric dimensionful parameter (the only one in the MSSM), it could take a value on the order of the Planck mass. We now present a solution to the  $\mu$ -problem, but we shall see that this solution will cause problems in gauge mediation due to the second relation, (2.101).

One way to address this  $\mu$ -problem is to forbid it in the supersymmetric limit and then assume that it is generated through the SUSY breaking sector. For example, a global Peccei-Quinn (PQ) symmetry,

$$H_u \rightarrow e^{i\alpha} H_u \quad (2.122)$$

$$H_d \rightarrow e^{i\alpha} H_d, \quad (2.123)$$

prohibits the  $\mu$  term in the superpotential. Gravity, however, is believed to explicitly break global symmetries. Indeed, gravity mediation of SUSY breaking will produce a  $\mu$  term. Consider, for example, the higher order Kähler potential term that couples the SUSY breaking superfield  $X$  to the Higgses [29],

$$\int d^4\theta \frac{X^\dagger H_u \cdot H_d}{M_{\text{Pl}}} + \text{h.c.} \quad (2.124)$$

When  $\langle X \rangle \sim \theta^2 F$ , one generates an effective  $\mu$  term of order  $\mu \sim F/M_{\text{Pl}}$ . This neatly addresses the  $\mu$ -problem and ties the  $\mu$  term to the SUSY breaking masses. The  $B_\mu$  term that is generated comes from

$$\int d^4\theta \frac{X^\dagger X H_u \cdot H_d}{M_{\text{Pl}}^2} \quad (2.125)$$

and thus is of the same order as  $\mu^2$ . This is consistent with the observation in (2.101) that  $B_\mu$ ,  $\mu$ , and the soft breaking terms seem to want to be the same order. We remark that this is no longer true in gauge mediation since  $F \ll 10^{11}$  GeV, the  $\mu$  and  $B_\mu$  terms generated from gravitational breaking are far too small. This must be addressed separately in such theories.

## 2.17 Variations beyond the MSSM

The MSSM is under pressure from the LHC. For a review of the status after Run I of the LHC, see [30]. There are two main issues:

1. The Higgs mass  $m_h = 125$  GeV is hard to achieve in the MSSM since it requires a large radiative correction to the tree level upper bound of  $m_h = M_Z$ .
2. There are no signs of superpartners. With the simplest assumption that  $m_{\tilde{q}} \sim m_{\tilde{g}}$ , the LHC pushes the scale of colored superpartners to be over 1.2 TeV. This appears to no longer be natural.

In this section we present some model-building directions that the LHC data may be suggesting.

### 2.17.1 Additional $D$ -term contributions

One simple direction to increase the tree-level Higgs mass is to add extra  $D$ -terms to increase the Higgs quartic coupling [31–35]. This requires charging the Higgs under an additional  $U(1)_X$  gauge group which one must break above the weak scale. This technique is able to indeed push the tree-level Higgs mass up to the observed value, but one is constrained by changes to Higgs decay branching ratios, particularly  $h \rightarrow b\bar{b}$  [36, 37].

### 2.17.2 The NMSSM

At the cost of adding an additional singlet superfield  $S$  to the MSSM sector, one may solve the  $\mu$  problem and also raise the Higgs mass by enhancing its quartic coupling [38–42]. The Higgs sector superpotential for this “next-to-minimal” supersymmetric SM (NMSSM, see [43, 44] for reviews) is

$$W_{\text{NMSSM}} = y_u H_u Q \bar{U} + y_d H_d Q \bar{D} + y_e H_d \bar{E} + \lambda S H_u H_d + \frac{1}{3} \kappa S^3. \quad (2.126)$$

The  $\kappa$  term breaks the Peccei-Quinn symmetry, (2.122 – 2.123), to a  $\mathbb{Z}_3$ . Since  $S$  is a gauge singlet, the  $D$ -term potential is unchanged from (2.91). Note, however, that there is no longer a  $\mu$  term in the superpotential, instead the  $S H_u H_d$  coupling has taken its place. Thus the  $F$ -term potential differs from that of the MSSM, (2.92), and is instead

$$V_{F,\text{NMSSM}} = \lambda^2 |S|^2 (|H_u|^2 + |H_d|^2) + \lambda^2 |H_u H_d|^2. \quad (2.127)$$

We observe that the combination  $\lambda \langle S \rangle$  plays the role of an effective  $\mu$  term and solves the  $\mu$ -problem. Finally, there are additional soft terms allowed which augment  $V_{\text{soft}}$  in (2.93),

$$\Delta V_{\text{soft,NMSSM}} = m_S^2 |S|^2 + \lambda A_\lambda (S H_u H_d + \text{h.c.}) + \frac{1}{3} \kappa A_\kappa (S^3 + \text{h.c.}). \quad (2.128)$$

The resulting expression for the Higgs mass is approximately

$$m_h^2 \approx M_Z^2 \cos^2 2\beta + \lambda^2 v^2 \sin^2 2\beta - \frac{\lambda^2 v^2}{\kappa^2} (\lambda - \kappa \sin 2\beta)^2 + \frac{3m_t^4}{4\pi^2 v^2} \left[ \ln \left( \frac{m_{\tilde{t}}}{m_t} \right) + \frac{A_t^2}{m_{\tilde{t}}} \left( 1 - \frac{A_t^2}{12m_{\tilde{t}}} \right) \right]. \quad (2.129)$$

This can be larger than the value in the MSSM depending on the value of  $\lambda$ . There are limits on the size of  $\lambda$  coming from perturbativity, but lifting the Higgs mass to 125 GeV is fine. The singlet  $S$  contributes an additional complex scalar to the Higgs sector and an additional neutralino.

### 2.17.3 Natural SUSY

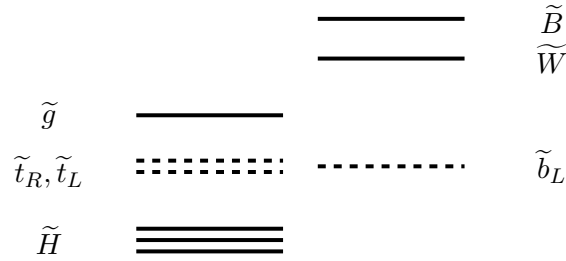
The simplest choices for the MSSM parameters—those that treat all the flavors universally, as preferred by the flavor problem—are tightly constrained by the non-observation of new physics at the LHC. Because the LHC is a proton-proton collider, it is easy for it to produce colored superpartners such as squarks and gluinos. These, in turn, are expected to show up as events with many jets and missing energy as the heavy colored states decay into the LSP. The fact that no significant excesses have been found pushes one to consider other parts of the large MSSM parameter space.

Instead of biasing our parameter preferences by simplicity, one may take a different approach and ask what is the minimal sparticle content required for naturalness? In other words, which superpartners are absolutely required to cancel quadratic divergences? Once these are identified, one may decouple the remaining particles and check the experimental constraints on the resulting spectrum. The ingredients of a ‘minimally’ natural MSSM spectrum are [45, 46] (see [47–50] for a re-examination from the early LHC run)

1. **Light stops.** The largest SM contribution to the Higgs quadratic UV sensitivity is the top loop. Naturalness thus requires that its superpartner, the stop, is also accessible to cancel these loops. Since the stop lives in both the  $U_R$  and  $Q_L$  superfields, this typically also suggests that the left-handed sbottom is also light.
2. **Light Higgsinos.** In order to preserve natural electroweak symmetry breaking,  $\mu$  should be on the order of the electroweak scale. This is the same parameter that determines the Higgsino mass, so the Higgsinos should also be light.
3. **Not-too-heavy gluinos.** The stop is a scalar particle which is, itself, quadratically UV sensitive at face value. The main contribution to the stop mass comes from gluon loops so that naturalness requires ‘not-too-heavy’ ( $\sim 1.5$  TeV) gluinos to cancel these loops. In other words, the gluino feeds into the Higgs mass at two-loop order since it keeps the light-stop light enough to cancel the Higgs’ one-loop UV sensitivity.
4. **Light-ish electroweak-inos** (optional). Finally, if one insists on grand unification, the scale of the gluinos imposes a mass spectrum on the electroweak gauginos with  $M_{\text{EW-ino}} < M_{\text{gluino}}$ . As a rough estimate, Majorana gluinos should have mass  $\lesssim 2m_t$  while Dirac gluinos should have mass  $\lesssim 4m_t$ .
5. **All other particles decoupled.** All of the other squarks and sleptons are assumed to be well above the TeV scale and effectively inaccessible at Run-I of the LHC.

These are shown in Fig. 3.

The simplest models have a light stop  $\tilde{t}_L$  which decays either to a top and neutralino/gravitino,  $t + \tilde{N}$ , or a bottom and a chargino,  $b + \tilde{C}$ . Bounds on these decays depend on the  $\tilde{N}$  ( $\tilde{C}$ ) mass. The ‘stealthy’ region near  $m_{\tilde{N}} = 0$  and the ‘compressed’ region near  $m_{\tilde{N}} \approx m_t$  are especially difficult to probe kinematically.



**Fig. 3:** Heuristic picture of a natural SUSY spectrum. All other superpartners are assumed to have masses well above the TeV scale and decouple.

#### 2.17.4 $R$ -parity violation

One of the main ways to search for ‘vanilla’ SUSY signatures is to trigger on the large amount of missing energy (MET or  $\cancel{E}_T$ ) expected from the neutral LSP. Underlying this assumption is  $R$ -parity, which forces the LSP to be stable.

Recall that  $R$ -parity was something that we embraced because it killed the supersymmetric terms in the superpotential (2.50) that would violate lepton and baryon number and would be severely constrained by experiments, most notably proton decay. If, however, there were another way to suppress these dangerous operators, then perhaps we could avoid the experimental bounds while giving the LSP a way to decay into non-supersymmetric particles. This would allow us to consider models with  $R$ -parity violation (RPV) with no missing energy signal [51–55], see [56] for a review. Such models would be immune to the usual MET-based SUSY search strategies.

The simplest way to do this is to turn on only the  $\lambda_4 \bar{U} \bar{D} \bar{D}$  term. This violates baryon number but preserves lepton number so that protons remain stable. Motivated by naturalness, we may now allow the stop to be the LSP since this is no longer a dark matter candidate. The RPV coupling would allow a decay  $\tilde{t} \rightarrow \bar{b} \bar{s}$ , which would be hidden in the large QCD di-jet background.

One still has to worry about the effects of this RPV coupling on the partners of the light squarks. Phenomenologically, the strictest bounds come from neutron–anti-neutron oscillation and dinucleon decay. Indeed, most of the flavor bounds on the MSSM come from the first two generations of sparticles. One interesting model-building tool is to invoke **minimal flavor violation**, which posits that the flavor structure of the entire MSSM is carried by the Yukawa matrices [57]. This then implies that the coefficient of the  $\bar{U}^i \bar{D}^j \bar{D}^k$  RPV coupling is proportional to a product of Yukawa elements depending on the generations  $i, j$ , and  $k$ . This gives a natural explanation for why the RPV couplings of the first two generation squarks are much smaller than the stop.

### 3 Extra Dimensions

The original proposal for extra dimensions by Kaluza [58], Klein [59], and later Einstein [60] were attempts to unify electromagnetism with gravitation. Several decades later the development of string theory—originally as a dual theory to explain the Regge trajectories of hadronic physics—led physicists to revisit the idea of compact extra dimensions [61–63]. In early models, the non-observation of an additional spatial direction was explained by requiring the compactification radius to be too small for macroscopic objects.

**Further reading:** Two of the authors’ favorite reviews on this subject are [64] and [65]. This lecture is meant to be largely complementary. Additional references include [66–71], which focus on different aspects.

### 3.1 Kaluza-Klein decomposition

The simplest example to begin with is a real scalar field in 5D where the fifth dimension is compactified to a circle of radius  $R$ . The details of the compactification do not change the qualitative behavior of the theory at low energies. The Lagrangian is

$$S = \int d^5x \frac{1}{2} \partial_M \phi(x, y) \partial^M \phi(x, y) = \int d^5x \frac{1}{2} \left[ \partial_\mu \phi(x, y) \partial^\mu \phi(x, y) - (\partial_y \phi(x, y))^2 \right], \quad (3.1)$$

where  $M = 0, \dots, 5$  and  $x_5 = y$ . Since  $y$  is compact, we may identify energy eigenstates by doing a Fourier decomposition in the extra dimension,

$$\phi(x, y) = \frac{1}{\sqrt{2\pi R}} \sum_{n=-\infty}^{\infty} \phi^{(n)}(x) e^{i \frac{n}{R} y}. \quad (3.2)$$

Since  $\phi$  is real,  $(\phi^{(n)})^\dagger = \phi^{(-n)}$ . Plugging this expansion into the action allows us to use the orthogonality of the Fourier terms to perform the  $dy$  integral. This leaves us with an expression for the action that is an integral over only the non-compact dimensions, but written in terms of the KK modes  $\phi^{(n)}(x)$ ,

$$S = \int d^4x \sum_{mn} \left( \int dy \frac{1}{2\pi R} e^{i \frac{(m+n)}{R} y} \right) \frac{1}{2} \left[ \partial_\mu \phi^{(m)}(x) \partial^\mu \phi^{(n)}(x) + \frac{mn}{R^2} \phi^{(m)}(x) \phi^{(n)}(x) \right] \quad (3.3)$$

$$= \frac{1}{2} \int d^4x \sum_n \left[ \partial_\mu \phi^{(-n)} \partial^\mu \phi^{(n)} - \frac{n^2}{R^2} \phi^{(-n)} \phi^{(n)} \right] \quad (3.4)$$

$$= \int d^4x \sum_{n>0} \left[ \left( \partial_\mu \phi^{(n)} \right)^\dagger \partial^\mu \phi^{(n)} - \frac{n^2}{R^2} \left| \phi^{(n)} \right|^2 \right]. \quad (3.5)$$

From the 4D point of view, a single 5D scalar becomes a ‘Kaluza-Klein (KK) tower’ of 4D particles, each with mass  $n/R$ . If there were more than one extra dimension, for example if one compactified on an  $k$ -dimensional torus with radii  $R_5, R_6, \dots$ , then the KK tower would have  $k$  indices and masses

$$m_{n_5, n_6, \dots, n_k}^2 = m_0^2 + \frac{n_5^2}{R_5^2} + \frac{n_6^2}{R_6^2} + \dots + \frac{n_k^2}{R_k^2}, \quad (3.6)$$

where  $m_0^2$  is the higher dimensional mass of the field.

### 3.2 Gauge fields

A more complicated example is a gauge field. We know that gauge fields are associated with vector particles, but in 5D the vector now carries five components,  $A_M$ . We perform the same KK decomposition for each component  $M$ ,

$$A_M(x, y) = \frac{1}{\sqrt{2\pi R}} \sum_n A_M^{(n)}(x) e^{i \frac{n}{R} y}. \quad (3.7)$$

Note that this decomposes into a KK tower of 4D vectors,  $A_\mu^{(n)}$ , and a KK tower of 4D scalars,  $A_5^{(n)}$ . Similarly, the field strengths are antisymmetric with respect to indices  $M$  and  $N$  so that the action is decomposed according to

$$S = \int d^4x dy \left( -\frac{1}{4} F_{MN} F^{MN} \right) \quad (3.8)$$

$$= \int d^4x dy \left[ -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} (\partial_\mu A_5 - \partial_5 A_\mu) (\partial^\mu A_5 - \partial_5 A^\mu) \right] \quad (3.9)$$

$$= \int d^4x \sum_n -\frac{1}{4} F_{\mu\nu}^{(-n)} F^{(n)\mu\nu} + \frac{1}{2} \left( \partial_\mu A_5^{(-n)} - \partial_5 A_\mu^{(-n)} \right) \left( \partial^\mu A_5^{(n)} - \partial_5 A^{(n)\mu} \right). \quad (3.10)$$

This looks complicated because there is an odd mixing between the 4D vector,  $A_\mu^{(n)}$ , and the 4D scalar  $A_5^{(n)}$ . Fortunately, this mixing term can be removed by fixing to 5D axial gauge,

$$A_\mu^{(n)} \rightarrow A_\mu^{(n)} - \frac{i}{n/R} \partial_\mu A_5^{(n)} \quad A_5^{(n)} \rightarrow 0, \quad (3.11)$$

for  $n \neq 0$ . Note that for  $n = 0$  there's no scalar–vector mixing anyway. The resulting action takes a much nicer form,

$$S = \int d^4x -\frac{1}{4} \left( F_{\mu\nu}^{(0)} \right)^2 + \frac{1}{2} \left( \partial_\mu A_5^{(0)} \right)^2 + \sum_{n \geq 1} 2 \left( -\frac{1}{4} F_{\mu\nu}^{(-n)} F^{(n)\mu\nu} + \frac{1}{2} \frac{n^2}{R^2} A_\mu^{(-n)} A^{(n)\mu} \right). \quad (3.12)$$

The spectrum includes a tower of massive vector particles as well as a massless (zero mode) gauge boson and scalar.

Recall the usual expression for the number of degrees of freedom in a massless 4D gauge boson:

$$(4 \text{ components in } A_\mu) - (\text{longitudinal mode}) - (\text{gauge redundancy}). \quad (3.13)$$

When the gauge boson becomes massive, it picks up a longitudinal mode from eating a scalar by the Goldstone mechanism. This is precisely what has happened to our KK gauge bosons,  $A_\mu^{(n)}$ : they pick up a mass by eating the scalar KK modes,  $A_5^{(n)}$ .

In a theory with  $(4+n)$  dimensions, the  $(4+n)$ -component vector  $A_M$  decomposes into a massless gauge boson,  $n$  massless scalars, a tower of massive KK vectors  $A_\mu$ , and a tower of  $(n-1)$  massive KK scalars.

One may similarly generalize to spin-2 particles such as the graviton. In  $(4+n)$  dimensions these are represented by an antisymmetric  $(4+n) \times (4+n)$  tensor,

$$g_{MN} = \begin{pmatrix} g_{\mu\nu} & \vdots & A_\mu \\ \vdots & \ddots & \vdots \\ \varphi & \vdots & \varphi \end{pmatrix}. \quad (3.14)$$

The massless 4D zero modes include the usual 4D graviton, a vector, and a scalar. At the massive level, there is a KK tower of gravitons with  $(n-1)$  gauge fields and  $[\frac{1}{2}n(n+1) - n]$  scalars. Here we observe the graviton and vector eating the required degrees of freedom to become massive.

### 3.3 Matching of couplings

It is important to notice that the mass dimension of couplings and fields depend on the number of space-time dimensions. The action is dimensionless,  $[S] = 0$ , since it is exponentiated in the partition function. Then, in  $(4+n)$  dimensions, the kinetic term for a boson gives

$$\left[ d^{(4+n)}x (\partial\phi)^2 \right] = -(4+n) + 2 + 2[\phi] = 0 \quad \Rightarrow \quad [\phi] = 1 + \frac{n}{2}. \quad (3.15)$$

Note that this is consistent with the dimensions in the KK expansion (3.2). The 5D scalar contains the 4D scalars with a prefactor  $\sim R^{-1/2}$  that has mass dimension  $1/2$ . Similarly, for fermions,  $[\psi] = \frac{3}{2} + \frac{n}{2}$ . With this information, dimensions of the Lagrangian couplings can be read off straightforwardly. For example, the 5D gauge field lives in the covariant derivative,

$$D_\mu = \partial_\mu - ig_5 A_\mu = \partial_\mu - i \frac{g_5}{\sqrt{2\pi R}} A_\mu^{(0)} + \dots. \quad (3.16)$$

We see that  $[g_5] = -1/2$  since  $[\partial] = 1$  and  $[A_M] = 3/2$ . Further, we find an explicit relation between the 5D parameter  $g_5$  and the observed 4D gauge coupling,

$$g_4 = \frac{g_5}{\sqrt{2\pi R}}. \quad (3.17)$$

More generally, in  $(4+n)$  dimensions the 4D coupling is related to the higher dimensional coupling by the volume of the extra dimensional space,

$$g_4^2 = \frac{g_{(4+n)}^2}{\text{Vol}_n}. \quad (3.18)$$

One can read off the matching of the gravitational coupling by looking at the prefactor of the Ricci term in the action,

$$S_{(4+n)} = -M_{(4+n)}^{2+n} \int d^{4+n}x \sqrt{g} R_{(4+n)} = -M_{(4+n)}^{2+n} V_n \int d^4x \sqrt{g_{(4)}} R_{(4)} + \dots, \quad (3.19)$$

where we've written  $g$  for the determinant of the metric. From this we identify 4D Planck mass  $M_{\text{Pl}}$  from the fundamental higher dimensional Planck mass,  $M_{(4+n)}$ ,

$$M_{\text{Pl}}^2 = M_{(4+n)}^{2+n} V_n. \quad (3.20)$$

The higher dimensional Planck mass is a good choice for a fundamental mass scale for the theory,

$$M_* = M_{(4+n)}. \quad (3.21)$$

In a  $(4+n)$  dimensional theory where the characteristic mass scale is  $M_*$  and a compactification radius  $R$ . Then dimensional analysis tells us that the higher dimensional gauge couplings, which are dimensionful, characteristically scale like

$$g_{(4+n)} \sim M_*^{-n/2}. \quad (3.22)$$

Relating this to the 4D couplings with (3.18) and relating  $M_*$  to the 4D Planck mass with (3.20) gives

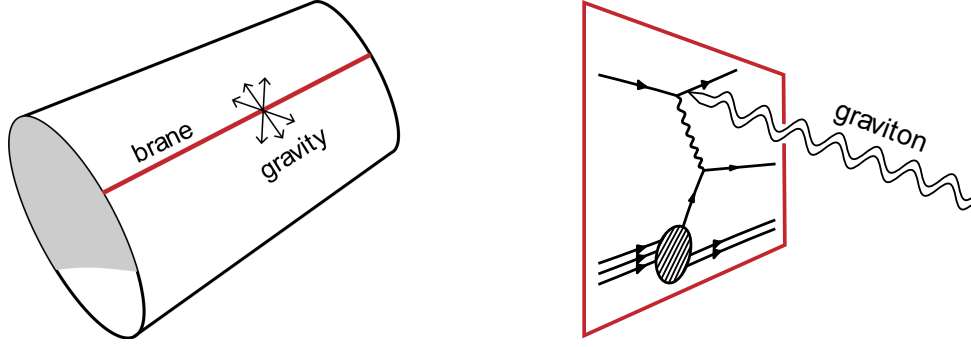
$$R \sim \frac{1}{M_{\text{Pl}}} g_4^{(n+2)/N}. \quad (3.23)$$

Plugging in the observed SM gauge couplings on the right hand side gives a compactification radius which is far too small to be relevant at colliders—the first KK modes will be near the Planck scale.

### 3.4 Branes and Large Extra Dimensions

In the mid '90s, developments in string theory led to a new ingredient that renewed interest in extra dimensions that might be accessible at collider scales. The key idea is that **branes**, solitonic objects which form lower dimensional subspaces, can trap fields. In other words, not all fields have to propagate in all dimensions. This was introduced by Rubakov and Shaposhnikov [72], who showed that instead of a very small radius of compactification, it may be that our observed universe is constrained to live in a  $(3+1)$ -dimensional subspace of a higher dimensional spacetime.

**Terminology.** Models that make use of branes to localize fields are known as **braneworld** models and are distinguished from models where all fields propagate in the extra dimensions, known as **universal extra dimensions**. In braneworld models, fields which are allowed to propagate in the full space are said to live in the **bulk**.



**Fig. 4:** Cartoon pictures of a (3+1) dimensional brane in a compact 5D space. (LEFT) The brane (red line) as a subspace. Gravity propagates in the entire space ‘diluting’ its field lines relative to forces localized on the brane. (RIGHT) SM processes localized on the brane, now with an additional dimension drawn, emitting a graviton into the bulk.

Allowing the fields to be brane-localized buys us quite a lot. It allows us to separate particle physics from gravity. One can, for example, force the SM fields to be truly four-dimensional objects that are stuck to a (3+1)-dimensional brane. This avoids the bound on the size of the extra dimension in (3.23), since that relied on the SM propagating in the bulk.

With this in mind, one could allow the volume of the extra dimensions to actually be quite large. This idea was explored by Arkani-Hamed, Dimopoulos, and Dvali in the ADD or **large extra dimension** scenario [73]. If this were feasible, then (3.20) gives a new way to address the Hierarchy problem. The large volume factor allows the fundamental scale of nature to be much smaller than the observed Planck mass,  $M_* \ll M_{\text{Pl}}$ . If, for example,  $M_* \sim 1 \text{ TeV}$ , then there is no Hierarchy problem. Gravity appears to be weaker at short distances because its flux is diluted by the extra dimensions. As one accesses scales smaller than  $R$ , however, one notices that gravity actually propagates in  $(4 + n)$  dimensions. A cartoon of the braneworld scenario is shown in Fig. 4.

How large can this extra dimension be? Doing a rough matching and using  $\text{Vol}_n = r^n$  in (3.20) gives

$$R = \frac{1}{M_*} \left( \frac{M_{\text{Pl}}}{M_*} \right)^{2/n}. \quad (3.24)$$

Pushing the fundamental scale to  $M_* \sim \text{TeV}$  requires

$$R = 10^{32/n} \text{ TeV}^{-1} = 2 \cdot 10^{-17} 10^{32/n} \text{ cm}, \quad (3.25)$$

using  $\text{GeV}^{-1} = 2 \cdot 10^{-14} \text{ cm}$ . We make the important caveat that this is specifically for the ADD model. Considering different numbers of extra dimensions,

- $n = 1$ . For a single extra dimension we have  $R = 10^{15} \text{ cm}$ , which is roughly the size of the solar system and is quickly ruled out.
- $n = 2$ . Two extra dimensions brings us down to  $R \approx 0.1 \text{ cm}$ , which is barely ruled out by gravitational Cavendish experiments.
- $n = 3$ . Three extra dimensions pushes us down to  $R < 10^{-6} \text{ cm}$ .

How much do we know about gravity at short distances? Surprisingly little, actually. Cavendish experiments (e.g. Eöt-Wash<sup>5</sup>) test the  $r^{-2}$  law down to  $10^{-4} \text{ m}$ . These set a direct bound on the  $n = 2$  case that  $R < 37 \mu\text{m}$  and  $M_* > 1.4 \text{ TeV}$ . For larger  $n$  one is allowed to have  $M_* = \text{TeV}$ .

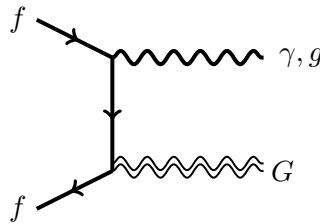
<sup>5</sup>The name is a play on the Eötvös experiment by University of Washington researchers.



One might have objected that one cannot say that  $M_*$  is the fundamental scale while allowing  $R$ , itself a dimensionful quantity, float to take on any value. Indeed, in a completely natural theory, one expects  $R \sim 1/M_*$  so that  $R \sim \text{TeV}^{-1}$ . This is quite different from what we wrote in (3.24). Indeed, what we have done here is swapped the hierarchy in mass scales to a Hierarchy between  $R$  and  $M_*^{-1}$ . In other words, we have reformulated the Hierarchy problem to a problem of **radius stabilization**. This is indeed very difficult to solve in ADD.

Nevertheless, we may explore the phenomenological consequences of an ADD type model at colliders and through astrophysical observations.

- The first thing to consider is the production of KK gravitons.



The KK graviton couples too weakly to interact with the detector so it appears as missing energy. By itself, however, missing energy is difficult to disentangle from, say, neutrino production. Thus it's useful to have a handle for the hardness of the event (more energetic than  $Z \rightarrow \nu\bar{\nu}$ ) so one can look for processes that emit a hard photon or gluon. Thus a reasonable search is a jet or photon with missing energy. It is worth noting that this is the same search used for searching for dark matter, which is also typically a massive particle which appears as missing energy.

- Alternately, one may search for  $s$ -channel virtual graviton exchange in processes like  $e^+e^- \rightarrow f\bar{f}$ . One expects a resonance at the KK graviton mass.
- Supernovae can cool due to the emission of gravitons. This is similar to the supernovae cooling bounds on axions. The strongest bounds on  $n = 2$  theories push  $M_* \gtrsim 100 \text{ TeV}$ .
- An additional byproduct of lowering the fundamental gravitational scale is that one may form microscopic black holes at energies kinematically accessible to the LHC and cosmic rays. For  $E_{\text{CM}} > M_*$  black holes are formed with a radius

$$R_S \sim \frac{1}{M_*} \left( \frac{M_{\text{BH}}}{M_*} \right)^{\frac{1}{n+1}}. \quad (3.26)$$

the cross section is roughly the geometric value,  $\sigma_{\text{BH}} \sim \pi R_S^2$  and can be as large as 400 pb. These microscopic black holes decay via Hawking radiation,

$$T_H \sim \frac{1}{R_S} \quad (3.27)$$

with this energy distributed equally to all degrees of freedom, for example 10% going to leptons, 2% going to photons, and 75% going to many jets.

### 3.5 Warped extra dimensions

We've seen that the framework of large extra dimensions leads to interesting phenomenology, but the ADD realization leaves the size of the radius unexplained and is therefore not a complete solution to the Hierarchy problem. The Randall-Sundrum (RS) proposal for a warped extra dimension offers a more

interesting possibility [74]. The set up differs from ADD in that the space between the two branes has a non-factorizable metric that depends on the extra space coordinate,  $z$ ,

$$ds^2 = \left(\frac{R}{z}\right)^2 (\eta_{\mu\nu} dx^\mu dx^\nu - dz^2). \quad (3.28)$$

This is the metric of anti-de Sitter space (AdS) with curvature  $k = 1/R$ . There are two branes located at  $z = R$  (the UV brane) and  $z = R' > R$  (the IR brane) that truncate the extra dimension; in this sense the RS background is often described as a ‘slice of AdS.’ We see that  $1/R$  is naturally a fundamental UV scale of the theory. The metric (3.28) warps down the natural physical scale as a function of the position along the extra dimension. In particular, when  $R' \gg R$  one finds that near  $z = R'$ , the scales are warped down to much smaller values. Note the different notation from the ADD case: the size of the extra dimension is  $R' - R \approx R'$ , while  $R$  should be identified with the radius of curvature.

To see how this works, suppose that the Higgs is localized to live on the IR brane at  $z = R'$ . The action on this brane depends on the 4D induced metric  $\hat{g}_{\mu\nu}$  (note that  $\sqrt{\hat{g}} = \sqrt{g}/\sqrt{g_{55}}$ ),

$$S = \int d^4x \sqrt{\hat{g}} \left[ \partial_\mu H \partial_\nu H \hat{g}^{\mu\nu} - \left( |H|^2 - \frac{v^2}{2} \right)^2 \right] \quad (3.29)$$

We assume that the Higgs VEV is on the order of the UV scale,  $v = 1/R$ , since this is the fundamental 5D scale. Plugging in the metric gives

$$S = \int d^4x \left(\frac{R}{R'}\right)^4 \left[ \partial_\mu H \partial^\mu H \left(\frac{R'}{R}\right)^2 - \left( |H|^2 - \frac{v^2}{2} \right)^2 \right]_{z=R'} \quad (3.30)$$

where indices are implicitly raised with respect to the Minkowski metric. Canonically normalizing the kinetic term via

$$\hat{H} = \frac{R}{R'} H, \quad (3.31)$$

allows us to write the action in the form,

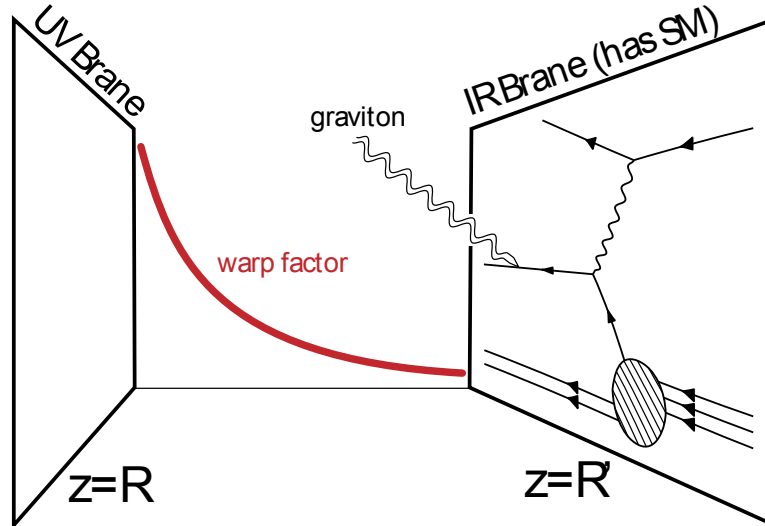
$$S = \int d^4x \left( \partial_\mu \hat{H} \right)^2 - \lambda \left[ |\hat{H}|^2 - \frac{1}{2} \left( v \frac{R}{R'} \right)^2 \right]^2, \quad (3.32)$$

where we see that the canonically normalized Higgs picks up a VEV that is warped down to the TeV scale. One can further imagine that the cutoff for loops contributing to the Higgs mass are similarly warped down to, say, the TeV scale. In this way, the warped extra dimension gives a new handle for generating hierarchies. Readers should be skeptical that we’re not just hiding the Hierarchy problem in some fine tuning of the IR scale  $R'$  relative to the fundamental scale  $R$ . Indeed, the real solution to the Hierarchy problem requires a mechanism for radius stabilization, which we present below. Note that typically  $R \approx M_{\text{Pl}}^{-1}$  and  $R' \approx \text{TeV}^{-1}$  so that  $R'$  is roughly the size of the extra dimension. A cartoon of this scenario is shown in Fig. 5.

In the remainder of this lecture we’ll focus on the RS background. In the appendices we present some additional technical results that may be useful for building RS models. Further details of the RS gravitational background are discussed in Appendices A.1 and A.2. Details of bulk matter fields are discussed in Appendices A.3 and A.4

### 3.6 The Planck scale and hierarchy in RS

We have seen how the AdS curvature can warp mass scales to be much smaller than the fundamental 5D scale  $1/R$ . It is instructive to also check the observed Planck scale. With respect to the fundamental



**Fig. 5:** Cartoon of the RS scenario with a brane-localized SM. The warp factor,  $(R/z)^2$ , causes energy scales to be scaled down towards the IR brane.

Planck scale  $M_*$  (ostensibly  $M_* \sim 1/R$ ), the gravitational action is

$$S_g = M_*^3 \int_R^{R'} dz \int d^4x \sqrt{g_{(5)}} \mathcal{R}_{(5)}, \quad (3.33)$$

where the quantities with subscripts are the determinant of the 5D metric and the 5D Ricci scalar, respectively. By performing the  $dz$  integral one finds the effective 4D gravitational action,

$$S_g = M_*^3 \int_R^{R'} dz \left(\frac{R}{z}\right)^3 \int d^4x \sqrt{g_{(4)}} \mathcal{R}_{(4)} = M_*^3 \left[1 - \left(\frac{R}{R'}\right)^2\right] \int d^4x \sqrt{g_{(4)}} \mathcal{R}_{(4)}. \quad (3.34)$$

We can thus identify the effective 4D Planck mass by reading off the coefficient,

$$M_{\text{Pl}}^3 = M_*^3 \left[1 - \left(\frac{R}{R'}\right)^2\right] \approx M_*^3, \quad (3.35)$$

so that for a large extra dimension,  $R' \gg R$ , the 4D Planck mass is insensitive to  $R'$  and is fixed by the 5D Planck mass,  $M_* \sim 1/R$ . This is precisely what we have set out to construct: assuming there is a dynamical reason for  $R' \gg R$ , we are able to warp down masses to the TeV scale by forcing particles to localize on the IR brane while simultaneously maintaining that 4D observers will measure a Planck mass that is much heavier.

An alternate way of saying this is that the Hierarchy problem is solved because the SM Higgs is peaked towards the IR brane while gravity is peaked towards the UV brane. What we mean by the latter part of this statement is that the graviton zero mode has a bulk profile that is peaked towards the UV brane. Recall that in flat space, zero modes have flat profiles since they carry no momentum in the extra dimension. In RS, the warping of the space also warps the shape of the graviton zero mode towards the UV brane; the weakness of gravity is explained by the smallness of the graviton zero mode profile where the Standard Model particles live. This should be compared to the case of a flat interval where the zero mode wave function decouples as the size of the extra dimension increases. In this case the coupling with the IR brane indeed becomes weaker, but the graviton KK modes become accessible and can spoil the appearance of 4D gravity. In RS the zero mode doesn't decouple and one doesn't need to appeal to a dilution of the gravitational flux into the extra dimensions as in the ADD model. See [64] for more explicit calculations in this picture.

### 3.7 Bulk scalar profiles in RS

In the original RS model, only gravity propagates in the bulk and has KK modes. However, it is instructive to derive the KK properties of a bulk scalar.

- This serves as a simple template for how to KK reduce more complicated fields, such as the graviton, in a warped background.
- We’re anticipating the ‘modern’ incarnations of the RS where gauge and matter fields are pulled into the bulk.
- As mentioned above, the solution to the Hierarchy problem depends on stabilizing the position of the IR brane,  $z = R'$ , relative to the UV brane,  $z = R$ . The standard technique for doing this requires a bulk scalar.

Start with a bulk complex scalar  $\Phi(x, z)$  with a bulk mass parameter  $m$ . The bulk action is

$$S = \int_R^{R'} dz \int d^4x \sqrt{g} [(\partial_M \Phi)^* \partial^M \Phi - m^2 \Phi^* \Phi]. \quad (3.36)$$

In principle one may have additional brane-localized interactions proportional to  $\delta(z - R')$  or  $\delta(z - R)$ . We use  $M, N$  to index 5D coordinates while  $\mu, \nu$  only run over 4D coordinates. Varying with respect to  $\Phi^*$  yields an equation of motion

$$\partial_M (\sqrt{g} g^{MN} \partial_N \Phi) - \sqrt{g} m^2 \Phi = 0. \quad (3.37)$$

In writing this we have dropped an overall surface term that we picked up when integrating by parts. Specializing to the RS metric, this amounts to picking boundary conditions such that

$$\Phi^*(z) \partial_z \Phi(z) |_{R, R'} = 0, \quad (3.38)$$

with the appropriate modifications if there are brane-localized terms. We see that we have a choice of Dirichlet and Neumann boundary conditions. We now plug in the Kaluza-Klein decomposition in terms of yet-unknown basis functions  $f^{(n)}(z)$  which encode the profile of the  $n^{\text{th}}$  mode in the extra dimension:

$$\Phi(x, z) = \sum_n^{\infty} \phi^{(n)}(x) f^{(n)}(z). \quad (3.39)$$

By assumption the  $\phi^{(n)}$  are eigenstates of  $\eta^{\mu\nu} \partial_\mu \partial_\nu$  with eigenvalue  $-m_{(n)}^2$ , the KK mass. We are thus left with a differential equation for  $f^{(n)}(z)$ ,

$$\left[ \left( \frac{R}{z} \right)^3 m_{(n)}^2 - \frac{3}{z} \left( \frac{R}{z} \right)^3 \partial_z + \left( \frac{R}{z} \right)^3 \partial_z^2 - \left( \frac{R}{z} \right)^5 m^2 \right] f^{(n)}(z) = 0. \quad (3.40)$$

This is a Sturm-Liouville equation with real eigenvalues and real, orthonormal eigenfunctions,

$$\int_R^{R'} dz \left( \frac{R}{z} \right)^3 f^{(n)}(z) f^{(m)}(z) = \delta^{mn}. \quad (3.41)$$

Just as we saw in Section 3.1 for a flat extra dimension, this orthonormality relation diagonalizes the KK kinetic terms. One may now solve (3.40) by observing that through suitable redefinitions this is simply a Bessel equation. The result is a general solution for  $n > 0$  of the form

$$f^{(n)}(z) = c_1 z^2 J_\alpha(m_{(n)} z) + c_2 z^2 Y_\alpha(m_{(n)} z), \quad (3.42)$$

where  $J, Y$  are the familiar Bessel functions and  $\alpha = \sqrt{4 + m^2 R^2}$ . The integration constants  $c_{1,2}$  and the spectrum of KK masses  $m_{(n)}^2$  can be found using boundary conditions on each brane and the orthonormality relation (3.41). The states have a discrete spectrum with spacing of approximately  $R'^{-1} \sim \text{TeV}$  with profiles peaked towards the IR brane.

For  $n = 0$  the zero mode profile is

$$f^{(0)}(z) = c_1 z^{2 - \sqrt{4 + m^2 R^2}} + c_2 z^{2 + \sqrt{4 + m^2 R^2}}. \quad (3.43)$$

We use this result in the Goldberger-Wise mechanism discussed below, but let us remark that the zero mode is neither consistent with Neumann nor Dirichlet boundary conditions and requires brane localized terms to generate boundary conditions that permit a zero mode.

The same general procedure can be used to find the profiles of higher spin bulk fields. In Appendices A.3 and A.4 we work through the additional subtleties coming from fermions and gauge bosons. A Standard Model field is associated with the zero mode of a 5D field, where the SM mass is a correction from electroweak symmetry breaking on the zero mass from the KK decomposition. Note that the meaning of the 5D profile is that a 4D particle, even though it is localized and pointlike in the four Minkowski dimensions, is an extended plane wave in the fifth dimension. The boundary conditions imposed by the branes mean that this system is essentially identical to a waveguide in electrodynamics<sup>6</sup>.

### 3.8 Radius stabilization

We've now shifted the Hierarchy problem to a question of why the IR scale  $R'$  is so much larger than the UV scale  $R$ . In fact, one should think about  $R'$  as the expectation value of a dynamical degree of freedom,  $R' = \langle r(x, z) \rangle$ , called the **radion**. This is identified with the 4D scalar arising from the dimensional decomposition of the 5D metric. This isn't surprising since the metric is, of course, the quantity which measures distances. Thus far in our description of the RS framework, the radion is a **modulus**—it has no potential and could take any value. This is problematic since excitations of this field would be massless and lead to long-range modifications to gravity. It is thus important to find a mechanism that dynamically fixes  $R' \sim \text{TeV}^{-1}$  to (1) provide a complete solution to the Hierarchy problem and (2) avoid constraints from modifications to gravity.

**Don't be fooled by coordinate choices.** The original RS literature used variables such that the metric explicitly contained an exponential warping  $ds^2 = e^{-2ky} dx^2 - dy^2$  so that an  $\mathcal{O}(10)$  value of  $k\pi R'$  leads to large hierarchies. Do not confuse this variable choice with a solution to the Hierarchy problem—it just shifts the fine tuning into a parameter to which the theory is exponentially sensitive. The reason why the exponential hierarchy is actually physical in RS (with a dynamically stabilized radius) is that fields propagating in the space are redshifted as they 'fall' towards the IR brane in the gravitational well of the AdS background.

A standard solution in the RS model is the Goldberger-Wise mechanism<sup>7</sup> [76, 77], where radion kinetic and potential energy terms conspire against one another to select vacuum with finite  $R'$ . To do this, we introduce a massive bulk scalar field  $\Phi(x, z)$  of the type in Section 3.7. We introduce brane-localized potentials for this field which force it to obtain a different VEV at each brane,  $\varphi_{UV} \neq \varphi_{IR}$ ,

$$\Delta\mathcal{L} = -\lambda\delta(z - R) (\Phi^2 - \varphi_{UV}^2)^2 - \lambda\delta(z - R') (\Phi^2 - \varphi_{IR}^2)^2 \quad \lambda \rightarrow \infty. \quad (3.44)$$

This causes the scalar to pick up a  $z$ -dependent VEV that interpolates between  $\varphi_{UV}$  and  $\varphi_{IR}$ ,

$$\langle \Phi(x, z) \rangle = \varphi(z) \quad \varphi(R) = \varphi_{UV} \quad \varphi(R') = \varphi_{IR}. \quad (3.45)$$

<sup>6</sup>This should have been no surprise given the appearance of Bessel functions.

<sup>7</sup>While the Goldberger-Wise mechanism is just one simple option to stabilize the size of the extra dimension, it is close to what actually happens in string compactifications that tacitly UV complete the RS scenario [75].

The general form of  $\varphi(z)$  is precisely the zero mode profile in (3.43) since the VEV carries zero momentum in the Minkowski directions. One may now consider the terms in the action of  $\Phi(x, z)$  (evaluated on the VEV  $\varphi(z)$ ) as contributions to the potential for the radion via  $R' = \langle r(x, z) \rangle$ . The kinetic term for  $\Phi(x, z)$  contributes a potential to  $r(x, z)$  that goes like  $\varphi'(z)^2$ .

1. This gradient energy is minimized when  $\varphi(z)$  has a large distance to interpolate between  $\varphi_{\text{UV,IR}}$  since larger  $R'$  allows a smaller slope.
2. On the other hand, the bulk mass for  $\Phi(x, z)$  gives an energy per unit length in the  $z$ -direction when  $\varphi(z) \neq 0$ . Thus the energy from this term is minimized when  $R'$  is small.

By balancing these two effects, one is able to dynamically fix a value for  $R'$ . A pedagogical derivation of this is presented in [66]. The main idea is that for small values of the bulk  $\Phi(x, z)$  mass,  $m^2 \ll R'^{-2}$ , one may write the  $\Phi(x, z)$  VEV as

$$\varphi = c_1 z^{-\varepsilon} + c_2 z^{4+\varepsilon}, \quad (3.46)$$

where  $\varepsilon = \alpha - 2 = \sqrt{4 + m^2 R'^2} - 2 \approx m^2 R'^2 / 4$  is small. The coefficients  $c_{1,2}$  are determined by the boundary conditions (3.45). The potential takes the form

$$V[R'] = \varepsilon \frac{\varphi_{\text{UV}}^2}{R} + \frac{R^3}{R'^4} \left[ (4 + 2\varepsilon) \left( \varphi_{\text{IR}} - \varphi_{\text{UV}} \left( \frac{R}{R'} \right)^\varepsilon \right)^2 - \varepsilon \varphi_{\text{IR}} R'^{-4} \right] + \mathcal{O} \left( \frac{R^4}{R'^8} \right), \quad (3.47)$$

where judicious checkers of dimensions will recall that the dimension of the 5D scalar is  $[\varphi(x, z)] = \frac{3}{2}$ . The minimum of this potential is

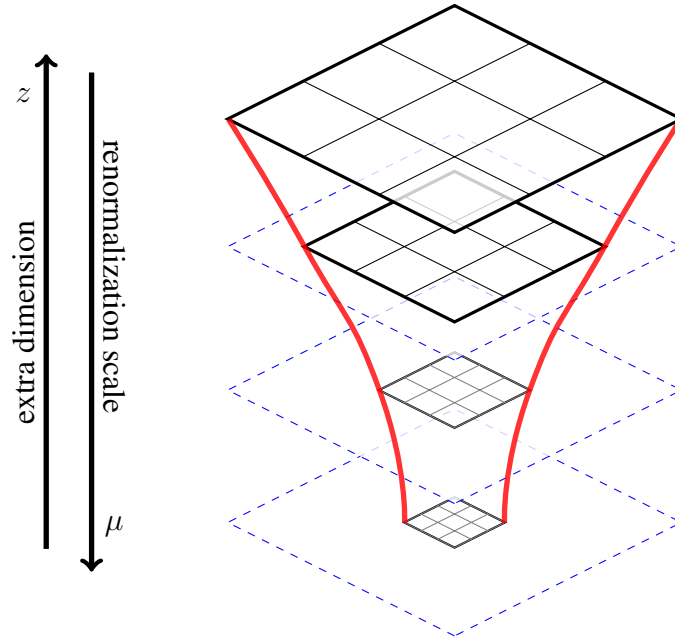
$$R' = R \left( \frac{\varphi_{\text{UV}}}{\varphi_{\text{IR}}} \right)^{1/\varepsilon}. \quad (3.48)$$

We can generate the Planck-weak hierarchy with  $1/\varepsilon \sim 20$  and  $\varphi_{\text{UV}}/\varphi_{\text{IR}} \sim 10$ . A key point here is that we may write the radius in terms of a characteristic energy scale,  $R' \sim 1/\mu$ , and the potential for  $\mu$  carries terms that go like  $\mu^4$  times a polynomial in  $\mu^\varepsilon$ . This is reminiscent of dimensional transmutation and, indeed, we explain below that the RS scenario can be understood as a dual description of strongly coupled 4D dynamics.

The above description of the Goldberger-Wise mechanism neglects the effect of the background  $\Phi$  field on the RS geometry. For example, one may wonder if the RS metric is even compatible with the  $\Phi$  VEV. In order to account for this gravitational backreaction, one must solve the  $\Phi$  equation of motion combined with the Einstein equation as a function of the metric (discussed in Appendix A.1) in the presence of the  $\Phi$  VEV. This set of coupled second order differential equations is generically very difficult to solve. Fortunately, there exists a ‘superpotential’<sup>8</sup> trick that one may apply to solve the system exactly. This method is described and demonstrated pedagogically for the Goldberger-Wise field in [64, 67]. One finds that it is indeed possible to maintain the RS background in the presence of the bulk field necessary to stabilize the radius.

### 3.9 Holographic interpretation

<sup>8</sup>The trick was inspired by similar calculations in supergravity but otherwise is only related to SUSY in the sense that the ‘superpotential’ here also allows one to write first order equations of motion [78].



**Fig. 6:** Cartoon of the AdS/CFT correspondence. The isometries of the extra dimensional space enforce the conformal symmetry of the 4D theory. Moving in the  $z$  direction corresponds to a renormalization group transformation (rescaling) of the 4D theory.

**Gauge/gravity duality** is a way to understand the physics of a warped extra dimension as the dual to a strongly coupled 4D theory. Our goal here is to develop the intuition to use and understand the AdS/CFT dictionary as an interpretational tool. The most rigorous explicit derivations of this duality are often presented in the language of string theory. This idea is presented pedagogically in the language of 4D quantum field theory (rather than string theory) in [69, 70, 79–83]. Those interested in presentations that connect to supergravity and string theory may explore [?, ?, ?, ?, 84], listed roughly in order of increasing formal theory sophistication starting from very little assumed background. We also point out [85] which is an excellent presentation of dualities between 4D supersymmetric gauge theories that are analogous to the gauge/gravity correspondence.

We now introduce a way to re-interpret the observables of RS scenario in terms of the dynamics of a purely four-dimensional theory in its non-perturbative regime. The idea is that the symmetries of the bulk AdS space enforce the symmetries of a conformal theory in 4D—this latter theory approximates a strongly coupled theory near a fixed point. Combined with the observation that a shift in  $z$  causes an overall rescaling of the AdS metric (3.28), we can identify slices of constant  $z$  as scale transformations of the 4D [approximately] conformal theory. In this way, the 5D AdS theory ‘geometrizes’ the renormalization group flow of the 4D theory. One then interprets the physics on the UV brane as a 4D conformal theory that sets the boundary conditions for the 5D fields. Slices of constant  $z$  describe the RG evolution of this theory at lower energies,  $\mu \sim 1/z$ . Because the higher-dimensional theory encodes information about the behavior of a lower-dimensional theory on its boundary, this identification is known as the **holographic** interpretation of warped extra dimensions. This interpretation is sketched in Fig. 6.

### 3.9.1 Plausibility check from an experimentalist’s perspective

As a very rough check of why this would be plausible, consider the types of spectra one expects from an extra dimensional theory versus a strongly coupled 4D theory. In other words, consider the first thing that an experimentalists might want to check about either theory. The theory with an extra dimension predicts a tower of Kaluza-Klein excitations for each particle. The strongly coupled gauge theory predicts

a similar tower of bound states such as the various meson resonances in QCD. From the experimentalist’s point of view, these two theories are qualitatively very similar.

### 3.9.2 Sketch of a more formal description

We can better motivate the holographic interpretation by appealing to more formal arguments. One of the most powerful developments in theoretical physics over the past two decades is the AdS/CFT correspondence—more generally, the **holographic principle** or the **gauge/gravity correspondence** [79, 86–88]. The conjecture states that type IIB string theory on  $\text{AdS}_5 \times S^5$  is equivalent to 4D  $\mathcal{N} = 4$  superconformal  $\text{SU}(N)$  theory on Minkowski space in the large  $N$  limit:

$$\text{AdS}_5 \times S^5 \quad \iff \quad \mathcal{N} = 4 \text{ super Yang-Mills.} \quad (3.49)$$

The essence of this duality is the observation that a stack of  $N$  so-called  $D3$ -branes in string theory can be interpreted at low energies in two ways:

1. A solitonic configuration of closed strings which manifests itself as an extended black hole-like object for which  $\text{AdS}_5 \times S^5$  is a solution.
2. Dirichlet boundary conditions for open strings which admit a non-Abelian  $U(N)$  gauge symmetry associated with the  $N$  coincident  $D3$ -branes.

These correspond to the left- and right-hand side of (3.49) and form the basis of the AdS/CFT correspondence.

The key for us is that the  $\text{AdS}_5 \times S^5$  extra dimension ‘geometrizes’ the renormalization group flow of the strongly coupled theory by relating the position in the extra dimension  $z$  with the RG scale  $\mu$ . An operator  $\mathcal{O}_i$  in the 4D theory has a source  $j_i(x, \mu)$  that satisfies an RG equation

$$\mu \frac{\partial}{\partial \mu} j_i(x, \mu) = \beta_i(j_j(x, \mu), \mu). \quad (3.50)$$

The gauge/gravity correspondence identifies this source as the value of a bulk field  $j_i(x, \mu) \Leftrightarrow \Phi_i(x, z)$  at the UV boundary of the  $\text{AdS}_5$  extra dimension. The profile of  $\Phi_i$  in the extra dimension is associated with the RG flow of  $j_i(x, \mu)$ . Each Minkowski slice of  $\text{AdS}_5$  represents a picture of the 4D theory probed at a different energy scale  $\mu \sim 1/z$ .

More concretely, the duality gives a prescription by which the correlation functions of one theory are identified with correlation functions of the other. The parameters of these two theories are related by

$$\frac{R^4}{\ell^4} = 4\pi g^2 N, \quad (3.51)$$

where  $R$  is the AdS curvature,  $\ell$  is the string length, and  $g$  is the Yang Mills coupling. Here we see why AdS/CFT is such a powerful tool. In the limit of small string coupling  $\alpha' \sim \ell^2$  where string theory can be described by classical supergravity, the dual gauge theory is strongly coupled and very ‘quantum’. The correlation functions of that theory are non-perturbative and difficult to calculate, whereas the dual description is weakly coupled. The duality gives a handle to calculate observables in theories outside the regime where our usual tools are applicable.

### 3.9.3 What it means to geometrize the RG flow

For our purposes, it is only important that we understand the warped extra dimension as the renormalization group flow of a strongly coupled 4D gauge theory. To see how this RG flow is ‘geometrized,’ we consider the internal symmetries of the two theories.



- The isometry of the  $S^5$  space is  $\text{SO}(6) \cong \text{SU}(4)$ . This is precisely the  $R$ -symmetry group of the  $\mathcal{N} = 4$  gauge theory.
- The isometry of the  $\text{AdS}_5$  space is  $\text{SO}(4, 2)$ , which exactly matches the spacetime symmetries of a 4D conformal theory.

Since RS only has a slice of the AdS space without the  $S^5$ , we expect it to be dual to a conformal theory without supersymmetry. Steps towards formalizing the holographic interpretation of Randall-Sundrum are reviewed in [85].

Armed with this background, we can develop a working understanding of how to interpret RS models as a picture of a strong, four-dimensional dynamics. Observe that in the conformal coordinates that we've chosen, the metric has a manifest scale symmetry

$$z \rightarrow \alpha z \qquad x \rightarrow \alpha x. \qquad (3.52)$$

Consider 4D cross sections perpendicular to the  $z$  direction. Moving this cross section to another position  $z \rightarrow \alpha z$  is equivalent to a rescaling of the 4D length scales. Increasing  $z$  thus corresponds to a decrease in 4D energy scales. In this way, the AdS space gives us a holographic handle on the renormalization group behavior of the strongly coupled theory.

### 3.9.4 What it means to take a slice of Anti-de Sitter

The RS scenario differs from  $\text{AdS}_5$  due to the presence of the UV and IR branes which truncate the extra dimension. Since flows along the extra dimension correspond to scale transformations, the branes represent scales at which conformal symmetry is broken. The UV brane corresponds to an explicit UV cutoff for the 4D conformal theory. The IR brane sets the scale of the KK modes. We heuristically identified these with bound states of the strongly coupled theory, and so we can identify the IR brane as a scale where conformal symmetry is spontaneously broken, the theory confines, and one finds a spectrum of bound states. Recall that the bound state profiles are localized toward the IR brane; this is an indication that these bound states only exist as one approaches the confinement scale. The picture of the RS 'slice of AdS' is thus of a theory which is nearly conformal in the UV that runs slowly under RG flow down to the IR scale where it produces bound state resonances.

The SM, and in particular the Higgs, exist on the IR brane and are thus identified with composite states of the strongly coupled theory. In the extra dimensional picture, we argued that the Higgs mass is natural because the UV cutoff was warped down to the TeV scale. In the dual theory, the solution to the Hierarchy problem is compositeness (much like in technicolor): the scalar mass is natural because above the confinement scale the scalar disappears and one accesses its strongly coupled constituents. By comparison, a state stuck on the UV brane is identified with an elementary (non-composite) field that couples to the CFT.

### 3.9.5 The meaning of 5D calculations

At the level presented, it may seem like the AdS/CFT correspondence is a magic wand for describing strong coupling perturbatively—and indeed, if you have started to believe this, it behooves you to always know the limits of your favorite tools. A 5D calculation includes entire towers of 4D strongly coupled bound states—in what sense are we doing a perturbation expansion? First of all, we underscore that the AdS/CFT correspondence assumes the 't Hooft large  $N$  limit, where  $N$  is the rank of the gauge group [?]. Further, whether in four or five dimensions, a scattering calculation assumes a gap in the particle spectrum. This gap in the 5D mass is translated into a gap in the scaling dimension  $\Delta$  of the 4D CFT operators. Thus one of the implicit assumptions of a holographic calculation is that the spectrum of the CFT has a gap in scaling dimensions. More practically, a scattering process in 5D include 4D fields with large KK masses. We can say definite things about these large KK mass states, but only as long as these questions include a sum over the entire tower.

### 3.10 The RS Radion is a Dilaton

We have already met the **radion** as the dynamical field whose VEV sets the distance between the UV and IR brane. Excitations of the radion about this VEV correspond to fluctuations in the position of the IR brane. From its origin as a part of the 5D dynamical metric, it couples to the trace of the energy-momentum tensor,

$$\frac{r}{\Lambda_{\text{IR}}} T_{\mu}^{\mu}. \quad (3.53)$$

Observe that this is very similar to the coupling of the SM Higgs except that it is scaled by a factor of  $\frac{v}{\Lambda_{\text{IR}}}$  and there are additional couplings due to the trace anomaly—for example, a coupling to gluons of the form [77, 89]

$$\left[ \frac{r}{\Lambda_{\text{IR}}} - \frac{1}{2} \frac{r}{\Lambda_{\text{IR}}} F_{1/2}(m_t) \right] \frac{\alpha_s}{8\pi} (G_{\mu\nu}^a)^2, \quad (3.54)$$

where  $F_{1/2}(m_t) = -8m_t^2/m_h^2 + \dots$  is a triangle diagram function, see e.g. (2.17) of [90].

Why should the radion coupling be so similar to the Higgs? Before one stabilizes the radion VEV (e.g. as in Section 3.8), the radion is a modulus and has a flat potential. In the holographic 4D dual, the radion corresponds to the Goldstone boson from the spontaneous breaking of conformal symmetry by the confining dynamics at the IR scale. In other words, in the 4D theory, the radion is a **dilaton**. This is the reason why it is so similar to the SM Higgs: the Higgs is *also* a dilaton in a simple limit of the Standard Model.

In the SM the only dimensionful parameter is that of the Higgs mass,

$$V(H) = \lambda \left( H^\dagger H - \frac{v^2}{2} \right)^2. \quad (3.55)$$

In the limit when  $\lambda \rightarrow 0$ , the Standard Model thus enjoys an approximate scale invariance. If we maintain  $v \neq 0$  while taking  $\lambda \rightarrow 0$ , that is, we leave the Higgs VEV on, then:

- Electroweak breaking  $\text{SU}(2) \times \text{U}(1) \rightarrow \text{U}(1)$  gives the usual three Goldstone bosons eaten by the  $W^\pm$  and  $Z$
- The breaking of scale invariance gives an additional Goldstone boson, which is precisely the Higgs.

Indeed, the Higgs couples to the sources of scale invariance breaking: the masses of the fundamental SM particles,

$$\frac{h}{v} (m_f \bar{\Psi} \Psi + M_W^2 W_\mu W^\mu + \dots). \quad (3.56)$$

This observation leads to an interesting possibility: could one construct a complete model with no elementary scalar Higgs, but where a condensate breaks electroweak symmetry and scale invariance? Then the dilaton of this theory may have the properties of the SM Higgs. If one can reproduce the observed Higgs mass then it could be very difficult to tell the scenario apart from the SM [91].

### 3.11 Realistic Randall-Sundrum Models

While the original RS model is sometimes used as a template by LHC experiments to put bounds on KK gravitons, most theorists usually refer to RS to mean a more modern variant than the model presented thus far. In the so-called ‘realistic’ version of Randall-Sundrum, all of the Standard Model fields are allowed to propagate in the bulk [92–94]. Doing this allows one to use other features of the RS framework to address other model building issues. For example, pulling the gauge fields into the bulk can help for

grand unification, but this typically leads to unacceptably large corrections to the Peskin-Takeuchi  $S$ -parameter. One way to control this is to also allow the fermions to live in the bulk. We explain below that the bulk fermions open up a powerful new way to use the RS background to generate the hierarchies in the Yukawa matrix.

Solving the Hierarchy problem requires the Higgs to either be stuck on the IR brane or otherwise have a bulk profile that is highly peaked towards it. Allowing the fermions and gauge fields to propagate in the bulk introduces a tower of KK modes for each state. These tend to be peaked towards the IR brane and, as we learned above, are identified with bound states of the strongly coupled holographic dual. The Standard Model matter and gauge content are identified with the zero modes of the bulk fields. These carry zero KK mass and pick up small non-zero masses from their interaction with the Higgs. When boundary conditions permit them, zero mode profiles can have different types of behavior:

- Fermion zero modes<sup>9</sup> are either exponentially peaked toward the IR brane or the UV brane. The parameter controlling this behavior is the bulk mass<sup>10</sup>, see (A.40).
- Gauge boson zero modes are flat in the extra dimension, though electroweak symmetry breaking on the IR brane distorts this a bit, see (3.61).

The holographic interpretation of a Standard Model field with a bulk profile is that the SM state is **partially composite**. That is to say that it is an admixture of elementary and composite states. This is analogous to the mixing between the  $\rho$  meson and the photon in QCD. States whose profiles are peaked towards the UV brane are mostly elementary, states peaked toward the IR brane are mostly composite, and states with flat profiles are an equal admixture.

The effective 4D coupling between states depends on the overlap integral of their extra dimensional profiles. This gives a way to understand the hierarchies in the Yukawa matrices, since these are couplings to the Higgs, which is mostly localized on the IR brane [93–99]. This is a realization of the split fermion scenario<sup>11</sup> [100–103]. The zero-mode fermions that couple to the Higgs, on the other hand, can be peaked on either brane. We can see that even with  $\mathcal{O}(1)$  5D couplings, if the zero-mode fermions are peaked away from the Higgs, the  $dz$  overlap integral of their profiles will produce an exponentially small prefactor. We can thus identify heavier quarks as those whose bulk mass parameters cause them to lean towards the Higgs, while the lighter quarks are those whose bulk mass parameters cause them to lean away from the Higgs. Because the 5D couplings can be arbitrary  $\mathcal{O}(1)$  numbers, this is often called **flavor anarchy**. This scenario is sketched in Fig. 7.

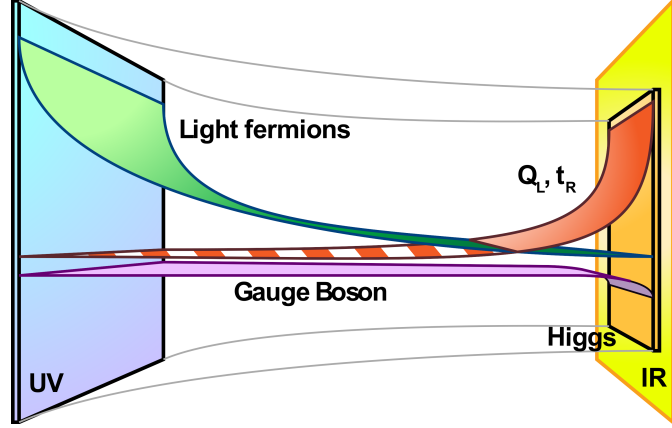
This framework tells us how to search for ‘realistic’ RS models. Unlike the original RS model, whose main experimental signature were KK gravitons decaying to SM states like leptons, the profiles of our SM fields tell us what we expect realistic RS to produce. The most abundantly produced new states are those with strong coupling, say the KK gluon. Like all of the RS KK states, this is peaked towards the IR brane. The SM field which couples the most to this state are the right-handed tops. This is because we want the tops to have a large Yukawa coupling, and the left-handed top cannot be too peaked on the IR brane or else the bottom quark—part of the same electroweak doublet—would become heavy. These KK gluons are expected to have a mass  $\gtrsim 3$  TeV, so we expect these tops to be very boosted. This suggests experimental techniques like jet substructure (see [104–106] for reviews).

There are additional features that one may add to the RS scenario to make it even more realistic. From the picture above, the electroweak gauge KK modes lean towards the IR brane where the Higgs can cause large mixing with the SM  $W$  and  $Z$ . This causes large corrections to the Peskin-Takeuchi

<sup>9</sup>One immediate concern with bulk fermions is that in 5D the basic spinor representation is Dirac rather than Weyl. Thus one does not automatically obtain a chiral spectrum of the type observed in the SM. While heavy KK states indeed appear as Dirac fermions, one may pick boundary conditions for the bulk fermion field that project out the ‘wrong chirality’ zero-mode state. See Appendix A.3.6.

<sup>10</sup>Observe that this is a manifestation of our identification of bulk masses and scaling dimension in Sec. 3.9.5.

<sup>11</sup>Note that the use of an extra dimension to explain flavor hierarchies does not require warping.



**Fig. 7:** A cartoon of the zero mode profiles of various SM particles in the ‘realistic’ RS scenario.

$T$  parameter which seems to push up the compactification scale, causing a reintroduction of tuning. A second issue is that the third generation SM fermions also have a large overlap with the Higgs and can induce a large  $Z\bar{b}b$  coupling through the neutral Goldstone. This coupling is well measured and would also require some tuning in the couplings. It turns out, however, that imposing custodial symmetry in the bulk can address both of these problems [107, 108]. The symmetry is typically gauged and broken on the IR brane so that it is holographically identified with a global symmetry of the 4D theory—just as in the SM. This introduces several new states in the theory, many of which are required to have boundary conditions that prevent zero modes.

### 3.12 A sketch of RS flavor

Let us assume that the Higgs is effectively IR brane-localized. The effective 4D Yukawa coupling between a left-handed quark doublet and a right-handed quark singlet is given by the  $\mathcal{O}(1)$  anarchic (non hierarchial) 5D Yukawa coupling multiplied by the zero-mode fermion profiles evaluated on the IR brane,  $\epsilon$ ,

$$y_i^{uj} \sim \mathcal{O}(1)_{ij} \times \epsilon_i^Q \epsilon_j^{uR}. \quad (3.57)$$

Here we have implicitly treated the Higgs boson profile as a  $\delta$ -function on the IR brane and integrated over the profiles. In the 4D mass eigenstate basis,  $y_t \sim 1$ , we can write  $\epsilon_3^{uR} \sim \epsilon_3^Q \sim 1$ . For a choice of these parameters, one may then use the bottom mass to determine the value of  $\epsilon_3^{dR}$ . This, in turn, may be used in conjunction with the CKM matrix,

$$V_{\text{CKM} \leq ij} \sim \mathcal{O} \left( \frac{\epsilon_i^Q}{\epsilon_j^Q} \right), \quad (3.58)$$

to determine the  $\epsilon$ s of lower generations and so forth. One automatically obtains a hierarchical pattern of mixing.

Neutrino zero modes, on the other hand, must be highly peaked on the UV brane. In fact, these are typically even more peaked on the UV brane than the Higgs is peaked on the IR brane. In other words, one should no longer treat the Higgs as purely brane localized<sup>12</sup> and rather as a profile which is exponentially small on the UV brane. In this limit, one can treat the right-handed neutrinos as each having a  $\delta$ -function profile on the UV brane. Even with  $\mathcal{O}(1)$  anarchic Yukawa couplings, the smallness

<sup>12</sup>This itself causes some conceptual issues since the interactions of a purely brane Higgs is incompatible with the boundary conditions required to make the fermion zero modes chiral [109].

of the Higgs profile then suppresses the neutrino mass to automatically be small. Further, since each neutrino Yukawa coupling has the same Higgs mass, one finds larger mixing than in the quark sector, as phenomenologically observed.

### 3.13 Example: the coupling of the $Z$ in RS

As a sample calculation, consider the coupling of the  $Z$  boson in RS. We first derive the effective 4D (SM) coupling of the  $Z$  in terms of the 5D parameters and then calculate the FCNC induced by the zero mode  $Z$ . In the SM the  $Z$  is, of course, flavor universal and flavor-changing coupling. Indeed, at zeroth order, RS also prevents such a FCNC since the gauge boson zero mode profile is flat and therefore universal. We will see, however, that the correction to the  $Z$  profile induces a small FCNC term.

Let us first state some results that are derived in the appendix. The localization of the normalized zero mode fermion profile is controlled by the dimensionless parameter  $c$ ,

$$\Psi_c^{(0)}(x, z) = \frac{1}{\sqrt{R'}} \left(\frac{z}{R}\right)^2 \left(\frac{z}{R'}\right)^{-c} f_c P_L \Psi_c^{(0)}(x), \quad (3.59)$$

where  $c/R$  is the fermion bulk mass and  $P_L$  is the left-chiral projection operator. Right chiral states differ by  $P_L \rightarrow P_R$  and  $c \rightarrow -c$ . We have also used the RS flavor function characterizing the fermion profile on the IR brane (larger  $f$  means larger overlap with the Higgs),

$$f_c = \sqrt{\frac{1-2c}{1-(R/R')^{1-2c}}}. \quad (3.60)$$

Each SM fermion has a different bulk mass  $c$  which according to the size of its SM Yukawa coupling. For simplicity of notation, we will simultaneously use  $c$  as the bulk mass parameter and as a flavor index rather than  $c_i$ . Further, the profile for the zero mode  $Z$  boson is

$$h_Z^{(0)}(z) = \frac{1}{\sqrt{R \log R'/R}} \left[ 1 - \frac{M_Z^2}{4} \left( z^2 - 2z^2 \log \frac{z}{R} \right) \right], \quad (3.61)$$

Starting in the canonical 5D basis where the bulk masses ( $c$  parameters) are diagonal, the zero mode fermion coupling to the zero mode  $Z$  is

$$g_{4D} Z_\mu^{(0)}(x) \bar{\Psi}_c^{(0)}(x) \gamma^\mu \Psi_c^{(0)}(x) + \dots = \int dz \left(\frac{R}{z}\right)^5 g_{5D} Z_M^{(0)}(x, z) \bar{\Psi}_c^{(0)}(x, z) \Gamma^M \Psi_c^{(0)}(x, z), \quad (3.62)$$

where  $\Gamma^M = \frac{z}{R} \gamma^M$ , the prefactor coming from the vielbein. Plugging in the profiles gives

$$g_{4D}^{cc} = g_{5D} \int_R^{R'} dz \frac{1}{R'} \left(\frac{z}{R}\right)^{-2c} f_c^2 \frac{1}{\sqrt{R \log R'/R}} \left[ 1 + \frac{M_Z}{4} \left( z^2 - 2z^2 \log \frac{z}{R} \right) \right], \quad (3.63)$$

where the  $cc$  superscripts index fermion flavor. We write  $g_{4D}^{cc} = g_{\text{SM}} + g_{\text{FCNC}}^{cc}$  in anticipation that the term in the bracket proportional to  $M_Z$  is non-universal and will contribute a FCNC. The leading term, on the other hand, gives the usual SM coupling. Performing the  $dz$  integral for that term gives

$$g_{\text{SM}} = \frac{g_5 f_c^2 (R')^{2c}}{R' \sqrt{R \log R'/R}} \frac{R'}{1-2c} \left[ 1 - \left(\frac{R}{R'}\right)^{1-2c} \right] = \frac{g_5}{\sqrt{R \log R'/R}}. \quad (3.64)$$

This is indeed flavor-universal since it is independent of  $c$  so that upon diagonalization of the zero mode mass matrix with respect to the Yukawa matrices, this contribution remains unchanged.

On the other hand, the term proportional to  $M_Z$  gives a non-universal contribution. Performing a change of variables to  $y = z/R$  and performing the  $dy$  integral gives

$$g_{\text{FCNC}}^{cc} = -g_5 \frac{(M_Z R')^2 \log R'/R}{2(3-2c)} f_c^2, \quad (3.65)$$

where we've dropped a subleading term that doesn't have the  $\log R'/R$  enhancement. Consider, for example, the coupling between a muon and an electron through the zero mode  $Z$ . The unitary transformation that diagonalizes the Yukawa mass matrix goes like  $f_i/f_j$  so that

$$g_{\text{FCNC}}^{Z_0\mu e} = \left( U^\dagger g^{ee} U \right)_{\mu e} \sim -\frac{f_e}{f_\mu} \left( \frac{f_\mu^2}{3-2c_\mu} - \frac{f_e^2}{3-2c_e} \right) (M_Z R')^2 \frac{1}{2} \log \frac{R'}{R} g_{\text{SM}}. \quad (3.66)$$

We can drop the second term since flavor anarchy requires  $f_e^2 \ll f_\mu^2$ . The result is

$$g_{\text{FCNC}}^{Z_0\mu e} = -g_{\text{SM}} \frac{(M_Z R')^2}{2(3-2c_\mu)} \log \frac{R'}{R} f_\mu f_e. \quad (3.67)$$

The observation that the coupling is suppressed by  $(M_Z R')^2$  is sometimes called the ‘**RS GIM mechanism**.’ Note that in order to do a full calculation, one must also include the non-universal contribution from KK  $Z$  bosons. These couplings do not have a  $(M_Z R')^2$  suppression, but FCNC diagrams with these KK modes are suppressed by the  $Z^{(n)}$  mass.

## 4 The Higgs from Strong Dynamics

**Further reading:** The original phenomenological Lagrangian papers lay the foundation for the general treatment of Goldstone bosons [110, 111]. See §19.6 of [112] for a slightly more pedagogical treatment that maintains much of the rigor of [110, 111], or Donoghue, *et al.* for a discussion tied closely to QCD [?]. Very readable discussions can be found in [?, ?]. For a rather comprehensive review that emphasizes the role of ‘gauge’ symmetries, see [113]. For the composite Higgs see [114] or the 2012 ICTP “School on Strongly Coupled Physics Beyond the Standard Model” [115] for a modern set of lectures and [116] for a phenomenological review. Finally, see [117, 118] for reviews of the little Higgs scenario.

For our last topic we explore models where strong dynamics at a scale  $\Lambda \sim 10$  TeV produces a light, composite Higgs. The solution to the Hierarchy problem is that there is no elementary scalar—beyond  $\Lambda$  one becomes sensitive to the underlying ‘partons’ that make up the Higgs. Through the holographic principle, we have already discussed many broad features of this paradigm in the context of warped extra dimensions above.

One key question to address is the lightness of the Higgs mass. If  $\Lambda \sim 10$  TeV, how is it that the Higgs appears at 125 GeV? By comparison, the strong coupling scale for quantum chromodynamics is  $\Lambda_{\text{QCD}} \sim \mathcal{O}(300 \text{ MeV})$  while most QCD states, such as the  $\rho$  meson and proton are at least as heavy as this<sup>13</sup>. Those who are sharp with their meson spectroscopy will quickly observe that there is a counterexample in QCD: the pions are all *lighter* than  $\Lambda_{\text{QCD}}$ , albeit by only an  $\mathcal{O}(1)$  factor.

The reason that the pions can be appreciably lighter than the other QCD states is the well-known story of **chiral perturbation theory**, a subset of the more general **nonlinear  $\Sigma$  model** ( $\text{NL}\Sigma\text{M}$ ) construction. The pions are the Goldstone bosons of the spontaneously broken  $\text{SU}(2)_L \times \text{SU}(2)_R$  flavor symmetry coming from chiral rotations of the up and down quarks. Small explicit breaking of this symmetry generates a mass for the pions so that they are **pseudo-Goldstone** modes. In the composite models that we consider in this section, we assume a similar structure where the Higgs is a pseudo-Goldstone boson of

<sup>13</sup>A better comparison is  $\Lambda = 4\pi f_\pi \sim \mathcal{O}(\text{GeV})$ , where  $f_\pi$  is the pion decay constant. ‘Typical’ QCD states such as the  $\rho$  meson have masses of at least this value,  $m_\rho \sim \Lambda$ . We explain the distinction in Section 4.3.7.

some symmetry for which  $\Lambda \approx 4\pi f$  with breaking scale  $f \approx 1$  TeV. We show that the generic composite Higgs set up still requires some tuning between the electroweak scale  $v$  and the symmetry breaking scale  $f$ . One way to generate this ‘little hierarchy’ is through the mechanism of collective symmetry breaking. We close this section by drawing connections to models of an extra dimension and by providing a phenomenological taxonomy of composite Higgs models to help clarify nomenclature.

#### 4.1 Pions as Goldstone bosons

Before exploring composite Higgs models in earnest, it is useful to review strong electroweak symmetry breaking in QCD since this gives a concrete example of the effective theory of Goldstone bosons. It is also useful because electroweak symmetry breaking in QCD formed the motivation for technicolor models that have since fallen out of favor—it is useful to see why this is, and how composite Higgs models are different from a revival of technicolor.

First, consider the Lagrangian for pure QCD: a theory of vector-like quarks and gluons, where ‘vector-like’ mean the left- and right-handed quarks come in conjugate representations,

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} \bar{q}(i\not{D} - m)q. \quad (4.1)$$

This is a theory which becomes strongly coupled and confines at low energies, leading to a spectrum of composite states. This makes it a good template for our own explorations into compositeness. We can already guess that at low energies the effective theory is described by Goldstone bosons, the pions. In anticipation, we examine the global symmetries of the theory.

We focus only on the three lightest quarks with masses  $m_i \ll \Lambda_{\text{QCD}}$ . In the chiral limit,  $m \rightarrow 0$ , the physical quarks are Weyl spinors and have an enhanced  $U(3)_L \times U(3)_R$  global flavor symmetry acting separately on the left- and right-handed quarks,

$$q_L^i \rightarrow (U_L)^i_j q_L^j \quad (4.2)$$

$$q_R^i \rightarrow (U_R)^i_j q_R^j. \quad (4.3)$$

One may write the currents for this global symmetry. For compactness we move back to Dirac spinors and write in terms of the vector ( $U_L = U_R$ ) and axial ( $U_L = U_R^\dagger$ ) transformations:

$$(j_V^a)^\mu = \bar{q}\gamma^\mu T^a q \quad (j_A^a)^\mu = \bar{q}\gamma^\mu \gamma_5 T^a q \quad (4.4)$$

$$(j_V)^\mu = \bar{q}\gamma^\mu q \quad (j_A)^\mu = \bar{q}\gamma^\mu \gamma_5 q, \quad (4.5)$$

where the  $T^a$  are the generators of SU(3). We can identify  $j_V$  with baryon number, which is conserved in QCD, and we note that  $j_A$  is anomalous so that it is not a good symmetry and we don’t expect to see it at low energies<sup>14</sup>. The vectorial SU(3), with current  $j_V^a$ , is precisely the symmetry of Gell-Mann’s eightfold way and can be used to classify the light hadrons. What do we make of the axial SU(3),  $j_A^a$ ?

Phenomenologically we can observe that the axial SU(3) is not a symmetry of the low energy spectrum, otherwise we would expect a parity doubling of all the ‘eightfold way’ multiplets. There is one way out: this symmetry must be spontaneously broken. What could possibly enact this breaking in a theory with no Higgs boson? It turns out that QCD itself can do the job! We assume that the axial SU(3)<sub>A</sub> is broken spontaneously by a quark–anti-quark condensate,

$$\langle \bar{q}q \rangle = \langle \bar{q}_L^i q_{Ri} + \text{h.c.} \rangle \neq 0 \quad (4.6)$$

<sup>14</sup>What happens to this symmetry at low energies is rather subtle and was known as the ‘U(1) problem.’ There is a lot more to the story than simply saying that the axial U(1) is anomalous and so does not appear at low energies. One can construct a current out of  $j_A$  and a Chern-Simons (topological) current that is anomaly-free and spontaneously broken. This current indeed has a Goldstone pole. However, Kogut and Susskind showed that this current is not gauge invariant. There are actually two Goldstone bosons that cancel in any gauge invariant operator [119].

in such a way that the vector  $SU(3)_V$  is preserved. This is the unique combination that preserves Lorentz invariance and breaks  $SU(3)_A$ . By dimensional analysis, this ‘chiral condensate’ takes the form  $\langle \bar{q}^i q_j \rangle \sim \delta_{ij} \Lambda_{\text{QCD}}^3$ . Given that QCD is strongly interacting in the IR, the existence of this non-trivial vacuum condensate should not be surprising and is indeed supported by lattice calculations. However, the exact mechanism by which this condensate forms is non-perturbative and not fully understood. This also gives a robust prediction: we should have eight pseudoscalar Goldstone bosons as light excitations. These are precisely the pions, kaons, and  $\eta$ . Because  $SU(3)_A$  is only a symmetry in the chiral  $m \rightarrow 0$  limit, these are not exactly Goldstone bosons as the symmetry is explicitly broken by the quark masses and electromagnetism. However, because this explicit breaking is small relative to  $\Lambda_{\text{QCD}}$ , these excitations are still very light  $m_\pi \ll \Lambda_{\text{QCD}}$  and are often referred to as **pseudo-Goldstone bosons** (sometimes pseudo-Nambu–Goldstone bosons, pNGB, in the literature).

Note that the electroweak group sits inside the QCD flavor symmetry<sup>15</sup>,

$$SU(3)_L \times SU(3)_R \times U(1)_B \supset SU(2)_L \times U(1)_Y. \quad (4.7)$$

We can see this since an  $SU(3)_L$  fundamental contains  $(u_L, d_L, s_L)$ , where the first two components form the usual  $SU(2)_L$  first generation quark doublet. In this way,  $SU(2)_L$  is simply the upper left  $2 \times 2$  component of the  $SU(3)_L$  generators. Similarly, hypercharge is a combination of the diagonal generators,

$$Y = T_{R3} + \frac{B}{2}. \quad (4.8)$$

We say that the electroweak group is **weakly gauged** with respect to low energy QCD. By this we mean that the gauge couplings are perturbative in all energy scales of interest. This weak gauging is a small explicit breaking of the QCD flavor symmetries and accounts for the mass splitting between the  $\pi^0$  and  $\pi^\pm$ .

## 4.2 A farewell to technicolor

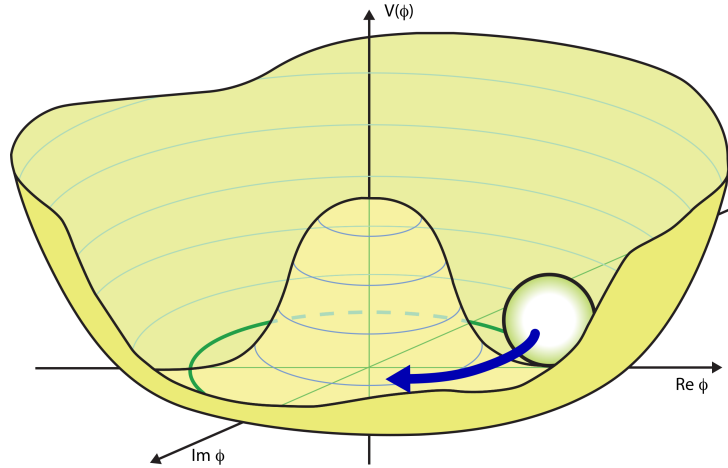
Because of (4.7), the spontaneous breaking of  $SU(3)_A$  by the chiral condensate  $\langle \bar{q}q \rangle$  also breaks electroweak symmetry. This is an important observation: even if there were no Higgs boson, electroweak symmetry would still be broken and  $W$  and  $Z$  bosons would still be massive, albeit with a much smaller mass. This mass comes from ‘eating’ part of the appropriately charged pseudo-Goldstone bosons. We will see this in slightly more detail below. Readers unfamiliar with this story are encouraged to follow the treatment in [120].

The observation that strong dynamics can—and indeed, *does* in QCD—break electroweak symmetry led to the development of **technicolor** theories where the SM is extended by a confining sector [121–125]. Note that by the holographic interpretation of extra dimensions, this type of electroweak symmetry breaking is analogous to the RS scenario where a brane-localized Higgs picks up a VEV. The large hierarchy between the Planck and electroweak scales is then understood to be a result of dimensional transmutation. The simplest constructions of these models, however, suffer from several issues. These include the requirement for an additional mechanism to generate fermion masses [126, 127] and generically large deviations in flavor and electroweak precision observables [127–129]. However, the nail in the coffin for most of these models is observation of the Higgs boson at 125 GeV, much lighter than the compositeness scale. Such a state—even if it is not the *Standard Model* Higgs—is very difficult to explain in the context of these models.

As such, even though the models we consider here encode strong dynamics, they are distinct from the pre-Higgs technicolor strong dynamics of the past. All of these models require a Higgs. We will borrow from the above story, however, the importance of the effective theory of pseudo-Goldstone bosons. By identifying the Higgs as one of these pion-like states, we can explain its lightness.

<sup>15</sup>It has to be true that the electroweak gauge group sits in the full QCD global symmetry group in order for some of the quarks to have non-trivial electroweak charges.





**Fig. 8:** Cartoon of the Goldstone excitation for a ‘Mexican hat’ potential. Image from [?].

### 4.3 Chiral perturbation theory

In this section we review the main framework for describing Goldstone bosons of chiral symmetry breaking, known as **chiral perturbation theory**. Many of the results highlight general principles that appear in any theory of Goldstone bosons, known as **nonlinear sigma models**. A completely general treatment of spontaneously broken global symmetries is captured in the the so-called Callan-Coleman-Wess-Zumino (CCWZ) construction, which we present in Appendix B.

The importance of having a Lagrangian theory of Goldstone bosons is clear from the success of SM predictions before the Higgs discovery. Naïvely, one might wonder how we knew so much about the Standard Model before the Higgs discovery—isn’t the Higgs a very central piece to the theory? As we saw above, the key feature is actually electroweak symmetry breaking: whether or not there is a Higgs, one always has the Goldstone bosons which are eaten by the  $W^\pm$  and  $Z$  to become massive. It is this nonlinear sigma model that pre-Higgs experiments had studied so carefully. The discovery of the Higgs is a statement that the nonlinear sigma model is UV completed into a linear sigma model.

#### 4.3.1 Framework

We begin with the concrete example of low-energy QCD that we described above. Given that the chiral condensate  $\langle \bar{q}q \rangle$  breaks  $SU(3)_A$ , we proceed to write down the effective theory describing the interaction of the resulting Goldstone bosons. Let us write  $U_0$  to refer to the direction in field space associated with the chiral condensate,  $U_0 \sim \langle \bar{q}q \rangle$ . This transforms as a **bifundamental** with respect to  $SU(3)_L \times SU(3)_R$ ,

$$U(x) \rightarrow U_L U(x) U_R^\dagger, \quad (4.9)$$

where  $U_L$  and  $U_R$  are the transformation matrices under the  $SU(3)_L$  and  $SU(3)_R$  respectively. The observation that  $SU(3)_A$  is broken corresponds to  $U_0 = \mathbb{1}$ . Note that this indeed preserves the  $SU(3)_V$  transformations  $U_L = U_R$ .

We now consider the fluctuations  $U(x)$  about  $U_0$ —these are what we identify with the Goldstone bosons. Recall the picture of spontaneous symmetry breaking through the ‘Mexican hat’ potential in Fig. 8. The action of an unbroken symmetry does not affect the VEV (represented by the ball), while broken symmetries shift the VEV along the vacuum manifold. This gives an intuitive picture of how to identify the Goldstone modes:

1. Identify a convenient VEV,  $U_0$

2. Act on that VEV with the broken group elements
3. Promote the transformation parameter to a field, identify these with the Goldstones.

For the chiral Lagrangian, our broken symmetries are those for which  $U_L = U_R^\dagger$ . Writing  $U_L = \exp(i\epsilon^a T^a)$ , we act on  $U_0 = \mathbb{1}$ ,

$$e^{i\epsilon^a T^a} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} e^{i\epsilon^a T^a} = e^{2i\epsilon^a T^a}. \quad (4.10)$$

We now promote the transformation parameter  $\epsilon^a$  to Goldstone fields,  $\epsilon^a \sim \pi^a(x)$ . Since  $\epsilon^a$  is dimensionless, in order for  $\pi^a$  to have canonical scaling dimension we should rescale by the **decay constant**<sup>16</sup>  $f$ . We may understand the physical meaning of  $f$  if we recall Fig. 8, since we want  $\epsilon$  to be an angle that parameterizes the position along the vacuum circle: the Goldstone is a periodic variable with period  $2\pi f$ , so that  $f$  is identified with the value of the symmetry breaking VEV. The angle  $\epsilon$  is then  $\pi(x)/f$ . We thus promote  $\epsilon^a \rightarrow \pi(x)/f$  so that we may define the field  $U(x)$ ,

$$U(x) = e^{i\frac{\pi^a(x)}{f} T^a} U_0 e^{i\frac{\pi^a(x)}{f} T^a} = e^{2i\frac{\pi^a(x)}{f} T^a}. \quad (4.11)$$

We now have an object  $U(x)$  which packages the Goldstone fields,  $\pi^a(x)$ . Note that  $U(x)$  transforms linearly under the full  $SU(3)_L \times SU(3)_R$  group,  $U(x) \rightarrow U_L U(x) U_R^\dagger$ , but the fields that actually describe the low energy spectrum are related in a non-trivial way to  $U(x)$ .

### 4.3.2 How pions transform

We can determine the transformation of the pions  $\pi^a$  by using the transformation of the linear field  $U(x)$ . Under the  $SU(3)_V$  (unbroken) symmetry,  $U_L = U_R = U_V$ , we have

$$U(x) \rightarrow U_V U(x) U_V^\dagger = U_V \left( 1 + 2i \frac{\pi^a(x)}{f} T^a + \dots \right) U_V^\dagger, \quad (4.12)$$

where we can see from the first term that  $\pi^a(x) T^a \rightarrow U_V \pi^a(x) T^a U_V^\dagger$ . In other words,  $\pi^a(x)$  transforms *linearly* under the unbroken symmetry. Note that the higher order terms also obey this by trivially inserting factors of  $U_V^\dagger U_V = \mathbb{1}$ . Indeed, we expected this result because we know that Gell-Mann's eightfold way is precisely a realization of  $SU(3)_V$ , so our pions must transform as octets.

Things are not as simple for the broken symmetry,  $U_L = U_R^\dagger = U_A$ . In this case the transformation is

$$U(x) \rightarrow U_A U(x) U_A \equiv e^{2i\frac{\pi'^a(x)}{f} T^a}. \quad (4.13)$$

In this case the pion does *not* transform in a nice, linear way<sup>17</sup>. Unlike the above case, there is no sense in which this looks like  $\pi^a(x) T^a \rightarrow U_A \pi^a(x) T^a U_A^\dagger$ . The best we can do is say that we have moved  $U_0$  to a new point on the vacuum manifold, which we parameterize by an angle  $2\pi'^a(x)/f$ . The transformation  $\pi^a(x) \rightarrow \pi'^a(x)$  is nonlinear. To leading order,

$$1 + 2i\frac{\pi'^a(x)}{f} T^a = (1 + ic^a T^a) \left( 1 + 2i\frac{\pi^a(x)}{f} T^a \right) (1 + ic^a T^a) \quad (4.14)$$

<sup>16</sup>The name comes from identifying the appearance of this factor in the matrix element for pion decays, e.g.  $\langle 0 | \bar{u} \gamma^\mu \gamma_5 d | \pi^- \rangle \equiv i f p^\mu$ .

<sup>17</sup>This may seem confusing since  $U(x)$  transforms as a bifundamental under  $SU(3)_L \times SU(3)_R$ . However, components of  $U(x)$  are not independent due to the nonlinear constraints of being unitary and having unit determinant.

so that

$$\pi'^a(x)T^a = \pi^a(x)T^a + fc^aT^a. \quad (4.15)$$

In other words, to leading order the pion shifts  $\pi^a \rightarrow fc^a$ . This shift symmetry in the nonlinear realization is precisely why the pion is massless; the only non-trivial pion Lagrangian terms must carry derivatives.

**Coset space description.** In anticipation of the more general CCWZ construction, let us restate the above arguments in a more compact way. The symmetry breaking pattern is the coset  $SU(3)_L \times SU(3)_R / SU(3)_V$ . Using the notation above, this means that group elements of the full symmetry  $U_{L,R}$  can be written as a product of elements of the unbroken group  $U_V$  and the [left] coset  $U_A \in SU(3)_L \times SU(3)_R / SU(3)_V$ ,

$$U_L = U_A U_V \qquad U_R = U_A^\dagger U_V. \quad (4.16)$$

One can check that this matches the above cases when one sets  $U_A = \mathbb{1}$  or  $U_V = \mathbb{1}$ . The general transformation of the linear packaging of the pions,  $U(x) = \exp(2i\pi^a(x)T^a/f)$ , is

$$U(x) \rightarrow U_A \left( U_V U(x) U_V^\dagger \right) U_A. \quad (4.17)$$

From here it is clear that  $SU(3)_V$  is realized linearly while  $SU(3)_A$  is realized non-linearly.

**$SU(3)_A$  is not a subgroup of  $SU(3)_L \times SU(3)_R$ .** While one can divide the algebra of  $SU(3)_L \times SU(3)_R$  into axial and vector generators, one should note that there is no such thing as an ‘axial subgroup’ of  $SU(3)_L \times SU(3)_R$ . One can check that the commutation relations of axial generators include vector generators so that the  $SU(3)_A$  algebra doesn’t close by itself.

### 4.3.3 Lagrangian description

Thus far we have found a convenient way to package the Goldstone fields  $\pi^a(x)$  into a linear realization of the full  $SU(3)_L \times SU(3)_R$  symmetry. We would like to write down a Lagrangian describing the dynamics of the Goldstones. Our strategy will be to write the lowest order terms in  $U(x)$  that are  $SU(3)_L \times SU(3)_R$  invariant and then expand  $U(x)$  in Goldstone excitations about  $U_0$ . One can see that many invariants, such as  $U(x)^\dagger U(x)$ , are independent of the Goldstones. In fact, only derivative terms contain the Goldstone fields. This is consistent with our argument that Goldstones must have derivative couplings. The lowest order non-trivial term is

$$\mathcal{L} = \frac{f^2}{4} \text{Tr} \left[ \left( \partial^\mu U^\dagger(x) \right) \partial_\mu U(x) \right] \quad (4.18)$$

The pre-factor is fixed by expanding  $U(x) = 1 + 2i\frac{\pi^a(x)}{f}T^a + \dots$  and ensuring that the kinetic term for  $\pi^a(x)$  is canonically normalized. We have used the normalization that  $\text{Tr} T^a T^b = \frac{1}{2}\delta^{ab}$ . The higher order terms in the expansion of  $U$  yield a series of non-renormalizable pion–pion interactions.

Next we weakly gauge the electroweak group. Recall that this sits in  $SU(3)_L \times SU(3)_R \times U(1)_B$ . The left- and right-chiral quarks are fundamentals under  $SU(3)_L$  and  $SU(3)_R$  respectively and have baryon number  $1/3$ . This information, combined with knowing how  $SU(2)_L$  sits in  $SU(3)_L$  and (4.8), determines the quantum numbers of the linear field  $U(x)$ , which transforms as a  $\bar{\mathbf{3}} \times \mathbf{3} \times 0$  under  $SU(3)_L \times SU(3)_R \times U(1)_B$ . To ‘turn on’ the electroweak gauge interactions, we simply promote derivatives to covariant derivatives  $\partial_\mu \rightarrow D_\mu$  where

$$D_\mu U(x)^i_j = \partial_\mu U(x)^i_j - igW_\mu^a \frac{1}{2} (\tau^a)^i_k U(x)^k_j + ig'B_\mu \frac{1}{2} (T_R^3)^k_i U(x)^i_k. \quad (4.19)$$

We have written the  $SU(2)_L$  generators as

$$\frac{1}{2}\tau^a = \frac{1}{2} \left( \begin{array}{c|c} \tau^a & 0 \\ \hline 0 & 0 \end{array} \right) \subset SU(3)_L. \quad (4.20)$$

Promoting  $\partial_\mu \rightarrow D_\mu$  in (4.18) yields

$$\mathcal{L} = \frac{f^2}{4} \text{Tr} \left| \left( \partial_\mu - \frac{ig}{2} W_\mu^a(x) \tau^a \right) U(x) \right|^2 + \dots, \quad (4.21)$$

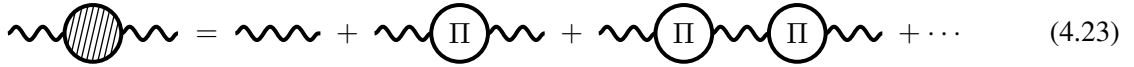
where we leave the similar term with  $B_\mu(x)$  implicit.

#### 4.3.4 Electroweak symmetry breaking

One may check that (4.21) has terms that are linear in  $W(x)$  such as  $\frac{g}{2} f W_\mu^+(x) \partial^\mu \pi^-(x) + \text{h.c.}$  This is precisely a mixing term between the  $\pi^+(x)$  and the  $W_\mu^+(x)$ . In other words, the  $W$  has eaten the Goldstone boson to pick up a longitudinal polarization. This is precisely electroweak symmetry breaking at work. Note that similar terms mixing the  $W_\mu^3$  and  $B_\mu$  with the  $\pi^0$ . As usual, the masses of the heavy gauge bosons come from the gauge fields acting on the  $U_0$  ‘VEV’ part of  $U(x)$ , the resulting spectrum is

$$\Delta\mathcal{L} = \frac{g^2 f^2}{4} W^+ W^- + \frac{g^2 + g'^2}{4} f^2 \frac{Z^2}{2}. \quad (4.22)$$

The characteristic mass scale is 100 MeV, much smaller than the actual  $W$  and  $Z$  since most of the mass contribution to those fields comes from the Higgs VEV. Diagrammatically, we can imagine the mixing as follows:



$$\text{wavy line with shaded circle} = \text{wavy line} + \text{wavy line with } \Pi \text{ circle} + \text{wavy line with } \Pi \text{ circles} + \dots \quad (4.23)$$

We have parameterized the strong dynamics in terms of a momentum-dependent form factor  $\Pi(q^2)$ . What the  $W$  boson is really coupling to is the  $SU(2)_L$  current formed from the quarks,



$$\mu \text{ wavy line with } \Pi \text{ circle } \nu = \mu \text{ wavy line with QCD triangle } \nu \quad (4.24)$$

where the  $W$  bosons are coupling to quarks which then interact strongly with one another. In other words,

$$i\Pi_{\mu\nu}(q) = \langle J_\mu^+(q) J_\nu^-( -q) \rangle. \quad (4.25)$$

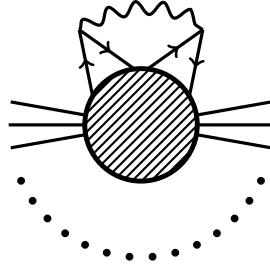
The QCD corrected  $W$  propagator  $\Delta_{\mu\nu}(q)$  from resumming the diagrams in (4.23) is

$$\Delta_{\mu\nu}(q) = \frac{-i}{q^2 - g^2 \Pi(q^2)/2} \quad \Pi_{\mu\nu}(q) = \left( \eta_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) \Pi(q^2). \quad (4.26)$$

The observation that a charged pion has been ‘eaten’ to make the  $W$  massive is the statement that  $\Pi_{\mu\nu}(q^2)$  has a zero-momentum pole. Indeed,  $\langle 0 | J_\mu^+ | \pi^-(p) \rangle = i f_\pi p_\mu / \sqrt{2}$ . The QCD blobs in (4.23) also encode, however, the effects of heavier resonances and has poles at the masses of these states. In the ‘large  $N$ ’ limit (large number of colors) one may write the current-current correlation function as a sum of resonances [?, ?, 130],

$$\left( \eta_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) \Pi(q^2) = (q^2 \eta_{\mu\nu} - q_\mu q_\nu) \sum_n \frac{f_n^2}{q^2 - m_n^2}, \quad (4.27)$$

where the Goldstone pole appears for  $m_0 = 0$ .



**Fig. 9:** ‘Cat diagram’ adapted from [131]. Despite the silly appearance, the key point is that the photon couples to the electric current  $J_\mu = e\bar{\Psi}\gamma_\mu\Psi$  (‘ears’) formed from interactions with fundamental quarks in the strongly coupled sector. The ‘whiskers’ are the pseudo-Goldstone external states when expanding the  $U(x)$  field in (4.28). The contribution to the charged meson masses come from the ‘two whisker’ diagram.

#### 4.3.5 Electromagnetic mass splitting

In addition to the spontaneous chiral symmetry breaking by strong dynamics, the  $SU(3)_L \times SU(3)_R$  group is also broken *explicitly* from the gauging of  $U(1)_{EM} \subset SU(3)_V$ . The neutral Goldstones (pions, kaons, and the  $\eta$ ) are unaffected by this. The charged Goldstones, on the other hand, pick up masses from photon loop diagrams of the form in Fig. 9. These diagrams contribute to an operator that gives a shift in the [pseudo-]Goldstone mass,

$$\Delta\mathcal{L} \sim e^2 \text{Tr} \left[ QU(x)^\dagger QU(x) \right], \quad (4.28)$$

where  $Q = \frac{1}{3}\text{diag}(2, -1, -1)$  is the matrix of quark electric charges. Since the electromagnetic force does not distinguish between the down and strange quarks, this diagram gives an equal shift to both the charged pions (e.g.  $u\bar{d}$ ) and kaons (e.g.  $u\bar{s}$ ). Since the up and anti-down/strange quark have the same charge, the bound state is more energetic than the neutral mesons and we expect the shift in the mass-squared to be positive [131, 132]. Note that the contribution to the charged pion mass is quadratically sensitive to the chiral symmetry breaking scale, though it is also suppressed by the smallness of  $\alpha_{EM}$ .

#### 4.3.6 Explicit breaking from quark spectrum

One can add quark masses that constitute a small ( $m_q \ll \Lambda_{QCD}$ ) explicit breaking of the global symmetry and generate small masses to the pseudo-Goldstone bosons. One can write this as a spurion  $M = \text{diag}(m_u, m_d, m_s)$  which has the same quantum numbers as  $U(x)$ . One can add these terms to the effective Lagrangian by forming the appropriate global symmetry group invariant. In particular, we add to the Lagrangian

$$\Delta\mathcal{L} \sim \text{Tr} [MU(x)] \sim \text{Tr} \left[ M \left( \frac{\pi^a(x)}{f} T^a \right)^2 \right] + \dots \quad (4.29)$$

In the limit where  $m_u = m_d$  and ignoring the electromagnetic splitting above, one may identify the masses for the pions, kaons, and  $\eta$  (different components of  $\pi^a$ ) to derive the Gell-Mann–Okubo relation,

$$m_\eta^2 + m_\pi^2 = 4m_K^2. \quad (4.30)$$

#### 4.3.7 NDA: When the theory breaks down

Finally, let us note that the effective Lagrangian for pions is non-renormalizable, so we should say something about the cutoff for this theory. At tree-level, the two-to-two scattering of pions with characteristic

momentum  $p$  goes like  $p^2/f^2$  from (4.18). Using naïve dimensional analysis (NDA) [133–135], we see that the loop contributions go like

$$\text{Loop Diagram} \sim \int d^4k \left( \frac{p^2 k^2}{f^4} \right) \frac{1}{k^4} \sim \frac{\Lambda^2 p^2}{16\pi^2 f^4} \sim \text{Tree Diagram} \times \frac{\Lambda^2}{16\pi^2 f^2}, \quad (4.31)$$

where we have used the shift symmetry (the full  $SU(3)^2$  group structure) to tell us that at the numerator of the integrand carries at least two powers of the external momenta. Validity of our loop expansion thus requires that  $\Lambda \sim 4\pi f \sim \text{GeV}$ , and this is indeed the scale at which additional QCD states appear. Note that this cutoff, based on perturbativity of the  $1/f$  couplings in the chiral Lagrangian, is slightly different from  $\Lambda_{\text{QCD}} \sim \mathcal{O}(300 \text{ MeV})$ , which is the scale where  $\alpha_s$  becomes non-perturbative.

Indeed, this UV behavior of the theory of Goldstones is one of the reasons why we expected either the Higgs or something new to be manifest at the LHC: the SM without a Higgs is simply a nonlinear sigma model. By the Goldstone equivalence theorem, the scattering cross section for longitudinal  $W$  boson scattering grows linearly with the center of mass energy. In order to maintain unitarity, one requires that either there is a Higgs boson (a linearization of the nonlinear sigma model) or that the theory becomes strongly coupled so that higher order terms can cancel the unphysical behavior.

#### 4.3.8 NDA: Characteristic couplings

To show the power of NDA, let's consider the generic behavior of a strongly coupled theory beyond the Goldstone modes. At the level of dimensional analysis, there is one relevant mass scale: the mass of the lowest non-Goldstone resonances,  $m_\rho$ , where we use the  $\rho$  meson as an example. Let us identify the separation between the mass of the  $\rho$  and the compositeness scale  $f$  with the parameter  $g_\rho = m_\rho/f$  that describes the coupling of  $\rho$  to the strong sector.

For a strong sector field  $\phi$ , define the dimensionless combinations

$$x = g_\rho \frac{h}{m_\rho} \qquad y = \frac{\partial}{m_\rho}. \quad (4.32)$$

We'd like to build an NDA Lagrangian to estimate the size of couplings. We start by writing some dimensionless function  $f(x, y)$ . In order to obtain the correct mass dimension of a Lagrangian, we further define  $F(x, y) = m_\rho^4 f(x, y)$ . This function is assumed to contain a kinetic term,

$$F(x, y) \supset m_\rho^4 x^2 y^2 = g_\rho^2 \mathcal{O}(\partial^2, \phi^2). \quad (4.33)$$

We see that we have to rescale by  $g_\rho^{-2}$  to obtain a canonically normalized Lagrangian,

$$L = \frac{1}{g_\rho^2} F(x, y) = \frac{m_\rho^4}{g_\rho^2} f(x, y) = m_\rho^2 f^2 f(x, y). \quad (4.34)$$

As an example that is useful below, let us use this to determine the expected size of a quartic coupling of strong sector fields. This comes from the  $\mathcal{O}(x^4)$  term in the expansion of  $f(x, y)$  so that

$$L \supset \frac{m_\rho^2}{g_\rho^2} g_\rho^4 \frac{\phi^4}{m_\rho^4} = g_\rho^2 \phi^4. \quad (4.35)$$

Thus we expect the quartic coupling of the  $\phi$  to go like  $g_\rho^2 \sim m_\rho^2/f^2$ .

## 4.4 Composite, pseudo-Goldstone Higgs

The main idea for composite pseudo-Goldstone Higgs models is that the Higgs mass parameter is protected against quadratic quantum corrections up to the compositeness scale because it is a pseudo-Goldstone boson. Above the scale of compositeness, it is simply not an elementary scalar. This should be contrasted with the solutions to the Hierarchy problem already discussed:

- SUPERSYMMETRY: due to the extended spacetime symmetry, there is a cancellation of the quadratic corrections through the introduction of different-spin partners.
- TECHNICOLOR/HIGGS-LESS: there is no elementary Higgs and electroweak symmetry breaking proceeds through a Fermi condensate. This is now excluded.
- WARPED EXTRA DIMENSIONS: the Higgs itself is a composite state so that above the compositeness scale it no longer behaves like a fundamental scalar. However, there is no explanation for why the Higgs is lighter than the confinement scale.

Note, in particular, that the composite Higgs scenario that we’re interested in is distinct from technicolor: the pseudo-Goldstone nature of the Higgs is an explanation for why the Higgs mass is so much lighter than the other bound states in the strongly coupled sector.

Goldstone bosons, however, behave very differently from the Standard Model Higgs. We saw that Goldstone bosons have derivative couplings owing to their shift symmetry. The Higgs, on the other hand, has Yukawa couplings and the all important electroweak symmetry-breaking potential. Our goal in this section is to see how to construct a theory of Goldstones which can produce a Higgs particle that has all of the required couplings of the SM Higgs.

We shall closely follow the discussion in [114] and refer the reader there for further details and references.

#### 4.4.1 The framework

Start with a large global symmetry group  $G$ , analogous to the ‘large’  $SU(3)_L \times SU(3)_R$  global symmetry of low energy QCD. We will break this symmetry in two ways:

1. We assume that the strong dynamics spontaneously breaks  $G$  to a subgroup  $H_{\text{global}}$ . This is analogous to chiral symmetry breaking in QCD,  $SU(3)_L \times SU(3)_R \rightarrow SU(3)_V$ .
2. In addition to this, we will explicitly break  $G$  by weakly gauging a subgroup  $H_{\text{gauge}}$  which contains the SM electroweak group  $SU(2)_L \times U(1)_Y$ . This is analogous to the gauging of  $U(1)_{\text{EM}}$ .

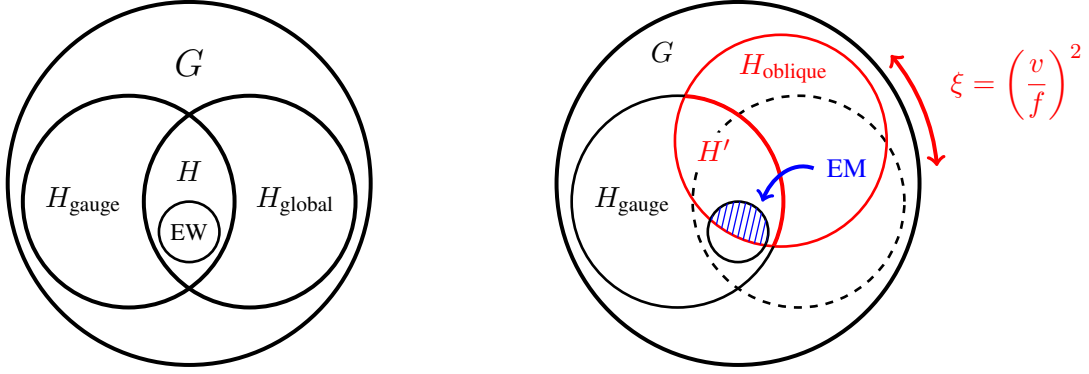
We assume that the SM electroweak group is a subgroup of  $H = H_{\text{gauge}} \cup H_{\text{global}}$  so that it is gauged and preserved by the strong dynamics. This is shown on the left of Fig. 10. This results in  $\dim H_{\text{gauge}}$  transverse gauge bosons and  $(\dim G - \dim H_{\text{global}})$  Goldstone bosons. The breaking  $G \rightarrow H_{\text{global}}$  also breaks some of the gauge group so that there are a total of  $(\dim H_{\text{gauge}} - H)$  massive gauge bosons and  $(\dim G - \dim H_{\text{global}}) - (\dim H_{\text{gauge}} - \dim H)$  ‘uneaten’ massless Goldstones.

Now we address the white elephant of the Higgs interactions—can we bequeath to our Goldstone bosons the necessary non-derivative interactions to make one of them a realistic Higgs candidate? This is indeed possible through **vacuum misalignment**, which we illustrate on the right of Fig. 10. The gauging of  $H_{\text{gauge}}$  gives loop-level corrections to the dynamical symmetry breaking pattern since this is an explicit breaking of the global symmetry. This is analogous to how the gauged  $U(1)_{\text{EM}}$  splits the masses of the charged and neutral pions through a photon loop. Loops of SM gauge bosons can generate an electroweak symmetry breaking potential for the Higgs. We illustrate this below.

One key point here is that since the Higgs potential is generated dynamically through SM gauge interactions, the electroweak scale  $v$  is distinct from the  $G \rightarrow H$  symmetry breaking scale  $f$ . The ‘angle’

$$\xi = \left(\frac{v}{f}\right)^2 \quad (4.36)$$

parameterizes this separation of scales and quantifies the degree of vacuum misalignment. Note that this is a separation of scales which does *not* exist in technicolor and is the key to parameterizing how the



**Fig. 10:** Pattern of symmetry breaking. (LEFT, tree level) Strong dynamics breaks  $G \rightarrow H_{\text{global}}$  spontaneously, while  $H_{\text{gauge}} \subset G$  is explicitly broken through gauging. The unbroken group  $H = H_{\text{gauge}} \cap H_{\text{global}}$  contains the SM electroweak group,  $SU(2)_L \times U(1)_Y$ . (RIGHT, loop level) Vacuum misalignment from SM interactions shifts the unbroken group  $H \rightarrow H'$  and breaks the electroweak group to  $U(1)_{\text{EM}}$ . The degree of misalignment is parametrized by  $\xi$ , the squared ratio of the EWSB VEV to the  $G \rightarrow H$  VEV. Adapted from [?].

Higgs remains light relative to the heavier resonances despite not being a ‘true’ Goldstone boson. The limits  $\xi \rightarrow 0$  and  $\xi \rightarrow 1$  correspond to the SM (heavy states completely decoupled) and technicolor, respectively. We note that this parameter is also a source of tuning in realistic composite Higgs models. Once the pseudo-Goldstone Higgs state is given non-derivative interactions, these interactions generically introduce quadratic divergences at loop level which would lead to an expected  $\mathcal{O}(1\%)$  tuning. To avoid this, one needs to introduce a smart way of dealing with these explicit breaking terms called **collective symmetry breaking** which we discuss below. First, however, we focus on the effects of gauge bosons on the Higgs potential.

We have the following constraints for picking a symmetry breaking pattern:

1. The SM electroweak group is a subgroup of the unbroken group,  $SU(2)_L \times U(1)_Y \subset H$ . In fact, it is better to have the full custodial  $SU(2)_L \times SU(2)_R \cong SO(4)$  group embedded in  $H$  since this will protect against large contributions to the  $\rho$ -parameter.
2. There is at least one pseudo-Goldstone boson with the quantum numbers of the SM Higgs. To protect the  $\rho$ -parameter, it is better to have a  $(\mathbf{2}, \mathbf{2})$  under the custodial group.

At this point we have said nothing about the SM fermions. These, too, will have to couple to the strong sector to generate Yukawa couplings with the Higgs. We show below that a reasonable way to do this is to allow the SM fermions to be **partially composite**, a scheme that we had already seen in the holographic interpretation of the RS scenario. Indeed, extra dimensions provide a natural language to construct composite Higgs models.

#### 4.4.2 Minimal Composite Higgs: set up

We now consider an explicit example, the **minimal composite Higgs model**, which was explored in [136, 137] using the intuition from the RS framework. Following the guidelines set above, we would like to choose  $H_{\text{global}} = SO(4)$ , the custodial group which is the minimal choice to protect the  $\rho$ -parameter. However, the  $SO(4) = SU(2)_L \times SU(2)_R$  charge assignments don’t give the correct  $U(1)_Y$  charges, as is well known in left-right symmetric models. Thus our ‘minimal’ choice for  $H_{\text{global}}$  requires an additional  $U(1)_X$  so that one may include hypercharge in the unbroken group,  $H$ ,

$$Y = (T^R)^3 + X. \quad (4.37)$$



We then choose  $G = \text{SO}(5) \times \text{U}(1)_X$  and introduce a linear field  $\Sigma$  that is an  $\text{SO}(5)$  fundamental and uncharged under  $\text{U}(1)_X$ . Note that we can ignore the  $\text{U}(1)_X$  charge in our spontaneous symmetry breaking analysis since it's really just 'coming along for the ride' at this point.  $\Sigma$  acquires a VEV to break  $\text{SO}(5) \rightarrow \text{SO}(4)$ ,

$$\langle \Sigma \rangle = (0, 0, 0, 0, 1)^T. \quad (4.38)$$

This is analogous to the QCD chiral condensate. We can now follow the intuition we developed with chiral perturbation theory. The Goldstone bosons of this breaking are given by transforming this VEV by the broken generators. A useful parameterization of the four broken generators is

$$T_{ij}^{\hat{a}} = \frac{i}{\sqrt{2}} \left( \delta_i^k \delta_j^5 - \delta_j^k \delta_i^5 \right), \quad (4.39)$$

where  $\hat{a} \in \{1, \dots, 4\}$ . We refer to the unbroken generators with an undecorated index:  $T^a$ . The  $\text{SO}(5)$  group element that acts non-trivially on the VEV,  $\exp(ih^{\hat{a}}T^{\hat{a}}/f)$ , can be written in terms of sines and cosines by separately summing the odd and even terms of the exponential. The linear field  $\Sigma$  can then be decomposed into the Goldstone pieces  $h^{\hat{a}}(x)$  and a radial component  $h(x) = \sqrt{h^{\hat{a}}(x)h^{\hat{a}}(x)}$ ,

$$\Sigma = e^{ih^{\hat{a}}(x)T^{\hat{a}}/f} \langle \Sigma \rangle = \frac{\sin(h/f)}{h} (h^1, h^2, h^3, h^4, h \cot(h/f)). \quad (4.40)$$

With this parameterization, the SM Higgs doublet is

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} h^1 + ih^2 \\ h^3 + ih^4 \end{pmatrix}. \quad (4.41)$$

#### 4.4.3 Gauge couplings

We would like to write down a Lagrangian for this theory and parameterize the effects of the strong sector on the SM couplings. A useful trick for this is to pretend that the global  $\text{SO}(5) \times \text{U}(1)_X$  symmetry is gauged and then 'demote' the additional gauge fields to spurions—i.e. turn them off. We can then parameterize the quadratic part of the Lagrangian for the full set of  $\text{SO}(5)$  [partially spurious] gauge bosons,  $V_\mu = A_\mu^a T^a + A_\mu^{\hat{a}} T^{\hat{a}}$ , and the  $\text{U}(1)_X$  gauge boson,  $X$ , by writing down the leading  $\text{SO}(5) \times \text{U}(1)_X$ -invariant operators:

$$\Delta \mathcal{L} = \frac{1}{2} \left( \eta^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right) \left[ \Pi_X(q^2) X_\mu X_\nu + \Pi_0(q^2) \text{Tr}(A_\mu A_\nu) + \Pi_1(q^2) \text{Tr}(\Sigma A_\mu A_\nu \Sigma^T) \right]. \quad (4.42)$$

Where the form factors are completely analogous to (4.25) and (4.26). Contained in this expression are the kinetic and mass terms of the SM electroweak gauge bosons. To extract them, we must expand the form factors  $\Pi(q^2)$  in momenta and identify the  $\mathcal{O}(q^0)$  terms as mass terms and the  $\mathcal{O}(q^2)$  terms as kinetic terms. Since the  $\Pi_X$  and  $\Pi_0$  terms include gauge fields in the unbroken directions, they should vanish at  $q^2 = 0$ , otherwise masses would be generated for those directions. The  $\Pi_1$  term, however, selects out the broken direction upon inserting the  $\Sigma \rightarrow \Sigma_0$  and thus contains the Goldstone pole, (4.27). We thus find

$$\Pi_0(0) = \Pi_X(0) = 0 \quad \Pi_1(0) = f^2. \quad (4.43)$$

Assuming that the Higgs obtains a VEV, one may rotate it into a convenient location  $(h^1, \dots, h^4) = (0, 0, v/\sqrt{2}, 0)$  corresponding to the usual SM Higgs VEV parameterization. We now assume that  $H_{\text{gauge}}$  is the SM electroweak group and drop all spurious gauge bosons. Using (4.40), the strong sector contribution to the Lagrangian of these gauge bosons to  $\mathcal{O}(q^2)$  is

$$\Delta \mathcal{L}_{q^0} = \left( \eta^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right) \frac{1}{2} \left( \frac{f^2}{4} \sin^2 \frac{\langle h \rangle}{f} \right) (B_\mu B_\nu + W_\mu^3 W_\nu^3 - 2W_\mu^3 B_\nu + 2W_\mu^+ W_\nu^-) \quad (4.44)$$

$$\Delta\mathcal{L}_{q^2} = \frac{q^2}{2} [\Pi'_0(0)W_\mu^a W_\nu^a + (\Pi'_0(0) + \Pi'_X(0))B_\mu B_\nu], \quad (4.45)$$

where we have used the choice of  $SO(5)$  generators in the appendix of [138].  $\Delta\mathcal{L}_{q^2}$  gives contributions to the kinetic terms of the gauge bosons. Observe that these are not canonically normalized, but instead can be thought of as shifts in the gauge coupling,

$$\Delta\left(\frac{1}{g^2}\right) = -\Pi'_0(0) \quad \Delta\left(\frac{1}{g'^2}\right) = -(\Pi'_0(0) + \Pi'_X(0)). \quad (4.46)$$

Thus if the  $SU(2)_L$  gauge bosons have a ‘pure’ gauge coupling  $g_0$  when one turns off the strongly coupled sector, the full observed  $SU(2)_L$  gauge coupling is

$$\frac{1}{g_{SM}^2} = \frac{1}{g_0^2} - \Pi'_0(0), \quad (4.47)$$

and similarly for  $g'_{SM}$ .

$\Delta\mathcal{L}_{q^0}$  corresponds to contributions the masses of the heavy electroweak gauge bosons. Taking into account the need to canonically normalize with respect to  $\Delta\mathcal{L}_{q^2}$ , we obtain the usual  $W^\pm$  and  $Z$  masses by identifying the SM Higgs VEV as  $v = f \sin(\langle h \rangle / f)$ . We see the appearance of the misalignment angle,

$$\xi = \sin^2 \frac{\langle h \rangle}{f} \equiv \frac{v^2}{f^2}. \quad (4.48)$$

Finally, by restoring  $\langle h \rangle \rightarrow h(x)$  in (4.44) we may determine the composite Higgs couplings to the gauge bosons<sup>18</sup>. The key is the expansion

$$f^2 \sin^2 \frac{h(x)}{f} = v^2 + 2v\sqrt{1-\xi}h(x) + (1+2\xi)h(x)^2 + \dots \quad (4.49)$$

From this we can make a prediction for the  $SO(5)/SO(4)$  composite Higgs couplings to the heavy electroweak gauge bosons  $V = W^\pm, Z$  relative to their SM values,

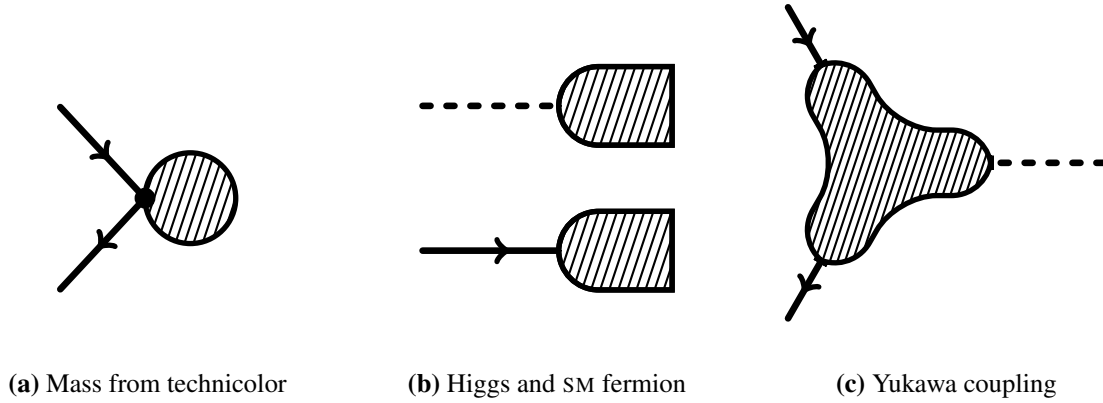
$$g_{VVh} = \sqrt{1-\xi}g_{VVh}^{SM} \quad g_{VVhh} = (1-2\xi)g_{VVhh}^{SM}. \quad (4.50)$$

At this point, these couplings introduce gauge boson loops which are quadratically divergent. These loops go like

$$\text{---} \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \end{array} \sim \frac{g^2}{16\pi^2} \Lambda^2 \sim g_{SM}^2 (1-\xi) f^2, \quad (4.51)$$

where we have used the dimensional analysis limit  $\Lambda = 4\pi f$ . We see that having explained the lightness of the Higgs by appealing to the Goldstone shift symmetry, reintroducing the Higgs couplings to the gauge bosons breaks this shift symmetry and wants to push the Higgs mass back up towards the symmetry breaking scale. In order to avoid this, one additional ingredient called collective breaking (along with light gauge and top partners) is necessary. We present this in Section 4.5.

<sup>18</sup>This is a trivial use of the Higgs low-energy theorem: the low-momentum Higgs couplings are equivalent to promoting the VEV to  $h(x)$  like [139, 140] This theorem can be used, for example, to calculate the Higgs coupling to photons by evaluating the mass dependence of the running of the QED gauge coupling. The application of the theorem to composite Higgs models is explored in [141].



**Fig. 11:** Fermion couplings to the composite sector, represented by shaded blobs. (a): Bilinear coupling of fermions to the composite sector (4.52) lead to fermion masses from the condensate of techniquarks. (b): Partial compositeness scenario. In addition to the Higgs being part of the strong sector, the elementary SM fermions mix linearly with strong sector operators with the same quantum numbers. (c): Yukawa interactions are generated through the strong sector dynamics. Adapted from [138].

#### 4.4.4 Partial compositeness

Having introduced the Higgs couplings to the gauge bosons, we can move on to finding a way to incorporate the Yukawa couplings into composite Higgs models. The way this is done in technicolor is to introduce a four-Fermi interaction that is bilinear in SM fields, e.g.

$$\Delta\mathcal{L} \sim (\bar{Q}_L u_R)(\bar{\psi}_{TC}\psi_{TC}) \quad (4.52)$$

where the  $(\bar{\psi}_{TC}\psi_{TC})$  are bilinears of the techni-quarks. The resulting fermion mass is shown in Fig. 11a. This strategy typically runs afoul of constraints on CP violation and flavor-changing neutral currents since one can imagine the composite sector similarly generating a four-fermion operator between SM states unless elaborate flavor symmetry schemes are assumed.

Instead of connecting the strong sector to a SM fermion bilinear, we can consider a *linear* connection. This is known as **partial compositeness** and is shown in Fig. 11b. We assume that instead of (4.52), the elementary fermions mix with a fermionic composite operator,

$$\Delta L \sim \bar{Q}_L \mathcal{O}_{Q_L}, \quad (4.53)$$

where  $\mathcal{O}_{Q_L}$  is a strong sector operator that interpolates a composite quark doublet. We assume similar mixing terms for each of the other SM fermions. In order to preserve the SM quantum numbers we must assume that the SM gauge group is a weakly gauged subgroup of the strongly coupled sector's flavor symmetries. Note that the gauge bosons are also partially composite<sup>19</sup>, as we saw in (4.23). The resulting Yukawa interactions are shown in Fig. 11c.

The degree of mixing is now a freedom in our description. Let us parameterize the elementary–composite mixing by ‘angles’  $\epsilon$ ,

$$|\text{observed particle}\rangle \sim |\text{elementary}\rangle + \epsilon|\text{composite}\rangle. \quad (4.54)$$

We can use this degree of compositeness to control flavor violation. Since the strongest flavor constraints are for the first two generations, we assume that the first two generations have very small mixing with

<sup>19</sup>In this framework the longitudinal modes of the massive SM gauge bosons pick up this partial compositeness from the Higgs. It is also possible to have a scenario where the transverse modes are partially composite, see [142, 143] for explicit realizations.

the composite sector. This suppresses dangerous flavor-violating four-fermion operators. On the other hand, we may assume that the third particles are more composite than the first two generations,

$$\epsilon_3 \gg \epsilon_{1,2}. \quad (4.55)$$

Since the degree of compositeness also controls the interaction with the Higgs, this means that the third generation particles have a larger Yukawa coupling and, upon electroweak symmetry breaking, have heavier masses. The astute reader will note that this is exactly the same as the flavor structure of the ‘realistic’ Randall-Sundrum models in Sec. 3.11-3.12. The observation that light fermions can automatically avoid flavor bounds is precisely what we called the ‘RS GIM mechanism.’ This is no surprise since the holographic interpretation of the RS model is indeed one where the Higgs is composite.

#### 4.4.5 Breaking electroweak symmetry

Having addressed the Higgs couplings to both the SM gauge bosons and fermions, we move on to the Higgs self-couplings. Until now we have simply assumed that the strong sector generates an electroweak-symmetry breaking potential. We now check that this assumption is plausible by arguing that loops involving the third generation quarks generate such a potential; this is similar to the Nambu–Jona-Lasinio model [144, 145].

The SM fermions do not form complete representations of the global group  $G = \text{SO}(5) \times U(1)_X$ . We thus follow the same strategy that we used for the gauge bosons in Sec. 4.4.3. Let us promote the SM fermions to full  $\text{SO}(5)$  spinor representations,

$$\Psi_Q = \begin{pmatrix} Q_L \\ \chi_Q \end{pmatrix} \quad \Psi_u = \begin{pmatrix} \psi_u \\ u_R \\ \chi_u \end{pmatrix} \quad \Psi_d = \begin{pmatrix} \psi_d \\ \chi_d \\ d_R \end{pmatrix}, \quad (4.56)$$

where the dashed line separates the  $\text{SU}(2)_L \times \text{SU}(2)_R$  parts of  $\text{SO}(4) \subset \text{SO}(5)$ . The  $\psi$  and  $\chi$  fields are spurions. Recall from Sec. A.3.1 that the fundamental spinor representation for  $\text{SO}(5)$  is a Dirac spinor which decomposes into two Weyl spinors. Do not confuse these Weyl spinors (4.56) with Poincaré representations—these are representations of the global  $\text{SO}(5)$  internal group. In other words, the entire  $\Psi$  multiplet are Weyl spinors with respect to Poincaré symmetry but are Dirac spinors with respect to the internal  $\text{SO}(5)$  symmetry. The upper half of the Dirac  $\Psi$  spinors are charged under  $\text{SU}(2)_L$  while the lower half is charged under  $\text{SU}(2)_R$ . This imposes a  $U(1)_X$  charge of  $1/6$  on the  $\Psi$  fields to give the correct hypercharge assignments on the SM fields.

Now let us parameterize the strong sector dynamics in the couplings of the  $\text{SO}(5)$  fermions  $\Psi$  and the linear field  $\Sigma$  in (4.40) that encodes the composite Higgs. Since the  $\Sigma$  is an  $\text{SO}(5)$  vector, it can appear in a fermion bilinear as  $\Sigma_i \Gamma^i$ , where the  $\Gamma$  are the 5D Euclidean space representation of the Clifford algebra. The effective SM fermion bilinear terms are

$$\mathcal{L} = \sum_{r=Q,u,d} \bar{\Psi}_r \not{p} [\Pi_{r0} + \Pi_{r1}(\Gamma \cdot \Sigma)] \Psi_r + \sum_{r=u,d} \bar{\Psi}_Q \not{p} [M_{r0} + M_{r1}(\Gamma \cdot \Sigma)] \Psi_r + \text{h.c.} \quad (4.57)$$

where, as before, the form factors  $\Pi$  and  $M$  are momentum-dependent. We shall focus on only the  $Q_L$  and  $t_R$  pieces since they have the largest coupling to the strong sector.

**Keeping track of conjugate fields.** One should be careful with the conjugate fields in the above expression. For the Lorentz group in four and five dimensions,  $\text{SO}(3,1)$  and  $\text{SO}(4,1)$ , we use the Dirac conjugate  $\bar{\Psi} \equiv \Psi^\dagger \gamma^0$  to form Lorentz invariants. Recall that this is because objects like  $\Psi^\dagger \Psi$  are not necessarily invariant because representations of the Lorentz group are not unitary—boosts acting on the spinor representation do not satisfy  $U^\dagger U = 1$ . This is due to the relative sign between the time-like and space-like directions in the Minkowski metric. The Dirac conjugate is a way around this. For the case of the  $G = \text{SO}(5)$  internal symmetry, however, there is no issue of non-unitarity. Hence no additional  $\Gamma^0$  (acting

on the internal  $SO(5)$  space) is necessary in the Lagrangian. To be clear, we can write out the spacetime  $\gamma$  and internal  $\Gamma$  matrices explicitly:

$$\bar{\Psi} = \Psi^\dagger \gamma^0 \neq \Psi^\dagger \gamma^0 \Gamma^0. \quad (4.58)$$

The matrix  $\Gamma \cdot \Sigma$  takes the form

$$\Gamma \cdot \Sigma = \frac{1}{h} \begin{pmatrix} h \cos(h/f) & \not{h} \sin(h/f) \\ \not{h} \sin(h/f) & -h \cos(h/f) \end{pmatrix}, \quad (4.59)$$

where  $\not{h}$  and  $\bar{\not{h}}$  are appropriate contractions with Pauli matrices. With the above caveat that there is no  $\Gamma^0$  acting on the  $SO(5)$  conjugate, we may write out the Lagrangian for  $Q_L$  and  $t_R$  by dropping the spurious components of the  $\Psi$  fields,

$$\mathcal{L} = \bar{Q}_L \not{p} \left[ \Pi_{Q0} + \Pi_{Q1} \cos \frac{h}{f} \right] Q_L + \bar{t}_R \not{p} \left[ \Pi_{t0} + \Pi_{t1} \cos \frac{h}{f} \right] t_R + \bar{Q}_L M_{u1} \left[ h \sin \frac{h}{f} H^c \right] t_R, \quad (4.60)$$

where  $H^c = i\sigma^2 H$  is the usual conjugate Higgs doublet in the  $SM^{20}$ . Observe that upon canonical normalization, the top mass can be read off the Yukawa term,

$$m_t^2 = \left( \frac{v}{f} \right)^2 \frac{M_{t1}^2}{(\Pi_{Q0} + \Pi_{Q1})(\Pi_{t0} - \Pi_{t1})}, \quad (4.61)$$

where the form factors are evaluated at zero momentum. One may write similar expressions for the other fermions.

In order to determine whether electroweak symmetry is broken, we can now plug this information into the Coleman-Weinberg potential for the Higgs, also known as the [quantum] effective potential. This is the potential term in the effective action after taking into account quantum corrections from integrating out the top quarks. In other words, it is the potential that determines the vacuum expectation value of fields. The result is

$$V_{CW} = -6 \int \bar{d}^4 p \log \left( \Pi_{Q0} + \Pi_{Q1} \cos \frac{h}{f} \right) + \log \left[ p^2 \left( \Pi_{Q0} + \Pi_{Q1} \cos \frac{h}{f} \right) \left( \Pi_{t0} - \Pi_{t1} \cos \frac{h}{f} \right) M_{t1}^2 \sin^2 \frac{v}{f} \right]. \quad (4.62)$$

Expanding this to first order and keeping the leading order terms in the Higgs gives

$$V_{CW}(h) = \alpha \cos \frac{h}{f} - \beta \sin^2 \frac{h}{f}, \quad (4.63)$$

where  $\alpha$  and  $\beta$  are integrals over functions of the form factors where  $\beta$  is typically of the order the top Yukawa. If  $\alpha \leq 2\beta$ , then the Higgs acquires a VEV parameterized by

$$\xi \equiv \sin^2 \frac{\langle h \rangle}{f} = 1 - \left( \frac{\alpha}{2\beta} \right)^2. \quad (4.64)$$

This means that a small  $\xi$  typically requires a cancellation between  $\alpha$  and  $\beta$ . Since these come from different sources, this is generically a tuning in the theory.

One can also ask if it was necessary to rely on the top quark. For example, we know that the gauge sector also breaks the Goldstone shift symmetry so that loops of gauge bosons can generate quadratic

<sup>20</sup>Here we have used the  $SO(5)$  basis in [114].

and quartic terms in the Higgs potential. However, for a vector-like strong sector, gauge loops contribute with the wrong sign to the  $\beta$  term and pushes to align—rather than misalign—the vacuum [146, 147].

There are, however, alternate mechanisms to enact electroweak symmetry breaking for a composite Higgs. For example,

- mixing the composite Higgs with an elementary state [131],
- making use of an explicitly broken global symmetry [148]
- enlarging the  $H_{\text{gauge}}$  so that it cannot be completely embedded into  $H_{\text{global}}$  [149–151].

#### 4.5 Collective symmetry breaking

The general composite Higgs is a useful framework for working with the Higgs as a pseudo-Goldstone boson. However, we saw in Section 4.4.1 and equation (4.51) that this is not enough to avoid tuning. The source is clear: a pure Goldstone Higgs is protected from quadratic corrections to its mass because of its shift symmetry. This very same shift symmetry prevents the required Higgs couplings to gauge bosons, fermions, and itself. One must break this shift symmetry in order to endow the Higgs with these couplings; this generically reintroduces a dependence on the cutoff,  $\Lambda = 4\pi f$ .

This may make it seem like a no-go theorem for any realistic model of a pseudo-Goldstone Higgs. However, there is a nice way out of this apparent boondoggle called **collective symmetry breaking** that was originally introduced in ‘little Higgs’ models [152–154] (see [117, 118] for reviews) and is now an a key ingredient in composite Higgs models<sup>21</sup>. The idea is that one can separate the scales  $v$  and  $f$  by introducing new particles which cancel the quadratic divergences at one-loop order. Unlike supersymmetry, these partner particles carry the same spin as the Standard Model particles whose virtual contributions are to be cancelled. Further, this cancellation only occurs for one-loop diagrams: higher loop diagrams are expected to contribute quadratically at their naïve dimensional analysis size, but these are suppressed relative to the leading term.

The general principle that allows this cancellation is that the shift symmetry is redundantly protected. A process is only sensitive to explicit symmetry breaking—as necessary for SM-like Higgs couplings—if this explicit breaking is communicated by at least two different sectors of the theory. More concretely, the symmetry is only explicitly broken if multiple couplings are non-zero in the theory so that any diagram that encodes this explicit breaking must include insertions from at least two different couplings. This softens the cutoff sensitivity of various operators by requiring additional field insertions that decrease the degree of divergence of loop diagrams.

##### 4.5.1 Collective breaking in action

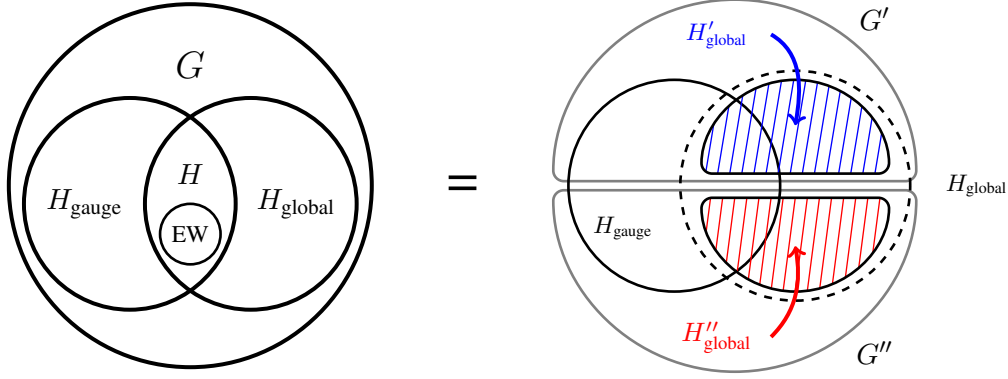
We now demonstrate collective symmetry breaking in a model based on the ‘anatomy’ in Fig. 12. The reader may find it useful to refer to the explicit example of a simple little Higgs model in Section 4.5.2 below. Instead of a simple global group  $G$ , suppose that  $G = G' \times G''$ . Each of these factors breaks spontaneously to subgroups  $H'_{\text{global}}$  and  $H''_{\text{global}}$ , respectively. The spontaneous symmetry breaking pattern is thus

$$G = G' \times G'' \rightarrow H'_{\text{global}} \times H''_{\text{global}}. \quad (4.65)$$

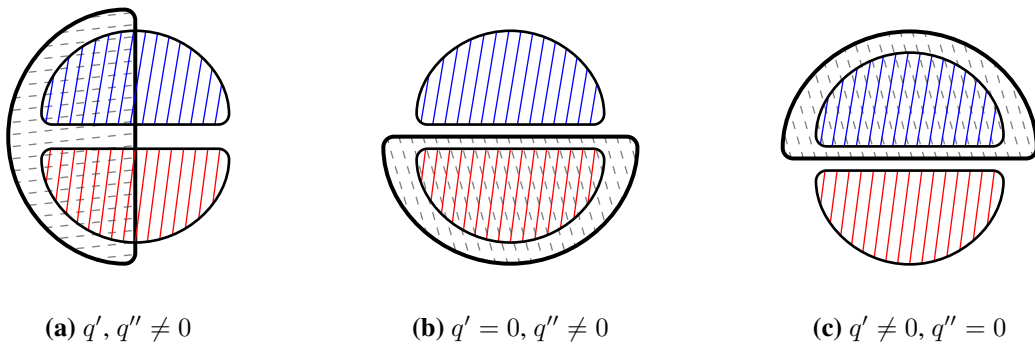
This gives us two linear fields  $\Sigma'$  and  $\Sigma''$  analogous to (4.40) so that there are two separate sets of Goldstone bosons.

We explicitly break  $G$  by gauging  $H_{\text{gauge}} \subset G$ . Suppose that both  $H'_{\text{global}}$  and  $H''_{\text{global}}$  are subgroups of  $H_{\text{gauge}}$  in such a way that both  $\Sigma'$  and  $\Sigma''$  are charged under  $H_{\text{gauge}}$  with nonzero charges  $q'$  and  $q''$

<sup>21</sup>In Section 4.7.4 we present an alternate protection mechanism based on a  $\mathbb{Z}_2$  symmetry.



**Fig. 12:** Anatomy of collective symmetry breaking, following the conventions in Fig. 10.



**Fig. 13:** Collective symmetry breaking. Upper (blue) and lower (red) blobs represent  $H'$  and  $H''$  in Fig. 12. The thick black line represents the gauged symmetry  $H_{\text{gauge}}$  under which  $\Sigma'$  has charge  $q'$  and  $\Sigma''$  has charge  $q''$ . When either  $q'$  or  $q''$  vanishes, the unbroken group is  $H'_{\text{global}} \times H''_{\text{global}}$ .

respectively. A piece of each subgroup is gauged, as shown in Fig. 13a.  $H'_{\text{global}} \times H''_{\text{global}}$  is then explicitly broken to a smaller subgroup, for example a vectorial subgroup identified by the gauging,  $H$ .

On the other hand, when either  $q'$  or  $q''$  is set to zero, only one of the global subgroups is gauged, as shown in Fig. 13b and 13c. In either of these cases, the resulting global symmetry group is still  $H'_{\text{global}} \times H''_{\text{global}}$ . In other words, one requires both  $q'$  and  $q''$  to explicitly break  $H_{\text{global}} = H'_{\text{global}} \times H''_{\text{global}}$ .

When one of the global subgroups is uncharged under the gauged subgroup, say  $H''_{\text{global}}$ , those Goldstone bosons pick up no mass from the gauge sector. For the other global subgroup which is charged under the gauge group, say  $H'_{\text{global}}$ , there are two possibilities:

1. If  $H_{\text{gauge}} \subseteq H'_{\text{global}}$ , then loops of the gauge bosons will feed into the mass of the pseudo-Goldstone bosons. In the absence of collective symmetry breaking, this gives a contribution that is quadratic in the cutoff.
2. If, on the other hand<sup>22</sup>,  $G' \subset H_{\text{gauge}}$ , then the would-be Goldstone bosons from  $G' \rightarrow H'_{\text{global}}$  are eaten by the  $(G/H'_{\text{global}}) \cap H_{\text{gauge}}$  gauge bosons. There is no quadratic sensitivity to the cutoff.

In the second case, the Higgs mechanism removed the  $\Lambda^2$  contribution to the pseudo-Goldstone mass, but it also got rid of the pseudo-Goldstones themselves.

<sup>22</sup>It is sufficient to consider some subgroup  $\tilde{G}' \subseteq G'$  that contains  $H'_{\text{global}}$  as a proper subgroup

This leads us to consider the case when both  $G'$  and  $G''$  (not just their  $H_{\text{global}}$  subgroups) are charged under the gauged symmetry. For simplicity, suppose  $G' = G'' = H_{\text{gauge}}$  so that one gauges the vectorial combination. In this case, both the  $\Sigma'$  and  $\Sigma''$  fields carrying our Goldstone bosons are charged under the gauge group. The gauge fields become massive by the Higgs mechanism, but there are twice as many Goldstone bosons than it can eat<sup>23</sup>. Indeed, the ‘axial’ combination of  $G'$  and  $G''$  furnishes a set of Goldstone bosons that remain uneaten and are sensitive to explicit breaking effects so that they are formally pseudo-Goldstones. Any contribution to the pseudo-Goldstone mass, however, must be proportional to  $(gq')(gq'')$ , where  $g$  is the gauge coupling. In other words, it requires interactions from both  $\Sigma'$  and  $\Sigma''$ . The resulting mass term is suppressed since this requires factors of the  $\Sigma'$  and  $\Sigma''$  VEVs to soak up additional boson legs. We now demonstrate this with an explicit example.

**Why can't you just rotate to a different basis?** Based on Fig. 13, one might wonder if we can repartition  $G = G' \times G''$  so that the  $H'_{\text{global}}$  and  $H''_{\text{global}}$  subgroups are always both gauged. Alternately, perhaps one can repartition  $G$  so that only one subgroup is ever gauged. This cannot be done, even when  $q' = q''$ . The reason is precisely what we pointed out above Sec. 4.3.3: the axial combination of two groups is not itself a group since its algebra doesn't close.

#### 4.5.2 Explicit example: $(SU(3) \rightarrow SU(2))^2$

Let us see how this fits together in a simple little Higgs model—though we emphasize that collective symmetry breaking is a generic feature of all realistic composite Higgs models, not just those of little Higgs type. We classify composite Higgs models in Section 4.7 to clarify this disambiguation. Consider the case where  $G' = G'' = SU(3)$  and  $H'_{\text{global}} = H''_{\text{global}} = SU(2)$ . We thus have two fields which are linear representations of  $SU(3)$  and carry the Goldstone bosons,

$$\Sigma' = \exp \left[ \frac{i}{f'} \begin{pmatrix} 0_{2 \times 2} & H' \\ -\bar{H}'^\dagger & 0 \end{pmatrix} \right] \begin{pmatrix} 0 \\ 0 \\ f' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ f' \end{pmatrix} + i \begin{pmatrix} H' \\ \bar{0} \end{pmatrix} - \frac{1}{2f'} \begin{pmatrix} 0 \\ \bar{H}'^\dagger \bar{H}' \end{pmatrix}, \quad (4.66)$$

and similarly for  $\Sigma''$ . For simplicity let us set  $f' = f'' \equiv f$ . The kinetic terms for the  $\Sigma$  fields are

$$\mathcal{L} = |D_\mu \Sigma'|^2 + |D_\mu \Sigma''|^2 = \dots + (gq')^2 |V_\mu^a T'^a \Sigma'|^2 + (gq'')^2 |V_\mu^a T''^a \Sigma''|^2, \quad (4.67)$$

where  $T'^a = T''^a$  are the generators of the gauged group. To see the contribution to the Higgs mass, one can Wick contract the two gauge bosons in these terms—this is precisely the analog of the ‘cat diagram’ in Fig. 9. This contraction ties together the gauge boson indices so that the resulting term goes like

$$[\text{loop factor}] (gq')^2 \Sigma'^\dagger T'^a T'^a \Sigma' = [\text{loop factor}] \frac{(gq')^2}{2} C_2 \Sigma'^\dagger \mathbf{1}_{\text{gauge}} \Sigma', \quad (4.68)$$

and similarly for  $\Sigma''$ . Here the loop factor contains the quadratic dependence on the cutoff,  $[\text{loop factor}] \sim \Lambda^2/16\pi^2$ , and the factor  $\mathbf{1}_{\text{gauge}}$  is the identity matrix in the appropriate gauged subgroup. Here we have used  $T^a T^a = C_2 \mathbf{1}$ , where  $C_2$  is the quadratic Casimir operator of the representation<sup>24</sup>. Now let's explicitly demonstrate how collective breaking works.

- If only the  $SU(2) = H'_{\text{global}} = H''_{\text{global}}$  parts of  $G'$  and  $G''$  were gauged, then there would be two separate sets of pseudo-Goldstone bosons  $H'$  and  $H''$ . We plug in the expansion of  $\Sigma'$  (4.66) into (4.68) and note that in this case,

$$\mathbf{1}_{\text{gauge}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (4.69)$$

<sup>23</sup>This is a manifestation of general outdoors advice: if you (a Goldstone boson) are being chased by a hungry bear (a gauge boson), it is not necessary for your survival that you can outrun it (have zero coupling). It is sufficient that you are with friends whom you can outrun. Collective breaking is, in part, the requirement that you have more slow friends than hungry bears.

<sup>24</sup> $C_2(\text{fundamental}) = (N^2 - 1)/2N$  for  $SU(N)$ .



This picks up the Goldstones in the second term on the right-hand side of (4.66) so that there is indeed a Goldstone mass term proportional to  $\Lambda^2 = (4\pi f)^2$  for each set of Goldstones.

- On the other hand, in the case where  $G'$  and  $G''$  are both gauged with  $q' = q''$ , the matrix  $\mathbf{1}_{\text{gauge}}$  becomes a true identity operator,

$$\mathbf{1}_{\text{gauge}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.70)$$

Now the global symmetry breaking VEVs  $\langle \Sigma' \rangle$  and  $\langle \Sigma'' \rangle$  break part of the gauge symmetry and the Higgs mechanism tells us that there are gauge bosons that eat would-be Goldstones. Indeed, the first term on the right-hand side of (4.66)—which is no longer projected out by  $\mathbf{1}_{\text{gauge}}$ —encodes the mass picked up by the gauge bosons. Observe, however, what has happened to the  $\Lambda^2$  mass contribution in the previous scenario: it is now cancelled by the cross term between the first and third terms on the right hand side of (4.66). In other words, the terms which gave the quadratic sensitivity to the cutoff have vanished.

If we were only considering a single  $SU(3) \rightarrow SU(2)$  global symmetry breaking, then we would still be out of luck since the massive gauge bosons would have eaten all of our Goldstone bosons—so even though we got rid of the  $\Lambda^2$  sensitivity of the pseudo-Goldstone masses, we also would have gotten rid of the pseudo-Goldstones themselves. With foresight, however, we have followed the advice of footnote 23: we have more Goldstones than our gauge bosons can possibly eat.

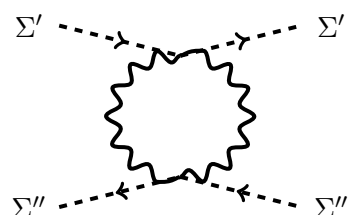
A useful way to parameterize our Goldstones is to follow the convention in (4.16):

$$\Sigma' = \exp \left[ \frac{i}{f} \begin{pmatrix} 0_{2 \times 2} & V \\ -\bar{V}^\dagger & 0 \end{pmatrix} \right] \exp \left[ \frac{i}{f} \begin{pmatrix} 0_{2 \times 2} & H \\ -\bar{H}^\dagger & 0 \end{pmatrix} \right] \begin{pmatrix} 0 \\ 0 \\ f \end{pmatrix} \quad (4.71)$$

$$\Sigma'' = \exp \left[ \frac{i}{f} \begin{pmatrix} 0_{2 \times 2} & V \\ -\bar{V}^\dagger & 0 \end{pmatrix} \right] \exp \left[ \frac{-i}{f} \begin{pmatrix} 0_{2 \times 2} & H \\ -\bar{H}^\dagger & 0 \end{pmatrix} \right] \begin{pmatrix} 0 \\ 0 \\ f \end{pmatrix}, \quad (4.72)$$

where we have identified the Higgs as the axial combination of global shifts, while the vector combination of Goldstones,  $V$ , is eaten by the gauge bosons to become massive.

Now the  $H$  pseudo-Goldstones only pick up mass from diagrams that involve both the  $(gq')$  and the  $(gq'')$  couplings. In other words, it requires a combination of the  $\Sigma'$  and the  $\Sigma''$  fields. The leading order contribution comes from diagrams of the form



$$\sim \frac{g^4}{16\pi^2} \log \Lambda^2 \left| \Sigma'^\dagger \Sigma'' \right|^2. \quad (4.73)$$

Since  $\Sigma'^\dagger \Sigma'' = f^2 - 2H^\dagger H + \dots$ , we see that the leading term in the Higgs mass is only logarithmically sensitive to  $\Lambda$  because it required one power each of the  $\Sigma'$  and  $\Sigma''$  VEVs. The Higgs mass sets the electroweak scale to be on the order of  $f/(4\pi)$ . This is a factor of  $(4\pi)$  suppressed compared to the global symmetry breaking scale  $f$ —generating the hierarchy in  $\xi$  that we wanted—and also a further factor of  $(4\pi)$  from the cutoff  $\Lambda = 4\pi f$ . In this sense, collective symmetry breaking shows us what we can buy for factors of  $(4\pi)$  and why those factors are important in naïve dimensional analysis.

### 4.5.3 Top partners

As before, the largest contribution to the Higgs mass comes from the top quark. In the simple scenario above, we have extended our gauge group<sup>25</sup> from  $SU(2)_L$  to  $H_{\text{gauge}} = SU(3)$  so we'll need to also extend the usual top doublet to include a partner  $T_L$

$$Q = \begin{pmatrix} t_L \\ b_L \end{pmatrix} \rightarrow Q = \begin{pmatrix} t_L \\ b_L \\ T_L \end{pmatrix}. \quad (4.74)$$

We must also include a right-handed  $SU(3)$  singlet  $T'_R$  as a partner for the  $T_L$ , in parallel to the usual right-handed  $t'_R$  partner of the SM  $t_L$ . The prime on the  $t'_R$ —what is normally called  $t_R$  in the SM—is for future convenience. The Yukawa terms for the top quarks are,

$$\mathcal{L}_{\text{top}} = \lambda' \Sigma'^{\dagger} Q t'_R{}^{\dagger} + \lambda'' \Sigma''^{\dagger} Q T'_R{}^{\dagger} + \text{h.c.} \quad (4.75)$$

where the fermions are written in terms of Weyl spinors. Other terms, such as  $\Sigma'^{\dagger} Q T'_R{}^{\dagger}$  or  $\Sigma''^{\dagger} Q t'_R{}^{\dagger}$ , can typically be prohibited by invoking chiral symmetries. Observe that the  $\lambda'$  term is invariant under  $G'$  if  $Q$  is a fundamental under  $G'$ . Similarly, the  $\lambda''$  term is invariant under  $G''$  if  $Q$  is a fundamental under  $G''$ . This is indeed consistent since  $Q$  is a fundamental under  $H_{\text{gauge}}$  which is the diagonal subgroup of  $G' \times G''$ . This shows us how collective symmetry breaking is embedded in the Yukawa sector. When only one of the  $\lambda$  terms is nonzero,  $\mathcal{L}_{\text{top}}$  is  $G' \times G''$  invariant. However, when both are turned on, the global symmetry is broken down to the diagonal subgroup.

This collective breaking is similar to the breaking of the global  $U(3)_Q \times U(3)_U \times U(3)_D$  flavor symmetry to  $U(3)$  by the up- and down-type Yukawas in the Standard Model. If  $y_u = 0$  and  $y_d \neq 0$ , then the flavor symmetry would be enhanced to  $U(3)^2$  since the right-handed up-type quarks could be rotated independently of the other fields.

We can now plug in the expansion (4.71 – 4.72) into the Yukawa terms (4.75), ignoring the  $V$  terms since we now know those are eaten by the gauge bosons. Expanding the resulting product gives

$$\mathcal{L}_{\text{top}} = i H^{\dagger} Q (\lambda'' T'_R{}^{\dagger} - \lambda' t'_R{}^{\dagger}) + \left( f - \frac{H^{\dagger} H}{2f} \right) T_L (\lambda' t'_R{}^{\dagger} + \lambda'' T'_R{}^{\dagger}). \quad (4.76)$$

From this we can write out the right-handed top eigenstates

$$T_R = \frac{\lambda' t'_R + \lambda'' T'_R}{\sqrt{\lambda'^2 + \lambda''^2}} \quad t_R = i \frac{\lambda'' T'_R - \lambda' t'_R}{\sqrt{\lambda'^2 + \lambda''^2}} \quad (4.77)$$

and the resulting top Yukawa, top partner mass, and top partner coupling to  $H^{\dagger} H$ ,

$$\mathcal{L}_{\text{top}} = \lambda_t H^{\dagger} Q t'_R{}^{\dagger} + \lambda_t f T_L T'_R{}^{\dagger} - \frac{\lambda_t}{2f} H^{\dagger} H T_L T'_R{}^{\dagger}, \quad (4.78)$$

where we see that all of the couplings are simply related to the SM top Yukawa,  $\lambda_t = \sqrt{\lambda'^2 + \lambda''^2}$ . These relations ensure the cancellation between diagrams that give a  $\Lambda^2$  contribution to the Higgs mass,

$$h \text{---} \frac{t}{\lambda_t} \text{---} \text{---} h + \text{---} \frac{T}{-\lambda_t/f} \text{---} \text{---} h = \mathcal{O}(\log \Lambda). \quad (4.79)$$

<sup>25</sup>For simplicity we ignore the  $U(1)_Y$  factor, it is straightforward to assign charges appropriately.

Note the symmetry factor of  $1/2$  in the  $h^2 T_L T_R^\dagger$  Feynman rule. For simplicity we also drop an overall  $\sqrt{2}$  in the normalization of the  $h$  field which is irrelevant for the  $\Lambda^2$  cancellation. We see that indeed collective symmetry breaking can protect against the reintroduction of quadratic sensitivity to the cutoff by the Yukawa interactions.

Just as in the case of natural SUSY, an important signature of this class of models is to look for the ‘partner top’ particles which are responsible for the softening of the cutoff dependence of Higgs mass from the top sector. One can search for these objects at the LHC through either pair production,

$$q\bar{q}/gg \rightarrow T\bar{T}, \tag{4.80}$$

or through single production in association with a SM quark,

$$bq \rightarrow Tq' \qquad qq' \rightarrow Tb. \tag{4.81}$$

The top partner decays are fixed by the Goldstone equivalence theorem. The partner top decays approximately 50% of the time to  $bW$ , with the remaining decay products split evenly between  $tZ$  and  $th$  [155]. The lower bound on the top partner mass from vector-like heavy top (also referred to as fourth generation) searches is  $\gtrsim 700$  GeV [156].

One can continue to calculate the Coleman-Weinberg potential in this scenario to check for electroweak symmetry breaking and further study the phenomenology of these models. We refer the reader to the excellent reviews [117, 118] for a pedagogical introduction in the context of the little Higgs. See [157–159] for a more general discussion of experimental bounds on top partners.

#### 4.6 Deconstruction and moose models

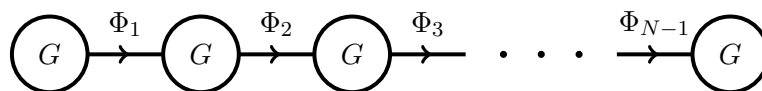
We now briefly mention some connections with extra dimensional models and introduce a diagrammatical language that is sometimes used to describe the symmetry breaking pattern in composite models.

In Section 3.9 we introduced the holographic principle as a connection between strongly coupled 4D theories and weakly coupled theories on a curved spacetime with an extra spatial dimension. This turns out to be a natural tool to get a handle for some of the strong dynamics encoded into the form factors. Indeed, the minimal composite Higgs model described above was developed using these insights [136].

There is, however, another way to connect 5D models to 4D models. 5D models have dimensionful couplings and are manifestly non-renormalizable. One proposal for a UV completion is to discretize (‘lattice’) the extra dimension [154, 160, 161]. In this picture, the extra dimension is split into  $N$  discrete sites which should no longer be thought of as discrete spacetimes, but rather as nodes in a ‘theory space’ that describe a gauge symmetry structure on a single 4D spacetime. The bulk gauge symmetry  $G$  lattice into a 4D gauged  $G$  on each of the  $N$  nodes,



At this level the nodes are just  $N$  separate gauge groups; after all, this is precisely what we mean by a local symmetry (see [162, 163] for a discussion in depth). We next introduce a set of  $(N - 1)$  scalar **link fields**  $\Phi_i$  which are in the bifundamental representation with respect to the  $N^{\text{th}}$  and  $(N + 1)^{\text{th}}$  gauge groups:  $(\mathbf{N}_i, \bar{\mathbf{N}}_{i+1})$ . We may draw these link fields as lines between the nodes,



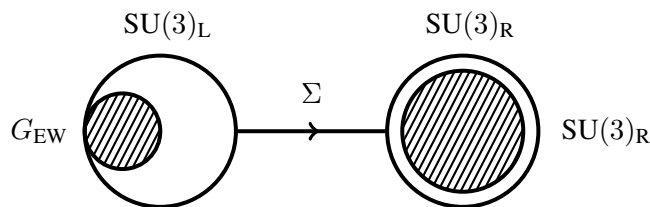
The arrow on the link field keeps track of the representation with respect to a group:

- Arrows leaving a node are fundamental with respect to that group.
- Arrows entering a node are anti-fundamental with respect to that group.

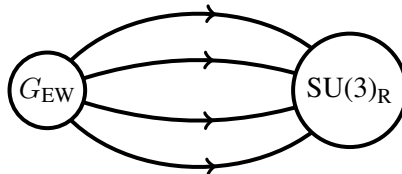
Now suppose each of these link fields acquires a VEV proportional to  $\mathbf{1}$  in their respective  $G_i \times G_{i+1}$  internal spaces. Each link field would spontaneously break the symmetry  $G_i \times G_{i+1} \rightarrow G_{\text{diag}}$ . The symmetries are broken down to  $G$ . One can diagonalize the mass matrix for the gauge boson—a problem that is mathematically identical to solving the waves in a system of  $N - 1$  springs in series [164]—to find that the spectrum looks like a tower of Kaluza-Klein modes. In fact, the link fields can be identified with the KK modes of the fifth component of the bulk gauge field  $A_5$ . This construction also shows explicitly that the Kaluza-Klein gauge fields in 5D acquire their masses from eating the KK modes of the  $A_5$ , which are here manifestly would-be Goldstone bosons. By coupling matter appropriately, one constructs a UV complete 4D model of a product of gauge groups that gives the same ‘low’ energy physics as an extra dimension. We refer the reader to the original literature for details [154, 160, 161] or [68] for a brief summary.

Rather than just way to UV complete extra dimensions, deconstructions are also a useful tool for motivating models of chiral symmetry breaking. In fact, they are a manifestation of a more general tool for composite models called **moose diagrams**<sup>26</sup> [166, 167]. One can use this diagrammatic language to construct little Higgs models; indeed, this was the original inspiration for the development of collective symmetry breaking paradigm in Section 4.5. The topology of these diagrams encodes information about spectrum of Goldstone modes [168]. From the dimensional deconstruction of an extra dimension, it’s clear that all of the Goldstones are eaten by the KK modes of gauge bosons. More general connections between nodes, however, allow more Goldstones to survive hungry gauge bosons.

As an example, we present the ‘minimal moose’ little Higgs model from [169]. The basic building block is the coset for chiral symmetry breaking,  $SU(3)_L \times SU(3)_R / SU(3)_V$ . We gauge the electroweak subgroup  $G_{EW}$  of  $SU(3)_L$  and the entire  $SU(3)_R$ , which we represent schematically with shaded blobs:

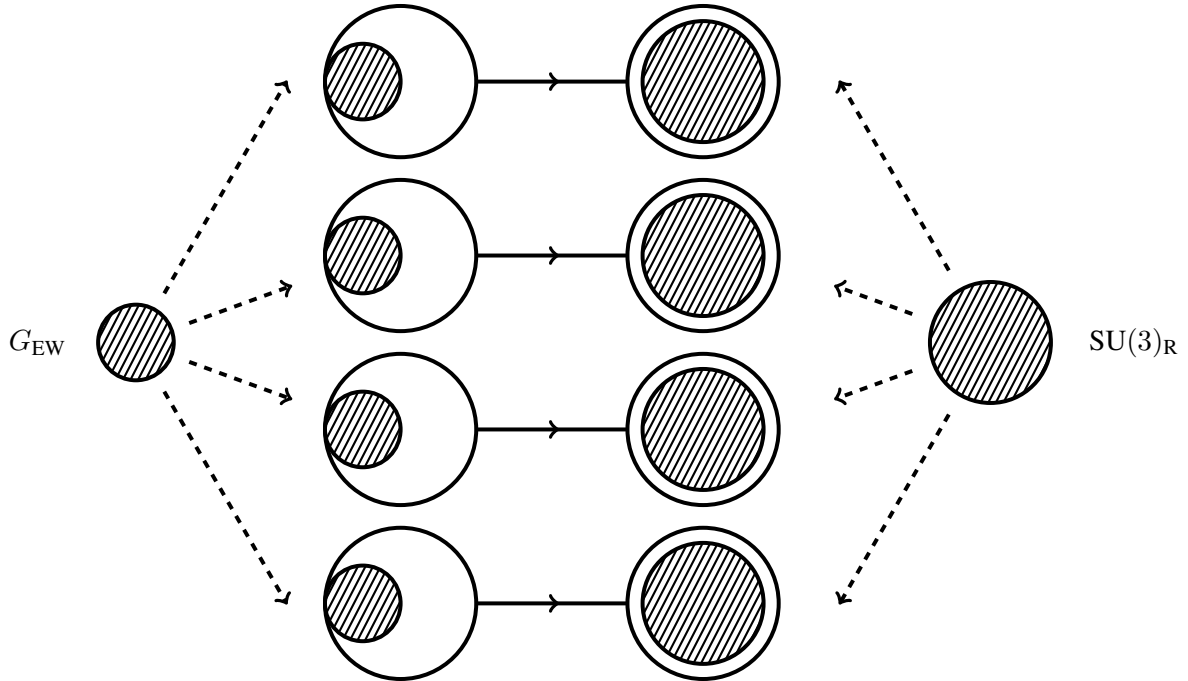


The minimal moose model actually requires four copies of this basic structure. As before, we only gauge the vectorial  $G_{EW}$  of each of the  $SU(3)_L$  factors and similarly for the  $SU(3)_R$  factors. In other words, the theory only has two gauge couplings. This is shown schematically in Fig. 14. We note that typically one only draws nodes for the gauge groups so that the usual moose diagram for this model is:



See §4.1 of [118] for a review of this particular model. A full discussion of these moose-based little Higgs models is outside of the scope of these lectures. In addition to the reviews mentioned above [117, 118], see [170] for the self-described ‘bestest’ little Higgs model and [157, 171] for a discussion of the status of composite Higgs models after the first run of the LHC.

<sup>26</sup>These diagrams are also called **quiver diagrams** by string theorists [165].



**Fig. 14:** Full symmetry structure of the minimal moose little Higgs model. Shaded blobs represent gauged subgroups. We explicitly show that only the ‘diagonal’ subgroups are gauged.

#### 4.7 A taxonomy of composite Higgs models

Having surveyed the main features of composite Higgs models, let us classify the landscape of such theories. This section is meant to clarify the distinctions between what is colloquially called a ‘composite Higgs’ versus a ‘little Higgs’ or a ‘holographic composite Higgs’ versus a ‘dilaton Higgs.’ We closely follow the discussion in Sections 2 – 3 of [116], to which we refer the reader for further details and references.

As a warm up and review, recall the Standard Model Higgs potential

$$V(h) = -\mu^2|H|^2 + \lambda|H|^4 \quad \longrightarrow \quad -\frac{1}{2}\mu^2 h^2 + \frac{\lambda}{4}h^4. \quad (4.82)$$

Minimizing the potential and matching to experiment yields

$$v = \langle h \rangle = \frac{\mu^2}{\lambda} = 246 \text{ GeV} \quad m_h^2 = 2\mu^2 = (125 \text{ GeV})^2, \quad (4.83)$$

where  $v$  has long been known from the masses and couplings of the electroweak gauge bosons, but  $m_h^2$  is new data from 2012. This new information tells us that  $\mu = 89 \text{ GeV}$  and, from the expression for  $v$ , that  $\lambda = 0.13$ .

Let us now map this onto a convenient parameterization of the Higgs potential in composite Higgs models.

$$V(h) = \frac{g_{\text{SM}}^2 M^2}{16\pi^2} \left( -ah^2 + \frac{b}{2f^2}h^4 \right). \quad (4.84)$$

One can compare this to (4.63). Here  $g_{\text{SM}}$  is a characteristic Standard Model coupling, such as  $g_{\text{SM}}^2 = N_c y_t^2$ . Implicit in this parameterization is the expectation that the Higgs potential is radiatively generated, giving a  $g_{\text{SM}}^2/16\pi^2$  prefactor. With this normalization, tree-level contributions appear as coefficients  $a, b$

MODEL	$\mathcal{O}(a)$	$\mathcal{O}(b)$	$\mathcal{O}(g_*)$	COMMENTS
Bona-fide composite Higgs	1	1	$4\pi$	Requires tuning of both $a$ and $b$ .
Little Higgs	1	$\frac{16\pi^2}{g_*^2}$	$\ll 4\pi$	Tree level quartic, $h$ too heavy.
Holographic Higgs	1	1	$\ll 4\pi$	$\sim$ little Higgs with loop-level quartic.
Twin Higgs	1	$1 - \frac{16\pi^2}{g_*^2}$	$g_{\text{SM}}$	$\mathbb{Z}_2$ rather than collective breaking.
Dilatonic Higgs		SEE TEXT		Related to RS radion Higgs.

Table 2: Taxonomy of composite Higgs models according to the couplings in (4.84) and (4.85); based on [116]. Models must be tuned when phenomenology requires values of the couplings that are very different from the expected magnitudes shown here.

that go like  $16\pi^2/g_{\text{SM}}^2$ . The mass scale  $M$  is typically that of the new states (e.g. top partners) that cut off the quadratic divergence introduced by the explicit breaking of the Goldstone shift symmetry, as discussed in Section 4.4.3. It is useful to parameterize this in terms of the coupling of these new states to the strong sector  $g_*$ ,

$$M = g_* f. \quad (4.85)$$

These states are typically lighter than the cutoff,  $4\pi f$ , to help with the little hierarchy problem. We expect the lighter mass comes from a weaker coupling to the strong sector,  $g_*$ , motivating the definition (4.85). The experimental information that the SM quartic is  $\lambda = 0.13$  is strongly suggestive of a loop induced coupling. Using the NDA scaling of a strong sector quartic (4.35) and a proportionality factor from an explicit global symmetry breaking SM loop,  $g_{\text{SM}}^2/16\pi^2$ , we estimate

$$\lambda_{\text{loop}} \approx 2 \frac{1}{16\pi^2} g_{\text{SM}}^2 g_*^2 \approx 0.15 \times \left( \frac{g_{\text{SM}}}{\sqrt{N_c} y_t} \right)^2 \left( \frac{g_*}{2} \right)^2. \quad (4.86)$$

Here the factor of 2 comes from two top partner polarizations and the scaling with respect to  $g_* = M/f$  comes from NDA [135]. Thus the coupling of the new state is  $g_* \ll 4\pi$  and is expected to be *weakly coupled*. This is a more quantitative version of the statement that the discovery of the 125 GeV Higgs signaled the death of technicolor, as we explained qualitatively in Section 4.2. The other implication of this weak coupling is that the new particles that cancel the quadratic sensitivity of the Higgs potential have masses well below the strong coupling scale,  $M \ll \Lambda = 4\pi f$ ; where we recall the NDA cutoff from Section 4.3.7.

Comparing (4.84) to (4.82 – 4.83) gives

$$v^2 = \frac{a}{b} f^2 = (246 \text{ GeV})^2 \quad m_h^2 = 2 \frac{g_{\text{SM}}^2}{16\pi^2} M^2 a = 4v^2 \frac{g_{\text{SM}}^2 g_*^2}{16\pi^2} b = (125 \text{ GeV})^2. \quad (4.87)$$

In the remainder of this section we examine five classes of composite Higgs models and classify them according to their natural expectations for  $a$ ,  $b$ , and  $g_*$ . These are summarized in Table 2.

#### 4.7.1 Bona-Fide Composite Higgs

The ‘bona-fide composite Higgs’ models in the first row of Table 2 are the simplest realizations of the Higgs as pseudo-Nambu–Goldstone boson idea: a strongly coupled sector has a global symmetry which is spontaneously broken and yields a Goldstone with the quantum numbers of the Higgs. The Higgs potential is assumed to be radiatively generated by explicit breaking terms so that in the parameterization (4.84),  $a \sim b \sim \mathcal{O}(1)$ . From the left-side equation of (4.87), a parametric separation between  $v$  and  $f$  requires  $a$  to be tuned small by an amount  $\xi$  in (4.36).

Even with this, however, this is a second tuning required on  $b$  since the new states are expected to couple to the strong sector with strong couplings,  $g_* \sim 4\pi$ . Thus one finds that the quartic coupling is too large in (4.86) compared to  $\lambda = 0.13$ . In other words, one predicts a Higgs mass that is heavier than observed in (4.87). This is mapped onto a tuning of  $b$ .

#### 4.7.2 Little Higgs

In little Higgs models, collective symmetry breaking naturally gives a hierarchy

$$\xi = \frac{v^2}{f^2} \sim \frac{g_*^2}{16\pi^2} \ll 1. \quad (4.88)$$

The quartic coupling appears at tree-level,  $\lambda \sim g_{\text{SM}}$ . This is shown as  $b \sim 16\pi^2/g_*^2$  in Table 2. Prior to the Higgs discovery, this set up was seen to be a feature: one explains the separation between  $v$  and  $f$ . However, (4.87) shows that this predicts a Higgs mass that is on the order of 500 GeV for  $g_{\text{SM}} \sim 1$ .

#### 4.7.3 Holographic Higgs

These models are motivated by AdS/CFT duals of warped extra dimensional models, as we discussed in Section 3.9. Like the ‘bona-fide composite Higgs,’ the entire potential for these models are radiatively generated. This thus suffers the same  $\mathcal{O}(\xi = v^2/f^2)$  to obtain the correct electroweak symmetry breaking scale. Unlike the ‘bona-fide composite Higgs,’ however,  $g_*$  is still weak and thus no additional tuning is required to keep the Higgs light.

The holographic Higgs also has a version of collective symmetry breaking that is a result of locality in 5D [162]. Unlike the little Higgs models above, however, holographic Higgs models have radiative quartics. These models have the minimal amount of tuning: just  $\xi$ , which is a tuning of a few percent.

#### 4.7.4 Twin Higgs and neutral naturalness

Twin Higgs models [172, 173] have received a lot of interest after the non-discovery of any top-partners at Run I of the LHC. Rather than protecting the pseudo-Goldstone Higgs from quadratic corrections with collective symmetry breaking, these models impose a  $\mathbb{Z}_2$  symmetry that protects the Higgs potential. The key phenomenological feature of this framework is that the partner particles that enact this protection are uncharged under the Standard Model. Since the top partners aren’t colored, one no longer expects a large production cross section at the LHC and one avoids the Run 1 bounds. These models are thus often referred to under the banner of ‘neutral naturalness’ and are considered a last bastion for naturalness against collider bounds.

We illustrate the twin mechanism with the toy example presented in [172]; the interested reader is encouraged to read the succinct paper in its entirety. Suppose a theory has a global  $G = \text{SU}(4)$  symmetry and a field  $H$  in the fundamental representation with a symmetry-breaking potential,

$$V(H) = -\mu^2 |H|^2 + \lambda |H|^4. \quad (4.89)$$

The field develops a VEV  $\langle |H| \rangle = m/\sqrt{2\lambda} \equiv f$  and breaks  $\text{SU}(4) \rightarrow \text{SU}(3)$ . Now let us gauge a subgroup  $\text{SU}(2)_A \times \text{SU}(2)_B$  of the global symmetry. We decompose  $H$  into a doublet under each gauge group,  $H_A$  and  $H_B$ . We may identify  $A$  with the Standard Model  $\text{SU}(2)_L$ . As we saw in Section 4.4.3, this gauging generates mass terms for the would-be Goldstone bosons,

$$V \supset \frac{9\Lambda^2}{64\pi^2} (g_A^2 |H_A|^2 + g_B^2 |H_B|^2). \quad (4.90)$$

Next impose a  $\mathbb{Z}_2$  ‘twin’ symmetry which swaps  $A \leftrightarrow B$ . This imposes  $g_A = g_B$  so that the quadratic potential becomes,

$$V \supset \frac{9g^2\Lambda^2}{64\pi^2} |H|^2 + \dots, \quad (4.91)$$

which respects the original SU(4) symmetry of the theory and thus does not contribute to the mass of the Goldstone bosons. The higher order terms still introduce logarithmically divergent terms that break this SU(4) symmetry.

We can see the ‘twin’ cancellation in the top couplings:

$$\mathcal{L} \supset -y_t H_A \bar{t}_L^{(A)} t_R^{(A)} - y_t H_B \bar{t}_L^{(B)} t_R^{(B)}. \quad (4.92)$$

The SU(4)→SU(3) breaking imposes  $\langle h_a \rangle^2 + \langle h_b \rangle^2 = f^2$ . Expanding the SU(4) fundamental  $H$  analogously to (4.40), one may expand to  $\mathcal{O}(h^2/f^2)$ ,

$$H_A \rightarrow h, \quad H_B \rightarrow f - \frac{h^2}{2f}. \quad (4.93)$$

Inserting this into (4.92) yields a cancellation that is diagrammatically identical to (4.79) with the important difference that the  $t^{(B)}$  and  $\bar{t}^{(B)}$  are charged under a ‘twin’ QCD, but not ordinary QCD.

Having demonstrated the basic principle, we refer the reader to the original literature for a demonstration of a complete model. In our phenomenological taxonomy of composite Higgs models, we have written  $b \sim \mathcal{O}(1 - 16\pi^2/g_*^2)$  reflecting that the original twin Higgs models included a tree-level quartic put in by hand to generate the  $v \ll f$  hierarchy. As we have discussed, the observed  $\lambda = 0.13$  disfavors the inclusion of this tree-level term.

#### 4.7.5 Dilatonic Higgs

Rather than being a pseudo-Goldstone of an internal global symmetry, this scenario assumes that the Higgs is a **dilaton** coming from the spontaneous breaking of scale invariance [91, 174–179]. We have already explored this scenario in Section 3.10, where we identified the **radion** in a warped extra dimension as a state which is holographically dual to the dilaton. This is distinct from the ‘holographic Higgs’ scenario where the Higgs is the Goldstone of an internal global symmetry.

In this scenario the dilaton VEV,  $f$ , sets the scale of the potential and is unrelated to the electroweak VEV. Thus the parameterization in (4.84) is not relevant for comparison with the other composite Higgs models discussed: the dilaton is a completely different type of beast. The explicit breaking terms in the dilaton potential come from the explicit breaking from the running of the couplings. The scale  $f$  is naturally separated from the UV scale only when  $g_{\text{SM}} \sim 4\pi$  at the condensation scale, implying that the symmetry-breaking potential is driven by either the top Yukawa or new physics.

Unlike the other composite Higgs scenarios where phenomenology prefers  $f \gg v$ , the dilaton resembles the SM Higgs when  $f \approx v$ . Calculations using AdS/CFT, however, imply that  $f \gg v$  in the large  $N$  limit where these calculations are valid.

### 4.8 Parameterization of phenomenology

We see that the composite Higgs can be probed through its deviations from the SM Higgs couplings. A phenomenological parameterization of the space of light, composite Higgs models is presented in [180] as the ‘strongly interacting light Higgs’ (SILH, pronounced “silch”) effective Lagrangian and revised in [181]. We briefly review the main results and refer to [180, 181] for a detailed discussion including matching to specific composite Higgs models. We now examine a convenient parameterization of the phenomenological Lagrangian,

$$\begin{aligned} \mathcal{L}_{\text{SILH}} = & \mathcal{L}_H^{\text{SM}} + \frac{\bar{c}_H}{2v^2} \partial^\mu (H^\dagger H) \partial_\mu (H^\dagger H) + \frac{\bar{c}_T}{2v^2} \left( H^\dagger \overleftrightarrow{D} H \right)^2 - \frac{\bar{c}_6}{v^2} \lambda (H^\dagger H)^3 \\ & + \left( \frac{\bar{c}_u}{2v^2} y_u H^\dagger H \bar{Q}_L \tilde{H} u_R + \dots \right) + \frac{i\bar{c}_W g}{2M_W^2} \left( H^\dagger \sigma^i \overleftrightarrow{H} \right) (D^\nu W_{\nu\mu})^i + \dots, \end{aligned} \quad (4.94)$$



where  $H^\dagger \overleftrightarrow{D}_\mu H = H^\dagger D_\mu H - D_\mu H^\dagger H$  and the  $\dots$  represent similar terms for the other fermions and gauge bosons. The expected sizes of these coefficients are

$$\bar{c}_H, \bar{c}_T, \bar{c}_6, \bar{c}_\psi \sim \frac{v^2}{f^2} \qquad \bar{c}_{W,B} \sim \frac{M_W^2}{g_\rho^2 f^2}. \quad (4.95)$$

Here  $g_\rho$  parameterizes the size of the non-Goldstone composite states, for example in chiral perturbation theory the first such state is the spin-1  $\rho$  meson where  $g_\rho$  is defined to be

$$m_\rho = g_\rho f. \quad (4.96)$$

In this sense  $g_\rho$  is a ratio of mass scales, but when one includes this state in chiral perturbation theory (using the CCWZ formalism introduced in Appendix B), this ratio is manifestly the value of the  $\rho\pi\pi$  coupling. Following [181], the operators in (4.94) are normalized with respect to the Higgs VEV  $v$  rather than the scale  $f$  in [180]; this is why the expected values of the barred couplings  $\bar{c}_i$  differ by factors of  $v/f$  from the couplings  $c_i \sim 1$  in equation (15) of [180].

The phenomenological Lagrangian (4.94) can be constructed systematically from the non-linear sigma model including symmetry breaking terms which we assume are parameterized by the SM couplings that break those symmetries: the Higgs quartic coupling  $\lambda$  (breaking the pseudo-Goldstone Higgs shift symmetry) and the Yukawas (breaking shift and flavor symmetries). See Appendix B of [182] for a detailed discussion in terms of naïve dimensional analysis. The general strategy is to write

$$\mathcal{L}_{\text{SILH}} = \frac{m_\rho^4}{g_\rho^2} \tilde{\mathcal{L}} \left( U, \frac{\partial}{m_\rho} \right), \quad (4.97)$$

where  $U$  is the dimensionless linear field containing the Goldstones, analogous to (4.11) and the partial derivative carries a factor of  $m_\rho^{-1}$  to make it dimensionless. Note that the expansion of  $U$  in Goldstones  $\pi$  automatically comes with factors of  $f^{-1}$  to keep each term dimensionless. In this way the prefactor  $m_\rho^4 g_\rho^{-2} = m_\rho^2 f^2$  carries the dimension of the Lagrangian. Note that this is analogous to the prefactor  $\Lambda^2 f^2$  in naïve dimensional analysis [133–135] except that we replace  $\Lambda$  with a scale which exists in the effective theory,  $m_\rho < \Lambda$ . This inequality is equivalent to  $g_\rho < 4\pi$  and parameterizes the regime in which the NLSM is weakly coupled. We then take the dimensionless function  $\tilde{\mathcal{L}}$  to be a derivative expansion analogous to (4.18); the SILH interactions appear at higher order from the term

$$\mathcal{L}_{\text{SILH}} = \frac{m_\rho^4}{g_\rho^2} \left( \dots + \frac{1}{3} \left( \frac{\pi(x)}{f} \frac{\overleftrightarrow{\partial}}{m_\rho} \frac{\pi(x)}{f} \right)^2 + \dots \right) \quad (4.98)$$

where  $\pi(x)$  is identified with the Higgs doublet  $H(x)$  and the partial derivatives  $\partial_\mu$  are promoted to SM gauge covariant derivatives. Gauge field strengths are included with factors of  $m_\rho^{-2}$  since  $F_{\mu\nu} \sim [D_\mu, D_\nu]$ . The  $\bar{c}_H$  and  $\bar{c}_T$  terms encode the  $\mathcal{O}(H^4, \partial^2)$  interactions after shifting the Higgs by a factor proportional to  $(H^\dagger H)H/f^2$  (see [180]). The  $\bar{c}_6$ ,  $\bar{c}_u$ ,  $\bar{c}_W$  (and analogous terms) break the shift symmetries of the NLSM and carry explicit factors of the SM couplings that break those symmetries: the Higgs quartic interaction, the Yukawas, or SM gauge couplings, respectively. Electroweak precision observables<sup>27</sup> set bounds on composite Higgs models [181, 188]; at  $2\sigma$ :

$$-1.5 \times 10^{-3} < \bar{c}_T < 2.2 \times 10^{-3} \quad (4.99)$$

$$-1.4 \times 10^{-3} < \bar{c}_W + \bar{c}_B < 1.9 \times 10^{-3}, \quad (4.100)$$

coming from the  $\hat{T}$  and  $\hat{S}$  parameters respectively. The former condition reflects the requirement of custodial symmetry [189] (see [190] for an introduction) which is assumed in the latter bound. The

<sup>27</sup>See [183, 184] for pedagogical reviews and [185–187] for details.

observation of the 125 GeV Higgs and the opportunity to measure its couplings offers additional data to fit the phenomenological Lagrangian. For example, the  $\bar{c}_H$  and  $c_f$  ( $f$  running over the SM fermions) are related to each other via the couplings of the Higgs to  $W$  bosons [191]. The Higgs mass sets (at  $3\sigma$ )

$$\bar{c}_H \leq 0.16. \quad (4.101)$$

This and other bounds on composite Higgs models coming from Higgs observables are reviewed in [192, 193] using a slightly different effective theory parameterization introduced in [138]. In that notation, (4.101) comes from  $a^2 \geq 0.84$ . Further phenomenological bounds and their relations to specific models can be found in [116, 180, 181, 194]. At  $2\sigma$ , Higgs data constraints the minimal composite Higgs model to satisfy [195]

$$\frac{v}{f} \lesssim 0.5. \quad (4.102)$$

The bounds on composite Higgs models coming from Higgs observables are reviewed in [181, 192, 193]. Further phenomenological bounds and their relations to specific models can be found in [116, 180, 194].

## 5 Closing Thoughts

We briefly review interconnections between some of the salient ideas in these lectures, acknowledge topics omitted, and point to directions of further study. One of the themes in the latter part of these lectures were weakly coupled descriptions of strong dynamics and we close by highlighting this common thread.

### 5.1 Covariant Derivatives

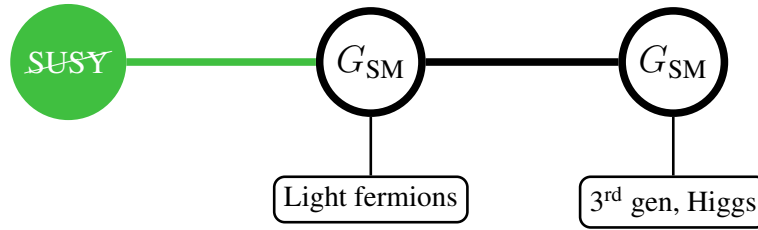
Each of the scenarios that we explored carries its own sense of covariant derivative. The most explicit example is in a warped extra dimension, where spacetime is explicitly curved. The holographic principle made use of this geometry: the isometries of AdS match the conformal symmetries of the strongly coupled theory near a fixed point. The system is so constrained by these symmetries that the behavior of 5D fields could be identified with the renormalization group flow of 4D operators. Even in supersymmetry—where superspace can be thought of a ‘fermionic’ extra dimension—we introduced a SUSY covariant derivative. Even though superspace is flat, there is a covariant derivative came from it being torsion-free [?]. Finally, the nonlinear realizations we used for composite Higgs models also has a geometric structure coming from the coset space. This is seen in the CCWZ formalism reviewed in Appendix B, where one identifies covariant derivative and gauge field for the coset space that are necessary to construct invariant Lagrangians.

### 5.2 Nonlinear realizations

The simplest handle on strong dynamics is to work in an effective theory of pseudo-Goldstone bosons given by the pattern of global symmetry breaking in the strong sector. In composite Higgs models, one addresses the Hierarchy problem by assuming that the Higgs is a pseudo-Goldstone boson associated with the dynamics of a strongly coupled sector that break global symmetries at a scale  $f$ . We saw that generically the SM interactions required for a Higgs boson tend to push its mass back up towards the compositeness scale,  $\Lambda \sim 4\pi f$ . One way to push the Higgs mass back down is to invoke collective symmetry breaking, which can often be described succinctly using ‘moose’ diagrams.

### 5.3 Holographic and deconstructed extra dimensions

We saw that the holographic principle is an alternative way to describe the dynamics of a strongly coupled sector through the use of a higher dimensional theory with a non-trivial geometry. From the extra dimensional perspective, the Higgs mass is natural because it is localized towards the IR brane where



**Fig. 15:** Schematic moose diagram for natural SUSY.

the Planck scale is warped down to TeV scale. Holographically, this is interpreted as the Higgs being a composite state. Indeed, the minimal composite Higgs model [136] was constructed using holography as a guiding principle.

Further, the little Higgs models that first highlighted collective symmetry breaking were motivated by deconstructions of an extra dimension. This picture allowed us to clearly see that the Goldstone modes of the spontaneously broken global symmetries can be identified with the scalar component of a 5D gauge field. In the deconstruction, the gauge bosons from each copy of the gauge group eat these Goldstone modes to become the spectrum of heavy KK modes. In this sense, little Higgs and holographic composite Higgs constructions are similar to gauge-Higgs unification scenarios in 5D where the Higgs is the zero mode of a bulk gauge field, see [196] and references therein. One way to interpret the lightness of the Higgs mass is via locality in the deconstructed extra dimension: the symmetries are only broken on the boundaries and one needs a loop that stretches between the boundaries to generate a Higgs potential. This implies that the loop cannot be shrunk to zero and that the Higgs potential is finite since it can have no short-distance divergences. The natural cutoff is set by the size of the extra dimension.

Deconstruction itself, however, is rooted in the idea of a **hidden local symmetry** in nonlinear models. See [113] for a comprehensive review. A 5D version of the little Higgs in AdS was presented in [197]. Shortly after, [162] connected the holographic composite Higgs to a little Higgs theory, relating the CCWZ formalism of Appendix B to the hidden local symmetry construction.

#### 5.4 Natural SUSY and partial compositeness

We began these lectures with what appeared to be a completely different subject: supersymmetry. We saw that the natural setting for SUSY is superspace, which is superficially an ‘extra quantum dimension’ that is both Grassmannian and spinorial. One way to see how SUSY solves the Hierarchy problem is to observe that it requires the existence of superpartners (differing by half integer spin) that cancel the loop contributions of particles to superpotential parameters such as the Higgs mass. We saw a similar cancellation when invoking collective symmetry breaking (or the twin Higgs mechanism) in composite Higgs theories with the notable difference that the partner particles had the same spin as their SM counterparts.

SUSY, however, must be broken. These effects feed into the large parameter space of the minimal supersymmetric Standard Model and are required (in the MSSM) for electroweak symmetry breaking. The LHC puts tight bounds on the simplest MSSM spectra and leads us to consider ways to hide SUSY. One of these solutions is ‘natural SUSY’ where one only maintains the minimal spectrum of superpartners required for the naturalness of the Higgs mass. Among the predictions of natural SUSY is a light stop and heavy first and second generation quarks. This type of spectrum, however, is automatic when supersymmetrizing the RS model with anarchic flavor<sup>28</sup>. When SUSY is broken on the UV brane<sup>29</sup>, 5D *superfields* which are localized near the UV brane are more sensitive to the splitting between the SM and

<sup>28</sup>One should note that because 5D spinors are Dirac,  $\mathcal{N} = 1$  SUSY in 5D corresponds to  $\mathcal{N} = 2$  SUSY in 4D.  $\mathcal{N} = 2$  was used in [198] to generate Dirac gaugino masses, which can help soften the two-loop quadratic corrections to the Higgs mass. See [199] for a recent analysis of prospects.

<sup>29</sup>This is of the ways to interpret anomaly mediation of SUSY breaking [200].

FIELD	5D LOCALIZATION	4D INTERPRETATION	SUPERPARTNER
Higgs	IR localized	Composite state	$\approx$ degenerate, mixes with $\tilde{B}, \tilde{W}$
Top	Peaked toward IR	Mostly composite	Slightly heavier than top
Light quarks	Peaked toward UV	Mostly elementary	Large SUSY breaking masses

Table 3: Holographic picture of natural SUSY spectra. Superfields localized near the IR brane have a large overlap with the Higgs so that the SM component of the superfield picks up a large mass. Superfields localized near the UV brane have a large overlap with SUSY breaking so that the ‘superpartner’ component of the superfield picks up a large mass. Thus light SM fermions have heavy superpartners and vice versa.

superpartner masses. Invoking what we know about the anarchic flavor 5D mass spectrum (i.e. localization of the fermion profiles), we come to the conclusions in Table 3. Holographically this is interpreted as supersymmetry being an accidental symmetry in the IR. That is, the strong sector flows to a fixed point that is supersymmetric, even though the theory at the UV is not manifestly supersymmetric. As a particle becomes more composite, it becomes more degenerate in mass with its superpartner. A schematic moose diagram is shown in Fig. 15; note that one of the sacrifices of this realization of natural SUSY is conventional unification, see e.g. [201].

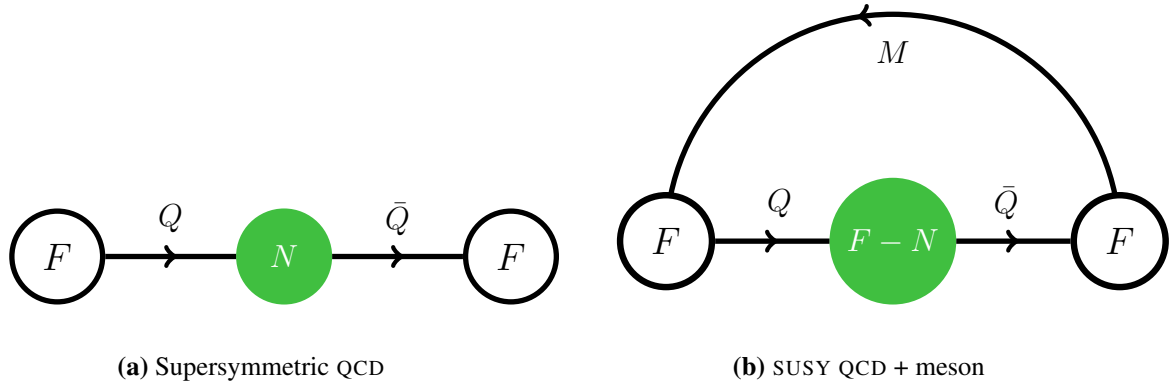
### 5.5 Naturalness and top partners

The three classes of physics beyond the Standard Model that we have explored all generically predict new particles accessible at high energy colliders. For supersymmetry and extra dimensions, these particles were a manifestation of the extended spacetime symmetry under which the SM particles must transform. For a composite Higgs, this reflected a larger global symmetry breaking pattern and included additional fermions that appear necessary to generate an SM-like Higgs potential. At a technical level, we needed new particles to run in Higgs loops to soften the quadratic sensitivity to the cutoff. Since the top quark has the largest coupling to the Higgs, a generic prediction for naturalness are light (i.e. accessible at the LHC) states to cancel the top loop. While these particles may have different spin, the examples we’ve explored focused on the case where they have the same SM quantum numbers as the top. The color charge of these new particles make them easy to produce at the LHC so that their non-observation is particularly disconcerting. One model building direction out of this puzzle is to consider models where the top partner is not color charged. We saw this in the twin Higgs model in Section 4.7.4. A supersymmetric cousin on these models go under the name of folded SUSY [202], where the top partners are uncolored but still carry electroweak charges.

### 5.6 Seiberg duality

In the composite Higgs models, same-spin partners cancelled the leading SM particle contributions to the quadratic cutoff contributions to the Higgs mass. We saw that this was not coincidental, but rather imposed by the structure of collective symmetry breaking. In the same way, the protection of the Higgs mass in SUSY is most clearly understood from the tremendous constraints put on the theory by supersymmetry. Among other things, these constraints imposed that the holomorphy of the superpotential which, in turn, prevents the perturbative renormalization of any of the superpotential terms—the Higgs mass being just one example. A derivation of this important result is beyond the scope of these lectures, but is explored—along with further implications of SUSY—in the reviews already mentioned.

Supersymmetry turns out to also be a powerful constraint on the behavior of gauge theories. In fact, they allow one to map out the entire phase structure of the supersymmetric generalization of QCD, SQCD. This, in itself, is a topic of depth and elegance which is covered very well in [?, 2, 5–7]. One key outcome of this exploration in the 1990s was the observation that two distinct supersymmetric non-



**Fig. 16:** Moose diagrams for a pair of Seiberg duals. Green nodes are gauged symmetries while white nodes are global symmetries. Note that the lines now represent superfields.

Abelian gauge theories, shown in Fig. 16, flow to the same IR fixed point. One theory, SQCD, is a standard  $SU(N)$  supersymmetric gauge theory with  $F$  flavors such that

$$F > N + 1. \quad (5.1)$$

In the case where  $N + 1 < F < \frac{3}{2}N$ , this theory is asymptotically free and confines in the IR. The other theory is an  $SU(F - N)$  gauge theory with  $F$  flavors and an additional color singlet ‘meson’ which is a bifundamental under the  $SU(F) \times SU(F)$  flavor symmetry. This theory has a superpotential,

$$W \sim \bar{Q}MQ, \quad (5.2)$$

which can be understood as a loop in the moose diagram since all indices are contracted. In the case where  $N + 1 < F < \frac{3}{2}N$ , the dual theory is IR free and is perturbative at the fixed point.

The fact that these two *a priori* unrelated theories flow to the same fixed point suggest a compelling interpretation: the asymptotically free theory confines at low energies and is replaced by the effective description of the IR free theory. This is an ‘electromagnetic’ duality in the sense of exchanging strongly and weakly coupled descriptions of the same physics, similar to the AdS/CFT correspondence.

This **Seiberg duality** is a powerful handle on strongly coupled physics via a weakly coupled 4D dual description. One popular application was to simplify the construction of models with dynamical SUSY breaking, see [7, 203] for reviews. In some sense this is completely analogous to using chiral perturbation theory to describe low-energy QCD. However, unlike QCD, the low energy (‘magnetic’) theory is not composed of gauge singlets. In fact, there is an *emergent*  $SU(F - N)$  gauge symmetry that appears to have nothing to do with the original  $SU(N)$  gauge symmetry of the ‘electric’ theory. One recent interpretation, however, is that this magnetic gauge group can be identified with the ‘hidden local symmetry’ in nonlinear models [204], which we previously mentioned in the context of deconstruction and moose models. In this construction, the  $\rho$  meson in QCD (the lightest spin-1 meson) is identified as the massive gauge boson of a spontaneously broken gauge symmetry present in the nonlinear Lagrangian.

One can also relate Seiberg duality to the AdS/CFT correspondence through explicit string realizations. Note that (5.1) is typically a different regime from the large  $N$  limit invoked in AdS/CFT. From a purely field theoretical point of view, the AdS/CFT correspondence can be understood as a **duality cascade** where a SUSY gauge theory has a renormalization trajectory that zig-zags between a series of fixed points. This is reviewed pedagogically for a field theory audience in [85].

### 5.7 Multiple guises of strong dynamics

In this final section we have touched on multiple ways in which we can address strong dynamics in field theory: nonlinear realizations based on the symmetry breaking structure, holographic extra dimensions, and Seiberg duality in SUSY. The lesson to take away from this overview is that one should be flexible to think about strong dynamics in different languages. Often the intuition from one understanding of strong dynamics can shed light on constructions based on a different description.

One example is the use of Seiberg duality to describe a [partially-]composite electroweak sector based on the ‘fat Higgs’ model [205]. The idea is to take super-QCD with  $F = N + 2$  flavors so that the magnetic gauge group can be linked with  $SU(2)_L$ . The realization of this idea in [142] described this in terms of moose diagrams where the magnetic gauge group is ‘color-flavor locked’ with an externally gauged  $SU(2)_L$ . This mixes the magnetic gauge bosons with the external gauge bosons so that the observed  $W$  and  $Z$  are partially composite. Independently, a similar model was presented in [143] where the nature of this mixing was explained in terms of the intuition from a warped extra dimension. In particular, one hope that one could directly identify the magnetic  $SU(2)$  with the electroweak  $SU(2)_L$ . This, however, is not possible since—as we know from composite model building—at the compositeness scale the naïve dimensional analysis expectation is that the composite vector boson couples strongly:  $g \sim 4\pi/\sqrt{N}$ . In other words, if the electroweak gauge bosons are strongly coupled bound states, then one would expect a large residual interaction with other strongly coupled bound states. In the RS language, a composite  $W$  and  $Z$  would have IR brane localized profiles and this would typically predict very strong couplings. This would require a very large running to squeeze the profile on the IR brane. In the Seiberg dual picture, this requires a very large number of flavors if one maintains that the  $W$  and  $Z$  are purely composite but have the observed SM couplings, leading one to prefer partial compositeness of these particles. This general framework was later used to construct a model of natural SUSY in which follows the general deconstruction/moose in Fig. 15 [206].

### 5.8 Omissions

We have necessarily been limited in scope. Even among the topics discussed, we have omitted an exploration of SUSY gauge theories (leading up to Seiberg duality), variants of the ‘realistic’ RS models (as well as ‘universal extra dimension’ models), the virtues of different cosets for composite Higgs model building, and an overview of product space (moose-y) little Higgs models. Many explicit calculations were left out and are left to the dilligent reader as exercises, and we only made cursory nods to the phenomenology of these models. In addition to the three major topics covered in these lectures, there are various other extensions to the Standard Model that we have not discussed. Our preference focused on models that address the Higgs hierarchy problem, and as such we have omitted discussions of many important topics such as grand unification, dark matter, flavor, strong CP, cosmology (of which the cosmological constant is the most extreme fine tuning problem), or any of the phenomenology of interpreting possible experimental signals from colliders/telescopes/underground experiments/etc. We have only presented very cursory comparisons to current data; we refer the reader to the appropriate experiments’ results pages and conference proceedings for the latest bounds. See also [207] for an overview of the Run I searches for new physics. All of these topics—and perhaps many others—are, in some combination, key parts of a model builder’s toolbox in the LHC era.

### Acknowledgements

C.C. thanks the organizers of the 2013 European School of High-Energy Physics for the invitation to give these lectures. P.T. is grateful to Mohammad Abdullah, Brando Bellazzini, Jack Collins, Nathaniel Craig, Anthony DiFranzo, Roni Harnik, Javi Serra, Yuri Shirman, Tim M.P. Tait, and Daniel Stolarski for useful comments and discussions. We thank Gilad Perez and Martijn Mulders for their patience with this manuscript. We thank the Aspen Center for Physics (NSF grant #1066293) for its hospitality during a

period where part of this work was completed. C.C. is supported in part by the NSF grant PHY-1316222. P.T. is supported in part by the NSF grant PHY-1316792 and a UCI Chancellor's ADVANCE fellowship.

## Appendices

### A Appendix: Extra Dimensions

#### A.1 The RS gravitational background

We have assumed the metric (3.28). In this appendix we derive it from the assumption of a non-factorizable metric of the form

$$ds^2 = e^{-A(z)} (\eta_{\mu\nu} dx^\mu dx^\nu - dz^2) \quad (\text{A.1})$$

and check the conditions for which a flat 4D background exists. This generic form of the metric is useful since it is an overall rescaling of the flat metric, that is, it is conformally flat. We can thus use a convenient relation between the Einstein tensors  $G_{MN} = R_{MN} - \frac{1}{2}g_{MN}R$  of two conformally equivalent metrics  $g_{MN} = e^{-A(x)}\tilde{g}_{MN}$  in  $d$  dimensions [?],

$$G_{MN} = \tilde{G}_{MN} + \frac{d-2}{2} \left[ \frac{1}{2} \tilde{\nabla}_M A \tilde{\nabla}_N A + \tilde{\nabla}_M \tilde{\nabla}_N A - \tilde{g}_{MN} \left( \tilde{\nabla}_K \tilde{\nabla}^K A - \frac{d-3}{4} \tilde{\nabla}_K A \tilde{\nabla}^K A \right) \right]. \quad (\text{A.2})$$

When  $A = A(z)$  this is straightforward to calculate by hand for  $\tilde{g}_{MN} = \eta_{MN}$ . Alternately, one may use a computer algebra system to geometric quantities for general metrics, e.g. [?]. We assume a bulk cosmological constant  $\Lambda$  so that the 5D bulk Einstein action is

$$S = - \int d^5x \sqrt{g} (M_*^3 R + \Lambda). \quad (\text{A.3})$$

The Einstein equation is  $G_{MN} = (M_*)^{-3} T_{MN}$ . The  $MN = 55$  component gives

$$\frac{3}{2} A'^2 = \frac{1}{2M_*^3} \Lambda e^{-A}. \quad (\text{A.4})$$

This only has a solution for  $\Lambda < 0$  so that we're forced to consider AdS spaces. This equation is separable with the general solution,

$$e^{-A(z)} = \frac{1}{(kz + \text{constant})^2} \quad k = \frac{-\Lambda}{12M_*^3}. \quad (\text{A.5})$$

To recover (3.28) we identify  $R = 1/k$  and impose  $A = 0$  at  $z = R$ , setting the constant to zero. The latter choice simply sets the warp factor at the UV brane to be 1.

We must remember that the RS space is finite—and has branes at its endpoints—when we solve the  $MN = \mu\nu$  Einstein equations. These equations depend on the second derivative of  $A(z)$  and one should be concerned that this may be sensitive to the energy densities on the branes. This is analogous to the Poisson equation in electrostatics where a second derivative picks up the  $\delta$ -function of a point charge. In general the branes carry tensions which appear as 4D cosmological constants,  $\Lambda_{\text{IR,UV}}$ . Recalling the form of the induced metric  $\sqrt{\hat{g}} = \sqrt{g/g_{55}}$ , these appear in the action as

$$\int d^5x \sqrt{\frac{g}{g_{55}}} \Lambda_{\text{ir,uv}} \delta(z - R') \quad \Rightarrow \quad T_{\mu\nu} = \frac{1}{\sqrt{g}} \frac{\delta S}{\delta g^{\mu\nu}} = \frac{g_{\mu\nu}}{2\sqrt{g_{55}}} (\Lambda_{\text{IR}} \delta(z - R') + \Lambda_{\text{UV}} \delta(z - R)). \quad (\text{A.6})$$

## A.2 RS as an orbifold

To better understand the physics of the brane cosmological constants, it is useful to represent the interval with an orbifold  $S^1/\mathbb{Z}_2$ . This is simply the circle  $y \in [-\pi, \pi]$  with the identification  $y = -y$ . While this may sound somewhat exotic, such compactifications are common in string theory, and was the original formulation of the RS scenario. Note that  $y$  can take any value due to the periodic identification of the circle, while the fixed points at  $y = 0, \pi$  demarcate the physical RS space.

The orbifold identification forces us to modify (A.5) by replacing  $z \rightarrow |z|$  to preserve the  $z \leftrightarrow -z$  symmetry. This absolute value, in turn, leads to  $\delta$  functions in  $A''(z)$  at the fixed points,

$$A''(z) = -\frac{2k^2}{(k|z| + \text{const})^2} + \frac{4k}{k|z| + \text{const}} (\delta(z - R) - \delta(z - R')). \quad (\text{A.7})$$

The  $\mu\nu$  Einstein equation then implies

$$-\frac{3}{2}\eta_{\mu\nu} \left[ -\frac{4k(\delta(z - R) - \delta(z - R'))}{k|z| + \text{const}} \right] = \frac{\eta_{\mu\nu}}{2M_*^3} \left[ \frac{\Lambda_{\text{UV}}\delta(z - R) + \Lambda_{\text{IR}}\delta(z - R')}{k|z| + \text{const}} \right]. \quad (\text{A.8})$$

From this we see that the brane cosmological constants must have opposite values,

$$\Lambda_{\text{UV}} = -\Lambda_{\text{IR}} = 12kM_*^3. \quad (\text{A.9})$$

Recall, further, that  $k$  is related to the bulk cosmological constant by (A.5), so that this represents a tuning of the bulk and brane cosmological constants. This is a necessary condition for a static, gravitational solution. Physically, we see that the brane and the bulk cosmological constants are balanced against one another to cause the brane to be flat.

## A.3 Bulk Fermions in RS

The properties of fermions in a curved space can be subtle. In particular, it's not clear how to generalize the usual Dirac operator, Dirac operator  $i\gamma^\mu\partial_\mu$ . In this appendix we review properties of fermions in an extra dimension and then derive the form of the fermion action in RS.

### A.3.1 The fifth $\gamma$ matrix

Firstly, unlike in 4D where the fundamental fermion representation is a Weyl spinor, 5D Lorentz invariance requires that fermions appear as Dirac spinors. A simple heuristic way of seeing this is to note that in 4D one can construct a  $\gamma^5 \sim \gamma^0 \cdots \gamma^3$  as a linearly independent chirality operator. In 5D, however,  $\gamma^5$ , is part of the 5D Clifford algebra and is just a normal  $\gamma$  matrix in the  $z$ -direction. Note that the normalization of  $\gamma^5$  is fixed by  $\{\gamma^5, \gamma^5\} = 2\eta^{55}$  and has a factor of  $i$  compared to the usual definition in 4D. One should immediately be concerned: if the 5D fermions are Dirac, then how does one generate the *chiral* spectrum of the Standard Model matter? As we show below, this follows from a choice of boundary conditions. An excellent reference for the properties of fermions in arbitrary dimension is [208].

### A.3.2 Vielbeins

In order to write down the fermionic action, we first need to establish some differential geometry so that we may write the appropriate covariant derivative for the spinor representation. We will be necessarily brief here, but refer to [209–211] for the interested reader<sup>30</sup>.

The familiar  $\gamma$  matrices which obey the Clifford algebra are only defined for flat spaces. That is to say that they live on the tangent space (locally inertial frame) of our spacetime manifold. In order

<sup>30</sup>Those with slightly less mathematical background can refer to [212] or their favorite ‘grown up’ general relativity textbook as a starting point.



to define curved-space generalizations of objects like the Dirac operator, we need a way to convert spacetime indices  $M$  to tangent space indices  $a$ . **Vielbeins**,  $e_\mu^a(x)$ , are the geometric objects which do this. The completeness relations associated with vielbeins allow them to be interpreted as a sort of “square root” of the metric in the sense that

$$g_{MN}(x) = e_M^a(x)e_N^b(x)\eta_{ab}, \quad (\text{A.10})$$

where  $\eta_{ab} = \text{diag}(+, -, \dots, -)$  is the Minkowski metric on the tangent space. For our particular purposes we need the inverse vielbein,  $e_a^M(x)$ , defined such that

$$e_a^M(x)e_N^a(x) = \delta_N^M \quad e_a^M(x)e_N^b(x) = \delta_a^b. \quad (\text{A.11})$$

Spacetime indices are raised and lowered using the spacetime metric  $g_{MN}(x)$  while tangent space indices are raised and lowered using the flat (tangent space) metric  $\eta_{ab}(x)$ .

Physically we may think of the vielbein in terms of reference frames. The equivalence principle states that at any point one can always set up a coordinate system such that the metric is flat (Minkowski) at that point. Thus for each point  $x$  in space there exists a family of coordinate systems that are flat at  $x$ . For each point we may choose one such coordinate system, which we call a frame. By general covariance one may define a map that transforms to this flat coordinate system at each point. This is the vielbein. One can see that it is a kind of local gauge transformation, and indeed this is the basis for treating gravity as a gauge theory built upon diffeomorphism invariance.

### A.3.3 Spin covariant derivative

The covariant derivative is composed of a partial derivative term plus connection terms which depend on the particular object being differentiated. For example, the covariant derivative on a spacetime vector  $V^\mu$  is

$$D_M V^N = \partial_M V^N + \Gamma_{ML}^N V^L. \quad (\text{A.12})$$

The vielbein allows us to work with objects with a tangent space index,  $a$ , instead of just spacetime indices,  $\mu$ . The  $\gamma$  matrices allow us to further convert tangent space indices to spinor indices. We would then define a covariant derivative acting on the tangent space vector  $V^a$ ,

$$D_M V^a = \partial_M V^a + \omega_{Mb}^a V^b, \quad (\text{A.13})$$

where the quantity  $\omega_{Mb}^a$  is called the spin covariant derivative. Consistency of the two equations implies

$$D_M V^a = e_N^a D_M V^N. \quad (\text{A.14})$$

This is sufficient to determine the spin connection. It is a fact from differential geometry that the spin connection is expressed in terms of the vielbeins via [213]

$$\omega_M^{ab} = \frac{1}{2}g^{RP}e_R^{[a}\partial_{[M}e_{P]}^{b]} + \frac{1}{4}g^{RP}g^{TS}e_R^{[a}e_T^{b]}\partial_{[S}e_{P]}^c e_M^d \eta_{cd} \quad (\text{A.15})$$

$$= \frac{1}{2}e^{Na} \left( \partial_M e_N^b - \partial_N e_M^b \right) - \frac{1}{2}e^{Nb} \left( \partial_M e_N^a - \partial_N e_M^a \right) - \frac{1}{2}e^{Pa}e^{Rb} \left( \partial_P e_{Rc} - \partial_R e_{Pc} \right) e_M^c. \quad (\text{A.16})$$

When acting on spinors one needs the appropriate structure to convert the  $a, b$  tangent space indices into spinor indices. This is provided by

$$\sigma_{ab} = \frac{1}{4} [\gamma_a, \gamma_b] \quad (\text{A.17})$$

so that the appropriate spin covariant derivative is

$$D_M = \partial_M + \frac{1}{2}\omega_M^{ab}\sigma_{ab}. \quad (\text{A.18})$$

### A.3.4 Antisymmetrization and Hermiticity

The fermionic action on a  $d$ -dimensional curved background is

$$S = \int d^d x \sqrt{|g_d|} \bar{\Psi} \left( i e_a^M \gamma^a \overleftarrow{D}_M - m \right) \Psi, \quad (\text{A.19})$$

where the antisymmetrized covariant derivative is defined by a difference of right- and left-acting derivatives

$$\overleftarrow{D}_M = \frac{1}{2} D_M - \frac{1}{2} \overleftarrow{D}_M. \quad (\text{A.20})$$

This is somewhat subtle. The canonical form of the fermionic action must be antisymmetric in this derivative in order for the operator to be Hermitian and thus for the action to be real. In flat space we are free to integrate by parts in order to write the action exclusively in terms of a right-acting Dirac operator. Hermiticity is defined with respect to an inner product. The inner product in this case is given by

$$\langle \Psi_1 | \mathcal{O} \Psi_2 \rangle = \int d^5 x \sqrt{g} \bar{\Psi}_1 \mathcal{O} \Psi_2. \quad (\text{A.21})$$

A manifestly Hermitian operator is  $\mathcal{O}_H = \frac{1}{2} (\mathcal{O} + \mathcal{O}^\dagger)$ , where we recall that

$$\langle \Psi_1 | \mathcal{O}^\dagger \Psi_2 \rangle = \langle \mathcal{O} \Psi_1 | \Psi_2 \rangle = \int d^5 x \sqrt{g} \overline{\mathcal{O} \Psi_1} \Psi_2. \quad (\text{A.22})$$

The definition of an inner product on the phase space of a quantum field theory can be nontrivial on curved spacetimes. However, since our spacetime is not warped in the time direction there is no ambiguity in picking a canonical Cauchy surface to quantize our fields and we may follow the usual procedure of Minkowski space quantization with the usual Minkowski spinor inner product.

As a sanity-check, consider the case of the partial derivative operator  $\partial_\mu$  on flat space time. The Hermitian conjugate of the operator is the left-acting derivative,  $\overleftarrow{\partial}_\mu$ , by which we really mean

$$\int d^d x \bar{\Psi}_1 \partial^\dagger \Psi_2 = \langle \Psi_1 | \partial^\dagger \Psi_2 \rangle = \langle \partial_\mu \Psi_1 | \Psi_2 \rangle = \int d^d x \overline{\partial_\mu \Psi_1} \Psi_2 = \int d^d x \bar{\Psi}_1 \overleftarrow{\partial}_\mu \Psi_2 = \int d^d x \bar{\Psi}_1 (-\partial_\mu) \Psi_2.$$

In the last step we've integrated by parts and dropped the boundary term. We see that the Hermitian conjugate of the partial derivative is negative itself. Thus the partial derivative is not a Hermitian operator. This is why the momentum operator is given by  $\hat{P}_\mu = i \partial_\mu$ , since the above analysis then yields  $\hat{P}_\mu^\dagger = \hat{P}_\mu$ , where we again drop the boundary term and recall that the  $i$  flips sign under the bar.

Now we can be explicit in what we mean by the left-acting derivative in (A.19). The operator  $i e_a^M \gamma^a D_M$  is not Hermitian and needs to be made Hermitian by writing it in the form  $\mathcal{O}_H = \frac{1}{2} (\mathcal{O} + \mathcal{O}^\dagger)$ . Thus we may write a manifestly Hermitian Dirac operator as,

$$\bar{\Psi} (\text{Dirac}) \Psi = \bar{\Psi} \left[ \frac{1}{2} (i e_a^M \gamma^a D_M) + \frac{1}{2} (i e_a^M \gamma^a D_M)^\dagger \right] \Psi \quad (\text{A.23})$$

$$= \bar{\Psi} \frac{i}{2} e_a^M \gamma^a D_M \Psi - \frac{i}{2} e_a^M \overline{\gamma^a D_M \Psi} \Psi, \quad (\text{A.24})$$

where we've used the fact that  $e_a^M$  is a real function with no spinor indices. The second term on the right-hand side can be massaged further,

$$\overline{\gamma^a D_M \Psi} \Psi = \Psi^\dagger \overleftarrow{D}_M^\dagger \gamma^{a\dagger} \gamma^0 \Psi = \Psi^\dagger (\overleftarrow{\partial}_M + \omega_M^{bc} \sigma^{bc\dagger}) \gamma^0 \gamma^a \Psi = \bar{\Psi} \overleftarrow{D}_M \gamma^a \Psi = \bar{\Psi} \gamma^a \overleftarrow{D}_M \Psi. \quad (\text{A.25})$$

Note that we have used that  $\gamma^{M\dagger} = \gamma^0 \gamma^M \gamma^0$  and, in the last line, that  $[\sigma^{bc}, \gamma^a] = 0$ . Putting this all together, we can write down our manifestly real fermion action as in (A.19),

$$S = \int d^d x \sqrt{|g|} \bar{\Psi} \left( i e_a^M \gamma^a \overleftrightarrow{D}_M - m \right) \Psi \quad (\text{A.26})$$

$$= \int d^d x \sqrt{|g|} \left( \frac{i}{2} \bar{\Psi} e_a^M \gamma^a D_M \Psi - \frac{i}{2} \overline{D}_M \bar{\Psi} e_a^M \gamma^a \Psi - m \bar{\Psi} \Psi \right). \quad (\text{A.27})$$

All of this may seem overly pedantic since integration by parts allows one to go back and forth between the ‘canonical’ form and the usual ‘right-acting only’ form of the fermion kinetic operator. Our interest, however, is to apply this to the Randall-Sundrum background where integration by parts introduces boundary terms and so it is crucial to take the canonical form of the Dirac operator as the starting point.

### A.3.5 Application to the RS background

We now apply this machinery to the RS background. The vielbein and inverse vielbein are

$$e_M^a(z) = \frac{R}{z} \delta_M^a \quad e_a^M(z) = \frac{z}{R} \delta_a^M. \quad (\text{A.28})$$

We may write out the spin connection term of the covariant derivative as

$$\omega_M^{ab} = \underbrace{\frac{1}{2} g^{RP} e_R^{[a} \partial_{[M} e_{P]}^{b]}}_{\omega_M^{ab(1)}} + \underbrace{\frac{1}{4} g^{RP} g^{TS} e_R^{[a} e_T^{b]} \partial_{[S} e_{P]}^d e_M^d \eta_{cd}}_{\omega_M^{ab(2)}}. \quad (\text{A.29})$$

This can be simplified using the fact that the vielbein only depends on  $z$ . The first part is

$$\omega_M^{ab(1)} = \frac{1}{2z} \delta_M^{[a} \delta_5^{b]}, \quad (\text{A.30})$$

where we’ve used  $\partial_M e_P^b = -\frac{1}{z} e_P^b \delta_M^5$  and the completeness relation  $g^{MN} e_M^a e_N^b = \eta^{ab}$ . Similarly, with some effort the second part is given by

$$\omega_M^{ab(2)} = \frac{1}{2z} \delta_M^{[a} \delta_5^{b]}. \quad (\text{A.31})$$

These vanish identically for  $M = 5$ . We can now write out the spin-connection part of the covariant derivative,

$$\frac{1}{2} \omega_M^{ab} \sigma_{ab} = \frac{1}{2} \left( \frac{1}{z} \delta_M^{[a} \delta_5^{b]} \right)_{M \neq 5} \frac{1}{4} [\gamma_a, \gamma_b] = \frac{1}{4z} (\gamma_M \gamma_5 + \delta_M^5), \quad (\text{A.32})$$

where we’ve inserted a factor of  $\delta_M^5$  to cancel the  $(\gamma_5)^2$  when  $M = 5$ . Finally, the spin connection part of the covariant derivative is

$$\frac{1}{2} \omega_M^{ab} \sigma_{ab} = \frac{1}{4z} (\gamma_M \gamma_5 + \delta_M^5) \quad (\text{A.33})$$

so that the spin covariant derivative is

$$D_M = \begin{cases} \partial_\mu + \frac{1}{4z} \gamma_\mu \gamma_5 & \text{if } M = \mu \\ \partial_5 & \text{if } M = 5. \end{cases} \quad (\text{A.34})$$

For all of the geometric heavy lifting we've done, we are led to an anticlimactic result: the spin connection drops out of the action,

$$S = \int d^5x \frac{i}{2} \left(\frac{R}{z}\right)^4 \left( \bar{\Psi} \gamma^M \overleftrightarrow{\partial}_M \Psi + \frac{1}{4z} \bar{\Psi} \gamma_\mu \gamma_5 \gamma^\mu \Psi - \frac{1}{4z} \overline{\gamma_\mu \gamma_5 \gamma^\mu \Psi} \Psi \right), \quad (\text{A.35})$$

The two spin connection terms cancel since  $\overline{\gamma_\mu \gamma_5 \gamma^\mu \Psi} \Psi = \bar{\Psi} \gamma_\mu \gamma_5 \gamma^\mu \Psi$ , so that upon including a bulk mass term,

$$S = \int d^5x \frac{i}{2} \left(\frac{R}{z}\right)^4 \bar{\Psi} \gamma^M \overleftrightarrow{\partial}_M \Psi - \int d^5x \frac{i}{2} \left(\frac{R}{z}\right)^5 m \bar{\Psi} \Psi = \int d^5x \frac{i}{2} \left(\frac{R}{z}\right)^4 \bar{\Psi} \left( \gamma^M \overleftrightarrow{\partial}_M - \frac{c}{z} \right) \Psi, \quad (\text{A.36})$$

where  $c = mR = m/k$  is a dimensionless parameter that is the ratio of the bulk mass to the curvature and

Before we can dimensionally reduce the action straightforwardly, we must write the Dirac operator to be right-acting, i.e. acting on  $\Psi$ , so that we can vary with respect to  $\bar{\Psi}$  to get an operator equation for  $\Psi$ . Obtaining this is from (A.36) is now a straightforward matter of integration by parts of the left-acting derivative term. Note that it is crucially important that we pick up a derivative acting on the metric/vielbein factor  $(R/z)^4$ . We would have missed this term if we had mistakenly written our original 'canonical action,' (A.19), as being right-acting only.

The integration by parts for the  $M = \mu = 0, \dots, 4$  terms proceeds trivially since these directions have no boundary and the metric/vielbein factor is independent of  $x^\mu$ . Performing the  $M = 5$  integration by parts we find

$$S = \int d^4x \int_{R'}^R dz \left(\frac{R}{z}\right)^4 \bar{\Psi} \left( i\cancel{\partial} + i\gamma^5 \partial_5 - i\frac{2}{z} \gamma^5 - \frac{c}{z} \right) \Psi + (\text{boundary term}) \Big|_{R'}^R. \quad (\text{A.37})$$

The term in the parenthesis can be identified with the Dirac operator for the Randall-Sundrum model with bulk fermions. The boundary term is

$$(\text{boundary}) = (R/z)^4 (\psi \chi - \bar{\chi} \bar{\psi}) \Big|_{R'}^R, \quad (\text{A.38})$$

where we've written out the Dirac spinor  $\Psi$  in terms of two-component Weyl spinors  $\chi$  and  $\psi$ . This term vanishes when we impose chiral boundary conditions, which we review in the next section. In terms of Weyl spinors this gives

$$S = \int d^4x \int_{R'}^R dz \left(\frac{R}{z}\right)^4 (\psi \quad \bar{\chi}) \begin{pmatrix} -\partial_5 + \frac{2-c}{z} & i\cancel{\partial} \\ i\cancel{\partial} & \partial_5 - \frac{2+c}{z} \end{pmatrix} \begin{pmatrix} \chi \\ \bar{\psi} \end{pmatrix}, \quad (\text{A.39})$$

where we use the two-component slash convention  $\cancel{\partial} = v_\mu \bar{\sigma}^\mu$ ,  $\psi = v_\mu \sigma^\mu$ . From here one may perform a straightforward dimensional reduction to obtain, among other things, the profile of a bulk fermion in RS,

$$\Psi_c^{(0)}(x, z) = \frac{1}{\sqrt{R'}} \left(\frac{z}{R}\right)^2 \left(\frac{z}{R'}\right)^{-c} \sqrt{\frac{1-2c}{1-(R/R')^{1-2c}}} P_L \Psi_c^{(0)}(x), \quad (\text{A.40})$$

where  $\Psi_c^{(0)}(x)$  is a canonically normalized 4D field and  $P_L$  is the usual left-chiral projector. The term in the square root is a flavor factor that is often written as  $f_c$ .

### A.3.6 Chiral boundary conditions

The vector-like (Dirac) nature of 5D spinors is an immediate problem for model-building since the Standard Model is manifestly chiral and there appears to be no way to write down a chiral fermion without immediately introducing a partner fermion of opposite chirality and the same couplings. To get around this problem, we can require that only the zero modes of the 5D fermions—those which are identified with Standard Model states—to be chiral. We show that one chirality of zero modes can indeed be projected out, while the heavier Kaluza-Klein excitations are vector-like but massive.

We can project out the zero modes of the wrong-chirality components of a bulk Dirac 5D fermion by imposing chiral boundary conditions that these states vanish on the branes. Since zero modes have trivial profiles, these boundary conditions force the mode to be identically zero everywhere. For left-chiral boundary conditions,  $\psi = 0$  on the branes, while for right-chiral boundary conditions  $\chi = 0$  on the branes. Thus we are guaranteed that both terms in (A.38) vanish at  $z = R, R'$  for either chirality.

Imposing these chiral boundary conditions is equivalent to the statement that the compactified extra dimension is an orbifold. This treatment of boundary conditions for interval compact spaces was first discussed from this viewpoint in [109].

## A.4 Gauge fields in RS

We now move on to the case of bulk gauge fields. We follow the approach of [214], though we adapt it to follow the same type of derivation espoused above for the fermion propagator. The bulk action is

$$S_5 = \int d^4x dz \sqrt{g} \left[ -\frac{1}{4} F_{MN} F^{MN} + (\text{brane}) + (\text{gauge fixing}) \right] \quad (\text{A.41})$$

To derive the propagator, we would like to write the kinetic term in the form  $A_M \mathcal{O}^{MN} A_N$  so that we may invert the quadratic differential operator  $\mathcal{O}^{MN}$ . This requires judicious integration by parts including the  $(R/z)$  factors from the metric and the measure,  $\sqrt{g}$ . The relevant integration is

$$\frac{R}{4z} F^{MN} F_{MN} = -\frac{R}{2} A^N \partial^M \left( \frac{1}{z} \partial_M \right) A_N + \frac{R}{2} A^N \partial^M \left( \frac{1}{z} \partial_N \right) A_N + \frac{R}{2} \partial^M \left( \frac{1}{z} A^N \partial_{[M} A_{N]} \right), \quad (\text{A.42})$$

where the last term integrates to a boundary term. Observe that this boundary term vanishes for both Dirichlet and Neumann boundary conditions so that it vanishes for  $\mu \rightarrow \nu$  and 5th component scalar propagators. It does not vanish, however, for the case of vector–scalar mixing. For simplicity, we will drop the term here in anticipation that it will be removed by gauge fixing. With this caveat, the above integration becomes

$$\frac{R}{4z} F^{MN} F_{MN} = A_\mu \left[ \frac{R}{2z} \partial^2 \eta^{\mu\nu} - \frac{R}{2} \partial_z \left( \frac{1}{z} \partial_z \right) \eta^{\mu\nu} - \frac{R}{2z} \partial^\mu \partial^\nu \right] A_\nu + A_5 \frac{R}{z} \partial_z \partial^\mu A_\mu - A_5 \frac{R}{2z} \partial^2 A_5. \quad (\text{A.43})$$

This is now in the desired form: we can read off the quadratic differential operators which encode the propagation of the 5D gauge bosons. Observe that we have a term that connects the 4D vector  $A_\mu$  to the 4D scalar  $A_5$ . In our mixed position-momentum space formalism, we prefer to leave these as separate fields. This term is removed by a judicious choice of gauge fixing.

We must now gauge fix to remove the gauge redundancy which otherwise appears as unphysical states in the propagator. Ideally we would like to pick a gauge where the scalar vanishes  $A_5 = 0$  and the vector has a convenient gauge, say, Lorenz gauge  $\partial_\mu A^\mu = 0$ . Unfortunately, these gauges are incompatible. Intuitively this is because we only have a single gauge fixing functional to work with in the path integral so that we are allowed to set at most one expression to vanish. Instead, motivated by the

potential for vector–scalar mixing from the boundary term of (A.43), we choose a gauge fixing functional which cancels this mixing term,

$$\mathcal{L}_{\text{gauge fix}} = - \left( \frac{R}{z} \right) \frac{1}{2\xi} \left[ \partial_\mu A^\mu - \xi z \partial_z \left( \frac{1}{z} A_5 \right) \right]^2 \quad (\text{A.44})$$

We have introduced a gauge fixing parameter  $\xi$  which will play the role of the ordinary  $R_\xi$  gauge fixing parameter in 4D. We can integrate by parts to convert this to the form  $A_M \mathcal{O}_{\text{gauge fix}}^{MN} A_N$ ,

$$\mathcal{L}_{\text{gauge fix}} = A_\mu \frac{1}{2\xi} \frac{R}{z} \partial^\mu \partial_\nu A_\nu - A_5 \frac{R}{z} \partial_z \partial^\mu A_\mu + A_5 \frac{\xi R}{2} \frac{R}{z} \partial_z \left[ z \partial_z \left( \frac{1}{z} A_5 \right) \right]. \quad (\text{A.45})$$

Observe that the second term here cancels the unwanted mixing term in (A.43). Summing this together with the gauge kinetic term gives a clean separation for the kinetic terms for the gauge vector and scalar:

$$\begin{aligned} \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{gauge fix}} &= A_\mu \left[ \frac{R}{2z} \partial^2 \eta^{\mu\nu} - \frac{R}{2} \partial_z \left( \frac{1}{z} \partial_z \right) \eta^{\mu\nu} - \left( 1 - \frac{1}{\xi} \right) \frac{R}{2z} \partial^\mu \partial^\nu \right] A_\nu \\ &\quad + A_5 \frac{R}{2z} \left[ -\partial^2 + \xi \left( \frac{1}{z^2} - \frac{1}{z} \partial_z + \partial_z^2 \right) \right] A_5 \end{aligned} \quad (\text{A.46})$$

$$\equiv A_\mu \mathcal{O}^{\mu\nu} A_\nu + A_5 \mathcal{O}_5 A_5. \quad (\text{A.47})$$

As above, now that we have the action written in terms of right-acting operators on the gauge fields, we may proceed to do a KK reduction to determine the KK mode properties,  $A_\mu^{(n)}(x, z) = A_\mu^{(n)}(x) h^{(n)}(z)$ .

The general solution for the  $n^{\text{th}}$  KK mode profile of a bulk gauge field is

$$h^{(n)} = a J_1(M^{(n)} z) + b Y_1(M^{(n)} z), \quad (\text{A.48})$$

where  $J_\alpha$  and  $Y_\alpha$  are Bessel functions. A SM gauge field must have a zero mode (which is identified with the SM state) so that it must have Neumann boundary conditions (BC). Using the formulae for derivatives of Bessel functions, we find

$$Y_0(M^{(n)} R) J_0(M^{(n)} R') = J_0(M^{(n)} R) Y_0^{(n)}(R'), \quad (\text{A.49})$$

where  $M^{(n)}$  is the mass of the  $n^{\text{th}}$  KK mode. We know that  $M^{(n)} \sim n/R'$  and that  $R \ll R'$ . Thus  $M^{(n)} R \approx 0$  for reasonable  $n$ . Now invoke two important properties of the  $J_0$  and  $Y_0$  Bessel functions:

1.  $J_0(0) = 1$  and  $J_0(x > 0)$  is under control, i.e.  $|J_0(x)| < 1$ .
2.  $Y_0(0) = -\infty$  and  $Y_0(x > y_1)$  is similarly under control, where  $y_1$  is the first zero of  $Y_0(x)$ .

From this we see that the left-hand side of (A.49) is very large and negative due to the  $Y_0(M^{(n)} R)$  term while the right-hand side is a product of ‘under control’ terms that are  $\mathcal{O}(1)$  or less. This implies that  $J_0(M^{(n)} R') \approx 0$ . In other words, the KK masses are given by the zeros of  $J_0$ . The first zero is  $x_1 = 2.405$  so that the first KK gauge boson excitation has mass  $M^{(n)} \approx 2.4/R'$ . The solution for the  $n^{\text{th}}$  KK mode profile of a SM gauge field is thus [215]

$$h^{(n)}(z) = \mathcal{N} z \left[ Y_0(M^{(n)} R) J_1(M^{(n)} z) - J_0(M^{(n)} R) Y_1(M^{(n)} z) \right]. \quad (\text{A.50})$$

The normalization is fixed by performing the  $dz$  integral and requiring canonical normalization of the zero mode 4D kinetic term,

$$\int d^4x dz \sqrt{g} F_{MN} F_{PQ} g^{MP} g^{NQ} = \int d^4x dz \frac{R}{z} F^{(0)}(x)_{\mu\nu} F^{(0)}(x)^{\mu\nu} \left[ h^{(0)}(z) \right]^2 + \dots \quad (\text{A.51})$$

This gives  $\mathcal{N}^{-1/2} = R \log R'/R$ .

Finally, we note that for the  $W$  and  $Z$  bosons, the Higgs VEV on the IR brane changes the boundary conditions so that the zero mode profile is not flat. Heuristically it introduces a kink on the profile near the IR brane. Since  $M_Z \ll M^{(1)}$ , we may treat this as a perturbation to  $M^{(0)} = 0$  so that the  $Z$  boson profile is

$$h_Z^{(0)}(z) = \frac{1}{\sqrt{R \log R'/R}} \left[ 1 - \frac{M_Z^2}{4} \left( z^2 - 2z^2 \log \frac{z}{R} \right) \right], \quad (\text{A.52})$$

and similarly for the  $W$ .

### A.5 Caution with finite loops

One should be careful when calculating loop diagrams in theories with extra dimensions. When one calculates a finite loop, say a dipole operator, naïve application of effective field theory suggests taking only the lowest KK mode and letting the 4D loop momentum go to  $k \rightarrow \infty$ . This, however, can lead to erroneous results since the loop integral runs over all momenta, including those in the fifth dimension. Only integrating over the 4D directions removes terms that scale like  $k^2/M_{\text{KK}}^2$  which would otherwise make an  $\mathcal{O}(1)$  finite contribution. This can appear as a dependence on the order in which one does the 4D loop integral versus KK sum; this discrepancy has appeared in the RS  $gg \rightarrow h$  production calculations [216]. One way to avoid this problem is to work in mixed position-momentum space [214]. This was used to calculate RS constraints from  $f \rightarrow f'\gamma$  [217, 218] and the muon magnetic moment in [219]. These references include Feynman rules for performing mixed space calculations. For a recent explanation of the subtleties of 5D dipoles and the resolution to puzzles in the previous literature, see [220]. In particular, Section 3 of that paper shows how to quickly estimate the size of 4D couplings from overlap integrals.

## B Appendix: Compositeness; the CCWZ Construction

The general theory of Goldstone bosons is described in the papers by Callan, Coleman, Wess, and Zumino (CCWZ) [110, 111]. In this appendix we present relevant aspects for generalizations to the composite Higgs models of interest, while making connections to the chiral Lagrangian above as an explicit example of their abstract procedure. See §19.5 – 19.7 of [112] for a more pedagogical and explicit discussion, the relevant sections of [113], or [?, 221] for more depth on how this procedure is applied to the chiral Lagrangian.

### B.1 Preliminaries

Suppose a Lagrangian is invariant under a global symmetry  $G$ , but that  $G$  is spontaneously broken to a subgroup  $H \subset G$  by the VEV of a field  $\langle \psi \rangle = \psi_0$  that is in a linear representation of  $G$ . This means that for any  $h \in H$ ,  $h\psi_0 = \psi_0$ . The spontaneous symmetry breaking pattern  $G \rightarrow H$  implies the existence of  $\dim G - \dim H$  Goldstone bosons that take values on a vacuum manifold. This manifold can be identified with the **coset space**  $G/H$  ( $G \bmod H$ ). In particular, the *left* coset space  $G/H$  is an equivalence class of elements  $g \in G$  modulo elements  $h \in H$ ,  $g \sim gh$ . In other words, any element in  $g$  is equivalent to another element  $g'$  if there exists an  $h$  such that  $g' = gh$ . Note that in general  $G/H$  is not a group.

### B.2 Decomposition of the Algebra

The generators of  $G$  can be divided between two classes:  $T^i$  which generate the unbroken group  $H$ , and  $X^a$  which do not. This is called the **Cartan decomposition**. The generators satisfy the following commutation relations:

$$[T^i, T^j] = i f^{ijk} T^k \quad (\text{B.1})$$

$$[T^i, X^\alpha] = if^{i\alpha\beta} X^\beta \quad (\text{B.2})$$

$$[X^\alpha, X^\beta] = if^{\alpha\beta k} T^k + if^{\alpha\beta\gamma} X^\gamma, \quad (\text{B.3})$$

where the (B.2) comes from explicitly checking that  $\langle T^i | [T^j, X^\alpha] \rangle = 0$ , where the inner product is  $\langle A | B \rangle = \text{Tr}(AB)$ . Observe that this means that the  $X$ s furnish a linear representation of  $H$ . If, additionally, there exists a parity transformation  $P$  such that  $P^2 = 1$  and  $P([g_1, g_2]) = [P(g_1), P(g_2)]$  and further such that  $P(x) = -X$  and  $P(T) = +T$ , then one can further restrict

$$[X^\alpha, X^\beta] = if^{\alpha\beta k} T^k. \quad (\text{B.4})$$

In this case, the coset  $G/H$  is a **symmetric space**.

### B.3 Decomposition of the Group

Without loss of generality, we may write any  $g \in G$  as

$$g(\xi, u) = e^{i\xi^\alpha X^\alpha} e^{iu^i T^i} \equiv \hat{g}(\xi)h(u). \quad (\text{B.5})$$

Further, note that the distinct  $G$  elements  $gh_1, gh_2, gh_3, \dots \in G$  are all identified with the *same* element of  $G/H$ . For each element of  $G/H$ , it is useful to pick a representative element of  $G$ , which we can choose to be  $\hat{g}(\xi)$ . This is simply the decomposition (4.16) when applied to chiral perturbation theory.

### B.4 Decomposition of the Linear Representation

We may further write the linearly represented field  $\psi(x)$  with respect to any non-trivial reference value such as  $\psi_0$  by defining  $\gamma(x) \in G$  to be the transformation from  $\psi_0 \rightarrow \psi(x)$ ,

$$\psi_i(x) = \gamma_{ij}(x) (\psi_0)_j. \quad (\text{B.6})$$

By the invariance of  $\psi_0$  under  $H$  transformations,  $\gamma(x)$  is only defined up to right multiplication by any  $h \in H$ . In other words, we may identify  $\gamma(x)$  with the representative element  $\hat{\gamma}(x)$  which is chosen to be the exponentiation of only broken generators, analogously to  $\hat{g}$  above. (For the moment we are ignoring radial excitations.) We may now drop the hat on  $\hat{\gamma}$  for notational clarity. Let us suggestively call the transformation parameter  $\pi(x)$ ,

$$\gamma(\pi(x)) = e^{i\pi^a(x)X^a}. \quad (\text{B.7})$$

The  $\pi^a(x)$  are to be identified with the Goldstone bosons (pions). We leave it dimensionless, remember that the pion field with canonical mass dimension can be restored with  $\pi^a(x) \rightarrow \pi^a(x)_{\text{can}}/f$ .

Suppose the Lagrangian of the theory with respect to the linearly represented field  $\psi(x)$  is written in terms of  $\psi(x)$  and  $\partial\psi(x)$ . The former don't contain the Goldstone fields, while the latter can be written in terms of  $\psi_0$  and the Goldstone fields using (B.6) and (B.7),

$$\partial_\mu\psi(x) = \gamma [\partial_\mu + \gamma^{-1}(\partial_\mu\gamma)] \psi_0, \quad (\text{B.8})$$

where we've suppressed the  $x$  dependence of  $\gamma$ . Without loss of generality, we can write  $\gamma^{-1}\partial_\mu\gamma$  in terms of the broken and unbroken generators,

$$\gamma^{-1}\partial_\mu\gamma = iD_\mu^\alpha X^\alpha + iE_\mu^i T^i \quad (\text{B.9})$$

$$D_\mu^\alpha = D^{\alpha\beta}(\pi)\partial_\mu\pi^\beta \quad (\text{B.10})$$

$$E_\mu^i = E^{i\beta}(\pi)\partial_\mu\pi^\beta. \quad (\text{B.11})$$

The proof of this expression uses the trick

$$\partial_\mu e^{i\pi \cdot X} = i\partial_\mu\pi^\alpha \int_0^1 ds e^{i(1-s)\pi \cdot X} X^\alpha e^{is\pi \cdot X} \quad (\text{B.12})$$

and then the Baker–Campbell–Hausdorff (BCH) relation to show that you end up with an expansion in the  $X$ s and  $T$ s.



### B.5 Transformation of the Goldstones

We would like to see how the  $\pi^a$  transform under the global group  $G$ . We can derive this from the transformation of the linear field  $\psi(X)$  using (B.6) and (B.7). Let us *define*  $\pi'$  and  $u'$  such that

$$g\psi(x) = ge^{i\pi \cdot X}\psi_0 \equiv e^{i\pi' \cdot X}e^{iu' \cdot T}\psi_0, \quad (\text{B.13})$$

where the primed fields are *nonlinear* functions of  $g$  and  $\pi(x)$ , that is  $\pi' = \pi'(g, \pi)$ . It is cleaner to write this after peeling off the  $\psi_0$ ,

$$g\gamma(\pi) = \gamma(\pi')h(u'). \quad (\text{B.14})$$

The element  $h(u') \in H$  leaves  $\psi_0$  invariant. The non-linear dependence of  $\pi'(x)$  on  $\pi(x)$  for generic  $g$  is the key result.

In the case when we transform  $\psi(x)$  by an element  $h \in H$ ,

$$h\psi(x) = e^{iu^i T^i}e^{i\pi(x) \cdot X}\psi_0 \equiv e^{i\pi' \cdot X}\psi_0, \quad (\text{B.15})$$

where we have used (B.2) and the BCH formula. In other words,

$$h\gamma(\pi) = \gamma(\pi') \quad \pi^a X^a \rightarrow \pi'^a X^a. \quad (\text{B.16})$$

In this case, we see that  $\pi(x)$  has transformed *linearly*, in contrast to (B.14) or (B.13). The main difference is that due to the general decomposition (B.5), we simply observed that  $g\gamma(\pi(x)) \in G$  and so that there must exist *some*  $\pi'$  and  $u'$  satisfies (B.13). The actual expression for  $\pi'$  as a function of  $\pi$  is messy. On the other hand, in (B.15) we used BCH to derive the  $\pi'$  explicitly and one can see that this is a linear transformation.

### B.6 From Linear to Non-Linear

From our assumed linear UV theory, we have now identified the Goldstone fields and can integrate out the massive ‘radial’ modes to obtain a low-energy Lagrangian. The radial modes can be identified with excitations along the VEV direction  $\psi_0$ . So let us define the radial field  $\psi_r(x) = r(x)\psi_0$  with a VEV  $\langle r(x) \rangle = 1$ . From (B.13) with  $\psi_0 \rightarrow \psi_r(x)$ , we see that the radial field transforms as

$$\psi_r \rightarrow h(u'(\pi))\psi_r. \quad (\text{B.17})$$

Thus in order to build  $G$  invariants out of the radial fields  $\psi_r$ , it is sufficient to construct  $H$  invariants. Said differently, the decomposition  $\psi(x) = \gamma(\pi)\psi_r$  is a tool for converting  $G$ -linear representations  $\psi$  into  $H$ -linear representations  $\psi_r$ .

Effective field theory tells us that this UV theory wasn’t necessary to construct the low energy theory of Goldstone bosons. So we can now integrate out the radial modes and remain agnostic about the UV completion of the theory. Prior to the discovery of the Higgs boson—a linear UV completion of the theory of the Goldstone bosons eaten by  $W^\pm$  and  $Z$ —the reason why experiments like LEP could make precision measurements of the SM without knowing the details of the Higgs is that the precision measurements asked precise questions about the non-linear sigma model (NL $\Sigma$ M) of Goldstones that were insensitive to the particular UV completion, linear or otherwise.

### B.7 A Low-Energy Lagrangian without the UV

Let us now return to (B.8) since we know from the Goldstone shift symmetry that the Goldstones only appear in derivative interactions. The object  $g^{-1}\partial g$ , where  $g = \gamma$  in (B.8), is called the **Mauer-Cartan form**, it takes an element of the group  $g \in G$ , differentiates it—pulling out the Lie algebra element based

at  $g$ , and then pulls that generator back to the group identity so that one can compare elements of the algebra on the same tangent space.

The expansion of the Maurer-Cartan form into broken and unbroken generators is given in (B.9). Differentiate the transformation rule for  $\gamma$ , (B.14),

$$g\partial_\mu\gamma(\pi) = [\partial_\mu\gamma(\pi')] h + \gamma(\pi')\partial_\mu h. \quad (\text{B.18})$$

Now multiply each side of this equation by the respective side in the inverse of (B.14),

$$\gamma^{-1}(\pi)\partial_\mu\gamma(\pi) = h^{-1}\gamma^{-1}(\pi')([\partial_\mu\gamma(\pi')] h + \gamma(\pi')\partial_\mu h). \quad (\text{B.19})$$

Comparing this to (B.9), we find

$$iD_\mu^\alpha(\pi)X^\alpha + iE_\mu^i(\pi)T^i = ih^{-1}D_\mu^\alpha(\pi')X^\alpha h^{-1} + ih[E_\mu^i(\pi')T^i + i\partial_\mu] h^{-1}. \quad (\text{B.20})$$

In other words, the objects  $D$  and  $E$  defined in (B.9) transform under  $g \in G$  as

$$D_\mu^\alpha \rightarrow hD_\mu^\alpha h^{-1} \quad (\text{B.21})$$

$$E_\mu^i \rightarrow hE_\mu^i h^{-1} - ih\partial_\mu h^{-1}, \quad (\text{B.22})$$

where here  $h = h(u'(\pi, g))$  in (B.14). This should look very familiar:  $D$  transforms linearly and  $E$  transforms like a gauge field. Both transform under  $G$  with respect to the subgroup  $H$  rather than the whole group  $G$ . This realizes the observation in Section B.6: to write Lagrangians for nonlinear realizations of  $G/H$ , we need to construct  $H$  invariants. The linear object  $D$  can indeed be used to construct a simple lowest-order Lagrangian,

$$\mathcal{L} = \frac{f^2}{4} \text{Tr}(D_\mu D^\mu), \quad (\text{B.23})$$

where we've introduced the compositeness scale  $f$  to preserve dimensionality.

What about the curious object  $E_\mu$ ? This appears to transform as the gauge field of a local symmetry. The locality of this symmetry is inherited from the  $x$ -dependence of the Goldstone fields  $\pi(x)$  and is unsurprising since the coset identification  $g \sim gh$  is local.  $E_\mu$  is thus a 'gauge potential' with respect to the unbroken symmetry  $H$ . Indeed, differentiating (B.17)—recalling that  $h(u')$  depends on  $x$  through its implicit dependence on  $\pi(x)$ —shows that derivatives of the non-Goldstone fields transform inhomogeneously under  $G$ . Promoting the partial derivative to a covariant derivative,  $\mathcal{D}_\mu = \partial_\mu \rightarrow \partial_\mu + iE_\mu$ , ensures that  $\mathcal{D}_\mu\psi_r(x)$  transforms homogeneously under  $H$ .

**When did  $H$  become gauged?** The appearance of a covariant derivative and a gauge symmetry may seem surprising in a system where *global* symmetry  $G$  is spontaneously broken to a subgroup  $H$ . The appearance of a local symmetry, however, is not surprising since the resulting coset space  $G/H$  precisely describes a gauge redundancy. Mathematically, the description of a 'gauged' symmetry is identical to that of a spontaneously broken global symmetry. For the mathematically inclined, details of the geometric structure of these theories are presented in [222] and [223].

The punchline is that one can construct a Goldstone boson Lagrangian which is invariant under the full, nonlinearly realized group  $G$ , by constructing an  $H$ -invariant Lagrangian out of  $D_\mu$ . One can further introduce non-Goldstone fields  $\psi_r$  (not necessarily related to the linear field that gets a VEV) so long as one uses the appropriate  $H$  covariant derivative. In this way one may include, for example, 'nucleon' excitations to the effective theory.

The description above is based on a 'standard realization' of the nonlinearly realized symmetry, (B.13). One of the main results of the CCWZ papers was the observation that every non-linear realization can be brought to this standard realization [110, 111]. Physically, this means that no matter how one imposes the  $G/H$  restriction, the  $S$ -matrix elements for the low-energy dynamics will be identical. Explicit examples of this are presented in chapter IV of [?].

## References

- [1] A. de Gouvea, D. Hernandez, and T. M. P. Tait, “Criteria for Natural Hierarchies,” arXiv:1402.2658 [hep-ph].
- [2] M. J. Strassler, “An Unorthodox Introduction to Supersymmetric Gauge Theory,” arXiv:hep-th/0309149 [hep-th].
- [3] M. A. Shifman, “The Many Faces of the Superworld: Yuri Golfand Memorial Volume,”.
- [4] G. L. Kane and M. Shifman, “The Supersymmetric World: the Beginning of the Theory,”.
- [5] M. E. Peskin, “Duality in Supersymmetric Yang-Mills Theory,” arXiv:hep-th/9702094 [hep-th].
- [6] N. Seiberg, “Electric–magnetic Duality in Supersymmetric Nonabelian Gauge Theories,” *Nucl.Phys.* **B435** (1995) 129–146, arXiv:hep-th/9411149 [hep-th].
- [7] K. A. Intriligator and N. Seiberg, “Lectures on Supersymmetry Breaking,” *Class.Quant.Grav.* **24** (2007) S741–S772, arXiv:hep-ph/0702069 [hep-ph].
- [8] S. R. Coleman and J. Mandula, “All Possible Symmetries of the S Matrix,” *Phys.Rev.* **159** (1967) 1251–1256.
- [9] R. Haag, J. T. Lopuszanski, and M. Sohnius, “All Possible Generators of Supersymmetries of the S Matrix,” *Nucl.Phys.* **B88** (1975) 257.
- [10] H. K. Dreiner, H. E. Haber, and S. P. Martin, “Two-Component Spinor Techniques and Feynman Rules for Quantum Field Theory and Supersymmetry,” *Phys.Rept.* **494** (2010) 1–196, arXiv:0812.1594 [hep-ph].
- [11] F. Gieres, “Geometry of Supersymmetric Gauge Theories: Including an Introduction to Brs Differential Algebras and Anomalies,” *Lect.Notes Phys.* **302** (1988) 1–189.
- [12] K. A. Intriligator and N. Seiberg, “Lectures on Supersymmetric Gauge Theories and Electric–magnetic Duality,” *Nucl.Phys.Proc.Suppl.* **45BC** (1996) 1–28, arXiv:hep-th/9509066 [hep-th].
- [13] N. Arkani-Hamed, M. A. Luty, and J. Terning, “Composite Quarks and Leptons from Dynamical Supersymmetry Breaking without Messengers,” *Phys.Rev.* **D58** (1998) 015004, arXiv:hep-ph/9712389 [hep-ph].
- [14] M. A. Luty and J. Terning, “Improved Single Sector Supersymmetry Breaking,” *Phys.Rev.* **D62** (2000) 075006, arXiv:hep-ph/9812290 [hep-ph].
- [15] Z. Komargodski and D. Shih, “Notes on SUSY and R-Symmetry Breaking in Wess-Zumino Models,” *JHEP* **0904** (2009) 093, arXiv:0902.0030 [hep-th].
- [16] A. E. Nelson and N. Seiberg, “R Symmetry Breaking Versus Supersymmetry Breaking,” *Nucl.Phys.* **B416** (1994) 46–62, arXiv:hep-ph/9309299 [hep-ph].
- [17] M. Drees, R. Godbole, and P. Roy, “Theory and phenomenology of sparticles: An account of four-dimensional N=1 supersymmetry in high energy physics,”.
- [18] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model,” *Phys.Rept.* **459** (2008) 1–241, arXiv:hep-ph/0503173 [hep-ph].
- [19] H. E. Haber, “The Status of the Minimal Supersymmetric Standard Model and Beyond,” *Nucl.Phys.Proc.Suppl.* **62** (1998) 469–484, arXiv:hep-ph/9709450 [hep-ph].
- [20] M. Dine and W. Fischler, “A Supersymmetric GUT,” *Nucl.Phys.* **B204** (1982) 346.
- [21] M. Dine, A. E. Nelson, Y. Nir, and Y. Shirman, “New Tools for Low-Energy Dynamical Supersymmetry Breaking,” *Phys.Rev.* **D53** (1996) 2658–2669, arXiv:hep-ph/9507378 [hep-ph].
- [22] M. Dine, A. E. Nelson, and Y. Shirman, “Low-Energy Dynamical Supersymmetry Breaking Simplified,” *Phys.Rev.* **D51** (1995) 1362–1370, arXiv:hep-ph/9408384 [hep-ph].

- [23] M. Dine, Y. Nir, and Y. Shirman, “Variations on Minimal Gauge Mediated Supersymmetry Breaking,” *Phys.Rev.* **D55** (1997) 1501–1508, arXiv:hep-ph/9607397 [hep-ph].
- [24] G. Giudice and R. Rattazzi, “Theories with Gauge Mediated Supersymmetry Breaking,” *Phys.Rept.* **322** (1999) 419–499, arXiv:hep-ph/9801271 [hep-ph].
- [25] G. Giudice and R. Rattazzi, “Extracting supersymmetry breaking effects from wave function renormalization,” *Nucl.Phys.* **B511** (1998) 25–44, arXiv:hep-ph/9706540 [hep-ph].
- [26] N. Arkani-Hamed, G. F. Giudice, M. A. Luty, and R. Rattazzi, “Supersymmetry breaking loops from analytic continuation into superspace,” *Phys.Rev.* **D58** (1998) 115005, arXiv:hep-ph/9803290 [hep-ph].
- [27] P. Meade, N. Seiberg, and D. Shih, “General Gauge Mediation,” *Prog.Theor.Phys.Suppl.* **177** (2009) 143–158, arXiv:0801.3278 [hep-ph].
- [28] L. M. Carpenter, M. Dine, G. Festuccia, and J. D. Mason, “Implementing General Gauge Mediation,” *Phys.Rev.* **D79** (2009) 035002, arXiv:0805.2944 [hep-ph].
- [29] G. Giudice and A. Masiero, “A Natural Solution to the Mu Problem in Supergravity Theories,” *Phys.Lett.* **B206** (1988) 480–484.
- [30] N. Craig, “The State of Supersymmetry After Run I of the Lhc,” arXiv:1309.0528 [hep-ph].
- [31] P. Batra, A. Delgado, D. E. Kaplan, and T. M. Tait, “The Higgs Mass Bound in Gauge Extensions of the Minimal Supersymmetric Standard Model,” *JHEP* **0402** (2004) 043, arXiv:hep-ph/0309149 [hep-ph].
- [32] A. Maloney, A. Pierce, and J. G. Wacker, “D-Terms, Unification, and the Higgs Mass,” *JHEP* **0606** (2006) 034, arXiv:hep-ph/0409127 [hep-ph].
- [33] A. D. Medina, N. R. Shah, and C. E. Wagner, “A Heavy Higgs and a Light Sneutrino NLSP in the MSSM with Enhanced SU(2) D-terms,” *Phys.Rev.* **D80** (2009) 015001, arXiv:0904.1625 [hep-ph].
- [34] B. Bellazzini, C. Csaki, A. Delgado, and A. Weiler, “SUSY without the Little Hierarchy,” *Phys.Rev.* **D79** (2009) 095003, arXiv:0902.0015 [hep-ph].
- [35] C. Cheung and H. L. Roberts, “Higgs Mass from D-Terms: a Litmus Test,” *JHEP* **1312** (2013) 018, arXiv:1207.0234 [hep-ph].
- [36] K. Blum, R. T. D’Agnolo, and J. Fan, “Natural SUSY Predicts: Higgs Couplings,” *JHEP* **1301** (2013) 057, arXiv:1206.5303 [hep-ph].
- [37] N. Craig and A. Katz, “A Supersymmetric Higgs Sector with Chiral D-terms,” *JHEP* **1305** (2013) 015, arXiv:1212.2635 [hep-ph].
- [38] P. Fayet, “Supergauge Invariant Extension of the Higgs Mechanism and a Model for the Electron and Its Neutrino,” *Nucl.Phys.* **B90** (1975) 104–124.
- [39] M. Dine, W. Fischler, and M. Srednicki, “A Simple Solution to the Strong CP Problem with a Harmless Axion,” *Phys.Lett.* **B104** (1981) 199.
- [40] H. P. Nilles, M. Srednicki, and D. Wyler, “Weak Interaction Breakdown Induced by Supergravity,” *Phys.Lett.* **B120** (1983) 346.
- [41] J. R. Ellis, J. Gunion, H. E. Haber, L. Roszkowski, and F. Zwirner, “Higgs Bosons in a Nonminimal Supersymmetric Model,” *Phys.Rev.* **D39** (1989) 844.
- [42] U. Ellwanger, M. Rausch de Traubenberg, and C. A. Savoy, “Particle Spectrum in Supersymmetric Models with a Gauge Singlet,” *Phys.Lett.* **B315** (1993) 331–337, arXiv:hep-ph/9307322 [hep-ph].
- [43] U. Ellwanger, C. Hugonie, and A. M. Teixeira, “The Next-To-Minimal Supersymmetric Standard Model,” *Phys.Rept.* **496** (2010) 1–77, arXiv:0910.1785 [hep-ph].
- [44] M. Maniatis, “The Next-To-Minimal Supersymmetric Extension of the Standard Model Reviewed,” *Int.J.Mod.Phys.* **A25** (2010) 3505–3602, arXiv:0906.0777 [hep-ph].

- [45] S. Dimopoulos and G. Giudice, “Naturalness Constraints in Supersymmetric Theories with Nonuniversal Soft Terms,” *Phys.Lett.* **B357** (1995) 573–578, arXiv:hep-ph/9507282 [hep-ph].
- [46] A. G. Cohen, D. Kaplan, and A. Nelson, “The More Minimal Supersymmetric Standard Model,” *Phys.Lett.* **B388** (1996) 588–598, arXiv:hep-ph/9607394 [hep-ph].
- [47] Y. Kats, P. Meade, M. Reece, and D. Shih, “The Status of Gmsb After  $1/\text{Fb}$  at the Lhc,” *JHEP* **1202** (2012) 115, arXiv:1110.6444 [hep-ph].
- [48] C. Brust, A. Katz, S. Lawrence, and R. Sundrum, “Susy, the Third Generation and the Lhc,” *JHEP* **1203** (2012) 103, arXiv:1110.6670 [hep-ph].
- [49] M. Papucci, J. T. Ruderman, and A. Weiler, “Natural SUSY Endures,” *JHEP* **1209** (2012) 035, arXiv:1110.6926 [hep-ph].
- [50] L. J. Hall, D. Pinner, and J. T. Ruderman, “A Natural SUSY Higgs Near 126 GeV,” *JHEP* **1204** (2012) 131, arXiv:1112.2703 [hep-ph].
- [51] L. J. Hall and M. Suzuki, “Explicit R-Parity Breaking in Supersymmetric Models,” *Nucl.Phys.* **B231** (1984) 419.
- [52] G. G. Ross and J. Valle, “Supersymmetric Models without R-Parity,” *Phys.Lett.* **B151** (1985) 375.
- [53] V. D. Barger, G. Giudice, and T. Han, “Some New Aspects of Supersymmetry R-Parity Violating Interactions,” *Phys.Rev.* **D40** (1989) 2987.
- [54] G. Bhattacharyya, “A Brief Review of R-Parity Violating Couplings,” arXiv:hep-ph/9709395 [hep-ph].
- [55] H. K. Dreiner, “An Introduction to Explicit R-Parity Violation,” arXiv:hep-ph/9707435 [hep-ph].
- [56] R. Barbier, C. Berat, M. Besancon, M. Chemtob, A. Deandrea, *et al.*, “R-Parity Violating Supersymmetry,” *Phys.Rept.* **420** (2005) 1–202, arXiv:hep-ph/0406039 [hep-ph].
- [57] C. Csaki, Y. Grossman, and B. Heidenreich, “Mfv Susy: a Natural Theory for R-Parity Violation,” *Phys.Rev.* **D85** (2012) 095009, arXiv:1111.1239 [hep-ph].
- [58] T. Kaluza, “On the Problem of Unity in Physics,” *Sitzungsber.Preuss.Akad.Wiss.Berlin (Math.Phys.)* **1921** (1921) 966–972.
- [59] O. Klein, “Quantum Theory and Five-Dimensional Theory of Relativity. (In German and English),” *Z.Phys.* **37** (1926) 895–906.
- [60] A. Einstein and P. Bergmann, “On a Generalization of Kaluza’s Theory of Electricity,” *Annals Math.* **39** (1938) 683–701.
- [61] J. Scherk and J. H. Schwarz, “Dual Field Theory of Quarks and Gluons,” *Phys.Lett.* **B57** (1975) 463–466.
- [62] E. Cremmer and J. Scherk, “Dual Models in Four-Dimensions with Internal Symmetries,” *Nucl.Phys.* **B103** (1976) 399.
- [63] E. Cremmer and J. Scherk, “Spontaneous Compactification of Space in an Einstein Yang-Mills Higgs Model,” *Nucl.Phys.* **B108** (1976) 409.
- [64] C. Csaki, “Tasi Lectures on Extra Dimensions and Branes,” arXiv:hep-ph/0404096.
- [65] C. Csaki, J. Hubisz, and P. Meade, “Tasi Lectures on Electroweak Symmetry Breaking from Extra Dimensions,” arXiv:hep-ph/0510275 [hep-ph].
- [66] R. Rattazzi, “Cargese Lectures on Extra-Dimensions,” arXiv:hep-ph/0607055 [hep-ph].
- [67] E. Ponton, “Tasi 2011: Four Lectures on TeV Scale Extra Dimensions,” arXiv:1207.3827 [hep-ph].
- [68] H.-C. Cheng, “2009 Tasi Lecture – Introduction to Extra Dimensions,” arXiv:1003.1162 [hep-ph].

- [69] T. Gherghetta, “Tasi Lectures on a Holographic View of Beyond the Standard Model Physics,” arXiv:1008.2570 [hep-ph].
- [70] R. Sundrum, “From Fixed Points to the Fifth Dimension,” arXiv:1106.4501 [hep-th].
- [71] R. Sundrum, “Tasi 2004 Lectures: to the Fifth Dimension and Back,” arXiv:hep-th/0508134 [hep-th].
- [72] V. Rubakov and M. Shaposhnikov, “Do We Live Inside a Domain Wall?,” *Phys.Lett.* **B125** (1983) 136–138.
- [73] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, “Phenomenology, Astrophysics and Cosmology of Theories with Submillimeter Dimensions and TeV Scale Quantum Gravity,” *Phys.Rev.* **D59** (1999) 086004, arXiv:hep-ph/9807344 [hep-ph].
- [74] L. Randall and R. Sundrum, “A Large Mass Hierarchy from a Small Extra Dimension,” *Phys.Rev.Lett.* **83** (1999) 3370–3373, arXiv:hep-ph/9905221 [hep-ph].
- [75] F. Brummer, A. Hebecker, and E. Trincherini, “The Throat as a Randall-Sundrum Model with Goldberger-Wise Stabilization,” *Nucl.Phys.* **B738** (2006) 283–305, arXiv:hep-th/0510113 [hep-th].
- [76] W. D. Goldberger and M. B. Wise, “Modulus Stabilization with Bulk Fields,” *Phys.Rev.Lett.* **83** (1999) 4922–4925, arXiv:hep-ph/9907447 [hep-ph].
- [77] W. D. Goldberger and M. B. Wise, “Phenomenology of a Stabilized Modulus,” *Phys.Lett.* **B475** (2000) 275–279, arXiv:hep-ph/9911457 [hep-ph].
- [78] O. DeWolfe, D. Freedman, S. Gubser, and A. Karch, “Modeling the Fifth-Dimension with Scalars and Gravity,” *Phys.Rev.* **D62** (2000) 046008, arXiv:hep-th/9909134 [hep-th].
- [79] N. Arkani-Hamed, M. Porrati, and L. Randall, “Holography and Phenomenology,” *JHEP* **08** (2001) 017, arXiv:hep-th/0012148.
- [80] T. Gherghetta, “Warped Models and Holography,” arXiv:hep-ph/0601213.
- [81] R. Rattazzi and A. Zaffaroni, “Comments on the Holographic Picture of the Randall-Sundrum Model,” *JHEP* **04** (2001) 021, arXiv:hep-th/0012248.
- [82] K. Skenderis, “Lecture Notes on Holographic Renormalization,” *Class.Quant.Grav.* **19** (2002) 5849–5876, arXiv:hep-th/0209067 [hep-th].
- [83] J. McGreevy, “Holographic Duality with a View Toward Many-Body Physics,” *Adv.High Energy Phys.* **2010** (2010) 723105, arXiv:0909.0518 [hep-th].
- [84] H. Nastase, “Introduction to AdS-CFT,” arXiv:0712.0689 [hep-th].
- [85] M. J. Strassler, “The Duality Cascade,” arXiv:hep-th/0505153 [hep-th].
- [86] J. M. Maldacena, “The Large  $N$  Limit of Superconformal Field Theories and Supergravity,” *Adv.Theor.Math.Phys.* **2** (1998) 231–252, arXiv:hep-th/9711200.
- [87] S. Gubser, I. R. Klebanov, and A. M. Polyakov, “Gauge Theory Correlators from Noncritical String Theory,” *Phys.Lett.* **B428** (1998) 105–114, arXiv:hep-th/9802109 [hep-th].
- [88] E. Witten, “Anti-de Sitter Space and Holography,” *Adv.Theor.Math.Phys.* **2** (1998) 253–291, arXiv:hep-th/9802150 [hep-th].
- [89] C. Csaki, M. Graesser, L. Randall, and J. Terning, “Cosmology of Brane Models with Radion Stabilization,” *Phys.Rev.* **D62** (2000) 045015, arXiv:hep-ph/9911406 [hep-ph].
- [90] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, “The Higgs Hunter’s Guide,” *Front.Phys.* **80** (2000) 1–448.
- [91] B. Bellazzini, C. Csaki, J. Hubisz, J. Serra, and J. Terning, “A Higgslike Dilaton,” *Eur.Phys.J.* **C73** (2013) no. 2, 2333, arXiv:1209.3299 [hep-ph].
- [92] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, “Bulk Gauge Fields in the Randall-Sundrum Model,” *Phys. Lett.* **B473** (2000) 43–49, arXiv:hep-ph/9911262.

- [93] Y. Grossman and M. Neubert, “Neutrino Masses and Mixings in Non-Factorizable Geometry,” *Phys. Lett.* **B474** (2000) 361–371, arXiv:hep-ph/9912408.
- [94] T. Gherghetta and A. Pomarol, “Bulk Fields and Supersymmetry in a Slice of AdS,” *Nucl. Phys.* **B586** (2000) 141–162, arXiv:hep-ph/0003129.
- [95] S. J. Huber and Q. Shafi, “Fermion Masses, Mixings and Proton Decay in a Randall-Sundrum Model,” *Phys.Lett.* **B498** (2001) 256–262, arXiv:hep-ph/0010195 [hep-ph].
- [96] S. J. Huber, “Flavor Violation and Warped Geometry,” *Nucl.Phys.* **B666** (2003) 269–288, arXiv:hep-ph/0303183 [hep-ph].
- [97] G. Burdman, “Constraints on the Bulk Standard Model in the Randall-Sundrum Scenario,” *Phys.Rev.* **D66** (2002) 076003, arXiv:hep-ph/0205329 [hep-ph].
- [98] K. Agashe, G. Perez, and A. Soni, “Flavor Structure of Warped Extra Dimension Models,” *Phys.Rev.* **D71** (2005) 016002, arXiv:hep-ph/0408134 [hep-ph].
- [99] K. Agashe, G. Perez, and A. Soni, “B-Factory Signals for a Warped Extra Dimension,” *Phys.Rev.Lett.* **93** (2004) 201804, arXiv:hep-ph/0406101 [hep-ph].
- [100] N. Arkani-Hamed and M. Schmaltz, “Hierarchies without Symmetries from Extra Dimensions,” *Phys. Rev.* **D61** (2000) 033005, arXiv:hep-ph/9903417.
- [101] D. E. Kaplan and T. M. Tait, “Supersymmetry Breaking, Fermion Masses and a Small Extra Dimension,” *JHEP* **0006** (2000) 020, arXiv:hep-ph/0004200 [hep-ph].
- [102] D. E. Kaplan and T. M. Tait, “New Tools for Fermion Masses from Extra Dimensions,” *JHEP* **0111** (2001) 051, arXiv:hep-ph/0110126 [hep-ph].
- [103] Y. Grossman and G. Perez, “Realistic Construction of Split Fermion Models,” *Phys.Rev.* **D67** (2003) 015011, arXiv:hep-ph/0210053 [hep-ph].
- [104] A. Altheimer, S. Arora, L. Asquith, G. Brooijmans, J. Butterworth, *et al.*, “Jet Substructure at the Tevatron and Lhc: New Results, New Tools, New Benchmarks,” *J.Phys.* **G39** (2012) 063001, arXiv:1201.0008 [hep-ph].
- [105] J. Shelton, “Tasi Lectures on Jet Substructure,” arXiv:1302.0260 [hep-ph].
- [106] G. P. Salam, “Towards Jetography,” *Eur.Phys.J.* **C67** (2010) 637–686, arXiv:0906.1833 [hep-ph].
- [107] K. Agashe, A. Delgado, M. J. May, and R. Sundrum, “Rs1, Custodial Isospin and Precision Tests,” *JHEP* **08** (2003) 050, arXiv:hep-ph/0308036.
- [108] K. Agashe, R. Contino, L. Da Rold, and A. Pomarol, “A Custodial Symmetry for Z B Anti-B,” *Phys. Lett.* **B641** (2006) 62–66, arXiv:hep-ph/0605341.
- [109] C. Csaki, C. Grojean, J. Hubisz, Y. Shirman, and J. Terning, “Fermions on an Interval: Quark and Lepton Masses without a Higgs,” *Phys. Rev.* **D70** (2004) 015012, arXiv:hep-ph/0310355.
- [110] S. R. Coleman, J. Wess, and B. Zumino, “Structure of Phenomenological Lagrangians. 1.,” *Phys.Rev.* **177** (1969) 2239–2247.
- [111] J. Callan, Curtis G., S. R. Coleman, J. Wess, and B. Zumino, “Structure of Phenomenological Lagrangians. 2.,” *Phys.Rev.* **177** (1969) 2247–2250.
- [112] S. Weinberg, “The Quantum Theory of Fields. Vol. 2: Modern Applications,”.
- [113] M. Bando, T. Kugo, and K. Yamawaki, “Nonlinear Realization and Hidden Local Symmetries,” *Phys.Rept.* **164** (1988) 217–314.
- [114] R. Contino, “The Higgs as a Composite Nambu-Goldstone Boson,” arXiv:1005.4269 [hep-ph].
- [115] K. Agashe, C. Grojean, R. Rattazzi, and M. Serone (organizers), “School on Strongly Coupled Physics Beyond the Standard Model.” ICTP SMR 2326. Lecture notes and videos available online, <http://indico.ictp.it/event/a11150/overview>, January, 2012.
- [116] B. Bellazzini, C. Csaki, and J. Serra, “Composite Higgses,” arXiv:1401.2457 [hep-ph].

- [117] M. Perelstein, “Little Higgs Models and Their Phenomenology,” *Prog.Part.Nucl.Phys.* **58** (2007) 247–291, arXiv:hep-ph/0512128 [hep-ph].
- [118] M. Schmaltz and D. Tucker-Smith, “Little Higgs Review,” *Ann.Rev.Nucl.Part.Sci.* **55** (2005) 229–270, arXiv:hep-ph/0502182 [hep-ph].
- [119] J. B. Kogut and L. Susskind, “Quark Confinement and the Puzzle of the Ninth Axial Current,” *Phys.Rev.* **D10** (1974) 3468–3475.
- [120] C. Quigg and R. Shrock, “Gedanken Worlds without Higgs: QCD-Induced Electroweak Symmetry Breaking,” *Phys.Rev.* **D79** (2009) 096002, arXiv:0901.3958 [hep-ph].
- [121] R. Jackiw and K. Johnson, “Dynamical Model of Spontaneously Broken Gauge Symmetries,” *Phys.Rev.* **D8** (1973) 2386–2398.
- [122] J. Cornwall and R. Norton, “Spontaneous Symmetry Breaking without Scalar Mesons,” *Phys.Rev.* **D8** (1973) 3338–3346.
- [123] M. Weinstein, “Conserved Currents, Their Commutators and the Symmetry Structure of Renormalizable Theories of Electromagnetic, Weak and Strong Interactions,” *Phys.Rev.* **D8** (1973) 2511.
- [124] S. Weinberg, “Implications of Dynamical Symmetry Breaking: an Addendum,” *Phys.Rev.* **D19** (1979) 1277–1280.
- [125] L. Susskind, “Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory,” *Phys.Rev.* **D20** (1979) 2619–2625.
- [126] S. Dimopoulos and L. Susskind, “Mass without Scalars,” *Nucl.Phys.* **B155** (1979) 237–252.
- [127] E. Eichten and K. D. Lane, “Dynamical Breaking of Weak Interaction Symmetries,” *Phys.Lett.* **B90** (1980) 125–130.
- [128] R. F. Dashen, “Chiral  $SU(3) \times SU(3)$  as a Symmetry of the Strong Interactions,” *Phys.Rev.* **183** (1969) 1245–1260.
- [129] R. F. Dashen, “Some Features of Chiral Symmetry Breaking,” *Phys.Rev.* **D3** (1971) 1879–1889.
- [130] E. Witten, “Baryons in the  $1/N$  Expansion,” *Nucl.Phys.* **B160** (1979) 57.
- [131] D. B. Kaplan and H. Georgi, “ $SU(2) \times U(1)$  Breaking by Vacuum Misalignment,” *Phys.Lett.* **B136** (1984) 183.
- [132] H. Georgi, “The Higgs as a Pseudo-Goldstone Boson,” *Comptes Rendus Physique* **8** (2007) 1029–1047.
- [133] H. Georgi and L. Randall, “Flavor Conserving CP Violation in Invisible Axion Models,” *Nucl.Phys.* **B276** (1986) 241.
- [134] A. Manohar and H. Georgi, “Chiral Quarks and the Nonrelativistic Quark Model,” *Nucl.Phys.* **B234** (1984) 189.
- [135] H. Georgi, “Generalized Dimensional Analysis,” *Phys.Lett.* **B298** (1993) 187–189, arXiv:hep-ph/9207278 [hep-ph].
- [136] K. Agashe, R. Contino, and A. Pomarol, “The Minimal Composite Higgs Model,” *Nucl.Phys.* **B719** (2005) 165–187, arXiv:hep-ph/0412089 [hep-ph].
- [137] R. Contino, L. Da Rold, and A. Pomarol, “Light Custodians in Natural Composite Higgs Models,” *Phys.Rev.* **D75** (2007) 055014, arXiv:hep-ph/0612048 [hep-ph].
- [138] R. Contino, C. Grojean, M. Moretti, F. Piccinini, and R. Rattazzi, “Strong Double Higgs Production at the Lhc,” *JHEP* **1005** (2010) 089, arXiv:1002.1011 [hep-ph].
- [139] J. R. Ellis, M. K. Gaillard, and D. V. Nanopoulos, “A Phenomenological Profile of the Higgs Boson,” *Nucl.Phys.* **B106** (1976) 292.
- [140] M. A. Shifman, A. Vainshtein, and V. I. Zakharov, “Remarks on Higgs Boson Interactions with Nucleons,” *Phys.Lett.* **B78** (1978) 443.



- [141] M. Gillioz, R. Grober, C. Grojean, M. Muhlleitner, and E. Salvioni, “Higgs Low-Energy Theorem (And Its Corrections) in Composite Models,” *JHEP* **1210** (2012) 004, arXiv:1206.7120 [hep-ph].
- [142] N. Craig, D. Stolarski, and J. Thaler, “A Fat Higgs with a Magnetic Personality,” *JHEP* **1111** (2011) 145, arXiv:1106.2164 [hep-ph].
- [143] C. Csaki, Y. Shirman, and J. Terning, “A Seiberg Dual for the Mssm: Partially Composite W and Z,” *Phys.Rev.* **D84** (2011) 095011, arXiv:1106.3074 [hep-ph].
- [144] Y. Nambu and G. Jona-Lasinio, “Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. II,” *Phys.Rev.* **124** (1961) 246–254.
- [145] Y. Nambu and G. Jona-Lasinio, “Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I.,” *Phys.Rev.* **122** (1961) 345–358.
- [146] E. Witten, “Some Inequalities among Hadron Masses,” *Phys.Rev.Lett.* **51** (1983) 2351.
- [147] C. Vafa and E. Witten, “Restrictions on Symmetry Breaking in Vector-Like Gauge Theories,” *Nucl.Phys.* **B234** (1984) 173.
- [148] D. B. Kaplan, H. Georgi, and S. Dimopoulos, “Composite Higgs Scalars,” *Phys.Lett.* **B136** (1984) 187.
- [149] T. Banks, “Constraints on  $SU(2) \times U(1)$  Breaking by Vacuum Misalignment,” *Nucl.Phys.* **B243** (1984) 125.
- [150] H. Georgi, D. B. Kaplan, and P. Galison, “Calculation of the Composite Higgs Mass,” *Phys.Lett.* **B143** (1984) 152.
- [151] H. Georgi and D. B. Kaplan, “Composite Higgs and Custodial  $SU(2)$ ,” *Phys.Lett.* **B145** (1984) 216.
- [152] H. Georgi and A. Pais, “Calculability and Naturalness in Gauge Theories,” *Phys.Rev.* **D10** (1974) 539.
- [153] H. Georgi and A. Pais, “Vacuum Symmetry and the Pseudogoldstone Phenomenon,” *Phys.Rev.* **D12** (1975) 508.
- [154] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, “Electroweak Symmetry Breaking from Dimensional Deconstruction,” *Phys.Lett.* **B513** (2001) 232–240, arXiv:hep-ph/0105239 [hep-ph].
- [155] M. Perelstein, M. E. Peskin, and A. Pierce, “Top quarks and electroweak symmetry breaking in little Higgs models,” *Phys.Rev.* **D69** (2004) 075002, arXiv:hep-ph/0310039 [hep-ph].
- [156] CMS Collaboration, S. Chatrchyan *et al.*, “Inclusive search for a vector-like T quark with charge  $\frac{2}{3}$  in pp collisions at  $\sqrt{s} = 8$  TeV,” *Phys.Lett.* **B729** (2014) 149–171, arXiv:1311.7667 [hep-ex].
- [157] A. De Simone, O. Matsedonskyi, R. Rattazzi, and A. Wulzer, “A First Top Partner Hunter’s Guide,” *JHEP* **1304** (2013) 004, arXiv:1211.5663 [hep-ph].
- [158] J. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, and M. Perez-Victoria, “Handbook of Vectorlike Quarks: Mixing and Single Production,” *Phys.Rev.* **D88** (2013) no. 9, 094010, arXiv:1306.0572 [hep-ph].
- [159] M. Buchkremer, G. Cacciapaglia, A. Deandrea, and L. Panizzi, “Model Independent Framework for Searches of Top Partners,” *Nucl.Phys.* **B876** (2013) 376–417, arXiv:1305.4172 [hep-ph].
- [160] C. T. Hill, S. Pokorski, and J. Wang, “Gauge Invariant Effective Lagrangian for Kaluza-Klein Modes,” *Phys.Rev.* **D64** (2001) 105005, arXiv:hep-th/0104035 [hep-th].
- [161] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, “(De)Constructing Dimensions,” *Phys.Rev.Lett.* **86** (2001) 4757–4761, arXiv:hep-th/0104005 [hep-th].
- [162] J. Thaler, “Little Technicolor,” *JHEP* **0507** (2005) 024, arXiv:hep-ph/0502175 [hep-ph].
- [163] Y. Kahn and J. Thaler, “Locality in Theory Space,” *JHEP* **1207** (2012) 007, arXiv:1202.5491

- [hep-ph].
- [164] H. Georgi, “Fun with Higgsless Theories,” *Phys.Rev.* **D71** (2005) 015016, arXiv:hep-ph/0408067 [hep-ph].
- [165] M. R. Douglas and G. W. Moore, “D-Branes, Quivers, and ALE Instantons,” arXiv:hep-th/9603167 [hep-th].
- [166] H. Georgi, “A Tool Kit for Builders of Composite Models,” *Nucl.Phys.* **B266** (1986) 274.
- [167] H. Georgi, “ $SU(2) \times U(1)$  Breaking, Compositeness, Flavor and Guts,” *Conf.Proc.* **C850701** (1985) 339–418.
- [168] T. Gregoire and J. G. Wacker, “Moose, Topology and Higgs,” *JHEP* **0208** (2002) 019, arXiv:hep-ph/0206023 [hep-ph].
- [169] N. Arkani-Hamed, A. Cohen, E. Katz, A. Nelson, T. Gregoire, *et al.*, “The Minimal Moose for a Little Higgs,” *JHEP* **0208** (2002) 021, arXiv:hep-ph/0206020 [hep-ph].
- [170] M. Schmaltz, D. Stolarski, and J. Thaler, “The Bestest Little Higgs,” *JHEP* **1009** (2010) 018, arXiv:1006.1356 [hep-ph].
- [171] L. Vecchi, “The Natural Composite Higgs,” arXiv:1304.4579 [hep-ph].
- [172] Z. Chacko, H.-S. Goh, and R. Harnik, “The Twin Higgs: Natural Electroweak Breaking from Mirror Symmetry,” *Phys.Rev.Lett.* **96** (2006) 231802, arXiv:hep-ph/0506256 [hep-ph].
- [173] R. Barbieri, T. Gregoire, and L. J. Hall, “Mirror World at the Large Hadron Collider,” arXiv:hep-ph/0509242 [hep-ph].
- [174] W. D. Goldberger, B. Grinstein, and W. Skiba, “Distinguishing the Higgs boson from the dilaton at the Large Hadron Collider,” *Phys.Rev.Lett.* **100** (2008) 111802, arXiv:0708.1463 [hep-ph].
- [175] J. Fan, W. D. Goldberger, A. Ross, and W. Skiba, “Standard Model couplings and collider signatures of a light scalar,” *Phys.Rev.* **D79** (2009) 035017, arXiv:0803.2040 [hep-ph].
- [176] L. Vecchi, “Phenomenology of a light scalar: the dilaton,” *Phys.Rev.* **D82** (2010) 076009, arXiv:1002.1721 [hep-ph].
- [177] Z. Chacko and R. K. Mishra, “Effective Theory of a Light Dilaton,” *Phys.Rev.* **D87** (2013) no. 11, 115006, arXiv:1209.3022 [hep-ph].
- [178] B. Bellazzini, C. Csaki, J. Hubisz, J. Serra, and J. Terning, “A Naturally Light Dilaton and a Small Cosmological Constant,” *Eur.Phys.J.* **C74** (2014) 2790, arXiv:1305.3919 [hep-th].
- [179] F. Coradeschi, P. Lodone, D. Pappadopulo, R. Rattazzi, and L. Vitale, “A naturally light dilaton,” *JHEP* **1311** (2013) 057, arXiv:1306.4601 [hep-th].
- [180] G. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi, “The Strongly-Interacting Light Higgs,” *JHEP* **0706** (2007) 045, arXiv:hep-ph/0703164 [hep-ph].
- [181] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner, and M. Spira, “Effective Lagrangian for a Light Higgs-Like Scalar,” *JHEP* **1307** (2013) 035, arXiv:1303.3876 [hep-ph].
- [182] G. Panico and A. Wulzer, “The Discrete Composite Higgs Model,” *JHEP* **1109** (2011) 135, arXiv:1106.2719 [hep-ph].
- [183] W. Skiba, “Tasi Lectures on Effective Field Theory and Precision Electroweak Measurements,” arXiv:1006.2142 [hep-ph].
- [184] J. D. Wells, “Tasi Lecture Notes: Introduction to Precision Electroweak Analysis,” arXiv:hep-ph/0512342 [hep-ph].
- [185] G. Cacciapaglia, C. Csaki, G. Marandella, and A. Strumia, “The Minimal Set of Electroweak Precision Parameters,” *Phys.Rev.* **D74** (2006) 033011, arXiv:hep-ph/0604111 [hep-ph].
- [186] Z. Han and W. Skiba, “Effective Theory Analysis of Precision Electroweak Data,” *Phys.Rev.* **D71** (2005) 075009, arXiv:hep-ph/0412166 [hep-ph].

- [187] R. Barbieri, A. Pomarol, R. Rattazzi, and A. Strumia, “Electroweak Symmetry Breaking After Lep-1 and Lep-2,” *Nucl.Phys.* **B703** (2004) 127–146, arXiv:hep-ph/0405040 [hep-ph].
- [188] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, *et al.*, “The Electroweak Fit of the Standard Model After the Discovery of a New Boson at the Lhc,” *Eur.Phys.J.* **C72** (2012) 2205, arXiv:1209.2716 [hep-ph].
- [189] P. Sikivie, L. Susskind, M. B. Voloshin, and V. I. Zakharov, “Isospin Breaking in Technicolor Models,” *Nucl.Phys.* **B173** (1980) 189.
- [190] S. Willenbrock, “Symmetries of the Standard Model,” arXiv:hep-ph/0410370 [hep-ph].
- [191] R. Barbieri, B. Bellazzini, V. S. Rychkov, and A. Varagnolo, “The Higgs Boson from an Extended Symmetry,” *Phys.Rev.* **D76** (2007) 115008, arXiv:0706.0432 [hep-ph].
- [192] A. Azatov, R. Contino, and J. Galloway, “Model-Independent Bounds on a Light Higgs,” *JHEP* **1204** (2012) 127, arXiv:1202.3415 [hep-ph].
- [193] A. Azatov and J. Galloway, “Electroweak Symmetry Breaking and the Higgs Boson: Confronting Theories at Colliders,” *Int.J.Mod.Phys.* **A28** (2013) 1330004, arXiv:1212.1380.
- [194] B. Bellazzini, C. Csaki, J. Hubisz, J. Serra, and J. Terning, “Composite Higgs Sketch,” *JHEP* **1211** (2012) 003, arXiv:1205.4032 [hep-ph].
- [195] A. Falkowski, F. Riva, and A. Urbano, “Higgs at Last,” *JHEP* **1311** (2013) 111, arXiv:1303.1812 [hep-ph].
- [196] I. Gogoladze, N. Okada, and Q. Shafi, “125 GeV Higgs Boson from Gauge-Higgs Unification: a Snowmass White Paper,” arXiv:1307.5079 [hep-ph].
- [197] J. Thaler and I. Yavin, “The Littlest Higgs in Anti-de Sitter Space,” *JHEP* **0508** (2005) 022, arXiv:hep-ph/0501036 [hep-ph].
- [198] P. J. Fox, A. E. Nelson, and N. Weiner, “Dirac Gaugino Masses and Supersoft Supersymmetry Breaking,” *JHEP* **0208** (2002) 035, arXiv:hep-ph/0206096 [hep-ph].
- [199] C. Csaki, J. Goodman, R. Pavesi, and Y. Shirman, “The  $M_D$ - $B_M$  Problem of Dirac Gauginos and Its Solutions,” *Phys.Rev.* **D89** (2014) 055005, arXiv:1310.4504 [hep-ph].
- [200] L. Randall and R. Sundrum, “Out of This World Supersymmetry Breaking,” *Nucl.Phys.* **B557** (1999) 79–118, arXiv:hep-th/9810155 [hep-th].
- [201] N. Weiner, “Unification without Unification,” *Phys.Rev.Lett.* (2001), arXiv:hep-ph/0106097 [hep-ph].
- [202] G. Burdman, Z. Chacko, H.-S. Goh, and R. Harnik, “Folded Supersymmetry and the Lep Paradox,” *JHEP* **0702** (2007) 009, arXiv:hep-ph/0609152 [hep-ph].
- [203] Y. Shirman, “Tasi 2008 Lectures: Introduction to Supersymmetry and Supersymmetry Breaking,” arXiv:0907.0039 [hep-ph].
- [204] Z. Komargodski, “Vector Mesons and an Interpretation of Seiberg Duality,” *JHEP* **1102** (2011) 019, arXiv:1010.4105 [hep-th].
- [205] R. Harnik, G. D. Kribs, D. T. Larson, and H. Murayama, “The Minimal Supersymmetric Fat Higgs Model,” *Phys.Rev.* **D70** (2004) 015002, arXiv:hep-ph/0311349 [hep-ph].
- [206] C. Csaki, L. Randall, and J. Terning, “Light Stops from Seiberg Duality,” *Phys.Rev.* **D86** (2012) 075009, arXiv:1201.1293 [hep-ph].
- [207] E. Halkiadakis, G. Redlinger, and D. Shih, “Status and Implications of Beyond-the-Standard-Model Searches at the LHC,” *Ann.Rev.Nucl.Part.Sci.* **64** (2014) 319–342, arXiv:1411.1427 [hep-ex].
- [208] J. Polchinski, *String Theory: Volume 2, Superstring Theory and Beyond*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1998.
- [209] R. Bertlmann, *Anomalies in Quantum Field Theory*. International Series of Monographs on Physics. Clarendon Press, 2000.

- [210] T. Frankel, *The Geometry of Physics: An Introduction*. Cambridge University Press, 2011.
- [211] T. Schücker, *Differential Geometry, Gauge Theories, and Gravity*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1989.
- [212] A. Collinucci and A. Wijn, “Topology of Fibre Bundles and Global Aspects of Gauge Theories,” arXiv:hep-th/0611201 [hep-th].
- [213] R. Sundrum, “Effective Field Theory for a Three-Brane Universe,” *Phys.Rev.* **D59** (1999) 085009, arXiv:hep-ph/9805471 [hep-ph].
- [214] L. Randall and M. D. Schwartz, “Quantum Field Theory and Unification in AdS5,” *JHEP* **0111** (2001) 003, arXiv:hep-th/0108114 [hep-th].
- [215] C. Csaki, J. Erlich, and J. Terning, “The Effective Lagrangian in the Randall-Sundrum Model and Electroweak Physics,” *Phys.Rev.* **D66** (2002) 064021, arXiv:hep-ph/0203034 [hep-ph].
- [216] M. Carena, S. Casagrande, F. Goertz, U. Haisch, and M. Neubert, “Higgs Production in a Warped Extra Dimension,” *JHEP* **1208** (2012) 156, arXiv:1204.0008 [hep-ph].
- [217] C. Csaki, Y. Grossman, P. Tanedo, and Y. Tsai, “Warped Penguin Diagrams,” *Phys.Rev.* **D83** (2011) 073002, arXiv:1004.2037 [hep-ph].
- [218] M. Blanke, B. Shakya, P. Tanedo, and Y. Tsai, “The Birds and the Bs in Rs: the B to S Gamma Penguin in a Warped Extra Dimension,” arXiv:1203.6650 [hep-ph].
- [219] M. Beneke, P. Dey, and J. Rohrwild, “The Muon Anomalous Magnetic Moment in the Randall-Sundrum Model,” *JHEP* **1308** (2013) 010, arXiv:1209.5897 [hep-ph].
- [220] K. Agashe, A. Azatov, Y. Cui, L. Randall, and M. Son, “Warped Dipole Completed, with a Tower of Higgs Bosons,” arXiv:1412.6468 [hep-ph].
- [221] C. Burgess, “Goldstone and Pseudogoldstone Bosons in Nuclear, Particle and Condensed Matter Physics,” *Phys.Rept.* **330** (2000) 193–261, arXiv:hep-th/9808176 [hep-th].
- [222] R. Coquereaux, “Comments About the Geometry of Nonlinear Sigma Models,” *Int.J.Mod.Phys.* **A2** (1987) 1763.
- [223] C. Yastremiz, “Nonlinear Realizations and Extended Objects,”. DAMTP-R-91-10, GTCRG-91-7.

## Flavor Physics and CP Violation

Z. Ligeti

Ernest Orlando Lawrence Berkeley National Laboratory,  
University of California, Berkeley, CA 94720

### Abstract

These notes overlap with lectures given at the TASI summer schools in 2014 and 2011, as well as at the European School of High Energy Physics in 2013. This is primarily an attempt at transcribing my hand-written notes, with emphasis on topics and ideas discussed in the lectures. It is not a comprehensive introduction or review of the field, nor does it include a complete list of references. I hope, however, that some may find it useful to better understand the reasons for excitement about recent progress and future opportunities in flavor physics.

### Preface

There are many books and reviews on flavor physics (e.g., Refs. [1–9]). The main points I would like to explain in these lectures are:

- *CP* violation and flavor-changing neutral currents (FCNC) are sensitive probes of short-distance physics, both in the standard model (SM) and in beyond standard model (BSM) scenarios.
- The data taught us a lot about not directly seen physics in the past, and are likely crucial to understand LHC new physics (NP) signals.
- In most FCNC processes  $\text{BSM/SM} \sim \mathcal{O}(20\%)$  is still allowed today, the sensitivity will improve to the few percent level in the future.
- Measurements are sensitive to very high scales, and might find unambiguous signals of BSM physics, even outside the LHC reach.
- There is a healthy and fun interplay of theoretical and experimental progress, with many open questions and important problems.

Flavor physics is interesting because there is a lot we do not understand yet. The “standard model flavor puzzle” refers to our lack of understanding of why and how the 6 quark and 6 lepton flavors differ, why masses and quark mixing are hierarchical, but lepton mixing is not. The “new physics flavor puzzle” is the tension between the relatively low scale required to solve the fine tuning problem (also suggested by the WIMP paradigm), and the high scale that is seemingly required to suppress the non-SM contributions to flavor changing processes. If there is NP at the TeV scale, we need to understand why and in what way its flavor structure is non-generic.

The key questions and prospects that make the future interesting are [7]

- What is the achievable experimental precision?  
The LHCb, Belle II, NA62, KOTO,  $\mu \rightarrow e\gamma$ ,  $\mu 2e$ , etc., experiments will improve the sensitivity in many modes by orders of magnitude.
- What are the theoretical uncertainties?  
In many key measurements, the theory uncertainty is well below future experimental sensitivity; while in some cases theoretical improvements are needed (so you can make an impact!).
- How large deviations from SM can we expect due to TeV-scale NP?  
New physics with generic flavor structure is ruled out; observable effects near current bounds are possible, many models predict some.

- What will the measurements teach us?

In all scenarios there is complementarity with high- $p_T$  measurements, and synergy in understanding the structure of any NP seen.

Another simple way to get a sense of (a lower bound on) the next 10–15 years of  $B$  physics progress is to consider the expected increase in data,

$$\frac{(\text{LHCb upgrade})}{(\text{LHCb } 1 \text{ fb}^{-1})} \sim \frac{(\text{Belle II data set})}{(\text{Belle data set})} \sim \frac{(\text{2009 BaBar data set})}{(\text{1999 CLEO data set})} \sim 50.$$

This will yield a  $\sqrt[4]{50} \sim 2.5$  increase in sensitivity to higher mass scales, even just by redoing existing measurements. More data has always motivated new theory ideas, yielding even faster progress. This is a comparable increase in reach as going from LHC7–8  $\rightarrow$  LHC13–14.

## Outline

The topics these lectures will cover include a brief introduction to flavor physics in the SM, testing the flavor structure in neutral meson mixing and  $CP$  violation, and examples of how to get theoretically clean information on short-distance physics. After a glimpse at the ingredients of the SM CKM fit, we discuss how sizable new physics contributions are still allowed in neutral meson mixing, and how this will improve in the future. Then we explain some implications of the heavy quark limit, tidbits of heavy quark symmetry, the operator product expansion and inclusive decays, to try to give an impression of what makes some hadronic physics tractable. The last lecture discusses some topics in TeV-scale flavor physics, top quark physics, Higgs flavor physics, bits of the interplay between searches for supersymmetry and flavor, and comments on minimal flavor violation. Some questions one may enjoy thinking about are in the footnotes.

## 1 Introduction to Flavor Physics and $CP$ Violation

Most of the experimentally observed particle physics phenomena are consistent with the standard model (SM). Evidence that the minimal SM is incomplete comes from the lack of a dark matter candidate, the baryon asymmetry of the Universe, its accelerating expansion, and nonzero neutrino masses. The baryon asymmetry and neutrino mixing are certainly connected to  $CP$  violation and flavor physics, and so may be dark matter. The hierarchy problem and seeking to identify the particle nature of dark matter strongly motivate TeV-scale new physics.

Studying flavor physics and  $CP$  violation provides a rich program to probe the SM and search for NP, with sensitivity to the  $1 - 10^5$  TeV scales, depending on details of the models. As we shall see, the sensitivity to BSM contributions to the dimension-6 four-quark operators mediating  $K$ ,  $D$ ,  $B_d$ , and  $B_s$  mixing, when parametrized by coefficients  $1/\Lambda^2$ , corresponds to scales  $\Lambda \sim 10^2 - 10^5$  TeV (see Table 1 and the related discussion below).

Understanding the origin of this sensitivity and how it can be improved, requires going into the details of a variety of flavor physics measurements.

### 1.1 Baryon asymmetry requires $CP$ violation beyond SM

The baryon asymmetry of the Universe is the measurement of

$$\frac{n_B - n_{\bar{B}}}{s} \approx 10^{-10}, \quad (1)$$

where  $n_B$  ( $n_{\bar{B}}$ ) is the number density of (anti-)baryons and  $s$  is the entropy density. This means that  $10^{-6}$  seconds after the Big Bang, when the temperature was  $T > 1$  GeV, and quarks and antiquarks were in thermal equilibrium, there was a corresponding asymmetry between quarks and antiquarks. Sakharov

pointed out [10] that for a theory to generate such an asymmetry in the course of its evolution from a hot Big Bang (assuming inflation washed out any possible prior asymmetry), it must contain:

1. baryon number violating interactions;
2.  $C$  and  $CP$  violation;
3. deviation from thermal equilibrium.

Interestingly, the SM contains 1–2–3, but (i)  $CP$  violation is too small, and (ii) the deviation from thermal equilibrium is too small at the electroweak phase transition. The SM expectation is many orders of magnitude below the observation, due to the suppression of  $CP$  violation by

$$[\prod_{u_i \neq u_j} (m_{u_i}^2 - m_{u_j}^2)] [\prod_{d_i \neq d_j} (m_{d_i}^2 - m_{d_j}^2)] / m_W^{12}, \quad (2)$$

and  $m_W$  indicates a typical weak interaction scale here.<sup>1</sup>

Therefore,  $CP$  violation beyond the SM must exist. While this argument does not tell us the scale of the corresponding new physics, it motivates searching for new sources of  $CP$  violation. (It may occur only in flavor-diagonal processes, such as EDMs, or only in the lepton sector, as in leptogenesis.) In any case, we want to understand the microscopic origin of  $CP$  violation, and how precisely we can test those  $CP$ -violating processes that we can measure.

Equally important is that almost all TeV-scale new physics models contain new sources of  $CP$  violation. Baryogenesis at the electroweak scale may still be viable, and the LHC will probe the remaining parameter space.

## 1.2 The SM and flavor

The SM is defined by the gauge interactions,

$$SU(3)_c \times SU(2)_L \times U(1)_Y, \quad (3)$$

the particle content, i.e., three generations of the fermion representations,

$$Q_L(3, 2)_{1/6}, \quad u_R(3, 1)_{2/3}, \quad d_R(3, 1)_{-1/3}, \quad L_L(1, 2)_{-1/2}, \quad \ell_R(1, 1)_{-1}, \quad (4)$$

and electroweak symmetry breaking. A condensate  $\langle \phi \rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}$  breaks  $SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM}$ , the dynamics of which we now know is well approximated by a seemingly elementary SM-like scalar Higgs field.

The kinetic terms in the SM Lagrangian are

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4} \sum_{\text{groups}} (F_{\mu\nu}^a)^2 + \sum_{\text{rep's}} \bar{\psi} i \not{D} \psi. \quad (5)$$

These are always  $CP$  conserving, as long as we neglect a possible  $F\tilde{F}$  term. The ‘‘strong  $CP$  problem’’ [11] is the issue of why the coefficient of the  $F\tilde{F}$  term for QCD is tiny. Its solution is an open question; however, we know that it is negligible for flavor-changing processes. The Higgs terms,

$$\mathcal{L}_{\text{Higgs}} = |D_\mu \phi|^2 + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2, \quad (6)$$

are  $CP$  conserving in the SM, but can be  $CP$  violating with an extended Higgs sector (already with two Higgs doublets; three are needed if natural flavor conservation is imposed [12]). Finally, the Yukawa couplings are,

$$\mathcal{L}_Y = -Y_{ij}^d \overline{Q_{Li}^I} \phi d_{Rj}^I - Y_{ij}^u \overline{Q_{Li}^I} \tilde{\phi} u_{Rj}^I - Y_{ij}^\ell \overline{L_{Li}^I} \phi \ell_{Rj}^I + \text{h.c.} \quad (7)$$

<sup>1</sup>Why is this suppression a product of all up and down quark mass differences, while fewer factors of mass splittings suppress  $CP$  violation in hadron decays and meson mixings?

The  $Y_{u,d}^{ij}$  are  $3 \times 3$  complex matrices,  $i, j$  are generation indices,  $\tilde{\phi} = i\sigma_2\phi^*$ .

After electroweak symmetry breaking, Eq. (7) gives quark mass terms,

$$\begin{aligned}\mathcal{L}_{\text{mass}} &= -\overline{d_{Li}^I}(M_d)_{ij}d_{Rj}^I - \overline{u_{Li}^I}(M_u)_{ij}u_{Rj}^I + \text{h.c.} \\ &= -(\overline{d_L^I}V_{dL}^\dagger)(V_{dL}M_dV_{dR}^\dagger)(V_{dR}d_R^I) \\ &\quad - (\overline{u_L^I}V_{uL}^\dagger)(V_{uL}M_uV_{uR}^\dagger)(V_{uR}u_R^I) + \text{h.c.},\end{aligned}\quad (8)$$

where  $M_f = (v/\sqrt{2})Y^f$ . The last two lines show the diagonalization of the mass matrices necessary to obtain the physical mass eigenstates,

$$M_f^{\text{diag}} \equiv V_{fL}M_fV_{fR}^\dagger, \quad f_{Li} \equiv V_{fL}^{ij}f_{Lj}^I, \quad f_{Ri} \equiv V_{fR}^{ij}f_{Rj}^I, \quad (9)$$

where  $f = u, d$  denote up- and down-type quarks. The diagonalization is different for  $u_{Li}$  and  $d_{Li}$ , which are in the same  $SU(2)_L$  doublet,

$$\begin{pmatrix} u_{Li}^I \\ d_{Li}^I \end{pmatrix} = (V_{uL}^\dagger)_{ij} \begin{pmatrix} u_{Lj} \\ (V_{uL}V_{dL}^\dagger)_{jk}d_{Lk} \end{pmatrix}. \quad (10)$$

The ‘‘misalignment’’ between these two transformations,

$$V_{\text{CKM}} \equiv V_{uL}V_{dL}^\dagger, \quad (11)$$

is the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix. By virtue of Eq. (11), it is unitary.

Eq. (10) shows that the charged current weak interactions, which arise from the  $\bar{\psi}i\not{D}\psi$  terms in Eq. (5), become non-diagonal in the mass basis

$$-\frac{g}{2}\overline{Q_{Li}^I}\gamma^\mu W_\mu^a\tau^a Q_{Li}^I + \text{h.c.} \Rightarrow -\frac{g}{\sqrt{2}}(\overline{u_L}, \overline{c_L}, \overline{t_L})\gamma^\mu W_\mu^+ V_{\text{CKM}} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.}, \quad (12)$$

where  $W_\mu^\pm = (W_\mu^1 \mp W_\mu^2)/\sqrt{2}$ . Thus, charged-current weak interactions change flavor, and this is the only flavor-changing interaction in the SM.

In the absence of Yukawa couplings, the SM has a global  $[U(3)]^5$  symmetry ( $[U(3)]^3$  in the quark and  $[U(3)]^2$  in the lepton sector), rotating the 3 generations of the 5 fields in Eq. (4). This is broken by the Yukawa interactions in Eq. (7). In the quark sector the breaking is

$$U(3)_Q \times U(3)_u \times U(3)_d \rightarrow U(1)_B, \quad (13)$$

In the lepton sector, we do not yet know if  $U(3)_L \times U(3)_\ell$  is fully broken.

### 1.3 Flavor and $CP$ violation in the SM

Since the  $Z$  couples flavor diagonally,<sup>2</sup> there are no tree-level flavor-changing neutral currents, such as  $K_L \rightarrow \mu^+\mu^-$ . This led GIM [13] to predict the existence of the charm quark. Similarly,  $K^0 - \bar{K}^0$  mixing vanishes at tree-level, which allowed the prediction of  $m_c$  [14, 15] before the discovery of the charm quark. In the previous examples, because of the unitarity of the CKM matrix,

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0. \quad (14)$$

Expanding the loop functions, e.g., in a FCNC kaon decay amplitude,

$$V_{ud}V_{us}^*f(m_u) + V_{cd}V_{cs}^*f(m_c) + V_{td}V_{ts}^*f(m_t), \quad (15)$$

<sup>2</sup>Show that there are no tree-level flavor-changing  $Z$  couplings in the SM. What if, besides doublets, there were a left-handed  $SU(2)$  singlet quark field as well?



the result is always proportional to the up-quark mass-squared differences,

$$\frac{m_i^2 - m_j^2}{m_W^2}. \quad (16)$$

So FCNCs probe directly the differences between the generations.

One can also see that  $CP$  violation is related to irremovable phases of Yukawa couplings. Starting from a term in Eq. (7),

$$Y_{ij} \overline{\psi_{Li}} \phi \psi_{Rj} + Y_{ij}^* \overline{\psi_{Rj}} \phi^\dagger \psi_{Li} \xrightarrow{CP} Y_{ij} \overline{\psi_{Rj}} \phi^\dagger \psi_{Li} + Y_{ij}^* \overline{\psi_{Li}} \phi \psi_{Rj}. \quad (17)$$

The two expressions are identical if and only if a basis for the quark fields can be chosen such that  $Y_{ij} = Y_{ij}^*$ , i.e., that  $Y_{ij}$  are real.

#### 1.4 Counting flavor parameters

Most parameters of the SM (and also of many of its extensions) are related to flavor. In the CKM matrix, due to unitarity, 9 complex elements depend on 9 real parameters. Of these 5 phases can be absorbed by redefining the quark fields, leaving 4 physical parameters, 3 mixing angles and 1  $CP$  violating phase. This is the only source of  $CP$  violation in flavor-changing transitions in the SM.

A more general way to account for all flavor parameters is to consider that the two Yukawa matrices,  $Y_{i,j}^{u,d}$  in Eq. (7), contain 18 real and 18 imaginary parameters. They break the global  $[U(3)]^3 \rightarrow U(1)_B$ , see Eq. (13), so there are 26 broken generators (9 real and 17 imaginary). This leaves 10 physical quark flavor parameters: 9 real ones (the 6 quark masses and 3 mixing angles) and 1 complex  $CP$  violating phase.<sup>3</sup>

#### 1.5 Neutrino masses

How does lepton flavor differ? With the particle content in Eq. (4), it is not possible to write down a renormalizable mass term for neutrinos. It would require introducing a  $\nu_R(1, 1)_0$  field, a singlet under all SM gauge groups, to be light, which is unexpected. Such a particle is sometimes called a sterile neutrino, as it has no SM interactions. Whether there are such fields can only be decided experimentally.

Viewing the SM as a low energy effective theory, there is a single type of dimension-5 gauge invariant term made of SM fields,

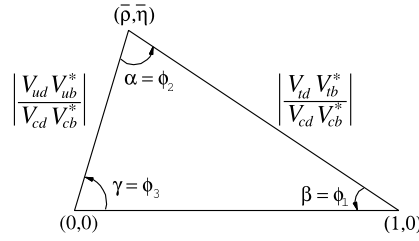
$$\mathcal{L}_Y = -\frac{Y_\nu^{ij}}{\Lambda_{\text{NP}}} L_{Li}^I L_{Lj}^I \phi \phi. \quad (18)$$

This term gives rise to neutrino masses and also violates lepton number. Its suppression cannot be the electroweak scale,  $1/v$  (instead of  $1/\Lambda_{\text{NP}}$ ), because such a term in the Lagrangian cannot be generated from SM fields at arbitrary loop level, or even nonperturbatively. [Eq. (18) violates  $B - L$ , which is an accidental symmetry of the SM that is not anomalous.] The above mass term is called a Majorana mass, as it couples  $\overline{\nu_L}$  to  $(\nu_L)^c$  instead of  $\nu_R$  [the latter occurs for Dirac mass terms, see Eq. (8)]. The key distinction is whether lepton number is violated or conserved. In the presence of Eq. (18) and the charged lepton Yukawa coupling in the last term in Eq. (7), the global  $U(3)_L \times U(3)_\ell$  symmetry is completely broken, and the counting of lepton flavor parameters is<sup>4</sup>

$$(12 + 18 \text{ couplings}) - (18 \text{ broken sym.}) \Rightarrow 12 \text{ physical parameters}. \quad (19)$$

<sup>3</sup>Show that for  $N$  generations, the CKM matrix depends on  $N(N-1)/2$  mixing angles and  $(N-1)(N-2)/2$   $CP$  violating phases. So the 2-generation SM conserves  $CP$ .

<sup>4</sup>Show that the Yukawa matrix in Eq. (18) is symmetric,  $Y_\nu^{ij} = Y_\nu^{ji}$ . Derive that for  $N$  such generations there are  $N(N-1)/2$   $CP$  violating phases.



**Fig. 1:** The unitarity triangle.

These are the 6 masses, 3 mixing angles, and 3  $CP$  violating phases, of which one is the analog of the CKM phase measurable in oscillation experiments, while two additional “Majorana phases” only contribute to lepton number violating processes, such as neutrinoless double beta decay.<sup>5</sup>

### 1.6 The CKM matrix

Quark mixing is observed to be approximately flavor diagonal. The Wolfenstein parametrization conveniently exhibits this,

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \dots, \quad (20)$$

where  $\lambda \simeq 0.23$  may be viewed as an “expansion parameter”. It is a useful book-keeping of the magnitudes of the CKM matrix elements, but it hides which combination of CKM elements are phase-convention independent. Sometimes it can be useful to think of  $V_{ub}$  and  $V_{td}$  as the ones with  $\mathcal{O}(1)$   $CP$  violating phases, but it is important that any  $CP$  violating observable in the SM must depend on at least four CKM elements.<sup>6</sup>

In any case, the interesting question is not primarily measuring CKM elements, but testing how precisely the SM description of flavor and  $CP$  violation holds. This can be done by “redundant” measurements, which in the SM relate to some combination of flavor parameters, but are sensitive to different BSM physics, thus testing for (in)consistency. Since there are many experimental constraints, a simple way to compare different measurements can be very useful. Recall that CKM unitarity implies

$$\sum_k V_{ik} V_{jk}^* = \sum_k V_{ki} V_{kj}^* = \delta_{ij}, \quad (21)$$

and the 6 vanishing relations can be represented as triangles in a complex plane. The most often used such “unitarity triangle” (shown in Fig. 1) arises from the scalar product of the 1st and 3rd columns,

$$V_{ud} V_{ub}^* + V_{cd} V_{cb}^* + V_{td} V_{tb}^* = 0. \quad (22)$$

(Unitarity triangles constructed from neighboring columns or rows are “squashed”.) We define the  $\alpha$ ,  $\beta$ ,  $\gamma$  angles of this triangle, and two more,

$$\begin{aligned} \alpha &\equiv \arg\left(-\frac{V_{td} V_{tb}^*}{V_{ud} V_{ub}^*}\right), & \beta &\equiv \arg\left(-\frac{V_{cd} V_{cb}^*}{V_{td} V_{tb}^*}\right), & \gamma &\equiv \arg\left(-\frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*}\right), \\ \beta_s &\equiv \arg\left(-\frac{V_{ts} V_{tb}^*}{V_{cs} V_{cb}^*}\right), & \beta_K &\equiv \arg\left(-\frac{V_{cs} V_{cd}^*}{V_{us} V_{ud}^*}\right). \end{aligned} \quad (23)$$

<sup>5</sup>Can you think of ways to get sensitivity to another linear combination of the two  $CP$  violating Majorana phases, besides the one that enters neutrinoless double beta decay?

<sup>6</sup>Prove this statement. Are there constraints on which four?

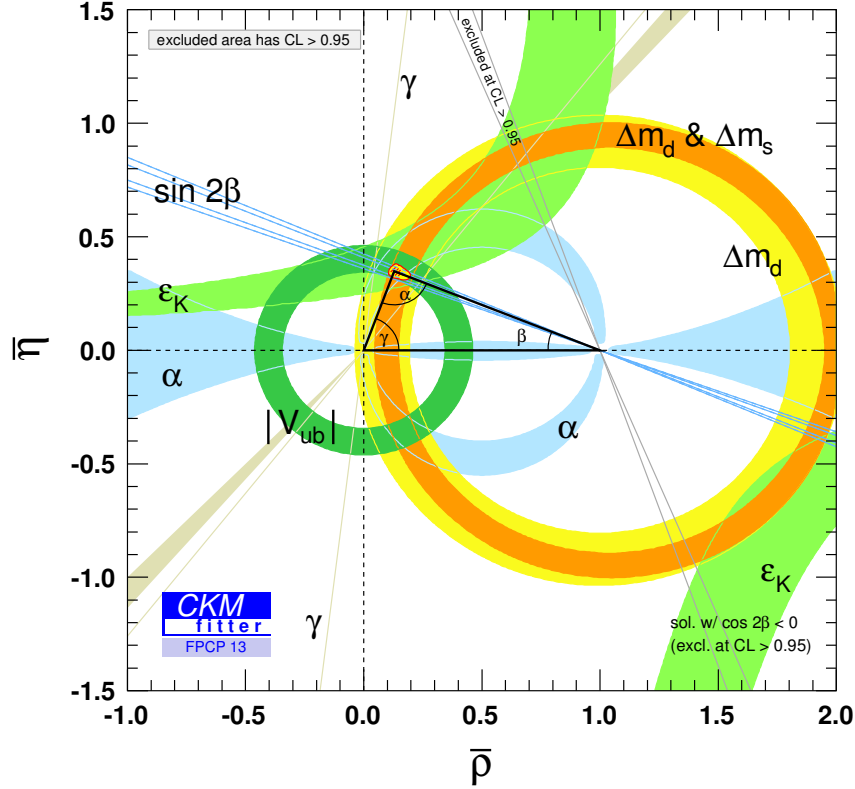


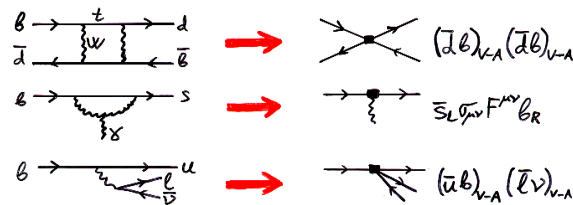
Fig. 2: The SM CKM fit, and individual constraints (colored regions show 95% CL).

On different continents the  $\phi_1 = \beta$ ,  $\phi_2 = \alpha$ ,  $\phi_3 = \gamma$ , and/or the  $\phi_s = -2\beta_s$  notations are used. Here  $\beta_s$  ( $\beta_K$ ), of order  $\lambda^2$  ( $\lambda^4$ ), is the small angle of a “squashed” unitarity triangle obtained by multiplying the 2nd column of the CKM matrix with the 3rd (1st) column.

The magnitudes of CKM elements determine the sides of the unitarity triangle. They are mainly extracted from semileptonic and leptonic  $K$  and  $B$  decays, and  $B_{d,s}$  mixing. Any constraint which renders the area of the unitarity triangle nonzero, such as angles, has to measure  $CP$  violation. Some of the most important constraints are shown in Fig. 2, together with the CKM fit in the SM. (Using  $\bar{\rho}$ ,  $\bar{\eta}$  instead of  $\rho$ ,  $\eta$  simply corresponds to a small modification of the parametrization, to keep unitarity exact.)

### 1.7 The low energy effective field theory (EFT) viewpoint

At the few GeV scale, relevant for  $B$ ,  $D$ , and some  $K$  decays, all flavor changing processes (both tree and loop level) are mediated by dozens of higher dimension local operators. They arise from integrating out heavy particles,  $W$  and  $Z$  bosons and the  $t$  quark in the SM, or not yet observed heavy states (see Fig. 3). Since the coefficients of a large number of operators depend on just a few parameters in the SM, there are many correlations between decays of hadrons containing  $s$ ,  $c$ ,  $b$  quarks, which NP may violate. From this point of view there is no difference between flavor-changing neutral currents and  $\Delta F = 1$  processes, as all flavor-changing processes are due to heavy particles with masses  $\gg m_{s,c,b}$ . Thus, one can test the SM in many ways by asking (i) does NP modify the coefficients of dimension-6 operators? (ii) does NP generate operators absent in the SM (e.g., right-handed couplings)?



**Fig. 3:** Diagrams at the electroweak scale (left) and operators at the scale  $m_b$  (right).

**Table 1:** Bounds on some  $\Delta F = 2$  operators,  $(C/\Lambda^2) \mathcal{O}$ , with  $\mathcal{O}$  given in the first column. The bounds on  $\Lambda$  assume  $C = 1$ , the bounds on  $C$  assume  $\Lambda = 1$  TeV. (From Ref. [19].)

Operator	Bound on $\Lambda$ [TeV] ( $C = 1$ )		Bound on $C$ ( $\Lambda = 1$ TeV)		Observables
	Re	Im	Re	Im	
$(\bar{s}_L \gamma^\mu d_L)^2$	$9.8 \times 10^2$	$1.6 \times 10^4$	$9.0 \times 10^{-7}$	$3.4 \times 10^{-9}$	$\Delta m_K; \epsilon_K$
$(\bar{s}_R d_L)(\bar{s}_L d_R)$	$1.8 \times 10^4$	$3.2 \times 10^5$	$6.9 \times 10^{-9}$	$2.6 \times 10^{-11}$	$\Delta m_K; \epsilon_K$
$(\bar{c}_L \gamma^\mu u_L)^2$	$1.2 \times 10^3$	$2.9 \times 10^3$	$5.6 \times 10^{-7}$	$1.0 \times 10^{-7}$	$\Delta m_D;  q/p , \phi_D$
$(\bar{c}_R u_L)(\bar{c}_L u_R)$	$6.2 \times 10^3$	$1.5 \times 10^4$	$5.7 \times 10^{-8}$	$1.1 \times 10^{-8}$	$\Delta m_D;  q/p , \phi_D$
$(\bar{b}_L \gamma^\mu d_L)^2$	$6.6 \times 10^2$	$9.3 \times 10^2$	$2.3 \times 10^{-6}$	$1.1 \times 10^{-6}$	$\Delta m_{B_d}; S_{\psi K_S}$
$(\bar{b}_R d_L)(\bar{b}_L d_R)$	$2.5 \times 10^3$	$3.6 \times 10^3$	$3.9 \times 10^{-7}$	$1.9 \times 10^{-7}$	$\Delta m_{B_d}; S_{\psi K_S}$
$(\bar{b}_L \gamma^\mu s_L)^2$	$1.4 \times 10^2$	$2.5 \times 10^2$	$5.0 \times 10^{-5}$	$1.7 \times 10^{-5}$	$\Delta m_{B_s}; S_{\psi\phi}$
$(\bar{b}_R s_L)(\bar{b}_L s_R)$	$4.8 \times 10^2$	$8.3 \times 10^2$	$8.8 \times 10^{-6}$	$2.9 \times 10^{-6}$	$\Delta m_{B_s}; S_{\psi\phi}$

### 1.8 Neutral meson mixing

Let us first sketch a back-of-an-envelope estimate of the mass difference in  $K^0 - \bar{K}^0$  mixing. In the SM,

$$\Delta m_K \sim \alpha_w^2 |V_{cs} V_{cd}|^2 \frac{m_c^2 - m_u^2}{m_W^4} f_K^2 m_K. \quad (24)$$

The result is suppressed by CKM angles, a loop factor, the weak coupling, and the GIM mechanism. If a heavy particle,  $X$ , contributes  $\mathcal{O}(1)$  to  $\Delta m_K$ ,

$$\left| \frac{\Delta m_K^{(X)}}{\Delta m_K^{(\text{exp})}} \right| \sim \left| \frac{g^2 \Lambda_{\text{QCD}}^3}{M_X^2 \Delta m_K^{(\text{exp})}} \right| \Rightarrow \frac{M_X}{g} \gtrsim 2 \times 10^3 \text{ TeV}. \quad (25)$$

So even TeV-scale particles with loop-suppressed couplings [ $g \sim \mathcal{O}(10^{-3})$ ] can give observable effects. This illustrates that flavor physics measurements indeed probe the TeV scale if NP has SM-like flavor structure, and much higher scales if the NP flavor structure is generic.

A more careful evaluation of the bounds in all four neutral meson systems is shown in Table 1. (See Sec. 2 for the definitions of the observables in the  $B$  meson systems.) If  $\Lambda = \mathcal{O}(1 \text{ TeV})$  then  $C \ll 1$ , and if  $C = \mathcal{O}(1)$  then  $\Lambda \gg 1 \text{ TeV}$ . The bounds are weakest for  $B_{(s)}$  mesons, as mixing is the least suppressed in the SM in that case. The bounds on many NP models are the strongest from  $\Delta m_K$  and  $\epsilon_K$ , since so are the SM suppressions. These are built into NP models since the 1970s, otherwise the models are immediately excluded. In the SM, larger FCNCs and  $CP$  violating effects occur in  $B$  mesons, which can be measured precisely. In many BSM models the 3rd generation is significantly different than the first two, motivated by the large top Yukawa, and may give larger signals in the  $B$  sector.

### 1.9 A few more words on kaons

With recent lattice QCD progress on  $B_K$  and  $f_K$  [16],  $\epsilon_K$  has become a fairly precise constraint on the SM. However,  $\epsilon'_K$  is notoriously hard to calculate, involving cancellation between two comparable terms,

each with sizable uncertainties. (Lattice QCD calculations of the hadronic matrix elements for  $\epsilon'_K$  may be reliably computed in the future.) At present, we cannot prove nor rule out that a large part of the observed value of  $\epsilon'_K$  is due to BSM. Thus, to test  $CP$  violation, one had to consider other systems; it was realized in the 1980s that many precise measurements of  $CP$  violation are possible in  $B$  decays.

In the kaon sector, precise calculations of rare decays involving neutrinos (see Fig. 4) are possible, and the SM predictions are [17]

$$\mathcal{B}(K^+ \rightarrow \pi^+ \nu \bar{\nu}) = (8.4 \pm 1.0) \times 10^{-11}, \quad \mathcal{B}(K_L^0 \rightarrow \pi^0 \nu \bar{\nu}) = (3.4 \pm 0.6) \times 10^{-11}. \quad (26)$$

The  $K_L^0$  decay is  $CP$  violating, and therefore it is under especially good theoretical control, since it is determined by the top quark loop contribution, and the  $CP$  conserving charm quark contribution is absent (which enters  $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ , and is subject to some hadronic uncertainty).

The E787/E949 measurement is  $\mathcal{B}(K \rightarrow \pi^+ \nu \bar{\nu}) = (17.3_{-10.5}^{+11.5}) \times 10^{-11}$  [18], whereas in the  $K_L$  mode the experimental upper bound is still many times the SM rate. NA62 at CERN aims to measure the  $K^+$  rate with 10% uncertainty, and will start to have dozens of events in 2015. The  $K_L$  mode will probably be first observed by the KOTO experiment at J-PARC.

## 2 Theory of Some Important $B$ Decays

Studying FCNC and  $CP$  violation is particularly interesting in  $B$  meson decays, because many measurements are possible with clean interpretations.

The main theoretical reasons are: (i)  $t$  quark loops are neither GIM nor CKM suppressed; (ii) large  $CP$  violating effects are possible; (iii) some of the hadronic physics is understandable model independently ( $m_b \gg \Lambda_{\text{QCD}}$ ).

The main experimental reasons are: (i) the long  $B$  lifetime (small  $|V_{cb}|$ ); (ii) the  $\Upsilon(4S)$  is a clean source of  $B$  mesons at  $e^+e^-$  colliders; (iii) for  $B_d$ , the ratio  $\Delta m/\Gamma = \mathcal{O}(1)$ .

### 2.1 Neutral meson mixing formalism

Similar to neutral kaons, there are two neutral  $B^0$  meson flavor eigenstates,

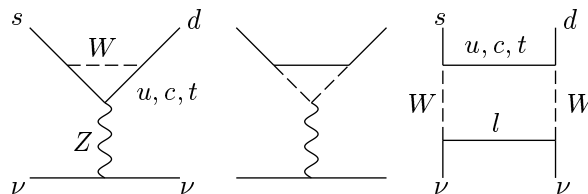
$$|B^0\rangle = |\bar{b}d\rangle, \quad |\bar{B}^0\rangle = |b\bar{d}\rangle. \quad (27)$$

They mix in the SM due to weak interactions (see Fig. 5). The time evolutions of the two states are described by the Schrödinger equation,

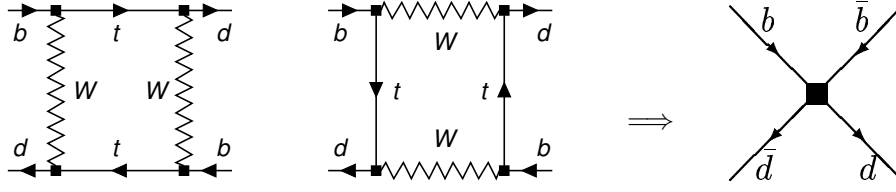
$$i \frac{d}{dt} \begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix} = \left( M - \frac{i}{2} \Gamma \right) \begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix}, \quad (28)$$

where the mass ( $M$ ) and the decay ( $\Gamma$ ) mixing matrices are  $2 \times 2$  Hermitian matrices.  $CPT$  invariance implies  $M_{11} = M_{22}$  and  $\Gamma_{11} = \Gamma_{22}$ . The heavier and lighter mass eigenstates are the eigenvectors of  $M - i\Gamma/2$ ,

$$|B_{H,L}\rangle = p |B^0\rangle \mp q |\bar{B}^0\rangle, \quad (29)$$



**Fig. 4:** Diagrams contributing to  $K \rightarrow \pi \nu \bar{\nu}$  decay.



**Fig. 5:** Left: box diagrams that give rise to the  $B^0 - \bar{B}^0$  mass difference; Right: operator in the effective theory below  $m_W$  whose  $B$  meson matrix element determines  $\Delta m_B$ .

and their time dependence is

$$|B_{H,L}(t)\rangle = e^{-(im_{H,L} + \Gamma_{H,L}/2)t} |B_{H,L}\rangle. \quad (30)$$

Here  $\Delta m \equiv m_H - m_L$  and  $\Delta\Gamma = \Gamma_L - \Gamma_H$  are the mass and width differences. This defines  $\Delta m$  to be positive, but the sign of  $\Delta\Gamma$  is physical. Note that  $m_{H,L}$  ( $\Gamma_{H,L}$ ) are not the eigenvalues of  $M$  ( $\Gamma$ ).<sup>7</sup> The off-diagonal elements,  $M_{12}$  and  $\Gamma_{12}$ , arise from virtual and on-shell intermediate states, respectively. In the SM,  $M_{12}$  is dominated by the top-quark box diagrams in Fig. 5. Thus,  $M_{12}$  is determined by short-distance physics, it is calculable with good accuracy, and is sensitive to high scales. (This is the complication for  $D$  mixing: the  $W$  can always be shrunk to a point, but the  $d$  and  $s$  quarks in the box diagrams cannot, so long-distance effects are important.) The width difference  $\Gamma_{12}$  is determined by on-shell states to which both  $B^0$  and  $\bar{B}^0$  can decay, corresponding to  $c$  and  $u$  quarks in the box diagrams.

The solution of the eigenvalue equation is

$$\begin{aligned} (\Delta m)^2 - \frac{(\Delta\Gamma)^2}{4} &= 4|M_{12}|^2 - |\Gamma_{12}|^2, & \Delta m \Delta\Gamma &= -4 \operatorname{Re}(M_{12}\Gamma_{12}^*), \\ \frac{q}{p} &= -\frac{\Delta m + i\Delta\Gamma/2}{2M_{12} - i\Gamma_{12}} = -\frac{2M_{12}^* - i\Gamma_{12}^*}{\Delta m + i\Delta\Gamma/2}. \end{aligned} \quad (33)$$

The physical observables that are measurable in neutral meson mixing are

$$x = \frac{\Delta m}{\Gamma}, \quad y = \frac{\Delta\Gamma}{2\Gamma}, \quad \left| \frac{q}{p} \right| - 1. \quad (34)$$

The orders of magnitudes of the SM predictions are shown in Table 2. That  $x \neq 0$  is established in the  $K$ ,  $B$ , and  $B_s$  mixing;  $y \neq 0$  in the  $K$ ,  $D$ , and  $B_s$  mixing;  $|q/p| \neq 1$  in  $K$  mixing. The significance of  $x_D \neq 0$  is  $\sim 2\sigma$ , and in  $B_{d,s}$  mixing there is an unconfirmed  $D\emptyset$  signal for  $|q/p| \neq 1$ ; more below.

Simpler approximate solutions can be obtained expanding about the limit  $|\Gamma_{12}| \ll |M_{12}|$ . This is a good approximation in both  $B_d$  and  $B_s$  systems.  $|\Gamma_{12}| < \Gamma$  always holds, because  $\Gamma_{12}$  arises from decays to final states common to  $B^0$  and  $\bar{B}^0$ . For  $B_s$  mixing the world average is  $\Delta\Gamma_s/\Gamma_s = 0.138 \pm 0.012$  [20], while  $\Delta\Gamma_d$  is expected to be  $\sim 20$  times smaller and is not yet measured. Up to higher order terms in  $|\Gamma_{12}/M_{12}|$ , Eqs. (33) become

$$\Delta m = 2|M_{12}|, \quad \Delta\Gamma = -2 \frac{\operatorname{Re}(M_{12}\Gamma_{12}^*)}{|M_{12}|},$$

<sup>7</sup>Derive that the time evolutions of mesons that are  $B^0$  and  $\bar{B}^0$  at  $t = 0$  are given by

$$|B^0(t)\rangle = g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\bar{B}^0\rangle, \quad |\bar{B}^0(t)\rangle = \frac{p}{q}g_-(t)|B^0\rangle + g_+(t)|\bar{B}^0\rangle, \quad (31)$$

where, denoting  $m = (m_H + m_L)/2$  and  $\Gamma = (\Gamma_H + \Gamma_L)/2$ ,

$$\begin{aligned} g_+(t) &= e^{-it(m-i\Gamma/2)} \left( \cosh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} - i \sinh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \right), \\ g_-(t) &= e^{-it(m-i\Gamma/2)} \left( -\sinh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} + i \cosh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \right). \end{aligned} \quad (32)$$

**Table 2:** Orders of magnitudes of the SM predictions for mixing parameters. The uncertainty of  $(|q/p| - 1)_D$  is especially large.

meson	$x = \Delta m/\Gamma$	$y = \Delta\Gamma/(2\Gamma)$	$ q/p  - 1$
$K$	1	1	$10^{-3}$
$D$	$10^{-2}$	$10^{-2}$	$10^{-3}$
$B_d$	1	$10^{-2}$	$10^{-4}$
$B_s$	$10^1$	$10^{-1}$	$10^{-5}$

$$\frac{q}{p} = -\frac{M_{12}^*}{|M_{12}|} \left( 1 - \frac{1}{2} \text{Im} \frac{\Gamma_{12}}{M_{12}} \right), \quad (35)$$

where we kept the second term in  $q/p$ , as it will be needed later.

## 2.2 CP violation in decay

This is any form of  $CP$  violation that cannot be absorbed in a neutral meson mixing amplitude (also called direct  $CP$  violation). It can occur in any hadron decay, as opposed to those specific to neutral mesons discussed below. For a given final state,  $f$ , the  $B \rightarrow f$  and  $\bar{B} \rightarrow \bar{f}$  decay amplitudes can, in general, receive several contributions

$$A_f = \langle f | \mathcal{H} | B \rangle = \sum_k A_k e^{i\delta_k} e^{i\phi_k}, \quad \bar{A}_{\bar{f}} = \langle \bar{f} | \mathcal{H} | \bar{B} \rangle = \sum_k A_k e^{i\delta_k} e^{-i\phi_k}. \quad (36)$$

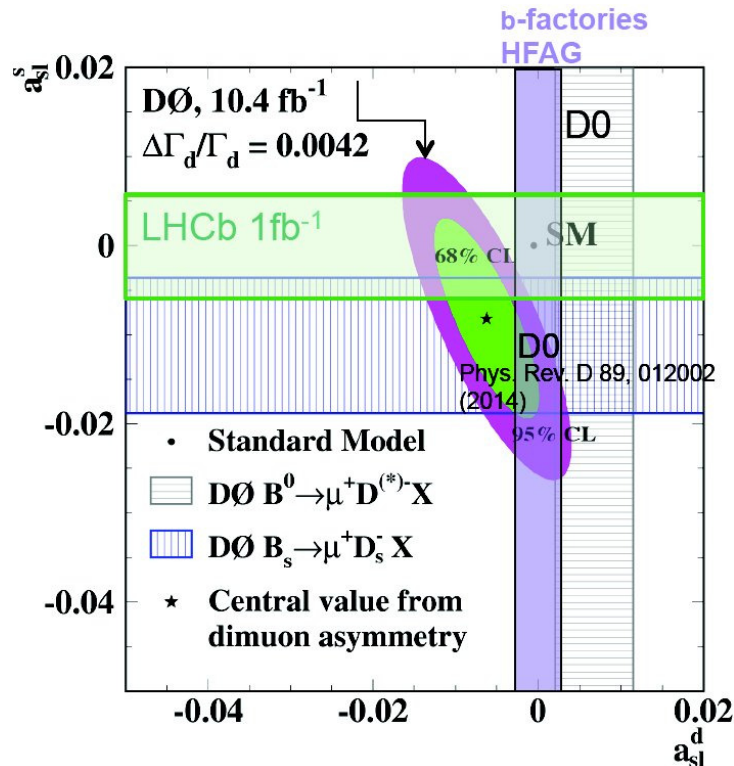
There are two types of complex phases. Complex parameters in the Lagrangian which enter a decay amplitude also enter the  $CP$  conjugate amplitude but in complex conjugate form. In the SM such “weak phases”,  $\phi_k$ , only occur in the CKM matrix. Another type of phase is due to absorptive parts of decay amplitudes, and gives rise to  $CP$  conserving “strong phases”,  $\delta_k$ . These phases arise from on-shell intermediate states rescattering into the desired final state, and they are the same for an amplitude and its  $CP$  conjugate. The individual phases  $\delta_k$  and  $\phi_k$  are convention dependent, but the phase differences,  $\delta_i - \delta_j$  and  $\phi_i - \phi_j$ , and therefore  $|\bar{A}_{\bar{f}}|$  and  $|A_f|$ , are physical. Clearly, if  $|\bar{A}_{\bar{f}}| \neq |A_f|$  then  $CP$  is violated; this is called  $CP$  violation in decay, or direct  $CP$  violation.<sup>8</sup>

There are many measurements of direct  $CP$  violation. While some give strong constraints on NP models which evade the SM suppressions (e.g.,  $\epsilon'_K$ , the first direct  $CP$  violation measured with high significance), at present no single direct  $CP$  violation measurement gives a precise test of the SM, due to the lack of reliable calculations of relevant strong phases. For all observations of direct  $CP$  violation in a single decay mode, viewed in isolation [see the caveat near Eq. (42)], it is possible that, say, half of the measured value is from BSM. For  $\epsilon'_K$ , lattice QCD may yield progress in the future. In certain  $B$  decays we may better understand the implications of the heavy quark limit; so far  $A_{K^+\pi^0} - A_{K^+\pi^-} = 0.12 \pm 0.02$  [20], the “ $K\pi$  puzzle”, is poorly understood.

## 2.3 CP violation in mixing

If  $CP$  were conserved, the mass and  $CP$  eigenstates would coincide, and the mass eigenstates would be proportional to  $|B^0\rangle \pm |\bar{B}^0\rangle$ , up to phases; i.e.,  $|q/p| = 1$  and  $\arg(M_{12}/\Gamma_{12}) = 0$ . If  $|q/p| \neq 1$ , then  $CP$  is violated. This is called  $CP$  violation in mixing. It follows from Eq. (29) that  $\langle B_H | B_L \rangle = |p|^2 - |q|^2$ , so if  $CP$  is violated in mixing, the physical states are not orthogonal. (This illustrates again that  $CP$  violation is a quantum mechanical effect, impossible in a classical system.) The simplest example is the

<sup>8</sup>Derive that direct  $CP$  violation requires interference of at least two contributing amplitudes with different strong and weak phases,  $|\bar{A}|^2 - |A|^2 = 4A_1A_2 \sin(\delta_1 - \delta_2) \sin(\phi_1 - \phi_2)$ .



**Fig. 6:** Status of  $A_{\text{SL}}$  measurements (from M. Artuso, talk at FPCP 2014). The  $D\bar{0}$  result is in a  $3.6\sigma$  tension with the SM expectation.

$CP$  asymmetry in semileptonic decay of neutral mesons to “wrong sign” leptons (Fig. 6 summarizes the data),

$$A_{\text{SL}}(t) = \frac{\Gamma(\bar{B}^0(t) \rightarrow \ell^+ X) - \Gamma(B^0(t) \rightarrow \ell^- X)}{\Gamma(\bar{B}^0(t) \rightarrow \ell^+ X) + \Gamma(B^0(t) \rightarrow \ell^- X)} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \simeq \text{Im} \frac{\Gamma_{12}}{M_{12}}. \quad (37)$$

To obtain the right-hand side, use Eqs. (31) and (32) for the time evolution, and Eq. (35) for  $|q/p|$ . In kaon decays this asymmetry is measured [21], in agreement with the SM prediction,  $4 \text{Re} \epsilon_K$ . In  $B_d$  and  $B_s$  decays the asymmetry is expected to be [22]

$$A_{\text{SL}}^d \approx -4 \times 10^{-4}, \quad A_{\text{SL}}^s \approx 2 \times 10^{-5}. \quad (38)$$

The calculation of  $\text{Im}(\Gamma_{12}/M_{12})$  requires calculating inclusive nonleptonic decay rates, which can be addressed using an operator product expansion in the  $m_b \gg \Lambda_{\text{QCD}}$  limit. Such a calculation has sizable hadronic uncertainties, the details of which would lead to a long discussion. The constraints on new physics are significant nevertheless [23], as the  $m_c^2/m_b^2$  suppression of  $A_{\text{SL}}$  in the SM can be avoided in the presence of new physics.

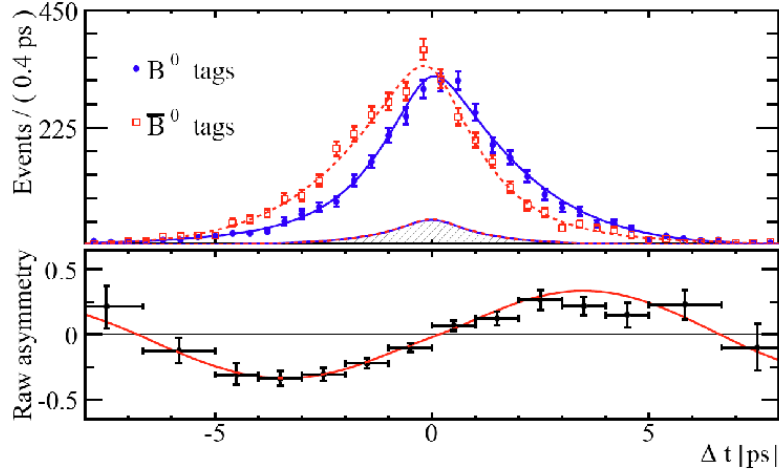
#### 2.4 $CP$ violation in the interference of decay with and without mixing

A third type of  $CP$  violation is possible when both  $B^0$  and  $\bar{B}^0$  can decay to a final state,  $f$ . In the simplest cases, when  $f$  is a  $CP$  eigenstate, define

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f}. \quad (39)$$

If there is no direct  $CP$  violation in a given mode, then  $\bar{A}_f = \eta_f \bar{A}_{\bar{f}}$ , where  $\eta_f = \pm 1$  is the  $CP$  eigenvalue of  $f$  [ $+1$  ( $-1$ ) for  $CP$ -even (-odd) states]. This is useful, because  $A_f$  and  $\bar{A}_{\bar{f}}$  are related by





**Fig. 7:** Time dependence of tagged  $B \rightarrow \psi K$  decays (top);  $CP$  asymmetry (below) [24].

a  $CP$  transformation. If  $CP$  were conserved, then not only  $|q/p| = 1$  and  $|\bar{A}_f/A_f| = 1$ , but the relative phase between  $q/p$  and  $\bar{A}_f/A_f$  also vanishes, hence  $\lambda_f = \pm 1$ .

The experimentally measurable  $CP$  violating observable is<sup>9</sup>

$$\begin{aligned}
 a_f &= \frac{\Gamma[\bar{B}^0(t) \rightarrow f] - \Gamma[B^0(t) \rightarrow f]}{\Gamma[\bar{B}^0(t) \rightarrow f] + \Gamma[B^0(t) \rightarrow f]} \\
 &= \frac{(1 - |\lambda_f|^2) \cos(\Delta m t) - 2 \operatorname{Im} \lambda_f \sin(\Delta m t)}{1 + |\lambda_f|^2} \\
 &\equiv S_f \sin(\Delta m t) - C_f \cos(\Delta m t),
 \end{aligned} \tag{40}$$

where we have neglected  $\Delta\Gamma$  (it is important in the  $B_s$  system). The last line defines the  $S$  and  $C$  coefficients, which are fit to the experimental data (see Fig. 7). If  $\operatorname{Im} \lambda_f \neq 0$ , then  $CP$  violation arises in the interference between the decay  $B^0 \rightarrow f$ , and mixing followed by decay,  $B^0 \rightarrow \bar{B}^0 \rightarrow f$ .

This asymmetry can be nonzero if any type of  $CP$  violation occurs. In particular, in both the  $B_d$  and  $B_s$  systems  $||q/p| - 1| < \mathcal{O}(10^{-2})$  model independently, and it is much smaller in the SM [see, Eq. (38)]. If, in addition, amplitudes with a single weak phase dominate a decay, then  $|\bar{A}_f/A_f| \simeq 1$ , and  $\arg(\bar{A}_f/A_f)$  is just (twice) the weak phase, determined by short-distance physics. It is then possible that  $\operatorname{Im} \lambda_f \neq 0$ ,  $|\lambda_f| \simeq 1$ , and although we cannot compute the decay amplitude, we can extract the weak phase difference between  $B^0 \rightarrow f$  and  $B^0 \rightarrow \bar{B}^0 \rightarrow f$  in a theoretically clean way from the measurement of

$$a_f = \operatorname{Im} \lambda_f \sin(\Delta m t). \tag{41}$$

There is an interesting subtlety. Consider two final states,  $f_{1,2}$ . It is possible that direct  $CP$  violation in each channel,  $|\lambda_{f_1}| - 1$  and  $|\lambda_{f_2}| - 1$ , is unmeasurably small, but direct  $CP$  violation is detectable nevertheless. If

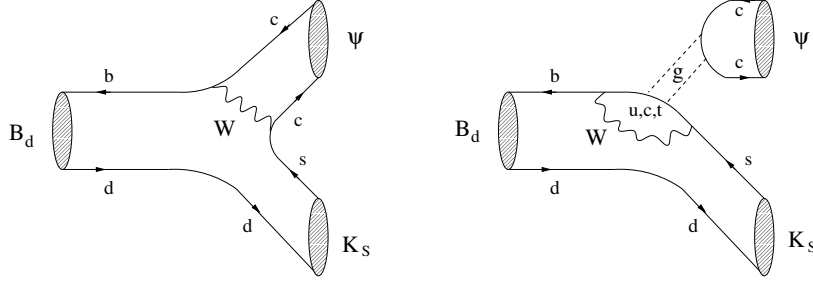
$$\eta_{f_1} \operatorname{Im}(\lambda_{f_1}) \neq \eta_{f_2} \operatorname{Im}(\lambda_{f_2}), \tag{42}$$

then  $CP$  violation must occur outside the mixing amplitude, even though it may be invisible in the data on any one final state.

## 2.5 $\sin 2\beta$ from $B \rightarrow \psi K_{S,L}$

This is one of the cleanest examples of  $CP$  violation in the interference between decay with and without mixing, and one of the theoretically cleanest measurements of a CKM parameter.

<sup>9</sup>Derive the  $CP$  asymmetry in Eq. (40) using Eq. (31). For extra credit, keep  $\Delta\Gamma \neq 0$ .



**Fig. 8:** “Tree” (left) and “penguin” (right) contributions to  $B \rightarrow \psi K_S$  (from Ref. [25]).

There are “tree” and “penguin” contributions to  $B \rightarrow \psi K_{S,L}$ , with different weak and strong phases (see Fig. 8). The tree contribution is dominated by the  $b \rightarrow c\bar{c}s$  transition, while there are penguin contributions with three different combinations of CKM elements,

$$\bar{A}_T = V_{cb}V_{cs}^* T_{c\bar{c}s}, \quad \bar{A}_P = V_{tb}V_{ts}^* P_t + V_{cb}V_{cs}^* P_c + V_{ub}V_{us}^* P_u. \quad (43)$$

( $P_u$  can be defined to absorb the  $V_{ub}V_{us}^* T_{u\bar{u}s}$  “tree” contribution.) We can rewrite the decay amplitude using  $V_{tb}V_{ts}^* + V_{cb}V_{cs}^* + V_{ub}V_{us}^* = 0$  to obtain

$$\begin{aligned} \bar{A} &= V_{cb}V_{cs}^* (T_{c\bar{c}s} + P_c - P_t) + V_{ub}V_{us}^* (P_u - P_t) \\ &\equiv V_{cb}V_{cs}^* T + V_{ub}V_{us}^* P, \end{aligned} \quad (44)$$

where the second line defines  $T$  and  $P$ . Since  $|(V_{ub}V_{us}^*)/(V_{cb}V_{cs}^*)| \approx 0.02$ , the  $T$  amplitude with  $V_{cb}V_{cs}^*$  weak phase dominates. Thus,

$$\lambda_{\psi K_{S,L}} = \mp \left( \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \right) \left( \frac{V_{cb} V_{cs}^*}{V_{cb}^* V_{cs}} \right) \left( \frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}} \right) = \mp e^{-2i\beta}, \quad (45)$$

and so  $\text{Im}\lambda_{\psi K_{S,L}} = \pm \sin 2\beta$ . The first term is the SM value of  $q/p$  in  $B_d$  mixing, the second is  $\bar{A}/A$ , the last one is  $p/q$  in the  $K^0$  system, and  $\eta_{\psi K_{S,L}} = \mp 1$ . Note that without  $K^0 - \bar{K}^0$  mixing there would be no interference between  $\bar{B}^0 \rightarrow \psi \bar{K}^0$  and  $B^0 \rightarrow \psi K^0$ . The accuracy of the relation between  $\lambda_{\psi K_{S,L}}$  and  $\sin 2\beta$  depends on model dependent estimates of  $|P/T|$ , which are below unity, so one expects it to be of order

$$\left| \frac{V_{ub}V_{us}^* P}{V_{cb}V_{cs}^* T} \right| \lesssim 10^{-2}. \quad (46)$$

The absence of detectable direct  $CP$  violation does not in itself bound this. To fully utilize future LHCb and Belle II data, better estimates are needed.

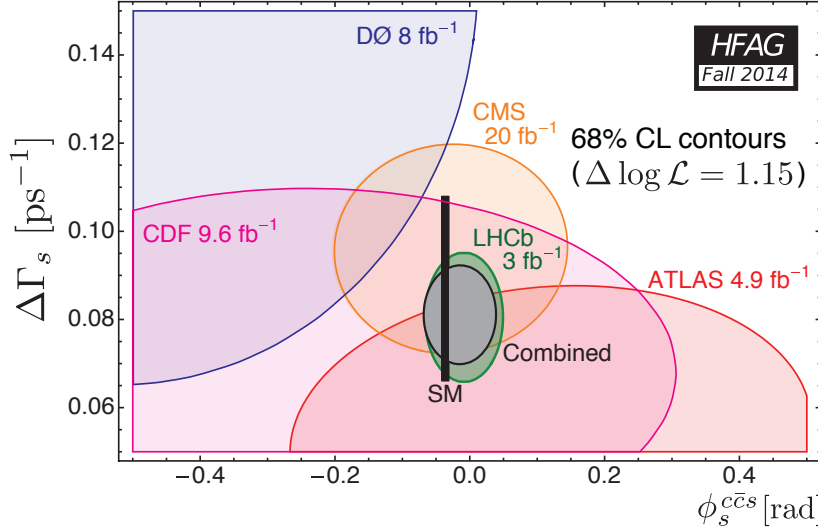
The first evidence for  $CP$  violation outside the kaon sector was the BaBar and Belle measurements of  $S_{\psi K}$ . The current world average is [20]

$$\sin 2\beta = 0.682 \pm 0.019. \quad (47)$$

This is consistent with other constraints, and shows that  $CP$  violation in quark mixing is an  $\mathcal{O}(1)$  effect, which is simply suppressed in  $K$  decays by small flavor violation suppressing the third generation’s contributions.

## 2.6 $\phi_s \equiv -2\beta_s$ from $B_s \rightarrow \psi\phi$

The analogous  $CP$  asymmetry in  $B_s$  decay, sensitive to BSM contributions to  $B_s - \bar{B}_s$  mixing, is  $B_s \rightarrow \psi\phi$ . Since the final state consists of two vector mesons, it is a combination of  $CP$ -even ( $L = 0, 2$ ) and  $CP$ -odd ( $L = 1$ ) partial waves. What is actually measured is the time-dependent  $CP$  asymmetry



**Fig. 9:** Measurements of  $CP$  violation in  $B_s \rightarrow \psi\phi$  and  $\Delta\Gamma_s$  (from Ref. [20]).

for each  $CP$  component of the  $\psi K^+ K^-$  and  $\psi\pi^+\pi^-$  final states. The SM prediction is suppressed compared to  $\beta$  by  $\lambda^2$ , and is rather precise,  $\beta_s = 0.0182_{-0.0006}^{+0.0007}$  [26]. The latest LHCb result using  $3\text{ fb}^{-1}$  data is [27] (Fig. 9 shows all measurements)

$$\phi_s \equiv -2\beta_s = -0.010 \pm 0.039, \quad (48)$$

which has an uncertainty approaching that of  $2\beta$ , suggesting that the “room for new physics” in  $B_s$  mixing is no longer larger than in  $B_d$  (more below).

## 2.7 “Penguin-dominated” measurements of $\beta_{(s)}$

Time dependent  $CP$  violation in  $b \rightarrow s$  dominated decays is a sensitive probe of new physics. Tree-level contributions to  $b \rightarrow s\bar{s}s$  transitions are expected to be small, and the penguin contributions to  $B \rightarrow \phi K_S$  (left diagram in Fig. 10) are

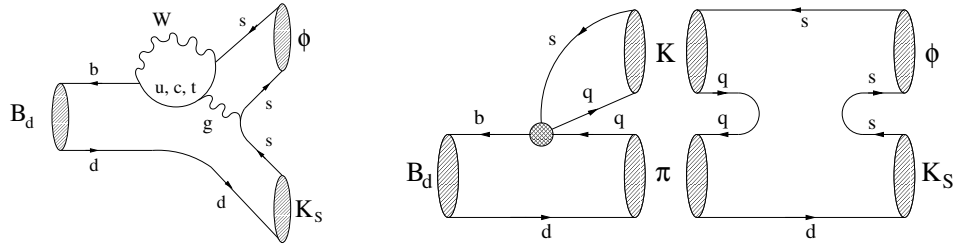
$$\bar{A}_P = V_{cb}V_{cs}^*(P_c - P_t) + V_{ub}V_{us}^*(P_u - P_t). \quad (49)$$

Due to  $|(V_{ub}V_{us}^*)/(V_{cb}V_{cs}^*)| \approx 0.02$  and expecting  $|P_c - P_t|/|P_u - P_t| = \mathcal{O}(1)$ , the  $B \rightarrow \phi K_S$  amplitude is also dominated by a single weak phase,  $V_{cb}V_{cs}^*$ . Therefore, the theory uncertainty relating  $S_{\phi K_S}$  to  $\sin 2\beta$  is small, although larger than in  $B \rightarrow \psi K_S$ . There is also a “tree” contribution from  $b \rightarrow u\bar{u}s$  followed by  $u\bar{u} \rightarrow s\bar{s}$  rescattering (right diagram in Fig. 10). This amplitude is proportional to the suppressed CKM combination,  $V_{ub}V_{us}^*$ , and it is actually not separable from  $P_u - P_t$ . Unless its matrix element is largely enhanced, it should not upset the  $\text{Im}\lambda_{\phi K_S} = \sin 2\beta + \mathcal{O}(\lambda^2)$  expectation in the SM. Similar reasons make many other modes, such as  $B \rightarrow \eta^{(\prime)} K_S$ ,  $B_s \rightarrow \phi\phi$ , etc., interesting and promising to study.

## 2.8 The determinations of $\gamma$ and $\alpha$

By virtue of Eq. (23),  $\gamma$  does not depend on CKM elements involving the top quark, so it can be measured in tree-level  $B$  decays. This is an important distinction from  $\alpha$  and  $\beta$ , and implies that  $\gamma$  is less likely to be affected by BSM physics.

Most measurements of  $\gamma$  utilize the fact that interference of  $B^- \rightarrow D^0 K^-$  ( $b \rightarrow c\bar{u}s$ ) and  $B^- \rightarrow \bar{D}^0 K^-$  ( $b \rightarrow u\bar{c}s$ ) transitions can be studied in final states accessible in both  $D^0$  and  $\bar{D}^0$  decays [28]. (A notable exception is the measurement from the four time-dependent  $\bar{B}_s$  and  $B_s \rightarrow D_s^\pm K^\mp$  rates,



**Fig. 10:** “Penguin” (left) and “tree” (right) contributions to  $B \rightarrow \phi K_S$  (from Ref. [25]).

which is possible at LHCb.) It is possible to measure the  $B$  and  $D$  decay amplitudes, their relative strong phases, and the weak phase  $\gamma$  from the data. There are many variants, based on different  $D$  decay channels [29–34]. The best current measurement comes from  $D^0, \bar{D}^0 \rightarrow K_S \pi^+ \pi^-$  [33, 34], in which case both amplitudes are Cabibbo allowed, and the analysis can be optimized by studying the Dalitz plot dependence of the interference. The world average of all  $\gamma$  measurements is [26]

$$\gamma = (73.2_{-7.0}^{+6.3})^\circ. \quad (50)$$

Most importantly, the theory uncertainty in the SM measurement is smaller than the accuracy of any planned or imaginable future experiment.

The measurements usually referred to as determining  $\alpha$ , measure  $\pi - \beta - \gamma$ , the third angle of the unitarity triangle in any model in which the unitarity of the  $3 \times 3$  CKM matrix is maintained. These measurements are in time-dependent  $CP$  asymmetries in  $B \rightarrow \pi\pi$ ,  $\rho\rho$ , and  $\rho\pi$  decays. In these decays the  $b \rightarrow u\bar{u}d$  “tree” amplitudes are not much larger than the  $b \rightarrow \sum_q q\bar{q}d$  “penguin” contributions, which have different weak phases.<sup>10</sup> The tree contributions change isospin by  $\Delta I = 3/2$  or  $1/2$ , while the penguin contribution is  $\Delta I = 1/2$  only. It is possible to use isospin symmetry of the strong interaction to isolate  $CP$  violation in the  $\Delta I = 3/2$  channel, eliminating the penguin contributions [35–37], yielding [26]

$$\alpha = (87.7_{-3.3}^{+3.5})^\circ. \quad (51)$$

Thus, the measurements of  $\alpha$  are sensitive to new physics in  $B^0 - \bar{B}^0$  mixing and via possible  $\Delta I = 3/2$  (or  $\Delta I = 5/2$ ) contributions [38].

## 2.9 New physics in $B_d$ and $B_s$ mixing

Although the SM CKM fit in Fig. 2 shows impressive and nontrivial consistency, the implications of the level of agreement are often overstated. Allowing new physics contributions, there are a larger number of parameters related to  $CP$  and flavor violation, and the fits become less constraining. This is shown in the left plot in Fig. 11 where the allowed region is indeed significantly larger than in Fig. 2 (the 95% CL combined fit regions are indicated on both plots).

It has been known for decades that the mixing of neutral mesons is particularly sensitive to new physics, and probes some of the highest scales. In a large class of models, NP has a negligible impact on tree-level SM transitions, and the  $3 \times 3$  CKM matrix remains unitary. (In such models  $\alpha + \beta + \gamma = \pi$  is maintained, and independent measurements of  $\pi - \beta - \alpha$  and  $\gamma$  can be averaged.) We can parametrize the NP contributions to neutral meson mixing as

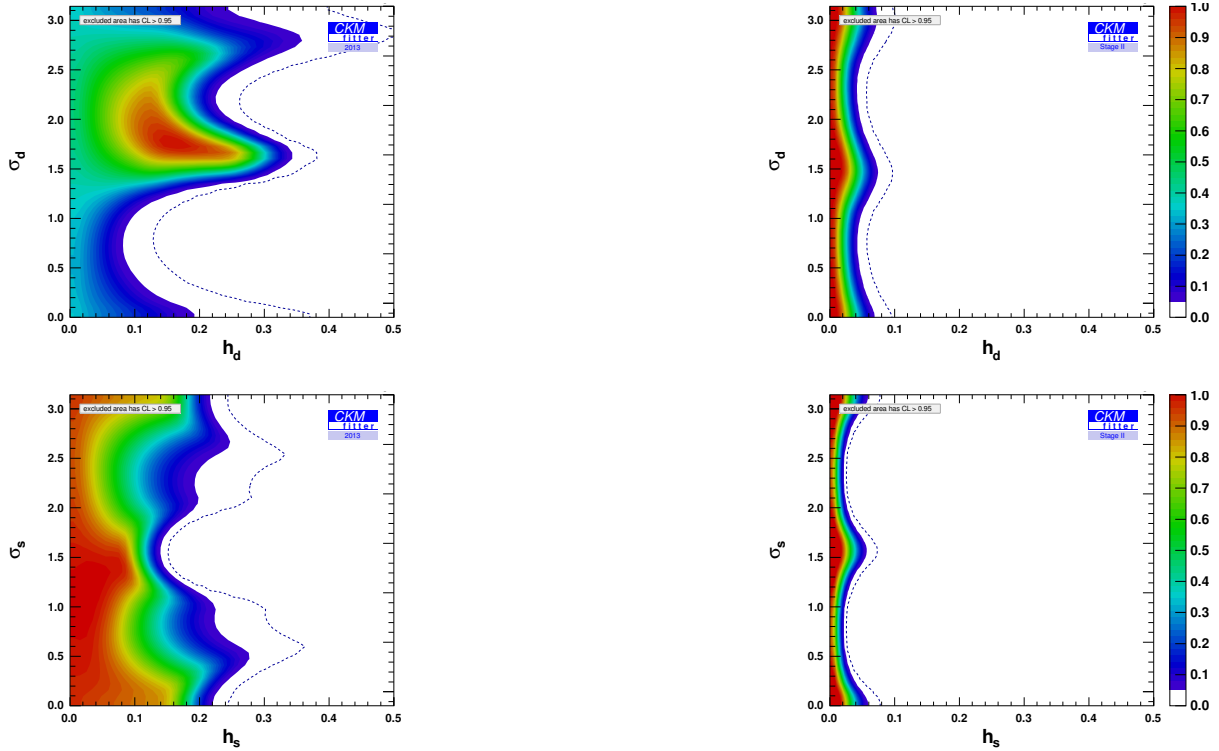
$$M_{12} = M_{12}^{\text{SM}}(1 + h_q e^{2i\sigma_q}), \quad q = d, s. \quad (52)$$

The constraints on  $h_q$  and  $\sigma_q$  in the  $B_d^0$  and  $B_s^0$  systems are shown in the top and bottom rows of Fig. 12, respectively.

<sup>10</sup>Show that if the “tree” amplitudes dominated these decays then  $\lambda_{\pi\pi}^{(\text{tree})} = e^{2i\alpha}$ .



**Fig. 11:** Constraints on  $\bar{\rho} - \bar{\eta}$ , allowing new physics in the  $B_{d,s}$  mixing amplitudes. Left plot shows the current constraints, right plot is the expectation using  $50 \text{ ab}^{-1}$  Belle II and  $50 \text{ fb}^{-1}$  LHCb data. Colored regions show 95% CL, as in Fig. 2. (From Ref. [39].)



**Fig. 12:** Constraints on the  $h_d - \sigma_d$  (top row) and  $h_s - \sigma_s$  parameters (bottom row). Left plots show the current constraints, right plots show those estimated to be achievable using  $50 \text{ ab}^{-1}$  Belle II and  $50 \text{ fb}^{-1}$  LHCb data. Colored regions show  $2\sigma$  limits with the colors indicating CL as shown, while the dashed lines show  $3\sigma$  limits. (From Ref. [39].)

For example, if NP modifies the SM operator describing  $B$  mixing, by

$$\frac{C_q^2}{\Lambda^2} (\bar{b}_L \gamma^\mu q_L)^2, \quad (53)$$

then one finds

$$h_q \simeq \frac{|C_q|^2}{|V_{tb}^* V_{tq}|^2} \left( \frac{4.5 \text{ TeV}}{\Lambda} \right)^2. \quad (54)$$

We can then translate the plots in Fig. 12 to the scale of new physics probed. The summary of expected

**Table 3:** The scale of the operator in Eq. (53) probed by  $B_d^0$  and  $B_s^0$  mixings with  $50 \text{ ab}^{-1}$  Belle II and  $50 \text{ fb}^{-1}$  LHCb data. The differences due to CKM-like hierarchy of couplings and/or loop suppression is indicated. (From Ref. [39].)

Couplings	NP loop order	Scales (TeV) probed by	
		$B_d$ mixing	$B_s$ mixing
$ C_q  =  V_{tb}V_{tq}^* $ (CKM-like)	tree level	17	19
	one loop	1.4	1.5
$ C_q  = 1$ (no hierarchy)	tree level	$2 \times 10^3$	$5 \times 10^2$
	one loop	$2 \times 10^2$	40

sensitivities are shown in Table 3. The sensitivities, even with SM-like loop- and CKM-suppressed coefficients, are comparable to the scales probed by the LHC.

### 3 Some Implications of the Heavy Quark Limit

We have not directly discussed so far that most quark flavor physics processes (other than top quark decays) involve strong interactions in a regime where perturbation theory is not (or not necessarily) reliable. The running of the QCD coupling at lowest order is

$$\alpha_s(\mu) = \frac{\alpha_s(\Lambda)}{1 + \frac{\alpha_s}{2\pi} \beta_0 \ln \frac{\mu}{\Lambda}}, \quad (55)$$

where  $\beta_0 = 11 - 2n_f/3$  and  $n_f$  is the number of light quark flavors. Even in  $B$  decays, the typical energy scale of certain processes can be a fraction of  $m_b$ , possibly around or below a GeV. The ways I know how to deal with this in a tractable way are (i) symmetries of QCD, exact, or approximate in some limits (CP invariance, heavy quark symmetry, chiral symmetry); (ii) the operator product expansion (for inclusive decays); (iii) lattice QCD (for certain hadronic matrix elements). An example of (i) is the determination of  $\sin 2\beta$  from  $B \rightarrow \psi K_S$ , see Eq. (46). So is the determination of  $|V_{cb}|$  from  $B \rightarrow D^* \ell \bar{\nu}$ , see Eq. (73) below. An example of (ii) is the analysis of inclusive  $B \rightarrow X_s \gamma$  decay rates discussed below, which provides some of the strongest constraints on many TeV-scale BSM scenarios.

The role of (strong interaction) model-independent measurements cannot be overstated. To establish that a discrepancy between experiment and theory is a sign of new physics, model-independent predictions are crucial. Results that rely on modeling nonperturbative strong interaction effects will not disprove the SM. Most model-independent predictions are of the form,

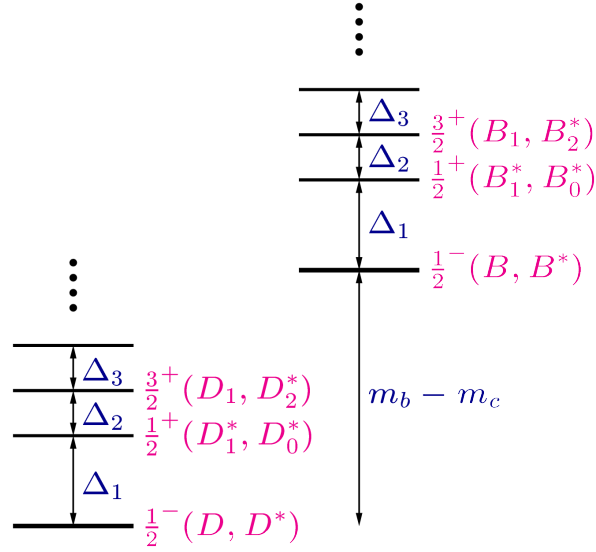
$$\text{Observable} = (\text{calculable terms}) \times \left\{ 1 + \sum_{i,k} [(\text{small parameters})_i]^k \right\}, \quad (56)$$

where the small parameters can be  $\Lambda_{\text{QCD}}/m_b$ ,  $m_s/\Lambda_{\text{QCD}}$ ,  $\alpha_s(m_b)$ , etc. For the purpose of these lectures, strong-interaction model-independent means that the theoretical uncertainty is suppressed by small parameters, so that theorists argue about  $\mathcal{O}(1) \times (\text{small numbers})$  instead of  $\mathcal{O}(1)$  effects. There are always theoretical uncertainties suppressed by some (small parameter)<sup>n</sup>, which cannot be calculated from first principles. If the goal is to test the SM, one must assign  $\mathcal{O}(1)$  uncertainties in such terms.

In addition, besides formal suppressions of certain corrections in some limits, experimental guidance is always needed to establish how well an expansion works; for example,  $f_\pi$ ,  $m_\rho$ , and  $m_K^2/m_s$  are all of order  $\Lambda_{\text{QCD}}$ , but their numerical values span an order of magnitude.

#### 3.1 Heavy quark symmetry (HQS)

In hadrons composed of heavy quarks the dynamics of QCD simplifies. Mesons containing a heavy quark – heavy antiquark pair,  $Q\bar{Q}$ , form positronium-type bound states, which become perturbative in



**Fig. 13:** Spectroscopy of  $B$  and  $D$  mesons. For each doublet level, the spin-parity of the light degrees of freedom,  $s_l^{\pi_l}$ , and the names of the physical states are indicated.

the limit  $m_Q \gg \Lambda_{\text{QCD}}$  [40]. In mesons composed of a heavy quark,  $Q$ , and a light antiquark,  $\bar{q}$  (and gluons and  $q\bar{q}$  pairs), the heavy quark acts as a static color source with fixed four-velocity,  $v^\mu$ , and the wave function of the light degrees of freedom (the “brown muck”) become insensitive to the spin and mass (flavor) of the heavy quark, resulting in heavy quark spin-flavor symmetries [41].

The physical picture is similar to atomic physics, where simplifications occur due to the fact that the electron mass,  $m_e$ , is much smaller than the nucleon mass,  $m_N$ . The analog of flavor symmetry is that isotopes have similar chemistry, because the electrons’ wave functions become independent of  $m_N$  in the  $m_N \gg m_e$  limit. The analog of spin symmetry is that hyperfine levels are almost degenerate, because the interaction of the electron and nucleon spin diminishes in the  $m_N \gg m_e$  limit.

### 3.2 Spectroscopy of heavy-light mesons

The spectroscopy of heavy hadrons simplifies due to heavy quark symmetry. We can write the angular momentum of a heavy-light meson as  $J = \vec{s}_Q + \vec{s}_l$ , where  $\vec{s}_l$  is the total angular momentum of the light degrees of freedom. Angular momentum conservation,  $[\vec{J}, \mathcal{H}] = 0$ , and heavy quark symmetry,  $[\vec{s}_Q, \mathcal{H}] = 0$ , imply  $[\vec{s}_l, \mathcal{H}] = 0$ . In the  $m_Q \gg \Lambda_{\text{QCD}}$  limit, the spin of the heavy quark and the total angular momentum of light degrees of freedom are separately conserved, modified only by subleading interactions suppressed by  $\Lambda_{\text{QCD}}/m_Q$ .

Thus, hadrons containing a single heavy quark can be labeled with  $s_l$ , and for any value of  $s_l$  there are two (almost) degenerate states with total angular momentum  $J_\pm = s_l \pm \frac{1}{2}$ . (An exception occurs for the lightest baryons containing a heavy quark, when  $s_l = 0$ , and there is a single state with  $J = \frac{1}{2}$ , the  $\Lambda_b$  and  $\Lambda_c$ .) The ground state mesons with  $Q\bar{q}$  flavor quantum numbers contain light degrees of freedom with spin-parity  $s_l^{\pi_l} = \frac{1}{2}^-$ , giving a doublet containing a spin zero and spin one meson. For  $Q = c$  these are the  $D$  and  $D^*$ , while  $Q = b$  gives the  $B$  and  $B^*$  mesons.

The mass splittings between the doublets,  $\Delta_i$ , are of order  $\Lambda_{\text{QCD}}$ , and are the same in the  $B$  and  $D$  sectors at leading order in  $\Lambda_{\text{QCD}}/m_Q$ , as illustrated in Fig. 13. The mass splittings within each doublet are of order  $\Lambda_{\text{QCD}}^2/m_Q$ . This is supported by experimental data; e.g., for the  $s_l^{\pi_l} = \frac{1}{2}^-$  ground state doublets  $m_{D^*} - m_D \approx 140 \text{ MeV}$  while  $m_{B^*} - m_B \approx 45 \text{ MeV}$ , and their ratio, 0.3, is consistent with  $m_c/m_b$ .

Let us mention a puzzle. The mass splitting of the lightest vector and pseudoscalar mesons being

$\mathcal{O}(\Lambda_{\text{QCD}}^2/m_Q)$  implies that  $m_V^2 - m_P^2$  is approximately constant. This argument relies on  $m_Q \gg \Lambda_{\text{QCD}}$ . The data are

$$\begin{aligned} m_{B^*}^2 - m_B^2 &= 0.49 \text{ GeV}^2, & m_{B_s^*}^2 - m_{B_s}^2 &= 0.50 \text{ GeV}^2, \\ m_{D^*}^2 - m_D^2 &= 0.54 \text{ GeV}^2, & m_{D_s^*}^2 - m_{D_s}^2 &= 0.58 \text{ GeV}^2, \\ m_\rho^2 - m_\pi^2 &= 0.57 \text{ GeV}^2, & m_{K^*}^2 - m_K^2 &= 0.55 \text{ GeV}^2. \end{aligned} \quad (57)$$

It is not understood why the light meson mass splittings in the last line are so close numerically. (It is expected in the nonrelativistic constituent quark model, which fails to account for several properties of these mesons.) There must be something more going on than heavy quark symmetry, and if this were its only prediction, we could not say that there is strong evidence that it is useful. So in general, to understand a theory, it is not only important how well it works, but also how it breaks down outside its range of validity.

### 3.3 Heavy quark effective theory (HQET)

The consequences of heavy quark symmetry and the corrections to the symmetry limit can be studied by constructing an effective theory which makes the consequences of heavy quark symmetry explicit. The heavy quark in a heavy-light meson is almost on-shell, so we can expand its momentum as  $p_Q^\mu = m_Q v^\mu + k^\mu$ , where  $|k| = \mathcal{O}(\Lambda_{\text{QCD}})$  and  $v^2 = 1$ . Expanding the heavy quark propagator,

$$\frac{i}{\not{p} - m_Q} = \frac{i(\not{p} + m_Q)}{p^2 - m_Q^2} = \frac{i(m_Q \not{v} + \not{k} + m_Q)}{2m_Q v \cdot k + k^2} = \frac{i}{v \cdot k} \frac{1 + \not{v}}{2} + \dots \quad (58)$$

it becomes independent of the heavy quark mass, a manifestation of heavy quark flavor symmetry. Hence the Feynman rules simplify,

$$\begin{array}{c} \text{-----} \\ \frac{i}{\not{p} - m_Q} \end{array} \quad \longrightarrow \quad \begin{array}{c} \text{====} \\ \frac{i}{v \cdot k} P_+(v), \end{array} \quad (59)$$

where  $P_\pm = (1 \pm \not{v})/2$  are projection operators, and the double line denotes the heavy quark propagator. In the rest frame of the heavy quark,  $P_+ = (1 + \gamma^0)/2$  projects onto the heavy quark (rather than anti-quark) components. The coupling of a heavy quark to gluons simplifies due to

$$P_+ \gamma^\mu P_+ = P_+ v^\mu P_+ = v^\mu P_+, \quad (60)$$

hence we can replace

$$\begin{array}{c} \text{-----} \\ \text{oooo} \end{array} \quad \longrightarrow \quad \begin{array}{c} \text{====} \\ \text{oooo} \end{array} \quad (61)$$

$$ig\gamma^\mu \frac{\lambda^a}{2} \quad \longrightarrow \quad igv^\mu \frac{\lambda^a}{2}.$$

The lack of any  $\gamma$  matrix is a manifestation of heavy quark spin symmetry.

To derive the effective Lagrangian of HQET, it is convenient to decompose the four-component Dirac spinor as

$$Q(x) = e^{-im_Q v \cdot x} [Q_v(x) + \mathcal{Q}_v(x)], \quad (62)$$

where

$$Q_v(x) = e^{im_Q v \cdot x} P_+(v) Q(x), \quad \mathcal{Q}_v(x) = e^{im_Q v \cdot x} P_-(v) Q(x). \quad (63)$$

The  $e^{im_Q v \cdot x}$  factor subtracts  $m_Q v$  from the heavy quark momentum. At leading order only  $Q_v$  contributes, and the effects of  $\mathcal{Q}_v$  are suppressed by powers of  $\Lambda_{\text{QCD}}/m_Q$ . The heavy quark velocity,  $v$ , acts



as a label of the heavy quark fields [42], because  $v$  cannot be changed by soft interactions. In terms of these fields the QCD Lagrangian simplifies,

$$\mathcal{L} = \bar{Q}(i\not{D} - m_Q)Q = \bar{Q}_v i\not{D} Q_v + \dots = \bar{Q}_v (iv \cdot D) Q_v + \dots, \quad (64)$$

where the ellipses denote terms suppressed by powers of  $\Lambda_{\text{QCD}}/m_Q$ . The absence of any Dirac matrix is a consequence of heavy quark symmetry, which implies that the heavy quark's propagator and its coupling to gluons are independent of the heavy quark spin. This effective theory provides a framework to calculate perturbative  $\mathcal{O}(\alpha_s)$  corrections and to parametrize nonperturbative  $\mathcal{O}(\Lambda_{\text{QCD}}/m_Q)$  terms.

### 3.4 Semileptonic $B \rightarrow D^{(*)} \ell \bar{\nu}$ decays and $|V_{cb}|$

Heavy quark symmetry is particularly predictive for these decays. In the  $m_{b,c} \gg \Lambda_{\text{QCD}}$  limit, the configuration of the brown muck only depends on the four-velocity of the heavy quark, but not on its mass and spin. So when the weak current changes suddenly (on a time scale  $\ll \Lambda_{\text{QCD}}^{-1}$ ) the flavor  $b \rightarrow c$ , the momentum  $\vec{p}_b \rightarrow \vec{p}_c$ , and possibly flips the spin,  $\vec{s}_b \rightarrow \vec{s}_c$ , the brown muck only feels that the four-velocity of the static color source changed,  $v_b \rightarrow v_c$ . Therefore, the matrix elements that describe the transition probabilities from the initial to the final state are independent of the Dirac structure of weak current, and can only depend on a scalar quantity,  $w \equiv v_b \cdot v_c$ .

The ground-state pseudoscalar and vector mesons for each heavy quark flavor (the spin symmetry doublets  $D^{(*)}$  and  $B^{(*)}$ ) can be represented by a ‘‘superfield’’, combining fields with different spins, that has the right transformation property under heavy quark and Lorentz symmetry,

$$\mathcal{M}_v^{(Q)} = \frac{1 + \not{v}}{2} \left[ \gamma^\mu M_\mu^{*(Q)}(v, \varepsilon) - i\gamma_5 M^{(Q)}(v) \right]. \quad (65)$$

The  $B^{(*)} \rightarrow D^{(*)}$  matrix element of any current can be parametrized as

$$\langle M^{(c)}(v') | \bar{c}_{v'} \Gamma b_v | M^{(b)}(v) \rangle = \text{Tr} \left[ F(v, v') \bar{\mathcal{M}}_{v'}^{(c)} \Gamma \mathcal{M}_v^{(b)} \right]. \quad (66)$$

Because of heavy quark symmetry, there cannot be other Dirac matrices between the  $\bar{\mathcal{M}}_{v'}^{(c)}$  and  $\mathcal{M}_v^{(b)}$  fields. The most general form of  $F$  is

$$F(v, v') = f_1(w) + f_2(w)\not{v} + f_3(w)\not{v}' + f_4(w)\not{v}\not{v}'. \quad (67)$$

As stated above,  $w \equiv v \cdot v'$  is the only possible scalar, simply related to  $q^2 = (p_B - p_{D^{(*)}})^2 = m_B^2 + m_{D^{(*)}}^2 - 2m_B m_{D^{(*)}} w$ . Using  $\mathcal{M}_v^{(Q)} = P_+(v) \mathcal{M}_v^{(Q)} P_-(v)$ , we can write

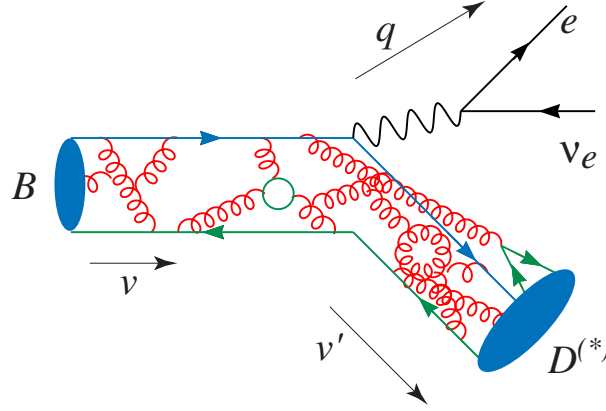
$$\begin{aligned} F &\doteq P_-(v) F P_-(v') = [f_1(w) - f_2(w) - f_3(w) + f_4(w)] P_-(v) P_-(v') \\ &= \xi(w) P_-(v) P_-(v') \doteq \xi(w). \end{aligned} \quad (68)$$

This defines the Isgur-Wise function,  $\xi(w)$ , and  $\doteq$  denotes relations valid when evaluated inside the trace in Eq. (66).

Since only weak interactions change  $b$ -quark number, the matrix element of  $\bar{b}\gamma_0 b$ , the  $b$ -quark number current, is  $\langle B(v) | \bar{b}\gamma_0 b | B(v) \rangle = 2m_B v_0$ . Comparing it with the result obtained using Eq. (66),

$$\langle B(v) | \bar{b}\gamma_\mu b | B(v) \rangle = 2m_B v_\mu \xi(1), \quad (69)$$

implies that  $\xi(1) = 1$ . That is, at  $w = 1$ , the ‘‘zero recoil’’ point, when the  $D^{(*)}$  is at rest in the rest-frame of the decaying  $B$  meson, the configuration of the brown muck does not change at all, and heavy quark symmetry determines the hadronic matrix element (see Fig. 14). Moreover, the six form factors that



**Fig. 14:** Illustration of strong interactions parametrized by the Isgur-Wise function.

describe semileptonic  $B \rightarrow D^{(*)} \ell \bar{\nu}$  decays are related to this universal function, which contains all the low energy nonperturbative hadronic physics relevant for these decays.<sup>11</sup>

The determination of  $|V_{cb}|$  from  $B \rightarrow D^{(*)} \ell \bar{\nu}$  decays use fits to the decay distributions to measure the rates near zero recoil,  $w = 1$ . The rates can be schematically written as

$$\frac{d\Gamma(B \rightarrow D^{(*)} \ell \bar{\nu})}{dw} = (\text{calculable}) |V_{cb}|^2 \begin{cases} (w^2 - 1)^{1/2} \mathcal{F}_*^2(w), & \text{for } B \rightarrow D^*, \\ (w^2 - 1)^{3/2} \mathcal{F}^2(w), & \text{for } B \rightarrow D. \end{cases} \quad (72)$$

Both  $\mathcal{F}(w)$  and  $\mathcal{F}_*(w)$  are equal to the Isgur-Wise function in the  $m_Q \rightarrow \infty$  limit, and  $\mathcal{F}_{(*)}(1) = 1$  is the basis for a model-independent determination of  $|V_{cb}|$ . There are calculable corrections in powers of  $\alpha_s(m_{c,b})$ , as well as terms suppressed by  $\Lambda_{\text{QCD}}/m_{c,b}$ , which can only be parametrized, and that is where hadronic uncertainties enter. Schematically,

$$\begin{aligned} \mathcal{F}_*(1) &= 1_{\text{(Isgur-Wise)}} + c_A(\alpha_s) + \frac{0_{\text{(Luke)}}}{m_{c,b}} + \frac{\text{(lattice or models)}}{m_{c,b}^2} + \dots, \\ \mathcal{F}(1) &= 1_{\text{(Isgur-Wise)}} + c_V(\alpha_s) + \frac{\text{(lattice or models)}}{m_{c,b}} + \dots \end{aligned} \quad (73)$$

The absence of the  $\mathcal{O}(\Lambda_{\text{QCD}}/m_{c,b})$  term for  $B \rightarrow D^* \ell \bar{\nu}$  at zero recoil is a consequence of Luke's theorem [43]. Calculating corrections to the heavy quark limit in these decays is a vast subject. Heavy quark symmetry also makes model-independent predictions for  $B$  decays to excited  $D$  mesons [44]. It is due to heavy quark symmetry that the SM predictions for the recently observed anomalies in the  $B \rightarrow D^{(*)} \tau \bar{\nu}$  branching ratios [45] are under good theoretical control.

<sup>11</sup>Using only Lorentz invariance, six form factors parametrize  $B \rightarrow D^{(*)} \ell \bar{\nu}$  decay,

$$\begin{aligned} \langle D(v') | V_\nu | B(v) \rangle &= \sqrt{m_B m_D} [h_+(v + v')_\nu + h_-(v - v')_\nu], \\ \langle D^*(v') | V_\nu | B(v) \rangle &= i\sqrt{m_B m_{D^*}} h_V \epsilon_{\nu\alpha\beta\gamma} \epsilon^{*\alpha} v'^\beta v^\gamma, \\ \langle D(v') | A_\nu | B(v) \rangle &= 0, \\ \langle D^*(v') | A_\nu | B(v) \rangle &= \sqrt{m_B m_{D^*}} [h_{A_1}(w + 1) \epsilon_\nu^* - h_{A_2}(\epsilon^* \cdot v) v_\nu - h_{A_3}(\epsilon^* \cdot v) v'_\nu], \end{aligned} \quad (70)$$

where  $V_\nu = \bar{c}\gamma_\nu b$ ,  $A_\nu = \bar{c}\gamma_\nu\gamma_5 b$ , and  $h_i$  are functions of  $w$ . Show that this is indeed the most general form of these matrix elements, and that at leading order in  $\Lambda_{\text{QCD}}/m_Q$ ,

$$h_+(w) = h_V(w) = h_{A_1}(w) = h_{A_3}(w) = \xi(w), \quad h_-(w) = h_{A_2}(w) = 0. \quad (71)$$

### 3.5 Inclusive semileptonic decays and $B \rightarrow X_s \gamma$

Instead of identifying all final-state particles in a decay, sometimes it is useful to sum over final-state hadrons that can be produced by the strong interaction, subject to constraints determined by short-distance physics, e.g., the energy of a photon or a charged lepton. Although hadronization is nonperturbative, it occurs on much longer distance (and time) scales than the underlying weak decay. Typically we are interested in a quark-level transition, such as  $b \rightarrow c \ell \bar{\nu}$ ,  $b \rightarrow s \gamma$ , etc., and we would like to extract from the data short distance parameters,  $|V_{cb}|$ ,  $C_7(m_b)$ , etc. To do this, we need to relate the quark-level operators to the measurable decay rates.

For example, consider inclusive semileptonic  $b \rightarrow c$  decay mediated by

$$O_{\text{sl}} = -\frac{4G_F}{\sqrt{2}} V_{cb} (J_{bc})^\alpha (J_{\ell\nu})_\alpha, \quad (74)$$

where  $J_{bc}^\alpha = \bar{c} \gamma^\alpha P_L b$  and  $J_{\ell\nu}^\beta = \bar{\ell} \gamma^\beta P_L \nu$ . The decay rate is given by the square of the matrix element, integrated over phase space, and summed over final states,

$$\Gamma(B \rightarrow X_c \ell \bar{\nu}) \sim \sum_{X_c} \int d[\text{PS}] |\langle X_c \ell \bar{\nu} | O_{\text{sl}} | B \rangle|^2. \quad (75)$$

Since leptons have no strong interaction, the squared matrix element and phase space factorize into  $B \rightarrow X_c W^*$  and a perturbatively calculable leptonic part,  $W^* \rightarrow \ell \bar{\nu}$ . The nontrivial part is the hadronic tensor,

$$\begin{aligned} W^{\mu\nu} &= \sum_{X_c} (2\pi)^3 \delta^4(p_B - q - p_X) \langle B | J_{bc}^{\mu\dagger} | X_c \rangle \langle X_c | J_{bc}^\nu | B \rangle \\ &= \frac{1}{\pi} \text{Im} \int d^4x e^{-iq \cdot x} \langle B | T \{ J_{bc}^{\mu\dagger}(x) J_{bc}^\nu(0) \} | B \rangle, \end{aligned} \quad (76)$$

where the second line is obtained using the optical theorem, and  $T$  denotes here the time-ordered product of the operators. It is this time-ordered product that can be expanded in an operator product expansion (OPE) [46–49]. In the  $m_b \gg \Lambda_{\text{QCD}}$  limit, the time-ordered product is dominated by short distances,  $x \ll \Lambda_{\text{QCD}}^{-1}$ , and one can express the hadronic tensor  $W^{\mu\nu}$  as a sum of matrix elements of local operators. Schematically,

$$= \text{[Diagrammatic expansion]} + \frac{0}{m_b} \text{[Diagram]} + \frac{1}{m_b^2} \text{[Diagram]} + \dots \quad (77)$$

This is analogous to the multipole expansion. At leading order in  $\Lambda_{\text{QCD}}/m_b$  the lowest dimension operator is  $\bar{b} \Gamma b$ , where  $\Gamma$  is some (process-dependent) Dirac matrix. Its matrix element is determined by the  $b$  quark content of the initial state using Eqs. (66) and (69); therefore, inclusive  $B$  decay rates in the  $m_b \gg \Lambda_{\text{QCD}}$  limit are equal to the  $b$  quark decay rates. Subleading effects are parametrized by matrix elements of operators with increasing number of derivatives, which are sensitive to the distribution of chromomagnetic and chromoelectric fields. There are no  $\mathcal{O}(\Lambda_{\text{QCD}}/m_b)$  corrections, because the  $B$  meson matrix element of any dimension-4 operator vanishes,  $\langle B(v) | \bar{Q}_v^{(b)} i D_\alpha \Gamma Q_v^{(b)} | B(v) \rangle = 0$ . The leading nonperturbative effects, suppressed by  $\Lambda_{\text{QCD}}^2/m_b^2$ , are parametrized by two HQET matrix elements, denoted by  $\lambda_{1,2}$ . This is the basis of the model-independent determinations of  $m_b$  and  $|V_{cb}|$  from inclusive semileptonic  $B$  decays.

Some important applications, such as  $B \rightarrow X_s \gamma$  [50] or  $B \rightarrow X_u \ell \bar{\nu}$ , are more complicated. Near boundaries of phase space, the energy release to the hadronic final state may not be large. One can think of the OPE as an expansion in the residual momentum of the  $b$  quark,  $k$ , shown in Eq. (77),

$$\frac{1}{(m_b v + k - q)^2 - m_q^2} = \frac{1}{[(m_b v - q)^2 - m_q^2] + [2k \cdot (m_b v - q)] + k^2}. \quad (78)$$

For the expansion in  $k$  to converge, the final state phase space can only be restricted in a way that allows hadronic final states,  $X$ , to contribute with

$$m_X^2 - m_q^2 \gg E_X \Lambda_{\text{QCD}} \gg \Lambda_{\text{QCD}}^2. \quad (79)$$

In  $B \rightarrow X_s \gamma$  when an experimental lower cut is imposed on  $E_\gamma$  to reject backgrounds, the left-most inequality can be violated. The same occurs in  $B \rightarrow X_u \ell \bar{\nu}$  when experimental cuts are used to suppress  $B \rightarrow X_c \ell \bar{\nu}$  backgrounds. If the right-most inequality in Eq. (79) is satisfied, a more complicated OPE in terms of nonlocal operators is still possible [51, 52].

## 4 Top, Higgs, and New Physics Flavor

### 4.1 The scale of new physics

In the absence of direct observation of BSM particles so far, viewing the standard model as a low energy effective theory, the search for new physics amounts to seeking evidence for higher dimension operators invariant under the SM gauge symmetries.

Possible dimension-6 operators include baryon and lepton number violating operators, such as  $\frac{1}{\Lambda^2} Q Q Q L$ . Limits on the proton lifetime imply  $\Lambda \gtrsim 10^{16}$  GeV. Non-SM flavor and  $CP$  violation could arise from  $\frac{1}{\Lambda^2} Q Q \bar{Q} \bar{Q}$ . The bounds on the scale of such operators are  $\Lambda \gtrsim 10^{4\dots 7}$  GeV, depending on the generation index of the quark fields. Precision electroweak measurements constrain operators of the form  $\frac{1}{\Lambda^2} (\phi D_\mu \phi)^2$  to have  $\Lambda \gtrsim 10^{3\dots 4}$  GeV. These constraints are remarkable, because flavor,  $CP$ , and custodial symmetry are broken by the SM itself, so it is unlikely for new physics to have a symmetry reason to avoid introducing additional contributions.

As mentioned earlier, there is a single type of gauge invariant dimension-5 operators made of SM fields, which give rise to neutrino masses, see Eq. (18). The observed neutrino mass square differences hint at scales  $\Lambda > 10^{10}$  GeV for these  $\frac{1}{\Lambda} (L\phi)^2$  type operators (in many models  $\Lambda \sim 10^{15}$  GeV). Such mass terms violate lepton number. It is an experimental question to determine the nature of neutrino masses, which is what makes the search for neutrinoless double beta decay (and determining the neutrino mass hierarchy) so important.

### 4.2 Charged lepton flavor violation (CLFV)

The SM with vanishing neutrino masses would have predicted lepton flavor conservation. We now know that this is not the case, hence there is no reason to impose it on possible new physics scenarios. In particular, if there are TeV-scale new particles that carry lepton number (e.g., sleptons), then they have their own mixing matrices, which could give rise to CLFV signals. While the one-loop SM contributions to processes such as  $\mu \rightarrow e \gamma$  are suppressed by the neutrino mass-squared differences<sup>12</sup>, the NP contributions have a-priori no such suppressions, other than the somewhat heavier scales and being generated at one-loop in most BSM scenarios.

Within the next decade, the CLFV sensitivity will improve by about 4 orders of magnitude, corresponding to an increase in the new physics scale probed by an order of magnitude, possibly the largest such gain in sensitivity achievable soon. If any CLFV signal is discovered, we would want to measure many processes to map out the underlying patterns, including  $\mu \rightarrow e \gamma$ ,  $\mu \rightarrow 3e$ ,  $\tau \rightarrow e \gamma$ ,  $\tau \rightarrow 3e$ ,  $\tau \rightarrow \mu \gamma$ ,  $\tau \rightarrow 3\mu$ , etc.

<sup>12</sup>Estimate the  $\mu \rightarrow e \gamma$  rate in the SM.

### 4.3 Electric dipole moments (EDM)

The experimental bound on the neutron EDM implies that a possible dimension-4 term in the SM Lagrangian,  $\theta_{\text{QCD}} F\tilde{F}/(16\pi^2)$ , has a coefficient  $\theta_{\text{QCD}} \lesssim 10^{-10}$ . While there are plausible explanations [11], we do not yet know the resolution with certainty. Neglecting this term,  $CP$  violation in the CKM matrix only gives rise to quark EDMs at three-loop order, and lepton EDMs at four-loop level, resulting in EDMs below near future experimental sensitivities. On the other hand, new physics (e.g., supersymmetry) could generate both quark and lepton EDMs at the one-loop level, so even if the scale of new physics is 10–100 TeV, observable effects could arise.

### 4.4 Top quark flavor physics

Well before the LHC turned on, it was already certain that it was going to be a top quark factory; the HL-LHC is expected to produce a few times  $10^9$   $t\bar{t}$  pairs. In the SM, top quarks almost exclusively decay to  $Wb$ , as  $||V_{tb}| - 1| \approx 10^{-3}$ . The current bounds on FCNC top decays are at the  $10^{-3}$  level, and the ultimate LHC sensitivity is expected to reach the  $10^{-5}$  to  $10^{-6}$  level, depending on the decay mode. The SM rates are much smaller<sup>13</sup>, so observation of any FCNC top decay signal would be clear evidence for new physics.

There is obvious complementarity between FCNC searches in the top sector and low energy flavor physics bounds. Since  $t_L$  is in the same  $SU(2)$  doublet as  $b_L$ , several operators have correlated effects in  $t$  and  $b$  decays. For some operators, mainly those involving left-handed quark fields, the low energy constraints already exclude a detectable LHC signal, whereas other operators may still have large enough coefficients to yield detectable effects in top FCNCs at the LHC (see, e.g., Ref. [53]).

The  $t\bar{t}$  forward-backward asymmetry provided a clear example recently of the interplay between flavor physics and anomalies in the high energy collider data (even those that may seem little to do with flavor at first). The CDF measurement in 2011,  $A_{t\bar{t}}^{\text{FB}}(m_{t\bar{t}} > 450 \text{ GeV}) = 0.475 \pm 0.114$  [54], was stated to be  $3.4\sigma$  above the NLO SM prediction. At the LHC, the same underlying physics would produce a rapidity asymmetry.<sup>14</sup> It became quickly apparent that models that could account for this signal faced severe flavor constraints. This provides an example (with hundreds of papers in the literature) that flavor physics will likely be crucial to understand what the explanation of a high- $p_T$  LHC anomaly can be, and also what it cannot be. By now this excitement has subsided, because the significance of the Tevatron anomaly decreased and because the LHC has not seen any anomalies in the top production data predicted by most models (see, e.g., Ref. [55]) built to explain the Tevatron signal.

### 4.5 Higgs flavor physics

With the discovery of a SM-like Higgs boson at the LHC, it is now clear that the LHC is also a Higgs factory. Understanding the properties of this particle entails both the precision measurements of its observed (and not yet seen) couplings predicted by the SM, and the search for possible decays forbidden in the SM.

The source of Higgs flavor physics, obviously, is the same set of Yukawa couplings whose structure and consequences we also seek to understand in low energy flavor physics measurements. While in terms of SUSY model building  $m_h \approx 125 \text{ GeV}$  is challenging to understand, this mass allows experimentally probing many Higgs production and decay channels. The fact that ultimately the LHC will be able to probe Higgs production via (i) gluon fusion ( $gg \rightarrow h$ ), (ii) vector boson fusion ( $q\bar{q} \rightarrow q\bar{q}h$ ), (iii)  $W/Z$  associated production ( $q\bar{q} \rightarrow hZ$  or  $hW$ ), (iv)  $b/t$  associated production ( $gg \rightarrow hb\bar{b}$  or  $ht\bar{t}$ ) sensitively depend on the Yukawa couplings and  $m_h$ .<sup>15</sup>

<sup>13</sup>Estimate the  $t \rightarrow cZ$  and  $t \rightarrow c\gamma$  branching ratios in the SM.

<sup>14</sup>Show that if in  $t\bar{t}$  production at the Tevatron more  $t$  goes in the  $p$  than in the  $\bar{p}$  direction, then at the LHC the mean magnitude of the  $t$  quark rapidity is greater than that of the  $\bar{t}$ .

<sup>15</sup>How would Higgs production and decay change if  $m_t$  were, say, 50 GeV?

If we allow new physics to contribute to Higgs-related processes, which is especially well motivated for loop-induced production (e.g., the dominant  $gg \rightarrow h$ ) and decay (e.g.,  $h \rightarrow \gamma\gamma$ ) channels, then the first evidence for non-universal Higgs couplings to fermions was the bound on  $h \rightarrow \mu^+\mu^-$  below  $10 \times$  (SM prediction), combined with the observations of  $h \rightarrow \tau^+\tau^-$  at the SM level, implicitly bounding  $\mathcal{B}(h \rightarrow \mu^+\mu^-)/\mathcal{B}(h \rightarrow \tau^+\tau^-) \lesssim 0.03$ .

There is an obvious interplay between the search for flavor non-diagonal Higgs decays and indirect bounds from flavor-changing quark transitions and bounds on CLFV in the lepton sector. For example,  $y_{e\mu} \neq 0$  would generate a one-loop contribution to  $\mu \rightarrow e\gamma$ ,  $y_{uc} \neq 0$  would generate  $D^0 - \bar{D}^0$  mixing, etc. [56]. In some cases the flavor physics constraints imply that there is no chance to detect a particular flavor-violating Higgs decay, while signals in some modes may be above future direct search sensitivities. The interplay between measurements and constraints on flavor-diagonal and flavor-changing Higgs decay modes can provide additional insight on which flavor models are viable (see, e.g., Ref. [57]).

#### 4.6 Supersymmetry and flavor

While I hope the LHC will discover something unexpected, of the known BSM scenarios, supersymmetry is particularly interesting, and its signals have been worked out in great detail. The minimal supersymmetric standard model (MSSM) contains 44  $CP$  violating phases and 80 other  $CP$  conserving flavor parameters [58].<sup>16</sup> It has long been known that flavor physics (neutral meson mixings,  $\epsilon'_K$ ,  $\mu \rightarrow e\gamma$ ,  $B \rightarrow X_s\gamma$ , etc.) imposes strong constraints on the SUSY parameter space. The MSSM also contains flavor-diagonal  $CP$  violation (in addition to  $\theta_{\text{QCD}}$ ), and the constraints from the bounds on electric dipole moments are fairly strong on these phases if the mass scale is near 1 TeV.

As an example, consider the  $K_L - K_S$  mass difference. The squark–gluino box contribution compared to the data contains terms, roughly,

$$\frac{\Delta m_K^{(\text{SUSY})}}{\Delta m_K^{(\text{exp})}} \sim 10^4 \left( \frac{1 \text{ TeV}}{\tilde{m}} \right)^2 \left( \frac{\Delta \tilde{m}^2}{\tilde{m}^2} \right)^2 \text{Re}[(K_L^d)_{12}(K_R^d)_{12}], \quad (80)$$

where  $K_L^d$  ( $K_R^d$ ) are the mixing matrices in the gluino couplings to left-handed (right-handed) down quarks and their scalar partners [3]. The constraint from  $\epsilon_K$  corresponds to replacing  $10^4 \text{Re}[(K_L^d)_{12}(K_R^d)_{12}]$  with  $10^6 \text{Im}[(K_L^d)_{12}(K_R^d)_{12}]$ . The simplest supersymmetric frameworks with parameters in the ballpark of  $\tilde{m} = \mathcal{O}(1 \text{ TeV})$ ,  $\Delta \tilde{m}^2/\tilde{m}^2 = \mathcal{O}(0.1)$ , and  $(K_{L,R}^d)_{ij} = \mathcal{O}(1)$  are excluded by orders of magnitude.

There are several ways to address the supersymmetric flavor problems. There are classes of models that suppress each of the terms in Eq. (80): (i) heavy squarks, when  $\tilde{m} \gg 1 \text{ TeV}$  (e.g., split SUSY); (ii) universality, when  $\Delta \tilde{m}_{\tilde{Q},\tilde{D}}^2 \ll \tilde{m}^2$  (e.g., gauge mediation); (iii) alignment, when  $(K_{L,R}^d)_{12} \ll 1$  (e.g., horizontal symmetry). All viable models incorporate some of these ingredients in order not to violate the experimental bounds. Conversely, if SUSY is discovered, mapping out its flavor structure will help to answer important questions about even higher scales, e.g., the mechanism of SUSY breaking, how it is communicated to the MSSM, etc.

A special role in constraining SUSY models is played by  $D^0 - \bar{D}^0$  mixing, which was the first observed FCNC process in the up-quark sector. It is a special probe of BSM physics, because it is the only neutral meson system in which mixing is generated by intermediate down-type quarks in the SM, or intermediate up-type squarks in SUSY. The constraints are thus complementary to FCNC processes involving  $K$  and  $B$  mesons.  $D^0 - \bar{D}^0$  mixing and FCNC in the up-quark sector are particularly important in constraining scenarios utilizing quark-squark alignment [59, 60].

Another important implication for SUSY searches is that the LHC constraints on squark masses are sensitive to the level of (non-)degeneracy of squarks required to satisfy flavor constraints. Most SUSY searches assume that the first two generation squarks,  $\tilde{u}_{L,R}$ ,  $\tilde{d}_{L,R}$ ,  $\tilde{s}_{L,R}$ ,  $\tilde{c}_{L,R}$ , are all degenerate,

<sup>16</sup>Check this, using the counting of couplings and broken global symmetries.

which increases signal cross sections. Relaxing this assumption consistent with flavor bounds [60, 61], results in substantially weaker squark mass limits from Run 1, as low as around the 500 GeV scale [62].

It is apparent from the above discussion that there is a tight interplay between the implications of the non-observation of new physics at the LHC so far, and the non-observation of deviations from the SM in flavor physics. If there is new physics at the TeV scale, which we hope the LHC will discover in its next run, then we know already that its flavor structure must be rather non-generic to suppress FCNCs, and the combination of all data will contain plenty of additional information about the structure of new physics. The higher the scale of new physics, the less severe the flavor constraints are. If NP is beyond the reach of the LHC, flavor physics experiments may still observe robust deviations from the SM predictions, which would point to an upper bound on the next scale to probe.

#### 4.7 Minimal flavor violation (MFV)

The standard model without Yukawa couplings has a global  $[U(3)]^5$  symmetry ( $[U(3)]^3$  in the quark and  $[U(3)]^2$  in the lepton sector), rotating the 3 generations of the 5 fields in Eq. (4). This is broken by the Yukawa interactions in Eq. (7). One may view the Yukawa couplings as spurions, fields which transform under  $[U(3)]^5$  in a way that makes the Lagrangian invariant, and then the global flavor symmetry is broken by the background values of the Yukawas. BSM scenarios in which there are no new sources of flavor violation beyond the Yukawa matrices are called minimal flavor violation [63–65]. Since the SM breaks the  $[U(3)]^5$  flavor symmetry already, MFV gives a framework to characterize “minimal reasonable” deviations from the SM predictions.

Let us focus on the quark sector. Under  $U(3)_Q \times U(3)_u \times U(3)_d$  the transformation properties are

$$Q_L(3, 1, 1), \quad u_R(1, 3, 1), \quad d_R(1, 1, 3), \quad Y_u(3, \bar{3}, 1), \quad Y_d(3, 1, \bar{3}). \quad (81)$$

One can choose a basis in which

$$Y_d = \text{diag}(y_d, y_s, y_b), \quad Y_u = V_{\text{CKM}}^\dagger \text{diag}(y_u, y_c, y_t). \quad (82)$$

To generate a flavor-changing transition, requires constructing  $[U(3)]^3$  singlet terms that connect the required fields. For example, in the down-quark sector, the simplest terms are [65]

$$\bar{Q}_L Y_u Y_u^\dagger Q_L, \quad \bar{d}_R Y_d^\dagger Y_u Y_u^\dagger Q_L, \quad \bar{d}_R Y_d^\dagger Y_u Y_u^\dagger Y_d d_R. \quad (83)$$

A useful feature of this approach is that it allows EFT-like analyses.

Consider  $B \rightarrow X_s \gamma$  as an example. We are interested in the magnitude of a possible NP contribution to the Wilson coefficient of the operator  $\frac{X}{\Lambda} (\bar{s}_L \sigma_{\mu\nu} F^{\mu\nu} b_R)$ . A term  $\bar{Q}_L b_R$  is not invariant under  $[U(3)]^3$ . A term  $\bar{Q}_L Y_d d_R$  is  $[U(3)]^3$  invariant, but it is diagonal, so it only connects same generation fields. The first non-vanishing contribution comes from  $\bar{Q}_L Y_u Y_u^\dagger Y_d d_R$ , which has a  $V_{tb} V_{ts}^* y_t^2 y_b (\bar{s}_L b_R)$  component. We learn that in MFV models, in general,  $X \propto y_b V_{tb} V_{ts}^*$ , as is the case in the SM.

Thus, in MFV models, most flavor-changing operators “automatically” have their SM-like suppressions, proportional to the same CKM elements, quark masses from chirality flips, etc. Therefore, the scale of MFV models can be  $\mathcal{O}(1 \text{ TeV})$  without violating flavor physics bounds, thus solving the new physics flavor puzzle. Originally introduced for technicolor models [63], gauge-mediated supersymmetry breaking provides another well known scenario in which MFV is expected to be a good approximation.

MFV models have important implications for new particle searches, too. Since the only quark flavor-changing parameters are the CKM elements, and the ones that couple the third generation to the lighter ones are very small, in MFV models new particles that decay to a single final quark (and other particles) decay to either a third generation quark or to quarks from the first two generations, but (to a good approximation) not to both [66].

The MFV ansatz can be incorporated into models that do not contain explicitly flavor breaking unrelated to Yukawa couplings. MFV is not expected to be an exact symmetry, but it may be a useful organizing principle to understand details of the new physics we soon hope to get a glimpse of.

## 5 Summary

An essential feature of flavor physics is its ability to probe very high scales, beyond the masses of particles that can be produced on-shell in colliders. Flavor physics can also teach us about properties of TeV-scale new physics, that cannot be learned from the direct production of new particles.

Some of the main points I tried to explain in these lectures were:

- Flavor-changing neutral currents and meson mixing probe scales well above the masses of particles colliders can produce, and provide strong constraints on TeV-scale new physics.
- $CP$  violation is always the result of interference phenomena, without a classical analog.
- The KM phase has been established as the dominant source of  $CP$  violation in flavor-changing processes.
- Tremendous progress will continue: Until  $\sim 10$  years ago, more than  $\mathcal{O}(1)$  deviations from the SM were possible; at present  $\mathcal{O}(20\%)$  corrections to most FCNC processes are still allowed; in the future, sensitivities of a few percent will be reached.
- The future goal is not measuring SM parameters better, but to search for corrections to the SM, and to learn about NP as much as possible.
- Direct information on new particles and their influence on flavor-changing processes will both be crucial to understand the underlying physics.
- The sensitivity of future experiments in a number of important processes is only limited by statistics, not theory.
- The interesting (and fun) interplay between theoretical and experimental developments in flavor physics will continue.

At present, both direct production and flavor physics experiments only give bounds on new physics. The constraints imply that if new physics is accessible at the LHC, it is likely to have flavor suppression factors similar to the SM. In many models (e.g., the MSSM), measurements or bounds on FCNC transitions constrain the product of certain mass splittings times mixing parameters divided by the square of the new physics scale. If the LHC discovers new physics, then in principle the mass splittings and mixing parameters can be measured separately. If flavor physics experiments establish a deviation from the SM in a related process, the combination of LHC and flavor data can be very powerful to discriminate between models. The consistency of measurements could ultimately tell us that we understand the flavor structure of new physics and how the new physics flavor puzzle is solved. The present situation and an (optimistic) future scenario for supersymmetry are shown in Fig. 15. Let's hope that we shall have the privilege to think about such questions, motivated by data, in the coming years.

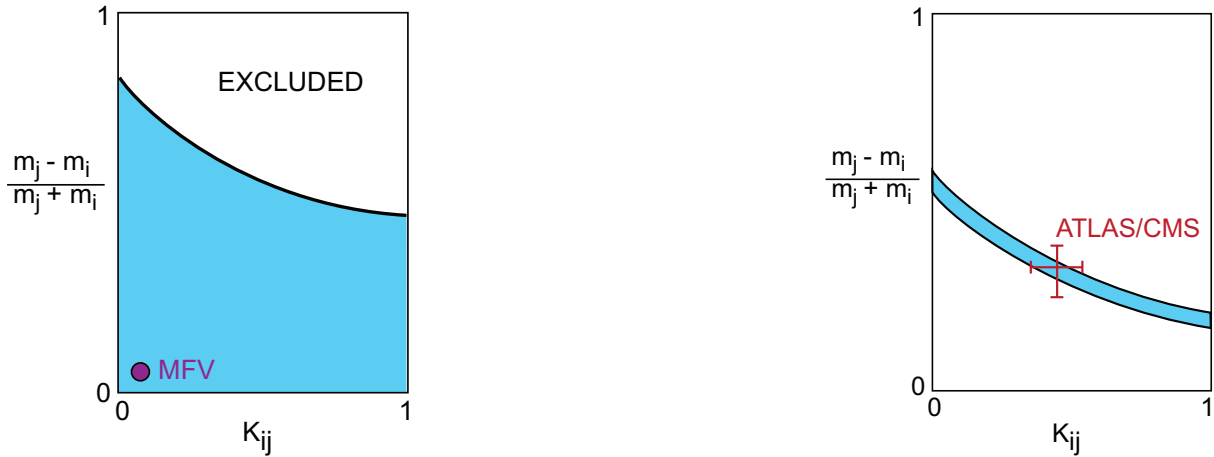
## Acknowledgments

I thank Lance Dixon and Frank Petriello for surprisingly successful arm-twisting (and patience), so that these notes got written up, and Marat Freytsis, Yonit Hochberg, and Dean Robinson for helpful comments. This work was supported by the Office of Science, Office of High Energy Physics, of the U.S. Department of Energy under contract DE-AC02-05CH11231.

## References

- [1] G. C. Branco, L. Lavoura and J. P. Silva, "CP Violation," Int. Ser. Monogr. Phys. **103**, 1 (1999).
- [2] A. V. Manohar and M. B. Wise, "Heavy quark physics," Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **10**, 1 (2000).
- [3] Z. Ligeti and Y. Nir, "New physics and future  $B$  factories," Nucl. Phys. Proc. Suppl. **111**, 82 (2002) [hep-ph/0202117].





**Fig. 15:** Schematic description of the constraints on the mass splitting,  $(m_i - m_j)/(m_i + m_j)$ , and mixing angle,  $K_{ij}$ , between squarks (or sleptons). Left: typical constraint from not observing deviations from the SM. The fact that  $\mathcal{O}(1)$  splittings and mixings are excluded constitutes the new physics flavor puzzle. Right: possible future scenario where ATLAS/CMS measurements fit flavor physics signals of NP. (From Ref. [7].)

- [4] Z. Ligeti, “Introduction to heavy meson decays and CP asymmetries,” eConf C **020805** (2002) L02 [hep-ph/0302031].
- [5] Y. Nir, “CP violation in meson decays,” hep-ph/0510413.
- [6] A. Höcker and Z. Ligeti, “CP violation and the CKM matrix,” Ann. Rev. Nucl. Part. Sci. **56**, 501 (2006) [hep-ph/0605217].
- [7] Y. Grossman, Z. Ligeti and Y. Nir, “Future prospects of B physics,” Prog. Theor. Phys. **122**, 125 (2009) [arXiv:0904.4262].
- [8] B. Grinstein, “TASI-2013 Lectures on Flavor Physics,” arXiv:1501.05283.
- [9] See the reviews on the “Cabibbo-Kobayashi-Maskawa quark mixing matrix” and on “CP violation in the quark sector” in: K. A. Olive *et al.* [Particle Data Group], Chin. Phys. C **38**, 090001 (2014).
- [10] A. D. Sakharov, “Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe,” Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967) [JETP Lett. **5**, 24 (1967)].
- [11] M. Dine, “TASI lectures on the strong CP problem,” hep-ph/0011376.
- [12] S. L. Glashow and S. Weinberg, “Natural Conservation Laws for Neutral Currents,” Phys. Rev. D **15**, 1958 (1977).
- [13] S. L. Glashow, J. Iliopoulos and L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry,” Phys. Rev. D **2**, 1285 (1970).
- [14] A. I. Vainshtein and I. B. Khriplovich, “Restrictions on masses of supercharged hadrons in the Weinberg model,” Pisma Zh. Eksp. Teor. Fiz. **18**, 141 (1973) [JETP Lett. **18**, 83 (1973)].
- [15] M. K. Gaillard and B. W. Lee, “Rare Decay Modes of the K-Mesons in Gauge Theories,” Phys. Rev. D **10**, 897 (1974).
- [16] S. Aoki *et al.*, “Review of lattice results concerning low-energy particle physics,” Eur. Phys. J. C **74**, no. 9, 2890 (2014) [arXiv:1310.8555].
- [17] A. J. Buras, D. Buttazzo, J. Girrbach-Noe and R. Knegjens, “ $K^+ \rightarrow \pi^+ \nu \bar{\nu}$  and  $K_L \rightarrow \pi^0 \nu \bar{\nu}$  in the Standard Model: Status and Perspectives,” arXiv:1503.02693.
- [18] A. V. Artamonov *et al.* [E949 Collaboration], “New measurement of the  $K^+ \rightarrow \pi^+ \nu \bar{\nu}$  branching ratio,” Phys. Rev. Lett. **101**, 191802 (2008) [arXiv:0808.2459].
- [19] G. Isidori, Y. Nir and G. Perez, “Flavor Physics Constraints for Physics Beyond the Standard Model,” Ann. Rev. Nucl. Part. Sci. **60**, 355 (2010) [arXiv:1002.0900]; updated in G. Isidori, “Flavor

- physics and CP violation,” arXiv:1302.0661.
- [20] Heavy Flavor Averaging Group, Y. Amhis *et al.*, “Averages of  $b$ -hadron,  $c$ -hadron, and  $\tau$ -lepton properties as of summer 2014,” arXiv:1412.7515; and updates at <http://www.slac.stanford.edu/xorg/hfag/>.
- [21] A. Angelopoulos *et al.* [CPLEAR Collaboration], “First direct observation of time reversal noninvariance in the neutral kaon system,” Phys. Lett. B **444**, 43 (1998).
- [22] A. Lenz, “ $B$ -mixing in and beyond the Standard model,” arXiv:1409.6963.
- [23] S. Laplace, Z. Ligeti, Y. Nir and G. Perez, “Implications of the  $CP$  asymmetry in semileptonic  $B$  decay,” Phys. Rev. D **65**, 094040 (2002) [hep-ph/0202010].
- [24] B. Aubert *et al.* [BaBar Collaboration], “Measurement of Time-Dependent  $CP$  Asymmetry in  $B^0 \rightarrow c\bar{c}K^{(*)0}$  Decays,” Phys. Rev. D **79**, 072009 (2009) [arXiv:0902.1708]; I. Adachi *et al.* [Belle Collaboration], “Precise measurement of the  $CP$  violation parameter  $\sin 2\phi_1$  in  $B^0 \rightarrow (c\bar{c})K^0$  decays,” Phys. Rev. Lett. **108**, 171802 (2012) [arXiv:1201.4643].
- [25] R. Fleischer, “ $CP$  violation in the  $B$  system and relations to  $K \rightarrow \pi\nu\bar{\nu}$  decays,” Phys. Rept. **370**, 537 (2002) [hep-ph/0207108].
- [26] A. Höcker, H. Lacker, S. Laplace and F. Le Diberder, “A New approach to a global fit of the CKM matrix,” Eur. Phys. J. C **21**, 225 (2001) [hep-ph/0104062]; J. Charles *et al.*, “ $CP$  violation and the CKM matrix: Assessing the impact of the asymmetric  $B$  factories,” Eur. Phys. J. C **41** (2005) 1 [hep-ph/0406184]; and updates at <http://ckmfitter.in2p3.fr/>.
- [27] R. Aaij *et al.* [LHCb Collaboration], “Precision measurement of  $CP$  violation in  $B_s^0 \rightarrow J/\psi K^+ K^-$  decays,” arXiv:1411.3104.
- [28] A. B. Carter and A. I. Sanda, “ $CP$  Violation in Cascade Decays of  $B$  Mesons,” Phys. Rev. Lett. **45**, 952 (1980).
- [29] M. Gronau and D. London, “How to determine all the angles of the unitarity triangle from  $B_d^0 \rightarrow DK_S$  and  $B_s^0 \rightarrow D^0\phi$ ,” Phys. Lett. B **253**, 483 (1991).
- [30] M. Gronau and D. Wyler, “On determining a weak phase from  $CP$  asymmetries in charged  $B$  decays,” Phys. Lett. B **265**, 172 (1991).
- [31] D. Atwood, I. Dunietz and A. Soni, “Enhanced  $CP$  violation with  $B \rightarrow KD^0(\bar{D}^0)$  modes and extraction of the CKM angle  $\gamma$ ,” Phys. Rev. Lett. **78**, 3257 (1997) [hep-ph/9612433].
- [32] Y. Grossman, Z. Ligeti and A. Soffer, “Measuring  $\gamma$  in  $B^\pm \rightarrow K^\pm(KK^*)(D)$  decays,” Phys. Rev. D **67**, 071301 (2003) [hep-ph/0210433].
- [33] A. Bondar, talk at the BELLE analysis workshop, Novosibirsk, Sept. 2002.
- [34] A. Giri, Y. Grossman, A. Soffer and J. Zupan, “Determining  $\gamma$  using  $B^\pm \rightarrow DK^\pm$  with multibody  $D$  decays,” Phys. Rev. D **68**, 054018 (2003) [hep-ph/0303187].
- [35] M. Gronau and D. London, “Isospin analysis of  $CP$  asymmetries in  $B$  decays,” Phys. Rev. Lett. **65**, 3381 (1990).
- [36] H. J. Lipkin, Y. Nir, H. R. Quinn and A. Snyder, “Penguin trapping with isospin analysis and  $CP$  asymmetries in  $B$  decays,” Phys. Rev. D **44**, 1454 (1991).
- [37] A. F. Falk, Z. Ligeti, Y. Nir and H. Quinn, “Comment on extracting  $\alpha$  from  $B \rightarrow \rho\rho$ ,” Phys. Rev. D **69**, 011502 (2004) [hep-ph/0310242].
- [38] S. Baek, F. J. Botella, D. London and J. P. Silva, “Can one detect new physics in  $I = 0$  and/or  $I = 2$  contributions to the decays  $B \rightarrow \pi\pi$ ?,” Phys. Rev. D **72**, 036004 (2005) [hep-ph/0506075]; F. J. Botella, D. London and J. P. Silva, “Looking for  $\Delta I = 5/2$  amplitude components in  $B \rightarrow \pi\pi$  and  $B \rightarrow \rho\rho$  experiments,” Phys. Rev. D **73**, 071501 (2006) [hep-ph/0602060].
- [39] J. Charles, S. Descotes-Genon, Z. Ligeti, S. Monteil, M. Papucci and K. Trabelsi, “Future sensitivity to new physics in  $B_d$ ,  $B_s$  and  $K$  mixings,” Phys. Rev. D **89**, 033016 (2014) [arXiv:1309.2293].
- [40] T. Appelquist and H. D. Politzer, “Orthocharmonium and  $e^+e^-$  Annihilation,” Phys. Rev. Lett. **34**,

- 43 (1975).
- [41] N. Isgur and M. B. Wise, “Weak Decays of Heavy Mesons in the Static Quark Approximation,” *Phys. Lett. B* **232**, 113 (1989); “Weak Transition Form-factors Between Heavy Mesons,” *Phys. Lett. B* **237**, 527 (1990).
- [42] H. Georgi, “An Effective Field Theory for Heavy Quarks at Low-energies,” *Phys. Lett. B* **240**, 447 (1990).
- [43] M. E. Luke, “Effects of subleading operators in the heavy quark effective theory,” *Phys. Lett. B* **252**, 447 (1990).
- [44] A. K. Leibovich, Z. Ligeti, I. W. Stewart and M. B. Wise, “Model independent results for  $B \rightarrow D_1(2420)\ell\bar{\nu}$  and  $B \rightarrow D_2^*(2460)\ell\bar{\nu}$  at order  $\Lambda_{\text{QCD}}/m_{c,b}$ ,” *Phys. Rev. Lett.* **78**, 3995 (1997) [hep-ph/9703213]; “Semileptonic  $B$  decays to excited charmed mesons,” *Phys. Rev. D* **57**, 308 (1998) [hep-ph/9705467].
- [45] J. P. Lees *et al.* [BaBar Collaboration], “Evidence for an excess of  $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$  decays,” *Phys. Rev. Lett.* **109**, 101802 (2012) [arXiv:1205.5442]; A. Bozek *et al.* [Belle Collaboration], “Observation of  $B^+ \rightarrow \bar{D}^{*0}\tau^+\nu_\tau$  and Evidence for  $B^+ \rightarrow \bar{D}^0\tau^+\nu_\tau$  at Belle,” *Phys. Rev. D* **82**, 072005 (2010) [arXiv:1005.2302].
- [46] J. Chay, H. Georgi and B. Grinstein, “Lepton energy distributions in heavy meson decays from QCD,” *Phys. Lett. B* **247**, 399 (1990).
- [47] I. I. Y. Bigi, N. G. Uraltsev and A. I. Vainshtein, “Nonperturbative corrections to inclusive beauty and charm decays: QCD versus phenomenological models,” *Phys. Lett. B* **293**, 430 (1992) [Erratum-ibid. *B* **297**, 477 (1993)] [hep-ph/9207214].
- [48] I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, “QCD predictions for lepton spectra in inclusive heavy flavor decays,” *Phys. Rev. Lett.* **71**, 496 (1993) [hep-ph/9304225].
- [49] A. V. Manohar and M. B. Wise, “Inclusive semileptonic  $B$  and polarized  $\Lambda_b$  decays from QCD,” *Phys. Rev. D* **49**, 1310 (1994) [hep-ph/9308246].
- [50] M. Misiak *et al.*, “Estimate of  $\mathcal{B}(\bar{B} \rightarrow X_s\gamma)$  at  $\mathcal{O}(\alpha_s^2)$ ,” *Phys. Rev. Lett.* **98**, 022002 (2007) [hep-ph/0609232]; and references therein.
- [51] I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, “On the motion of heavy quarks inside hadrons: Universal distributions and inclusive decays,” *Int. J. Mod. Phys. A* **9**, 2467 (1994) [hep-ph/9312359].
- [52] M. Neubert, “Analysis of the photon spectrum in inclusive  $B \rightarrow X_s\gamma$  decays,” *Phys. Rev. D* **49**, 4623 (1994) [hep-ph/9312311].
- [53] P. J. Fox, Z. Ligeti, M. Papucci, G. Perez and M. D. Schwartz, “Deciphering top flavor violation at the LHC with  $B$  factories,” *Phys. Rev. D* **78**, 054008 (2008) [arXiv:0704.1482].
- [54] T. Aaltonen *et al.* [CDF Collaboration], “Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production,” *Phys. Rev. D* **83**, 112003 (2011) [arXiv:1101.0034].
- [55] M. I. Gresham, I. W. Kim and K. M. Zurek, “Tevatron Top  $A_{FB}$  Versus LHC Top Physics,” *Phys. Rev. D* **85**, 014022 (2012) [arXiv:1107.4364].
- [56] R. Harnik, J. Kopp and J. Zupan, “Flavor Violating Higgs Decays,” *JHEP* **1303**, 026 (2013) [arXiv:1209.1397].
- [57] A. Dery, A. Efrati, Y. Hochberg and Y. Nir, “What if  $\text{BR}(h \rightarrow \mu\mu)/\text{BR}(h \rightarrow \tau\tau)$  does not equal  $m_\mu^2/m_\tau^2$ ?” *JHEP* **1305**, 039 (2013) [arXiv:1302.3229].
- [58] H. E. Haber, “The Status of the minimal supersymmetric standard model and beyond,” *Nucl. Phys. Proc. Suppl.* **62**, 469 (1998) [hep-ph/9709450].
- [59] Y. Nir and N. Seiberg, “Should squarks be degenerate?,” *Phys. Lett. B* **309**, 337 (1993) [hep-ph/9304307].
- [60] O. Gedalia, J. F. Kamenik, Z. Ligeti and G. Perez, “On the Universality of  $CP$  Violation in  $\Delta F = 1$

- Processes,” Phys. Lett. B **714**, 55 (2012) [arXiv:1202.5038].
- [61] A. Crivellin and M. Davidkov, “Do squarks have to be degenerate? Constraining the mass splitting with  $K - \bar{K}$  and  $D - \bar{D}$  mixing,” Phys. Rev. D **81**, 095004 (2010) [arXiv:1002.2653].
- [62] R. Mahbubani, M. Papucci, G. Perez, J. T. Ruderman and A. Weiler, “Light Nondegenerate Squarks at the LHC,” Phys. Rev. Lett. **110**, no. 15, 151804 (2013) [arXiv:1212.3328].
- [63] R. S. Chivukula and H. Georgi, “Composite Technicolor Standard Model,” Phys. Lett. B **188**, 99 (1987).
- [64] L. J. Hall and L. Randall, “Weak scale effective supersymmetry,” Phys. Rev. Lett. **65**, 2939 (1990).
- [65] G. D’Ambrosio, G. F. Giudice, G. Isidori and A. Strumia, “Minimal flavour violation: An effective field theory approach,” Nucl. Phys. B **645**, 155 (2002) [hep-ph/0207036].
- [66] Y. Grossman, Y. Nir, J. Thaler, T. Volansky and J. Zupan, “Probing minimal flavor violation at the LHC,” Phys. Rev. D **76**, 096006 (2007) [arXiv:0706.1845].

# Practical Statistics for Particle Physicists

Harrison B. Prosper

Florida State University, Department of Physics, Tallahassee, USA

## Abstract

These lectures introduce the basic ideas and practices of statistical analysis for particle physicists, using a real-world example to illustrate how the abstractions on which statistics is based are translated into practical application.

## 1 Introduction

The day-to-day task of particle physicists is to suggest, build, test, discard and, or, refine models of the observed regularities in Nature with the ultimate goal of building a comprehensive model that answers all the scientific questions we might think to ask. One goal of experimental particle physicists is to make quantitative statements about the parameters  $\theta$  of a model given a set of experimental observations  $X$ . However, in order to make such statements, the connection between the observations and the model parameters must itself be modeled, and herein lies a difficulty. While there is general agreement about how to connect model parameters to data, there is long history [1] of disagreement about the best way to solve the inverse problem, that is, to go from observations to model parameters. The solution of this inverse problem requires a theory of inference.

These lectures introduce to two broad classes of theories of inference, the frequentist and Bayesian approaches. While our focus is on the practical, we do not shy away from brief discussions of foundations. We do so in order to make two points. The first is that when it comes to statistics, there is no such thing as “the” answer; rather there are answers based on assumptions, or proposals, on which reasonable people may disagree for purely intellectual reasons. Second, none of the current theories of inference is perfect. It is worth appreciating these points, even superficially, if only to avoid fruitless arguments that cannot be resolved because they are ultimately about intellectual taste rather than mathematical correctness.

For more in-depth expositions of the topics here covered, and different points of view, we highly recommend the excellent textbooks on statistics written for physicists, by physicists [2–4].

## 2 Lecture 1: Descriptive Statistics, Probability and Likelihood

### 2.1 Descriptive Statistics

Suppose we have a sample of  $N$  data  $X = x_1, x_2, \dots, x_N$ . It is often useful to summarize these data with a few numbers called statistics. A **statistic** is any number that can be calculated from the data and known parameters. For example,  $t = (x_1 + x_N)/2$  is a statistic, but if the value of  $\theta$  is unknown  $t = (x_1 - \theta)^2$  is not. However, a word of caution is in order: we particle physicists are prone to misuse the jargon of professional statisticians. For example, we tend to refer to *any* function of the data as a statistic including those that contain unknown parameters.

The two most important statistics are

the <b>sample mean</b> (or average)	$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i,$	(1)
and the <b>sample variance</b>	$s^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2,$	

$$\begin{aligned}
&= \frac{1}{N} \sum_{i=1}^N x_i^2 - \bar{x}^2, \\
&= \overline{x^2} - \bar{x}^2.
\end{aligned} \tag{2}$$

The sample average is a measure of the center of the distribution of the data, while the sample variance is a measure of its spread. Statistics that merely characterize the data are called **descriptive statistics**, of which the sample average and variance are the most important. If we order the data, say from the smallest value to the largest, we can compute another interesting statistic  $t_k \equiv x_{(k)}$ , where  $1 \leq k \leq N$  and  $x_{(k)}$  denotes the datum at the  $k^{\text{th}}$  position. The statistic  $t_k$  is called the  $k^{\text{th}}$  **order statistic** and is a measure of the value of outlying data.

The average and variance, Eqs. (1) and (2), are numbers that can always be calculated given a data sample  $X$ . But now we consider numbers that cannot be calculated from the data alone. Imagine the repetition, infinitely many times, of whatever data generating system yielded our data sample  $X$  thereby creating an infinite sequence of data sets. We shall refer to the data generating system as an experiment and the infinite sequence as an infinite ensemble. The latter, together with all the mathematical operations we may wish to apply to it, are abstractions. After all, it is not possible to realize an infinite ensemble. The ensemble and all the operations on it exist in the same sense that the number  $\pi$  exists along with all valid mathematical operations on  $\pi$ .

The most common operation to perform on an ensemble is to compute the average of the statistics. This **ensemble average** suggests several potentially useful characteristics of the ensemble, which we list below.

Ensemble average	$\langle x \rangle$
Mean	$\mu$
Error	$\epsilon = x - \mu$
Bias	$b = \langle x \rangle - \mu$
Variance	$V = \langle (x - \langle x \rangle)^2 \rangle$
Standard deviation	$\sigma = \sqrt{V}$
Mean square error	$\text{MSE} = \langle (x - \mu)^2 \rangle$
Root MSE	$\text{RMS} = \sqrt{\text{MSE}}$

(3)

Notice that none of these numbers can be calculated in practice because the data required to do so do not concretely exist. Even in an experiment simulated on a computer, there are very few of these numbers we can calculate. If we know the mean  $\mu$ , perhaps because we have chosen its value — for example, we may have chosen the mass of the Higgs boson in our simulation, we can certainly calculate the error  $\epsilon$  for any simulated datum  $x$ . But, we can only *approximate* the ensemble average  $\langle x \rangle$ , bias  $b$ , variance  $V$ , and MSE, since our virtual ensemble is always finite. The point is this: the numbers that characterize the infinite ensemble are also abstractions, albeit useful ones. For example, the MSE is the most widely used measure of the closeness of an ensemble of numbers to some parameter  $\mu$ . The square root of the MSE is called the root mean square (RMS)<sup>1</sup>. The MSE can be written as

$$\text{MSE} = V + b^2. \tag{4}$$

**Exercise 1:** Show this

The MSE is the sum of the variance and the square of the bias, a very important result with practical consequences. For example, suppose that  $\mu$  represents the mass of the Higgs boson and  $x$  represents

<sup>1</sup>Sometimes, the RMS and standard deviation are using interchangeably. However, the RMS is computed with respect to  $\mu$ , while the standard deviation is computed with respect to the ensemble average  $\langle x \rangle$ . The RMS and standard deviations are identical only if the bias is zero.

some (typically very complicated) statistic that is considered an **estimator** of the mass. An estimator is any function, which when data are entered into it, yields an **estimate** of the quantity of interest, which we may take to be a measurement.

Words are important; “bias” is a case in point. It is an unfortunate choice for the difference  $\langle x \rangle - \mu$  because the word “bias” biases attitudes towards bias! Something that, or someone who, is biased is surely bad and needs to be corrected. Perhaps. But, it would be wasteful of data to make the bias zero if the net effect is to make the MSE larger than an MSE in which the bias is non-zero. The price for achieving  $b = 0$  in our example would be not only throwing away expensive data — which is bad enough — but also measuring a mass that is more likely to be further away from the Higgs boson mass. This may, or may not, be what we want to achieve.

As noted, many of the numbers listed in Eq. (3) cannot be calculated because the information needed is unknown. This is true, in particular, of the bias. However, sometimes it is possible to relate the bias to another ensemble quantity. Consider the ensemble average of the sample variance, Eq. (2),

$$\begin{aligned} \langle s^2 \rangle &= \langle \bar{x}^2 \rangle - \langle \bar{x} \rangle^2, \\ &= V - \frac{V}{N}, \end{aligned}$$

**Exercise 2a:** Show this

The sample variance has a bias of  $b = -V/N$ , which many argue should be corrected. Unfortunately, we cannot calculate the bias because it depends on an unknown parameter, namely, the variance  $V$ . However, if we replace the sample variance by  $s'^2 = cs^2$ , where the correction factor  $c = N/(N - 1)$ , we find that for the corrected variance estimator  $s'^2$  the bias is zero. Surely the world is now a better place? Well, not necessarily. Consider the ratio of  $MSE'$  to  $MSE$ , where  $MSE' = \langle (s'^2 - V)^2 \rangle$ ,  $MSE = \langle \delta^2 \rangle$  with  $\delta = s^2 - V$ , and  $b = -V/N$ ,

$$\begin{aligned} MSE'/MSE &= \langle (cs^2 - V)^2 \rangle / \langle \delta^2 \rangle, \\ &= c^2 \langle (s^2 - V/c)^2 \rangle / \langle \delta^2 \rangle, \\ &= c^2 \langle (\delta - b)^2 \rangle / \langle \delta^2 \rangle, \\ &= c^2 (1 - b^2 / \langle \delta^2 \rangle), \\ &= c^2 [1 - b^2 / (b^2 + \langle s^4 \rangle - (V + b)^2)]. \end{aligned}$$

From this we deduce that if  $\langle s^4 \rangle / [(V + b)^2 + b^2 / (c^2 - 1)] > 1$ , the unbiased estimate will be further away on average from  $V$  than the biased estimate. This is the case, for example, for a uniform distribution.

**Exercise 2b:** Use the method `Rndm()` of the Root class `TRandom3` to verify that  $MSE' > MSE$ .

## 2.2 Probability

When the weather forecast specifies that there is a 80% chance of rain tomorrow, most people have an intuitive sense of what this means. Likewise, most people have an intuitive understanding of what it means to say that there is a 50-50 chance for a tossed coin to land heads up. Probabilistic ideas are thousands of years old, but, starting in the sixteenth century these ideas were formalized into increasingly rigorous mathematical theories of probability. In the theory formulated by Kolmogorov in 1933,  $\Omega$  is some fixed mathematical space,  $E_1, E_2, \dots \subset \Omega$  are subsets (called events) defined in some reasonable way<sup>2</sup>, and  $P(E_j)$  is a number associated with subset  $E_j$ . These numbers satisfy the

### Kolmogorov Axioms

---

<sup>2</sup>If  $E_1, E_2, \dots$  are meaningful subsets of  $\Omega$ , so to is the complement  $\bar{E}_1, \bar{E}_2, \dots$  of each, as are countable unions and intersections of these subsets.

1.  $P(E_j) \geq 0$
2.  $P(E_1 + E_2 + \dots) = P(E_1) + P(E_2) + \dots$  for disjoint subsets
3.  $P(\Omega) = 1$ .

Consider two subsets  $A = E_1$  and  $B = E_2$ . The quantity  $AB$  means *A and B*, while  $A + B$  means *A or B*, with associated probabilities  $P(AB)$  and  $P(A + B)$ , respectively. Kolmogorov assumed, not unreasonably given the intuitive origins of probability, that probabilities sum to unity; hence the axiom  $P(\Omega) = 1$ . However, this assumption can be dropped so that probabilities remain meaningful even if  $P(\Omega) = \infty$  [5].

Figure 1 suggests another probability, namely, the number  $P(A|B) = P(AB)/P(B)$ , called the **conditional probability** of *A* given *B*. This permits statements such as: “the probability that this track was created by an electron given the measured track parameters” or “the probability to observe 17 events given that the mean background is 3.8 events”. Conditional probability is a very powerful idea, but the term itself is misleading. It implies that there are two kinds of probability: conditional and unconditional. In fact, *all* probabilities are conditional in that they always depend on a specific set of conditions, namely, those that define the space  $\Omega$ . It is entirely possible to embed a family of subsets of  $\Omega$  into another space  $\Omega'$  which assigns to each family member a different probability  $P'$ . A probability is defined only relative to some space of possibilities  $\Omega$ .

*A* and *B* are said to be mutually exclusive if  $P(AB) = 0$ , that is, if the truth of one denies the truth of the other. They are said to be exhaustive if  $P(A) + P(B) = 1$ . Figure 1 suggests the theorem

$$P(A + B) = P(A) + P(B) - P(AB), \quad (5)$$

**Exercise 3:** Prove theorem

which can be deduced from the rules given above. Another useful theorem is an immediate consequence of the commutativity of “anding”  $P(AB) = P(BA)$  and the definition of  $P(A|B)$ , namely,

**Bayes Theorem**

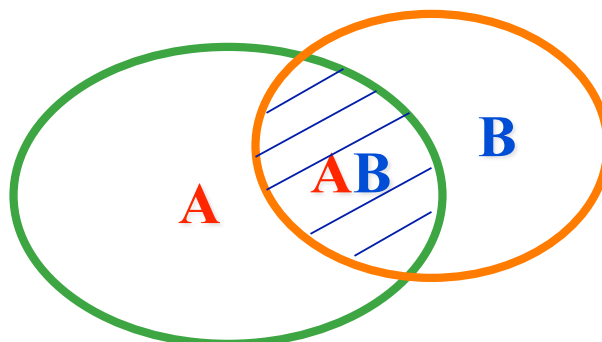
$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}, \quad (6)$$

which provides a way to convert the probability  $P(A|B)$  to the probability  $P(B|A)$ . Using Bayes theorem, we can, for example, deduce the probability  $P(e|x)$  that a particle is an electron, *e*, given a set of measurements, *x*, from the probability  $P(x|e)$  of a set of measurements given that the particle is an electron.

**2.2.1 Probability Distributions**

In this section, we illustrate the use of these rules to derive more complicated probabilities. First we start with a definition:

A **Bernoulli trial**, named after the Swiss mathematician Jacob Bernoulli (1654 – 1705), is an experiment with only two possible outcomes: *S* = success or *F* = failure.



**Fig. 1:** Venn diagram of the sets *A*, *B*, and *AB*.  $P(A)$  is the probability of *A*, while  $P(A|B) = P(AB)/P(B)$  is the probability of *AB* relative to that of *B*, i.e., the probability of *A* given the condition *B*.



*Example*

Each collision between protons at the Large Hadron Collider (LHC) is a Bernoulli trial in which something interesting happens ( $S$ ) or does not ( $F$ ). Let  $p$  be the probability of a success, which is assumed to be the *same for each trial*. Since  $S$  and  $F$  are exhaustive, the probability of a failure is  $1 - p$ . For a given order  $O$  of  $n$  proton-proton collisions and exactly  $k$  successes, and therefore exactly  $n - k$  failures, the probability  $P(k, O, n, p)$  is given by

$$P(k, O, n, p) = p^k(1 - p)^{n-k}. \tag{7}$$

If the order  $O$  of successes and failures is judged to be irrelevant, we can eliminate the order from the problem by summing over all possible orders,

$$P(k, n, p) = \sum_O P(k, O, n, p) = \sum_O p^k(1 - p)^{n-k}. \tag{8}$$

This procedure is called **marginalization**. It is one of the most important operations in probability calculations. Every term in the sum in Eq. (8) is identical and there are  $\binom{n}{k}$  of them. This yields the **binomial distribution**,

$$\text{Binomial}(k, n, p) \equiv \binom{n}{k} p^k(1 - p)^{n-k}. \tag{9}$$

By definition, the mean number of successes  $a$  is given by

$$\begin{aligned} a &= \sum_{k=0}^n k \text{Binomial}(k, n, p), \\ &= pn. \end{aligned} \tag{10}$$

**Exercise 4:** Show this

At the LHC  $n$  is a number in the trillions, while for successes of interest such as the creation of a Higgs boson the probability  $p \ll 1$ . In this case, it proves convenient to consider the limit  $p \rightarrow 0, n \rightarrow \infty$  in such a way that  $a$  remains constant. In this limit

$$\begin{aligned} \text{Binomial}(k, n, p) &\rightarrow e^{-a} a^k / k!, \\ &\equiv \text{Poisson}(k, a). \end{aligned} \tag{11}$$

**Exercise 5:** Show this

Below we list the most common probability distributions.

**Discrete distributions**

Binomial( $k, n, p$ )	$\binom{n}{k} p^k(1 - p)^{n-k}$
Poisson( $k, a$ )	$a^k \exp(-a) / k!$
Multinomial( $k, n, p$ )	$\frac{n!}{k_1! \cdots k_K!} \prod_{i=1}^K p_i^{k_i}, \quad \sum_{i=1}^K p_i = 1, \quad \sum_{i=1}^K k_i = n$

**Continuous densities**

Uniform( $x, a$ )	$1/a$
Gaussian( $x, \mu, \sigma$ )	$\exp[-(x - \mu)^2 / (2\sigma^2)] / (\sigma\sqrt{2\pi})$

(also known as the Normal density)

$$\begin{array}{ll}
 \text{LogNormal}(x, \mu, \sigma) & \exp[-(\ln x - \mu)^2 / (2\sigma^2)] / (x\sigma\sqrt{2\pi}) \\
 \text{Chisq}(x, n) & x^{n/2-1} \exp(-x/2) / [2^{n/2}\Gamma(n/2)] \\
 \text{Gamma}(x, a, b) & x^{a-1} a^b \exp(-ax) / \Gamma(b) \\
 \text{Exp}(x, a) & a \exp(-ax) \\
 \text{Beta}(x, n, m) & \frac{\Gamma(n+m)}{\Gamma(m)\Gamma(n)} x^{n-1} (1-x)^{m-1} \quad (12)
 \end{array}$$

Particle physicists tend to use the term probability distribution for both discrete and continuous functions, such as the Poisson and Gaussian distributions, respectively. But, strictly speaking, the continuous functions are probability *densities*, not probability distributions. In order to compute a probability from a density we need to integrate the density over a finite set in  $x$ .

### Discussion

Probability is the foundation for models of non-deterministic data generating mechanisms, such as particle collisions at the LHC. A **probability model** is the probability distribution together with all the assumptions on which the distribution is based. For example, suppose we wish to count, during a given period of time, the number of entries  $N$  in a given transverse momentum ( $p_T$ ) bin due to particles created in proton-proton collisions at the LHC; that is, suppose we wish to perform a counting experiment. If we assume that the probability to obtain a count in this bin is very small and that the number of proton-proton collisions is very large, then it is common practice to use a Poisson distribution to model the data generating mechanism, which yields the bin count  $N$ . If we have multiple independent bins, we may choose to model the data generating mechanism as a product of Poisson distributions. Or, perhaps, we may prefer to model the possible counts conditional on a fixed total count in which case a multinomial distribution would be appropriate.

So far, we have assumed the meaning of the word probability to be self-evident. However, the meaning of probability [6] has been the subject of debate for more than two centuries and there is no sign that the debate will end anytime soon. Probability, in spite of its intuitive beginnings, is an abstraction. Therefore, for it to be of practical use it must be *interpreted*. The two most widely used interpretations of probability are:

1. **degree of belief** in, or plausibility of, a proposition, for example, "It will snow at CERN on December 18th", and the
2. **relative frequency** of outcomes in an *infinite* ensemble of trials, for example, the relative frequency of Higgs boson creation in an infinite number of proton-proton collisions.

The first interpretation is the older, while the second was championed by influential mathematicians and logicians starting in the mid-nineteenth century and became the dominant interpretation. Of the two interpretations, however, the older is the more general in that it encompasses the latter and can be used in contexts in which the latter makes no sense. The relative frequency, or **frequentist**, interpretation is useful for situations in which one can contemplate counting the number of times  $k$  a given outcome is realized in  $n$  trials, as in the example of a counting experiment. The relative frequency  $r = k/n$  is expected to converge, in a subtle but well-defined sense, to some number  $p$  that satisfies the rules of probability. It should be noted, however, that the numbers  $k/n$  and  $p$  are conceptually distinct. The former is something we can actually calculate, while there is no *finite* operational way to calculate the latter from data. The probability  $p$ , even when interpreted as a relative frequency, remains an abstraction.

On the other hand, the degrees of belief, which is the basis of the *Bayesian* approach to statistics (see Lecture 2), are just that: the degree to which a rational being *ought* to believe in the veracity of a given statement. The word "ought" in the last sentence is important: probability theory, with probabilities

interpreted as degrees of belief, is *not* a model of how human beings actually reason in situations of uncertainty; rather probability theory when interpreted this way is a normative theory in that it specifies how an idealized reasoning being, or system, ought to reason when faced with uncertainty.

There is a school of thought that argues that degrees of belief should be an individual’s own assessment of her or his degree of belief in a statement, which are then to be updated using the probability rules. The problem with this position is that it presupposes probability theory to be a model of human reasoning, which we argue it is not — a position confirmed by numerous psychological experiments. It is perhaps better to think of degrees of belief as numbers that inform one’s reasoning rather than as numbers that describe it, and relative frequencies as numbers that characterize stochastic data generation mechanisms. Both are probabilities and both are useful.

### 2.3 Likelihood

Let us assume that  $p(x|\theta)$  is a **probability density function** (pdf) such that  $P(A|\theta) = \int_A p(x|\theta) dx$  is the probability of the statement  $A = x \in R_x$ , where  $x$  denotes possible data,  $\theta$  the parameters that characterize the probability model, and  $R_x$  is a finite set. If  $x$  is discrete, then both  $p(x|\theta)$  and  $P(A|\theta)$  are probabilities. The **likelihood function** is simply the probability model  $p(x|\theta)$  evaluated at the data  $x_O$  actually obtained, i.e., the function  $p(x_O|\theta)$ . The following are examples of likelihoods.

*Example 1*

In 1995, CDF and DØ discovered the top quark [7, 8] at Fermilab. The DØ Collaboration found  $x = D$  events ( $D = 17$ ). For a counting experiment, the datum can be modeled using

$$\begin{aligned} p(x|d) &= \text{Poisson}(x, d) && \text{probability to get } x \text{ events} \\ p(D|d) &= \text{Poisson}(D, d) && \text{likelihood of observation } D \text{ events} \\ &= d^D \exp(-d)/D! \end{aligned}$$

We shall analyze this example in detail in Lectures 2 and 3.

*Example 2*

Figure 2 shows the transverse momentum spectrum of jets in  $pp \rightarrow \text{jet} + X$  events measured by the CMS Collaboration [9]. The spectrum has  $K = 20$  bins with total count  $N$  that was modeled using the likelihood

$$\begin{aligned} p(D|p) &= \text{Multinomial}(D, N, p), \quad D = D_1, \dots, D_K, \quad p = p_1, \dots, p_K \\ \sum_{i=1}^K D_i &= N. \end{aligned}$$

This is an example of a *binned* likelihood.

**Exercise 6a:** Show that a multi-Poisson likelihood can be written as the product of a multinomial and a Poisson with count  $N$

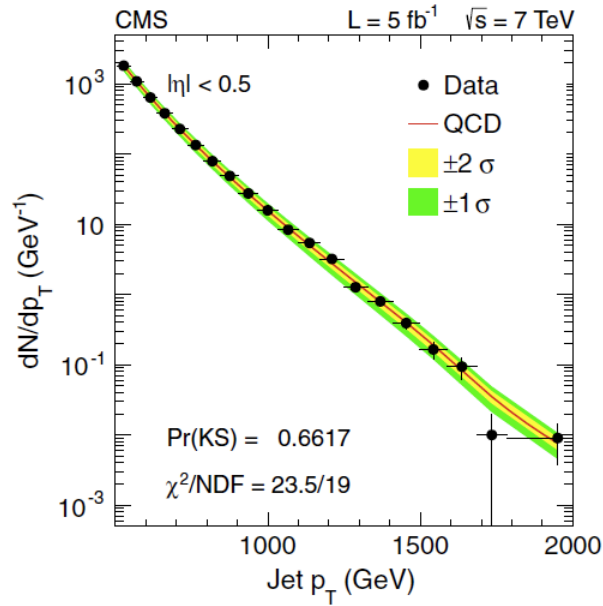
*Example 3*

Figure 3 shows a plot of the distance modulus versus redshift for  $N = 580$  Type 1a supernovae [10]. These heteroscedastic data<sup>3</sup>  $\{z_i, x_i \pm \sigma_i\}$  are modeled using the likelihood

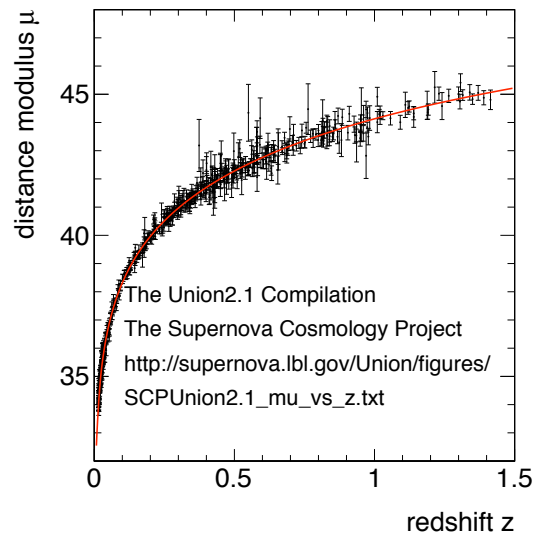
$$p(D|\Omega_M, \Omega_\Lambda, Q) = \prod_{i=1}^N \text{Gaussian}(x_i, \mu_i, \sigma_i),$$

---

<sup>3</sup>Data in which each item,  $x_i$ , or group of items has a different uncertainty.



**Fig. 2:** Transverse momentum spectrum of jets in  $pp \rightarrow \text{jet} + X$  events measured by CMS compared with the QCD prediction at next-to-leading order. This spectrum was used to search for evidence of contact interactions [9] (Courtesy CMS Collaboration).



**Fig. 3:** Plot of the data points  $(z_i, x_i \pm \sigma_i)$  for 580 Type 1a supernovae [10] showing a fit of the standard cosmological model (with a cosmological constant) to these data (curve).

which is an example of an *un-binned* likelihood. The cosmological model is encoded in the distance modulus function  $\mu_i$ , which depends on the redshift  $z_i$  and the matter density and cosmological constant parameters  $\Omega_M$  and  $\Omega_\Lambda$ , respectively. (See Ref. [11] for an accessible introduction to the analysis of these data.)

*Example 4*

The discovery of a neutral Higgs boson in 2012 by ATLAS [12] and CMS [13] in the di-photon final state ( $pp \rightarrow H \rightarrow \gamma\gamma$ ) made use of an un-binned likelihood of the form,

$$p(x|s, m, w, b) = \exp[-(s + b)] \prod_{i=1}^N [s f_s(x_i|m, w) + b f_b(x_i)]$$

where  $x$  = di-photon masses

$m$  = mass of boson

$w$  = width of resonance

$s$  = expected (i.e., mean) signal count

$b$  = expected background count

$f_s$  = signal probability density

$f_b$  = background probability density

**Exercise 6b:** Show that a binned multi-Poisson likelihood yields an un-binned likelihood of this form as the bin widths go to zero

The likelihood function is arguably the most important quantity in a statistical analysis. Because it can be used to answer questions such as the following.

1. How do I estimate a parameter?
2. How do I quantify its accuracy?
3. How do I test an hypothesis?
4. How do I quantify the significance of a result?

Writing down the likelihood function requires:

1. identifying all that is *known*, e.g., the observations,
2. identifying all that is *unknown*, e.g., the parameters,
3. constructing a probability model for *both*.

Many analyses in particle physics do not use likelihood functions explicitly. However, it is worth spending time to think about them because doing so encourages a deeper reflection on what is being done, a more systematic approach to the statistical analysis, and ultimately leads to better answers.

Being explicit about what is and is not known in an analysis problem may seem a pointless exercise; surely these things are obvious. Consider the  $D\bar{D}$  top quark discovery data [8],  $D = 17$  events observed with a background estimate of  $B = 3.8 \pm 0.6$  events. The uncertainty in 17 is invariably said to be  $\sqrt{17} = 4.1$ . Not so! The count 17 is perfectly known: it is 17. What we are uncertain about is the mean count  $d$ , that is, the parameter of the probability model, which we take to be a Poisson distribution. The  $\pm 4.1$  must somehow be a statement not about 17 but rather about the unknown parameter  $d$ . We shall explain what the  $\pm 4.1$  means in Lecture 2.

### 3 Lecture 2: The Frequentist and Bayesian Approaches

In this lecture, we consider the two most important approaches to statistical inference, frequentist and Bayesian. Both are needed to make sense of statistical inference, though this is not the dominant opinion

in particle physics. Most particle physicists, if pressed, will say they are frequentist in their approach. The typical reason given is that this approach is objective, whereas the Bayesian approach is not. Moreover, they would argue, the frequentist approach is less arbitrary whereas the Bayesian approach is plagued with arbitrariness that renders its results suspect. We wish, however, to focus on the practical, therefore, we shall sidestep this debate and assume a pragmatic attitude to both approaches. We begin with a description of salient features of the frequentist approach, followed by a description of the Bayesian approach.

### 3.1 The Frequentist Approach

The most important principle in this approach is that enunciated by the Polish statistician Jerzy Neyman in the 1930s, namely,

#### The Frequentist Principle

The goal of a frequentist analysis is to construct statements so that a fraction  $f \geq p$  of them are guaranteed to be true over an infinite ensemble of statements.

The fraction  $f$  is called the **coverage probability**, or coverage for short, and  $p$  is called the **confidence level** (C.L.). A procedure which satisfies the frequentist principle is said to *cover*. The confidence level as well as the coverage is a property of the ensemble of statements. Consequently, the confidence level may change if the ensemble changes. Here is an example of the frequentist principle in action.

#### Example

Over the course of a long career, a doctor sees thousands of patients. For each patient he issues one of two conclusions: “you are sick” or “you are well” depending on the results of diagnostic measurements. Because he is a frequentist, he has devised an approach to medicine in which although he does not know which of his conclusions were correct, he can at least retire happy in the knowledge that he was correct at least 75% of the time!

In a seminal paper published in 1937, Neyman [14] invented the concept of the confidence interval, a way to quantify uncertainty, that respects the frequentist principle. The confidence interval is such an important idea, and its meaning so different from the superficially similar Bayesian concept of a credible interval, that it is worth working through the concept in detail.

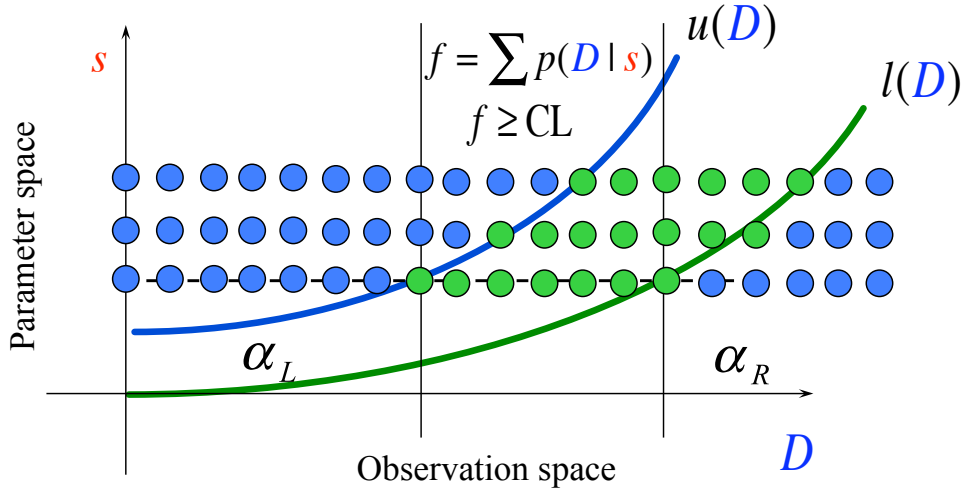
#### 3.1.1 Confidence Intervals

The confidence interval is a concept best explained by example. Consider an experiment that observes  $D$  events with expected (that is, mean) signal  $s$  and no background. Neyman devised a way to make statements of the form

$$s \in [l(D), u(D)], \quad (13)$$

with the *a priori* guarantee that at least a fraction  $p$  of them will be true, as required by the frequentist principle. A procedure for constructing such intervals is called a **Neyman construction**. The frequentist principle must hold for any ensemble of experiments, not necessarily all making the same kind of observations and statements. For simplicity, however, we shall presume the experiments to be of the same kind and to be completely specified by a single unknown parameter  $s$ . The Neyman construction is illustrated in Fig. 4.

The construction proceeds as follows. Choose a value of  $s$  and use some rule to find an interval in the space of observations (or, more generally, a region), for example, the interval defined by the two



**Fig. 4:** The Neyman construction. Plotted is the Cartesian product of the parameter space, with parameter  $s$ , and the space of observations with potential observations  $D$ . For a given value of  $s$ , the observation space is partitioned into three disjoint intervals, such that the probability to observe a count  $D$  within the interval demarcated by the two vertical lines is  $f \geq p$ , where  $p = \text{C.L.}$  is the desired confidence level. The inequality is needed because, for discrete data, it may not be possible to find an interval with  $f = p$  exactly.

vertical lines in the center of the figure, such that the probability to obtain a count in this interval is  $f \geq p$ , where  $p$  is the desired confidence level. We move to another value of  $s$  and repeat the procedure. The procedure is repeated for a sufficiently dense set of points in the parameter space over a sufficiently large range. When this is done, as illustrated in Fig. 4, the intervals of probability content  $f$  will form a band in the Cartesian product of the parameter space and the observation space. The upper edge of this band defines the curve  $u(D)$ , while the lower edge defines the curve  $l(D)$ . These curves are the product of the Neyman construction.

For a given value of the parameter of interest  $s$ , the interval with probability content  $f$  in the space of observations is not unique; different rules for choosing the interval will, in general, yield different intervals. Neyman suggested choosing the interval so that the probability to obtain an observation below or above the interval are the same. The Neyman rule yields the so-called **central intervals**. One virtue of central intervals is that their boundaries can be more efficiently calculated by solving the equations,

$$\begin{aligned} P(x \leq D|u) &= \alpha_L, \\ P(x \geq D|l) &= \alpha_R, \end{aligned} \quad (14)$$

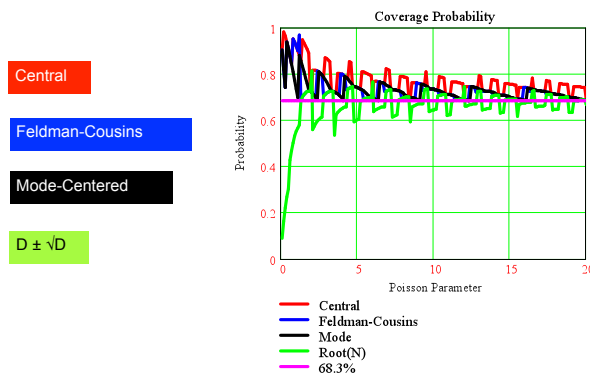
a mathematical fact that becomes clear after staring at Fig. 4 long enough.

Another rule was suggested by Feldman and Cousins [15]. For our example, the Feldman-Cousins rule requires that the potential observations  $\{D\}$  be ordered in descending order,  $D_{(1)}, D_{(2)}, \dots$ , of the likelihood ratio  $p(D|s)/p(D|\hat{s})$ , where  $\hat{s}$  is the maximum likelihood estimator (see Sec. 3.1.2) of the parameter  $s$ . Once ordered, we compute the running sum  $f = \sum_j p(D_{(j)}|s)$  until  $f$  equals or just exceeds the desired confidence level  $p$ . This rule does not guarantee that the potential observations  $D$  are contiguous, but this does not matter because we simply take the minimum element of the set  $\{D_{(j)}\}$  to be the lower bound of the interval and its maximum element to be the upper bound.

Another simple rule is the mode-centered rule: order  $D$  in descending order of  $p(D|s)$  and proceed as with the Feldman-Cousins rule. In principle, absent criteria for choosing a rule, there is nothing to prevent the use of *ordering rules* randomly chosen for different values of  $s$ ! Figure 5 compares the widths of the intervals  $[l(D), u(D)]$  for three different ordering rules, central, Feldman-Cousins, and mode-centered as a function of the count  $D$ . It is instructive to compare these widths with those provided

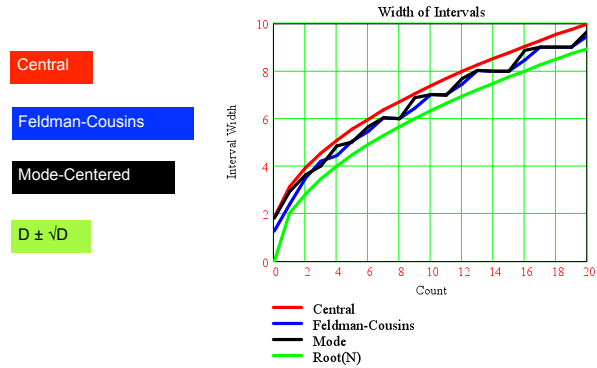
by the well-known root(N) interval,  $l(D) = D - \sqrt{D}$  and  $u(D) = D + \sqrt{D}$ . Of the three sets of intervals, the ones suggested by Neyman are the widest, the Feldman-Cousins and mode-centered ones are of similar width, while the root(N) intervals are the shortest. So why are we going through all the trouble of the Neyman construction? We shall return to this question shortly.

Having completed the Neyman construction and found the curves  $u(D)$  and  $l(D)$  we can use the latter to make statements of the form  $s \in [l(D), u(D)]$ : for a given observation  $D$ , we simply read off the interval  $[l(D), u(D)]$  from the curves. For example, suppose in Fig. 4 that the true value of  $s$  is represented by the horizontal line that intersects the curves  $u(D)$  and  $l(D)$  and which therefore defines the interval demarcated by the two vertical lines. If the observation  $D$  happens to fall in the interval to the left of the left vertical line, or to the right of the right vertical line, then the interval  $[l(D), u(D)]$  will not bracket  $s$ . However, if  $D$  falls between the two vertical lines, the interval  $[l(D), u(D)]$  will bracket  $s$ . Moreover, by virtue of the Neyman construction, a fraction  $f$  of the intervals  $[l(D), u(D)]$  will bracket the value of  $s$  whatever its value happens to be, which brings us back to the question about the root(N) intervals. Figure 6 shows the coverage probability over the parameter space of  $s$ . As expected, the three rules, Neyman's, that of Feldman-Cousins, and the mode-centered, satisfy the condition coverage probability  $\geq$  confidence level over all values of  $s$  that are possible *a priori*; that is, the intervals cover. However, the root(N) intervals do not and indeed fail badly for  $s < 2$ .



**Fig. 6:** Interval widths as a function of count  $D$  for four sets of intervals.

presume it is about that particular interval. It is not because  $p$ , as noted, is a property of the ensemble to which this statement belongs. The precise statement is this:  $s \in [l(D), u(D)]$  is a member of an (infinite) ensemble of statements a fraction  $f \geq p$  of which are true. This mathematical fact is the principal reason why the frequentist approach is described as objective; the probability  $p$  is something for which there seems, in principle, to be an operational definition: we just count how many statements of the form  $s \in [l(D), u(D)]$  are true and divide by the total number of statements. Unfortunately, in



**Fig. 5:** Interval widths as a function of count  $D$  for four sets of intervals.

However, notice that the coverage probability of the root(N) intervals bounces around the (68%) confidence level for values of  $s > 2$ . Therefore, if we knew for sure that  $s > 2$ , it would seem that using the root(N) intervals may not be that bad after all. Whether it is or not depends entirely on one's attitude towards the frequentist principle. Some will lift mountains and carry them to the Moon in order to achieve exact coverage, while others, including the author, is entirely happy with coverage that bounces around a little.

*Discussion*

We may summarize the content of the Neyman construction with a statement of the form: there is a probability of at least  $p$  that  $s \in [l(D), u(D)]$ .

But it would be a misreading of the statement to



the real world this procedure cannot be realized because in general we are not privy to which statements are true and, even if we came down from a mountain with the requisite knowledge, we would need to examine an infinite number of statements, which is impossible. Nevertheless, the Neyman construction is a remarkable procedure that always yields exact coverage for any problem that depends on a *single* unknown parameter.

Matters quickly become less tidy, however, when a probability model contains more than one unknown parameter. In almost every particle physics experiment there is background that is usually not known precisely. Consequently, even for the simplest experiment we must contend with at least two parameters, the expected signal  $s$  and the expected background  $b$ , neither of which is known. Neyman required a procedure to cover whatever the value of *all* the parameters be they known or unknown. This is a very tall order, which cannot be met in general. In practice, we resort to approximations, the most widely used of which is the profile likelihood to which we now turn.

### 3.1.2 The Profile Likelihood

As noted in Sec. 2.3, likelihood functions can be used to estimate the parameters on which they depend. The method of choice to do so, in a frequentist analysis, is called **maximum likelihood**, a method first used by Karl Frederick Gauss, *The Prince of Mathematics*, but developed into a formidable statistical tool in the 1930s by Sir Ronald A. Fisher [16], perhaps the most influential statistician of the twentieth century.

Fisher showed that a good way to estimate the parameters of a likelihood function is to pick the value that maximizes it. Such estimates are called maximum likelihood estimates (MLE). In general, a function into which data can be inserted to yield an MLE of a parameter is called a maximum likelihood estimator. For simplicity, we shall use the same abbreviation MLE to mean both the estimate and the estimator and we shall not be too picky about distinguishing the two. The DØ top quark discovery example illustrates the method.

#### *Example: Top Quark Discovery Revisited*

We start by listing

##### **the knowns**

$D = N, B$  where

$N = 17$  observed events

$B = 3.8$  estimated background events with uncertainty  $\delta B = 0.6$

##### **and the unknowns**

$b$  mean background count

$s$  mean signal count.

Next, we construct a probability model for the data  $D = N, B$  assuming that  $N$  and  $B$  are statistically independent. Since this is a counting experiment, we shall assume that  $p(x|s, b)$  is a Poisson distribution with mean count  $s + b$ . In the absence of details about how the background  $B$  was arrived at, the standard assumption is that data of the form  $y \pm \delta y$  can be modeled with a Gaussian (or normal) density. However, we can do a bit better. Background estimates are usually based on auxiliary experiments, either real or simulated, that define control regions.

Suppose that the observed count in the control region is  $Q$  and the mean count is  $bk$ , where  $k$  (ideally) is the known scale factor between the control and signal regions. We can model these data with a Poisson distribution with count  $Q$  and mean  $bk$ . But, we are given  $B$  and

$\delta B$  rather than  $Q$  and  $k$ , so we need a model to relate the two pairs of numbers. The simplest model is  $B = Q/k$  and  $\delta B = \sqrt{Q}/k$  from which we can infer an effective count  $Q$  using  $Q = (B/\delta B)^2$ . What of the scale factor  $k$ ? Well, since it is not given, it must be estimated. The obvious estimate is  $Q/B = B/\delta B^2$ . With these assumptions, our likelihood function is

$$p(D|s, b) = \text{Poisson}(N, s + b) \text{Poisson}(Q, bk), \quad (15)$$

where

$$\begin{aligned} Q &= (B/\delta B)^2 = 41.11, \\ k &= B/\delta B^2 = 10.56. \end{aligned}$$

The first term in Eq. (15) is the likelihood for the count  $N = 17$ , while the second term is the likelihood for  $B = 3.8$ , or equivalently the count  $Q$ . The fact that  $Q$  is not an integer causes no difficulty: we merely write the Poisson distribution as  $(bk)^Q \exp(-bk)/\Gamma(Q+1)$ , which permits continuation to non-integer counts  $Q$ .

The maximum likelihood estimators for  $s$  and  $b$  are found by maximizing Eq. (15), that is, by solving the equations

$$\begin{aligned} \frac{\partial \ln p(D|s, b)}{\partial s} &= 0 \quad \text{leading to } \hat{s} = N - B, \\ \frac{\partial \ln p(D|s, b)}{\partial b} &= 0 \quad \text{leading to } \hat{b} = B, \end{aligned}$$

as expected.

A more complete analysis would account for the uncertainty in  $k$ . One way is to introduce two more control regions with observed counts  $V$  and  $W$  and mean counts  $v$  and  $wk$ , respectively, and extend Eq. (15) with two more Poisson distributions.

The maximum likelihood method is the most widely used method for estimating parameters because it generally leads to reasonable estimates. But the method has features, or encourages practices, which, somewhat uncharitably, we label the good, the bad, and the ugly!

– *The Good*

- Maximum likelihood estimators are consistent: the RMS goes to zero as more and more data are included in the likelihood. This is an extremely important property, which basically says it makes sense to take more data because we shall get more accurate results. One would not knowingly use an inconsistent estimator!
- If an unbiased estimator for a parameter exists the maximum likelihood method will find it.
- Given the MLE for  $s$ , the MLE for any function  $y = g(s)$  of  $s$  is, very conveniently, just  $\hat{y} = g(\hat{s})$ . This is a very nice practical feature which makes it possible to maximize the likelihood using the most convenient parameterization of it and then transform back to the parameter of interest at the end.

– *The Bad (according to some!)*

- In general, MLEs are biased.

**Exercise 7:** Show this

Hint: Taylor expand  $y = g(\hat{s} + h)$  about the MLE  $\hat{s}$ , then consider its ensemble average.

- *The Ugly (according to some!)*
  - The fact that most MLEs are biased encourages the routine application of bias correction, which can waste data and, sometimes, yield absurdities.

Here is an example of the seriously ugly.

*Example*

For a discrete probability distribution  $p(k)$ , the **moment generating function** is the ensemble average

$$G(x) = \langle e^{xk} \rangle = \sum_k e^{xk} p(k).$$

For the binomial, with parameters  $p$  and  $n$ , this is

$$G(x) = (e^x p + 1 - p)^n, \quad \boxed{\text{Exercise 8a: Show this}}$$

which is useful for calculating **moments**

$$\mu_r = \left. \frac{d^r G}{dx^r} \right|_{x=0} = \sum_k k^r p(k),$$

e.g.,  $\mu_2 = (np)^2 + np - np^2$  for the binomial distribution. Given that  $k$  events out of  $n$  pass a set of cuts, the MLE of the event selection efficiency is the obvious estimate  $\hat{p} = k/n$ . The equally obvious estimate of  $p^2$  is  $(k/n)^2$ . But,

$$\langle (k/n)^2 \rangle = p^2 + V/n, \quad \boxed{\text{Exercise 8b: Show this}}$$

so  $(k/n)^2$  is a biased estimate of  $p^2$  with positive bias  $V/n$ . The unbiased estimate of  $p^2$  is

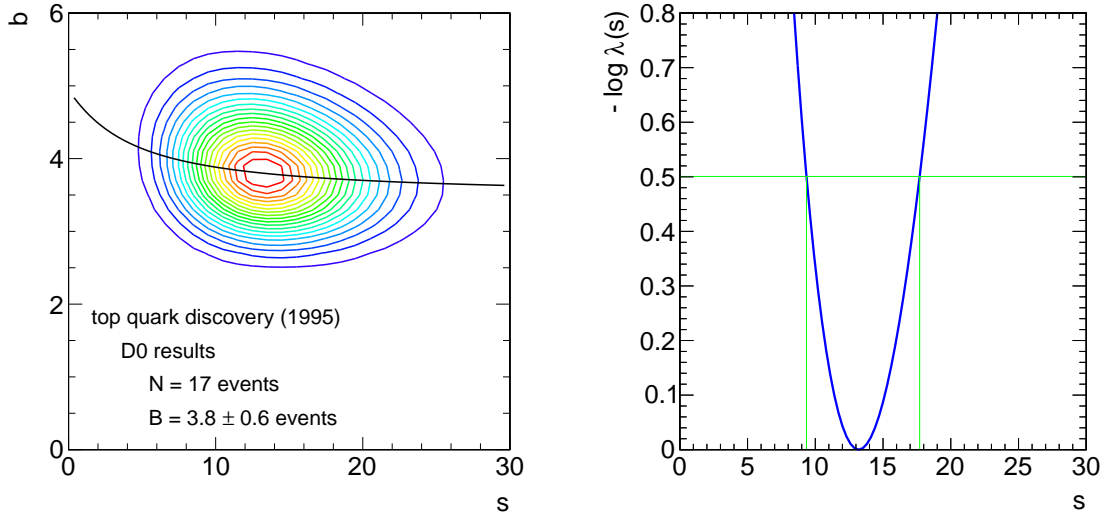
$$k(k-1)/[n(n-1)], \quad \boxed{\text{Exercise 8c: Show this}}$$

which, for a single success, i.e.,  $k = 1$ , yields the sensible estimate  $\hat{p} = 1/n$ , but the less than helpful one  $\hat{p}^2 = 0!$

In order to infer a value for the parameter of interest, for example, the signal  $s$  in our 2-parameter likelihood function in Eq. (15), the likelihood must be reduced to one involving the parameter of interest only, here  $s$ , by somehow getting rid of all the **nuisance** parameters, here the background parameter  $b$ . A nuisance parameter is simply a parameter that is not of current interest. In a strict frequentist calculation, this reduction to the parameter of interest must be done in such a way as to respect the frequentist principle: *coverage probability*  $\geq$  *confidence level*. In general, this is very difficult to do exactly.

In practice, we replace all nuisance parameters by their **conditional maximum likelihood estimates** (CMLE). The CMLE is the maximum likelihood estimate conditional on a *given* value of the current parameter (or parameters) of interest. In the top discovery example, we construct an estimator of  $b$  as a function of  $s$ ,  $\hat{b}(s)$ , and replace  $b$  in the likelihood  $p(D|s, b)$  by  $\hat{b}(s)$  to yield a function  $p_{PL}(D|s)$  called the **profile likelihood**.

*Since the profile likelihood entails an approximation, namely, replacing unknown parameters by their conditional estimates, it is not the likelihood but rather an approximation to it. Consequently, the frequentist principle is not guaranteed to be satisfied exactly.*



**Fig. 7:** (a) Contours of the DØ top discovery likelihood and the graph of  $\hat{b}(s)$ . (b) Plot of  $-\ln \lambda(17, s)$  versus the expected signal  $s$ . The vertical lines show the boundaries of the approximate 68% interval.

This does not seem to be much progress. However, things are much better than they may appear because of an important theorem proved by Wilks in 1938. If certain conditions are met, roughly that the MLEs do not occur on the boundary of the parameter space and the likelihood becomes ever more Gaussian as the data become more numerous — that is, in the so-called **asymptotic limit**, then if the true density of  $x$  is  $p(x|s, b)$  the random number

$$t(x, s) = -2 \ln \lambda(x, s), \quad (16)$$

$$\text{where } \lambda(x, s) = \frac{p_{PL}(x|s)}{p_{PL}(x|\hat{s})}, \quad (17)$$

has a probability density that converges to a  $\chi^2$  density with one degree of freedom. More generally, if the numerator of  $\lambda$  contains  $m$  free parameters the asymptotic density of  $t$  is a  $\chi^2$  density with  $m$  degrees of freedom. Therefore, we may take  $t(D, s)$  to be a  $\chi^2$  variate, at least approximately, and solve  $t(D, s) = n^2$  for  $s$  to get approximate  $n$ -standard deviation confidence intervals. In particular, if we solve  $t(D, s) = 1$ , we obtain approximate 68% intervals. This calculation is what `Minuit`, and now `TMinuit`, has done countless times since the 1970s! Wilks' theorem provides the main justification for using the profile likelihood. We again use the top discovery example to illustrate the procedure.

#### *Example: Top Quark Discovery Revisited Again*

The conditional MLE of  $b$  is found to be

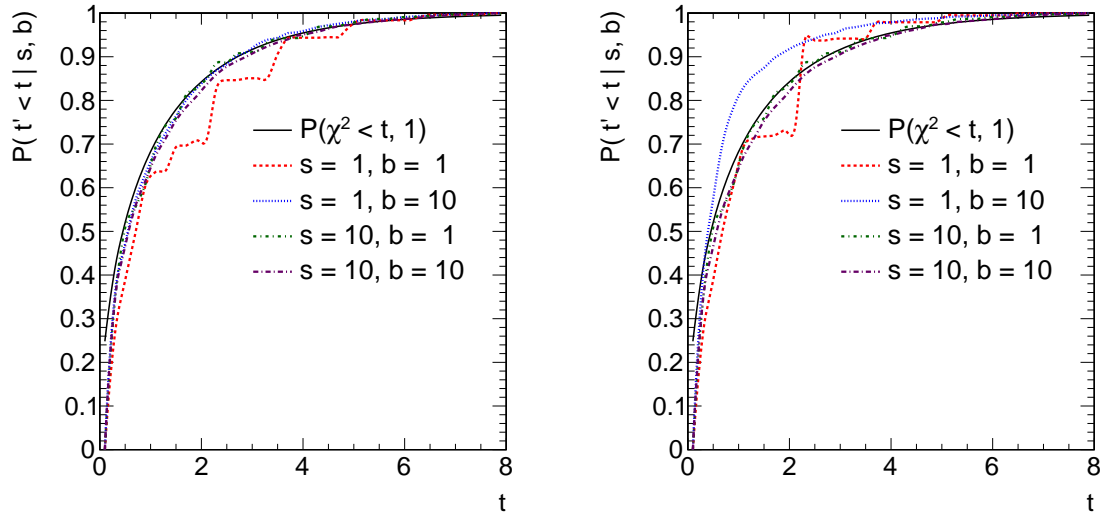
$$\hat{b}(s) = \frac{g + \sqrt{g^2 + 4(1+k)Qs}}{2(1+k)}, \quad (18)$$

where

$$g = N + Q - (1+k)s.$$

The likelihood  $p(D|s, b)$  is shown in Fig. 7(a) together with the graph of  $\hat{b}(s)$ . The mode (i.e. the peak) occurs at  $s = \hat{s} = N - B$ . By solving

$$-2 \ln \frac{p_{PL}(17|s)}{p_{PL}(17|17-3.8)} = 1$$



**Fig. 8:** Plots of the cumulative distribution function (cdf),  $P(\chi^2 < t, 1)$ , of the  $\chi^2$  density for one degree of freedom compared with the cdf  $P(t' < t | s, b)$  for four different values of the mean signal and background,  $s$  and  $b$ . The left plot shows that even for a mean signal or background count as low as 10, the density  $p(t|s, b)$  is already close to  $p(\chi^2, 1)$  and therefore largely independent of  $s$  and  $b$ . This is true, however, only if most of the time the maximum of the likelihood occurs away from the boundary of the parameter space. In the left plot, the signal is estimated using  $\hat{s} = N - B$ , which can, in principle, be arbitrarily negative. But, if we choose to set  $\hat{s} = 0$  whenever  $B > N$  in order to avoid negative signal estimates, we obtain the curves in the right plot. We see that for small signals,  $p(t|s, b)$  still depends on the parameters.

for  $s$  we get two solutions  $s = 9.4$  and  $s = 17.7$ . Therefore, we can make the statement  $s \in [9.4, 17.7]$  at approximately 68% C.L. Figure 7(b) shows a plot of  $-\ln \lambda(17, s)$  created using the RooFit [17] and RooStats [18] packages.

**Exercise 9:** Verify this interval using the RooFit/RooStats package

Intervals constructed this way are not guaranteed to satisfy the frequentist principle. In practice, however, their coverage is very good for the typical probability models used in particle physics, even for modest amounts of data. This is illustrated in Fig. 8, which shows how rapidly the density of  $t(x, s)$  converges to a  $\chi^2$  density for the probability distribution  $p(x, y | s, b) = \text{Poisson}(x | s + b) \text{Poisson}(y | b)$ <sup>4</sup>. The figure also shows what happens if we impose the restriction  $\hat{s} \geq 0$ , that is, we forbid negative signal estimates.

### 3.1.3 Hypothesis Tests

It is hardly possible in experimental particle physics to avoid the testing of hypotheses, testing that invariably leads to decisions. For example, electron identification entails hypothesis testing; given data  $D$  we ask: is this particle an isolated electron or is it not an isolated electron? Then we decide whether or not it is and proceed on the basis of the decision that has been made. In the discovery of the Higgs boson, we had to test whether, given the data available in early summer 2012, the Standard Model without a Higgs boson, a somewhat ill-founded background-only model, or the Standard Model with a Higgs boson, the background + signal model, was the preferred hypothesis. We decided that the latter model

<sup>4</sup>It was the difficulty of extracting information from this distribution that compelled the author (against his will) to repair his parlous knowledge of statistics [19]!

was preferred and announced the discovery of a new boson. Given the ubiquity of hypothesis testing, it is important to have a grasp of the methods that have been invented to implement it.

One method was due to Fisher [16], another was invented by Neyman, and a third (Bayesian) method was proposed by Sir Harold Jeffreys, all around the same time. Today, we tend to merge the approaches of Fisher and Neyman, and we hardly ever use the method of Jeffreys even though in several respects the method of Jeffreys and their modern variants are arguably more natural. In particle physics, we regard our Fisher/Neyman hybrid as sacrosanct, witness the near-religious adherence to the  $5\sigma$  discovery rule. However, the pioneers disagreed strongly with each other about how to test hypotheses, which suggests that the topic is considerably more subtle than it seems. We first describe the method of Fisher, then follow with a description of the method of Neyman. For concreteness, we consider the problem of deciding between a background-only model and a background + signal model.

### 3.1.3.1 Fisher's Approach

In Fisher's approach, we construct a **null hypothesis**, often denoted by  $H_0$ , and *reject* it should some measure be judged small enough to cast doubt on the validity of this hypothesis. In our example, the null hypothesis is the background-only model, for example, the SM without a Higgs boson. The measure is called a **p-value** and is defined by

$$\text{p-value}(x_0) = P(x > x_0 | H_0), \quad (19)$$

where  $x$  is a statistic designed so that large values indicate departure from the null hypothesis. This is illustrated in Fig. 9, which shows the location

of the observed value  $x_0$  of  $x$ . The p-value is the probability that  $x$  could have been higher than the  $x$  actually observed. It is argued that a small p-value implies that either the null hypothesis is false or something rare has occurred. If the p-value is extremely small, say  $\sim 3 \times 10^{-7}$ , then of the two possibilities the most common response is to presume the null to be false. If we apply this method to the  $D\bar{O}$  top quark discovery data, and neglect the uncertainty in null hypothesis, we find

$$\text{p-value} = \sum_{D=17}^{\infty} \text{Poisson}(D, 3.8) = 5.7 \times 10^{-7}.$$

In order to report a more intuitive number, the common practice is to map the p-value to the  $Z$  scale defined by

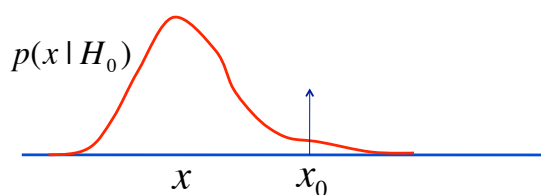
$$Z = \sqrt{2} \operatorname{erf}^{-1}(1 - 2\text{p-value}). \quad (20)$$

This is the number of Gaussian standard deviations away from the mean<sup>5</sup>. A p-value of  $5.7 \times 10^{-7}$  corresponds to a  $Z$  of  $4.9\sigma$ . The  $Z$ -value can be calculated using the Root function

$$Z = \text{-TMath::NormQuantile}(\text{p-value}).$$

### 3.1.3.2 Neyman's Approach

<sup>5</sup> $\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^x \exp(-t^2) dt$  is the error function.



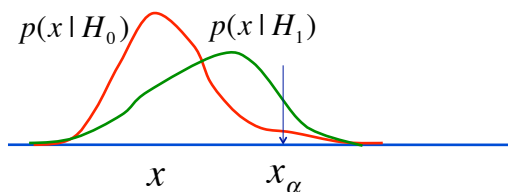
**Fig. 9:** The p-value is the tail-probability,  $P(x > x_0 | H_0)$ , calculated from the probability density under the null hypothesis,  $H_0$ . Consequently, the probability density of the p-value under the null hypothesis is  $\text{Uniform}(x, 1)$ .

In Neyman’s approach *two* hypotheses are considered, the null hypothesis  $H_0$  and an alternative hypothesis  $H_1$ . This is illustrated in Fig. 10. In our example, the null is the same as before but the alternative hypothesis is the SM with a Higgs boson. Again, one generally chooses  $x$  so that large values would cast doubt on the validity of  $H_0$ . However, the Neyman test is specifically designed to respect the frequentist principle, which is done as follows. A *fixed* probability  $\alpha$  is chosen, which corresponds to some threshold value  $x_\alpha$  defined by

$$\alpha = P(x > x_\alpha | H_0), \quad (21)$$

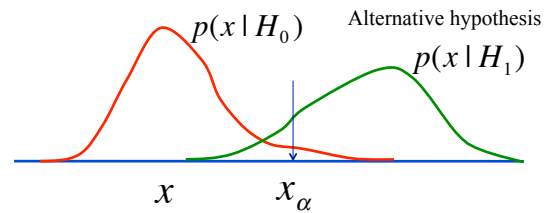
called the significance (or size) of the test. Should the observed value  $x_0 > x_\alpha$ , or equivalently,  $p\text{-value}(x_0) < \alpha$ , the hypothesis  $H_0$  is rejected in favor of the alternative. In particle physics, in addition to applying the Neyman hypothesis test, we also report the p-value. This is sensible because there is a more information in the p-value than merely reporting the fact that a null hypothesis was rejected at a significance level of  $\alpha$ .

The Neyman method satisfies the frequentist principle by construction. Since the significance of the test is fixed,  $\alpha$  is the relative frequency with which true null hypotheses would be rejected and is called the **Type I** error rate.



**Fig. 11:** See Fig. 10 for details. Unlike the case in Fig. 10, the two hypotheses  $H_0$  and  $H_1$  are not that different. It is then not clear whether it makes practical sense to reject  $H_0$  when  $x > x_\alpha$  only to replace it with an hypothesis  $H_1$  that is not much better.

simple hypotheses — that is, hypotheses in which all parameters have well-defined values — the optimal statistic  $t$  to use in the hypothesis test is the likelihood ratio  $t = p(x|H_1)/p(x|H_0)$ . Maximizing the power seems sensible. Consider Fig. 11. The significance of the test in this figure is the same as that in Fig. 10, so the Type I error rate is identical. However, the Type II error rate is much greater in Fig. 11 than in Fig. 10, that is, the power of the test is considerably weaker in the former. In that case, there may be no compelling reason to reject the null since the alternative is not that much better. This insight was one source of Neyman’s disagreement with Fisher. Neyman objected to possibility that one might reject a null hypothesis regardless of whether it made sense to do so. Neyman insisted that the task is always one of deciding between competing hypotheses. Fisher’s counter argument was that an alternative hypothesis may not be available, but we may nonetheless wish to know whether the only hypothesis that is available is worth keeping. As we shall see, the Bayesian approach also requires an alternative, in agreement with Neyman, but in a way that neither he nor Fisher agreed with!



**Fig. 10:** Distribution of a test statistic  $x$  for two hypotheses, the null  $H_0$  and the alternative  $H_1$ . In Neyman’s approach to testing,  $\alpha = P(x > x_\alpha | H_0)$  is a *fixed* probability called the significance of the test, which for a given class of experiments corresponds to the threshold  $x_\alpha$ . The hypothesis  $H_0$  is rejected if  $x > x_\alpha$ .

However, since we have specified an alternative hypothesis there is more that can be said. Figure 10 shows that we can also calculate

$$\beta = P(x \leq x_\alpha | H_1), \quad (22)$$

which is the relative frequency with which we would reject the hypothesis  $H_1$  if it is true. This mistake is called a Type II error. The quantity  $1 - \beta$  is called the **power** of the test and is the relative frequency with which we would accept the hypothesis  $H_1$  if it is true. Obviously, for a given  $\alpha$  we want to maximize the power. Indeed, this is the basis of the Neyman-Pearson lemma (see for example Ref. [2]), which asserts that given two

We have assumed that the hypotheses  $H_0$  and  $H_1$  are simple, that is, fully specified. Unfortunately, most of the hypotheses that arise in realistic particle physics analyses are not of this kind. In the Higgs boson discovery analyses by ATLAS and CMS the probability models depend on many nuisance parameters for which only estimates are available. Consequently, neither the background-only nor the background + signal hypotheses are fully specified. Such hypotheses are called **compound hypotheses**. In order to illustrate how hypothesis testing proceeds in this case, we again turn again to the top discovery example.

### Example

As we saw in Sec. 3.1.2, the standard way to handle nuisance parameters in the frequentist approach is to replace them by their conditional MLEs and thereby reduce the likelihood function to the profile likelihood. In the top discovery example, we obtain a function  $p_{PL}(D|s)$  that depends on the single parameter,  $s$ . We now treat this function as if it were a likelihood and invoke both the Neyman-Pearson lemma, which suggests the use of likelihood ratios, and Wilks' theorem to motivate the use of the function  $t(x, s)$  given in Eq. (17) to distinguish between two hypotheses: the hypothesis  $H_1$  in which  $s = \hat{s} = N - B$  and the hypothesis  $H_0$  in which  $s \neq \hat{s}$ , for example, the background-only hypothesis  $s = 0$ . In the context of testing,  $t(x, s)$  is called a **test statistic**, which, unlike a statistic as we have defined it (see Sec. 2.1), usually depends on at least one unknown parameter.

In principle, the next step is the computationally arduous task of simulating the distribution of the statistic  $t(x, s)$ . The task is arduous because *a priori* the probability density  $p(t|s, b)$  can depend on *all* the parameters that exist in the original likelihood. If this is really the case, then after all this effort we seem to have achieved a pyrrhic victory! But, this is where Wilks' theorem saves the day, at least approximately. We can avoid the burden of simulating  $t(x, s)$  because the latter is approximately a  $\chi^2$  variate.

Using  $N = 17$  and  $s = 0$ , we find  $t_0 = t(N = 17, s = 0) = 4.6$ . According to the results shown in Fig. (7)(a),  $N = 17$  may be considered "a lot of data"; therefore, we may use  $t_0$  to implement a hypothesis test by comparing  $t_0$  with a fixed value  $t_\alpha$  corresponding to the significance level  $\alpha$  of the test.

## 4 Lecture 3: The Bayesian Approach

In this lecture, we introduce the Bayesian approach to inference starting with a description of its salient features and ending with a detailed example, again using the top quark discovery data from DØ.

The main point to be understood about the Bayesian approach is that it is merely applied probability theory (see Sec. 2.2). A method is Bayesian if

- it is based on the degree of belief interpretation of probability and
- it uses Bayes theorem

$$p(\theta, \omega|D) = \frac{p(D|\theta, \omega) \pi(\theta, \omega)}{p(D)}, \quad (23)$$

where

$D$  = observed data,

$\theta$  = parameters of interest,

$\omega$  = nuisance parameters,

$p(\theta, \omega|D)$  = posterior density,

$\pi(\theta, \omega)$  = *prior density (or prior for short)*.



for *all* inferences. The result of a Bayesian inference is the posterior density  $p(\theta, \omega|D)$  from which, if desired, various summaries can be extracted. The parameters can be discrete or continuous and nuisance parameters are eliminated by marginalization,

$$\begin{aligned} p(\theta|D) &= \int p(\theta, \omega|D) d\omega, \\ &\propto \int p(D|\theta, \omega) \pi(\theta, \omega) d\omega. \end{aligned} \quad (24)$$

The function  $\pi(\theta, \omega)$ , called the prior, encodes whatever information we have about the parameters  $\theta$  and  $\omega$  independently of the data  $D$ . A key feature of the Bayesian approach is recursion; the use of the posterior density  $p(\theta, \omega|D)$  or one, or more, of its marginals as the prior in a subsequent analysis.

These simple rules yield an extremely powerful and general inference model. Why then is the Bayesian approach not more widely used in particle physics? The answer is partly historical: the frequentist approach was dominant at the dawn of particle physics. It is also partly the widespread perception that the Bayesian approach is too subjective to be useful for scientific work. However, there is published evidence that this view is mistaken, witness the success of Bayesian methods in high-profile analyses in particle physics such as the discovery of single top quark production at the Tevatron [20, 21].

#### 4.1 Model Selection

Conceptually, hypothesis testing in the Bayesian approach (also called model selection) proceeds exactly the same way as any other Bayesian calculation: we compute the posterior density,

$$p(\theta, \omega, H|D) = \frac{p(D|\theta, \omega, H) \pi(\theta, \omega, H)}{p(D)}, \quad (25)$$

and marginalize it with respect to all parameters except the ones that label the hypotheses or models,  $H$ ,

$$p(H|D) = \int p(\theta, \omega, H|D) d\theta d\omega. \quad (26)$$

Equation (26) is the probability of hypothesis  $H$  given the observed data  $D$ . In principle, the parameters  $\omega$  could also depend on  $H$ . For example, suppose that  $H$  labels different parton distribution function (PDF) models, say CT10, MSTW, and NNPDF, then  $\omega$  would indeed depend on the PDF model and should be written as  $\omega_H$ .

It is usually more convenient to arrive at the probability  $p(H|D)$  in stages.

1. Factorize the prior in the most convenient form,

$$\begin{aligned} \pi(\theta, \omega_H, H) &= \pi(\theta, \omega_H|H) \pi(H), \\ &= \pi(\theta|\omega_H, H) \pi(\omega_H|H) \pi(H), \end{aligned} \quad (27)$$

or

$$= \pi(\omega_H|\theta, H) \pi(\theta|H) \pi(H). \quad (28)$$

Often, we can assume that the parameters of interest  $\theta$  are independent, *a priori*, of both the nuisance parameters  $\omega_H$  and the model label  $H$ , in which case we can write,  $\pi(\theta, \omega_H, H) = \pi(\theta) \pi(\omega_H|H) \pi(H)$ .

2. Then, for each hypothesis,  $H$ , compute the function

$$p(D|H) = \int p(D|\theta, \omega_H, H) \pi(\theta, \omega_H|H) d\theta d\omega_H. \quad (29)$$

3. Then, compute the probability of each hypothesis,

$$p(H|D) = \frac{p(D|H) \pi(H)}{\sum_H p(D|H) \pi(H)}. \quad (30)$$

Clearly, in order to compute  $p(H|D)$  it is necessary to specify the priors  $\pi(\theta, \omega|H)$  and  $\pi(H)$ . With some effort, it is possible to arrive at an acceptable form for  $\pi(\theta, \omega|H)$ , however, it is highly unlikely that consensus could ever be reached on the discrete prior  $\pi(H)$ . At best, one may be able to adopt a convention. For example, if by convention two hypotheses  $H_0$  and  $H_1$  are to be regarded as equally likely, *a priori*, then it would make sense to assign  $\pi(H_0) = \pi(H_1) = 0.5$ .

One way to circumvent the specification of the prior  $\pi(H)$  is to compare the probabilities,

$$\frac{p(H_1|D)}{p(H_0|D)} = \left[ \frac{p(D|H_1)}{p(D|H_0)} \right] \frac{\pi(H_1)}{\pi(H_0)}. \quad (31)$$

and use only the term in brackets, called the global **Bayes factor**,  $B_{10}$ , as a way to compare hypotheses. The Bayes factor specifies by how much the relative probabilities of two hypotheses changes as a result of incorporating new data,  $D$ . The word global indicates that we have marginalized over all the parameters of the two models. The *local* Bayes factor,  $B_{10}(\theta)$  is defined by

$$B_{10}(\theta) = \frac{p(D|\theta, H_1)}{p(D|\theta, H_0)}, \quad (32)$$

where,

$$p(D|\theta, H_1) \equiv \int p(D|\theta, \omega_{H_1}, H_1) \pi(\omega_{H_1}|H_1) d\omega_{H_1}, \quad (33)$$

are the **marginal** or integrated likelihoods in which we have assumed the *a priori* independence of  $\theta$  and  $\omega_{H_1}$ . We have further assumed that the marginal likelihood  $H_0$  is independent of  $\theta$ , which is a very common situation. For example,  $\theta$  could be the expected signal count  $s$ , while  $\omega_{H_1} = \omega$  could be the expected background  $b$ . In this case, the hypothesis  $H_0$  is a special case of  $H_1$ , namely, it is the same as  $H_1$  with  $s = 0$ . An hypothesis that is a special case of another is said to be **nested** in the more general hypothesis. The Bayesian example, discussed below, will make this clearer. There is a subtlety that may be missed: because of the way we have defined  $p(D|\theta, H)$ , we need to multiply  $p(D|\theta, H)$  by the prior  $\pi(\theta)$  and then integrate with respect to  $\theta$  in order to calculate  $p(D|H)$ .

#### 4.1.1 A Word About Priors

Constructing a prior for nuisance parameters is generally neither controversial (for most parameters) nor problematic. Such difficulties as do arise occur when the priors must, of necessity, depend on expert judgement. For example, one theorist may insist that a uniform prior within a finite interval is a reasonable prior for the factorization scale in a QCD calculation, while in the expert judgement of another the interval should be twice as large. Clearly, in this case, there is no getting around the fact that the prior for this parameter is unavoidably subjective. However, once a choice is made, a prior  $\pi(\omega_H|H)$  that integrates to one can be constructed.

The Achilles heal of the Bayesian approach is the need to specify the prior  $\pi(\theta)$ , for the parameters of interest, at the start of the inference chain when we know almost nothing about these parameters. Careless specification of this prior can yield results that are unreliable or even nonsensical. The mandatory requirement is that the posterior density be proper, that is integrate to unity. Ideally, the same should hold for priors. A very extensive literature exists on the topic of prior specification when the available information is extremely limited. However, a discussion of this topic is beyond the scope of these lectures; but, we shall make a few remarks.

For model selection, we need to proceed with caution because Bayes factors are sensitive to the choice of priors and therefore less robust than posterior densities. Suppose that the prior  $\pi(\theta) = Cf(\theta)$ , where  $C$  is a normalization constant. The global Bayes factor for the two hypotheses  $H_1$  and  $H_0$  can be written as

$$B_{10} = C \frac{\int p(D|\theta, H_1) f(\theta) d\theta}{p(D|H_0)}. \tag{34}$$

Therefore, if the constant  $C$  is ill defined, typically because  $\int f(\theta) d\theta = \infty$ , the Bayes factor will likewise be ill defined. For this reason, it is generally recommended that an improper prior not be used for parameters  $\theta$  that occur only in one hypothesis, here  $H_1$ . However, for parameters that are common to all hypotheses, it is permissible to use improper priors because the ill defined constant cancels in the Bayes factor.

The discussion so far has been somewhat abstract. The next section therefore works through a detailed example of a possible Bayesian analysis of the  $D\bar{O}$  top discovery data.

### 4.2 The Top Quark Discovery: A Bayesian Analysis

In this section we shall perform the following calculations as a way to illustrate a typical Bayesian analysis,

1. compute the posterior density  $p(s|D)$ ,
2. compute a 68% credible interval  $[l(D), u(D)]$ , and
3. compute the global Bayes factor  $B_{10} = p(D|H_1)/p(D|H_0)$ .

#### Probability model

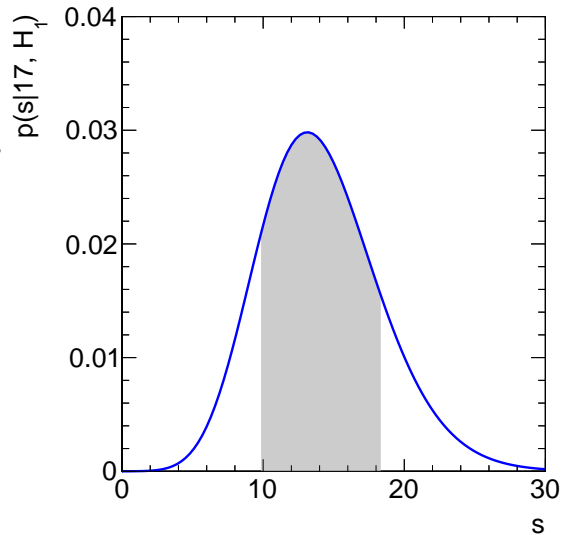
The first step in any serious statistical analysis is to think deeply about what has been done in the physics analysis; for example, to trace in detail the steps that led to the background estimates, determine the independent systematic effects and identify explicitly what is known about them. Although, by tradition, we tend to think of potential data  $x$  separately from the parameters  $s$  and  $b$ , it should be recognized that this is done for convenience. The full probability model is the joint probability

$$p(x, s, b|I),$$

which, as is true of *all* probability models, is conditional on the information and assumptions,  $I$ , that define the abstract space  $\Omega$  (see Sec. 2.2). In these lectures, we have omitted the conditioning data  $I$ , and will continue to do so here, but it should not be forgotten that it is always present and may differ from one probability model to another.

The full probability model  $p(x, s, b)$  can be factorized in several ways, all of which are mathematically valid. However, we find it convenient to factorize the model in the following way

$$p(x, s, b) = p(x|s, b) \pi(s, b), \tag{35}$$



**Fig. 12:** Posterior density computed for  $D\bar{O}$  top quark discovery data. The shaded area is the 68% central credible interval.

where we have introduced the symbol  $\pi$  in order to highlight the distinction we choose to make between this part of the model and the remainder. We are entirely free to decide how much of the model we place in  $p(x|s, b)$  and how much in  $\pi(s, b)$ ; what matters is the form of the full model  $p(x, s, b)$ . In the frequentist analysis of the top quark discovery data, we took  $N$  and  $B$  to be the data  $D$ . We did so because in the frequentist approach, the function  $\pi(s, b)$  does not exist and consequently we have no choice but to include everything in the function  $p(x|s, b)$ . One virtue of a Bayesian perspective is that we are not bound by this stricture. To make the point explicitly, we take the probability distribution,  $p(x|s, b)$ , to be

$$p(x|s, b) = \text{Poisson}(x, s + b). \quad (36)$$

The interpretation of  $p(x|s, b)$  is clear: it is the probability to observe  $x$  events *given* that the mean event count is  $s + b$ . What does  $\pi(s, b)$  represent? This function is the **prior** that encodes what we *know*, or *assume*, about the mean background and signal independently of the potential observations  $x$ . The prior  $\pi(s, b)$  can be factored in two ways,

$$\begin{aligned} \pi(s, b) &= \pi(s|b) \pi(b), \\ &= \pi(b|s) \pi(s), \end{aligned} \quad (37)$$

both of which accord with the probability rules. The factorizations remind us that the parameters  $s$  and  $b$  may not be probabilistically independent. However, we shall assume that they are, at least at this stage of the analysis, in which case it is permissible to write,

$$\pi(s, b) = \pi(s) \pi(b). \quad (38)$$

We first consider the background prior  $\pi(b)$  and ask: what do we know about the background? We know the count  $Q$  in the control region and we have an estimate of the control region to signal region scale factor  $k$ . The likelihood for  $Q$  is taken to be

$$p(Q|k, b) = \text{Poisson}(Q, kb), \quad (39)$$

from which, together with a prior  $\pi(k, b)$ , we can compute the posterior density

$$p(b|Q, k) = p(Q|k, b) \pi(k, b) / p(Q). \quad (40)$$

As usual, we factorize the prior,  $\pi(k, b) = \pi(k|b)\pi_0(b)$ , where we have introduced the subscript 0 to distinguish  $\pi_0(b)$  from the background prior associated with Eq. (36). Then, we consider the separate factors  $\pi_0(b)$  and  $\pi(k|b)$ .

What do we know about  $b$  at this stage? Clearly,  $b \geq 0$ . But, that is all we know apart from the background likelihood, Eq. (39). Today, after a century of argument and discussion, the consensus amongst statisticians is that there is no unique way to represent such vague information. However, well founded ways to construct such priors are available, see for example Ref. [22] and references therein; but for simplicity we take the prior  $\pi_0(b) = 1$ , that is, the **flat prior**. If the uncertainty in  $k$  can be neglected, the (proper!) prior for  $k$  is  $\pi(k|b) = \delta(k - Q/B)$ , which amounts to replacing  $k$  in Eq. (40) by  $Q/B$ . When the dust settles, we find

$$p(b|Q, k) = \text{Gamma}(kb, 1, Q + 1) = \frac{e^{-kb}(kb)^Q}{\Gamma(Q + 1)}, \quad (41)$$

for the posterior density of  $b$ , which can serve as the prior  $\pi(b)$  associated with Eq. (36).

By construction,  $p(x, s, b)$  is identical in form to the likelihood in Eq. (15); we have simply availed ourselves of the freedom to factorize  $p(x, s, b)$  as we wish and therefore to reinterpret the factors. This

freedom is useful because it makes it possible to keep the likelihood simple while relegating the complexity to the prior. This may not seem, at first, to be terribly helpful; after all, we arrived at the same mathematical form as Eq. (15). However, the complexity can be substantially mitigated through the numerical treatment of the prior, as discussed at the end of the next section. The likelihood, as we have conceptualized the problem, is given by

$$p(D|s, b) = \frac{e^{-(s+b)}(s+b)^D}{D!}, \quad (42)$$

where  $D = 17$  events.

The final ingredient is the prior  $\pi(s)$ . At this stage, all we know is that  $s \geq 0$ . Again, there is no unique way to specify  $\pi(s)$ , though as noted there are well founded methods to construct it. We shall variously assume either the improper prior  $\pi(s) = 1$  or the proper prior  $\pi(s) = \delta(s - 14)$ .

### **Marginal likelihood**

After this somewhat discursive discussion of the probability model, we have done the hard part: building the full probability model. Hereafter, the rest of the Bayesian analysis is mere computation.

It is convenient to eliminate the nuisance parameter  $b$ ,

$$\begin{aligned} p(D|s, H_1) &= \int_0^\infty p(D|s, b) \pi(b) d(kb), \\ &= \frac{1}{Q} (1-x)^2 \sum_{r=0}^N \text{Beta}(x, r+1, Q) \text{Poisson}(N-r|s), \end{aligned} \quad (43)$$

where  $x = 1/(1+k)$ ,

**Exercise 10:** Show this

and thereby arrive at the marginal likelihood  $p(D|s, H_1)$ . This example, the **Poisson-gamma** model is particularly simple and lends itself to exact calculation. However, the complexity rapidly increases as the prior becomes more and more complicated. In the probability model that is used in the Higgs boson analyses at the LHC, the part we would consider the prior,  $\pi(\mu, m_H, \omega)$ , is of enormous complexity. However, the part that we would call the likelihood,  $p(D|\mu, m_H, \omega)$ , is relatively simple. The parameter  $\mu$  denotes one or more signal strengths — the ratio of the cross section times branching fraction to that predicted by the Standard Model (SM), and  $m_H$  is the Higgs boson mass. The parameter  $\omega$  represent the expected (and therefore unknown) SM signal predictions and the expected backgrounds. When faced with such complexity, it proves useful to use a **hierarchical Bayesian model**. Briefly, the prior  $\pi(\mu, m_H, \omega)$  is written as

$$\begin{aligned} \pi(\mu, m_H, \omega) &= \pi(\omega|\mu, m_H) \pi(\mu, m_H), \\ \text{where } \pi(\omega|\mu, m_H) &= \int \pi(\omega|\phi, \mu, m_H) \pi(\phi|\mu, m_H) d\phi. \end{aligned}$$

The prior  $\pi(\phi|\mu, m_H)$  models the lowest level systematic parameters that define quantities such as the jet energy scale, lepton efficiencies, trigger efficiencies, and the parton distribution functions. It is usually straightforward to sample from this prior. Moreover, the function  $\pi(\omega|\phi, \mu, m_H)$  is nothing more than prior for the expected signal and background parameters  $\omega$ , which through estimates  $\hat{\omega}$  depend implicitly on the parameters  $\phi$ . The prior  $\pi(\omega|\phi, \mu, m_H)$  is generally quite simple; for binned data it is just a product of gamma (or gamma mixture) densities; more generally, it is a product of gamma, Gaussian, or log-normal densities. Consequently, the marginalizations over  $\omega$  can be done in two steps: first generate

a point  $\phi_i$  from  $\pi(\phi|\mu, m_H)$ , then generate a point  $\omega_i$  from  $\pi(\omega|\phi_i, \mu, m_H)$ . In that way, the enormous complexity of explicitly modeling the dependence of  $\omega$  on  $\phi$  is avoided, with the added benefit that all, possibly very complicated, correlations (in principle, to all orders) are accounted for automatically. The marginal likelihood can be approximated by

$$p(D|\mu, m_H) \approx \frac{1}{M} \sum_{m=1}^M p(D|\mu, m_H, \omega_m). \quad (44)$$

What we have just described is merely integration via a Monte Carlo approximation. The point is that the sampling required to compute  $p(D|\mu, m_H)$  can be run in  $M$  parallel analysis jobs, each of which is given a different random number seed in order to sample a single pair of points  $\phi_m$  and  $\omega_m$ . The results of such a Bayesian analysis would be the likelihood  $p(D|\mu, m_H, \omega)$  and an ensemble of points  $\{\omega_m\}$ .

**Posterior density**

Given the marginal likelihood  $p(D|s, H_1)$  and a prior  $\pi(s)$  we can compute the posterior density,

$$p(s|D, H_1) = p(D|s, H_1) \pi(s) / p(D|H_1), \quad (45)$$

where,

$$p(D|H_1) = \int_0^\infty p(D|s, H_1) \pi(s) ds.$$

Again, for simplicity, we assume a flat prior for the signal,  $\pi(s) = 1$  and find

$$p(s|D, H_1) = \frac{\sum_{r=0}^N \text{Beta}(x, r + 1, Q) \text{Poisson}(N - r|s)}{\sum_{r=0}^N \text{Beta}(x, r + 1, Q)}, \quad (46)$$

**Exercise 11:** Derive an expression for  $p(s|D, H_1)$  assuming  $\pi(s) = \text{Gamma}(qs, 1, M + 1)$  where  $q$  and  $M$  are constants

from which we can compute the central **credible interval** [9.9, 18.4] for  $s$  at 68% C.L., which is shown in Fig. 12.

**4.2.1 Bayes factor**

As noted, the number  $p(D|H_1)$  can be used to perform a hypothesis test. But, as argued above, we need to use a proper prior for the signal, that is, a prior that integrates to one. The simplest such prior is a  $\delta$ -function, e.g.,  $\pi(s) = \delta(s - 14)$ . Using this prior, we find

$$p(D|H_1) = p(D|14, H_1) = 9.28 \times 10^{-2}.$$

Since the background-only hypothesis  $H_0$  is nested in  $H_1$ , and defined by  $s = 0$ , the number  $p(D|H_0)$  is given by  $p(D|0, H_1)$ , which yields

$$p(D|H_0) = p(D|0, H_1) = 3.86 \times 10^{-6}.$$

We conclude that the hypothesis  $s = 14$  is favored over  $s = 0$  by a Bayes factor of 24,000. In order to avoid large numbers, the Bayes factor can be mapped into a (signed) measure akin to the frequentist "n-sigma" [23],

$$Z = \text{sign}(\ln B_{10}) \sqrt{2|\ln B_{10}|}, \quad (47)$$

which gives  $Z = 4.5$ . Negative values of  $Z$  correspond to hypotheses that are excluded.

## Summary

These lectures gave an overview of the main ideas of statistical inference in a form directly applicable to statistical analysis in particle physics. Two widely used approaches were covered, frequentist and Bayesian. While we tried to focus on the practical, our hope is that we have given just enough commentary about the topics to place them in some intellectual context. We hope that the take away message is that it is worth learning a bit more about statistics if only to avoid fruitless arguments and discussions with co-workers. Statistics is not physics. Nature is the ultimate arbiter of which physics ideas are “correct”. Unfortunately, the ultimate arbiter of statistical ideas, apart from the mundanity of mathematical correctness, is intellectual taste. Therefore, the other take home message is

“Have the courage to you use your own understanding”

Immanuel Kant

## Acknowledgement

I thank Nick Ellis, Martijn Mulders, Kate Ross, and their counterparts from JINR, for organizing and hosting a very enjoyable school, and the students for their keen participation and youthful enthusiasm. These lectures were supported in part by US Department of Energy grant DE-FG02-13ER41942.

## References

- [1] S. K. Chatterjee, *Statistical Thought: A Perspective and History*, Oxford University Press, Oxford (2003).
- [2] F. James, *Statistical Methods in Experimental Physics*, 2nd Edition, World Scientific, Singapore (2006).
- [3] G. Cowan, *Statistical Data Analysis*, Oxford University Press, Oxford (1998).
- [4] R. J. Barlow, *Statistics: A Guide To The Use Of Statistical Methods In The Physical Sciences*, The Manchester Physics Series, John Wiley and Sons, New York (1989).
- [5] G. Taraldsen and B.H. Lindqvist, “Improper Priors Are Not Improper,” *The American Statistician*, Vol. 64, Issue 2, 154 (2010).
- [6] L. Daston, “How Probability Came To Be Objective And Subjective,” *Hist. Math.* 21, 330 (1994).
- [7] F. Abe *et al.* [CDF Collaboration], “Observation of top quark production in  $\bar{p}p$  collisions,” *Phys. Rev. Lett.* **74**, 2626 (1995) [hep-ex/9503002].
- [8] S. Abachi *et al.* [D0 Collaboration], “Observation of the top quark,” *Phys. Rev. Lett.* **74**, 2632 (1995) [hep-ex/9503003].
- [9] S. Chatrchyan *et al.* [CMS Collaboration], “Search for contact interactions using the inclusive jet  $p_T$  spectrum in  $pp$  collisions at  $\sqrt{s} = 7$  TeV,” *Phys. Rev. D* **87**, 052017 (2013) [arXiv:1301.5023 [hep-ex]].
- [10] N. Suzuki, D. Rubin, C. Lidman, G. Aldering, R. Amanullah, K. Barbary, L. F. Barrientos and J. Botyanszki *et al.*, “The Hubble Space Telescope Cluster Supernova Survey: V. Improving the Dark Energy Constraints Above  $z > 1$  and Building an Early-Type-Hosted Supernova Sample,” *Astrophys. J.* **746**, 85 (2012) [arXiv:1105.3470 [astro-ph.CO]].
- [11] R. Dungan and H. B. Prosper, “Varying-G Cosmology with Type Ia Supernovae,” arXiv:0909.5416 [astro-ph.CO].
- [12] G. Aad *et al.* [ATLAS Collaboration], “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys. Lett. B* **716**, 1 (2012) [arXiv:1207.7214 [hep-ex]].
- [13] S. Chatrchyan *et al.* [CMS Collaboration], “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett. B* **716**, 30 (2012) [arXiv:1207.7235 [hep-ex]].

- [14] J. Neyman, "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability," *Phil. Trans. R. Soc. London A236*, 333 (1937).
- [15] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals," *Phys. Rev. D57*, 3873 (1998).
- [16] S. E. Fienberg and D. V. Hinkley, eds., *R.A. Fisher: An Appreciation*, Lecture Notes on Statistics, Volume 1, Springer Verlag (1990).
- [17] W. Verkerke and D. Kirkby, RooFit, <http://roofit.sourceforge.net>.
- [18] K. Cranmer, G. Schott, L. Moneta and W. Verkerke, RooStats, <https://twiki.cern.ch/twiki/bin/view/RooStats>
- [19] G. Fidecaro *et al.* [CERN-Rutherford-ILL-Sussex-Padua (CRISP) Collaboration], "Experimental Search For Neutron Anti-neutron Transitions With Free Neutrons," *Phys. Lett. B* **156**, 122 (1985).
- [20] V. M. Abazov *et al.* [D0 Collaboration], "Observation of Single Top Quark Production," *Phys. Rev. Lett.* **103**, 092001 (2009) [arXiv:0903.0850 [hep-ex]].
- [21] T. Aaltonen *et al.* [CDF Collaboration], "First Observation of Electroweak Single Top Quark Production," *Phys. Rev. Lett.* **103**, 092002 (2009) [arXiv:0903.0885 [hep-ex]].
- [22] L. Demortier, S. Jain and H. B. Prosper, "Reference priors for high energy physics," *Phys. Rev. D* **82**, 034002 (2010) [arXiv:1002.1111 [stat.AP]].
- [23] S. Sekmen *et al.*, "Phenomenological MSSM interpretation of the CMS 2011 5fb-1 results," CMS Physics Analysis Summary, CMS-PAS-SUS-12-030, CERN (2012).



## Organizing Committee

T. Donskova (Schools Administrator, JINR)  
N. Ellis (CERN)  
H. Haller (Schools Administrator, CERN)  
M. Mulders (CERN)  
A. Olchevsky (JINR)  
G. Perez (CERN and Weizmann Institute, Israel)  
K. Ross (Schools Administrator, CERN)

## Local Organizing Committee

C. Hajdu (Wigner RCP, Hungary)  
G. Hamar (Wigner RCP, Hungary)  
D. Horváth (Wigner RCP, Hungary)  
J. Karacsi (Univ. Debrecen, Hungary)  
P. Lévai (Wigner RCP, Hungary)  
F. Silklér (Wigner RCP, Hungary)  
B. Ujvári (Univ. Debrecen, Hungary)  
V. Veszprémi (Wigner RCP, Hungary)  
A. Zsigmond (Wigner RCP, Hungary)

## Lecturers

E. Boos (Moscow State University, Russia)  
C. Csaki (Cornell University, USA)  
J. Ellis (King's College London, UK and CERN)  
B. Gavela (Univ. Autonoma and IFT UAM/CSIC, Spain)  
D. Gorbunov (INR, Russia)  
Z. Ligeti (Berkeley, USA)  
H. Prosper (Florida State University, USA)  
K. Rajagopal (MIT, USA and CERN)  
Z. Trócsányi (Univ. Debrecen, Hungary)  
T. Virdee (Imperial College London, UK)

## Discussion Leaders

A. Arbuzov (JINR)  
T. Biro (Wigner RCP, Hungary)  
M. Blanke (CERN)  
G. Cynolter (Roland Eotvos Univ., Hungary)  
A. De Simone (SISSA, Italy and CERN)  
A. Gladyshev (JINR)

## Students

Ehab ABBAS	Ahmed A HASIB	Bedrich ROSKOVEC
Rosemarie ABEN	Veerle HEIJNE	Nils RUTHMANN
Vincent ANDRIEUX	Kristin HEINE	Arseniy RYBNIKOV
Nikolay ANFIMOV	Gregor HELLWIG	Antonio SALVUCCI
Spyridon ARGYROPOULOS	Geoffrey HERBERT	Ruth SANDBACH
Ivan ASIN CRUZ	Rebekka HOEING	James SAXON
Vadim BABKIN	Julia HOFMANN	Yan-Jie SCHNELLBACH
Petr BALEK	Tova HOLMES	Martin SCHULTENS
Ahmed BASSALAT	Jacob HOWARD	Federico SCUTTI
Andrey BELYAEV	Diedi HU	Claudia SEITZ
Gyula BENCEDI	Artem IVANOV	Hale SERT
Joram BERGER	Manuel KAMBEITZ	Paul SEYFERT
Aurthem BORODENKO	Bastian KARGOLL	Edmund SMITH
Christian BÖSER	Jure KLUCAR	Livia SOFFI
Jessica BRINSON	Lucy KOGAN	Johannes STILLER
Alexander BYLINKIN	Ilse KRÄTSCHMER	Antonella SUCCURRO
Stefano CASASSO	Lawrence LEE	Maximilian SWIATLOWSKI
Veronika CHOBANOVA	Guillaume LEFEBVRE	Florian THIBAUD
Bonnie CHOW	Bing LI	Elodie TIOUCHICHINE
Predrag CIRKOVIC	Jessica LIEBAL	Emrah TIRAS
Claudio CORTI	Miaoyuan LIU	Marco TRESCH
Mykhailo DALCHENKO	Jonathan David LONG	Monica TROVATELLI
Aurélien DEMILLY	Joern MAHLSTEDT	Alexander TUNA
Charles DIETZ	Krisztina MARTON	Mark TURNER
Agnieszka DZIURDA	Michal MEREŠ	Karoly UERMOESSY
Andrew EDMONDS	Francesco MICHELI	Michael UGHETTO
Julia FISCHER	Swagata MUKHERJEE	Francesca UNGARO
Stefan GADATSCH	Marion MURUMAA	Anna USANOVA
Bruno GALHARDO	Benjamin NACHMAN	Nika VALENCIC
Lucie GAUTHIER	Eduardo NAVARRO DE MARTINO	Gagik VARDANYAN
Claudia GIULIANI	Thomas NEEP	Tobias VERLAGE
Maximilian GOBLIRSCH-KOLB	Francesco NUTI	Mario VORMSTEIN
Reza GOLDOUZIAN	Katherine PACHAL	Christian VOSS
Ed GREENING	Priscilla PANI	Elena YATSENKO
Zara GROUT	Lucia PERRINI	Maria ZANGOLI
Emine GURPINAR	Ivan RAZUMOV	Radek ZLEBCEK
Marion HABIBY ALAOUI	Emanuele RIPICCINI	Anna ZSIGMOND
Jon HARRISON	Elmar RITSCH	Giovanni ZURZOLO

## Posters

Author	Poster title
ANDRIEUX, V.; IVANOV, A.	Investigation into the nucleon spin structure
ARGYROPOULOS, S.	Studying jets in the presence of pile-up with the ATLAS detector
ASIN CRUZ, I.	Measurement of Normalized Differential $t\bar{t}$ Production Cross-Section
BABKIN, V.	MRPC for the Time-of-Flight System of the MPD/NICA project.
BASSALAT, A.	Testing and Ranking the Pixel IBL- Staves
BERGER, J.	Calibration of the Jet Energy Scale with $Z(\rightarrow \mu\mu) + \text{Jet}$ Events
BYLINKIN, A.	A new approach to describe hadroproduction in high energy particle collisions
BÖSER, C.	Search for the Higgs-like boson decaying into bottom quarks in the WH channel
CIRKOVIC, P.	B-flavour Charge Tagging in CP Violation Measurements in ATLAS Experiment
CORTI, C.	The proton rejection with the ECAL in AMS-02
DEMILLY, A.	Top Quark Mass Measurement in the multi lepton $e\mu$ channel with the Matrix Element Method with the ATLAS detector at the LHC
EDMONDS, A.	A Search for Charged Lepton Flavour Violation with COMET
GADATSCH, S.	Combined Measurements of the couplings of the Higgs boson in ATLAS
GALHARDO, B.	Search for the rare top-quark decay $t \rightarrow Zq$
GAUTHIER, L.	Search for supersymmetry with the Razor variables, multijet final state, with the CMS experiment at 8TeV
GIULIANI, C.	Search for direct stop production in final states with jets and missing transverse momentum with the ATLAS detector
GOBLIRSCH-KOLB, M.	Search for R-Parity violating Supersymmetry in multilepton events with the ATLAS detector
GREENING, E.	Search for $D(s)^+ \rightarrow \pi^+\mu^+\mu^-$ and $D(s)^+ \rightarrow \pi\mu^+\mu^+$ decays

Author	Poster title
GROUT, Z.	R-parity Violating Supersymmetry Search in Four Lepton Events at ATLAS
HARRISON, J.	Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb
HASIB, A.	Top quark pair production cross section measurement in tau + jets final state
HEINE, K.	Search for new physics in the multijet and missing transverse momentum final state
HELLWIG, G.	Light NMSSM Higgs Boson Search in bb Final States with the CMS Experiment
HOFMANN, J.; LIU M.	Measurement of Wyy final states with ATLAS
HU, D.	Search for pair production of vector-like bottom quarks in the lepton-plus-jets final state at $\sqrt{s}=8$ TeV with the ATLAS detector
KARGOLL, B.	Higgs searches based on hadronic tau triggers at the CMS experiment
KOGAN, L.	The Jet Energy Scale uncertainty of large R jets from gamma jet events
KRÄTSCHMER, I.	Measurement of Prompt Psi(2S) Polarization With CMS
LEE, L.; SWIATLOWSKI, M.	Quark vs Gluon Jet Discrimination in ATLAS
LEFEBVRE, G.	Jet Calibration in ATLAS
LONG, J. D.	Measurements of Higgs Boson Production Rates in the WW* $\rightarrow l\nu l\nu$ Channel in ATLAS
MAHLSTEDT, J.	Limit Setting on new BSM models
MUKHERJEE, S.	Electron to Photon Fakerate measurement from data collected by CMS detector
NACHMAN, B.	Top Squark Pair Production Top + LSP Decay, Single Lepton Final State
NEEP, T.	Measurements of spin correlation in ttbar events using the ATLAS detector
PANI, P.	Search for a heavy SUSY partner of the top quark decaying to a chargino and a b-jet in ATLAS

<b>Author</b>	<b>Poster title</b>
PERRINI, L.	Search for the Standard Model Scalar boson decaying to tau pairs produced in association with a Z boson
RIPICCINI, E.	Active Target
RUTHMANN, N.	Search for a SM Higgs Boson in the decay channel H-tau lep tau had with ATLAS at the LHC
SALVUCCI, A.	The new Higgs-like particle in the H→ZZ→4l searches with the ATLAS detector
SEITZ, C.	Search for Light- and Heavy-flavor Three-jet Resonances in Multijet Final States with the CMS Detector at sqrt(s) = 8 TeV
SERT, H.	Light Higgsino Precision Measurement at the ILC
SEYFERT, P.	Search for Lepton Flavour Violation in $\tau \rightarrow \mu\mu\mu$ in LHCb
SMITH, E.	CP Violation in Bd to DK* decays
SOFFI, L.	Use of ECAL Time in physics analyses at CMS
STILLER, J.	Full Reconstruction of B Mesons with the ALICE Inner Tracker Upgrade
SUCCURRO, A.	Search for pair-produced vector-like quarks with the ATLAS detector
USANOVA, A.	Angular Analysis of $B_d \rightarrow K^*0\mu\mu$ with the ATLAS experiment
VARDANYAN, G.	Jet energy calibration and inclusive jet cross-section measurement with the ATLAS detector.
VORMSTEIN, M.	The NA62 Experiment and Simulation of the Muon Veto Detector
YATSENKO, E.	Drell Yan production measurements and constrains on proton structure
ZANGOLI, M.	Direct CP-asymmetries in $B^0 \rightarrow K\pi$ and $B_s^0 \rightarrow \pi K$ decays at LHCb
ZSIGMOND, A. J.	Z boson production in PbPb collisions