ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE

CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

2005 European School of High-Energy Physics

Kitzbühel, Austria
21 August–3 September 2005

Proceedings

Editor: R. Fleischer

Geneva
2006

Abstract

The European School of High-Energy Physics is intended to give young experimental physicists an introduction to the theoretical aspects of recent advances in elementary particle physics. These proceedings contain lectures notes on field theory and the Standard Model, quantum chromodynamics, flavour physics and CP violation, experimental aspects of CP violation in K and B decays, relativistic heavy-ion physics, and the scientific programme of the Joint Institute for Nuclear Research. These core scientific topics are complemented by a lecture about the physics of ski jumping.

Preface

Ninety-seven students coming from thirty-three different countries attended the thirteenth in the new series of the European School of High-Energy Physics which took place in Kitzbühel, Austria, from 21 August to 3 September 2005. The School was hosted in the beautiful Hotel Kitzhof, a Tyrolean style hotel, two hundred metres from the city centre. According to the tradition of the School, the students shared twin rooms, mixing nationalities and in particular Eastern participants with Western ones.

Emmerich Kneringer from the University of Innsbruck was the local director of the School, co-chaired by Laurenz Widhalm from the Austrian Academy of Sciences, Vienna, with Manfred Jeitler, also from Vienna, as the third member of the local committee. In addition Andreas Salzburger acted as a very efficient 'assistant' to the local team.

Meinhart Regler together with Dietmar Kuhn helped in the initial phases of setting up the School and in providing some of the funding.

Mr Otto Langer from the Kitzbühel Rotary Club helped to bring the School to Kitzbühel and provided important links to the town authorities.

Our thanks go to the very efficient local team for their help and assistance and whose efforts contributed in major ways to the success of the School.

Our thanks are also 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 their personal contribution in answering questions and explaining points of theory.

We are very grateful to Danielle Métral and Tatyana Donskova for their untiring efforts in the lengthy preparations for and the day-to-day care of the School. Their efficient teamwork and continuous care of the students and their needs were highly appreciated. Our special thanks also go to the hotel manager, Markus Frischknecht, and to Susan and the hotel staff who were always ready to assist the school participants in a most friendly manner.

The participants spent their free time hiking in the impressive mountains around Kitzbühel or swimming in the Schwartzee, a few excursions were also organized. The first Wednesday afternoon featured a guided tour of Kitzbühel. The tour, which was organized as a travel in time, was very interesting allowing the participants to learn about the region's history and the beautiful architecture of the town. Saturday saw a visit to Salzburg. The city of Salzburg—and especially its historic city centre—is in fact one of the loveliest places in Europe, winning international acclaim in 1997 when it was designated a World Heritage site by UNESCO.

Thanks go to the Austrian Academy of Sciences who sponsored the School and to Professor Rudi Grimm, Dean of the University of Innsbruck, who offered the welcome drink. Our thanks are also due to Andreas and his team who had set up and assisted in the running of an excellent PC centre for the School.

However, the success of the 2005 School was to a large extent due to the students themselves. Their posters set up in and around the auditorium were of excellent quality both technically and in content, and throughout the School they participated actively during the lectures, in the discussion sessions, and with genuine interest in the different activities and excursions.

Egil Lillestøl
on behalf of the Organizing Committee
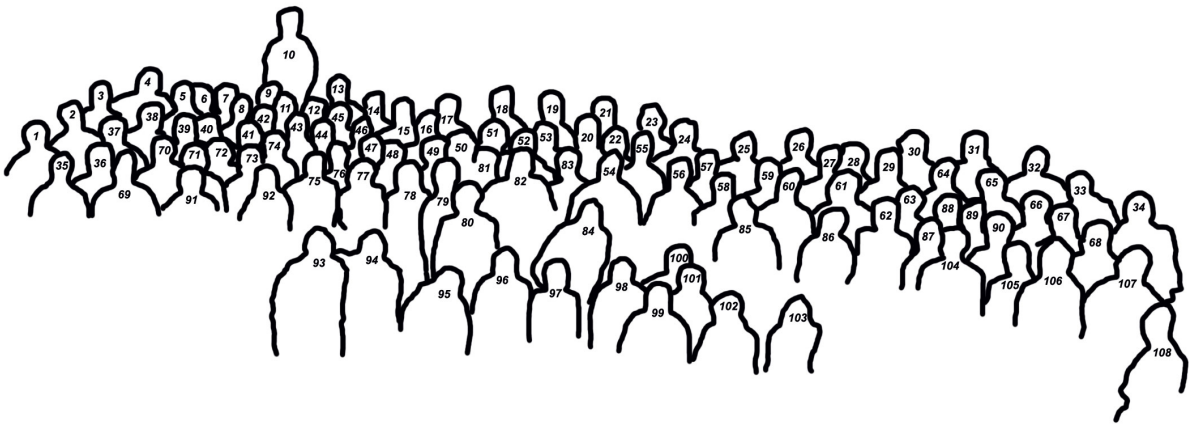
People in the photograph

Name	Number	Name	Number
Cyril Adamuscin	43	Stephan Huber	89
Cristina Adorisio	102	Alejandro Ibarra	8
Paolo Adragna	62	Vincenzo Izzo	75
Philip Allfrey	19	Christian Jacoby	17
Thomas Atkinson	26	Manfred Jeitler	4
Giuseppe Avolio	99	Christopher Jung	90
Konstanstin Babich	57	Vojtech Juranek	11
Mark Bell	81	Patrick Jussel	21
Michele Bianco	40	Nicolas Kerschen	59
Marc-Oliver Boenig	50	Graham Kilvington	9
Richard Brauer	52	Thomas Kittelmann	25
Nicholas Brett	18	Emmerich Kneringer	104
Wilfried Buchmüller	2	Rocky Kolb	107
Philippe Calfayan	1	Jens Konrath	15
Nuno Filipe Castro	82	Truls Martin Larsen	32
Teh Lee Cheng	79	Alfio Lazzaro	106
Yaw Ming Chia	10	Sabina Lehocka	76
Claudio Chiri	38	Debora Leone	92
Christian Holm Christensen	30	Michael Rudolf Lerchster	20
Thijs Cornelissen	7	Egil Lillestøl	94
Jonathan Cox	67	Phillip Litchfield	68
Sergei Demidov	91	Christoph Luedeling	22
Daniel Dobos	66	Jaroslaw Lukasik	100
Tatyana Donskova	80	Rasmus Mackeprang	31
Uladzimir Druhakou	6	Tuula Maki	29
Michael Duehrssen	64	Evelina Marinova	63
Viacheslav Duk	69	Harald Markum	3
Helmut Eberl	53	Larry McLerran	73
Gerhard Ecker	28	Danielle Métral	97
Christina Edgar	48	Marine Michaut	86
Saöa Fratina	47	Thomas Millet	14
Esteban Fullana Torregrosa	85	Jovan Mitrevski	61
Vincent Giangiobbe	35	Ciprian Mitu	33
Alexei Gladyshev	96	Jukka Nysten	88
Carlos Gonzalez	36	Fadmar Osmic	65
Elena Guardincerri	87	Siarhei Padolski	37
Hayk Hakobyan	84	Sung Jin Park	95
Per Hansson	12	Shabnaz Pashapour	72
Sam Harper	83	Peicho Petkov	34
Zdenek Hubacek	13	Anton Poluektov	27

Name	Number	Name	Number
Paul Prideaux	49	Oliver Stelzer-Chilton	44
Sebastien Procureur	23	Tadeusz Szymocha	54
Tobias Raufer	105	Fabien Tarrade	46
Stefan Rieke	60	Giovanni Francesco Tassielli	39
Belen Salvachua	103	Barbara Toczek	98
Andreas Salzburger	108	Sergey Troitsky	71
Pavel Sazhin	55	Niels van Eldik	16
Gerolf Schlager	5	Filipe Veloso	58
Sezen Sekmen	77	Sara Vigano	78
Serhiy Senyukov	45	Ian Vollrath	41
Anna Sfyrla	51	Dmytro Volyanskyy	70
Vincent Siccardi	74	Markus Warsinsky	24
Alexey Stadnik	56	Laurenz Widhalm	93
Bernd Stelzer	42	Pasquale Federico Zema	101

Contents

Field theory and the Standard Model

W. Buchmüller and C. Lüdeling
Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany

Abstract
We give a short introduction to the Standard Model and the underlying concepts of quantum field theory.

1 Introduction

In these lectures we shall give a short introduction to the Standard Model of particle physics with emphasis on the electroweak theory and the Higgs sector, and we shall also attempt to explain the underlying concepts of quantum field theory.

The Standard Model of particle physics has the following key features:

- As a theory of elementary particles, it incorporates relativity and quantum mechanics, and therefore it is based on quantum field theory.
- Its predictive power rests on the regularization of divergent quantum corrections and the renormalization procedure which introduces scale-dependent 'running couplings'.
- Electromagnetic, weak, strong and also gravitational interactions are all related to local symmetries and described by Abelian and non-Abelian gauge theories.
- The masses of all particles are generated by two mechanisms: confinement and spontaneous symmetry breaking.

In the following chapters we shall explain these points one by one. Finally, instead of a summary, we briefly recall the history of 'The making of the Standard Model' [1].

From the theoretical perspective, the Standard Model has a simple and elegant structure: it is a chiral gauge theory. Spelling out the details reveals a rich phenomenology which can account for strong and electroweak interactions, confinement and spontaneous symmetry breaking, hadronic and leptonic flavour physics etc. [2, 3]. The study of all these aspects has kept theorists and experimenters busy for three decades. Let us briefly consider these two sides of the Standard Model before we discuss the details.

1.1 Theoretical perspective

The Standard Model is a theory of fields with spins 0, $\frac{1}{2}$ and 1. The fermions (matter fields) can be arranged in a big vector containing left-handed spinors only:

$$\Psi_L^T = \big(\underbrace{q_{L1}, u_{R1}^C, e_{R1}^C, d_{R1}^C, l_{L1}, \left(n_{R1}^C\right)}_{\text{1st family}}, \underbrace{q_{L2}, \ldots,}_{\text{2nd}} \underbrace{\ldots, \left(n_{R3}^C\right)}_{\text{3rd}} \big) \,, \tag{1}$$

where the fields are the quarks and leptons, all in threefold family replication. The quarks come in triplets of colour, i.e., they carry an index α, $\alpha = 1, 2, 3$, which we suppressed in the above expression. The left-handed quarks and leptons come in doublets of weak isospin,

$$q_{Li}^\alpha = \begin{pmatrix} u_{Li}^\alpha \\ d_{Li}^\alpha \end{pmatrix} \quad \text{and} \quad l_{Li} = \begin{pmatrix} \nu_{Li} \\ e_{Li} \end{pmatrix} \,,$$

where i is the family index $i = 1, 2, 3$. We have included a right-handed neutrino n_R because there is evidence for neutrino masses from neutrino oscillation experiments.

The subscripts L and R denote left- and right-handed fields, respectively, which are eigenstates of the chiral projection operators P_L or P_R. The superscript C indicates the charge conjugate field (the antiparticle). Note that the charge conjugate of a right-handed field is left-handed:

$$P_L\psi_L \equiv \frac{1-\gamma^5}{2}\psi_L = \psi_L \,, \qquad P_L\psi_R^C = \psi_R^C \,, \qquad P_L\psi_R = P_L\psi_L^C = 0 \,, \qquad (2)$$

$$P_R\psi_R \equiv \frac{1+\gamma^5}{2}\psi_R = \psi_R \,, \qquad P_R\psi_L^C = \psi_L^C \,, \qquad P_R\psi_L = P_R\psi_R^C = 0 \,. \qquad (3)$$

So all fields in the big column vector of fermions have been chosen left-handed. Altogether there are 48 chiral fermions. The fact that left- and right-handed fermions carry different weak isospin makes the Standard Model a chiral gauge theory. The threefold replication of quark-lepton families is one of the puzzles whose explanation requires physics beyond the Standard Model [4].

The spin-1 particles are the gauge bosons associated with the fundamental interactions in the Standard Model,

$$G_\mu^A, \ A = 1,\dots,8: \qquad \text{the gluons of the strong interactions} \,, \qquad (4)$$

$$W_\mu^I, I = 1,2,3, \ B_\mu: \qquad \text{the } W \text{ and } B \text{ bosons of the electroweak interactions.} \qquad (5)$$

These forces are gauge interactions, associated with the symmetry group

$$G_{\text{SM}} = \mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y \,, \qquad (6)$$

where the subscripts C, W, and Y denote colour, weak isospin and hypercharge, respectively.

The gauge group acts on the fermions via the covariant derivative D_μ, which is an ordinary partial derivative, plus a big matrix A_μ built out of the gauge bosons and the generators of the gauge group:

$$D_\mu\Psi_L = \left(\partial_\mu\mathbb{1} + gA_\mu\right)\Psi_L \,. \qquad (7)$$

From the covariant derivative we can also construct the field strength tensor,

$$F_{\mu\nu} = -\frac{\mathrm{i}}{g}\left[D_\mu, D_\nu\right] \,, \qquad (8)$$

which is a matrix-valued object as well.

The last ingredient of the Standard Model is the Higgs field Φ, the only spin-0 field in the theory. It is a complex scalar field and a doublet of weak isospin. It couples left- and right-handed fermions together.

Written in terms of these fields, the Lagrangian of the theory is rather simple:

$$\begin{aligned}
\mathscr{L} = &-\frac{1}{2}\operatorname{tr}\left[F_{\mu\nu}F^{\mu\nu}\right] + \overline{\Psi}_L\mathrm{i}\gamma^\mu D_\mu\Psi_L + \operatorname{tr}\left[(D_\mu\Phi)^\dagger D^\mu\Phi\right] \\
&+ \mu^2\,\Phi^\dagger\Phi - \frac{1}{2}\lambda\left(\Phi^\dagger\Phi\right)^2 + \left(\frac{1}{2}\Psi_L^T Ch\Phi\Psi_L + \text{h.c.}\right) \,.
\end{aligned} \qquad (9)$$

The matrix C in the last term is the charge conjugation matrix acting on the spinors, h is a matrix of Yukawa couplings. All coupling constants are dimensionless, in particular, there is no mass term for any quark, lepton or vector boson. All masses are generated via the Higgs mechanism which gives a vacuum expectation value to the Higgs field,

$$\langle\Phi\rangle \equiv v = 174\,\text{GeV} \,. \qquad (10)$$

The Higgs boson associated with the Higgs mechanism has not yet been found, but its discovery is generally expected at the LHC.

1.2 Phenomenological aspects

The Standard Model Lagrangian (9) has a rich structure which has led to different areas of research in particle physics:

– The gauge group is composed of three subgroups with different properties:

– The $\mathsf{SU}(3)$ part leads to quantum chromodynamics, the theory of strong interactions [5]. Here the most important phenomena are asymptotic freedom and confinement: The quarks and gluons appear as free particles only at very short distances, probed in deep-inelastic scattering, but are confined into mesons and baryons at large distances.

– The $\mathsf{SU}(2) \times \mathsf{U}(1)$ subgroup describes the electroweak sector of the Standard Model. It gets broken down to the $\mathsf{U}(1)_{\mathrm{em}}$ subgroup of quantum electrodynamics by the Higgs mechanism, leading to massive W and Z bosons which are responsible for charged and neutral-current weak interactions, respectively.

– The Yukawa interaction term can be split into different pieces for quarks and leptons:

$$\frac{1}{2}\Psi_L^T C h \Phi \Psi_L = h_{u\,ij}\bar{u}_{Ri}q_{Lj}\Phi + h_{d\,ij}\bar{d}_{Ri}q_{Lj}\widetilde{\Phi} + h_{e\,ij}\bar{e}_{Ri}l_{Lj}\widetilde{\Phi} + h_{n\,ij}\bar{n}_{Ri}l_{Lj}\Phi \,, \qquad (11)$$

where $i, j = 1, 2, 3$ label the families and $\widetilde{\Phi}_a = \epsilon_{ab}\Phi_b^*$. When the Higgs field develops a vacuum expectation value $\langle\Phi\rangle = v$, the Yukawa interactions generate mass terms. The first two terms, mass terms for up-type- and down-type-quarks, respectively, cannot be diagonalized simultaneously, and this misalignment leads to the CKM matrix and flavour physics [6]. Similarly, the last two terms give rise to lepton masses and neutrino mixings [7].

2 Quantization of fields

In this section we cover some basics of quantum field theory (QFT). For a more in-depth treatment, there are many excellent books on QFT and its application in particle physics, such as Refs. [2,3].

2.1 Why fields?

2.1.1 Quantization in quantum mechanics

Quantum mechanics is obtained from classical mechanics by a method called quantization. Consider, for example, a particle moving in one dimension along a trajectory $q(t)$, with velocity $\dot{q}(t)$ (see Fig. 1). Its motion can be calculated in the Lagrangian or the Hamiltonian approach. The Lagrange function $L(q, \dot{q})$ is a function of the position and the velocity of the particle, usually just the kinetic minus the potential energy. The equation of motion is obtained by requiring that the action, the time integral of the Lagrange function, be extremal, or, in other words, that its variation under arbitrary perturbations around the trajectory vanishes:
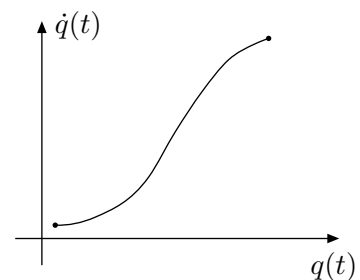
Fig. 1: Particle moving in one dimension

$$\delta S = \delta \int \mathrm{d}t L\left(q(t), \dot{q}(t)\right) = 0 \,. \qquad (12)$$

The Hamiltonian of the system, which corresponds to the total energy, depends on the coordinate q and its conjugate momentum p rather than $\dot{q}$:

$$H(p, q) = p\dot{q} - L\left(q, \dot{q}\right) \,, \qquad p = \frac{\partial L}{\partial \dot{q}} \,. \qquad (13)$$

3

To quantize the system, one replaces the coordinate and the momentum by operators q and p acting on some Hilbert space of states that we shall specify later. In the Heisenberg picture, the states are time-independent and the operators change with time as

$$q(t) = e^{iHt} q(0) e^{-iHt} \ . \tag{14}$$

Since p and q are now operators, they need not commute, and one postulates the commutation relation

$$[p(0), q(0)] = -i\hbar \ , \tag{15}$$

where $h = 2\pi\hbar$ is Planck's constant. In the following we shall use units where $\hbar = c = 1$. The commutator (15) leads to the uncertainty relation

$$\Delta q \cdot \Delta p \geq \frac{1}{2} \ . \tag{16}$$

Note that on Schrödinger wave functions the operator q is just the coordinate itself and p is $-i\partial/\partial q$. In this way the commutation relation (15) is satisfied.

As an example example of a quantum mechanical system, consider the harmonic oscillator with the Hamiltonian

$$H = \frac{1}{2} \left(p^2 + \omega^2 q^2 \right) \ , \tag{17}$$

which corresponds to a particle (with mass 1) moving in a quadratic potential with a strength characterized by ω^2. Classically, H is simply the sum of kinetic and potential energy. In the quantum system, we can define new operators as linear combinations of p and q:

$$q = \frac{1}{\sqrt{2\omega}} \left(a + a^\dagger \right) \ , \qquad\qquad p = -i\sqrt{\frac{\omega}{2}} \left(a - a^\dagger \right) \ , \tag{18a}$$

$$\text{i.e.} \ , \quad a = \sqrt{\frac{\omega}{2}} q + i\sqrt{\frac{1}{2\omega}} p \ , \qquad\qquad a^\dagger = \sqrt{\frac{\omega}{2}} q - i\sqrt{\frac{1}{2\omega}} p \ . \tag{18b}$$

a and $a^\dagger$ satisfy the commutation relations

$$\left[a, a^\dagger \right] = 1 \ . \tag{19}$$

In terms of a and $a^\dagger$ the Hamiltonian is given by

$$H = \frac{\omega}{2} \left(aa^\dagger + a^\dagger a \right) \ . \tag{20}$$

Since Eqs. (18) are linear transformations, the new operators a and $a^\dagger$ enjoy the same time evolution as q and p:

$$a(t) = e^{iHt} a(0) e^{-iHt} = a(0) e^{-i\omega t} \ , \tag{21}$$

where the last equality follows from the commutator of a with the Hamiltonian,

$$[H, a] = -\omega a \ , \quad \left[H, a^\dagger \right] = \omega a^\dagger \ . \tag{22}$$

We can now construct the Hilbert space of states that the operators act on. We first notice that the commutators (22) imply that a and $a^\dagger$ decrease and increase the energy of a state, respectively. To see this, suppose we have a state $|E\rangle$ with fixed energy, $H|E\rangle = E|E\rangle$. Then

$$Ha|E\rangle = (aH + [H, a])|E\rangle = aE|E\rangle - \omega a|E\rangle = (E - \omega) a|E\rangle \ . \tag{23}$$

i.e., the energy of the state $a|E\rangle$ is $(E - \omega)$. In the same way one finds $Ha^\dagger|E\rangle = (E + \omega)|E\rangle$. From the form of H we can also see that its eigenvalues must be positive. This suggests constructing the space of states starting from a lowest-energy state $|0\rangle$, the vacuum or no-particle state. This state needs to satisfy

$$a|0\rangle = 0 , \tag{24}$$

so its energy is $\omega/2$. States with more 'particles', i.e., higher excitations, are obtained by successive application of $a^\dagger$:

$$|n\rangle = \left(a^\dagger\right)^n |0\rangle , \quad \text{with} \quad H|n\rangle = \left(n + \frac{1}{2}\right)\omega|n\rangle . \tag{25}$$

2.1.2 Special relativity requires antiparticles

So far, we have considered non-relativistic quantum mechanics. A theory of elementary particles, however, has to incorporate special relativity. It is very remarkable that quantum mechanics together with special relativity implies the existence of antiparticles. To see this (following an argument in Ref. [8]), consider two systems (e.g., atoms) A_1 and A_2 at positions $\vec{x}_1$ and $\vec{x}_2$. Assume that at time t_1 atom A_1 emits an electron and turns into B_1. So the charge of B_1 is one unit higher than that of A_1. At a later time t_2 the electron is absorbed by atom A_2 which turns into B_2 with charge lower by one unit. This is illustrated in Fig. 2.

$$A_2 + e^- \to B_2$$
$$\Delta Q = -1$$
$$(t_2, \vec{x}_2)$$

$$(t_1, \vec{x}_1)$$

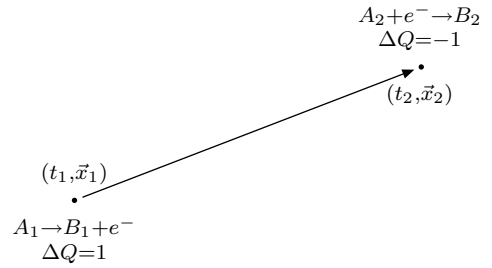$$A_1 \to B_1 + e^-$$
$$\Delta Q = 1$$

Fig. 2: Electron moving from A_1 to A_2

According to special relativity, we can also watch the system from a frame moving with relative velocity $\vec{v}$. One might now worry whether the process is still causal, i.e., whether the emission still precedes the absorption. In the boosted frame (with primed coordinates), one has

$$t_2' - t_1' = \gamma (t_2 - t_1) + \gamma \vec{v}(\vec{x}_2 - \vec{x}_1) , \quad \gamma = \frac{1}{\sqrt{1 - \vec{v}^2}} . \tag{26}$$

Here $t_2' - t_1'$ must be positive for the process to remain causal. Since $|\vec{v}| < 1$, $t_2' - t_1'$ can only be negative for spacelike distances, i.e., $(t_2 - t_1)^2 - (\vec{x}_1 - \vec{x}_2)^2 < 0$. This, however, would mean that the electron travelled faster than the speed of light, which is not possible according to special relativity. Hence, within classical physics, causality is not violated.

This is where quantum mechanics comes in. The uncertainty relation leads to a 'fuzzy' light cone, which gives a non-negligible propagation probability for the electron even for slightly spacelike distances, as long as

$$(t_2 - t_1)^2 - (\vec{x}_1 - \vec{x}_2)^2 \gtrsim -\frac{\hbar^2}{m^2} . \tag{27}$$

Does this mean causality is violated?

$$A_1 + e^+ \to B_1$$
$$\Delta Q = 1$$
$$(t_1', \vec{x}_1')$$

$$(t_2', \vec{x}_2')$$
$$A_2 \to B_2 + e^+$$
$$\Delta Q = -1$$

Fig. 3: Positron moving from A_2 to A_1

Fortunately, there is a way out: The antiparticle. In the moving frame, one can consider the whole process as the emission of a positron at $t = t_2'$, followed by its absorption at a later time $t = t_1'$ (see Fig. 3). So we see that quantum mechanics together with special relativity requires the existence of antiparticles for consistency. In addition, particle and antiparticle need to have the same mass.
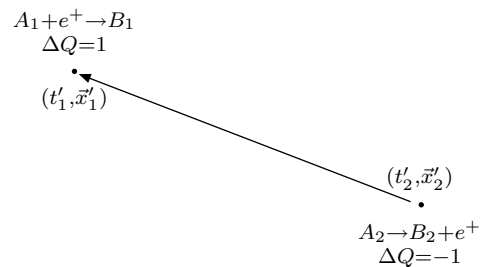
In a relativistic theory, the uncertainty relation (16) also implies that particles cannot be localized below their Compton wavelength

$$\Delta x \geq \frac{\hbar}{mc} \,. \tag{28}$$

For shorter distances the momentum uncertainty $\Delta p > mc$ allows for contributions from multiparticle states, and one can no longer talk about a single particle.

2.2 Multiparticle states and fields

In the previous section we saw that the combination of quantum mechanics and special relativity has important consequences. First, we need antiparticles, and second, particle number is not well defined. These properties can be conveniently described by means of fields. A field here is a collection of infinitely many harmonic oscillators, corresponding to different momenta. For each oscillator, we can construct operators and states just as before in the quantum mechanical case. These operators will then be combined into a field operator, the quantum analogue of the classical field. These results can be obtained by applying the method of canonical quantization to fields.

2.2.1 States, creation and annihilation

The starting point is a continuous set of harmonic oscillators which are labelled by the spatial momentum $\vec{k}$. We want to construct the quantum fields for particles of mass m, so we can combine each momentum $\vec{k}$ with the associated energy $\omega_k = k^0 = \sqrt{\vec{k}^2 + m^2}$ to form the momentum 4-vector k. This 4-vector satisfies $k^2 \equiv k^\mu k_\mu = m^2$. For each k we define creation and annihilation operators, both for particles $(a, a^\dagger)$ and antiparticles $(b, b^\dagger)$, and construct the space of states just as we did for the harmonic oscillator in the previous section.

For the states we again postulate the vacuum state, which is annihilated by both particle and antiparticle annihilation operators. Each creation operator $a^\dagger(k)$ $(b^\dagger(k))$ creates a (anti)particle with momentum k, so the space of states is

$$\text{vacuum: } |0\rangle \,, \quad a(k)|0\rangle = b(k)|0\rangle = 0$$
$$\text{one-particle states: } a^\dagger(k)|0\rangle \,, \ b^\dagger(k)|0\rangle$$
$$\text{two-particle states: } a^\dagger(k_1)a^\dagger(k_2)|0\rangle \,, \ a^\dagger(k_1)b^\dagger(k_2)|0\rangle \,, \ b^\dagger(k_1)b^\dagger(k_2)|0\rangle$$
$$\vdots$$

Like in the harmonic oscillator case, we also have to postulate the commutation relations of these operators, and we choose them in a similar way: operators with different momenta correspond to different harmonic oscillators and hence they commute. Furthermore, particle and antiparticle operators should commute with each other. Hence, there are only two non-vanishing commutators ('canonical commutation relations'):

$$\left[a(k), a^\dagger(k') \right] = \left[b(k), b^\dagger(k') \right] = (2\pi)^3 \, 2\omega_k \, \delta^3\!\left(\vec{k} - \vec{k}' \right) \,, \tag{29}$$

which are the counterparts of relation (19). The expression on the right-hand side is the Lorentz-invariant way to say that only operators with the same momentum do not commute [the $(2\pi)^3$ is just convention].

Since we now have a continuous label for the creation and annihilation operators, we need a Lorentz-invariant way to sum over operators with different momentum. The four components of k are not independent, but satisfy $k^2 \equiv k_\mu k^\mu = m^2$, and we also require positive energy, that is $k^0 = \omega_k > 0$.

Taking these things into account, one is led to the integration measure

$$
\begin{aligned}
\int \overline{\mathrm{d}k} &\equiv \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \, 2\pi \, \delta\!\left(k^2 - m^2\right) \Theta\!\left(k^0\right) \\
&= \int \frac{\mathrm{d}^4 k}{(2\pi)^3} \delta\!\left(\left(k^0 - \omega_k\right)\left(k^0 + \omega_k\right)\right) \Theta\!\left(k^0\right) \\
&= \int \frac{\mathrm{d}^4 k}{(2\pi)^3} \frac{1}{2\omega_k} \left(\delta\!\left(k^0 - \omega_k\right) + \delta\!\left(k^0 + \omega_k\right)\right) \Theta\!\left(k^0\right) \\
&= \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{1}{2\omega_k} \,.
\end{aligned}
\tag{30}
$$

The numerical factors are chosen such that they match those in Eq. (29) for the commutator of $a(k)$ and $a^\dagger(k)$.

2.2.2 Charge and momentum

Now we have the necessary tools to construct operators which express some properties of fields and states. The first one is the operator of 4-momentum, i.e., of spatial momentum and energy. Its construction is obvious, since we interpret $a^\dagger(k)$ as a creation operator for a state with 4-momentum k. That means we just have to count the number of particles with each momentum and sum the contributions:

$$
P^\mu = \int \overline{\mathrm{d}k} \, k^\mu \left(a^\dagger(k)a(k) + b^\dagger(k)b(k)\right) .
\tag{31}
$$

This gives the correct commutation relations:

$$
\left[P^\mu, a^\dagger(k)\right] = k^\mu a^\dagger(k) , \qquad\qquad \left[P^\mu, b^\dagger(k)\right] = k^\mu b^\dagger(k) ,
\tag{32a}
$$

$$
\left[P^\mu, a(k)\right] = -k^\mu a(k) , \qquad\qquad \left[P^\mu, b(k)\right] = -k^\mu b(k) .
\tag{32b}
$$

Another important operator is the charge. Since particles and antiparticles have opposite charge, the net charge of a state is proportional to the number of particles minus the number of antiparticles:

$$
Q = \int \overline{\mathrm{d}k} \left(a^\dagger(k)a(k) - b^\dagger(k)b(k)\right) ,
\tag{33}
$$

and one easily verifies

$$
\left[Q, a^\dagger(k)\right] = a^\dagger(k) , \qquad \left[Q, b^\dagger(k)\right] = -b^\dagger(k) .
\tag{34}
$$

We have now confirmed our intuition that $a^\dagger(k)$ $\left(b^\dagger(k)\right)$ creates a particle with 4-momentum k and charge +1 (−1). Both momentum and charge are conserved: The time derivative of an operator is equal to the commutator of the operator with the Hamiltonian, which is the 0-component of P^μ. This obviously commutes with the momentum operator, but also with the charge:

$$
\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}Q = [Q, H] = 0 .
\tag{35}
$$

So far, this construction applied to the case of a complex field. For the special case of neutral particles, one has $a = b$ and $Q = 0$, i.e., the field is real.

2.2.3 Field operator

We are now ready to introduce field operators, which can be thought of as the Fourier transform of creation and annihilation operators:

$$\phi(x) = \int \overline{dk}\left(e^{-ikx}a(k) + e^{ikx}b^\dagger(k)\right) . \tag{36}$$

A spacetime translation is generated by the 4-momentum in the following way:

$$e^{iyP}\phi(x)e^{-iyP} = \phi(x+y) . \tag{37}$$

This transformation can be derived from the transformation of the a's:

$$e^{iyP}a^\dagger(k)e^{-iyP} = a^\dagger(k) + iy_\mu\left[P^\mu, a^\dagger(k)\right] + \mathcal{O}\left(y^2\right) \tag{38}$$

$$= (1 + iyk + \cdots)\, a^\dagger(k) \tag{39}$$

$$= e^{iyk}a^\dagger(k) . \tag{40}$$

The commutator with the charge operator is

$$[Q, \phi(x)] = -\phi(x) , \qquad\qquad \left[Q, \phi^\dagger\right] = \phi^\dagger . \tag{41}$$

The field operator obeys the (free) field equation,

$$\left(\Box + m^2\right)\phi(x) = \int \overline{dk}\left(-k^2 + m^2\right)\left(e^{-ikx}a(k) + e^{ikx}b^\dagger(k)\right) = 0 , \tag{42}$$

where $\Box = \partial^2/\partial t^2 - \vec{\nabla}^2$ is the d'Alambert operator.

2.2.4 Propagator

Now we can tackle the problem of causal propagation that led us to introduce antiparticles. We consider the causal propagation of a charged particle between $x_1^\mu = (t_1, \vec{x}_1)$ and $x_2^\mu = (t_2, \vec{x}_2)$, see Fig. 4. The field operator creates a state with charge ± 1 'at position $(t, \vec{x})$',

$$Q\,\phi(t, \vec{x})|0\rangle = -\phi(t, \vec{x})|0\rangle , \tag{43}$$

$$Q\,\phi^\dagger(t, \vec{x})|0\rangle = \phi^\dagger(t, \vec{x})|0\rangle . \tag{44}$$

Fig. 4: Propagation of a particle or an antiparticle, depending on the temporal order

Depending on the temporal order of x_1 and x_2, we interpret the propagation of charge either as a particle going from x_1 to x_2 or an antiparticle going the other way. Formally, this is expressed as the time-ordered product [using the Θ-function, $\Theta(\tau) = 1$ for $\tau > 0$ and $\Theta(\tau) = 0$ for $\tau < 0$]:

$$\mathrm{T}\,\phi(x_2)\phi^\dagger(x_1) = \Theta(t_2 - t_1)\phi(x_2)\phi^\dagger(x_1) + \Theta(t_1 - t_2)\phi^\dagger(x_1)\phi(x_2) . \tag{45}$$

The vacuum expectation value of this expression is the Feynman propagator:

$$i\Delta_{\mathrm{F}}(x_2 - x_1) = \left\langle 0\,\middle|\,\mathrm{T}\,\phi(x_2)\phi^\dagger(x_1)\,\middle|\,0\right\rangle$$

$$= i\int \frac{d^4k}{(2\pi)^4}\frac{e^{ik(x_2 - x_1)}}{k^2 - m^2 + i\varepsilon} , \tag{46}$$

where we used the Θ-function representation

$$\Theta(\tau) = -\frac{1}{2\pi\mathrm{i}} \int_{-\infty}^{\infty} \mathrm{d}\omega \frac{e^{-\mathrm{i}\omega\tau}}{\omega + \mathrm{i}\epsilon} \, . \tag{47}$$

This Feynman propagator is a Green function for the field equation,

$$\left(\Box + m^2\right) \Delta_{\mathrm{F}}(x_2 - x_1) = \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{\left(-p^2 + m^2\right)}{p^2 - m^2 + \mathrm{i}\varepsilon} e^{-\mathrm{i}p(x_2 - x_1)} = -\delta^4 \left(x_2 - x_1\right) \, . \tag{48}$$

It is causal, i.e., it propagates particles into the future and antiparticles into the past.

2.3 Canonical quantization

All the results from the previous section can be derived in a more rigorous manner by using the method of canonical quantization which provides the step from classical to quantum mechanics. We now start from classical field theory, where the field at point $\vec{x}$ corresponds to the position q in classical mechanics, and we again have to construct the conjugate momentum variables and impose commutation relations among them.

Let us consider the Lagrange density for a complex scalar field ϕ. Like the Lagrangian in classical mechanics, the free Lagrange density is just the kinetic minus the potential energy density,

$$\mathscr{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi \, . \tag{49}$$

The Lagrangian has a $\mathsf{U}(1)$-symmetry, i.e., under the transformation of the field

$$\phi \to \phi' = e^{\mathrm{i}\alpha}\phi \, , \quad \alpha = \mathrm{const.} \, , \tag{50}$$

it stays invariant. From Noether's theorem, there is a conserved current j_μ associated with this symmetry,

$$j^\mu = \mathrm{i}\phi^\dagger \overleftrightarrow{\partial}^\mu \phi = \mathrm{i}\left(\phi^\dagger \partial^\mu \phi - \partial^\mu \phi^\dagger \phi\right) \, , \qquad\qquad \partial_\mu j^\mu = 0 \, . \tag{51}$$

The space integral of the time component of this current is conserved in time:

$$Q = \int \mathrm{d}^3 x \, \mathrm{i}\phi^\dagger \overleftrightarrow{\partial}^0 \phi \, , \quad \partial_0 Q = 0 \, . \tag{52}$$

The time derivative vanishes because we can interchange derivation and integration and then replace $\partial_0 j^0$ by $\partial_i j^i$ since $\partial_\mu j^\mu = \partial_0 j^0 + \partial_i j^i = 0$. So we are left with an integral of a total derivative which we can transform into a surface integral via Gauss's theorem. Since we always assume that all fields vanish at spatial infinity, the surface term vanishes.

Now we need to construct the 'momentum' $\pi(x)$ conjugate to the field ϕ. Like in classical mechanics, it is given by the derivative of the Lagrangian with respect to the time derivative of the field,

$$\pi(x) = \frac{\partial \mathscr{L}}{\partial \dot{\phi}(x)} = \dot{\phi}^\dagger(x) \, , \qquad\qquad \pi^\dagger(x) = \frac{\partial \mathscr{L}}{\partial \dot{\phi}^\dagger(x)} = \dot{\phi} \, . \tag{53}$$

At this point, we again replace the classical fields by operators which act on some Hilbert space of states and which obey certain commutation relations. The commutation relations we have to impose are analogous to Eq. (15). The only non-vanishing commutators are the ones between field and conjugate momentum, at different spatial points but at equal times,

$$\left[\pi(t, \vec{x}), \phi(t, \vec{x}')\right] = \left[\pi^\dagger(t, \vec{x}), \phi^\dagger(t, \vec{x}')\right] = -\mathrm{i}\delta^3 \left(\vec{x} - \vec{x}'\right) \, , \tag{54}$$

all other commutators vanish.

These relations are satisfied by the field operator defined in Eq. (36) via the (anti)particle creation and annihilation operators. Its field equation can be derived from the Lagrangian,

$$\partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathscr{L}}{\partial \phi} = \left(\Box + m^2\right)\phi^\dagger = 0\,. \tag{55}$$

From the Lagrangian and the momentum, we can also construct the Hamiltonian density,

$$\mathscr{H} = \pi \dot{\phi} + \pi^\dagger \dot{\phi}^\dagger - \mathscr{L} = \pi^\dagger \pi + \left(\vec{\nabla}\phi^\dagger\right)\left(\vec{\nabla}\phi\right) + m^2 \phi^\dagger \phi\,. \tag{56}$$

Note that canonical quantization yields Lorentz-invariant results, although it requires the choice of a particular time direction.

2.4 Fermions

Fermions make calculations unpleasant.

In the previous section we considered a scalar field which describes particles with spin 0. In the Standard Model, there is just one fundamental scalar field, the Higgs field, which still remains to be discovered. There are other bosonic fields, gauge fields which carry spin 1 (photons, $\mathrm{W}^\pm$, Z^0 and the gluons). Those are described by vector fields which will be discussed in Section 3. Furthermore, there are the matter fields, fermions with spin $\frac{1}{2}$, the quarks and leptons.

To describe fermionic particles, we need to introduce new quantities, spinor fields. These are four-component objects (but not vectors!) ψ, which are defined via a set of γ matrices. These four-by-four matrices are labelled by a vector index and act on spinor indices. They fulfil the anticommutation relations (the Clifford or Dirac algebra),

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}\mathbb{1}\,, \tag{57}$$

with the metric $g_{\mu\nu} = \mathrm{diag}(+,-,-,-)$. The numerical form of the γ matrices is not fixed, rather, one can choose among different possible representations. A common representation is the so-called chiral or Weyl representation, which is constructed from the Pauli matrices:

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix}\,, \qquad\qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\,. \tag{58}$$

This representation is particularly useful when one considers spinors of given chiralities. However, for other purposes, other representations are more convenient. Various rules and identities related to γ matrices are collected in Appendix A.

The Lagrangian for a free fermion contains, just as for a scalar, the kinetic term and the mass:

$$\mathscr{L} = \overline{\psi}\mathrm{i}\slashed{\partial}\psi - m\overline{\psi}\psi\,. \tag{59}$$

The kinetic term contains only a first-order derivative, the operator $\slashed{\partial} \equiv \gamma^\mu \partial_\mu$. The adjoint spinor $\overline{\psi}$ is defined as $\overline{\psi} \equiv \psi^\dagger \gamma^0$. (The first guess $\psi^\dagger \psi$ is not Lorentz invariant.) To derive the field equation, one has to treat ψ and $\overline{\psi}$ as independent variables. The Euler–Lagrange equation for $\overline{\psi}$ is the familiar Dirac equation:

$$0 = \frac{\partial \mathscr{L}}{\partial \overline{\psi}} = \left(\mathrm{i}\slashed{\partial} - m\right)\psi\,, \tag{60}$$

since $\mathscr{L}$ does not depend on derivatives of $\overline{\psi}$ [1].

---

[1] Of course one can shift the derivative from ψ to $\overline{\psi}$ via integration by parts. This slightly modifies the computation, but the result is still the same.

The Lagrangian again has a $U(1)$ symmetry, the multiplication of ψ by a constant phase,

$$\psi \rightarrow \psi' = e^{i\alpha}\psi\,, \qquad \overline{\psi} \rightarrow \overline{\psi}' = e^{-i\alpha}\overline{\psi}\,, \tag{61}$$

which leads to a conserved current and, correspondingly, to a conserved charge,

$$j^\mu = \overline{\psi}\gamma^\mu\psi\,, \quad \partial^\mu j_\mu = 0\,, \quad Q = \int \mathrm{d}^3x\,\overline{\psi}\gamma^0\psi\,. \tag{62}$$

2.4.1 Canonical quantization of fermions

Quantization proceeds along similar lines as in the scalar case. One first defines the momentum π_α conjugate to the field ψ_α ($\alpha = 1, \ldots, 4$),

$$\pi_\alpha = \frac{\partial \mathscr{L}}{\partial \dot{\psi}_\alpha} = i\left(\overline{\psi}\gamma^0\right)_\alpha = i\psi_\alpha^\dagger\,. \tag{63}$$

Instead of imposing commutation relations, however, for fermions one has to impose anticommutation relations. This is a manifestation of the Pauli exclusion principle which can be derived from the spin-statistics theorem. The relations are again postulated at equal times ('canonical anticommutation relations'):

$$\left\{\pi_\alpha(t,\vec{x}), \psi_\beta(t,\vec{x}')\right\} = -i\delta_{\alpha\beta}\delta^3\left(\vec{x} - \vec{x}'\right)\,, \tag{64a}$$

$$\left\{\pi_\alpha(t,\vec{x}), \pi_\beta(t,\vec{x}')\right\} = \left\{\psi_\alpha(t,\vec{x}), \psi_\beta(t,\vec{x}')\right\} = 0\,. \tag{64b}$$

In order to obtain creation and annihilation operators, we again expand the field operator in terms of plane waves. Because of the four-component nature of the field, now a spinor $u(p)$ occurs, where p is the momentum four-vector of the plane wave:

$$\left(i\slashed{\partial} - m\right)u(p)e^{-ipx} = 0\,, \tag{65}$$

which implies

$$\left(\slashed{p} - m\right)u(p) = 0\,. \tag{66}$$

This is an eigenvalue equation for the 4×4 matrix $p_\mu\gamma^\mu$, which has two solutions for $p^2 = m^2$ and $p^0 > 0$. They are denoted $u^{(1,2)}(p)$ and represent positive energy particles. Taking a positive sign in the exponential in Eq. (65), which is equivalent to considering $p^0 < 0$, we obtain two more solutions, $v^{(1,2)}(p)$ that can be interpreted as antiparticles. The form of these solutions depends on the representation of the γ matrices. For the Weyl representation they are given in the Appendix.

The eigenspinors determined from the equations ($i = 1, 2$),

$$\left(\slashed{p} - m\right)u^{(i)}(p) = 0\,, \qquad\qquad \left(\slashed{p} + m\right)v^{(i)}(p) = 0\,, \tag{67}$$

obey the identities:

$$\overline{u}^{(i)}(p)u^{(j)}(p) = -\overline{v}^{(i)}(p)v^{(j)}(p) = 2m\delta_{ij}\,, \tag{68}$$

$$\sum_i u_\alpha^{(i)}(p)\overline{u}_\beta^{(i)}(p) = \left(\slashed{p} + m\right)_{\alpha\beta}\,, \qquad \sum_i v_\alpha^{(i)}(p)\overline{v}_\beta^{(i)}(p) = \left(\slashed{p} - m\right)_{\alpha\beta}\,. \tag{69}$$

These are the ingredients we need to define creation and annihilation operators in terms of the spinor field $\psi(x)$ and its conjugate $\overline{\psi}(x)$:

$$\psi(x) = \int \overline{\mathrm{d}p} \sum_i \left(b_i(p)u^{(i)}(p)e^{-ipx} + d_i^\dagger(p)v^{(i)}(p)e^{ipx}\right)\,, \tag{70a}$$

$$\overline{\psi}(x) = \int \overline{\mathrm{d}p} \sum_i \left(b_i^\dagger(p)\overline{u}^{(i)}(p)e^{ipx} + d_i(p)\overline{v}^{(i)}(p)e^{-ipx} \right) . \tag{70b}$$

Here, as before,

$$\overline{\mathrm{d}p} = \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2E_p}, \quad E_p = \sqrt{\vec{p}^2 + m^2}. \tag{71}$$

Inverting Eq. (70a) one obtains

$$b_i(p) = \int \mathrm{d}^3 x \, \overline{u}^{(i)}(p)e^{ipx}\gamma^0\psi(x), \tag{72}$$

and similar equations for the other operators.

The creation and annihilation operators inherit the anticommutator algebra from the field operators,

$$\left\{ b_i(\vec{p}), b_j^\dagger(\vec{p}') \right\} = \left\{ d_i(\vec{p}), d_j^\dagger(\vec{p}') \right\} = (2\pi)^3 2E_p \delta^3 \left(\vec{p} - \vec{p}' \right) , \tag{73a}$$

$$\left\{ b_i(\vec{p}), d_j(\vec{p}') \right\} = \text{all other anticommutators} = 0. \tag{73b}$$

The momentum and charge operators are again constructed from the creation and annihilation operators by 'counting' the number of particles in each state and summing over all states,

$$P^\mu = \int \overline{\mathrm{d}k} \, k^\mu \left(b^\dagger(k)b(k) + d^\dagger(k)d(k) \right) , \tag{74}$$

$$Q = \int \overline{\mathrm{d}k} \left(b^\dagger(k)b(k) - d^\dagger(k)d(k) \right) . \tag{75}$$

These operators have the correct algebraic relations, which involve commutators, since P^μ and Q are bosonic operators (not changing the number of fermions in a given state):

$$\left[P^\mu, b_i^\dagger(p) \right] = p^\mu b_i^\dagger(p), \qquad\qquad \left[P^\mu, d_i^\dagger(p) \right] = p^\mu d_i^\dagger(p), \tag{76}$$

$$\left[Q, b_i^\dagger(p) \right] = b_i^\dagger(p), \qquad\qquad \left[Q, d_i^\dagger(p) \right] = -d_i^\dagger(p). \tag{77}$$

An operator we did not encounter in the scalar case is the spin operator $\vec{\Sigma}$. It has three components, corresponding to the three components of an angular momentum vector[2]. Only one combination of these components is, however, measurable. This is specified by a choice of quantization axis, i.e., a spatial unit vector $\vec{s}$. The operator that measures the spin of a particle is given by the scalar product $\vec{s} \cdot \vec{\Sigma}$. Creation operators for particles with definite spin satisfy the commutation relations

$$\left[\vec{s} \cdot \vec{\Sigma}, d_\pm^\dagger(p) \right] = \mp \frac{1}{2} d_\pm^\dagger(p), \qquad\qquad \left[\vec{s} \cdot \vec{\Sigma}, b_\pm^\dagger(p) \right] = \pm \frac{1}{2} b_\pm^\dagger(p). \tag{78}$$

In summary, all these commutation relations tell us how to interpret the operators $d_\pm^\dagger(p)$ ($b_\pm^\dagger(p)$): They create spin-$\frac{1}{2}$ fermions with four-momentum p^μ, charge $+1$ (-1) and spin orientation $\pm\frac{1}{2}$ ($\mp\frac{1}{2}$) relative to the chosen axis $\vec{s}$. Their conjugates $d_\pm(p)$ and $b_\pm(p)$ annihilate those particles.

This immediately leads to the construction of the Fock space of fermions: We again start from a vacuum state $|0\rangle$, which is annihilated by the annihilation operators, and construct particle states by successive application of creation operators:

$$\text{vacuum:} \ |0\rangle , \quad b_i(p)|0\rangle = d_i(p)|0\rangle = 0$$

---

[2] Actually, Σ is constructed as a commutator of γ matrices and as such has six independent components. But three of these correspond to Lorentz boosts which mix time and spatial directions. $\vec{\Sigma}$ is the spin operator in the rest frame.

$$\text{one-particle states: } b_i^\dagger(p)|0\rangle \ , \ d_i^\dagger(p)|0\rangle$$

$$\text{two-particle states: } b_i^\dagger(p_1)d_j^\dagger(p_2)|0\rangle \ ,\ldots$$

$$\vdots$$

At this point we can verify that the Pauli principle is indeed satisfied, on account of the choice of anticommutation relations in Eq. (64). For a state of two fermions with identical quantum numbers, we would get

$$\underbrace{b_i^\dagger(p)\,b_i^\dagger(p)}_{\text{anticommuting}}|X\rangle = -b_i^\dagger(p)\,b_i^\dagger(p)|X\rangle = 0\,, \tag{79}$$

where $|X\rangle$ is an arbitrary state. Had we quantized the theory with commutation relations instead, the fermions would have the wrong (i.e., Bose) statistics.

The final expression we need for the further discussion is the propagator. By the same reasoning as in the scalar case, it is obtained as the time-ordered product of two field operators. The Feynman propagator S_F for fermions, which is now a matrix-valued object, is given by

$$\begin{aligned}
\mathrm{i}S_\mathrm{F}(x_1 - x_2)_{\alpha\beta} &= \big\langle 0 \,\big|\, \mathrm{T}\,\psi_\alpha(x_1)\overline{\psi}_\beta(x_2)\,\big|\, 0 \big\rangle \\
&= \mathrm{i}\int \frac{\mathrm{d}^4p}{(2\pi)^4} \frac{(\not{p} + m)_{\alpha\beta}}{p^2 - m^2 + \mathrm{i}\varepsilon} e^{-\mathrm{i}p(x_1 - x_2)}\,.
\end{aligned} \tag{80}$$

This completes our discussion on the quantization of free scalar and spinor fields.

2.5 Interactions

So far, we have considered free particles and their propagation. A theory of elementary particles obviously needs interactions. Unfortunately, they are much more difficult to handle, and little is known rigorously (except in two dimensions). Hence, we have to look for approximations.

By far the most important approximation method is perturbation theory where one treats the interaction as a small effect, a perturbation, to the free theory. The interaction strength is quantified by a numerical parameter, the coupling constant, and one expresses physical quantities as power series in this parameter. This approach has been very successful and has led to many celebrated results, like the precise prediction of the anomalous magnetic moment of the electron, despite the fact that important conceptual problems still remain to be resolved.

2.5.1 ϕ^4 theory

Let us consider the simplest example of an interacting theory, involving only one real scalar field with a quartic self-interaction (a cubic term would look even simpler, but then the theory would not have a ground state since the energy would not be bounded from below):

$$\begin{aligned}
\mathscr{L} &= \mathscr{L}_0 + \mathscr{L}_\mathrm{I} \\
&= \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\,.
\end{aligned} \tag{81}$$

$\mathscr{L}_0$ is the free Lagrangian, containing kinetic and mass term, while $\mathscr{L}_\mathrm{I}$ is the interaction term, whose strength is given by the dimensionless coupling constant λ.

In perturbation theory we can calculate various physical quantities, in particular scattering cross-sections for processes like the one in Fig. 5: n particles with momenta p_i interact, resulting in m particles
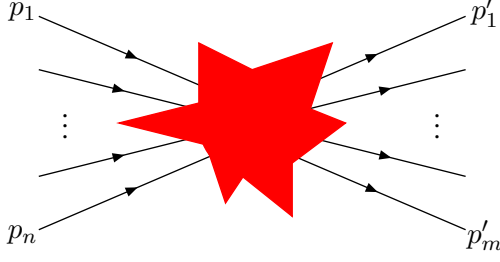
Fig. 5: Scattering of n incoming particles, producing m outgoing ones with momenta $p_1, \ldots, p_n$ and $p'_1, \ldots, p'_m$, respectively
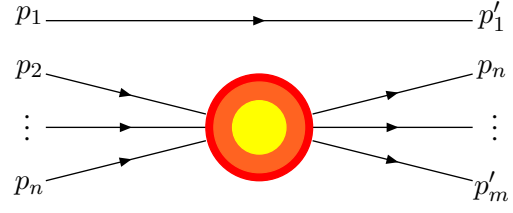

Fig. 6: A disconnected diagram: One particle does not participate in the interaction

with momenta p'_j. Since the interaction is localized in a region of spacetime, particles are free at infinite past and future. In other words, we have free asymptotic states

$$\left| p_1, \ldots, p_n, \text{in} \right\rangle \text{ at } , t = -\infty \quad \text{and} \quad \left| p'_1, \ldots, p'_m, \text{out} \right\rangle \text{ at } t = +\infty. \tag{82}$$

The transition amplitude for the scattering process is determined by the scalar product of incoming and outgoing states, which defines a unitary matrix, the so-called S-matrix (S for scattering),

$$\left\langle p'_1, \ldots, p'_m, \text{out} \,\middle|\, p_1, \ldots, p_n, \text{in} \right\rangle = \left\langle p'_1, \ldots, p'_m \,\middle|\, S \,\middle|\, p_1, \ldots, p_n \right\rangle. \tag{83}$$

Detailed techniques have been developed to obtain a perturbative expansion for the S-matrix from the definition (83). The basis is Wick's theorem and the LSZ formalism. One starts from a generalization of the propagator, the time-ordered product of k fields,

$$\tau(x_1, \ldots, x_k) = \left\langle 0 \,\middle|\, \text{T}\, \phi(x_1), \ldots \phi(x_k) \middle| 0 \right\rangle. \tag{84}$$

First, disconnected pieces involving non-interacting particles have to be subtracted (see Fig. 6), and the blob in Fig. 5 decomposes into a smaller blob and straight lines just passing from the left to the right side. From the Fourier transform

$$\tau(x'_1, \ldots, x'_m, x_1, \ldots, x_n) \xrightarrow{\text{F.T.}} \tilde{\tau}(p'_1, \ldots, p'_m, p_1, \ldots, p_n) \tag{85}$$
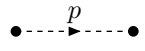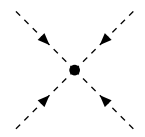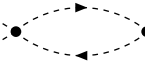
one then obtains the amplitude for the scattering process

$$\left\langle p'_1, \ldots, p'_m \,\middle|\, S \,\middle|\, p_1, \ldots, p_n \right\rangle = (2\pi)^4\, \delta^4\left(\sum_{\text{out}} p'_i - \sum_{\text{in}} p_i \right) i\mathcal{M}, \tag{86}$$

where the matrix element $\mathcal{M}$ contains all the dynamics of the interaction. Because of the translational invariance of the theory, the total momentum is conserved. The matrix element can be calculated perturbatively up to the desired order in the coupling λ via a set of Feynman rules. To calculate the cross-section for a particular process, one first draws all possible Feynman diagrams with a given number of vertices and then translates them into an analytic expression using the Feynman rules.

For the ϕ^4 theory, the Feynman diagrams are all composed out of three building blocks: External lines corresponding to incoming or outgoing particles, propagators, and 4-vertices. The Feynman rules read:

i.   1   External lines: For each external line, multiply by 1 (i.e., external lines do not contribute to the matrix element in this theory). However, one needs to keep track of the momentum of each particle entering or leaving the interaction. The momentum direction is indicated by the arrow.

ii. $$\frac{i}{p^2 - m^2 + i\varepsilon}$$ Propagators between vertices are free propagators corresponding to the momentum of the particle. Note that particles of internal lines need not be on-shell, i.e., $p^2 = m^2$ need not hold!

iii. $-i\lambda$ Vertices yield a factor of the coupling constant. In this theory, there is only one species of particles, and the interaction term does not contain derivatives, so there is only one vertex, and it does not depend on the momenta.

iv. $$\int \frac{\mathrm{d}^4 p}{(2\pi)^4}$$ The momenta of internal loops are not fixed by the incoming momenta. For each undetermined loop momentum p, one integrates over all values of p.

As an example, let us calculate the matrix element for the $2 \to 2$ scattering process to second order in λ. The relevant diagrams are collected in Fig. 7.

(a) Tree graph
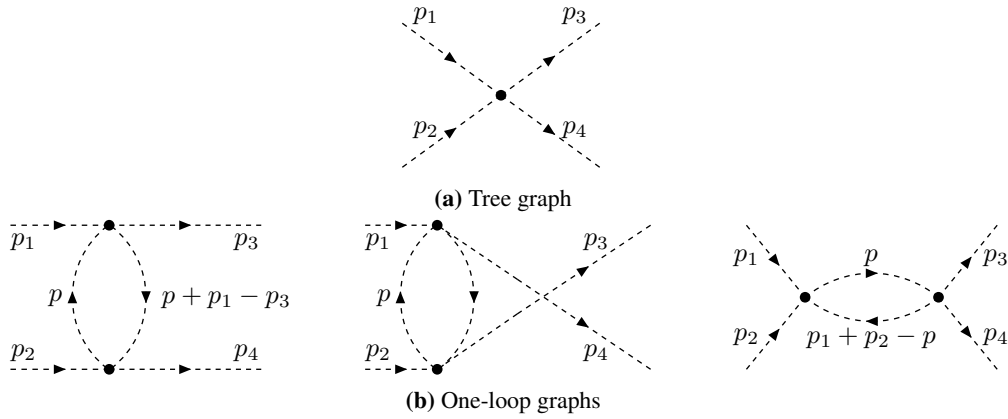
(b) One-loop graphs

Fig. 7: Feynman graphs for $2 \to 2$ scattering in ϕ^4 theory to second order. The one-loop graphs are all invariant under the interchange of the internal lines and hence get a symmetry factor of $\frac{1}{2}$.

The first-order diagram simply contributes a factor of $-i\lambda$, while the second-order diagrams involve an integration:

$$
\begin{aligned}
i\mathcal{M} = -i\lambda &+ \frac{1}{2}\left(-i\lambda\right)^2 \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2} \frac{i}{\left(p + p_1 - p_3\right)^2 - m^2} \\
&+ \frac{1}{2}\left(-i\lambda\right)^2 \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2} \frac{i}{\left(p + p_1 - p_4\right)^2 - m^2} \\
&+ \frac{1}{2}\left(-i\lambda\right)^2 \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2} \frac{i}{\left(p_1 + p_2 - p\right)^2 - m^2} + \mathcal{O}\left(\lambda^3\right).
\end{aligned}
\tag{87}
$$

The factors of $\frac{1}{2}$ are symmetry factors which arise if a diagram is invariant under interchange of internal lines. The expression for $\mathcal{M}$ has a serious problem: The integrals do not converge. This can be seen by counting the powers of the integration variable p. For p much larger than incoming momenta and the mass, the integrand behaves like p^{-4}. That means that the integral depends logarithmically on the upper integration limit,

$$
\int^\Lambda \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2} \frac{i}{\left(p + p_1 - p_3\right)^2 - m^2} \xrightarrow{p \gg p_i, m} \int^\Lambda \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{-1}{p^4} \propto \ln \Lambda.
\tag{88}
$$

15

Divergent loop diagrams are ubiquitous in quantum field theory. They can be cured by regularization, i.e., making the integrals finite by introducing some cutoff parameter, and renormalization, where this additional parameter is removed in the end, yielding finite results for observables. This will be discussed in more detail in the section on quantum corrections.

2.5.2 Fermions

We can augment the theory by adding a fermionic field ψ, with a Lagrangian including an interaction with the scalar ϕ,

$$\mathscr{L}_\psi = \underbrace{\overline{\psi}\left(i\slashed{\partial} - m\right)\psi}_{\text{free Lagrangian}} - \underbrace{g\overline{\psi}\phi\psi}_{\text{interaction}} \;. \tag{89}$$

There are additional Feynman rules for fermions. The lines carry two arrows, one for the momentum as for the scalars and one for the fermion number flow, which basically distinguishes particles and antiparticles. The additional rules are:

i. $u(p)$ Incoming or outgoing particles get a factor of $u(p)$ or $\overline{u}(p)$, respectively.

              $\overline{u}(p)$

ii. $\overline{v}(p)$ Incoming or outgoing antiparticles get a factor of $\overline{v}(p)$ or $v(p)$, respectively.

              $v(p)$

iii. $\dfrac{i\left(\slashed{p} + m\right)}{p^2 - m^2 + i\varepsilon}$ Free propagator for fermion with momentum p.

iv. $-ig$ The fermion-fermion-scalar vertex yields a factor of the coupling constant. Again, there is no momentum dependence.

3 Gauge theories

In addition to spin-0 and spin-$\frac{1}{2}$ particles, the standard model contains spin-1 particles. They are the quanta of vector fields which can describe strong and electroweak interactions. The corresponding theories come with a local ("gauge") symmetry and are called gauge theories.

3.1 Global symmetries versus gauge symmetries

Consider a complex scalar field with the Lagrangian

$$\mathscr{L} = \partial_\mu \phi^\dagger \partial^\mu - V\left(\phi^\dagger \phi\right), \tag{90}$$

which is a generalization of the one considered in Eq. (49). This theory has a $\mathsf{U}(1)$ symmetry under which $\phi \to \phi' = \exp\{i\alpha\}\phi$ with constant parameter α. Usually it is sufficient to consider the variation of the fields and the Lagrangian under infinitesimal transformations,

$$\delta\phi = \phi' - \phi = i\alpha\phi, \qquad \delta\phi^\dagger = -i\alpha\phi^\dagger, \tag{91}$$

where terms $\mathcal{O}\left(\alpha^2\right)$ have been neglected. To derive the Noether current, Eq. (51), we compute the variation of the Lagrangian under such a transformation:

$$
\begin{aligned}
\delta\mathscr{L} &= \frac{\partial\mathscr{L}}{\partial\phi}\delta\phi + \frac{\partial\mathscr{L}}{\partial\phi^{\dagger}}\delta\phi^{\dagger} + \frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi\right)}\underbrace{\delta\left(\partial_{\mu}\phi\right)}_{=\partial_{\mu}\delta\phi} + \frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi^{\dagger}\right)}\delta\left(\partial_{\mu}\phi^{\dagger}\right) \\
&= \underbrace{\left(\frac{\partial\mathscr{L}}{\partial\phi} - \partial_{\mu}\frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi\right)}\right)}_{=0 \text{ by equation of motion}}\delta\phi + \underbrace{\left(\frac{\partial\mathscr{L}}{\partial\phi^{\dagger}} - \partial_{\mu}\frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi^{\dagger}\right)}\right)}_{=0}\delta\phi^{\dagger} \\
&\quad + \partial_{\mu}\left(\frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi\right)}\delta\phi + \frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi^{\dagger}\right)}\delta\phi^{\dagger}\right) \\
&= \alpha\partial_{\mu}\left(\mathrm{i}\partial^{\mu}\phi^{\dagger}\phi - \mathrm{i}\phi^{\dagger}\partial^{\mu}\phi\right) \\
&= -\alpha\partial_{\mu}j^{\mu}\,.
\end{aligned}
\tag{92}
$$

Since the Lagrangian is invariant, $\delta\mathscr{L} = 0$, we obtain a conserved current for solutions of the equations of motion,

$$
\partial_{\mu}j^{\mu} = 0\,.
\tag{93}
$$

From the first to the second line we have used

$$
\frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi\right)}\partial_{\mu}\delta\phi = \partial_{\mu}\left(\frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi\right)}\delta\phi\right) - \left(\partial_{\mu}\frac{\partial\mathscr{L}}{\partial\left(\partial_{\mu}\phi\right)}\right)\delta\phi
\tag{94}
$$

by the Leibniz rule.

The above procedure can be generalized to more complicated Lagrangians and symmetries. The derivation does not depend on the precise form of $\mathscr{L}$, and up to the second line of (92), it is independent of the form of $\delta\phi$. As a general result, a symmetry of the Lagrangian always implies a conserved current, which in turn gives a conserved quantity (often referred to as charge, but it can be angular momentum or energy as well).

What is the meaning of such a symmetry? Loosely speaking, it states that "physics does not change" under such a transformation. This, however, does not mean that the solutions to the equations of motion derived from this Lagrangian are invariant under such a transformation. Indeed, generically they are not, and only $\phi \equiv 0$ is invariant.

As an example, consider the Mexican hat potential,

$$
V(\phi^{\dagger}\phi) = -\mu^2\phi^{\dagger}\phi + \lambda\left(\phi^{\dagger}\phi\right)^2\,.
\tag{95}
$$

This potential has a ring of minima, namely all fields for which $|\phi|^2 = \mu^2/(2\lambda)$. This means that any constant ϕ with this modulus is a solution to the equation of motion,

$$
\Box\phi + \frac{\partial V}{\partial\phi}\left(\phi, \phi^{\dagger}\right) = \Box\phi - \phi^{\dagger}\left(\mu^2 - 2\lambda\phi^{\dagger}\phi\right) = 0\,.
\tag{96}
$$

These solutions are not invariant under $\mathsf{U}(1)$ phase rotations. On the other hand, it is obvious that any solution to the equations of motion will be mapped into another solution under such a transformation.

This situation is analogous to the Kepler problem: A planet moving around a stationary (very massive) star. The setup is invariant under spatial rotations around the star, i.e., the symmetries form the group $\mathsf{SO}(3)$. This group is three-dimensional (meaning that any rotation can be built from three independent rotations, e.g., around the three axes of a Cartesian coordinate system). Thus there are three

conserved charges which correspond to the three components of angular momentum. The solutions of this problem—the planet's orbits—are ellipses in a plane, so they are not at all invariant under spatial rotations, not even under rotations in the plane of motion. Rotated solutions, however, are again solutions.

In particle physics, most experiments are scattering experiments at colliders. For those, the statement that "physics does not change" translates into "transformed initial states lead to transformed final states": If one applies the transformation to the initial state and performs the experiment, the result will be the same as if one had done the experiment with the untransformed state and transformed the result.

There is a subtle, but important, difference between this and another type of symmetry, gauge symmetry. A gauge transformation is also a transformation which leaves the Lagrangian invariant, but it does relate identical states which describe exactly the same physics.

This might be familiar from electrodynamics. One formulation uses electric and magnetic fields $\vec{E}$ and $\vec{B}$, together with charge and current densities ρ and $\vec{j}$. These fields and sources are related by Maxwell's equations:

$$\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0\,, \qquad\qquad \vec{\nabla} \cdot \vec{B} = 0\,, \qquad (97a)$$

$$\vec{\nabla} \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{j}\,, \qquad\qquad \vec{\nabla} \cdot \vec{E} = \rho\,. \qquad (97b)$$

The first two of these can be identically solved by introducing the potentials ϕ and $\vec{A}$, which yield $\vec{E}$ and $\vec{B}$ via

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}\,, \qquad\qquad \vec{B} = \vec{\nabla} \times \vec{A}\,. \qquad (98)$$

So we have reduced the six components of $\vec{E}$ and $\vec{B}$ down to the four independent ones ϕ and $\vec{A}$. However, the correspondence between the physical fields and the potentials is not unique. If some potentials ϕ and $\vec{A}$ lead to certain $\vec{E}$ and $\vec{B}$ fields, the transformed potentials

$$\vec{A}' = \vec{A} + \vec{\nabla}\Lambda\,, \qquad\qquad \phi' = \phi - \frac{\partial \Lambda}{\partial t}\,, \qquad (99)$$

where Λ is a scalar field, give the same electric and magnetic fields.

This transformation (98) is called gauge transformation. It is a symmetry of the theory, but it is different from the global symmetries we considered before. First, it is a local transformation, i.e., the transformation parameter Λ varies in space and time. Second, it relates physically indistinguishable field configurations, since solutions of the equations of motion for electric and magnetic fields are invariant. It is important to note that this gauge transformation is inhomogeneous, i.e., the variation is not multiplicative, but can generate non-vanishing potentials from zero. Potentials that are related to $\phi = 0$ and $\vec{A} = 0$ by a gauge transformation are called pure gauge.

Phrased differently, we have expressed the physical fields $\vec{E}$ and $\vec{B}$ in terms of the potentials ϕ and $\vec{A}$. These potentials still contain too many degrees of freedom for the physical fields $\vec{E}$ and $\vec{B}$, since different potentials can lead to the same $\vec{E}$ and $\vec{B}$ fields. So the description in terms of potentials is redundant, and the gauge transformation (99) quantifies just this redundancy. Physical states and observables have to be invariant under gauge transformations.

3.2 Abelian gauge theories

The easiest way to come up with a gauge symmetry is to start from a global symmetry and promote it to a gauge one, that is, demand invariance of the Lagrangian under local transformations (where the transformation parameter is a function of spacetime). To see this, recall the Lagrangian with the global $U(1)$ symmetry from the preceding section,

$$\mathscr{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - V(\phi^\dagger \phi)\,,$$

and the transformation

$$\phi \to \phi' = e^{i\alpha}\phi\,, \quad \delta\phi = \phi' - \phi = i\alpha\phi\,.$$

If we now allow spacetime-dependent parameters $\alpha(x)$, the Lagrangian is no longer invariant. The potential part still is, but the kinetic term picks up derivatives of $\alpha(x)$, so the variation of the Lagrangian is

$$\delta\mathscr{L} = i\partial_\mu\alpha\left(\partial^\mu\phi^\dagger\phi - \phi^\dagger\partial^\mu\phi\right) = -\partial_\mu\alpha\, j^\mu\,, \tag{100}$$

the derivative of α times the Noether current of the global symmetry derived before.

The way to restore invariance of the Lagrangian is to add another field, the gauge field, with a gauge transformation just like the electromagnetic potentials in the previous section, combined into a four-vector $A^\mu = (\phi, \vec{A})$:

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x)\,. \tag{101}$$

The factor $\frac{1}{e}$ is included for later convenience. We can now combine the inhomogeneous transformation of A_μ with the inhomogeneous transformation of the derivative in a covariant derivative D_μ:

$$D_\mu\phi = (\partial_\mu + ieA_\mu)\,\phi\,. \tag{102}$$

This is called covariant derivative because the differentiated object $D_\mu\phi$ transforms in the same way as the original field,

$$\begin{aligned}
D_\mu\phi \longrightarrow (D_\mu\phi)' &= \left(\partial_\mu + ieA'_\mu\right)\phi' \\
&= \partial_\mu\left(e^{i\alpha(x)}\phi\right) + ie\left(A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x)\right)e^{i\alpha(x)}\phi \\
&= e^{i\alpha(x)}D_\mu\phi\,.
\end{aligned} \tag{103}$$

So we can construct an invariant Lagrangian from the field and its covariant derivative:

$$\mathscr{L} = (D_\mu\phi)^\dagger\left(D^\mu\phi\right) - V\left(\phi^\dagger\phi\right)\,. \tag{104}$$

So far this is a theory of a complex scalar with $\mathsf{U}(1)$ gauge invariance. The gauge field A_μ, however, is not a dynamical field, i.e., there is no kinetic term for it. This kinetic term should be gauge invariant and contain derivatives up to second order. In order to find such a kinetic term, we first construct the field strength tensor from the commutator of two covariant derivatives:

$$\begin{aligned}
F_{\mu\nu} &= -\frac{i}{e}\left[D_\mu, D_\nu\right] = -\frac{i}{e}\left[\left(\partial_\mu + ieA_\mu\right), \left(\partial_\nu + ieA_\nu\right)\right] \\
&= -\frac{i}{e}\left(\left[\partial_\mu, \partial_\nu\right] + \left[\partial_\nu, ieA_\nu\right] + \left[ieA_\mu, \partial_\nu\right] - e^2\left[A_\mu, A_\nu\right]\right) \\
&= \partial_\mu A_\nu - \partial_\nu A_\mu\,.
\end{aligned} \tag{105}$$

To check that this is a sensible object to construct, we can redecompose A_μ into the scalar and vector potential ϕ and $\vec{A}$ and spell out the field strength tensor in electric and magnetic fields,

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_2 & 0 \end{pmatrix}\,. \tag{106}$$

This shows that the field strength is gauge invariant, as are $\vec{E}$ and $\vec{B}$. Of course, this can also be shown by straightforward calculation,

$$\delta F_{\mu\nu} = \partial_\mu \delta A_\nu - \partial_\nu \delta A_\mu = -\frac{1}{e} \left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \alpha(x) = 0 \,, \tag{107}$$

so it is just the antisymmetry in μ and ν that ensures gauge invariance.

The desired kinetic term is now just the square of the field strength tensor,

$$\mathscr{L}_{\text{gaugekin}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,, \tag{108a}$$

or, in terms of $\vec{E}$ and $\vec{B}$ fields,

$$\mathscr{L} = \frac{1}{2} \left(\vec{E}^2 - \vec{B}^2 \right) \,. \tag{108b}$$

The coupling to scalar fields via the covariant derivative can also be applied to fermions. To couple a fermion ψ to the gauge field, one simply imposes the gauge transformation

$$\psi \rightarrow \psi' = e^{\mathrm{i}\alpha} \psi \,. \tag{109}$$

In the Lagrangian, one again replaces the ordinary derivative with the covariant one. The Lagrangian for a fermion coupled to a $\mathsf{U}(1)$ gauge field is quantum electrodynamics (QED), if we call the fields electron and photon:

$$\mathscr{L}_{\text{QED}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \overline{\psi} \left(\mathrm{i} \slashed{D} - m \right) \psi \,. \tag{110}$$

Finally, let us note that for a $\mathsf{U}(1)$ gauge theory, different fields may have different charges under the gauge group (as, for example, quarks and leptons indeed do). For fields with charge q (in units of elementary charge), we have to replace the gauge transformations and consequently the covariant derivative as follows:

$$\psi_q \rightarrow \psi_q' = e^{\mathrm{i}q\alpha} \psi_q \,, \qquad\qquad D_\mu^{(q)} \psi_q = \left(\partial_\mu + \mathrm{i}qeA_\mu \right) \psi_q \,. \tag{111}$$

What have we done so far? We started from a Lagrangian, Eq. (90) with a global $\mathsf{U}(1)$ symmetry (91). We imposed invariance under local transformations, so we had to introduce a new field, the gauge field A_μ. This field transformed inhomogeneously under gauge transformations, just in a way to make a covariant derivative. This covariant derivative was the object that coupled the gauge field to the other fields of the theory. To make this into a dynamical theory, we added a kinetic term for the gauge field, using the field strength tensor. Alternatively, we could have started with the gauge field and tried to couple it to other fields, and we would have been led to the transformation properties (91). This is all we need to construct the Lagrangian for QED. For QCD and the electroweak theory, however, we need a richer structure: non-Abelian gauge theories.

3.3 Non-Abelian gauge theories

To construct non-Abelian theories in the same way as before, we first have to discuss non-Abelian groups, i.e., groups whose elements do not commute. We shall focus on the groups $\mathsf{SU}(n)$, since they are most relevant for the Standard Model. $\mathsf{SU}(n)$ is the group of $n \times n$ complex unitary matrices with determinant 1. To see how many degrees of freedom there are, we have to count: A $n \times n$ complex matrix U has n^2 complex entries, equivalent to $2n^2$ real ones. The unitarity constraint, $U^\dagger U = \mathbb{1}$, is a matrix equation, but not all component equations are independent. Actually, $U^\dagger U$ is Hermitian, $\left(U^\dagger U \right)^\dagger = U^\dagger U$, so

the diagonal entries are real and the lower triangle is the complex conjugate of the upper one. Thus, there are $n + 2 \cdot \frac{1}{2}n(n-1)$ real constraints. Finally, by taking the determinant of the unitarity constraint, $\det\left(U^\dagger U\right) = |\det U|^2 = 1$. Hence, restricting to $\det U = 1$ eliminates one more real degree of freedom. All in all, we have $2n^2 - n - 2 \cdot \frac{1}{2}n(n-1) - 1 = n^2 - 1$ real degrees of freedom in the elements of $\mathsf{SU}(n)$.

This means that any $U \in \mathsf{SU}(n)$ can be specified by $n^2 - 1$ real parameters α_a. The group elements are usually written in terms of these parameters and $n^2 - 1$ matrices T^a, the generators of the group, as an exponential

$$U = \exp\left\{\mathrm{i}\alpha_a T^a\right\} = \mathbb{1} + \mathrm{i}\alpha_a T^a + \mathcal{O}\left(\alpha^2\right), \tag{112}$$

and one often considers only infinitesimal parameters.

The generators are usually chosen as Hermitian matrices[3]. The product of group elements translates into commutation relations for the generators,

$$\left[T^a, T^b\right] = \mathrm{i}f^{abc}T^c, \tag{113}$$

with the antisymmetric structure constants f^{abc}, which of course also depend on the choice of generators.

In the Standard Model, the relevant groups are $\mathsf{SU}(2)$ for the electroweak theory and $\mathsf{SU}(3)$ for QCD. $\mathsf{SU}(2)$ has three parameters. The generators are usually chosen to be the Pauli matrices, $T^a = \frac{1}{2}\sigma^a$, whose commutation relations are $\left[\sigma^a, \sigma^b\right] = \mathrm{i}\varepsilon^{abc}\sigma^c$. The common generators of $\mathsf{SU}(3)$ are the eight Gell-Mann matrices, $T^a = \frac{1}{2}\lambda^a$.

To construct a model with a global $\mathsf{SU}(n)$ symmetry, we consider not a single field, but an n-component vector Φ_i, $i = 1, \ldots, n$ (called a multiplet of $\mathsf{SU}(n)$), on which the matrices of $\mathsf{SU}(n)$ act by multiplication:

$$\Phi = \begin{pmatrix} \Phi_1 \\ \vdots \\ \Phi_n \end{pmatrix} \longrightarrow \Phi' = U\Phi, \quad \Phi^\dagger = \left(\Phi_1^\dagger, \cdots, \Phi_n^\dagger\right) \longrightarrow \left(\Phi^\dagger\right)' = \Phi^\dagger U^\dagger. \tag{114}$$

Now we see why we want unitary matrices U: A product $\Phi^\dagger\Phi$ is invariant under such a transformation. This means that we can generalize the Lagrangian (90) in a straightforward way to include a non-Abelian symmetry:

$$\mathscr{L} = (\partial_\mu\Phi)^\dagger(\partial^\mu\Phi) - V\left(\Phi^\dagger\Phi\right). \tag{115}$$

If we allow for local transformations $U = U(x)$, we immediately encounter the same problem as before: The derivative term is not invariant, because the derivatives act on the matrix U as well,

$$\partial_\mu\Phi \rightarrow \partial_\mu\Phi' = \partial_\mu\left(U\Phi\right) = U\partial_\mu\Phi + \left(\partial_\mu U\right)\Phi. \tag{116}$$

To save the day, we again need to introduce a covariant derivative consisting of a partial derivative plus a gauge field. This time, however, the vector field needs to be matrix-valued, i.e., $A_\mu = A_\mu^a T^a$, where T^a are the generators of the group. We clearly need one vector field per generator, as each generator represents an independent transformation in the group.

The transformation law of A_μ is chosen such that the covariant derivative is covariant,

$$\begin{aligned} (D_\mu\Phi)' &= \left[\left(\partial_\mu + \mathrm{i}gA_\mu\right)\Phi\right]' \\ &= \left(\partial_\mu + \mathrm{i}gA_\mu'\right)(U\Phi) \\ &= U\left(\partial_\mu + U^{-1}\left(\partial_\mu U\right) + \mathrm{i}gU^{-1}A_\mu' U\right)\Phi \\ &\overset{!}{=} UD_\mu\Phi. \end{aligned} \tag{117}$$

---

[3]Actually, the generators live in the Lie algebra of the group, and so one can choose any basis one likes, Hermitian or not.

This requirement fixes the transformation of A_μ to be

$$A'_\mu = U A_\mu U^{-1} - \frac{\mathrm{i}}{g} U \partial_\mu U^{-1} \,. \tag{118}$$

In the Abelian case this reduces to the known transformation law, Eq. (101).

For infinitesimal parameters $\alpha = \alpha^a T^a$, the matrix $U = \exp\{\mathrm{i}\alpha\} = 1 + \mathrm{i}\alpha$, and Eq. (118) becomes

$$A'_\mu = A_\mu - \frac{1}{g}\partial_\mu \alpha + \mathrm{i}\,[\alpha, A_\mu] \,, \tag{119}$$

or for each component

$$A^{a'}_\mu = A^a_\mu - \frac{1}{g}\partial_\mu \alpha^a - f^{abc}\alpha^b A^c_\mu \,. \tag{120}$$

Sometimes it is convenient to write down the covariant derivative in component form:

$$(D_\mu \Phi)_i = \left(\partial_\mu \delta_{ij} + \mathrm{i}g T^a_{ij} A^a_\mu\right)\Phi_j \,. \tag{121}$$

Next we need a kinetic term, which again involves the field strength, the commutator of covariant derivatives:

$$\begin{aligned}
F_{\mu\nu} &= -\frac{\mathrm{i}}{g}\,[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + \mathrm{i}g\,[A_\mu, a_\nu] = F^a_{\mu\nu}T^a \,, \\
F^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f^{abc} A^b_\mu A^c_\nu \,.
\end{aligned} \tag{122}$$

Now we see that the field strength is more that just the derivative: There is a quadratic term in the potentials. This leads to a self-interaction of gauge fields, like in QCD, where the gluons interact with each other. This is the basic reason for confinement, unlike in QED, where the photon is not charged.

Furthermore, when we calculate the transformation of the field strength, we find that it is not invariant, but transforms as

$$F_{\mu\nu} \to F'_{\mu\nu} = U F_{\mu\nu} U^{-1} \,, \tag{123}$$

i.e., it is covariant. There is an easy way to produce an invariant quantity out of this: the trace. Since $\operatorname{tr} AB = \operatorname{tr} BA$, the Lagrangian

$$\mathscr{L} = -\frac{1}{2}\operatorname{tr}\left(F_{\mu\nu}F^{\mu\nu}\right) = -\frac{1}{4}F^a_{\mu\nu}F^{a\,\mu\nu} \tag{124}$$

is indeed invariant, as $\operatorname{tr}\left(U F^2 U^{-1}\right) = \operatorname{tr}\left(U^{-1} U F^2\right) = \operatorname{tr} F^2$. In the second step we have used a normalization convention,

$$\operatorname{tr}\left(T^a T^b\right) = \frac{1}{2}\delta^{ab} \,, \tag{125}$$

and every generator is necessarily traceless. The factor $\frac{1}{2}$ is arbitrary and could be chosen differently, with compensating changes in the coefficient of the kinetic term.

By choosing the gauge group $\mathsf{SU}(3)$ and coupling the gauge field to fermions, the quarks, we can write down the Lagrangian of quantum chromodynamics (QCD):

$$\mathscr{L}_{\mathrm{QCD}} = -\frac{1}{4}G^a_{\mu\nu}G^{a\,\mu\nu} + \overline{q}\left(\mathrm{i}\slashed{D} - m\right)q \,, \tag{126}$$

where $a = 1,\ldots,8$ counts the gluons and q is a three-component (i.e. three-colour) quark.

3.4 Quantization

So far we have discussed only classical gauge theories. If we want to quantize the theory and find the Feynman rules for diagrams involving gauge fields, we have a problem: we have to make sure we do not count field configurations of A_μ which are pure gauge, nor count separately fields which differ only by a gauge transformation, since those are meant to be physically identical. On the more technical side, the naïve Green function for the free equation of motion does not exist. In the Abelian case, the equation is

$$\partial_\mu F^{\mu\nu} = \Box A^\nu - \partial^\nu \partial_\mu A^\mu = (\Box g^{\mu\nu} - \partial^\nu \partial^\mu) A_\mu = 0. \tag{127}$$

The Green function should be the inverse of the differential operator in brackets, but the operator is not invertible. Indeed, it annihilates every pure gauge mode, as it should,

$$(\Box g^{\mu\nu} - \partial^\nu \partial^\mu) \partial_\mu \Lambda = 0, \tag{128}$$

so it has zero eigenvalues. Hence, the propagator must be defined in a more clever way.
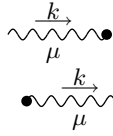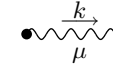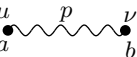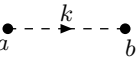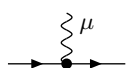
One way out would be to fix the gauge, i.e., simply demand a certain gauge condition like $\vec{\nabla} \cdot \vec{A} = 0$ (Coulomb gauge) or $n_\mu A^\mu = 0$ with a fixed 4-vector (axial gauge). It turns out, however, that the loss of Lorentz invariance causes many problems in calculations.
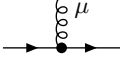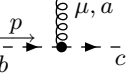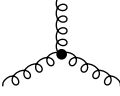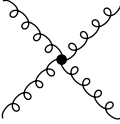
A better way makes use of Faddeev–Popov ghosts. In this approach, we add two terms to the Lagrangian, the gauge-fixing term and the ghost term. The gauge-fixing term is not gauge invariant, but rather represents a certain gauge condition which can be chosen freely. The fact that it is not gauge invariant means that now the propagator is well defined, but the price to pay is that it propagates too many degrees of freedom, namely gauge modes. This is compensated by the propagation of ghosts, strange fields which are scalars but which anticommute and do not show up as physical states but only as internal lines in loop calculations. It turns out that gauge invariance is not lost but rather traded for a different symmetry, BRST symmetry, which ensures that we get physically sensible results.

For external states, we have to restrict to physical states, of which there are two for massless bosons. They are labelled by two polarization vectors $\epsilon_\mu^\pm$ which are transverse, i.e., orthogonal to the momentum four-vector and the spatial momentum, $k_\mu \epsilon^\mu = \vec{k}\vec{\epsilon} = 0$.

The form of the gauge fixing and ghost terms depends on the gauge condition we want to take. A common (class of) gauge is the covariant gauge which depends on a parameter ξ, which becomes Feynman gauge (Landau/Lorenz gauge) for $\xi = 1$ ($\xi = 0$).

We now list the Feynman rules for a non-Abelian gauge theory (QCD) coupled to fermions (quarks) and ghosts. The fermionic external states and propagators are listed in Section 2.5.2.

i.	$\epsilon_\mu(k)$	For each external line one has a polarization vector.
	$\epsilon_\mu^*(k)$	
ii.	$\dfrac{-\mathrm{i}\delta^{ab}}{k^2 + \mathrm{i}\varepsilon}$ $\times \left(g_{\mu\nu} + (1-\xi)\dfrac{k_\mu k_\nu}{k^2}\right)$	The propagator for gauge bosons contains the parameter ξ.
iii.	$\dfrac{-\mathrm{i}\delta^{ab}}{k^2 + \mathrm{i}\varepsilon}$	The propagator for ghosts is the one of scalar particles. There are no external ghost states.
iv.	$\mathrm{i}e\gamma^\mu$	In QED, there is just one vertex between photon and electron.

v.		$i\frac{g}{2}\gamma^\mu\lambda^a$	In QCD, the basic quark–quark–gluon vertex involves the Gell-Mann matrices.
vi.		$-gf^{abc}p^\mu$	The ghosts couple to the gauge field.
vii.		$gf^{abc}k_\mu +$ permutations	Three-gluon self-interaction.
viii.		$-\frac{1}{4}g^2f^{abc}f^{ade}g^{\mu\nu}g^{\rho\sigma}$ $+$permutations	Four-gluon self-interaction.

4 Quantum corrections

Now that we have the Feynman rules, we are ready to calculate quantum corrections [3, 5, 9]. As a first example we consider the anomalous magnetic moment of the electron at one-loop order. This was historically, and still is today, one of the most important tests of quantum field theory. The calculation is still quite simple because the one-loop expression is finite. In most cases, however, one encounters divergent loop integrals. In the following sections we shall study these divergences and show how to remove them by renormalization. Finally, as an application, we shall discuss the running of coupling constants and asymptotic freedom.

4.1 Anomalous magnetic moment

The magnetic moment of the electron determines its energy in a magnetic field,

$$H_{\text{mag}} = -\vec{\mu}\cdot\vec{B}\,. \tag{129}$$

For a particle with spin $\vec{s}$, the magnetic moment is aligned in the direction of $\vec{s}$, and for a classical spinning particle of mass m and charge e, its magnitude would be the Bohr magneton, $e/(2m)$. In the quantum theory, the magnetic moment is different, which is expressed by the Landé factor g_e,

$$\vec{\mu}_e = g_e\frac{e}{2m}\vec{s}\,. \tag{130}$$

We now want to calculate g_e in QED. To lowest order, this just means solving the Dirac equation in an external electromagnetic field $A^\mu = (\phi, \vec{A})$,

$$\left(i\slashed{D} - m\right)\psi = \left[\gamma^\mu\left(i\partial_\mu - eA_\mu\right) - m\right]\psi = 0\,. \tag{131}$$

For a bound non-relativistic electron a stationary solution takes the form

$$\psi(x) = \begin{pmatrix}\varphi(\vec{x})\\\chi(\vec{x})\end{pmatrix}e^{-iEt}\,,\quad\text{with}\quad\frac{E-m}{m}\ll 1\,. \tag{132}$$

It is convenient to use the following representation of the Dirac matrices:

$$\gamma^0 = \begin{pmatrix}\mathbb{1} & 0\\0 & -\mathbb{1}\end{pmatrix},\qquad\qquad\gamma^i = \begin{pmatrix}0 & \sigma^i\\-\sigma^i & 0\end{pmatrix}\,. \tag{133}$$
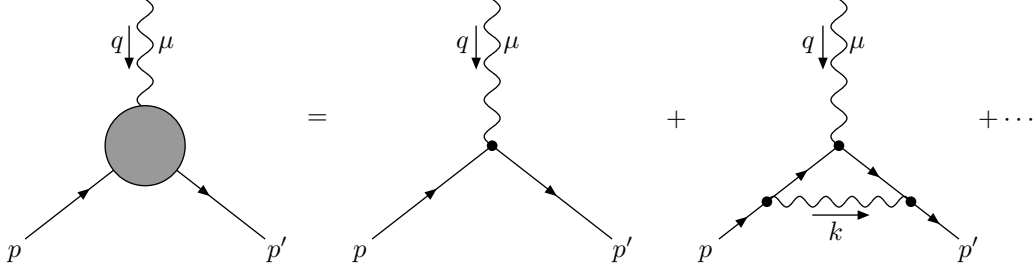
Fig. 8: Tree-level and one-loop diagram for the magnetic moment

One then obtains the two coupled equations

$$[(E - e\phi) - m]\,\varphi - \left(-\mathrm{i}\vec{\nabla} - e\vec{A}\right) \cdot \vec{\sigma}\chi = 0\,, \tag{134a}$$

$$\big[\underbrace{-(E - e\phi) - m}_{\approx -2m}\big]\chi + \left(-\mathrm{i}\vec{\nabla} - e\vec{A}\right) \cdot \vec{\sigma}\varphi = 0\,. \tag{134b}$$

The coefficient of χ in the second equation is approximately independent of ϕ, so we can solve the equation to determine χ in terms of φ,

$$\chi = \frac{1}{m}\left(-\mathrm{i}\vec{\nabla} - e\vec{A}\right) \cdot \vec{\sigma}\varphi\,. \tag{135}$$

Inserting this into (134a), we get the Pauli equation,

$$\left[\frac{1}{2m}\left(-\mathrm{i}\vec{\nabla} - e\vec{A}\right)^2 + e\phi - \frac{e}{2m}\vec{B}\cdot\vec{\sigma}\right]\varphi = (E - m)\,\varphi\,. \tag{136}$$

This is a Schrödinger-like equation which implies (since $\vec{s} = \frac{1}{2}\vec{\sigma}$),

$$H_{\mathrm{mag}} = -2\frac{e}{2m}\vec{s}\vec{B}\,. \tag{137}$$

Hence, the Landé factor of the electron is $g_e = 2$.

In QED, the magnetic moment is modified by quantum corrections. The magnetic moment is the spin-dependent coupling of the electron to a photon in the limit of vanishing photon momentum. Diagrammatically, it is contained in the blob on the left side of Fig. 8, which denotes the complete electron–photon coupling. The tree-level diagram is the fundamental electron–photon coupling. There are several one-loop corrections to this diagram, but only the so-called vertex correction, where an internal photon connects the two electron lines, gives a contribution to the magnetic moment. All other one-loop diagrams concern only external legs, such as an electron–positron bubble on the incoming photon, and will be removed by wave-function renormalization.

The expression for the tree-level diagram is

$$\mathrm{i}\bar{u}(p')e\gamma^\mu u(p)\,. \tag{138}$$

Note that the photon becomes on-shell only for $q \to 0$, so no polarization vector is included. The matrix element of the electromagnetic current can be decomposed via the Gordon identity into convection and spin currents,

$$\bar{u}(p')\gamma^\mu u(p) = \bar{u}(p')\left(\frac{(p + p')^\mu}{2m} + \frac{\mathrm{i}}{2m}\sigma^{\mu\nu}\left(p' - p\right)_\nu\right)u(p)\,. \tag{139}$$

Here the first term can be viewed as the net flow of charged particles, the second one is the spin current. Only this one is relevant for the magnetic moment, since it gives the spin-dependent coupling of the electron.

In order to isolate the magnetic moment from the loop diagram, we first note that the corresponding expression will contain the same external states, so it can be written as

$$i\overline{u}(p')e\Gamma^\mu(p,q)u(p)\,, \quad q = p' - p\,, \tag{140}$$

where $\Gamma^\mu(p,q)$ is the correction to the vertex due to the exchange of the photon. We can now decompose Γ^μ into different parts according to index structure and extract the term $\propto \sigma^{\mu\nu}$. Using the Feynman rules, we find for Γ^μ in Feynman gauge ($\xi = 1$),

$$
\begin{aligned}
i e \Gamma^\mu(p,q) = (-ie)^3 \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{-i g_{\rho\sigma}}{k^2 + i\varepsilon} \gamma^\rho \frac{i\left(\slashed{p}' - \slashed{k} + m\right)}{(p'-k)^2 - m^2 + i\varepsilon} \gamma^\mu \\
\times \frac{i\left(\slashed{p} - \slashed{k} + m\right)}{(p-k)^2 - m^2 + i\varepsilon} \gamma^\sigma\,.
\end{aligned}
\tag{141}
$$

This integral is logarithmically divergent, as can be seen by power counting, since the leading term is $\propto k^2$ in the numerator and $\propto k^6$ in the denominator.

On the other hand, the part $\propto \sigma^{\mu\nu} q_\mu$ is finite and can be extracted via some tricks:

– Consider first the denominator of the integral (141). It is the product of three terms of the form $(\text{momentum})^2 - m^2$, which can be transformed into a sum at the expense of further integrations over the so-called Feynman parameters x_1 and x_2,

$$\frac{1}{A_1 A_2 A_3} = 2 \int_0^1 \mathrm{d}x_1 \int_0^{1-x_1} \mathrm{d}x_2 \frac{1}{\left[A_1 x_1 + A_2 x_2 + A_3 \left(1 - x_1 - x_2\right)\right]^3}\,. \tag{142}$$

– After introducing the Feynman parameters, the next trick is to shift the integration momentum $k \to k'$, where

$$A_1 x_1 + A_2 x_2 + A_3 \left(1 - x_1 - x_2\right) = \underbrace{\left(k - x_1 p' - x_2 p\right)^2}_{k'} - \left(x_1 p' + x_2 p\right)^2 + i\varepsilon\,. \tag{143}$$

Note that one must be careful when manipulating divergent integrals. In principle, one should first regularize them and then perform the shifts on the regularized integrals, but in this case, there is no problem.

– For the numerator, the important part is the Dirac algebra of γ matrices. A standard calculation gives (see Appendix)

$$
\begin{aligned}
\gamma^\nu \left(\slashed{p}' - \slashed{k} + m\right) \gamma^\mu \left(\slashed{p} - \slashed{k} + m\right) \gamma_\nu \\
= -2m^2 \gamma_\mu - 4im\sigma^{\mu\nu}\left(p' - p\right)_\nu - 2\slashed{p}\gamma_\mu\slashed{p}' + \mathcal{O}(k) + \mathcal{O}\left(k^2\right)\,.
\end{aligned}
\tag{144}
$$

Here we have used again the Gordon formula to trade $(p + p')_\nu$ for $\sigma_{\nu\rho}q^\rho$, which is allowed only if the expression is sandwiched between on-shell spinors $\overline{u}(p')$ and $u(p)$.

– Now the numerator is split into pieces independent of k, linear and quadratic in k. The linear term can be dropped under the integral. The quadratic piece leads to a divergent contribution which we shall discuss later. The integral over the k-independent part in the limit $q^\mu \to 0$ yields

$$\int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{1}{\left[k^2 - (x_1 + x_2)^2 m^2 + i\varepsilon\right]^3} = -\frac{i}{32\pi^2} \frac{1}{(x_1 + x_2)^2 m^2}\,. \tag{145}$$

Now all that is left are the parameter integrals over x_1 and x_2.

Finally, one obtains the result, usually expressed in terms of the fine structure constant $\alpha = e^2/(4\pi)$,

$$\mathrm{i}e\overline{u}(p')\Gamma^\mu u(p) = +\mathrm{i}e\overline{u}(p')\left(\frac{\alpha}{2\pi}\frac{\mathrm{i}}{2m}\sigma^{\mu\nu}q_\nu + \cdots\right)u(p), \tag{146}$$

where the dots represent contributions which are not $\propto \sigma^{\mu\nu}q_\nu$.

Comparison with the Gordon decomposition (139) gives the one-loop correction to the Landé factor,

$$g = 2\left(1 + \frac{\alpha}{2\pi}\right). \tag{147}$$

This correction was first calculated by Schwinger in 1948. It is often expressed as the anomalous magnetic moment a_e,

$$a_e = \frac{g-2}{2}. \tag{148}$$

Today, a_e is known up to three loops analytically and to four loops numerically [10]. The agreement of theory and experiment is impressive:

$$\begin{aligned} a_e^{\mathrm{exp}} &= (1159652185.9 \pm 3.8) \cdot 10^{-12}, \\ a_e^{\mathrm{th}} &= (1159652175.9 \pm 8.5) \cdot 10^{-12}. \end{aligned} \tag{149}$$

This is one of the cornerstones of our confidence in quantum field theory.

4.2 Divergences

The anomalous magnetic moment we calculated in the last section was tedious work, but at least the result was finite. Most other expressions, however, have divergent momentum integrals. One such example is the vertex function Γ^μ we already considered. It has contributions which are logarithmically divergent. We can isolate these by setting $q = 0$, which yields

$$\Gamma^\mu(p,0) = -2\mathrm{i}e^2\int_0^1 \mathrm{d}x_1 \int_0^{1-x_1}\mathrm{d}x_2 \int \frac{\mathrm{d}^4k}{(2\pi)^2}\frac{\gamma^\nu\slashed{k}\gamma_\mu\slashed{k}\gamma_\nu}{\left[k^2 - (x_1+x_2)^2 m^2 + \mathrm{i}\varepsilon\right]^3}. \tag{150}$$

This expression is treated in two steps:

– First we make the integral finite in a step called regularization. In this step, we have to introduce a new parameter of mass dimension 1. An obvious choice would be a cutoff Λ which serves as an upper bound for the momentum integration. One might even argue that there should be a cutoff at a scale where quantum gravity becomes important, although a regularization parameter has generally no direct physical meaning.

– The second step is renormalization. The divergences are absorbed into the parameters of the theory. The key idea is that the 'bare' parameters which appear in the Lagrangian are not physical, so they can be divergent. Their divergences are chosen such as to cancel the divergences coming from the divergent integrals.

– Finally, the regulator is removed. Since all divergences have been absorbed into the parameters of the theory, the results remain finite for infinite regulator. Of course, one has to make sure the results do not depend on the regularization method.

The cutoff regularization, while conceptually simple, is not a convenient method, as it breaks Lorentz and gauge invariance. Symmetries, however, are very important for all calculations, so a good regularization scheme should preserve as many symmetries as possible. We shall restrict ourselves to dimensional regularization, which is the most common scheme used nowadays.

4.2.1 Dimensional regularization

The key idea is to define the theory not in four, but in $d = 4 - \epsilon$ dimensions [9]. If ϵ is not an integer, the integrals do converge. Non-integer dimensionality might seem weird, but in the end we shall take the limit of $\epsilon \to 0$ and return to four dimensions. This procedure is well defined and just an intermediate step in the calculation.

Let us consider some technical issues. In d dimensions, the Lorentz indices 'range from 0 to d', in the sense that

$$g^{\mu\nu} g_{\nu\mu} = d \,, \tag{151}$$

and there are d γ-matrices obeying the usual algebra,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1} \,, \quad \mathrm{tr}\,(\mathbb{1}) = 4 \,. \tag{152}$$

The γ-matrix contractions are also modified due to the change in the trace of $g_{\mu\nu}$, such as

$$\gamma^\mu \gamma^\nu \gamma_\mu = -\left(2 - \epsilon\right)\gamma^\nu \,, \quad \gamma^\mu \gamma^\nu \gamma^\rho \gamma_\mu = 4g^{\nu\rho} - \epsilon \gamma^\nu \gamma^\rho \,. \tag{153}$$

The tensor structure of diagrams can be simplified as follows. If a momentum integral over k contains a factor of $k_\mu k_\nu$, this must be proportional to $g_{\mu\nu} k^2$, since it is of second order in k and symmetric in $(\mu\nu)$. The only symmetric tensor we have is the metric (as long as the remaining integrand depends only on the square of k and the squares of the external momenta p_i), and the coefficient can be obtained by contracting with $g^{\mu\nu}$ to yield

$$\int \frac{\mathrm{d}^4 k}{(2\pi)^4} k_\mu k_\nu f\left(k^2, p_i^2\right) = \frac{1}{d}\, g_{\mu\nu} \int \frac{\mathrm{d}^4 k}{(2\pi)^4} k^2 f\left(k^2, p_i^2\right) \,. \tag{154}$$

The measure of an integral changes from $\mathrm{d}^4 k$ to $\mathrm{d}^d k$. Since k is a dimensionful quantity[4] (of mass dimension 1), we need to compensate the change in dimensionality by a factor of μ^ϵ, where μ is an arbitrary parameter of mass dimension 1. The mass dimensions of fields and parameters also change. They can be derived from the condition that the action, which is the d-dimensional integral over the Lagrangian, be dimensionless. Schematically (i.e., without all numerical factors), a Lagrangian of gauge fields, scalars and fermions reads

$$\begin{aligned} \mathscr{L} &= (\partial_\mu A_\nu)^2 + e\partial_\mu A^\mu A_\nu A^\nu + e^2 \left(A_\mu A^\mu\right)^2 \\ &+ (\partial_\mu \phi)^2 + \overline{\psi}\left(\mathrm{i}\slashed{\partial} - m\right)\psi + e\overline{\psi}\slashed{A}\psi + m^2\phi^2 + \cdots \,. \end{aligned} \tag{155}$$

The condition of dimensionless action, $[S] = 0$, translates into $[\mathscr{L}] = d$, since $\left[\mathrm{d}^d x\right] = -d$. Derivatives have mass dimension 1, and so do masses. That implies for the dimensions of the fields (and the limit as $d \to 4$),

$$[A_\mu] = \frac{d-2}{2} \to 1 \,, \qquad\qquad [\phi] = \frac{d-2}{2} \to 1 \,, \tag{156}$$

$$[\psi] = \frac{d-1}{2} \to \frac{3}{2} \,, \qquad\qquad [e] = 2 - \frac{d}{2} \to 0 \,. \tag{157}$$

How do we evaluate a d-dimensional integral? One first transforms to Euclidean space replacing k^0 by $\mathrm{i}k_4$, so that the Lorentzian measure $\mathrm{d}^d k$ becomes $\mathrm{d}^d k_\mathrm{E}$. In Euclidean space, one can easily convert

---

[4]In our units where $\hbar = c = 1$, the only dimension is mass, so everything can be expressed in powers of GeV. The basic quantities have $[\mathrm{mass}] = [\mathrm{energy}] = [\mathrm{momentum}] = 1$ and $[\mathrm{length}] = [\mathrm{time}] = -1$, so $[\mathrm{d}x^\mu] = -1$ and $[\partial_\mu] = 1$.

(a) The electron self-energy

(b) The vacuum polarization
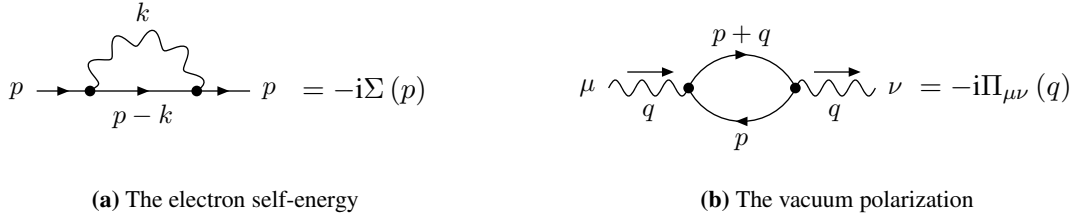
Fig. 9: One-loop corrections to the propagators of electron and photon

to spherical coordinates and perform the integral over the angular variables, which gives the 'area' of the d-dimensional 'unit sphere',

$$\int \frac{\mathrm{d}^d k_{\mathrm{E}}}{(2\pi)^d} f\left(k^2\right) = \underbrace{\int \frac{\mathrm{d}\Omega_d}{(2\pi)^d}}_{\frac{1}{2^{d-1}\pi^{d/2}} \frac{1}{\Gamma(d/2)}} \int_0^\infty \mathrm{d}k_{\mathrm{E}}\, k_{\mathrm{E}}^{d-1} f\left(k^2\right) \ . \tag{158}$$

The remaining integral can then be evaluated, again often using Γ functions. The result is finite for $d \neq 4$, but as we let $d \to 4$, the original divergence appears again in the form of $\Gamma(2 - d/2)$. The Γ function has poles at negative integers and at zero, so the integral exists for non-integer dimension. In the limit $d \to 4$, or equivalently, $\epsilon \to 0$, one has

$$\Gamma\left(2 - \frac{d}{2}\right) = \Gamma\left(\frac{\epsilon}{2}\right) = \frac{2}{\epsilon} - \gamma_{\mathrm{E}} + \mathcal{O}(\epsilon) \ , \tag{159}$$

with the Euler constant $\gamma_{\mathrm{E}} \simeq 0.58$.

As an example, consider the logarithmically divergent integral [cf. (150)]

$$\int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{1}{\left(k^2 + C\right)^2} \ , \tag{160}$$

where $C = (x_1 + x_2)^2\, m^2$. In d Euclidean dimensions, this becomes

$$\mu^\epsilon \int \frac{\mathrm{d}^4 k_{\mathrm{E}}}{(2\pi)^4} \frac{1}{\left(k_{\mathrm{E}}^2 + C\right)^2} = \frac{\mu^\epsilon \Gamma\left(2 - \frac{d}{2}\right)}{(4\pi)^{d/2}\, \Gamma(2)} \frac{1}{C^{2-d/2}} = \frac{1}{8\pi^2} \frac{1}{\epsilon} + \cdots \tag{161}$$

For the original expression (150) we thus obtain

$$\Gamma^\mu(p, 0) = \frac{\alpha}{2\pi} \frac{1}{\epsilon} \gamma^\mu + \mathcal{O}(1) \ . \tag{162}$$

What have we achieved? In four dimensions, the result is still divergent. However, the situation is better than before: We have separated the divergent part from the finite one and can take care of the divergence before taking the limit $\epsilon \to 0$. This is done in the procedure of renormalization.

There are more divergent one-loop graphs where we can achieve the same: the electron self-energy Σ in Fig. 9(a) (linearly divergent) and the photon self-energy or vacuum polarization $\Pi_{\mu\nu}$ in Fig. 9(b) (quadratically divergent). The self-energy graph has two divergent terms,

$$\Sigma(p) = \frac{3\alpha}{2\pi} \frac{1}{\epsilon} m - \frac{\alpha}{2\pi} \frac{1}{\epsilon} \left(\not{p} - m\right) + \mathcal{O}(1), \tag{163}$$

which contribute to the mass renormalization and the wave function renormalization, respectively. The vacuum polarization seems more complicated since it is a second rank tensor. However, the tensor structure is fixed by gauge invariance which requires

$$q^\mu \Pi_{\mu\nu}(q) = 0 \, . \tag{164}$$

Therefore, because of Lorentz invariance,

$$\Pi_{\mu\nu}(q) = \left(g_{\mu\nu}q^2 - q_\mu q_\nu\right)\Pi\!\left(q^2\right) \, . \tag{165}$$

The remaining scalar quantity $\Pi(q^2)$ has the divergent part

$$\Pi\!\left(q^2\right) = \frac{2\alpha}{3\pi}\frac{1}{\epsilon} + \mathcal{O}(1) \, . \tag{166}$$

4.2.2 Renormalization

So far we have isolated the divergences, but they are still there. How do we get rid of them? The crucial insight is that the parameters of the Lagrangean, the 'bare' parameters, are not observable. Rather, the sum of bare parameters and loop-induced corrections are physical. Hence, divergencies of bare parameters can cancel against divergent loop corrections, leaving physical observables finite.

To make this more explicit, let us express, as an example, the QED Lagrangian in terms of bare fields A_0^μ and ψ_0 and bare parameters m_0 and e_0,

$$\mathscr{L} = -\frac{1}{4}\left(\partial_\mu A_{0\,\nu} - \partial_\nu A_{0\,\mu}\right)\left(\partial^\mu A^{0\,\nu} - \partial^\nu A^{0\,\mu}\right) + \overline{\psi}_0\left(\gamma^\mu\left(\mathrm{i}\partial_\mu - e_0 A_{0\,\mu}\right) - m_0\right)\psi_0 \, . \tag{167}$$

The 'renormalized fields' A_μ and ψ and the 'renormalized parameters' e and m are then obtained from the bare ones by multiplicative rescaling,

$$A_{0\,\mu} = \sqrt{Z_3}A_\mu \, , \qquad\qquad \psi_0 = \sqrt{Z_2}\psi \, , \tag{168}$$

$$m_0 = \frac{Z_m}{Z_2}m \, , \qquad\qquad e_0 = \frac{Z_1}{Z_2\sqrt{Z_3}}\mu^{2-d/2}e \, . \tag{169}$$

Note that coupling and electron mass now depend on the mass parameter μ,

$$e = e(\mu) \, , \quad m = m(\mu) \, . \tag{170}$$
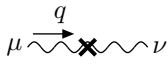
In terms of the renormalized fields and parameters the Lagrangian (167) reads

$$\mathscr{L} = -\frac{1}{4}\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) + \overline{\psi}\left(\gamma^\mu\left(\mathrm{i}\partial_\mu - eA_\mu\right) - m\right)\psi + \Delta\mathscr{L} \, , \tag{171}$$

where $\Delta\mathscr{L}$ contains the divergent counterterms,

$$\begin{aligned}
\Delta\mathscr{L} = &- (Z_3 - 1)\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (Z_2 - 1)\overline{\psi}\mathrm{i}\slashed{\partial}\psi \\
&- (Z_m - 1)m\overline{\psi}\psi - (Z_1 - 1)e\overline{\psi}\slashed{A}\psi \, .
\end{aligned} \tag{172}$$

The counter-terms have the same structure as the original Lagrangian and lead to new vertices in the Feynman rules:

i. $\quad -\mathrm{i}\,(Z_3 - 1)$ $\times\left(g_{\mu\nu}q^2 - q_\mu q_\nu\right)$ Photon wave function counter-term (counter-terms are generically denoted by ✗). It has the same tensor structure as the vacuum polarization.

ii. $-\mathrm{i}\,(Z_2 - 1)\,\not{p}$ Electron wave function counter-term.

iii. $-\mathrm{i}\,(Z_m - 1)\,m$ Electron mass counter-term.

iv. $-\mathrm{i}e\,(Z_1 - 1)\,\gamma^\mu$ Vertex counter-term.

The renormalization constants Z_i are determined by requiring that the counter-terms cancel the divergences. They can be determined as power series in α. The lowest order counter-terms are $\mathcal{O}\,(\alpha)$ and have to be added to the one-loop diagrams. Calculating, for example, the $\mathcal{O}\,(\alpha)$ correction to the electron–photon vertex, one has

$$\text{} + \text{} = -\mathrm{i}e\gamma^\mu \left(\frac{\alpha}{2\pi} \frac{1}{\epsilon} + (Z_1 - 1) + \mathcal{O}\,(1) \right) . \tag{173}$$

Demanding that the whole expression be finite determines the divergent part of Z_1,

$$Z_1 = 1 - \frac{\alpha}{2\pi} \frac{1}{\epsilon} + \mathcal{O}\,(1) . \tag{174}$$

Similarly, the $\mathcal{O}\,(\alpha)$ vacuum polarization now has two contributions,

$$\text{} + \text{} = -\mathrm{i}\,\left(g_{\mu\nu}q^2 - q_\mu q_\nu\right) \left(\frac{2\alpha}{3\pi} \frac{1}{\epsilon} + (Z_3 - 1) + \mathcal{O}\,(1) \right) , \tag{175}$$

which yields

$$Z_3 = 1 - \frac{2\alpha}{3\pi} \frac{1}{\epsilon} + \mathcal{O}\,(1) . \tag{176}$$

The other constants Z_2 and Z_m are fixed analogously. A Ward identity, which follows from gauge invariance, yields the important relation $Z_1 = Z_2$. The finite parts of the renormalization constants are still undetermined. There are different ways to fix them, corresponding to different renormalization schemes. All schemes give the same results for physical quantities, but differ at intermediate steps.

Having absorbed the divergences into the renormalized parameters and fields, we can safely take the limit $\epsilon \to 0$. The theory now yields well-defined relations between physical observables. Divergences can be removed to all orders in the loop expansion for renormalizable theories [3, 9]. Quantum electrodynamics and the Standard Model belong to this class. The proof is highly non-trivial and has been a major achievement in quantum field theory!

4.2.3   *Running coupling in QED*

Contrary to the bare coupling e_0, the renormalized coupling $e(\mu)$ depends on the renormalization scale μ [cf. (169)],

$$e_0 = \frac{Z_1}{Z_2\sqrt{Z_3}}\mu^{-2+d/2}e(\mu) = e(\mu)\mu^{-\epsilon/2}Z_3^{-\frac{1}{2}} ,$$

where we have used the Ward identity $Z_1 = Z_2$. It is very remarkable that the scale dependence is determined by the divergences. To see this, expand Eq. (169) in ϵ and $e(\mu)$,

$$\begin{aligned} e_0 &= e(\mu) \left(1 - \frac{\epsilon}{2}\ln\mu + \cdots \right) \left(1 + \frac{1}{\epsilon}\frac{\alpha}{3\pi} + \cdots \right) \\ &= e(\mu) \left(\frac{1}{\epsilon}\frac{e^2(\mu)}{12\pi^2} + 1 - \frac{e^2(\mu)}{24\pi^2}\ln\mu + \mathcal{O}\left(\epsilon, e^4(\mu)\right) \right) , \end{aligned} \tag{177}$$

31

where we have used $\alpha = e^2/(4\pi)$. Since the bare mass e_0 does not depend on μ, differentiation with respect to μ yields

$$0 = \mu\frac{\partial}{\partial\mu}e_0 = \mu\frac{\partial}{\partial\mu}e - \frac{e^3}{24\pi^2} + \mathcal{O}\left(e^5\right) \ , \tag{178}$$

and therefore

$$\mu\frac{\partial}{\partial\mu}e = \frac{e^3}{24\pi^2} + \mathcal{O}\left(e^5\right) \equiv \beta(e) \ . \tag{179}$$

This equation is known as the renormalization group equation, and the function on the right-hand side of Eq. (179) is the so-called β function,

$$\beta(e) = \frac{b_0}{(4\pi)^2}e^3 + \mathcal{O}\left(e^5\right) \ , \quad \text{with} \quad b_0 = \frac{2}{3} \ . \tag{180}$$

The differential equation (179) can easily be integrated. Using a given value of e at a scale μ_1, the coupling α at another scale μ is given by

$$\alpha(\mu) = \frac{\alpha\left(\mu_1\right)}{1 - \alpha\left(\mu_1\right)\frac{b_0}{(2\pi)}\ln\frac{\mu}{\mu_1}} \ . \tag{181}$$

Since $b_0 > 0$, the coupling increases with μ until it approaches the so-called Landau pole where the denominator vanishes and perturbation theory breaks down.

What is the meaning of a scale-dependent coupling? This becomes clear when one calculates physical quantities, such as a scattering amplitude at some momentum transfer q^2. In the perturbative expansion one then finds terms $\propto e^2(\mu)\log\left(q^2/\mu^2\right)$. Such terms make the expansion unreliable unless one chooses $\mu^2 \sim q^2$. Hence, $e^2\left(q^2\right)$ represents the effective interaction strength at a momentum (or energy) scale q^2 or, alternatively, at a distance of $r \sim 1/q$.

The positive β function in QED implies that the effective coupling strength decreases at large distances. Qualitatively, this can be understood as the effect of 'vacuum polarization': Electron–positron pairs etc. from the vacuum screen any bare charge at distances larger than the corresponding Compton wavelength. Quantitatively, one finds that the value $\alpha(0) = \frac{1}{137}$, measured in Thompson scattering, increases to $\alpha\left(M_Z^2\right) = \frac{1}{127}$, the value conveniently used in electroweak precision tests.

4.2.4 Running coupling in QCD

Everything we did so far for QED can be extended to non-Abelian gauge theories, in particular to QCD [5]. It is, however, much more complicated, since there are more diagrams to calculate, and we shall not be able to discuss this in detail. The additional diagrams contain gluon self-interactions and ghosts, and they lead to similar divergences, which again are absorbed by renormalization constants. Schematically, these are

$$\tag{182}$$

$$\tag{183}$$

$$\tag{184}$$

The renormalized coupling can again be defined as in QED, Eq. (169),

$$g_0 = \frac{Z_1}{Z_2 \sqrt{Z_3}} \mu^{-2+d/2} g \,. \tag{185}$$

The coefficients of the $1/\epsilon$-divergences depend on the gauge group and on the number of different fermions. For a $\mathsf{SU}(N_c)$ gauge group with N_f flavours of fermions, one obtains the β function for the gauge coupling g,

$$\mu \frac{\partial}{\partial \mu} g = \frac{b_0}{(4\pi)^2} g^3 + \mathcal{O}\left(g^5\right) \,, \quad b_0 = -\left(\frac{11}{3} N_c - \frac{4}{3} N_f\right) \,. \tag{186}$$

Note that for $N_f < 11 N_c/4$ the coefficient is negative! Hence, the coupling decreases at high-momentum transfers or short distances. The calculation of this coefficient earned the Nobel Prize in 2004 for Gross, Politzer and Wilczek. The decrease of the coupling at short distances is the famous phenomenon of asymptotic freedom. As a consequence, one can treat in deep-inelastic scattering quarks inside the proton as quasi-free particles, which is the basis of the parton model.

The coupling at a scale μ can again be expressed in terms of the coupling at a reference scale μ_1,

$$\alpha(\mu) = \frac{\alpha\left(\mu_1\right)}{1 + \alpha\left(\mu_1\right) \frac{|b_0|}{(2\pi)} \ln\frac{\mu}{\mu_1}} \,. \tag{187}$$

The analogue of the Landau pole now occurs at small μ or large distances. For QCD with $N_c = 3$ and $N_f = 6$, the pole is at the 'QCD scale' $\Lambda_{\mathrm{QCD}} \simeq 300$ MeV. At the QCD scale gluons and quarks are strongly coupled and colour is confined [5]. Correspondingly, the inverse of Λ_{QCD} gives roughly the size of hadrons, $r_{\mathrm{had}} \sim \Lambda_{\mathrm{QCD}}^{-1} \sim 0.7$ fm.

5 Electroweak theory

So far we have studied QED, the simplest gauge theory, and QCD, the prime example of a non-Abelian gauge theory. But there also are the weak interactions, which seem rather different. They are short-ranged, which requires massive messenger particles, seemingly inconsistent with gauge invariance. Furthermore, weak interactions come in two types, charged and neutral current–current interactions, which couple quarks and leptons differently. Charged-current interactions, mediated by the $W^\pm$ bosons, only involve left-handed fermions and readily change flavour, as in the strange quark decay $s \to u e^- \overline{\nu}_e$. Neutral-current interactions, on the other hand, couple both left- and right-handed fermions, and flavour-changing neutral currents are strongly suppressed.

Despite these differences from QED and QCD, weak interactions also turn out to be described by a non-Abelian gauge theory. Yet the electroweak theory is different for two reasons: it is a chiral gauge theory, and the gauge symmetry is spontaneously broken.

5.1 Quantum numbers

In a chiral gauge theory, the building blocks are massless left- and right-handed fermions,

$$\psi_L = \frac{1}{2}\left(1 - \gamma^5\right) \psi_L \,, \qquad \psi_R = \frac{1}{2}\left(1 + \gamma^5\right) \psi_R \,, \tag{188}$$

with different gauge quantum numbers. For one generation of Standard Model particles, we have seven chiral spinors: Two each for up- and down-type quark and charged lepton, and just one for the neutrino which we shall treat as massless in this section, i.e., we omit the right-handed one. The electroweak gauge group is a product of two groups, $\mathsf{G}_{EW} = \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$. Here the subscript W stands for 'weak isospin', which is the quantum number associated with the $\mathsf{SU}(2)_W$ factor, and the $\mathsf{U}(1)$ charge is the hypercharge Y.

The assignment of quantum numbers, which corresponds to the grouping into representations of the gauge group, is obtained as follows: The non-Abelian group $SU(2)_W$ has a chargeless one-dimensional singlet (**1**) representation and charged multidimensional representations, starting with the two-dimensional doublet (**2**) representation[5]. We are not allowed to mix quarks and leptons, since weak interactions do not change colour, nor left- and right-handed fields, which would violate Lorentz symmetry. The $U(1)_Y$ factor is Abelian, so it has only one-dimensional representations. This means we can assign different hypercharges to the various singlets and doublets of $SU(2)_W$.

Furthermore, we know that charged currents connect up- with down-type quarks and charged leptons with neutrinos, and that the $W^\pm$ bosons couple only to left-handed fermions. This suggests to form doublets from u_L and d_L, and from e_L and ν_L, and to keep the right-handed fields as singlets. So we obtain the $SU(2)_W$ multiplets

$$q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \qquad u_R, \qquad d_R, \qquad l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \qquad e_R, \qquad (189)$$

with the hypercharges (which we shall justify later)

$$
\begin{array}{cccccc}
\text{field:} & q_L & u_R & d_R & l_L & e_R \\
\text{hypercharge:} & \frac{1}{6} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{2} & -1
\end{array}
\qquad . \qquad (190)
$$

With these representations, we can write down the covariant derivatives. The $SU(2)_W$ has three generators, which we choose to be the Pauli matrices, and therefore three gauge fields W_μ^I, $I = 1, 2, 3$. The $U(1)_Y$ gauge field is B_μ, and the coupling constants are g and g', respectively. The covariant derivatives acting on the left-handed fields are

$$D_\mu \psi_L = \left(\partial_\mu + ig W_\mu + ig' Y B_\mu \right) \psi_L, \quad \text{where } W_\mu = \tfrac{1}{2} \sigma^I W_\mu^I, \qquad (191)$$

while the right-handed fields are singlets under $SU(2)_W$, and hence do not couple to the W bosons,

$$D_\mu \psi_R = \left(\partial_\mu + ig' Y B_\mu \right) \psi_R. \qquad (192)$$

From the explicit form of the Pauli matrices,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad (193)$$

we see that W_μ^1 and W_μ^2 mix up- and down-type quarks, while W_μ^3 does not, like the $U(1)$ boson B_μ.

It is often convenient to split the Lagrangian into the free (kinetic) part and the interaction Lagrangian, which takes the form (current) · (vector field). In the electroweak theory, one has

$$\mathscr{L}_{\text{int}} = -g J_{W,\mu}^I W^{I\,\mu} - g' J_{Y,\mu} B^\mu, \qquad (194)$$

with the currents

$$J_{W,\mu}^I = \overline{q}_L \gamma_\mu \tfrac{1}{2} \sigma^I q_L + \overline{l}_L \gamma_\mu \tfrac{1}{2} \sigma^I l_L, \qquad (195)$$

$$J_{Y,\mu} = \frac{1}{6} \overline{q}_L \gamma_\mu q_L - \frac{1}{2} \overline{l}_L \gamma_\mu l_L + \frac{2}{3} \overline{u}_R \gamma_\mu u_R - \frac{1}{3} \overline{d}_R \gamma_\mu d_R - \overline{e}_R \gamma_\mu e_R. \qquad (196)$$

These currents have to be conserved, $\partial_\mu J^\mu = 0$, to allow a consistent coupling to gauge bosons.

---

[5]Here we use 'representation' as meaning 'irreducible representation'. Of course we can build reducible representations of any dimension.

5.1.1 Anomalies

Before considering the Higgs mechanism which will lead to the identification of the physical $W^\pm$, Z and γ bosons of the Standard Model, let us briefly discuss anomalies. We shall see that the choice of hypercharges in (190) is severely constrained by the consistency of the theory.

Suppose we have a classical field theory with a certain symmetry and associated conserved current. After quantizing the theory, the resulting quantum field theory might not have that symmetry any more, which means the current is no longer conserved. This is called an anomaly. Anomalies are not a problem for global symmetries, where the quantized theory just lacks that particular symmetry. For gauge symmetries, however, the currents have to be conserved, otherwise the theory is inconsistent.
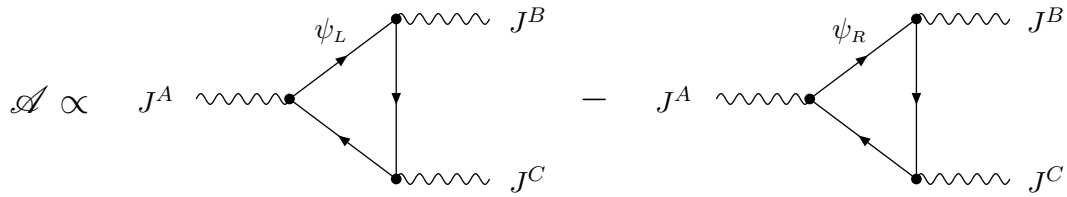
Fig. 10: The gauge anomaly is given by triangle diagrams with chiral fermions in the loop

Anomalies are caused by certain one-loop diagrams, the so-called triangle diagrams (see Fig. 10). The left- and right-handed fermions contribute with different sign, so if they have the same quantum numbers, the anomaly vanishes. This is the case in QED and QCD, which thus are automatically anomaly free. In general, for currents J^A, J^B and J^C, the anomaly $\mathscr{A}$ is the difference of the traces of the generators T^A, T^B and T^C in the left- and right-handed sectors,

$$\mathscr{A} = \mathrm{tr}\left[\{T^A, T^B\}\,T^C\right]_L - \mathrm{tr}\left[\{T^A, T^B\}\,T^C\right]_R \overset{!}{=} 0 \,. \tag{197}$$

Here the trace is taken over all fermions. For the electroweak theory, in principle there are four combinations of currents, containing three, two, one or no $\mathsf{SU}(2)_W$ current. However, the trace of any odd number of σ^I matrices vanishes, so we only have to check the $\mathsf{SU}(2)_W^2 \mathsf{U}(1)_Y$ and $\mathsf{U}(1)_Y^3$ anomalies.

The $\mathsf{SU}(2)_W$ generators are $\frac{1}{2}\sigma^I$, whose anticommutator is $\{\frac{1}{2}\sigma^I, \frac{1}{2}\sigma^J\} = \frac{1}{2}\delta^{IJ}$. Furthermore, only the left-handed fields contribute, since the right-handed ones are $\mathsf{SU}(2)_W$ singlets. Hence the $\mathsf{SU}(2)_W^2 \mathsf{U}(1)_Y$ anomaly is

$$\mathscr{A} = \mathrm{tr}\left[\left\{\tfrac{1}{2}\sigma^I, \tfrac{1}{2}\sigma^J\right\} Y\right]_L = \tfrac{1}{2}\delta^{IJ}\,\mathrm{tr}\,[Y]_L = \tfrac{1}{2}\delta^{IJ}\left(\underbrace{3}_{N_c}\cdot\tfrac{1}{6} - \tfrac{1}{2}\right) = 0 \,. \tag{198}$$

We see that it vanishes only if quarks come in three colours!

The $\mathsf{U}(1)_Y^3$ anomaly also vanishes:

$$\mathscr{A} = \mathrm{tr}\left[\{Y, Y\}\,Y\right]_L - \mathrm{tr}\left[\{Y, Y\}\,Y\right]_R = 2\left(\mathrm{tr}\left[Y^3\right]_L - \mathrm{tr}\left[Y^3\right]_R\right)$$

$$= 2\left(3\cdot 2\left(\tfrac{1}{6}\right)^3 + 2\left(-\tfrac{1}{2}\right)^3 - 3\left(\tfrac{2}{3}\right)^3 - 3\left(-\tfrac{1}{3}\right)^3 - (-1)^3\right) \tag{199}$$

$$= 0 \,.$$

This vanishing of the anomaly is again related to the number of colours. It does not vanish in either the left- or right-handed sector, nor in the quark and lepton sector individually. Hence the vanishing of

anomalies provides a deep connection between quarks and leptons in the Standard Model, which is a hint to grand unified theories where anomaly cancellation is often automatic.

Anomaly cancellation is not restricted to the electroweak gauge currents, but applies to the strong force and gravity as well: Mixed $SU(3)_C$-$U(1)_Y$ anomalies vanish by the same argument as above: Only the $SU(3)_C^2 U(1)_Y$ triangle contributes, but it is $\text{tr}\,[Y]_L - \text{tr}\,[Y]_R = 0$. The same is true for the last possible anomaly, the gravitational one, where two non-Abelian gauge currents are replaced by the energy-momentum tensor $T_{\mu\nu}$.

Hence, the Standard Model is anomaly free, as it should be. For this, all particles of one generation with their strange hypercharges have to conspire to cancel the different anomalies. A 'Standard Model' without quarks, for instance, would not be a consistent theory, nor would a 'Standard Model' with four colours of quarks. Note that a right-handed neutrino, suggested by neutrino masses, does not pose any problem, since it is a complete singlet, without any charge, and thus it does not contribute to any gauge anomaly.

5.2 Higgs mechanism

The electroweak model discussed so far bears little resemblance to the physics of weak interactions. The gauge bosons W_μ^I and B_μ are massless, implying long-range forces, because a mass term $m^2 W_\mu W^\mu$ would violate gauge invariance. Furthermore, the fermions are massless as well, again because of gauge invariance: a mass term mixes left- and right-handed fermions,

$$m\overline{\psi}\psi = m\left(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L\right) , \tag{200}$$

and since these have different gauge quantum numbers, such a term is not gauge invariant. The way out is the celebrated Higgs mechanism: Spontaneous symmetry breaking generates masses for the gauge bosons and fermions without destroying gauge invariance. A simpler version of this effect is what happens in superconductors: the condensate of Cooper pairs induces an effective mass for the photon, so that electromagnetic interactions become short-ranged, leading to the Meissner–Ochsenfeld effect where external magnetic fields are expelled from the superconductor, levitating it.

The key ingredient for the Higgs mechanism is a complex scalar field Φ, which is a doublet under $SU(2)_W$ with hypercharge $-\frac{1}{2}$, which has four real degrees of freedom. The crucial feature of the Higgs field is its potential, which is of the Mexican hat form:

$$\mathscr{L} = (D_\mu\Phi)^\dagger (D^\mu\Phi) - V\left(\Phi^\dagger\Phi\right) , \tag{201}$$

with

$$D_\mu\Phi = \left(\partial_\mu + igW_\mu - \frac{i}{2}g'B_\mu\right)\Phi ,$$

$$V\left(\Phi^\dagger\Phi\right) = -\mu^2\,\Phi^\dagger\Phi + \frac{1}{2}\lambda\left(\Phi^\dagger\Phi\right)^2 , \quad \mu^2 > 0 . \tag{202}$$

This potential has a minimum away from the origin, at $\Phi^\dagger\Phi = v^2 \equiv \mu^2/\lambda$. In the vacuum, the Higgs field settles in this minimum. At first sight, the minimization of the potential only fixes the modulus $\Phi^\dagger\Phi$, i.e., one of the four degrees of freedom. The other three, however, can be eliminated by a gauge transformation, and we can choose the following form of Φ, which is often referred to as unitary gauge:

$$\Phi = \begin{pmatrix} 0 \\ v + \frac{1}{\sqrt{2}}H(x) \end{pmatrix} , \quad H = H^* . \tag{203}$$

Here we have eliminated the upper component and the imaginary part of the lower one. We have also shifted the lower component to the vacuum value, so that the dynamical field $H(x)$ vanishes in the vacuum.

In unitary gauge, the Higgs Lagrangian (201) becomes

$$
\begin{aligned}
\mathscr{L} = &\ \frac{\lambda}{2}v^4 \\
&+ \frac{1}{2}\partial_\mu H\, \partial^\mu H - \lambda v^2 H^2 + \frac{\lambda}{\sqrt{2}}vH^3 + \frac{\lambda}{8}H^4 \\
&+ \frac{1}{4}\left(v + \frac{1}{\sqrt{2}}H\right)^2 \left(W_\mu^1, W_\mu^2, W_\mu^3, B_\mu\right)
\begin{pmatrix}
g^2 & 0 & & 0 \\
0 & g^2 & & \\
& & g^2 & gg' \\
0 & & gg' & g'^2
\end{pmatrix}
\begin{pmatrix}
W^{1\mu} \\
W^{2\mu} \\
W^{3\mu} \\
B^\mu
\end{pmatrix} .
\end{aligned}
\tag{204}
$$

The first line could be interpreted as vacuum energy density, i.e., a cosmological constant. However, such an interpretation is on shaky grounds in quantum field theory, so we shall ignore this term[6]. The second line describes a real scalar field H of mass $m_H^2 = 2\lambda v^2$ with cubic and quartic self-interactions. The most important line, however, is the last one: It contains mass terms for the vector bosons! A closer look at the mass matrix reveals that it only is of rank three, so it has one zero eigenvalue, and the three remaining ones are g^2, g^2, and $(g^2 + g'^2)$. In other words, it describes one massless particle, two of equal non-zero mass, and one which is even heavier, i.e., we have identified the physical γ, $\mathrm{W}^\pm$ and Z bosons.

The massless eigenstate of the mass matrix, i.e., the photon, is the linear combination $A_\mu = -\sin\theta_{\mathrm{W}}W_\mu^3 + \cos\theta_{\mathrm{W}}B_\mu$, the orthogonal combination is the Z boson, $Z_\mu = \cos\theta_{\mathrm{W}}W_\mu^3 + \sin\theta_{\mathrm{W}}B_\mu$. Here we have introduced the Weinberg angle θ_{W}, which is defined by

$$
\sin\theta_{\mathrm{W}} = \frac{g'}{\sqrt{g^2 + g'^2}}\,, \quad \cos\theta_{\mathrm{W}} = \frac{g}{\sqrt{g^2 + g'^2}}\,.
\tag{205}
$$

To summarize, the theory contains the following mass eigenstates:

- Two charged vector bosons $\mathrm{W}^\pm$ with mass $M_{\mathrm{W}}^2 = \frac{1}{2}g^2 v^2$,
- two neutral vector bosons with masses $M_{\mathrm{Z}} = \frac{1}{2}\left(g^2 + g'^2\right)v^2 = M_W^2 \cos^{-2}\theta_{\mathrm{W}}$ and $M_\gamma = 0$,
- and one neutral Higgs boson with mass $m_H^2 = 2\lambda v^2$.

The Higgs mechanism and the diagonalization of the vector boson mass matrix allow us to rewrite the interaction Lagrangian (194), which was given in terms of the old fields W_μ^I and B_μ and their currents (195) and (196), in terms of the physical field. The associated currents are separated into a charged current (for $W_\mu^\pm$) and neutral currents (for A_μ and Z_μ):

$$
\mathscr{L}_{\mathrm{CC}} = -\frac{g}{\sqrt{2}}\sum_{i=1,2,3}\left(\overline{u}_{Li}\gamma^\mu d_{Li} + \overline{\nu}_{Li}\gamma^\mu e_{Li}\right)W_\mu^+ + \mathrm{h.c.}\,,
\tag{206}
$$

$$
\begin{aligned}
\mathscr{L}_{\mathrm{NC}} &= -gJ_\mu^3 W^{3\,\mu} - g'J_{Y\,\mu}B^\mu \\
&= -eJ_{\mathrm{em}\,\mu}A^m - \frac{e}{\sin 2\theta_{\mathrm{W}}}J_{Z\,\mu}Z^\mu\,,
\end{aligned}
\tag{207}
$$

with the electromagnetic and Z currents

$$
J_{\mathrm{em}\,\mu} = \sum_{\substack{i=u,d,c,\\ s,t,b,e,\mu,\tau}}\overline{\psi}_i\gamma_\mu Q_i\psi_i\,, \quad \text{with the electric charge}\quad Q_i = T_i^3 + Y_i\,,
\tag{208}
$$

---

[6]Generally, nothing prevents us from adding an arbitrary constant to the Lagrangian, obtaining any desired 'vacuum energy'. For example, the Higgs potential is often written as $\left(\Phi^\dagger\Phi - v^2\right)^2$, so that its expectation value vanishes in the vacuum. These potentials just differ by the shift $\sim v^4$, and are indistinguishable within QFT.

$$J_{Z\,\mu} = \sum_{\substack{i=u,d,c,s,t,b \\ e,\mu,\tau,\nu_e,\nu_\mu,\nu_\tau}} \overline{\psi}_i \gamma_\mu \left(v_i - a_i \gamma^5 \right) \psi_i \; . \tag{209}$$

Here the fermions ψ_i are the sum of left- and right-handed fields,

$$\psi_i = \psi_{Li} + \psi_{Ri} \; . \tag{210}$$

The coupling to the photon, the electric charge Q, is given by the sum of the third component of weak isospin T^3 ($\pm\frac{1}{2}$ for doublets, zero for singlets) and the hypercharge Y. This reproduces the known electric charges of quarks and leptons, which justifies the hypercharge assignments in (190). The coupling constant e is related to the original couplings and the weak mixing angle:

$$e = g \sin\theta_{\mathrm{W}} = g' \cos\theta_{\mathrm{W}} \; . \tag{211}$$

The photon couples only vector-like, i.e., it does not distinguish between different chiralities. The Z boson, on the other hand, couples to the vector and axial-vector currents of different fermions ψ_i (i.e., their left-and right-handed components) with different strengths. They are given by the respective couplings v_i and a_i, which are universal for all families. In particular, the Z couples in the same way to all leptons, a fact known as lepton universality.

The Higgs mechanism described above is also called spontaneous symmetry breaking. This term, however, is somewhat misleading: Gauge symmetries are never broken, but only hidden. The Lagrangian (204) has only a manifest $\mathsf{U}(1)$ symmetry associated with the massless vector field, so it seems we have lost three gauge symmetries. This, however, is just a consequence of choosing the unitary gauge. The Higgs mechanism can also be described in a manifestly gauge-invariant way, and all currents remain conserved.

The 'spontaneous breaking of gauge invariance' reshuffles the degrees of freedom of the theory: Before symmetry breaking, we have the complex Higgs doublet (four real degrees of freedom) and four massless vector fields with two degrees of freedom each, so twelve in total. After symmetry breaking (and going to unitary gauge), three Higgs degrees of freedom are gone (one remaining), but they have resurfaced as extra components of three massive vector fields[7] (nine), and one vector field stays massless (another two). So there are still twelve degrees of freedom.

5.3 Fermion masses and mixings

The Higgs mechanism generates masses not only for the gauge bosons, but also for the fermions. As already emphasized, direct mass terms are not allowed in the Standard Model. There are, however, allowed Yukawa couplings of the Higgs doublet to two fermions. They come in three classes, couplings to quark doublets and either up- or down-type quark singlets, and to lepton doublet and charged lepton singlets. Each term is parametrized by a 3×3 matrix in generation space,

$$\mathscr{L}_{\mathrm{Y}} = (h_u)_{ij}\, \overline{q}_{L\,i} u_{R\,j} \Phi + (h_d)_{ij}\, \overline{q}_{L\,i} d_{R\,j} \widetilde{\Phi} + (h_e)_{ij}\, \overline{l}_{L\,i} e_{R\,j} \widetilde{\Phi} + \text{h.c.} \; , \tag{212}$$

where $\widetilde{\Phi}$ is given by $\widetilde{\Phi}_a = \epsilon_{ab} \Phi_b^*$.

These Yukawa couplings effectively turn into mass terms once the electroweak symmetry is spontaneously broken: A vacuum expectation value $\langle \Phi_X \rangle = v$ inserted in the Lagrangian (212) yields

$$\mathscr{L}_m = (m_u)_{ij}\, \overline{u}_{L\,i} u_{R\,j} + (m_d)_{ij}\, \overline{d}_{L\,i} d_{R\,j} + (m_e)_{ij}\, \overline{e}_{L\,i} e_{R\,j} + \text{h.c.} \; . \tag{213}$$

Here the mass matrices are $m_u = h_u v$ etc., and u_L, d_L and e_L denote the respective components of the quark and lepton doublets q_L and l_L.

---

[7] Remember that a massless vector has only two (transverse) degrees of freedom, while a massive one has a third, longitudinal, mode.

The mass matrices thus obtained are in general not diagonal in the basis where the charged current is diagonal. They can be diagonalized by bi-unitary transformations,

$$V^{(u)\dagger} m_u \widetilde{V}^{(u)} = \text{diag}(m_u, m_c, m_t) \ , \tag{214a}$$

$$V^{(d)\dagger} m_d \widetilde{V}^{(d)} = \text{diag}(m_d, m_s, m_b) \ , \tag{214b}$$

$$V^{(e)\dagger} m_e \widetilde{V}^{(e)} = \text{diag}(m_e, m_\mu, m_\tau) \ , \tag{214c}$$

with unitary matrices V,

$$V^{(u)\dagger} V^{(u)} = \mathbb{1} , \quad \text{etc.}$$

This amounts to a change of basis from the weak eigenstates (indices $i, j, \ldots$) to mass eigenstates (with indices $\alpha, \beta, \ldots$):

$$u_{L\,i} = V^{(u)}_{i\alpha} u_{L\,\alpha}, \qquad d_{L\,i} = V^{(d)}_{i\alpha} d_{L,\alpha}, \qquad u_{R\,i} = \widetilde{V}^{(u)}_{i\alpha} u_{R\,\alpha}, \qquad d_{R\,i} = \widetilde{V}^{(d)}_{i\alpha} d_{R\,\alpha}. \tag{215}$$

The up- and down-type matrices $V^{(u)}$ and $V^{(d)}$ are not identical, which has an important consequence: The charged-current couplings are now no longer diagonal, but rather

$$\mathscr{L}_{\text{CC}} = -\frac{g}{\sqrt{2}} V_{\alpha\beta} \overline{u}_{L\,\alpha} \gamma^\mu d_{L\,\beta} W^+_\mu + \text{h.c.} \ , \tag{216}$$

with the CKM matrix

$$V_{\alpha\beta} = V^{(u)\dagger}_{\alpha i} V^{(d)}_{i\beta} \ , \tag{217}$$

which carries the information about flavour mixing in charged-current interactions. Because of the unitarity of the transformations, there is no flavour mixing in the neutral current.

We saw that the Higgs mechanism generates fermion masses since direct mass terms are not allowed due to gauge invariance. There is one possible exception: a right-handed neutrino, which one may add to the Standard Model to have also neutrino masses. It is a singlet of the Standard Model gauge group and can therefore have a Majorana mass term which involves the charge conjugate fermion

$$\psi^C = C \overline{\psi}^T \ , \tag{218}$$

where $C = \mathrm{i}\gamma^2\gamma^0$ is the charge conjugation matrix. As the name suggests, the charge conjugate spinor has charges opposite to the original one. It also has opposite chirality, $P_L \psi^C_R = \psi_R$. Thus we can produce a mass term $\overline{\psi}^C \psi$ (remember that a mass term always requires both chiralities), which is gauge invariant only for singlet fields.

So a right-handed neutrino ν_R can have the usual Higgs coupling and a Majorana mass term,

$$\mathscr{L}_{\nu,\text{mass}} = h_{\nu\,ij} \overline{l}_{L\,i} \nu_{R\,j} \Phi + \frac{1}{2} M_{ij} \overline{\nu}_{R\,i} \nu_{R\,j} + \text{h.c.} \ , \tag{219}$$

where i, j again are family indices.

The Higgs vacuum expectation value v turns the coupling matrix h_ν into the Dirac mass matrix $m_D = h_\nu v$. The eigenvalues of the Majorana mass matrix M can be much larger than the Dirac masses, and a diagonalization of the (ν_L, ν_R) system leads to three light modes ν_i with the mass matrix

$$m_\nu = -m_D M^{-1} m_D^T \ . \tag{220}$$

Large Majorana masses naturally appear in grand unified theories. For $M \sim 10^{15}$ GeV, and $m_D \sim m_t \sim 100$ GeV for the largest Dirac mass, one finds $m_\nu \sim 10^{-2}$ eV, which is consistent with results from neutrino oscillation experiments. This 'see-saw mechanism', which explains the smallness of neutrino masses masses as a consequence of large Majorana mass terms, successfully relates neutrino physics to grand unified theories.

5.4 Predictions

The electroweak theory contains four parameters, the two gauge couplings and the two parameters of the Higgs potential: g, g', μ^2 and λ. They can be traded for four other parameters, which are more easily measured: The fine-structure constant α, the Fermi constant G_F and the Z boson mass M_Z, which are known to great accuracy, and the Higgs mass m_H which is not yet known.
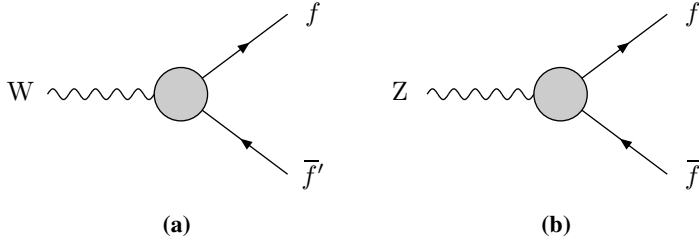
Fig. 11: Decays of the W and Z bosons into two fermions. In W decays, the fermion and antifermion can have different flavour. The grey blobs indicate higher order corrections which must be included to match the experimental precision.

At LEP, W and Z bosons were produced in huge numbers. There are many observables related to their production and decay (Fig. 11). These include

- The W mass M_W and the decay widths Γ_W and Γ_Z.
- Ratios of partial decay widths, for example, the ratio of the partial Z width into bottom quarks to that into all hadrons,

$$R_b = \frac{1}{\Gamma(Z \to \text{hadrons})}\Gamma\left(Z \to b\overline{b}\right) . \tag{221}$$

- Forward–backward asymmetries: In $e^+ + Pem \to Z/\gamma \to f\overline{f}$ reactions, the direction of the outgoing fermion is correlated with the incoming electron. This is quantified by the asymmetries A_{fb}^f,

$$A_{\text{fb}}^f = \frac{\sigma_{\text{f}}^f - \sigma_{\text{b}}^f}{\sigma_{\text{f}}^f + \sigma_{\text{b}}^f}, \quad \text{for } f = \mu, \tau, b, c , \tag{222}$$

where σ_{f}^f is the cross-section for an outgoing fermion in the forward direction, i.e., $\theta \in [0, \pi/2]$ in Fig. 12, while σ_{b}^f is the cross-section for backward scattering.
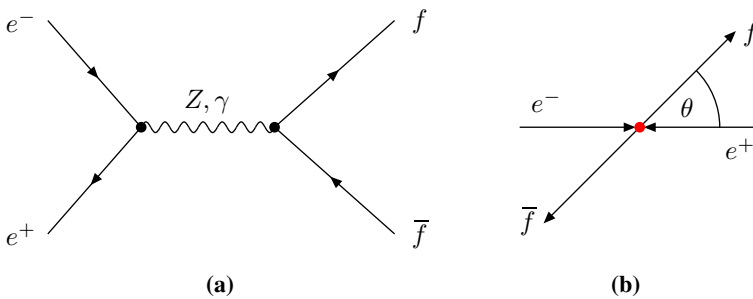
Fig. 12: The forward–backward asymmetry A_{fb}: In the process $e^+e^- \to Z/\gamma \to f\overline{f}$, there is a correlation between the directions of the outgoing fermion and the incoming electron. This asymmetry has been measured for several types of final-state fermions, mostly at LEP with centre-of-mass energy $\sqrt{s} = M_Z$.

Also important are double, left–right and forward–backward asymmetries,

$$A_{\text{LR}}^{\text{fb}} = \frac{\sigma_{\text{Lf}}^f - \sigma_{\text{Lb}}^f - \sigma_{\text{Rf}}^f + \sigma_{\text{Rb}}^f}{\sigma_{\text{Lf}}^f + \sigma_{\text{Lb}}^f + \sigma_{\text{Rf}}^f + \sigma_{\text{Rb}}^f} \equiv \frac{3}{4}A_f . \tag{223}$$

The reason for these asymmetries is the presence of the axial couplings a_i in the Z boson current (209), which lead to different cross-sections for the processes $Z \rightarrow f_L \bar{f}_R$ and $Z \rightarrow f_R \bar{f}_L$. Thus, one can deduce the a_i and v_i couplings for fermions from the forward–backward asymmetries, and finally the weak mixing angle, on which the vector and axial-vector couplings of the Z boson depend,

$$\sin^2 \theta_{\text{eff}}^{\text{lept}} = \frac{1}{4} \left(1 - \frac{v_l}{a_l} \right) . \tag{224}$$

– Electroweak measurements by now are very precise, and require the inclusion of W boson loops in theoretical calculations, so that they test the non-Abelian nature of the electroweak theory. The theoretical predictions critically depend on the the electromagnetic coupling at the electroweak scale, $\alpha(m_Z)$, which differs from the low energy value $\alpha(0)$ in particular by hadronic corrections, $\Delta \alpha_{\text{had}}(m_Z)$.

An important observable is the ρ parameter, defined by

$$\rho = \frac{M_W^2}{M_Z^2 \cos^2 \theta_W} . \tag{225}$$

At tree level, $\rho = 1$. Loop corrections to the masses of the gauge bosons, and therefore to ρ, due to quark or Higgs boson loops as in Fig. 13, are an important prediction of the electroweak theory.

(a) Heavy quark corrections

(b) Higgs corrections
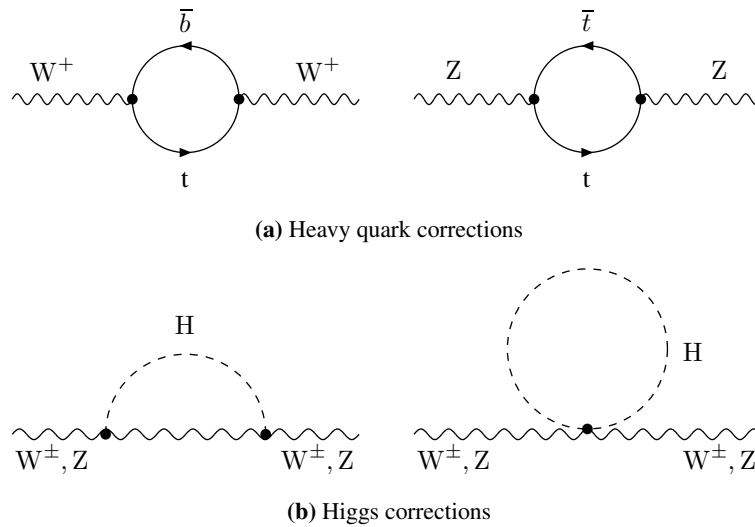
Fig. 13: Radiative corrections to the masses of the W and Z bosons, which depend on the masses of the particles in the loop. Diagrams with gauge boson self-interactions have been omitted.

The tree-level value $\rho = 1$ is protected by an approximate $\mathsf{SU}(2)$ symmetry, called custodial symmetry, which is only broken by the $\mathsf{U}(1)_\mathsf{Y}$ gauge interaction and by Yukawa couplings. Thus the corrections depend on the fermion masses, and are dominated by the top quark, as in Fig 13(a). The leading correction is

$$\Delta \rho^{(t)} = \frac{3 G_F m_t^2}{8 \pi^2 \sqrt{2}} \propto \frac{m_t^2}{M_W^2} . \tag{226}$$

This led to the correct prediction of the top mass from electroweak precision data before the top quark was discovered at the Tevatron.
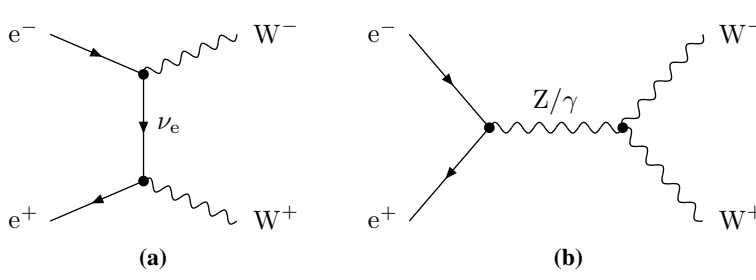
Fig. 14: The process $e^+e^- \to W^+W^-$. The diagrams of panel (b) contain triple gauge boson vertices, γWW and ZWW.

The correction due to the Higgs boson diagrams in Fig. 13(b) again depends on the Higgs mass, but this time the effect is only logarithmic:

$$\Delta\rho^{(H)} = -C \ln\frac{m_H^2}{M_W^2} \ . \tag{227}$$

From this relation, one can obtain a prediction for the mass of the Higgs boson. Clearly, the accuracy of this prediction strongly depends on the experimental error on the top mass, which affects ρ quadratically.

However, the Higgs mass (weakly) influences many other quantities, and from precision measurements one can obtain a fit for the Higgs mass. This is shown in the famous blue-band plot, Fig. 21.

- A characteristic prediction of any non-Abelian gauge theory is the self-interaction of the gauge bosons. In the electroweak theory, this can be seen in the process $e^+e^- \to W^+W^-$.

(a) ... cross-section at LEP2. Pre- ... W vertex deviate substan-

(b) For Z pair production, there is no triple Z vertex, which agrees well with the experimental result.
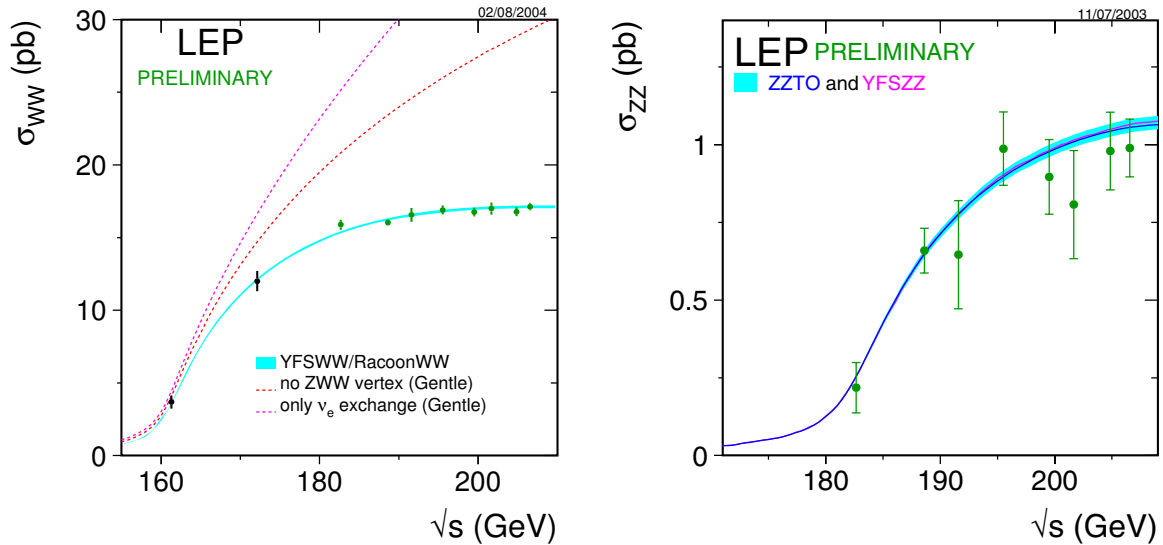
Fig. 15: Gauge boson pair production cross-sections at LEP2 energies (from Ref. [11])

The tree-level diagrams are given in Fig. 14, and Fig. 15(a) shows the measured cross-section from LEP, compared with theoretical predictions. Clearly, the full calculation including all diagrams agrees well with data, while the omission of the γWW and ZWW vertices leads to large discrepancies. For the process $e^+e^- \to ZZ$, on the other hand, there is no triple gauge boson (ZZZ or γZZ) vertex, so at tree level one only has the t-channel diagram which is similar to the diagram

in Fig. 14(a), but with an electron instead of the neutrino. The agreement between theory and data is evident from Fig. 15(b).

5.4.1 Fermi theory

The exchange of a W boson with momentum q in a Feynman diagram contributes a factor of $\left(M_{\mathrm{W}}^2 - q^2\right)^{-2}$ to the amplitude. For low-energy processes like muon decay (see Fig. 16), the momentum transfer is much smaller than the mass of the W boson. Hence to good approximation one can ignore q^2 and replace the propagator by M_{W}^{-2}. This amounts to introducing an effective four-fermion vertex (see Fig. 17),

Fig. 16: μ decay

$$\mathscr{L}_{\mathrm{CC}}^{\mathrm{eff}} = -\frac{G_{\mathrm{F}}}{\sqrt{2}} J_{\mathrm{CC}}^{\mu} J_{\mathrm{CC}\mu}^{\dagger} \,, \tag{228}$$

where G_{F} is Fermi's constant,

$$G_{\mathrm{F}} = \frac{g^2}{4\sqrt{2}\,M_{\mathrm{W}}^2} = \frac{1}{2\sqrt{2}\,v^2} \,, \tag{229}$$

which is inversely proportional to the Higgs vacuum expectation value v^2. A four-fermion theory for the weak interactions was first introduced by Fermi in 1934. Since it is not renormalizable, it cannot be considered a fundamental theory. However, one can use it as an effective theory at energies small compared to the W mass. This is sufficient for many applications in flavour physics, where the energy scale is set by the masses of leptons, kaons and B mesons.
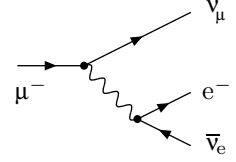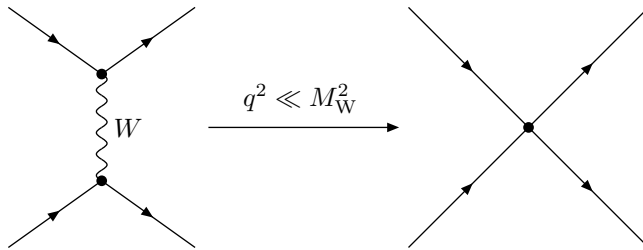
Fig. 17: W boson exchange can be described in terms of the Fermi theory, an effective theory for momentum transfers small compared to the W mass. The W propagator is replaced by a four-fermion vertex $\propto G_{\mathrm{F}}$.

5.5 Summary

The electroweak theory is a chiral gauge theory with gauge group $\mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$. This symmetry is spontaneously broken down to $\mathsf{U}(1)_{\mathrm{em}}$ by the Higgs mechanism which generates the gauge boson and Higgs masses, and also all fermion masses, since direct mass terms are forbidden by gauge invariance.

The electroweak theory is extremely well tested experimentally, to the level of 0.1%, which probes loop effects of the non-Abelian gauge theory. The results of a global electroweak fit are shown in Fig. 18. There is one deviation of almost 3σ, all other quantities agree within less than 2σ.

This impressive agreement is only possible because of two properties of the electroweak interactions: they can be tested in lepton–lepton collisions, which allow for very precise measurements, and they can be reliably calculated in perturbation theory. QCD, on the other hand, requires hadronic processes which are known experimentally with less accuracy and are also theoretically subject to larger uncertainties.

	Measurement	Fit	$|O^{meas}-O^{fit}|/\sigma^{meas}$
			0 1 2 3
$\Delta\alpha_{had}^{(5)}(m_Z)$	0.02758 ± 0.00035	0.02767	
m_Z [GeV]	91.1875 ± 0.0021	91.1874	
Γ_Z [GeV]	2.4952 ± 0.0023	2.4959	
σ_{had}^0 [nb]	41.540 ± 0.037	41.478	
R_l	20.767 ± 0.025	20.743	
$A_{fb}^{0,l}$	0.01714 ± 0.00095	0.01643	
$A_l(P_\tau)$	0.1465 ± 0.0032	0.1480	
R_b	0.21629 ± 0.00066	0.21581	
R_c	0.1721 ± 0.0030	0.1722	
$A_{fb}^{0,b}$	0.0992 ± 0.0016	0.1037	
$A_{fb}^{0,c}$	0.0707 ± 0.0035	0.0742	
A_b	0.923 ± 0.020	0.935	
A_c	0.670 ± 0.027	0.668	
$A_l(SLD)$	0.1513 ± 0.0021	0.1480	
$\sin^2\theta_{eff}^{lept}(Q_{fb})$	0.2324 ± 0.0012	0.2314	
m_W [GeV]	80.404 ± 0.030	80.376	
Γ_W [GeV]	2.115 ± 0.058	2.092	
m_t [GeV]	172.5 ± 2.3	172.9	
			0 1 2 3

Fig. 18: Results of a global fit to electroweak precision data. The right column shows the deviation of the fit from measured values in units of the standard deviation. From Ref. [11].

6   The Higgs profile

The only missing building block of the Standard Model is the Higgs boson. Spontaneously broken electroweak symmetry, however, is a cornerstone of the Standard Model, and so the discovery of the Higgs boson and the detailed study of its interactions is a topic of prime importance for the LHC and the ILC.

The investigation of the Higgs sector can be expected to to give important insight also on physics beyond the Standard Model. Since the Higgs is a scalar particle, its mass is subject to quadratically divergent quantum corrections, and an enormous 'fine-tuning' of the tree-level mass term is needed to keep the Higgs light (this is usually referred to as the 'naturalness problem' of the Higgs sector). Such considerations have motivated various extensions of the Standard Model:

- Supersymmetry retains an elementary scalar Higgs (and actually adds four more), while radiative corrections with opposite signs from bosons and fermions cancel.
- Technicolour theories model the Higgs as a composite particle of size $1/\Lambda_{\text{TC}}$, where $\Lambda_{\text{TC}} \sim 1$ TeV is the confinement scale of a new non-Abelian gauge interaction. These theories generically have problems with electroweak precision tests and the generation of fermion masses.
- A related idea regards the Higgs as a pseudo-Goldstone boson of some approximate global symmetry spontaneously broken at an energy scale above the electroweak scale. The Higgs mass is then related to the explicit breaking of this symmetry.
- In theories with large extra dimensions new degrees of freedom occur, and the Higgs field can be identified, for instance, as the fifth component of a five-dimensional vector field.

All such ideas can be tested at the LHC and the ILC, since the unitarity of WW scattering implies that the Standard Model Higgs and/or other effects related to electroweak symmetry breaking become manifest at energies below ~ 1 TeV.

6.1   Higgs couplings and decay

Suppose a resonance is found at the LHC with a mass above 114 GeV and zero charge. How can one establish that it indeed is the Higgs?

The Higgs boson can be distinguished from other scalar particles as they occur, for instance, in supersymmetric theories, by its special couplings to Standard Model particles. All couplings are proportional to the mass of the particle, since it is generated by the Higgs mechanism. Hence the Higgs decays dominantly into the heaviest particles kinematically allowed, which are $t\bar{t}$ or, for a light Higgs, $b\bar{b}$ and $t\bar{t}$ pairs. It also has a strong coupling $\propto m_{\text{H}}$ to the longitudinal component of W and Z bosons. The tree-level diagrams are given in Figs. 19(a) and 19(b). In addition, there are important loop-induced couplings to massless gluons and photons [see Fig. 19(c)].

The tree-level decay widths in the approximation $m_{\text{H}} \gg m_{\text{f}}, M_{\text{W}}$ are given by

$$\Gamma\big(\text{H} \to f\overline{f}\big) = \frac{G_{\text{F}} m_{\text{H}} m_{\text{f}}^2}{4\pi\sqrt{2}} N_{\text{c}} , \tag{230a}$$

$$\Gamma(\text{H} \to \text{Z}_{\text{L}}\text{Z}_{\text{L}}) = \frac{1}{2}\Gamma(\text{H} \to \text{W}_{\text{L}}\text{W}_{\text{L}}) = \frac{G_{\text{F}} m_{\text{H}}^3}{32\pi\sqrt{2}} . \tag{230b}$$

The branching fractions of the Higgs into different decay products strongly depend on the Higgs mass, as shown in Fig. 20. For a heavy Higgs, with $m_{\text{H}} > 2M_{\text{W}}$, the decay into a pair of W bosons dominates. At the threshold the width increases by two orders of magnitude, and it almost equals the Higgs mass at $m_{\text{H}} \sim 1$ TeV where the Higgs dynamics becomes non-perturbative. For a light Higgs with a mass just above the present experimental limit, $m_H > 114$ GeV, the decay into two photons might be the best possible detection channel given the large QCD background for the decay into two gluons at the LHC. It

45
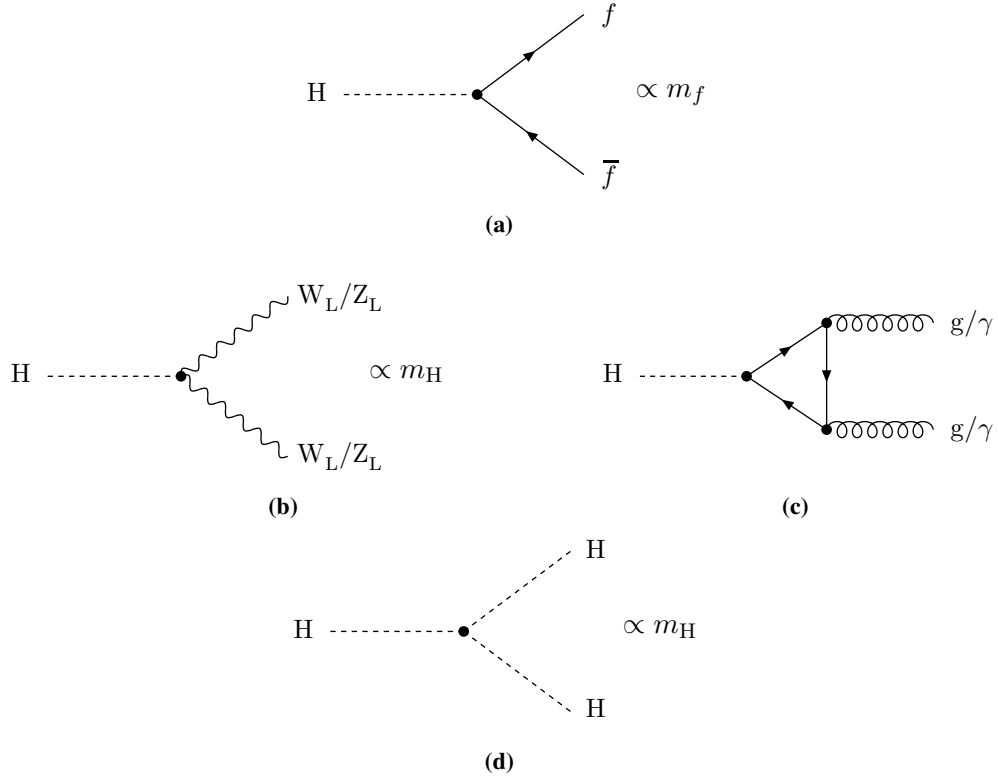
Fig. 19: Higgs boson decays. Tree-level couplings are proportional to masses, but there also are loop-induced decays into massless particles. The cubic Higgs self-coupling can be probed at the ILC and possibly at the LHC.

is clearly an experimental challenge to establish the mass dependence of the Higgs couplings, so the true discovery of the Higgs is likely to take several years of LHC data!
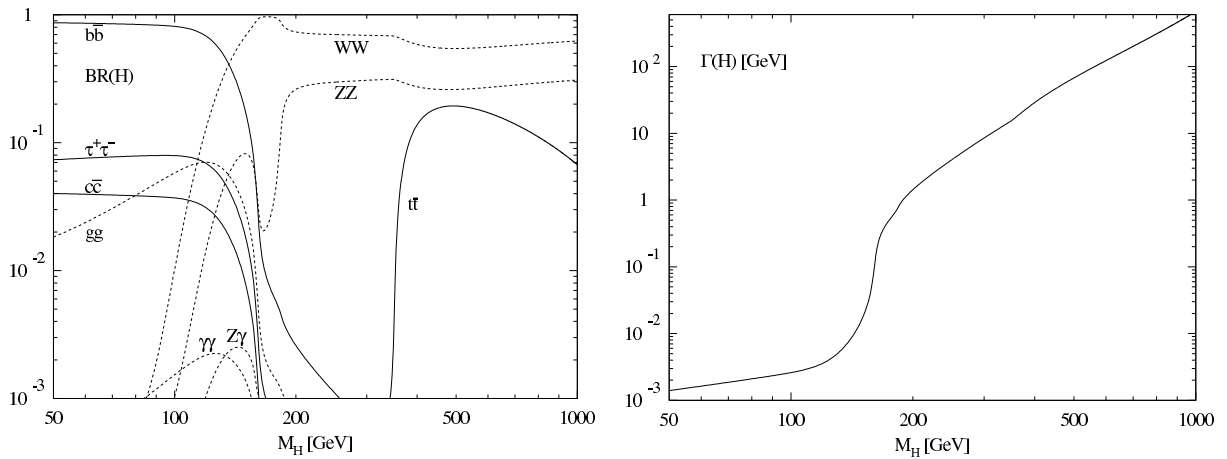
Fig. 20: Left: Higgs branching ratios as function of the Higgs mass. Right: Higgs decay width as function of the Higgs mass. It increases by two orders of magnitude at the WW threshold. From Ref. [12].

6.2 Higgs mass bounds

We now turn to the issue of the Higgs mass. Within the Standard Model, $m_H^2 = 2\lambda v^2$ is a free parameter which cannot be predicted. There are, however, theoretical consistency arguments which yield stringent upper and lower bounds on the Higgs mass.

Before we present these arguments, we first recall the experimental bounds:

– The Higgs has not been seen at LEP. This gives a lower bound on the mass, $m_H > 114$ GeV.
– The Higgs contributes to radiative corrections, in particular for the ρ parameter. Hence precision measurements yield indirect constraints on the Higgs mass. The result of a global fit is shown in the blue-band plot, Fig. 21. The current 95% confidence level upper bound is $m_H < 185$ GeV, an impressive result! One should keep in mind, however, that the loop corrections used to determine the Higgs mass strongly depend on the top mass as well. A shift of a few GeV in the top mass, well within the current uncertainties, can shift the Higgs mass best fit by several tens of GeV, as can be seen by comparing the plots in Fig. 21.

Theoretical bounds on the Higgs mass arise, even in the Standard Model, from two consistency requirements: (Non-)Triviality and vacuum stability. In the Minimal Supersymmetric Standard Model (MSSM), on the other hand, the Higgs self-coupling is given by the gauge couplings, which implies the upper bound $m_H \lesssim 135$ GeV.

The mass bounds in the Standard Model arise from the scale dependence of couplings, as explained in Section 4. Most relevant are the quartic Higgs self-coupling λ and the top quark Yukawa coupling h_t which gives the top mass via $m_t = h_t v$. Other Yukawa couplings are much smaller and can be ignored. The renormalization group equations for the couplings $\lambda(\mu)$ and $h_t(\mu)$ are

$$\mu\frac{\partial}{\partial\mu}\lambda(\mu) = \beta_\lambda\big(\lambda, h_t\big) = \frac{1}{(4\pi)^2}\left(12\lambda^2 - 12h_t^4 + \ldots\right) , \tag{231a}$$

$$\mu\frac{\partial}{\partial\mu}h_t(\mu) = \beta_\lambda\big(\lambda, h_t\big) = \frac{h_t}{(4\pi)^2}\left(\frac{9}{2}h_t^2 - 8g_s^2 + \ldots\right) . \tag{231b}$$

These equations imply that h_t decreases with increasing μ whereas the behaviour of $\lambda(\mu)$ depends on the initial condition $\lambda(v)$, i.e., on the Higgs mass.

For the Standard Model to be a consistent theory from the electroweak scale v up to some high-energy cutoff Λ, one needs to satisfy the following two conditions in the range $v < \mu < \Lambda$:

– The triviality bound: $\lambda(\mu) < \infty$. If λ would hit the Landau pole at some scale $\mu_L < \Lambda$, a finite value $\lambda(\mu_L)$ would require $\lambda(v) = 0$, i.e., the theory would be 'trivial'.
– The vacuum stability bound: $\lambda(\mu) > 0$. If λ would become negative, the Higgs potential would not be bounded from below any more, and the electroweak vacuum would no longer be the ground state of the theory.

These two requirements define allowed regions in the m_H–m_t plane as a function of the cutoff Λ [see Fig. (22a)]. For a given top mass, this translates into an upper and lower bound on the Higgs mass. For increasing Λ, the allowed region shrinks, and for the known top quark mass and $\Lambda \sim \Lambda_{GUT} \sim 10^{16}$ GeV, the Higgs mass is constrained to lie in a narrow region, 130 GeV $< m_H < 180$ GeV [see Fig. (22b)].

The impressively narrow band of allowed Higgs masses, which one obtains from the triviality and vacuum stability bounds, assumes that the Standard Model is valid up to Λ_{GUT}, the scale of grand unification. This might seem a bold extrapolation, given the fact that our present experimental knowledge ends at the electroweak scale, $\sim 10^2$ GeV. There are, however, two indications for such a 'desert' between the electroweak scale and the GUT scale: First, the gauge couplings empirically unify at the GUT scale, especially in the supersymmetric Standard Model, if there are no new particles between $\sim 10^2$ GeV and Λ_{GUT}; second, via the see-saw mechanism, the evidence for small neutrino masses is also consistent with an extrapolation to Λ_{GUT} without new physics at intermediate scales.
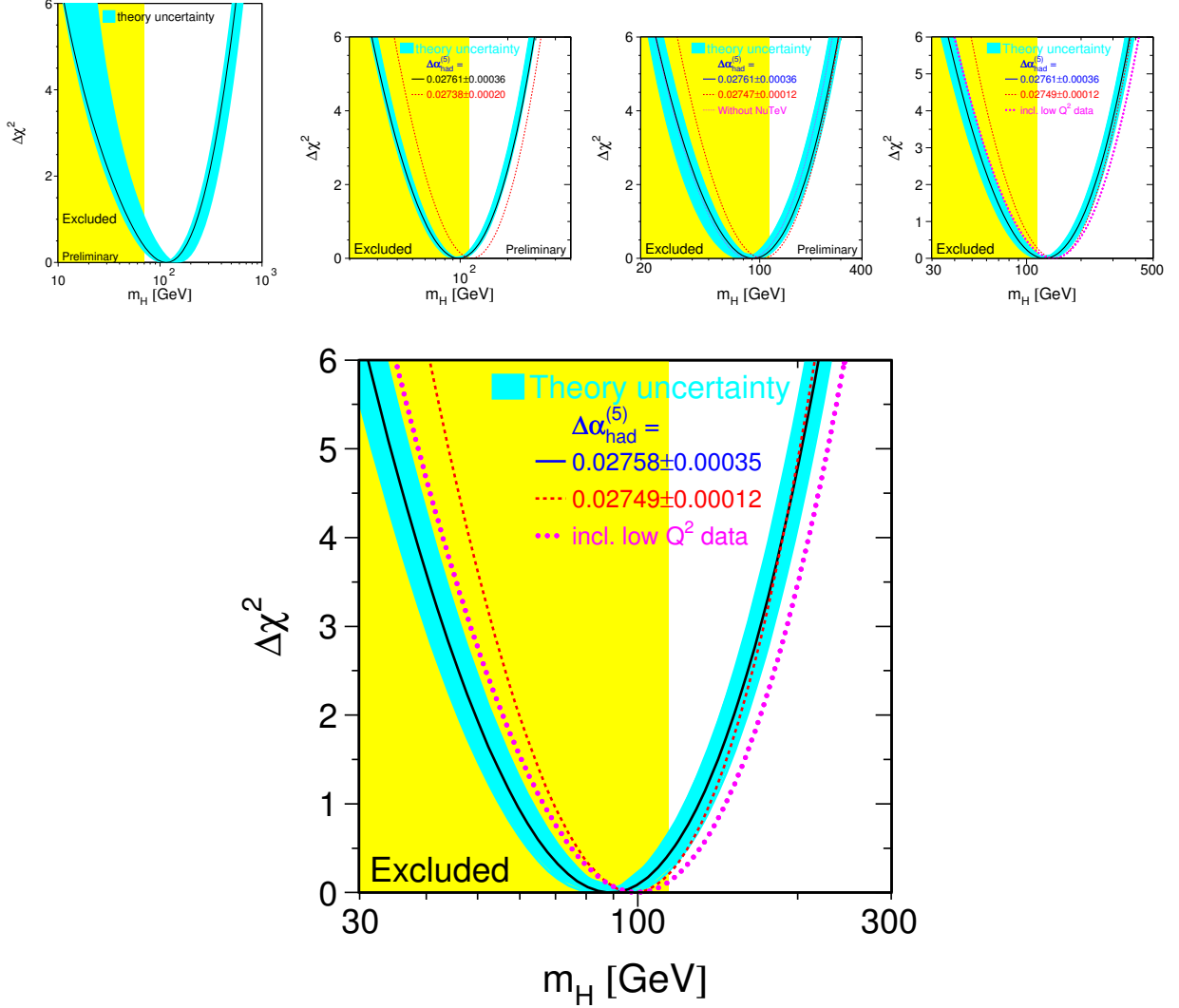
Fig. 21: The blue-band plot showing the constraints on the Higgs mass from precision measurements. The small plots show the same plot from winter conferences of different years: 1997, 2001, 2003 and 2005 (left to right). The big plot dates from winter 2006. The best fit and the width of the parabola vary, most notable due to shifts in the top mass and its uncertainty. From Ref. [11].

7 History and outlook

Finally, instead of a summary, we shall briefly recall the history of 'The making of the Standard Model' following a review by S. Weinberg [1]. It is very instructive to look at this process as the interplay of some 'good ideas' and some 'misunderstandings' which often prevented progress for many years.

1. A 'good idea' was the quark model, proposed in 1964 independently by Gell-Mann and Zweig. The hypothesis that hadrons are made out of three quarks and antiquarks allowed one to understand their quantum numbers and mass spectrum in terms of an approximate SU(3) flavour symmetry, the 'eightfold way'. Furthermore, the deep-inelastic scattering experiments at SLAC in 1968 could be interpreted as elastic scattering of electrons off point-like partons inside the proton, and it was natural to identify these partons with quarks.

   But were quarks real or just some mathematical entities? Many physicists did not believe in quarks

48
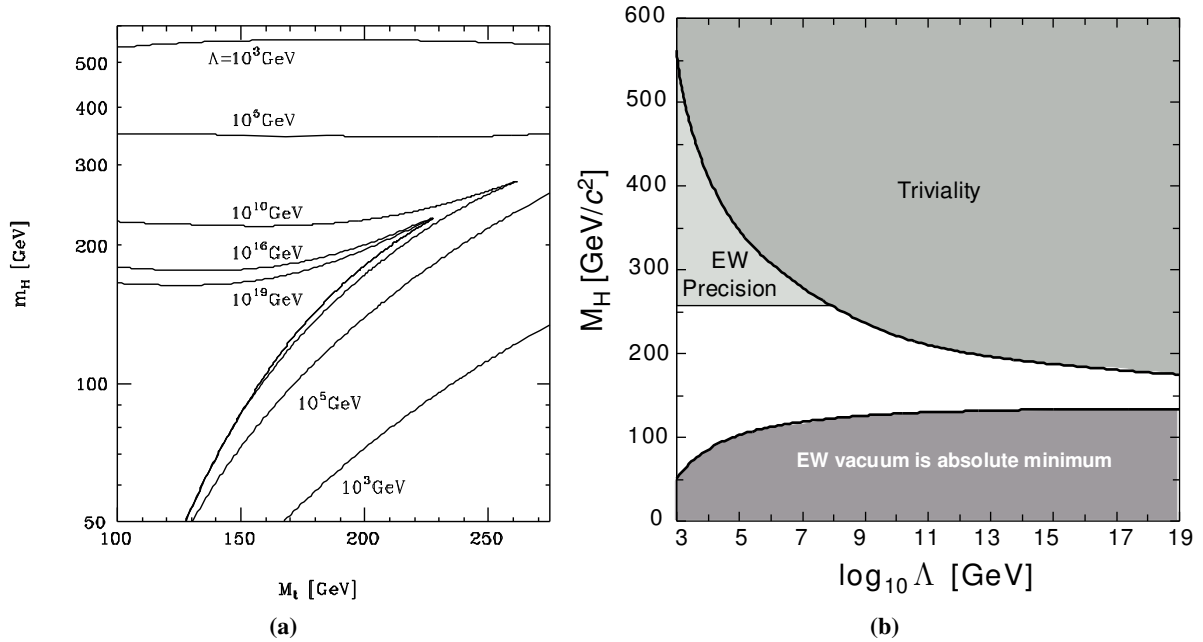
Fig. 22: Bounds on the Higgs and top mass from triviality and vacuum stability. Panel (a) shows the combined bounds for different values of Λ (from Ref. [13]). Panel (b) gives the bounds on the Higgs mass for the known top mass (from Ref. [14]).

since no particles with third integer charges were found despite many experimental searches.

2. Another 'good idea' was the invention of non-Abelian gauge theories by Yang and Mills in 1954. The local symmetry was the isospin group $SU(2)$, and one hoped to obtain in this way a theory of strong interactions with the ρ-mesons as gauge bosons. Only several years later, after the $V - A$ structure of the weak interactions had been identified, did Bludman, Glashow, Salam and Ward and others develop theories of the weak interactions with intermediate vector bosons.

But all physical applications of non-Abelian gauge theories seemed to require massive vector bosons because no massless ones had been found, neither in strong nor weak interactions. Such mass terms had to be inserted by hand, breaking explicitly the local gauge symmetry and thereby destroying the rationale for introducing non-Abelian local symmetries in the first place. Furthermore, it was realized that non-Abelian gauge theories with mass terms would be non-renormalizable, plagued by the same divergences as the four-fermion theory of weak interactions.

3. A further 'good idea' was spontaneous symmetry breaking: There can be symmetries of the Lagrangian that are not symmetries of the vacuum. According to the Goldstone theorem there must be a massless spinless particle for every spontaneously broken global symmetry. On the other hand, there is no experimental evidence for any massless scalar with strong or weak interactions. In 1964 Higgs and Englert and Brout found a way to circumvent Goldstone's theorem: The theorem does not apply if the symmetry is a gauge symmetry as in electrodynamics or the non-Abelian Yang–Mills theory. Then the Goldstone boson becomes the helicity-zero part of the gauge boson, which thereby acquires a mass.

But again, these new developments were applied to broken symmetries in strong interactions, and in 1967 Weinberg still considered the chiral $SU(2)_L \times SU(2)_R$ symmetry of strong interactions to be a gauge theory with the ρ and $a1$ mesons as gauge bosons. In the same year, however, he then applied the idea of spontaneous symmetry breaking to the weak interactions of the leptons of the first family, (ν_L, e_L) and e_R (he did not believe in quarks!). This led to the gauge group

SU(2) × U(1), massive W and Z bosons, a massless photon and the Higgs boson!

The next steps on the way to the Standard Model are well known: The proof by 't Hooft and Veltman that non-Abelian gauge theories are renormalizable and the discovery of asymptotic freedom by Gross and Wilczek and Politzer. Finally, it was realized that the infrared properties of non-Abelian gauge theories lead to the confinement of quarks and massless gluons, and the generation of hadron masses. So, by 1973 'The making of the Standard Model' was completed!

Since 1973 many important experiments have confirmed that the Standard Model is indeed the correct theory of elementary particles:

- 1973: discovery of neutral currents
- 1979: discovery of the gluon
- 1983: discovery of the W and Z bosons
- 1975–2000: discovery of the third family, τ, b, t and ν_τ
- During the past decade impressive quantitative tests have been performed of the electroweak theory at LEP, SLC and Tevatron, and of QCD at LEP, HERA and Tevatron.

Today, there are also a number of 'good ideas' on the market, which lead beyond the Standard Model. These include grand unification, dynamical symmetry breaking, supersymmetry and string theory. Very likely, there are again some 'misunderstandings' among theorists, but we can soon hope for clarifications from the results of the LHC.

Acknowledgements

We would like to thank the participants of the school for stimulating questions and the organizers for arranging an enjoyable and fruitful meeting in Kitzbühel.

Appendix

A Vectors, spinors and γ algebra

A.1 Metric conventions

Our spacetime metric is mostly minus,

$$g_{\mu\nu} = \mathrm{diag}(+,-,-,-) \,, \tag{A.1}$$

so timelike vectors v^μ have positive norm $v_\mu v^\mu > 0$. The coordinate four-vector is $x^\mu = (t, \vec{x})$ (with upper index), and derivatives with respect to x^μ are denoted by

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \vec{\nabla} \right) \,. \tag{A.2}$$

Greek indices $\mu, \nu, \rho, \ldots$ run from 0 to 3, purely spatial vectors are indicated by a vector arrow.

A.2 γ matrices

In four dimensions, the γ matrices are defined by their anticommutators,

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu} \mathbb{1} \,, \quad \mu = 0, \ldots, 3 \,. \tag{A.3}$$

In addition, $\gamma_0 = \gamma_0^\dagger$ is Hermitian while the $\gamma_i = -\gamma_i^\dagger$ are anti-Hermitian, and all γ^μ are traceless. The matrix form of the γ matrices is not fixed by the algebra, and there are several common representations,

like the Dirac and Weyl representations, Eqs. (133) and (58), respectively. However, the following identities hold regardless of the representation.

The product of all γ matrices is

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{A.4}$$

which is Hermitian, squares to one and anticommutes with all γ matrices,

$$\{\gamma^5, \gamma^\mu\} = 0. \tag{A.5}$$

The chiral projectors $P_{L/R}$ are defined as

$$P_{L/R} = \frac{1}{2}\left(1 \pm \gamma^5\right), \quad P_L P_R = P_R P_L = 0, \quad P_{L/R}^2 = P_{L/R}. \tag{A.6}$$

To evaluate Feynman diagrams like for the anomalous magnetic moment, one often needs to contract several γ matrices such as

$$\gamma^\mu \gamma_\mu = 4 \tag{A.7a}$$
$$\gamma^\mu \gamma^\nu \gamma_\mu = -2\gamma^\nu \tag{A.7b}$$
$$\gamma^\mu \gamma^\nu \gamma^\rho \gamma_\mu = 4g^{\nu\rho} \tag{A.7c}$$
$$\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma_\mu = -2\gamma^\sigma \gamma^\rho \gamma^\nu \quad \text{etc.} \tag{A.7d}$$

For a vector v^μ we sometimes use the slash $\slashed{v} = \gamma^\mu v_\mu$.

A.3 Dirac, Weyl and Majorana spinors

The solutions of the Dirac equation in momentum space are fixed by the equations

$$\left(\slashed{p} - m\right)u^{(i)}(p) = 0 \qquad\qquad \left(\slashed{p} + m\right)v^{(i)}(p) = 0. \tag{A.8}$$

Here it is convenient to choose the Weyl representation (58) of the Dirac matrices,

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \Rightarrow \quad \gamma^5 = \begin{pmatrix} -\mathbb{1}_2 & 0 \\ 0 & \mathbb{1}_2 \end{pmatrix}.$$

In this basis, the spinors $u(p)$ and $v(p)$ are given by

$$u^s(p) = \begin{pmatrix} \sqrt{E\mathbb{1}_2 + \vec{p}\cdot\vec{\sigma}}\,\xi^s \\ \sqrt{E\mathbb{1}_2 - \vec{p}\cdot\vec{\sigma}}\,\xi^s \end{pmatrix}, \quad v^s(p) = \begin{pmatrix} \sqrt{E\mathbb{1}_2 + \vec{p}\cdot\vec{\sigma}}\,\eta^s \\ -\sqrt{E\mathbb{1}_2 - \vec{p}\cdot\vec{\sigma}}\,\eta^s \end{pmatrix}. \tag{A.9}$$

Here ξ and η are two-component unit spinors. Choosing the momentum along the z-axis and e.g. $\xi = (1,0)^T$, the positive-energy spinor becomes

$$u^+ = \begin{pmatrix} \sqrt{E + p_z} \\ 0 \\ \sqrt{E - p_z} \\ 0 \end{pmatrix}, \tag{A.10}$$

which has spin $+\frac{1}{2}$ along the z-axis. For $\xi = (0,1)^T$, the spin is reversed, and similar for η and the negative energy spinors.

The spinors considered so far are called Dirac spinors: They are restricted only by the Dirac equation and have four degrees of freedom (particle and antiparticle, spin up and spin down). There are two restricted classes of spinors, Weyl and Majorana spinors, which have only two degrees of freedom.

Weyl or chiral spinors are subject to the constraint

$$P_L \psi_L = \psi_L \quad \text{or} \quad P_R \psi_R = \psi_R \tag{A.11}$$

and correspond to purely left- or right-handed fermions. In the language of u's and v's, chiral spinors correspond to sums $u \pm \gamma^5 v$. Chiral spinors can have a kinetic term, but no usual mass term, since

$$\overline{(\psi_L)} = \overline{P_L \psi_L} = (P_L \psi_L)^\dagger \gamma^0 = \psi_L^\dagger P_L \gamma^0 = \overline{\psi}_L P_R \tag{A.12}$$

and hence

$$\overline{\psi_L}\, \psi_L = \overline{\psi}_L \underbrace{P_R P_L}_{=0} \psi_L = 0 \,. \tag{A.13}$$

However, there is the possibility of a Majorana mass term via the charge conjugate spinor ψ^C:

$$\psi^C = C \overline{\psi}^T \quad \text{with the charge conjugation matrix} \quad C = \mathrm{i}\gamma^0 \gamma^2 \,. \tag{A.14}$$

ψ^C is of opposite chirality to ψ, so it can be used to build a bilinear $\overline{\psi}^C \psi$ for a mass term. However, this term violates all symmetries under which ψ is charged, so it is acceptable only for complete singlets, like right-handed neutrinos.

References

[1] S. Weinberg, The making of the Standard Model, Eur. Phys. J. C **34** (2004) 5 [arXiv:hep-ph/0401010].

[2] O. Nachtmann, Elementary Particle Physics: Concepts and Phenomena (Springer, Berlin, 1990).

[3] M. E. Peskin and D. V. Schroeder, An Introduction to Quantum Field Theory (Addison-Wesley, Reading, MA, 1995).

[4] J. Ellis, lecture given at this School (unpublished).

[5] G. Ecker, lecture given at this School, see these proceedings.

[6] R. Fleischer, lecture given at this School, see these proceedings.

[7] M. Lindner, lecture given at this School (unpublished).

[8] S. Weinberg, Gravitation and Cosmology (Wiley, New York, 1972), pp. 61–63.

[9] G. 't Hooft and M. Veltman, Diagrammar, CERN report 73–9 (1973).

[10] T. Kinoshita, Nucl. Phys. Proc. Suppl. **157** (2006) 101.

[11] LEP Electroweak Working Group, http://lepewwg.web.cern.ch.

[12] M. Spira, Fortschr. Phys. **46** (1998) 203 [arXiv:hep-ph/9705337].

[13] C. D. Froggatt, Surveys High Energ. Phys. **18** (2003) 77 [arXiv:hep-ph/0307138].

[14] C. Quigg, Acta Phys. Polon. B **30** (1999) 2145 [arXiv:hep-ph/9905369].

Index

Quantum chromodynamics

G. Ecker

Institute for Theoretical Physics, University of Vienna, Austria

Abstract

After a brief historical review of the emergence of QCD as the quantum field theory of strong interactions, the basic notions of colour and gauge invariance are introduced leading to the QCD Lagrangian. The second lecture is devoted to perturbative QCD, from tree-level processes to higher-order corrections in renormalized perturbation theory, including jet production in e^+e^- annihilation, hadronic τ decays and deep inelastic scattering. The final two lectures treat various aspects of QCD beyond perturbation theory. The main theme is effective field theories, from heavy quarks to the light quark sector where the spontaneously broken chiral symmetry of QCD plays a crucial role.

1 Introduction

Why do we still study QCD after more than 30 years?

- By decision of the Nobel Prize Committee in 2004 [1], QCD is the correct theory of the strong interactions.
- The parameters of QCD, the coupling strength α_s and the quark masses, need to be measured as precisely as possible.
- Electroweak processes of hadrons necessarily involve the strong interactions.
- In searches for new physics at present and future accelerators, the 'QCD background' must be understood quantitatively.
- Although QCD is under control for high-energy processes, many open questions remain in the nonperturbative domain (confinement, chiral symmetry breaking, hadronization, etc.).
- Last but not least, QCD is a fascinating part of modern physics. The lectures will therefore start with a brief historical review of the developments in particle physics in the 1960s and early 1970s.

The following lectures were given to an audience of young experimental particle physicists. Although the lectures emphasize some of the theoretical aspects of QCD, the mathematical level was kept reasonably low. The first two lectures cover the basics of QCD, from the concepts of colour and gauge invariance to some applications of perturbative QCD. The last two lectures treat aspects of QCD beyond perturbation theory. The main theme is effective field theories, from heavy quarks to the light quark sector where spontaneously broken chiral symmetry plays a crucial role.

1.1 Historical background

Particle physics in the early 1960s was not in a very satisfactory state. Only for the electromagnetic interactions of leptons was a full-fledged quantum field theory (QFT) available. Quantum electrodynamics (QED) produced increasingly precise predictions that were confirmed experimentally. Nevertheless, the methodology of renormalization, an essential aspect of the perturbative treatment of QED, was not universally accepted. Even among the founding fathers of QFT, the dissatisfaction with "sweeping the infinities under the rug" was widespread. At the Solvay Conference of 1961, Feynman confessed [2] that he did not "subscribe to the philosophy of renormalization".

What is the essence of this controversial procedure of renormalization that has turned out to be crucial for the shaping of both QCD and the Standard Model? In the case of QED three main steps are important.

   – Amplitudes $A(p_i; e_0, m_0; \Lambda)$ depend on the momenta p_i of the particles involved, on the parameters e_0, m_0 of the QED Lagrangian and on a cut-off Λ that cuts off the high-momentum modes of the theory. The cut-off is essential because $A(p_i; e_0, m_0; \Lambda)$ diverges for $\Lambda \to \infty$, rendering the result meaningless.

   – With the help of measurable quantities (cross-sections, particle four-momenta) one defines physical parameters $e(\mu), m(\mu)$ that depend in general on an arbitrary renormalization scale μ. One then trades e_0, m_0 for the physical $e(\mu), m(\mu)$ to a given order in perturbation theory.

   – The limit $\lim_{\Lambda \to \infty} A(p_i; e_0(e, m, \Lambda), m_0(e, m, \Lambda); \Lambda) = \hat{A}(p_i; e(\mu), m(\mu))$ is now finite and unambiguous for the chosen definitions of $e(\mu), m(\mu)$.

Based on this procedure, the agreement between theory and experiment was steadily improving.

For the weak interactions, the Fermi theory (in the $V - A$ version) was quite successful for weak decays, but

   – higher-order corrections were not calculable;

   – for scattering processes the theory became inconsistent for energies $E \gtrsim 300$ GeV (unitarity problem).

The rescue came at the end of the 1960s in the form of the electroweak gauge theory of the Standard Model.

Of all the fundamental interactions, the strong interactions were in the most deplorable state. Although the rapidly increasing number of hadrons could be classified successfully by the quark model of Gell-Mann and Zweig [3], the dynamics behind the quark model was a complete mystery. A perturbative treatment was clearly hopeless and the conviction gained ground that QFT might not be adequate for the strong interactions.

This conviction was spelled out explicitly by the proponents of the bootstrap philosophy (Chew *et al.*). Under the banner of nuclear democracy, all hadrons were declared to be equal. Instead of looking for more fundamental constituents of hadrons, the S-matrix for strong processes was investigated directly without invoking any quantum field theory. Although the expectations were high, nuclear democracy shared the fate of the student movement of the late 1960s: the promises could not be fulfilled.

A less radical approach assumed that QFT could still be useful as a kind of toy model. The main proponent of this approach was Gell-Mann who suggested to abstract algebraic relations from a Lagrangian field theory model but then throw away the model ("French cuisine program" [4]). The usefulness of this approach had been demonstrated by Gell-Mann himself: current algebra and the quark model were impressive examples. Until the early 1970s, Gell-Mann took his programme seriously in declaring the quarks to be purely mathematical entities without any physical reality, a view shared by many particle physicists of the time.

The decisive clue came from experiment. Started by the MIT–SLAC Collaboration at the end of the 1960s, deep inelastic scattering of leptons on nucleons and nuclei produced unexpected results. Whereas at low energies the cross-sections were characterized by baryon resonance production, the behaviour at large energies and momentum transfer was surprisingly simple: the nucleons seemed to consist of noninteracting partons (Feynman). Obvious candidates for the partons were the quarks but this idea led to a seeming paradox. How could the quarks be quasi-free at high energies and yet be permanently bound in hadrons, a low-energy manifestation?

That the strength of an interaction could be energy dependent was not really new to theorists. In QED, the vacuum acts like a polarizable medium leading to the phenomenon of charge screening. However, contrary to what the deep inelastic experiments seemed to suggest for the strong interactions, the effective charge in QED increases with energy: QED is ultraviolet unstable.

To understand the phenomenon of an energy-dependent interaction, we consider the dimensionless

ratio of cross-sections

$$R_{e^+e^-} = \frac{\sigma(e^+ + e^- \to \text{hadrons})}{\sigma(e^+ + e^- \to \mu^+ + \mu^-)} . \tag{1}$$

Beyond the leading-order value R_0 (cf. Section 2.1), one finds to lowest order in the strong coupling constants g_s

$$R_{e^+e^-} = R_0 \left(1 + \frac{g_s^2}{4\pi^2}\right) . \tag{2}$$

The general form to any order in g_s (neglecting quark masses) is

$$R_{e^+e^-} = R_{e^+e^-}(E, \mu, g_s(\mu)) \tag{3}$$

where E is the centre-of-mass energy and μ is the renormalization scale. Since $R_{e^+e^-}$ is a measurable quantity, it must be independent of the arbitrary scale μ:

$$\mu \frac{d}{d\mu} R_{e^+e^-}(E, \mu, g_s(\mu)) = 0 \qquad \longrightarrow \qquad \left(\mu \frac{\partial}{\partial \mu} + \beta(g_s) \frac{\partial}{\partial g_s}\right) R_{e^+e^-} = 0 , \tag{4}$$

with the beta function

$$\beta(g_s) = \mu \frac{dg_s(\mu)}{d\mu} . \tag{5}$$

Dimensional analysis tells us that the dimensionless ratio $R_{e^+e^-}$ must be of the form

$$R_{e^+e^-}(E, \mu, g_s(\mu)) = f(\frac{E}{\mu}, g_s(\mu)) . \tag{6}$$

The seemingly uninteresting dependence on μ can therefore be traded for the dependence on energy or on the dimensionless ratio $z = E/\mu$:

$$\left(-\frac{\partial}{\partial \log z} + \beta(g_s) \frac{\partial}{\partial g_s}\right) f(z, g_s(\mu)) = 0 . \tag{7}$$

The general solution of this renormalization group equation is

$$f(z, g_s(\mu)) = \hat{f}(\overline{g_s}(z, g_s)) , \tag{8}$$

i.e., a function of a single variable, the energy-dependent (running) coupling constant $\overline{g_s}(z, g_s)$ satisfying

$$\frac{\partial \overline{g_s}}{\partial \log z} = \beta(\overline{g_s}) \tag{9}$$

with the boundary condition $\overline{g_s}(1, g_s) = g_s$. For any gauge coupling, the leading one-loop result for the β function is

$$\beta(x) = -\frac{\beta_0}{(4\pi)^2} x^3 \tag{10}$$

implying

$$\overline{g_s}^2(E/\mu, g_s(\mu)) = \frac{g_s^2(\mu)}{1 + \frac{\beta_0}{(4\pi)^2} g_s^2(\mu) \log E^2/\mu^2} . \tag{11}$$

Expanding the denominator, we observe that the renormalization group equation has allowed us to sum the leading logs $\left(g_s^2(\mu) \log E^2/\mu^2\right)^n$ of all orders in perturbation theory. Even more importantly, the energy dependence of the running coupling constant is determined by the sign of β_0 in Eq. (10):

$\beta_0 < 0$:	$\displaystyle\lim_{E\to 0} \overline{g}(E) = 0$	infrared stable (QED)
$\beta_0 > 0$:	$\displaystyle\lim_{E\to\infty} \overline{g}(E) = 0$	ultraviolet stable (QCD)

For the cross-section ratio $R_{e^+e^-}$ we get finally

$$R_{e^+e^-} = R_0 \left(1 + \frac{g_s^2(E)}{4\pi^2} + O(g_s^4(E)) \right) \tag{12}$$

in terms of $g_s^2(E) \equiv \overline{g_s}^2(E/\mu, g_s(\mu))$.

The crucial question in the early 1970s was therefore whether QFT was compatible with ultraviolet stability (asymptotic freedom)? The majority view was expressed in a paper by Zee [5]: " ... we conjecture that there are no asymptotically free quantum field theories in four dimensions." While Coleman and Gross set out to prove that conjecture, their graduate students Politzer and Wilczek (together with Gross) tried to close a loophole: the β function for non-Abelian gauge theories (Yang–Mills theories) was still unpublished and probably unknown to everybody except t'Hooft. In the spring of 1973, the Nobel prize winning work of Politzer, Gross, and Wilczek [6] demonstrated that Yang–Mills theories are indeed asymptotically free.

The crucial difference between QED and QCD is that photons are electrically neutral whereas the gluons as carriers of the strong interactions are coloured. Further physical insight can be obtained by taking up an analogy with the electrodynamics of continuous media [7]. Because of Lorentz invariance, the vacuum of a relativistic QFT is characterized by

$$\varepsilon\mu = 1 \tag{13}$$

for the product of permittivity ε and permeability μ. In QED, charge screening implies $\varepsilon > 1$ so that the vacuum of QED acts like a diamagnet ($\mu < 1$). In QCD, the colour charge screening of quarks ($\varepsilon > 1$) is overcompensated by gluons (spin 1) acting as permanent colour dipoles ($\mu > 1$). Because

$$\beta_0^{\mathrm{QCD}} = \frac{1}{3}\left(11 N_c - 2 N_F\right) , \tag{14}$$

the QCD vacuum is a (colour) paramagnet for $N_F < 11\,N_c/2 < 17$ quark flavours (for $N_c = 3$). Because of the general relation (13) this can also be interpreted as anti-screening ($\varepsilon < 1$).

The existence of three colours was already widely accepted at that time. Gell-Mann and collaborators had been investigating a model of coloured quarks interacting via a singlet gluon (not asymptotically free). In a contribution of Fritzsch and Gell-Mann in the Proceedings of the International Conference on High-Energy Physics in Chicago in 1972 [8] one finds probably the first reference to non-Abelian gluons: "Now the interesting question has been raised lately whether we should regard the gluons as well as the quarks as being non-singlets with respect to colour (J. Wess, private communication to B. Zumino)". Although Gell-Mann is generally credited for the name QCD, the first published occurrence of QCD is much less known (cf., for example, Refs. [1]). My own investigation of the early literature has produced a footnote in a paper of Fritzsch, Gell-Mann and Minkowski in 1975 [9] suggesting "A good name for this theory is quantum chromodynamics".

1.2 Colour

Already before the arrival of QCD, there were a number of indications for the colour degree of freedom.

- Triality problem
  In the original quark model, often called the naive quark model, the three quarks u, d, s give

rise to mesonic bound states of the form $\overline{q}q$. All of the nine expected bound states had already been observed suggesting an attractive force between all quarks and antiquarks. The baryons fit nicely into qqq bound states. If the strong force is purely attractive why do antiquarks not bind to baryons? The resulting objects of the form $\overline{q}qqq$ would have fractional charge and have never been observed. Introducing three colours for each quark and antiquark allows for $9 \times 9 = 81$ combinations of $\overline{q}q$, only nine of which had been found. The remaining 72 combinations are not bound states invalidating the previous argument.

- Spin-statistics problem

  Consider the state

  $$|\Delta^{++}(S_z = 3/2)\rangle \sim |u \uparrow u \uparrow u \uparrow\rangle . \qquad (15)$$

  Since the spin-flavour content is completely symmetric, Fermi statistics for quarks seems to require an antisymmetric spatial wave function. On the other hand, for every reasonable potential the ground state is symmetric in the space variables. Colour solves this problem because the state (15) is totally antisymmetric in the colour indices respecting the generalized Pauli principle with a spatially symmetric wave function.

- Renormalizability of the Standard Model

  With the usual charges of quarks and leptons, the Standard Model is a consistent gauge invariant QFT only if there are three species of quarks in order to cancel the so-called gauge anomalies.

- $\pi^0 \to 2\,\gamma$ decay

  The by far dominant contribution to the decay amplitude is due to the chiral anomaly (exact for massless quarks). The observed rate can only be understood if there are again three species of quarks.

- Momentum balance in deep inelastic scattering

  The momentum sum rule indicates that only about 50% of the nucleon momentum is carried by valence quarks. The remainder is mainly carried by gluons.

- Quark counting

  After the advent of QCD, many more direct confirmations of the colour degree of freedom were obtained. One of the first confirmations was provided by the total cross-section $\sigma(e^+ + e^- \to \text{hadrons})$ already discussed in the previous subsection.

- Hadronic τ decays

  As we shall discuss in the next lecture, hadronic τ decays not only give clear evidence for $N_c = 3$ but they also provide an excellent opportunity for extracting $\alpha_s = g_s^2/4\pi$.
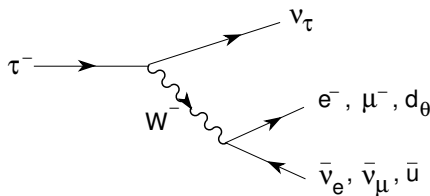
$$R_\tau = \frac{\Gamma(\tau^- \to \nu_\tau + \text{hadrons})}{\Gamma(\tau^- \to \nu_\tau e^- \overline{\nu_e})} \qquad (16)$$
$$= N_c \left(|V_{ud}|^2 + |V_{us}|^2\right)(1 + O(\alpha_s)) .$$

Fig. 1: Feynman diagram for $\tau^- \to \nu_\tau + X$

1.3 Gauge invariance

Gauge invariance is a main ingredient not only of QCD but of the Standard Model as a whole. We start with the Lagrangian for a single free Dirac fermion:

$$\mathcal{L}_0 = \overline{\psi}(x)\, i\, \partial\!\!\!/\, \psi(x) - m\, \overline{\psi}(x)\, \psi(x) , \qquad\qquad \partial\!\!\!/ := \gamma^\mu a_\mu . \qquad (17)$$

This Lagrangian and the resulting Dirac equation are invariant under a phase transformation (global $U(1)$)

$$\psi(x) \longrightarrow \psi'(x) = e^{-iQ\,\varepsilon}\psi(x) \tag{18}$$

with $Q\varepsilon$ an arbitrary real constant. One may now pose the question whether the phase in the transformation law (18) must really be the same here and 'behind the moon', as is the case in (18) with a space–time-independent phase $Q\varepsilon$. Instead of experimenting behind the moon, we replace the constant ε with an arbitrary real function $\varepsilon(x)$ and see what happens. As is easily checked, the mass term in (17) remains invariant but not the kinetic term because

$$\partial_\mu\psi(x) \longrightarrow e^{-iQ\varepsilon(x)}\left(\partial_\mu - iQ\partial_\mu\varepsilon(x)\right)\psi(x)\,. \tag{19}$$

The conclusion is that the phase must indeed be the same here and behind the moon for the theory to be invariant under transformations of the form (18).

However, there is a well-known procedure for enforcing local invariance, i.e., invariance for a completely arbitrary space–time-dependent phase $\varepsilon(x)$. We enlarge the theory by introducing a spin-1 vector field A_μ that has precisely the right transformation property to cancel the obnoxious piece in (19) with $\partial_\mu\varepsilon(x)$. The idea is to replace the ordinary derivative ∂_μ by a covariant derivative D_μ:

$$D_\mu\psi(x) \;=\; \left(\partial_\mu + iQA_\mu(x)\right)\psi(x)\,, \tag{20}$$

with

$$A_\mu(x) \longrightarrow A'_\mu(x) \;=\; A_\mu(x) + \partial_\mu\varepsilon(x)\,. \tag{21}$$

It is easy to check that $D_\mu\psi$ transforms covariantly,

$$D_\mu\psi(x) \longrightarrow (D_\mu\psi)'(x) = e^{-iQ\varepsilon(x)}D_\mu\psi(x)\,, \tag{22}$$

so that the enlarged Lagrangian

$$\mathcal{L} = \overline{\psi}(x)\left(i\,\slashed{D} - m\right)\psi(x) = \mathcal{L}_0 - QA_\mu(x)\overline{\psi}(x)\,\gamma^\mu\psi(x) \tag{23}$$

is invariant under local $U(1)$ transformations (gauge invariance).

The most important feature of this exercise is that the requirement of gauge invariance has generated an interaction between the fermion field ψ and the gauge field A_μ. Introducing a kinetic term for the gauge field to promote it to a propagating quantum field, the full Lagrangian

$$\mathcal{L} = \overline{\psi}\left(i\,\slashed{D} - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{24}$$

is still gauge invariant because the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is automatically gauge invariant. Setting $Q = -e$ for the electron field, we have 'deduced' QED from the free electron theory and the requirement of gauge invariance. For completeness, we take note that a mass term of the form $M_\gamma^2 A_\mu A^\mu$ is forbidden by gauge invariance, implying massless photons.

1.4 $SU(3)_c$ and the QCD Lagrangian

As we have seen, quarks come in three colours. Suppressing all space–time dependence, the free quark Lagrangian for a single flavour has the form

$$\mathcal{L}_0 = \sum_{i=1}^{3}\overline{q_i}\left(i\,\slashed{\partial} - m_q\right)q_i\,. \tag{25}$$

Assuming the three quarks with different colours to have the same mass m_q (different flavours have different masses, of course), we can ask for the global invariances of $\mathcal{L}_0$. All transformations that leave $\mathcal{L}_0$ invariant are of the form

$$q_i \longrightarrow q_i' = U_{ij}q_j \,, \qquad U\,U^\dagger = U^\dagger\,U = \mathbb{1} \tag{26}$$

with arbitrary unitary matrices U_{ij}. Splitting off a common phase transformation $q_i \to e^{-i\varepsilon} q_i$ treated previously and generating electromagnetic interactions that we know to be colour-blind, we are left with the special unitary group $SU(3)$ comprising all three-dimensional unitary matrices with unit determinant.

In contrast to the $U(1)$ case treated before, we now have eight independent transformations in $SU(3)$ (an 8-parameter Lie group). Therefore, continuing in the same spirit as before, it is not just a question of demanding gauge invariance but we also have to find out which part of $SU(3)$ should be gauged. With hindsight, the following two criteria lead to a unique solution.

  i. The three colours are not like three arbitrary electric charges but are instead intimately connected through gauge transformations. This requires the quarks to be in an irreducible three-dimensional representation leaving only two possibilities: either all of $SU(3)$ or one of the $SU(2)$ subgroups must be gauged.

  ii. Quarks and antiquarks transform differently under gauge transformations ($\underline{3} \neq \underline{3}^*$). This closes the case and implies that all $SU(3)$ transformations must be gauged.

The above requirements guarantee that both $\overline{q}q$ and qqq contain colour singlets = hadrons,

$$\underline{3}^* \otimes \underline{3} = \underline{1} \oplus \underline{8}, \qquad \underline{3} \otimes \underline{3} \otimes \underline{3} = \underline{1} \oplus \underline{8} \oplus \underline{8} \oplus \underline{10} \,, \tag{27}$$

but neither qq nor $\overline{q}qqq$ or other exotic combinations.

Every three-dimensional unitary matrix with $\det U = 1$ can be written as

$$U(\varepsilon_a) = \exp\{-i\sum_{a=1}^{8} \varepsilon_a \frac{\lambda_a}{2}\} \tag{28}$$

with eight parameters ε_a and with eight traceless Hermitian Gell-Mann matrices λ_a. Their commutation relations define the Lie algebra of $SU(3)$:

$$[\lambda_a, \lambda_b] = 2i\,f_{abc}\lambda_c \,, \tag{29}$$

with real, totally antisymmetric structure constants f_{abc}.

The gauge principle demands invariance of the theory for arbitrary space–time-dependent functions $\varepsilon_a(x)$. Instead of a single gauge field A_μ, we now need eight vector fields $G_a^\mu(x)$ entering the covariant derivative

$$(D^\mu q)_i = \left(\partial^\mu \delta_{ij} + i\,g_s \sum_{a=1}^{8} G_a^\mu \frac{\lambda_{a,ij}}{2}\right) q_j =: \{(\partial^\mu + ig_s G^\mu)q\}_i \,. \tag{30}$$

The real coupling constant g_s measures the strength of the quark–gluon interaction just as the charge Q is a measure of the electromagnetic interaction. Using the convenient matrix notation (summation convention implied)

$$G_{ij}^\mu := G_a^\mu \frac{\lambda_{a,ij}}{2} \,, \tag{31}$$

the covariant derivative now transforms as

$$G_\mu \longrightarrow G_\mu' = U(\varepsilon)G_\mu U^\dagger(\varepsilon) + \frac{i}{g_s}\left(\partial_\mu U(\varepsilon)\right)U^\dagger(\varepsilon) \tag{32}$$

in order for $(D_\mu q)_i$ to transform like the q_i themselves (covariance requirement).

Because of the non-Abelian character of $SU(3)$, the transformation laws are more complicated than in the electromagnetic case. The differences can already be seen in the infinitesimal transformations of the gluon fields G_a^μ following from (32):

$$G_a^\mu \longrightarrow G_a^{\mu'} = G_a^\mu + \frac{1}{g_s}\partial^\mu \varepsilon_a + f_{abc}\varepsilon_b G_c^\mu + O(\varepsilon^2) . \tag{33}$$

In order to have propagating gluon fields, we need an analogue of the electromagnetic field strength tensor $F_{\mu\nu}$. The simplest approach is to calculate the commutator of two covariant derivatives:

$$[D_\mu, D_\nu] = [\partial_\mu + i\,g_s G_\mu, \partial_\nu + i\,g_s G_\nu] =: i\,g_s G_{\mu\nu} . \tag{34}$$

The non-Abelian field strength tensor $G^{\mu\nu} = G_a^{\mu\nu}\frac{\lambda_a}{2}$ has the explicit form

$$\begin{aligned} G^{\mu\nu} &= \partial^\mu G^\nu - \partial^\nu G^\mu + ig_s\,[G^\mu, G^\nu] \\ G_a^{\mu\nu} &= \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f_{abc} G_b^\mu G_c^\nu \end{aligned} \tag{35}$$

and it transforms covariantly under $SU(3)$ gauge transformations:

$$G_{\mu\nu} \longrightarrow G'_{\mu\nu} = U(\varepsilon)G_{\mu\nu}U^\dagger(\varepsilon) . \tag{36}$$

The gauge-invariant colour trace

$$\mathrm{tr}(G_{\mu\nu}G^{\mu\nu}) = \frac{1}{2}G_a^{\mu\nu}G_{\mu\nu}^a \tag{37}$$

has the right structure for a gluon kinetic term leading immediately to the $SU(3)_c$ invariant QCD Lagrangian for $f = 1, \ldots, N_F$ quark flavours:

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{2}\mathrm{tr}(G_{\mu\nu}G^{\mu\nu}) + \sum_{f=1}^{N_F} \overline{q}_f \left(i\slashed{D} - m_f \mathbb{1}_c\right) q_f . \tag{38}$$

As in the $U(1)$ case, gauge invariance requires massless gluons. Writing out the Lagrangian (38) in detail, one finds three types of vertices instead of a single one for QED:

$$\overline{q}qG \qquad\qquad GGG \qquad\qquad GGGG$$
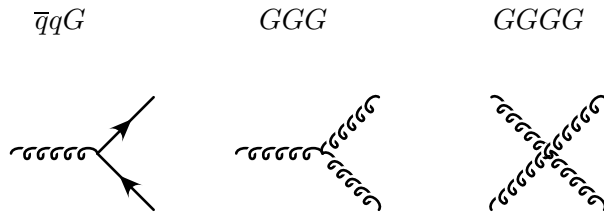
Fig. 2: Basic vertices of QCD

In addition to the quark masses, QCD has a single parameter describing the strength of the strong interactions, the strong coupling constant g_s ($\alpha_s = g_s^2/4\pi$.

$$\boxed{\text{Experimental group theory}}$$

Can experimentalists determine more than a single coupling strength α_s in a strong process? All the information is contained in the vertices: in addition to g_s, the vertices also contain the two matrices

$$(t_a^F)_{ij} = \frac{1}{2}(\lambda_a)_{ij} , \qquad\qquad (t_a^A)_{bc} = -i f_{abc} , \qquad\qquad (39)$$

defining the fundamental and adjoint representations of (the Lie algebra of) $SU(3)$:

$$[t_a, t_b] = i\, f_{abc} t_c . \qquad\qquad (40)$$

Let us pretend for a moment that we do not know that there are three colours and eight gluons. For a general (compact Lie) group of symmetry transformations, the vertices are again determined by quark and gluon representation matrices t_a^F, t_a^A. The combinations that actually appear in measurable quantities are the following traces and sums:

$$\text{tr}(t_a^R t_b^R) = T_R \delta_{ab}, \qquad \sum_a (t_a^R)_{ij}(t_a^R)_{jk} = C_R \delta_{ik} \qquad (R = F, A) , \qquad\qquad (41)$$

with
$$\begin{aligned} &T_R: && \text{Dynkin index for the representation R;} \\ &C_R: && \text{(quadratic) Casimir for R.} \end{aligned}$$

For a d_R-dimensional representation, one derives from the definitions (41) the general relation

$$d_R\, C_R = n_G\, T_R \qquad\qquad (42)$$

where n_G is the number of independent parameters of G. Restricting the discussion to $SU(n)$, the two cases of interest are

$R = A$ (adjoint representation):
$$d_A = n_G \qquad \longrightarrow \qquad C_A = T_A = n \ \ \text{for } SU(n);$$

$R = F$ (fundamental representation of $SU(n)$):
$$d_F = n, \ n_G = n^2 - 1, \ T_F = 1/2 \qquad \longrightarrow \qquad C_F = \frac{n^2 - 1}{2n}$$

and for the special case of $SU(3)$: $\qquad C_F = \dfrac{4}{3}, \ C_A = T_A = N_c = 3 .$

The independent quantities that can be measured are C_F and C_A. A combined jet analysis in $e^+ e^-$ annihilation at LEP found [10]

$$C_F = 1.30 \pm 0.01(\text{stat}) \pm 0.09(\text{sys}), \qquad C_A = 2.89 \pm 0.03(\text{stat}) \pm 0.21(\text{sys}) \qquad\qquad (43)$$

in manifest agreement with $SU(3)$.

Feynman diagrams are constructed with the vertices and propagators of the QCD Lagrangian (38). The problem here is the same as in QED: because of the gauge invariance of (38), the gluon propagator does not exist. At least for perturbation theory, the inescapable consequence is that gauge invariance must be broken in the Lagrangian! Or, in a more euphemistic manner of speaking, the gauge must be fixed. In the simplest and widely used version (covariant gauge with real parameter ξ) the Lagrangian (38) is replaced by

$$\mathcal{L}_{\text{QCD}} \qquad \longrightarrow \qquad \mathcal{L}_{\text{QCD}} - \frac{\xi}{2}(\partial_\mu G_a^\mu)^2 + \mathcal{L}_{\text{ghost}} . \qquad\qquad (44)$$

The gluon propagator now exists ($\xi = 1$: Feynman gauge):

$$\Delta_{ab}^{\mu\nu}(k) = \delta_{ab} \frac{-i}{k^2 + i\epsilon}\left(g^{\mu\nu} + (\xi^{-1} - 1)\frac{k^\mu k^\nu}{k^2}\right) \overset{\xi=1}{=} \delta_{ab}\frac{-i\, g^{\mu\nu}}{k^2 + i\epsilon} . \qquad\qquad (45)$$

The additional ghost Lagrangian $\mathcal{L}_{\text{ghost}}$ repairs the damage done by gauge fixing: although Green functions are now gauge dependent, observable S-matrix elements are still gauge invariant and therefore independent of ξ (Feynman, Faddeev, Popov, BRST, etc.).

2 Perturbative QCD

2.1 QCD at tree level

The calculation of tree amplitudes in QCD is straightforward but

- to compare theory with experiment, we must have hadrons rather than quarks and gluons in the initial and final states;
- amplitudes and cross-sections are in general infrared divergent for massless gluons.

The general recipe is to consider infrared safe quantities, the more inclusive the better. A good example is once more $e^+e^- \rightarrow$ hadrons. The sum over all hadronic final states can be expressed in terms of the
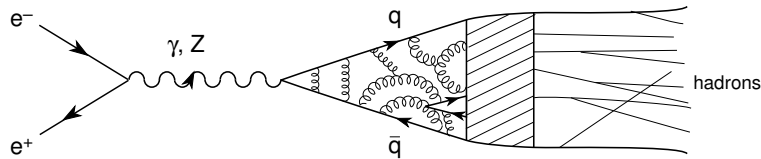
Fig. 3: $e^+e^- \rightarrow$ hadrons

imaginary (absorptive) part of the two-point function of electromagnetic currents (photonic case), the hadronic vacuum polarization:

$$\Pi^{\mu\nu}_{\mathrm{em}}(q) = i \int d^4x \, e^{iq \cdot x} \langle 0| \, T J^\mu_{\mathrm{em}}(x) J^\nu_{\mathrm{em}}(0)|0\rangle = \left(-g^{\mu\nu}q^2 + q^\mu q^\nu\right) \Pi_{\mathrm{em}}(q^2) \, . \tag{46}$$

The sum over all intermediate states can be performed either with quarks and gluons or with hadrons.
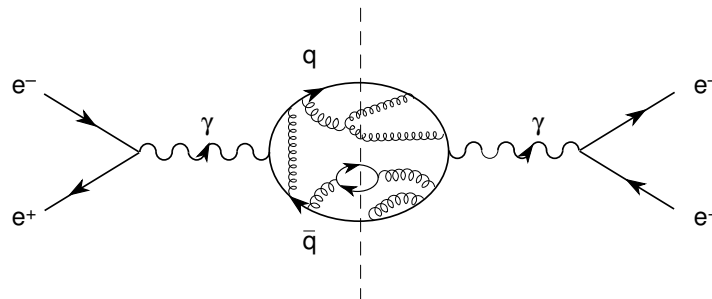
Fig. 4: Hadronic vacuum polarization

Since there are no massless hadrons, the hadronic vacuum polarization is infrared safe.

To lowest order in QCD, the amplitude for

$$e^+e^- \rightarrow \gamma^*(Z^*) \rightarrow \overline{q}q \tag{47}$$

is in fact independent of the strong coupling constant g_s. Except for the charges, masses and multiplicities of quarks, it is the same calculation as for $e^+e^- \rightarrow \mu^+\mu^-$ in QED. Therefore, for quarks with given flavour f and colour i the amplitude is (neglecting m_e, m_μ, m_q)

$$A(e^+e^- \rightarrow \overline{q}^i_f q^i_f) = \frac{Q_f}{e} A(e^+e^- \rightarrow \mu^+\mu^-) \, . \tag{48}$$

Quarks and antiquarks with different colour and flavour are in principle distinguishable so that the total hadronic cross-section, normalized to $\sigma(e^+e^- \to \mu^+\mu^-)$, is

$$R_{e^+e^-} = \frac{\sigma(e^+e^- \to \text{ hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = \sum_{i,f} Q_f^2/e^2 = N_c \sum_f Q_f^2/e^2 . \tag{49}$$

As shown in Fig. 5, this is a good approximation to the experimental data between quark thresholds.
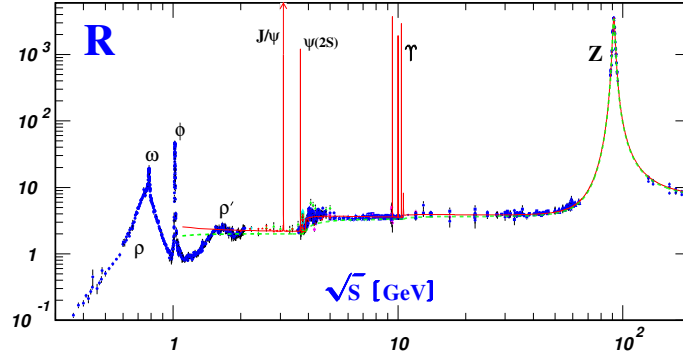
Fig. 5: Experimental data for $R_{e^+e^-}$ taken from Ref. [11]

A similar result is obtained for the hadronic width of the Z:

$$R_Z = \Gamma(Z \to \text{ hadrons})/\Gamma(Z \to e^+e^-) = N_c(1 + \delta_{\text{EW}}) \sum_f (v_f^2 + a_f^2)/(v_e^2 + a_e^2) , \tag{50}$$

where v_F, a_F are the (axial-)vector couplings for $Z \to \overline{F}F$.

$$\boxed{e^+ e^- \to \text{jets}}$$

At high energies, the two-jet structure from $e^+e^- \to \overline{q}q$ dominates, being the only process at $O(\alpha_s^0)$. At $O(\alpha_s)$ and omitting Z exchange, we have in addition gluon bremsstrahlung off quarks giving rise to a three-jet structure:

$$e^+(q_1)e^-(q_2) \to q(p_1)\overline{q}(p_2)G(p_3) . \tag{51}$$
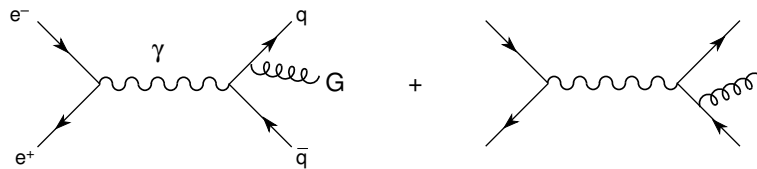
Fig. 6: Leading-order diagrams for three-jet production

The calculation is again identical to QED bremsstrahlung except for a factor (sum over all final states in the rate)

$$\sum_a \text{tr}(t_a^F t_a^F) = T_F \sum_a \delta_{aa} = T_F\, n_{SU(3)} = d_F C_F = 3\, C_F = 4 . \tag{52}$$

With the kinematics specified by

$$s = (q_1 + q_2)^2, \quad (p_i + p_j)^2 = (q_1 + q_2 - p_k)^2 =: s(1 - x_k) \tag{53}$$

$$x_1 + x_2 + x_3 = 2, \qquad \text{CMS} : x_i = 2E_i/\sqrt{s} \,,$$

the double differential cross-section (for massless quarks) is found to be

$$\frac{d^2\sigma}{dx_1 \, dx_2} = \frac{2\alpha_s \sigma_0}{3\pi} \frac{x_1^2 + x_2^2}{(1 - x_1)(1 - x_2)} \qquad \text{with} \qquad \sigma_0 = \frac{4\pi\alpha^2}{s} \sum_f (Q_f/e)^2 \,. \tag{54}$$

The problem with this cross-section is that it diverges for $x_i \to 1$ ($i = 1, 2$). This infrared divergence is due to the singular behaviour of the quark propagator and it happens even for massive quarks:

$$(p_2 + p_3)^2 - m_q^2 = 2p_2 \cdot p_3 = s(1 - x_1) \,. \tag{55}$$

$m_q > 0 :$ $x_1 \to 1$ only possible for $p_3 \to 0$ (soft gluon singularity);

$m_q = 0 :$ $x_1 = 1$ also possible for $p_3 \, \| \, p_2$ (collinear singularity) .

To understand the origin of infrared divergences, we first take the viewpoint of an experimentalist measuring three-jet events where the jets stand for the quarks and the gluon in the final state.

- Depending on the detector resolution, a quark and a soft gluon cannot be distinguished from a single quark. In that case, the event will be counted as a two-jet event.
- Two collinear massless particles can never be resolved: they always stay together.

From the viewpoint of a theorist, we recall that perturbation theory is built on the assumption that particles do not interact when they are sufficiently far apart. This assumption is not really satisfied for massless quanta like photons or gluons that give rise to long-range forces. In other words, an electron (a quark) can never be separated from its cloud of soft photons (gluons).

The practitioner's solution of the infrared problem is well understood:

- One must define criteria to distinguish between (in the present case) two- and three-jet events (jet algorithms).
- Virtual gluon (loop) corrections for the process $e^+e^- \to \bar{q}q$ must be included.

2.2 Higher-order corrections and renormalization

The loop corrections of $O(\alpha_s)$ for $e^+e^- \to \bar{q}q$ are calculated from the Feynman diagrams below. The resulting amplitudes are both infrared and ultraviolet divergent.
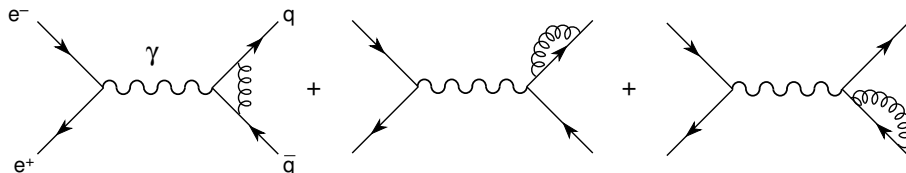
Fig. 7: One-loop diagrams for $e^+e^- \to \bar{q}q$

In contrast to infrared divergences, ultraviolet divergences are due to the high-momentum components of the particles in loops. Common sense tells us that those components can not influence physics at

low energies. If this were the case we would have to give up all hopes of being able to make predictions at currently accessible energies.

The recipe to handle ultraviolet divergences is also well understood. One first has to choose a method to cut off the high-momentum components. There are infinitely many ways to do that, so the question is legitimate whether the final amplitudes will depend on that procedure rendering the result completely arbitrary. The answer is that the cut-off procedure (regularization) must always be accompanied by renormalization. Before choosing a suitable regularization procedure let us therefore try to understand the idea of renormalization, using the most naive regularization method.

To simplify matters as much as possible, we consider the elastic scattering of two particles in massless scalar ϕ^4 theory ($\mathcal{L}_{\text{int}} \sim \lambda \phi^4$):

$$\phi\phi \to \phi\phi \,, \tag{56}$$

with scattering amplitude $A(s,t)$ in terms of the usual Mandelstam variables. We now define what we mean by the physical (renormalized) coupling constant. The definition should be applicable at every order of perturbation theory and it should coincide with the constant λ in the Lagrangian at tree level. A possible definition in scalar ϕ^4 theory is

$$\lambda_r(\mu) := A(s = -t = \mu^2) \tag{57}$$

with an arbitrary renormalization scale μ. At tree level, $A(s,t)$ is momentum independent and with the proper normalization we have indeed $\lambda_r(\mu) = \lambda$.

Beyond tree level, the amplitude has an ultraviolet divergence that we regularize with a simple momentum cut-off Λ here. The relevant diagrams up to one loop are shown in Fig. 8.
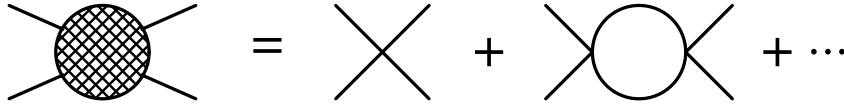
Fig. 8: Scattering amplitude for $\phi\phi \to \phi\phi$ to one-loop order

Setting $s = -t = \mu^2$, one finds (β_0 is a constant)

$$\begin{aligned}
\lambda_r(\mu) = A(s = -t = \mu^2) &= \lambda + \beta_0 \lambda^2 \log \Lambda/\mu \\
&+ \mu-\text{independent terms of } O(\lambda^2) + O(\lambda^3) \,.
\end{aligned} \tag{58}$$

Since $\lambda_r(\mu)$ is finite, being equal to the physical scattering amplitude at some fixed point in phase space, the bare coupling λ diverges as the cut-off $\Lambda \to \infty$. However, the bare coupling is not related to any physical quantity. Therefore, we are free to "sweep the infinities under the rug" as long as this is done in a transparent and controllable way.

To do this, we change the renormalization scale by a small amount $\delta\mu$:

$$\begin{aligned}
\lambda_r(\mu + \delta\mu) - \lambda_r(\mu) &= \beta_0 \lambda^2 \log \left(\frac{\Lambda}{\mu + \delta\mu} \frac{\mu}{\Lambda} \right) + O(\lambda^3) \\
&= \beta_0 \lambda_r^2 \log \frac{\mu}{\mu + \delta\mu} + O(\lambda_r^3) = -\beta_0 \lambda_r^2 \frac{\delta\mu}{\mu} + O[(\delta\mu)^2] + O(\lambda_r^3) \,.
\end{aligned} \tag{59}$$

The bare coupling and the cut-off have disappeared in the last equation. Expanding $\lambda_r(\mu + \delta\mu)$ around μ and letting $\delta\mu \to 0$, we recover the β function of ϕ^4 theory to one-loop order:

$$\mu \frac{d\lambda_r(\mu)}{d\mu} = -\beta_0 \lambda_r^2(\mu) + O(\lambda_r^3) = \beta(\lambda_r(\mu)) \,. \tag{60}$$

Unlike in Yang–Mills theories, $\beta_0 < 0$ so that ϕ^4 theory is ultraviolet unstable like all quantum field theories except non-Abelian gauge theories [5,12]. However, for understanding the essence of renormalization the important observation is that physical quantities do not depend on the bare coupling constant λ nor on the cut-off Λ (for $\Lambda \to \infty$) but only on the renormalized coupling $\lambda_r(\mu)$. For the purpose of comparing theory with experiment at present energies, we shall never notice the stuff that was swept under the rug.

We now turn to the choice of a regularization scheme. Although there are infinitely many possibilities, some choices are clearly better than others. The main criteria are

- The regularization method should respect symmetries of the theory as much as possible. In this respect, the previously employed momentum cut-off is as bad as it gets violating Poincaré symmetry, gauge invariance, etc.
- The scheme should violate only those symmetries that are necessarily violated by quantum effects (anomalies).
- The method should be simple to handle in practice.

From the practitioner's point of view, dimensional regularization is the almost unique choice fulfilling these criteria. Let us demonstrate the method with a simple example, electronic vacuum polarization (setting $m_e = 0$).
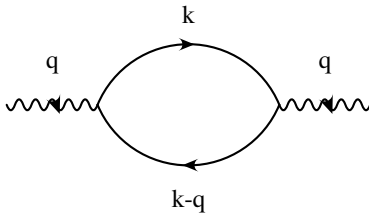
Fig. 9: Vacuum polarization at one loop

Gauge invariance, guaranteed by dimensional regularization, implies the same structure as in the hadronic case:

$$\Pi^{\mu\nu}(q) = \left(-g^{\mu\nu}q^2 + q^\mu q^\nu\right)\Pi(q^2) \quad (61)$$

$$\Pi(q^2) = \frac{8e^2\Gamma(\varepsilon)}{(4\pi)^{2-\varepsilon}}\int_0^1 \frac{dx\,x(1-x)}{[-q^2 x(1-x)]^\varepsilon}$$

$$\Gamma(x) = 1/x - \gamma + O(x), \quad 2\varepsilon = 4 - d\,.$$

Since dimensional regularization works in d dimensions, there is a small problem here: there is no scale for $\log\left(-q^2\right)$ that will appear in the explicit form of $\Pi(q^2)$. To solve the problem, we insert unity (the scale μ and the constant c are completely arbitrary) in the expression and expand the second factor in ε:

$$1 = (c\mu)^{-2\varepsilon}(c\mu)^{2\varepsilon} = (c\mu)^{-2\varepsilon}\left[1 + \varepsilon\log\mu^2 + 2\varepsilon\log c + O(\varepsilon^2)\right]\,. \quad (62)$$

Various schemes on the market differ by the constant c:

$$\text{MS} \qquad\qquad c = 1$$
$$\overline{\text{MS}} \qquad\qquad \log c = (\gamma - \log 4\pi)/2\,.$$

Using the most popular scheme ($\overline{\text{MS}}$), the final result is

$$\Pi(q^2) = \frac{e^2}{12\pi^2}\left\{\frac{(c\mu)^{-2\varepsilon}}{\varepsilon} - \log\left(-q^2/\mu^2\right) + \frac{5}{3}\right\} + O(\varepsilon)$$

$$= \Pi_{\text{div}}^{\overline{\text{MS}}}(\varepsilon,\mu) - \frac{e^2}{12\pi^2}\left\{\log\left(-q^2/\mu^2\right) - \frac{5}{3}\right\}\,. \quad (63)$$

The divergent part $\Pi_{\text{div}}^{\overline{\text{MS}}}(\varepsilon,\mu)$ has been isolated and it will be absorbed by wave function renormalization of the photon field contributing to charge renormalization. We also notice that the coefficients of $1/\varepsilon$ and $-\log\left(-q^2/\mu^2\right)$ are identical: the β function can be extracted from the divergent part. The stuff under the rug is useful after all.

Back to the loop corrections of Fig. 7, we observe that the diagrams give rise to an amplitude proportional to g_s^2, whereas $A(e^+e^- \to \overline{q}qG) \sim g_s$. How can the infrared divergences cancel among amplitudes of different order in g_s? The answer is that they can not cancel on the level of amplitudes because the final states are different. Not the amplitudes but the rates must be added. Interference with the tree amplitude produces an $O(\alpha_s)$ term in $\sigma(e^+e^- \to \overline{q}q)$ that can and will cancel the infrared divergence in $\sigma(e^+e^- \to \overline{q}qG)$. For details of the calculation I refer to the monograph [13], an excellent source for applications of perturbative QCD in general.

The easier part are the loop corrections for $\sigma(e^+e^- \to \overline{q}q)$. With dimensional regularization to regularize the infrared divergences, one finds

$$\sigma_{\overline{q}q}^{\text{interference}} = \sigma_0 C_F \frac{\alpha_s}{4\pi} H(\varepsilon) \left\{ -\frac{4}{\varepsilon^2} - \frac{6}{\varepsilon} - 16 + O(\varepsilon) \right\} , \qquad H(0) = 1 . \qquad (64)$$

The less familiar part is the three-body phase space integration in d dimensions giving rise to

$$\sigma_{\overline{q}qG} = \sigma_0 C_F \frac{\alpha_s}{4\pi} H(\varepsilon) \left\{ \frac{4}{\varepsilon^2} + \frac{6}{\varepsilon} + 19 + O(\varepsilon) \right\} . \qquad (65)$$

The infrared divergences cancel as expected.

Adding the lowest-order cross-section, one obtains finally

$$\sigma(e^+e^- \to \text{hadrons}) = \sigma_0 \left(1 + 3C_F \frac{\alpha_s}{4\pi} + O(\alpha_s^2) \right) = \sigma_0 \left(1 + \frac{\alpha_s}{\pi} + O(\alpha_s^2) \right) , \qquad (66)$$

with σ_0 defined in Eq. (54). The cross-section $\sigma(e^+e^- \to \text{hadrons})$ is nowadays known up to $O(\alpha_s^3)$. Replacing α_s by the renormalization group improved running coupling $\alpha_s(\sqrt{s})$, the general result for $R_{e^+e^-}$ can be written

$$\begin{aligned} R_{e^+e^-}(s) &= N_c \sum_f Q_f^2/e^2 \left\{ 1 + \sum_{n \geq 1} C_n \left(\frac{\alpha_s(\sqrt{s})}{\pi} \right)^n \right\} \qquad (67) \\ &= R_{e^+e^-}^{(0)} \left\{ 1 + C_1 \frac{\alpha_s(\mu)}{\pi} + \left[C_2 - C_1 \frac{\beta_0}{4} \log(s/\mu^2) \right] \left(\frac{\alpha_s(\mu)}{\pi} \right)^2 + \dots \right\} . \end{aligned}$$

The normalization is such that $C_1 = 1$, the coefficients C_2, C_3 being also known.

In principle, $R_{e^+e^-}(s)$ is independent of the arbitrary scale μ by construction. In reality, the unavoidable truncation of the perturbative series introduces a scale dependence. Although there is no unique prescription for the optimal choice of μ, the obvious choice here is $\mu^2 = s$ to avoid large logarithms. If the perturbative expansion is to make sense, we expect higher orders to mitigate the scale dependence. This is nicely demonstrated in Fig. 10 taken from Ref. [13] where the deviation (in per cent) of $R_{e^+e^-}(\sqrt{s} = 33 \text{ GeV})$ from $R_{e^+e^-}^{(0)}$ is plotted as a function of μ. As expected, $\mu^2 = s$ is indeed a very reasonable choice already at $O(\alpha_s)$ (L).

2.3 Measurements of α_s

How should one characterize the coupling strength of QCD? After all, the scale μ is arbitrary and, in addition, $\alpha_s(\mu)$ is in general scheme dependent. For qualitative purposes, one may introduce a scale Λ_{QCD} that is independent of the renormalization scale μ. The drawback is that this quantity is scheme independent only at leading (one-loop) order:

$$\alpha_s(E) = \frac{4\pi}{\beta_0 \log(E^2/\Lambda_{\text{QCD}}^2)} . \qquad (68)$$
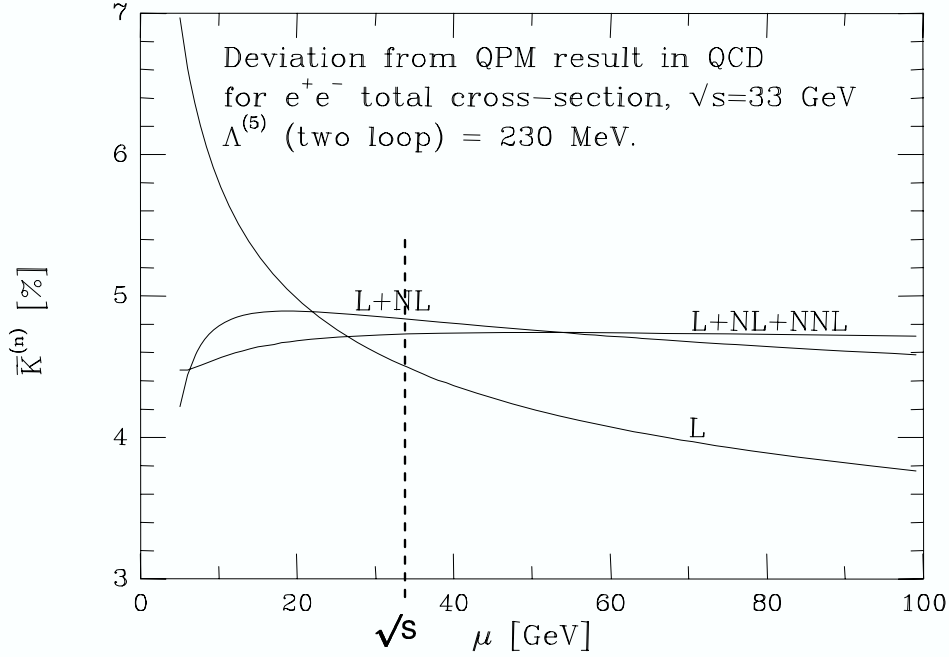
Fig. 10: Improvement of the scale dependence in higher orders of perturbation theory for $R_{e^+e^-}(\sqrt{s} = 33 \text{ GeV})$ (taken from the book of Ellis, Stirling and Webber [13]).

The coefficient β_0 is defined by rewriting the β function for α_s (instead of for g_s as in Section 1.1):

$$\mu \frac{d\alpha_s(\mu)}{d\mu} = 2\beta(\alpha_s) = -\frac{\beta_0}{2\pi}\alpha_s^2 - \frac{\beta_1}{4\pi^2}\alpha_s^3 + \dots \tag{69}$$

The β function is known up to four loops (coefficient β_3) but only the first two coefficients

$$\beta_0 = 11 - 2N_F/3 \,, \qquad\qquad \beta_1 = 51 - 19N_F/3 \tag{70}$$

are scheme and gauge independent.

Since the scheme dependence is unavoidable, the coupling strength is nowadays usually given in the form of $\alpha_s^{\overline{\text{MS}}}(M_Z)$. Of course, this fixes $\alpha_s^{\overline{\text{MS}}}(\mu)$ at any scale via the integral

$$\log\left(\mu_2^2/\mu_1^2\right) = \int_{\alpha_s(\mu_1)}^{\alpha_s(\mu_2)} \frac{dx}{\beta(x)} \,. \tag{71}$$

To get a first rough estimate of α_s, consider the leading-order prediction

$$R_{e^+e^-}(M_Z) = R_{e^+e^-}^{(0)}(M_Z)\left(1 + \frac{\alpha_s(M_Z)}{\pi}\right) \,. \tag{72}$$

Comparing the combined LEP result [14] $R_{e^+e^-}(M_Z) = 20.767 \pm 0.025$ with the tree-level prediction $R_{e^+e^-}^{(0)}(M_Z) = 19.984$, one obtains

$$\alpha_s(M_Z) = 0.123 \pm 0.004 \,, \tag{73}$$

close to the full three-loop result and not bad at all for a first estimate. A compilation of results can be found in the Review of Particle Physics [14]. In Figs. 11 and 12, the most recent data compiled by Bethke [15] are shown.

The first impression is the remarkable agreement among experiments and with theory. However, for determining the best value of $\alpha_s(M_Z)$, the following two problems must be kept in mind:
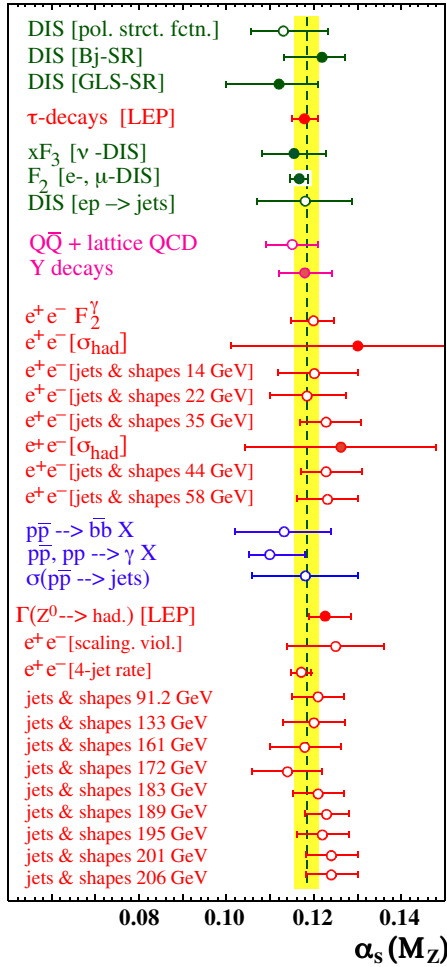
Fig. 11: Compilation of data for the extraction of $\alpha_s^{\overline{\mathrm{MS}}}(M_Z)$ by Bethke [15]
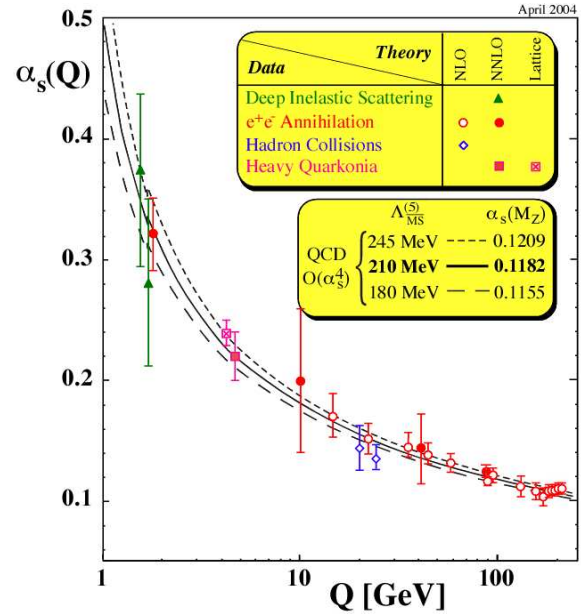
Fig. 12: Energy dependence of the running coupling constant $\alpha_s^{\overline{\mathrm{MS}}}(Q)$ [15]

– Different observables are known with different theoretical accuracy: next-to-leading order (NLO) vs. next-to-next-to-leading order (NNLO). Not only the scale dependence but also different scheme dependences must be taken into account.

– Theoretical errors are not normally distributed.

Using only NNLO results, Bethke found [15]

$$\alpha_s(M_Z) = 0.1182 \pm 0.0027 \,, \tag{74}$$

very similar to the PDG average [14] (using a different procedure)

$$\alpha_s(M_Z) = 0.1187 \pm 0.0020 \,. \tag{75}$$

All values in this paragraph refer to $\alpha_s^{\overline{\mathrm{MS}}}(M_Z)$.

2.4 Hadronic τ decays

A remarkably precise value for $\alpha_s^{\overline{\mathrm{MS}}}(M_Z)$ comes from hadronic τ decays. At first sight, this is quite surprising because at the natural scale $\mu = m_\tau$ one has approximately $\alpha_s(m_\tau) \simeq 0.35$. Can one expect reasonable convergence of the perturbative series for such a large coupling and how big are the nonperturbative corrections?

The first systematic investigation of $R_\tau = \Gamma(\tau^- \to \nu_\tau + \text{hadrons})/\Gamma(\tau^- \to \nu_\tau e^- \overline{\nu_e})$ was performed by Braaten, Narison and Pich [16]. The analysis is similar to the one for $R_{e^+e^-}$, with obvious modifications: the electromagnetic current (coupling to e^+e^-) must be replaced by the charged weak current (coupling to $\tau\nu_\tau$).

We start again with the two-point function (of weak currents $L^\mu = \overline{u}\gamma^\mu(1-\gamma_5)d_\theta$):

$$
\begin{aligned}
\Pi_L^{\mu\nu}(q) &= i \int d^4x\, e^{iq \cdot x} \langle 0|\, TL^\mu(x)L^\nu(0)^\dagger|0\rangle \\
&= \left(-g^{\mu\nu}q^2 + q^\mu q^\nu\right) \Pi_L^{(1)}(q^2) + q^\mu q^\nu \Pi_L^{(0)}(q^2) \,,
\end{aligned}
\tag{76}
$$

with d_θ the Cabibbo-rotated d-quark field. One major difference to the electromagnetic case is that one has to integrate over the neutrino energy or, equivalently, over the hadronic invariant mass s:

$$
R_\tau = 12\pi \int_0^{m_\tau^2} \frac{ds}{m_\tau^2} \left(1 - \frac{s}{m_\tau^2}\right)^2 \left\{ \left(1 + 2\frac{s}{m_\tau^2}\right) \text{Im}\Pi_L^{(1)}(s) + \text{Im}\Pi_L^{(0)}(s) \right\} \,.
\tag{77}
$$

The problem is that the integration extends all the way down to $s = 0$ (for $m_u = m_d = 0$) where perturbation theory is certainly not applicable.

However, QFT provides information about the analytic structure of two-point functions that can be used in a standard manner to circumvent the problem. The invariant functions $\Pi_L^{(0,1)}(s)$ are known to be analytic in the complex s-plane with a cut on the positive real axis. Therefore, Cauchy's theorem tells us that the contour integral in Fig. 13 vanishes.
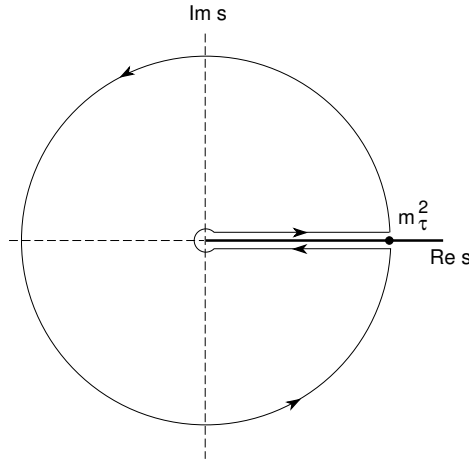
Fig. 13: Contour in the complex s-plane for the two-point functions $\Pi_L^{(0,1)}(s)$

One can now trade the integral along the cut of

$$
\text{Im}\Pi_L^{(0,1)}(s) = \frac{1}{2i}\left[\Pi_L^{(0,1)}(s+i\varepsilon) - \Pi_L^{(0,1)}(s-i\varepsilon)\right]
\tag{78}
$$

for an integral along the circle $|s| = m_\tau^2$ in the complex s-plane. It turns out that the nonperturbative corrections are now manageable, being suppressed as $(\Lambda_{\text{QCD}}/m_\tau)^6$. Very helpful in this respect is the factor $\left(1 - \frac{s}{m_\tau^2}\right)^2$ in the integrand that suppresses potentially big contributions near the endpoint of the cut. The final result can be written in the form [16]

$$
R_\tau = 3\left(|V_{ud}|^2 + |V_{us}|^2\right) S_{\text{EW}} \left\{1 + \delta'_{\text{EW}} + \delta_{\text{pert}} + \delta_{\text{nonpert}}\right\}
\tag{79}
$$

with leading and nonleading electroweak corrections $S_{\mathrm{EW}} = 1.0194$ and $\delta'_{\mathrm{EW}} = 0.0010$, respectively. The perturbative QCD corrections of interest for the extraction of α_s are contained in

$$
\begin{aligned}
\delta_{\mathrm{pert}} &= \frac{\alpha_s(m_\tau)}{\pi} + \left(C_2 + \frac{19}{48}\beta_0\right)\left(\frac{\alpha_s(m_\tau)}{\pi}\right)^2 + \ldots \\
&= \frac{\alpha_s(m_\tau)}{\pi} + 5.2\left(\frac{\alpha_s(m_\tau)}{\pi}\right)^2 + 26.4\left(\frac{\alpha_s(m_\tau)}{\pi}\right)^3 + O(\alpha_s(m_\tau)^4) \, .
\end{aligned}
\tag{80}
$$

Finally, the best estimates of nonperturbative contributions, using QCD sum rules and experimental input, yield

$$
\delta_{\mathrm{nonpert}} = -0.014 \pm 0.005 \, .
\tag{81}
$$

From the PDG fit for R_τ one then obtains $\alpha_s(m_\tau) = 0.35 \pm 0.03$. More interestingly, running this value down to M_Z with the help of the four-loop β function, one finds

$$
\alpha_s(M_Z) = 0.121 \pm 0.0007(\mathrm{exp}) \pm 0.003(\mathrm{th}),
\tag{82}
$$

not only compatible but in fact very much competitive with other high-precision determinations.

2.5 Deep inelastic scattering

From the conception of QCD till today, deep inelastic scattering of leptons on hadrons has had an enormous impact on the field. It is also a classic example for the factorization between long- and short-distance contributions.

Let us start with the kinematics of (in)elastic electron–proton scattering $e^-(k) + p(p) \rightarrow e^-(k') + X(p_X)$ shown in Fig. 14. In the case of elastic scattering ($X = p$), we have

$$
W^2 = m^2 \, , \qquad Q^2 = 2m\nu \, , \qquad x = 1
\tag{83}
$$

and the usual two variables are $s = (p + k)^2, Q^2$ with the differential cross-section $d\sigma(s, Q^2)/dQ^2$.

For the inclusive scattering, there is a third independent variable: a convenient choice is s, x, y with $0 \leq x \leq 1, \ 0 \leq y \leq 1$. In general, we distinguish different types of deep inelastic scattering:

type	exchange
neutral current (NC) DIS	$\gamma, \ Z, \ \gamma Z$-interference
charged current (CC) DIS	$W^\pm$

For these lectures I restrict the discussion to photon exchange and to unpolarized (spin-averaged) DIS.

The matrix element for the diagram in Fig. 14 has the structure $e\, l(\mathrm{epton})_\mu \, \dfrac{g^{\mu\nu}}{Q^2}\, e\, h(\mathrm{adron})_\nu$, with leptonic and hadronic current matrix elements $l(\mathrm{epton})_\mu$ and $h(\mathrm{adron})_\nu$, respectively. The resulting double differential cross-section is of the form

$$
\begin{aligned}
\frac{d^2\sigma}{dx\,dy} &= x(s - m^2)\frac{d^2\sigma}{dx\,dQ^2} = \frac{2\pi y\alpha^2}{Q^4}L_{\mu\nu}\, H^{\mu\nu} \\
L_{\mu\nu} &= 2(k_\mu k'_\nu + k'_\mu k_\nu - k\cdot k'\, g_{\mu\nu}) \\
H^{\mu\nu}(p, q) &= \frac{1}{4\pi}\int d^4z\, e^{iq\cdot z}\langle p, s|\left[J_{\mathrm{elm}}^\mu(z), J_{\mathrm{elm}}^\nu(0)\right]|p, s\rangle \, .
\end{aligned}
\tag{84}
$$
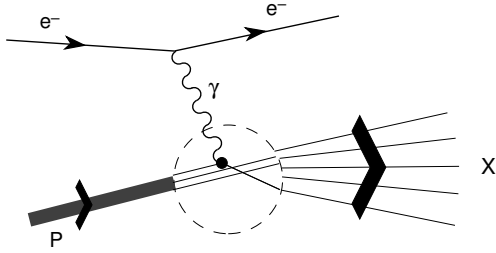
$$q = k - k', \quad Q^2 = -q^2 > 0, \quad p^2 = m^2$$

$$\nu = p \cdot q / m = E - E' \quad \text{(target rest frame)}$$

$$x = \frac{Q^2}{2m\nu}, \quad y = \frac{p \cdot q}{p \cdot k} = 1 - E'/E$$

$$W^2 = p_X^2 = (p + q)^2 = m^2 + 2m\nu - Q^2 \geq m^2 \; .$$

Fig. 14: Deep inelastic scattering

One now performs a Lorentz decomposition of the hadronic tensor $H^{\mu\nu}$ and contracts it with the leptonic tensor $L_{\mu\nu}$ (setting $m_e = 0$). The differential cross-section then depends on two invariant structure functions F_1, F_2 (in the photon case), which are themselves functions of the scalars $p.q$, q^2 or ν, Q^2 or x, Q^2:

$$\frac{d^2\sigma}{dxdy} = \frac{Q^2}{y} \frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{xyQ^2} \left\{ \left(1 - y - \frac{x^2 y^2 m^2}{Q^2}\right) F_2(x, Q^2) + y^2 x F_1(x, Q^2) \right\} \; . \tag{85}$$

Deep inelastic scattering corresponds to $Q^2 \gg m^2$ and $W^2 \gg m^2$. While the cross-section shows a rather complicated behaviour at low and intermediate momentum transfer, the structure functions exhibit an originally unexpected simple behaviour in the so-called Bjorken limit

$$Q^2 \gg m^2, \quad \nu \gg m \quad \text{with} \quad x = \frac{Q^2}{2m\nu} \quad \text{fixed} \; . \tag{86}$$

As shown for F_2 in Fig. 15, in the Bjorken limit the structure functions seem to depend on the variable x only:

$$F_i(x, Q^2) \longrightarrow F_i(x) \; . \tag{87}$$

This scaling behaviour suggested that the photon scatters off point-like constituents (no scale) giving rise to the quark parton model (QPM).
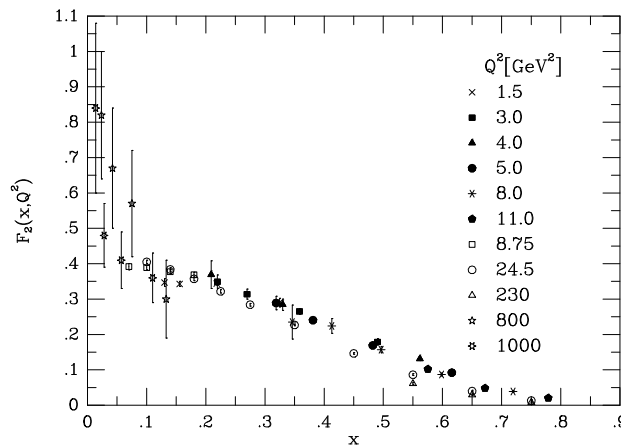
Fig. 15: Evidence for Bjorken scaling taken from Ref. [13]

QPM in the Breit frame

The characteristics of the QPM can best be visualized in the so-called Breit frame where the proton and the virtual photon collide head-on.

The nucleon is pictured as a bunch of partons with negligible transverse momenta. Each parton

Breit frame

$$p \simeq (P, P, 0, 0) \text{ with } P \gg m$$

$$q = (0, -\sqrt{Q^2}, 0, 0)$$
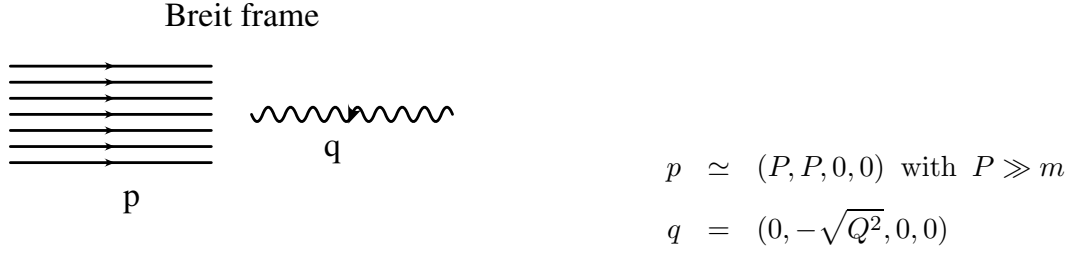
Fig. 16: DIS in the Breit frame

carries a fractional momentum ξp. Since the scattered quark is massless (compared to P), we have

$$(q + \xi p)^2 \simeq -Q^2 + 2\xi p.q = 0 \tag{88}$$

and therefore

$$\xi = x , \qquad P = \frac{\sqrt{Q^2}}{2x} , \qquad q + xp = (xP, -\sqrt{Q^2}/2, 0, 0) . \tag{89}$$

The struck parton scatters with momentum $q + xp$ backwards, i.e., in the direction of the virtual photon, justifying a major assumption of the QPM: the virtual photon scatters incoherently on the partons.

The fundamental process of the QPM is elastic electron–quark scattering

$$e^-(k) + q(\xi p) \rightarrow e^-(k') + q(\xi p + q) . \tag{90}$$

Since there are now only two independent variables, the double differential cross-section in Eq. (85) contains a δ-function setting x equal to ξ:

$$\frac{d^2\sigma_{(q)}}{dxdy} = \frac{4\pi\alpha^2}{yQ^2} \left[1 + (1-y)^2 \right] \frac{Q_q^2}{2} \delta(x - \xi) . \tag{91}$$

In the notation of Eq. (85),

$$F_{2(q)} = xQ_q^2\delta(x - \xi) = 2xF_{1(q)} . \tag{92}$$

The incoherent sum of partonic cross-sections amounts to an integral over quark distribution functions $q(\xi), \overline{q}(\xi)$:

$$F_2(x) = \sum_{q,\overline{q}} \int_0^1 d\xi \, q(\xi) \, x \, Q_q^2 \, \delta(x - \xi) = \sum_{q,\overline{q}} Q_q^2 \, x \, q(x) , \tag{93}$$

implying the Callan–Gross relation [17]

$$F_2(x) = 2xF_1(x) \tag{94}$$

that is due to the spin-1/2 nature of quarks. The so-called longitudinal structure function $F_L = F_2 - 2xF_1$ therefore vanishes in the QPM.

It was already known at the beginning of the seventies, before the advent of QCD, that exact scaling in the sense of the QPM was incompatible with a nontrivial QFT. QCD must therefore account for the systematic deviation from scaling that is clearly seen in the data (e.g., in Fig. 17): with increasing Q^2, the structure function F_2 increases (decreases) at small (large) x. Qualitatively, scaling violation is due to the radiation of (hard) gluons generating transverse momenta for the quarks. More gluons are radiated off when Q^2 increases, leading to logarithmic scaling violations in the structure functions and to scale-dependent parton distribution functions (pdf) $q_i(x, \mu^2)$, $g(x, \mu^2)$ as we shall now discuss in more detail.
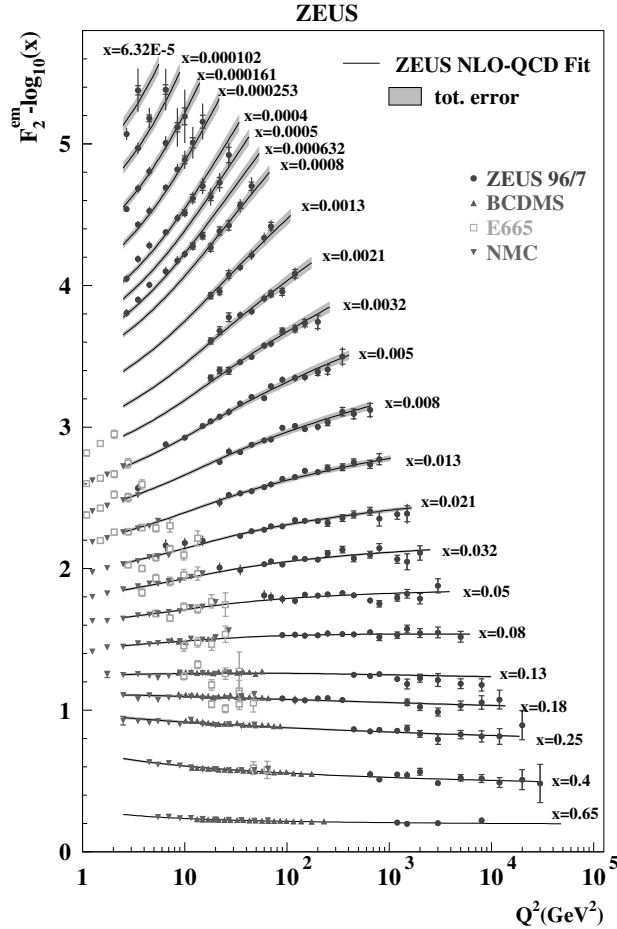
Fig. 17: Experimental evidence for the violation of Bjorken scaling taken from Ref. [18]

DIS in QCD

In QCD at leading order in g_s, the same diagrams need to be considered as the ones in Figs. 6, 7 relevant for $e^+e^- \to$ hadrons except for crossing.

Previously, the sum of real and virtual gluon emission was infrared finite because a sum over all final-state quarks and gluons was involved. In DIS the situation is different because the initial state contains a quark. Since different incoming quark momenta are in principle distinguishable, gluons collinear with the incoming quark generate in fact an infrared divergence. How to get rid of this divergence will be discussed later but for now we regulate the infrared divergence with a cut-off on the transverse quark momentum $k_\perp^2 \geq \kappa^2$. Adding the contribution to $F_{2(q)}$ from real gluon emission ($\xi = 1$ for simplicity) one finds

$$F_{2(q)}(x, Q^2) = Q_q^2\, x \left[\delta(1-x) + \frac{\alpha_s}{2\pi} \left(P_{qq}(x) \log \frac{Q^2}{\kappa^2} + C_q(x) \right) \right] . \qquad (95)$$

Introduction of the cut-off κ has produced a logarithmic dependence on Q^2. This dependence is governed by the so-called quark–quark splitting function $P_{qq}(x)$. This function is universal, in contrast to the non-logarithmic coefficient function $C_q(x)$ that is scheme dependent.

The origin of the scheme-independent $\log Q^2$ can be understood as follows. The struck quark acquires a transverse momentum $k_\perp$ with probability $\alpha_s \dfrac{dk_\perp^2}{k_\perp^2}$. Since $k_\perp^2$ cannot be bigger than Q^2, integrating over all $k_\perp$ produces the term $\alpha_s \log Q^2/\kappa^2$. Virtual gluon contributions (diagrams in Fig. 7)

must still be added and they are ultraviolet finite as before. Since this is a contribution from elastic scattering it must be proportional to $\delta(1 - x)$. Altogether, the quark–quark splitting function at leading order is

$$P_{qq}(x) = \frac{4}{3}\left(\frac{1 + x^2}{[1 - x]_+}\right) + 2\,\delta(1 - x)\,. \tag{96}$$

Also the first part of $P_{qq}(x)$ is actually a distribution. The distribution $[F(x)]_+$ is defined in such a way that for every sufficiently regular (test) function $f(x)$ one has

$$\int_0^1 dx\, f(x)\,[F(x)]_+ = \int_0^1 dx\,(f(x) - f(1))\,F(x)\,. \tag{97}$$

It is then straightforward to show that $P_{qq}(x)$ in Eq. (96) can also be written as

$$P_{qq}(x) = \frac{4}{3}\left[\frac{1 + x^2}{(1 - x)}\right]_+\,. \tag{98}$$

The problem remains how to interpret (or rather get rid of) the infrared cut-off κ. Up to now, we have only considered the quark structure function $F_{2(q)}(x, Q^2)$. To get $F_2(x, Q^2)$ for the nucleon, we convolute $F_{2(q)}(\frac{x}{\xi}, Q^2)$ with a (bare) pdf $q_0(\xi)$:

$$F_2(x, Q^2) = x\sum_{q,\overline{q}} Q_q^2\left[q_0(x) + \frac{\alpha_s}{2\pi}\int_x^1 \frac{dy}{y} q_0(y)\left\{P_{qq}(x/y)\log\frac{Q^2}{\kappa^2} + C_q(x/y)\right\}\right]\,. \tag{99}$$

One now absorbs the collinear singularity $\sim \log\kappa^2$ into $q_0(x)$ at a factorization scale μ to define a renormalized pdf $q(x, \mu^2)$:

$$q(x, \mu^2) = q_0(x) + \frac{\alpha_s}{2\pi}\int_x^1 \frac{dy}{y} q_0(y)\left\{P_{qq}(x/y)\log\frac{\mu^2}{\kappa^2} + C_q'(x/y)\right\}\,. \tag{100}$$

The interpretation of $q(x, \mu^2)$ is straightforward: the soft part $k_\perp^2 \leq \mu^2$ is now included in the pdf. Since the scale μ is arbitrary, the renormalized pdf is necessarily scale dependent, in complete analogy with the renormalization of ultraviolet divergences. As a small aside, we note that the coefficient function C_q' need not be the same as C_q in Eq. (95), both being scheme dependent. The final form for the nucleon structure function F_2 in the $\overline{\mathrm{MS}}$ scheme (except for a contribution from the gluon pdf) is then

$$
\begin{aligned}
F_2(x, Q^2) &= x\sum_{q,\overline{q}} Q_q^2\int_x^1 \frac{dy}{y} q(y, \mu^2)\left[\delta(1 - x/y) + \frac{\alpha_s}{2\pi}\left\{P_{qq}(x/y)\log\frac{Q^2}{\mu^2} + C_q^{\overline{\mathrm{MS}}}(x/y)\right\}\right] \\
&= x\sum_{q,\overline{q}} Q_q^2\int_x^1 \frac{dy}{y} q(y, Q^2)\left[\delta(1 - x/y) + \frac{\alpha_s}{2\pi}C_q^{\overline{\mathrm{MS}}}(x/y)\right]\,.
\end{aligned}
\tag{101}
$$

This factorization formula can be proven to all orders in α_s, separating the calculable hard part from the soft part contained in the scale-dependent pdfs. The pdfs $q(x, \mu^2)$, $\overline{q}(x, \mu^2)$, $g(x, \mu^2)$ describe the composition of nucleons and are, of course, not calculable in perturbation theory. They can be extracted from experimental data with the help of appropriate parametrizations (cf., for example, Ref. [13]) but the question remains what the factorization result (101) is actually good for.

The answer is that even though the functional dependence of the pdfs can not be calculated their scale dependence is calculable in QCD perturbation theory. The derivation of the so-called DGLAP evolution equations [19] is very similar to the derivation of the β function in Section 1.1, starting from the observation that the measurable structure function must be scale independent:

$$\mu^2\frac{dF_2(x, Q^2)}{d\mu^2} = 0 \qquad \longrightarrow \qquad \mu^2\frac{dq(x, \mu^2)}{d\mu^2} = \frac{\alpha_s(\mu)}{2\pi}\int_x^1 \frac{dy}{y} P_{qq}(x/y, \alpha_s(\mu))q(y, \mu^2) \tag{102}$$

$$P_{qq}(x, \alpha_s(\mu)) = P_{qq}^{(0)}(x) + \frac{\alpha_s(\mu)}{2\pi} P_{qq}^{(1)}(x) + \ldots \tag{103}$$

With $P_{qq}^{(0)}(x)$ given by Eq. (98), the evolution equation at leading order takes the explicit form

$$\mu^2 \frac{dq(x, \mu^2)}{d\mu^2} = \frac{2\alpha_s(\mu)}{3\pi} \int_x^1 \frac{dz}{z} q(x/z, \mu^2) \frac{1+z^2}{1-z} - \frac{2\alpha_s(\mu)}{3\pi} q(x, \mu^2) \int_0^1 dz \frac{1+z^2}{1-z} . \tag{104}$$

Owing to soft gluons, both terms on the right-hand side are divergent: the first term with positive sign is due to quarks with momentum fraction larger than x radiating off gluons whereas the second term leads to a decrease from quarks with given x that radiate gluons. The overall result is finite.

At $O(\alpha_s)$ also the gluon pdf enters via $\gamma^* + g \to q + \overline{q}$. At any order, the DGLAP equations are in general $(2N_F + 1)$-dimensional matrix equations for $q_i(x, \mu^2)$, $\overline{q_i}(x, \mu^2)$ $(i = 1, \ldots, N_F)$, $g(x, \mu^2)$ with splitting functions $P_{qq}(x)$, $P_{qg}(x)$, $P_{gq}(x)$ and $P_{gg}(x)$. The analytic calculation of these splitting functions to next-to-next-to-leading order (three loops) has just been completed [20] allowing for precise tests of scaling violations. Finally, we note that the longitudinal structure function $F_L(x, Q^2)$ that was zero in the QPM is generated in QCD already at $O(\alpha_s)$.

For more applications of perturbative QCD I refer once again to the book of Ellis, Stirling and Webber [13]: jets in e^+e^- and hadroproduction, vector boson production (Drell–Yan), heavy quark production and decays, Higgs production at the LHC, etc.

3 Heavy and light quarks

3.1 Effective field theories

Unlike QED, QCD is valid down to shortest distances because of asymptotic freedom. However, at long distances where quarks and gluons are practically invisible, perturbative QCD is not applicable and a nonperturbative approach is needed. Many models can be found in the literature that are more or less inspired by QCD. Qualitative insights into the structure of the strong interactions have been found from model studies but quantitative predictions require methods that can be related directly to QCD. There are essentially only two approaches that satisfy this criterion.

- Lattice QCD has already scored impressive results and may in the long run be the most predictive method. At present, the range of applicability is still limited.
- Effective field theories (EFTs) are the quantum field theoretical formulation of the "quantum ladder": the relevant degrees of freedom depend on the typical energy of the problem. EFTs become practical tools for phenomenology when the characteristic energy scales are well separated.

Let a given step of the quantum ladder be characterized by an energy scale Λ. The region $E > \Lambda$ is the short-distance region where the fundamental theory is applicable. At long distances ($E < \Lambda$), on the other hand, an effective QFT can and sometimes must be used. By definition, the notions "fundamental" and "effective" only make sense for a given energy scale Λ. As we probe deeper into the physics at short distances, today's fundamental theory will become an effective description of an even more fundamental underlying theory.

To understand the different effective field theories that are being used in particle physics, it is useful to classify them according to the structure of the transition between the fundamental and the effective level.

- Complete decoupling
  The heavy degrees of freedom (heavy with respect to Λ) are integrated out, i.e., they disappear from the spectrum of states that can be produced with energies $< \Lambda$. Correspondingly, the effective

Lagrangian contains only light fields (once again, light stands for masses $< \Lambda$) and may be written symbolically as

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{d \leq 4} + \sum_{d > 4} \frac{1}{\Lambda^{d-4}} \sum_{i_d} g_{i_d} O_{i_d} . \tag{105}$$

The first part $\mathcal{L}_{d \leq 4}$ contains all the renormalizable couplings for the given set of fields. The best-known example for such a Lagrangian is the Standard Model itself. The second part of the Lagrangian (105) contains the nonrenormalizable couplings having operator dimension $d > 4$. The best-known example here is the Fermi theory of weak interactions with $d = 6$ and $\Lambda = M_W$. For the Standard Model, on the other hand, we do not know the scale where new physics will appear. Present experimental evidence implies $\Lambda > 100$ GeV but there are good reasons to expect new physics around $\Lambda \sim 1$ TeV.

– Partial decoupling

In this case, heavy fields do not disappear completely in the EFT. Via so-called field redefinitions, only the high-momentum modes are integrated out. The main application of this scenario in particle physics is for heavy quark physics.

– Spontaneous symmetry breaking (SSB)

In the previous two classes, the transition from the fundamental to the effective level was smooth. Some of the fields or at least their high-energy modes just drop out and the effective description involves the remaining fields only. In the present case, the transition is more dramatic and involves a phase transition: SSB generates new degrees of freedom, the (pseudo-)Goldstone bosons associated with spontaneously broken symmetries (to be discussed in more detail in Section 3.4). The prefix pseudo accounts for the frequent case where the symmetry in question is not only spontaneously but also explicitly broken. Goldstone bosons in the strict sense are massless and the associated SSB relates processes with different numbers of Goldstone bosons. As a consequence, the distinction in the Lagrangian (105) between renormalizable ($d \leq 4$) and nonrenormalizable ($d > 4$) terms becomes meaningless. Therefore, EFTs in this category are generically nonrenormalizable. An important but maybe too simple exception is the Higgs model for electroweak SSB. The generic EFT Lagrangian is organized in the number of derivatives of Goldstone fields and in the number of terms with explicit symmetry breaking. An important concept is universality: it turns out that EFTs describing different physical situations have very similar structure. In QCD, the symmetry in question is chiral symmetry that becomes exact in the limit of massless quarks. In the real world, SSB of chiral symmetry generates pseudo-Goldstone bosons that are identified with the pseudoscalar mesons π, K, η.

We are used to deriving quantitative predictions from renormalizable QFTs in the framework of perturbation theory but how should we treat nonrenormalizable EFTs? The clue to the answer is Goldstone's theorem [21] that makes two crucial predictions. The first prediction is well known: SSB implies the existence of massless Goldstone bosons. The second consequence of Goldstone's theorem is not that well known but very important as well: Goldstone bosons decouple when their energies tend to zero. In other words, independently of the strength of the underlying interaction (the strong interaction in our case), Goldstone bosons interact only weakly at low energies. This important feature allows for a systematic expansion of strong amplitudes even in the confinement regime, which is precisely the low-energy regime of QCD.

However, in contrast to the decoupling case (e.g., in heavy quark physics), the coupling constants of the low-energy EFT cannot be obtained by perturbative matching with the underlying theory of QCD. Other methods have to be used to get access to the low-energy couplings.

3.2 Heavy quarks

Quarks can not be put on a balance, or more realistically, their energies and momenta can not be measured directly. How do we then determine their masses? Two methods have been used.

– The first approach ignores confinement and calculates the pole of the quark propagator just as we determine, at least in theory, the mass of the electron. This looks rather artificial because the full quark propagator should have no pole because of confinement. Going ahead nevertheless, one expects those pole masses to be very much affected by nonperturbative infrared effects. In practice, this method is only used for the top quark with [14]

$$m_t = 174.3 \pm 5.1 \text{ GeV} . \tag{106}$$

– Quark masses are parameters of the QCD Lagrangian just like the strong coupling constant g_s. One therefore studies the influence of these parameters on measurable quantities and extracts specific values for the masses by comparison with experimental measurements. As for the strong coupling constant, renormalization is crucial and introduces a scale dependence also for quark masses. The scale dependence is governed by a differential equation very similar to the renormalization group equations (9) or (69) for the strong coupling:

$$\mu \frac{dm_q(\mu)}{d\mu} = -\gamma(\alpha_s(\mu)) m_q(\mu) \tag{107}$$

where the anomalous dimension γ is nowadays known up to four-loop accuracy:

$$\gamma(\alpha_s) = \sum_{n=1}^{4} \gamma_n \left(\frac{\alpha_s}{\pi}\right)^n . \tag{108}$$

The solution of this renormalization group equation for $m_q(\mu)$ is flavour independent (in the $\overline{\text{MS}}$ scheme):

$$m_q(\mu_2) = m_q(\mu_1) \exp\left\{ -\int_{\alpha_s(\mu_1)}^{\alpha_s(\mu_2)} dx \frac{\gamma(x)}{2\beta(x)} \right\} . \tag{109}$$

Since $\gamma(\alpha_s)$ is positive, quark masses decrease with increasing scale μ, e.g.,

$$\frac{m_q(1 \text{ GeV})}{m_q(M_Z)} = 2.30 \pm 0.05 . \tag{110}$$

Different methods have been used to determine the quark masses: H(eavy) Q(uark) E(ffective) T(heory) (see below), QCD sum rules (Section 3.3), lattice QCD, etc. The current state is summarized in Fig. 18 taken from Ref. [22]. All values correspond to the $\overline{\text{MS}}$ scheme: light quarks are given at the scale 2 GeV whereas the heavy quarks m_c, m_b are listed as $m_q(m_q)$.

Heavy quark effective theory (HQET)

Why should one use an EFT for b-quark physics? After all, QCD is still accessible in perturbation theory for $\mu = m_b$. The main arguments in favour of HQET are the following.

– Although the hard effects are calculable in QCD perturbation theory, there are inevitably incalculable soft effects because hadrons rather than quarks and gluons appear in the final states of B decays. The necessary separation between perturbative and nonperturbative contributions is much easier to achieve in an EFT description. The keyword is the same as in deep inelastic scattering: factorization.
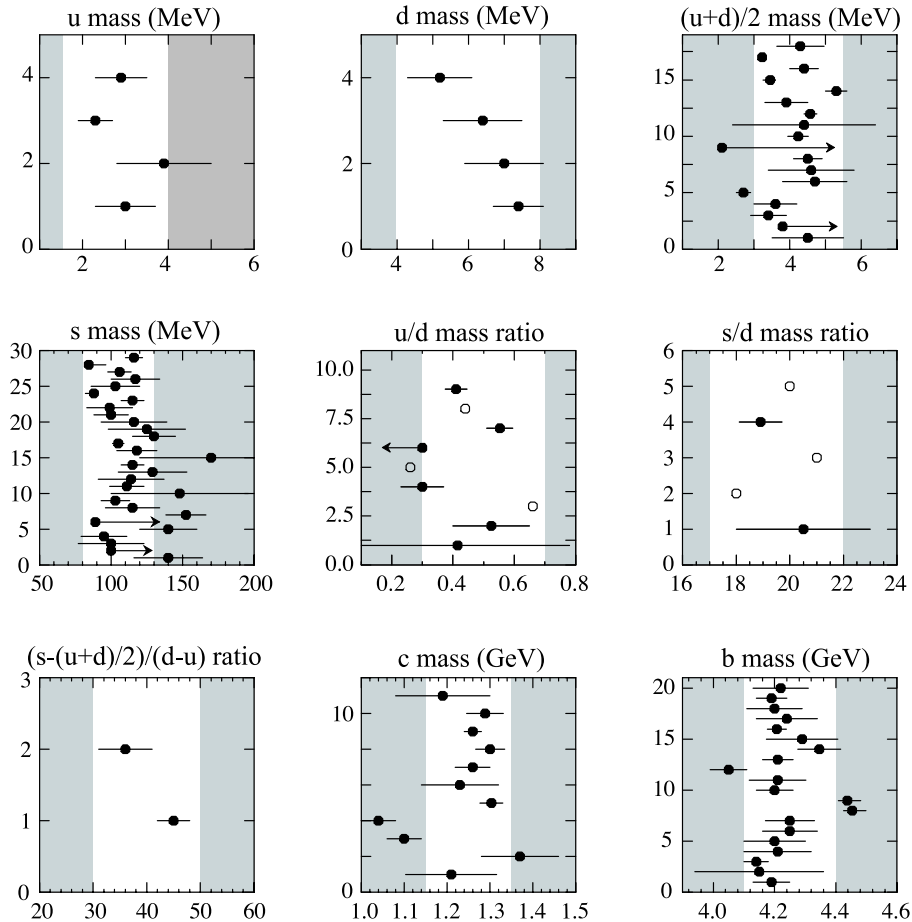
Fig. 18: The values of quark masses taken from the 2004 Review of Particle Physics [22]. The most recent data points are at the top of each plot.

- Approximate symmetries that are hidden in full QCD become manifest in an expansion in $1/m_Q$.
- Explicit calculations simplify in general, in particular the resummation of large logs via renormalization group equations.

The physics behind HQET can be understood by an analogy with atomic physics. The spectrum of the hydrogen atom is to a good approximation insensitive to the proton mass. In fact, the atomic spectra of hydrogen and deuterium are practically identical. The implementation of this analogy is most straightforward for hadrons with a single heavy quark (b or c). In the hadron rest frame the heavy quark 'just sits there' acting as a colour source just as the proton acts as a Coulomb source in the hydrogen atom.

We decompose the momentum of the heavy quark as

$$p^\mu = m_Q v^\mu + k^\mu \,, \tag{111}$$

where v is the hadron velocity normalized to $v^2 = 1$ ($v = (1,0,0,0)$ in the hadron rest frame). The residual quark momentum k is then expected to be of $O(\Lambda_{\rm QCD})$ only. Starting from the QCD Lagrangian for a heavy quark Q,

$$\mathcal{L}_Q = \overline{Q}(i\slashed{D} - m_Q)Q \,, \tag{112}$$

we decompose the quark field $Q(x)$ into two fields $h_v(x)$ and $H_v(x)$ by using energy projectors $P_v^\pm = (1 \pm \slashed{v})/2$ and applying (shift) factors $e^{im_Q v \cdot x}$:

$$Q(x) = e^{-im_Q v \cdot x}\left(h_v(x) + H_v(x)\right) \tag{113}$$

$$h_v(x) = e^{im_Q v \cdot x} P_v^+ Q(x) \,, \qquad H_v(x) = e^{im_Q v \cdot x} P_v^- Q(x) \,.$$

It is easy to check that in the hadron rest frame the fields h_v and H_v are just the upper (big) and lower (small) components of the spinor field Q, respectively. Expressing $\mathcal{L}_Q$ in terms of h_v and H_v, one finds

$$\begin{aligned} \mathcal{L}_Q &= \overline{Q}(i\not{D} - m_Q)Q \\ &= \overline{h_v}\, iv \cdot D\, h_v - \overline{H_v}(iv \cdot D + 2m_Q)H_v + \text{mixed terms} \,. \end{aligned} \tag{114}$$

For the purpose of illustration, we use the field equation $(i\not{D} - m_Q)Q = 0$ to eliminate H_v in this Lagrangian:

$$\mathcal{L}_Q = \overline{h_v}\, iv \cdot D\, h_v + \overline{h_v}\, i\not{D}_\perp \frac{1}{iv \cdot D + 2m_Q - i\epsilon} i\not{D}_\perp h_v \quad \text{with} \quad D_\perp^\mu = (g^{\mu\nu} - v^\mu v^\nu)D_\nu \,. \tag{115}$$

The heavy quark mass m_Q has disappeared from the kinetic term of the shifted field h_v and has moved to the denominator of a nonlocal Lagrangian that is in fact the starting point for a systematic expansion in $1/m_Q$.

The Lagrangian for b and c quarks to leading order in $1/m_Q$ is therefore

$$\mathcal{L}_{b,c} = \overline{b_v}\, iv \cdot D\, b_v + \overline{c_v}\, iv \cdot D\, c_v \,. \tag{116}$$

This Lagrangian exhibits two important symmetries. The symmetries are only approximate because the Lagrangian (116) is not full QCD but the first approximation in an expansion in $1/m_Q$. The symmetries are manifest in (116) but they are hidden in full QCD.

- A heavy-flavour symmetry $SU(2)$ relates b and c quarks moving with the same velocity.
- Because there is no Dirac matrix in the Lagrangian (116), both spin degrees of freedom couple to gluons in the same way. Together with the flavour symmetry, this leads to an overall spin-flavour symmetry $SU(4)$.

The simplest spin-symmetry multiplet $\mathcal{M}$ consists of a pseudoscalar M and a vector meson M^*. One of the first important applications of spin-flavour symmetry was for the semileptonic decays $B \to D^{(*)}l\nu_l$. In general, there are several form factors governing the two matrix elements (for D and D^*). To leading order in $1/m_Q$, all those form factors are given up to Clebsch–Gordan coefficients by a single function $\xi(v \cdot v')$ called Isgur–Wise function:

$$\langle \mathcal{M}(v')|\overline{h_{v'}}\Gamma h_v|\mathcal{M}(v)\rangle \sim \xi(v \cdot v') \,. \tag{117}$$

Moreover, $\overline{h_v}\gamma^\mu h_v$ is the conserved current of heavy-flavour symmetry. Similar to electromagnetic form factors that are normalized at $q^2 = 0$ due to charge conservation, the Isgur–Wise function is fixed in the no-recoil limit $v = v'$ to be

$$\xi(v \cdot v' = 1) = 1 \,. \tag{118}$$

Of course, there are corrections to this result valid only in the symmetry limit, both of $O(\alpha_s)$ and in general of $O(1/m_Q)$. For the decay $B \to D^* l\nu_l$, the leading mass corrections turn out to be of $O(1/m_Q^2)$ only. HQET provides the standard method for the determination of the CKM matrix element V_{cb} (see Ref. [23] for reviews).

HQET has been extended in several directions.

- Soft collinear effective theory (SCET)

  HQET can not be applied to decays like $B \to X_s\gamma$ or $B \to \pi\pi$ where the light particles in the final state can have momenta of $O(m_Q)$. SCET accounts for those energetic light states but it is more complicated than HQET. Because of the presence of several scales, several effective fields must be introduced by successive field transformations. A major achievement is again the proof of factorization that is for instance crucial for a reliable extraction of the CKM matrix element V_{ub} from experiment.

– Nonrelativistic QCD (NRQCD)

This extension of HQET includes quartic interactions to treat heavy quarkonia $\overline{b}\,b$ and $\overline{c}\,c$. In this case, three widely separate scales are involved: the heavy mass m_Q, the bound-state momentum $p \sim m_Q v$ ($v \ll 1$) and the kinetic energy $E \sim m_Q v^2$. Applications include the analysis of the $q\,\overline{q}$ potential and the production and decay of quarkonia [24].

3.3   QCD sum rules

The general idea of QCD sum rules is to use the analyticity properties of current correlation functions to relate low-energy hadronic quantities to calculable QCD contributions at high energies. We recall the example of the two-point functions $\Pi_L^{(0,1)}(q^2)$ in hadronic τ decays discussed in Section 2.4.

In general, the QCD contribution consists itself of two different parts,

– a purely perturbative part and
– a partly nonperturbative part that is important at intermediate energies and makes use of the operator product expansion (OPE, another case of factorization).

Altogether, a typical two-point function (QCD sum rules are not restricted only to two-point functions, however) has the form

$$\Pi(q^2) = \Pi_{\text{pert}}(q^2) + \sum_d C_d(q^2)\langle 0|O_d|0\rangle \ . \tag{119}$$

The so-called Wilson coefficients $C_d(q^2)$ are calculable perturbatively and they depend on the two-point function under consideration. Up to logs, they decrease for large $|q^2|$ as $(q^2)^{-n(d)}$ with some positive integer $n(d)$. The vacuum condensates $\langle 0|O_d|0\rangle$, on the other hand, are universal and they absorb long-distance contributions with characteristic momenta $< \sqrt{|q^2|}$.

Three main types of applications of QCD sum rules can be distinguished.

– Using experimental data as input for the low-energy hadronic part, one can extract universal QCD parameters: α_s, quark masses, condensates, etc.
– With QCD parameters known, one can predict hadronic quantities: hadron masses, decay constants, amplitudes, etc.
– The compatibility of low-energy data with QCD can be checked. I shall discuss a recent example of topical interest, the spectral data relevant for the leading hadronic contribution to the anomalous magnetic moment of the muon.

$$\boxed{(\mathbf{g} - \mathbf{2})_\mu}$$

The biggest source of uncertainty in the calculation of the anomalous magnetic moment of the muon $a_\mu = (g_\mu - 2)/2$ in the Standard Model is at present the lowest-order hadronic vacuum polarization

$$a_\mu^{\text{had,LO}} = a_\mu^{\text{vac.pol.}} = \int_{4M_\pi^2}^{\infty} dt K(t)\sigma_0(e^+e^- \rightarrow \text{hadrons})(t) \tag{120}$$

depicted in Fig. 19. The kernel $K(t)$ is a known function [25]. Although the integral extends from threshold to infinity, about 73% of $a_\mu^{\text{had,LO}}$ are due to the $\pi\pi$ intermediate state in Fig. 19, governed by the pion form factor. Moreover, 70% of $a_\mu^{\text{had,LO}}$ come from the region $t \leq 0.8$ GeV^2. Therefore, by far the most important part is not calculable in QCD perturbation theory.
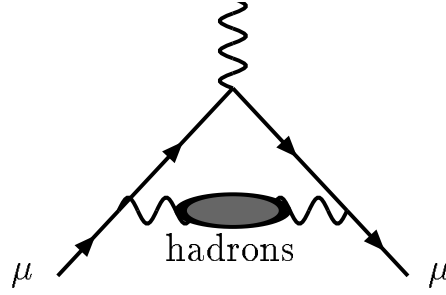
Fig. 19: Lowest-order hadronic vacuum polarization contribution to the anomalous magnetic moment of the muon

A few years ago, Alemany, Davier and Höcker [26] suggested to use not only data from $e^+e^- \to \pi^+\pi^-$ to extract the pion form factor but also from the decay $\tau^- \to \pi^0\pi^-\nu_\tau$. In the isospin limit, it is straightforward to derive the relation

$$\sigma_0(e^+e^- \to \pi^+\pi^-)(t) = h(t)\frac{d\Gamma(\tau^- \to \pi^0\pi^-\nu_\tau)}{dt} \tag{121}$$

with a known function $h(t)$ where t is the two-pion invariant mass squared.

At the level of precision required for comparison with experiment (better than 1% for $a_\mu^{\text{had,LO}}$), isospin violating and electromagnetic corrections are mandatory [27, 28]. The status until recently was summarized by Höcker at the High-Energy Physics Conference in Beijing [29].

– There was a significant discrepancy between the τ and e^+e^- data, mainly above the ρ resonance region, as shown in Fig. 20.
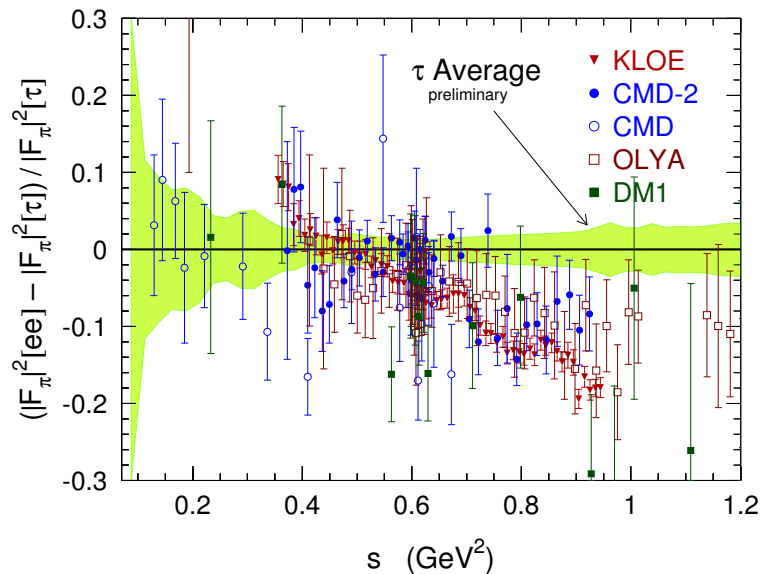
Fig. 20: Comparison between e^+e^- and τ data for the pion form factor from Ref. [29]; plotted is the difference of the form factors squared normalized to the τ data

– Isospin violation cannot explain the difference.
– The e^+e^- data from the KLOE experiment confirm the previous trend of the CMD-2 data although the agreement among the e^+e^- data is not impressive.

The widely accepted recommendation at the Beijing Conference was to ignore the τ data until the origin of the discrepancy is understood [29].

The situation has changed both on the theoretical and on the experimental side. The theoretical clarification is due to Maltman [30] who checked the consistency of experimental data with QCD by investigating sum rule constraints for the two spectral functions relevant for the e^+e^- and the τ case, respectively:

$$\rho_{\text{em}}(s) = \text{Im}\,\Pi_{\text{elm}}(s) \qquad \text{and} \qquad \rho_V^{I=1}(s) = \text{Im}\,\Pi_{L,ud}(s) \ .$$

By using a contour integral in the complex s-plane as in Fig. 13, with m_τ^2 replaced by an a priori arbitrary s_0, one derives a so-called F(inite)E(nergy)S(um)R(ule)

$$\int_0^{s_0} w(s)\rho(s)ds = -\frac{1}{2\pi} \oint_{|s|=s_0} w(s)\Pi(s)ds \ . \tag{122}$$

$\Pi(s), \rho(s)$ refer to the two cases considered (electromagnetic currents in the e^+e^- case, charged currents for τ) and $w(s)$ is an analytic function (actually a polynomial) that will be chosen conveniently.

The right-hand side of the FESR can be estimated in QCD as exemplified by Eq. (119). The purely perturbative part is known up to α_s^3, with estimates of the $O(\alpha_s^4)$ contribution available. For not too small s_0 the purely perturbative part dominates the right-hand side depending only on $\alpha_s(M_Z)$. The $d = 2$ part is completely negligible in the τ case and it depends only on the mass of the strange quark in the electromagnetic case. The $d = 4$ contributions involve the quark condensates $\langle 0|m_q\bar{q}q|0\rangle (q = u, d, s)$ (very well known from chiral perturbation theory, cf. Section 3.5) and the gluon condensate $\langle 0|\alpha_s G^a_{\mu\nu}G_a^{\mu\nu}|0\rangle$ (less well but sufficiently well known from charmonium sum rules). For $d \geq 6$, the relevant condensates are practically unknown. However, by using the pinching trick ($w(s_0) = 0$) appropriately, Maltman eliminates the $d = 6$ OPE contributions to $\Pi(s)$ altogether. The negligible effect of $d \geq 8$ contributions can be checked by varying s_0.

Turning now to the left-hand side of the FESR (122), Maltman uses the most precise experimental data for the spectral functions: ALEPH (compatible with CLEO and OPAL) for $\rho_V^{I=1}(s)$, CMD-2 for $\rho_{\text{em}}(s)$. As a first test, he fits $\alpha_s(M_Z)$ (keeping the remaining OPE input fixed) from the experimentally determined left-hand side of the FESR. The outcome is, quite independently of the weight function $w(s)$ that is always chosen positive and monotonically increasing for $0 \leq s \leq s_0$, that the fitted value of $\alpha_s(M_Z)$ is systematically lower than the high-energy determination dominated by LEP in the electromagnetic case while there is perfect agreement in the τ case. This is a first indication that the electromagnetic spectral density is too low.

A second test, largely independent of the value of $\alpha_s(M_Z)$, compares the slopes of the OPE parts and spectral integrals with respect to s_0. The results are shown in Figs. 21, 22.

The situation is similar to before. While the slopes differ between data and QCD by $\gtrsim 2.5\,\sigma$ in the electromagnetic case, there is perfect agreement in the τ case. The previous conclusion is reinforced: the e^+e^- spectral data are systematically too low whereas the τ data are completely consistent with QCD, both in normalization and in the slopes. The QCD sum rule tests clearly favour the τ over the e^+e^- data for the pion form factor.

The most recent development is again an experimental one. Only two months before the School new data on $e^+e^- \to \pi^+\pi^-$ were released by the SND Collaboration from Novosibirsk. Their results disagree with both CMD-2 and KLOE and they go in the right direction towards reconciling the e^+e^- with the τ data. The discrepancies between the three data sets in e^+e^- annihilation remain to be understood. Based on τ data for the 2π and 4π channels in the hadronic vacuum polarization, the calculation of the anomalous magnetic moment of the muon in the Standard Model [31] compares well with the measured value [32]:

$$a_\mu^{\text{exp}} - a_\mu^{\text{SM}} = (7.6 \pm 8.9) \cdot 10^{-10} \ . \tag{123}$$

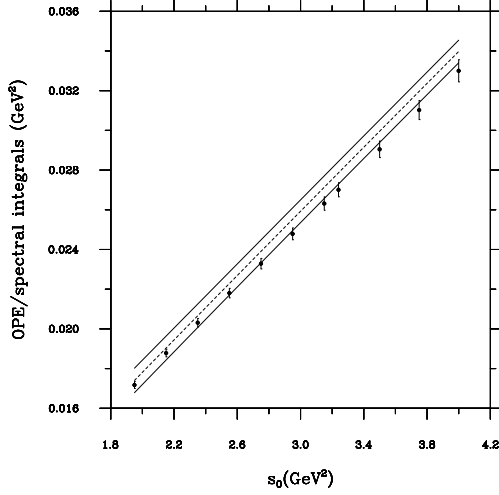There is at present no evidence for a discrepancy between the Standard Model and experiment.

Fig. 21: Slopes with respect to s_0 in the $e^+ e^-$ case for a specific weight function $w_6(s)$. The dashed lines denote the central values for the OPE input and the solid lines indicate the error bands. The spectral integrals are shown for several points, error bars included.
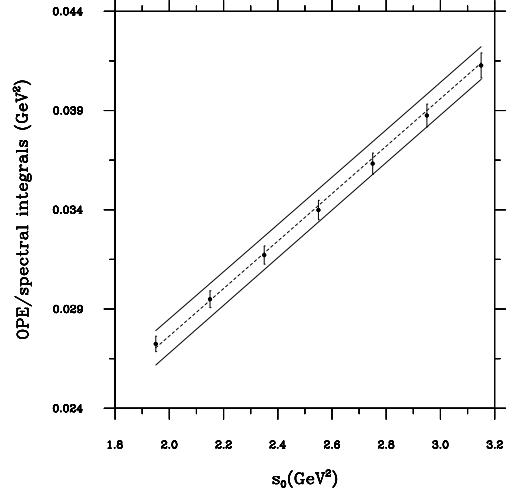
Fig. 22: Same as in the previous figure for the τ data. Both figures are taken from Ref. [30].

3.4 Chiral symmetry

By construction, QCD is a gauge theory with gauge group $SU(3)_c$. However, the QCD Lagrangian

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{2}\mathrm{tr}(G_{\mu\nu}G^{\mu\nu}) + \sum_{f=1}^{N_F} \overline{q}_f \left(i\not{D} - m_f \mathbb{1}_c\right) q_f \tag{124}$$

possesses additional symmetries. As in QED, the theory is parity invariant because of the absence of γ_5 (vector couplings only). Moreover, the coupling constant g_s and the quark masses m_f are real so that QCD conserves also CP, ignoring the so-called strong CP problem here.

Are there still additional symmetries in the QCD Lagrangian (124)? To answer this question, we first have a look at the quark kinetic term only (with $N_F = 6$ flavours):

$$\mathcal{L}_{\mathrm{kin}} = i\sum_{f=1}^{6} \overline{q}_f \not{D} q_f = i\sum_{f=1}^{6} \left\{ \overline{q}_{fL} \not{D} q_{fL} + \overline{q}_{fR} \not{D} q_{fR} \right\} , \tag{125}$$

with chiral components

$$q_L = \frac{1}{2}(1 - \gamma_5)q, \qquad\qquad q_R = \frac{1}{2}(1 + \gamma_5)q . \tag{126}$$

Since the q_{fL} and the q_{fR} do not talk to each other in (125), they can be rotated separately implying the maximal global flavour symmetry $U(6) \times U(6)$. However, this is a symmetry of the kinetic term only. In the full quark Lagrangian (colour indices will be suppressed from now on)

$$\mathcal{L}_q = \sum_{f=1}^{6} \left\{ \overline{q}_{fL}\, i\not{D}\, q_{fL} + \overline{q}_{fR}\, i\not{D}\, q_{fR} - m_f(\overline{q}_{fR}q_{fL} + \overline{q}_{fL}q_{fR}) \right\} \tag{127}$$

q_{fL} and q_{fR} can in general not be rotated separately any longer because of the quark masses. The actual flavour symmetry therefore depends on the quark mass matrix

$$\mathcal{M}_q = \mathrm{diag}\,(m_u, m_d, m_s, m_c, m_b, m_t) . \tag{128}$$

In order to find all even only approximate symmetries of QCD, we distinguish several cases depending on the specific values of the quark masses.

i. In the real world, all quark masses are non-zero and they are all different from each other. In this case, the remaining flavour symmetry amounts to the phase transformations $q_{fL,R} \rightarrow e^{-i\varepsilon_f} q_{fL,R}$ ($f = 1, \ldots, 6$) where the phase ε_f for a given flavour must be the same for q_{fL} and q_{fR}. The symmetry group reduces to the product $U(1) \times U(1) \times \cdots \times U(1) = U(1)^6$ leading to the well-known conserved flavour quantum numbers N_u, N_d, N_s, N_c, N_b and N_t. All these symmetries are broken by the weak interactions, except their sum (baryon number)

$$B = (N_u + N_d + N_s + N_c + N_b + N_t)/3 . \tag{129}$$

ii. In some approximation, the quark masses are still non-zero but n_F of them are equal ($n_F < N_F = 6$). In this case, the maximal symmetry group $U(6) \times U(6)$ reduces to

$$U(n_F) \times U(1)^{6-n_F} \simeq SU(n_F) \times U(1) \times U(1)^{6-n_F} . \tag{130}$$

The only realistic cases are $n_F = 2$ or 3 and they lead to well-known approximate symmetries:

$$n_F = 2: \qquad m_u = m_d \qquad \longrightarrow \qquad \text{isospin } SU(2)$$
$$n_F = 3: \qquad m_u = m_d = m_s \qquad \longrightarrow \qquad \text{flavour } SU(3) .$$

iii. A much more radical approximation consists in setting some of the quark masses to zero: $m_f = 0$ ($f = 1, \ldots, n_F$). In this case, n_F of the q_{fL} and q_{fR} can again be rotated separately implying the chiral symmetry

$$SU(n_F)_L \times SU(n_F)_R \times U(1)_V \times U(1)_A \left[\times U(1)^{6-n_F} \right] . \tag{131}$$

To set $n_F = 2$ of the quark masses to zero is a reasonable approximation in view of $m_{u,d} \ll \Lambda_{\mathrm{QCD}}$, whereas $n_F = 3$ (setting also $m_s = 0$) is certainly more daring. $U(1)_V$ is again responsible for baryon number conservation. The factor $U(1)_A$ is actually not a symmetry of full QCD at the quantum level (Abelian anomaly).

We are familiar with isospin and flavour $SU(3)$ that we see at least approximately realized in the hadron spectrum. But what are the consequences of the approximate chiral symmetry of QCD? If chiral symmetry would manifest itself in the hadron spectrum, each hadron would have to have a partner of opposite parity of approximately the same mass. That is obviously not the case: chiral symmetry appears to be more hidden than isospin, for example. In order to understand the manifestations of chiral symmetry, we have to recall the main features of

$$\boxed{\text{Spontaneous symmetry breaking}}$$

There are many examples of SSB in physics and the best-known example in particle physics is the spontaneously broken electroweak symmetry (see the lectures of W. Buchmüller at this School).

The mechanism was first realized in condensed-matter physics and a good example for our purpose is the ferromagnet. The underlying theory of the ferromagnet (eventually QED) is certainly rotationally invariant. The Hamiltonian (e.g., of the Heisenberg model) does not single out any direction in space. Nevertheless, in the ground state of the ferromagnet the spins align in a certain direction. The direction is arbitrary and there is no trace of it in the Hamiltonian. In this sense rotational symmetry is 'spontaneously' broken. It is certainly not manifest in the ground state but it has other important consequences such as the existence of excitations (called magnons or spin waves in this case) with a very special dispersion law. The dispersion law is the dependence of the frequency ω on the wave length λ or on the wave

number $k = 2\pi/\lambda$. SSB in condensed-matter physics implies that for some excitations the frequency tends to zero for infinite wave length:

$$\lim_{k\to 0} \omega(k) = \lim_{\lambda\to\infty} \omega(k) = 0 \ . \tag{132}$$

This property of magnons in particular is easy to visualize as shown in Fig. 23. In the ground state of the ferromagnet spins are aligned. A typical spin wave is displayed in the second line: the wave length is the distance between spins pointing in the same direction. In the limit $\lambda \to \infty$, the spins become again aligned, albeit in a different direction in general. Since by the assumed rotational symmetry of the theory each direction is as good as any other, the configurations in the first and in the last line must have the same energy as expressed by Eq. (132). Put in another way, magnons do not have an energy gap in their spectrum. What is the analogy in a relativistic QFT like QCD? The frequency is replaced by the

ground state

spin wave: finite λ
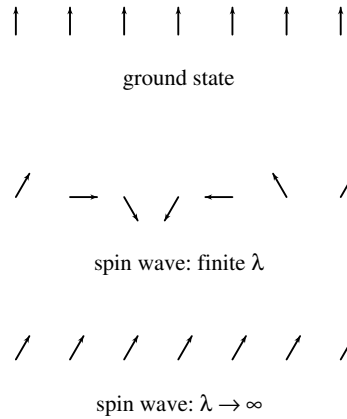
spin wave: $\lambda \to \infty$

Fig. 23: Spin directions in the (one-dimensional) ferromagnet: ground state, spin wave with finite wave length and a spin wave with infinite wave length

energy of the particle, with the three-momentum instead of the wave number. The dispersion law is just the relativistic energy-momentum relation $E = \sqrt{p^2 + m^2}$. The energy tends to zero for $p \to 0$ if and only if the particle is massless:

$$\lim_{p\to 0} E = \lim_{p\to 0} \sqrt{p^2 + m^2} = 0 \qquad \Longleftrightarrow \qquad m = 0 \ . \tag{133}$$

How can we distinguish if a symmetry is realized in the usual way (like isospin) or if it is spontaneously broken? The crucial question is what the conserved charge $Q = \int d^3x J^0(x)$ associated with a symmetry current $J^\mu(x)$ ($\partial_\mu J^\mu = 0$) does when applied to the ground state (vacuum). In a relativistic quantum field theory, there are only two options:

<u>Goldstone alternative</u>

$Q\lvert 0\rangle = 0$	$\lVert Q\lvert 0\rangle\rVert = \infty$
Wigner–Weyl	Nambu–Goldstone
linear representation	nonlinear realization
degenerate multiplets	massless Goldstone bosons
exact symmetry	spontaneously broken symmetry

The left column describes the more familiar case (Wigner–Weyl) where states are grouped in multiplets (irreducible representations of the symmetry group). The vacuum is annihilated by the charge and the states in a given multiplet all have the same mass (degeneracy). In the other possible realization (Nambu–Goldstone), applying the charge to the vacuum is, strictly speaking, not defined. There are no degenerate

multiplets (therefore we do not see the symmetry in the spectrum) but there must be massless particles in the theory (Goldstone bosons). Although the charge can not be applied to the vacuum directly, the following matrix element called an order parameter of SSB may be well defined:

$$\langle 0|[Q, A]|0\rangle \tag{134}$$

where A is some operator. If we can find an order parameter that is different from zero, the symmetry associated with the charge Q is necessarily spontaneously broken. This is easy to understand: because of the commutator in (134) the order parameter vanishes if Q annihilates the vacuum. In the familiar example of electroweak symmetry breaking, scalar field operators φ_i take the place of A with $[Q, \varphi_i] = c_{ij}\varphi_j$. If $c_{ij}\langle 0|\varphi_j|0\rangle \neq 0$ (Higgs vacuum expectation value), the electroweak symmetry is spontaneously broken. This is also a good example that the argument goes in one direction only. Even if all scalars have zero vacuum expectation values or if there are no scalar fields at all the symmetry may still be spontaneously broken. The mechanism could involve some other operator A in the order parameter (134).

For each spontaneously broken symmetry Goldstone's theorem implies the existence of a massless state $|\mathrm{G}\rangle$ with

$$\langle 0|J^0(0)|\mathrm{G}\rangle\langle \mathrm{G}|A|0\rangle \neq 0 \ . \tag{135}$$

A necessary and sufficient condition for SSB is that the

$$\boxed{\text{Goldstone matrix element } \langle 0|J^0(0)|\mathrm{G}\rangle \neq 0}$$

implying also that the Goldstone state $|\mathrm{G}\rangle$ has the same quantum numbers as $J^0(0)|0\rangle$. The following remarks are useful:

- The state $|\mathrm{G}\rangle$ need not correspond to a physical particle. This can only happen in the case of a spontaneously broken gauge symmetry as in electroweak theory.
- $J^0(0)$ is usually a rotationally invariant bosonic operator and thus $|\mathrm{G}\rangle$ carries spin 0.
- Spontaneous breaking of discrete symmetries does not give rise to Goldstone bosons.

The main features of SSB can be discussed in the original Goldstone model. It has a single complex scalar field $\phi(x)$ with the Lagrangian

$$\mathcal{L}_{\text{Goldstone}} = \partial_\mu \phi \partial^\mu \phi^\dagger - \lambda(\phi\phi^\dagger - \frac{v^2}{2})^2 \qquad (\lambda, \ v \text{ real and positive}) \ , \tag{136}$$

symmetric with respect to global $U(1)$ transformations $\phi(x) \to e^{i\alpha}\phi(x)$. The minimum of the Mexican hat potential occurs at $\phi\phi^\dagger = \dfrac{v^2}{2}$. Decomposing the complex field $\phi(x)$ into two Hermitian fields $R(x), G(x)$ with

$$\phi(x) \quad = \quad (R(x) + iG(x))/\sqrt{2} \tag{137}$$

$$\langle 0|R(x)|0\rangle = v, \qquad \langle 0|G(x)|0\rangle = 0 \ ,$$

the Lagrangian expressed in terms of the fields $R(x), G(x)$ displays the following spectrum at tree level:

$$\begin{array}{ll} \text{Goldstone boson field } G(x) & M_G = 0 \\ \text{massive field } H(x) = R(x) - v & M_H = \sqrt{2\lambda}\, v \ . \end{array}$$

Denoting the four-momenta of Goldstone particles generically as p_G, one finds an unexpected behaviour for scattering amplitudes: they vanish for $p_G \to 0$, e.g.,

$$A(GG \to GG) = O(p_G^4), \qquad A(GH \to GH) = O(p_G^2) \tag{138}$$

for arbitrary values of non-Goldstone momenta p_H. More generally, Goldstone bosons decouple when their energies tend to zero.

This behaviour looks mysterious at first, but it can be understood by choosing a different set of fields. Instead of the fields $G(x)$ and $H(x)$, we choose another representation (polar decomposition) that may be familiar from electroweak theory:

$$\phi(x) = \frac{1}{\sqrt{2}}[h(x) + v]e^{ig(x)/v} . \tag{139}$$

In terms of the Hermitian fields $g(x), h(x)$ the Goldstone Lagrangian takes the form

$$\begin{aligned} \mathcal{L}_{\text{Goldstone}} &= \frac{1}{2}(\partial_\mu g)^2 + \frac{1}{2v^2}(h^2 + 2vh)(\partial_\mu g)^2 \\ &+ \frac{1}{2}(\partial_\mu h)^2 - \lambda v^2 h^2 - \frac{\lambda}{4}(h^4 + 4vh^3) . \end{aligned} \tag{140}$$

A general theorem of QFT ensures that the fields G, H on one side and g, h on the other side produce the same S-matrix elements although the Green functions are in general different.

The main consequences are the following.

– The particle spectrum is unchanged:

Goldstone field $g(x)$ $\qquad M_g = 0$
massive field $h(x)$ $\qquad M_h = \sqrt{2\lambda}\, v$.

– The Goldstone field g has only derivative couplings implying for the scattering amplitudes considered previously:

$$\lim_{p_G \to 0} A(p_G) = 0 . \tag{141}$$

The important lesson is very general and not restricted to the Goldstone model. S-matrix elements with only Goldstone states vanish for $p_G \to 0$. When other non-Goldstone particles participate in the initial and final states, the statement remains true for some matrix elements like for elastic scattering $G\,H \to G\,H$. In general, the interactions of Goldstone bosons among themselves and with other matter become arbitrarily weak for small momenta.

3.5 Chiral perturbation theory

We start from a theorist's paradise (copyright H. Leutwyler), QCD in the chiral limit where $n_F = 2$ [or 3] quarks $u, d\,[, s]$ are massless:

$$\mathcal{L}_{\text{QCD}}^0 = \overline{q_L}i\slashed{D}q_L + \overline{q_R}i\slashed{D}q_R + \mathcal{L}_{\text{heavy quarks}} + \mathcal{L}_{\text{gauge}} \tag{142}$$

with

$$q^\top = (u\; d\; [s]) .$$

As explained in the previous section, this Lagrangian has a global symmetry

$$SU(n_F)_L \times SU(n_F)_R \times U(1)_V \times U(1)_A \left[\times U(1)^{6-n_F}\right] . \tag{143}$$

The non-Abelian factor $G = SU(n_F)_L \times SU(n_F)_R$ is called the chiral group.

Although not yet proven from QCD alone, there is strong evidence, both from phenomenology and from theory, that chiral symmetry is spontaneously broken. The spontaneous breaking does not affect all of G but, roughly speaking, half of it: $G \longrightarrow H = SU(n_F)_V$. The so-called vectorial subgroup H of G is nothing but isospin (for $n_F = 2$) or flavour $SU(3)$ (for $n_F = 3$) and it is realized in the familiar way à la Wigner–Weyl. Some arguments in favour of this spontaneous breakdown are as follows.

- As already emphasized before, there are no parity doublets in the hadron spectrum.

- There is no other convincing argument why the pseudoscalar mesons are the lightest hadrons. Spontaneous chiral symmetry breaking implies that they would be massless in the chiral limit (pseudo-Goldstone bosons).

- The vector and axial-vector spectral functions are very different (ρ vs. a_1).

- The so-called anomaly matching conditions together with confinement require that G be spontaneously broken for $n_F \geq 3$.

- Under very reasonable assumptions, $SU(n_F)_V$ is not spontaneously broken. It is of course explicitly broken by the differences between quark masses.

Even if it has not been possible so far to prove directly from QCD that chiral symmetry is spontaneously broken, we can ask for possible order parameters. It turns out (more details can be found in Ref. [33], for instance) that the simplest such order parameter involves the pseudoscalar operators $A_b = \overline{q}\gamma_5\lambda_b q$ ($a = 1, \ldots, 8$) giving rise to the quark condensate $\langle 0|\overline{q}\, q|0\rangle$. There is evidence both from lattice QCD and from phenomenology that the quark condensate is non-zero, implying spontaneous chiral symmetry breaking. As will be discussed in Section 3.7, the quark condensate is in fact the dominant order parameter of spontaneous chiral symmetry breaking, in a sense to be specified later.

From Goldstone's theorem we know (still in the chiral limit) that there are $n_F^2 - 1$ massless Goldstone bosons:

n_F	$n_F^2 - 1$	Goldstone bosons
2	3	π
3	8	π, K, η

Although the real world is not a theorist's paradise, we still expect low-energy amplitudes to be dominated by the exchange of pseudoscalar mesons, which are the lightest hadrons also in the real world. In order to calculate such amplitudes, we construct an effective field theory with only Goldstone fields. As already explained in Section 3.1, the Lagrangian of Goldstone fields is nonrenormalizable and it is in fact even nonpolynomial. The underlying physical reason is that we can add any number of sufficiently soft pions (still massless!) to a hadron state without appreciably changing its energy. Therefore, we have degenerate states with different numbers of particles that are related by chiral symmetry transformations. For the Lagrangian this argument implies that the symmetry transformations are nonlinear in the pion fields. Starting from a Lagrangian with fixed powers in the Goldstone fields, successive nonlinear transformations generate any number of fields in the Lagrangian. Since the Lagrangian is to be symmetric under such transformations it must necessarily be nonpolynomial.

The basic building block of chiral Lagrangians is therefore a nonpolynomial matrix function of the Goldstone fields, e.g., the exponential function (for $n_F = 3$)

$$U(\phi) = \exp\left(i\sqrt{2}\Phi/F\right), \qquad \Phi = \begin{pmatrix} \dfrac{\pi^0}{\sqrt{2}} + \dfrac{\eta_8}{\sqrt{6}} & \pi^+ & K^+ \\[2mm] \pi^- & -\dfrac{\pi^0}{\sqrt{2}} + \dfrac{\eta_8}{\sqrt{6}} & K^0 \\[2mm] K^- & \overline{K^0} & -\dfrac{2\eta_8}{\sqrt{6}} \end{pmatrix} \tag{144}$$

where F is the pion decay constant in the chiral limit characterizing the size of the Goldstone matrix element $\langle 0|J^0(0)|\mathrm{G}\rangle$.

Chiral Lagrangians are organized according to the number of derivatives of the fields. The unique lowest-order Lagrangian of $O(p^2)$ with two derivatives is the so-called nonlinear σ model:

$$\mathcal{L}_2^{(0)} = \frac{F^2}{4}\mathrm{tr}_{n_F}\left(\partial_\mu U \partial^\mu U^\dagger\right) =: \frac{F^2}{4}\langle\partial_\mu U \partial^\mu U^\dagger\rangle = \partial_\mu\pi^+\partial^\mu\pi^- + \frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 + O(\pi^4), \tag{145}$$

using a bracket notation for n_F-dimensional traces.

So much for the paradise. Back to reality, we must admit that there is no chiral symmetry in nature! In the Standard Model, it is explicitly broken in two different ways.

– Chiral symmetry is explicitly broken by nonvanishing quark masses. This should be a small modification for two, a more pronounced one for three flavours:

$$
\begin{array}{llll}
m_u, m_d & \ll & M_\rho & n_F = 2 \\
m_s & < & M_\rho & n_F = 3 \, .
\end{array}
$$

– Also the electroweak interactions break chiral symmetry. If electroweak effects are to be included, they can be taken into account perturbatively in α, G_F.

The main assumption of chiral perturbation theory (CHPT) is that an expansion around the chiral limit (the theorist's paradise) makes sense. Therefore, even in the absence of electroweak interactions, chiral Lagrangians are organized in a two-fold expansion.

i. Spontaneous chiral symmetry breaking gives rise to an expansion in derivatives of the fields leading to an expansion of amplitudes in the momenta of pseudo-Goldstone bosons.

ii. Explicit symmetry breaking suggests an expansion also in the quark masses m_q.

The two expansions can be related via the meson masses. As will be discussed in the next subsection, the squares of the meson masses start out linear in the quark masses:

$$
M_M^2 \sim m_q + O(m_q^2) \, . \tag{146}
$$

The standard chiral counting therefore amounts to treating quark masses like the second power of momenta:

$$
m_q = O(M_M^2) = O(p^2) \, . \tag{147}
$$

The effective Lagrangian (for pseudoscalar mesons) is therefore of the form [34]

$$
\begin{aligned}
\mathcal{L}_{\text{eff}} &= \mathcal{L}_2 + \mathcal{L}_4 + \mathcal{L}_6 + \dots \\
\mathcal{L}_2 &= \frac{F^2}{4} \langle \partial_\mu U \partial^\mu U^\dagger + \chi U^\dagger + \chi^\dagger U \rangle
\end{aligned} \tag{148}
$$

where χ represents the quark masses: $\chi = 2B\mathcal{M}_q = 2B \operatorname{diag}(m_u, m_d[, m_s])$. The lowest-order Lagrangian contains only two parameters F, B that are related to physical quantities as

$$
F_\pi = F \left[1 + O(m_q)\right] \, , \qquad \langle 0|\overline{u}u|0\rangle = -F^2 B \left[1 + O(m_q)\right] \, . \tag{149}
$$

The lowest-order amplitudes of CHPT are of $O(p^2)$ and they correspond to the current algebra amplitudes of 40 years ago. The tree-level amplitudes can be read off directly from the Lagrangian (148) depending only on F_π and M_M^2 $[M_\pi^2 = B(m_u + m_d), \dots]$. For instance, the elastic $\pi\pi$ scattering amplitude of $O(p^2)$ is given by

$$
A_2(s, t, u) = \frac{s - M_\pi^2}{F_\pi^2} \, . \tag{150}
$$

Contrary to symmetries like isospin that relate different amplitudes, the spontaneously broken chiral symmetry makes an absolute prediction for this scattering amplitude. It was left as an exercise to the audience to explain why that is possible.

The lowest-order results we have been discussing so far were already known in the late 1960s and early 1970s (current algebra, phenomenological Lagrangians). After an influential paper by Weinberg [35], but especially with the work of Gasser and Leutwyler [34] the systematic treatment of QCD at

low energies became a respectable theory. The first step was to construct the Lagrangian of next-to-leading order $\mathcal{L}_4$ that contains 10 (7) additional coupling constants (usually called LECs for low-energy constants) for $SU(3)$ ($SU(2)$). With a Hermitian Lagrangian tree amplitudes are necessarily real but we know that unitarity and analyticity require complex amplitudes. It is not difficult to convince oneself that imaginary parts occur first at $O(p^4)$. The consequence is that a systematic low-energy expansion requires a loop expansion beyond lowest order [35]. But loop amplitudes have a tendency to be divergent. For a bona fide QFT we therefore need both regularization and renormalization. As strange as it may sound, nonrenormalizable theories also can and in fact must be renormalized to qualify as respectable QFTs.

A nonrenormalizable QFT like CHPT has many common features with the more standard renormalizable theories.

- Divergences are absorbed by the coupling constants in the higher-order Lagrangians $\mathcal{L}_4, \mathcal{L}_6, \ldots$. Unlike in renormalizable theories, new LECs occur at every order of the chiral expansion.
- The renormalized LECs are scale dependent just like the strong coupling constant $g_s(\mu)$. They can be interpreted as describing the effect of all heavy hadronic states that are not represented by explicit fields in the Lagrangian.
- Renormalization ensures that there is no dependence on some artificial cut-off.

For phenomenological applications, we have to know the values of the various LECs. In principle, QCD fixes those constants but a matching between QCD and CHPT is not possible in perturbation theory. This was already discussed in Section 3.1 for general EFTs with SSB but it is also easy to understand in the present case: CHPT can only be applied for energies $E \ll M_\rho$ whereas perturbative QCD only makes sense for $E \gtrsim 1.5$ GeV. The most successful approaches bridging this gap to get information on the LECs are resonance saturation (based on the properties of QCD for large N_c, i.e., for a fictitious world with many colours) and lattice QCD.

The chiral expansion is an expansion in $p^2/(4\pi F)^2$ where p is a characteristic momentum for the process in question. Therefore, the chiral expansion should and does work better for $SU(2)$ than for $SU(3)$:

$$n_F = 2 : \frac{p^2}{(4\pi F)^2} = 0.014 \frac{p^2}{M_\pi^2} , \qquad n_F = 3 : \frac{p^2}{(4\pi F)^2} = 0.18 \frac{p^2}{M_K^2} . \qquad (151)$$

Most amplitudes and form factors for realistic processes have been calculated at least to next-to-leading order. There is an easy-to-use Mathematica program to generate both strong and nonleptonic weak transitions up to $O(p^4)$ that is described in Ref. [36] and is available for general use. The state of the art is next-to-next-to-leading order or $O(p^6)$. A short introduction can be found in Ref. [33].

3.6 Light quark masses

In CHPT, the light quark masses always appear in the combination $B\,m_q \sim m_q\,\langle 0|\overline{u}\,u|0\rangle$. Since there are no quarks or gluons in CHPT, only QCD scale-invariant quantities can appear. As we discussed in Section 3.2, quark masses are scale dependent whereas the product $m_q\,\langle 0|\overline{u}\,u|0\rangle$ is not. The consequence is that CHPT can only provide methods for extracting the ratios of quark masses.

The lowest-order expressions for the meson masses in terms of quark masses can be read off directly from the Lagrangian (148):

$$M_{\pi^+}^2 = 2\hat{m}B , \qquad M_{\pi^0}^2 = 2\hat{m}B + O\left[(m_u - m_d)^2/(m_s - \hat{m})\right]$$

$$M_{K^+}^2 = (m_u + m_s)B , \qquad M_{\eta_8}^2 = \frac{2}{3}(\hat{m} + 2m_s)B + O\left[(m_u - m_d)^2/(m_s - \hat{m})\right]$$

$$M_{K^0}^2 = (m_d + m_s)B , \qquad \hat{m} := \frac{1}{2}(m_u + m_d) . \qquad (152)$$

Several well-known relations follow from these expressions:

| Gell-Mann–Oakes–Renner | $F_\pi^2 M_\pi^2 = -2\hat{m}\langle 0|\bar{u}\,u|0\rangle$ |

Gell-Mann–Oakes–Renner $\qquad F_\pi^2 M_\pi^2 = -2\hat{m}\langle 0|\bar{u}\,u|0\rangle$

GMOR, Weinberg $\qquad B = \dfrac{M_\pi^2}{2\hat{m}} = \dfrac{M_{K^+}^2}{m_s + m_u} = \dfrac{M_{K^0}^2}{m_s + m_d}$

Gell-Mann–Okubo $\qquad 3M_{\eta_8}^2 = 4M_K^2 - M_\pi^2 \qquad$ (isospin limit)

The relations (152) are also the basis for the so-called current algebra mass ratios

$$\frac{m_u}{m_d} = 0.55\,, \qquad \frac{m_s}{m_d} = 20.1\,, \qquad \frac{m_s}{\hat{m}} = 25.9\,. \qquad (153)$$

These ratios are subject to higher-order corrections, most importantly of $O(p^4) = O(m_q^2)$ and $O(e^2 m_s)$. Because of an accidental symmetry at $O(p^4)$, the ratios m_s/m_d, m_u/m_d cannot be extracted separately from S-matrix elements but only in the combination known as Leutwyler's ellipse [37] shown in Fig. 24.
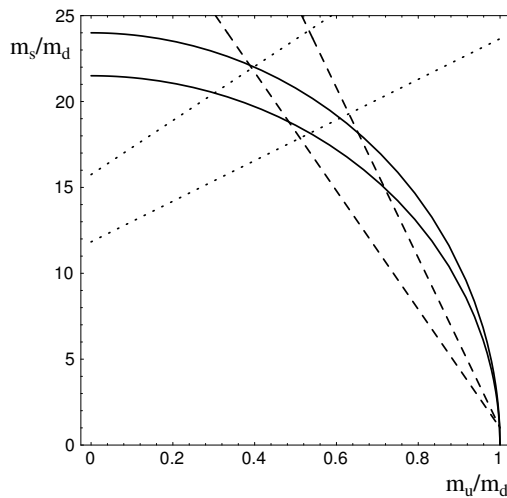
Fig. 24: Constraints on light quark mass ratios [37]

In addition to the full boundaries following directly from CHPT (the difference is due to uncertainties in the electromagnetic corrections), information is also available from $\eta - \eta'$ mixing (dotted lines), baryon mass splittings and $\rho - \omega$ mixing (dashed lines). The overall conclusion is that the corrections of $O(p^4)$ are small for the ratios. The next-to-next-to-leading corrections of $O(p^6)$ are also known [38] but there are at the moment too many unknown LECs for quantitative conclusions.

Table 1: Quark mass ratios to $O(p^4)$ [37]

	m_u/m_d	m_s/m_d	$m_s/\hat{m}$
$O(p^2)$	0.55	20.1	25.9
$O(p^4)$	0.55 ± 0.04	18.9 ± 0.8	24.4 ± 1.5

Absolute values of the quark masses are less well known than the ratios. The main methods are QCD sum rules and lattice simulations, most recently with full (unquenched) QCD. From the Review of Particle Physics [14], the combined result of lattice determinations of the strange quark mass is $m_s(2\,\text{GeV}) = (100 \pm 25)$ MeV. More results are reproduced in Fig. 18.

94

3.7 Pion–pion scattering

Pion–pion scattering has a privileged status in CHPT. It is **the** fundamental scattering process of CHPT and it involves only pions. The low-energy expansion can therefore be set up in the framework of chiral $SU(2)$ and it can be expected to converge well. The scattering amplitude is very sensitive to the mechanism of spontaneous chiral symmetry breaking giving access to the quark condensate in particular.

The following review of recent developments will be restricted to the isospin limit ($m_u = m_d$) in the absence of electromagnetic corrections. In this case, the information for all possible scattering channels is contained in a single amplitude $A(s, t, u)$ (with $s + t + u = 4M_\pi^2$).

The lowest-order amplitude of $O(p^2)$ was already shown in Eq. (150):

$$A_2(s, t, u) = \frac{s - M_\pi^2}{F_\pi^2} .$$

At the same order, the quark mass ratio $r = \dfrac{m_s}{\hat{m}} = \dfrac{2M_K^2}{M_\pi^2} - 1 \simeq 26$, as also shown in Table 1. As the mass ratio r, the S-waves also are very sensitive to the quark condensate. In a modified version of CHPT (Generalized CHPT [39]), one can tune the quark condensate. As an example, I show the leading-order results for the $I = 0$ S-wave scattering length a_0^0 and for the quark mass ratio r for both the standard and for a very small value of the quark condensate:

a_0^0	r	$B(\nu = 1 \text{ GeV})$	
0.16	26	1.4 GeV	(standard value)
0.26	10	F_π	

At next-to-leading order, the scattering amplitude was calculated in 1983 [40]. It turns out that especially the S-wave scattering lengths are quite sensitive to chiral corrections (chiral logs). For instance, a_0^0 increases from 0.16 at $O(p^2)$ to 0.20 at $O(p^4)$, an increase of 25% and thus quite a bit larger than the natural estimate in Eq. (151). Since the favoured experimental value of a_0^0 at that time was 0.26 (with a 25% error), it seemed mandatory to perform one more step in the chiral expansion. From the amplitude to $O(p^6)$ [41] it was clear that a value $a_0^0 = 0.26$ was not compatible with QCD. To narrow down the uncertainties related to the LECs appearing in the $\pi\pi$ amplitude, the chiral amplitude was finally combined with dispersion theory (Roy equations) [42].

CHPT together with dispersion theory predicts not only the S-wave scattering lengths with amazing precision [42],

$$a_0^0 = 0.220 \pm 0.005 , \qquad a_0^2 = -0.0444 \pm 0.0010 , \tag{154}$$

but also the S- and P-wave phase shifts. As a by-product, the pionium lifetime is predicted to be $\tau = (2.9 \pm 0.1) \cdot 10^{-15}$. A short history of the S-wave scattering lengths is shown in Fig. 25.

There is a small caveat here. All the results have been derived in the standard framework of CHPT that assumes a large quark condensate. With recent experimental information on the pion–pion phase shift difference $\delta_0^0 - \delta_1^1$ from K_{e4} decays, even this last loophole could be closed. Using the correlation between a_0^2 and a_0^0 implied by the Roy equations, the measured phase shift difference [43] can be used to determine a_0^0 as shown in Fig. 26. The fitted value [44] $a_0^0 = 0.221 \pm 0.026$ is in perfect agreement with the prediction (154) from the combined analysis of CHPT and Roy equations.

The precise determination of the $\pi\pi$ scattering amplitude has several important implications. One application concerns the chiral expansion of the pion mass in terms of the light quark masses m_u, m_d:

$$M_\pi^2 = M^2 - \frac{\bar{l}_3}{32\pi^2 F^2} M^4 + O(M^6) \tag{155}$$

$$M^2 = (m_u + m_d)|\langle 0|\bar{u}u|0\rangle|/F^2 .$$

A by-product of the analysis of $\pi\pi$ scattering is a precise determination of LECs like $\bar{l}_3$, which implies in turn that more than 94% of M_π are in fact due to the leading term (the Gell-Mann–Oakes–Renner term) confirming the standard mechanism of spontaneous chiral symmetry breaking [44]. In other words, the quark condensate is indeed the dominant order parameter of chiral symmetry breaking.
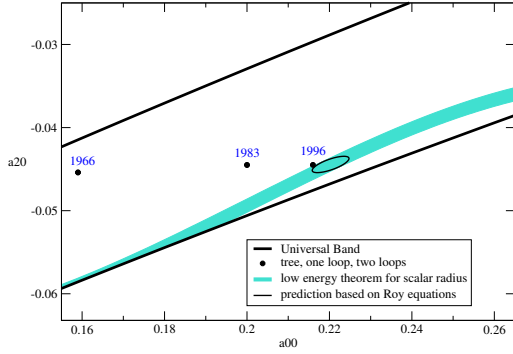
Fig. 25: Predictions for S-wave scattering lengths from current algebra till today, taken from Ref. [42]
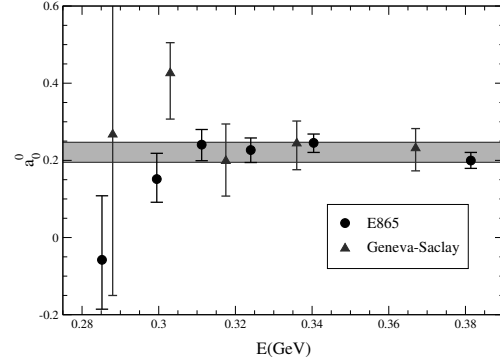
Fig. 26: K_{e4} data translated into a determination of the $I = 0$ S-wave scattering length a_0^0 [44]

The precise knowledge of the $\pi\pi$ scattering amplitude from CHPT and dispersion theory has more recently (after the School) produced another important insight. The much discussed scalar isoscalar state now called $f_0(600)$ by the Particle Data Group [14], but more commonly known as the σ meson, was analysed by Caprini, Colangelo and Leutwyler [45] by extending the Roy equations to complex values of the Mandelstam variable s. The result is an astonishingly precise determination of mass and width of the lightest hadronic resonance:

$$M_\sigma = 441 \, ^{+16}_{-8} \text{ MeV} , \qquad \Gamma_\sigma = 544 \, ^{+25}_{-18} \text{ MeV} . \tag{156}$$

The σ resonance has the quantum numbers of the vacuum and it corresponds to an unambiguous pole on the second sheet of the scalar isoscalar partial wave. Its real part is close to threshold but the pole is quite far from the real axis. Although not as straightforward as its position in the complex energy plane, the most appealing interpretation of the σ is a quasi-bound state of two pions, quite different from a member of a standard $\overline{q}q$ meson nonet [42, 45].

3.8 K_{l3} decays and V_{us}

The Cabibbo–Kobayashi–Maskawa (CKM) matrix V_{ij} determines the structure of the hadronic charged weak current. The matrix elements are fundamental parameters of the Standard Model. Together with the quark and lepton masses, the CKM matrix and the corresponding lepton mixing matrix contain information about the mechanism of mass generation. Anticipating forthcoming LHC data to unveil the secrets of mass generation (Higgs sector), both masses and mixing angles should be measured as precisely as possible.

With three generations of quarks, the CKM matrix must be a unitary matrix. For some time, a possible problem with CKM unitarity has been discussed. With the PDG values of 2004 [14],

$$|V_{ud}| = 0.9738(5) , \qquad |V_{us}| = 0.2200(26) , \tag{157}$$

unitarity appeared to be violated at the 2.2 σ level by the elements of the first row V_{uj} $(j = d, s, b)$:

$$\sum_{j=d,s,b} |V_{uj}|^2 - 1 = -0.0033(15) . \tag{158}$$

At this level of accuracy, the element V_{ub} is still negligible. On the other hand, the required precision for V_{ud} and V_{us} calls for reliable isospin violating and electromagnetic corrections.

The most accurate determination of V_{us}, both experimentally and theoretically, comes from K_{l3} decays that can be treated in the framework of CHPT. The decay amplitude is governed by two form factors $f_+(t)$ and $f_-(t)$ with $t = (p_K - p_\pi)^2$:

$$\langle \pi^-(p_\pi)|\bar{s}\gamma_\mu u|K^0(p_K)\rangle = f_+^{K^0\pi^-}(t)\,(p_K + p_\pi)_\mu + f_-^{K^0\pi^-}(t)\,(p_K - p_\pi)_\mu\,. \tag{159}$$

Both form factors are known to next-to-next-to-leading order in CHPT. For the extraction of V_{us}, the crucial quantity is $f_+(0)$. The chiral expansion is of the form

$$f_+^{K^0\pi^-}(0) = 1 + f_{p^4} + f_{e^2 p^2} + f_{p^6} + O[(m_u - m_d)p^4, e^2 p^4]\,. \tag{160}$$

The status of the various contributions is as follows:

f_{p^4}	-0.0227 (no uncertainty)	Gasser, Leutwyler [46]
$f_{e^2 p^2}$	radiative corrections	Cirigliano, Neufeld, Pichl [47]
f_{p^6}	loop contributions	Bijnens, Talavera [48]; Post, Schilcher [49]
	tree contributions	LECs L_5^2, $C_{12} + C_{34}$

As a first test, we look at the ratio

$$r_{+0} = \left(\frac{2\,\Gamma(K^+_{e3(\gamma)})\,M_{K^0}^5\,I_{K^0}}{\Gamma(K^0_{e3(\gamma)})\,M_{K^+}^5\,I_{K^+}}\right)^{1/2} = \frac{|f_+^{K^+\pi^0}(0)|}{|f_+^{K^0\pi^-}(0)|}\,. \tag{161}$$

This ratio is independent of f_{p^6} in Eq. (160) and it can therefore be predicted quite accurately [47, 50]:

$$r_{+0}^{\text{th}} = 1.023 \pm 0.003\,, \tag{162}$$

to be compared with the experimental value [51]:

$$r_{+0}^{\text{exp}} = 1.036 \pm 0.008\,. \tag{163}$$

What could be the origin of a possible discrepancy that is also suggested by the compilation of recent data in Fig. 27?

– In the past, radiative corrections have not always been state of the art. Nowadays, also experimentalists should only use the modern CHPT treatment [47].
– Measurements of the K^+ and K_L lifetimes should still be improved.
– On the theory side, an unlikely but in principle still possible explanation could be that the error due to effects of $O[(m_u - m_d)p^4, e^2 p^4]$ is underestimated.

The contribution f_{p^6} is the sum of a loop and of a tree-level part:

$$f_{p^6}^{L=1,2}(M_\rho) = 0.0093 \pm 0.0005 \qquad\qquad \text{Bijnens, Talavera [46]} \tag{164}$$

$$
\begin{aligned}
f_{p^6}^{\text{tree}}(M_\rho) &= 8\frac{(M_K^2 - M_\pi^2)^2}{F_\pi^2}\left[\frac{(L_5^r(M_\rho))^2}{F_\pi^2} - C_{12}^r(M_\rho) - C_{34}^r(M_\rho)\right] \\
&= -\frac{(M_K^2 - M_\pi^2)^2}{2\,M_S^4}\left(1 - \frac{M_S^2}{M_P^2}\right)^2\,. \qquad \begin{array}{l}\text{large-}N_c\text{ matching}\\ \text{Cirigliano }\textit{et al.}\text{ [52]}\end{array}
\end{aligned}
\tag{165}
$$

The last equation is based on a large-N_c estimate of Ref. [52]. As can be seen in Fig. 28, two terms interfere, destructively weakening the overall dependence on the scalar resonance mass M_S. The same interference leads to a rather modest scale dependence of the result for $M_\eta \le \mu \le 1$ GeV.
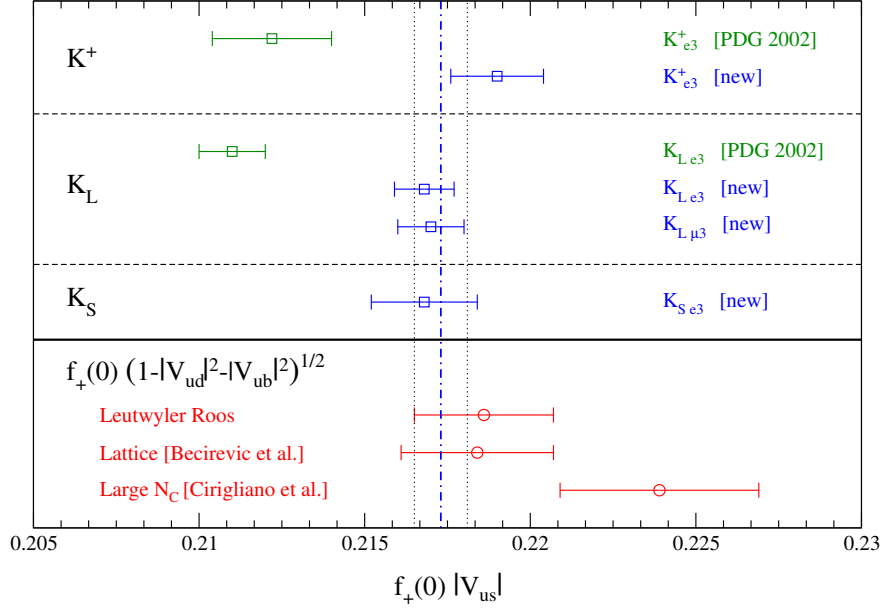
Fig. 27: Present experimental and theoretical status of $f_+(0)\,|V_{us}|$ taken from Ref. [53]
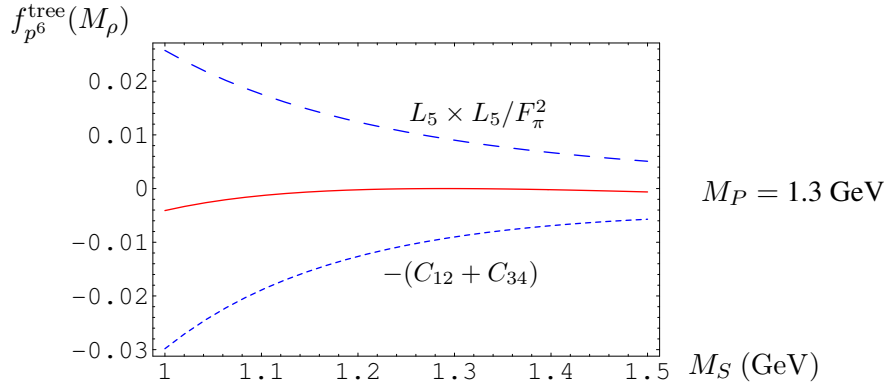
Fig. 28: Tree-level contribution of $O(p^6)$ to $f_+^{K^0\pi^-}(0)$ and its two components [52]

The final result, including the uncertainty due to a possible second pseudoscalar multiplet P', is [52]

$$
\begin{aligned}
f_{p^6}^{\mathrm{tree}}(M_\rho) &= -0.002 \pm 0.008_{\,1/N_c} \pm 0.002_{\,M_S}{}^{+0.000}_{-0.002}{}_{\,P'} \\
f_{p^6} &= 0.007 \pm 0.012 \\
f_+^{K^0\pi^-}(0) &= 0.984 \pm 0.012 \, .
\end{aligned}
\tag{166}
$$

With our large-N_c estimate of the tree-level contribution of $O(p^6)$, f_{p^6} is dominated by the loop part (164). It exhibits less $SU(3)$ breaking than the well-known result of Leutwyler and Roos [54] and a recent lattice estimate [55]. Taking the most recent value of V_{ud} and assuming unitarity of the CKM matrix, the predictions can be compared directly with the experimental results as shown in Fig. 27. A new result for the neutron lifetime [56] would imply a shift of all theoretical values in Fig. 27 to the left but the corresponding accuracy of V_{ud} is not yet competitive with super-allowed nuclear Fermi transitions. Another way to read Fig. 27 is that the estimate of Ref. [52] leads to a value of V_{us},

$$
|V_{us}| = 0.2208 \pm 0.0027_{f_+(0)} \pm 0.0008_{\mathrm{exp}} \, ,
\tag{167}
$$

that is a bit smaller than the unitarity prediction.

Finally, the slope of the scalar form factor that also depends on the LECs C_{12}, C_{34} can also be predicted [52] and it is in good agreement with the recent most precise experimental value [57]:

$$
\begin{aligned}
\lambda_0 &= (13 \pm 3) \cdot 10^{-3} &&\text{Cirigliano } \textit{et al.} \quad [52] \\
\lambda_0 &= (13.72 \pm 1.31) \cdot 10^{-3} &&\text{KTeV } [57]
\end{aligned}
\tag{168}
$$

4 Summary and epilogue

There is an amazing richness contained in the simple Lagrangian that we 'derived' from the existence of colour and from the gauge principle:

$$
\mathcal{L}_{\text{QCD}} = -\frac{1}{2}\text{tr}(G_{\mu\nu}G^{\mu\nu}) + \sum_{f=1}^{N_F} \overline{q}_f \left(i\slashed{D} - m_f \mathbb{1}_c \right) q_f \, .
$$

There is in fact no better summary for these lectures.

It may seem a long way from the naive quark model to QCD but it all happened in less than ten years. On the asymptotically free side, perturbative QCD is a complete success and it will be especially needed for understanding the background for new physics at the LHC and beyond.

There is much more left to be understood at the other end of the scale. If confinement is really forever, we would witness the first case in the history of physics when new constituents of matter have been identified beyond reasonable doubt and yet they can never be isolated. The question sounds preposterous but it is well supported: have we already arrived at the innermost structure of hadrons? Or put in a different way, is there no further structure to be expected before strings and quantum gravity eventually take over? In a short time, once the LHC produces its first results, we may learn how to rephrase the question.

Acknowledgements

References

[1] D. J. Gross, Proc. Nat. Acad. Sci. **102** (2005) 9099;
H. D. Politzer, Proc. Nat. Acad. Sci. **102** (2005) 7789;
F. Wilczek, Proc. Nat. Acad. Sci. **102** (2005) 8403 [arXiv:hep-ph/0502113].

[2] R. P. Feynman, in *The Quantum Theory of Fields - The 12th Solvay Conference* (Interscience, New York, 1961).

[3] M. Gell-Mann, Phys. Lett. **8** (1964) 214;
G. Zweig, CERN-TH-401,412 (1964) (unpublished).

[4] M. Gell-Mann, Physics **1** (1964) 63.

[5] A. Zee, Phys. Rev. D **7** (1973) 3630.

[6] D. J. Gross and F. Wilczek, Phys. Rev. Lett. **30** (1973) 1343;
H. D. Politzer, Phys. Rev. Lett. **30** (1973) 1346.

[7] N. K. Nielsen, Am. J. Phys. **49** (1981) 1171.

[8] H. Fritzsch and M. Gell-Mann, Current algebra: quarks and what else?, Proc. 16th International Conference on High-Energy Physics, Chicago, 1972, Vol. 2, J. D. Jackson, A. Roberts and R. Donaldson (Eds.) (NAL, Batavia, IL, 1973), p. 135 [arXiv:hep-ph/0208010].

[9] H. Fritzsch, M. Gell-Mann and P. Minkowski, Phys. Lett. B **59** (1975) 256.

[10] S. Kluth, Nucl. Phys. Proc. Suppl. **133** (2004) 36 [arXiv:hep-ex/0309070].

[11] V. V. Ezhela, S. B. Lugovsky and O. V. Zenin, Hadronic part of the muon g–2 estimated on the sigma(total)(2003)(e+ e- → hadrons) evaluated data compilation, arXiv:hep-ph/0312114.

[12] S. R. Coleman and D. J. Gross, Phys. Rev. Lett. **31** (1973) 851.

[13] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics* (Cambridge University Press, Cambridge, 2003).

[14] S. Eidelman *et al.* [Particle Data Group], Phys. Lett. B **592** (2004) 1.

[15] S. Bethke, Nucl. Phys. Proc. Suppl. **135** (2004) 345 [arXiv:hep-ex/0407021].

[16] E. Braaten, S. Narison and A. Pich, Nucl. Phys. B **373** (1992) 581.

[17] C. G. Callan and D. J. Gross, Phys. Rev. Lett. **22** (1969) 156.

[18] A. Kappes [ZEUS Collaboration], Structure function results from ZEUS, arXiv:hep-ex/0210032.

[19] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 438 [Yad. Fiz. **15** (1972) 781];
L. N. Lipatov, Sov. J. Nucl. Phys. **20** (1975) 94 [Yad. Fiz. **20** (1974) 181];
G. Altarelli and G. Parisi, Nucl. Phys. B **126** (1977) 298;
Y. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641 [Zh. Eksp. Teor. Fiz. **73** (1977) 1216].

[20] J. A. M. Vermaseren, A. Vogt and S. Moch, Nucl. Phys. B **724** (2005) 3 [arXiv:hep-ph/0504242].

[21] J. Goldstone, Nuovo Cimento **19** (1961) 154.

[22] A. V. Manohar and C. T. Sachrajda, Quark masses, in Ref. [14].

[23] A. V. Manohar and M. B. Wise, *Heavy quark physics*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **10** (2000) 1;
T. Mannel, *Effective field theories in flavour physics*, Springer Tracts in Modern Physics **203** (Springer, Berlin, 2004).

[24] N. Brambilla *et al.*, Heavy quarkonium physics, arXiv:hep-ph/0412158.

[25] M. Gourdin and E. de Rafael, Nucl. Phys. B **10** (1969) 667.

[26] R. Alemany, M. Davier and A. Höcker, Eur. Phys. J. C **2** (1998) 123 [arXiv:hep-ph/9703220].

[27] M. Davier, S. Eidelman, A. Höcker and Z. Zhang, Eur. Phys. J. C **31** (2003) 503 [arXiv:hep-ph/0308213].

[28] V. Cirigliano, G. Ecker and H. Neufeld, Phys. Lett. B **513**, 361 (2001) [arXiv:hep-ph/0104267]; JHEP **0208** (2002) 002 [arXiv:hep-ph/0207310].

[29] A. Höcker, The hadronic contribution to $(g - 2)_\mu$, arXiv:hep-ph/0410081.

[30] K. Maltman, Phys. Lett. B **633** (2006) 512 [arXiv:hep-ph/0504201].

[31] M. Davier and W. J. Marciano, Annu. Rev. Nucl. Part. Sci. **54** (2004) 115.

[32] G. W. Bennett *et al.* [Muon g–2 Collaboration], Phys. Rev. Lett. **92** (2004) 161802 [arXiv:hep-ex/0401008].

[33] G. Ecker, Chiral perturbation theory, Lectures given at the Topical Seminar on Frontiers of Particle Physics: QCD and Light Hadrons, Beijing, China, Sept. 2004,
`http://homepage.univie.ac.at/Gerhard.Ecker/beijingge.pdf`.

[34] J. Gasser and H. Leutwyler, Ann. Phys. **158** (1984) 142; Nucl. Phys. B **250** (1985) 465.

[35] S. Weinberg, Physica A **96** (1979) 327.

[36] R. Unterdorfer and G. Ecker, JHEP **0510** (2005) 017 [arXiv:hep-ph/0507173].

[37] H. Leutwyler, Phys. Lett. B **378** (1996) 313 [arXiv:hep-ph/9602366].

[38] G. Amoros, J. Bijnens and P. Talavera, Nucl. Phys. B **602** (2001) 87 [arXiv:hep-ph/0101127].

[39] J. Stern, Light quark masses and condensates in QCD, arXiv:hep-ph/9712438 and references therein.

[40] J. Gasser and H. Leutwyler, Phys. Lett. B **125** (1983) 325.

[41] J. Bijnens, G. Colangelo, G. Ecker, J. Gasser and M. E. Sainio, Phys. Lett. B **374** (1996) 210 [arXiv:hep-ph/9511397]; Nucl. Phys. B **508** (1997) 263 [Erratum-ibid. B **517** (1998) 639] [arXiv:hep-ph/9707291].

[42] B. Ananthanarayan, G. Colangelo, J. Gasser and H. Leutwyler, Phys. Rep. **353** (2001) 207 [arXiv:hep-ph/0005297]; G. Colangelo, J. Gasser and H. Leutwyler, Nucl. Phys. B **603** (2001) 125 [arXiv:hep-ph/0103088].

[43] S. Pislak *et al.* [BNL-E865 Collaboration], Phys. Rev. Lett. **87** (2001) 221801 [arXiv:hep-ex/0106071]; Phys. Rev. D **67** (2003) 072004 [arXiv:hep-ex/0301040].

[44] G. Colangelo, J. Gasser and H. Leutwyler, Phys. Rev. Lett. **86** (2001) 5008 [arXiv:hep-ph/0103063].

[45] I. Caprini, G. Colangelo and H. Leutwyler, Mass and width of the lowest resonance in QCD, arXiv:hep-ph/0512364.

[46] J. Gasser and H. Leutwyler, Nucl. Phys. B **250** (1985) 517.

[47] V. Cirigliano, H. Neufeld and H. Pichl, Eur. Phys. J. C **35** (2004) 53 [arXiv:hep-ph/0401173].

[48] J. Bijnens and P. Talavera, Nucl. Phys. B **669** (2003) 341 [arXiv:hep-ph/0303103].

[49] P. Post and K. Schilcher, Eur. Phys. J. C **25** (2002) 427 [arXiv:hep-ph/0112352].

[50] S. Descotes-Genon and B. Moussallam, Eur. Phys. J. C **42** (2005) 403 [arXiv:hep-ph/0505077].

[51] H. Neufeld, private communication.

[52] V. Cirigliano, G. Ecker, M. Eidemüller, R. Kaiser, A. Pich and J. Portolés, JHEP **0504** (2005) 006 [arXiv:hep-ph/0503108].

[53] E. Blucher *et al.*, Status of the Cabibbo angle (CKM2005 - WG 1), arXiv:hep-ph/0512039.

[54] H. Leutwyler and M. Roos, Z. Phys. C **25** (1984) 91.

[55] D. Becirevic *et al.*, Nucl. Phys. B **705** (2005) 339 [arXiv:hep-ph/0403217].

[56] A. Serebrov *et al.*, Phys. Lett. B **605** (2005) 72 [arXiv:nucl-ex/0408009].

[57] T. Alexopoulos *et al.* [KTeV Collaboration], Phys. Rev. D **70** (2004) 092007 [arXiv:hep-ex/0406003].

[58] A. Pich, Aspects of quantum chromodynamics, arXiv:hep-ph/0001118.

Bibliography

M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, MA, 1995).

J. F. Donoghue, E. Golowich and B. R. Holstein, *Dynamics of the Standard Model* (Cambridge University Press, Cambridge, 1994).

M. H. Seymour, *Quantum Chromodynamics*, Proc. 2004 European School of High-Energy Physics, Sant Feliu de Guíxols, Spain, Ed. R. Fleischer, CERN 2006-003, p. 49. arXiv:hep-ph/0505192.

A. Khodjamirian, *QCD and Hadrons: an Elementary Introduction*, Proc. 2003 European School of High-Energy Physics, Tsakhkadzor, Armenia, Ed. A. Olchevski, CERN 2005-007, p. 173. arXiv:hep-ph/0403145.

Yu. L. Dokshitzer, *QCD Phenomenology*, Proc. 2002 European School of High-Energy Physics, Pylos, Greece, Eds. N. Ellis and R. Fleischer, CERN 2004-001, p. 1. http://doc.cern.ch/yellowrep/2004/2004-001/p1.pdf.

Flavour physics and CP violation

R. Fleischer
CERN, Geneva, Switzerland

Abstract

The starting point of these lectures is an introduction to the weak interactions of quarks and the Standard-Model description of CP violation, where the central role is played by the Cabibbo–Kobayashi–Maskawa matrix and the corresponding unitarity triangles. Since the B-meson system will govern the stage of (quark) flavour physics and CP violation in this decade, it will be our main focus. We shall classify B-meson decays, introduce the theoretical tools to deal with them, investigate the requirements for non-vanishing CP-violating asymmetries, and discuss the main strategies to explore CP violation and the preferred avenues to enter for physics beyond the Standard Model. This formalism is then applied to discuss the status of important B-factory benchmark modes, where we focus on puzzling patterns in the data that may indicate new-physics effects, as well as the prospects for B-decay studies at the LHC.

1 Introduction

The history of CP violation, i.e., the non-invariance of the weak interactions with respect to a combined charge-conjugation (C) and parity (P) transformation, goes back to 1964 when this phenomenon was discovered through the observation of $K_{\mathrm{L}} \to \pi^+\pi^-$ decays [1] which exhibit a branching ratio at the 10^{-3} level. This surprising effect is a manifestation of *indirect* CP violation, which arises from the fact that the mass eigenstates $K_{\mathrm{L,S}}$ of the neutral kaon system, which shows K^0–$\bar{K}^0$ mixing, are not eigenstates of the CP operator. In particular, the K_{L} state is governed by the CP-odd eigenstate, but has also a tiny admixture of the CP-even eigenstate, which may decay through CP-conserving interactions into the $\pi^+\pi^-$ final state. These CP-violating effects are described by the following observable:

$$\varepsilon_K = (2.280 \pm 0.013) \times 10^{-3} \times e^{i\pi/4} \, . \tag{1}$$

On the other hand, CP-violating effects may also arise directly at the decay-amplitude level, thereby yielding *direct* CP violation. This phenomenon, which leads to a non-vanishing value of a quantity $\mathrm{Re}(\varepsilon'_K/\varepsilon_K)$, was established in 1999 through the NA48 (CERN) and KTeV (FNAL) Collaborations [2]; the final results of the corresponding measurements are given by

$$\mathrm{Re}(\varepsilon'_K/\varepsilon_K) = \left\{ \begin{array}{ll} (14.7 \pm 2.2) \times 10^{-4} & \text{(NA48 [3])} \\ (20.7 \pm 2.8) \times 10^{-4} & \text{(KTeV [4])} \, . \end{array} \right. \tag{2}$$

In this decade, there are huge experimental efforts to further explore CP violation and the quark-flavour sector of the Standard Model (SM). In these studies, the main actor is the B-meson system, where we distinguish between charged and neutral B mesons, which are characterized by the following valence-quark contents:

$$\begin{array}{lllll} B^+ \sim u\bar{b}, & B_c^+ \sim c\bar{b}, & B_d^0 \sim d\bar{b}, & B_s^0 \sim s\bar{b}, \\ B^- \sim \bar{u}b, & B_c^- \sim \bar{c}b, & \bar{B}_d^0 \sim \bar{d}b, & \bar{B}_s^0 \sim \bar{s}b \, . \end{array} \tag{3}$$

In contrast to the charged B mesons, their neutral counterparts B_q ($q \in \{d, s\}$) show —in analogy to K^0–$\bar{K}^0$ mixing—the phenomenon of B_q^0–$\bar{B}_q^0$ mixing. The asymmetric e^+e^- B factories at SLAC and KEK with their detectors BaBar and Belle, respectively, can only produce B^+ and B_d^0 mesons (and

their anti-particles) since they operate at the $\Upsilon(4S)$ resonance, and have already collected $\mathcal{O}(10^8)$ $B\bar{B}$ pairs of this kind. Moreover, first B-physics results from Run II of the Tevatron were reported from the CDF and D0 Collaborations, including also B_c^+ and B_s^0 studies, and second-generation B-decay studies will become possible at the Large Hadron Collider (LHC) at CERN, in particular thanks to the LHCb experiment, starting in the autumn of 2007. For the more distant future, an e^+–e^- 'super-B factory' is under consideration, with an increase of luminosity by up to two orders of magnitude with respect to the currently operating machines. Moreover, there are plans to measure the very 'rare' kaon decays $K^+ \to \pi^+ \nu\bar{\nu}$ and $K_{\mathrm{L}} \to \pi^0 \nu\bar{\nu}$, which are absent at the tree level in the Standard Model (SM), at CERN and KEK/J-PARC.

In 2001, CP-violating effects were discovered in B decays with the help of $B_d \to J/\psi K_{\mathrm{S}}$ modes by the BaBar and Belle Collaborations [5], representing the first observation of CP violation outside the kaon system. This particular kind of CP violation, which is by now well established, originates from the interference between B_d^0–$\bar{B}_d^0$ mixing and $B_d^0 \to J/\psi K_{\mathrm{S}}$, $\bar{B}_d^0 \to J/\psi K_{\mathrm{S}}$ decay processes, and is referred to as 'mixing-induced' CP violation. In the summer of 2004, direct CP violation was detected in $B_d \to \pi^\mp K^\pm$ decays [6], thereby complementing the measurement of a non-zero value of $\mathrm{Re}(\varepsilon_K'/\varepsilon_K)$.

Studies of CP violation and flavour physics are particularly interesting since 'new physics' (NP), i.e., physics lying beyond the SM, typically leads to new sources of flavour and CP violation. Furthermore, the origin of the fermion masses, flavour mixing, CP violation etc. lies completely in the dark and is expected to involve NP, too. Interestingly, CP violation offers also a link to cosmology. One of the key features of our Universe is the cosmological baryon asymmetry of $\mathcal{O}(10^{-10})$. As was pointed out by Sakharov [7], the necessary conditions for the generation of such an asymmetry include also the requirement that elementary interactions violate CP (and C). Model calculations of the baryon asymmetry indicate, however, that the CP violation present in the SM seems to be too small to generate the observed asymmetry [8]. On the one hand, the required new sources of CP violation could be associated with very high energy scales, as in 'leptogenesis', where new CP-violating effects appear in decays of heavy Majorana neutrinos [9]. On the other hand, new sources of CP violation could also be accessible in the laboratory, as they arise naturally when going beyond the SM.

Before searching for NP, it is essential to understand first the picture of flavour physics and CP violation arising in the framework of the SM, where the Cabibbo–Kobayashi–Maskawa (CKM) matrix—the quark-mixing matrix—plays the key role [10, 11]. The corresponding phenomenology is extremely rich [12]. In general, the key problem for the theoretical interpretation is related to strong interactions, i.e., to 'hadronic' uncertainties. A famous example is $\mathrm{Re}(\varepsilon_K'/\varepsilon_K)$, where we have to deal with a subtle interplay between different contributions which largely cancel [13]. Although the non-vanishing value of this quantity has unambiguously ruled out 'superweak' models of CP violation [14], it currently does not allow a stringent test of the SM.

In the B-meson system, there are various strategies to eliminate the hadronic uncertainties in the exploration of CP violation (simply speaking, there are many B decays). Moreover, we may also search for relations and/or correlations that hold in the SM but could well be spoiled by NP. These topics will be the focus of this lecture, which is complemented by the dedicated lectures on the experimental aspects of K- and B-meson decays in Refs. [15] and [16], respectively. The outline is as follows: in Section 2, we discuss the quark mixing in the SM by having a closer look at the CKM matrix and the associated unitarity triangles. The main actors of this lecture—the B mesons and their weak decays—will then be introduced in Section 3. There we shall also move towards studies of CP violation and shall classify the main strategies for its exploration, using amplitude relations and the phenomenon of B_q^0–$\bar{B}_q^0$ mixing ($q \in \{d, s\}$). In Section 4, we illustrate the former kind of methods by having a closer look at clean amplitude relations between $B^\pm \to K^\pm D$ and $B_c^\pm \to D_s^\pm D$ decays, whereas we discuss features of neutral B_q mesons in Section 5. In Section 6, we address the question of how NP could enter, and then apply these considerations in Section 7 to the B-factory benchmark modes $B_d^0 \to J/\psi K_{\mathrm{S}}$, $B_d^0 \to \phi K_{\mathrm{S}}$ and $B_d^0 \to \pi^+ \pi^-$. Since the data for certain $B \to \pi K$ decays have shown a puzzling pattern for several
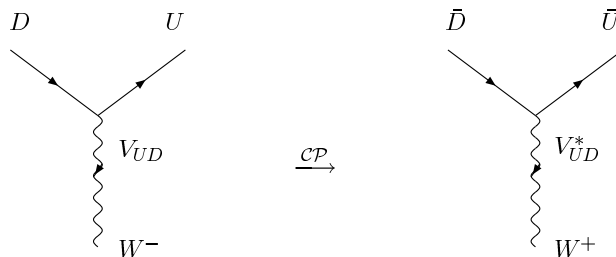
Fig. 1: CP-conjugate charged-current quark-level interaction processes in the SM

years, we have devoted Section 8 to a detailed discussion of this '$B \to \pi K$ puzzle' and its interplay with rare K and B decays. In Section 9, we focus on $b \to d$ penguin processes, which are now coming within experimental reach at the B factories, thereby offering an exciting new playground. Finally, in Section 10, we discuss B-decay studies at the LHC, where the physics potential of the B_s^0-meson system can be fully exploited. The conclusions and a brief outlook are given in Section 11.

For detailed discussions and textbooks dealing with flavour physics and CP violation, the reader is referred to Refs. [17]– [21], alternative lecture notes can be found in Refs. [22, 23], and a selection of more compact recent reviews is given in Refs. [24]– [26]. The data used in these lectures refer to the situation in the spring of 2006.

2 CP violation in the Standard Model

2.1 Weak interactions of quarks and the quark-mixing matrix

In the framework of the Standard Model of electroweak interactions [27, 28], which is based on the spontaneously broken gauge group

$$SU(2)_{\mathrm{L}} \times U(1)_{\mathrm{Y}} \xrightarrow{\mathrm{SSB}} U(1)_{\mathrm{em}} , \tag{4}$$

CP-violating effects may originate from the charged-current interactions of quarks, having the structure

$$D \to U W^- . \tag{5}$$

Here $D \in \{d, s, b\}$ and $U \in \{u, c, t\}$ denote down- and up-type quark flavours, respectively, whereas the W^- is the usual $SU(2)_{\mathrm{L}}$ gauge boson. From a phenomenological point of view, it is convenient to collect the generic 'coupling strengths' V_{UD} of the charged-current processes in (5) in the form of the following matrix:

$$\hat{V}_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} , \tag{6}$$

which is referred to as the Cabibbo–Kobayashi–Maskawa (CKM) matrix [10, 11].

From a theoretical point of view, this matrix connects the electroweak states (d', s', b') of the down, strange and bottom quarks with their mass eigenstates (d, s, b) through the following unitary transformation [27]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{7}$$

Consequently, $\hat{V}_{\mathrm{CKM}}$ is actually a *unitary* matrix. This feature ensures the absence of flavour-changing neutral-current (FCNC) processes at the tree level in the SM, and is hence at the basis of the famous

Glashow–Iliopoulos–Maiani (GIM) mechanism [29]. We shall return to the unitarity of the CKM matrix in Section 2.6, discussing the 'unitarity triangles'. If we express the non-leptonic charged-current interaction Lagrangian in terms of the mass eigenstates appearing in (7), we arrive at

$$\mathcal{L}_{\text{int}}^{\text{CC}} = -\frac{g_2}{\sqrt{2}} \left(\begin{array}{ccc} \bar{u}_{\text{L}}, & \bar{c}_{\text{L}}, & \bar{t}_{\text{L}} \end{array} \right) \gamma^\mu \hat{V}_{\text{CKM}} \left(\begin{array}{c} d_{\text{L}} \\ s_{\text{L}} \\ b_{\text{L}} \end{array} \right) W_\mu^\dagger + \text{h.c.,} \tag{8}$$

where the gauge coupling g_2 is related to the gauge group $SU(2)_{\text{L}}$, and the $W_\mu^{(\dagger)}$ field corresponds to the charged W bosons. Looking at the interaction vertices following from (8), we observe that the elements of the CKM matrix describe in fact the generic strengths of the associated charged-current processes, as we have noted above.

In Fig. 1, we show the $D \to UW^-$ vertex and its CP conjugate. Since the corresponding CP transformation involves the replacement

$$V_{UD} \xrightarrow{\mathcal{CP}} V_{UD}^* , \tag{9}$$

CP violation could—in principle—be accommodated in the SM through complex phases in the CKM matrix. The crucial question in this context is, of course, whether we may actually have physical complex phases in that matrix.

2.2 Phase structure of the CKM matrix

We have the freedom to redefine the up- and down-type quark fields in the following manner:

$$U \to \exp(i\xi_U)U, \quad D \to \exp(i\xi_D)D . \tag{10}$$

If we perform such transformations in (8), the invariance of the charged-current interaction Lagrangian implies the following phase transformations of the CKM matrix elements:

$$V_{UD} \to \exp(i\xi_U)V_{UD}\exp(-i\xi_D) . \tag{11}$$

Using these transformations to eliminate unphysical phases, it can be shown that the parametrization of the general $N \times N$ quark-mixing matrix, where N denotes the number of fermion generations, involves the following parameters:

$$\underbrace{\frac{1}{2}N(N-1)}_{\text{Euler angles}} + \underbrace{\frac{1}{2}(N-1)(N-2)}_{\text{complex phases}} = (N-1)^2. \tag{12}$$

If we apply this expression to the case of $N = 2$ generations, we observe that only one rotation angle—the Cabibbo angle θ_{C} [10]—is required for the parametrization of the 2×2 quark-mixing matrix, which can be written in the following form:

$$\hat{V}_{\text{C}} = \left(\begin{array}{cc} \cos\theta_{\text{C}} & \sin\theta_{\text{C}} \\ -\sin\theta_{\text{C}} & \cos\theta_{\text{C}} \end{array} \right), \tag{13}$$

where $\sin\theta_{\text{C}} = 0.22$ can be determined from $K \to \pi\ell\bar{\nu}$ decays. On the other hand, in the case of $N = 3$ generations, the parametrization of the corresponding 3×3 quark-mixing matrix involves three Euler-type angles and a single *complex* phase. This complex phase allows us to accommodate CP violation in the SM, as was pointed out by Kobayashi and Maskawa in 1973 [11]. The corresponding picture is referred to as the Kobayashi–Maskawa (KM) mechanism of CP violation.

In the 'standard parametrization' advocated by the Particle Data Group (PDG) [30], the three-generation CKM matrix takes the following form:

$$\hat{V}_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \qquad (14)$$

where $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$. Performing appropriate redefinitions of the quark-field phases, the real angles θ_{12}, θ_{23} and θ_{13} can all be made to lie in the first quadrant. The advantage of this parametrization is that the generation labels $i, j = 1, 2, 3$ are introduced in such a manner that the mixing between two chosen generations vanishes if the corresponding mixing angle θ_{ij} is set to zero. In particular, for $\theta_{23} = \theta_{13} = 0$, the third generation decouples, and the 2×2 submatrix describing the mixing between the first and second generations takes the same form as (13).

Another interesting parametrization of the CKM matrix was proposed by Fritzsch and Xing [31]:

$$\hat{V}_{\text{CKM}} = \begin{pmatrix} s_{\text{u}}s_{\text{d}}c + c_{\text{u}}c_{\text{d}}e^{-i\varphi} & s_{\text{u}}c_{\text{d}}c - c_{\text{u}}s_{\text{d}}e^{-i\varphi} & s_{\text{u}}s \\ c_{\text{u}}s_{\text{d}}c - s_{\text{u}}c_{\text{d}}e^{-i\varphi} & c_{\text{u}}c_{\text{d}}c + s_{\text{u}}s_{\text{d}}e^{-i\varphi} & c_{\text{u}}s \\ -s_{\text{d}}s & -c_{\text{d}}s & c \end{pmatrix}. \qquad (15)$$

It is inspired by the hierarchical structure of the quark-mass spectrum and is particularly useful in the context of models for fermion masses and mixings. The characteristic feature of this parametrization is that the complex phase arises only in the 2×2 submatrix involving the up, down, strange and charm quarks.

Let us finally note that physical observables, for instance CP-violating asymmetries, *cannot* depend on the chosen parametrization of the CKM matrix, i.e., have to be invariant under the phase transformations specified in (11).

2.3   Further requirements for CP violation

As we have just seen, in order to be able to accommodate CP violation within the framework of the SM through a complex phase in the CKM matrix, at least three generations are required. However, this feature is not sufficient for observable CP-violating effects. To this end, further conditions have to be satisfied, which can be summarized as follows [32, 33]:

$$(m_t^2 - m_c^2)(m_t^2 - m_u^2)(m_c^2 - m_u^2)(m_b^2 - m_s^2)(m_b^2 - m_d^2)(m_s^2 - m_d^2) \times J_{\text{CP}} \neq 0, \qquad (16)$$

where

$$J_{\text{CP}} = |\text{Im}(V_{i\alpha}V_{j\beta}V_{i\beta}^*V_{j\alpha}^*)| \quad (i \neq j,\ \alpha \neq \beta). \qquad (17)$$

The mass factors in (16) are related to the fact that the CP-violating phase of the CKM matrix could be eliminated through an appropriate unitary transformation of the quark fields if any two quarks with the same charge had the same mass. Consequently, the origin of CP violation is closely related to the 'flavour problem' in elementary particle physics, and cannot be understood in a deeper way, unless we have fundamental insights into the hierarchy of quark masses and the number of fermion generations.

The second element of (16), the 'Jarlskog parameter' J_{CP} [32], can be interpreted as a measure of the strength of CP violation in the SM. It does not depend on the chosen quark-field parametrization, i.e., it is invariant under (11), and the unitarity of the CKM matrix implies that all combinations $|\text{Im}(V_{i\alpha}V_{j\beta}V_{i\beta}^*V_{j\alpha}^*)|$ are equal to one another. Using the standard parametrization of the CKM matrix introduced in (14), we obtain

$$J_{\text{CP}} = s_{12}s_{13}s_{23}c_{12}c_{23}c_{13}^2 \sin\delta_{13} . \qquad (18)$$

The experimental information on the CKM parameters implies $J_{\text{CP}} = \mathcal{O}(10^{-5})$, so that CP-violating phenomena are hard to observe. However, new complex couplings are typically present in scenarios for NP [34]. Such additional sources for CP violation could be detected through flavour experiments.
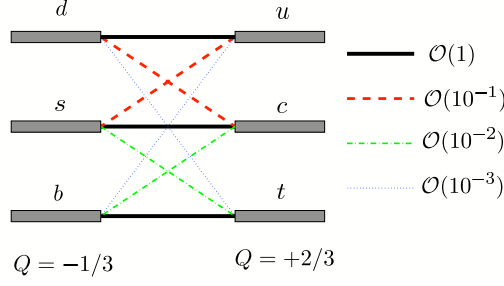
Fig. 2: Hierarchy of the quark transitions mediated through charged-current processes

2.4 Experimental information on $|V_{\mathrm{CKM}}|$

In order to determine the magnitudes $|V_{ij}|$ of the elements of the CKM matrix, we may use the following tree-level processes:

- Nuclear beta decays, neutron decays $\Rightarrow |V_{ud}|$.
- $K \to \pi \ell \bar{\nu}$ decays $\Rightarrow |V_{us}|$.
- ν production of charm off valence d quarks $\Rightarrow |V_{cd}|$.
- Charm-tagged W decays (as well as ν production and semileptonic D decays) $\Rightarrow |V_{cs}|$.
- Exclusive and inclusive $b \to c \ell \bar{\nu}$ decays $\Rightarrow |V_{cb}|$.
- Exclusive and inclusive $b \to u \ell \bar{\nu}$ decays $\Rightarrow |V_{ub}|$.
- $\bar{t} \to \bar{b} \ell \bar{\nu}$ processes $\Rightarrow$ (crude direct determination of) $|V_{tb}|$.

If we use the corresponding experimental information, together with the CKM unitarity condition, and assume that there are only three generations, we arrive at the following 90% C.L. limits for the $|V_{ij}|$ [30]:

$$|\hat{V}_{\mathrm{CKM}}| = \begin{pmatrix} 0.9739\text{–}0.9751 & 0.221\text{–}0.227 & 0.0029\text{–}0.0045 \\ 0.221\text{–}0.227 & 0.9730\text{–}0.9744 & 0.039\text{–}0.044 \\ 0.0048\text{–}0.014 & 0.037\text{–}0.043 & 0.9990\text{–}0.9992 \end{pmatrix} . \tag{19}$$

In Fig. 2, we have illustrated the resulting hierarchy of the strengths of the charged-current quark-level processes: transitions within the same generation are governed by CKM matrix elements of $\mathcal{O}(1)$, those between the first and the second generation are suppressed by CKM factors of $\mathcal{O}(10^{-1})$, those between the second and the third generation are suppressed by $\mathcal{O}(10^{-2})$, and the transitions between the first and the third generation are even suppressed by CKM factors of $\mathcal{O}(10^{-3})$. In the standard parametrization (14), this hierarchy is reflected by

$$s_{12} = 0.22 \gg s_{23} = \mathcal{O}(10^{-2}) \gg , s_{13} = \mathcal{O}(10^{-3}) . \tag{20}$$

2.5 Wolfenstein parametrization of the CKM matrix

For phenomenological applications, it would be useful to have a parametrization of the CKM matrix available that makes the hierarchy arising in (19)—and illustrated in Fig. 2—explicit [35]. In order to derive such a parametrization, we introduce a set of new parameters, λ, A, ρ and η, by imposing the following relations [36]:

$$s_{12} \equiv \lambda = 0.22, \quad s_{23} \equiv A\lambda^2, \quad s_{13} e^{-i\delta_{13}} \equiv A\lambda^3(\rho - i\eta) . \tag{21}$$

If we now go back to the standard parametrization (14), we obtain an *exact* parametrization of the CKM matrix as a function of λ (and A, ρ, η), allowing us to expand each CKM element in powers of the small

parameter λ. If we neglect terms of $\mathcal{O}(\lambda^4)$, we arrive at the famous 'Wolfenstein parametrization' [35]:

$$\hat{V}_{\text{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \,, \tag{22}$$

which makes the hierarchical structure of the CKM matrix very transparent and is an important tool for phenomenological considerations, as we shall see throughout these lectures.

For several applications, next-to-leading order corrections in λ play an important role. Using the exact parametrization following from (14) and (21), they can be calculated straightforwardly by expanding each CKM element to the desired accuracy in λ [36, 37]:

$$V_{ud} = 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 + \mathcal{O}(\lambda^6) \,, \quad V_{us} = \lambda + \mathcal{O}(\lambda^7) \,, \quad V_{ub} = A\lambda^3(\rho - i\,\eta) \,,$$

$$V_{cd} = -\lambda + \frac{1}{2}A^2\lambda^5\left[1 - 2(\rho + i\eta)\right] + \mathcal{O}(\lambda^7) \,,$$

$$V_{cs} = 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) + \mathcal{O}(\lambda^6) \,, \tag{23}$$

$$V_{cb} = A\lambda^2 + \mathcal{O}(\lambda^8) \,, \quad V_{td} = A\lambda^3\left[1 - (\rho + i\eta)\left(1 - \frac{1}{2}\lambda^2\right)\right] + \mathcal{O}(\lambda^7) \,,$$

$$V_{ts} = -A\lambda^2 + \frac{1}{2}A(1 - 2\rho)\lambda^4 - i\eta A\lambda^4 + \mathcal{O}(\lambda^6) \,, \quad V_{tb} = 1 - \frac{1}{2}A^2\lambda^4 + \mathcal{O}(\lambda^6) \,.$$

It should be noted that

$$V_{ub} \equiv A\lambda^3(\rho - i\eta) \tag{24}$$

receives *by definition* no power corrections in λ within this prescription. If we follow Ref. [36] and introduce the generalized Wolfenstein parameters

$$\bar{\rho} \equiv \rho\left(1 - \frac{1}{2}\lambda^2\right) \,, \quad \bar{\eta} \equiv \eta\left(1 - \frac{1}{2}\lambda^2\right) \,, \tag{25}$$

we may simply write, up to corrections of $\mathcal{O}(\lambda^7)$,

$$V_{td} = A\lambda^3(1 - \bar{\rho} - i\,\bar{\eta}) \,. \tag{26}$$

Moreover, we have to an excellent accuracy

$$V_{us} = \lambda \quad \text{and} \quad V_{cb} = A\lambda^2 \,, \tag{27}$$

as these quantities receive only corrections at the λ^7 and λ^8 levels, respectively. In comparison with other generalizations of the Wolfenstein parametrization found in the literature, the advantage of (23) is the absence of relevant corrections to V_{us} and V_{cb}, and that V_{ub} and V_{td} take forms similar to those in (22). As far as the Jarlskog parameter introduced in (17) is concerned, we obtain the simple expression

$$J_{\text{CP}} = \lambda^6 A^2 \eta \,, \tag{28}$$

which should be compared with (18).

2.6 Unitarity triangles of the CKM matrix

The unitarity of the CKM matrix, which is described by

$$\hat{V}_{\text{CKM}}^{\dagger} \cdot \hat{V}_{\text{CKM}} = \hat{1} = \hat{V}_{\text{CKM}} \cdot \hat{V}_{\text{CKM}}^{\dagger} , \tag{29}$$

leads to a set of 12 equations, consisting of 6 normalization and 6 orthogonality relations. The latter can be represented as 6 triangles in the complex plane [38], all having the same area, $2A_{\Delta} = J_{\text{CP}}$ [39]. Let us now have a closer look at these relations: those describing the orthogonality of different columns of the CKM matrix are given by

$$\underbrace{V_{ud}V_{us}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{cd}V_{cs}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{td}V_{ts}^*}_{\mathcal{O}(\lambda^5)} = 0 \tag{30}$$

$$\underbrace{V_{us}V_{ub}^*}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}V_{cb}^*}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{ts}V_{tb}^*}_{\mathcal{O}(\lambda^2)} = 0 \tag{31}$$

$$\underbrace{V_{ud}V_{ub}^*}_{(\rho+i\eta)A\lambda^3} + \underbrace{V_{cd}V_{cb}^*}_{-A\lambda^3} + \underbrace{V_{td}V_{tb}^*}_{(1-\rho-i\eta)A\lambda^3} = 0 , \tag{32}$$

whereas those associated with the orthogonality of different rows take the following form:

$$\underbrace{V_{ud}^*V_{cd}}_{\mathcal{O}(\lambda)} + \underbrace{V_{us}^*V_{cs}}_{\mathcal{O}(\lambda)} + \underbrace{V_{ub}^*V_{cb}}_{\mathcal{O}(\lambda^5)} = 0 \tag{33}$$

$$\underbrace{V_{cd}^*V_{td}}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}^*V_{ts}}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{cb}^*V_{tb}}_{\mathcal{O}(\lambda^2)} = 0 \tag{34}$$

$$\underbrace{V_{ud}^*V_{td}}_{(1-\rho-i\eta)A\lambda^3} + \underbrace{V_{us}^*V_{ts}}_{-A\lambda^3} + \underbrace{V_{ub}^*V_{tb}}_{(\rho+i\eta)A\lambda^3} = 0 . \tag{35}$$

Here we have also indicated the structures that arise if we apply the Wolfenstein parametrization by keeping just the leading, non-vanishing terms. We observe that only in (32) and (35), which describe the orthogonality of the first and third columns and of the first and third rows, respectively, are all three sides of comparable magnitude, $\mathcal{O}(\lambda^3)$, while in the remaining relations, one side is suppressed with respect to the others by factors of $\mathcal{O}(\lambda^2)$ or $\mathcal{O}(\lambda^4)$. Consequently, we have to deal with only *two* non-squashed unitarity triangles in the complex plane. However, as we have already indicated in (32) and (35), the corresponding orthogonality relations agree with each other at the λ^3 level, yielding

$$[(\rho + i\eta) + (-1) + (1 - \rho - i\eta)] A\lambda^3 = 0 . \tag{36}$$

Consequently, they describe the same triangle, which is usually referred to as *the* unitarity triangle of the CKM matrix [39, 40].

Concerning second-generation B-decay studies in the LHC era, the experimental accuracy will be so tremendous that we shall also have to take the next-to-leading order terms of the Wolfenstein expansion into account, and shall have to distinguish between the unitarity triangles following from (32) and (35). Let us first have a closer look at the former relation. Including terms of $\mathcal{O}(\lambda^5)$, we obtain the following generalization of (36):

$$[(\bar{\rho} + i\bar{\eta}) + (-1) + (1 - \bar{\rho} - i\bar{\eta})] A\lambda^3 + \mathcal{O}(\lambda^7) = 0 , \tag{37}$$

where $\bar{\rho}$ and $\bar{\eta}$ are as defined in (25). If we divide this relation by the overall normalization factor $A\lambda^3$, and introduce

$$R_b \equiv \sqrt{\bar{\rho}^2 + \bar{\eta}^2} = \left(1 - \frac{\lambda^2}{2}\right) \frac{1}{\lambda} \left|\frac{V_{ub}}{V_{cb}}\right| \tag{38}$$
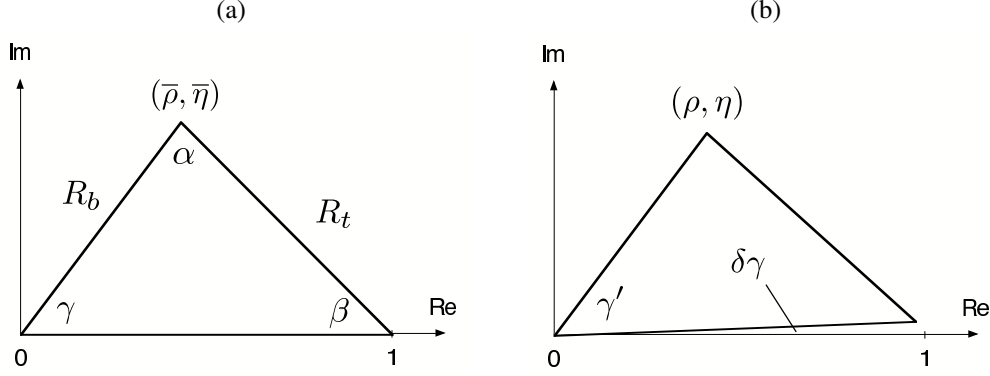
Fig. 3: The two non-squashed unitarity triangles of the CKM matrix, as explained in the text: (a) and (b) correspond to the orthogonality relations (32) and (35), respectively. In Asia, the notation $\phi_1 \equiv \beta$, $\phi_2 \equiv \alpha$, and $\phi_3 \equiv \gamma$ is used for the angles of the triangle shown in (a).

$$R_t \equiv \sqrt{(1 - \bar{\rho})^2 + \bar{\eta}^2} = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right|, \tag{39}$$

we arrive at the unitarity triangle illustrated in Fig. 3 (a). It is a straightforward generalization of the leading-order case described by (36): instead of (ρ, η), the apex is now simply given by $(\bar{\rho}, \bar{\eta})$ [36]. The two sides R_b and R_t, as well as the three angles α, β and γ, will show up at several places throughout these lectures. Moreover, the relations

$$V_{ub} = A\lambda^3\left(\frac{R_b}{1 - \lambda^2/2}\right)e^{-i\gamma}, \quad V_{td} = A\lambda^3 R_t e^{-i\beta} \tag{40}$$

are also useful for phenomenological applications, since they make the dependences of γ and β explicit; they correspond to the phase convention chosen both in the standard parametrization (14) and in the generalized Wolfenstein parametrization (23). Finally, if we take also (21) into account, we obtain

$$\delta_{13} = \gamma. \tag{41}$$

Let us now turn to (35). Here we arrive at an expression that is more complicated than (37):

$$\left[\left\{1 - \frac{\lambda^2}{2} - (1 - \lambda^2)\rho - i(1 - \lambda^2)\eta\right\} + \left\{-1 + \left(\frac{1}{2} - \rho\right)\lambda^2 - i\eta\lambda^2\right\} + \{\rho + i\eta\}\right]A\lambda^3 + \mathcal{O}(\lambda^7) = 0. \tag{42}$$

If we divide again by $A\lambda^3$, we obtain the unitarity triangle sketched in Fig. 3 (b), where the apex is given by (ρ, η) and *not* by $(\bar{\rho}, \bar{\eta})$. On the other hand, we encounter a tiny angle

$$\delta\gamma \equiv \lambda^2\eta = \mathcal{O}(1°) \tag{43}$$

between real axis and basis of the triangle, which satisfies

$$\gamma = \gamma' + \delta\gamma, \tag{44}$$

where γ coincides with the corresponding angle in Fig. 3 (a).

Whenever we refer to a 'unitarity triangle' (UT) in the following discussion, we mean the one illustrated in Fig. 3 (a), which is the generic generalization of the leading-order case described by (36). As we shall see below, the UT is the central target of the experimental tests of the SM description of CP violation. Interestingly, the tiny angle $\delta\gamma$ also can be probed directly through certain CP-violating effects that can be explored at hadron colliders, in particular at the LHC.
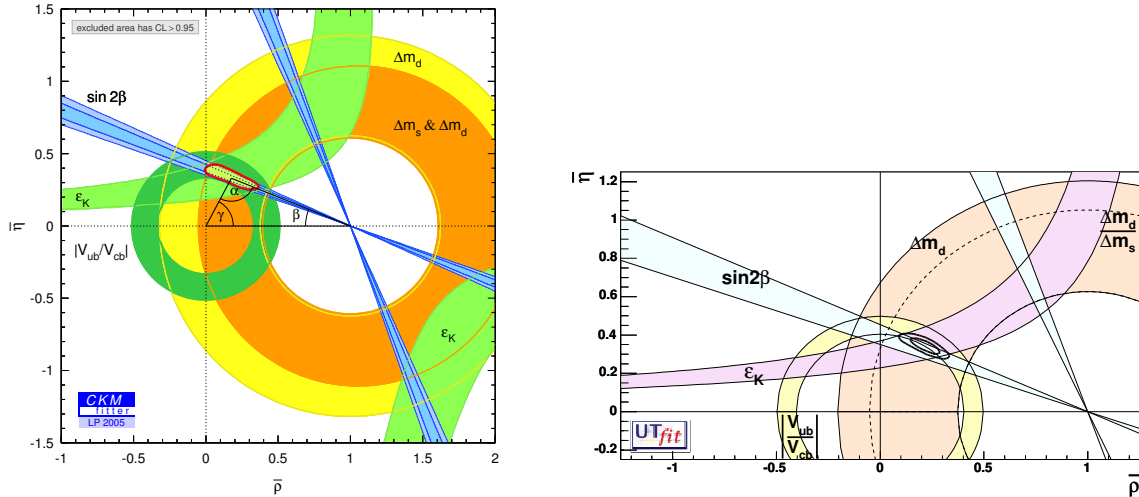
Fig. 4: Analyses of the CKMfitter and UTfit Collaborations [41,42]

2.7 The determination of the unitarity triangle

The next obvious question is how to determine the UT. There are two conceptually different avenues that we may follow:

(i) In the 'CKM fits', theory is used to convert experimental data into contours in the $\bar{\rho}$–$\bar{\eta}$ plane. In particular, semi-leptonic $b \to u\ell\bar{\nu}_\ell$, $c\ell\bar{\nu}_\ell$ decays and B_q^0–$\bar{B}_q^0$ mixing ($q \in \{d, s\}$) allow us to determine the UT sides R_b and R_t, respectively, i.e., to fix two circles in the $\bar{\rho}$–$\bar{\eta}$ plane. Furthermore, the indirect CP violation in the neutral kaon system described by ε_K can be transformed into a hyperbola.

(ii) Theoretical considerations allow us to convert measurements of CP-violating effects in B-meson decays into direct information on the UT angles. The most prominent example is the determination of $\sin 2\beta$ through CP violation in $B_d^0 \to J/\psi K_S$ decays, but several other strategies were proposed.

The goal is to 'overconstrain' the UT as much as possible. In the future, additional contours can be fixed in the $\bar{\rho}$–$\bar{\eta}$ plane through the measurement of rare decays.

In Fig. 4, we show examples of the comprehensive analyses of the UT that are performed (and continuously updated) by the 'CKM Fitter Group' [41] and the 'UTfit Collaboration' [42]. In these figures, we can nicely see the circles that are determined through the semi-leptonic B decays and the ε_K hyperbolas. Moreover, the straight lines following from the direct measurement of $\sin 2\beta$ with the help of $B_d^0 \to J/\psi K_S$ modes are also shown. We observe that the global consistency is very good. However, looking closer, we also see that the most recent average for $(\sin 2\beta)_{\psi K_S}$ is now on the lower side, so that the situation in the $\bar{\rho}$–$\bar{\eta}$ plane is no longer 'perfect'. As we shall discuss in detail in the course of these lectures, there are certain puzzles in the B-factory data, and several important aspects have not yet been addressed experimentally and are hence still essentially unexplored. Consequently, we may hope that flavour studies will eventually establish deviations from the SM description of CP violation. Since B mesons play a key role in these explorations, let us next have a closer look at them.

3 Decays of B mesons

The B-meson system consists of charged and neutral B mesons, which are characterized by the valence quark contents in (3). The characteristic feature of the neutral B_q ($q \in \{d, s\}$) mesons is the phenomenon of B_q^0–$\bar{B}_q^0$ mixing, which will be discussed in Section 5. As far as the weak decays of B mesons are concerned, we distinguish between leptonic, semileptonic, and non-leptonic transitions.
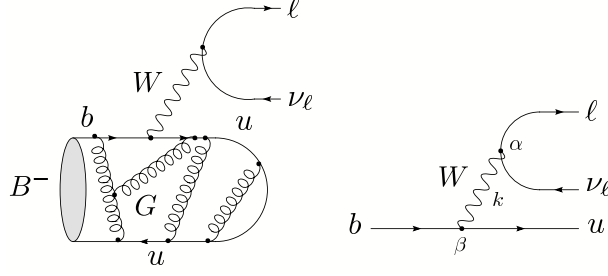
Fig. 5: Feynman diagrams contributing to the leptonic decay $B^- \to \ell\bar{\nu}_\ell$

3.1 Leptonic decays

The simplest B-meson decay class is given by leptonic decays of the kind $B^- \to \ell\bar{\nu}_\ell$, as illustrated in Fig. 5. If we evaluate the corresponding Feynman diagram, we arrive at the following transition amplitude:

$$T_{fi} = -\frac{g_2^2}{8}V_{ub}\underbrace{[\bar{u}_\ell\gamma^\alpha(1-\gamma_5)v_\nu]}_{\text{Dirac spinors}}\left[\frac{g_{\alpha\beta}}{k^2-M_W^2}\right]\underbrace{\langle 0|\bar{u}\gamma^\beta(1-\gamma_5)b|B^-\rangle}_{\text{hadronic ME}}\,, \tag{45}$$

where g_2 is the $SU(2)_{\mathrm{L}}$ gauge coupling, V_{ub} the corresponding element of the CKM matrix, α and β are Lorentz indices, and M_W denotes the mass of the W gauge boson. Since the four-momentum k that is carried by the W satisfies $k^2 = M_B^2 \ll M_W^2$, we may write

$$\frac{g_{\alpha\beta}}{k^2-M_W^2} \longrightarrow -\frac{g_{\alpha\beta}}{M_W^2} \equiv -\left(\frac{8G_{\mathrm{F}}}{\sqrt{2}g_2^2}\right)g_{\alpha\beta}\,, \tag{46}$$

where G_{F} is Fermi's constant. Consequently, we may 'integrate out' the W boson in (45), which yields

$$T_{fi} = \frac{G_{\mathrm{F}}}{\sqrt{2}}V_{ub}\left[\bar{u}_\ell\gamma^\alpha(1-\gamma_5)v_\nu\right]\langle 0|\bar{u}\gamma_\alpha(1-\gamma_5)b|B^-\rangle\,. \tag{47}$$

In this simple expression, *all* the hadronic physics is encoded in the *hadronic matrix element*

$$\langle 0|\bar{u}\gamma_\alpha(1-\gamma_5)b|B^-\rangle\,,$$

i.e., there are no other strong-interaction QCD effects (for a detailed discussion of QCD, see Ref. [43]). Since the B^- meson is a pseudoscalar particle, we have

$$\langle 0|\bar{u}\gamma_\alpha b|B^-\rangle = 0\,, \tag{48}$$

and may write

$$\langle 0|\bar{u}\gamma_\alpha\gamma_5 b|B^-(q)\rangle = if_B q_\alpha\,, \tag{49}$$

where f_B is the B-meson *decay constant*, which is an important input for phenomenological studies. In order to determine this quantity, which is a very challenging task, non-perturbative techniques, such as QCD sum-rule analyses [44] or lattice studies, where a numerical evaluation of the QCD path integral is performed with the help of a space-time lattice, [45]– [47], are required. If we use (47) with (48) and (49), and perform the corresponding phase-space integrations, we obtain the following decay rate:

$$\Gamma(B^- \to \ell\bar{\nu}_\ell) = \frac{G_{\mathrm{F}}^2}{8\pi}M_B m_\ell^2\left(1-\frac{m_\ell^2}{M_B^2}\right)^2 f_B^2|V_{ub}|^2\,, \tag{50}$$
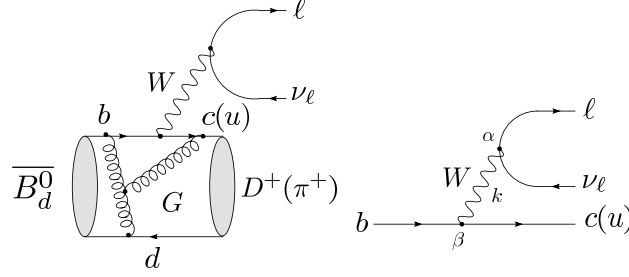
Fig. 6: Feynman diagrams contributing to semileptonic $\bar{B}_d^0 \to D^+(\pi^+)\ell\bar{\nu}_\ell$ decays

where M_B and m_ℓ denote the masses of the B^- and ℓ, respectively. Because of the tiny value of $|V_{ub}| \propto \lambda^3$ and a helicity-suppression mechanism, we obtain unfortunately very small branching ratios of $\mathcal{O}(10^{-10})$ and $\mathcal{O}(10^{-7})$ for $\ell = e$ and $\ell = \mu$, respectively [48]. The helicity suppression is not effective for $\ell = \tau$, but—because of the required τ reconstruction—these modes are also very challenging from an experimental point of view. Nevertheless, the Belle experiment has recently reported the first evidence for the purely leptonic decay $B^- \to \tau^-\bar{\nu}_\tau$, with the following branching ratio [49]:

$$\mathrm{BR}(B^- \to \tau^-\bar{\nu}_\tau) = \left[1.06^{+0.34}_{-0.28}\,\text{,(stat)}\,^{+0.18}_{-0.16}\,\text{(syst)}\right] \times 10^{-4}\,, \tag{51}$$

which corresponds to a significance of 4.2 standard deviations. Using the SM expression for this branching ratio and the measured values of G_F, M_B, m_τ and the B-meson lifetime, the product of the B-meson decay constant f_B and the magnitude of the CKM matrix element $|V_{ub}|$ is obtained as

$$f_B|V_{ub}| = \left[7.73^{+1.24}_{-1.02}\,\text{(stat)}\,^{+0.66}_{-0.58}\,\text{(syst)}\right] \times 10^{-4}\,\text{GeV}\,. \tag{52}$$

The determination of this quantity is very interesting, as knowledge of $|V_{ub}|$ allows us to extract f_B, thereby providing tests of non-perturbative calculations of this important parameter.

Before discussing the determination of $|V_{ub}|$ from semileptonic B decays in the next subsection, let us have a look at the leptonic D-meson decay $D^+ \to \mu^+\nu$. It is governed by the CKM factor

$$|V_{cd}| = |V_{us}| + \mathcal{O}(\lambda^5) = \lambda[1 + \mathcal{O}(\lambda^4)]\,, \tag{53}$$

whereas $B^- \to \mu^-\bar{\nu}$ involves $|V_{ub}| = \lambda^3 R_b$. Consequently, we win a factor of $\mathcal{O}(\lambda^4)$ in the decay rate, so that $D^+ \to \mu^+\nu$ is accessible at the CLEO-c experiment [50]. Since the corresponding CKM factor is well known, the decay constant f_{D^+} defined in analogy to (49) can be extracted, allowing another interesting testing ground for lattice calculations. Thanks to recent progress in these techniques [51], the 'quenched' approximation, which had to be applied for many many years and ingnores quark loops, is no longer required for the calculation of f_{D^+}. In the summer of 2005, there was a first showdown between the corresponding theoretical prediction and experiment: the lattice result of $f_{D^+} = (201 \pm 3 \pm 17)$MeV was reported [52], while CLEO-c announced the measurement of $f_{D^+} = (222.6 \pm 16.7^{+2.8}_{-3.4})$ MeV [53]. Both numbers agree well within the uncertainties, and it will be interesting to stay tuned for future results.

3.2 Semileptonic decays

3.2.1 General structure

Semileptonic B-meson decays of the kind shown in Fig. 6 have a structure that is more complicated than the one of the leptonic transitions. If we evaluate the corresponding Feynman diagram for the $b \to c$ case, we obtain

$$T_{fi} = -\frac{g_2^2}{8}V_{cb}\underbrace{\left[\bar{u}_\ell\gamma^\alpha(1-\gamma_5)v_\nu\right]}_{\text{Dirac spinors}}\left[\frac{g_{\alpha\beta}}{k^2 - M_W^2}\right]\underbrace{\langle D^+|\bar{c}\gamma^\beta(1-\gamma_5)b|\bar{B}_d^0\rangle}_{\text{hadronic ME}}\,. \tag{54}$$

Because of $k^2 \sim M_B^2 \ll M_W^2$, we may again—as in (45)—integrate out the W boson with the help of (46), which yields

$$T_{fi} = \frac{G_{\mathrm{F}}}{\sqrt{2}} V_{cb} \left[\bar{u}_\ell \gamma^\alpha (1 - \gamma_5) v_\nu\right] \langle D^+ | \bar{c} \gamma_\alpha (1 - \gamma_5) b | \bar{B}_d^0 \rangle \,, \tag{55}$$

where *all* the hadronic physics is encoded in the hadronic matrix element

$$\langle D^+ | \bar{c} \gamma_\alpha (1 - \gamma_5) b | \bar{B}_d^0 \rangle \,,$$

i.e., there are *no* other QCD effects. Since the $\bar{B}_d^0$ and D^+ are pseudoscalar mesons, we have

$$\langle D^+ | \bar{c} \gamma_\alpha \gamma_5 b | \bar{B}_d^0 \rangle = 0 \,, \tag{56}$$

and may write

$$\langle D^+(k) | \bar{c} \gamma_\alpha b | \bar{B}_d^0(p) \rangle = F_1(q^2) \left[(p + k)_\alpha - \left(\frac{M_B^2 - M_D^2}{q^2}\right) q_\alpha\right] + F_0(q^2) \left(\frac{M_B^2 - M_D^2}{q^2}\right) q_\alpha \,, \tag{57}$$

where $q \equiv p - k$, and the $F_{1,0}(q^2)$ denote the *form factors* of the $\bar{B} \to D$ transitions. Consequently, in contrast to the simple case of the leptonic transitions, semileptonic decays involve *two* hadronic form factors instead of the decay constant f_B. In order to calculate these parameters, which depend on the momentum transfer q, again non-perturbative techniques (QCD sum rules, lattice, etc.) are required.

3.2.2 Aspects of the heavy-quark effective theory

If the mass m_Q of a quark Q is much larger than the QCD scale parameter $\Lambda_{\mathrm{QCD}} = \mathcal{O}(100\,\mathrm{MeV})$ [43], it is referred to as a 'heavy' quark. Since the bottom and charm quarks have masses at the level of $5\,\mathrm{GeV}$ and $1\,\mathrm{GeV}$, respectively, they belong to this important category. As far as the extremely heavy top quark, with $m_t \sim 170\,\mathrm{GeV}$ is concerned, it decays unfortunately through weak interactions before a hadron can be formed. Let us now consider a heavy quark that is bound inside a hadron, i.e., a bottom or a charm quark. The heavy quark then moves almost with the hadron's four velocity v and is almost on-shell, so that

$$p_Q^\mu = m_Q v^\mu + k^\mu \,, \tag{58}$$

where $v^2 = 1$ and $k \ll m_Q$ is the 'residual' momentum. Owing to the interactions of the heavy quark with the light degrees of freedom of the hadron, the residual momentum may only change by $\Delta k \sim \Lambda_{\mathrm{QCD}}$, and $\Delta v \to 0$ for $\Lambda_{\mathrm{QCD}}/m_Q \to 0$.

It is now instructive to have a look at the elastic scattering process $\bar{B}(v) \to \bar{B}(v')$ in the limit of $\Lambda_{\mathrm{QCD}}/m_b \to 0$, which is characterized by the following matrix element:

$$\frac{1}{M_B} \langle \bar{B}(v') | \bar{b}_{v'} \gamma_\alpha b_v | \bar{B}(v) \rangle = \xi(v' \cdot v)(v + v')_\alpha \,. \tag{59}$$

Since the contraction of this matrix element with $(v - v')^\alpha$ has to vanish because of $\slashed{v} b_v = b_v$ and $\bar{b}_v \slashed{v}' = \bar{b}_{v'}$, no $(v - v')_\alpha$ term arises in the parametrization in (59). On the other hand, the $1/M_B$ factor is related to the normalization of states, i.e., the right-hand side of

$$\left(\frac{1}{\sqrt{M_B}} \langle \bar{B}(p') |\right) \left(|\bar{B}(p)\rangle \frac{1}{\sqrt{M_B}}\right) = 2v^0 (2\pi)^3 \delta^3(\vec{p} - \vec{p}') \tag{60}$$

does not depend on M_B. Finally, current conservation implies the following normalization condition:

$$\xi(v' \cdot v = 1) = 1 \,, \tag{61}$$

where the 'Isgur–Wise' function $\xi(v' \cdot v)$ does *not* depend on the flavour of the heavy quark (heavy-quark symmetry) [54]. Consequently, for $\Lambda_{\mathrm{QCD}}/m_{b,c} \to 0$, we may write

$$\frac{1}{\sqrt{M_D M_B}} \langle D(v') | \bar{c}_{v'} \gamma_\alpha b_v | \bar{B}(v) \rangle = \xi(v' \cdot v)(v + v')_\alpha \, , \tag{62}$$

and observe that this transition amplitude is governed—in the heavy-quark limit—by *one* hadronic form factor $\xi(v' \cdot v)$, which satisfies $\xi(1) = 1$. If we now compare (62) with (57), we obtain

$$F_1(q^2) = \frac{M_D + M_B}{2\sqrt{M_D M_B}} \xi(w) \tag{63}$$

$$F_0(q^2) = \frac{2\sqrt{M_D M_B}}{M_D + M_B} \left[\frac{1 + w}{2} \right] \xi(w) \, , \tag{64}$$

with

$$w \equiv v_D \cdot v_B = \frac{M_D^2 + M_B^2 - q^2}{2 M_D M_B} \, . \tag{65}$$

Similar relations hold for the $\bar{B} \to D^*$ form factors because of the heavy-quark spin symmetry, since the D^* is related to the D by a rotation of the heavy-quark spin. A detailed discussion of these interesting features and the associated 'heavy-quark effective theory' (HQET) is beyond the scope of these lectures. For a detailed overview, we refer the reader to Ref. [55], where also a comprehensive list of original references can be found. For a more phenomenological discussion, Ref. [56] is very useful.

3.2.3 Applications

An important application of the formalism sketched above is the extraction of the CKM element $|V_{cb}|$. To this end, $\bar{B} \to D^* \ell \bar{\nu}$ decays are particularly promising. The corresponding rate can be written as

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}w} = G_{\mathrm{F}}^2 K(M_B, M_{D^*}, w) F(w)^2 |V_{cb}|^2 \, , \tag{66}$$

where $K(M_B, M_{D^*}, w)$ is a known kinematic function, and $F(w)$ agrees with the Isgur–Wise function, up to perturbative QCD corrections and $\Lambda_{\mathrm{QCD}}/m_{b,c}$ terms. The form factor $F(w)$ is a non-perturbative quantity. However, it satisfies the following normalization condition:

$$F(1) = \eta_A(\alpha_s) \left[1 + \frac{0}{m_c} + \frac{0}{m_b} + \mathcal{O}(\Lambda_{\mathrm{QCD}}^2/m_{b,c}^2) \right] \, , \tag{67}$$

where $\eta_A(\alpha_s)$ is a perturbatively calculable short-distance QCD factor, and the $\Lambda_{\mathrm{QCD}}/m_{b,c}$ corrections *vanish* [55, 57]. The important latter feature is an implication of Luke's theorem [58]. Consequently, if we extract $F(w)|V_{cb}|$ from a measurement of (66) as a function of w and extrapolate to the 'zero-recoil point' $w = 1$ (where the rate vanishes), we may determine $|V_{cb}|$. In the case of $\bar{B} \to D\ell\bar{\nu}$ decays, we have $\mathcal{O}(\Lambda_{\mathrm{QCD}}/m_{b,c})$ corrections to the corresponding rate $\mathrm{d}\Gamma/\mathrm{d}w$ at $w = 1$. In order to determine $|V_{cb}|$, inclusive $B \to X_c \ell \bar{\nu}$ decays offer also very attractive avenues. As becomes obvious from (27) and the considerations in Section 2.6, $|V_{cb}|$ fixes the normalization of the UT. Moreover, this quantity is an important input parameter for various theoretical calculations. The CKM matrix element $|V_{cb}|$ is currently known with 2% precision; performing an analysis of leptonic and hadronic moments in inclusive $b \to c\ell\bar{\nu}$ processes [59], the following value was extracted from the B-factory data [60]:

$$|V_{cb}| = (42.0 \pm 0.7) \times 10^{-3} \, , \tag{68}$$

which agrees with that from exclusive decays.

Let us now turn to $\bar{B} \to \pi\ell\bar{\nu}, \rho\ell\bar{\nu}$ decays, which originate from $b \to u\ell\bar{\nu}$ quark-level processes, as can be seen in Fig. 6, and provide access to $|V_{ub}|$. If we complement this CKM matrix element with $|V_{cb}|$,

we may determine the side R_b of the UT with the help of (38). The determination of $|V_{ub}|$ is hence a very important aspect of flavour physics. Since the π and ρ are 'light' mesons, the HQET symmetry relations cannot be applied to the $\bar{B} \to \pi\ell\bar{\nu}, \rho\ell\bar{\nu}$ modes. Consequently, in order to determine $|V_{ub}|$ from these exclusive channels, the corresponding heavy-to-light form factors have to be described by models. An important alternative is provided by inclusive decays. The corresponding decay rate takes the following form:

$$\Gamma(\bar{B} \to X_u\ell\bar{\nu}) = \frac{G_{\mathrm{F}}^2|V_{ub}|^2}{192\pi^3}m_b^5\left[1 - 2.41\frac{\alpha_s}{\pi} + \frac{\lambda_1 - 9\lambda_2}{2m_b^2} + \ldots\right],\tag{69}$$

where λ_1 and λ_2 are non-perturbative parameters, which describe the hadronic matrix elements of certain 'kinetic' and 'chromomagnetic' operators appearing within the framework of the HQET. Using the heavy-quark expansions

$$M_B = m_b + \bar{\Lambda} - \frac{\lambda_1 + 3\lambda_2}{2m_b} + \ldots, \quad M_{B^*} = m_b + \bar{\Lambda} - \frac{\lambda_1 - \lambda_2}{2m_b} + \ldots\tag{70}$$

for the $B^{(*)}$-meson masses, where $\bar{\Lambda} \sim \Lambda_{\mathrm{QCD}}$ is another non-perturbative parameter that is related to the light degrees of freedom, the parameter λ_2 can be determined from the measured values of the $M_{B^{(*)}}$. The strong dependence of (69) on m_b is a significant source of uncertainty. On the other hand, the $1/m_b^2$ corrections can be better controlled than in the exclusive case (67), where we have, moreover, to deal with $1/m_c^2$ corrections. From an experimental point of view, we have to struggle with large backgrounds, which originate from $b \to c\ell\bar{\nu}$ processes and require also a model-dependent treatment. The determination of $|V_{ub}|$ from B-meson decays caused by $b \to u\ell\bar{\nu}$ quark-level processes is therefore a very challenging issue, and the situation is less favourable than with $|V_{cb}|$: there is a $1\,\sigma$ discrepancy between the values from inclusive and exclusive transitions [61]:

$$|V_{ub}|_{\mathrm{incl}} = (4.4 \pm 0.3) \times 10^{-3}, \quad |V_{ub}|_{\mathrm{excl}} = (3.8 \pm 0.6) \times 10^{-3},\tag{71}$$

which has to be settled in the future. The error on $|V_{ub}|_{\mathrm{excl}}$ is dominated by the theoretical uncertainty of lattice and light-cone sum rule calculations of $B \to \pi$ and $B \to \rho$ transition form factors [62, 63], whereas for $|V_{ub}|_{\mathrm{incl}}$ experimental and theoretical errors are at par. Using the values of $|V_{cb}|$ and $|V_{ub}|$ given above and $\lambda = 0.225 \pm 0.001$ [64], we obtain

$$R_b^{\mathrm{incl}} = 0.45 \pm 0.03, \qquad R_b^{\mathrm{excl}} = 0.39 \pm 0.06,\tag{72}$$

where the labels 'incl' and 'excl' refer to the determinations of $|V_{ub}|$ through inclusive and exclusive $b \to u\ell\bar{\nu}_\ell$ transitions, respectively.

For a much more detailed discussion of the determinations of $|V_{cb}|$ and $|V_{ub}|$, addressing also various recent developments and the future prospects, we refer the reader to Ref. [12], where also the references to the vast original literature can be found. Another excellent presentation is given in Ref. [56].

3.3 Non-leptonic decays

3.3.1 Classification

The most complicated B decays are the non-leptonic transitions, which are mediated by $b \to q_1\bar{q}_2 d\,(s)$ quark-level processes, with $q_1, q_2 \in \{u, d, c, s\}$. There are two kinds of topologies contributing to such decays: tree-diagram-like and 'penguin' topologies. The latter consist of gluonic (QCD) and electroweak (EW) penguins. In Fig. 7, the corresponding leading-order Feynman diagrams are shown. Depending on the flavour content of their final states, we may classify $b \to q_1\bar{q}_2 d\,(s)$ decays as follows:

- $q_1 \neq q_2 \in \{u, c\}$: *only* tree diagrams contribute.
- $q_1 = q_2 \in \{u, c\}$: tree *and* penguin diagrams contribute.
- $q_1 = q_2 \in \{d, s\}$: *only* penguin diagrams contribute.

(a)

(b)

(c)
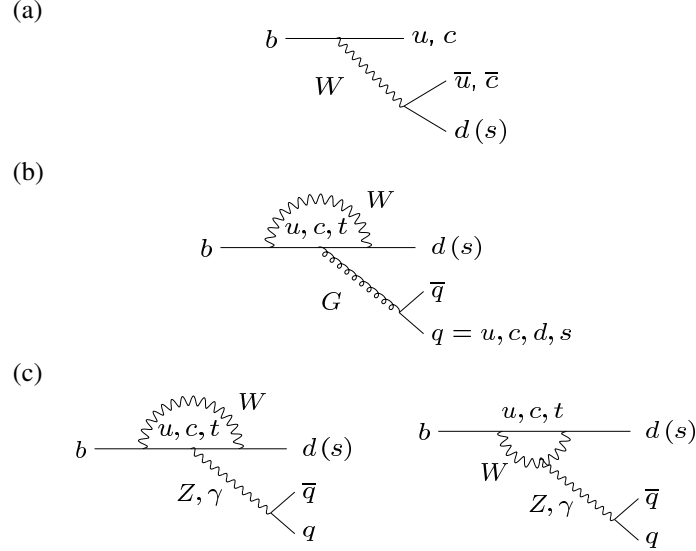
Fig. 7: Feynman diagrams of the topologies characterizing non-leptonic B decays: (a) trees, (b) QCD penguins, and (c) electroweak penguins
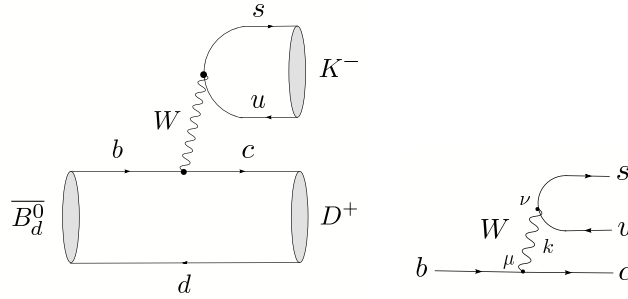
Fig. 8: Feynman diagrams contributing to the non-leptonic $\bar{B}_d^0 \to D^+ K^-$ decay
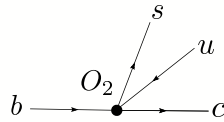
Fig. 9: The description of the $b \to d\bar{u}s$ process through the four-quark operator O_2 in the effective theory after the W boson has been integrated out

3.3.2 Low-energy effective Hamiltonians

In order to analyse non-leptonic B decays theoretically, one uses low-energy effective Hamiltonians, which are calculated by making use of the 'operator product expansion', yielding transition matrix elements of the following structure:

$$\langle f|\mathcal{H}_{\text{eff}}|i\rangle = \frac{G_{\text{F}}}{\sqrt{2}}\lambda_{\text{CKM}}\sum_k C_k(\mu)\langle f|Q_k(\mu)|i\rangle\,. \tag{73}$$

The technique of the operator product expansion allows us to separate the short-distance contributions to this transition amplitude from the long-distance ones, which are described by perturbative quantities $C_k(\mu)$ ('Wilson coefficient functions') and non-perturbative quantities $\langle f|Q_k(\mu)|i\rangle$ ('hadronic matrix elements'), respectively. As before, G_{F} is the Fermi constant, whereas λ_{CKM} is a CKM factor and μ
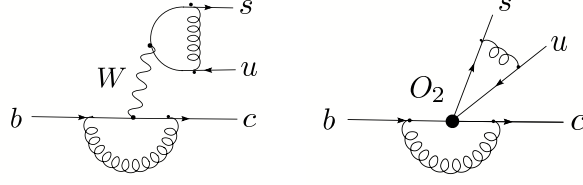
Fig. 10: Factorizable QCD corrections in the full and effective theories
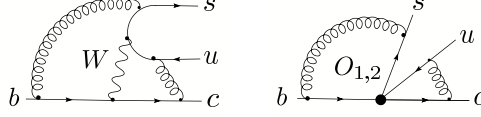
Fig. 11: Non-factorizable QCD corrections in the full and effective theories

denotes an appropriate renormalization scale. The Q_k are local operators, which are generated by electroweak interactions and QCD, and govern 'effectively' the decay in question. The Wilson coefficients $C_k(\mu)$ can be considered as scale-dependent couplings related to the vertices described by the Q_k.

In order to illustrate this rather abstract formalism, let us consider the decay $\bar{B}_d^0 \to D^+ K^-$, which allows a transparent discussion of the evaluation of the corresponding low-energy effective Hamiltonian. Since this transition originates from a $b \to c\bar{u}s$ quark-level process, it is—as we have seen in our classification in Subsection 3.3.1—a pure 'tree' decay, i.e., we do not have to deal with penguin topologies, which simplifies the analysis considerably. The leading-order Feynman diagram contributing to $\bar{B}_d^0 \to D^+ K^-$ can straightforwardly be obtained from Fig. 6 by substituting ℓ and ν_ℓ by s and u, respectively, as can be seen in Fig. 8. Consequently, the lepton current is simply replaced by a quark current, which will have important implications shown below. Evaluating the corresponding Feynman diagram yields

$$-\frac{g_2^2}{8}V_{us}^* V_{cb}\left[\bar{s}\gamma^\nu(1-\gamma_5)u\right]\left[\frac{g_{\nu\mu}}{k^2-M_W^2}\right]\left[\bar{c}\gamma^\mu(1-\gamma_5)b\right].\tag{74}$$

Because of $k^2 \sim m_b^2 \ll M_W^2$, we may—as in (54)—'integrate out' the W boson with the help of (46), and arrive at

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}}V_{us}^* V_{cb}\left[\bar{s}_\alpha\gamma_\mu(1-\gamma_5)u_\alpha\right]\left[\bar{c}_\beta\gamma^\mu(1-\gamma_5)b_\beta\right]$$
$$= \frac{G_F}{\sqrt{2}}V_{us}^* V_{cb}(\bar{s}_\alpha u_\alpha)_{\text{V-A}}(\bar{c}_\beta b_\beta)_{\text{V-A}} \equiv \frac{G_F}{\sqrt{2}}V_{us}^* V_{cb}O_2 ,\tag{75}$$

where α and β denote the colour indices of the $SU(3)_{\text{C}}$ gauge group of QCD. Effectively, our $b \to c\bar{u}s$ decay process is now described by the 'current–current' operator O_2, as is illustrated in Fig. 9.

So far, we neglected QCD corrections. Their important impact is twofold: thanks to *factorizable* QCD corrections as shown in Fig. 10, the Wilson coefficient C_2 acquires a renormalization-scale dependence, i.e., $C_2(\mu) \neq 1$. On the other hand, *non-factorizable* QCD corrections as illustrated in Fig. 11 generate a second current–current operator through 'operator mixing', which is given by

$$O_1 \equiv \left[\bar{s}_\alpha\gamma_\mu(1-\gamma_5)u_\beta\right]\left[\bar{c}_\beta\gamma^\mu(1-\gamma_5)b_\alpha\right].\tag{76}$$

Consequently, we eventually arrive at a low-energy effective Hamiltonian of the following structure:

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}}V_{us}^* V_{cb}\left[C_1(\mu)O_1 + C_2(\mu)O_2\right].\tag{77}$$

In order to evaluate the Wilson coefficients $C_1(\mu) \neq 0$ and $C_2(\mu) \neq 1$ [65], we must first calculate the QCD corrections to the decay processes both in the full theory, i.e., with W exchange, and in the effective theory, where the W is integrated out (see Figs. 10 and 11), and have then to express the QCD-corrected transition amplitude in terms of QCD-corrected matrix elements and Wilson coefficients as in (73). This procedure is called 'matching' between the full and the effective theory. The results for the $C_k(\mu)$ thus obtained contain terms of $\log(\mu/M_W)$, which become large for $\mu = \mathcal{O}(m_b)$, the scale governing the hadronic matrix elements of the O_k. Making use of the renormalization group, which exploits the fact that the transition amplitude (73) cannot depend on the chosen renormalization scale μ, we may sum up the following terms of the Wilson coefficients:

$$\alpha_s^n \left[\log\left(\frac{\mu}{M_W} \right) \right]^n \quad \text{(LO)}, \quad \alpha_s^n \left[\log\left(\frac{\mu}{M_W} \right) \right]^{n-1} \quad \text{(NLO)}, \quad \ldots \; ; \tag{78}$$

detailed discussions of these rather technical aspects can be found in Refs. [66, 67].

For the exploration of CP violation, the class of non-leptonic B decays that receives contributions both from tree and from penguin topologies plays a key role. In this important case, the operator basis is much larger than in our example (77), where we considered a pure 'tree' decay. If we apply the relation

$$V_{ur}^* V_{ub} + V_{cr}^* V_{cb} + V_{tr}^* V_{tb} = 0 \quad (r \in \{d, s\}), \tag{79}$$

which follows from the unitarity of the CKM matrix, and 'integrate out' the top quark (which enters through the penguin loop processes) and the W boson, we may write

$$\mathcal{H}_{\text{eff}} = \frac{G_{\text{F}}}{\sqrt{2}} \left[\sum_{j=u,c} V_{jr}^* V_{jb} \left\{ \sum_{k=1}^{2} C_k(\mu)\, Q_k^{jr} + \sum_{k=3}^{10} C_k(\mu)\, Q_k^r \right\} \right]. \tag{80}$$

Here we have introduced another quark-flavour label $j \in \{u, c\}$, and the Q_k^{jr} can be divided as follows:

– Current–current operators:
$$\begin{aligned} Q_1^{jr} &= (\bar{r}_\alpha j_\beta)_{\text{V–A}} (\bar{j}_\beta b_\alpha)_{\text{V–A}} \\ Q_2^{jr} &= (\bar{r}_\alpha j_\alpha)_{\text{V–A}} (\bar{j}_\beta b_\beta)_{\text{V–A}}. \end{aligned} \tag{81}$$

– QCD penguin operators:
$$\begin{aligned} Q_3^r &= (\bar{r}_\alpha b_\alpha)_{\text{V–A}} \sum_{q'} (\bar{q}'_\beta q'_\beta)_{\text{V–A}} \\ Q_4^r &= (\bar{r}_\alpha b_\beta)_{\text{V–A}} \sum_{q'} (\bar{q}'_\beta q'_\alpha)_{\text{V–A}} \\ Q_5^r &= (\bar{r}_\alpha b_\alpha)_{\text{V–A}} \sum_{q'} (\bar{q}'_\beta q'_\beta)_{\text{V+A}} \\ Q_6^r &= (\bar{r}_\alpha b_\beta)_{\text{V–A}} \sum_{q'} (\bar{q}'_\beta q'_\alpha)_{\text{V+A}}. \end{aligned} \tag{82}$$

– EW penguin operators (the $e_{q'}$ denote the electrical quark charges):
$$\begin{aligned} Q_7^r &= \tfrac{3}{2}(\bar{r}_\alpha b_\alpha)_{\text{V–A}} \sum_{q'} e_{q'} (\bar{q}'_\beta q'_\beta)_{\text{V+A}} \\ Q_8^r &= \tfrac{3}{2}(\bar{r}_\alpha b_\beta)_{\text{V–A}} \sum_{q'} e_{q'} (\bar{q}'_\beta q'_\alpha)_{\text{V+A}} \\ Q_9^r &= \tfrac{3}{2}(\bar{r}_\alpha b_\alpha)_{\text{V–A}} \sum_{q'} e_{q'} (\bar{q}'_\beta q'_\beta)_{\text{V–A}} \\ Q_{10}^r &= \tfrac{3}{2}(\bar{r}_\alpha b_\beta)_{\text{V–A}} \sum_{q'} e_{q'} (\bar{q}'_\beta q'_\alpha)_{\text{V–A}}. \end{aligned} \tag{83}$$

The current–current, QCD and EW penguin operators are related to the tree, QCD and EW penguin processes shown in Fig. 7. At a renormalization scale $\mu = \mathcal{O}(m_b)$, the Wilson coefficients of the current–current operators are $C_1(\mu) = \mathcal{O}(10^{-1})$ and $C_2(\mu) = \mathcal{O}(1)$, whereas those of the penguin operators are $\mathcal{O}(10^{-2})$ [66, 67]. Note that penguin topologies with internal charm- and up-quark exchanges [68] are described in this framework by penguin-like matrix elements of the corresponding current–current operators [69], and may also have important phenomenological consequences [70, 71].

Since the ratio $\alpha/\alpha_s = \mathcal{O}(10^{-2})$ of the QED and QCD couplings is very small, we would expect naïvely that EW penguins should play a minor role in comparison with QCD penguins. This would actually be the case if the top quark was not 'heavy'. However, since the Wilson coefficient C_9 increases strongly with m_t, we obtain interesting EW penguin effects in several B decays: $B \to K\phi$ modes are affected significantly by EW penguins, whereas $B \to \pi\phi$ and $B_s \to \pi^0\phi$ transitions are even *dominated* by such topologies [72, 73]. EW penguins also have an important impact on the $B \to \pi K$ system [74].

The low-energy effective Hamiltonians discussed above apply to all B decays that are caused by the same quark-level transition, i.e., they are 'universal'. Consequently, the differences between the various exclusive modes of a given decay class arise within this formalism only through the hadronic matrix elements of the relevant four-quark operators. Unfortunately, the evaluation of such matrix elements is associated with large uncertainties and is a very challenging task. In this context, 'factorization' is a widely used concept, which is our next topic.

3.3.3 *Factorization of hadronic matrix elements*

In order to discuss 'factorization', let us consider once more the decay $\bar{B}_d^0 \to D^+ K^-$. Evaluating the corresponding transition amplitude, we encounter the hadronic matrix elements of the $O_{1,2}$ operators between the $\langle K^- D^+ |$ final and the $|\bar{B}_d^0\rangle$ initial states. If we use the well-known $SU(N_{\rm C})$ colour-algebra relation

$$T^a_{\alpha\beta} T^a_{\gamma\delta} = \frac{1}{2}\left(\delta_{\alpha\delta}\delta_{\beta\gamma} - \frac{1}{N_{\rm C}}\delta_{\alpha\beta}\delta_{\gamma\delta}\right) \tag{84}$$

to rewrite the operator O_1, we obtain

$$\langle K^- D^+ |\mathcal{H}_{\rm eff}|\bar{B}_d^0\rangle = \frac{G_{\rm F}}{\sqrt{2}} V_{us}^* V_{cb}\Big[a_1\langle K^- D^+ |(\bar{s}_\alpha u_\alpha)_{\rm V\text{-}A}(\bar{c}_\beta b_\beta)_{\rm V\text{-}A}|\bar{B}_d^0\rangle$$

$$+2\,C_1\langle K^- D^+ |(\bar{s}_\alpha T^a_{\alpha\beta} u_\beta)_{\rm V\text{-}A}(\bar{c}_\gamma T^a_{\gamma\delta} b_\delta)_{\rm V\text{-}A}|\bar{B}_d^0\rangle\Big],$$

with

$$a_1 = C_1/N_{\rm C} + C_2 \sim 1 \,. \tag{85}$$

It is now straightforward to 'factorize' the hadronic matrix elements in (85):

$$\langle K^- D^+ |(\bar{s}_\alpha u_\alpha)_{\rm V\text{-}A}(\bar{c}_\beta b_\beta)_{\rm V\text{-}A}|\bar{B}_d^0\rangle\big|_{\rm fact}$$
$$= \langle K^- | [\bar{s}_\alpha \gamma_\mu(1-\gamma_5)u_\alpha] |0\rangle\langle D^+ | [\bar{c}_\beta \gamma^\mu(1-\gamma_5)b_\beta] |\bar{B}_d^0\rangle$$
$$= \underbrace{if_K}_{\text{decay constant}} \times \underbrace{F_0^{(BD)}(M_K^2)}_{B \to D \text{ form factor}} \times \underbrace{(M_B^2 - M_D^2)}_{\text{kinematical factor}}, \tag{86}$$

$$\langle K^- D^+ |(\bar{s}_\alpha T^a_{\alpha\beta} u_\beta)_{\rm V\text{-}A}(\bar{c}_\gamma T^a_{\gamma\delta} b_\delta)_{\rm V\text{-}A}|\bar{B}_d^0\rangle\big|_{\rm fact} = 0 \,. \tag{87}$$

The quantity a_1 is a phenomenological 'colour factor', which governs 'colour-allowed' decays; the decay $\bar{B}_d^0 \to D^+ K^-$ belongs to this category, since the colour indices of the K^- meson and the $\bar{B}_d^0$–D^+ system run independently from each other in the corresponding leading-order diagram shown in Fig. 8. On the other hand, in the case of 'colour-suppressed' modes, for instance $\bar{B}_d^0 \to \pi^0 D^0$, where only one colour index runs through the whole diagram, we have to deal with the combination

$$a_2 = C_1 + C_2/N_{\rm C} \sim 0.25 \,. \tag{88}$$

The concept of factorizing the hadronic matrix elements of four-quark operators into the product of hadronic matrix elements of quark currents has a long history [75], and can be justified, for example, in the large-$N_{\rm C}$ limit [76]. Interesting recent developments are the following:

- 'QCD factorization' [77], which is in accordance with the old picture that factorization should hold for certain decays in the limit of $m_b \gg \Lambda_{\text{QCD}}$ [78], provides a formalism to calculate the relevant amplitudes at the leading order of a Λ_{QCD}/m_b expansion. The resulting expression for the transition amplitudes incorporates elements both of the naïve factorization approach sketched above and of the hard-scattering picture. Let us consider a decay $\bar{B} \to M_1 M_2$, where M_1 picks up the spectator quark. If M_1 is either a heavy (D) or a light (π, K) meson, and M_2 a light (π, K) meson, QCD factorization gives a transition amplitude of the following structure:

$$A(\bar{B} \to M_1 M_2) = [\text{'naïve factorization'}] \times [1 + \mathcal{O}(\alpha_s) + \mathcal{O}(\Lambda_{\text{QCD}}/m_b)]. \tag{89}$$

  While the $\mathcal{O}(\alpha_s)$ terms, i.e., the radiative non-factorizable corrections, can be calculated systematically, the main limitation of the theoretical accuracy originates from the $\mathcal{O}(\Lambda_{\text{QCD}}/m_b)$ terms.
- Another QCD approach to deal with non-leptonic B-meson decays—the 'perturbative hard-scattering approach' (PQCD)—was developed independently in Ref. [79], and differs from the QCD factorization formalism in some technical aspects.
- An interesting technique for 'factorization proofs' is provided by the framework of the 'soft collinear effective theory' (SCET) [80], which has received a lot of attention in the recent literature and led to various applications.
- Non-leptonic B decays can also be studied within QCD light-cone sum-rule approaches [81].

A detailed presentation of these topics would be very technical and is beyond the scope of these lectures. However, for the discussion of the CP-violating effects in the B-meson system, we must only be familiar with the general structure of the non-leptonic B decay amplitudes and not enter the details of the techniques to deal with the corresponding hadronic matrix elements. Let us finally note that the B-factory data will eventually decide how well factorization and the new concepts sketched above are actually working. For example, data on the $B \to \pi\pi$ system point towards large non-factorizable corrections [82, 83], to which we shall return in Section 8.2.

3.4 Towards studies of CP violation

As we have seen above, leptonic and semileptonic B-meson decays involve only a single weak (CKM) amplitude. On the other hand, the structure of non-leptonic transitions is considerably more complicated. Let us consider a non-leptonic decay $\bar{B} \to \bar{f}$ that is described by the low-energy effective Hamiltonian in (80). The corresponding decay amplitude is then given as follows:

$$
\begin{aligned}
A(\bar{B} \to \bar{f}) &= \langle \bar{f} | \mathcal{H}_{\text{eff}} | \bar{B} \rangle \\
&= \frac{G_{\text{F}}}{\sqrt{2}} \left[\sum_{j=u,c} V_{jr}^* V_{jb} \left\{ \sum_{k=1}^{2} C_k(\mu) \langle \bar{f} | Q_k^{jr}(\mu) | \bar{B} \rangle + \sum_{k=3}^{10} C_k(\mu) \langle \bar{f} | Q_k^{r}(\mu) | \bar{B} \rangle \right\} \right].
\end{aligned}
\tag{90}
$$

Concerning the CP-conjugate process $B \to f$, we have

$$
\begin{aligned}
A(B \to f) &= \langle f | \mathcal{H}_{\text{eff}}^{\dagger} | B \rangle \\
&= \frac{G_{\text{F}}}{\sqrt{2}} \left[\sum_{j=u,c} V_{jr} V_{jb}^* \left\{ \sum_{k=1}^{2} C_k(\mu) \langle f | Q_k^{jr\dagger}(\mu) | B \rangle + \sum_{k=3}^{10} C_k(\mu) \langle f | Q_k^{r\dagger}(\mu) | B \rangle \right\} \right].
\end{aligned}
\tag{91}
$$

If we use now that strong interactions are invariant under CP transformations, insert $(\mathcal{CP})^{\dagger}(\mathcal{CP}) = \hat{1}$ both after the $\langle f |$ and in front of the $| B \rangle$, and take the relation

$$(\mathcal{CP}) Q_k^{jr\dagger} (\mathcal{CP})^{\dagger} = Q_k^{jr} \tag{92}$$

into account, we arrive at

$$A(B \to f) = e^{i[\phi_{\text{CP}}(B) - \phi_{\text{CP}}(f)]}$$

$$\times \frac{G_{\text{F}}}{\sqrt{2}} \left[\sum_{j=u,c} V_{jr} V_{jb}^* \left\{ \sum_{k=1}^{2} C_k(\mu) \langle \bar{f} | Q_k^{jr}(\mu) | \bar{B} \rangle + \sum_{k=3}^{10} C_k(\mu) \langle \bar{f} | Q_k^{r}(\mu) | \bar{B} \rangle \right\} \right], \tag{93}$$

where the convention-dependent phases $\phi_{\text{CP}}(B)$ and $\phi_{\text{CP}}(f)$ are defined through

$$(\mathcal{CP})|B\rangle = e^{i\phi_{\text{CP}}(B)}|\bar{B}\rangle, \quad (\mathcal{CP})|f\rangle = e^{i\phi_{\text{CP}}(f)}|\bar{f}\rangle. \tag{94}$$

Consequently, we may write

$$A(\bar{B} \to \bar{f}) = e^{+i\varphi_1}|A_1|e^{i\delta_1} + e^{+i\varphi_2}|A_2|e^{i\delta_2} \tag{95}$$

$$A(B \to f) = e^{i[\phi_{\text{CP}}(B) - \phi_{\text{CP}}(f)]} \left[e^{-i\varphi_1}|A_1|e^{i\delta_1} + e^{-i\varphi_2}|A_2|e^{i\delta_2} \right]. \tag{96}$$

Here the CP-violating phases $\varphi_{1,2}$ originate from the CKM factors $V_{jr}^* V_{jb}$, and the CP-conserving 'strong' amplitudes $|A_{1,2}|e^{i\delta_{1,2}}$ involve the hadronic matrix elements of the four-quark operators. In fact, these expressions are the most general forms of any non-leptonic B-decay amplitude in the SM, i.e., they do not only refer to the $\Delta C = \Delta U = 0$ case described by (80). Using (95) and (96), we obtain the following CP asymmetry:

$$\mathcal{A}_{\text{CP}} \equiv \frac{\Gamma(B \to f) - \Gamma(\bar{B} \to \bar{f})}{\Gamma(B \to f) + \Gamma(\bar{B} \to \bar{f})} = \frac{|A(B \to f)|^2 - |A(\bar{B} \to \bar{f})|^2}{|A(B \to f)|^2 + |A(\bar{B} \to \bar{f})|^2}$$

$$= \frac{2|A_1||A_2|\sin(\delta_1 - \delta_2)\sin(\varphi_1 - \varphi_2)}{|A_1|^2 + 2|A_1||A_2|\cos(\delta_1 - \delta_2)\cos(\varphi_1 - \varphi_2) + |A_2|^2}. \tag{97}$$

We observe that a non-vanishing value can be generated through the interference between the two weak amplitudes, provided both a non-trivial weak phase difference $\varphi_1 - \varphi_2$ and a non-trivial strong phase difference $\delta_1 - \delta_2$ are present. This kind of CP violation is referred to as 'direct' CP violation, as it originates directly at the amplitude level of the considered decay. It is the B-meson counterpart of the effects that are probed through $\text{Re}(\varepsilon'/\varepsilon)$ in the neutral kaon system[1], and could recently be established with the help of $B_d \to \pi^{\mp} K^{\pm}$ decays [6], as we shall see in Section 7.3.

Since $\varphi_1 - \varphi_2$ is in general given by one of the UT angles—usually γ—the goal is to extract this quantity from the measured value of $\mathcal{A}_{\text{CP}}$. Unfortunately, hadronic uncertainties affect this determination through the poorly known hadronic matrix elements in (90). In order to deal with this problem, we may proceed along one of the following two avenues:

(i) Amplitude relations can be used to eliminate the hadronic matrix elements. We distinguish between exact relations, using pure 'tree' decays of the kind $B^{\pm} \to K^{\pm} D$ [84, 85] or $B_c^{\pm} \to D_s^{\pm} D$ [86], and relations which follow from the flavour symmetries of strong interactions, i.e., isospin or $SU(3)_{\text{F}}$, and involve $B_{(s)} \to \pi\pi, \pi K, KK$ modes [87].

(ii) In decays of neutral B_q mesons, interference effects between B_q^0–$\bar{B}_q^0$ mixing and decay processes may induce 'mixing-induced CP violation'. If a single CKM amplitude governs the decay, the hadronic matrix elements cancel in the corresponding CP asymmetries; otherwise we again have to use amplitude relations. The most important example is the decay $B_d^0 \to J/\psi K_{\text{S}}$ [88].

Before discussing the features of neutral B_q mesons and B_q^0–$\bar{B}_q^0$ mixing in detail in Section 5, let us illustrate the use of amplitude relations for clean extractions of the UT angle γ from decays of charged B_u and B_c mesons.

---

[1]In order to calculate this quantity, an appropriate low-energy effective Hamiltonian having the same structure as (80) is used. The large theoretical uncertainties mentioned in Section 1 originate from a strong cancellation between the contributions of the QCD and EW penguins (caused by the large top-quark mass) and the associated hadronic matrix elements.

4 Amplitude relations

4.1 $B^\pm \to K^\pm D$

The prototype of the strategies using theoretically clean amplitude relations is provided by $B^\pm \to K^\pm D$ decays [84]. Looking at Fig. 12, we observe that $B^+ \to K^+ \bar{D}^0$ and $B^+ \to K^+ D^0$ are pure 'tree' decays. If we consider, in addition, the transition $B^+ \to D_+^0 K^+$, where D_+^0 denotes the CP eigenstate of the neutral D-meson system with eigenvalue $+1$,

$$|D_+^0\rangle = \frac{1}{\sqrt{2}} \left[|D^0\rangle + |\bar{D}^0\rangle \right] , \tag{98}$$

we obtain interference effects, which are described by

$$\sqrt{2} A(B^+ \to K^+ D_+^0) = A(B^+ \to K^+ D^0) + A(B^+ \to K^+ \bar{D}^0) \tag{99}$$
$$\sqrt{2} A(B^- \to K^- D_+^0) = A(B^- \to K^- \bar{D}^0) + A(B^- \to K^- D^0) . \tag{100}$$

These relations can be represented as two triangles in the complex plane. Since we have only to deal with tree-diagram-like topologies, we have moreover

$$A(B^+ \to K^+ \bar{D}^0) = A(B^- \to K^- D^0) \tag{101}$$
$$A(B^+ \to K^+ D^0) = A(B^- \to K^- \bar{D}^0) \times e^{2i\gamma} , \tag{102}$$

allowing a *theoretically clean* extraction of γ, as shown in Fig. 13. Unfortunately, these triangles are very squashed, since $B^+ \to K^+ D^0$ is colour-suppressed with respect to $B^+ \to K^+ \bar{D}^0$:

$$\left| \frac{A(B^+ \to K^+ D^0)}{A(B^+ \to K^+ \bar{D}^0)} \right| = \left| \frac{A(B^- \to K^- \bar{D}^0)}{A(B^- \to K^- D^0)} \right| \approx \frac{1}{\lambda} \frac{|V_{ub}|}{|V_{cb}|} \times \frac{a_2}{a_1} \approx 0.4 \times 0.3 = \mathcal{O}(0.1) , \tag{103}$$

where the phenomenological 'colour' factors were introduced in Subsection 3.3.3.

Another—more subtle—problem is related to the measurement of BR$(B^+ \to K^+ D^0)$. From the theoretical point of view, $D^0 \to K^- \ell^+ \nu$ would be ideal to measure this tiny branching ratio. However, because of the huge background from semileptonic B decays, we must rely on Cabibbo-allowed hadronic $D^0 \to f_{\rm NE}$ decays, such as $f_{\rm NE} = \pi^+ K^-$, $\rho^+ K^-$, ..., i.e., have to measure

$$B^+ \to K^+ D^0 [\to f_{\rm NE}] . \tag{104}$$

Unfortunately, we then encounter another decay path into the *same* final state $K^+ f_{\rm NE}$ through

$$B^+ \to K^+ \bar{D}^0 [\to f_{\rm NE}] , \tag{105}$$

where BR$(B^+ \to K^+ \bar{D}^0)$ is *larger* than BR$(B^+ \to K^+ D^0)$ by a factor of $\mathcal{O}(10^2)$, while $\bar{D}^0 \to f_{\rm NE}$ is doubly Cabibbo-suppressed, i.e., the corresponding branching ratio is suppressed with respect to the one of $D^0 \to f_{\rm NE}$ by a factor of $\mathcal{O}(10^{-2})$. Consequently, we obtain interference effects of $\mathcal{O}(1)$ between the decay chains in (104) and (105). However, if two different final states $f_{\rm NE}$ are considered, γ can be extracted [85], although this determination is then more involved than the original triangle approach presented in [84].

The angle γ can be determined in a variety of ways through CP-violating effects in pure tree decays of type $B \to D^{(*)} K^{(*)}$ [89]. Using the present B-factory data, the following results were obtained through a combination of various methods:

$$\gamma|_{D^{(*)} K^{(*)}} = \begin{cases} (62^{+35}_{-25})° & \text{(CKMfitter Collaboration [41]) ,} \\ (65 \pm 20)° & \text{(UTfit Collaboration [42]) .} \end{cases} \tag{106}$$

Here we have discarded a second solution given by $180° + \gamma|_{D^{(*)} K^{(*)}}$ in the third quadrant of the $\bar{\rho}$–$\bar{\eta}$ plane, as it is disfavoured by the global fits of the UT, and by the data for mixing-induced CP violation in pure tree decays of type $B_d \to D^\pm \pi^\mp$, $D^{*\pm} \pi^\mp$, ... [90]. A similar comment applies to the information from $B \to \pi\pi, \pi K$ modes [91].
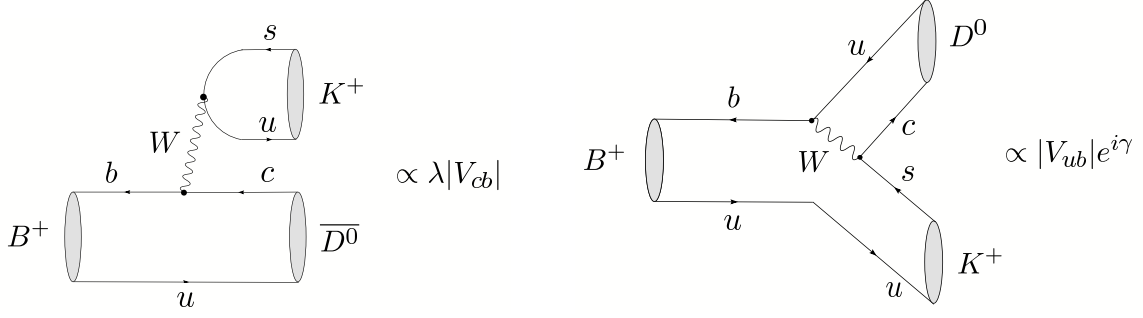
Fig. 12: Feynman diagrams contributing to $B^+ \to K^+ \bar{D}^0$ and $B^+ \to K^+ D^0$
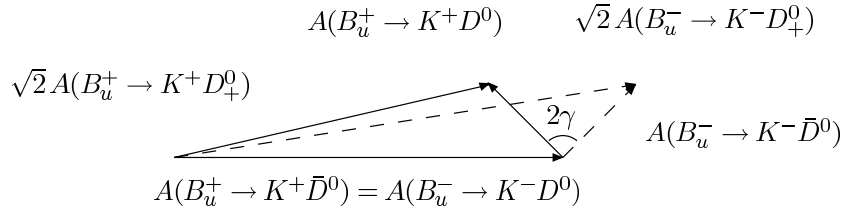
Fig. 13: The extraction of γ from $B^\pm \to K^\pm \{D^0, \bar{D}^0, D^0_+\}$ decays
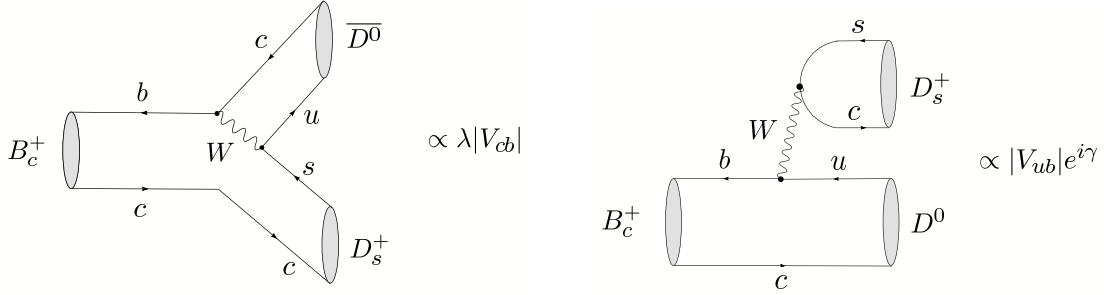
Fig. 14: Feynman diagrams contributing to $B^+_c \to D^+_s \bar{D}^0$ and $B^+ \to D^+_s D^0$

4.2 $B^\pm_c \to D^\pm_s D$

In addition to the 'conventional' $B^\pm_u$ mesons, there is yet another species of charged B mesons, the B_c-meson system, which consists of $B^+_c \sim c\bar{b}$ and $B^-_c \sim b\bar{c}$. These mesons were observed by the CDF Collaboration through their decay $B^+_c \to J/\psi \ell^+ \nu$, with the following mass and lifetime [92]:

$$M_{B_c} = (6.40 \pm 0.39 \pm 0.13)\,\text{GeV}, \quad \tau_{B_c} = (0.46^{+0.18}_{-0.16} \pm 0.03)\,\text{ps} . \tag{107}$$

Meanwhile, the D0 Collaboration observed the $B^+_c \to J/\psi\,\mu^+ X$ mode [93], which led to the following B_c mass and lifetime determinations:

$$M_{B_c} = (5.95^{+0.14}_{-0.13} \pm 0.34)\,\text{GeV}, \quad \tau_{B_c} = (0.448^{+0.123}_{-0.096} \pm 0.121)\,\text{ps} , \tag{108}$$

and CDF reported evidence for the $B^+_c \to J/\psi\pi^+$ channel [94], implying

$$M_{B_c} = (6.2870 \pm 0.0048 \pm 0.0011)\,\text{GeV} . \tag{109}$$

Since Run II of the Tevatron will provide further insights into B_c physics and a huge number of B_c mesons will be produced at LHCb, the natural question of how to explore CP violation with charged B_c

125
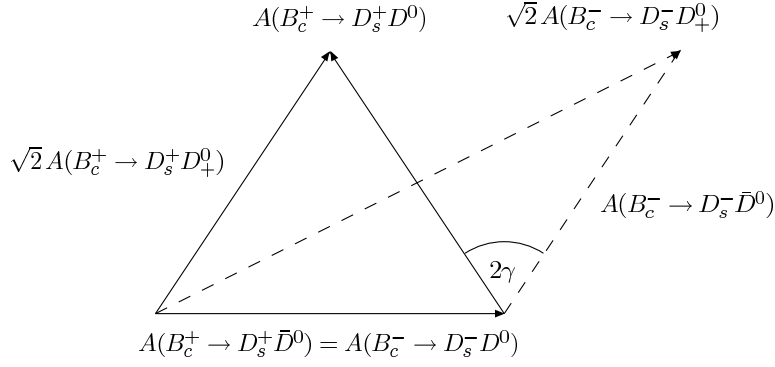
Fig. 15: The extraction of γ from $B_c^{\pm} \to D_s^{\pm}\{D^0, \bar{D}^0, D_+^0\}$ decays

decays arises, in particular whether an extraction of γ with the help of the triangle approach is possible. Such a determination is actually offered by $B_c^{\pm} \to D_s^{\pm}D$ decays, which are the B_c counterparts of the $B_u^{\pm} \to K^{\pm}D$ modes (see Fig. 14), and satisfy the following amplitude relations [95]:

$$\sqrt{2}A(B_c^+ \to D_s^+ D_+^0) = A(B_c^+ \to D_s^+ D^0) + A(B_c^+ \to D_s^+ \bar{D}^0) \tag{110}$$

$$\sqrt{2}A(B_c^- \to D_s^- D_+^0) = A(B_c^- \to D_s^- \bar{D}^0) + A(B_c^- \to D_s^- D^0) , \tag{111}$$

with

$$A(B_c^+ \to D_s^+ \bar{D}^0) = A(B_c^- \to D_s^- D^0) \tag{112}$$

$$A(B_c^+ \to D_s^+ D^0) = A(B_c^- \to D_s^- \bar{D}^0) \times e^{2i\gamma} . \tag{113}$$

At first sight, everything is completely analogous to the $B_u^{\pm} \to K^{\pm}D$ case. However, there is an important difference [86], which becomes obvious by comparing the Feynman diagrams shown in Figs. 12 and 14: in the $B_c^{\pm} \to D_s^{\pm}D$ system, the amplitude with the rather small CKM matrix element V_{ub} is not colour-suppressed, while the larger element V_{cb} comes with a colour-suppression factor. Therefore, we obtain

$$\left| \frac{A(B_c^+ \to D_s^+ D^0)}{A(B_c^+ \to D_s^+ \bar{D}^0)} \right| = \left| \frac{A(B_c^- \to D_s^- \bar{D}^0)}{A(B_c^- \to D_s^- D^0)} \right| \approx \frac{1}{\lambda} \frac{|V_{ub}|}{|V_{cb}|} \times \frac{a_1}{a_2} \approx 0.4 \times 3 = \mathcal{O}(1) , \tag{114}$$

and conclude that the two amplitudes are similar in size. In contrast to this favourable situation, in the decays $B_u^{\pm} \to K^{\pm}D$, the matrix element V_{ub} comes with the colour-suppression factor, resulting in a very stretched triangle. The extraction of γ from the $B_c^{\pm} \to D_s^{\pm}D$ triangles is illustrated in Fig. 15, which should be compared with the squashed $B_u^{\pm} \to K^{\pm}D$ triangles shown in Fig. 13. Another important advantage is that the interference effects arising from $D^0, \bar{D}^0 \to \pi^+ K^-$ are practically unimportant for the measurement of $\mathrm{BR}(B_c^+ \to D_s^+ D^0)$ and $\mathrm{BR}(B_c^+ \to D_s^+ \bar{D}^0)$ since the B_c-decay amplitudes are of the same order of magnitude. Consequently, the $B_c^{\pm} \to D_s^{\pm}D$ decays provide—from the theoretical point of view—the ideal realization of the 'triangle' approach to determine γ. On the other hand, the practical implementation still appears to be challenging, although detailed experimental feasibility studies for LHCb are strongly encouraged. The corresponding branching ratios were estimated in Ref. [96], with a pattern in accordance with (114).

5 Features of neutral B mesons

5.1 Schrödinger equation for B_q^0–$\bar{B}_q^0$ mixing

Within the SM, B_q^0–$\bar{B}_q^0$ mixing ($q \in \{d, s\}$) arises from the box diagrams shown in Fig. 16. Because of this phenomenon, an initially, i.e., at time $t = 0$, present B_q^0-meson state evolves into a time-dependent
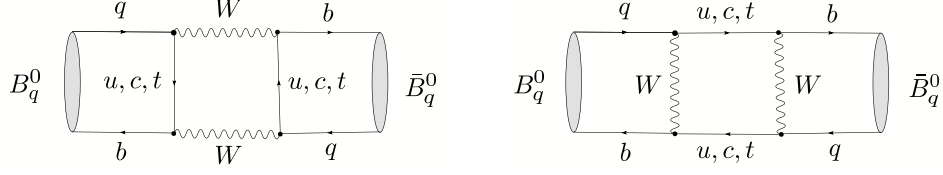
Fig. 16: Box diagrams contributing to B_q^0–$\bar{B}_q^0$ mixing in the SM ($q \in \{d, s\}$)

linear combination of B_q^0 and $\bar{B}_q^0$ states:

$$|B_q(t)\rangle = a(t)|B_q^0\rangle + b(t)|\bar{B}_q^0\rangle, \tag{115}$$

where $a(t)$ and $b(t)$ are governed by a Schrödinger equation of the following form:

$$i\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = H \cdot \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \equiv \left[\underbrace{\begin{pmatrix} M_0^{(q)} & M_{12}^{(q)} \\ M_{12}^{(q)*} & M_0^{(q)} \end{pmatrix}}_{\text{mass matrix}} - \frac{i}{2}\underbrace{\begin{pmatrix} \Gamma_0^{(q)} & \Gamma_{12}^{(q)} \\ \Gamma_{12}^{(q)*} & \Gamma_0^{(q)} \end{pmatrix}}_{\text{decay matrix}} \right] \cdot \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}.$$

The special form $H_{11} = H_{22}$ of the Hamiltonian H is an implication of the CPT theorem, i.e., of the invariance under combined CP and time-reversal (T) transformations.

It is straightforward to calculate the eigenstates $|B_\pm^{(q)}\rangle$ and eigenvalues $\lambda_\pm^{(q)}$ of (116):

$$|B_\pm^{(q)}\rangle = \frac{1}{\sqrt{1 + |\alpha_q|^2}}\left(|B_q^0\rangle \pm \alpha_q|\bar{B}_q^0\rangle\right) \tag{116}$$

$$\lambda_\pm^{(q)} = \left(M_0^{(q)} - \frac{i}{2}\Gamma_0^{(q)}\right) \pm \left(M_{12}^{(q)} - \frac{i}{2}\Gamma_{12}^{(q)}\right)\alpha_q , \tag{117}$$

where

$$\alpha_q e^{+i\left(\Theta_{\Gamma_{12}}^{(q)} + n'\pi\right)} = \sqrt{\frac{4|M_{12}^{(q)}|^2 e^{-i2\delta\Theta_{M/\Gamma}^{(q)}} + |\Gamma_{12}^{(q)}|^2}{4|M_{12}^{(q)}|^2 + |\Gamma_{12}^{(q)}|^2 - 4|M_{12}^{(q)}||\Gamma_{12}^{(q)}|\sin\delta\Theta_{M/\Gamma}^{(q)}}} . \tag{118}$$

Here we have written

$$M_{12}^{(q)} \equiv e^{i\Theta_{M_{12}}^{(q)}}|M_{12}^{(q)}|, \quad \Gamma_{12}^{(q)} \equiv e^{i\Theta_{\Gamma_{12}}^{(q)}}|\Gamma_{12}^{(q)}|, \quad \delta\Theta_{M/\Gamma}^{(q)} \equiv \Theta_{M_{12}}^{(q)} - \Theta_{\Gamma_{12}}^{(q)} , \tag{119}$$

and have introduced the quantity $n' \in \{0, 1\}$ to parametrize the sign of the square root in (118).

Evaluating the dispersive parts of the box diagrams shown in Fig. 16, which are dominated by internal top-quark exchanges, yields (for a more detailed discussion, see Ref. [17]):

$$M_{12}^{(q)} = \frac{G_F^2 M_W^2}{12\pi^2}\eta_B M_{B_q} f_{B_q}^2 \hat{B}_{B_q}\left(V_{tq}^* V_{tb}\right)^2 S_0(x_t)e^{i(\pi - \phi_{\mathrm{CP}}(B_q))} , \tag{120}$$

where $\phi_{\mathrm{CP}}(B_q)$ is a convention-dependent phase, which is defined in analogy to (94). The short-distance physics is encoded in the 'Inami–Lim' function $S_0(x_t \equiv m_t^2/M_W^2)$ [97], which can be written—to a good approximation—in the SM as [98]

$$S_0(x_t) = 2.40 \times \left[\frac{m_t}{167\,\mathrm{GeV}}\right]^{1.52} , \tag{121}$$

and in the perturbative QCD correction factor $\eta_B = 0.55 \pm 0.01$ [99], which does *not* depend on $q \in \{d, s\}$, i.e., is the same for B_d and B_s mesons. On the other hand, the non-perturbative physics is

described by the quantities $f_{B_q}\hat{B}_{B_q}^{1/2}$, involving—in addition to the B_q decay constant f_{B_q}—the 'bag' parameter $\hat{B}_{B_q}$, which is related to the hadronic matrix element $\langle \bar{B}_q^0|(\bar{b}q)_{\mathrm{V-A}}(\bar{b}q)_{\mathrm{V-A}}|B_q^0\rangle$. These non-perturbative parameters can be determined through QCD sum-rule calculations [100] or lattice studies. Concerning the latter analyses, the front runners are now unquenched calculations with 2 or 3 dynamical quarks. Despite tremendous progress, the results still suffer from several uncertainties. For the analysis of the mixing parameters discussed below [101], we use two sets of parameters from the JLQCD [102] and HPQCD [103] lattice collaborations:

$$
\begin{aligned}
f_{B_d}\hat{B}_{B_d}^{1/2}\Big|_{\mathrm{JLQCD}} &= (0.215 \pm 0.019^{+0}_{-0.023})\,\mathrm{GeV} \\
f_{B_s}\hat{B}_{B_s}^{1/2}\Big|_{\mathrm{JLQCD}} &= (0.245 \pm 0.021^{+0.003}_{-0.002})\,\mathrm{GeV}\,,
\end{aligned}
\tag{122}
$$

which were obtained for two flavours of dynamical light ('Wilson') quarks, and

$$
\begin{aligned}
f_{B_d}\hat{B}_{B_d}^{1/2}\Big|_{\mathrm{(HP+JL)QCD}} &= (0.244 \pm 0.026)\,\mathrm{GeV} \\
f_{B_s}\hat{B}_{B_s}^{1/2}\Big|_{\mathrm{(HP+JL)QCD}} &= (0.295 \pm 0.036)\,\mathrm{GeV}\,,
\end{aligned}
\tag{123}
$$

where f_{B_q} comes from HPQCD (3 dynamical flavours) and $\hat{B}_{B_q}$ from JLQCD as no value for this parameter is available from the former collaboration [104].

If we calculate also the absorptive parts of the box diagrams in Fig 16, we obtain

$$
\frac{\Gamma_{12}^{(q)}}{M_{12}^{(q)}} \approx -\frac{3\pi}{2S_0(x_t)}\left(\frac{m_b^2}{M_W^2}\right) = \mathcal{O}(m_b^2/m_t^2) \ll 1\,.
\tag{124}
$$

Consequently, we may expand (118) in $\Gamma_{12}^{(q)}/M_{12}^{(q)}$. Neglecting second-order terms, we arrive at

$$
\alpha_q = \left[1 + \frac{1}{2}\left|\frac{\Gamma_{12}^{(q)}}{M_{12}^{(q)}}\right|\sin\delta\Theta_{M/\Gamma}^{(q)}\right]e^{-i\left(\Theta_{M_{12}}^{(q)}+n'\pi\right)}\,.
\tag{125}
$$

The deviation of $|\alpha_q|$ from 1 measures CP violation in B_q^0–$\bar{B}_q^0$ oscillations, and can be probed through the following 'wrong-charge' lepton asymmetries:

$$
\mathcal{A}_{\mathrm{SL}}^{(q)} \equiv \frac{\Gamma(B_q^0(t)\to\ell^-\bar{\nu}X) - \Gamma(\bar{B}_q^0(t)\to\ell^+\nu X)}{\Gamma(B_q^0(t)\to\ell^-\bar{\nu}X) + \Gamma(\bar{B}_q^0(t)\to\ell^+\nu X)} = \frac{|\alpha_q|^4 - 1}{|\alpha_q|^4 + 1} \approx \left|\frac{\Gamma_{12}^{(q)}}{M_{12}^{(q)}}\right|\sin\delta\Theta_{M/\Gamma}^{(q)}\,.
\tag{126}
$$

Because of $|\Gamma_{12}^{(q)}|/|M_{12}^{(q)}| \propto m_b^2/m_t^2$ and $\sin\delta\Theta_{M/\Gamma}^{(q)} \propto m_c^2/m_b^2$, the asymmetry $\mathcal{A}_{\mathrm{SL}}^{(q)}$ is suppressed by a factor of $m_c^2/m_t^2 = \mathcal{O}(10^{-4})$ and is hence tiny in the SM. However, this observable may be enhanced through NP effects, thereby representing an interesing probe for physics beyond the SM [105, 106]. The current experimental average for the B_d-meson system compiled by the 'Heavy Flavour Averaging Group' [61] reads as follows:

$$
\mathcal{A}_{\mathrm{SL}}^{(d)} = 0.0030 \pm 0.0078\,,
\tag{127}
$$

and does not indicate any non-vanishing effect.

5.2 Mixing parameters

Let us denote the masses of the eigenstates of (116) by $M_{\mathrm{H}}^{(q)}$ ('heavy') and $M_{\mathrm{L}}^{(q)}$ ('light'). It is then useful to introduce

$$
M_q \equiv \frac{M_{\mathrm{H}}^{(q)} + M_{\mathrm{L}}^{(q)}}{2} = M_0^{(q)}\,,
\tag{128}
$$

as well as the mass difference

$$\Delta M_q \equiv M_{\rm H}^{(q)} - M_{\rm L}^{(q)} = 2|M_{12}^{(q)}| > 0 \,, \tag{129}$$

which is by definition *positive*. While B_d^0–$\bar{B}_d^0$ mixing is well established and

$$\Delta M_d = (0.507 \pm 0.004)\,{\rm ps}^{-1} \tag{130}$$

known with impressive experimental accuracy [61], only lower bounds on ΔM_s were available, for many years, from the LEP (CERN) experiments and SLD (SLAC) [107]. In the spring of 2006, ΔM_s could eventually be pinned down at the Tevatron: the D0 Collaboration reported a two-sided bound

$$17\,{\rm ps}^{-1} < \Delta M_s < 21\,{\rm ps}^{-1} \quad (90\%\ {\rm C.L.})\,, \tag{131}$$

corresponding to a 2.5 σ signal at $\Delta M_s = 19\,{\rm ps}^{-1}$ [108], and CDF announced the following result [109]:

$$\Delta M_s = \left[17.31^{+0.33}_{-0.18}({\rm stat}) \pm 0.07({\rm syst})\right]{\rm ps}^{-1}\,. \tag{132}$$

The decay widths $\Gamma_{\rm H}^{(q)}$ and $\Gamma_{\rm L}^{(q)}$ of the mass eigenstates, which correspond to $M_{\rm H}^{(q)}$ and $M_{\rm L}^{(q)}$, respectively, satisfy

$$\Delta\Gamma_q \equiv \Gamma_{\rm H}^{(q)} - \Gamma_{\rm L}^{(q)} = \frac{4\,{\rm Re}\left[M_{12}^{(q)}\Gamma_{12}^{(q)*}\right]}{\Delta M_q}\,, \tag{133}$$

whereas

$$\Gamma_q \equiv \frac{\Gamma_{\rm H}^{(q)} + \Gamma_{\rm L}^{(q)}}{2} = \Gamma_0^{(q)}\,. \tag{134}$$

There is the following interesting relation:

$$\frac{\Delta\Gamma_q}{\Gamma_q} \approx -\frac{3\pi}{2S_0(x_t)}\left(\frac{m_b^2}{M_W^2}\right)x_q = -\mathcal{O}(10^{-2}) \times x_q\,, \tag{135}$$

where

$$x_q \equiv \frac{\Delta M_q}{\Gamma_q} = \begin{cases} 0.771 \pm 0.012 & (q = d) \\ \mathcal{O}(20) & (q = s) \end{cases} \tag{136}$$

is often referred to as *the B_q^0–$\bar{B}_q^0$ 'mixing parameter'*[2]. Consequently, we observe that $\Delta\Gamma_d/\Gamma_d \sim 10^{-2}$ is negligibly small, while $\Delta\Gamma_s/\Gamma_s \sim 10^{-1}$ may be sizeable. In fact, as was reviewed in Ref. [110], the state of the art of calculations of these quantities is given as follows:

$$\frac{|\Delta\Gamma_d|}{\Gamma_d} = (3 \pm 1.2) \times 10^{-3}, \quad \frac{|\Delta\Gamma_s|}{\Gamma_s} = 0.12 \pm 0.05\,. \tag{137}$$

Recently, the first results for $\Delta\Gamma_s$ were reported from the Tevatron, using the $B_s^0 \to J/\psi\phi$ channel [111]:

$$\frac{|\Delta\Gamma_s|}{\Gamma_s} = \begin{cases} 0.65^{+0.25}_{-0.33} \pm 0.01 & ({\rm CDF}\ [112]) \\ 0.24^{+0.28+0.03}_{-0.38-0.04} & ({\rm D0}\ [113])\,. \end{cases} \tag{138}$$

It will be interesting to follow the evolution of the data for this quantity.

In Sections 7.1 and 10.1, we give detailed discussions of the theoretical interpretation of the data for the B_q^0–$\bar{B}_q^0$ mixing parameters.

---

[2]Note that $\Delta\Gamma_q/\Gamma_q$ is negative in the SM because of the minus sign in (135).

5.3   Time-dependent decay rates

The time evolution of initially, i.e., at $t = 0$, pure B_q^0- and $\bar{B}_q^0$-meson states is given by

$$|B_q^0(t)\rangle = f_+^{(q)}(t)|B_q^0\rangle + \alpha_q f_-^{(q)}(t)|\bar{B}_q^0\rangle \qquad (139)$$

and

$$|\bar{B}_q^0(t)\rangle = \frac{1}{\alpha_q} f_-^{(q)}(t)|B_q^0\rangle + f_+^{(q)}(t)|\bar{B}_q^0\rangle \,, \qquad (140)$$

respectively, with

$$f_\pm^{(q)}(t) = \frac{1}{2}\left[e^{-i\lambda_+^{(q)}t} \pm e^{-i\lambda_-^{(q)}t}\right]. \qquad (141)$$

These time-dependent state vectors allow the calculation of the corresponding transition rates. To this end, it is useful to introduce

$$|g_\pm^{(q)}(t)|^2 = \frac{1}{4}\left[e^{-\Gamma_{\rm L}^{(q)}t} + e^{-\Gamma_{\rm H}^{(q)}t} \pm 2\, e^{-\Gamma_q t}\cos(\Delta M_q t)\right] \qquad (142)$$

$$g_-^{(q)}(t)\, g_+^{(q)}(t)^* = \frac{1}{4}\left[e^{-\Gamma_{\rm L}^{(q)}t} - e^{-\Gamma_{\rm H}^{(q)}t} + 2\, i\, e^{-\Gamma_q t}\sin(\Delta M_q t)\right], \qquad (143)$$

as well as

$$\xi_f^{(q)} = e^{-i\Theta_{M_{12}}^{(q)}}\frac{A(\bar{B}_q^0 \to f)}{A(B_q^0 \to f)}, \quad \xi_{\bar{f}}^{(q)} = e^{-i\Theta_{M_{12}}^{(q)}}\frac{A(\bar{B}_q^0 \to \bar{f})}{A(B_q^0 \to \bar{f})}\,. \qquad (144)$$

Looking at (120), we find

$$\Theta_{M_{12}}^{(q)} = \pi + 2\arg(V_{tq}^* V_{tb}) - \phi_{\rm CP}(B_q)\,, \qquad (145)$$

and observe that this phase depends on the chosen CKM and CP phase conventions specified in (11) and (94), respectively. However, these dependences are cancelled through the amplitude ratios in (144), so that $\xi_f^{(q)}$ and $\xi_{\bar{f}}^{(q)}$ are *convention-independent* observables. Whereas n' enters the functions in (141) through (117), the dependence on this parameter is cancelled in (142) and (143) through the introduction of the *positive* mass difference ΔM_q [see (129)]. Combining the formulae listed above, we eventually arrive at the following transition rates for decays of initially, i.e., at $t = 0$, present B_q^0 or $\bar{B}_q^0$ mesons:

$$\Gamma(\overset{\scriptscriptstyle(-)}{B_q^0}(t) \to f) = \left[|g_\mp^{(q)}(t)|^2 + |\xi_f^{(q)}|^2|g_\pm^{(q)}(t)|^2 - 2\,{\rm Re}\left\{\xi_f^{(q)}g_\pm^{(q)}(t)g_\mp^{(q)}(t)^*\right\}\right]\tilde{\Gamma}_f\,, \qquad (146)$$

where the time-independent rate $\tilde{\Gamma}_f$ corresponds to the 'unevolved' decay amplitude $A(B_q^0 \to f)$, and can be calculated by performing the usual phase-space integrations. The rates into the CP-conjugate final state $\bar{f}$ can straightforwardly be obtained from (146) by making the substitutions

$$\tilde{\Gamma}_f \;\to\; \tilde{\Gamma}_{\bar{f}}, \quad \xi_f^{(q)} \;\to\; \xi_{\bar{f}}^{(q)}\,. \qquad (147)$$

5.4   'Untagged' rates

The expected sizeable width difference $\Delta\Gamma_s$ may provide interesting studies of CP violation through 'untagged' B_s rates (see Ref. [111] and [114]– [117]), which are defined as

$$\langle \Gamma(B_s(t) \to f)\rangle \equiv \Gamma(B_s^0(t) \to f) + \Gamma(\bar{B}_s^0(t) \to f)\,, \qquad (148)$$

and are characterized by the feature that we do not distinguish between initially, i.e., at time $t = 0$, present B_s^0 or $\bar{B}_s^0$ mesons. If we consider a final state f to which both a B_s^0 and a $\bar{B}_s^0$ may decay, and use the expressions in (146), we find

$$\langle \Gamma(B_s(t) \to f)\rangle \propto [\cosh(\Delta\Gamma_s t/2) - \mathcal{A}_{\Delta\Gamma}(B_s \to f)\sinh(\Delta\Gamma_s t/2)]\, e^{-\Gamma_s t}\,, \qquad (149)$$

with

$$\mathcal{A}_{\Delta\Gamma}(B_s \to f) \equiv \frac{2\,\mathrm{Re}\,\xi_f^{(s)}}{1 + \left|\xi_f^{(s)}\right|^2} \,. \tag{150}$$

We observe that the rapidly oscillating $\Delta M_s t$ terms cancel, and that we may obtain information about the phase structure of the observable $\xi_f^{(s)}$, thereby providing valuable insights into CP violation.

Following these lines, for instance, the untagged observables offered by the angular distribution of the $B_s \to K^{*+}K^{*-}, K^{*0}\bar{K}^{*0}$ decay products allow a determination of the UT angle γ, provided $\Delta\Gamma_s$ is actually sizeable [115]. Untagged B_s-decay rates are interesting in terms of efficiency, acceptance and purity, and are already applied for the physics analyses at the Tevatron. Later on, they will help to fully exploit the physics potential of the B_s-meson system at the LHC.

5.5 CP asymmetries

A particularly simple—but also very interesting—situation arises if we restrict ourselves to decays of neutral B_q mesons into final states f that are eigenstates of the CP operator, i.e., satisfy the relation

$$(\mathcal{CP})|f\rangle = \pm|f\rangle \,. \tag{151}$$

Consequently, we have $\xi_f^{(q)} = \xi_{\bar{f}}^{(q)}$ in this case, as can be seen in (144). Using the decay rates in (146), we find that the corresponding time-dependent CP asymmetry is given by

$$\begin{aligned}
\mathcal{A}_{\mathrm{CP}}(t) &\equiv \frac{\Gamma(B_q^0(t) \to f) - \Gamma(\bar{B}_q^0(t) \to f)}{\Gamma(B_q^0(t) \to f) + \Gamma(\bar{B}_q^0(t) \to f)} \\
&= \left[\frac{\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_q \to f)\cos(\Delta M_q t) + \mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_q \to f)\sin(\Delta M_q t)}{\cosh(\Delta\Gamma_q t/2) - \mathcal{A}_{\Delta\Gamma}(B_q \to f)\sinh(\Delta\Gamma_q t/2)}\right],
\end{aligned} \tag{152}$$

with

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_q \to f) \equiv \frac{1 - \left|\xi_f^{(q)}\right|^2}{1 + \left|\xi_f^{(q)}\right|^2}, \qquad \mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_q \to f) \equiv \frac{2\,\mathrm{Im}\,\xi_f^{(q)}}{1 + \left|\xi_f^{(q)}\right|^2} \,. \tag{153}$$

Because of the relation

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_q \to f) = \frac{|A(B_q^0 \to f)|^2 - |A(\bar{B}_q^0 \to \bar{f})|^2}{|A(B_q^0 \to f)|^2 + |A(\bar{B}_q^0 \to \bar{f})|^2} \,, \tag{154}$$

this observable measures the direct CP violation in the decay $B_q \to f$, which originates from the interference between different weak amplitudes, as we have seen in (97). On the other hand, the interesting *new* aspect of (152) is due to $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_q \to f)$, which originates from interference effects between B_q^0–$\bar{B}_q^0$ mixing and decay processes, and describes 'mixing-induced' CP violation. Finally, the width difference $\Delta\Gamma_q$, which may be sizeable in the B_s-meson system, provides access to $\mathcal{A}_{\Delta\Gamma}(B_q \to f)$ introduced in (150). However, this observable is not independent from $\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_q \to f)$ and $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_q \to f)$, satisfying

$$\left[\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_q \to f)\right]^2 + \left[\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_q \to f)\right]^2 + \left[\mathcal{A}_{\Delta\Gamma}(B_q \to f)\right]^2 = 1 \,. \tag{155}$$

In order to calculate $\xi_f^{(q)}$, we use the general expressions (95) and (96), where $e^{-i\phi_{\mathrm{CP}}(f)} = \pm 1$ because of (151), and $\phi_{\mathrm{CP}}(B) = \phi_{\mathrm{CP}}(B_q)$. If we insert these amplitude parametrizations into (144) and take (145) into account, we observe that the phase-convention-dependent quantity $\phi_{\mathrm{CP}}(B_q)$ cancels, and finally arrive at

$$\xi_f^{(q)} = \mp e^{-i\phi_q}\left[\frac{e^{+i\varphi_1}|A_1|e^{i\delta_1} + e^{+i\varphi_2}|A_2|e^{i\delta_2}}{e^{-i\varphi_1}|A_1|e^{i\delta_1} + e^{-i\varphi_2}|A_2|e^{i\delta_2}}\right], \tag{156}$$

where

$$\phi_q \equiv 2 \arg(V_{tq}^* V_{tb}) = \left\{ \begin{array}{ll} +2\beta & (q = d) \\ -2\delta\gamma & (q = s) \end{array} \right. \tag{157}$$

is associated with the CP-violating weak B_q^0–$\bar{B}_q^0$ mixing phase arising in the SM; β and $\delta\gamma$ refer to the corresponding angles in the unitarity triangles shown in Fig. 3.

In analogy to (97), the caclulation of $\xi_f^{(q)}$ is—in general—also affected by large hadronic uncertainties. However, if one CKM amplitude plays the dominant role in the $B_q \to f$ transition, we obtain

$$\xi_f^{(q)} = \mp e^{-i\phi_q} \left[\frac{e^{+i\phi_f/2}|M_f|e^{i\delta_f}}{e^{-i\phi_f/2}|M_f|e^{i\delta_f}} \right] = \mp e^{-i(\phi_q - \phi_f)} , \tag{158}$$

and observe that the hadronic matrix element $|M_f|e^{i\delta_f}$ cancels in this expression. Since the requirements for direct CP violation discussed above are no longer satisfied, direct CP violation vanishes in this important special case, i.e., $\mathcal{A}_{\text{CP}}^{\text{dir}}(B_q \to f) = 0$. On the other hand, this is *not* the case for the mixing-induced CP asymmetry. In particular,

$$\mathcal{A}_{\text{CP}}^{\text{mix}}(B_q \to f) = \pm \sin\phi \tag{159}$$

is now governed by the CP-violating weak phase difference $\phi \equiv \phi_q - \phi_f$ and is not affected by hadronic uncertainties. The corresponding time-dependent CP asymmetry takes then the simple form

$$\left. \frac{\Gamma(B_q^0(t) \to f) - \Gamma(\bar{B}_q^0(t) \to \bar{f})}{\Gamma(B_q^0(t) \to f) + \Gamma(\bar{B}_q^0(t) \to \bar{f})} \right|_{\Delta\Gamma_q = 0} = \pm \sin\phi \, \sin(\Delta M_q t) , \tag{160}$$

and allows an elegant determination of $\sin\phi$.

6 How could new physics enter?

Using the concept of the low-energy effective Hamiltonians introduced in Section 3.3.2, we may address this important question in a systematic manner [118]:

– NP may modify the 'strength' of the SM operators through new short-distance functions which depend on the NP parameters, such as the masses of charginos, squarks, charged Higgs particles and $\tan\bar{\beta} \equiv v_2/v_1$ in the 'minimal supersymmetric SM' (MSSM). The NP particles may enter in box and penguin topologies, and are 'integrated out' as the W boson and top quark in the SM. Consequently, the initial conditions for the renormalization-group evolution take the following form:

$$C_k \to C_k^{\text{SM}} + C_k^{\text{NP}} . \tag{161}$$

It should be emphasized that the NP pieces C_k^{NP} may also involve new CP-violating phases which are *not* related to the CKM matrix.

– NP may enhance the operator basis:

$$\{Q_k\} \to \{Q_k^{\text{SM}}, Q_l^{\text{NP}}\} , \tag{162}$$

so that operators which are not present (or strongly suppressed) in the SM may actually play an important role. In this case, we encounter, in general, also new sources for flavour and CP violation.

The B-meson system offers a variety of processes and strategies for the exploration of CP violation [12, 119], as we have illustrated in Fig. 17 through a collection of prominent examples. We see that there are processes with a very *different* dynamics that are—in the SM—sensitive to the *same* angles of the UT. Moreover, rare B- and K-meson decays [120], which originate from loop effects in the SM,
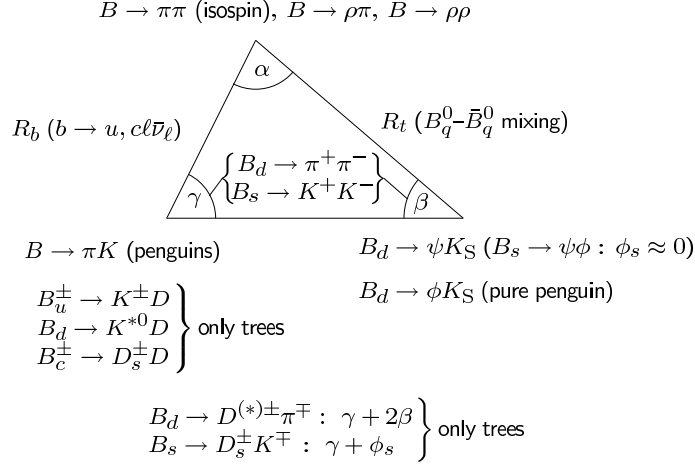
Fig. 17: A brief roadmap of B-decay strategies for the exploration of CP violation

provide complementary insights into flavour physics and interesting correlations with the CP-B sector; key examples are $B \to X_s \gamma$ and the exclusive modes $B \to K^* \gamma$, $B \to \rho \gamma$, as well as $B_{s,d} \to \mu^+ \mu^-$ and $K^+ \to \pi^+ \nu \bar{\nu}$, $K_{\rm L} \to \pi^0 \nu \bar{\nu}$.

In the presence of NP contributions, the subtle interplay between the different processes could well be disturbed. There are two popular avenues for NP to enter the roadmap of quark-flavour physics:

– B_q^0–$\bar{B}_q^0$ *mixing:* NP could enter through the exchange of new particles in the box diagrams, or through new contributions at the tree level. In general, we may write

$$M_{12}^{(q)} = M_{12}^{q,\rm SM} \left(1 + \kappa_q e^{i \sigma_q} \right), \tag{163}$$

where the expression for $M_{12}^{q,\rm SM}$ can be found in (120). Consequently, we obtain

$$\Delta M_q = \Delta M_q^{\rm SM} + \Delta M_q^{\rm NP} = \Delta M_q^{\rm SM} \left| 1 + \kappa_q e^{i \sigma_q} \right|, \tag{164}$$

$$\phi_q = \phi_q^{\rm SM} + \phi_q^{\rm NP} = \phi_q^{\rm SM} + \arg(1 + \kappa_q e^{i \sigma_q}), \tag{165}$$

with $\Delta M_q^{\rm SM}$ and $\phi_q^{\rm SM}$ given in (129) and (157), respectively. Using dimensional arguments borrowed from effective field theory [121, 122], it can be shown that $\Delta M_q^{\rm NP}/\Delta M_q^{\rm SM} \sim 1$ and $\phi_q^{\rm NP}/\phi_q^{\rm SM} \sim 1$ could—in principle—be possible for a NP scale $\Lambda_{\rm NP}$ in the TeV regime; such a pattern may also arise in specific NP scenarios. Introducing

$$\rho_q \equiv \left| \frac{\Delta M_q}{\Delta M_q^{\rm SM}} \right| = \sqrt{1 + 2\kappa_q \cos \sigma_q + \kappa_q^2}, \tag{166}$$

the measured values of the mass differences ΔM_q can be converted into constraints in NP parameter space through the contours shown in Fig. 18. Further constraints are implied by the NP phases $\phi_q^{\rm NP}$, which can be probed through mixing-induced CP asymmetries, through the curves in the σ_q–κ_q plane shown in Fig. 19. Interestingly, κ_q is bounded from below for any value of $\phi_q^{\rm NP} \neq 0$. For example, even a small phase $|\phi_q^{\rm NP}| = 10°$ implies a clean lower bound of $\kappa_q \geq 0.17$, i.e., NP contributions of at most 17% [101].

– *Decay amplitudes:* NP has typically a small effect if SM tree processes play the dominant role. However, NP could well have a significant impact on the FCNC sector: new particles may enter in penguin or box diagrams, or new FCNC contributions may even be generated at the tree level. In fact, sizeable contributions arise generically in field-theoretical estimates with $\Lambda_{\rm NP} \sim$ TeV [123], as well as in specific NP models.
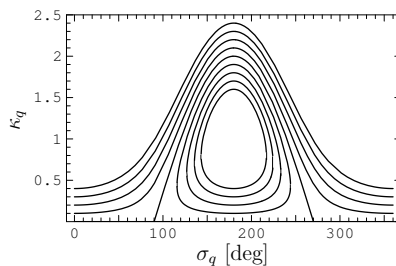
Fig. 18: The dependence of κ_q on σ_q for values of ρ_q varied between 1.4 (most upper curve) and 0.6 (most inner curve), in steps of 0.1
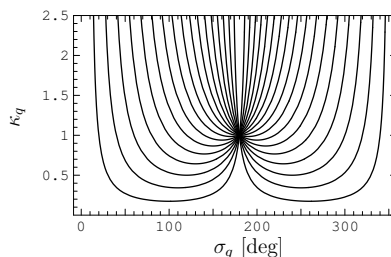
Fig. 19: The dependence of κ_q on σ_q for values of ϕ_q^{NP} varied between $\pm 10°$ (lower curves) and $\pm 170°$ in steps of $10°$: the curves for $0° < \sigma_q < 180°$ and $180° < \sigma_q < 360°$ correspond to positive and negative values of ϕ_q^{NP}, respectively.

Concerning model-dependent NP analyses, SUSY scenarios in particular have received a lot of attention; for a selection of recent studies, see Refs. [124]– [129]. Examples of other fashionable NP scenarios are left–right-symmetric models [130], scenarios with extra dimensions [131], models with an extra Z' [132], 'little Higgs' scenarios [133], and models with a fourth generation [134].

The simplest extension of the SM is given by models with 'minimal flavour violation' (MFV). Following the characterization given in Ref. [135], the flavour-changing processes are here still governed by the CKM matrix—in particular there are no new sources for CP violation—and the only relevant operators are those present in the SM (for an alternative definition, see Ref. [136]). Specific examples are the Two-Higgs Doublet Model II, the MSSM without new sources of flavour violation and $\tan \bar{\beta}$ not too large, models with one extra universal dimension and the simplest little Higgs models. Because of their simplicity, the extensions of the SM with MFV show several correlations between various observables, thereby allowing for powerful tests of this scenario [137]. A systematic discussion of models with 'next-to-minimal flavour violation' was recently given in Ref. [138].

There are other fascinating probes for the search of NP. Important examples are the D-meson system [139], electric dipole moments [140], or flavour-violating charged lepton decays [141]. Since a discussion of these topics is beyond the scope of these lectures, the interested reader should consult the corresponding references. Let us next have a closer look at prominent B decays, with a particular emphasis of the impact of NP.

7 Status of important B-factory benchmark modes

7.1 $B_d^0 \rightarrow J/\psi K_{\mathrm{S}}$

7.1.1 Basic formulae

This decay has a CP-odd final state, and originates from $\bar{b} \rightarrow \bar{c}c\bar{s}$ quark-level transitions. Consequently, as we discussed in Section 3.3.1, it receives contributions both from tree and from penguin topologies,
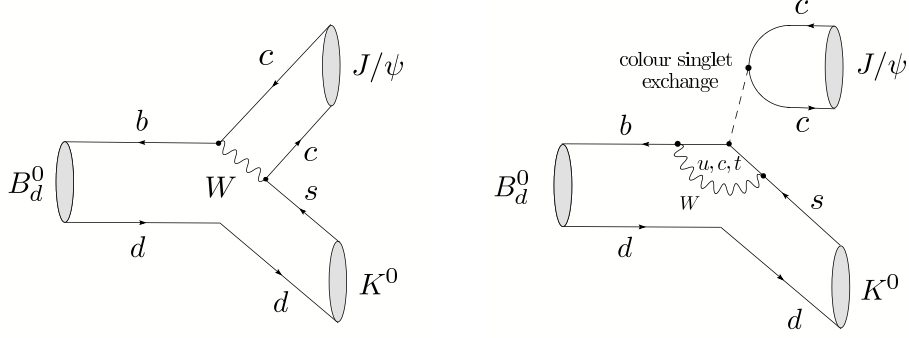
Fig. 20: Feynman diagrams contributing to $B_d^0 \to J/\psi K^0$ decays

as can be seen in Fig. 20. In the SM, the decay amplitude can hence be written as follows [142]:

$$A(B_d^0 \to J/\psi K_{\rm S}) = \lambda_c^{(s)} \left(A_{\rm T}^{c'} + A_{\rm P}^{c'} \right) + \lambda_u^{(s)} A_{\rm P}^{u'} + \lambda_t^{(s)} A_{\rm P}^{t'} \,. \tag{167}$$

Here the

$$\lambda_q^{(s)} \equiv V_{qs} V_{qb}^* \tag{168}$$

are CKM factors, $A_{\rm T}^{c'}$ is the CP-conserving strong tree amplitude, while the $A_{\rm P}^{q'}$ describe the penguin topologies with internal q quarks ($q \in \{u, c, t\}$), including QCD and EW penguins; the primes remind us that we are dealing with a $\bar{b} \to \bar{s}$ transition. If we eliminate now $\lambda_t^{(s)}$ through (79) and apply the Wolfenstein parametrization, we obtain

$$A(B_d^0 \to J/\psi K_{\rm S}) \propto \left[1 + \lambda^2 a e^{i\theta} e^{i\gamma} \right] \,, \tag{169}$$

where

$$a e^{i\vartheta} \equiv \left(\frac{R_b}{1 - \lambda^2} \right) \left[\frac{A_{\rm P}^{u'} - A_{\rm P}^{t'}}{A_{\rm T}^{c'} + A_{\rm P}^{c'} - A_{\rm P}^{t'}} \right] \tag{170}$$

is a hadronic parameter. Using now the formalism of Section 5.5 yields

$$\xi_{\psi K_{\rm S}}^{(d)} = +e^{-i\phi_d} \left[\frac{1 + \lambda^2 a e^{i\vartheta} e^{-i\gamma}}{1 + \lambda^2 a e^{i\vartheta} e^{+i\gamma}} \right]. \tag{171}$$

Unfortunately, $a e^{i\vartheta}$, which is a measure for the ratio of the $B_d^0 \to J/\psi K_{\rm S}$ penguin to tree contributions, can only be estimated with large hadronic uncertainties. However, since this parameter enters (171) in a doubly Cabibbo-suppressed way, its impact on the CP-violating observables is practically negligible. We can put this important statement on a more quantitative basis by making the plausible assumption that $a = \mathcal{O}(\bar{\lambda}) = \mathcal{O}(0.2) = \mathcal{O}(\lambda)$, where $\bar{\lambda}$ is a 'generic' expansion parameter:

$$\mathcal{A}_{\rm CP}^{\rm dir}(B_d \to J/\psi K_{\rm S}) = 0 + \mathcal{O}(\bar{\lambda}^3) \tag{172}$$

$$\mathcal{A}_{\rm CP}^{\rm mix}(B_d \to J/\psi K_{\rm S}) = -\sin\phi_d + \mathcal{O}(\bar{\lambda}^3) \overset{\rm SM}{=} -\sin 2\beta + \mathcal{O}(\bar{\lambda}^3) \,. \tag{173}$$

Consequently, (173) allows an essentially *clean* determination of $\sin 2\beta$ [88].

7.1.2 Experimental status

Since the CKM fits performed within the SM pointed to a large value of $\sin 2\beta$, $B_d^0 \to J/\psi K_{\rm S}$ offered the exciting perspective of exhibiting *large* mixing-induced CP violation. In 2001, the measurement of $\mathcal{A}_{\rm CP}^{\rm mix}(B_d \to J/\psi K_{\rm S})$ allowed indeed the first observation of CP violation *outside* the K-meson

system [5]. The most recent data are still not showing any signal for *direct* CP violation in $B_d^0 \to J/\psi K_S$ within the current uncertainties, as is expected from (172). The current world average reads [61]

$$\mathcal{A}_{CP}^{dir}(B_d \to J/\psi K_S) = 0.026 \pm 0.041. \tag{174}$$

As far as (173) is concerned, we have

$$(\sin 2\beta)_{\psi K_S} \equiv -\mathcal{A}_{CP}^{mix}(B_d \to J/\psi K_S) = \begin{cases} 0.722 \pm 0.040 \pm 0.023 & \text{(BaBar [143])} \\ 0.652 \pm 0.039 \pm 0.020 & \text{(Belle [144])}, \end{cases} \tag{175}$$

which gives the following world average [61]:

$$(\sin 2\beta)_{\psi K_S} = 0.687 \pm 0.032 . \tag{176}$$

In the SM, the theoretical uncertainties are generically expected to be below the 0.01 level; significantly smaller effects are found in Ref. [145], whereas a fit performed in Ref. [146] yields a theoretical penguin uncertainty comparable to the present experimental systematic error. A possibility to control these uncertainties is provided by the $B_s^0 \to J/\psi K_S$ channel [142], which can be explored at the LHC [147].

In Ref. [121], a set of observables to search for NP contributions to the $B \to J/\psi K$ decay amplitudes was introduced. It uses also the charged $B^\pm \to J/\psi K^\pm$ decay, and is given by

$$\mathcal{B}_{\psi K} \equiv \frac{1 - \mathcal{A}_{\psi K}}{1 + \mathcal{A}_{\psi K}} , \tag{177}$$

with

$$\mathcal{A}_{\psi K} \equiv \left[\frac{BR(B^+ \to J/\psi K^+) + BR(B^- \to J/\psi K^-)}{BR(B_d^0 \to J/\psi K^0) + BR(\bar{B}_d^0 \to J/\psi \bar{K}^0)} \right] \left[\frac{\tau_{B_d^0}}{\tau_{B^+}} \right] , \tag{178}$$

and

$$\mathcal{D}_{\psi K}^\pm \equiv \frac{1}{2} \left[\mathcal{A}_{CP}^{dir}(B_d \to J/\psi K_S) \pm \mathcal{A}_{CP}^{dir}(B^\pm \to J/\psi K^\pm) \right] . \tag{179}$$

As discussed in detail in Refs. [119, 121], the observables $\mathcal{B}_{\psi K}$ and $\mathcal{D}_{\psi K}^-$ are sensitive to NP in the $I = 1$ isospin sector, whereas a non-vanishing value of $\mathcal{D}_{\psi K}^+$ would signal NP in the $I = 0$ isospin sector. Moreover, the NP contributions with $I = 1$ are expected to be dynamically suppressed with respect to the $I = 0$ case because of their flavour structure. The most recent B-factory results yield

$$\mathcal{B}_{\psi K} = -0.035 \pm 0.037, \quad \mathcal{D}_{\psi K}^- = 0.010 \pm 0.023, \quad \mathcal{D}_{\psi K}^+ = 0.017 \pm 0.023 . \tag{180}$$

Consequently, NP effects of $\mathcal{O}(10\%)$ in the $I = 1$ sector of the $B \to J/\psi K$ decay amplitudes are already disfavoured by the data for $\mathcal{B}_{\psi K}$ and $\mathcal{D}_{\psi K}^-$. However, since a non-vanishing value of $\mathcal{D}_{\psi K}^+$ requires also a large CP-conserving strong phase, this observable still leaves room for sizeable $I = 0$ NP contributions.

7.1.3 *A closer look at new-physics effects*

Thanks to the new Belle result listed in (175), the average for $(\sin 2\beta)_{\psi K_S}$ went down by about 1σ, which was a somewhat surprising development of the summer of 2005. Consequently, the comparison of (176) with the CKM fits in the $\bar{\rho}$–$\bar{\eta}$ plane no longer looks 'perfect', as we saw in Fig. 4. Let us have a closer look at this feature. If we use γ determined from non-leptonic $B \to D^{(*)}K^{(*)}$ tree modes and R_b from semileptonic decays, we may calculate the 'true' value of β with the help of the relations

$$\sin \beta = \frac{R_b \sin \gamma}{\sqrt{1 - 2R_b \cos \gamma + R_b^2}} , \quad \cos \beta = \frac{1 - R_b \cos \gamma}{\sqrt{1 - 2R_b \cos \gamma + R_b^2}} , \tag{181}$$

which follow from the unitarity of the CKM matrix; the UTfit value

$$\gamma = (65 \pm 20)^\circ \tag{182}$$

in (106) and the inclusive and exclusive values of R_b in (72) yield

$$\beta_{\text{incl}} = (26.7 \pm 1.9)^\circ \, , \quad \beta_{\text{excl}} = (22.9 \pm 3.8)^\circ \, , \tag{183}$$

which can be converted into

$$\sin 2\beta|_{\text{incl}} = 0.80 \pm 0.04, \quad \sin 2\beta|_{\text{excl}} = 0.71 \pm 0.09 \, . \tag{184}$$

Consequently, we find

$$\mathcal{S}_{\psi K} \equiv (\sin 2\beta)_{\psi K_{\text{S}}} - \sin 2\beta = \left\{ \begin{array}{ll} -0.11 \pm 0.05 & (\text{incl}) \\ -0.02 \pm 0.10 & (\text{excl}) \, , \end{array} \right. \tag{185}$$

and see nicely the discrepancy arising for the inclusive determination of $|V_{ub}|$. As discussed in detail in Ref. [101], R_b is actually the key parameter for this possible discrepancy with the SM, whereas the situation is remarkably stable with respect to γ. There are two limiting cases of this possible discrepancy with the KM mechanism of CP violation:

- NP contributions to the $B \to J/\psi K$ decay amplitudes;
- NP effects entering through B_d^0–$\bar{B}_d^0$ mixing.

Let us first illustrate the former case. As the NP effects in the $I = 1$ sector are expected to be dynamically suppressed, we consider only NP in the $I = 0$ isospin sector, which implies $\mathcal{B}_{\psi K} = \mathcal{D}_{\psi K}^- = 0$, in accordance with (180). To simplify the discussion, we assume that there is effectively only a single NP contribution of this kind, so that we may write

$$A(B_d^0 \to J/\psi K^0) = A_0 \left[1 + v_0 e^{i(\Delta_0 + \phi_0)} \right] = A(B^+ \to J/\psi K^+) \, . \tag{186}$$

Here v_0 and the CP-conserving strong phase Δ_0 are hadronic parameters, whereas ϕ_0 denotes a CP-violating phase originating beyond the SM. An interesting specific scenario falling into this category arises if the NP effects enter through EW penguins. This kind of NP has recently received a lot of attention in the context of the $B \to \pi K$ puzzle, which we shall discuss in Section 8. Also within the SM, where ϕ_0 vanishes, EW penguins have a sizeable impact on the $B \to J/\psi K$ system [148]. Using factorization, the following estimate can be obtained [83]:

$$v_0 e^{i\Delta_0}\big|_{\text{fact}}^{\text{SM}} \approx -0.03 \, . \tag{187}$$

In Figs. 21 (a) and (b), we consider the inclusive value of (185), and show the situation in the $\mathcal{S}_{\psi K}$–$\mathcal{D}_{\psi K}^+$ plane for $\phi_0 = -90^\circ$ and $\phi_0 = +90^\circ$, respectively. The contours correspond to different values of v_0, and are obtained by varying Δ_0 between 0° and 360°; the experimental data are represented by the diamonds with the error bars. Since factorization gives $\Delta_0 = 180^\circ$, as can be seen in (187), the case of $\phi_0 = -90^\circ$ is disfavoured. On the other hand, in the case of $\phi_0 = +90^\circ$, the experimental region can straightforwardly be reached for Δ_0 not differing too much from the factorization result, although an enhancement of v_0 by a factor of $\mathcal{O}(3)$ with respect to the SM estimate in (187), which suffers from large uncertainties, would simultaneously be required in order to reach the central experimental value. Consequently, NP contributions to the EW penguin sector could, in principle, be at the origin of the possible discrepancy indicated by the inclusive value of (185). This scenario should be carefully monitored in the future.

Another explanation of (185) is provided by CP-violating NP contributions to B_d^0–$\bar{B}_d^0$ mixing, which affect the corresponding mixing phase as in (165), so that

$$\phi_d = 2\beta + \phi_d^{\text{NP}} \, . \tag{188}$$

(a) (b)
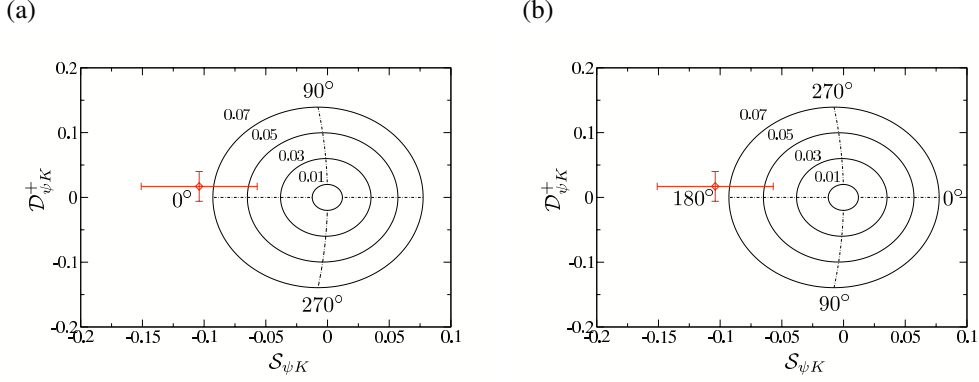
Fig. 21: The situation in the $\mathcal{S}_{\psi K}$–$\mathcal{D}^+_{\psi K}$ plane for NP contributions to the $B \to J/\psi K$ decay amplitudes in the $I = 0$ isospin sector for NP phases $\phi_0 = -90°$ (a) and $\phi_0 = +90°$ (b). The diamonds with the error bars represent the averages of the current data [for the inclusive value of (185)], whereas the numbers correspond to the values of Δ_0 and v_0.
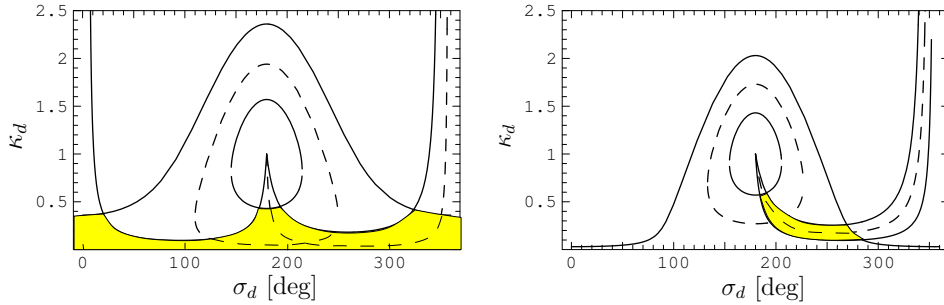
Fig. 22: Left panel: allowed region (yellow/grey) in the σ_d–κ_d plane in a scenario with the JLQCD lattice results (122) and $\phi^{\mathrm{NP}}_d\big|_{\mathrm{excl}}$. Dashed lines: central values of ρ_d and ϕ^{NP}_d, solid lines: $\pm 1\,\sigma$. Right panel: ditto for the scenario with the (HP+JL)QCD lattice results (123) and $\phi^{\mathrm{NP}}_d\big|_{\mathrm{incl}}$.

Assuming that the NP contributions to the $B \to J/\psi K$ amplitudes are negligible, (176) implies

$$\phi_d = (43.4 \pm 2.5)° \quad \vee \quad (136.6 \pm 2.5)° \,. \tag{189}$$

Here the latter solution would be in dramatic conflict with the CKM fits, and would require a large NP contribution to B^0_d–$\bar{B}^0_d$ mixing [122, 149]. Both solutions can be distinguished through the measurement of the sign of $\cos\phi_d$, where a positive value would select the SM-like branch. Using an angular analysis of the $B_d \to J/\psi[\to \ell^+\ell^-]K^*[\to \pi^0 K_{\mathrm{S}}]$ decay products, the BaBar Collaboration finds [150]

$$\cos\phi_d = 2.72^{+0.50}_{-0.79} \pm 0.27 \,, \tag{190}$$

thereby favouring the solution around $\phi_d = 43°$. Interestingly, this picture emerges also from the first data for CP-violating effects in $B_d \to D^{(*)\pm}\pi^\mp$ modes [90], and an analysis of the $B \to \pi\pi, \pi K$ system [83], although in an indirect manner. Recently, a new method has been proposed, which makes use of the interference pattern in $D \to K_{\mathrm{S}}\pi^+\pi^-$ decays emerging from $B_d \to D\pi^0$ and similar decays [151]. The results of this method are also consistent with the SM, so that a negative value of $\cos\phi_d$ is now ruled out with greater than 95% confidence [89].

Using the 'true' values of β in (183), the value of $\phi_d = (43.4 \pm 2.5)°$ implies

$$\phi^{\mathrm{NP}}_d\big|_{\mathrm{incl}} = -(10.1 \pm 4.6)° \,, \qquad \phi^{\mathrm{NP}}_d\big|_{\mathrm{excl}} = -(2.5 \pm 8.0)° \,; \tag{191}$$
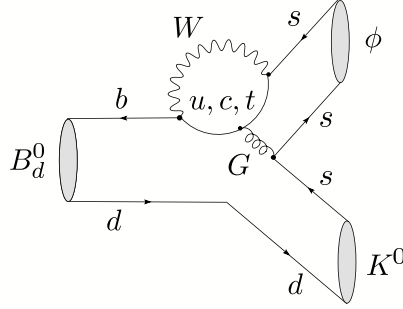
Fig. 23: Feynman diagrams contributing to $B_d^0 \to \phi K^0$ decays

results of $\phi_d^{\mathrm{NP}} \approx -10°$ were also recently obtained in Refs. [91, 152]. The contours in Fig. 19 allow us now to convert these numbers into constraints in the σ_d–κ_d plane. Further constraints can be obtained through the experimental value of ΔM_d in (130) with the help of the contours in Fig. 18, where ρ_d is introduced in (166). In addition to hadronic parameters, the SM prediction of ΔM_d involves also the CKM factor $|V_{td}^* V_{tb}|$, which can—if we use the unitarity of the CKM matrix—be expressed as

$$|V_{td}^* V_{tb}| = |V_{cb}|\lambda\sqrt{1 - 2R_b \cos\gamma + R_b^2} \,. \tag{192}$$

The values in (72) and (182), as well as the relevant lattice parameters in (122) and (123) yield then

$$\rho_d|_{\mathrm{JLQCD}} = 0.97 \pm 0.33^{-0.17}_{+0.26} \tag{193}$$

$$\rho_d|_{\mathrm{(HP+JL)QCD}} = 0.75 \pm 0.25 \pm 0.16 \,, \tag{194}$$

where the first and second errors are due to γ (and a small extent to R_b) and $f_{B_d}\hat{B}_{B_d}^{1/2}$, respectively [101]. These results are compatible with the SM value $\rho_d = 1$, but suffer from considerable uncertainties. In Fig. 22, we finally show the situation in the σ_d–κ_d plane. We see that the information about the CP-violating phase ϕ_d has a dramatic impact, reducing the allowed NP parameter space significantly.

The possibility of having a non-zero value of (185) could of course just be due to a statistical fluctuation. However, should it be confirmed, it could be due to CP-violating NP contributions to the $B_d^0 \to J/\psi K_\mathrm{S}$ decay amplitude or to B_d^0–$\bar{B}_d^0$ mixing, as we just saw. A tool to distinguish between these avenues is provided by decays of the kind $B_d \to D\pi^0, D\rho^0, ...$, which are pure 'tree' decays, i.e., they do *not* receive any penguin contributions. If the neutral D mesons are observed through their decays into CP eigenstates $D_\pm$, these decays allow extremely clean determinations of the 'true' value of $\sin\phi_d$ [153], as we shall discuss in more detail in Section 10.3. In view of (185), this would be very interesting, so that detailed feasibility studies for the exploration of the $B_d \to D\pi^0, D\rho^0, ...$ modes at a super-B factory are strongly encouraged.

7.2 $B_d^0 \to \phi K_\mathrm{S}$

Another important probe for the testing of the KM mechanism is offered by $B_d^0 \to \phi K_\mathrm{S}$, which is a decay into a CP-odd final state. As can be seen in Fig. 23, it originates from $\bar{b} \to \bar{s}s\bar{s}$ transitions and is, therefore, a pure penguin mode. This decay is described by the low-energy effective Hamiltonian in (80) with $r = s$, where the current–current operators may only contribute through penguin-like contractions, which describe the penguin topologies with internal up- and charm-quark exchanges. The dominant role is played by the QCD penguin operators [154]. However, thanks to the large top-quark mass, EW penguins have a sizeable impact as well [72, 155]. In the SM, we may write

$$A(B_d^0 \to \phi K_\mathrm{S}) = \lambda_u^{(s)}\tilde{A}_\mathrm{P}^{u'} + \lambda_c^{(s)}\tilde{A}_\mathrm{P}^{c'} + \lambda_t^{(s)}\tilde{A}_\mathrm{P}^{t'} \,, \tag{195}$$

(a)   (b)
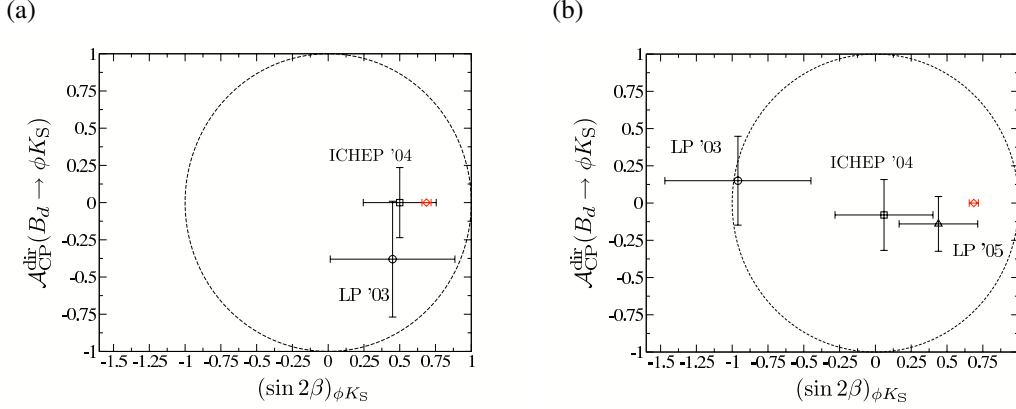
Fig. 24: The time evolution of the BaBar (a) and Belle (b) data for the CP violation in $B_d \to \phi K_S$. The diamonds represent the SM relations (199)–(201) with (176).

where we have applied the same notation as in Section 7.1. Eliminating the CKM factor $\lambda_t^{(s)}$ with the help of (79) yields

$$A(B_d^0 \to \phi K_S) \propto \left[1 + \lambda^2 b e^{i\Theta} e^{i\gamma} \right] , \tag{196}$$

where

$$b e^{i\Theta} \equiv \left(\frac{R_b}{1 - \lambda^2} \right) \left[\frac{\tilde{A}_P^{u'} - \tilde{A}_P^{t'}}{\tilde{A}_P^{c'} - \tilde{A}_P^{t'}} \right] . \tag{197}$$

Consequently, we obtain

$$\xi_{\phi K_S}^{(d)} = + e^{-i\phi_d} \left[\frac{1 + \lambda^2 b e^{i\Theta} e^{-i\gamma}}{1 + \lambda^2 b e^{i\Theta} e^{+i\gamma}} \right] . \tag{198}$$

The theoretical estimates of $b e^{i\Theta}$ suffer from large hadronic uncertainties. However, since this parameter enters (198) in a doubly Cabibbo-suppressed way, we obtain the following expressions [148]:

$$
\begin{aligned}
\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \phi K_S) &= & 0 + \mathcal{O}(\lambda^2) \tag{199} \\
\mathcal{A}_{\text{CP}}^{\text{mix}}(B_d \to \phi K_S) &= & -\sin\phi_d + \mathcal{O}(\lambda^2) , \tag{200}
\end{aligned}
$$

where we made the plausible assumption that $b = \mathcal{O}(1)$. On the other hand, the mixing-induced CP asymmetry of $B_d \to J/\psi K_S$ measures also $-\sin\phi_d$, as we saw in (173). We arrive therefore at the following relation [148, 156]:

$$-(\sin 2\beta)_{\phi K_S} \equiv \mathcal{A}_{\text{CP}}^{\text{mix}}(B_d \to \phi K_S) = \mathcal{A}_{\text{CP}}^{\text{mix}}(B_d \to J/\psi K_S) + \mathcal{O}(\lambda^2) , \tag{201}$$

which offers an interesting test of the SM. Since $B_d \to \phi K_S$ is governed by penguin processes in the SM, this decay may well be affected by NP. In fact, if we assume that NP arises generically in the TeV regime, it can be shown through field-theoretical estimates that the NP contributions to $b \to s\bar{s}s$ transitions may well lead to sizeable violations of (199) and (201) [119, 123]. Moreover, this is also the case for several specific NP scenarios; for examples, see Refs. [126, 128, 129, 157].

In Fig. 24, we show the time evolution of the B-factory data for the measurements of CP violation in $B_d \to \phi K_S$, using the results reported at the LP '03 [158], ICHEP '04 [159] and LP '05 [160] conferences. Because of (155), the corresponding observables have to lie inside a circle with radius one around the origin, which is represented by the dashed lines. The result announced by the Belle Collaboration in 2003 led to quite some excitement in the community. Meanwhile, the Babar [161] and Belle [162] results are in good agreement with each other, yielding the following averages [61]:

$$\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \phi K_S) = -0.09 \pm 0.14, \quad (\sin 2\beta)_{\phi K_S} = 0.47 \pm 0.19 . \tag{202}$$

140

(a) (b)
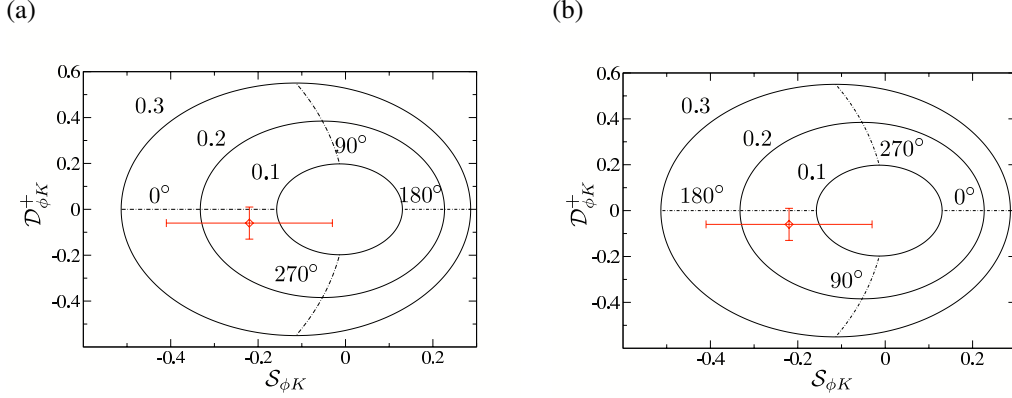
Fig. 25: The situation in the $\mathcal{S}_{\phi K}$–$\mathcal{D}_{\phi K}^+$ plane for NP contributions to the $B \to \phi K$ decay amplitudes in the $I = 0$ isospin sector for NP phases $\phi_0 = -90°$ (a) and $\phi_0 = +90°$ (b). The diamonds with the error bars represent the averages of the current data, whereas the numbers correspond to the values of $\tilde{\Delta}_0$ and $\tilde{v}_0$.

If we take (176) into account, we obtain the following result for the counterpart of (185):

$$\mathcal{S}_{\phi K} \equiv (\sin 2\beta)_{\phi K_{\mathrm{S}}} - (\sin 2\beta)_{\psi K_{\mathrm{S}}} = -0.22 \pm 0.19 \,. \tag{203}$$

This number still appears to be somewhat on the lower side, thereby indicating potential NP contributions to $b \to s\bar{s}s$ processes.

Further insights into the origin and the isospin structure of NP contributions can be obtained through a combined analysis of the neutral and charged $B \to \phi K$ modes with the help of observables $\mathcal{B}_{\phi K}$ and $\mathcal{D}_{\phi K}^\pm$ [123], which are defined in analogy to (177) and (179), respectively. The current experimental results read as follows:

$$\mathcal{B}_{\phi K} = 0.00 \pm 0.08 \,, \quad \mathcal{D}_{\phi K}^- = -0.03 \pm 0.07 \,, \quad \mathcal{D}_{\phi K}^+ = -0.06 \pm 0.07 \,. \tag{204}$$

As in the $B \to J/\psi K$ case, $\mathcal{B}_{\phi K}$ and $\mathcal{D}_{\phi K}^-$ probe NP effects in the $I = 1$ sector, which are expected to be dynamically suppressed, whereas $\mathcal{D}_{\phi K}^+$ is sensitive to NP in the $I = 0$ sector. The latter kind of NP could also manifest itself as a non-vanishing value of (203).

In order to illustrate these effects, let us consider again the case where NP enters only in the $I = 0$ isospin sector. An important example is given by EW penguins, which have a significant impact on $B \to \phi K$ decays [72]. In analogy to the discussion in Section 7.1, we may then write

$$A(B_d^0 \to \phi K^0) = \tilde{A}_0 \left[1 + \tilde{v}_0 e^{i(\tilde{\Delta}_0 + \phi_0)} \right] = A(B^+ \to \phi K^+) \,, \tag{205}$$

which implies $\mathcal{B}_{\phi K} = \mathcal{D}_{\phi K}^- = 0$, in accordance with (204). The notation corresponds to that of (186). Using the factorization approach to deal with the QCD and EW penguin contributions, we obtain the following estimate in the SM, where the CP-violating NP phase ϕ_0 vanishes [83]:

$$\left. \tilde{v}_0 e^{i\tilde{\Delta}_0} \right|_{\mathrm{fact}}^{\mathrm{SM}} \approx -0.2 \,. \tag{206}$$

In Figs. 25 (a) and (b), we show the situation in the $\mathcal{S}_{\phi K}$–$\mathcal{D}_{\phi K}^+$ plane for NP phases $\phi_0 = -90°$ and $\phi_0 = +90°$, respectively, and various values of $\tilde{v}_0$; each point of the contours is parametrized by $\tilde{\Delta}_0 \in [0°, 360°]$. We observe that the central values of the current experimental data, which are represented by the diamonds with the error bars, can straightforwardly be accommodated in this scenario in the case of $\phi_0 = +90°$ for strong phases satisfying $\cos \tilde{\Delta}_0 < 0$, as in factorization. Moreover, as can also be seen

in Fig. 25 (b), the EW penguin contributions would then have to be suppressed with respect to the SM estimate, which would be an interesting feature in view of the discussion of the $B \to \pi K$ puzzle and the rare decay constraints in Section 8.

It will be interesting to follow the evolution of the B-factory data, and to monitor also similar modes, such as $B_d^0 \to \pi^0 K_{\rm S}$ [163] and $B_d^0 \to \eta' K_{\rm S}$ [164]. For a compilation of the corresponding experimental results, see Ref. [61]; recent theoretical papers dealing with these channels can be found in Refs. [82, 83, 91, 165, 166]. We shall return to the CP asymmetries of the $B_d^0 \to \pi^0 K_{\rm S}$ channel in Section 8.

7.3 $B_d^0 \to \pi^+\pi^-$

This decay is a transition into a CP eigenstate with eigenvalue $+1$, and originates from $\bar{b} \to \bar{u}u\bar{d}$ processes, as can be seen in Fig. 26. In analogy to (167) and (195), its decay amplitude can be written as follows [167]:

$$A(B_d^0 \to \pi^+\pi^-) = \lambda_u^{(d)} \left(A_{\rm T}^u + A_{\rm P}^u \right) + \lambda_c^{(d)} A_{\rm P}^c + \lambda_t^{(d)} A_{\rm P}^t . \tag{207}$$

Using again (79) to eliminate the CKM factor $\lambda_t^{(d)} = V_{td}V_{tb}^*$ and applying once more the Wolfenstein parametrization yields

$$A(B_d^0 \to \pi^+\pi^-) = \mathcal{C} \left[e^{i\gamma} - d e^{i\theta} \right] , \tag{208}$$

where the overall normalization $\mathcal{C}$ and

$$d e^{i\theta} \equiv \frac{1}{R_b} \left[\frac{A_{\rm P}^c - A_{\rm P}^t}{A_{\rm T}^u + A_{\rm P}^u - A_{\rm P}^t} \right] \tag{209}$$

are hadronic parameters. The formalism discussed in Section 5.5 then implies

$$\xi_{\pi^+\pi^-}^{(d)} = -e^{-i\phi_d} \left[\frac{e^{-i\gamma} - d e^{i\theta}}{e^{+i\gamma} - d e^{i\theta}} \right] . \tag{210}$$

In contrast to the expressions (171) and (198) for the $B_d^0 \to J/\psi K_{\rm S}$ and $B_d^0 \to \phi K_{\rm S}$ counterparts, respectively, the hadronic parameter $d e^{i\theta}$, which suffers from large theoretical uncertainties, does *not* enter (210) in a doubly Cabibbo-suppressed way. This feature is at the basis of the famous 'penguin problem' in $B_d^0 \to \pi^+\pi^-$, which was addressed in many papers (see, for instance, [168]– [173]). If the penguin contributions to this channel were negligible, i.e., $d = 0$, its CP asymmetries were simply given by

$$\mathcal{A}_{\rm CP}^{\rm dir}(B_d \to \pi^+\pi^-) = 0 \tag{211}$$

$$\mathcal{A}_{\rm CP}^{\rm mix}(B_d \to \pi^+\pi^-) = \sin(\phi_d + 2\gamma) \stackrel{\rm SM}{=} \sin(\underbrace{2\beta + 2\gamma}_{2\pi - 2\alpha}) = -\sin 2\alpha . \tag{212}$$

Consequently, $\mathcal{A}_{\rm CP}^{\rm mix}(B_d \to \pi^+\pi^-)$ would then allow us to determine α. However, in the general case, we obtain expressions with the help of (153) and (210) of the form

$$\mathcal{A}_{\rm CP}^{\rm dir}(B_d \to \pi^+\pi^-) = G_1(d, \theta; \gamma) \tag{213}$$

$$\mathcal{A}_{\rm CP}^{\rm mix}(B_d \to \pi^+\pi^-) = G_2(d, \theta; \gamma, \phi_d) ; \tag{214}$$

for explicit formulae, see Ref. [167]. We observe that actually the phases ϕ_d and γ enter directly in the $B_d \to \pi^+\pi^-$ observables, and not α. Consequently, since ϕ_d can be fixed through the mixing-induced CP violation in the 'golden' mode $B_d \to J/\psi K_{\rm S}$, as we have seen in Subsection 7.1, we may use $B_d \to \pi^+\pi^-$ to probe γ.
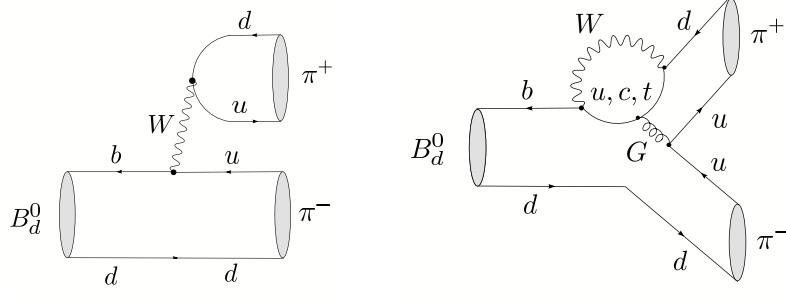
Fig. 26: Feynman diagrams contributing to $B_d^0 \to \pi^+\pi^-$ decays

The current measurements of the $B_d \to \pi^+\pi^-$ CP asymmetries are given as follows:

$$\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \pi^+\pi^-) \;=\; \begin{cases} -0.09 \pm 0.15 \pm 0.04 & \text{(BaBar [174])} \\ -0.56 \pm 0.12 \pm 0.06 & \text{(Belle [175])} \end{cases} \tag{215}$$

$$\mathcal{A}_{\text{CP}}^{\text{mix}}(B_d \to \pi^+\pi^-) \;=\; \begin{cases} +0.30 \pm 0.17 \pm 0.03 & \text{(BaBar [174])} \\ +0.67 \pm 0.16 \pm 0.06 & \text{(Belle [175]) .} \end{cases} \tag{216}$$

The BaBar and Belle results are still not fully consistent with each other, although the experiments are now in better agreement. In Ref. [61], the following averages were obtained:

$$\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \pi^+\pi^-) \;=\; -0.37 \pm 0.10 \tag{217}$$

$$\mathcal{A}_{\text{CP}}^{\text{mix}}(B_d \to \pi^+\pi^-) \;=\; +0.50 \pm 0.12 . \tag{218}$$

The central values of these averages are remarkably stable in time. Direct CP violation at this level would require large penguin contributions with large CP-conserving strong phases, thereby indicating large non-factorizable effects.

This picture is in fact supported by the direct CP violation in $B_d^0 \to \pi^- K^+$ modes that could be established by the B factories in the summer of 2004 [6]. Here the BaBar and Belle results agree nicely with each other, yielding the following average [61]:

$$\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \pi^{\mp} K^{\pm}) = 0.115 \pm 0.018 . \tag{219}$$

The diagrams contributing to $B_d^0 \to \pi^- K^+$ can straightforwardly be obtained from those in Fig. 26 by just replacing the anti-down quark emerging from the W boson through an anti-strange quark. Consequently, the hadronic matrix elements entering $B_d^0 \to \pi^+\pi^-$ and $B_d^0 \to \pi^- K^+$ can be related to one another through the $SU(3)$ flavour symmetry of strong interactions and the additional assumption that the penguin annihilation and exchange topologies contributing to $B_d^0 \to \pi^+\pi^-$, which have no counterpart in $B_d^0 \to \pi^- K^+$ and involve the 'spectator" down quark in Fig. 26, play actually a negligible role [176]. Following these lines, we obtain the following relation in the SM:

$$H_{\text{BR}} \equiv \underbrace{\frac{1}{\epsilon}\left(\frac{f_K}{f_\pi}\right)^2\left[\frac{\text{BR}(B_d \to \pi^+\pi^-)}{\text{BR}(B_d \to \pi^{\mp}K^{\pm})}\right]}_{7.5 \pm 0.7} = -\underbrace{\frac{1}{\epsilon}\left[\frac{\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \pi^{\mp}K^{\pm})}{\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \pi^+\pi^-)}\right]}_{6.7 \pm 2.0} \equiv H_{\mathcal{A}_{\text{CP}}^{\text{dir}}} , \tag{220}$$

where

$$\epsilon \equiv \frac{\lambda^2}{1 - \lambda^2} = 0.053 , \tag{221}$$

and the ratio $f_K/f_\pi = 160/131$ of the kaon and pion decay constants defined through

$$\langle 0|\bar{s}\gamma_\alpha\gamma_5 u|K^+(k)\rangle = if_K k_\alpha , \quad \langle 0|\bar{d}\gamma_\alpha\gamma_5 u|\pi^+(k)\rangle = if_\pi k_\alpha \tag{222}$$

143

describes factorizable $SU(3)$-breaking corrections. As usual, the CP-averaged branching ratios are defined as

$$\mathrm{BR} \equiv \frac{1}{2}\left[\mathrm{BR}(B \to f) + \mathrm{BR}(\bar{B} \to \bar{f})\right]. \tag{223}$$

In (220), we have also given the numerical values following from the data. Consequently, this relation is well satisfied within the experimental uncertainties, and does not show any anomalous behaviour. It supports therefore the SM description of the $B_d^0 \to \pi^- K^+$, $B_d^0 \to \pi^+ \pi^-$ decay amplitudes, and our working assumptions listed before (220).

The quantities H_{BR} and $H_{\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}}$ introduced in this relation can be written as follows:

$$H_{\mathrm{BR}} = G_3(d,\theta;\gamma) = H_{\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}}. \tag{224}$$

If we complement this expression with (213) and (214), and use [see (189)]

$$\phi_d = (43.4 \pm 2.5)^\circ, \tag{225}$$

we have sufficient information to determine γ, as well as (d,θ) [167,176,177]. In using (225), we assume that the possible discrepancy with the SM described by (185) is only due to NP in B_d^0–$\bar{B}_d^0$ mixing and not to effects entering through the $B_d^0 \to J/\psi K_{\mathrm{S}}$ decay amplitude. As was recently shown in Ref. [91], the results following from H_{BR} and $H_{\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}}$ give results that are in good agreement with one another. Since the avenue offered by $H_{\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}}$ is cleaner than the one provided by H_{BR}, it is preferable to use the former quantity to determine γ, yielding the following result [91]:

$$\gamma = (73.9^{+5.8}_{-6.5})^\circ. \tag{226}$$

Here a second solution around 42° was discarded, which can be excluded through an analysis of the whole $B \to \pi\pi, \pi K$ system [83]. As was recently discussed [91] (see also Refs. [176,177]), even large non-factorizable $SU(3)$-breaking corrections have a remarkably small impact on the numerical result in (226). The value of γ in (226) is somewhat higher than the central values in (106), but fully consistent within the large errors. An even larger value in the ballpark of 80° was recently extracted from the $B \to \pi\pi$ data with the help of SCET [178,179].

8 The $B \to \pi K$ puzzle and its relation to rare B and K decays

8.1 Preliminaries

We already made first contact with a $B \to \pi K$ decay in Section 7.3, the $B_d^0 \to \pi^- K^+$ channel. It receives contributions both from tree and from penguin topologies. Since this decay originates from a $\bar{b} \to \bar{s}$ transition, the tree amplitude is suppressed by a CKM factor $\lambda^2 R_b \sim 0.02$ with respect to the penguin amplitude. Consequently, $B_d^0 \to \pi^- K^+$ is governed by QCD penguins; the tree topologies contribute only at the 20% level to the decay amplitude. The feature of the dominance of QCD penguins applies to all $B \to \pi K$ modes, which can be classified with respect to their EW penguin contributions as follows (see Fig. 27):

(a) In the $B_d^0 \to \pi^- K^+$ and $B^+ \to \pi^+ K^0$ decays, EW penguins contribute in colour-suppressed form and are hence expected to play a minor role.

(b) In the $B_d^0 \to \pi^0 K^0$ and $B^+ \to \pi^0 K^+$ decays, EW penguins contribute in colour-allowed form and have therefore a significant impact on the decay amplitude, entering at the same order of magnitude as the tree contributions.

As we noted above, EW penguins offer an attractive avenue for NP to enter non-leptonic B decays, which is also the case for the $B \to \pi K$ system [180, 181]. Indeed, the decays of class (b) show a puzzling pattern, which may point towards such a NP scenario. This feature emerged already in 2000 [182],

(a)

(b)
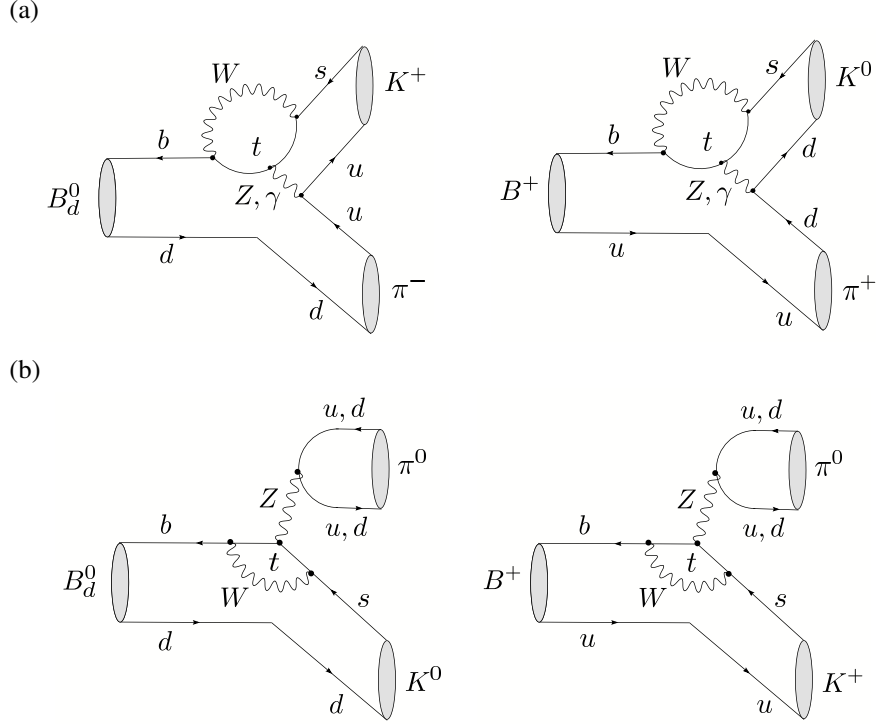
Fig. 27: Examples of the colour-suppressed (a) and colour-allowed (b) EW penguin contributions to the $B \to \pi K$ system

when the CLEO Collaboration reported the observation of the $B_d^0 \to \pi^0 K^0$ channel with a surprisingly prominent rate [183], and is still present in the most recent BaBar and Belle data, thereby receiving a lot of attention in the literature (see, for instance, Refs. [157] and [184]– [188]).

In the following discussion, we focus on the systematic strategy to explore the '$B \to \pi K$ puzzle' developed in Refs. [82,83]; all numerical results refer to the most recent analysis presented in Ref. [91]. The logical structure is very simple: the starting point is given by the values of ϕ_d and γ in (225) and (226), respectively, and by the $B \to \pi\pi$ system, which allows us to extract a set of hadronic parameters from the data with the help of the isospin symmetry of strong interactions. Then we make, in analogy to the determination of γ in Section 7.3, the following working hypotheses:

(i) $SU(3)$ flavour symmetry of strong interactions (but taking factorizable $SU(3)$-breaking corrections into account),

(ii) neglect of penguin annihilation and exchange topologies,

which allow us to fix the hadronic $B \to \pi K$ parameters through their $B \to \pi\pi$ counterparts. Interestingly, we may gain confidence in these assumptions through internal consistency checks [an example is relation (220)], which work nicely within the experimental uncertainties. Having the hadronic $B \to \pi K$ parameters at hand, we can predict the $B \to \pi K$ observables in the SM. The comparison of the corresponding picture with the B-factory data will then guide us to NP in the EW penguin sector, involving in particular a large CP-violating NP phase. In the final step, we explore the interplay of this NP scenario with rare K and B decays.

8.2 Extracting hadronic parameters from the $B \to \pi\pi$ system

In order to fully exploit the information that is provided by the whole $B \to \pi\pi$ system, we use— in addition to the two CP-violating $B_d^0 \to \pi^+\pi^-$ observables—the following ratios of CP-averaged

branching ratios:

$$R_{+-}^{\pi\pi} \equiv 2\left[\frac{\mathrm{BR}(B^+ \to \pi^+\pi^0) + \mathrm{BR}(B^- \to \pi^-\pi^0)}{\mathrm{BR}(B_d^0 \to \pi^+\pi^-) + \mathrm{BR}(\bar{B}_d^0 \to \pi^+\pi^-)}\right] = 2.04 \pm 0.28 \tag{227}$$

$$R_{00}^{\pi\pi} \equiv 2\left[\frac{\mathrm{BR}(B_d^0 \to \pi^0\pi^0) + \mathrm{BR}(\bar{B}_d^0 \to \pi^0\pi^0)}{\mathrm{BR}(B_d^0 \to \pi^+\pi^-) + \mathrm{BR}(\bar{B}_d^0 \to \pi^+\pi^-)}\right] = 0.58 \pm 0.13 \ . \tag{228}$$

The pattern of the experimental numbers in these expressions came as quite a surprise, as the central values calculated in QCDF gave $R_{+-}^{\pi\pi} = 1.24$ and $R_{00}^{\pi\pi} = 0.07$ [184]. As discussed in detail in Ref. [83], this '$B \to \pi\pi$ puzzle' can straightforwardly be accommodated in the SM through large non-factorizable hadronic interference effects, i.e., does not point towards NP. For recent SCET analyses, see Refs. [179, 189, 190].

Using the isospin symmetry of strong interactions, we can write

$$R_{+-}^{\pi\pi} = F_1(d, \theta, x, \Delta; \gamma) , \quad R_{00}^{\pi\pi} = F_2(d, \theta, x, \Delta; \gamma) , \tag{229}$$

where $xe^{i\Delta}$ is another hadronic parameter, which was introduced in Refs. [82, 83]. Using now, in addition, the CP-violating observables in (213) and (214), we arrive at the following set of haronic parameters:

$$d = 0.52_{-0.09}^{+0.09}, \quad \theta = (146_{-7.2}^{+7.0})^\circ, \quad x = 0.96_{-0.14}^{+0.13}, \quad \Delta = -(53_{-26}^{+18})^\circ \ . \tag{230}$$

In the extraction of these quantities, also the EW penguin effects in the $B \to \pi\pi$ system are included [191, 192], although these topologies have a tiny impact [163]. Let us emphasize that the results for the hadronic parameters listed above, which are consistent with the picture emerging in the analyses of other authors (see, for example, Refs. [193, 194]), are essentially clean and serve as a testing ground for calculations within QCD-related approaches. For instance, in recent QCDF [195] and PQCD [196] analyses, the following numbers were obtained:

$$d|_{\mathrm{QCDF}} = 0.29 \pm 0.09 , \quad \theta|_{\mathrm{QCDF}} = -(171.4 \pm 14.3)^\circ , \tag{231}$$

$$d|_{\mathrm{PQCD}} = 0.23_{-0.05}^{+0.07} , \quad +139^\circ < \theta|_{\mathrm{PQCD}} < +148^\circ , \tag{232}$$

which depart significantly from the pattern in (230) that is implied by the data.

Finally, we can predict the CP asymmetries of the decay $B_d \to \pi^0\pi^0$:

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_d \to \pi^0\pi^0) = -0.30_{-0.26}^{+0.48} , \quad \mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_d \to \pi^0\pi^0) = -0.87_{-0.19}^{+0.29} \ . \tag{233}$$

The current experimental value for the direct CP asymmetry is given as follows [61]:

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_d \to \pi^0\pi^0) = -0.28_{-0.39}^{+0.40} \ . \tag{234}$$

Consequently, no stringent test of the corresponding prediction in (233) is provided at this stage, although the indicated agreement is encouraging.

8.3 Analysis of the $B \to \pi K$ system

Let us begin the analysis of the $B \to \pi K$ system by having a closer look at the modes of class (a) introduced above, $B_d \to \pi^\mp K^\pm$ and $B^\pm \to \pi^\pm K$, which are only marginally affected by EW penguin contributions. We already used the branching ratio and direct CP asymmetry of the former channel in the $SU(3)$ relation (220), which is nicely satisfied by the current data, and in the extraction of γ with the help of the CP-violating $B_d \to \pi^+\pi^-$ observables, yielding the value in (226). The $B_d \to \pi^\mp K^\pm$ modes provide the CP-violating asymmetry

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B^\pm \to \pi^\pm K) \equiv \frac{\mathrm{BR}(B^+ \to \pi^+ K^0) - \mathrm{BR}(B^- \to \pi^- \bar{K}^0)}{\mathrm{BR}(B^+ \to \pi^+ K^0) + \mathrm{BR}(B^- \to \pi^- \bar{K}^0)} = 0.02 \pm 0.04 , \tag{235}$$
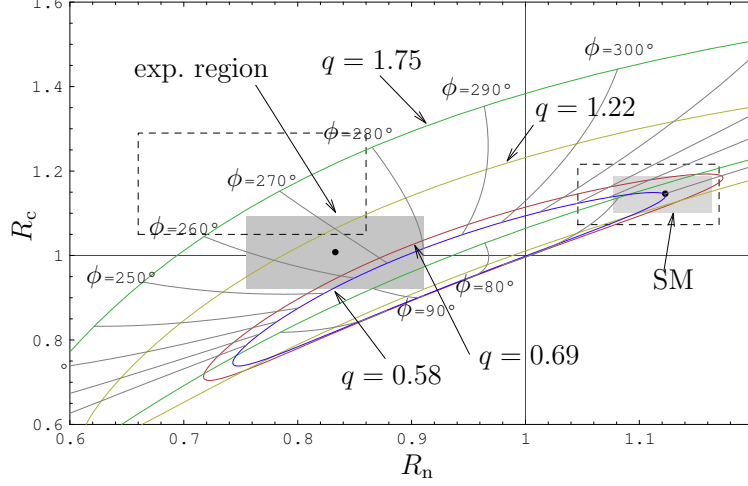
Fig. 28: The current situation in the $R_{\rm n}$–$R_{\rm c}$ plane: the shaded areas indicate the experimental and SM 1σ ranges, while the lines show the theory predictions for the central values of the hadronic parameters and various values of q with $\phi \in [0°, 360°]$

and enter in the following ratio [197]:

$$R \equiv \left[\frac{\mathrm{BR}(B_d^0 \to \pi^- K^+) + \mathrm{BR}(\bar{B}_d^0 \to \pi^+ K^-)}{\mathrm{BR}(B^+ \to \pi^+ K^0) + \mathrm{BR}(B^- \to \pi^- \bar{K}^0)}\right] \frac{\tau_{B^+}}{\tau_{B_d^0}} = 0.86 \pm 0.06 \; ; \tag{236}$$

the numerical values refer again to the most recent compilation in [61]. The $B^+ \to \pi^+ K^0$ channel involves another hadronic parameter $\rho_{\rm c} e^{i\theta_{\rm c}}$ which cannot be determined through the $B \to \pi\pi$ data [191, 198, 199]:

$$A(B^+ \to \pi^+ K^0) = -P' \left[1 + \rho_{\rm c} e^{i\theta_{\rm c}} e^{i\gamma}\right] \; ; \tag{237}$$

the overall normalization P' cancels in (235) and (236). Usually, it is assumed that the parameter $\rho_{\rm c} e^{i\theta_{\rm c}}$ can be neglected. In this case, the direct CP asymmetry in (235) vanishes, and R can be calculated through the $B \to \pi\pi$ data with the help of the assumptions specified in Section 8.1:

$$R|_{\rm SM} = 0.963^{+0.019}_{-0.022} \; . \tag{238}$$

This numerical result is 1.6σ larger than the experimental value in (236). As was discussed in detail in Ref. [200], the experimental range for the direct CP asymmetry in (235) and the first direct signals for the $B^{\pm} \to K^{\pm}K$ decays favour a value of $\theta_{\rm c}$ around $0°$. This feature allows us to essentially resolve the small discrepancy concerning R for values of $\rho_{\rm c}$ around 0.05. The remaining small numerical difference between the calculated value of R and the experimental result, if confirmed by future data, could be due to (small) colour-suppressed EW penguins, which enter R as well [83]. As was recently discussed in Ref. [91], even large non-factorizable $SU(3)$-breaking effects would have a small impact on the predicted value of R. In view of these results, it would not be a surprise to see an increase of the experimental value of R in the future.

Let us now turn to the $B^+ \to \pi^0 K^+$ and $B_d^0 \to \pi^0 K^0$ channels, which are the $B \to \pi K$ modes with significant contributions from EW penguin topologies. The key observables for the exploration of these modes are the following ratios of their CP-averaged branching ratios [182, 191]:

$$R_{\rm c} \equiv 2 \left[\frac{\mathrm{BR}(B^+ \to \pi^0 K^+) + \mathrm{BR}(B^- \to \pi^0 K^-)}{\mathrm{BR}(B^+ \to \pi^+ K^0) + \mathrm{BR}(B^- \to \pi^- \bar{K}^0)}\right] = 1.01 \pm 0.09 \tag{239}$$

$$R_{\rm n} \equiv \frac{1}{2}\left[\frac{{\rm BR}(B_d^0 \to \pi^- K^+) + {\rm BR}(\bar{B}_d^0 \to \pi^+ K^-)}{{\rm BR}(B_d^0 \to \pi^0 K^0) + {\rm BR}(\bar{B}_d^0 \to \pi^0 \bar{K}^0)}\right] = 0.83 \pm 0.08 \,, \tag{240}$$

where the overall normalization factors of the decay amplitudes cancel, as in (236). In order to describe the EW penguin effects, both a parameter q, which measures the strength of the EW penguins with respect to tree-like topologies, and a CP-violating phase ϕ are introduced. In the SM, this phase vanishes, and q can be calculated with the help of the $SU(3)$ flavour symmetry, yielding a value of $0.69 \times 0.086/|V_{ub}/V_{cb}| = 0.58$ [201]. Following the strategy described above yields the following SM predictions:

$$R_{\rm c}|_{\rm SM} = 1.15 \pm 0.05 \,, \quad R_{\rm n}|_{\rm SM} = 1.12 \pm 0.05 \,, \tag{241}$$

where in particular the value of $R_{\rm n}$ does not agree with the experimental number, which is a manifestation of the $B \to \pi K$ puzzle. As was recently discussed in Ref. [91], the internal consistency checks of the working assumptions listed in Subsection 8.1 are currently satisfied at the level of 25%, and can be systematically improved through better data. A detailed study of the numerical predictions in (241) (and those given below) shows that their sensitivity on non-factorizable $SU(3)$-breaking effects of this order of magnitude is surprisingly small. Consequently, it is very exciting to speculate that NP effects in the EW penguin sector, which are described effectively through (q, ϕ), are at the origin of the $B \to \pi K$ puzzle. Following Refs. [82, 83], we show the situation in the $R_{\rm n}$–$R_{\rm c}$ plane in Fig. 28, where—for the convenience of the reader—also the experimental range and the SM predictions at the time of the original analysis of Refs. [82, 83] are indicated through the dashed rectangles. We observe that although the central values of $R_{\rm n}$ and $R_{\rm c}$ have slightly moved towards each other, the puzzle is as prominent as ever. The experimental region can now be reached without an enhancement of q, but a large CP-violating phase ϕ of the order of $-90°$ is still required:

$$q = 0.99\,^{+0.66}_{-0.70} \,, \quad \phi = -(94\,^{+16}_{-17})° \,. \tag{242}$$

Interestingly, ϕ of the order of $+90°$ can now also bring us rather close to the experimental range of $R_{\rm n}$ and $R_{\rm c}$.

An interesting probe of the NP phase ϕ is also provided by the CP violation in $B_d^0 \to \pi^0 K_{\rm S}$. Within the SM, the corresponding observables are expected to satisfy the following relations [163]:

$$\mathcal{A}_{\rm CP}^{\rm dir}(B_d \to \pi^0 K_{\rm S}) \approx 0 \,, \quad \mathcal{A}_{\rm CP}^{\rm mix}(B_d \to \pi^0 K_{\rm S}) \approx \mathcal{A}_{\rm CP}^{\rm mix}(B_d \to \psi K_{\rm S}) \,. \tag{243}$$

The most recent Belle [162] and BaBar [202] measurements of these quantities are in agreement with each other, and lead to the following averages [61]:

$$\mathcal{A}_{\rm CP}^{\rm dir}(B_d \to \pi^0 K_{\rm S}) = -0.02 \pm 0.13 \tag{244}$$
$$\mathcal{A}_{\rm CP}^{\rm mix}(B_d \to \pi^0 K_{\rm S}) = -0.31 \pm 0.26 \equiv -(\sin 2\beta)_{\pi^0 K_{\rm S}} \,. \tag{245}$$

Taking (176) into account yields

$$\Delta S \equiv (\sin 2\beta)_{\pi^0 K_{\rm S}} - (\sin 2\beta)_{\psi K_{\rm S}} = -0.38 \pm 0.26 \,, \tag{246}$$

which may indicate a sizeable deviation of the experimentally measured value of $(\sin 2\beta)_{\pi^0 K_{\rm S}}$ from $(\sin 2\beta)_{\psi K_{\rm S}}$, and is therefore one of the recent hot topics. Since the strategy developed in Refs. [82, 83] allows us also to predict the CP-violating observables of the $B_d^0 \to \pi^0 K_{\rm S}$ channel both within the SM and within our scenario of NP, it allows us to address this issue, yielding

$$\mathcal{A}_{\rm CP}^{\rm dir}(B_d \to \pi^0 K_{\rm S})|_{\rm SM} = 0.06\,^{+0.09}_{-0.10} \,, \quad \Delta S|_{\rm SM} = 0.13 \pm 0.05 \,, \tag{247}$$

$$\mathcal{A}_{\rm CP}^{\rm dir}(B_d \to \pi^0 K_{\rm S})|_{\rm NP} = 0.01\,^{+0.14}_{-0.18} \,, \quad \Delta S|_{\rm NP} = 0.27\,^{+0.05}_{-0.09} \,, \tag{248}$$
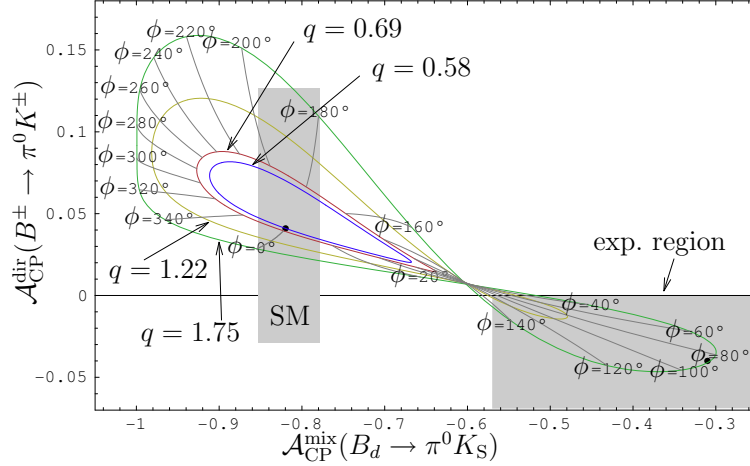
Fig. 29: The situation in the $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_d \to \pi^0 K_{\mathrm{S}})$–$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B^{\pm} \to \pi^0 K^{\pm})$ plane: the shaded regions represent the experimental and SM 1σ ranges, while the lines show the theory predictions for the central values of the hadronic parameters and various values of q with $\phi \in [0°, 360°]$.

where the NP results refer to the EW penguin parameters in (242). Consequently, ΔS is found to be *positive* in the SM. In the literature, values of $\Delta S|_{\mathrm{SM}} \sim 0.04$–$0.08$ can be found, which were obtained— in contrast to (247)—with the help of dynamical approaches such as QCDF [166] and SCET [179]. Moreover, bounds were derived with the help of the $SU(3)$ flavour symmetry [203]. Looking at (248), we see that the modified parameters (q, ϕ) in (242) imply an enhancement of ΔS with respect to the SM case. Consequently, the best values of (q, ϕ) that are favoured by the measurements of $R_{\mathrm{n,c}}$ make the potential $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_d \to \pi^0 K_{\mathrm{S}})$ discrepancy even larger than in the SM.

There is one CP asymmetry of the $B \to \pi K$ system left, which is measured as

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B^{\pm} \to \pi^0 K^{\pm}) = -0.04 \pm 0.04 . \tag{249}$$

In the limit of vanishing colour-suppressed tree and EW penguin topologies, it is expected to be equal to the direct CP asymmetry of the $B_d \to \pi^{\mp} K^{\pm}$ modes. Since the experimental value of the latter asymmetry in (219) does not agree with (249), the direct CP violation in $B^{\pm} \to \pi^0 K^{\pm}$ has also received a lot of attention. The lifted colour suppression described by the large value of x in (230) could, in principle, be responsible for a non-vanishing difference between (219) and (249),

$$\Delta A \equiv \mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B^{\pm} \to \pi^0 K^{\pm}) - \mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_d \to \pi^{\mp} K^{\pm}) \overset{\mathrm{exp}}{=} -0.16 \pm 0.04 . \tag{250}$$

However, applying once again the strategy described above yields

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B^{\pm} \to \pi^0 K^{\pm})|_{\mathrm{SM}} = 0.04 \, {}^{+0.09}_{-0.07} , \tag{251}$$

so that the SM still prefers a positive value of this CP asymmetry; the NP scenario characterized by (242) corresponds to

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B^{\pm} \to \pi^0 K^{\pm})|_{\mathrm{NP}} = 0.09 \, {}^{+0.20}_{-0.16} . \tag{252}$$

In view of the large uncertainties, no stringent test is provided at this point. Nevertheless, it is tempting to play a bit with the CP asymmetries of the $B^{\pm} \to \pi^0 K^{\pm}$ and $B_d \to \pi^0 K_{\mathrm{S}}$ decays. In Fig. 29, we show the situation in the $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_d \to \pi^0 K_{\mathrm{S}})$–$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B^{\pm} \to \pi^0 K^{\pm})$ plane for various values of q with $\phi \in [0°, 360°]$. We see that these observables seem to show a preference for positive values of ϕ around $+90°$. As we noted above, in this case, we can also get rather close to the experimental region in the R_{n}–R_{c} plane. It is now interesting to return to the discussion of the NP effects in the $B \to \phi K$ system given

in Section 7.2. In our scenario of NP in the EW penguin sector, we have just to identify the CP-violating phase ϕ_0 in (205) with the NP phase ϕ [83]. Unfortunately, we cannot determine the hadronic $B \to \phi K$ parameters $\tilde{v}_0$ and $\tilde{\Delta}_0$ through the $B \to \pi\pi$ data as in the case of the $B \to \pi K$ system. However, if we take into account that $\tilde{\Delta}_0 = 180°$ in factorization and look at Fig. 25, we see again that the case of $\phi \sim +90°$ would be favoured by the data for $\mathcal{S}_{\phi K}$. Alternatively, in the case of $\phi \sim -90°$, $\tilde{\Delta}_0 \sim 0°$ would be required to accommodate a negative value of $\mathcal{S}_{\phi K}$, which appears unlikely. Interestingly, a similar comment applies to the $B \to J/\psi K$ observables shown in Fig. 21, although here a dramatic enhancement of the EW penguin parameter v_0 relative to the SM estimate would be simultaneously needed to reach the central experimental values, in contract to the reduction of $\tilde{v}_0$ in the $B \to \phi K$ case. In view of rare decay constraints, the behaviour of the $B \to \phi K$ parameter $\tilde{v}_0$ appears much more likely, thereby supporting the assumption after (225).

8.4 The interplay with rare K and B decays and future scenarios

In order to explore the implications of the $B \to \pi K$ puzzle for rare K and B decays, we assume that the NP enters the EW penguin sector through Z^0 penguins with a new CP-violating phase. This scenario was already considered in the literature, where model-independent analyses and studies within SUSY can be found [204, 205]. In the strategy discussed here, the short-distance function C characterizing the Z^0 penguins is determined through the $B \to \pi K$ data [206]. Performing a renormalization-group analysis yields

$$C(\bar{q}) = 2.35\,\bar{q}e^{i\phi} - 0.82 \quad \text{with} \quad \bar{q} = q\left[\frac{|V_{ub}/V_{cb}|}{0.086}\right]. \tag{253}$$

Evaluating then the relevant box-diagram contributions in the SM and using (253), the short-distance functions

$$X = 2.35\,\bar{q}e^{i\phi} - 0.09 \quad \text{and} \quad Y = 2.35\,\bar{q}e^{i\phi} - 0.64 \tag{254}$$

can also be calculated, which govern the rare K, B decays with $\nu\bar{\nu}$ and $\ell^+\ell^-$ in the final states, respectively. In the SM, we have $C = 0.79$, $X = 1.53$ and $Y = 0.98$, with *vanishing* CP-violating phases. An analysis along these lines shows that the value of (q, ϕ) in (242), which is preferred by the $B \to \pi K$ observables $R_{\mathrm{n,c}}$, requires the following lower bounds for X and Y [91]:

$$|X|_{\min} \approx |Y|_{\min} \approx 2.2\,, \tag{255}$$

which appear to violate the 95% probability upper bounds

$$X \leq 1.95, \quad Y \leq 1.43 \tag{256}$$

that were recently obtained within the context of MFV [207]. Although we have to deal with CP-violating NP phases in our scenario, which goes therefore beyond the MFV framework, a closer look at $B \to X_s\ell^+\ell^-$ shows that the upper bound on $|Y|$ in (256) is difficult to avoid if NP enters only through EW penguins and the operator basis is the same as in the SM. A possible solution to the clash between (255) and (256) would be given by more complicated NP scenarios [91]. However, unless a specific model is chosen, the predictive power is then significantly reduced. For the exploration of the NP effects in rare decays, we shall therefore not follow this avenue.

Using an only slightly more generous bound on $|Y|$ by imposing $|Y| \leq 1.5$ and taking only those values of (242) that satisfy the constraint $|Y| = 1.5$ yields

$$q = 0.48 \pm 0.07\,, \quad \phi = -(93 \pm 17)°\,, \tag{257}$$

corresponding to a modest *suppression* of q relative to its updated SM value of 0.58. It is interesting to investigate the impact of various modifications of (q, ϕ), which allow us to satisfy the bounds in (256), for the $B \to \pi K$ observables and rare decays. To this end, three scenarios for the possible future evolution of the measurements of R_{n} and R_{c} were introduced in Ref. [91]:

Table 1: The $B \to \pi K$ observables for the three scenarios introduced in the text

Quantity	SM	Scen. A	Scen. B	Scen. C	Experiment
R_{n}	1.12	0.88	1.03	1	0.83 ± 0.08
R_{c}	1.15	0.96	1.13	1	1.01 ± 0.09
$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B^{\pm} \to \pi^0 K^{\pm})$	0.04	0.07	0.06	0.02	-0.04 ± 0.04
$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_d \to \pi^0 K_{\mathrm{S}})$	0.06	0.04	0.03	0.09	-0.02 ± 0.13
$\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_d \to \pi^0 K_{\mathrm{S}})$	-0.82	-0.89	-0.91	-0.70	-0.31 ± 0.26
ΔS	0.13	0.21	0.22	0.01	-0.38 ± 0.26
ΔA	-0.07	-0.04	-0.05	-0.09	-0.16 ± 0.04

Table 2: Rare decay branching ratios for the three scenarios introduced in the text. The $B_s \to \mu^+ \mu^-$ channel will be discussed in more detail in Section 10.5.

Decay	SM	Scen. A	Scen. B	Scen. C	Exp. bound (90% C.L.)
$\mathrm{BR}(K^+ \to \pi^+ \nu \bar{\nu})/10^{-11}$	9.3	2.7	8.3	8.4	$(14.7^{+13.0}_{-8.9})$
$\mathrm{BR}(K_{\mathrm{L}} \to \pi^0 \nu \bar{\nu})/10^{-11}$	4.4	11.6	27.9	7.2	$< 2.9 \times 10^4$
$\mathrm{BR}(K_{\mathrm{L}} \to \pi^0 e^+ e^-)/10^{-11}$	3.6	4.6	7.1	4.9	< 28
$\mathrm{BR}(B \to X_s \nu \bar{\nu})/10^{-5}$	3.6	2.8	4.8	3.3	< 64
$\mathrm{BR}(B_s \to \mu^+ \mu^-)/10^{-9}$	3.9	9.2	9.1	7.0	$< 1.5 \times 10^2$
$\mathrm{BR}(K_{\mathrm{L}} \to \mu^+ \mu^-)_{\mathrm{SD}}/10^{-9}$	0.9	0.9	0.001	0.6	< 2.5

– *Scenario A:* $q = 0.48$, $\phi = -93°$, which is in accordance with the currrent rare decay bounds and the $B \to \pi K$ data [see (257)].

– *Scenario B:* $q = 0.66$, $\phi = -50°$, which yields an increase of R_{n} to 1.03, and some interesting effects in rare decays. This could, for example, happen if radiative corrections to the $B_d^0 \to \pi^- K^+$ branching ratio enhance R_{n} [208], though this alone would probably account for only about 5%.

– *Scenario C:* here it is assumed that $R_{\mathrm{n}} = R_{\mathrm{c}} = 1$, which corresponds to $q = 0.54$ and $\phi = 61°$. The *positive* sign of ϕ distinguishes this scenario strongly from the others.

The patterns of the observables of the $B \to \pi K$ and rare decays corresponding to these scenarios are collected in Tables 1 and 2, respectively. We observe that the $K \to \pi \nu \bar{\nu}$ modes, which are theoretically very clean (for a recent review, see Ref. [209]), offer a particularly interesting probe for the different scenarios. Concerning the observables of the $B \to \pi K$ system, $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_d \to \pi^0 K_{\mathrm{S}})$ is very interesting: this CP asymmetry is found to be very large in Scenarios A and B, where the NP phase ϕ is negative. On the other hand, the positive sign of ϕ in Scenario C brings $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_d \to \pi^0 K_{\mathrm{S}})$ closer to the data, in agreement with the features discussed in Section 8.3. A similar comment applies to the direct CP asymmetry of $B^{\pm} \to \pi^0 K^{\pm}$.

In view of the large uncertainties, unfortunately no definite conclusions on the presence of NP can be drawn at this stage. However, the possible anomalies in the $B \to \pi K$ system complemented with the one in $B \to \phi K$ may actually indicate the effects of a modified EW penguin sector with a large CP-violating NP phase. As we just saw, rare K and B decays have an impressive power to reveal such a kind of NP. Finally, let us stress that the analysis of the $B \to \pi\pi$ modes, which signals large non-factorizable effects, and the determination of the UT angle γ described above are not affected by such NP effects. It will be interesting to monitor the evolution of the corresponding data with the help of the strategy discussed above.

9 Entering a new territory: $b \to d$ penguins

9.1 Preliminaries

Another hot topic which emerged recently is the exploration of $b \to d$ penguin processes. The non-leptonic decays belonging to this category, which are mediated by $b \to d\bar{s}s$ quark transitions (see the classification in Section 3.3.1), are now coming within experimental reach at the B factories. A similar comment applies to the radiative decays originating from $b \to d\gamma$ processes, whereas $b \to d\ell^+\ell^-$ modes are still far from being accessible. The B factories are therefore just entering a new territory, which is still essentially unexplored. Let us now have a closer look at the corresponding processes.

9.2 A prominent example: $B_d^0 \to K^0\bar{K}^0$

The Feynman diagrams contributing to this decay can be obtained from those for $B_d^0 \to \phi K^0$ shown in Fig. 23 by replacing the anti-strange quark emerging from the W boson through an anti-down quark. The $B_d^0 \to K^0\bar{K}^0$ decay is described by the low-energy effective Hamiltonian in (80) with $r = d$, where the current–current operators may only contribute through penguin-like contractions, corresponding to the penguin topologies with internal up- and charm-quark exchanges. The dominant role is played by QCD penguins; since EW penguins contribute only in colour-suppressed form, they have a minor impact on $B_d^0 \to K^0\bar{K}^0$, in contrast to the case of $B_d^0 \to \phi K^0$, where they may also contribute in colour-allowed form.

If we apply the notation introduced in Section 7, again make use of the unitarity of the CKM matrix and apply the Wolfenstein parametrization, we may write the $B_d^0 \to K^0\bar{K}^0$ amplitude as follows:

$$A(B_d^0 \to K^0\bar{K}^0) = \lambda^3 A(\tilde{A}_{\mathrm{P}}^t - \tilde{A}_{\mathrm{P}}^c)\left[1 - \rho_{KK}e^{i\theta_{KK}}e^{i\gamma}\right], \tag{258}$$

where

$$\rho_{KK}e^{i\theta_{KK}} \equiv R_b\left[\frac{\tilde{A}_{\mathrm{P}}^t - \tilde{A}_{\mathrm{P}}^u}{\tilde{A}_{\mathrm{P}}^t - \tilde{A}_{\mathrm{P}}^c}\right]. \tag{259}$$

This expression allows us to calculate the CP-violating asymmetries with the help of the formulae given in Section 5.5, taking the following form:

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}(B_d \to K^0\bar{K}^0) = D_1(\rho_{KK}, \theta_{KK}; \gamma) \tag{260}$$

$$\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_d \to K^0\bar{K}^0) = D_2(\rho_{KK}, \theta_{KK}; \gamma, \phi_d). \tag{261}$$

Let us assume, for a moment, that the penguin contributions are dominated by top-quark exchanges. In this case, (259) simplifies as

$$\rho_{KK}e^{i\theta_{KK}} \to R_b. \tag{262}$$

Since the CP-conserving strong phase θ_{KK} vanishes in this limit, the direct CP violation in $B_d^0 \to K^0\bar{K}^0$ vanishes, too. Moreover, if we take into account that $\phi_d = 2\beta$ in the SM and use trigonometrical relations which can be derived for the UT, we find that the mixing-induced CP asymmetry also would be zero. These features suggest an interesting test of the $b \to d$ flavour sector of the SM (see, for instance, Ref. [210]). However, contributions from penguins with internal up- and charm-quark exchanges are expected to yield sizeable CP asymmetries in $B_d^0 \to K^0\bar{K}^0$ even within the SM, so that the interpretation of these effects is much more complicated [211]; these contributions contain also possible long-distance rescattering effects [212], which are often referred to as 'GIM' and 'charming' penguins and recently received a lot of attention [213].

Despite this problem, interesting insights can be obtained through the $B_d^0 \to K^0\bar{K}^0$ observables [214]. By the time the CP-violating asymmetries in (260) and (261) can be measured, the angle
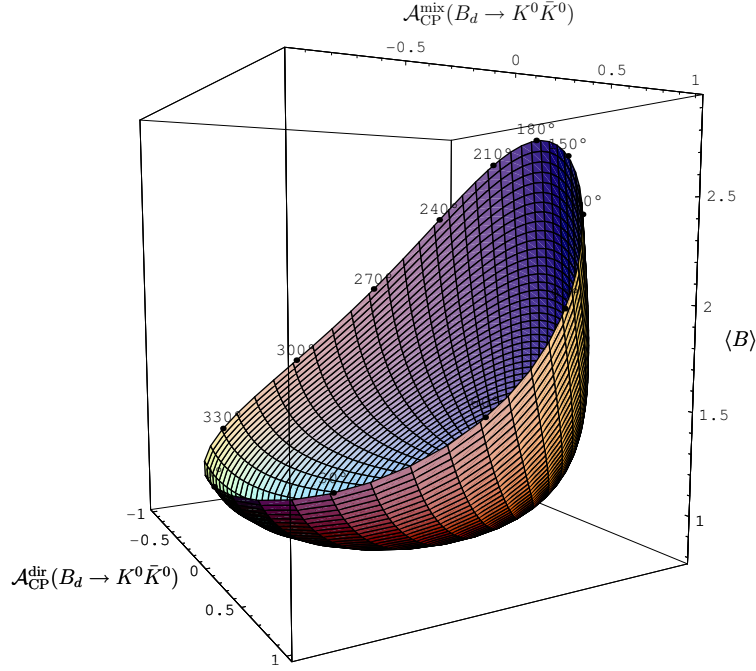
Fig. 30: Illustration of the surface in the $\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}$–$\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}$–$\langle B \rangle$ observable space characterizing the $B_d^0 \to K^0 \bar{K}^0$ decay in the SM. The intersecting lines on the surface correspond to constant values of ρ_{KK} and θ_{KK}; the numbers on the fringe indicate the value of θ_{KK}, while the fringe itself is defined by $\rho_{KK} = 1$.

γ of the UT will also be reliably known, in addition to the B_d^0–$\bar{B}_d^0$ mixing phase ϕ_d. The experimental values of the CP asymmetries can then be converted into ρ_{KK} and θ_{KK}, in analogy to the $B \to \pi\pi$ discussion in Section 8.2. Although these quantities are interesting to obtain insights into the $B \to \pi K$ parameter $\rho_c e^{i\theta_c}$ [see (237)] through $SU(3)$ arguments, and can be compared with theoretical predictions, for instance, those of QCDF, PQCD or SCET, they do not provide—by themselves—a test of the SM description of the FCNC processes mediating the decay $B_d^0 \to K^0 \bar{K}^0$. However, so far, we have not yet used the information offered by the CP-averaged branching ratio of this channel. It takes the following form:

$$\mathrm{BR}(B_d \to K^0 \bar{K}^0) = \frac{\tau_{B_d}}{16\pi M_{B_d}} \times \Phi_{KK} \times |\lambda^3 A \, \tilde{A}_{\mathrm{P}}^{tc}|^2 \langle B \rangle \,, \tag{263}$$

where Φ_{KK} denotes a two-body phase-space factor, $\tilde{A}_{\mathrm{P}}^{tc} \equiv \tilde{A}_{\mathrm{P}}^t - \tilde{A}_{\mathrm{P}}^c$, and

$$\langle B \rangle \equiv 1 - 2\rho_{KK} \cos\theta_{KK} \cos\gamma + \rho_{KK}^2 \,. \tag{264}$$

If we now use ϕ_d and the SM value of γ, we may characterize the decay $B_d^0 \to K^0 \bar{K}^0$—within the SM—through a surface in the observable space of $\mathcal{A}_{\mathrm{CP}}^{\mathrm{dir}}$, $\mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}$ and $\langle B \rangle$. In Fig. 30, we show this surface, where each point corresponds to a given value of ρ_{KK} and θ_{KK}. It should be emphasized that this surface is *theoretically clean* since it relies only on the general SM parametrization of $B_d^0 \to K^0 \bar{K}^0$. Consequently, should future measurements give a value in observable space that should *not* lie on the SM surface, we would have immediate evidence for NP contributions to $\bar{b} \to \bar{d} s \bar{s}$ processes.

Looking at Fig. 30, we see that $\langle B \rangle$ takes an absolute minimum. Indeed, if we keep ρ_{KK} and θ_{KK} as free parameters in (264), we find

$$\langle B \rangle \geq \sin^2 \gamma \,, \tag{265}$$

which yields a strong lower bound because of the favourably large value of γ. Whereas the direct and mixing-induced CP asymmetries can be extracted from a time-dependent rate asymmetry [see (152)], the determination of $\langle B \rangle$ requires further information to fix the overall normalization factor involving

the penguin amplitude $\tilde{A}_{\mathrm{P}}^{tc}$. The strategy developed in Refs. [82, 83] offers the following two avenues, using data for

i) $B \to \pi\pi$ decays, i.e., $b \to d$ transitions, implying the following lower bound:

$$\mathrm{BR}(B_d \to K^0\bar{K}^0)_{\min} = \Xi_\pi^K \times \left(1.39\,^{+1.54}_{-0.95}\right) \times 10^{-6}\,, \tag{266}$$

ii) $B \to \pi K$ decays, i.e., $b \to s$ transitions, which are complemented by the $B \to \pi\pi$ system to determine a small correction, implying the following lower bound:

$$\mathrm{BR}(B_d \to K^0\bar{K}^0)_{\min} = \Xi_\pi^K \times \left(1.36\,^{+0.18}_{-0.21}\right) \times 10^{-6}\,. \tag{267}$$

Here factorizable $SU(3)$-breaking corrections are included, as is made explicit through

$$\Xi_\pi^K = \left[\frac{f_0^K}{0.331}\frac{0.258}{f_0^\pi}\right]^2\,, \tag{268}$$

where the numerical values for the $B \to K, \pi$ form factors $f_0^{K,\pi}$ refer to a recent light-cone sum-rule analysis [215]. At the time of the derivation of these bounds, the B factories reported an experimental *upper* bound of $\mathrm{BR}(B_d \to K^0\bar{K}^0) < 1.5 \times 10^{-6}$ (90% C.L.). Consequently, the theoretical *lower* bounds given above suggested that the observation of this channel should just be ahead of us. Subsequently, the first signals were indeed announced, in accordance with (266) and (267):

$$\mathrm{BR}(B_d \to K^0\bar{K}^0) = \begin{cases} (1.19^{+0.40}_{-0.35} \pm 0.13) \times 10^{-6} & \text{(BaBar [216])}\,, \\ (0.8 \pm 0.3 \pm 0.1) \times 10^{-6} & \text{(Belle [217])}\,. \end{cases} \tag{269}$$

The SM description of $B_d^0 \to K^0\bar{K}^0$ has thus successfully passed its first test. However, the experimental errors are still very large, and the next crucial step—a measurement of the CP asymmetries—is still missing. Using QCDF, an analysis of NP effects in this channel was recently performed in the minimal supersymmetric standard model [218]. For further aspects of $B_d^0 \to K^0\bar{K}^0$, the reader is referred to Ref. [214].

9.3 Radiative $b \to d$ penguin decays: $\bar{B} \to \rho\gamma$

Another important tool to explore $b \to d$ penguins is provided by $\bar{B} \to \rho\gamma$ modes. In the SM, these decays are described by a Hamiltonian with the following structure [67]:

$$\mathcal{H}_{\mathrm{eff}}^{b \to d\gamma} = \frac{G_{\mathrm{F}}}{\sqrt{2}} \sum_{j=u,c} V_{jd}^* V_{jb} \left[\sum_{k=1}^{2} C_k Q_k^{jd} + \sum_{k=3}^{8} C_k Q_k^d\right]. \tag{270}$$

Here the $Q_{1,2}^{jd}$ denote the current–current operators, whereas the $Q_{3...6}^d$ are the QCD penguin operators, which govern the decay $\bar{B}_d^0 \to K^0\bar{K}^0$ together with the penguin-like contractions of $Q_{1,2}^{cd}$ and $Q_{1,2}^{ud}$. In contrast to these four-quark operators,

$$Q_{7,8}^d = \frac{1}{8\pi^2} m_b \bar{d}_i \sigma^{\mu\nu}(1 + \gamma_5)\left\{eb_i F_{\mu\nu}, g_{\mathrm{s}} T_{ij}^a b_j G_{\mu\nu}^a\right\} \tag{271}$$

are electro- and chromomagnetic penguin operators. The most important contributions to $\bar{B} \to \rho\gamma$ originate from $Q_{1,2}^{jd}$ and $Q_{7,8}^d$, whereas the QCD penguin operators play only a minor role, in contrast to $\bar{B}_d^0 \to K^0\bar{K}^0$. If we use again the unitarity of the CKM matrix and apply the Wolfenstein parametrization, we may write

$$A(\bar{B} \to \rho\gamma) = c_\rho \lambda^3 A \mathcal{P}_{tc}^{\rho\gamma}\left[1 - \rho_{\rho\gamma}e^{i\theta_{\rho\gamma}}e^{-i\gamma}\right]\,, \tag{272}$$

where $c_\rho = 1/\sqrt{2}$ and 1 for $\rho = \rho^0$ and $\rho^\pm$, respectively, $\mathcal{P}_{tc}^{\rho\gamma} \equiv \mathcal{P}_t^{\rho\gamma} - \mathcal{P}_c^{\rho\gamma}$, and

$$\rho_{\rho\gamma} e^{i\theta_{\rho\gamma}} \equiv R_b \left[\frac{\mathcal{P}_t^{\rho\gamma} - \mathcal{P}_u^{\rho\gamma}}{\mathcal{P}_t^{\rho\gamma} - \mathcal{P}_c^{\rho\gamma}} \right] . \tag{273}$$

Here we follow our previous notation, i.e., the $\mathcal{P}_j^{\rho\gamma}$ are strong amplitudes with the following interpretation: $\mathcal{P}_u^{\rho\gamma}$ and $\mathcal{P}_c^{\rho\gamma}$ refer to the matrix elements of $\sum_{k=1}^2 C_k Q_k^{ud}$ and $\sum_{k=1}^2 C_k Q_k^{cd}$, respectively, whereas $\mathcal{P}_t^{\rho\gamma}$ corresponds to $-\sum_{k=3}^8 C_k Q_k^d$. Consequently, $\mathcal{P}_u^{\rho\gamma}$ and $\mathcal{P}_c^{\rho\gamma}$ describe the penguin topologies with internal up- and charm-quark exchanges, respectively, whereas $\mathcal{P}_t^{\rho\gamma}$ corresponds to the penguins with the top quark running in the loop. Let us note that (272) refers to a given photon helicity. However, the b quarks couple predominantly to left-handed photons in the SM, so that the right-handed amplitude is usually neglected [219]; we shall return to this point below. Comparing (272) with (258), we observe that the structure of both amplitudes is the same. In analogy to $\rho_{KK} e^{i\theta_{KK}}$, $\rho_{\rho\gamma} e^{i\theta_{\rho\gamma}}$ may also be affected by long-distance effects, which represent a key uncertainty of $\bar{B} \to \rho\gamma$ decays [147, 219].

If we replace all down quarks in (270) by strange quarks, we obtain the Hamiltonian for $b \to s\gamma$ processes, which are already well established experimentally [61]:

$$\text{BR}(B^\pm \to K^{*\pm}\gamma) = (40.3 \pm 2.6) \times 10^{-6} \tag{274}$$
$$\text{BR}(B_d^0 \to K^{*0}\gamma) = (40.1 \pm 2.0) \times 10^{-6} . \tag{275}$$

In analogy to (272), we may write

$$A(\bar{B} \to K^*\gamma) = -\frac{\lambda^3 A \mathcal{P}_{tc}^{K^*\gamma}}{\sqrt{\epsilon}} \left[1 + \epsilon \rho_{K^*\gamma} e^{i\theta_{K^*\gamma}} e^{-i\gamma} \right], \tag{276}$$

where ϵ was introduced in (221). Thanks to the smallness of ϵ, the parameter $\rho_{K^*\gamma} e^{i\theta_{K^*\gamma}}$ plays an essentially negligible role for the $\bar{B} \to K^*\gamma$ transitions.

Let us have a look at the charged decays $B^\pm \to \rho^\pm\gamma$ and $B^\pm \to K^{*\pm}\gamma$ first. If we consider their CP-averaged branching ratios, we obtain

$$\frac{\text{BR}(B^\pm \to \rho^\pm\gamma)}{\text{BR}(B^\pm \to K^{*\pm}\gamma)} = \epsilon \left[\frac{\Phi_{\rho\gamma}}{\Phi_{K^*\gamma}} \right] \left| \frac{\mathcal{P}_{tc}^{\rho\gamma}}{\mathcal{P}_{tc}^{K^*\gamma}} \right|^2 H_{K^*\gamma}^{\rho\gamma} , \tag{277}$$

where $\Phi_{\rho\gamma}$ and $\Phi_{K^*\gamma}$ denote phase-space factors, and

$$H_{K^*\gamma}^{\rho\gamma} \equiv \frac{1 - 2\rho_{\rho\gamma}\cos\theta_{\rho\gamma}\cos\gamma + \rho_{\rho\gamma}^2}{1 + 2\epsilon\rho_{K^*\gamma}\cos\theta_{K^*\gamma}\cos\gamma + \epsilon^2\rho_{K^*\gamma}^2} . \tag{278}$$

Since $B^\pm \to \rho^\pm\gamma$ and $B^\pm \to K^{*\pm}\gamma$ are related through the interchange of all down and strange quarks, the U-spin flavour symmetry of strong interactions allows us to relate the corresponding hadronic amplitudes to each other; the U-spin symmetry is an $SU(2)$ subgroup of the full $SU(3)_{\rm F}$ flavour-symmetry group, which relates down and strange quarks in the same manner as the conventional strong isospin symmetry relates down and up quarks. Following these lines, we obtain

$$|\mathcal{P}_{tc}^{\rho\gamma}| = |\mathcal{P}_{tc}^{K^*\gamma}| \tag{279}$$

$$\rho_{\rho\gamma} e^{i\theta_{\rho\gamma}} = \rho_{K^*\gamma} e^{i\theta_{K^*\gamma}} \equiv \rho e^{i\theta} . \tag{280}$$

Although we may determine the ratio of the penguin amplitudes $|\mathcal{P}_{tc}|$ in (277) with the help of (279)—up to $SU(3)$-breaking effects to be discussed below—we are still left with the dependence on ρ and θ. However, keeping ρ and θ as free parameters, it can be shown that $H_{K^*\gamma}^{\rho\gamma}$ satisfies the following relation [220]:

$$H_{K^*\gamma}^{\rho\gamma} \geq \left[1 - 2\epsilon\cos^2\gamma + \mathcal{O}(\epsilon^2) \right] \sin^2\gamma , \tag{281}$$

where the term linear in ϵ gives a shift of about 1.9%.

Concerning possible $SU(3)$-breaking effects to (280), they may only enter this tiny correction and are negligible for our analysis. On the other hand, the $SU(3)$-breaking corrections to (279) have a sizeable impact. Following Refs. [221, 222], we write

$$\left[\frac{\Phi_{\rho\gamma}}{\Phi_{K^*\gamma}}\right]\left|\frac{\mathcal{P}_{tc}^{\rho\gamma}}{\mathcal{P}_{tc}^{K^*\gamma}}\right|^2 = \left[\frac{M_B^2 - M_\rho^2}{M_B^2 - M_{K^*}^2}\right]^3 \zeta^2 , \tag{282}$$

where $\zeta = F_\rho/F_{K^*}$ is the $SU(3)$-breaking ratio of the $B^\pm \to \rho^\pm\gamma$ and $B^\pm \to K^{*\pm}\gamma$ form factors; a light-cone sum-rule analysis gives $\zeta^{-1} = 1.31 \pm 0.13$ [223]. Consequently, (281) and (282) allow us to convert the measured $B^\pm \to K^{*\pm}\gamma$ branching ratio (274) into a *lower* SM bound for BR($B^\pm \to \rho^\pm\gamma$) with the help of (277) [220]:

$$\text{BR}(B^\pm \to \rho^\pm\gamma)_{\min} = \left(1.02\,^{+0.27}_{-0.23}\right) \times 10^{-6} . \tag{283}$$

A similar kind of reasoning holds also for the U-spin pairs $B^\pm \to K^\pm K, \pi^\pm K$ and $B^\pm \to K^\pm K^*, \pi^\pm K^*$, where the following lower bounds can be derived [220]:

$$\text{BR}(B^\pm \to K^\pm K)_{\min} = \Xi_\pi^K \times \left(1.69\,^{+0.21}_{-0.24}\right) \times 10^{-6} \tag{284}$$

$$\text{BR}(B^\pm \to K^\pm K^*)_{\min} = \Xi_\pi^K \times \left(0.68\,^{+0.11}_{-0.13}\right) \times 10^{-6} , \tag{285}$$

with Ξ_π^K given in (268). Thanks to the most recent B-factory data, we now also have evidence for $B^\pm \to K^\pm K$ decays:

$$\text{BR}(B^\pm \to K^\pm K) = \begin{cases} (1.5 \pm 0.5 \pm 0.1) \times 10^{-6} & \text{(BaBar [216])} \\ (1.0 \pm 0.4 \pm 0.1) \times 10^{-6} & \text{(Belle [217])} , \end{cases} \tag{286}$$

whereas the upper limit of 5.3×10^{-6} for $B^\pm \to K^\pm K^*$ still leaves a lot of space. Obviously, we may also consider the $B^\pm \to K^{*\pm}K, \rho^\pm K$ system [220]. However, since currently only the upper bound BR($B^\pm \to \rho^\pm K$) $< 48 \times 10^{-6}$ is available, we cannot yet give a number for the lower bound on BR($B^\pm \to K^{*\pm}K$). Experimental analyses of these modes are strongly encouraged.

Let us now turn to $\bar{B}_d^0 \to \rho^0\gamma$, which receives contributions from exchange and penguin annihilation topologies that are not present in $\bar{B}_d^0 \to \bar{K}^{*0}\gamma$; in the case of $B^\pm \to \rho^\pm\gamma$ and $B^\pm \to K^{*\pm}\gamma$, which are related by the U-spin symmetry, there is a one-to-one correspondence of topologies. Making the plausible assumption that the topologies involving the spectator quarks play a minor role, and taking the factor of $c_{\rho^0} = 1/\sqrt{2}$ in (272) into account, the counterpart of (283) is given by

$$\text{BR}(B_d \to \rho^0\gamma)_{\min} = \left(0.51\,^{+0.13}_{-0.11}\right) \times 10^{-6} . \tag{287}$$

At the time of the derivation of the *lower* bounds for the $B \to \rho\gamma$ branching ratios given above, the following experimental *upper* bounds (90% C.L.) were available:

$$\text{BR}(B^\pm \to \rho^\pm\gamma) < \begin{cases} 1.8 \times 10^{-6} & \text{(BaBar [224])} \\ 2.2 \times 10^{-6} & \text{(Belle [225])} \end{cases} \tag{288}$$

$$\text{BR}(B_d \to \rho^0\gamma) < \begin{cases} 0.4 \times 10^{-6} & \text{(BaBar [224])} \\ 0.8 \times 10^{-6} & \text{(Belle [225])} . \end{cases} \tag{289}$$

Consequently, it was expected that the $\bar{B} \to \rho\gamma$ modes should soon be discovered at the B factories [220]. Indeed, the Belle Collaboration recently reported the first observation of $b \to d\gamma$ processes [226]:

$$\text{BR}(B^\pm \to \rho^\pm\gamma) = \left(0.55\,^{+0.43+0.12}_{-0.37-0.11}\right) \times 10^{-6} \tag{290}$$

$$\mathrm{BR}(B_d \to \rho^0 \gamma) = \left(1.17^{+0.35+0.09}_{-0.31-0.08}\right) \times 10^{-6} \tag{291}$$

$$\mathrm{BR}(B \to (\rho, \omega)\gamma) = \left(1.34^{+0.34+0.14}_{-0.31-0.10}\right) \times 10^{-6} \, , \tag{292}$$

which was one of the hot topics of the 2005 summer conferences [227]. These measurements still suffer from large uncertainties, and the pattern of the central values of (290) and (291) would be in conflict with the expectation following from the isospin symmetry. It will be interesting to follow the evolution of the data. The next important conceptual step would be the measurement of the corresponding CP-violating observables, though this is still in the distant future.

An alternative avenue to confront the data for the $B \to \rho\gamma$ branching ratios with the SM is provided by converting them into information on the side R_t of the UT. To this end, the authors of Refs. [221, 222] also use (282), and calculate the CP-conserving (complex) parameter δa entering $\rho_{\rho\gamma} e^{i\theta_{\rho\gamma}} = R_b [1 + \delta a]$ in the QCDF approach. The corresponding result, which favours a small impact of δa, takes leading and next-to-leading order QCD corrections into account and holds to leading order in the heavy-quark limit [222]. In view of the remarks about possible long-distance effects made above and the B-factory data for the $B \to \pi\pi$ system, which indicate large corrections to the QCDF picture for non-leptonic B decays into two light pseudoscalar mesons (see Section 8.2), it is, however, not obvious that the impact of δa is actually small. The advantage of the bound following from (281) is that it is—by construction—*not* affected by $\rho_{\rho\gamma} e^{i\theta_{\rho\gamma}}$ at all.

9.4 General lower bounds for $b \to d$ penguin processes

Interestingly, the bounds discussed above are actually realizations of a general, model-independent bound that can be derived in the SM for $b \to d$ penguin processes [220]. If we consider such a decay, $\bar{B} \to \bar{f}_d$, we may—in analogy to (258) and (272)—write

$$A(\bar{B} \to \bar{f}_d) = A_d^{(0)} \left[1 - \varrho_d e^{i\theta_d} e^{-i\gamma}\right], \tag{293}$$

so that the CP-averaged amplitude square is given as follows:

$$\langle |A(B \to f_d)|^2 \rangle = |A_d^{(0)}|^2 \left[1 - 2\varrho_d \cos\theta_d \cos\gamma + \varrho_d^2\right] . \tag{294}$$

In general, ϱ_d and θ_d depend on the point in phase space considered. Consequently, the expression

$$\mathrm{BR}(B \to f_d) = \tau_B \left[\sum_{\mathrm{Pol}} \int d\,\mathrm{PS}\, \langle |A(B \to f_d)|^2 \rangle \right] \tag{295}$$

for the CP-averaged branching ratio, where the sum runs over possible polarization configurations of f_d, does *not* factorize into $|A_d^{(0)}|^2$ and $[1 - 2\varrho_d \cos\theta_d \cos\gamma + \varrho_d^2]$ as in the case of the two-body decays considered above. However, if we keep ϱ_d and θ_d as free, 'unknown' parameters at any given point in phase space, we obtain

$$\langle |A(B \to f_d)|^2 \rangle \geq |A_d^{(0)}|^2 \sin^2\gamma \, , \tag{296}$$

which implies

$$\mathrm{BR}(B \to f_d) \geq \tau_B \left[\sum_{\mathrm{Pol}} \int d\,\mathrm{PS}\, |A_d^{(0)}|^2 \right] \sin^2\gamma \, . \tag{297}$$

In order to deal with the term in square brackets, we use a $b \to s$ penguin decay $\bar{B} \to \bar{f}_s$, which is the counterpart of $\bar{B} \to \bar{f}_d$ in that the corresponding CP-conserving strong amplitudes can be related to one another through the $SU(3)$ flavour symmetry. In analogy to (276), we may then write

$$A(\bar{B} \to \bar{f}_s) = -\frac{A_s^{(0)}}{\sqrt{\epsilon}} \left[1 + \epsilon \varrho_s e^{i\theta_s} e^{-i\gamma}\right] . \tag{298}$$

If we neglect the term proportional to ϵ in the square bracket, we arrive at

$$\frac{\mathrm{BR}(B \to f_d)}{\mathrm{BR}(B \to f_s)} \geq \epsilon \left[\frac{\sum_{\mathrm{Pol}} \int d\,\mathrm{PS}\, |A_d^{(0)}|^2}{\sum_{\mathrm{Pol}} \int d\,\mathrm{PS}\, |A_s^{(0)}|^2} \right] \sin^2 \gamma \,. \tag{299}$$

Apart from the tiny ϵ correction, which gave a shift of about 1.9% in (281), (299) is valid exactly in the SM. If we now apply the $SU(3)$ flavour symmetry, we obtain

$$\frac{\sum_{\mathrm{Pol}} \int d\,\mathrm{PS}\, |A_d^{(0)}|^2}{\sum_{\mathrm{Pol}} \int d\,\mathrm{PS}\, |A_s^{(0)}|^2} \xrightarrow{SU(3)_{\mathrm{F}}} 1 \,. \tag{300}$$

Since $\sin^2 \gamma$ is favourably large in the SM and the decay $\bar{B} \to \bar{f}_s$ will be measured before its $b \to d$ counterpart—simply because of the CKM enhancement—(299) provides strong lower bounds for $\mathrm{BR}(B \to f_d)$.

It is instructive to return briefly to $B \to \rho\gamma$. If we look at (299), we observe immediately that the assumption that these modes are governed by a single photon helicity is no longer required. Consequently, (283) and (287) are actually very robust with respect to this issue, which may only affect the $SU(3)$-breaking corrections to a small extent. This feature is interesting in view of the recent discussion in Ref. [228], where the photon polarization in $B \to \rho\gamma$ and $B \to K^*\gamma$ decays was critically analysed.

We can now also derive a bound for the $B^{\pm} \to K^{*\pm}K^*, \rho^{\pm}K^*$ system, where we have to sum in (299) over three polarization configurations of the vector mesons. The analysis of the $SU(3)$-breaking corrections is more involved than in the case of the decays considered above, and the emerging lower bound of $\mathrm{BR}(B^{\pm} \to K^{*\pm}K^*)_{\mathrm{min}} \sim 0.6 \times 10^{-6}$ is still very far from the experimental upper bound of 71×10^{-6}. Interestingly, the theoretical lower bound would be reduced by ~ 0.6 in the strict $SU(3)$ limit, i.e., would be more conservative [220]. A similar comment applies to (266), (267) and (284), (285). On the other hand, the $B \to \rho\gamma$ bounds in (283) and (287) would be enhanced by ~ 1.7 in this case. However, here the theoretical situation is more favourable since we have not to rely on the factorization hypothesis to deal with the $SU(3)$-breaking effects as in the case of the non-leptonic decays.

Let us finally come to another application of (299), which is offered by decays of the kind $\bar{B} \to \pi \ell^+ \ell^-$ and $\bar{B} \to \rho \ell^+ \ell^-$. It is well known that the ρ_d terms complicate the interpretation of the corresponding data considerably [147]; the bound offers SM tests that are not affected by these contributions. The structure of the $b \to d \ell^+ \ell^-$ Hamiltonian is similar to (270), but involves the additional operators

$$Q_{9,10} = \frac{\alpha}{2\pi} (\bar{\ell}\ell)_{\mathrm{V,A}} (\bar{d}_i b_i)_{\mathrm{V-A}} \,. \tag{301}$$

The $b \to s\ell^+\ell^-$ modes $\bar{B} \to K\ell^+\ell^-$ and $\bar{B} \to K^*\ell^+\ell^-$ were already observed at the B factories, with branching ratios at the 0.6×10^{-6} and 1.4×10^{-6} levels [61], respectively, and received considerable theoretical attention (see, for example, Ref. [229]). For the application of (299), the charged decay combinations $B^{\pm} \to \pi^{\pm}\ell^+\ell^-, K^{\pm}\ell^+\ell^-$ and $B^{\pm} \to \rho^{\pm}\ell^+\ell^-, K^{*\pm}\ell^+\ell^-$ are suited best since the corresponding decay pairs are related to each other through the U-spin symmetry [230]. The numbers given above suggest

$$\mathrm{BR}(B^{\pm} \to \pi^{\pm}\ell^+\ell^-), \quad \mathrm{BR}(B^{\pm} \to \rho^{\pm}\ell^+\ell^-) \gtrsim 10^{-8} \,, \tag{302}$$

thereby leaving the exploration of these $b \to d$ penguin decays for the more distant future. Detailed studies of the associated $SU(3)$-breaking corrections are encouraged. It is hoped that by the time the $B^{\pm} \to \pi^{\pm}\ell^+\ell^-, \rho^{\pm}\ell^+\ell^-$ modes can be measured, we shall have a good picture of these effects.

It will be interesting to confront all of these bounds with experimental data. In the case of the non-leptonic $B_d \to K^0 \bar{K}^0$, $B^{\pm} \to K^{\pm}K$ modes and their radiative $B \to \rho\gamma$ counterparts, they have already provided a first successful test of the SM description of the corresponding FCNC processes, although the uncertainties are still very large in view of the fact that we are just at the beginning of the experimental

exploration of these channels. A couple of other non-leptonic decays of this kind may just be around the corner. It would be exciting if some bounds were significantly violated through destructive interference between SM and NP contributions. Since the different decay classes are governed by different operators, we could actually encounter surprises!

10 B-decay studies in the LHC era: fully exploiting the B_s system

10.1 In pursuit of new physics with ΔM_s

Concerning experimental information about this mass difference, only lower bounds were available for many years from the LEP experiments at CERN and SLD at SLAC [107]. Since the currently operating e^+e^- B factories run at the $\Upsilon(4S)$ resonance, which decays into $B_{u,d}$, but not into B_s mesons, the B_s system cannot be explored by the BaBar and Belle experiments[3]. However, plenty of B_s mesons are produced at the Tevatron (and will be later on at the LHC [232]), which—very recently—allowed the measurement of ΔM_s, as summarized in (131) and (132). These new results were one of the hot topics of spring 2006, and have already triggered several phenomenological papers (see, for example, Refs. [233]– [241]).

As in Section 6 and Section 7.1, we shall follow the analysis of Ref. [101]. In order to describe possible NP effects, we parametrize them through (164) and (165). The relevant CKM factor is $|V_{ts}^* V_{tb}|$. Using once again the unitarity of the CKM matrix and including next-to-leading order terms in the Wolfenstein expansion as given in Ref. [36], we have

$$\left| \frac{V_{ts}}{V_{cb}} \right| = 1 - \frac{1}{2} \left(1 - 2R_b \cos\gamma \right) \lambda^2 + \mathcal{O}(\lambda^4) \,. \tag{303}$$

Consequently, apart from the tiny correction in λ^2, the CKM factor for ΔM_s is independent of γ and R_b, which is an important advantage in comparison with the B_d-meson system. The accuracy of the SM prediction of ΔM_s is hence limited by the hadronic mixing parameter $f_{B_s} \hat{B}_{B_s}^{1/2}$. If we consider the ratio ρ_s introduced in (166) and use the CDF measurement in (132), we obtain

$$\rho_s|_{\mathrm{JLQCD}} = 1.08^{+0.03}_{-0.01}(\exp) \pm 0.19(\mathrm{th}) \tag{304}$$

$$\rho_s|_{\mathrm{(HP+JL)QCD}} = 0.74^{+0.02}_{-0.01}(\exp) \pm 0.18(\mathrm{th}) \,, \tag{305}$$

where we made the experimental and theoretical errors explicit. These numbers are consistent with the SM case $\rho_s = 1$, but suffer from significant theoretical uncertainties, which are much larger than the experimental errors. Nevertheless, it is interesting to note that the (HP+JL)QCD result is $1.5\,\sigma$ below the SM; a similar pattern arises in (193) and (194), though at the $1\,\sigma$ level. Any more precise statement about the presence or absence of NP requires the reduction of theoretical uncertainties.

In Fig. 31, we show the constraints in the σ_s–κ_s plane, which can be obtained from ρ_s with the help of the contours shown in Fig. 18. We see that upper bounds of $\kappa_s \lesssim 2.5$ arise from the measurement of ΔM_s. In the case of (305), σ_s would be constrainted to lie within the range $110° \leq \sigma_s \leq 250°$. Consequently, the CDF measurement of ΔM_s leaves ample space for the NP parameters σ_s and κ_s. As in the case of the B_d-meson system discussed in Section 7.1, this situation will change significantly as soon as information about CP violation in the B_s-meson system becomes available. We shall return to this topic in Section 10.2.

It is interesting to consider the ratio of ΔM_s and ΔM_d, which can be written as follows:

$$\frac{\Delta M_s}{\Delta M_d} = \frac{\rho_s}{\rho_d} \left| \frac{V_{ts}}{V_{td}} \right|^2 \frac{M_{B_s}}{M_{B_d}} \xi^2 \,, \tag{306}$$

---

[3]The asymmetric e^+e^- KEKB collider was recently also operated at the $\Upsilon(5S)$ resonance in an engineering run, allowing the Belle experiment to take first B_s data [231].
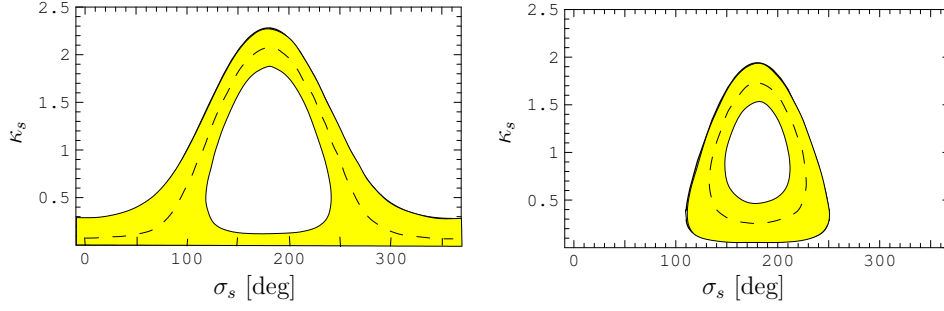
Fig. 31: The allowed regions (yellow/grey) in the σ_s–κ_s plane. Left panel: JLQCD lattice results (122). Right panel: (HP+JL)QCD lattice results (123).
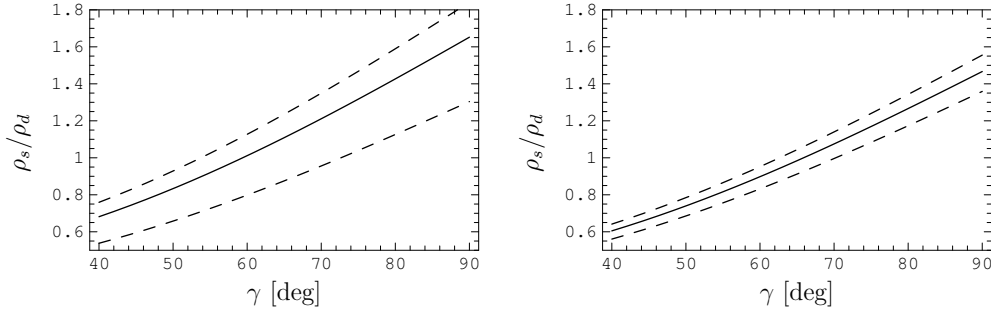
Fig. 32: The dependence of ρ_s/ρ_d on γ for the central values of $\Delta M_{d,s}$ in (130) and (132). Left panel: JLQCD results (308). Right panel: (HP+JL)QCD results (309). The plots are nearly independent of R_b.

where the hadronic $SU(3)$-breaking parameter ξ is defined through

$$\xi \equiv \frac{f_{B_s}\hat{B}_{B_s}^{1/2}}{f_{B_d}\hat{B}_{B_d}^{1/2}} \,. \tag{307}$$

In the class of NP models with 'minimal flavour violation' (see Section 6, and Ref. [237] for a recent analysis addressing also the ΔM_s measurement), we have $\rho_s/\rho_d = 1$, so that (306) allows the extraction of the CKM factor $|V_{ts}/V_{td}|$, and hence $|V_{td}|$, as $|V_{ts}|$ is known—to excellent accuracy—from (303). The advantage of this determination lies in the reduced theoretical uncertainty of ξ as compared to $f_{B_d}\hat{B}_{B_d}^{1/2}$. For the sets of lattice results in (122) and (123), we have

$$\xi_{\text{JLQCD}} = 1.14 \pm 0.06^{+0.13}_{-0} \tag{308}$$

$$\xi_{(\text{HP+JL})\text{QCD}} = 1.210^{+0.047}_{-0.035} \,. \tag{309}$$

Using the expression

$$R_t \equiv \frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right| = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{ts}}\right|\left[1 - \frac{1}{2}\left(1 - 2R_b\cos\gamma\right)\lambda^2 + \mathcal{O}(\lambda^4)\right], \tag{310}$$

we may convert the extracted value of $|V_{ts}/V_{td}|$ into a measurement of the UT side R_t. As we noted in Subsection 9.3, another determination of R_t can, in principle, be obtained from radiative decays, in particular the ratio of branching ratios $\mathcal{B}(B \to (\rho,\omega)\gamma)/\mathcal{B}(B \to K^*\gamma)$, but is presently limited by experimental statistics; see Ref. [242] for a recent analysis.

Alternatively, following Ref. [101], we may constrain the ratio ρ_s/ρ_d through the measured value of $\Delta M_s/\Delta M_d$. To this end, we express—in analogy to (192)—the UT side R_t in terms of R_b and γ:

$$R_t = \sqrt{1 - 2R_b \cos \gamma + R_b^2} \,, \tag{311}$$

allowing the determination of R_t through processes that are essentially unaffected by NP. The resulting value of R_t depends rather strongly on γ, which is the main source of uncertainty. Combining then (306) and (310), we obtain the following expression for ρ_s/ρ_d:

$$\frac{\rho_s}{\rho_d} = \lambda^2 \left[1 - 2R_b \cos \gamma + R_b^2 \right] \left[1 + (1 - 2R_b \cos \gamma)\lambda^2 + \mathcal{O}(\lambda^4) \right] \frac{1}{\xi^2} \frac{M_{B_d}}{M_{B_s}} \frac{\Delta M_s}{\Delta M_d} \,. \tag{312}$$

In Fig. 32, we plot this ratio for the central values of ΔM_d and ΔM_s in (130) and (132), respectively, as a function of the UT angle γ for the values of ξ given in (122) and (123). We find that the corresponding curves are nearly independent of R_b and that γ is actually the key CKM parameter for the determination of ρ_s/ρ_d. The corresponding numerical values are given by:

$$\left. \frac{\rho_s}{\rho_d} \right|_{\text{JLQCD}} = 1.11^{+0.02}_{-0.01}(\text{exp}) \pm 0.35(\gamma, R_b)^{+0.12}_{-0.28}(\xi) \tag{313}$$

$$\left. \frac{\rho_s}{\rho_d} \right|_{\text{(HP+JL)QCD}} = 0.99^{+0.02}_{-0.01}(\text{exp}) \pm 0.31(\gamma, R_b)^{+0.06}_{-0.08}(\xi) \,. \tag{314}$$

Because of the large range of allowed values of γ in (182), this ratio is currently not stringently constrained. This situation should, however, improve significantly in the LHC era thanks to the impressive determination of γ to be obtained at the LHCb experiment. In fact, a statistical accuracy of $\sigma_{\text{stat}}(\gamma) \approx 2.5°$ is expected at LHCb after five years of data taking [232].

Let us introduce a scenario for the year 2010 that is characterized by $\gamma = (70 \pm 5)°$ and the (HP+JL)QCD parameters in (123). We then find

$$\left. \frac{\rho_s}{\rho_d} \right|_{2010} = 1.07 \pm 0.09(\gamma, R_b)^{+0.06}_{-0.08}(\xi) = 1.07 \pm 0.12 \,, \tag{315}$$

where we made the errors arising from the uncertainties of γ and ξ explicit, and, in the last step, added them in quadrature. Consequently, the hadronic uncertainties and those induced by γ would now be of the same size, which should provide additional motivation for the lattice community to reduce the error of ξ even further. Despite the impressive reduction of uncertainty compared to the 2006 values in (313) and (314), the numerical value in (315) would still not allow a stringent test of whether ρ_s/ρ_d equals one: to establish a 3σ deviation from 1, central values of $\rho_s/\rho_d = 1.4$ or 0.7 would be needed. The assumed uncertainty of γ of $5°$ could also turn out to be too pessimistic, in which case even more progress would be needed from the lattice side to match the experimental accuracy.

The result in (315) would not necessarily suggest that there is no physics beyond the SM. In fact, the central values of $\rho_d = 0.69 \pm 0.16$ and $\rho_s = 0.74 \pm 0.18$ would both be smaller than 1, i.e., would both deviate from the SM picture, although the hadronic uncertainties would again not allow us to draw definite conclusions. In order to shed further light on these possible NP contributions, the exploration of CP-violating effects in the B_s-meson system is essential, which can be performed with the help of the 'golden' decay $B_s^0 \to J/\psi\phi$.

10.2 $B_s^0 \to J/\psi\phi$

As can be seen in Fig. 20, the decay $B_s^0 \to J/\psi\phi$ is simply related to $B_d^0 \to J/\psi K_S$ through a replacement of the down spectator quark by a strange quark. Consequently, the structure of the $B_s^0 \to J/\psi\phi$ decay amplitude is completely analogous to that of (169). On the other hand, the final state of $B_s^0 \to J/\psi\phi$
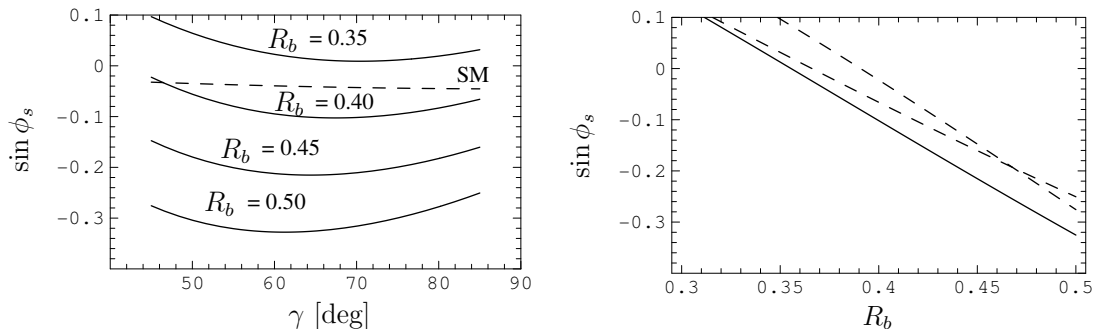
Fig. 33: $\sin \phi_s$ for a scenario with flavour-universal NP, i.e., $\phi_s^{\mathrm{NP}} = \phi_d^{\mathrm{NP}}$, as specified in Eq. (318), and $\phi_d = 43.4°$. Left panel: $\sin \phi_s$ as a function of γ for various values of R_b. Right panel: $\sin \phi_s$ as a function of R_b for various values of γ [solid line: $\gamma = 65°$, dashed lines: $\gamma = (45°, 85°)$].

consists of two vector mesons, and is hence an admixture of different CP eigenstates, which can, however, be disentangled through an angular analysis of the $B_s^0 \rightarrow J/\psi[\rightarrow \ell^+\ell^-]\phi[\rightarrow K^+K^-]$ decay products [111, 243]. The corresponding angular distribution exhibits tiny direct CP violation, and allows the extraction of

$$\sin \phi_s + \mathcal{O}(\overline{\lambda}^3) = \sin \phi_s + \mathcal{O}(10^{-3}) \tag{316}$$

through mixing-induced CP violation. Since we have $\phi_s = -2\delta\gamma = -2\lambda^2\eta \sim -2°$ in the SM, the determination of this phase from (316) is affected by hadronic uncertainties of $\mathcal{O}(10\%)$, which may become an issue for the LHC era. These uncertainties can be controlled with the help of flavour-symmetry arguments through the $B_d^0 \rightarrow J/\psi\rho^0$ decay [244].

Needless to say, the big hope is that large CP violation will be found in this channel. Since the CP-violating effects in $B_s^0 \rightarrow J/\psi\phi$ are tiny in the SM, such an observation would give us an unambiguous signal for NP [117, 245, 246]. As the situation for NP entering through the decay amplitude is similar to $B \rightarrow J/\psi K$, we would get evidence for CP-violating NP contributions to B_s^0–$\bar{B}_s^0$ mixing, and could extract the corresponding sizeable value of ϕ_s [117]. Such a scenario may generically arise in the presence of NP with $\Lambda_{\mathrm{NP}} \sim$ TeV [119], as well as in specific models, including supersymmetric frameworks and models with extra Z' bosons (see Ref. [101] and references therein).

Thanks to its nice experimental signature, $B_s^0 \rightarrow J/\psi\phi$ is very accessible at hadron colliders, and can be fully exploited at the LHC. After one year of data taking (which corresponds to 2 fb^{-1}), LHCb expects a measurement with the statistical accuracy $\sigma_{\mathrm{stat}}(\sin \phi_s) \approx 0.031$; adding modes such as $B_s \rightarrow J/\psi\eta, J/\psi\eta'$ and $\eta_c\phi$, $\sigma_{\mathrm{stat}}(\sin \phi_s) \approx 0.013$ is expected after five years [232]. Also ATLAS and CMS will contribute to the measurement of $\sin \phi_s$, expecting uncertainties at the 0.1 level after one year of data taking, which corresponds to 10 fb^{-1} [247, 248]. In order to illustrate the impact of NP effects on the quantity

$$\sin \phi_s = \sin(-2\lambda^2 R_b \sin \gamma + \phi_s^{\mathrm{NP}}) \,, \tag{317}$$

let us assume that the NP parameters satisfy the simple relation

$$\sigma_d = \sigma_s, \quad \kappa_d = \kappa_s \,, \tag{318}$$

i.e., that in particular $\phi_d^{\mathrm{NP}} = \phi_s^{\mathrm{NP}}$. This scenario would be supported by (315), although it would *not* belong to the class of models with MFV, as new sources of CP violation would be required. As we have seen in Section 7.1, the analysis of the B_d^0 data for $R_b^{\mathrm{incl}} = 0.45$ indicates a small NP phase around
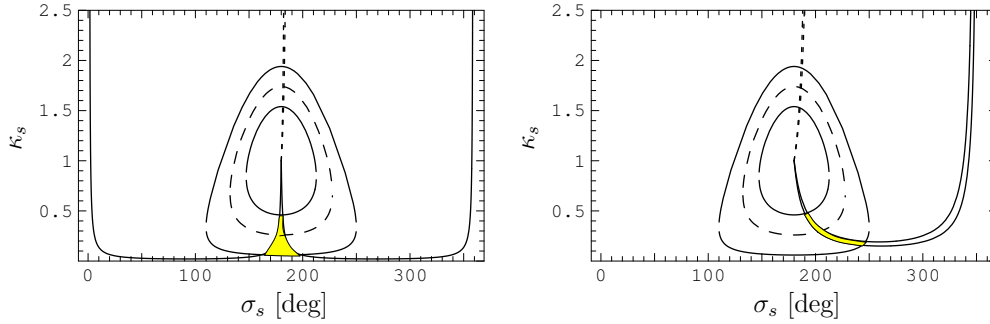
Fig. 34: Combined constraints for the allowed region (yellow/grey) in the σ_s–κ_s plane through ΔM_s in (132) for the (HP+JL)QCD results (123) and CP violation measurements. Left panel: the SM scenario $(\sin\phi_s)_{\text{exp}} = -0.04 \pm 0.02$. Right panel: a NP scenario with $(\sin\phi_s)_{\text{exp}} = -0.20 \pm 0.02$. The solid lines correspond to $\cos\phi_s > 0$, the dotted lines to $\cos\phi_s < 0$.

$-10°$ in the B_d system. In the above scenario, that would imply the presence of the same phase in the B_s system, which would interfere constructively with the small SM phase and result in CP asymmetries at the level of -20%. CP-violating effects of that size can easily be detected at the LHC. This exercise demonstrates again the great power of the B_s-meson system to reveal CP-violating NP contributions to B_q^0–$\bar{B}_q^0$ mixing. The presence of a small NP phase could actually be considerably magnified, as illustrated in Fig. 33.

Let us finally also discuss the impact of CP violation measurements on the allowed region in the σ_s–κ_s plane in our 2010 scenario. To this end, we consider two cases:

  i) $(\sin\phi_s)_{\text{exp}} = -0.04 \pm 0.02$, in accordance with the SM;
  ii) $(\sin\phi_s)_{\text{exp}} = -0.20 \pm 0.02$, in accordance with the NP scenario of Fig. 33.

The measurement of $\sin\phi_s$ implies a twofold solution for ϕ_s and, therefore, also for ϕ_s^{NP}. However, this ambiguity can be resolved through the determination of the sign of $\cos\phi_s$, which can be fixed through the strategies proposed in Ref. [117]. In Fig. 34, we show the situation in the σ_s–κ_s plane[4]. The dotted lines refer to negative values of $\cos\phi_s$. Assuming that these are experimentally excluded, we are left with strongly restricted regions, although κ_s could still take sizeable ranges, with upper bounds $\kappa_s \approx 0.5$. In the SM-like scenario, values of σ_s around $180°$ would arise, i.e., a NP contribution with a sign opposite to the SM. However, due to the absence of new CP-violating effects, the accuracy of lattice results would have to be considerably improved in order to allow the extraction of a value of κ_s incompatible with 0. On the other hand, a measurement of $(\sin\phi_s)_{\text{exp}} = -0.20 \pm 0.02$ would give a NP signal at the $10\,\sigma$ level, with $\kappa_s \gtrsim 0.2$. A determination of κ_s with 10% uncertainty requires $f_{B_s}\hat{B}_{B_s}^{1/2}$ with 5% accuracy, i.e., the corresponding error in (123) has to be reduced by a factor of 2.

Since our discussion does not refer to a specific model of NP, the question arises whether there are actually extensions of the SM that still allow large CP-violating NP phases in B_s^0–$\bar{B}_s^0$ mixing. This is in fact the case, also after the measurement of ΔM_s. In Ref. [101], where also a comprehensive guide to the relevant literature can be found, this exciting feature was illustrated by considering models with an extra Z' boson and SUSY scenarios with an approximate alignment of quark and squark masses.

Let us now continue our discussion of the B_s-meson system by having a closer look at other benchmark processes.

---

[4]The closed lines agree with those shown in the right panel of Fig. 31, as our 2010 scenario is based on the (HP+JL)QCD lattice results.
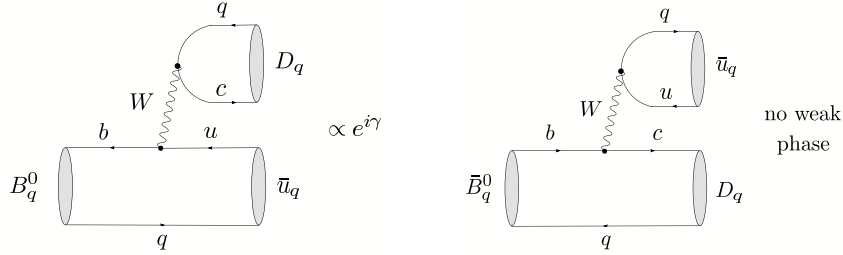
Fig. 35: Feynman diagrams contributing to $B_q^0 \to D_q \bar{u}_q$ and $\bar{B}_q^0 \to D_q \bar{u}_q$ decays
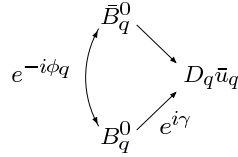
Fig. 36: Interference effects between $B_q^0 \to D_q \bar{u}_q$ and $\bar{B}_q^0 \to D_q \bar{u}_q$ decays

10.3 $\;\; B_s \to D_s^{\pm} K^{\mp}$ and $B_d \to D^{\pm} \pi^{\mp}$

The decays $B_s \to D_s^{\pm} K^{\mp}$ [249] and $B_d \to D^{\pm} \pi^{\mp}$ [250] can be treated on the same theoretical basis, and provide new strategies to determine γ [90]. Following this paper, we write these modes, which are pure 'tree' decays according to the classification of Section 3.3.1, generically as $B_q \to D_q \bar{u}_q$. As can be seen from the Feynman diagrams in Fig. 35, their characteristic feature is that both a B_q^0 and a $\bar{B}_q^0$ meson may decay into the same final state $D_q \bar{u}_q$. Consequently, as illustrated in Fig. 36, interference effects between B_q^0–$\bar{B}_q^0$ mixing and decay processes arise, which allow us to probe the weak phase $\phi_q + \gamma$ through measurements of the corresponding time-dependent decay rates.

In the case of $q = s$, i.e., $D_s \in \{D_s^+, D_s^{*+}, ...\}$ and $u_s \in \{K^+, K^{*+}, ...\}$, these interference effects are governed by a hadronic parameter $X_s e^{i\delta_s} \propto R_b \approx 0.4$, where $R_b \propto |V_{ub}/V_{cb}|$ is the usual UT side, and hence are large. On the other hand, for $q = d$, i.e., $D_d \in \{D^+, D^{*+}, ...\}$ and $u_d \in \{\pi^+, \rho^+, ...\}$, the interference effects are described by $X_d e^{i\delta_d} \propto -\lambda^2 R_b \approx -0.02$, and hence are tiny. In the following, we shall only consider $B_q \to D_q \bar{u}_q$ modes, where at least one of the D_q, $\bar{u}_q$ states is a pseudoscalar meson; otherwise a complicated angular analysis has to be performed.

The time-dependent rate asymmetries of these decays take the same form as (152). It is well known that they allow a *theoretically clean* determination of $\phi_q + \gamma$, where the 'conventional' approach works as follows [249, 250]: if we measure the observables $C(B_q \to D_q \bar{u}_q) \equiv C_q$ and $C(B_q \to \bar{D}_q u_q) \equiv \overline{C}_q$ provided by the $\cos(\Delta M_q t)$ pieces, we may determine the following quantities:

$$\langle C_q \rangle_+ \equiv \frac{1}{2} \left[\overline{C}_q + C_q \right] = 0, \quad \langle C_q \rangle_- \equiv \frac{1}{2} \left[\overline{C}_q - C_q \right] = \frac{1 - X_q^2}{1 + X_q^2} , \tag{319}$$

where $\langle C_q \rangle_-$ allows us to extract X_q. However, to this end we have to resolve terms entering at the X_q^2 level. In the case of $q = s$, we have $X_s = \mathcal{O}(R_b)$, implying $X_s^2 = \mathcal{O}(0.16)$, so that this should actually be possible, though challenging. On the other hand, $X_d = \mathcal{O}(-\lambda^2 R_b)$ is doubly Cabibbo-suppressed. Although it should be possible to resolve terms of $\mathcal{O}(X_d)$, this will be impossible for the vanishingly small $X_d^2 = \mathcal{O}(0.0004)$ terms, so that other approaches to fix X_d are required [250]. For the extraction of $\phi_q + \gamma$, the mixing-induced observables $S(B_q \to D_q \bar{u}_q) \equiv S_q$ and $S(B_q \to \bar{D}_q u_q) \equiv \overline{S}_q$ associated with the $\sin(\Delta M_q t)$ terms of the time-dependent rate asymmetry must be measured. In analogy to (319), it is convenient to introduce observable combinations $\langle S_q \rangle_\pm$. Assuming that X_q is known, we may consider the quantities

$$s_+ \equiv (-1)^L \left[\frac{1 + X_q^2}{2X_q} \right] \langle S_q \rangle_+ = + \cos \delta_q \sin(\phi_q + \gamma) \tag{320}$$

$$s_- \equiv (-1)^L \left[\frac{1 + X_q^2}{2X_q} \right] \langle S_q \rangle_- = - \sin \delta_q \cos(\phi_q + \gamma) , \tag{321}$$

which yield

$$\sin^2(\phi_q + \gamma) = \frac{1}{2} \left[(1 + s_+^2 - s_-^2) \pm \sqrt{(1 + s_+^2 - s_-^2)^2 - 4s_+^2} \right] , \tag{322}$$

implying an eightfold solution for $\phi_q + \gamma$. If we fix the sign of $\cos \delta_q$ through factorization, still a fourfold discrete ambiguity is left, which is limiting the power for the search of NP significantly. Note that this assumption allows us also to fix the sign of $\sin(\phi_q + \gamma)$ through $\langle S_q \rangle_+$. To this end, the factor $(-1)^L$, where L is the $D_q \bar{u}_q$ angular momentum, has to be properly taken into account. This is a crucial issue for the extraction of the sign of $\sin(\phi_d + \gamma)$ from $B_d \to D^{*\pm} \pi^\mp$ decays.

Let us now discuss new strategies to explore CP violation through $B_q \to D_q \bar{u}_q$ modes, following Ref. [90]. If $\Delta\Gamma_s$ is sizeable, the 'untagged' rates introduced in (149) allow us to measure $\mathcal{A}_{\Delta\Gamma}(B_s \to D_s \bar{u}_s) \equiv \mathcal{A}_{\Delta\Gamma_s}$ and $\mathcal{A}_{\Delta\Gamma}(B_s \to \bar{D}_s u_s) \equiv \overline{\mathcal{A}}_{\Delta\Gamma_s}$. Introducing, in analogy to (319), observable combinations $\langle \mathcal{A}_{\Delta\Gamma_s} \rangle_\pm$, we may derive the relations

$$\tan(\phi_s + \gamma) = - \left[\frac{\langle S_s \rangle_+}{\langle \mathcal{A}_{\Delta\Gamma_s} \rangle_+} \right] = + \left[\frac{\langle \mathcal{A}_{\Delta\Gamma_s} \rangle_-}{\langle S_s \rangle_-} \right] , \tag{323}$$

which allow an *unambiguous* extraction of $\phi_s + \gamma$ if we fix the sign of $\cos \delta_q$ through factorization. Another important advantage of (323) is that we do *not* have to rely on $\mathcal{O}(X_s^2)$ terms, as $\langle S_s \rangle_\pm$ and $\langle \mathcal{A}_{\Delta\Gamma_s} \rangle_\pm$ are proportional to X_s. On the other hand, a sizeable value of $\Delta\Gamma_s$ is of course needed.

If we keep the hadronic quantities X_q and δ_q as 'unknown', free parameters in the expressions for the $\langle S_q \rangle_\pm$, we may obtain bounds on $\phi_q + \gamma$ from

$$|\sin(\phi_q + \gamma)| \geq |\langle S_q \rangle_+| , \quad |\cos(\phi_q + \gamma)| \geq |\langle S_q \rangle_-| . \tag{324}$$

If X_q is known, stronger constraints are implied by

$$|\sin(\phi_q + \gamma)| \geq |s_+| , \quad |\cos(\phi_q + \gamma)| \geq |s_-| . \tag{325}$$

Once s_+ and s_- are known, we may of course determine $\phi_q + \gamma$ through the 'conventional' approach, using (322). However, the bounds following from (325) provide essentially the same information and are much simpler to implement. Moreover, as discussed in detail in Ref. [90] for several examples within the SM, the bounds following from the B_s and B_d modes may be highly complementary, thereby providing particularly narrow, theoretically clean ranges for γ.

Let us now further exploit the complementarity between the $B_s^0 \to D_s^{(*)+} K^-$ and $B_d^0 \to D^{(*)+} \pi^-$ processes. Looking at the corresponding decay topologies, we see that these channels are related to each other through an interchange of all down and strange quarks. Consequently, applying again the U-spin symmetry implies $a_s = a_d$ and $\delta_s = \delta_d$, where $a_s \equiv X_s/R_b$ and $a_d \equiv -X_d/(\lambda^2 R_b)$ are the ratios of the hadronic matrix elements entering X_s and X_d, respectively. There are various possibilities to implement these relations [90]. A particularly simple picture arises if we assume that $a_s = a_d$ *and* $\delta_s = \delta_d$, which yields

$$\tan \gamma = - \left[\frac{\sin \phi_d - S \sin \phi_s}{\cos \phi_d - S \cos \phi_s} \right] \overset{\phi_s = 0°}{=} - \left[\frac{\sin \phi_d}{\cos \phi_d - S} \right] . \tag{326}$$

Here we have introduced

$$S \equiv -R \left[\frac{\langle S_d \rangle_+}{\langle S_s \rangle_+} \right] \tag{327}$$

with

$$R \equiv \left(\frac{1 - \lambda^2}{\lambda^2} \right) \left[\frac{1}{1 + X_s^2} \right] , \tag{328}$$

where R can be fixed with the help of untagged B_s rates through

$$R = \left(\frac{f_K}{f_\pi} \right)^2 \left[\frac{\Gamma(\bar{B}_s^0 \to D_s^{(*)+}\pi^-) + \Gamma(B_s^0 \to D_s^{(*)-}\pi^+)}{\langle \Gamma(B_s \to D_s^{(*)+}K^-) \rangle + \langle \Gamma(B_s \to D_s^{(*)-}K^+) \rangle} \right] . \tag{329}$$

Alternatively, we can *only* assume that $\delta_s = \delta_d$ *or* that $a_s = a_d$ [90]. An important feature of this strategy is that it allows us to extract an *unambiguous* value of γ, which is crucial for the search of NP; first studies for LHCb are very promising in this respect [251]. Another advantage with respect to the 'conventional' approach is that X_q^2 terms have not to be resolved experimentally. In particular, X_d does *not* have to be fixed, and X_s may only enter through a $1 + X_s^2$ correction, which can straightforwardly be determined through untagged B_s rate measurements. In the most refined implementation of this strategy, the measurement of X_d/X_s would only be interesting for the inclusion of U-spin-breaking corrections in a_d/a_s. Moreover, we may obtain interesting insights into hadron dynamics and U-spin breaking.

The colour-suppressed counterparts of the $B_q \to D_q \bar{u}_q$ modes are also interesting for the exploration of CP violation. In the case of the $B_d \to D K_{S(L)}$, $B_s \to D\eta^{(')}, D\phi, \dots$ modes, the interference effects between B_q^0–$\bar{B}_q^0$ mixing and decay processes are governed by $x_{f_s} e^{i\delta_{f_s}} \propto R_b$. If we consider the CP eigenstates $D_\pm$ of the neutral D-meson system, we obtain additional interference effects at the amplitude level, which involve γ, and may introduce the following 'untagged' rate asymmetry [153]:

$$\Gamma_{+-}^{f_s} \equiv \frac{\langle \Gamma(B_q \to D_+ f_s) \rangle - \langle \Gamma(B_q \to D_- f_s) \rangle}{\langle \Gamma(B_q \to D_+ f_s) \rangle + \langle \Gamma(B_q \to D_- f_s) \rangle} , \tag{330}$$

which allows us to constrain γ through the relation

$$|\cos\gamma| \geq |\Gamma_{+-}^{f_s}| . \tag{331}$$

Moreover, if we complement $\Gamma_{+-}^{f_s}$ with

$$\langle S_{f_s} \rangle_\pm \equiv \frac{1}{2} \left[S_+^{f_s} \pm S_-^{f_s} \right] , \tag{332}$$

where $S_\pm^{f_s} \equiv \mathcal{A}_{\mathrm{CP}}^{\mathrm{mix}}(B_q \to D_\pm f_s)$, we may derive the following simple but *exact* relation:

$$\tan\gamma \cos\phi_q = \left[\frac{\eta_{f_s} \langle S_{f_s} \rangle_+}{\Gamma_{+-}^{f_s}} \right] + [\eta_{f_s} \langle S_{f_s} \rangle_- - \sin\phi_q] , \tag{333}$$

with $\eta_{f_s} \equiv (-1)^L \eta_{\mathrm{CP}}^{f_s}$. This expression allows a conceptually simple, theoretically clean and essentially unambiguous determination of γ [153]. Since the interference effects are governed by the tiny parameter $x_{f_d} e^{i\delta_{f_d}} \propto -\lambda^2 R_b$ in the case of $B_s \to D_\pm K_{S(L)}$, $B_d \to D_\pm \pi^0, D_\pm \rho^0, \dots$, these modes are not as interesting for the extraction of γ. However, they provide the relation

$$\eta_{f_d} \langle S_{f_d} \rangle_- = \sin\phi_q + \mathcal{O}(x_{f_d}^2) = \sin\phi_q + \mathcal{O}(4 \times 10^{-4}) , \tag{334}$$

allowing very interesting determinations of ϕ_q with theoretical accuracies one order of magnitude higher than those of the conventional $B_d^0 \to J/\psi K_S$ and $B_s^0 \to J/\psi\phi$ approaches [153]. As we pointed out in Section 7.1, these measurements would be very interesting in view of the new world average of $(\sin 2\beta)_{\psi K_S}$.

10.4 $B_s^0 \to K^+K^-$ and $B_d^0 \to \pi^+\pi^-$

The decay $B_s^0 \to K^+K^-$ is a $\bar{b} \to \bar{s}$ transition, and involves tree and penguin amplitudes, as the $B_d^0 \to \pi^+\pi^-$ mode [167]. However, because of the different CKM structure, the latter topologies actually play the dominant role in the $B_s^0 \to K^+K^-$ channel. In analogy to (208), we may write

$$A(B_s^0 \to K^+K^-) = \sqrt{\epsilon}\, \mathcal{C}' \left[e^{i\gamma} + \frac{1}{\epsilon} d' e^{i\theta'} \right] , \tag{335}$$

where ϵ was introduced in (221), and the CP-conserving hadronic parameters $\mathcal{C}'$ and $d'e^{i\theta'}$ correspond to $\mathcal{C}$ and $de^{i\theta}$, respectively. The corresponding observables then take the following generic form:

$$\mathcal{A}_{\text{CP}}^{\text{dir}}(B_s \to K^+K^-) = G_1'(d', \theta'; \gamma) \tag{336}$$
$$\mathcal{A}_{\text{CP}}^{\text{mix}}(B_s \to K^+K^-) = G_2'(d', \theta'; \gamma, \phi_s) , \tag{337}$$

in analogy to the expressions for the CP-violating $B_d^0 \to \pi^+\pi^-$ asymmetries in (213) and (214). Since $\phi_d = (43.4 \pm 2.5)°$ is already known (see Section 7.1) and ϕ_s is negligibly small in the SM—or can be determined through $B_s^0 \to J/\psi\phi$ should CP-violating NP contributions to B_s^0–$\bar{B}_s^0$ mixing make it sizeable—we may convert the measured values of $\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \pi^+\pi^-)$, $\mathcal{A}_{\text{CP}}^{\text{mix}}(B_d \to \pi^+\pi^-)$ and $\mathcal{A}_{\text{CP}}^{\text{dir}}(B_s \to K^+K^-)$, $\mathcal{A}_{\text{CP}}^{\text{mix}}(B_s \to K^+K^-)$ into *theoretically clean* contours in the γ–d and γ–d' planes, respectively. In Fig. 37, we show these contours for an example, which corresponds to the central values of (217) and (218) with the hadronic parameters (d, θ) in (230).

As can be seen in Fig. 26, the decay $B_d^0 \to \pi^+\pi^-$ is actually related to $B_s^0 \to K^+K^-$ through the interchange of *all* down and strange quarks. Consequently, each decay topology contributing to $B_d^0 \to \pi^+\pi^-$ has a counterpart in $B_s^0 \to K^+K^-$, and the corresponding hadronic parameters can be related to each other with the help of the U-spin flavour symmetry of strong interactions, implying the following relations [167]:

$$d' = d , \quad \theta' = \theta . \tag{338}$$

Applying the former, we may extract γ and d through the intersections of the theoretically clean γ–d and γ–d' contours. As discussed in Ref. [167], it is also possible to resolve straightforwardly the twofold ambiguity for (γ, d) arising in Fig. 37, thereby leaving us with the 'true' solution of $\gamma = 74°$ in this example. Moreover, we may determine θ and θ', which allow an interesting internal consistency check of the second U-spin relation in (338). An alternative avenue is provided if we eliminate d and d' through the CP-violating $B_d \to \pi^+\pi^-$ and $B_s \to K^+K^-$ observables, respectively, and then extract these parameters and γ through the U-spin relation $\theta' = \theta$.

As illustrated in Fig. 38, this strategy is very promising from an experimental point of view for the LHCb experiment, where an accuracy for γ of a few degrees can be achieved [147, 232, 252]. As far as possible U-spin-breaking corrections to $d' = d$ are concerned, they enter the determination of γ through a relative shift of the γ–d and γ–d' contours; their impact on the extracted value of γ therefore depends on the form of these curves, which is fixed through the measured observables. In the examples discussed in Refs. [119, 167], as well as in the one shown in Fig. 37, the extracted value of γ would be very stable under such effects. Let us also note that the U-spin relations in (338) are particularly robust since they involve only ratios of hadronic amplitudes, where all $SU(3)$-breaking decay constants and form factors cancel in factorization and also chirally enhanced terms would not lead to U-spin-breaking corrections [167]. On the other hand, the ratio $|\mathcal{C}'/\mathcal{C}|$, which equals 1 in the strict U-spin limit and enters the U-spin relation

$$\frac{\mathcal{A}_{\text{CP}}^{\text{mix}}(B_s \to K^+K^-)}{\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \pi^+\pi^-)} = - \left| \frac{\mathcal{C}'}{\mathcal{C}} \right|^2 \left[\frac{\text{BR}(B_d \to \pi^+\pi^-)}{\text{BR}(B_s \to K^+K^-)} \right] \frac{\tau_{B_s}}{\tau_{B_d}} , \tag{339}$$

is affected by U-spin-breaking effects within factorization. An estimate of the corresponding form factors was recently performed in Ref. [253] with the help of QCD sum rules, which is an important ingredient
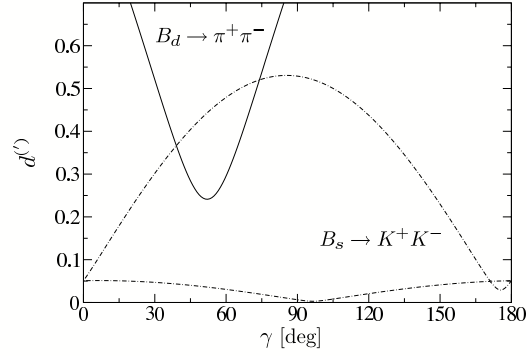
Fig. 37: The contours in the γ–$d^{(')}$ plane for an example with $d = d' = 0.52$, $\theta = \theta' = 146°$, $\phi_d = 43.4°$, $\phi_s = -2°$, $\gamma = 74°$, which corresponds to the CP asymmetries $\mathcal{A}_{\text{CP}}^{\text{dir}}(B_d \to \pi^+\pi^-) = -0.37$ and $\mathcal{A}_{\text{CP}}^{\text{mix}}(B_d \to \pi^+\pi^-) = +0.50$ (see Sections 7.3 and 8.2), as well as $\mathcal{A}_{\text{CP}}^{\text{dir}}(B_s \to K^+K^-) = +0.12$ and $\mathcal{A}_{\text{CP}}^{\text{mix}}(B_s \to K^+K^-) = -0.19$.
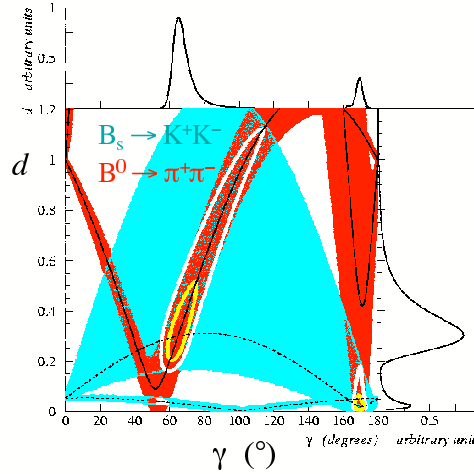
Fig. 38: Experimental LHCb feasibility study for the contours in the γ–$d^{(')}$ plane, as discussed in Ref. [252]

for a SM prediction of the CP-averaged $B_s \to K^+K^-$ branching ratio [83]. Following these lines, the prediction

$$\text{BR}(B_s \to K^+K^-) = (35 \pm 7) \times 10^{-6} \tag{340}$$

was obtained in Refs. [83, 200] from the CP-averaged $B_d \to \pi^\mp K^\pm$ branching ratio. On the other hand, the CDF Collaboration announced recently the observation of the $B_s \to K^+K^-$ channel, with the following branching ratio [254]:

$$\text{BR}(B_s \to K^+K^-) = (33 \pm 5.7 \pm 6.7) \times 10^{-6} \,, \tag{341}$$

which is in excellent accordance with (340). For other recent analyses of the $B_s \to K^+K^-$ decay, see Refs. [255, 256].

In addition to the $B_s \to K^+K^-$, $B_d \to \pi^+\pi^-$ and $B_s \to D_s^\pm K^\mp$, $B_d \to D^\pm \pi^\mp$ strategies discussed above, other U-spin methods for the extraction of γ were also proposed, using $B_{s(d)} \to J/\psi K_S$ or $B_{d(s)} \to D_{d(s)}^+ D_{d(s)}^-$ [142], $B_{d(s)} \to K^{0(*)}\bar{K}^{0(*)}$ [119, 244], $B_{(s)} \to \pi K$ [257], or $B_{s(d)} \to J/\psi\eta$ modes [258]. In a very recent paper [259], two-body decays of charged B mesons were also considered.
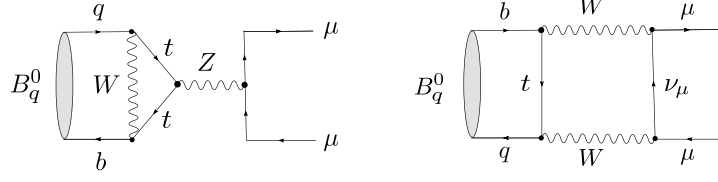
Fig. 39: Feynman diagrams contributing to $B_q^0 \to \mu^+\mu^-$ ($q \in \{s, d\}$)

10.5 $B_s^0 \to \mu^+\mu^-$ and $B_d^0 \to \mu^+\mu^-$

Let us finally have a closer look at the rare decay $B_s^0 \to \mu^+\mu^-$, which we already encountered briefly in Section 8.4. As can be seen in Fig. 39, this decay and its B_d-meson counterpart $B_d^0 \to \mu^+\mu^-$ originate from Z^0-penguin and box diagrams in the SM. The corresponding low-energy effective Hamiltonian is given as follows [67]:

$$\mathcal{H}_{\text{eff}} = -\frac{G_{\text{F}}}{\sqrt{2}}\left[\frac{\alpha}{2\pi\sin^2\Theta_{\text{W}}}\right] V_{tb}^* V_{tq}\, \eta_Y Y_0(x_t)(\bar{b}q)_{\text{V}-\text{A}}(\bar{\mu}\mu)_{\text{V}-\text{A}} + \text{h.c.}\,, \tag{342}$$

where α denotes the QED coupling and Θ_{W} is the Weinberg angle. The short-distance physics is described by $Y(x_t) \equiv \eta_Y Y_0(x_t)$, where $\eta_Y = 1.012$ is a perturbative QCD correction [260]– [262], and the Inami–Lim function $Y_0(x_t)$ describes the top-quark mass dependence. We observe that only the matrix element $\langle 0|(\bar{b}q)_{\text{V}-\text{A}}|B_q^0\rangle$ is required. Since here the vector-current piece vanishes, as the B_q^0 is a pseudoscalar meson, this matrix element is simply given by the decay constant f_{B_q}. Consequently, we arrive at a very favourable situation with respect to the hadronic matrix elements. Since, moreover, NLO QCD corrections were calculated, and long-distance contributions are expected to play a negligible role [260], the $B_q^0 \to \mu^+\mu^-$ modes belong to the cleanest rare B decays. The SM branching ratios can then be written in the following compact form [37]:

$$\text{BR}(B_s \to \mu^+\mu^-) = 4.1 \times 10^{-9}$$
$$\times \left[\frac{f_{B_s}}{0.24\,\text{GeV}}\right]^2 \left[\frac{|V_{ts}|}{0.040}\right]^2 \left[\frac{\tau_{B_s}}{1.5\,\text{ps}}\right] \left[\frac{m_t}{167\,\text{GeV}}\right]^{3.12} \tag{343}$$

$$\text{BR}(B_d \to \mu^+\mu^-) = 1.1 \times 10^{-10}$$
$$\times \left[\frac{f_{B_d}}{0.20\,\text{GeV}}\right]^2 \left[\frac{|V_{td}|}{0.008}\right]^2 \left[\frac{\tau_{B_d}}{1.5\,\text{ps}}\right] \left[\frac{m_t}{167\,\text{GeV}}\right]^{3.12}. \tag{344}$$

The most recent upper bounds (95% C.L.) from the CDF Collaboration read as follows [263]:

$$\text{BR}(B_s \to \mu^+\mu^-) < 1.0 \times 10^{-7}, \quad \text{BR}(B_d \to \mu^+\mu^-) < 3.0 \times 10^{-8}\,, \tag{345}$$

while the D0 Collaboration finds the following (95% C.L.) upper limit [264]:

$$\text{BR}(B_s \to \mu^+\mu^-) < 3.7 \times 10^{-7}\,. \tag{346}$$

Using again relation (310) and neglecting the tiny corrections entering at the λ^2 level, we find that the measurement of the ratio

$$\frac{\text{BR}(B_d \to \mu^+\mu^-)}{\text{BR}(B_s \to \mu^+\mu^-)} = \left[\frac{\tau_{B_d}}{\tau_{B_s}}\right]\left[\frac{M_{B_d}}{M_{B_s}}\right]\left[\frac{f_{B_d}}{f_{B_s}}\right]^2 \left|\frac{V_{td}}{V_{ts}}\right|^2 \tag{347}$$

would allow an extraction of the UT side R_t. Since the short-distance function Y cancels, this determination works not only in the SM, but also in the NP scenarios with MFV [137]. This strategy is complementary to that offered by the ratio $\Delta M_s/\Delta M_d$ discussed in the context of (306). If we look

at this expression in the MFV case, where $\rho_s/\rho_d = 1$, and (347), we see that the following relation is implied [265]:

$$\frac{\text{BR}(B_s \to \mu^+\mu^-)}{\text{BR}(B_d \to \mu^+\mu^-)} = \left[\frac{\tau_{B_s}}{\tau_{B_d}}\right]\left[\frac{\hat{B}_{B_d}}{\hat{B}_{B_s}}\right]\left[\frac{\Delta M_s}{\Delta M_d}\right], \tag{348}$$

which holds again in the context of MFV models, including the SM. Here the advantage is that the dependence on $(f_{B_d}/f_{B_s})^2$ cancels. Moreover, we may also use the data for the mass differences ΔM_q to reduce the hadronic uncertainties of the SM predictions of the $B_q \to \mu^+\mu^-$ branching ratios [265]:

$$\text{BR}(B_s \to \mu^+\mu^-) = (3.35 \pm 0.32) \times \times 10^{-9} \tag{349}$$

$$\text{BR}(B_d \to \mu^+\mu^-) = (1.03 \pm 0.09) \times 10^{-10}, \tag{350}$$

where (349) is another application of the recent ΔM_s measurement at the Tevatron [237].

The current experimental upper bounds in (345) and (346) are still about two orders of magnitude away from these numbers. Consequently, should the $B_q \to \mu^+\mu^-$ decays be governed by their SM contributions, we could only hope to observe them at the LHC [147]. On the other hand, since the $B_q \to \mu^+\mu^-$ transitions originate from FCNC processes, they are sensitive probes of NP. In particular, the branching ratios may be dramatically enhanced in specific NP (SUSY) scenarios, as was recently reviewed in Ref. [118]. Should this actually be the case, these decays may already be seen at Run II of the Tevatron, and the e^+e^- B factories could observe $B_d \to \mu^+\mu^-$. Let us finally emphasize that the experimental bounds on $B_s \to \mu^+\mu^-$ can also be converted into bounds on NP parameters in specific scenarios. In the context of the constrained minimal supersymmetric extension of the SM (CMSSM) with universal scalar masses, such constraints were recently critically discussed by the authors of Ref. [266].

11 Conclusions and outlook

CP violation is now well established in the B-meson system, thereby complementing the neutral K-meson system, where this phenomenon was discovered more than 40 years ago. The data of the e^+e^- B factories have provided valuable insights into the physics of strong and weak interactions. Concerning the former aspect, which is sometimes only considered as a by-product, the data give us important evidence for large non-factorizable effects in non-leptonic B-decays, so that the challenge for a reliable theoretical description within dynamical QCD approaches remains, despite interesting recent progress. As far as the latter aspect is concerned, the description of CP violation through the KM mechanism has successfully passed its first experimental tests, in particular through the comparison between the measurement of $\sin 2\beta$ with the help of $B_d^0 \to J/\psi K_S$ and the CKM fits. However, the most recent average for $(\sin 2\beta)_{\psi K_S}$ is now somewhat on the lower side, and there are a couple of puzzles in the B-factory data. It will be very interesting to monitor these effects, which could be first hints for physics beyond the SM, as the data improve. Moreover, it is crucial to refine the corresponding theoretical analyses further, to have a critical look at the underlying working assumptions and to check them through independent tests, and to explore correlations with other flavour probes.

Despite this impressive progress, there are still regions of the B-physics landscape left that are essentially unexplored. For instance, $b \to d$ penguin processes are now entering the stage, since lower bounds for the corresponding branching ratios that can be derived in the SM turn out to be very close to the corresponding experimental upper limits. Indeed, we have now evidence for the $B_d \to K^0\bar{K}^0$ and $B^\pm \to K^\pm K$ channels, and the first signals for the radiative $B \to \rho\gamma$ transitions were reported, representing one of the hot topics of the summer of 2005. These modes have now to be explored in much more detail, and several other decays are waiting to be observed.

Another very interesting aspect of future studies is the B_s-meson system. Although the mass difference ΔM_s was measured in the spring of 2006 at the Tevatron, many features of B_s physics are still essentially unexplored. Concerning the measurement of ΔM_s, NP may actually be hiding in this

quantity, but is currently obscured by parameter uncertainties. The smoking-gun signal for NP in B_s^0–$\bar{B}_s^0$ mixing would be the observation of sizeable CP violation in $B_s^0 \to J/\psi\phi$ and similar decays. Since there are various specific extensions of the SM where such effects arise (also when taking the ΔM_s constraints into account), we may hope that the LHC will detect them. Moreover, the B_s-meson system allows several determinations of the angle γ of the UT in an essentially unambiguous way, which are another key ingredient for the search of NP, and offers further tests of the SM through strongly suppressed rare decays. After new results from Run II of the Tevatron, the promising physics potential of the B_s-meson system can be fully exploited at the LHC, in particular by the LHCb experiment.

These studies can be nicely complemented through the kaon system, which governed the stage of CP violation for more than 35 years. The future lies now in rare decays, in particular on the $K^+ \to \pi^+\nu\bar{\nu}$ and $K_L \to \pi^0\nu\bar{\nu}$ modes; there is a proposal to measure the former channel at the CERN SPS, and efforts to explore the latter at KEK/J-PARC in Japan. Furthermore, flavour physics offers several other exciting topics. Important examples are top-quark physics, the D-meson system, the anomalous magnetic moment of the muon, electric dipole moments and flavour violation in the charged lepton and neutrino sectors.

The established neutrino oscillations as well as the evidence for dark matter and the baryon asymmetry of the Universe tell us that the SM is incomplete, and specific extensions usually contain also new sources of flavour and CP violation, which may manifest themselves at the flavour factories. Fortunately, the LHC is expected to go into operation in the autumn of 2007. This new accelerator will provide insights into electroweak symmetry breaking and, we hope, also give us direct evidence for physics beyond the SM through the production and subsequent decays of NP particles in the ATLAS and CMS detectors. It is obvious that there should be a very fruitful interplay between these 'direct' studies of NP, and the 'indirect' information provided by flavour physics[5]. I have no doubt that an exciting future is ahead of us!

Acknowledgements

References

[1] J.H. Christenson *et al.*, Phys. Rev. Lett. **13** (1964) 138.

[2] V. Fanti *et al.* [NA48 Collaboration], Phys. Lett. **B465** (1999) 335;
A. Alavi-Harati *et al.* [KTeV Collaboration], Phys. Rev. Lett. **83** (1999) 22.

[3] J.R. Batley *et al.* [NA48 Collaboration], Phys. Lett. **B544** (2002) 97.

[4] A. Alavi-Harati *et al.* [KTeV Collaboration], Phys. Rev. **D67** (2003) 012005.

[5] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **87** (2001) 091801;
K. Abe *et al.* [Belle Collaboration], Phys. Rev. Lett. **87** (2001) 091802.

[6] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **93** (2004) 131801;
Y. Chao *et al.* [Belle Collaboration], Phys. Rev. Lett. **93** (2004) 191802.

[7] A.D. Sakharov, JETP Lett. **5** (1967) 24.

[8] V.A. Rubako and M.E. Shaposhnikov, Usp. Fiz. Nauk **166** (1996) 493; Phys. Usp. **39** (1996) 461;
A. Riotto and M. Trodden, Annu. Rev. Nucl. Part. Sci. **49** (1999) 35.

[9] For a recent review, see W. Buchmüller, R.D. Peccei and T. Yanagida, Annu. Rev. Nucl. Part. Sci. **55** (2005) 311.

---

[5]This topic is currently addressed in detail within a CERN workshop: `http://cern.ch/flavlhc`.

[10] N. Cabibbo, Phys. Rev. Lett. **10** (1963) 531.

[11] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49** (1973) 652.

[12] M. Battaglia *et al.*, *The CKM Matrix and the Unitarity Triangle*, CERN 2003-002-corr. (CERN, Geneva, 2003) [hep-ph/0304132].

[13] For a recent review, see A.J. Buras and M. Jamin, JHEP **0401** (2004) 048.

[14] L. Wolfenstein, Phys. Rev. Lett. **13** (1964) 562.

[15] M. Jeitler, lecture given at this school, see these proceedings.

[16] L. Widhalm, lecture given at this school, see these proceedings.

[17] A.J. Buras and R. Fleischer, Adv. Ser. Direct. High Energy Phys. **15** (1998) 65.

[18] G. Branco, L. Lavoura and J. Silva, *CP Violation*, International Series of Monographs on Physics 103, Oxford Science Publications (Clarendon Press, Oxford, 1999).

[19] I.I. Bigi and A. I. Sanda, *CP Violation*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology (Cambridge University Press, Cambridge, 2000).

[20] K. Kleinknecht, Springer Tracts Mod. Phys. **195** (2003) 1.

[21] T. Mannel, Springer Tracts Mod. Phys. **203** (2004) 1.

[22] Y. Nir, hep-ph/0510413, lectures given at 3rd CERN–CLAF School of High Energy Physics, Malarguee, Mendoza, Argentina, 27 February to 12 March 2005.

[23] A.J. Buras, hep-ph/0505175, lectures given at 2004 European School of High-Energy Physics, Sant Feliu de Guíxols, Barcelona, Spain, 2004, R. Fleischer (Ed.), CERN-2006-003 (CERN, Geneva, 2006), p. 95.

[24] A. Ali, Int. J. Mod. Phys. **A20** (2005) 5080.

[25] M. Gronau, Nucl. Phys. Proc. Suppl. **156** (2006) 69.

[26] A. Höcker and Z. Ligeti, hep-ph/0605217.

[27] W. Buchmüller, lecture given at this school, see these proceedings.

[28] S.L. Glashow, Nucl. Phys. **22** (1961) 579;
S. Weinberg, Phys. Rev. Lett. **19** (1967) 1264;
A. Salam, in *Elementary Particle Theory*, Ed. N. Svartholm (Almqvist and Wiksell, Stockholm, 1968).

[29] S.L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. **D2** (1970) 1285.

[30] S. Eidelman *et al.* [Particle Data Group], Phys. Lett. **B592** (2004) 1.

[31] H. Fritzsch and Z.-Z. Xing, Phys. Lett. **B413** (1997) 396.

[32] C. Jarlskog, Phys. Rev. Lett. **55** (1985) 1039; Z. Phys. **C29** (1985) 491.

[33] J. Bernabeu, G. Branco and M. Gronau, Phys. Lett. **B169** (1986) 243.

[34] J. Ellis, lecture given at this school (unpublished).

[35] L. Wolfenstein, Phys. Rev. Lett. **51** (1983) 1945.

[36] A.J. Buras, M.E. Lautenbacher and G. Ostermaier, Phys. Rev. **D50** (1994) 3433.

[37] A.J. Buras, hep-ph/0101336, lectures given at Erice International School of Subnuclear Physics: Theory and Experiment Heading for New Physics, Erice, Italy, 27 August to 5 September 2000.

[38] R. Aleksan, B. Kayser and D. London, Phys. Rev. Lett. **73** (1994) 18.

[39] C. Jarlskog and R. Stora, Phys. Lett. **B208** (1988) 268.

[40] L.L. Chau and W.-Y. Keung, Phys. Rev. Lett. **53** (1984) 1802.

[41] J. Charles *et al.* [CKMfitter Group], Eur. Phys. J. C **41** (2005) 1; for the most recent updates, see http://ckmfitter.in2p3.fr/.

[42] M. Bona *et al.* [UTfit Collaboration], JHEP **0507** (2005) 028; for the most recent updates, see http://utfit.roma1.infn.it/.

[43] G. Ecker, lecture given at this school, hep-ph/0604165.

[44] A. Khodjamirian, lectures given at the 2003 European School of High-Energy Physics, Tsakhkad-zor, Armenia, 2003, A.G. Olshevskii (Ed.), CERN-2005-007 (CERN, Geneva, 2005), p. 173. [hep-ph/0403145].

[45] M. Lüscher, Annales Henri Poincaré **4** (2003) S197 [hep-ph/0211220].

[46] C. Davies, hep-lat/0509046, lectures given at 58th Scottish Universities Summer School in Physics (SUSSP58): A NATO Advanced Study Institute and EU Hadron Physics 13 Summer Institute, St. Andrews, Scotland, 22–29 August 2004.

[47] M. Lüscher, Proceedings of Science, 23rd International Symposium on Lattice Field, Dublin, 2005, PoS **LAT2005** (2006) paper 002.

[48] F. De Fazio, hep-ph/0010007.

[49] K. Ikado *et al.* [Belle Collaboration], hep-ex/0604018.

[50] D.G. Cassel, eConf **C0304052** (2003) WG501 [hep-ex/0307038].

[51] C. Davies, plenary talk at HEP2005 Europhysics Conference, Lisbon, Portugal, 21–27 July 2005, http://www.lip.pt/events/2005/hep2005/.

[52] C. Aubin *et al.*, Phys. Rev. Lett. **95** (2005) 122002.

[53] M. Artuso *et al.* [CLEO Collaboration], Phys. Rev. Lett. **95** (2005) 251801.

[54] N. Isgur and M.B. Wise, Phys. Lett. **B232** (1989) 113 and **B237** (1990) 527.

[55] M. Neubert, Phys. Rep. **245** (1994) 259.

[56] *The BaBar Physics Book*, eds. P. Harrison and H.R. Quinn, SLAC-R-504 (1998).

[57] M. Neubert, Phys. Lett. **B264** (1991) 455.

[58] M.E. Luke, Phys. Lett. **B252** (1990) 447.

[59] P. Gambino and N. Uraltsev, Eur. Phys. J. **C34** (2004) 181.

[60] O. Buchmüller and H. Flächer, Phys. Rev. **D73** (2006) 073008.

[61] Heavy Flavour Averaging Group [E. Barberio *et al.*], hep-ex/0603003; online updates: http://www.slac.stanford.edu/xorg/hfag.

[62] M. Okamoto *et al.*, Nucl. Phys. Proc. Suppl. **140** (2005) 461; E. Gulez *et al.*, Phys. Rev. **D73** (2006) 074502.

[63] A. Khodjamirian *et al.*, Phys. Rev. **D62** (2000) 114002; P. Ball and R. Zwicky, JHEP **0110** (2001) 019; Phys. Rev. **D71** (2005) 014015; Phys. Rev. **D71** (2005) 014029; Phys. Lett. **B625** (2005) 225.

[64] E. Blucher *et al.*, hep-ph/0512039.

[65] F.J. Gilman and M.B. Wise, Phys. Rev. **D20** (1979) 2392; G. Altarelli, G. Curci, G. Martinelli and S. Petrarca, Phys. Lett. **B99** (1981) 141; A.J. Buras and P.H. Weisz, Nucl. Phys. **B333** (1990) 66.

[66] G. Buchalla, A.J. Buras and M.E. Lautenbacher, Rev. Mod. Phys. **68** (1996) 1125.

[67] A.J. Buras, hep-ph/9806471, lectures given at Summer School on Theoretical Physics: Probing the Standard Model of Particle Interactions, Les Houches, France, 28 July–5 September 1997.

[68] M. Bander, D. Silverman and A. Soni, Phys. Rev. Lett. **43** (1979) 242.

[69] R. Fleischer, Z. Phys. **C58** (1993) 483.

[70] A.J. Buras and R. Fleischer, Phys. Lett. **B341** (1995) 379.

[71] M. Ciuchini, E. Franco, G. Martinelli, M. Pierini and L. Silvestrini, Phys. Lett. **B515** (2001) 33; C. Isola, M. Ladisa, G. Nardulli, T.N. Pham and P. Santorelli, Phys. Rev. **D65** (2002) 094005; C.W. Bauer, D. Pirjol, I.Z. Rothstein and I.W. Stewart, Phys. Rev. **D70** (2004) 054015.

[72] R. Fleischer, Z. Phys. **C62** (1994) 81; Phys. Lett. **B321** (1994) 259 and **B332** (1994) 419.

[73] R. Fleischer, Int. J. Mod. Phys. **A12** (1997) 2459.

[74] N.G. Deshpande and X.-G. He, Phys. Rev. Lett. **74** (1995) 26 [E: ibid., p. 4099];
M. Gronau, O.F. Hernandez, D. London and J.L. Rosner, Phys. Rev. **D52** (1995) 6374.

[75] M. Neubert, B. Stech, Adv. Ser. Direct. High Energy Phys. **15** (1998) 294, and references therein.

[76] A.J. Buras and J.-M. Gérard, Nucl. Phys. **B264** (1986) 371;
A.J. Buras, J.-M. Gérard and R. Rückl, Nucl. Phys. **B268** (1986) 16.

[77] M. Beneke, G. Buchalla, M. Neubert and C. Sachrajda, Phys. Rev. Lett. **83** (1999) 1914; Nucl. Phys. **B591** (2000) 313; Nucl. Phys. **B606** (2001) 245.

[78] J.D. Bjorken, Nucl. Phys. (Proc. Suppl.) **B11** (1989) 325;
M. Dugan and B. Grinstein, Phys. Lett. **B255** (1991) 583;
H.D. Politzer and M.B. Wise, Phys. Lett. **B257** (1991) 399.

[79] H.-n. Li and H.L. Yu, Phys. Rev. **D53** (1996) 2480;
Y.Y. Keum, H.-n. Li and A.I. Sanda, Phys. Lett. **B504** (2001) 6;
Y.Y. Keum and H.-n. Li, Phys. Rev. **D63** (2001) 074006;
Y.Y. Keum and A.I. Sanda, eConf **C0304052** (2003) WG420.

[80] C.W. Bauer, D. Pirjol and I.W. Stewart, Phys. Rev. Lett. **87** (2001) 201806, Phys. Rev. **D65** (2002) 054022; C.W. Bauer, S. Fleming, D. Pirjol, I.Z. Rothstein and I.W. Stewart, Phys. Rev. **D66** (2002) 014017; C.W. Bauer, B. Grinstein, D. Pirjol and I.W. Stewart, Phys. Rev. **D67** (2003) 014010.

[81] A. Khodjamirian, Nucl. Phys. **B605** (2001) 558;
A. Khodjamirian, T. Mannel and B. Melic, Phys. Lett. **B571** (2003) 75.

[82] A.J. Buras, R. Fleischer, S. Recksiegel and F. Schwab, Phys. Rev. Lett. **92** (2004) 101804.

[83] A.J. Buras, R. Fleischer, S. Recksiegel and F. Schwab, Nucl. Phys. **B697** (2004) 133.

[84] M. Gronau and D. Wyler, Phys. Lett. **B265** (1991) 172.

[85] D. Atwood, I. Dunietz, A. Soni, Phys. Rev. Lett. **78** (1997) 3257; Phys. Rev. **D63** (2001) 036005.

[86] R. Fleischer and D. Wyler, Phys. Rev. **D62** (2000) 057503.

[87] M. Gronau, J.L. Rosner and D. London, Phys. Rev. Lett. **73** (1994) 21;
M. Gronau, O.F. Hernandez, D. London and J.L. Rosner, Phys. Rev. **D50** (1994) 4529.

[88] A.B. Carter and A.I. Sanda, Phys. Rev. Lett. **45** (1980) 952; Phys. Rev. **D23** (1981) 1567;
I.I. Bigi and A.I. Sanda, Nucl. Phys. **B193** (1981) 85.

[89] G. Cavoto *et al.*, hep-ph/0603019.

[90] R. Fleischer, Nucl. Phys. **B671** (2003) 459.

[91] A.J. Buras, R. Fleischer, S. Recksiegel and F. Schwab, Eur. Phys. J. **C45** (2006) 701.

[92] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **81** (1998) 2432.

[93] D0 Collaboration, D0 Note 4539-CONF (August 2004).

[94] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **96** (2006) 082002.

[95] M. Masetti, Phys. Lett. **B286** (1992) 160.

[96] M.A. Ivanov, J.G. Körner and O.N. Pakhomova, Phys. Lett. **B555** (2003) 189.

[97] T. Inami and C.S. Lim, Prog. Theor. Phys. **65** (1981) 297 [E: ibid., p. 1772].

[98] A.J. Buras, hep-ph/0307203, lectures given at 41st International University School of Theoretical Physics: Flavour Physics (IUTP 41), Schladming, Styria, Austria, 22–28 February 2003.

[99] A.J. Buras, M. Jamin and P.H. Weisz, Nucl. Phys. **B347** (1990) 491;
J. Urban, F. Krauss, U. Jentschura and G. Soff, Nucl. Phys. **B523** (1998) 40.

[100] A.A. Penin and M. Steinhauser, Phys. Rev. **D65** (2002) 054006;
M. Jamin and B.O. Lange, Phys. Rev. **D65** (2002) 056005;
K. Hagiwara, S. Narison and D. Nomura, Phys. Lett. **B540** (2002) 233.

[101] P. Ball and R. Fleischer, hep-ph/0604249.

[102] S. Aoki *et al.* [JLQCD Collaboration], Phys. Rev. Lett. **91** (2003) 212001.

[103] A. Gray *et al.* [HPQCD Collaboration], Phys. Rev. Lett. **95** (2005) 212001.

[104] M. Okamoto, Proceedings of Science, 23rd International Symposium on Lattice Field, Dublin, 2005, PoS **LAT2005** (2005) paper 013.

[105] S. Laplace, Z. Ligeti, Y. Nir and G. Perez, Phys. Rev. **D65** (2002) 094040.

[106] M. Beneke, G. Buchalla, A. Lenz and U. Nierste, Phys. Lett. **B576** (2003) 173;
M. Ciuchini, E. Franco, V. Lubicz, F. Mescia and C. Tarantino, JHEP **0308** (2003) 031.

[107] *B* Oscillations Working Group: `http://lepbosc.web.cern.ch/LEPBOSC/`.

[108] V. M. Abazov *et al.* [D0 Collaboration], hep-ex/0603029.

[109] A. Abulencia *et al.* [CDF Collaboration], hep-ex/0606027.

[110] A. Lenz, hep-ph/0412007.

[111] A.S. Dighe, I. Dunietz and R. Fleischer, Eur. Phys. J. **C6** (1999) 647.

[112] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **94** (2005) 101803.

[113] V.M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **95** (2005) 171801.

[114] I. Dunietz, Phys. Rev. **D52** (1995) 3048.

[115] R. Fleischer and I. Dunietz, Phys. Rev. **D55** (1997) 259.

[116] R. Fleischer and I. Dunietz, Phys. Lett. **B387** (1996) 361.

[117] I. Dunietz, R. Fleischer and U. Nierste, Phys. Rev. **D63** (2001) 114015.

[118] A.J. Buras, hep-ph/0402191.

[119] R. Fleischer, Phys. Rep. **370** (2002) 537.

[120] For reviews, see, for instance, A. Ali, hep-ph/0412128;
G. Isidori, AIP Conf. Proc. **722** (2004) 181;
M. Misiak, Acta Phys. Polon. **B34** (2003) 4397.

[121] R. Fleischer and T. Mannel, Phys. Lett. **B506** (2001) 311.

[122] R. Fleischer, G. Isidori and J. Matias, JHEP **0305** (2003) 053.

[123] R. Fleischer and T. Mannel, Phys. Lett. **B511** (2001) 240.

[124] T. Goto *et al.*, Phys. Rev. **D70** (2004) 035012.

[125] S. Jäger and U. Nierste, Eur. Phys. J. **C33** (2004) S256.

[126] M. Ciuchini, E. Franco, A. Masiero and L. Silvestrini, eConf **C0304052** (2003) WG307 [J. Korean Phys. Soc. **45** (2004) S223].

[127] P. Ball, S. Khalil and E. Kou, Phys. Rev. **D69** (2004) 115011.

[128] P. Ko, J. Korean Phys. Soc. **45** (2004) S410.

[129] E. Gabrielli, K. Huitu and S. Khalil, hep-ph/0504168.

[130] P. Ball, J.M. Frere and J. Matias, Nucl. Phys. **B572** (2000) 3;
P. Ball and R. Fleischer, Phys. Lett. **B475** (2000) 111.

[131] A.J. Buras, M. Spranger and A. Weiler, Nucl. Phys. **B660** (2003) 225;
A.J. Buras, A. Poschenrieder, M. Spranger and A. Weiler, Nucl. Phys. **B678** (2004) 455;
K. Agashe, G. Perez and A. Soni, Phys. Rev. Lett. **93** (2004) 201804, Phys. Rev. **D71** (2005) 016002.

[132] V. Barger, C.W. Chiang, J. Jiang and P. Langacker, Phys. Lett. **B596** (2004) 229.

[133] S.R. Choudhury, N. Gaur, A. Goyal and N. Mahajan, Phys. Lett. **B601** (2004) 164;
A.J. Buras, A. Poschenrieder and S. Uhlig, Nucl. Phys. **B716** (2005) 173.

[134] W.S. Hou, M. Nagashima and A. Soddu, Phys. Rev. Lett. **95** (2005) 141601.

[135] A.J. Buras, P. Gambino, M. Gorbahn, S. Jäger and L. Silvestrini, Phys. Lett. **B500** (2001) 161.

[136] G. D'Ambrosio, G.F. Giudice, G. Isidori and A. Strumia, Nucl. Phys. **B645** (2002) 155.

[137] A.J. Buras, Acta Phys. Polon. **B34** (2003) 5615.

[138] K. Agashe, M. Papucci, G. Perez and D. Pirjol, hep-ph/0509117.

[139] For a review, see A.A. Petrov, Nucl. Phys. (Proc. Suppl.) **B142** (2005) 333.

[140] For a review, see M. Pospelov and A. Ritz, Ann. Phys. **318** (2005) 119.

[141] For a recent analysis, see P.H. Chankowski, J.R. Ellis, S. Pokorski, M. Raidal and K. Turzynski, Nucl. Phys. **B690** (2004) 279.

[142] R. Fleischer, Eur. Phys. J. **C10** (1999) 299.

[143] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **94** (2005) 161803.

[144] K. Abe *et al.* [Belle Collaboration], BELLE-CONF-0569 [hep-ex/0507037].

[145] H. Boos, T. Mannel and J. Reuter, Phys. Rev. **D70** (2004) 036006.

[146] M. Ciuchini, M. Pierini and L. Silvestrini, Phys. Rev. Lett. **95** (2005) 221804.

[147] P. Ball *et al.*, hep-ph/0003238, in 1999 CERN Workshop on Standard Model Physics (and More) at the LHC, Geneva, 1999, L.M. Mangano and G. Altarelli (Eds.), CERN-2000-004 (CERN, Geneva, 2000), p. 305.

[148] R. Fleischer, Int. J. Mod. Phys. **A12** (1997) 2459.

[149] R. Fleischer and J. Matias, Phys. Rev. **D66** (2002) 054009.

[150] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. **D71** (2005) 032005.

[151] A. Bondar, T. Gershon and P. Krokovny, Phys. Lett. **B624** (2005) 1.

[152] M. Bona *et al.* [UTfit Collaboration], JHEP **0603** (2006) 080.

[153] R. Fleischer, Phys. Lett. **B562** (2003) 234; Nucl. Phys. **B659** (2003) 321.

[154] D. London and R.D. Peccei, Phys. Lett. **B223** (1989) 257;
N.G. Deshpande and J. Trampetic, Phys. Rev. **D41** (1990) 895 and 2926;
J.-M. Gérard and W.-S. Hou, Phys. Rev. **D43** (1991) 2909; Phys. Lett. **B253** (1991) 478.

[155] N.G. Deshpande and X.-G. He, Phys. Lett. **B336** (1994) 471.

[156] Y. Grossman and M.P. Worah, Phys. Lett. **B395** (1997) 241.

[157] V. Barger, C.W. Chiang, P. Langacker and H.S. Lee, Phys. Lett. **B598** (2004) 218.

[158] XXI International Symposium on Lepton and Photon Interactions at High Energies (LP '03), 11–16 August 2003, Fermilab, Batavia, IL, USA, `http://conferences.fnal.gov/lp2003/`.

[159] 32nd International Conference on High-Energy Physics (ICHEP '04), 16–22 August 2004, Beijing, China, `http://ichep04.ihep.ac.cn/`.

[160] XXII International Symposium on Lepton and Photon Interactions at High Energies (LP '05), 30 June–5 July 2005, Uppsala, Sweden, `http://lp2005.tsl.uu.se/ lp2005/`.

[161] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. **D71** (2005) 091102.

[162] K. Abe *et al.* [Belle Collaboration], BELLE-CONF-0569 [hep-ex/0507037].

[163] R. Fleischer, Phys. Lett. **B365** (1996) 399.

[164] D. London and A. Soni, Phys. Lett. **B407** (1997) 61.

[165] M. Gronau, Y. Grossman and J.L. Rosner, Phys. Lett. **B579** (2004) 331.

[166] M. Beneke, Phys. Lett. **B620** (2005) 143.

[167] R. Fleischer, Phys. Lett. **B459** (1999) 306.

[168] M. Gronau and D. London, Phys. Rev. Lett. **65** (1990) 3381.

[169] J.P. Silva and L. Wolfenstein, Phys. Rev. **D49** (1994) 1151.

[170] R. Fleischer and T. Mannel, Phys. Lett. **B397** (1997) 269.

[171] Y. Grossman and H.R. Quinn, Phys. Rev. **D58** (1998) 017504.

[172] J. Charles, Phys. Rev. **D59** (1999) 054007.

[173] M. Gronau, D. London, N. Sinha and R. Sinha, Phys. Lett. **B514** (2001) 315.

[174] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **95** (2005) 151803.

[175] K. Abe *et al.* [Belle Collaboration], BELLE-CONF-0501 [hep-ex/0502035].

[176] R. Fleischer, Eur. Phys. J. **C16** (2000) 87.

[177] R. Fleischer and J. Matias, Phys. Rev. **D66** (2002) 054009.

[178] C.W. Bauer, I.Z. Rothstein and I.W. Stewart, Phys. Rev. Lett. **94** (2005) 231802.

[179] C.W. Bauer, I.Z. Rothstein and I.W. Stewart, hep-ph/0510241.

[180] R. Fleischer and T. Mannel, hep-ph/9706261.

[181] Y. Grossman, M. Neubert and A.L. Kagan, JHEP **9910** (1999) 029.

[182] A.J. Buras and R. Fleischer, Eur. Phys. J. **C16** (2000) 97.

[183] D. Cronin-Hennessy *et al.* [CLEO Collaboration], Phys. Rev. Lett. **85** (2000) 515.

[184] M. Beneke and M. Neubert, Nucl. Phys. **B675** (2003) 333.

[185] T. Yoshikawa, Phys. Rev. **D68** (2003) 054023.

[186] M. Gronau and J.L. Rosner, Phys. Lett. **B572** (2003) 43.

[187] S. Mishima and T. Yoshikawa, Phys. Rev. **D70** (2004) 094024.

[188] Y.L. Wu and Y.F. Zhou, Phys. Rev. **D71** (2005) 021701; Phys. Rev. **D72** (2005) 034037.

[189] C.W. Bauer, D. Pirjol, I.Z. Rothstein and I.W. Stewart, Phys. Rev. **D70** (2004) 054015.

[190] T. Feldmann and T. Hurth, JHEP **0411** (2004) 037.

[191] A.J. Buras and R. Fleischer, Eur. Phys. J. **C11** (1999) 93.

[192] M. Gronau, D. Pirjol and T.M. Yan, Phys. Rev. **D60** (1999) 034021 [E: **D69** (2004) 119901].

[193] A. Ali, E. Lunghi and A.Y. Parkhomenko, Eur. Phys. J. **C36** (2004) 183.

[194] C.W. Chiang, M. Gronau, J.L. Rosner and D.A. Suprun, Phys. Rev. **D70** (2004) 034020.

[195] G. Buchalla and A.S. Safir, Eur. Phys. J. **C45** (2006) 109.

[196] Y.Y. Keum and A.I. Sanda, eConf **C0304052** (2003) WG420.

[197] R. Fleischer and T. Mannel, Phys. Rev. **D57** (1998) 2752.

[198] R. Fleischer, Eur. Phys. J. **C6** (1999) 451.

[199] M. Neubert, JHEP **9902** (1999) 014.

[200] A.J. Buras, R. Fleischer, S. Recksiegel and F. Schwab, Acta Phys. Polon. **B36** (2005) 2015.

[201] M. Neubert and J.L. Rosner, Phys. Lett. **B441** (1998) 403; Phys. Rev. Lett. **81** (1998) 5076.

[202] B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. **D71** (2005) 111102.

[203] M. Gronau, Y. Grossman and J.L. Rosner, Phys. Lett. **B579** (2004) 331.

[204] A.J. Buras, and L. Silvestrini, Nucl. Phys. **B546** (1999) 299;
A.J. Buras, G. Colangelo, G. Isidori, A. Romanino, and L. Silvestrini, Nucl. Phys. **B566** (2000) 3; A.J. Buras, T. Ewerth, S. Jäger and J. Rosiek, Nucl. Phys. **B714** (2005) 103.

[205] G. Buchalla, G. Hiller and G. Isidori, Phys. Rev. **D63** (2001) 014015;
D. Atwood and G. Hiller, hep-ph/0307251.

[206] A.J. Buras, R. Fleischer, S. Recksiegel and F. Schwab, Eur. Phys. J. **C32** (2003) 45.

[207] C. Bobeth, M. Bona, A.J. Buras, T. Ewerth, M. Pierini, L. Silvestrini and A. Weiler, Nucl. Phys. **B726** (2005) 252.

[208] E. Baracchini and G. Isidori, Phys. Lett. **B633** (2006) 309.

[209] A.J. Buras, F. Schwab and S. Uhlig, TUM-HEP-547, hep-ph/0405132.

[210] H.R. Quinn, Nucl. Phys. Proc. Suppl. **37A** (1994) 21.

[211] R. Fleischer, Phys. Lett. **B341** (1994) 205.

[212] A.J. Buras, R. Fleischer and T. Mannel, Nucl. Phys. **B533** (1998) 3.

[213] M. Ciuchini, E. Franco, G. Martinelli and L. Silvestrini, Nucl. Phys. **B501** (1997) 271;

M. Ciuchini, E. Franco, G. Martinelli, M. Pierini and L. Silvestrini, Phys. Lett. **B515** (2001) 33;
C. Isola, M. Ladisa, G. Nardulli, T.N. Pham and P. Santorelli, Phys. Rev. **D65** (2002) 094005;
C.W. Bauer, D. Pirjol, I.Z. Rothstein and I.W. Stewart, Phys. Rev. **D70** (2004) 054015.

[214] R. Fleischer and S. Recksiegel, Eur. Phys. J. **C38** (2004) 251.

[215] P. Ball and R. Zwicky, Phys. Rev. **D71** (2005) 014015.

[216] B. Aubert *et al.* [BaBar Collaboration], BABAR-CONF-04-044 [hep-ex/0408080];
B. Aubert *et al.* [BaBar Collaboration], BABAR-PUB-05-035 [hep-ex/0507023].

[217] K. Abe *et al.* [Belle Collaboration], BELLE-CONF-0524 [hep-ex/0506080].

[218] A.K. Giri and R. Mohanta, JHEP **0411** (2004) 084.

[219] B. Grinstein and D. Pirjol, Phys. Rev. **D62** (2000) 093002.

[220] R. Fleischer and S. Recksiegel, Phys. Rev. **D71** (2005) 051501(R).

[221] A. Ali, E. Lunghi and A.Y. Parkhomenko, Phys. Lett. **B595** (2004) 323.

[222] S.W. Bosch and G. Buchalla, JHEP **0501** (2005) 035.

[223] P. Ball and V.M. Braun, Phys. Rev. **D58** (1998) 094016.

[224] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **94** (2005) 011801.

[225] K. Abe *et al.* [Belle Collaboration], BELLE-CONF-0401 [hep-ex/0408137].

[226] K. Abe *et al.*, BELLE-CONF-0520 [hep-ex/0506079].

[227] For the press release of the Belle collaboration about the observation of the $b \to d$ penguins, see
http://belle.kek.jp/hot/lp05_press.html.

[228] B. Grinstein, Y. Grossman, Z. Ligeti and D. Pirjol, Phys. Rev. **D71** (2005) 011504.

[229] A. Ali, P. Ball, L.T. Handoko and G. Hiller, Phys. Rev. **D61** (2000) 074024;
M. Beneke, T. Feldmann and D. Seidel, Nucl. Phys. **B612** (2001) 25;
A. Ali and A.S. Safir, Eur. Phys. J. **C25** (2002) 583.

[230] T. Hurth and T. Mannel, Phys. Lett. **B511** (2001) 196.

[231] A. Drutskoy (Belle Collaboration), hep-ex/0605110.

[232] O. Schneider, talk at the Flavour in the Era of the LHC Workshop, 7–10 November 2005, CERN,
http://cern.ch/flavlhc.

[233] M. Carena *et al.*, hep-ph/0603106.

[234] M. Ciuchini and L. Silvestrini, hep-ph/0603114.

[235] L. Velasco-Sevilla, hep-ph/0603115.

[236] M. Endo and S. Mishima, hep-ph/0603251.

[237] M. Blanke, A.J. Buras, D. Guadagnoli and C. Tarantino, hep-ph/0604057.

[238] Z. Ligeti, M. Papucci and G. Perez, hep-ph/0604112.

[239] J. Foster, K.I. Okumura and L. Roszkowski, hep-ph/0604121.

[240] K. Cheung, C.W. Chiang, N.G. Deshpande and J. Jiang, hep-ph/0604223.

[241] Y. Grossman, Y. Nir and G. Raz, hep-ph/0605028.

[242] P. Ball and R. Zwicky, JHEP **0604** (2006) 046.

[243] A.S. Dighe, I. Dunietz, H. Lipkin and J.L. Rosner, Phys. Lett. **B369** (1996) 144.

[244] R. Fleischer, Phys. Rev. **D60** (1999) 073008.

[245] Y. Nir and D.J. Silverman, Nucl. Phys. **B345** (1990) 301.

[246] G.C. Branco, T. Morozumi, P.A. Parada and M.N. Rebelo, Phys. Rev. **D48** (1993) 1167.

[247] M. Smizanska (ATLAS Collaboration), private communication.

[248] T. Speer (CMS Collaboration), private communication.

[249] R. Aleksan, I. Dunietz and B. Kayser, Z. Phys. **C54** (1992) 653.

[250] I. Dunietz and R.G. Sachs, Phys. Rev. **D37** (1988) 3186 [E: **D39** (1989) 3515];
I. Dunietz, Phys. Lett. **B427** (1998) 179;
D.A. Suprun, C.W. Chiang and J.L.Rosner, Phys. Rev. **D65** (2002) 054025.

[251] G. Wilkinson, talk at CKM 2005, 15–18 March 2005, San Diego, CA, USA,
`http://ckm2005.ucsd.edu/WG/WG5/thu2/Wilkinson-WG5-S3.pdf`.

[252] G. Balbi *et al.*, CERN-LHCb/2003-123 and 124;
R. Antunes Nobrega *et al.* [LHCb Collaboration], Reoptimized LHCb Detector, Design and Performance, Technical Design Report 9, CERN/LHCC 2003-030.

[253] A. Khodjamirian, T. Mannel and M. Melcher, Phys. Rev. **D68** (2003) 114007.

[254] A. Abulencia [CDF Collaboration], hep-ex/0607021.

[255] A.S. Safir, JHEP **0409** (2004) 053.

[256] S. Baek, D. London, J. Matias and J. Virto, JHEP **0602** (2006) 027.

[257] M. Gronau and J.L. Rosner, Phys. Lett. **B482** (2000) 71.

[258] P.Z. Skands, JHEP **0101** (2001) 008.

[259] A. Soni and D.A. Suprun, Phys. Lett. **B635** (2006) 330.

[260] G. Buchalla and A.J. Buras, Nucl. Phys. **B548** (1999) 309.

[261] G. Buchalla and A.J. Buras, Nucl. Phys. **B400** (1993) 225.

[262] M. Misiak and J. Urban, Phys. Lett. **B451** (1999) 161.

[263] CDF Collaboration, CDF Public Note 8176 (2006) [`http://www-cdf.fnal.gov`].

[264] D0 Collaboration, D0note 4733-CONF (2005) [`http://www-d0.fnal.gov`].

[265] A.J. Buras, Phys. Lett. **B566** (2003) 115.

[266] J.R. Ellis, K.A. Olive and V.C. Spanos, Phys. Lett. **B624** (2005) 47.

CP violation in K decays: experimental aspects

M. Jeitler

Institute for High Energy Physics, Vienna, Austria

Abstract

CP violation was originally discovered in neutral K mesons. Over the last few years, it has also been seen in B mesons, and most of the research in the field is currently concentrating on the B system. However, there are some parameters which could be best measured in kaons. In order to see to which extent our present understanding of CP violation within the framework of the CKM matrix is correct, one has to check for possible differences between the K system and the B system. After an historical overview, I discuss a few of the most important recent results, and give an outlook on experiments that are being prepared.

1   The discovery of CP violation

Symmetries are a salient feature of our world, but so is the breaking of approximate symmetries. Still, for a long time physicists believed that at the level of elementary particles, a high level of symmetry should prevail. In particular, it was expected that all fundamental interactions should be symmetric under the discrete transformations of spatial inversion (parity transformation P), substitution of antiparticles for particles (charge conjugation C), and time inversion (T). However, in 1956 Lee and Yang concluded from experimental data that the weak interactions might not be invariant under spatial inversion, in other words that parity might be violated. This was then explicitly shown in an experiment by Wu in 1957 [1].

For a few years, physicists were inclined to believe that although parity was broken, this symmetry violation was exactly compensated by the charge symmetry violation, and that the symmetry under a combined charge and parity transformation (CP) was exactly conserved. An obvious example was the helicity of the neutrino, which was always observed to be negative ('left-handed neutrino'). Parity transformation would transform it into a right-handed neutrino, which has not been observed in nature. In other words, it appears that parity is maximally violated. However, by performing charge conjugation in addition, one arrives at the right-handed anti-neutrino, which does exist in nature.

However, only a few years after the discovery of parity violation, it turned out that this so-called 'CP symmetry' was also violated, although to a much smaller extent than parity itself. In an experiment carried out at the Brookhaven Alternating Gradient Synchrotron (AGS), Christenson, Cronin, Fitch and Turlay found out that the longer-lived of the two neutral kaons, the K_L^0, which frequently decays into three pions and should therefore be assigned odd parity, in rare cases decayed into a parity-even two-pion state [2]. For some physicists, this was hard to believe and a number of possible explanations were looked at [3] before it was accepted that CP had to be broken at the per mil level.

While at first CP violation was regarded by some physicists as a sort of unwelcome guest and an unnecessary complication of nature, it later turned out that it is in fact vital for our very existence! According to the Big Bang model of the origin of the universe, particles and antiparticles were at first produced in equal numbers. We know that at present, however, the universe contains almost no antimatter. How could matter survive and not be annihilated right away by antimatter, in which case the universe would now be a fairly dull place made up largely of photons, without much of a structure and without physicists wondering about it? In 1967 Andrei Sakharov found three necessary conditions for creating such a 'baryon asymmetry' [4]:

– baryon number violation
– absence of thermal equilibrium
– CP violation.

While this in itself is certainly a very good reason to study CP violation, it must be said that the effects that have been found in the K and now also in the B system are far too weak to explain the size of the baryon asymmetry in the universe, and physicists are looking for new, stronger sources of CP violation.

After the experimental discovery of CP violation, various theories were developed, one of the first being the so-called 'superweak' theory developed by Lincoln Wolfenstein [5]. This theory introduced a fifth fundamental interaction (the 'superweak interaction') on top of the four known interactions: gravity, electromagnetism, strong interaction, and weak interaction. It is interesting to note that this theory was published less than two months after the publication of the experimental discovery but it took 35 years to decide by experiment if it really gave a satisfactory description of CP violation in nature. The paper concluded with the following remark:

"The most interesting point of the model discussed here lies in the possibility that the experiment ... may measure an interaction as much as 10^7 or 10^8 times weaker than the standard weak interactions. If this is the case it may prove extremely difficult to observe CP violation (or T violation) in independent ways."

Almost nine years went by before Makoto Kobayashi and Toshihide Maskawa discovered that CP violation could be described in an organic, natural way by a theory with at least three quark generations [6]. In a model with three generations, there is one physical complex phase (i.e., a phase which cannot be made to disappear, or 'rotate away', by phase conventions), and it is this very phase which gives rise to CP violation.

Soon this model was preferred by most theorists on 'aesthetic' grounds but it was not so easy to decide between these theories by experiment. In the 'superweak' theory CP violation is caused exclusively by state mixing, where the K_L^0 meson consists mostly of the CP-odd state K_2 but has a tiny admixture $\tilde{\epsilon}$ of the CP-even state K_1 while the K_S^0 meson corresponds to the CP-even state K_1 with only a small admixture of the K_2 state:

$$|K_L> \ \approx \ |K_2> + \tilde{\epsilon}|K_1> \tag{1}$$

$$|K_S> \ \approx \ |K_1> + \tilde{\epsilon}|K_2> \tag{2}$$

where

$$|K_1> \ = \ \frac{1}{\sqrt{2}}(|K^0> + |\bar{K}^0>) \tag{3}$$

and

$$|K_2> \ = \ \frac{1}{\sqrt{2}}(|K^0> - |\bar{K}^0>) \ . \tag{4}$$

If CP violation is, however, caused by the phase of the three-generation quark mixing matrix (which is now known as the Cabibbo–Kobayashi–Maskawa or CKM matrix) there should also exist a 'direct' violation of CP in the decay amplitude itself. From experiments it soon became clear that this effect had to be even much more suppressed than the CP violation due to state mixing.

2 The quest for direct CP violation

2.1 Direct CP violation in two-pion decays

If CP violation were exclusively due to state mixing as in Eq. (1), the amount of CP violation in any decay would only be determined by the mixing parameter $\tilde{\epsilon}$ and would therefore be the same for all decay channels. So, in the case of two-pion decays, the ratio of the CP-violating and the CP-conserving decay rates would be the same for the charged decay channel ($K^0 \to \pi^+\pi^-$) and for the corresponding neutral channel ($K^0 \to \pi^0\pi^0$). In other words, we should find

$$\frac{\Gamma(K_L \to \pi^0\pi^0)}{\Gamma(K_S \to \pi^0\pi^0)} = \frac{\Gamma(K_L \to \pi^+\pi^-)}{\Gamma(K_S \to \pi^+\pi^-)} \ . \tag{5}$$

In case of 'direct' CP violation in the decay amplitudes, however, this would not have to be so. The strength of direct CP violation is usually parametrized by a parameter ϵ' which can be obtained from the 'double ratio' R defined by the following equation:

$$R = \frac{\Gamma(K_L \to \pi^0\pi^0)}{\Gamma(K_S \to \pi^0\pi^0)} \Big/ \frac{\Gamma(K_L \to \pi^+\pi^-)}{\Gamma(K_S \to \pi^+\pi^-)} = 1 - 6 \times \mathrm{Re}(\epsilon'/\epsilon) \ . \tag{6}$$

Here, ϵ is a measure of mixing-induced CP violation and related to the mixing parameter $\tilde{\epsilon}$ introduced in Eqs. (1) and (2) by

$$\epsilon = \tilde{\epsilon} + i\frac{\mathrm{Im}(A_0)}{\mathrm{Re}(A_0)} \tag{7}$$

where A_0 is the isospin = 0 amplitude of the $K^0 \to \pi\pi$ decay.

Experimental results soon showed that the 'double ratio' R was very close to unity, and ϵ' had to be much smaller than ϵ. So, R had to be measured with very high precision, and this could only be obtained in a *relative* measurement of the four decay rates entering Eq. (6), which would allow one to reduce the systematic error. Over many years, a series of competing experiments tried to reach the precision needed for establishing a non-zero value of $\mathrm{Re}(\epsilon'/\epsilon)$. At the beginning of the 1990s, the NA31 experiment at CERN had found a more than three-sigma deviation of $\mathrm{Re}(\epsilon'/\epsilon)$ from zero ($\mathrm{Re}(\epsilon'/\epsilon) = (23.0 \pm 6.5) \times 10^{-4}$, Ref. [7]) while the E731 experiment at Fermilab had measured a value compatible both with zero and with NA31 ($\mathrm{Re}(\epsilon'/\epsilon) = (7.4 \pm 5.9) \times 10^{-4}$, Ref. [8]).

In the hope of finally finding a sign of direct CP violation, both laboratories built still more refined experiments: NA48 at CERN and KTeV at Fermilab. In order to minimize systematic errors due to acceptance or changes in the detector over time, both experiments aimed at measuring the four decay rates simultaneously in the same apparatus.

A fundamental problem in this measurement is the fact that K_L^0 and K_S^0 cannot be produced separately. At an accelerator, kaons are produced in strong interaction processes, and the eigenstates of neutral kaons from the point of view of strong interactions (the strangeness eigenstates) are not K_L^0 and K_S^0 but their linear combinations K^0 and $\bar{K}^0$. So, equal amounts of K_L^0 and K_S^0 are created. The large difference in lifetime (the K_L^0 lives 580 times longer than the K_S^0) allows, however, to obtain strongly enhanced samples of K_L^0 or K_S^0 decays.

Figure 1 shows the setup of the NA48 experiment [9]. At a first target, 450-GeV protons produce neutral kaons along with other particles. Charged particles are deflected by magnets while neutral particles continue along the beamline over 120 m. Most K_S^0 particles decay here while the K_L^0 mesons pass a final collimator and enter the fiducial volume of the experiment, which is observed by the detector. A proton beam of relatively low intensity continues along the same axis as the neutral kaon beam. Shortly before the collimator, the protons are deflected onto a second target where again neutral kaons (and other particles) are produced, which enter the detector's fiducial volume through a collimator close to the K_L^0 collimator. Within the fiducial volume of the detector, most of the K_S^0 mesons but only a tiny fraction of the K_L^0 mesons from this second beam decay. By detecting the individual protons which go to the
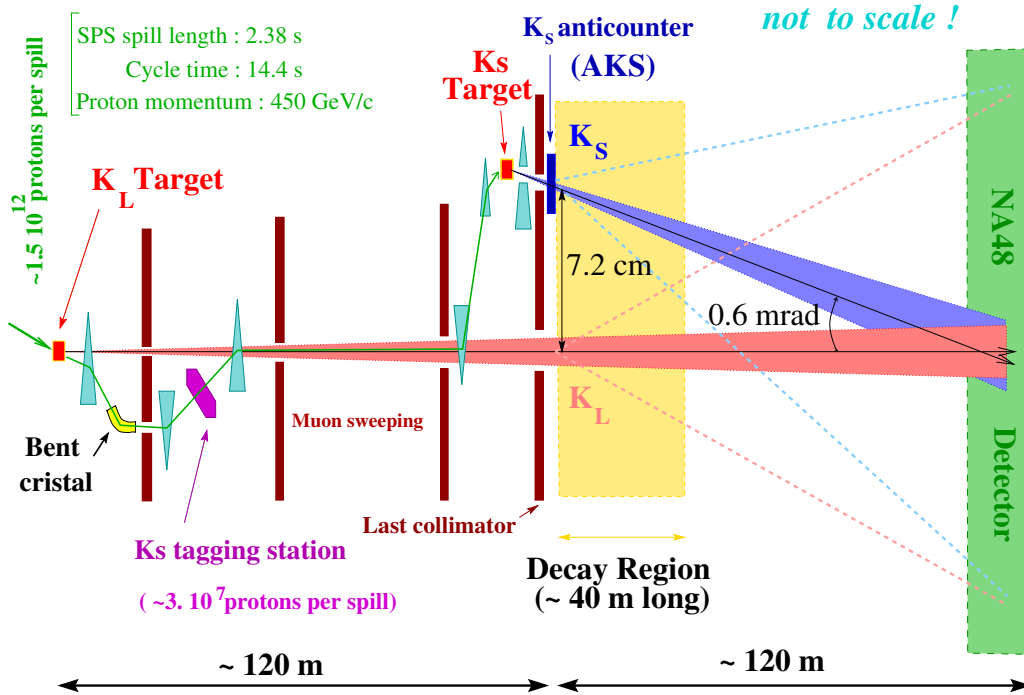
Fig. 1: Setup of the NA48 experiment

second target (by 'tagging' the decays), it can be decided if a particular decay stems from a K_L^0 or from a K_S^0 meson.

By identifying decays into two neutral or two charged pions, the decay rates of all four channels in Eq. (6) can be measured simultaneously, and the double ratio and thus the parameter Re ϵ'/ϵ can be computed. The KTeV experiment [10] used a similar setup. Instead of producing kaons at a second target, however, one of two K_L^0 beams hit a regenerator where its K_S^0 content was strongly enhanced.

These two experiments measured a value of Re ϵ'/ϵ about seven sigmas away from zero and thus established the existence of direct CP violation beyond any doubt [10, 9]. So, the CKM matrix and the Standard Model seem to explain the CP violation we observe in the kaon system, and the superweak model is excluded. However, on account of quantum chromodynamics effects it is very hard to calculate a theoretical value for ϵ' and thus to check how well the theory really describes experimental data. Before a major theoretical breakthrough is achieved (which might come from lattice QCD) it does not make much sense to improve the current experimental measurements of ϵ' (see Fig. 2).

2.2 Direct CP violation in three-pion decays of charged kaons

In order to really understand direct CP violation, it is important to also find it in other channels than in the two-pion decay of neutral kaons. One possibility could be the decays of charged kaons into three pions ($K^\pm \to \pi^\pm\pi^+\pi^-$ and $K^\pm \to \pi^\pm\pi^0\pi^0$). Differently from the neutral kaon system, K^+ and K^- cannot mix because of their different charge, and any difference in the decays for positive and negative kaons would be a sign of direct CP violation. The amount of CP violation in this channel as predicted by the Standard Model is very small and hardly measurable at present. However, certain theories have predicted a significant enhancement of the effect, which could be within the reach of present experiments (see Fig. 3).
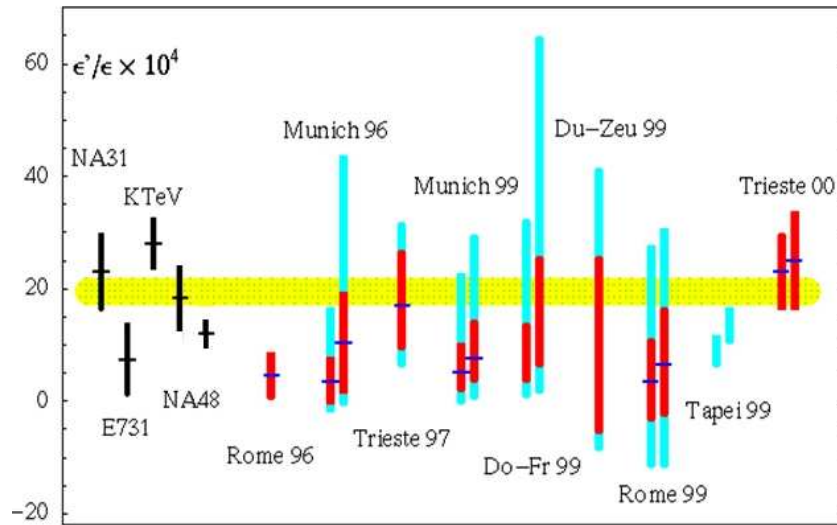
Fig. 2: Measurements of Re ϵ'/ϵ by the four most recent experiments (left) and calculations by various theory groups. The experimental errors are now far smaller than even the most optimistic theoretical errors. Over the last few years, theoretical calculations of Re ϵ'/ϵ have not achieved much progress, and at the moment it would be rather useless to carry out more refined experimental measurements of this quantity.
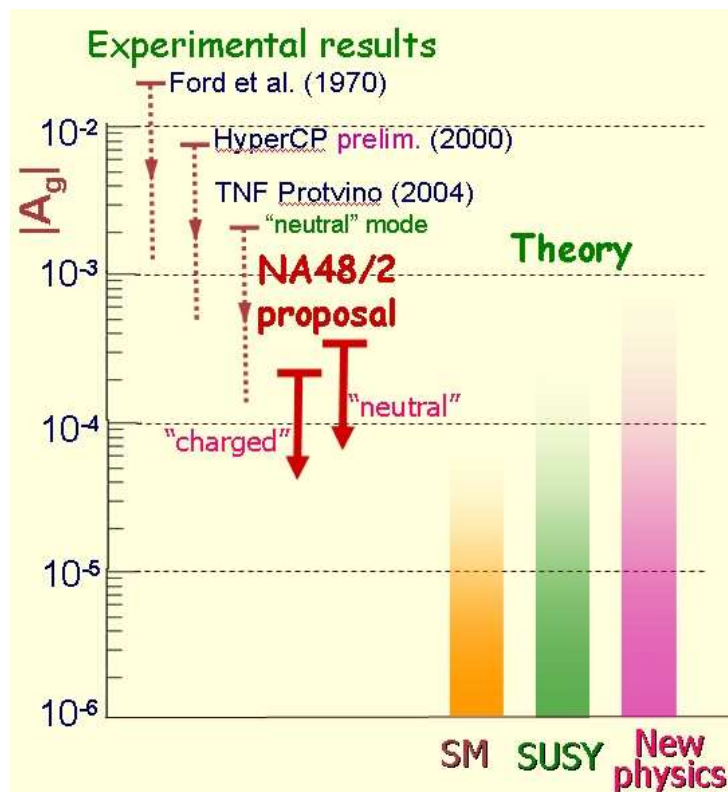
Fig. 3: Experimental limits and theoretical expectations for the size of CP violation visible in the charge asymmetry in the decay of charged kaons to three pions

Measuring a possible difference in the branching ratios of these decays does not look very promising. In all models such differences are predicted to be very small. Moreover, such a measurement would require an accurate knowledge of the kaon flux for both charge signs, which is very hard to achieve from the experimental point of view. However, somewhat larger differences are predicted for the distribution of the decay in phase space (the shape of the Dalitz plot), and these distributions can be measured independently of the kaon flux.

It is usual to parametrize the phase space of the kaon decay products in terms of the Dalitz-plot parameters u and v defined as

$$u = \frac{s_3 - s_0}{m_\pi^2} \qquad\qquad v = \frac{s_2 - s_1}{m_\pi^2} \qquad\qquad (8)$$

where $s_i = (p_K - p_i)^2$, p_K is the four-momentum of the decaying kaon and p_i are the four-momenta of the pions; p_3 corresponds to the 'odd' pion, i.e., the one that differs in charge from the other two, and $s_0 = (s_1 + s_2 + s_3)/3$. The matrix element can then be expanded as

$$|M|^2 \approx 1 + gu + hu^2 + kv^2 + ... \qquad\qquad (9)$$

with the linear g term being the dominant one (see Fig. 4).
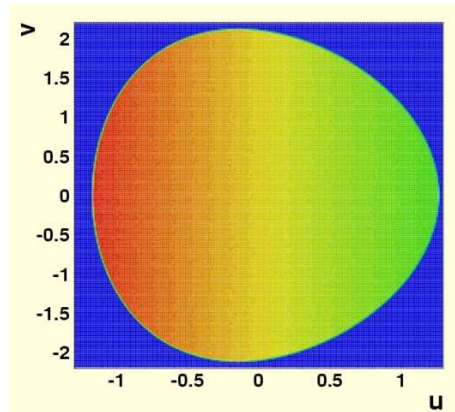
Fig. 4: The Dalitz plot for the decay $K^\pm \to \pi^\pm \pi^+ \pi^-$ describes the phase space distribution of the decay products in terms of the kinematic variables u and v defined in Eq. (8)

This allows to define an asymmetry parameter

$$A_g = \frac{g_+ - g_-}{g_+ + g_-} \qquad\qquad (10)$$

where g_+ and g_- are the values of g measured for positive and negative kaons.

Such measurements have been made by several experiments, the latest and by far the most accurate one being the NA48/2 experiment [11] at CERN. Owing to the smallness of the expected effect a large amount of data must be recorded, and great care has to be taken to minimize systematic errors. As in the case of the measurement in neutral kaons described above, simultaneous K^+ and K^- beams have been used to avoid systematic effects from variations in the detector and the magnetic fields over time. The fields in the beam and spectrometer magnets have been reversed at regular intervals to achieve cancellation of the effects of detector inefficiencies.

So far, part of the data has been analysed and no signal has been seen. The measured value of the asymmetry parameter from the data analysed so far is $A_g = (1.7 \pm 2.9) \times 10^{-4}$ (see Ref. [11]). This is in keeping with the Standard Model but excludes some theories that suggested a possible strong enhancement of the effect.

3 The decay $K_S \to \pi^0\pi^0\pi^0$ and CPT symmetry

Originally, CP violation was discovered in the decay of K_L^0 mesons into two pions. It is harder to study the analogous CP-violating decay of K_S^0 mesons into three pions for two reasons.

Firstly, the K_S^0 meson's decay constant is much larger (the lifetime is much shorter) than for the K_L^0 meson, so that comparable partial decay widths translate into much smaller branching ratios. In other words: while in K_L^0 the CP-violating decay (into two pions) is favoured by phase space over the CP-conserving decay (into three pions), in K_S^0 the decay into three pions is disfavoured both by being CP-violating *and* by the smaller phase space.

Secondly, it is impossible to produce K_S^0 without producing the same amount of K_L^0 mesons at the same time (while a rather pure K_L^0 sample can be obtained by waiting for the K_S^0 component to decay, as in the NA48 experiment described above). In fixed-target experiments, $K_S^0 \to \pi\pi\pi$ decays have been studied by investigating the interference of these decays with the corresponding K_L^0 decays (see Fig. 5). It is, however, also possible to carry out a direct search for such decays by 'tagging' the decaying neutral kaon: if, for example, a ϕ meson decays into two neutral kaons, and one of them is a K_L^0, the other one must be a K_S^0 because the two kaons form an entangled quantum mechanical system. So, a kaon can be identified by measuring the decay of its partner. This method is being used by the KLOE experiment at the DAΦNE e^+e^- collider in Frascati, Italy (see Fig. 6 and Ref. [12]).
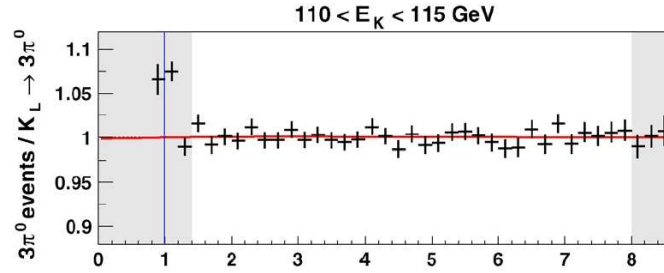
Fig. 5: Ratio of K_S^0/K_L^0 interference data over purely exponentially decaying K_L^0 data (from a target further upstream) as a function of kaon lifetime (in units of the K_S^0 lifetime; points with error bars). The almost constant line is the fit result for the interference signal.
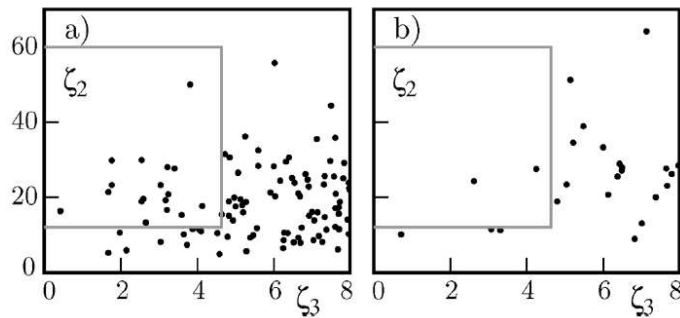
Fig. 6: Events in signal box from the KLOE experiment, where K_S^0 decays are tagged by measuring the K_L^0 which is produced together with the K_S^0. Left: Monte Carlo for 900 pb^{-1}, right: data, 450 pb^{-1}.

The decay $K_S^0 \to \pi^+\pi^-\pi^0$ may be CP violating or CP conserving, depending on the isospin (and thus the angular momentum) in the final state. The CP conserving component of this decay, which is suppressed by the higher angular momentum, has been measured (see, for example, Ref. [13]) and limits for the CP-violating component have been established [14].

For $K_S^0 \to \pi^0\pi^0\pi^0$, not all isospin states are possible that are accessible to $K_S^0 \to \pi^+\pi^-\pi^0$, and therefore this decay is always CP violating. It is of particular interest because of the so-called Bell–Steinberger relation which links possible CPT violation in the $K^0\bar{K}^0$ mixing matrix to CP violating amplitudes in K_L^0 and K_S^0 decays via unitarity (conservation of probability, see Ref. [15]). When parametrizing a possible violation of CPT by a parameter δ, this relation states that

$$(1 + i \tan \phi_{\text{SW}})[\text{Re}(\epsilon) - i\,\text{Im}(\delta)] = \sum A(K_L \to f)^* A(K_S \to f)/\Gamma_S . \tag{11}$$

Here ϕ_{SW} is the so-called 'superweak phase': $\tan \phi_{\text{SW}} = \frac{2\Delta m}{\Delta\Gamma}$; Δm and $\Delta\Gamma$ are the differences in mass and decay rate between K_L^0 and K_S^0. For some time, the uncertainty in the right-hand side of this equation was dominated by the uncertainty in the $K_S^0 \to \pi^0\pi^0\pi^0$ branching ratio.

Of course, there are very good theoretical reasons to believe in CPT symmetry. It is an almost inescapable consequence of Lorentz-invariant quantum field theories. There are, however, ways to theoretically envisage CPT violation, e.g., in superstring theories, which have a fundamentally non-local structure (cf. Ref. [16]). So, the experimental verification of CPT symmetry is not an academic exercise, but an important task!

Recent measurements of the CP-violating parameter

$$\eta_{000} \equiv \frac{A(K_S^0 \to \pi^0\pi^0\pi^0)}{A(K_L^0 \to \pi^0\pi^0\pi^0)} \tag{12}$$

(see Figs. 5, 6, 7, and Refs. [17], [12]) have not allowed one to see the decay $K_S^0 \to \pi^0\pi^0\pi^0$. In fact, the best limit on its branching ratio ($BR(K_S^0 \to \pi^0\pi^0\pi^0) < 1.2 \times 10^{-7}$, from Ref. [12]) is still almost two orders of magnitude away from the Standard Model prediction of $BR(K_S^0 \to \pi^0\pi^0\pi^0) \sim 1.9 \times 10^{-9}$. However, these recent experimental results have significantly reduced the *error* on the branching ratio and thus improved the constraint on CPT violation from the Bell–Steinberger relation. At present, the uncertainty in the right-hand side of Eq. (11) is dominated by the uncertainty on the decay into $\pi^+\pi^-$ — the very first decay in which CP violation was seen!
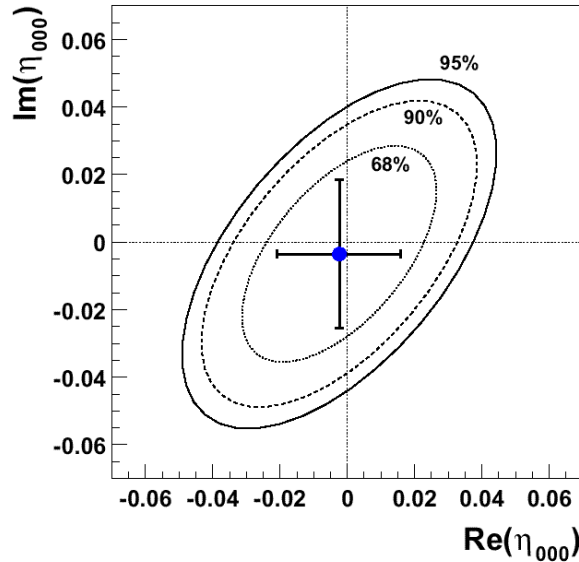
Fig. 7: The experimental result for the parameter η_{000} of the decay $K_S^0 \to 3\pi^0$ measured by the NA48 experiment

4 Time's arrow: the violation of T symmetry

From daily life, we are used to the fact that time always appears to move in the same direction. Travels into the past seem to be restricted to the realm of science fiction. If they were possible, this would completely upset our notion of causality, our understanding of how the world works. So, although for a long time philosophers have thought that space and time have something in common (cf. Ref. [18]), and this vague feeling has developed into the physical concept of spacetime in special relativity, we are convinced that there is a fundamental difference between space and time: we can move around freely in space, but time always progresses in the same direction, and there is nothing we can do about it.

So it might seem ridiculous to even think about something as symmetry in time. When we watch a film and see fragments of china scattered over the floor and coffee spilt over a carpet, and suddenly the fragments and the drops of coffee fly upwards and assemble into a nice cup with good hot coffee inside while the carpet below turns clean, we will be convinced that the film was taken in the reverse direction. But why is this so? After all, the fragments of the cup and the drops of coffee could move in any direction in space. It is the difference in the initial and final states which creates this asymmetry. For just one initial state of the cup being whole and the coffee being in it, there are billions of states for each fragment, for each drop of coffee being in a different place. One 'macroscopic' state is thus presented by an enormous multitude of different 'microscopic' states. If we accept that each microscopic state is equally likely, it becomes clear why the inverse transition between macroscopic states is never observed. So, the explanation for the obvious arrow of time we experience in everyday life lies in thermodynamics and the increase of entropy and has nothing to do with possible asymmetries in the interactions themselves.

If we watch a game of billiards with just three balls and look at the positions of the balls between the shots, it will not be obvious at all in which direction time is going, although the laws of mechanics should be the same as in the previous example. In fact, what we saw in the first example was not a time asymmetry in the interaction itself, but in the initial and final states. All the configurations of three balls on a billiard table are about equally likely, so that we cannot make out the direction in time in this example.

In fact, if we look at the interactions themselves, Newton's laws are symmetric in time. According to all observations, most of the basic interactions in nature—the strong, electromagnetic, and gravitational interactions—are all time symmetric. What about the weak interactions? For a long time it has been believed that the product of the three transformations of parity (P), particle–antiparticle exchange (C) and time inversion (T), in short 'CPT', is conserved under all interactions. This is true for any kind of interaction in a relativistic field theory (see Ref. [19]). Although recently it has been envisaged that CPT invariance might still be violated on a very small scale, it is an experimental fact that it is conserved to a very good approximation. So, CP violation should entail T violation, and one may say that in this indirect way, T asymmetry in weak interactions was discovered back in 1964 when CP violation was first observed.

Still, if weak interactions are really not symmetric under T, it is desirable to observe this in a more direct way. There are processes between particles whose inverse processes can also be observed. However, it is not straightforward to demonstrate T violation in this way. One major problem is due to finite-state interactions which may exist between the particles that are produced in the process. So, if in the decay of a particle more than one hadron is produced, they will interact strongly while they are sufficiently close to each other, thus influencing the rate of the decay. Likewise, charged particles produced in a decay or an interaction will continue to interact electromagnetically even at a distance, again influencing the rate for the process in question. So, different rates may be observed when looking at a process and its inverse, but this difference is not necessarily due to the basic interaction itself [20].

Again, neutral kaons have allowed us to carry out the first unequivocal measurement of T violation. $K^0 \leftrightarrow \bar{K}^0$ oscillations may serve to compare a process with its inverse, by comparing the number of neutral kaons that are created as K^0 and decay as $\bar{K}^0$ with the opposite process. For this, one needs

to know the flavour state of the neutral kaon (if it is a K^0 or a $\bar{K}^0$) both at production and at the time ofy decay. As stated above, the production of kaons at an accelerator is mediated by strong interactions, which conserve strangeness. At the CPLEAR experiment at CERN (see Ref. [21]) antiprotons impinged on a hydrogen target and the processes $p\bar{p} \to K^+\pi^-\bar{K}^0$ and $p\bar{p} \to K^-\pi^+K^0$ were selected. As K^+ and K^0 each contain a strange quark while K^- and $\bar{K}^0$ each contain an anti-strange quark, and a strange quark can only be produced together with an anti-strange quark, one can determine the flavour content of the neutral kaon by measuring the charge of the charged kaon.

The selected decay channels were the semileptonic channels $K^0 \to \pi^-e^+\nu$ and $\bar{K}^0 \to \pi^+e^-\bar{\nu}$ (so-called 'Ke3 decays'). Here, the charge of the pion allows one to determine the flavour state of the neutral kaon by means of the so-called '$\Delta S = \Delta Q$ rule', which is an expression of the experimental fact that no flavour-changing neutral currents are observed at tree level. This fact has been explained by the so-called 'GIM mechanism', which led to the prediction of the charm quark (see Ref. [22]). When a neutral kaon decays semileptonically, the s quark turns into a u quark (strangeness and charge change by +1), or the $\bar{s}$ quark into a $\bar{u}$ quark (strangeness and charge change by –1). The other quark (the $\bar{d}$ or the d quark) flies on as a 'spectator' (see Fig. 8). Owing to the absence of flavour-changing neutral currents it never happens, however, that the s quark transforms into a d quark (or the $\bar{s}$ quark into a $\bar{d}$ quark).
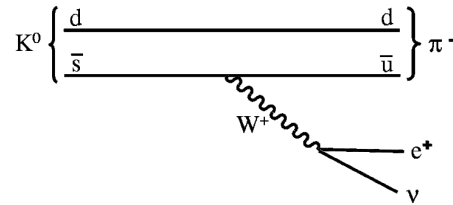
Fig. 8: Semileptonic decay of neutral kaons

So, the flavour of the neutral kaon can be determined both at production and at decay, and the difference in the rates of $K^0 \to \bar{K}^0$ and $\bar{K}^0 \to K^0$ can be measured:

$$A = \frac{R(\bar{K}^0 \to K^0) - R(K^0 \to \bar{K}^0)}{R(\bar{K}^0 \to K^0) + R(K^0 \to \bar{K}^0)} \; . \tag{13}$$

The measurement by the CPLEAR experiment at CERN yielded a value of $A = (6.6 \pm 1.3_{\text{stat}} \pm 1.0_{\text{syst}}) \times 10^{-3}$ (Ref. [21]), and T violation was thus established by a direct measurement, without making use of any assumptions on CPT symmetry.

It is not obvious that the analysis does not rely on implicit assumptions. An in-depth investigation [23] into the theoretical framework has shown that the measurement does not rely on the assumption of general CPT symmetry. It does, however, have to assume that semileptonic kaon decays are CPT symmetric, or that the Bell–Steinberger relation in its conventional form is valid [16]. This relation is a consequence of unitarity if we assume that all relevant kaon decay channels are known. As the experimental error on the branching ratios is of the order of a per cent, it would in principle be possible (although this may seem unlikely) that there exist hitherto unobserved decays with a branching fraction of 10^{-3}. Although this may appear somewhat far-fetched, this possibility is not necessarily more exotic than the possibility of CPT violation.

There are decays which show T-odd correlations between variables, which might be interpreted as a sign of T violation. One of them is the decay $K_L \to \pi^+\pi^-e^+e^-$, where a strong T-odd correlation has been measured [24]. However, the interpretation of this effect as T violation relies on the assumption of CPT conservation [16] although this is not as obvious as when simply deducing T violation from CP violation. Another experiment, E246 at KEK (Japan), is looking for T violation in the decays $K^+ \to \pi^0\mu^+\nu$ ('$K\mu3$') and $K^+ \to \mu^+\nu\gamma$ ('$K\mu2\gamma$'; see, for example, Ref. [25]). The expected Standard Model

branching ratio is small and the signal has to be seen against the background from electromagnetic final-state interactions. However, non-Standard-Model mechanisms of CP violation could possibly lead to a strong enhancement of the effect. So far, no signal has been seen.

5 Rare kaon decays: hard to measure but easier to calculate

As stated above in the discussion of the ϵ'/ϵ measurement, it is not always straightforward to derive a theoretical Standard Model prediction of decay rates. While the basic weak-interaction processes are thought to be under control, it is well known that strong interactions between the decay products give rise to large corrections which are very hard to calculate (cf. Ref. [26]). The Standard Model with three generations of quarks whose coupling is described by the Cabibbo–Kobayashi–Maskawa matrix, yields a plausible description of the phenomena of CP violation we have discovered. However, in most cases technical difficulties in the calculations do not allow us to make accurate predictions, so that possible limitations of the Standard Model that would require modifications in the theory may escape us. This is, however, exactly what physicists are looking for. There are good reasons to believe that the Standard Model cannot be 'the whole truth' and that there must be some sort of 'New Physics'. (With regard to CP violation, the observed baryon asymmetry in the universe discussed in Section 1 is one of these reasons.)

There are, however, a few rare kaon decays that can be calculated with much better accuracy. These decays feature only one strongly interacting particle in the final state, so that QCD corrections play a much smaller role. For the very same reason, and because of their small branching ratios, they are problematic from the experimental point of view. Their accurate measurement will be the main target of kaon physics over the coming years.

5.1 $K^0 \rightarrow \pi^0 l^+ l^-$

An accurate experimental determination of the directly CP-violating component of the decay $K_L^0 \rightarrow \pi^0 e^+ e^-$ (or $K_L^0 \rightarrow \pi^0 \mu^+ \mu^-$) would yield a value for the height of the so-called 'unitarity triangle' (designated by η), which is a measure of the overall strength of CP violation (see Fig. 9). One complication consists in the fact that these decays also have a CP-conserving part and an indirectly CP-violating component due to state mixing. The CP-conserving component is predicted by theory with good accuracy by making use of experimental data on the decay $K_L^0 \rightarrow \pi^0 \gamma \gamma$. For the electronic mode ($K_L^0 \rightarrow \pi^0 e^+ e^-$) it is negligible. The indirectly CP-violating contribution can be obtained by measuring the same decay for K_S^0 mesons.

The measurement of $K_L^0 \rightarrow \pi^0 e^+ e^-$ itself is complicated by the large background from $K_L^0 \rightarrow \gamma \gamma e^+ e^-$, whose branching ratio is $(5.95 \pm 0.33) \times 10^{-7}$ (Ref. [27]), while for $K_L^0 \rightarrow \pi^0 e^+ e^-$ the Standard Model predicts a branching ratio of only $10^{-12} - 10^{-11}$ (in some SUSY scenarios it could be up to 10^{-10}). As the π^0 decays almost instantaneously into two γ's, both decays show the same particles in the final state. Of course, one expects the invariant mass of the two γ's in the signal channel to be close to the known mass of the π^0, but due to the much higher rate of the background channel there may be some events in it where this also happens by accident (see Fig. 10). Simulation studies predict somewhat different distributions in a few kinematic variables for signal and background events but this is of limited use in a very rare decay where one might find only a handful of events. The best measurement so far found a number of events consistent with background expectations and allowed one to set an upper threshold on the branching ratio: $BR(K_L^0 \rightarrow \pi^0 e^+ e^-) < 2.8 \times 10^{-10}$ (see Ref. [28] and Fig. 11).

What has been measured are the (not so strongly suppressed, CP conserving) decay rates for the corresponding K_S^0 decays. The branching ratios are $BR(K_S^0 \rightarrow \pi^0 e^+ e^-) = (5.8^{+2.9}_{-2.4}) \times 10^{-9}$ (Ref. [29]) and $BR(K_S^0 \rightarrow \pi^0 \mu^+ \mu^-) = (2.9^{+1.5}_{-1.2}) \times 10^{-9}$ (Ref. [30]). As one sees from the number of identified signal events (seven events for the $K_S^0 \rightarrow \pi^0 e^+ e^-$ channel, see Fig. 12) this was not an easy measurement either, although the branching ratio is at least one order of magnitude higher than that of the corresponding K_L^0 decay.
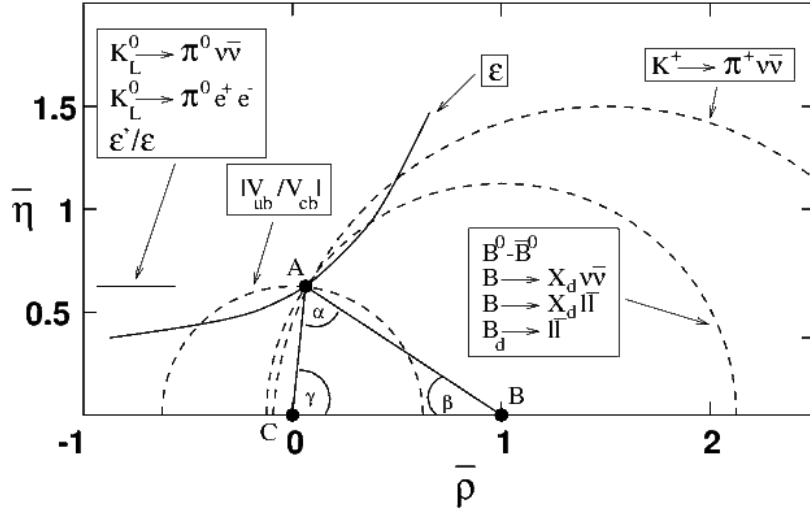
Fig. 9: The unitarity triangle and the various experimental ways to measure its parameters. If one assumes the CKM matrix to be unitary and multiplies it with its Hermitian conjugate, the off-diagonal elements must be zero. In the matrix multiplication, this means that certain sums of three products of CKM matrix elements add up to zero. When graphically representing these products as vectors in the complex plane, this yields a triangle, the so-called 'unitarity triangle'. By choosing the appropriate phase and normalization, two of its end points can be made to lie at (0,0) and (1,0) and the experimentalist's task is to determine the position of the third end point, the triangle's tip. Of course, if the CKM theory is not completely correct, the triangle may not close, and various measurements—in particular those derived from K physics and those derived from B physics—may yield contradictory results. Therefore it is very important to 'overconstrain' the unitarity triangle.
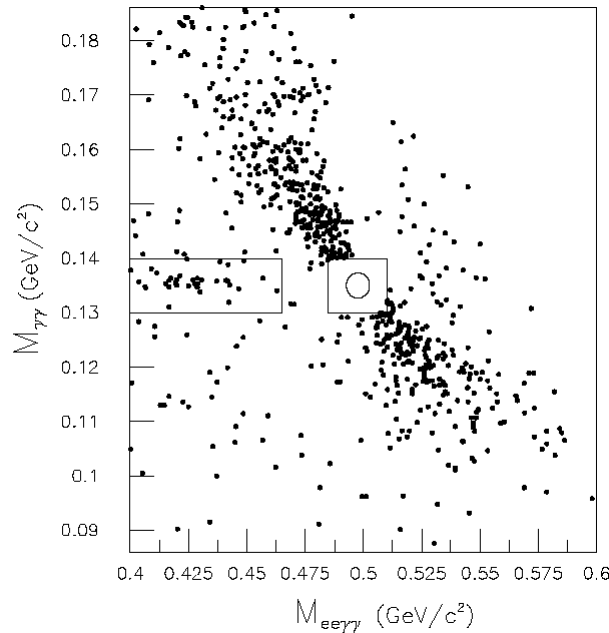
Fig. 10: The invariant mass of the photon pair (vertical axis) against the invariant $ee\gamma\gamma$ mass (horizontal axis). For K_L^0 decays where only two electrons and two photons are produced, the invariant $ee\gamma\gamma$ mass should be close to the K_L^0 mass (0.498 GeV/c^2). If the two photons have been produced in the decay of a neutral pion, their invariant mass should be close to the π^0 mass (0.135 GeV/c^2). The part of the plot where a signal from $K_L^0 \rightarrow \pi^0 e^+ e^-$ should be expected has been masked out by the circular 'signal region' and the square 'control region'. From looking at the rest of the plot, some background from other events is expected for the signal region.
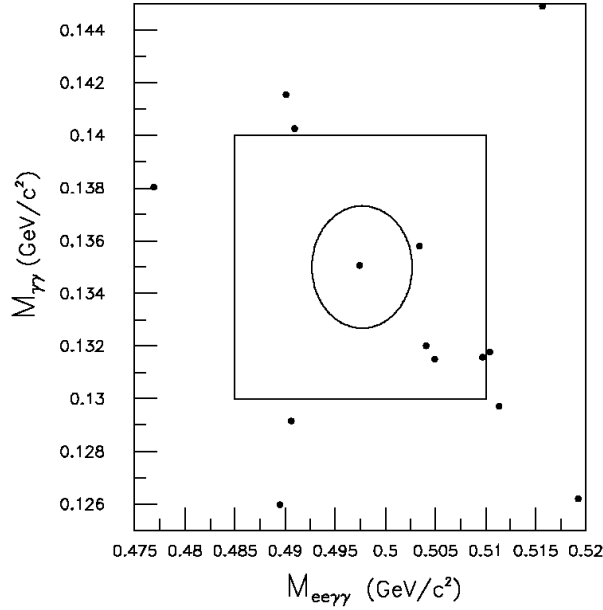
Fig. 11: The circular 'signal region' and the square 'control region' have been unmasked. The event in the signal region is compatible with the background expected from looking at the rest of the plot. So, it cannot be claimed that this should be a signal event.
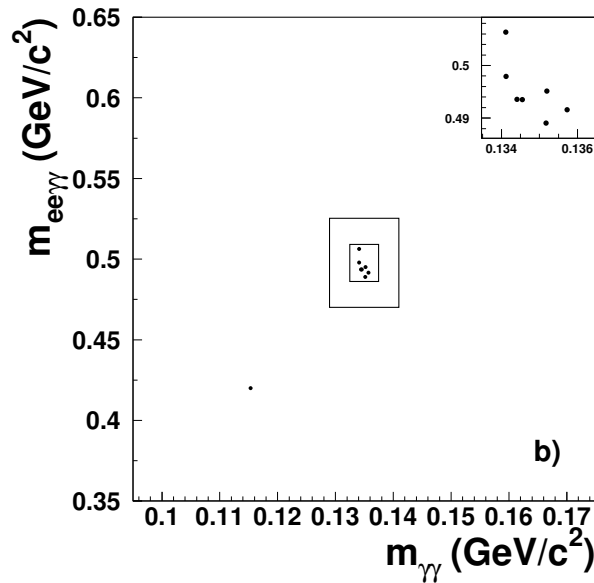
Fig. 12: The signal box for the $K_S^0 \to \pi^0 e^+ e^-$ decay (enlarged in the top right-hand corner). As in Fig. 10, for K_S^0 decays where only two electrons and two photons are produced, the invariant $ee\gamma\gamma$ mass should be close to the K_S^0 mass (0.498 GeV/c^2). If the two photons have been produced in the decay of a neutral pion, their invariant mass should be close to the π^0 mass (0.135 GeV/c^2). For this decay, it has been possible to choose cuts that suppress the background very efficiently, and the nearest background event is very far away from the signal box. The cuts were chosen while the 'signal box' in the centre and the 'control box' that surrounds it were masked. Then the control box was 'opened' to check if for some reason there was an accumulation of background close to the signal box. Only then was the signal box itself opened, thus giving confidence that no bias was introduced by a specific choice of cuts based on the experimentalists' expectations (or 'hopes') concerning the signal value.

Unfortunately, even when the branching ratio of $K_L^0 \to \pi^0 e^+ e^-$ is measured, these numbers by themselves will not be sufficient to determine the relative contribution of indirect and direct CP violation in $K_L^0 \to \pi^0 l^+ l^-$ because of the interference between these two decay amplitudes. Using chiral perturbation theory, the K_S^0 decay's branching ratio can be written as

$$BR(K_S^0 \to \pi^0 e^+ e^-) = 5.2 \times 10^{-9} a_s^2 \, , \tag{14}$$

while the branching ratio of the CP-violating component of the corresponding K_L^0 decay is written as [31]

$$BR_{\text{CPV}}(K_L^0 \to \pi^0 e^+ e^-) = \{15.3 a_s^2 - 6.8 a_s (\text{Im } \lambda_t \times 10^4) + 2.8 (\text{Im } \lambda_t \times 10^4)^2\} \times 10^{-12} \, . \tag{15}$$

While the measurement of the K_S^0 decay rates fixes the absolute size of the K_S^0 decay amplitudes, it does not tell us if the interference term is positive or negative (constructive or destructive interference), which will have to be decided by theory.

Owing to the above-mentioned experimental and theoretical difficulties, this decay channel does not appear to be the most promising for the near future.

5.2 $K^+ \to \pi^+ \nu \bar{\nu}$

This rare decay is not CP violating. When considering the unitarity triangle (see Fig. 9), the rate of this decay yields an ellipse around a point on the abscissa, so that the tip of the unitarity triangle should lie on this ellipse. Its measurement would allow one to derive in an independent way the length of the right side of the unitarity triangle, which has already been measured from $B - \bar{B}$ oscillations. The systematics which enter into these two kinds of measurement are different, so that they are complementary to each other. In case a significant difference in the results should be observed, this would be a strong hint towards new physics.

An advantage for the theoretical treatment of this decay is the fact that the hadronic matrix element can be calculated from other, measured processes, such as $K^+ \to \pi^0 e^+ \nu$. From the experimental point of view, this is of course a difficult measurement because two of the three decay products, the neutrinos, cannot be seen in the detector. The task of the experiment is thus to look for K^+ decays producing nothing but a π^+. The detector has to be completely hermetic in order to suppress other, much more frequent decay channels, such as $K^+ \to \pi^+ \pi^0$, which could be mistaken for a signal event if the π^0 were not observed. Excellent particle identification is needed to suppress decays such as $K^+ \to \mu^+ \nu_\mu$.

Experiments at Brookhaven running over many years found a total of three signal events with small background (see Fig. 13 and Ref. [32]). Figure 14 shows the detected events and the most important source of background from $K^+ \to \pi^+ \pi^0$. This is a good illustration of the virtues of the so-called 'blind analysis' technique in case of very rare decays. The 'signal box' is defined from background studies before events inside the box are looked at. Only when all experimental cuts have been defined is the signal box 'opened'. This ensures that expectations do not influence the result by tempting observers to arbitrarily change the values of the cuts.

The number of detected events (three events) has been enough to establish the decay but the accuracy to which the branching ratio has been measured ($BR(K^+ \to \pi^+ \nu \bar{\nu}) = 1.47^{+1.30}_{-0.89}$) is still too low to really verify the predictions of the Standard Model (see Fig. 15).

Because of funding problems, the experiments at Brookhaven have been discontinued. Other experiments using different techniques are in preparation (see, for example, Ref. [33]) and it is hoped that over the coming years a total of about 10^2 signal events might be observed.
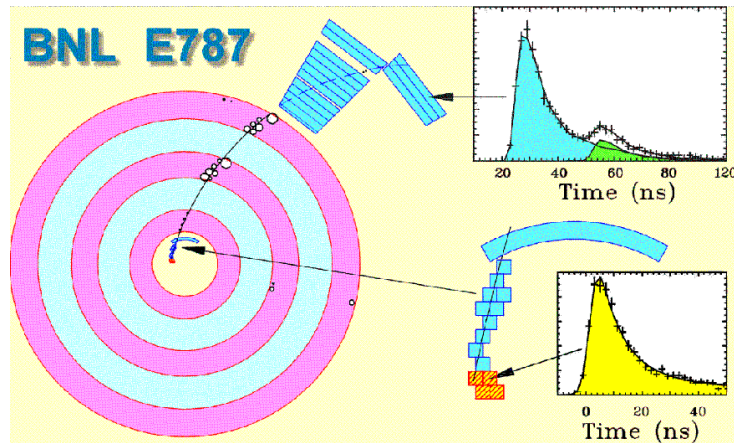
Fig. 13: A $K^+ \to \pi^+ \nu \bar{\nu}$ event in the E787 detector at Brookhaven. The kaon is stopped in the target and emits a signal (blow-up and signal shape at bottom right), which shows no extra activity. The only visible particle from the $K^+ \to \pi^+ \nu \bar{\nu}$ decay is the π^+, which travels towards the top right in the figure. The top-right graph shows the signal from the travelling pion and a second pulse caused by the $\pi^+ \to \mu^+ \nu_\mu$ decay.
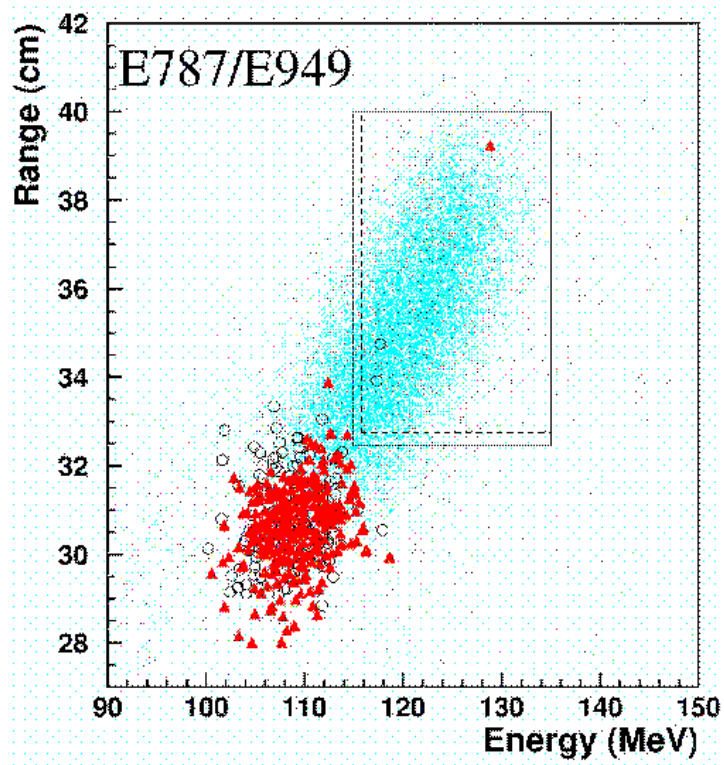
Fig. 14: Kaon range in scintillator versus kaon energy: Monte Carlo generated data for the decay $K^+ \to \pi^+ \nu \bar{\nu}$ (blue dots), and signal and $K^+ \to \pi^+ \pi^0$ background events measured in the two Brookhaven experiments E787 and E949 (circles and triangles). The signal boxes (frames containing the three signal events, at top right) for the two experiments were slightly different. This graph illustrates the virtue of a 'blind analysis' for such a rare decay, where a small change in the cut parameters (which define the signal box) may significantly influence the result.
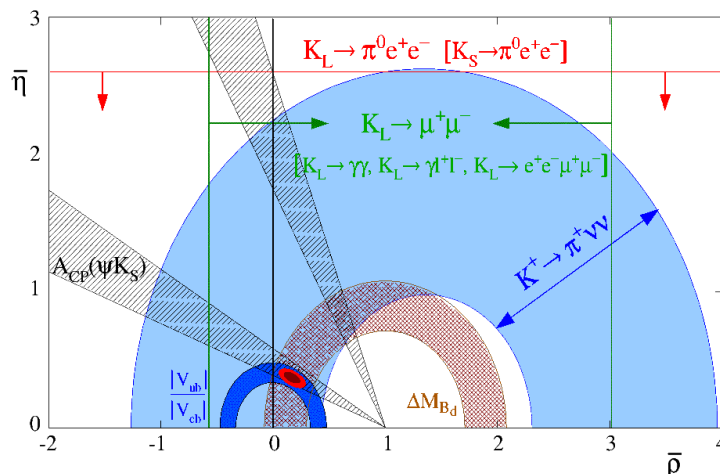
Fig. 15: The branching ratio of the decay $K^+ \to \pi^+ \nu \bar{\nu}$ defines an ellipse in the complex plane, on which the tip of the unitarity triangle should be located. The measurements available so far suffer from a large statistical error and yield a very broad band. By measuring around 100 events, the allowed region could be restricted to a narrow band. Should this band not cover the region of the unitarity triangle's tip as obtained from other measurements, this would be an unequivocal sign of new physics.

5.3 $K^0 \to \pi^0 \nu \bar{\nu}$

This is one of the potentially most instructive decays because it can be calculated with a very small theoretical error, so that any significant deviation between prediction and measurement would be an unequivocal sign of new physics. At the same time, the experiment is extremely challenging, so that this decay, dubbed the 'Holy Grail' of kaon physics, has provoked comments such as "a theorist's dream and an experimentalist's nightmare" where one attempts to measure "nothing goes to nothing plus nothing".

The measurement would directly yield the value of the height of the unitarity triangle, η (cf. Fig. 9). The decay is almost purely directly CP violating, so that its observation would show a second manifestation of direct CP violation in the kaon system. As for the preceding decay, the hadronic matrix element could be obtained from the measured rate of $K^0 \to \pi^+ e^- \nu$, and the total theoretical uncertainty is estimated to be only a few per cent.

Again, care must be taken to fight against the background, which dominates the signal (as expected from Standard Model calculations) by a factor of about 10^{10}. A large number of kaon decays must be observed with a completely hermetic detector, and all possible sources of background must be measured in a convincing way. Experiments were in preparation at BNL and at Fermilab but have been turned down because of funding problems. Hope remains with a new experiment which aims at measuring this decay in the J-PARC facility in Japan [34].

6 Conclusion

The discovery of CP violation in the decays of neutral kaons 41 years ago at first appeared as an unnecessary and unwanted complication of nature. This phenomenon has, however, proved extremely fruitful for our understanding of the world, and has turned out to be a vital ingredient of the universe as we know it. For a long time kaons remained the only particles where CP violation could be observed, but now mainstream research has shifted to the B system, where very promising results have been obtained over the last few years. There are, however, a few outstanding measurements in kaons which pose extreme experimental difficulties but whose results will be indispensable to obtaining a clear overall picture of CP violation.

References

[1] C.S. Wu *et al.*, Phys. Rev. **105** (1957) 1413.

[2] J.H. Christenson, J.W. Cronin, V.L. Fitch and R. Turlay, Phys. Rev. Lett. **13** (1964) 138.

[3] V.L. Fitch, Proc. of Kaon 2001, Frascati Physics Series vol. XXVI, Pisa (2001) 5.

[4] A.D. Sakharov, Violation of CP invariance, C asymmetry, and baryon asymmetry in the universe, JETP Lett. **5** (1966) 24.

[5] L. Wolfenstein, Violation of CP invariance and the possibility of very weak interactions, Phys. Rev. Lett. **13** (1964) 562.

[6] M. Kobayashi and T. Maskawa, CP violation in the renormalizable theory of weak interaction, Prog. Theor. Phys. **49** (1973) 652.

[7] G. Barr *et al.*, Phys. Lett. B **317** (1993) 233.

[8] L.K. Gibbons *et al.*, Phys. Rev. Lett. **70** (1993) 1203.

[9] J.R. Batley *et al.*, A precision measurement of direct CP violation in the decay of neutral kaons into two pions, Phys. Lett. B **544** (2002) 97.

[10] A. Alavi-Harati *et al.*, Measurements of direct CP violation, CPT symmetry, and other parameters in the neutral kaon system, Phys. Rev. D **67** (2003) 012005.

[11] J.R. Batley *et al.*, Search for direct CP violation in the decays $K^\pm \to 3\pi^\pm$, to be published in Phys. Lett. B (2006); I. Mikulec, Search for direct CP violation in $K^+ \to \pi^\pm \pi^+ \pi^-$ decays by NA48/2, Proc. of the XXXXth Recontres de Moriond, hep-ex/0505081 (2005).

[12] F. Ambrosino *et al.*, A direct search for the CP-violating decay $K_S \to 3\pi^0$ with the KLOE detector at DAΦNE, Phys. Lett. B **619** (2005) 61.

[13] J.R. Batley *et al.*, A measurement of the CP-conserving component of the decay $K_S^0 \to \pi^+ \pi^- \pi^0$, Phys. Lett. B **630** (2005) 31.

[14] A. Angelopoulos *et al.*, The neutral kaon decays to $\pi^+ \pi^- \pi^0$: a detailed analysis of the CPLEAR data, Eur. Phys. J. C **5** (1998) 389.

[15] J.S. Bell and J. Steinberger, Proc. of Oxford International Conference on Elementary Particles, Oxford, 1965 (Chilton, RAL, 1966), p. 195.
`http://cdsweb.cern.ch/search.py?recid=832870&`

[16] I.I. Bigi and A.I. Sanda, On limitations of T invariance in K decays, hep-ph/9904484.

[17] A. Lai *et al.*, Search for CP violation in $K^0 \to 3\pi^0$ decays, Phys. Lett. B **610** (2005) 165.

[18] I. Kant, Critique of Pure Reason (1781).

[19] G. Lüders, Dan. Mat. Fys. Medd. **28** (1954) 5, and Ann. Phys. (N.Y.) **2** (1957) 1.

[20] R.G. Sachs, The Physics of Time Reversal (University of Chicago Press, 1987).

[21] A. Angelopoulos *et al.*, First direct observation of time-reversal non-invariance in the neutral-kaon system, Phys. Lett. B **444** (1998) 43; A. Apostolakis *et al.*, Determination of the T- and CPT-violation parameters in the neutral kaon system using the Bell–Steinberger relation and data from CPLEAR, Phys. Lett. B **456** (1999) 297.

[22] S.L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D **2** (1970) 1285.

[23] L. Alvarez-Gaumé *et al.*, Violation of time-reversal invariance and CPLEAR measurements, Phys. Lett. B **458** (1999) 347.

[24] A. Alavi-Harati *et al.*, Observation of CP violation in $K_L \to \pi^+ \pi^- e^+ e^-$ decays, Phys. Rev. Lett. **84** (2000) 408; A. Lai *et al.*, Investigation of $K_{L,S} \to \pi^+ \pi^- e^+ e^-$ decays, Eur. Phys. J. C **30** (2003) 33.

[25] Yu.G. Kudenko, Search for T violation in $K^+ \to \pi^0 \mu^+ \nu$ and $K^+ \to \mu^+ \nu \gamma$ decays, hep-ex/0103007.

[26] I. Scimemi, E. Gámiz and J. Prades, CP violation in kaons: ϵ'_K/ϵ_K vs $K \to 3\pi$, hep-ph 0405204.

[27] Particle Data Group, Review of Particle Physics, Phys. Lett. B **592** (2004) 1.

[28] A. Alavi-Harati *et al.*, Search for the rare decay $K_L \to \pi^0 e^+ e^-$, Phys. Rev. Lett. **93** (2004) 021805.

[29] J.R. Batley *et al.*, Phys. Lett. B **576** (2003) 43.

[30] J.R. Batley *et al.*, Phys. Lett. B **599** (2004) 197.

[31] G. D'Ambrosio, G. Ecker, G. Isidori and J. Portoles, JHEP **08** (1998) 004.

[32] V.V. Anisimovsky *et al.*, Phys. Rev. Lett. **93** (2004) 031801.

[33] Proposal to measure the rare decay $K^+ \to \pi^+ \nu \bar{\nu}$ at the CERN SPS, CERN-SPSC-2005-013, SPSC-P-326, 11.6.2005;
`http://na48.web.cern.ch/NA48/NA48-3/`.

[34] Letter of Intent: Measurement of the $K_L \to \pi^0 \nu \bar{\nu}$ branching ratio,
`http://psux1.kek.jp/~jhf-np/LOIlist/pdf/L05.pdf`.

CP violation in B decays: experimental aspects

L. Widhalm
Institute for High-Energy Physics, Vienna, Austria

Abstract
This lecture, given at the 2005 European School of High-Energy Physics in Austria in succession of the series on CP Violation by Robert Fleischer, sheds light on the topic from a slightly different perspective, which is meant to be a link between theory and the daily work of experimentalists. An overview of B-meson experimental history and phenomenology is followed by a description of B-meson production techniques, facilities worldwide, and a list of important present and future experiments. Current analyses are discussed, and their latest results (as of summer of 2005) are given.

1 Introduction

Studies of CP violation, especially in the B-meson sector, are one of the hottest topics of today's high-energy physics; experiments like BaBar [1] in the USA, or Belle [2] in Japan, have recently started to produce large amounts of B mesons, and recent years have brought numerous new results of high relevance. From the theoretical point of view, several very firm and accurate predictions can be made within the Standard Model, and verified in experiment; New Physics can reveal itself in many places.

This lecture is divided in three main parts. Starting with an experimental history, the first part summarizes the properties of the B meson, with emphasis on how they show up in experiment. A comparison with the K meson stresses the common features as well as the differences (and the reason for these differences). The second part deals with B-meson production and detection techniques, and provides an overview of facilities and experiments worldwide. In the last part, the focus is on a selection of important analyses at these experiments. After a short summary of the theoretical background, recent results and their implications are discussed. The lecture closes with an outlook on the near (and not-that-near) future.

All given experimental numbers, unless some other reference is given, are taken from Particle Data Group (PDG) 2004 [3]. Charge conjugate modes of any given particles are also implied throughout, unless explicitly stated otherwise.

1.1 Disclaimer

The aim of this lecture is not to explain detector physics, experimental setups or analysis techniques in detail, although we give references for further reading about this. It is aimed at students who, after having heard the theoretical lecture on CP violation, would like to be reminded of the key features of B-meson phenomenology, and then get an overview about experimental designs, currently running experiments, and their most important analyses and results. Within the scope of a 90-minutes lecture, the emphasis is on the larger picture rather than on technical (or theoretical) details, addressing experimentalists who are not experts in this field, but would like to understand what is going on.

Although there is very interesting physics with heavier B mesons like the B_s, this lecture will concentrate mainly on the B^0, with some remarks about the $B^{\pm}$.

2 The B meson

2.1 Experimental history and properties

In 1977 the E288 fixed-target experiment at Fermilab (Batavia, USA), studying $\mu\mu$ events, discovered the b-quark in the $\Upsilon(1S)$ resonance [4], and marked the beginning of bottom physics (Table 1). What is

remarkable is the extraordinarily small width of this state of just 53.0 ± 1.5 keV, which is the consequence of the fact that its mass of 9460.30 ± 0.26 MeV is too small to allow the fast decay into two B mesons (composed of a b-quark and a light quark, with a mass of 5279 ± 0.5 MeV). While a decay into B mesons would preserve the b-quarks, and can be mediated by the strong interaction, instead both b-quarks have to decay weakly to lighter quarks, which is much slower and thus gives rise to the unusual observed width.

The same remains true even for the excitation states $\Upsilon(2S)$ (observed at DESY in 1978 [5]) at $m = 10023.26 \pm 0.31$ MeV and $\Upsilon(3S)$ at $m = 10355.2 \pm 0.5$ MeV. The excitation $\Upsilon(4S)$ with a mass of 10580.0 ± 3.5 MeV is the first which allows a decay into B mesons, which immediately broadens the width to $20\,000 \pm 2000 \pm 4000$ keV. This state was observed first in 1979 at the CLEO experiment (CESR, USA) [6]; a review of upsilon spectroscopy can be found, for example, in Ref. [7].

B mesons were discovered a year later, in 1980, also at CESR [8]. A fact that is experimentally quite important is the small mass difference between $\Upsilon(4S)$ and the sum of the two B mesons, which amounts to just 21 MeV. As a consequence, the relative velocity of the B mesons produced via this resonance is very small, and they are produced nearly at rest in the $\Upsilon(4S)$ centre-of-mass frame. Another consequence is that only the lightest B mesons, namely the B^+ (B^0) composed of an anti-b- and a u-quark (d-quark) can be produced via $\Upsilon(4S)$; the required energy for a production of a B_s^0 (with an s-quark instead of the d-quark) is 159 MeV larger than the $\Upsilon(4S)$ mass. Therefore, the production of B mesons at the $\Upsilon(4S)$ threshold is a very clean and efficient process, which is exploited in the so-called B factories discussed later in this lecture (Section 3.1). First evidence of the B_s meson was found much later, at ALEPH (CERN, Switzerland) in 1992 (Ref. [9], and the references therein); the discovery of B_c followed in 1998 at CDF (Fermilab, USA) [10].

1983 was the year of the first measurement of the b-quark inclusive lifetime at PEP (SLAC, USA) and PETRA (DESY, Germany) [11]; the exclusive lifetime of the B mesons was first measured considerably later (1994) at the DELPHI and ALEPH experiments at CERN (Geneva, Switzerland) [12]. Experimentally important is its product with the speed of light, which gives the scale for the spatial separation of primary and secondary decay vertices; it is known today as $c\tau = 501$ μm (461 μm) for B^+ (B^0), which is comparably rather large and is again caused by the fact that the B-meson decay is mediated by a weak transition of the b-quark to a lighter one, which is further suppressed by the CKM-hierarchy as described in Section 2.5.

In 1987, DESY was the first to observe $B^0\bar{B}^0$ oscillations [13]—a feature of neutral meson/anti-meson systems which was already known in the K^0 system, and anticipated also for the B^0. The reason for and the implications of these oscillations will be discussed in more detail in Section 2.4. In the following years, the knowledge of B physics was further improved also by experiments running at LEP (CERN, Geneva) and SLC (Stanford, USA).

The new millennium brought the age of CP violation in B physics: in 2001, the already anticipated large CP violation in the neutral B-meson system was found at both large B-factory experiments at PEP-II (BaBar) and KEKB (Belle) [14]. Only three years later, in 2004, the more subtle direct CP violation was established in B mesons [15] (for comparison: in K physics, three decades lie between the discovery of indirect and direct CP violation). The aspects of CP violation are discussed in Section 2.6.

2.2 The neutral B-meson system

The following discussion assumes B^0 mesons, but is valid generally for any neutral meson. Some important basic features of a system made of a neutral meson B^0 and its antiparticle $\bar{B}^0$ can be derived without the need to know (and independent of) the details of the underlying field theory if one studies the subspace of the complete Hilbert space comprised of just the states $|B^0>= |1>$ and $|\bar{B}^0>= |2>$; in this subspace, the projection of the Hamiltonian H is given by $H_{ij} :=< i|H|j >$, which forms a 2×2 matrix which is—in contrast to the full Hamiltonian—generally *not* Hermitian. Still, any matrix can be

Table 1: B-meson history

Year	Event
1977	Discovery of b-quark in $\Upsilon(1S)$ at FNAL (USA)
1978	$\Upsilon(1S)$ and $\Upsilon(2S)$ at DESY (Germany)
1979	Discovery of $\Upsilon(4S)$ at CESR (USA)
1980	First observation of B mesons at CESR (USA)
1983	Measurement of inclusive b lifetime at PEP and PETRA
1987	$B^0 \bar{B}^0$ oscillations discovered at DESY (Germany)
1992	Evidence of B_s
1993	Observation of time-dependent oscillations
1994	Measurement of exclusive B-meson lifetime
1998	Discovery of B_c
2001	CP violation found at PEP-II (USA) and KEKB (Japan)
2004	Direct CP violation established

decomposed into a Hermitian and an anti-Hermitian part. It turns out to be useful to write H as

$$H_{\text{eff}} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = M - \frac{i}{2}\Gamma = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}$$

where—by virtue of the pulled out i—both M and Γ are *Hermitian*, inferring $M_{ij} = M_{ji}^*$ and $\Gamma_{ij} = \Gamma_{ji}^*$. An additional constraint, namely $H_{11} = H_{22}$, comes from the CPT theorem, which states that the combined symmetry of Charge conjugation, Parity transformation and Time reversal holds in a very general class of quantum field theories.

The projected Schrödinger equation

$$H|B\rangle = i\frac{d}{dt}|B\rangle$$

yields the usual solution

$$|B_{H,L}\rangle(t) = \exp^{-iH_{H,L}t}|B_{H,L}\rangle(0)$$

where $H_{H,L}$ denotes the eigenvalues of H, which are under the assumption of CPT symmetry given as

$$H_{H,L} = H_{11} \pm \sqrt{H_{12}H_{21}}$$

and $|B_{H,L}\rangle$ are eigenstates of the form

$$|B_{H,L}\rangle = p|B^0\rangle \mp q|\bar{B}^0\rangle$$

with

$$\frac{q}{p} = -\frac{H_H - H_L}{2H_{12}} . \tag{1}$$

Rewriting the time-dependent solution using $H_{H,L} := M_{H,L} - \frac{i}{2}\Gamma_{H,L}$ with real M and Γ, one has

$$|B_{H,L}\rangle(t) = \exp^{-\frac{\Gamma_{H,L}}{2}t}\exp^{-iM_{H,L}t}|B_{H,L}\rangle(0)$$

which is interpreted as two neutral mesons (one Heavier with mass M_H, one Lighter with mass M_L), decaying with (generally different) decay constants $\Gamma_{H,L}$. It is conventional to introduce the mean mass $M := \frac{1}{2}(M_H + M_L)$ and $\Delta M := M_H - M_L$; similarly one introduces Γ and $\Delta\Gamma$.

Note that the above applies for all neutral mesons, while the specific experimental character of the B meson (which is quite different from that of the K meson, for example) is due to the specific values of the above parameters. This will be discussed in Section 2.3.

2.2.1 A short pre-discussion of CP violation

The topic of CP violation in the B-meson system will be discussed in greater depth later (Section 2.6), but it makes sense to discuss immediately some basic properties and connections referred to above.

The CP operator is defined as reversing both charge and parity of a state, i.e., a quark q with momentum p is transformed to $\bar{q}$ with momentum $-p$. It is an important point to understand that this does not imply that the CP operator acting on $|B^0\rangle$ gives $|\bar{B}^0\rangle$, all that can be concluded is that the result is *proportional* to $|\bar{B}^0\rangle$, which together with the normalization condition gives

$$\mathrm{CP}|B^0\rangle = \exp^{i\xi}|\bar{B}^0\rangle$$

where ξ is an arbitrary phase which can be *defined* to any value, but has to be chosen consistently. To keep formulas simple, we choose $\xi = 0$ which leads to

$$\mathrm{CP}|B^0\rangle = |\bar{B}^0\rangle$$

and similarly

$$\mathrm{CP}|\bar{B}^0\rangle = |B^0\rangle$$

but not without the warning to check for the definition of this phase whenever looking into different papers (to avoid misunderstandings, many theoretical papers explicitly state the phase wherever it occurs).

With this phase convention, the eigenstates of CP are easily found to be

$$|B_{\mathrm{CP+,CP-}}\rangle = |B^0\rangle \pm |\bar{B}^0\rangle \tag{2}$$

with eigenvalues ± 1. If one assumes CP symmetry, then CP eigenstates are also eigenstates of the Hamiltonian. Comparing the above result with Eq. (1) this infers that $q/p = \pm 1$ and consequently $H_{12} = H_{21}$.

2.3 The different experimental characters of neutral B and K mesons

While the underlying theory is the same, B and K mesons bear rather characteristic (and partly quite different) properties in experiment, especially with respect to CP violation. Table 2 compares the values of some basic parameters for B^0 and K^0. Remarkably, the lifetimes of the two mass eigenstates are practically the same for the B mesons, while they are almost three orders of magnitude different for the K mesons. This is the reason for some of the most characteristic experimental differences between B and K mesons.

For the K mesons, it is natural to think of them in terms of the mass eigenstates, for simple experimental reasons: by just waiting long enough (a couple of τ_L), the lighter component decays away, and one has a pure beam made of the heavier component. On the other hand, if one studies decays immediately after production ($t = \mathcal{O}(\tau_L)$), one mainly observes the lighter component, since it decays much faster. Instead of classifying them into heavier and lighter K mesons, it is usual to speak of a long-lived K_L (which—confusingly—is identified with the heavier kaon state formerly denoted K_H), and a short-lived K_S (formerly K_L).

As the B-meson mass eigenstates—owing to their similar lifetime—do not disentangle equivalently in experiment, frequently they are addressed in terms of the B^0 and $\bar{B}^0$ states (which are the strong eigenstates in which these particles are produced) rather than of the B_H and B_L states. However, it is important to realize that also the K mesons are produced in their strong eigenstates K^0 and $\bar{K}^0$, and the reason for preferring $K_{L,S}$ is only experimentally motivated—in some experiments, it is more useful to think of particles, and in others of waves; consequently, there are also experiments which are better understood in terms of K^0 and $\bar{K}^0$.

A notable property of the B meson is its large lifetime, which is rather surprising given the fact that it is much heavier than the K meson, with much more open decay channels. The reason for this

Table 2: Comparison of B meson with K meson

	B^0	K^0
Mean mass M	5279 MeV$/c^2$	497 MeV$/c^2$
Mass difference ΔM	$\approx 3.3 \times 10^{-10}$ MeV$/c^2$	$\approx 3.5 \times 10^{-12}$ MeV$/c^2$
Lifetime $(1/\Gamma)$	$\tau_H = 1.5$ ps	$\tau_H = 51800$ ps
	$\tau_L = 1.5$ ps	$\tau_L = 90$ ps
$|q/p|$	≈ 1	≈ 1
$\arg(q/p)$	$\mathcal{O}(\pi/2)$	$\mathcal{O}(10^{-3})$

experimentally very fortunate and welcome property (see, for example, the discussion in Section 5.1) is the CKM hierarchy discussed in Section 2.5, which suppresses transitions of b-quarks to lighter quarks one order of magnitude more than the comparable transition of an s-quark in a K-meson decay.

Concerning the eigenstate parameters p and q, their ratio is about 1 for both B and K mesons, but while for K mesons also the relative phase between p and q is small, it is large for B mesons. Comparing with Eq. (2) this means that the K mesons $K_{L,S}$ are almost CP eigenstates, while the B mesons are not.

Since K mesons are nearly CP eigenstates, the decay of the almost-CP-odd K_L into the kinematically favoured channel $\pi\pi$ is suppressed to a level of $\mathcal{O}(10^{-3})$. As the CP-odd decay into three pions is kinematically suppressed ($3m_\pi \approx m_K$), the K_L acquires its observed long lifetime.

To summarize the comparison: while the K mesons appear in experimentally distinct long- and short-lived mass eigenstates, which are almost CP eigenstates, B mesons bear a large phase relative to CP eigenstates and cannot as easily be separated experimentally into their mass eigenstates. Therefore as will be discussed in the following section, interference between the eigenstates plays a much more central role for B mesons than for K mesons.

2.4 How B mesons show up in experiment

B mesons are produced via strong interactions, therefore in the strong eigenstates B^0 and $\bar{B}^0$. As they cannot decay strongly, their further development with time is governed by weak interactions, and it is advisable to express the strong eigenstates as compositions of the weak mass eigenstates:

$$
\begin{aligned}
|B^0\rangle &\propto |B_H\rangle + |B_L\rangle \\
|\bar{B}^0\rangle &\propto |B_H\rangle - |B_L\rangle .
\end{aligned}
$$

Since time evolution is different for $|B_H\rangle$ and $|B_L\rangle$, generally *interference* occurs:

$$
\begin{aligned}
|B^0\rangle(t) &= \exp^{-\frac{t}{2\tau}} \exp^{-iMt} \left[|\cos(\frac{\Delta M}{2}t)|B^0\rangle + i\frac{q}{p}\sin(\frac{\Delta |M}{2}t)|\bar{B}^0\rangle \right] \\
\bar{B}^0\rangle(t) &= |\exp^{-\frac{t}{2\tau}} \exp^{-iMt} \left[i\frac{p}{q}|\sin(\frac{\Delta M}{2}t)|B^0\rangle + \cos(\frac{\Delta |M}{2}t)|\bar{B}^0\rangle \right]
\end{aligned}
$$

where τ is the mean lifetime, and the very small difference $\Delta\tau$ has been neglected. Obviously there is a damped *oscillation* between B^0 and $\bar{B}^0$ with a frequency determined by the mass difference ΔM. Figure 1 shows three characteristic cases for values of ΔM in relation to $1/\tau$. If ΔM is large compared to $1/\tau$, then many oscillations take place before the majority of the B mesons have been decayed; if ΔM is small, then most of the B mesons decay before the first oscillation occurs. Both cases are unfavourable in experiment; the former because of limits in time resolution, the latter because of limits in statistics. By some lucky coincidence, however, it turns out that for the neutral B meson $\Delta M \approx 0.7/\tau$, which makes one full oscillation well observable, as will be shown later in this lecture.
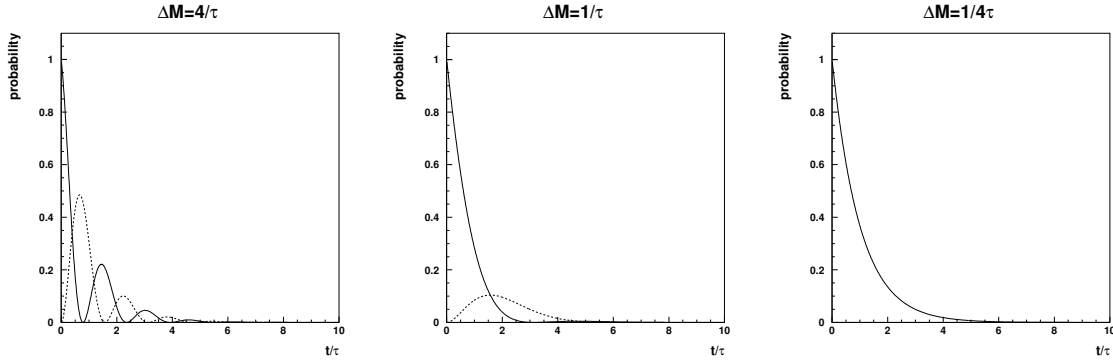
Fig. 1: Oscillation of B mesons: probability to find a B^0 (solid) or $\bar{B}^0$ (dashed) at time t, assuming a B^0 at $t = 0$

From the experimental point of view, it is interesting whether (and how) these interferences can be observed. One precondition is obviously that the *flavour* of the B meson (i.e., whether it is a B^0 or $\bar{B}^0$) can be measured. But it is not sufficient to measure the flavour at the time of decay: if the initial state (at $t = 0$) is unknown, then the interference cancels in the average.

The flavour can be *tagged* by looking for a channel which is only (or at least dominantly) open for either B^0 or $\bar{B}^0$, like, for example, the semileptonic decay $B^0 \rightarrow \ell^+ +$ anything and $\bar{B}^0 \rightarrow \ell^- +$ anything, where the charge of the lepton determines (tags) the flavour of the B meson. However, this determines the flavour only at one point in time and does not yet allow one to observe oscillations. To measure the flavour at a second point in time, one could go back to the production at $t = 0$, and determine the flavour at this time using the fact that b-quarks are produced in quark–anti-quark pairs; this works if the second b-quark went into a charged B meson, but only because the charged B meson does not oscillate between its production at $t = 0$ and its later decay, which reveals its charge. If a $B^0\bar{B}^0$ pair is produced at $t = 0$, both neutral B mesons oscillate, and the flavour of both mesons will change with time. However, as the pair builds an *entangled* state, a decay of one B meson at some time t_1 with a certain, tagged flavour forces the opposite flavour for the other B meson *at the same time* t_1. When this second B meson also decays at some later time t_2, flavour oscillation can be observed.

To avoid misunderstandings: the occurrence of oscillations in the B-meson system is *not* an effect of CP violation; as can be seen from the formulae above, it is governed by ΔM, which can be different from zero also when CP symmetry holds. As will be discussed later, the *amplitude* of the oscillations can be connected with the amount of CP violation *in certain decays*.

2.5 The larger picture: CP violation and CKM hierarchy in the Standard Model

To understand the specific properties of B mesons, it is necessary to look at the larger picture: the CKM matrix V_{CKM} describes the conversion between up- and down-type quarks [16]. Within the Standard Model (SM), it is unitary ($V^+V = VV^+ = 1$) which infers that it is defined by nine parameters (three angles and six phases); however, by redefining the quark phases, five of the phases (corresponding to the five relative phases between the quarks; a global phase change does not affect V_{CKM}) can be gauged to zero. Thus, only one phase with physical meaning remains.

It turns out that this single remaining phase is the source of CP violation within the SM (for the scope of this lecture, we do not discuss the strong CP problem [17]), and it is interesting to note that while for only two generations of quarks CP symmetry would necessarily hold, there is no reason for the CP-violating phase to be small in three generations. Experiment shows that it is not, i.e., CP symmetry is not a near miss. However, this does not imply that CP-violating effects are large—on the contrary, it turns out that either the effects are small, or the branching fractions involved.

The hierarchy of the CKM matrix can be illustrated in the approximation

$$V_{\text{CKM}} \approx \begin{pmatrix} 1 & V_{us} & V_{ub}e^{-i\gamma} \\ -V_{us} & 1 & V_{cb} \\ V_{us}V_{cb} - V_{ub}e^{i\gamma} & -V_{cb} & 1 \end{pmatrix}$$

with

$$|V_{us}| = 0.224 \pm 0.003 \ (\lambda) \, , |V_{cb}| = 0.041 \pm 0.002 \ (\mathcal{O}(\lambda^2)) \, , |V_{ub}| = 0.0037 \pm 0.0008 \ (\mathcal{O}(\lambda^3)) \, .$$

The dominant conversions are along the main axis, i.e., among quarks of the same generation. Conversions between the first and second generation are suppressed by about one order of magnitude, characterized by the parameter $\lambda \approx \mathcal{O}(10^{-1})$. Between the second and the third by two orders of magnitude, and finally those between the first and the third by three orders of magnitude. This hierarchy is important to understand certain properties of the B meson.
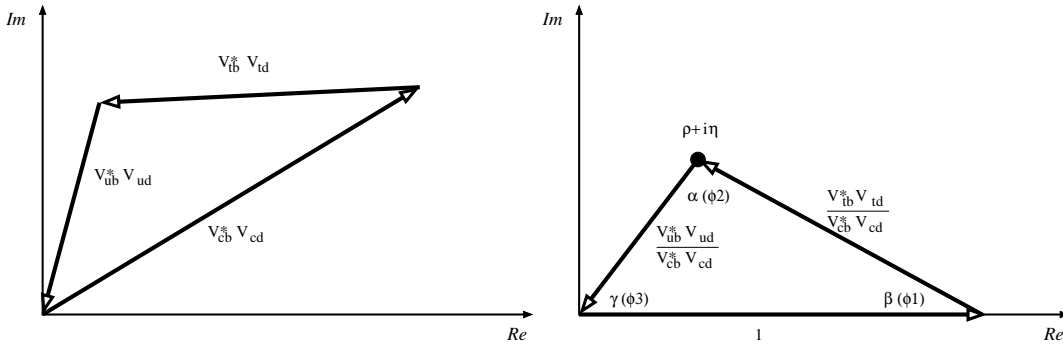
Fig. 2: Unitarity triangles in the complex plane; right the original one, left with one side normalized to 1

To characterize CP violation in the Standard Model, one can utilize the unitarity of the CKM matrix ($V^+V = 1$), which gives nine equations (three in the main diagonal, six off-diagonal). One off-diagonal equation is

$$V_{ub}^*V_{ud} + V_{cb}^*V_{cd} + V_{tb}^*V_{td} = 0 \, .$$

In the complex plane, the three terms of the above equation can be interpreted as the sides of a triangle (Fig. 2, left). It is conventional to normalize one side of the triangle to 1 by dividing the above equation by $V_{cb}^*V_{cd}$. Introducing the real parameters $\bar{\rho}$ and $\bar{\eta}$ as

$$\bar{\rho} + i\bar{\eta} := \frac{V_{ub}^*V_{ud}}{V_{cb}^*V_{cd}} \, ,$$

one arrives at the *unitarity triangle* (UT) (Fig. 2, right); note that there are six unitarity triangles, as there are six off-diagonal equations from CKM unitarity; however, most of the triangles are rather degenerated, i.e., very flat; the unitarity triangle discussed here is the most famous one, and the one usually referred to when speaking of *the* unitarity triangle. Its three angles are denoted α, β, γ (sometimes also ϕ_2, ϕ_1, ϕ_3).

The area of the unitarity triangle is a direct measure for the amount of CP violation in the CKM matrix [16], i.e., with no CP violation, it would be degenerated to a flat line.

2.6 CP violation in the neutral B-meson system

CP violation (CPV) can reveal itself in experiment in several ways, which we discuss here briefly:

CP violation in mixing is a manifestation of *indirect CP violation*; mixing means that mass eigenstates differ from the CP eigenstates (see discussion above in Section 2.2), and can be measured in

asymmetries of semileptonic decays, for example. Note: the mixing of B^0 and $\bar{B}^0$ (i.e., the fact that the mass eigenstates differ from the strong eigenstates) is *not* yet an effect of CP violation. Owing to $|q/p| \approx 1$ this asymmetry is *small* for B mesons (as for K mesons).

As the above asymmetry due to mixing does not require any CP violation in the B decay itself, *CP violation in decay* is another way for CP violation to reveal itself, a manifestation of *direct CP violation*. A non-zero asymmetry requires at least two terms in the amplitude of the decay (i.e., two different Feynman graphs) with a difference in both the strong and weak phases [16].

Since mass eigenstates differ from CP eigenstates, there is interference in the decay to CP eigenstates, and oscillations occur in the corresponding asymmetry (*CP violation in interference*); the oscillation frequency is again determined by ΔM, and the *amplitude* of the oscillation is a measure of CP violation. Generally, the asymmetry is given by $A_{\mathrm{CP}}^{\mathrm{dir}} \cos(\Delta M t) + A_{\mathrm{CP}}^{\mathrm{mix}} \sin(\Delta M t)$ [16] where because $|q/p| \approx 1$, $A_{\mathrm{CP}}^{\mathrm{dir}}$ is only different from zero for decay channels with direct CP violation (i.e., channels with at least two contributing Feynman graphs with different phases, see above), i.e., $A_{\mathrm{CP}}^{\mathrm{dir}}$ *is a measure for direct CP violation*. For channels without direct CP violation, $A_{\mathrm{CP}}^{\mathrm{mix}}$ depends only on purely electroweak parameters, i.e., $A_{\mathrm{CP}}^{\mathrm{mix}}$ *measures the unitarity triangle*. Note that the asymmetry cancels when integrated over time, so the time dependence has to be measured if one wants to study this form of CP violation.

Although all above forms of CP violation occur in the B-meson system, most attention is on the last kind (CP violation in interference), because a large effect directly related to angles in the unitarity triangle is predicted (and has already been confirmed experimentally, *cf.* Section 5.1).

Detailed discussions about the phenomenology of CP violation can be found in Ref. [16] or Refs. [18–20].

2.7 New Physics with the B meson

Although CP violation is already accommodated for within the Standard Model, the measured amount is far too small to explain the matter–antimatter asymmetry observed in the Universe [18]. There are many extensions to the Standard Model which predict sizeable differences in CP variables, e.g., Super-Symmetry (SUSY) brings in dozens of additional CP-violating phases. Therefore, the existence of New Physics is rather expected, and potentially can reveal itself in many places in B physics, and also especially in CP violation.

In the Standard Model, where CP violation is controlled by just one single phase, all possible different experiments are determined to give strongly correlated results: as the sides and angles of the unitarity triangle can be measured independently, and usually in more than one way, it is strongly over-determined. A disagreement would be a sign of physics beyond the Standard Model.

Good candidates to show New Physics are decay channels with sizeable contributions from Feynman graphs involving loops, since such loops may contain (heavy) new particles, which—like the top quark in many Standard Model loops—can have a measurable effect on the amplitudes.

3 Production of B mesons

Precision measurements with B mesons demand a large number of them, and special techniques have evolved with certain advantages and disadvantages, which are briefly discussed here.

3.1 B factories and hadron colliders

Lepton colliders have the advantage of clear environments, but are limited in energy because of synchrotron radiation. An early way to produce b-pairs was via the Z^0 particle, at an energy of about 90 GeV, as done in the LEP experiments at CERN (Geneva, Switzerland) and SLD (Stanford, USA). As

these experiments were not designed for the production of B mesons, the production capacity was not very large.

A more efficient way to produce B mesons is to exploit the $\Upsilon(4S)$ resonance at about 10 GeV, as discussed already in Section 2.1. Advantages are that the production is resonantly enhanced, and that the background is comparably small (in addition, it can be comfortably studied by reducing the collision energy to slightly below threshold). However, only B^0 and $B^\pm$ can be produced in this way, the B_s is too heavy. Still, this is currently the state-of-the-art way to produce large amounts of B mesons, as in the so-called *B factories* at PEP-II (Stanford, USA) and KEKB (Tsukuba, Japan); pioneers were DORIS II (Hamburg, Germany) and CESR (Cornell, USA).

An alternative way to produce B mesons is in hadron colliders; just by pure collision energy, B mesons of any kind (B^0, $B^\pm$, B_s, B_c) can be produced copiously in very large numbers, however, the background is very large. This method is currently used in the CDF and D0 experiments at Fermilab (Batavia, USA); and will be soon joined by the LHC experiments at CERN (Geneva, Switzerland).

A possible future linear collider will combine the advantages of lepton colliders (clean environment) with the higher energy of hadron colliders.

3.2 Symmetric vs asymmetric colliders
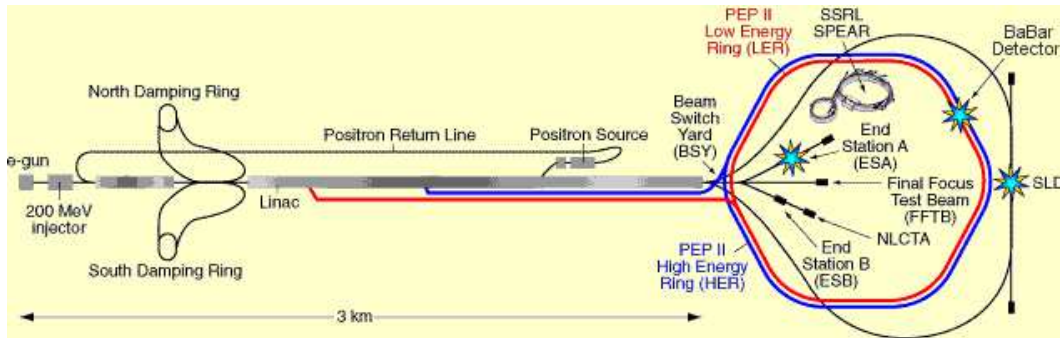
Fig. 3: The PEP-II B factory in Stanford, USA

When B mesons are produced at the $\Upsilon(4S)$ resonance, their relative momentum is very small, since the resonance lies only 24 MeV above the production threshold (*cf.* Section 2.1). In symmetric colliders, where both beams have the same energy, the B mesons are therefore produced nearly at rest, and it is hard (or impossible) to separate their decay vertices (which is important for the measurement of CP violation, see discussion below in Section 5.1). Historically, symmetric colliders were CESR (Cornell, USA) and DORIS (Hamburg, Germany).

Asymmetric colliders, where the beams have different energies, are technologically more demanding, but are essential for modern experiments. In such colliders, the B mesons are boosted in one favoured direction, and their lifetime can be measured via high-resolution vertex detectors. This technology is used in PEP-II (Stanford, USA), Fig. 3, and KEKB (Tsukuba, Japan).

4 B-meson experiments

4.1 B factories and hadron collider experiments

Since 2000, the BaBaR experiment [21] at SLAC (Stanford, USA, see Fig. 4) has been using the PEP-II B factory for precision studies of the B-meson system. A five-layer, double-sided silicon vertex tracker surrounds the interaction point and provides precise reconstruction of track angles and B-decay vertices. A 40-layer drift chamber provides measurements of the transverse momenta of charged particles. An
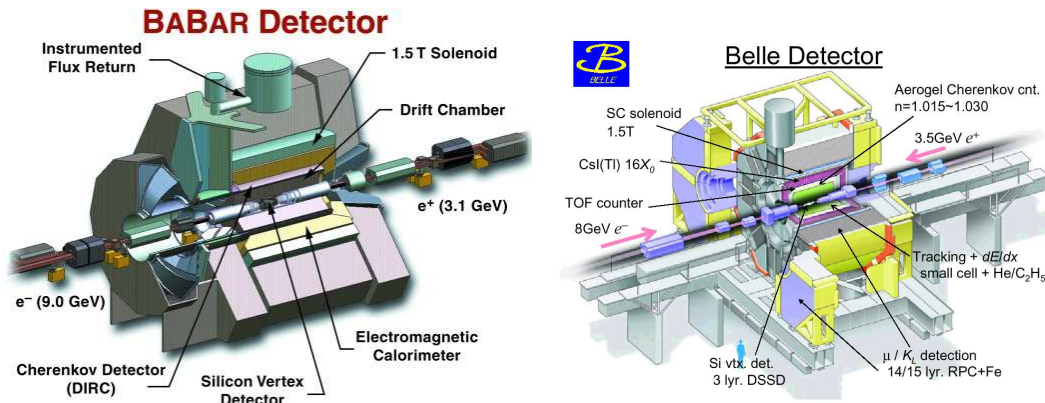
Fig. 4: The BaBaR and Belle detectors

internally reflecting ring-imaging Cherenkov detector is used for particle identification. A CsI(Tl) crystal electromagnetic calorimeter detects photons and electrons. The calorimeter is surrounded by a solenoidal magnet providing a 1.5 T field. The flux return is instrumented with resistive plate chambers used for muon and neutral-hadron identification.

Also since 2000, the Belle experiment at KEK (Tsukuba, Japan, see Fig. 4) with the KEKB B factory has similar goals. The Belle detector [22] is rather similar to that of BaBar: a large-solid-angle magnetic spectrometer that consists of a multilayer silicon vertex detector, a 50-layer central drift chamber, an array of aerogel threshold Cherenkov counters, a barrel-like arrangement of time-of-flight scintillation counters, and an electromagnetic calorimeter comprised of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is instrumented to detect K_L^0 mesons and to identify muons. The old inner detector configuration of a 2.0 cm radius beam pipe and a three-layer silicon vertex detector was upgraded in 2004 to a 1.5 cm radius beam pipe, and a four-layer silicon detector [23].

Besides these two B-factory experiments, there are also two important experiments running at a hadron collider, the Tevatron at Fermilab (Batavia, USA): the CDF II detector [24] consists of a charged-particle tracking system in a magnetic field of 1.4 T, segmented electromagnetic and hadronic calorimeters, and muon detectors. A silicon microstrip detector provides tracking over the radial range 1.5–28 cm and is used to detect displaced secondary vertices. The fiducial region of the silicon detector covers the pseudorapidity range $|\eta| < 2$, while the central tracking system and muon chambers provide coverage for $|\eta| < 1$.

The D0 detector [25] has a silicon microstrip tracker and a central fibre tracker located within a 2 T superconducting solenoidal magnet. The surrounding liquid-argon/uranium calorimeter has a central cryostat covering pseudorapidities $|\eta|$ up to 1.1, and two end-cryostats extending coverage to $|\eta| \approx 4$. A muon system resides beyond the calorimetry, and consists of a layer of tracking detectors and scintillation trigger counters before 1.8 T toroids, followed by two similar layers after the toroids.

Both the CDF II and D0 experiments have a rich physics programme which includes, in addition to B physics, top physics, electroweak physics, and QCD. They are collecting important new results on heavier B mesons like the B_s, which are not accessible to the B-factory experiments.

4.2 The future: experiments at the LHC and super factories

The Large Hadron Collider (LHC), a 14 TeV pp collider at CERN (Geneva, Switzerland), is scheduled to start in 2007. The goal is to get the luminosity to 10^{33} cm^{-2}s^{-1}. Later the luminosity will be increased to nominal 10^{34} cm^{-2}s^{-1}. The launch of the LHC will also bring a whole set of new experiments:
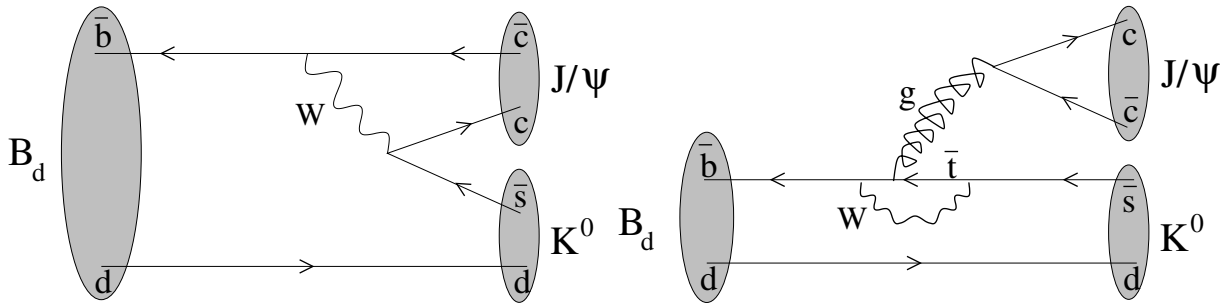
Fig. 5: Tree and penguin diagrams

while ATLAS and CMS have B physics as part of a wider programme, LHCb is a dedicated experiment. LHCb is a single-arm spectrometer covering the range $1.9 < \eta < 4.9$. It consists of a silicon vertex detector [26] which includes a pile-up system surrounding a beam pipe, a magnet and a tracking system, two RICH counters, a calorimeter system and a muon detector. Its construction has started and it will be ready to take data from the start of LHC operation. LHCb will benefit from an unprecedented source of b-hadrons, to substantially improve precision measurements of CP-violation parameters in many different channels. In particular, LHCb will also be capable of measuring CP-violation effects for the first time in decay modes involving B_s mesons.

For the more distant future (around 2010), there are already plans for a super B-factory: Super-Belle is a foreseen successor of the Belle experiment to run at a planned SuperKEKB collider [27], an asymmetric e^+e^- collider with a design luminosity of 5×10^{35} cm^{-2}s^{-1}, which is around 40 times larger than the peak luminosity achieved by the KEKB collider. The Belle detector will be upgraded to Super-Belle to take full advantage of the high luminosity of SuperKEKB. Despite large beam backgrounds, the detector performance will be at least as good as that of the present Belle detector and improvements in several aspects are envisaged.

5   The analyses

The theory behind the analyses is explained in greater detail elsewhere (e.g., Refs. [16, 18–20, 28]), and is only sketched here. For a basic understanding of the relevance of the certain decay channels which are being analysed, it is important to know the two typical Feynman diagrams which contribute to the decay, corresponding to two ways of the transition of the b-quark to a lighter quark: the *tree* diagram (Fig. 5, left), and the *penguin* diagram (Fig. 5, right), which contains a loop. As there are three possibilities for the quark inside the loop (u, c and t), and also three possibilities for the internal gauge boson (g, γ and Z^0—the first known as *strong penguin*, the latter two as *electroweak penguin*), there are in total ten different amplitudes from tree and penguin diagrams which can interfere with each other, and introduce different phases. However, depending on the decay channel, certain amplitudes may be suppressed, which is important for the way a measurement is interpreted.

In the following, a selection of the most important decay channels is briefly discussed, sorted by the amplitudes which dominate the decay.

5.1   Tree-dominated: the golden channel $B^0 \rightarrow \Psi K_{S,L}$

This channel is a decay to CP eigenstates, and therefore is expected to exhibit an asymmetry related to CP violation as discussed in Section 2.6. It is known as *golden channel* because it provides both a clear experimental signature, and a clean theoretical situation: it is dominated by the tree graph, there is no contribution from direct CP violation [16], and thus the asymmetry is a pure sine (*cf.* Section 2.6). The
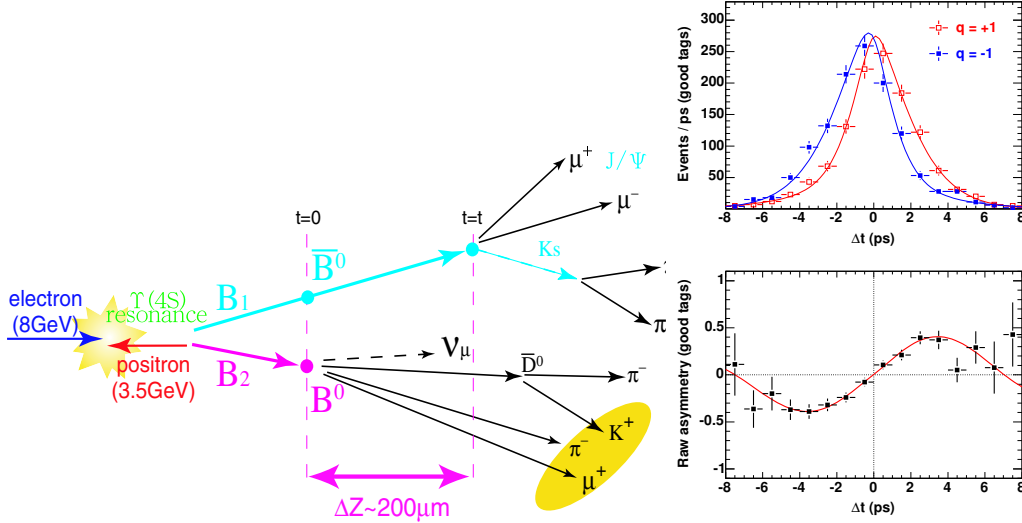
Fig. 6: Scheme of measurement of the time-dependent asymmetry (left), and experimental results from Belle 2005 (right)

amplitude of this sign turns out to be related to angle β of the unitarity triangle:

$$A_{\mathrm{CP}}^{\mathrm{mix}} = -\sin(2\beta) .$$

As the asymmetry cancels when integrated over time, the experimental challenge is to accurately measure the time difference between the decay of the tagging B meson (whose decay determines the flavours of the B meson at this time $t = 0$) and the B meson decaying into the golden channel at some other time t. Figure 6 schematically shows the procedure, which relies on an accurate measurement of the respective decay vertices. Because of the boost of the B mesons in the asymmetric collider, and the relatively long lifetime, the typical distance between the two decay vertices is $\Delta Z = 200\ \mu$m, which is enough to allow a sufficiently accurate measurement with the modern silicon vertex detectors of BaBar and Belle. A more detailed description of the reconstruction method can be found, for example, in Ref. [29].

First results came from both BaBar and Belle in 2001, and have been improved annually since. The 2005 values [14] are in very good agreement with each other (*cf.* Table 3), but also with the Standard Model prediction (using all available other information) of $\sin(2\beta) = 0.68 \pm 0.18$. No indication of a direct CP-violating amplitude was found.

Table 3: Overview of experimental results, as of summer 2005

Channel(s)	Measurement	BaBar	Belle
$B^0 \to \Psi K_{S,L}$	$A_{\mathrm{CP}}^{\mathrm{mix}}$	$0.722 \pm 0.040 \pm 0.023$	$0.625 \pm 0.039 \pm 0.020$
$B^0 \to \Phi K_{S,L}$	$A_{\mathrm{CP}}^{\mathrm{mix}}$	$0.50 \pm 0.25 \pm 0.07$	$0.44 \pm 0.27 \pm 0.05$
$B^0 \to \eta' K_{S,L}$	$A_{\mathrm{CP}}^{\mathrm{mix}}$	$0.36 \pm 0.13 \pm 0.03$	$0.62 \pm 0.12 \pm 0.04$
$B^0 \to \pi\pi$	$A_{\mathrm{CP}}^{\mathrm{mix}}$	$-0.30 \pm 0.17 \pm 0.03$	$-0.67 \pm 0.016 \pm 0.06$
	$A_{\mathrm{CP}}^{\mathrm{dir}}$	$-0.09 \pm 0.15 \pm 0.04$	$-0.56 \pm 0.12 \pm 0.06$
$B^0 \to K\pi$	a_d	$0.133 \pm 0.030 \pm 0.009$	$0.113 \pm 0.022 \pm 0.008$
$B^\pm \to D^0 K / \bar{D}^0 K$	γ	$(67 \pm 28 \pm 13 \pm 11)^\circ$	$(68 \pm 15 \pm 13 \pm 11)^\circ$

5.2 Penguin-dominated: $B^0 \to \Phi K_{S,L}$ and the like

Contrary to the golden channel $B^0 \to \Psi K_{S,L}$ discussed above, this channel is dominated by the penguin diagram. However, the direct CP-violating contribution is again suppressed, and the amplitude of the asymmetry is again related to the angle β:

$$A_{\text{CP}}^{\text{mix}} = \sin(2\beta) \, .$$

The interesting aspect of this channel is that it allows an independent second measurement of β—which tests the Standard Model. Furthermore, it is sensitive to New Physics, as it contains a loop (*cf.* Section 2.7).

Agreement between the BaBar and Belle results in this channel [30] has improved over the years, and no indication of direct CP violation has been found (*cf.* Table 3). Similarly, other channels $B^0 \to X K_{S,L}$ have also been analysed [31]; so far, results are not incompatible with the Standard Model, i.e., the precision result from the golden channel $B^0 \to \Psi K_{S,L}$. However, uncertainties are too large for final conclusions.

5.3 Both tree and penguin: $B^0 \to \pi\pi$ and the like

Although this channel is an example for a mode to which both tree and penguins contribute to a similar extent, it is instructive to discuss the case where the penguin contribution is neglected: in this case, there would be no contribution from direct CP violation, and the amplitude of the asymmetry would be related to another angle of the unitarity triangle:

$$A_{\text{CP}}^{\text{mix}} = \sin(2\alpha) \, .$$

However, the real situation is more complicated; there might be a non-zero amplitude of the cosine in the asymmetry, $A_{\text{CP}}^{\text{dir}}$, and the extraction of α is more difficult than in the corresponding measurement of β discussed above. It can be done, for example, by an isospin analysis (comparing the different isospin states of the $\pi\pi$ system) [32].

Starting with a large disagreement in the first 2001 results, agreement between BaBar and Belle has improved over the years, but is still not very good; in addition, no conclusive answer to the question of direct CP violation can be given yet. The latest results from BaBar (2004) and Belle (2005) [33] are given in Table 3.

5.4 Direct CP violation in $B^0 \to K\pi$

Whereas the above analyses studied time-dependent CP violation due to interference, this channel is used to measure direct CP violation. The corresponding asymmetry is given by

$$a_d = \frac{\Gamma(B^0 \to K^+\pi^-) - \Gamma(\bar{B}^0 \to K^-\pi^+)}{\Gamma(B^0 \to K^+\pi^-) + \Gamma(\bar{B}^0 \to K^-\pi^+)}$$

and is a simple counting experiment, with self-tagging modes. Yet, the effect is smaller than that found in interference.

The experimental establishment of direct CP violation in the B-meson system happened very recently, in 2004. Both experiments BaBar and Belle presented results [15] in good agreement with theoretical prediction [16], *cf.* Table 3.

5.5 Interference in production: $B^\pm \to D^0 K / \bar{D}^0 K$

Another kind of analysis studies this channel, which allows experimental access to the third angle of the unitarity triangle γ. While the previous analyses studied interference between B-meson states or during

decay, this one studies interference between two different end states, namely D^0 and $\bar{D}^0$. Interference between the D mesons can be observed in decay channels common to both. Experimentally, there are three established methods which differ by the studied common decay channel: while the GLW method [34], which uses decays into CP eigenstates like $\pi\pi$, and the ADS method [35] (decays into flavour-specific modes, e.g., $K^{\pm}\pi^{\mp}$) both suffer from small sensitivity for γ, the rather new Dalitz method [36] utilizing decays into three-body modes like, for example, $K\pi\pi$ is more promising. Still, the angle γ is certainly the hardest to measure, and the experimental errors are correspondingly large.

The 2005 results [37] listed in Table 3 show good agreement between BaBar and Belle when combining different channels, there is about one sigma difference in the single mode results.

The analyses presented here are only a—to some extent randomly—selected subset of results from BaBar and Belle; a complete listing of all studies would clearly be beyond the scope of this lecture.

6 Conclusions and outlook

As I hope I have demonstrated above, B physics is a very rich and active field; recent years have seen several results of high impact and relevance, deepening our understanding of the Standard Model, which has withstood another precision test: CP violation was found as large as expected in the B-meson system, and so far no disagreement with the Standard Model has been detected (see Fig. 7).
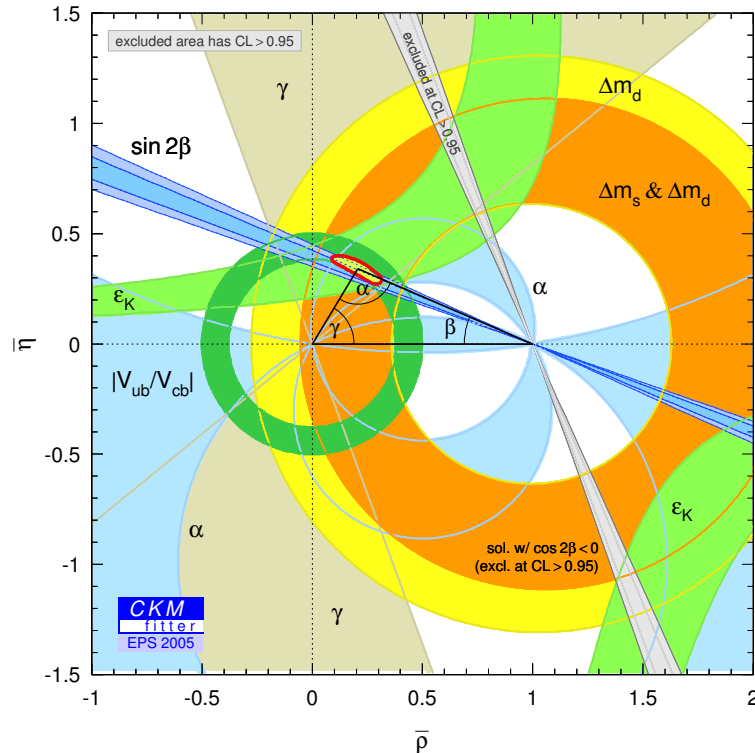
Fig. 7: Current knowledge about the CKM unitarity triangle [38]

However, as the results presented in the previous section show, in several fields no final conclusion can be drawn yet. There are some deviations from the Standard Model predictions which *could* be a hint for New Physics—but some patience is needed until the uncertainties can be further reduced.

Looking into the future, the big running experiments like BaBar, Belle, CDF II and D0 will continue to take data and will further improve their results. In 2007, when the LHC is scheduled to go into operation and bring some orders of magnitude more of data, the dedicated LHCb experiment will enter the game, and the other LHC experiments will also contribute to the field of B physics. This will enable

precision measurements of all three angles and the sides of the unitarity triangle, over-constraining it as a strong test of the Standard Model. The B_s meson will be studied in similar precision, and the search for New Physics will extend to rare decays.

For the next decade, the preparations have already started for a super B-factory, with well over one order of magnitude higher luminosity. It will allow precision measurements of whatever the LHC surprises us with.

Maybe the Standard Model will survive once again—but maybe there will be more. And perhaps, you will be there!

Acknowledgements

References

[1] http://www-public.slac.stanford.edu/babar/.

[2] http://belle.kek.jp .

[3] S. Eidelman *et al.* [Particle Data Group], Phys. Lett. B **592** (2004) 1.

[4] W. R. Innes *et al.*, Phys. Rev. Lett. **39** (1977) 1240 [Erratum ibid. **39** (1977) 1640].

[5] C. W. Darden *et al.*, Phys. Lett. B **78** (1978) 364.

[6] D. Andrews *et al.* [CLEO Collaboration], Phys. Rev. Lett. **45** (1980) 219;
G. Finocchiaro *et al.* [CUSB Collaboration], Phys. Rev. Lett. **45** (1980) 222.

[7] K. Berkelman, Phys. Rep. **98** (1983) 145.

[8] C. Bebek *et al.*, Phys. Rev. Lett. **46** (1981) 84.

[9] D. Buskulic *et al.* [ALEPH Collaboration], Phys. Lett. B **294** (1992) 145.

[10] P. P. Singh [CDF Collaboration], arXiv:hep-ex/9807022.

[11] N. Lockyer *et al.*, Phys. Rev. Lett. **51** (1983) 1316;
M. Althoff *et al.* [TASSO Collaboration], Phys. Lett. B **149** (1984) 524.

[12] P. Abreu *et al.* [DELPHI Collaboration], Z. Phys. C **63** (1994) 3;
D. Buskulic *et al.* [ALEPH Collaboration], Phys. Lett. B **322** (1994) 275.

[13] H. Albrecht *et al.* [ARGUS Collaboration], Phys. Lett. B **192** (1987) 245.

[14] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **94** (2005) 161803;
K. Abe *et al.* [Belle Collaboration], arXiv:hep-ex/0507037.

[15] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **93** (2004) 131801;
K. Abe *et al.*, arXiv:hep-ex/0507045.

[16] R. Fleischer, these proceedings or arXiv:hep-ph/0405091.

[17] M. Creutz, arXiv:hep-th/0303254.

[18] J. P. Silva, arXiv:hep-ph/0410351.

[19] A. J. Buras, Lect. Notes Phys. **629** (2004) 85, arXiv:hep-ph/0307203.

[20] Y. Nir, arXiv:hep-ph/0510413.

[21] B. Aubert *et al.* [BaBar Collaboration], Nucl. Instrum. Meth. A **479** (2002) 1.

[22] A. Abashian *et al.,* [Belle Collaboration], Nucl. Instrum. Meth. A **479** (2002) 117.

[23] Y. Ushiroda *et al.,* Nucl. Instrum. Meth. A **511** (2003) 6.

[24] D. Acosta *et al.* [CDF II Collaboration], Phys. Rev. D **71** (2005) 032001.

[25] S. Abachi *et al.* [D0 Collaboration], Nucl. Instrum. Methods Phys. Res. A **338** (1994) 185;
V. Abazov *et al.* [D0 Collaboration], The upgraded D0 detector, in preparation for submission to Nucl. Instrum. Meth. Phys. Res. A.

[26] S. Klous [LHCb Collaboration], Nucl. Instrum. Meth. A **549** (2005) 55;
L. B. A. Hommels [LHCb Outer Tracker Collaboration], CERN-LHCB-2005-014.

[27] A. G. Akeroyd *et al.* [SuperKEKB Physics Working Group], `arXiv:hep-ex/0406071`.

[28] H. Kakuno *et al.*, Nucl. Instrum. Meth. A **533** (2004) 516, `arXiv:hep-ex/0403022`.

[29] W. Trischuk [Belle Collaboration], eConf **C020805**, TW02 (2002) `arXiv:hep-ex/0212059`.

[30] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **87** (2001) 151801;
K. Abe *et al.* [Belle Collaboration], `arXiv:hep-ex/0507037`.

[31] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **87** (2001) 221802;
B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **94** (2005) 041802;
B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **93** (2004) 131805;
B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **93** (2004) 181805;
K. Abe *et al.* [Belle Collaboration], `arXiv:hep-ex/0507037`.

[32] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **94** (2005) 181802.

[33] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **95** (2005) 151803;
K. Abe *et al.* [Belle Collaboration], `arXiv:hep-ex/0502035`.

[34] M. Gronau and D. Wyler, Phys. Lett. B **265** (1991) 172;
M. Gronau and D. London., Phys. Lett. B **253** (1991) 483.

[35] D. Atwood, I. Dunietz and A. Soni, Phys. Rev. Lett. **78** (1997) 3257;
Phys. Rev. D **63** (2001) 036005.

[36] A. Giri, Y. Grossman, A. Soffer and J. Zupan, Phys. Rev. D **68** (2003) 054018, [`arXiv:hep-ph/0303187`].

[37] B. Aubert *et al.* [BaBar Collaboration], `arXiv:hep-ex/0408088`;
P. Krokovny [Belle Collaboration], `arXiv:hep-ex/0506033`.

[38] CKMfitter Group (J. Charles *et al.*), Eur. Phys. J. C **41** (2005) 1–131, `arXiv:hep-ph/0406184`, updated results and plots available at: `http://ckmfitter.in2p3.fr`.

Relativistic heavy-ion physics: three lectures

L. McLerran

Brookhaven National Laboratory, Upton, NY 11973, USA

Abstract

These lectures provide an introduction to the physics issues which are being studied in the the collisions of ultrarelativistic heavy ions. These issues are focused on the production of new states of matter. The quark–gluon plasma is thermal matter which once existed in the Big Bang. The colour glass condensate is a universal form of matter which controls the high-energy limit of strong interactions. I introduce the student to these topics, discuss results from experiments, and comment upon future opportunities.

1 Introduction

These lectures will introduce the student to the physics issues behind the study of new forms of matter, and the general issue of understanding the high-energy limit of QCD. The full programme of this study involves the collisions of protons on protons, deuterons on nuclei, and nuclei on nuclei. The reason for nuclei is that one can achieve extraordinary energy densities of matter, and because of the large size of nuclei relative to partons, more easily study effects associated with the bulk properties of matter. The highest energies are required, as this allows one to generate the highest energy densities, and as we shall see, at RHIC energies and higher, one can study novel effects associated with the high density of gluons in a hadron wavefunction.

Central to these experimental studies is the production of new forms of matter. This may be a Quark–Gluon Plasma (QGP) or a Colour Glass Condensate (CGC). The properties of these forms of matter are described below.

The outline of these lectures is

– **New states of matter**
In the first lecture, I describe the new forms of matter which may be produced in heavy-ion collisions. These are the quark–gluon plasma and the colour glass condensate.

– **Space–time dynamics**
This lecture describes the space–time dynamics of high-energy heavy-ion collisions. In this lecture, I illustrate how high energy density matter might be formed. I describe how the colour glass condensate may evolve into the quark–gluon plasma, and eventually to a gas of ordinary hadrons.

– **Experiment and theory**
In the final lecture, I show how various experimental measurements might teach us about the properties of matter. Topics discussed are multiplicities and the colour glass condensate, low-transverse-momentum particles and the quark–gluon plasma, heavy vector meson production and confinement, the flavour dependence of the quark–gluon plasma, high-transverse-momentum particles and what they tell us about the CGC and the QGP, and identical particle correlations and what they tell us about the space–time evolution of the matter produced in collisions.

2 Lecture I: High-density matter

2.1 The goals of RHIC

The goal of nuclear physics has traditionally been to study matter at densities of the order of those in the atomic nucleus

$$\epsilon \sim 0.15 \text{ GeV/fm}^3 . \tag{1}$$

High-energy nuclear physics has extended this study to energy densities several orders of magnitude higher. This extension includes the study of matter inside ordinary strongly interacting particles, such as the proton and the neutron, and producing new forms of matter at much higher energy densities in high-energy collisions of nuclei with nuclei, and various other probes.

There are at least three central issues of high-energy nuclear physics:

– **The production of matter at energy densities one to two orders of magnitude higher than that of nuclear matter and the study of its properties**
This matter is at such high densities that it is only simply described in terms of quarks and gluons and is generically referred to as the Quark–Gluon Plasma (QGP). The study of this matter may allow us to better understand the origin of the masses of ordinary particles such as nucleons, and of the confinement of quarks and gluons into hadrons. The QGP will be described below [1].

– **The study of the matter which controls high-energy strong interactions**
This matter is believed to be universal (independent of the hadron), and exists over sizes large compared to the typical microphysics size scales important for high-energy strong interactions. (The microphysics size scale here is about 1 fm and the microphysics time scale is the time it takes light to fly 1 fm, $t \sim 10^{-23}$ s.) It is called a Colour Glass Condensate (CGC) because it is composed of coloured particles, evolves on time scales long compared to microphysics time scales and therefore has properties similar to glasses, and a condensate since the phase-space density of gluons is very high. The study of this matter may allow us to better understand the typical features of strong interactions when they are truly strong, a problem which has eluded a basic understanding since strong interactions were first discovered. The CGC will be described below [2].

– **The study of the structure of the proton, most notably spin**
The structure of the proton and neutron is important as these particles form the ordinary matter from which we are composed. We would like to understand how valence quantum numbers such as baryon number, charge, and spin are distributed. RHIC has an active programme to study the spin of the proton [3].

Because I was asked to provide lectures on ultrarelativistic nuclear collisions, I shall discuss only the first two issues.

2.2 The quark–gluon plasma

This section describes the quark–gluon plasma, why it is important for astrophysics and cosmology, and why it provides a laboratory in which one can study the origin of mass and of confinement [1].

2.2.1 What is the quark–gluon plasma?

Matter at low energy densities is composed of electrons, protons, and neutrons. If we heat the system, we might produce thermal excitations which include light-mass strongly interacting particles such as the pion. Inside the protons, neutrons, and other strongly interacting particles are quarks and gluons. If we make the matter have high enough energy density, the protons, nucleons, and other particles overlap and get squeezed so tightly that their constituents are free to roam the system without being confined inside

Quark-Gluon Plasma $\longrightarrow$ **Hadron Gas**
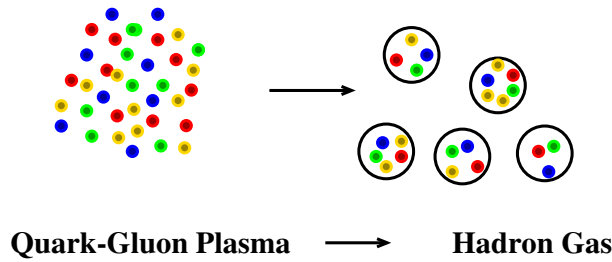
Fig. 1: As the energy density is decreased, the quark–gluon plasma condenses into a low-density gas of hadrons. Quarks are red, green, or blue and gluons are yellow.

hadrons [4]. At this density, there is deconfinement and the system is called a quark–gluon plasma. This is shown in Fig. 1.

As the energy density gets to be very large, the interactions between the quarks and gluons become weak. This is a consequence of the asymptotic freedom of strong interactions: at short distances the strong interactions become weak.

The QGP surely existed during the Big Bang. In Fig. 2, the various stages of evolution in the Big Bang are shown.
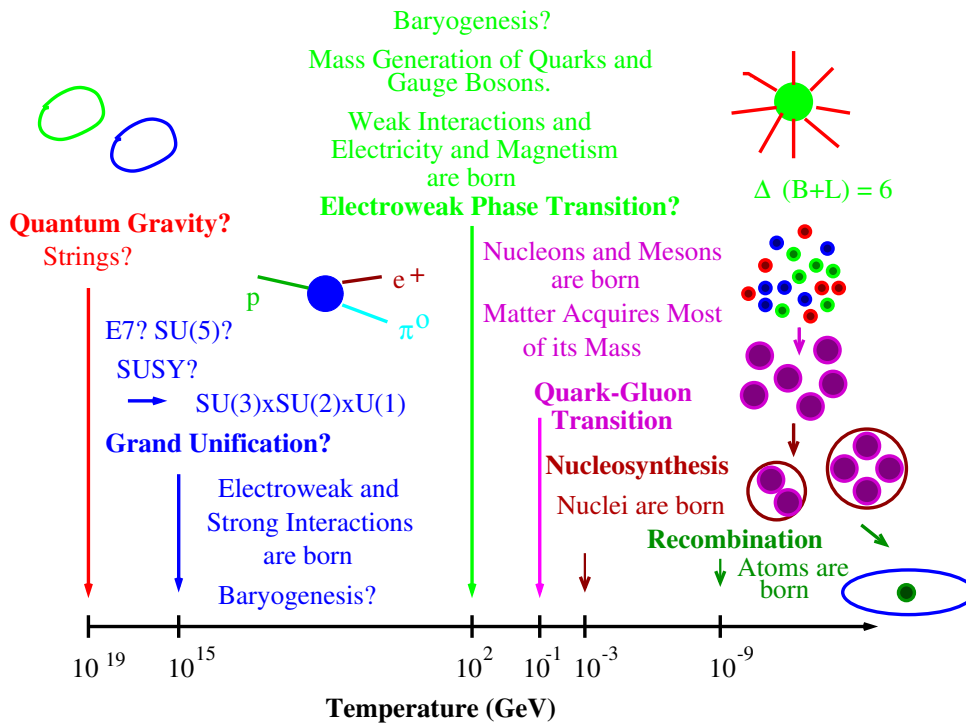
Fig. 2: The various forms of matter, and the types of physics which are probed during the Big Bang

At the earliest times in the Big Bang, temperatures are of order $T \sim 10^{19}$ GeV, quantum gravity is important, and despite the efforts of several generations of string theorists, we have little understanding. At somewhat lower temperatures, perhaps there is the grand unification of all the forces, except gravity. It might be possible that the baryon number of the universe is generated at this temperature scale. At much lower temperatures, of order $T \sim 100$ GeV, electroweak symmetry breaking takes place. It is possible here that the baryon asymmetry of the universe might be produced. At temperatures of order $T \sim 1$ GeV,

217

quarks and gluons become confined into hadrons. This is the temperature range appropriate for studies at RHIC and the LHC. At $T \sim 1$ MeV, the light elements are made. This temperature corresponds to an energy range which has been much studied, and is the realm of conventional nuclear physics. At temperatures of the order of an electronvolt, corresponding to the binding energies of electrons in atoms, the universe changes from an ionized gas to a lower-pressure gas of atoms, and structure begins to form.

The QGP is formed at energy densities of order 1 GeV/fm^3. Matter at such energy densities probably exists inside the cores of neutron stars as shown in Fig. 3. Neutron stars are objects of about
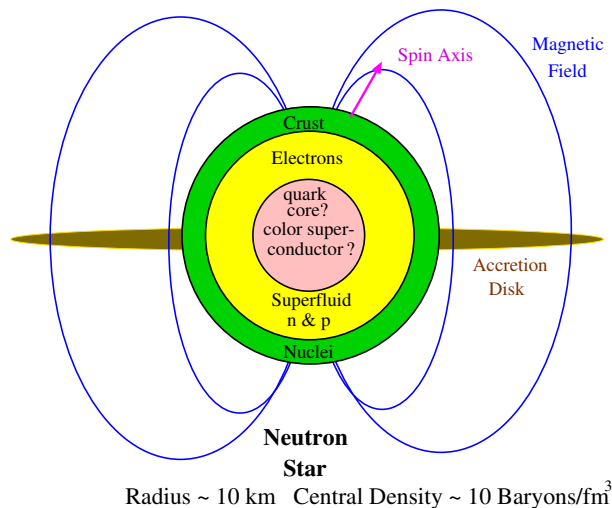
Fig. 3: A spinning neutron star

10 km in radius and are composed of extremely high energy density matter. The typical energy density in the core is of the order of 1 GeV/fm^3, and approaches zero at the surface. Unlike the matter in the Big Bang, this matter is cold and has temperature small compared to the Fermi energies of quarks. It is a cold, degenerate gas of quarks. At lower densities, this matter converts into a cold gas of nucleons.

Hot and dense matter with energy density of order 1 GeV/fm^3 may have occurred in the supernova explosion which led to the neutron star's formation. It may also occur in collisions of neutron stars and black holes, and may be the origin of the mysterious gamma-ray bursters. (Gamma-ray bursters are believed to be starlike objects which convert of the order of their entire mass into gamma rays.)

2.2.2 The quark–gluon plasma and ideal gases

At very high energy temperatures, the coupling constant of QCD becomes weak. A gas of particles should to a good approximation become an ideal gas. Each species of particle contributes to the energy density of an ideal gas as

$$\epsilon = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \sum_i \frac{E_i}{e^{\beta E_i} \pm 1} \tag{2}$$

where the $-$ is for bosons and the $+$ for fermions. The energy of each particle is E_i. At high temperatures, masses can be ignored, and the factor of ± 1 in the denominator turns out to make a small difference. One finds therefore that

$$\epsilon \sim \frac{\pi^2}{30} N T^4 \tag{3}$$

where N is the number of particle degrees of freedom. At low temperatures when masses are important, only the lowest mass, strongly interacting particle degree of freedom contributes; the pion, and the energy

density approaches zero as $\epsilon \sim e^{-m_\pi/T}$. For an ideal gas of pions, the number of pion degrees of freedom is three. For a QGP there are two helicities and eight colours for each gluon, and for each quark, three colours, two spins, and a quark–antiquark pair. The number of degrees of freedom is $N \sim 2 \times 8 + 4 \times 3 \times N_F$ where N_F is the number of important quark flavours, which is about three if the temperature is below the charm quark mass so that $N \sim 50$.

There is about an order of magnitude change in the number of degrees of freedom between a hadron gas and a QGP. This is because the degrees of freedom of the QGP include colour. In the large N_{colour} limit, the number of degrees of freedom of the plasma are proportional to N^2_{colour}, and in the confined phase is of order 1. In this limit, the energy density has an infinite discontinuity at the phase transition. There would be a limiting temperature for the hadronic world in the limit for which $N_{\text{colour}} \to \infty$, since at some temperature the energy density would go to infinity. This is the Hagedorn limiting temperature. (In the real world N_{colour} is three, and there is a temperature at which the energy density changes by an order of magnitude in a narrow range.)

2.2.3 *The quark–gluon plasma and fundamental physics issues*

The nature of matter at high densities is an issue of fundamental interest. Such matter occurred during the Big Bang, and it is the ultimate and universal state of matter at very high energy densities.

A hypothetical phase diagram for QCD is shown in Fig. 4. The vertical axis is temperature, and the horizontal is a measure of the matter or baryon number density, the baryon number chemical potential [5]. The solid lines indicate a first-order phase transition, and the dashed line a rapid cross-over. It is not known for sure whether or not the region marked cross-over is or is not a true first-order phase transition. There are analytic arguments for the properties of matter at high density, but numerical computations are of insufficient resolution. At high temperature and fixed baryon number density, there are both analytic arguments and numerical computations of good quality. At high density and fixed temperature, one goes into a superconducting phase, perhaps multiple phases of superconducting quark matter. At high temperature and fixed baryon number density, the degrees of freedom are those of a quark–gluon plasma.
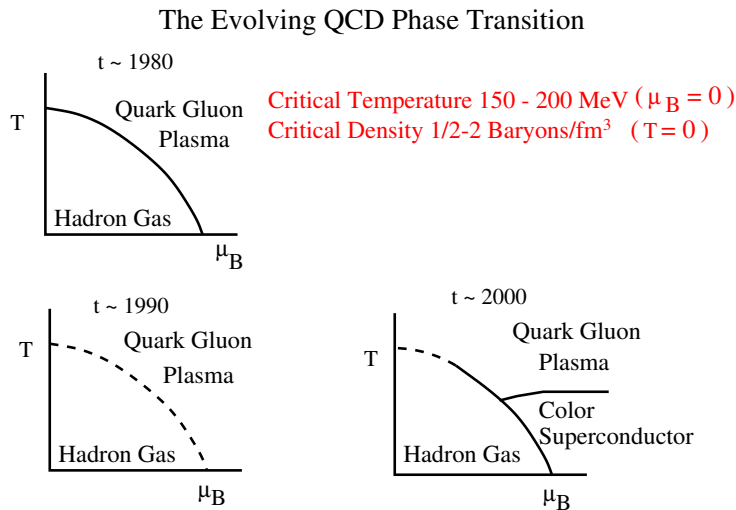
Fig. 4: A phase diagram for QCD collisions

I have shown this phase diagram as a function of time. What this means is that at various times people thought they knew what the phase diagram was. As time evolved, the picture changed. The latest ideas are marked with the date 2000. The point of doing this is to illustrate that theoretical ideas in the absence of experiment change with time. Physics is essentially an experimental science, and it is

very difficult to appreciate the richness which nature allows without knowing from experiment what is possible.

Much of the information we have about QCD at finite energy density comes from lattice gauge theory numerical simulation [5]. To see how lattice gauge theory works, recall that at finite temperature, the grand canonical ensemble is given by

$$Z = \mathrm{Tr}\, e^{-\beta H} \,. \tag{4}$$

This is similar to computing

$$Z = \langle e^{-itH} \rangle \tag{5}$$

where $-it = \beta$. That is we compute the expectation value of the time evolution operator for imaginary time. This object has a path integral representation, which has been described to you in your elementary field theory text books. Under the change of variables, the action becomes $iS = i \int \mathrm{d}t L \rightarrow S = -\int_0^\beta \mathrm{d}\tau L$. Here L is the Lagrangian.

The grand canonical ensemble has the representation

$$Z = \int [dA] e^{-S[A]} \tag{6}$$

for a system of pure gluons. The gluon fields satisfy periodic boundary conditions due to the trace in the definition of the grand canonical ensemble. (Fermions may also be included, although the path integral is more complicated, and the fermion fields are required to satisfy antiperiodic boundary conditions.) Expectation values are computed as

$$\langle 0 \rangle = \frac{\mathrm{Tr}\, O e^{-\beta H}}{\mathrm{Tr}\, e^{-\beta H}} \,. \tag{7}$$

The way that lattice Monte Carlo simulates the grand canonical ensemble is by placing all of the fields on a finite grid, so the path integral becomes finite dimensional. Then field configurations are selectively sampled, as weighted by their action. This works because the factor of $e^{-\beta H}$ is positive and real. (The method has essential complications for finite density systems, since there the action becomes complex.)

Lattice gauge theory numerical studies, and analytic studies have taught us much about the properties of these various phases of matter [5]. There have been detailed computations of the energy density as a function of temperature. In Fig. 5 the energy density scaled by T^4 is plotted. This is essentially the number of degrees of freedom of the system as a function of T. At a temperature of $T_c \sim 160\text{–}190$ MeV the number of degrees of freedom changes very rapidly, possibly discontinuously. This is the location of the transition from the hadron gas to the quark–gluon plasma.

In Fig. 6, the sound velocity is plotted as a function of temperature. The sound velocity increases at high temperature asymptoting to its ideal gas value of $v_{\mathrm{sound}}^2 \sim 1/3$. Near the phase transition, it becomes very small. This is because the energy density jumps at the transition temperature, but the pressure must be smooth and continuous. The sound velocity squared is $\mathrm{d}P/\mathrm{d}\epsilon$.

Lattice Monte Carlo simulation has also been used to study how the phase transition is related to the confining force. In a theory with only gluons, the potential for sources of fundamental representation colour charge grows linearly in the confined phase. (With dynamical fermions, the potential stops rising at some distance when it is energetically favourable to produce quark–antiquark pairs which short out the potential.)

We can understand how confinement might disappear at high temperature. At finite temperature, there is a symmetry of the pure gluon Yang–Mills system. Consider a Wilson line which propagates from
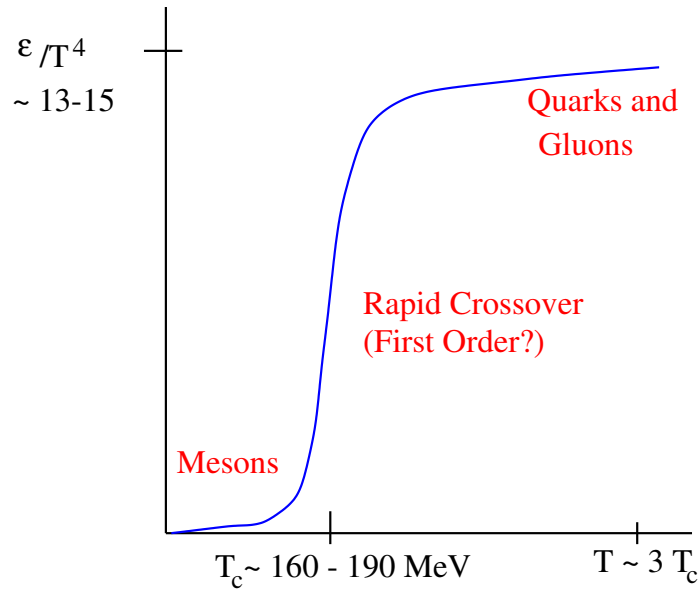
Fig. 5: The energy density scaled by T^4 as a function of temperature
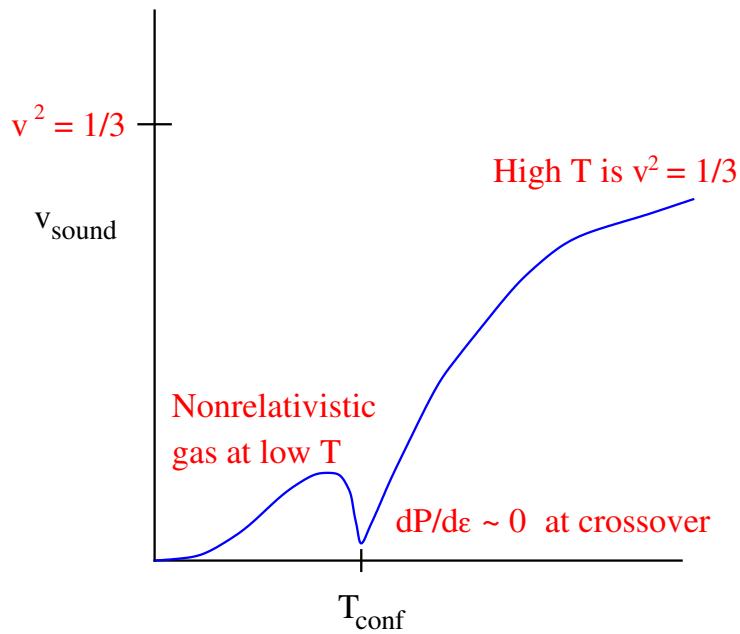
Fig. 6: The sound velocity as a function of temperature

$(0, \vec{x})$ to the point $(\beta, \vec{x})$. A Wilson line is a path-ordered phase,

$$L(x) = P \exp i \int_0^\beta dt\, A^0(t, \vec{x}) . \tag{8}$$

One can show that the expectation value of this line gives the free energy of an isolated quark:

$$e^{-\beta F} = \frac{1}{N_c} \langle \text{Tr}\, (L(x)) \rangle . \tag{9}$$

Now consider gauge transformations which maintain the periodic boundary conditions on the gauge

fields (required by the trace in the definition of the grand canonical ensemble). The most general gauge transformation which does this is not periodic but solves

$$U(\beta, \vec{x}) = ZU(0, \vec{x}) \ . \tag{10}$$

One can show that $[z, \tau^a] = 0$, and that $\nabla^i Z = 0$. Z is an element of the gauge group so that $\det Z = 1$. These conditions require that

$$Z = e^{2\pi i j/N_c} \ . \tag{11}$$

This symmetry under non-periodic gauge transformations is global, that is it does not depend upon the position in space. It may be broken. If it is realized, the free energy of a quark must be infinite since $L \to ZL$ under this transformation, and $\langle L \rangle = 0$. If the symmetry is broken, quarks can be free.

Lattice gauge computations have measured the quark–antiquark potential as a function of T, and at the deconfinement temperature, the potential changes from linear at infinity to constant. This is shown in Fig. 7.
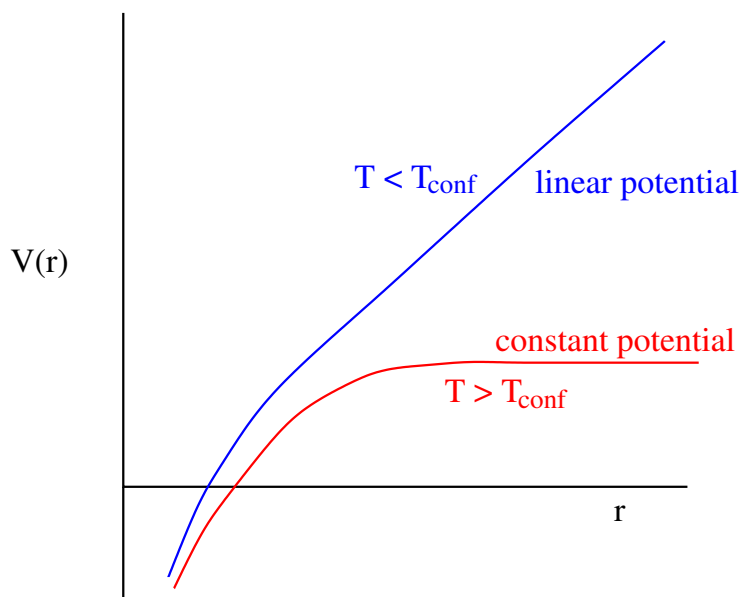
Fig. 7: The potential in pure gauge theory as a function of temperature

In addition to confinement–deconfinement, there is an additional symmetry which might occur at high temperatures. In nature, the up and down quark masses are almost zero. This leads to a chiral symmetry, which is the rotation of fermion fields by $e^{i\gamma_5\theta}$. This symmetry would require that either baryons are massless or occur in parity doublets. Neither arises in nature. The nucleon has a mass of about 1 GeV and has no opposite parity partner of almost equal mass. It is believed that this symmetry becomes broken, and as a consequence, the nucleon acquires mass, and that the pion becomes an almost massless Goldstone boson. It turns out that at the confinement–deconfinement phase transition, chiral symmetry is restored. This is seen in Fig. 8 where a quantity proportional to the nucleon mass is plotted as a function of T.

The chiral symmetry restoration phase transition can have interesting dynamical consequences. In the confined phase, the mass of a nucleon is of order $N_c\Lambda_{QCD}$, but in the deconfined phase is of order T. Therefore in the confined phase, the Boltzman weight $e^{-M/T}$ is very small. Imagine what happens as we go through the phase transition starting at a temperature above T_c. At first the system is entirely
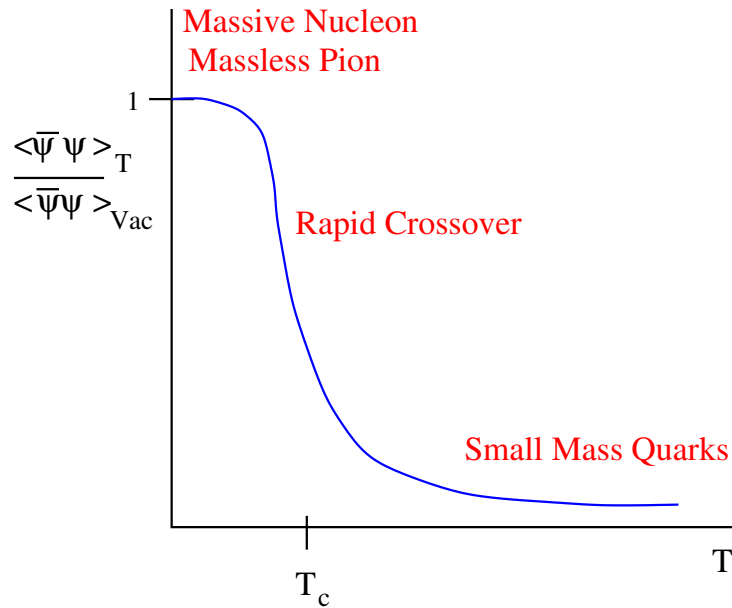
Fig. 8: The chiral order parameter $\langle \overline{\Psi} \Psi \rangle$ as a function of temperature

in QGP. As the system expands, a mixed phase of droplets of QGP and droplets of hadron gas forms. The nucleons like to stay in the QGP because their Boltzman weight is larger. As the system expands further, the droplets of QGP shrink, but most of the baryon number is concentrated in them. At the end of the mixed phase, one has made large-scale fluctuations in the baryon number. This scenario is shown in Fig. 9.

The confinement–deconfinement phase transition and the chiral symmetry restoration phase transition might be logically disconnected from one another. The confinement–deconfinement phase transition is related to a symmetry when the quark masses are infinite. The chiral transition is related to a symmetry when the quarks are massless. As a function of mass, one can follow the evolution of the phase transitions. At large and small masses there is a real phase transition marked by a discontinuity in physical quantities. At intermediate masses, there is probably a rapid transition, but not a real phase transition. It is believed that the real world has masses which make the transition closer to a cross-over than a phase transition, but the evidence from lattice Monte Carlo studies is very weak. In Fig. 10, the various possibilities are shown.

2.3   The colour glass condensate

This section describes the colour glass condensate, and why it is important for our understanding of basic properties of strong interactions [2], [6]. I argue that the colour glass condensate is a universal form of matter which controls the high-energy limit of all strong interaction processes and is the part of the hadron wavefunction important at such energies. Since the colour glass condensate is universal and controls the high-energy limit of all strong interactions, it is of fundamental importance.

2.3.1   *What is the colour glass condensate?*

The colour glass condensate is a new form of matter which controls the high-energy limit of strong interactions. It is universal and independent of the hadron which generated it. It should describe

– high-energy cross-sections
– distributions of produced particles

Macroscopic Baryon Number Fluctuations
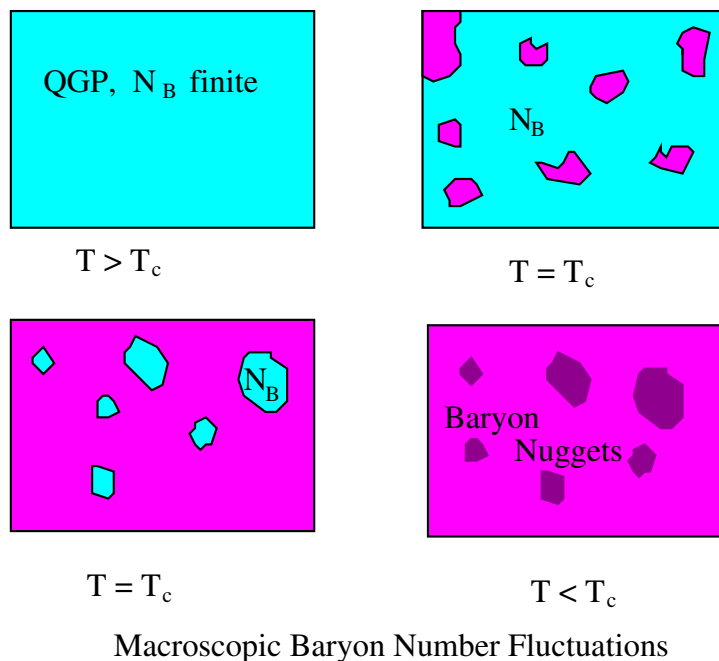
Fig. 9: Formation of large-scale baryon number fluctuations at the QCD phase transition

– the distribution of the small-x particles in a hadron
– initial conditions for heavy-ion collisions

A very-high-energy hadron has contributions to its wavefunction from gluons, quarks, and anti-quarks with energies up to that of the hadron and all the way down to energies of the order of the scale of light-mass hadron masses, $E \sim 200$ MeV. A convenient variable in which to think about these quark degrees of freedom is the typical energy of a constituent scaled by that of the hadron,

$$x = E_{\text{constituent}} / E_{\text{hadron}} . \tag{12}$$

Clearly the higher the energy of the hadron we consider, the lower the minimum x of a constituent. Sometimes it is also useful to consider the rapidity of a constituent which is $y \sim \ln(1/x)$.

The density of small-x partons is

$$\frac{\mathrm{d}N}{\mathrm{d}y} = xG(x, Q^2) . \tag{13}$$

The scale Q^2 appears because the number of constituents one measures depends (weakly) upon the resolution scale of the probe with which one measures. (Resolution scales are measured in units of the inverse momentum of the probe, which is usually taken to be a virtual photon.) A plot of $xG(x, Q^2)$ for gluons at various x and Q^2 measured at the HERA accelerator in protons [7] is shown in Fig. 11.

Note that the gluon density rises rapidly at small x in Fig. 11. This is the so-called small-x problem. It means that if we view the proton head-on at increasing energies, the low-momentum gluon density grows. This is shown in Fig. 12.

As the density of gluons per unit area per unit rapidity increases, the typical transverse separation of the gluons decreases. This means that the matter which controls high-energy strong interactions is very dense, and it means that the QCD interaction strength which is usually parametrized by the dimensionless scale α_s becomes small. The phase space density of these gluons $\rho \sim 1/\pi R^2 \, \mathrm{d}N/\mathrm{d}^2 p_T$ can become at

T_c

Confinement
Transition

Crossover?
Real Transition?

M_{quark}

Chiral Transition

T_{chiral}

T
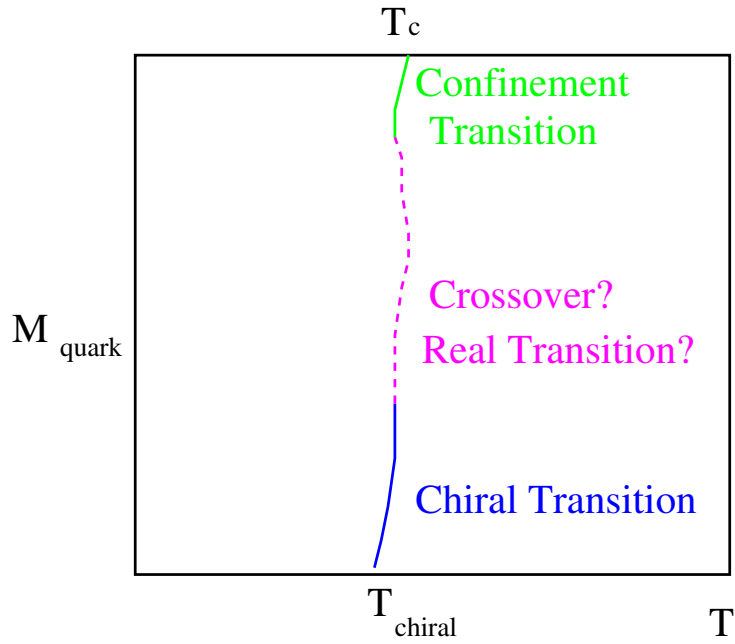
The Pisarski Plot

Fig. 10: The phase diagram of QCD as a function of fermion mass

$Q^2 = 20 \text{ GeV}^2$

$Q^2 = 200 \text{ GeV}^2$

$xG(x,Q^2)$

$Q^2 = 5 \text{ GeV}^2$

10^{-4} 10^{-3} 10^{-2} 10^{-1} x
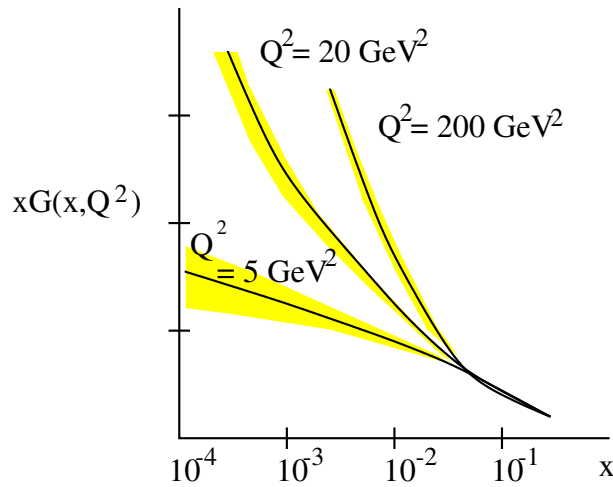
Fig. 11: The number of gluons in a proton per unit rapidity at various rapidities and Q^2 resolutions

most $1/\alpha_s$ since once this density is reached gluon interactions are important. This is characteristic of Bose condensation phenomena which are generated by an instability proportional to the density ρ and are compensated by interactions proportional to $\alpha_s \rho^2$, which become of the same order of magnitude when $\rho \sim 1/\alpha_s$. Thus the matter is a colour condensate.

The glassy nature of the condensate arises because the fields associated with the condensate are generated by constituents of the proton at higher momentum. These higher momentum constituents have their time scales Lorentz time dilated relative to those which would be measured in their rest frame. Therefore the fields associated with the low-momentum constituents also evolve on this long time scale. The low-momentum constituents are therefore glassy: their time evolution scale is unnaturally long
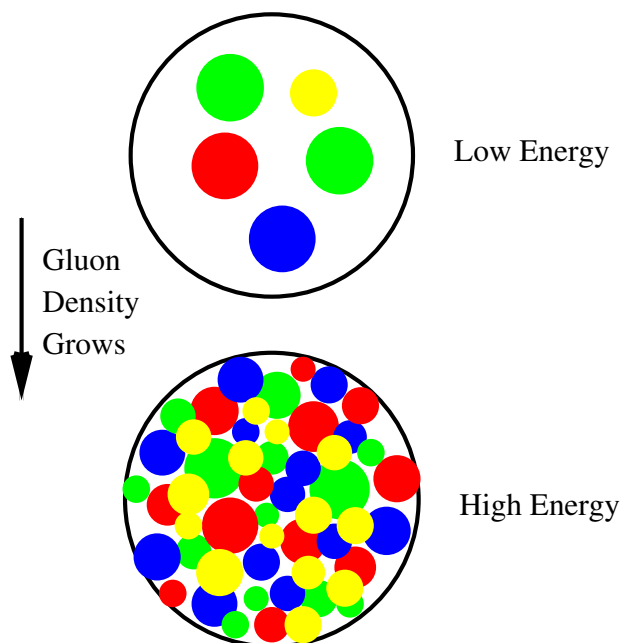
Fig. 12: The increasing density of wee partons as the energy increases

compared to their natural time scale. Hence the name colour glass condensate.

There is also a typical scale associated with the colour glass condensate: the saturation momentum. This is the typical momentum scale where the phase-space density of gluons becomes $\rho \leq 1/\alpha_{\mathrm{s}}$.

At very high momentum, the fields associated with the colour glass condensate can be treated as classical fields, like the fields of electricity and magnetism. Since they arise from fast moving partons, they are plane polarized, with mutually orthogonal colour electric and magnetic fields perpendicular to the direction of motion of the hadron. They are also random in two dimensions. This is shown in Fig. 13.
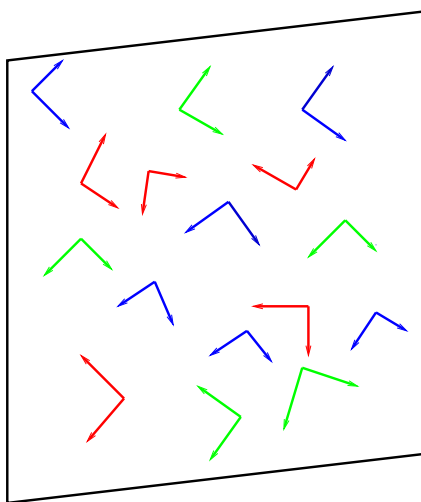
Fig. 13: The colour glass condensate as a high density of random gluon fields on a two-dimensional sheet travelling near the speed of light

2.3.2 Why is the colour glass condensate important?

Like nuclei and electrons compose atoms, and nucleons and protons compose nuclear matter, the colour glass condensate is the fundamental matter of which high-energy hadrons are composed. The colour glass condensate has the potential to allow for a first-principles description of the gross or typical properties of matter at high energies. For example, the total cross-section at high energies for proton–proton scattering, as shown in Fig. 14, has a simple form but for over 40 years has resisted simple explanation. (It has perhaps been understood recently in terms of the colour glass condensate or saturation ideas [8–11].)
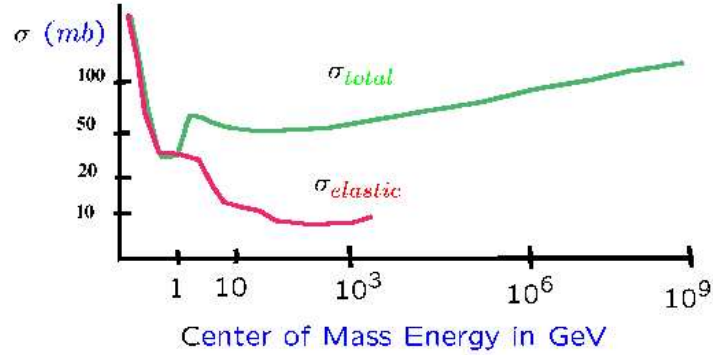
Fig. 14: The total cross-section for high-energy proton–proton interactions

The colour glass condensate forms the matter in the quantum mechanical state which describes a nucleus. In the earliest stages of nucleus–nucleus collisions, the matter must not be changed much from these quantum mechanical states. The colour glass condensate therefore provides the initial conditions for the quark–gluon plasma to form in these collisions. A space–time picture of nucleus–nucleus collisions is shown in Fig. 15. At very early times, the colour glass condensate evolves into a distribution of gluons. Later these gluons thermalize and may eventually form a quark–gluon plasma. At even later times, a mixed phase of plasma and hadronic gas may form.
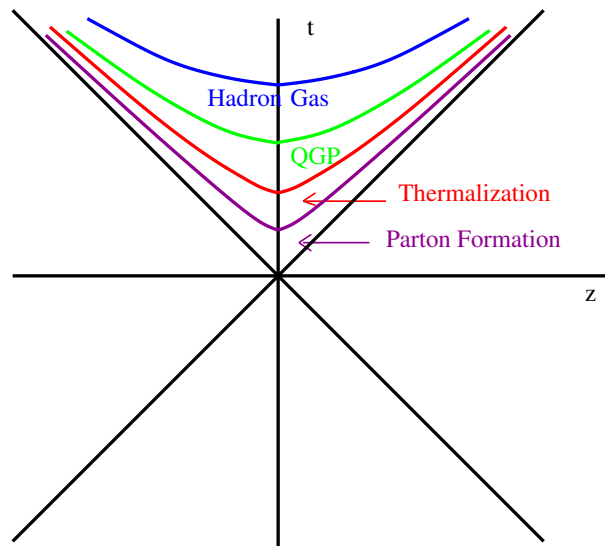
Fig. 15: A space–time diagram for the evolution of matter produced in heavy-ion collisions

3 Lecture II: Ultrarelativistic nuclear collisions

Heavy-ion collisions at ultrarelativistic energies are visualized in Fig. 16 as the collision of two sheets of coloured glass [12].

At ultrarelativistic energies, these sheets pass through one another. In their wake is left melting coloured glass, which eventually materializes as quarks and gluons. These quarks and gluons would naturally form in their rest frame on some natural microphysics time scale. For the saturated colour glass, this time scale is of the order of the inverse saturation momentum (again, we convert momentum into time by appropriate uses of Planck's constant and the speed of light), in the rest frame of the produced particle. When a particle has a large momentum along the beam axis, this time scale is Lorentz dilated. This means that the slow particles are produced first towards the centre of the collision regions and the fast particles are produced later further away from the collision region.
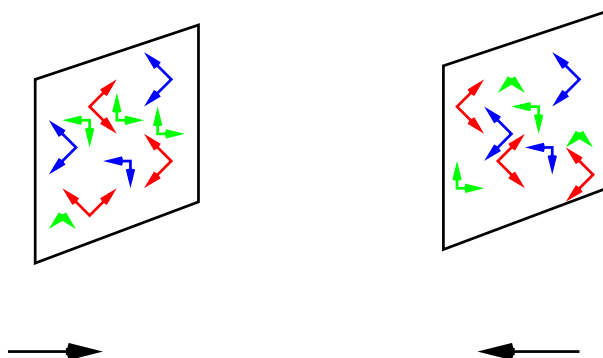
Fig. 16: The collision of two sheets of coloured glass

This correlation between space and momentum is similar to what happens to matter in Hubble expansion in cosmology. The stars which are further away have larger outward velocities. This means that this system, like the universe in cosmology is born expanding. This is shown in Fig. 17.

large p   small p   large p
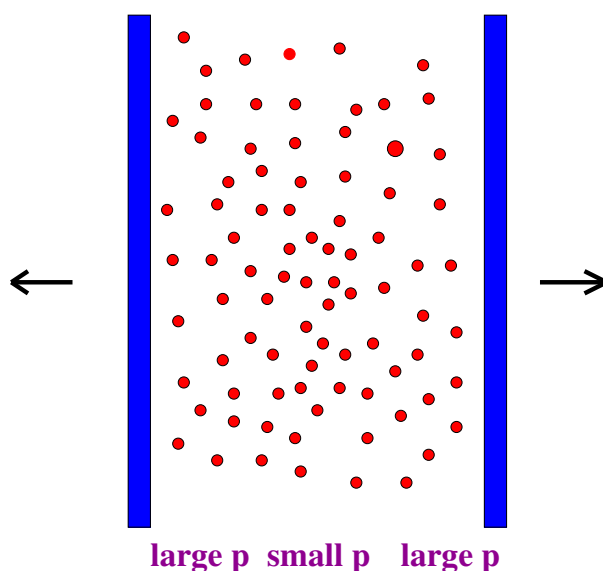
Fig. 17: Particles being produced after the collision of two nuclei

As this system expands, it cools. Presumably at some time the produced quarks and gluons ther-

malize. They then expand as a quark–gluon plasma and eventually as some mixture of hadrons and quarks and gluons. Finally, they may become a gas of only hadrons before they stop interacting and fly off to detectors.

In the last lecture, we shall describe the results from nucleus–nucleus collisions at RHIC in some detail. Before proceeding there, we need to learn a little bit more about the properties of high-energy hadrons. It is useful to introduce some kinematic variables which are useful in what will follow.

The light cone momenta are defined as

$$P^{\pm} = \frac{1}{\sqrt{2}}(E \pm p_z) \tag{14}$$

and light cone coordinates are

$$X^{\pm} = \frac{1}{\sqrt{2}}(t \pm z) . \tag{15}$$

The metric in these variables is

$$p \cdot x = p^+ x^- + p^- x^+ - p_{\mathrm{T}} \cdot x_{\mathrm{T}} . \tag{16}$$

Conjugate variables are $x^{\pm}\langle-\rangle p^{\mp}$. The square of the four momentum is

$$p^2 = 2P^+ P^- - P_{\mathrm{T}}^2 = M^2 . \tag{17}$$

The uncertainty principle is

$$\Delta x^{\pm} \Delta p^{\mp} \geq 1 . \tag{18}$$

Light cone variables are useful because in a high-energy collision, a left-moving particle has $p_z \sim E$, so that $p^+ \sim \sqrt{2}E$, but $p^- \sim m_{\mathrm{T}}^2/p_z \sim 0$. For the right-moving particles, it is p^- which is big and p^+ which is very small.

Light cone variables scale by a constant under Lorentz transformations along the collision axis. Ratios of light cone momentum are therefore invariant under such Lorentz boosts. The light cone momentum fraction $x = p_i^+/P^+$, where P^+ is that of the particle we probe and p_i^+ is that of the constituent of the probed hadron, satisfies $0 \leq x \leq 1$. It is the same as Bjorken x, and for a fast-moving hadron, it is almost Feynman $x_{\mathrm{Feynman}} = E_i/E$. This is the x variable one is using when one describes deep inelastic scattering. In this case the label i corresponds to a quark or gluon constituent of a hadron.

One can also define a rapidity variable:

$$y = \frac{1}{2}\ln\left\{\frac{p_i^+}{p_i^-}\right\} \sim \ln(2E_i/M_{\mathrm{T}}) . \tag{19}$$

Up to mass effects, the rapidity is in the range $-y_{\mathrm{proj}} \leq y \leq y_{\mathrm{proj}}$. When particles, like pions, are produced in high-energy hadronic collisions, one often plots them in terms of the rapidity variable. Distributions tend to be slowly varying functions of rapidity.

3.1 Is there simple behaviour at high energy?

A hint of the underlying simplicity of high-energy hadronic interactions comes from studying the rapidity distributions of produced particles for various collision energies. In Fig. 18, a generic plot of the rapidity distribution of produced pions is shown for two different energies. The rapidity distribution at lower energies has been cut in half and the particles associated with each of the projectiles have been displaced in rapidity so that their staring points in rapidity are the same. It is remarkable that, except for the slowest
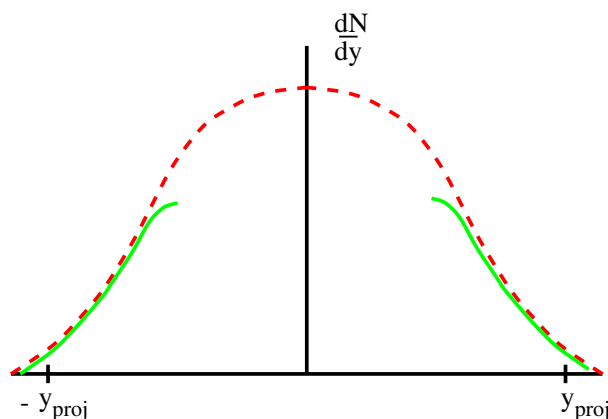
Fig. 18: The rapidity distributions of particles at two different energies
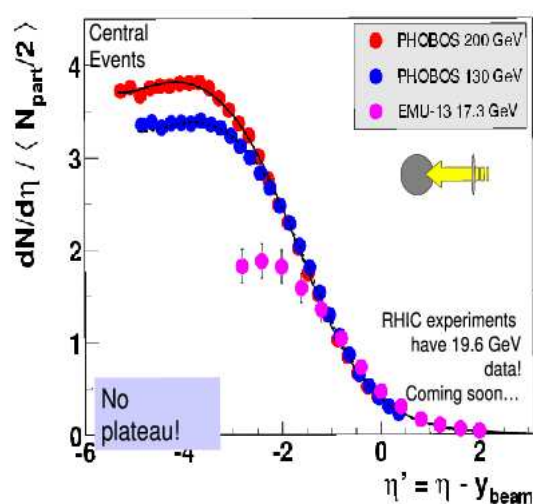
Fig. 19: Experimental evidence from the PHOBOS experiment at RHIC on limiting fragmentation

particles in the centre-of-mass frame, those located near $y \sim 0$, the distributions are almost identical [13]. This is shown for the data from RHIC in Fig. 19.

We conclude from this that going to higher energy adds in new degrees of freedom, the small-x part of the hadron wavefunction. At lower energies, these degrees of freedom are not kinematically relevant as they can never be produced. On the other hand, going to higher energy leaves the fast degrees of freedom of the hadron unchanged.

This suggests that there should be a renormalization group description of the hadrons. As we go to higher energy, the high-momentum degrees of freedom remain fixed. Integrating out the previous small-x degrees of freedom should incorporate them into what are now the high-energy degrees of freedom at the higher energy. This process generates an effective action for the new low-momentum degrees of freedom. Such a process, when done iteratively, is a renormalization group.

3.2 A single hadron

A plot of the rapidity distribution of the constituents of a hadron, the gluons, is shown generically in Fig. 20. I have used $y = y_{\text{hadron}} - \ln(1/x)$ as my definition of rapidity. This distribution is similar in
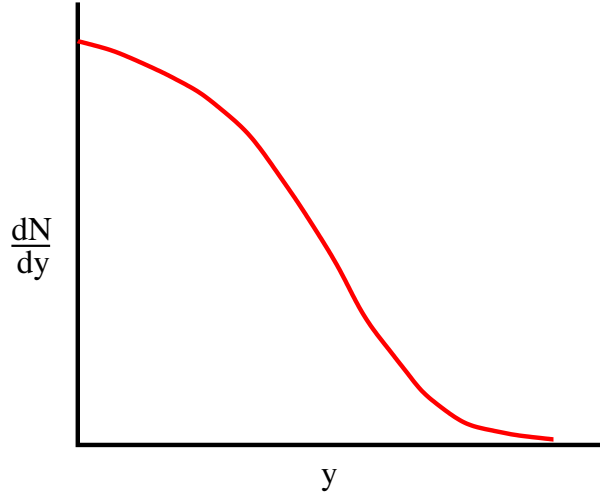
Fig. 20: The rapidity distribution of the constituents of a hadron

shape to that of the half of the rapidity distribution shown for hadron–hadron interactions in the centre-of-mass frame which has positive rapidity. The essential difference is that this distribution is for constituents where the hadron–hadron collision is for produced particles, mainly pions.

In the high-energy limit, as discussed in the previous section, the density of gluons grows rapidly. This suggests we introduce a density scale for the partons

$$\Lambda^2 = \frac{1}{\pi R^2} \frac{dN}{dy} \ . \tag{20}$$

One usually defines a saturation momentum to be $Q_{sat}^2 \sim \alpha_s \Lambda^2$, since this will turn out to be the typical momentum of particles in this high-density system. In fact, α_s is slowly varying compared to the variation of λ, so that for the purposes of the estimates we make here, whether or not there is a factor of α_s will not be so important. Note that α_s evaluated at the saturation scale will be $\alpha_s \ll 1$. The typical particle transverse momenta are of order $p_T^2 \sim Q_{sat}^2 \gg 1/R_{had}^2$. Therefore it is consistent to think of the parton distribution as a high-density, weakly coupled system which is localized in the transverse plane. The high-momentum partons, the degrees of freedom which should be frozen, can be thought of as sitting on an infinitesimally thin sheet. We shall study this system with a resolution size scale which is $\Delta x \ll 1/\Lambda_{QCD}$, so that we may use weak coupling methods. Such a thin sheet is shown in Fig. 21.

It is useful to discuss different types of rapidity variables before proceeding. The typical momentum space rapidity is

$$\begin{aligned}
y &= \frac{1}{2} \ln\left(\frac{p^+}{p^-}\right) \\
&= \ln\left(\frac{2p^+}{M_T}\right) \\
&= \ln\left(\frac{2p_{hadron}^+}{M_T}\right) + \ln\left(\frac{p^+}{p_{hadron}^+}\right) \\
&= y_{hadron} - \ln(1/x) \ . \tag{21}
\end{aligned}$$

Here M_T is a particle transverse mass, and we have made approximations which ignore overall shifts in rapidity by of order one unit. Within these approximations, the momentum space rapidity used to describe the production of particles is the same as that used to describe the constituents of hadrons.

231
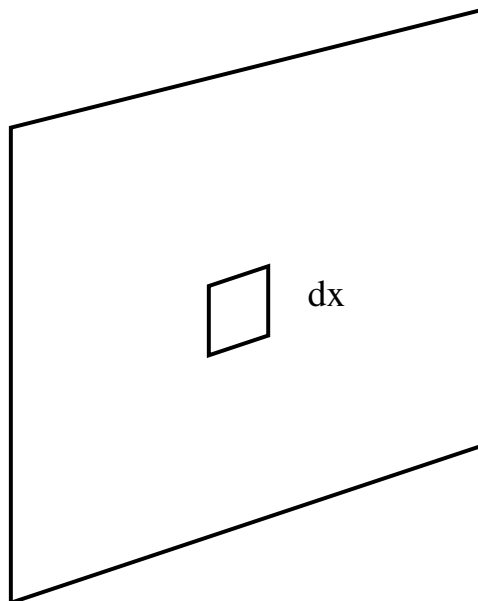
Fig. 21: A thin sheet travelling near light velocity. The transverse resolution scale is Δx.

Oftentimes a coordinate-space rapidity is introduced. With $\tau = \sqrt{t^2 - z^2}$,

$$y = \frac{1}{2} \ln \left(\frac{x^+}{x^-} \right) = \ln(2\tau/x^-) . \tag{22}$$

Taking τ to be a time scale of order $1/M_{\mathrm{T}}$, and using the uncertainty principle $x^\pm \sim 1/p^\mp$, we find that up to shifts in rapidity of order one, all the rapidities are the same. This implies that coordinate space and momentum space are highly correlated, and that one can identify momentum-space and coordinate-space rapidity with some uncertainty of order one unit.

If we plot the distribution of particles in a hadron in terms of the rapidity variable, the longitudinal dimension of the sheet is spread out. This is shown in Fig. 22. The longitudinal position is correlated

y_{max} y_{min}
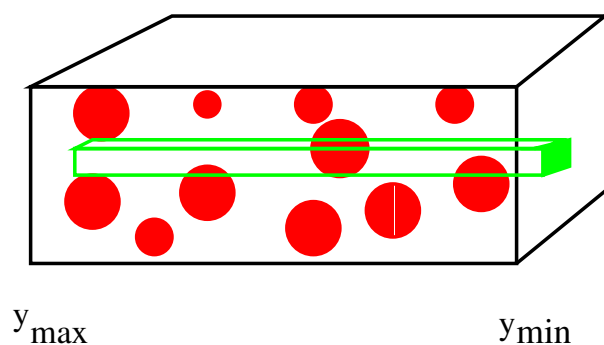
Fig. 22: The distribution of particles in a hadron in terms of rapidity variables

with the longitudinal momentum. The highest-rapidity particles are the fastest. In ordinary coordinate space, this means the fastest particles are those most Lorentz contracted. If we now look down a tube of transverse size $\Delta x << 1/\Lambda_{\mathrm{QCD}}$, we intersect the various constituents of the hadron only occasionally. The colour charge probed by this tube should therefore be random, until the transverse size scale becomes large enough so that it can probe the correlations. If the beam energy is large enough, or x is small

enough, there should be a large amount of colour charge in each tube of fixed size Δx. One can therefore treat the colour charge classically.

The physical picture we have generated is that there should be classical sources of to a good approximation random charges on a thin sheet. The current for this is

$$J_a^\mu = \delta^{\mu+}\delta(x^-)\rho_a(x_{\mathrm{T}}) . \tag{23}$$

The delta function approximation should be good for many purposes, but it may also be useful in some circumstances to insert the longitudinal structure

$$J_a^\mu = \delta^{\mu+}\rho_a(y, x_{\mathrm{T}}) \tag{24}$$

and to remember that the support of the source is for very large y.

3.3 The colour glass condensate

We now know how to write down a theory to describe the colour glass condensate. It is given by the path integral [6]

$$\int [dA][d\rho]\exp\left(iS[A, \rho] - W[\rho]\right) . \tag{25}$$

Here $S[A, \rho]$ is the Yang–Mills action in the presence of a source current as described above. The function W weights the various configurations of colour charge. In the simplest version of the colour glass condensate, this can be taken to be a Gaussian

$$W = \frac{1}{2}\int \mathrm{d}y\mathrm{d}^2x_{\mathrm{T}}\,\frac{\rho^2(y, x_{\mathrm{T}})}{\mu^2(y)} . \tag{26}$$

In this ansatz, $\mu^2(y)$ is the colour charge squared density per unit area per unit y scaled by $1/N_c^2 - 1$. The theory can be generalized to less trivial forms of the weight function, but this form works at small transverse resolution scales, $\Delta x \ll 1/Q_{\mathrm{sat}}$. As one increases the transverse resolution scale one needs a better determination of W. It turns out that at resolution scales of order $1/Q_{\mathrm{sat}} \ll \Delta x \ll 1/\Lambda_{\mathrm{QCD}}$, a Gaussian form is still valid.

The averaging over an external field makes the theory of the colour glass condensate similar to that of spin glasses. The solutions of the classical field equations also have $F^2 \sim 1/\alpha$, so the gluon fields are strong and have high occupation number, hence the word condensate.

The theory described above has an implicit longitudinal momentum cutoff scale. Particles with momentum above this scale are treated as sources, and those below it as fields. One computes physical quantities by first computing the classical fields and then averaging over sources ρ. This is a good approximation so long as the longitudinal momentum in the field is not too far below the longitudinal momentum cutoff Λ^+. If one computes quantum corrections, the expansion parameter is

$$\alpha_{\mathrm{s}}\,\ln(\Lambda^+/p^+) . \tag{27}$$

To generate a theory at smaller momenta $\overline{\Lambda}^+$ one first requires that $\alpha_{\mathrm{s}}\,\ln(\Lambda^+/\overline{\Lambda}^+) \ll 1$. Then one computes the quantum corrections in the presence of the background field. This turns out to change only the weight function W. Therefore the theory maps into itself under a change of scale. This is a renormalization group, and it determines the weight function W [6], [14, 15].

3.4 Colour glass fields

The form of the classical fields is easily inferred. On either side of the sheet the fields are zero. They have no time dependence, and in light cone gauge $A^+ = 0$. It is plausible to look for a solution which is purely transverse. On either side of the sheet, we have fields which are gauge transformations of zero field. It can be a different gauge transformation of zero field on different sides of the sheet. Continuity requires that $F^{ij} = 0$. F^{i-} is zero because of light cone time x^+ independence, and the assumption that $A^- = 0$. F^{i+} is non-zero $\sim \delta(x^-)$ because of the variation in x^- as one crosses the sheet. This means that $F^{i0} \sim -F^{iz}$, or that $E \perp B \perp \vec{z}$. These are transversely polarized Weiszacker–Williams fields. They are random in the two-dimensional plane because the source is random. This is shown in Fig. 13. The intensity of these fields is of order $1/\alpha_s$, and they are not at all stringlike.

3.5 The gluon distribution and saturation

The gluon distribution function is given by computing the expectation value of the number operator $\langle a^\dagger(p) a(p) \rangle$ and can be found from computing the gluon field expectation value $\langle A(p) A(-p) \rangle$. This is left as an exercise for the student. At large p_T, the distribution function scales as

$$\frac{\mathrm{d}N}{\mathrm{d}y \mathrm{d}^2 p_T} \sim \frac{1}{\alpha_s} \frac{Q_{\mathrm{sat}}^2}{p_T^2} \tag{28}$$

which is typical of a bremsstrahlung spectrum. At small p_T, the solution is $\sim \ln(Q_{\mathrm{sat}}^2/p_T^2)/\alpha_s$. The reason for this softer behaviour at smaller p_T is easy to understand. At small distances corresponding to large p_T, one sees point sources of charge, but at smaller p_T, the charges cancel one another and lead to a dipole field. The dipole field is less singular at large r, and when transformed into momentum space, one loses two powers of momentum in the distribution function. In terms of the colour field, the saturation phenomena is almost trivial to understand. (It is very difficult to understand if the gluons are treated as incoherently interacting particles.)

Now Q_{sat}^2 can grow with energy. In fact it turns out that Q_{sat}^2 never stops growing. The intrinsic transverse momentum grows without bound. Physically what is happening is that the low-momentum degrees of freedom below the saturation momentum grow very slowly, like $\ln(Q_{\mathrm{sat}}^2)$ because repulsive gluon interactions prevent more filling. On the other hand, one can always add more gluons at high momentum since the phase space is not filled there.

How is this consistent with unitarity? Unitarity is a statement about cross-sections at fixed Q^2. If Q^2 is above the saturation momentum, then the gluon distribution function grows rapidly with energy, as Q_{sat}^2. On the other hand, once the saturation momentum becomes larger than Q^2, the number of gluons one can probe

$$xG(x, Q^2) \sim \pi R^2 \int_0^{Q^2} \mathrm{d}^2 p_T \frac{\mathrm{d}N}{\mathrm{d}^2 p_T \mathrm{d}y} \tag{29}$$

varies only logarithmically. The number of gluons scales as the surface area. (At high Q^2, it is proportional to $R^2 Q_{\mathrm{sat}}^2$, and one expects that $Q_{\mathrm{sat}}^2 \sim A^{1/3}$ so that $xG(x, Q^2) \sim A$.)

3.6 Hadron collisions

In Fig. 16, the collision of two hadrons is represented as that of two sheets of coloured glass. Before the collisions, the left-moving hadron has fields

$$\begin{aligned}
F^{i+} &\sim \delta(x^-) \\
F^{ij} &\sim 0 \\
F^{i-} &\sim 0
\end{aligned} \tag{30}$$

and that of the right-moving fields is analogous to that of the above save that $\pm \rightarrow \mp$ in the indices and coordinates of all fields. The fields are of course different in each nucleus. We shall consider impact-parameter-zero head-on collisions in what follows.

These fields are plane polarized and have random colours. A solution of the classical Yang–Mills equation can be constructed by requiring that the fields be two-dimensional gauge transforms of vacuum everywhere but in the forward light cone. At the edges of the light cone, and at its tip $t = z = 0$, the equations are singular, and a global solution requires that the fields carry non-trivial energy and momentum in the forward light cone. At short times, these fields are highly non-linear. In a time of order $\tau \sim 1/Q_{\text{sat}}$, the fields linearize. When they linearize, we can identify the particle content of the classical radiation field.

This situation is much different than the case for quantum electrodynamics. Because of the gluon self-interaction, it is possible to classically convert the energy in the incident nuclei into radiation. In quantum electrodynamics, the charged particles are fermions, and they cannot be treated classically. Radiation is produced by annihilation or bremsstrahlung as quantum corrections to the initial value problem.

The solution to the field equation in the forward light cone is approximately boost invariant over an interval of rapidity of order $\Delta y \ll 1/\alpha_s$. At large momentum, the field equations can be solved in perturbation theory and the distribution is like that of a bremsstrahlung spectrum

$$\frac{dN}{dy d^2 p_{\text{T}}} \sim \frac{1}{\alpha_s} \pi R^2 \frac{Q_{\text{sat}}^4}{p_{\text{T}}^4} \ . \tag{31}$$

It can be shown that such a spectrum matches smoothly onto the result for high-momentum-transfer jet production.

One of the outstanding problems of particle production is computing the total multiplicity of produced gluons. In the CGC description, this problem is solved. When $p_{\text{T}} \leq Q_{\text{sat}}$, non-linearities of the field equations become important, and the field stops going as $1/p_{\text{T}}^4$. Instead it becomes of order

$$\frac{dN}{dy d^2 p_{\text{T}}} \sim \frac{1}{\alpha_s} \pi R^2 \ . \tag{32}$$

The total multiplicity is therefore of order

$$N \sim \frac{1}{\alpha_s} \pi R^2 Q_{\text{sat}}^2 \ . \tag{33}$$

If $Q_{\text{sat}}^2 \sim A^{1/3}$, then the total multiplicity goes as A, the high-p_{T} differential multiplicity goes as $A^{4/3}$, as we naively expect for hard processes since they should be incoherent, and the low-momentum differential multiplicity goes as $A^{2/3}$, since these particles arise from the region where the hadrons are black disks and the emission should take place from the surface.

In Fig. 23, the various kinematic regions for production of gluons are shown. In Fig. 24, the results of numerical simulation of gluon production are shown. At small p_{T}, it is amusing that the distribution is well described by a two-dimensional Bose–Einstein distribution. This is presumably a numerical accident, and in this computation has absolutely nothing to do with thermalized distributions.

3.7 Thermalization

As shown in Fig. 17, in a heavy-ion collision, the slow particles are produced first near the collision point and the slow particles are produced later far from the collision point. This introduces a gradient into the initial matter distribution, and the typical comoving volume element expands like $1/\tau$. To understand the factor of $1/\tau$ in the above equation, note that if we convert $dN/dz = dN/dy$, $dy/dz = dN/dy \ 1/t$, where we used our previous definition of space–time rapidity, and where we evaluated at $z = 0$. This is the physical rest frame density at $z = 0$.
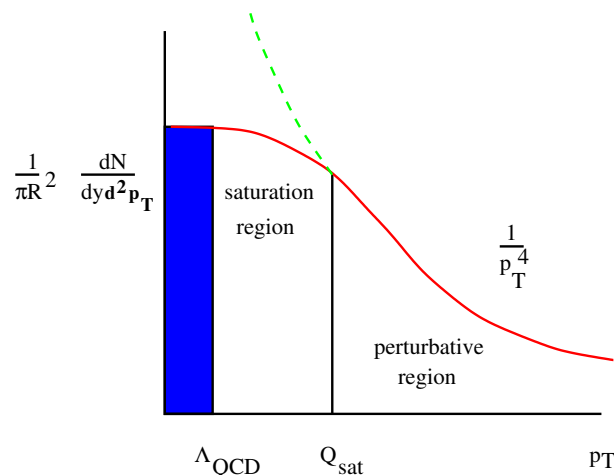
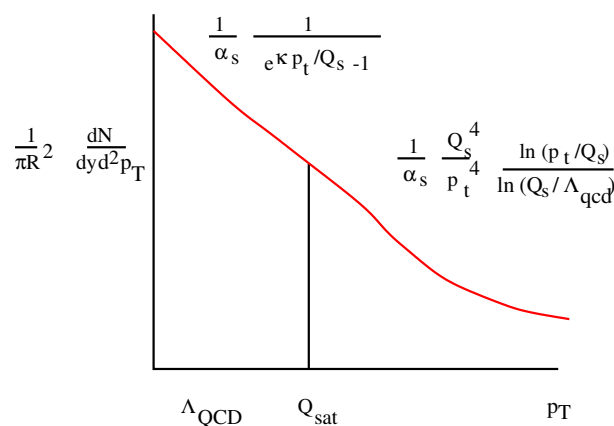**Fig. 23:** A cartoon representation of the various kinematic regions of gluon production

Fig. 24: The numerically computed distribution of produced gluons

If entropy is conserved, as is the case for a thermalized system with expansion time small compared to collision time,

$$S \sim T^3 \tau R^2 \tag{34}$$

is fixed so that $T \sim 1/\tau^{1/3}$. Therefore for a thermalized system, the energy density decreases as $\epsilon \sim 1/\tau^{4/3}$ for a system with no scattering so that the typical transverse momentum does not change, $\epsilon \sim 1/\tau$.

For the initial conditions typical of a colour glass condensate, thermalization is not so easy to do [16]. At the earliest times, the typical transverse momentum is large, of order of the saturation momentum. At this scale, the coupling is weak $\alpha_s(Q_{sat}) \ll 1$, at least for asymptotically large energy.

To estimate the typical scattering time, we need to know the density and the mean free path. At early times, the density is that in the transverse space diluted by the longitudinal expansion of the system,

$$\rho = \langle p_T^2 \rangle / \tau . \tag{35}$$

The scattering cross-section is on the other hand $\sigma \sim \alpha_s^2 \ln(\rho)/\langle p_T^2 \rangle$. The logarithmic cutoff comes about from Debye-screening the Coulomb cross-section. (The linear divergence can be shown to cancel for thermalization processes.)

236

Thermalization requires that $\tau \gg \tau_{\text{scat}}$, since τ itself is the characteristic expansion time. This requires that

$$\tau \geq \exp(c/\alpha_{\text{s}}) 1/Q_{\text{sat}} \ . \tag{36}$$

For practical purposes and for weakly coupled systems, there is never thermalization by elastic scattering!

Thermalization, if it in fact occurs, takes place by inelastic scattering. The physics of what is happening is easy to understand. Because the system begins its evolution with p_{T} at such a large typical scale Q_{sat}, the coupling is weak and the system does not easily thermalize by elastic scattering. It therefore expands and becomes overly dilute compared to the typical density associated with the transverse-momentum scale p_{T}^3. When a system is overly dilute, the Debye screening length becomes very large, since the Debye length scales inversely with the density. The Debye length is what cuts off scattering processes in the infrared. Multigluon production processes can be shown to diverge like the Debye screening length, whereas elastic processes only diverge like the logarithm of this length. Therefore, when the Debye screening length is of order $1/\alpha_{\text{s}}$, multigluon production begins to become more important than elastic scattering. This happens at a time $\tau \sim 1/(\alpha_{\text{s}} Q_{\text{sat}})$.

The details of how this thermalization occurs have not been fully worked out in detail. Current estimates of the time of thermalization matter produced in heavy-ion collisions at RHIC energies range from $0.3 \leq \tau \leq 3 \ \text{fm}/c$.

4 Lecture III: What we have learned from RHIC

In this lecture, I review results from RHIC and describe what we have learned so far about the production of new forms of matter in heavy-ion collisions. I shall make the case that we have produced matter of extremely high energy density, so high that it is silly not to think of it as composed of quarks and gluons. I also shall argue that this matter is strongly interacting with itself. The issue of the properties of this matter is still largely unresolved. For example, whether the various quantities measured are more properly described as arising from a colour glass condensate or from a quark–gluon plasma, although we can easily understand in most cases which form of matter should be most important.

The data presented here are taken from the RHIC whitepapers [13]. For references to the original publications, please look there.

4.1 The energy density is big

The particle multiplicity as a function of energy has been measured at RHIC, as shown in Fig. 25. Combining the multiplicity data together with the measurements of transverse energy or of typical particle transverse momenta, one can determine the energy density of the matter when it decouples [13]. One can then extrapolate backwards in time. We extrapolate backwards using one-dimensional expansion, since decoupling occurs when the matter first begins to expand three dimensionally. We can extrapolate backwards until the matter has melted from a colour glass.

To do this extrapolation we use that the density of particles falls as $N/V \sim 1/t$ during one-dimensional expansion. If the particles expand without interaction, then the energy per particle is constant. If the particles thermalize, then $E/N \sim T$, and since $N/V \sim T^3$ for a massless gas, the temperature falls as $T \sim t^{-1/3}$. For a gas which is not quite massless, the temperature falls somewhere in the range $T_o > T > T_o(t_o/t)^{1/3}$, that is the temperature is bracketed by the value corresponding to no interaction and to that of a massless relativistic gas. This one-dimensional expansion continues until the system begins to feel the effects of finite size in the transverse direction, and then rapidly cools through three-dimensional expansion. Very close to when three-dimensional expansion begins, the system decouples and particles free-stream to detectors without further interaction. We shall take a conservative overestimate of this time to be of order $t_{\text{melt}} \sim 0.3 \ \text{fm}/c$. The extrapolation backwards is bounded by $\epsilon_f(t_f/t) < \epsilon(t) < \epsilon_f(t_f/t)^{4/3}$. The lower bound is that assuming that the particles do not thermalize
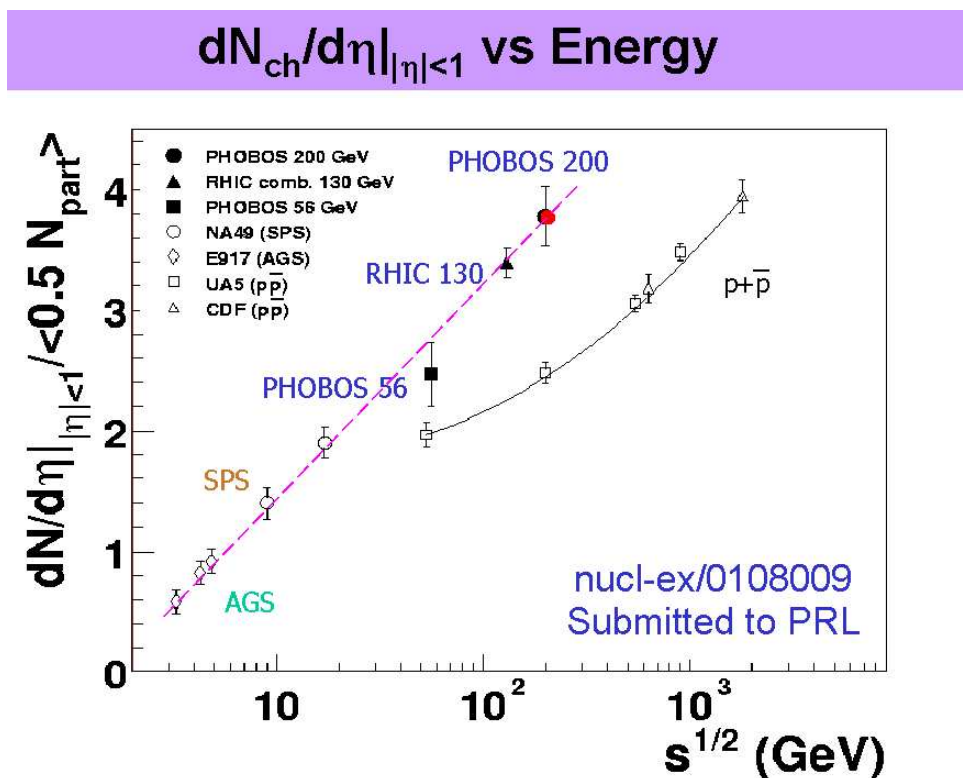
Fig. 25: Particle multiplicity as a function of energy as measured at RHIC

and their typical energy is frozen. The upper bound assumes that the system thermalizes as an ideal massless gas. We argued above that the true result is somewhere in between. When the time is as small as the melting time, then the energy density begins to decrease as time is further decreased.

This bound on the energy density is shown in Fig. 26. On the left axis is the energy density and on the bottom axis is time. The system begins as a colour glass condensate, then melts to quark–gluon matter which eventually thermalizes to a quark–gluon plasma. At a time of a few fm/c, the plasma becomes a mixture of quarks, gluons, and hadrons which expand together.

At a time of about 10 fm/c, the system falls apart and decouples. At a time of $t \sim 1$ fm/c, the estimate we make is identical to the Bjorken energy density estimate, and this provides a lower bound on the energy density achieved in the collision. (All estimates agree that by a time of order 1 fm/c, matter has been formed.) The upper bound corresponds to assuming that the system expands as a massless thermal gas from a melting time of 0.3 fm/c. (If the time was reduced, the upper bound would be increased yet further.) The bounds on the energy density are therefore

$$2\text{–}3 \ \text{GeV/fm}^3 < \epsilon < 20\text{–}30 \ \text{GeV/fm}^3 \tag{37}$$

where we included a greater range of uncertainty in the upper limit because of the uncertainty associated with the formation time. The energy density of nuclear matter is about 0.15 GeV/fm^3, and even the lowest energy densities in these collisions is in excess of this. At late times, the energy density is about that of the cores of neutron stars, $\epsilon \sim 1$ GeV/fm^3.

At such extremely high energy densities, the matter is most simply described in terms of its quark and gluon degrees of freedom.
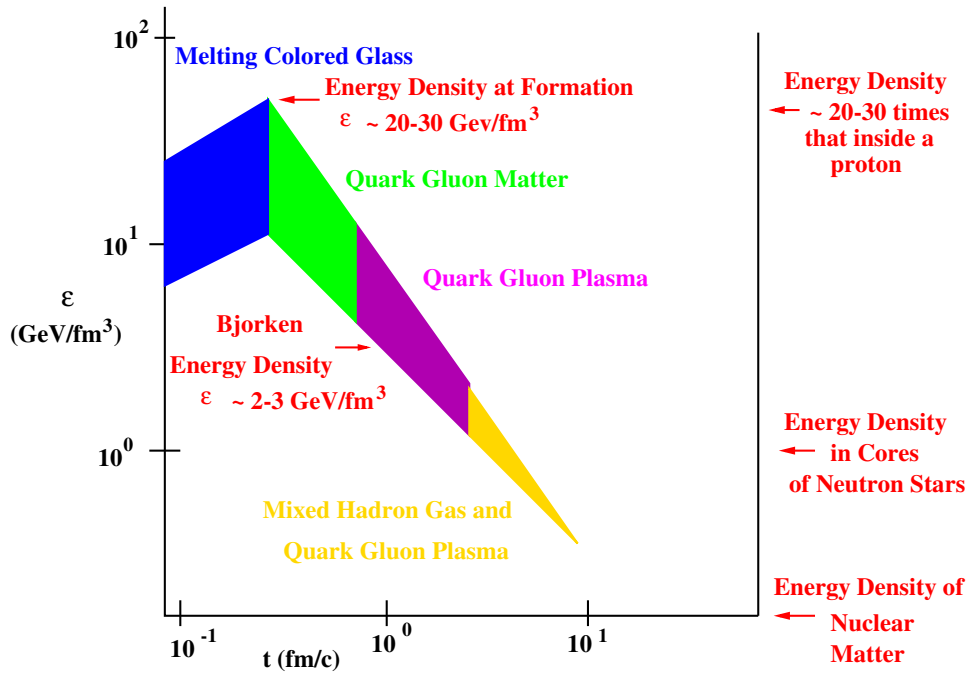
Fig. 26: Bounds on the energy density as a function of time in heavy-ion collisions
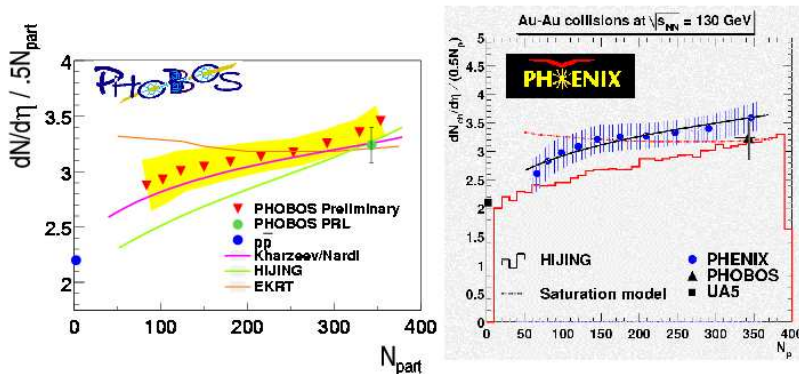
Fig. 27: The CGC description of the participant dependence of the multiplicity of produced particles

4.2 The gross features of multiplicity distributions are consistent with coloured glass

As argued by Kharzeev and Nardi [17], the density of produced particles per nucleon which participates in the collision N_{part} should be proportional to $1/\alpha_s(Q_{sat})$, and $Q_{sat}^2 \sim N_{part}$. This dependence follows from the $1/\alpha_s$ which characterizes the density of the colour glass condensate. In Fig. 27, we show the data from PHENIX and PHOBOS [13]. The Kharzeev–Nardi form fits the data well. Other attempts such as HIJING [18], or the so-called saturation model of Eskola–Kajantie–Ruuskanen–Tuominen [19] are also shown in the figure.
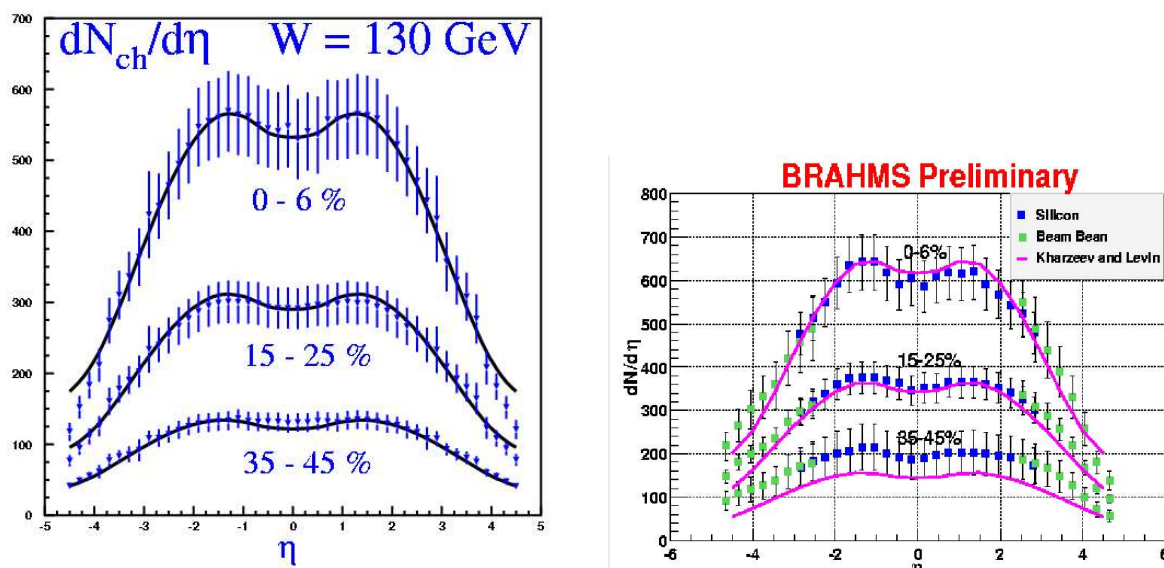
Fig. 28: Colour glass condensate fits to the rapidity density measured in the PHOBOS and BRAHMS experiments

Kharzeev and Levin have recently argued that the rapidity distributions as a function of centrality can be computed from the colour glass description [20]. This is shown in Fig. 28.

4.3 The CGC describes features of deep inelastic scattering

The colour glass condensate provides a theory of the hadron wavefunction at very small values of x. As such, it should describe features not only of high-energy nucleus–nucleus scattering, but also electron–hadron scattering. This includes inclusive scattering and diffraction. It indeed appears that there is such a successful phenomenology [21].

In these notes, I shall describe only one aspect of this phenomenology, geometric scaling [22]-[25]. The basic idea is that the cross-section for virtual photon scattering from a hadron should be, up to some trivial overall scale factor, a dimensionless function. If the saturation momentum is the only scale in the problem and the properties of the matter probed depend only upon the density of the matter, then

$$\sigma_{\gamma^* p} \sim F(Q^2/Q_{\text{sat}}^2(x)) \tag{38}$$

and is not an independent function of x and Q^2. The dependence of the saturation momentum on x can be computed [26], or can be determined from data. In Fig. 29, this cross-section is plotted as a function of $\tau = Q^2/Q_{\text{sat}}^2$ for values of $x \leq 10^{-2}$. Indeed, there appears to be such scaling.

While it is easy to understand this scaling for $Q^2 \leq Q_{\text{sat}}^2$, it is perhaps a little surprising that it works to much larger values of Q^2. One can show that one expects approximate scaling up to $Q^2 \sim Q_{\text{sat}}^4/\Lambda_{\text{QCD}}^2$. However, one should and can compute scaling violations [25].

4.4 The CGC provides a theory of shadowing

The naive expectation for the production of hard particles from a nucleus is that they should be generated by incoherent scattering. This is, however, modified because of multiparticle scattering, and because the gluon distribution function itself acquires a non-trivial dependence upon the nuclear baryon number. The colour glass condensate provides a theory of this modification [27–32]

240
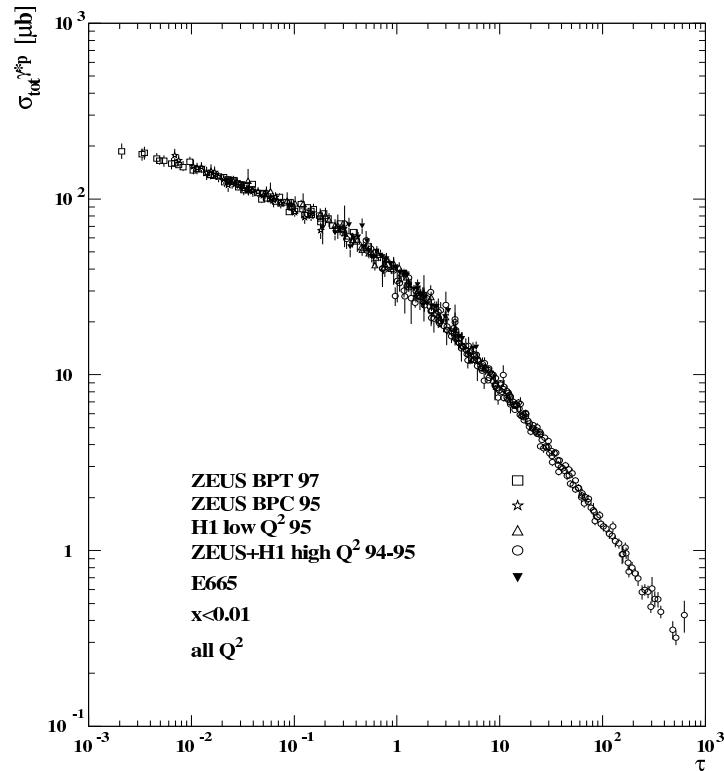
Fig. 29: The cross-section σ_{γ^*p} as a function of τ for $x \leq 10^{-2}$

One can understand how this works by first considering the effects of multiple elastic scattering. Such scattering does not change the number of particles. At very high p_T, the effects of multiple scattering should be small, since cross-sections are small. At intermediate p_T, the p_T distribution in a nucleus should broaden relative to that of incoherent scattering from nucleons. By conservation of probability, this requires a suppression at low p_T. By similar reasoning, one expects that these effects will be accentuated as one goes from peripheral to more central collisions. The results of one such computation of multiple scattering are shown in Fig. 30 [33].

One also expects that the effects of multiple scattering will be larger at small values of x because there are more degrees of freedom to scatter from.

Such multiple-scattering effects are included in the CGC description of the hadron collisions, but there is another effect which is larger at very small x. This is the quantum evolution of the hadron wavefunction. Because the saturation momentum is larger in nuclei than it is in protons, it is more difficult to produce glue at small x. Therefore as one goes to smaller values of x, there should be fewer particles at small x relative to the expectation from incoherent scattering. In Fig. 31, p_T distributions as a function of x are shown for the ratio of hadron–nucleus collisions to incoherent scattering. At large values of x there is a clear Cronin enhancement. At small values of x, there is a suppression as predicted by quantum evolution in the CGC. There is a similar suppression as the centrality of the collisions increases in distinction from the effects of multiple elastic scattering.

In the BRAHMS experiment, dAu collisions were used to study this effect. The results are shown in Fig. 32 [13]. Similar results have been found by STAR and PHENIX [13]. The results are qualitatively in accord with the CGC expectations, and also exhibit semi-quantitative agreement [34].
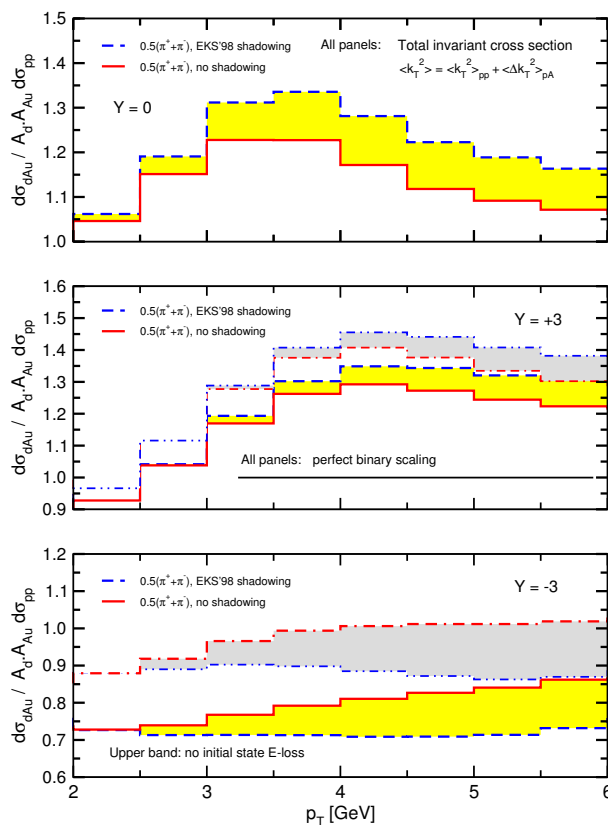
Fig. 30: The expectations of multiple scattering in dAu collisions in a multiple-scattering computation
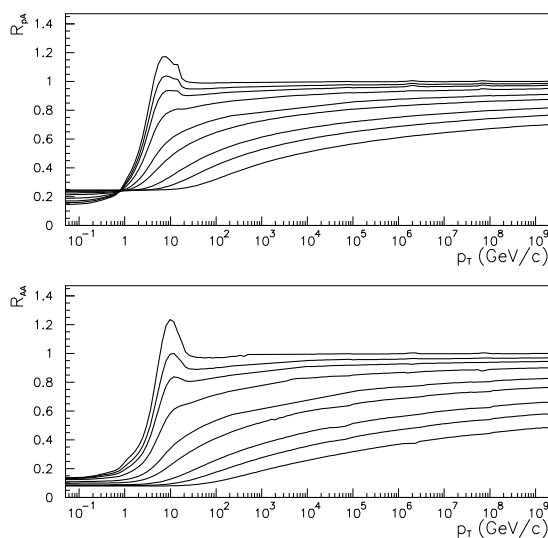
Fig. 31: The p_T distributions in hadron–nucleus collisions relative to incoherent scattering. Different curves correspond to different values of x.
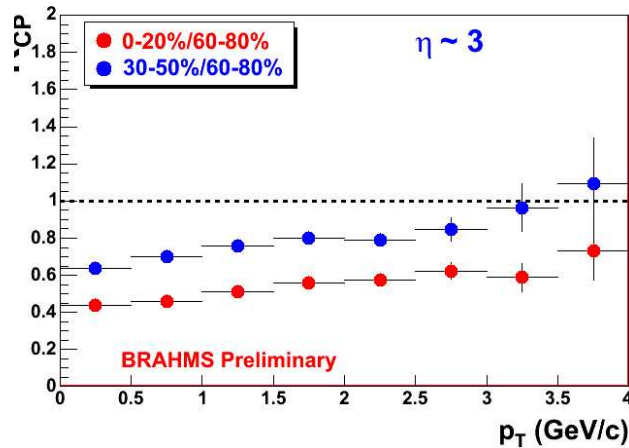
Fig. 32: The measurements from BRAHMS of the ratio of dAu high-p_T particle production to that of incoherent scattering as a function of x and centrality

4.5 Matter has been produced which interacts strongly with itself

In off-zero-impact-parameter heavy-ion collisions, the matter which overlaps has an asymmetry in density relative to the reaction plane. This is shown in the left-hand side of Fig. 33. Here the reaction plane is along the x axis. In the region of overlap there is an x–y asymmetry in the density of matter which overlaps. This means that there will be an asymmetry in the spatial gradients which will eventually transmute itself into an asymmetry in the momentum space distribution of particles, as shown in the right-hand side of Fig. 33.
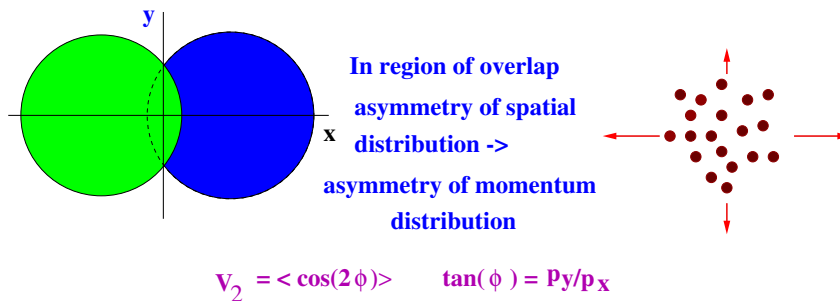
Fig. 33: The asymmetry in the distribution of matter in an off-centre collision is converted to an asymmetry of the momentum space distribution

This asymmetry is called elliptic flow and is quantified by the variable v_2. In all attempts to theoretically describe this effect, one needs very strong interactions among the quarks and gluons at very early times in the collision [35]. In Fig. 34, two different theoretical descriptions are fit to the data by STAR and PHOBOS [13]. On the left-hand side, a hydrodynamical model is used [36]. It is roughly of the correct order of magnitude and has roughly the correct shape to fit the data. This was not the case at lower energy. On the right-hand side are preliminary fits assuming colour glass [37]. Again it has roughly the correct shape and magnitude to describe the data. In the colour glass, the interactions are very strong essentially from $t = 0$, but unlike the hydrodynamic models it is field pressure rather than particle pressure which converts the spatial anisotropy into a momentum space-anisotropy.
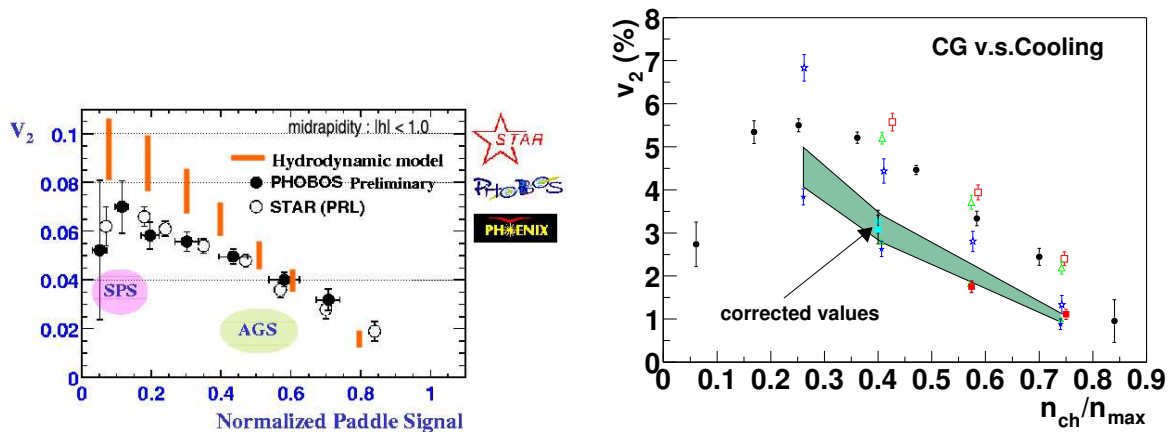
Fig. 34: (a) A hydrodynamic fit to v_2. (b) The coloured glass fit.

Probably the correct story for describing flow is complicated and will involve both the quark–gluon plasma and the colour glass condensate. Either description requires that matter be produced in the collisions and that it interact strongly with itself. In the limit of single scatterings for the partons in the two nuclei, no flow is generated.

Recent data on charm particles show that they too flow with the produced matter [38, 39]. Charm is a very heavy particle, and as such it requires many collisions with other particles before it can flow with the surrounding matter. The amount of flow seen experimentally exceeds the wildest expectations of theorists.

4.6 How strongly does the quark–gluon plasma interact?

4.6.1 Jets are quenched

One of the most interesting results from the RHIC experiments is the so-called 'jet quenching' [13], [40–43]. In Fig. 35(a), the single-particle hadron spectrum is scaled by the same result in pp collisions and scaled by the number of collisions. The number of collisions is the number of nucleon–nucleon interactions, which for central collisions should be almost all of the nucleons. One is assuming hard scattering in computing this number, so that a single nucleon can hard-scatter a number of times as it penetrates the other nucleus. The striking feature of this plot is that the ratio does not approach one at large p_T. This would be the value if these particles arose from hard scattering which produced jets of quarks and gluons, and the jets subsequently decayed.
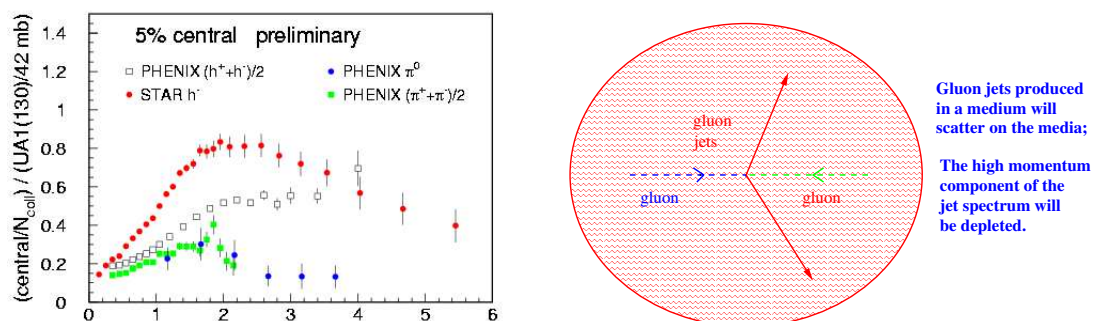
Fig. 35: (a) The p_T distribution of particles scaled by the data from pp collisions times the number of hard collisions inside the nuclei. (b) A pair of jets is produced in a hard collision and they propagate through the surrounding matter.

The explanation for this is shown in Fig. 35(b). Here a pair of jets is produced in a gluon–gluon collision. The jets are immersed in a quark–gluon plasma, and rescatter as they poke through the plasma. This shifts the transverse-momentum spectrum down, and the ratio to pp collisions, where there is no significant surrounding media, is reduced.

This suppression has been conclusively shown to be a final-state effect. One can measure the corresponding suppression in dA collisions, and at the central rapidity values where one sees strong suppression in jet production in AuAu collisions, there is little suppression or even enhancement seen in dA collisions. One can also look at the correlation in azimuthal angle of high-p_T-produced particles, as
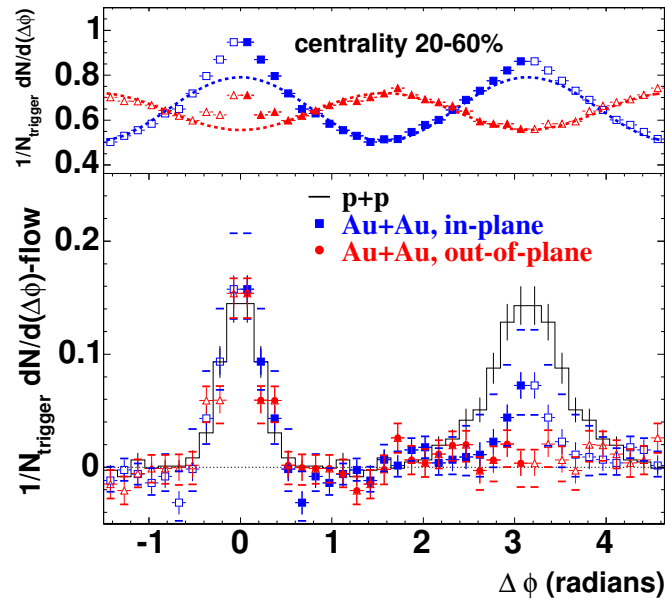
Fig. 36: The forward–backward correlation for high-p_T particles produced in STAR

shown in Fig. 36. In pp collisions, if there is a high-p_T particle produced, then at an azimuthal angle 180 degrees away, one expects to see an excess of hard particles. This was done and verified in STAR. In central AuAu collisions, one looks in the backward direction, and the peak has disappeared.

The essential problem with jet quenching is that it is much stronger than expected from QCD computation. Jet quenching apparently persists out to 20 GeV, and is present for charmed particles.

One of the reasons why jet quenching is so important for the RHIC programme is that it gives a good measure of the degree of thermalization in the collisions. If jets are strongly quenched by transverse momenta of 4 GeV, then because cross-sections go like $1/E^2$ for quarks and gluons, this would be strong evidence for thermalization at the lower energies typical of the emitted particles.

4.6.2 *The matter flows and is well approximated by perfect fluid hydrodynamics*

One can look for evidence of thermalization directly from the measured p_T distributions. Here one can do a hydrodynamic computation and in so far as it agrees with the results, one is encouraged to believe that there is thermalization. On the other hand, these distributions may have their origin in the initial conditions for the collision, the coloured glass. In reality, one will have to understand the tradeoff between both effects. The hydrodynamic models do a good job in describing the data for $p_T \leq 2$ GeV. Here there is approximate m_T scaling, a characteristic feature of hydrodynamic computations. This scaling arises naturally because hydrodynamic distributions are produced by flowing matter which has a characteristic transverse flow velocity with a well-defined local temperature. Particles with the same m_T should have arisen from regions with the same transverse-flow velocity and temperature.
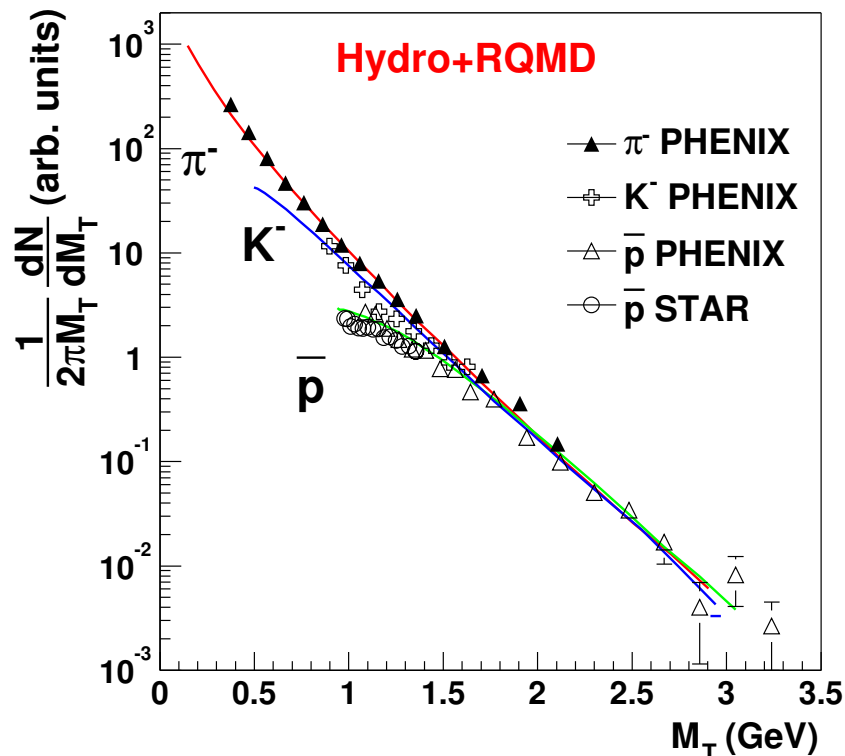
Fig. 37: The hydrodynamical model fits to the m_T spectra for the PHENIX and STAR data

Hydrodynamical models successfully describe the data on m_T distributions [44]. In Fig. 37 the results of the simulation by Shuryak and Teaney are shown compared to the STAR and PHENIX data [13]. The shape of the curve is a prediction of the hydrodynamic model, and is parametrized somewhat by the nature of the equation of state. Notice that the STAR data include protons near threshold, and here the m_T scaling breaks down. The hydrodynamic fits get this dependence correctly, and this is one of the best tests of this description. The hydrodynamic models do less well on fits to the more peripheral collisions. In general, a good place to test the hydrodynamic models' predictions is with massive particles close to threshold, since here one deviates in a computable way from the m_T scaling curve, which is arguably determined from parametrizing the equation of state.

If one can successfully argue that there is thermalization in the RHIC collisions, then the hydrodynamic computations become compelling. One should remember that hydrodynamics requires an equation of state plus initial conditions, and these initial conditions are determined by coloured glass. Presumably, a correct description will require the inclusion of both types of effects [45].

At present, hydrodynamical models do an excellent job of describing data on distributions of particles with $p_T \leq 2$ GeV. This uses perfect fluid hydrodynamics with no viscosity. This was not the case at CERN [46]. Estimates of the viscosity which is consistent with the experimental data give numbers which are quite small, leading some to conclude that the quark–gluon plasma is the most perfect fluid yet measured. There are of course some uncertainties in these conclusions, largely associated with the initial conditions for the hydrodynamic equations, uncertainty in the equation of state, and dispersion in the treatment of the matter at late times when the hydrodynamic description must break down. Nevertheless, the fact that the hydrodynamic computations seem to work well, and the existence of strong jet quenching, lead me to conclude that at a minimum, the matter produced is reasonably well approximated as a thermal system, and that is remarkably strongly self-interacting. This means that I believe that the semi-quantitative conclusions drawn from hydrodynamic simulation have substance.

4.7 Confinement and chiral symmetry restoration

We would like to know whether or not deconfinement has occurred in dense matter or whether chiral symmetry restoration has changed particle masses.
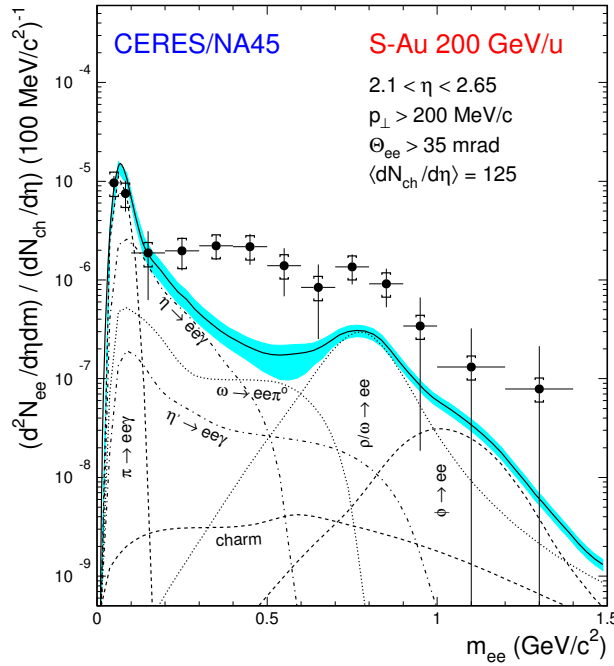

Fig. 38: The CERES data on low-mass electron–positron pairs. The expected contribution from ordinary hadrons is shown by the solid line. The data points are for the measured electron–positron pairs.

This can be studied in principle by measuring the spectrum of dileptons emitted from the heavy-ion collision. These particles probe the interior of the hot matter since electromagnetically interacting particles are not significantly attenuated by the hadronic matter. For electron–positron pairs, the mass distribution has been measured in the CERES experiment at CERN [47], and is shown in Fig. 38. Shown in the plot is the distribution predicted from extrapolating from pA collisions. There should be a clear ρ and ϕ peak, which has disappeared. This has been interpreted as a resonance mass shift [48], enhanced η' production [49], but is most probably collisional broadening of the resonances in the matter produced in the collisions [50]. Nevertheless, if one makes a plot such as this and the energy density is very high and there are no resonances at all, then this would be strong evidence that the matter is not hadronic, i.e., the hadrons have melted.

The resolution in the CERES experiment is unpleasantly large, making it difficult to unambiguously interpret the result. Whether or not such an experiment could be successfully run at RHIC, much less whether the resolution could be improved, is the subject of much internal debate among the RHIC experimentalists.

It has also been pointed out recently that the matter in the early stages after the collisions has remarkable properties [51]. At the earliest times, there is both longitudinal colour electric and colour magnetic fields. These fields evolve towards a thermalized system as the longitudinal fields evaporate into gluons. I call this early matter the glasma. The fields have a non-zero colour $\vec{E} \cdot \vec{B}$. This is an unusal situation and generates an anomaly in the axial vector current. This means that even very energetic quarks will flip their helicities in the presence of such fields, and generate chiral symmetry breaking. It has been conjectured that such helicity flip may ultimately be responsible for mass generation in QCD. The idea of the glasma with its anomalous fields is recent, and it is not yet possible to assess the experimental implications.

4.8 Confinement and J/Ψ suppression

In Fig. 39, the NA(50) data for J/Ψ production is shown [52]. In the first figure, the ratio of J/Ψ production cross-section to that of Drell–Yan is shown as a function of E_{T}, the transverse energy, for the lead–lead collisions at CERN. There is a clear suppression at large E_{T} which is greater than the hadronic
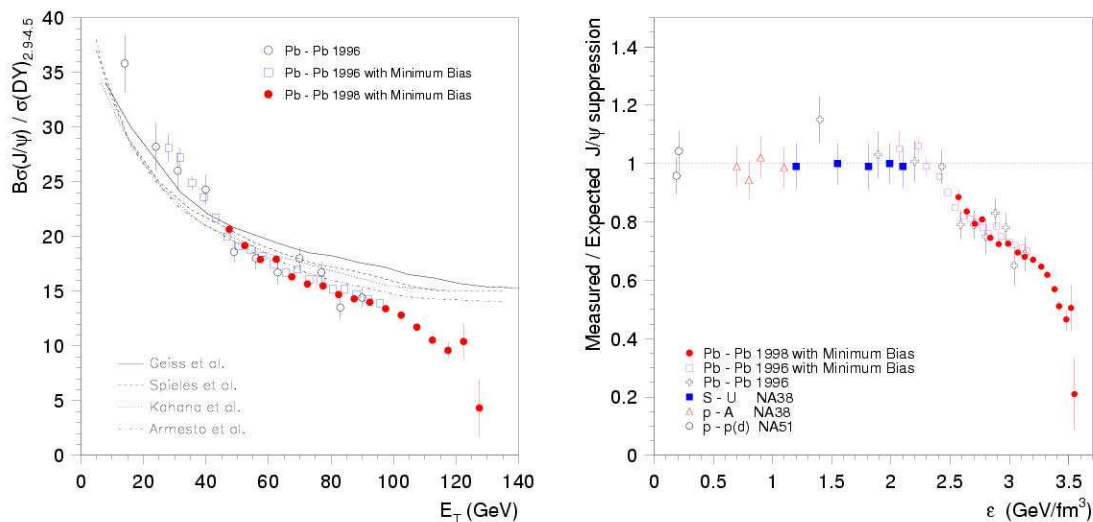

Fig. 39: (a) The ratio of produced J/Ψ pairs to Drell–Yan pairs as a function of transverse energy E_{T} at CERN energy. (b) The measured compared to the theoretically expected J/Ψ suppression as a function of the Bjorken energy density for various targets and projectiles.

absorption model calculations which are plotted with the data [53]. In the next figure, the theoretically expected J/Ψ suppression based on hadronic absorption models is compared to that measured as a function of the Bjorken energy density for various targets and projectiles. There appears to be a sharp increase in the amount of suppression for central lead–lead collisions.

Is this suppression due to Debye screening of the confinement potential in a high-density quark–gluon plasma [54–56]? Might it be due to higher twists, enhanced rescattering, or changes in the gluon distribution function [57, 58]? Might the J/ψ suppression be changed into an enhancement at RHIC energies and result from the recombination in the produced charm particles, many more of which are produced at RHIC energy [59–62]?

These various descriptions can be tested by using the capability at RHIC to do pp and pA as well as AA. Issues related to multiple scattering, higher twist effects, and changes in the gluon distribution function can be explored. A direct measurement of open charm will be important if final-state recombination of the produced open charm makes a significant amount of J/Ψ's.

The data from the PHENIX experiment show roughly the same pattern of suppression as seen at CERN [63]. This is a surprise since one naively expected that there should be more suppression at higher energy densities. This had led some to speculate that there may be significant recombination effects in the final state [59], [62]. This will be resolved after measurements of resonant states decaying into charm, J/Ψ flow, and more, as the programme at RHIC continues.

4.8.1 Direct photons

One of the first suggestions for a signal of the quark–gluon plasma was thermal radiation due to photons [64–68]. Produced photons are penetrating, and in principle can measure the properties of the hot matter at early times in the collision. The problem is that there are huge backgrounds from resonance decays.

At very large p_T, the resonance backgrounds are small, but there the dominant process for making photons is hard scattering of a gluon and a quark, and reflects the initial parton distribution functions. This has been measured at RHIC, and agrees with perturbative computations. More recently, there has been a claimed measurement from PHENIX of an excess seen at intermediate p_T [69].

The excess is surely interesting, but it is an excess relative to a theoretical computation, and these computations must be checked against the pp and dA data before one can be too excited about the result seen in AA. Of course, one has to check against all possible sources of photons for AA before one concludes that these photons arise from thermal radiation. As the results are new, these checks are not yet complete.

4.8.2 The lifetime and size of the matter produced

The measurement of correlated pion pairs, the so-called HBT pion interferometry, can measure properties of the space–time volume from which the hadronic matter emerges in heavy-ion collisions [70]. The quantities R_{long}, R_{side} and R_{out} shown in Fig. 40 measure the transverse size of the matter at decoupling and the decoupling time.
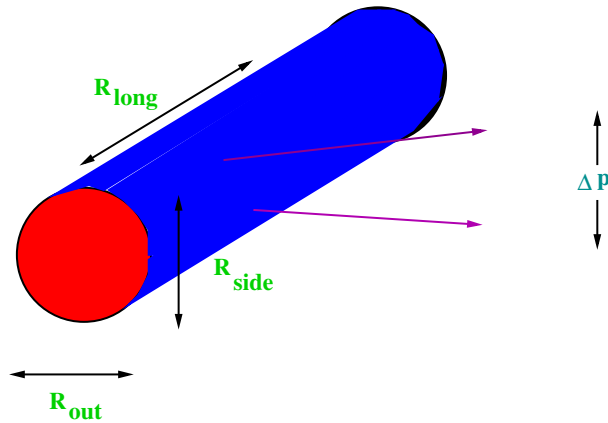
Fig. 40: The various radii used for HBT pion interferometry

In Fig. 41, the data from STAR and PHENIX are shown [13]. There is only a weak dependence on energy, and the results seem to be more or less what was observed at CERN energies. This is a surprise, since a longer time for decoupling is expected at RHIC. Perhaps the most surprising result is that R_{out}/R_{side} is close to 1, whereas most theoretical expectations give a value of about $R_{out}/R_{side} \sim 2$ [71, 72]. Perhaps this is due to greater than expected opacity of the emitting matter? At this time, there is no consistent theoretical description of the HBT data at RHIC. Is there something missing in our space–time picture?

4.8.3 The flavour composition of the quark–gluon plasma

The first signal proposed for the existence of a quark–gluon plasma in heavy-ion collisions was enhanced strangeness production [73]. This has led to a comprehensive programme in heavy-ion collisions to measure the ratios of abundances of various flavours of particles [74]. In Fig. 42(a), the ratios of flavour abundances is compared to a thermal model for the particle abundances [75–77]. The fit is quite good. In Fig. 42(b), the temperature and baryon chemical potential extracted from these fits is shown for experiments at various energies and with various types of nuclei. It seems to agree reasonably well with what might be expected for a phase boundary between hadronic matter and a quark–gluon plasma.

This would appear to be a compelling case for the production of a quark–gluon plasma. The problem is that the fits done for heavy ions to particle abundances work even better in $e^+ e^-$ collisions.
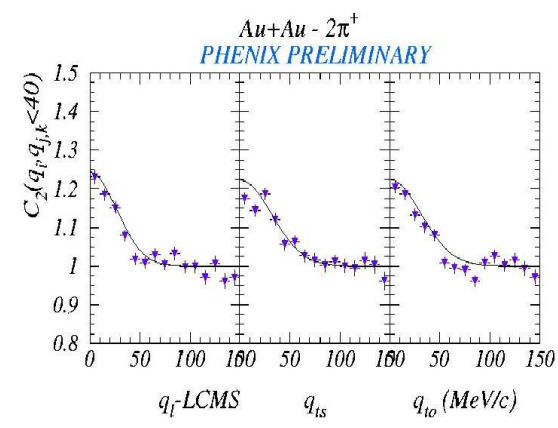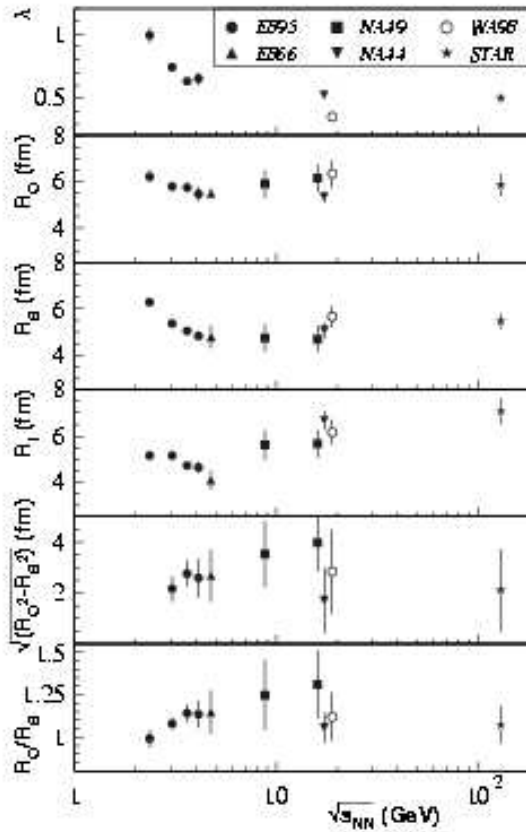
Fig. 41: (a) The various HBT radii measured in heavy-ion experiments including the new data from STAR. (b) The correlation functions which determine the radii as a function of the pair momenta measured in PHENIX.
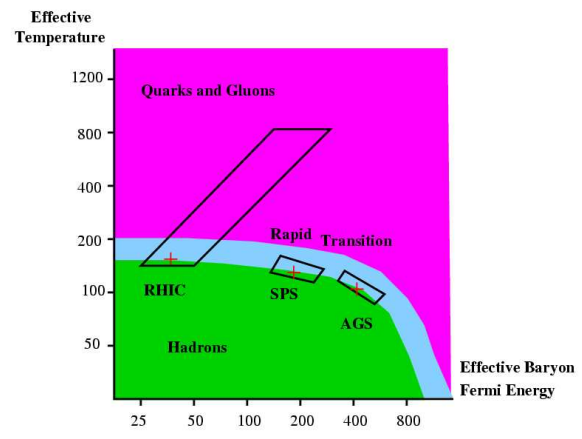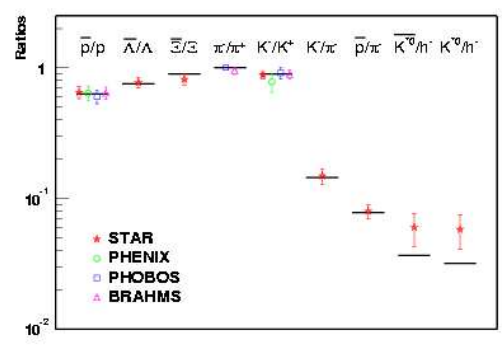

Fig. 42: (a) Various ratios of particle abundances and the RHIC data. The lines are the predictions of a thermal model. (b) The temperature vs baryon chemical potential for a thermal model which is fit to data at various energies. The dashed line is a hypothetical phase boundary between a quark–gluon plasma and a hadronic gas.

One definitely expects no quark–gluon plasma in $e^+ e^-$ collisions. There is a deep theoretical question to be understood here: How can thermal models work so well for non-thermal systems? Is there some simple saturation of phase space? The thermal model description can eventually be made compelling for heavy-ion collisions once the degree of thermalization in these collisions is understood.

References

[1] For a summary of recent results, see the excellent review talk by Jean-Paul Blaizot, Lectures at the 40th Internationale Universitätswochen für Kern- und Teilchenphysik: Dense Matter (IUKT 40), Schladming, Styria, Austria, 2001, Lectures on Quark Matter, L. Mathelitsch and W. Plessas [Eds.], Lect. Notes Phys. **583** (2002) 117–160, [hep-ph/0107131].

[2] For a state-of-the-art, up-to-date review on the properties of the colour glass condensate, see E. Iancu and R. Venugopalan, *in* Quark–Gluon Plasma 3, W. N. Wang and R. C. Hwa [Eds.] (World Scientific, Singapore, 2004), [hep-ph/0303204].

[3] For a review about the prospects for spin physics at RHIC and other accelerators, see G. Bunce, N. Saito, J. Soffer and W. Vogelsang, *Annu. Rev. Nucl. Part. Sci.*, **50** (2000) 525.

[4] J. C. Collins and M. J. Perry, *Phys. Rev. Lett.* **34** (1975) 1353.

[5] For a summary of recent results, see the excellent review talk by F. Karsch, Lectures at the 40th Internationale Universitätswochen für Kern- und Teilchenphysik: Dense Matter (IUKT 40), Schladming, Styria, Austria, 2001, Lectures on Quark Matter, L. Mathelitsch and W. Plessas [Eds.], Lect. Notes Phys. **583** (2002) 209–249, [hep-lat/0106019].

[6] L. V. Gribov, E. M. Levin and M. G. Ryskin, *Phys. Rep.* **100** (1983) 1; A. H. Mueller and Jian-wei Qiu, *Nucl. Phys.* **B268** (1986) 427; J.-P. Blaizot and A. H. Mueller, *Nucl. Phys.* **B289** (1987) 847; L. D. McLerran and R. Venugopalan, *Phys. Rev.* **D49** (1994) 2233; **D49** (1994) 3352; E. Iancu, A. Leonidov and L. D. McLerran, *Nucl. Phys.* **A692** (2001) 583; E. Ferreiro, E. Iancu, A. Leonidov and L. D. McLerran, hep-ph/0109115.

[7] J. Breitweg *et al., Eur. Phys. J.* **67** (1999) 609.

[8] M. Froissart, *Phys. Rev.* **123** (1961) 1053.

[9] A. Martin, *Nuovo Cim.* **42** (1966) 930; L. Lukaszuk and A. Martin, *Nuovo Cim.* **52** (1967) 122.

[10] E. Ferreiro, E. Iancu, K. Itakura and L. McLerran, *Nucl. Phys.* **A710** (2002) 373.

[11] A. Kovner and U. Wiedemann, *Phys. Rev.* **D66** (2002) 051502; *Phys. Lett.* **B551** (2203) 311.

[12] A. Kovner, L. D. McLerran and H. Weigert, *Phys. Rev.* **D52** (1995) 6231; **D52** (1995) 3809; A. Krasnitz and R. Venugopalan, *Phys. Rev. Lett* **84** (2000) 4309; *Nucl. Phys.* **B557** (1999) 237; A. Krasnitz, Y. Nara and R. Venugopalan, *Phys. Rev. Lett.* **87** (2001) 192302.

[13] See the Whitepapers published by the BRAHMS, PHENIX, PHOBOS and STAR experiments for the original data and references. *Nucl. Phys.* **A757** (2005) 1; *Nucl. Phys.* **A757** (2005) 27; *Nucl. Phys.* **A757** (2005) 102; *Nucl. Phys.* **A757** (2005) 184.

[14] J. Jalilian-Marian, A. Kovner, L. McLerran and H. Weigert, *Phys. Rev.* **D55** (1997) 5414.

[15] J. Jalilian-Marian, A. Kovner, A. Leonidov and H. Weigert, *Nucl. Phys.* **B504** (1997) 415; *Phys. Rev.* **D59** (1999) 014014.

[16] R. Baier, A. H. Mueller, D. Schiff and D.T. Son, *Phys. Lett.* **B502** (2001) 51.

[17] D. Kharzeev and M. Nardi, *Phys. Lett.* **B507** (2001) 121.

[18] M. Gyulassy and Xin-Nian Wang, *Comput. Phys. Commun.* **83** (1994) 307.

[19] K. Eskola, K. Kajantie, P. Ruuskanen and K. Tuominen, *Nucl. Phys.* **B570** (2000) 379; K. Eskola, K. Kajantie and K. Tuominen, *Phys. Lett.* **B497** (2001) 39.

[20] D. Kharzeev and E. Levin, nucl-th/0108006.

[21] L. McLerran, 12th International Workshop on Deep Inelastic Scattering, Štrbské Pleso, Slovakia, 2004, D. Bruncko, J. Ferencei and P. Strízenec [Eds.] (Slovak Acad. Sci. Inst. Exp. Phys., Kosice, 2004), p. 161.

[22] K. Golec-Biernat and M. Wustoff, *Phys. Rev.* **D60** (1999) 114023.

[23] A. M. Staśto, K. Golec-Biernat and J. Kwieceński, *Phys. Rev. Lett.* **86** (2001) 596.

[24] E. Iancu, K. Itakura and L. McLerran, *Nucl. Phys.* **A708** (2002) 327.

[25] E. Iancu, K. Itakura and S. Munier, *Phys. Lett.* **B590** (2004) 199.

[26] A. H. Mueller and D. Triantafyllopoulos, *Nucl. Phys.* **B640** (2002) 331; D. Triantafyllopoulos, *Nucl. Phys.* **B648** (2003) 294.

[27] D. Kharzeev, E. Levin and L. McLerran, *Phys. Lett.* **B561** (2003) 93.

[28] J. Jalilian-Marian, Y. Nara and R. Venugopalan, *Phys. Lett.* **B577** (2003) 54.

[29] D. Kharzeev, Y. Kovcehgov and K. Tuchin, *Phys. Rev.* **D68** (2003) 094013.

[30] R. Baier, A. Kovner and U. Wiedemann, *Phys. Rev.* **D68** (2003) 054009; J. Albacete, N. Armesto, C. Salgado and U. Wiedemann, *Phys. Rev. Lett.* **92** (2004) 082001.

[31] Jian-Wei Qiu and I. Vitev, *Phys. Lett.* **B632** (2006) 507.

[32] E. Iancu, K. Itakura and D. Triantafyllopoulos, *Nucl. Phys.* **A742** (2004) 182.

[33] I. Vitev, *Phys. Lett.* **B562** (2003) 36.

[34] A. Dumitru, A. Hayashigata and J. Jalilian-Marian, *Nucl. Phys.* **A765** (2006) 464.

[35] S. Voloshin and Y. Zhang, *Z. Phys.* **C70** (1996) 665; A. M. Poskhanzer and S. A. Voloshin, *Phys. Rev.* **C58** (1998) 1671; J. Y. Ollitrault, *Phys. Rev.* **D46** (1992) 229.

[36] P. F. Kolb, J. Sollfrank and U. Heinz, *Phys. Lett.* **B459** (1999) 667; P. F. Kolb, P. Huovinen, U. Heinz and H. Heiselberg, *Phys. Lett.* **B500** (2001) 232.

[37] A. Krasnitz, Y. Nara and R. Venugopalan, *Phys. Lett.* **B554** (2003) 21.

[38] J. Adams *et al., Phys. Rev. Lett.* **94** (2005) 06231.

[39] S. Adler *et al., Phys. Rev. Lett.* **94** (2005) 082301; *Phys. Rev. Lett.* **88** (2002) 192303.

[40] J. D. Bjorken, FERMILAB-PUB-82-059-THY.

[41] D. Appel, *Phys. Rev.* **D33** (1986) 717.

[42] J. P. Blaizot and L. McLerran, *Phys. Rev.* **D34** (1986) 2739.

[43] M. Gyulassy, P. Levai and I. Vitev, *Nucl. Phys.* **B571** (2000) 197; *Phys. Rev. Lett.* **85** (2000) 5535; *Nucl. Phys.* **B594** (2001) 371.

[44] D. Teaney and E. V. Shuryak, *Phys. Rev. Lett.* **83** (1999) 4951; D. Teaney, J. Lauret and E. V. Shuryak, nucl-th/0110037.

[45] D. K. Srivastava, *Phys. Rev.* **C64** (2001) 064901.

[46] M. Gyulassy and L. McLerran, *Nucl. Phys.* **A750** (2005) 30.

[47] G. Agakishiev *et al., Nucl. Phys.* **A638** (1998) 159.

[48] G. E. Brown and M. Rho, *Phys. Rep.* **269** (1996) 33.

[49] J. Kapusta, D. Kharzeev and L. D. McLerran, *Phys. Rev.* **D53** (1996) 5028.

[50] R. Rapp, G. Chanfry and J. Wambach, *Phys. Rev. Lett.* **76** (1996) 368.

[51] T. Lappi and L. McLerran, hep-ph/0602189.

[52] For the latest results, see M. C. Abreau *et al., Nucl. Phys.* **A661** (1999) 93.

[53] J. Geiss, E. Bratskaya, W. Cassing and C. Greiner, nucl-th/981005; C. Spieles, R. Vogt, L. Gerland, S. A. Bass, M. Bleicher, H. Stocker and W. Greiner, *Phys. Rev.* **C60** (1999) 054901; D. E. Kahana and S. H. Kahana, *Phys. Rev.* **C60** (1999) 065206; N. Armesto, A. Capella, E. Ferreiro, A. Kaidalov and D. Sousa, *Nucl. Phys.* **A698** (2002) 583.

[54] T. Matsui and H. Satz, *Phys. Lett.* **B178** (1986) 416.

[55] D. Kharzeev and H. Satz, *Phys. Lett.* **B334** (1994) 155.

[56] J.-P. Blaizot and J.-Y. Ollitrault, *Phys. Rev. Lett.* **77** (1996) 1703.

[57] J. Armesto and A. Capella, *Phys. Lett.* **B430** (1998) 23; A. Capella, E. G. Ferreiro and A. Kaidalov, *Phys. Rev. Lett.* **85** (2000) 2080.

[58] Jian-wei Qiu, J. P. Vary and Xiao-fei Zhang, hep-ph/9809442.

[59] R. Thews, M. Schroeder and J. Rafelski, *Phys. Rev.* **C63** (2001) 054905.

[60] P. Braun-Munzinger and J. Stachel, *Phys. Lett.* **B490** (2000) 196.

[61] P. Braun-Munzinger and K. Redlich, *Eur. Phys. J.* **C16** (2000) 519.

[62] M. Gorenstein and M. Gazdzicki, *Phys. Rev. Lett.* **83** (1999) 4009; M. Gorenstein, A. P. Kostyk, H. Stoecker and W. Greiner, *Phys. Lett.* **B509** (2001) 277; M. Gorenstein, A. Kostyk, L. McLerran, H. Stoecker and W. Greiner, hep-ph/0012292; M. Gorenstein, A. Kostyk, H. Stocker and W. Greiner, *Phys. Lett.* **B524** (2002) 265.

[63] S. Adler *et al., Phys. Rev.* **C69** (2004) 014901.

[64] E. Shuryak, *Phys. Lett.* **B78** (1978) 150.

[65] K. Kajantie and H. I. Miettinen, *Z. Phys.* **C9** (1981) 341.

[66] D. Srivastava and B. Sinha, *Phys. Rev. Lett.* **73** (1994) 2421.

[67] E. L. Feinberg, *Nuovo Cim.* **A34** (1976) 391.

[68] L. McLerran and T. Toimeal, *Phys. Rev.* **D31** (1985) 545.

[69] S. Adler *et al., Phys. Rev. Lett.* **94** (2005) 232301, see also the proceedings of Quark Matter 2005, Budapest, Hungary.

[70] M. Gyulassy, S. Kauffmann and L. Wilson, *Phys. Rev.* **C20** (1979) 2267.

[71] S. Chapman, P. Scotto and U. Heinz, *Phys. Rev. Lett.* **74** (1995) 4400; S. Chapman and U. Heinz, *Phys. Lett* **B340** (1994) 250.

[72] S. Soff, S. Bass and A. Dumitru, *Phys. Rev. Lett.* **86** (2001) 3981.

[73] B. Muller and J. Rafelski, *Phys. Rev. Lett.* **48** (1986) 1066; P. Koch, B. Muller and J. Rafelski, *Phys. Rep.* **142** (1986) 167.

[74] For a summary, see M. Kaneta and N. Xu, *J. Phys.* **G27** (2001) 589.

[75] For a state-of-the-art assessment review, see J. Cleymans, hep-ph/0201142; J. Cleymans and K. Redlich, *Phys. Rev. Lett.* **81** (1998) 5284; *Phys. Rev.* **C60** (1999) 054908.

[76] P. Braun-Munzinger, J. Stachel, J. P. Wessels and N. Xu, *Phys. Lett.* **B365** (1996) 1; P. Braun-Munzinger, I. Heppe and J. Stachel, *Phys. Lett.* **B465** (1999) 15.

[77] G. Yen and M. Gorenstein, *Phys. Rev.* **C59** (1999) 2788.

Scientific programme of the Joint Institute for Nuclear Research

A.N. Sissakian
Director-Designate, JINR, Dubna, Russia

Abstract

The Joint Institute for Nuclear Research (JINR) is a large, multidisciplinary, international, scientific centre in which fundamental research in modern elementary particle physics, nuclear physics, and condensed-matter physics is integrated with the development and application of the newest technologies and university education activity in related areas.

1 Introduction

1.1 General information about the Joint Institute for Nuclear Research (JINR)

The Joint Institute for Nuclear Research (JINR) in Dubna was established on the basis of the convention signed by the Plenipotentiaries of the governments of the Member States of JINR in March 1956 in Moscow. JINR was created in order to unify the intellectual and material potential of the Member States in order to study the fundamental properties of matter.

Dubna as a town of science was founded immediately after the end of World War II. In 1947 a group of scientists led by Academician I.V. Kurchatov initiated construction of the then largest accelerator of charged particles—the synchrocyclotron. The accelerator was commissioned in 1949. Extensive fundamental and applied investigations into the properties of nuclear matter immediately started at the newly established Institute for Nuclear Problems (INP) with its operating 680 MeV synchrocyclotron, headed by the young physicists M.G. Meshcheryakov and V.P. Dzhelepov, later world-renowned scientists.

After INP, the Electrophysical Laboratory of the USSR Academy of Sciences (EFLAN), headed by Academician V.I. Veksler, was set up in Dubna. A new accelerator, a synchrophasotron with record parameters for that time, was constructed at EFLAN.

In 1954 the European Organization for Nuclear Research (CERN) was established near Geneva to unite the efforts of Western European countries for studying the fundamental properties of matter.

About the same time, under the stimulus of the USSR Government, the countries then belonging to the socialist world took the decision to establish the Joint Institute for Nuclear Research in Dubna from the INP and EFLAN laboratories. The same year, specialists from 12 countries (Albania, Bulgaria, China, Czechoslovakia, East Germany, Hungary, Mongolia, N. Korea, Poland, Romania, USSR, and Vietnam) came to Dubna. The town became international, and investigations into many fields of nuclear physics of interest for research centres of the JINR Member States were launched there.

Many scientists and engineers from the Member States have been trained in the JINR scientific schools established by N.N. Bogoliubov, D.I. Blokhintsev (Fig. 1), G.N. Flerov, I.M. Frank, B.M. Pontecorvo, V.I. Veksler, and other outstanding physicists. The development of different scientific directions at JINR is connected with the names of A. Baldin, A. Logunov, M. Markov, D. Shirkov, A. Tavkhelidze, as well as L. Infeld and H. Niewodniczanski (Poland), G. Nadjakov (Bulgaria), H. Hulubei (Romania), L. Janossy (Hungary), N. Sodnom (Mongolia), Wang Gangchang (China), Nguyen Van Hieu (Vietnam), V. Votruba and J. Kozesnik (Czechoslovakia), H. Pose and K. Lanius (Germany), and others.

The Charter of the JINR was adopted in 1956, the text of which was revised in 1992 and more recently in 1999. In accordance with the Charter, the activities of the Institute are achieved on the basis of its openness, and the mutual and equal co-operation of all the interested parties to participate in research.

Fig. 1: N.N. Bogoliubov and D.I. Blokhintsev

The aim of the Institute is

– to carry out theoretical and experimental investigations on adopted scientific topics;
– to organize the exchange of experience when carrying out research and the exchange of information obtained as a result of these investigations through publication of scientific papers, holding of conferences, symposia, etc.;
– to promote the development of the intellectual and professional capabilities of the scientific personnel;
– to establish and maintain contacts with other national and international scientific organizations and institutes to ensure stable and mutual co-operation;
– to use the results of the investigations of an applied nature to provide supplementary financial resources for fundamental research by implementing them into industrial, medical, and technological developments.

The results of investigations carried out at the Institute can be used solely for peaceful purposes for the benefit of mankind. So until the late 1980s, Dubna was a centre that unified the efforts of leading research groups of nuclear sciences from socialist countries and the Soviet Union.

After the disintegration of the USSR, membership of JINR underwent the following changes: the majority of Eastern European countries, such as Poland, the Czech and Slovak Republics, Bulgaria, Romania, and others continue to be Member States and contribute to the budget. Germany remains as an observer and makes a substantial financial contribution. Most of the former Soviet Union republics, which became independent states, entered JINR as new members.

There are different ways to participate in the activities of the Institute: on the basis of full or associated membership, or bilateral and multilateral agreements in order to perform particular scientific programmes. The JINR Member States contribute financially to the Institute's activities and have equal rights in its management.

At present JINR has 18 Member States: Armenia, Azerbaijan, Belarus, Bulgaria, Cuba, Czech Republic, Georgia, Kazakhstan, D.P. Republic of Korea, Moldova, Mongolia, Poland, Romania, Russian Federation, Slovak Republic, Ukraine, Uzbekistan, Vietnam.

The JINR has special co-operation agreements concluded on the governmental level with Germany, Hungary, Italy, and recently with the Republic of South Africa. Positive preliminary negotiations with scientific officials of the USA, China, Greece, India and other countries are underway to conclude similar governmental agreements with these countries.

Among the major partners with whom JINR has long-term co-operation agreements are

– CERN, in the field of high-energy physics,
– IN2P3 (France), in the field of nuclear and particle physics,
– INFN (Italy), in the field of nuclear and particle physics,
– FNAL, BNL, SLAC and other research centres in the USA.

The activity of JINR in the territory of its host country, the Russian Federation, is carried out in accordance with the Russian Federal Law on Ratification of 'The Agreement between the Government of the Russian Federation and JINR on the Location and Terms of Activity of JINR in the Russian Federation'. The Federal Law was approved by the Federation Council on 22 December 1999 and signed by the President of the Russian Federation, V. Putin, on 2 January 2000. The Federal Law came into force from the date of its official publication—6 January 2000.

This Agreement underlines the international legal capacity of JINR and grants it privileges and immunities in compliance with established practice for international intergovernmental organizations.

1.2 Governing and advisory bodies of JINR

– Committee of Plenipotentiaries of the JINR Member States
– Finance Committee
– Scientific Council
– Programme Advisory Committee for Particle Physics
– Programme Advisory Committee for Nuclear Physics
– Programme Advisory Committee for Condensed Matter Physics.

The Committee of Plenipotentiaries of the Governments of the Institute Member States is the supreme body of the JINR and carries out its activities in the session order. Each member of the Institute has one representative in the Committee of Plenipotentiaries.

Scientific policy is developed and co-ordinated by the Scientific Council, whose members are eminent and well-known scientists from the Member States, as well as from CERN, Germany, Greece, France, Italy, Belgium, the Netherlands, China, India, and other countries.

Three Programme Advisory Committees (PACs) for Particle Physics, Nuclear Physics and Condensed Matter Physics are advisory bodies of the JINR Directorate and to the JINR Scientific Council in specific fields of research. The PACs hold their meetings twice a year.

1.3 JINR's structure and the main fields of research activities

The internal organization of the JINR is determined by scientific specialization. There are seven Laboratories at JINR and, by the scope of their scientific activities, each is comparable to a large research institution:

– Bogoliubov Laboratory of Theoretical Physics (BLTP)
– Veksler-Baldin Laboratory of High Energies (VBLHE)
– Laboratory of Particle Physics (LPP)
– Dzhelepov Laboratory of Nuclear Problems (DLNP)
– Flerov Laboratory of Nuclear Reactions (FLNR)
– Frank Laboratory of Neutron Physics (FLNP)
– Laboratory of Information Technologies (LIT).
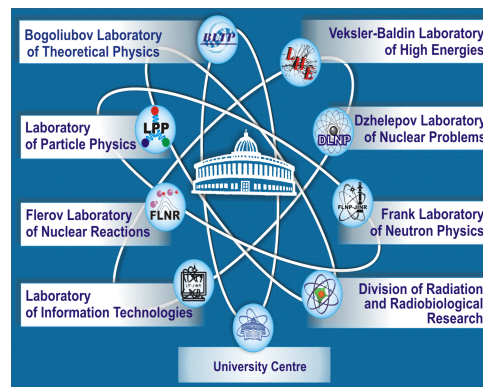
All-Institute subdivisions are

– Division of Radiation and Radiobiological Research
– University Centre.

Several associate experimental physics workshops are also part of the Institute. It is equipped with everything necessary for manufacturing large-sized, non-standard facilities, electronics, and has technological lines for constructing detectors for physics.

In the past four decades JINR has grown into a large multidisciplinary physics centre. It employs over 5500 people, including 1300 scientists and about 2000 engineers and technicians.

The main fields of the Institute's activities are as follows.

– **Theoretical Physics**
Quantum field theory and modern mathematical physics; Fundamental symmetries; Standard Model and beyond; Astroparticle physics; Nuclear structure near the drip line; Dynamics of few-body systems; Exotic properties of nuclear matter; Systems with strong correlations; Integrability; Self-organization; Disordered structures.

– **Elementary Particle Physics**
Origin of mass; Nature of spin; Fundamental symmetries (including chiral symmetry); Nature of dark matter; Neutrino mass; Deconfinement; Search for supersymmetry.

– **Relativistic Nuclear Physics**
Non-perturbative QCD; Spin effects in hadronic processes; Quark–gluon degrees of freedom; Asymptotics in nuclear collisions; Mechanism of hadronization and confinement; Heavy-ion collisions; Multiple production; Nuclear multifragmentation processes; Hypernuclei and η nuclei; Cumulative effects; Spin structure of light nuclei; Physics of resonances; Nuclotron; Superconducting magnet technology.

– **Heavy-Ion Physics**
Nuclear reactions induced by beams of stable and radioactive nuclei; Synthesis and properties of transuranium and superheavy nuclei; Properties of nuclei close to proton and neutron drip lines; Chemistry of actinides and transactinides; Radioanalytical investigations; Heavy-ion accelerators; Production of secondary beams; Interaction of heavy ions with condensed matter.

– **Low- and Intermediate-Energy Physics**
Fundamental physics phenomena and processes in nuclear physics; Rare decays of elementary particles and nuclei; Nonaccelerator particle physics; Nature and properties of the neutrino.

– **Nuclear Physics with Neutrons**
Violation of fundamental symmetries in neutron-induced reactions; Beta-decay and electromag-

netic properties of the neutron; Ultracold neutrons; Nuclear data for science and technology; Ecological study with neutrons.

– **Condensed-Matter Physics**

Strongly correlated electron systems; Low-dimensional systems; Heterostructures; Quantum wells and dots; Quantum liquids; Chaos; Self-organization; Disordered systems; Polymers; Biopolymers; Biomembranes; Nanomaterials; Physics of surfaces.

– **Radiation and Radiobiological Research**

Radiobiology; Radiation genetics; Mutagenesis; Chromosomal aberration; Biophysics of photobiological processes; Target therapy; Radiation protection; Dosimetry; Neutron spectrometry; Radiation transport through matter.

– **Networks, Computing and Computational Physics**

Distributed high-performance computing infrastructure; High-speed networking; Information, algorithmic, and software support; Modelling of physical systems; Data processing and analytic calculations for physics problems; JINR's Grid segment and global Grid structures.

– **Educational Programme**

University-type education; Continuous education 'secondary school–higher education institution–research institution'; Postgraduate programmes; Extension of specialties; Collective use centres; International schools; JINR-based educational departments.

A unique choice of experimental facilities is available at this Institute. Apart from the still operational early accelerator, the 680 MeV Phasotron, they include the Nuclotron, a new, superconducting synchrotron for nuclei and heavy ions up to 6 GeV/n intended for relativistic nuclear physics studies; the U400 and U400M cyclotrons used for experiments on the synthesis of heavy and exotic nuclei, on the studies of their properties and heavy-ion reaction mechanisms; the IBR-2 reactor (mean power 2 MW, peak power 1500 MW) used for nuclear physics research with neutrons and condensed-matter studies. Also, several new facilities are being constructed at JINR, namely IREN, a new source of resonance neutrons, and DRIBs, the Dubna Radioactive Ion Beams project.

2 JINR scientific policy and the road map

2.1 Worldwide recognized traditions of scientific schools

The Institute is proud of the prominent scientists who worked at JINR. They made outstanding contributions to JINR's research programme and established scientific schools at Dubna. Among them is the famous theoretical physics school founded by professors N. Bogoliubov, D. Blokhintsev and M. Markov.

Another school on neutrino physics was founded by B. Pontecorvo who made the supposition about neutrino oscillations. Professor V. Veksler, a distinguished scientist, is the author of the Phase Stability Principle, being a base for modern accelerators. Professor G. Flerov is a prominent scientist and founder of heavy-ion physics at JINR. Laureate of the Nobel and State Prizes, Professor I. Frank made important contributions to the formation and development of various directions in physics: electrodynamics of a moving charged relativistic particle; nuclear and especially neutron physics. Relativistic Nuclear Physics is a new scientific direction established at JINR by Professor A. Baldin.

A very big contribution to the formation of the JINR scientific schools was made by the prominent scientists from the JINR Member States (see introduction). Our obligation and key strategic goal is to preserve the traditions of JINR schools and to train young scientists in the spirit of these traditions.

2.2 International collaboration

Broad international co-operation is one of the most important principles of JINR's activities. Almost all investigations are carried out in close collaboration with the JINR Member States scientific centres, as well as international and national institutions and laboratories around the world. JINR collaborates with

nearly 700 research centres and universities in more than 60 countries, including Germany, France, Italy, Japan, Switzerland, and the USA.

Over its 49 years of existence, first-class theoretical and experimental research programmes accomplished at JINR have led to a significant enrichment of fundamental nuclear science. JINR accounts for a half (about 40) of the total number of discoveries in particle and nuclear physics, registered in the former Soviet Union. About 500 research papers and reports representing approximately 3000 authors are submitted every year by JINR to editorial boards of journals in many countries and to organizing committees of conferences. JINR publications are sent to over 50 countries.

The Institute carries out theoretical and experimental research using its own basic facilities and those of other major scientific centres throughout the world. These facilities provide ample and unique opportunities for research in high-energy physics as well as in low- and intermediate-energy physics. Widely used are novel information technologies and computational physics methods, which, on the whole, maintain a modern level of research.

Here I would like to recall the words of the great Russian writer A. Chekhov, who said: *"Science cannot be national, in the same way that a multiplication table cannot be national. If a science becomes national, it ceases to be a science."*

JINR is a perfect illustration of this idea.

2.2.1 Co-operation with CERN

Our long-standing scientific partner is CERN. For more than 40 years the co-operation between JINR and CERN has been very fruitful and mutually beneficial. Though the general agreement between JINR and CERN was signed in 1992, nevertheless, the real co-operation between the two international organizations has a long and rich history. JINR scientists and engineers are actively involved in the current CERN experiments as well as in the preparation of future LHC projects. Today JINR participates in twenty-seven CERN projects. Among them: ATLAS, CMS, ALICE, LHC (accelerator complex), DELPHI (data analysis), DIRAC, HARP, NA45, NA48/1, NA48/2, NA49, NA58 (COMPASS) and others.

2.2.1.1 Science bringing nations together

A special page in our co-operation with CERN is the joint poster exhibition 'Science Bringing Nations Together'.

Since 1997 when the first joint exhibition of this series was held at the University of Oslo (Norway), CERN and JINR have organized this exhibition every year. The exhibitions have been shown at UNESCO (1998), at the United Nation's office in Geneva (1999), in the European Parliament in Brussels (2000), in Russia State Duma (2001), in Romania (2002), in Armenia (2003), in Russian Diplomatic Academy (2003), and in Greece (2005).

The dominant theme of the exhibition is that joining of creative efforts and material resources by scientists from various countries has become another important way for fruitful communication of peoples and mutual understanding between them.

2.2.2 Co-operation with Germany

The Joint Institute has very fruitful relations with scientific centres in Germany. Since 1991, JINR has a special co-operation agreement concluded on the governmental level—namely BMBF—with Germany. At present JINR co-operates with 71 centres in Germany located in 47 cities.

Today the co-operation between JINR and German Scientific Groups is based on

– the BMBF–JINR Agreement, and
– Bilateral Agreements concluded between JINR and German Scientific Groups.

The co-operation in the field of High-Energy Physics at DESY was included in the JINR–BMBF Agreement in 1995. The Laboratory of Particle Physics of JINR and DESY co-operate in the experimental programme at the HERA electron–proton collider and in the R&D programme for the TESLA linear collider and Free-Electron Laser (FEL). The Dubna group has made important contributions in all fields of this collaboration.

In the HERMES Collaboration, the Dubna group has taken an active part in data analysis and detector upgrades. Many thousands of events for the deeply virtual Compton scattering processes have been collected at HERMES.

2.2.3 Co-operation with the USA

JINR is implementing a wide scientific collaborative programme with US research laboratories. This co-operation is carried out in the fields of high-energy physics, heavy-ion physics, nuclear physics with neutrons, and accelerator technologies. At present, the Joint Institute for Nuclear Research collaborates with 75 US scientific centres located in 66 cities.

JINR is developing a successful collaboration with FNAL. At this Laboratory, groups of JINR scientists have been taking part in the experiments using the CDF and D0 detectors.

Another good example is the active collaboration with the Brookhaven National Laboratory. JINR scientists are participating in the design and construction of the electromagnetic calorimeters for the STAR detector.

2.3 JINR's Road Map

First of all I would like to stress once more that fundamental studies remain the core mission of the Joint Institute. The 'old' challenging questions 'What is the world made of?' and 'What holds it together?' are still capturing the imagination and professional interests of physicist all over the world.

The evaluation of the Universe beginning from the Big Bang up to its modern state is a subject of particle physics studies (mainly at the early stages), as well as of nuclear and condensed-matter studies at later stages. The JINR research programme covers all three research directions. The investigations are being carried out both at in-home experiments and external experiments with JINR's participation. Thereby JINR enhances its role as a cluster scientific centre for its Member States, giving good opportunities for scientists to participate in research programmes of many other well-known centres in the world.

I would like to note that, since its very foundation, JINR has developed as a centre of Particle Physics. The first basic facilities of JINR—the synchrocyclotron and synchrophasotron—were at that time HEP accelerators with record parameters. Gradually the methods and approaches used in high-energy physics had an essential influence on other branches of science which were developed as new research directions at JINR: condensed-matter physics, nuclear physics with neutrons, radiobiology, and others.

We also consider our innovation activities to be of high importance. We must bring them to the commercial level, to get into the world high-tech market.

2.3.1 The Road Map in the field of Particle Physics

To ensure the scientific excellence of JINR and maximize its scientific and technological output, we recently started to elaborate the Road Map—a strategic plan of the Institute's development in the fields of particle physics, nuclear physics, and condensed-matter physics for the next 10–15 years. The development of the Road Map is based on JINR's seven-year Programme (2003–2009) and our annual document—the Topical Plan for Research and International Co-operation. At the same time we consider

the Road Map as an important step for strategic planning of the overall research activity of the Institute based on the three main scientific directions: particle physics, nuclear physics, and condensed-matter physics. We intend to focus our efforts on the implementation of really ambitious projects.

When developing first proposals for the Road Map in the field of particle physics (as well as in other fields) we proceeded from the priorities in this area accepted today by the international physical community and our real capabilities. In the light of these statements we single out the following priorities:

- the origin of mass;
- the properties of neutrinos;
- the properties of the strong interaction including properties of nuclear matter;
- the origin of the matter–antimatter asymmetry in the Universe;
- the unification of particles and forces including gravity.

JINR's research programme in the field of particle physics is aimed at solving these tasks based on the Nuclotron and external facilities.

Of special interest are studies on nucleon spin structure (Fig. 2) with JINR's participation in DESY (HERMES experiment) and CERN (Compass project). In the future this research will take place in Serpukhov at the U-70 accelerator and also in the PAX experiment—a spin physics experiment at FAIR (GSI, Germany).
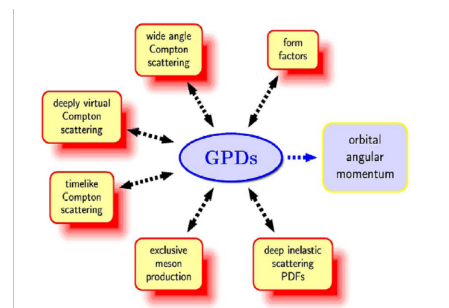
Fig. 2: Generalized Parton Distributions (GPD)

Neutrino physics is a traditional research direction for JINR and mainly connected with the name of Professor Bruno Pontecorvo, one of the founders of neutrino physics schools. The NOMAD and HARP experiments are almost completed and at present JINR is planning to participate actively in the OPERA experiment connected with tau neutrino appearance.

Another ambitious task in the field of particle physics is the CP-violation effect. Here we must mention first of all the NA48 experiment at CERN (Fig. 3). You know that the first evidence of direct CP violation has been obtained. This result was achieved in neutral kaon decays by the ε'/ε measurement with the world's highest precision. Professor V. Kekelidze from JINR occupies the position of the NA48/2 Spokesperson until the end of 2005.

The new estimate of the upper limit for K_L decay on pion and two neutrinos branching ratio was obtained 2×10^{-7}; confidence level is 90%. The JINR group contributed significantly to this experiment as well.

The verification of the Standard Model and the tasks beyond it are also traditionally in the focus of our interests. As an example I would like to mention a big success which has been achieved recently in the most precise top-mass measurement experiments (CDF and D0 projects) with the participation and remarkable contribution of JINR scientists (Fig. 4).

The JINR–INFN–FNAL team made a contribution of principal significance to the world's most precise M_{top} measurement in the so-called 'lepton + jets' topology. The efforts of the team are aimed at significantly reducing the error of the M_{top} up to the 2 GeV/c^2 level which will enable us to establish a new limit on the Higgs mass.
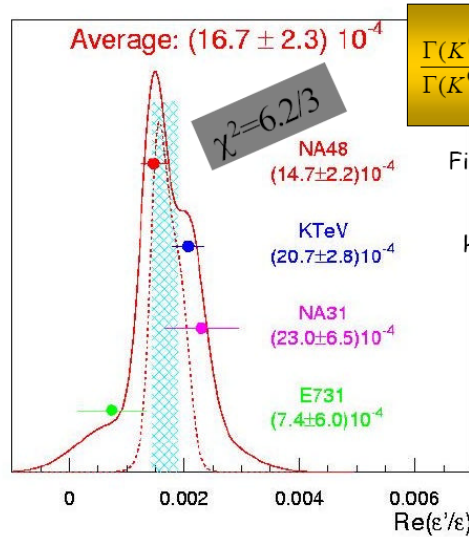
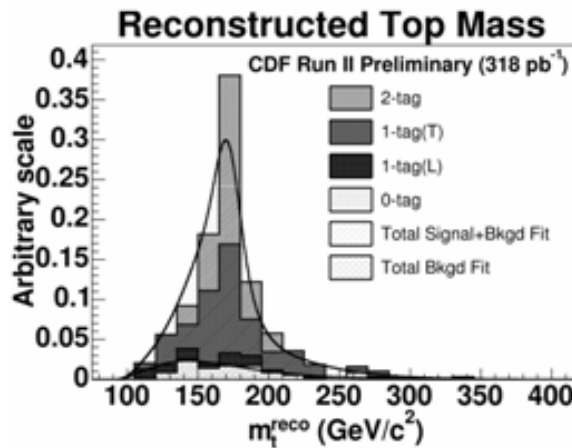**Fig. 3:** World best measurement of direct CP violation in K^0 decays

Fig. 4: World best: $M_{\text{top}} = 173.5 + 3.7/ - 3.6(\text{stat.}) \pm 1.3(\text{syst.})$ 'lepton + jets' mode

The Nuclotron accelerator complex

The Nuclotron acclerator (Fig. 5) is competitive with other world ion facilities for the acceleration of polarized deutrons and helium ions with an intensity higher than 10^{11} particles/cycle. It can accelerate heavy ions up to uranium.

In the next few years the Nuclotron will be equipped with new sources both for polarized light ions and for heavy ions. The energy will be increased up to 6 GeV/nucleon. That will allow one to investigate spin structure of nucleons and nuclei and to study phase transitions in nuclear matter.

Among the challenging tasks carried out at the Nuclotron, of special interest is a search for the mixed phase corresponding to the first-order chiral phase transition and probably to the deconfinement transition (Fig. 6).

It is predicted by a group of JINR theoreticians (A. Sissakian, A. Sorin, M. Soleimanov, V. Toneev, G. Zinovjev *et al.*) that at the highest possible energies of the Nuclotron heavy-ion beams the mixed phase formation may become possible. This will open a new perspective in the physics programme of the Nuclotron.

Fig. 5: Nuclotron (superconducting synchrotron)
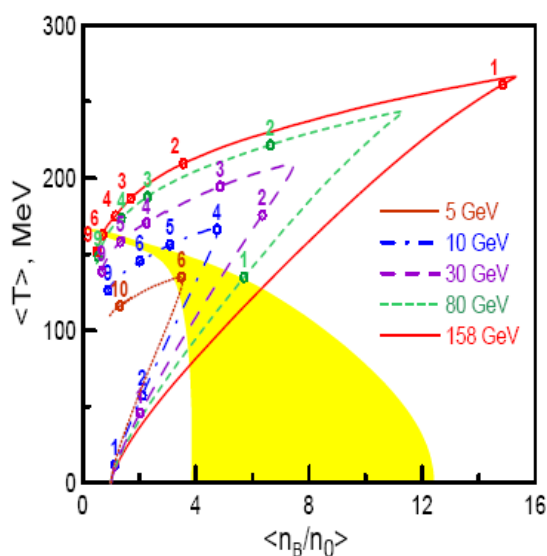
Fig. 6: Heavy-ion collisions in relativistic 3-fluid hydrodynamics

2.3.2 The Road Map in the field of Nuclear Physics

The priorities in heavy-ion physics are

- physics and chemistry investigations of the super heavy nuclei with $Z \geq 112$; structure and properties of the neutron reach light exotic nuclei;
- acceleration technology;
- heavy-ion interaction with matter and applied research.

To accomplish these tasks the JINR FLNR Cyclotron Complex will be upgraded for producing intense beams of accelerated ions of stable (^{48}Ca, ^{58}Fe, ^{64}Ni, ^{86}Kr) and radioactive (^{6}He, ^{8}He) isotopes.

The main experimental facility of the Flerov Laboratory of Nuclear Reactions is the complex of two heavy-ion cyclotrons—U400 and U400M. For the last decade, heavy-ion physics has been the most dynamically developing area of low- and intermediate-energy physics and JINR has become one of the leading scientific centres in heavy-ion physics in the world.

Let me remind you that at the end of 1998, scientists of JINR, in collaboration with colleagues from the Lawrence Livermore National Laboratory (USA), synthesized a new, long-lived superheavy element with atomic number 114. This discovery crowned 35 years of international research efforts in search of the 'stability island' for superheavy nuclei. A wide resonance followed new experiments accomplished in 1999–2005 at the JINR U400 accelerator on the synthesis of new elements with atomic numbers 116, 118, 115 and 113.
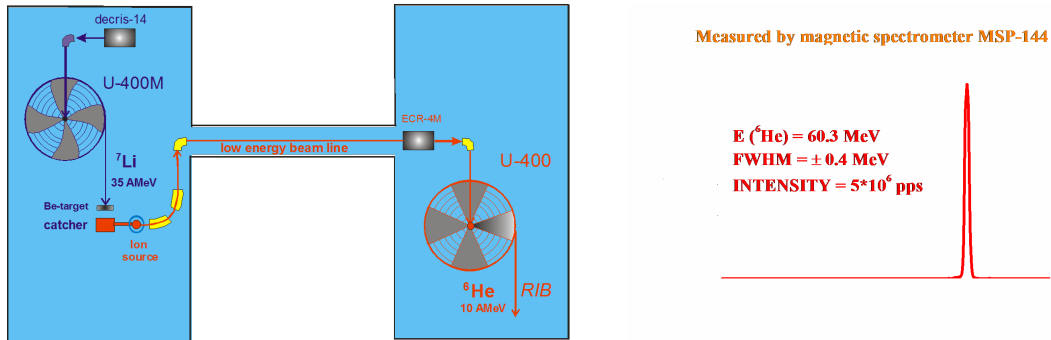
Fig. 7: Cyclotron complex of JINR FLNR - DRIBS-1

During the last few years, the DRIBs facility (Dubna Radioactive Ion Beams) has been developed at JINR. The main task of this machine is to produce intense beams of accelerated ions of stable and radioactive isotopes. The first experiment with the ^{6}He radioactive beam was carried out at DRIBs in 2005. The world record for intensity (5×10^6 pps) with this beam on the target was achieved (Fig. 7)

2.3.3 The Road Map in the field of Condensed-Matter Physics

The priorities are

– modernization of the IBR-2,
– condensed-matter research with neutrons,
– material science with heavy ions,
– radiobiological research.

The programme in the field of condensed-matter physics is oriented towards the use of nuclear-physical methods developed at JINR to solve topical problems of present-day natural sciences, concerned with the properties of matter in a condensed state.

The main facility here is the IBR-2 pulsed reactor (Fig. 8) which is in a phase of modernization. As a result JINR will have in operation a unique world-class pulsed neutron source—the only machine of this sort in the JINR Member States. Its parameters will be unique in many aspects which will allow a rich research programme for another 20–30 years. IBR-2 is included in a 20-year strategic programme of neutron scattering research in Europe.

Summarizing this part of my lecture I would like to note that while developing the Road Map we should take due account of the main supporting activities:

– theory of particle physics, nuclear physics, condensed-matter physics
– modern mathematical physics
– networking and computing
– training of young staff
– physics methods.

Fig. 8: IBR-2 pulsed reactor with neutron flux 10^{16}n cm^{-2}s^{-1}

Fig. 9: V.G. Kadyshevsky and A.N. Sissakian

3 Dubna as an educational centre

The educational programme plays an important role in the activities of JINR. It should be stressed that the concept of JINR's development is the integration of fundamental science, technological studies, and education. To achieve this task, in 1991 we established the JINR University Centre (Head: Dr S. Ivanova) and in 1994, together with the Russian Academy of Natural Sciences, the authorities and management of Moscow Region the 'Dubna' International University for Nature, Society, and Man (President: Academician V. Kadyshevsky, Vice-president: Professor A. Sissakian (Fig. 9), Rector: Professor O. Kuznetsov).

Since 1995, the University Centre of JINR has been offering postgraduate training. The University Centre offers graduate programmes in the fields of nuclear physics, elementary particle physics, condensed-matter physics, theoretical physics, technical physics, and radiobiology.

Our strategic plan is to develop JINR as a kind of 'superuniversity' centre with the aim of training specialists from the JINR Member States and other countries. These specialists will be engaged in the research activities in Dubna, as well as possibly joining future megaprojects like the LHC, TESLA, and others.

4 Innovation activities

Along with fundamental research, which is the main direction of JINR's development, we have a large number of applied studies and activities, including development of high technologies:

– New technological developments, R&D and construction of detectors and large facilities for scientific research and applied studies.
– Multichannel amplitude analysers, fixed and portable spectrometers of nuclear radiation, portal monitors.
– Development of IT, including Grid technology
– Ecological monitoring using methods of nuclear physics analysis.
– New materials based on track membranes.
– Medical beams at the Phasotron, JINR's Med-Nuclotron.
– Ultrapure radioisotopes for nuclear medicine and ecology.
– Detector 'DVIN' for identification of explosives, fissile materials, and narcotic substances.

At the present time, we have a good basis for developing the so-called 'innovation belt' around the Institute. Based on the huge intellectual and industrial–technological potential of JINR, it is proposed to actively participate in the creation of a 'Dubna' technopark (including development of the JINR 'innovation belt') as well as in the use of the mechanism of private and state partnership for creating a Special Economic Zone in Dubna, in accordance with the legislation of the host country of our Institute. JINR will play the role of scientific leader in this partnership. The well-known Russian company AFK 'Sistema' will act as business partner. The administration of the city, the 'Dubna' international university, and a number of Dubna enterprises are also engaged in the process of technopark creation.

5 JINR's long-term research programme

Elaborating on the JINR Road Map it is noteworthy to consider a proposal of an ambitious large-scale project to be created in Dubna. Among the possible megaprojects we are considering is the ILC project (International Linear Collider, see Fig. 10).
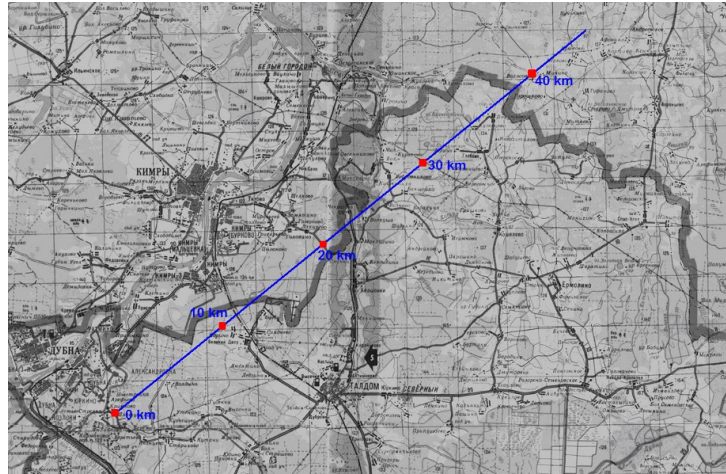

Fig. 10: Possible location for the ILC in the vicinity of Dubna (Russia)

ILC is a collaborative international project aimed at the design of a high-energy, large-scale positron–electron collider. The ILC would provide a tool for scientists to address many of the most challenging tasks of the 21st century. We consider the proposal on ILC (or may be some other proposals on megaprojects) as a 'maximum programme' of JINR's strategic development for approximately 20–25 years ahead.

6 Conclusion

We understand well that science is united. The methods, experience, and knowledge accumulated in high-energy physics research could be useful in other sciences too. We should think it over in order to find an appropriate interface with other actively developing branches of knowledge, for instance astroparticle physics and cosmology, informatics, nanotechnologies, biology and medicine, quantum computers, and others.

Acknowledgements

The physics of ski jumping

W. Müller
Department of Biophysics, Medical University of Graz, Austria

1   Introduction

The Olympian ideal of going faster, jumping further and leaping higher than the opposition is central to competitive sports. Winning or losing in sport is related to a number of factors, and biomechanics, anthropometrics, and aerodynamics play a major role in many sports. This lecture focuses on ski jumping.

Performance in ski jumping is determined not only by the motor abilities of the athlete, but also to a large extent by the aerodynamic features of the equipment used and by a low body weight. Many ski jumpers were extremely underweight to the point of having a body mass index (BMI = m/h^2 of 16.4 kgm^{-2} (height h = 1.73m, body mass m = 49 kg).

Severe eating disorders (e.g., anorexia nervosa, bulimia; [1]) were health problems of major concern in this sport. Strategies for improving the health, fairness, and safety of the athletes by modifying the regulations have been developed by the lecturer and his research team in close co-operation with the International Ski Federation (FIS), the International Olympic Committee (IOC), and the Austrian Research Funds (FWF). Based on our scientific studies the FIS has passed changes to the ski jumping regulations which relate relative body weight (in terms of BMI) to the maximum ski length permitted. Shorter skis (i.e., 'smaller wings') compensate for the advantage of very low weight and thus it is not attractive for the athletes to be underweight any more [2].

Our analyses of contemporary ski jumping employ field studies during World Cup and Olympic Games competitions, wind tunnel measurements, computational fluid dynamic (CFD) modelling of aerodynamic forces and torques, computer simulations of the flight trajectory, and computer-modelling-based design of jumping hills [3–7].

2   The dynamics of ski jumping: a brief description

Ski jumping puts high demands on the athlete's ability to control posture and movement. During the in-run the athlete tries to maximize acceleration by minimizing both the friction between skis and snow and the aerodynamic drag in order to obtain a maximum in-run speed v_0, which has a high degree of influence on the jump length. The friction between skis and snow is not well understood. The physics text book solutions to this problem do not reflect reality. The theoretical as well as the empirical basis for these complex problems are not sufficiently developed. The reduction of aerodynamic drag in the in-run phase is primarily a question of the athlete's posture and his dress. Owing to the curved form of the in-run just before the ramp, the athlete has to counteract the centrifugal force acting on him (as seen from the athlete's point of view) and this phase is immediately followed by the athlete's acceleration perpendicular to the ramp due to the muscular forces exerted. During this decisive phase of approximately 0.3 s duration the athlete has to produce a maximal momentum mv_{p0} perpendicular to the ramp (m: mass of the athlete plus equipment) through which an advantageous take-off angle has to be attained. The take-off velocity vector $\vec{v_{00}}$ is given by $\vec{v}_{00} = \vec{v}_0 + \vec{v}_{p0}$ with $\vec{v}_{p0}$ being the velocity perpendicular to the ramp due to the athlete's jump. Simultaneously, the athlete must produce an angular momentum forwards in order to obtain an advantageous angle of attack as soon as possible after leaving the ramp. During the jump phase the athlete must anticipate the magnitude of the backward torque due to the air-stream so that the forward rotation will be stopped at the right moment. If the forward angular momentum is too low, a disadvantageous flight position reduces velocity and, therefore, results in bad competitive performance. Worse is the production of too much forward angular momentum because this substantially increases the tumbling risk. During the flight the gravitational force F_g, the lift force F_l, and the drag force F_d act upon the athlete:

$$F_{\mathrm{g}} = mg\,; \quad F_{\mathrm{l}} = \frac{\rho}{2}v^2 c_{\mathrm{l}} A = \frac{\rho}{2}v^2 L\,; \quad F_{\mathrm{d}} = \frac{\rho}{2}v^2 c_{\mathrm{d}} A = \frac{\rho}{2}v^2 D\,.$$

The velocity of motion along the flight path v has the components $\dot{x}$ and $\dot{y}$

$$v^2 = \dot{x}^2 + \dot{y}^2\,.$$

The athlete can strongly influence the aerodynamic forces by changing his posture. He can affect the drag force, the lift force and the torque, and thereby significantly influence changes in his flight position relative to the air stream. The flight path is described by the following non-linear differential equations which can be solved numerically by using proper iterative procedures:

$$\dot{v}_x = \frac{(-F_{\mathrm{d}}\cos\varphi - f_{\mathrm{l}}\sin\varphi)}{m} \qquad\qquad \dot{v}_y = \frac{(-F_{\mathrm{d}}\sin\varphi - f_{\mathrm{l}}\cos\varphi)}{m} - g\,.$$

$$\dot{x} = v_x \qquad\qquad\qquad\qquad\qquad\quad \dot{y} = v_y$$

In order to achieve highly realistic computer simulations, it is necessary to be able to simulate changes in posture and position during the simulated flight, i.e., changes in the resulting aerodynamic forces. We developed such a computer model of ski jumping in 1995 [3,4].

2.1 Basic aerodynamic problems

Aerodynamic questions related to sports are complex and manifold. For this reason the influencing phenomena should be investigated by both theoretical and experimental approaches.

2.1.1 *Theoretical approach*

The Navier–Stokes equations which describe the dynamics of Newtonian fluids have inherent major mathematical difficulties. Exact solutions are only possible in special cases involving objects with simple geometries that do not exist in sports. The numerical solution becomes increasingly difficult as Reynolds numbers increase, even when supercomputers are used for the numerical solution of these non-linear partial differential equations. Owing to the non-linearity of the equations, a variation of the geometrical and fluid mechanical parameters can result in bifurcations and the non-linear fluid system can display deterministic chaos. Computational fluid dynamics is proving to be invaluable at the early stage of trend analysis prior to prototype testing in several kinds of sports (e.g., Formula 1, yachting). Initial studies of ski jumping have also been made. However, the thin boundary layers around moving bodies have to be resolved accurately and the associated physical effects are notoriously difficult to predict accurately in CFD. A combination of measurement and CFD is the state-of-the-art approach necessity when appropriate predictions for sports aerodynamics are desired. The best configurations found in wind tunnel tests still have to be tested in the field by the athletes before their competition debut. In ski jumping (and many other sports as well) the characteristic dimensions of the body and/or equipment and the typical velocities result in Reynolds numbers between $\mathrm{Re} = 10^4$ and $\mathrm{Re} = 10^6$ where pronounced changes in the drag coefficient may occur: $c_{\mathrm{d}} = c_{\mathrm{d}}(\mathrm{Re})$. This was already shown for the sphere in 1912 by G. Eiffel [8] and in 1914 by C. Wieselsberger [9]. The sensitivity of the transition from laminar to turbulent flow on the roughness of the surface (or on small surface obstacles) was also shown. The theoretical approach describing the lift forces is at least as difficult as the discussion of the drag, and science still does not have a complete understanding when turbulence phenomena occur. Micro effects can be the starting point for major flow changes in a system and to discount them in modern turbulence models or even to ignore them leads to completely inaccurate predictions of the whole process.

2.1.2 *Experimental approach*

Small changes in the lift and/or drag coefficients can have pronounced consequences for the sport concerned. Therefore, accurate measurements in the wind tunnel are necessary when these data are to be

used for subsequent mapping of 'real world' problems. A wind tunnel with a sufficiently large cross-section is necessary to minimize blockage effects. Because the athletes must be studied in various postures in full gear, adequate positioning devices are required. However, these devices can cause secondary errors even when the forces acting solely on the devices are considered. Positioning devices that lead to small disruptions in the air flow around the athletes need to be designed. The experimental problems are smaller for those questions where only relative changes in the aerodynamical parameters are to be considered. However, the design of experiments reflecting the special circumstances of different sports is not trivial at all. A reliable interpretation of the effects associated with different aerodynamical characteristics usually necessitates a computer-based analysis of the experimental data. So, for instance, a given increase of the drag area $D = c_{\mathrm{d}} A$ during the initial flight phase reduces the jump length much more than would be the case during the final phase of the flight. Analogously, a change in the drag coefficient of the athlete in the crouched position may occur during the in-run phase due to a Reynolds-number increase beyond the critical value (e.g., when the velocity increases). These phenomena may strongly depend on individual body dimensions of the athletes. Therefore, a computer-assisted analysis of the resulting effects based on experimental aerodynamic data is the only way of appropriate treatment (based on the equations of motion for the in-run phase). For all kinds of sports where a minimal aerodynamic drag is important, the factors influencing the shift of the drop in the drag coefficient within the critical Reynolds-number range are most important. These factors have not been sufficiently understood in the context of sports involving complex objects like human bodies. The surface structure (dress, jumping suit, downhill suit, etc.), temperature effects, vibration of surfaces like human skin, tension of the surface material etc. are all influential factors. Considering the significant influence of, for example, the geometry of a sphere versus an aerodynamically well-shaped object on the lift and drag coefficients (the difference can be a factor of 100; see, for example, H. Schlichting and K. Gersten [10]) the predominant importance of improvements in this direction is evident.

Very little is known about the drag and lift forces acting on the complex structure of a human body in an air-stream. From the aerodynamic data according to flight positions it is evident that the lift forces in ski jumping are of the same magnitude as the drag forces and that the flight length is very sensitive to changes in both. It is well known that pronounced changes of the lift coefficients of wings occur in the critical range of the Reynolds number, and it has been found from measurements in low-turbulence wind tunnels that the degree of turbulence of the outer flow has an important influence on the lift coefficient of a wing. Yet, no systematic study of these characteristics in relation to human bodies in the air-stream has been made. Our knowledge of wings cannot be adequately transferred to the flow around human bodies because the geometric form is not at all similar and wings usually do not work at angles of attack of up to 50 degrees.

Acknowledgements

References

[1] K. Sudi, D. Payerl, P. Baumgartl, K. Tauschmann and W. Müller, Anorexia athletica, Nutrition **20** (2004) 657–661.

[2] W. Müller, W. Gröschl, B. Schmölzer and K. Sudi, Body weight and performance in ski jumping: The low weight problem and a possible way to solve it, 7th IOC Olympic World Congress on Sport Science, Athens, 2003, 43 D.

[3] W. Müller, D. Platzer and B. Schmölzer, Scientific approach to ski safety, Nature 375 (1995) 455.

[4] W. Müller, D. Platzer and B. Schmölzer, Dynamics of human flight on skis: Improvements on safety and fairness in ski jumping, J. Biomech. **8** (1996) 1061–1068.

[5] W. Müller, Biomechanics of ski-jumping: Scientific jumping hill design, *in*: Müller E., Schwameder H., Kornexl E., Raschner C. (Eds.), Science and Skiing (Spon, London, 1997), pp. 36–48.

[6] B. Schmölzer and W. Müller, The importance of being light: Aerodynamic forces and weight in ski jumping, J. Biomech. **35** (2002) 1059–1069.

[7] B. Schmölzer and W. Müller, Individual flight styles in ski jumping: Results obtained during Olympic Games competitions, J. Biomech. **38** (2005) 1055–1065.

[8] G. Eiffel, Sur la resistance des sphères dans l'air en mouvement, Comptes Rendus **155** (1912) 1597–1599.

[9] C. Wieselsberger, Der Luftwiderstand von Kugeln, Z. Flugtech. Motor-Luftschiffahrt, **5** (1914) 140–144.

[10] H. Schlichting and K. Gersten: Grenzschicht-Theorie (Springer, 1997), p. 48.

[11] E. Reisenberger, W. Meile, G. Brenn and W. Müller, Aerodynamic behaviour of prismatic bodies with sharp and rounded edges, Exp. Fluids **37** (2004) 547–558.

Figures

Examples of experimental results
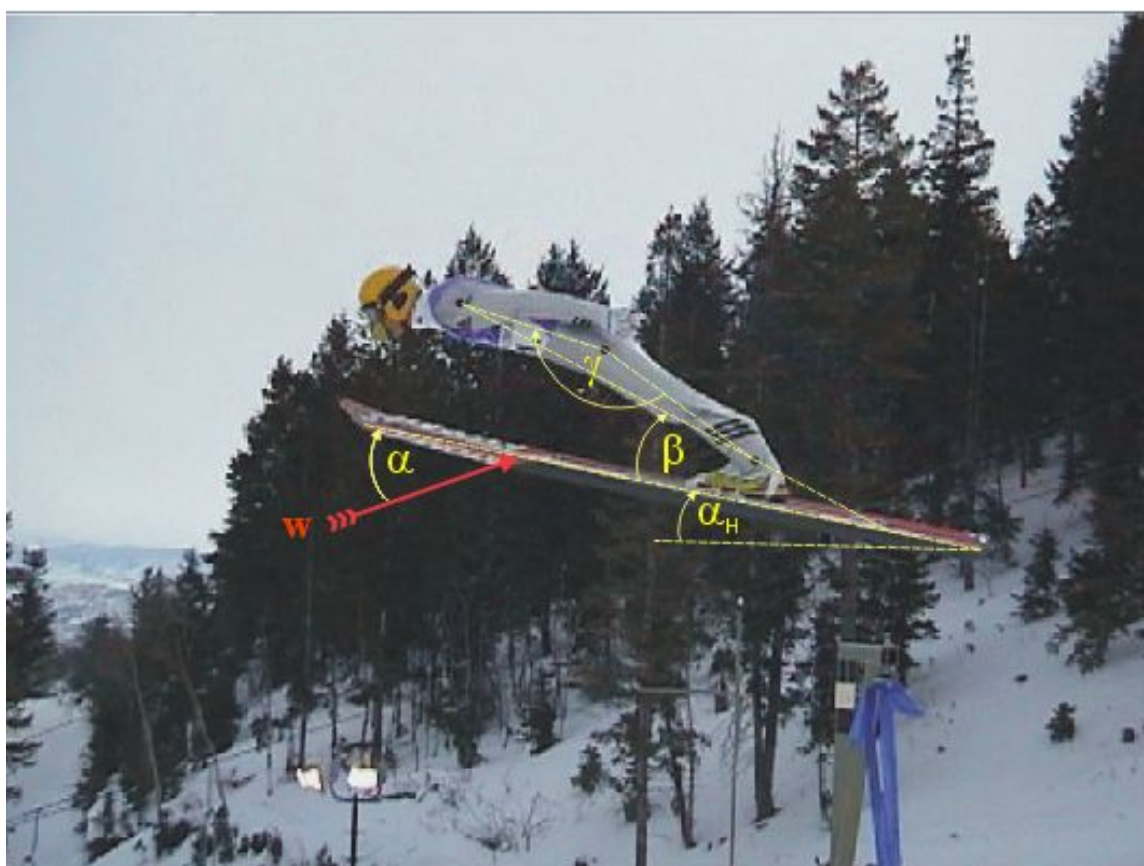
Fig. 1: A ski jumper at the $K = 120$ m jumping hill in Park City. The nomenclature used for the position angles is indicated, w is the resulting air stream vector.
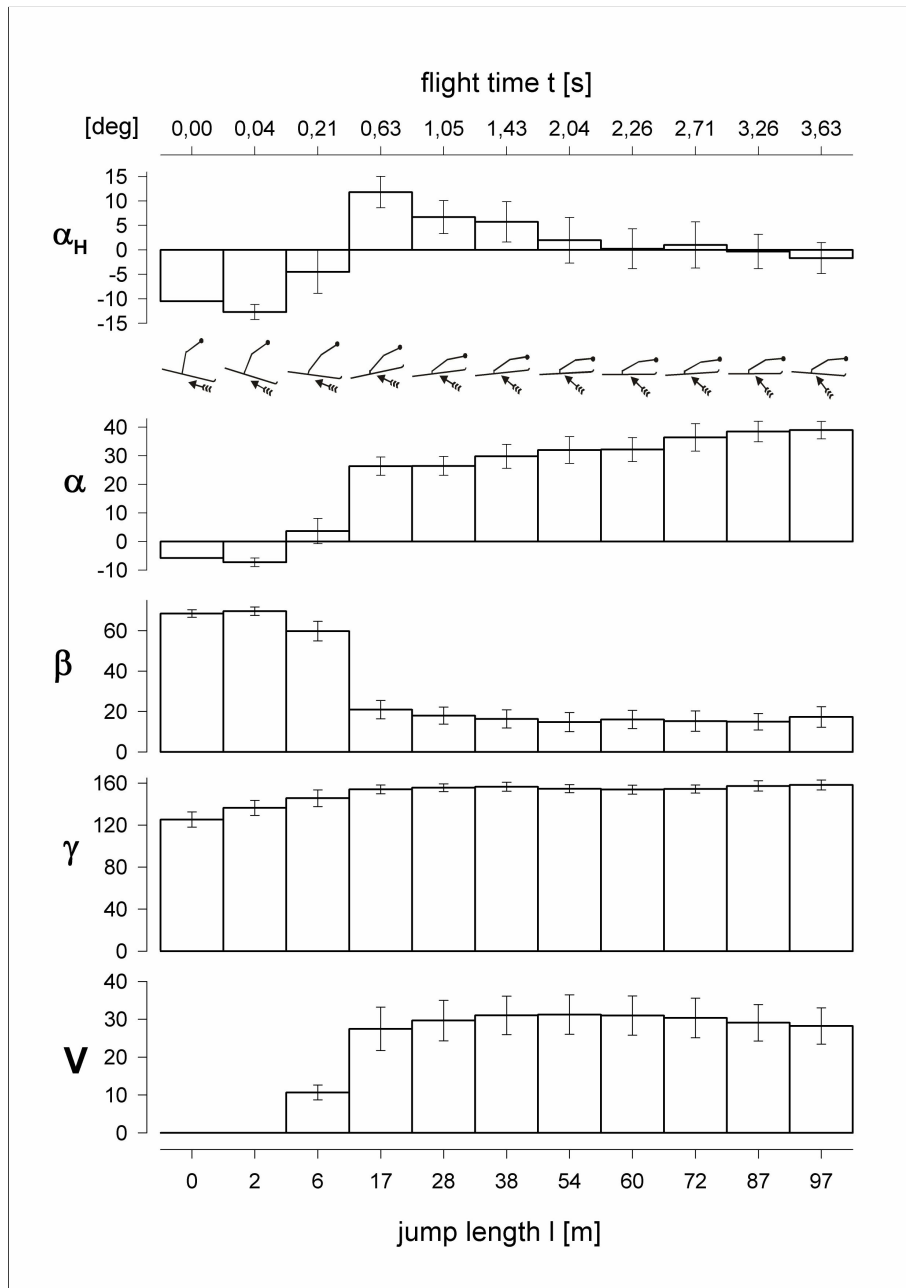
Fig. 2: Field research results obtained during the 19th Olympic Winter Games (Salt Lake City, 2002; venue in Park City). The histograms show the average values and standard deviations of position angles from the best ten athletes in each of the five runs at the $K = 120$ m jumping hill. The number of angle measurements ranged from 18 to 50 at each position. The angle V of the skis to each other was determined from digitized images taken from the end of the run out.
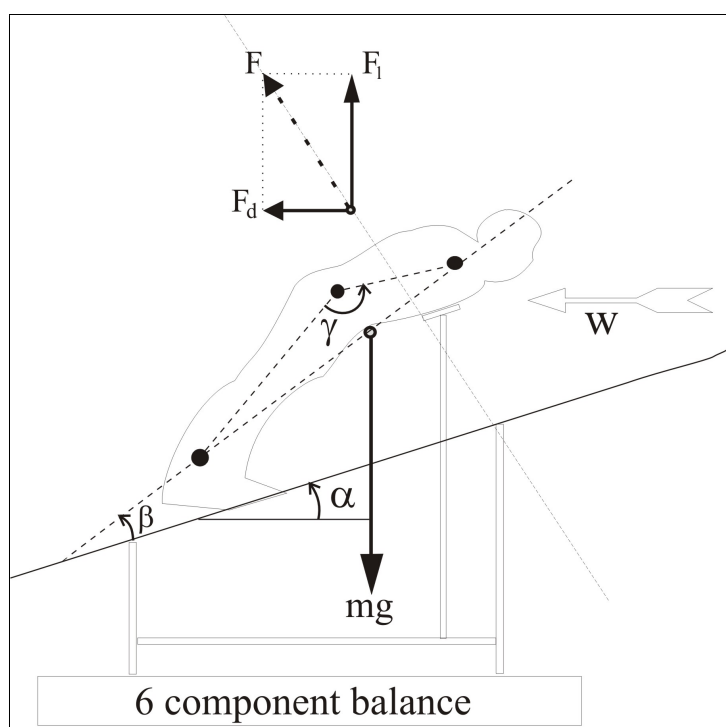
Fig. 3: Schematics of the wind tunnel measurements. The figure shows the apparatus, which enabled almost all postures of athletes and skis imaginable, and demonstrates the nomenclature used for the position angles. This study used the large wind tunnel at Arsenal Research in Vienna. The tunnel has a cross-section of 5×5 m^2. A 1.8 MW motor produces a maximum wind speed of 32 m s^{-1}.


Fig. 4: Wind tunnel measurements: The aerodynamic forces largely depend on the flight style. Large-scale wind tunnel in Vienna; Andreas Goldberger, wr 225 m (2000).
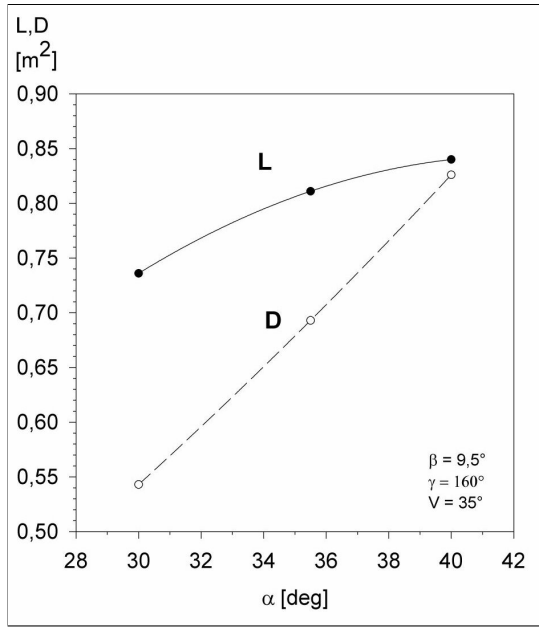
(a) Different angles of attack α. The body position was held constant ($\beta = 9.5°$ and $\gamma = 160°$) and the angles of attack were $30°$, $35.5°$ and $40°$. The opening of the skis was held constant at $V = 35°$. The interpolating functions are: $L = -0.43903 + 0.060743\alpha - 7.192 \times 10^{-4}\alpha^2$; $D = -0.032061 + 0.01232\alpha + 2.283x10^{-4}\alpha^2$.
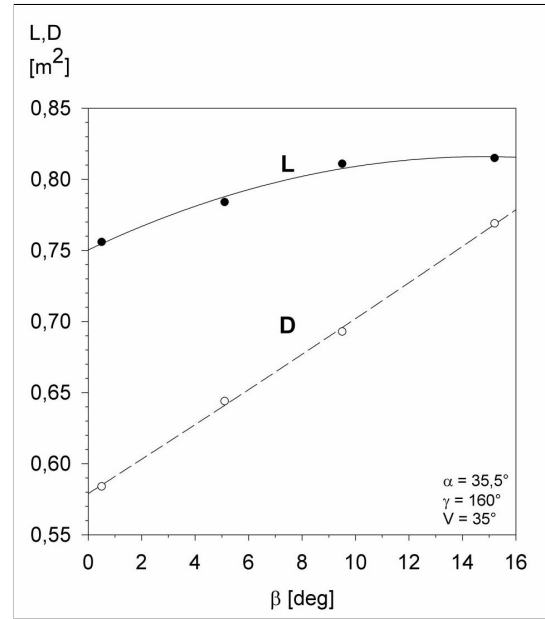
(b) L and D values depending on the body-to-ski angle β. The values shown have been taken at $\alpha = 35.5°$, $\gamma = 160°$ and $V = 35°$. The interpolating functions are: $L = -0.645718 + 0.0126185\beta - 3.348 \times 10^{-4}\beta^2$; $D = 0.408434 + 0.01364\beta + 3.9308 \times 10^{-5}\beta^2$.
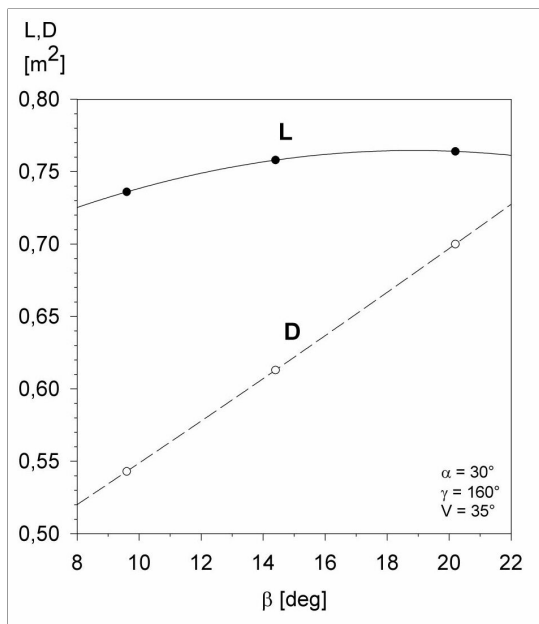
(c) L and D values depending on the body-to-ski angle β. The values shown here have been taken at $\alpha = 30°$, $\gamma = 160°$ and $V = 35°$. The interpolating functions are: $L = 0.75037 + 8.86746x10^{-3}\beta - 2.99665 \times 10^{-4}\beta^2$; $D = 0.578995 + 0.01201\beta + 2.91724 \times 10^{-5}\beta^2$.
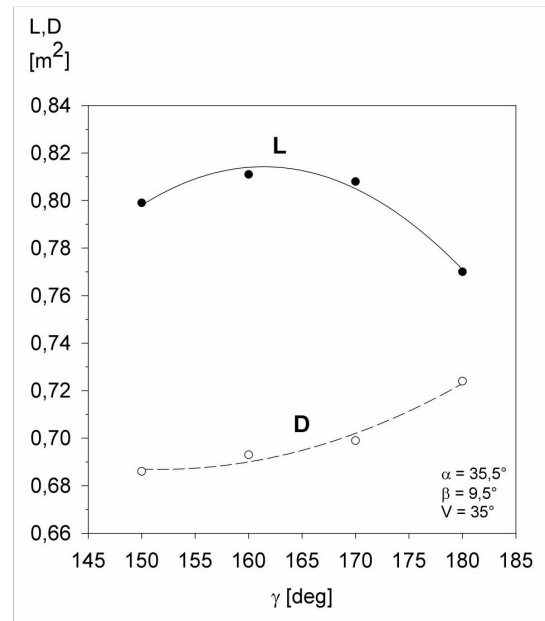
(d) L and D values depending on the hip angle γ. The body-to-ski angle β and the angle of attack α were held constant ($\beta = 9.5°$ and $\alpha = 35.5°$). The opening of the skis was held constant at $V = 35°$. The interpolating functions are: $L = -2.442 + 0.04035\gamma - 1.25 \times 10^{-4}\gamma^2$; $D = 1.722 - 0.01365\gamma + 4.5 \times 10^{-5}\gamma^2$.

Fig. 5: L and D values for model A (height $h = 1.78$ m)
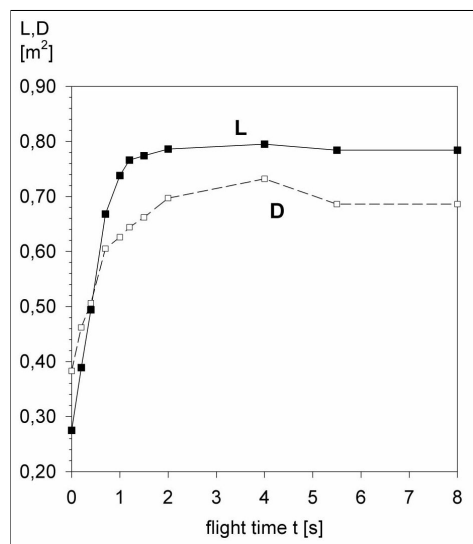
(a) Values are functions of time reflecting the athlete's position changes during the flight
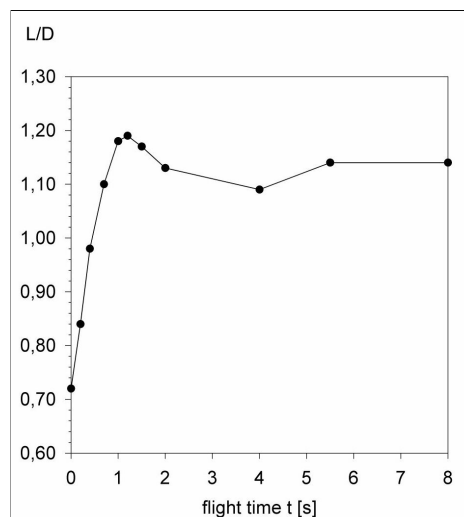
(b) L/D ratio for the reference jump with model A

Fig. 6: L and D values of reference jump of model A

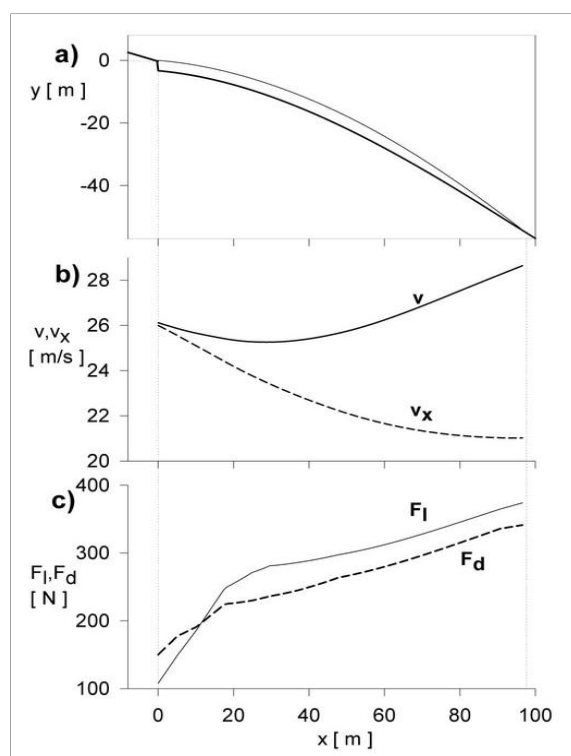## Examples of computer simulations of ski jumping

Fig. 7: Results with the reference jump for model A. **(a)** shows the profile of the jumping hill in Sapporo and the trajectory $y = y(x)$. Jumping hill parameters for Sapporo ($K = 120$ m): $a = 11°$, $b = -37°$, $c = 35°$, $H(K) = 59.449$ m, $N(K) = 103.391$ m, $r_1 = 105$ m, $R_2 = 120$ m, $M = 20$ m, $T = 7$ m, $S = 3.3$ m. **(b)** is the velocity of motion v (solid line) and the horizontal component of this velocity v_x (broken line). **(c)** shows the lift force F_l and drag force F_d acting on the athlete and his equipment. The air density was set to 1.15 kg m^{-3}; the mass of the athlete with equipment was 65 kg.
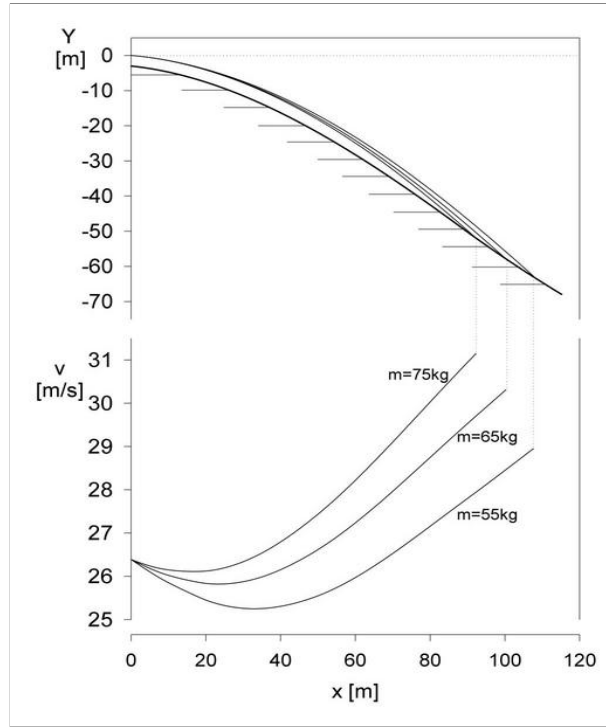
(a) Jumping hill parameters for Park City ($K = 120$ m): $\alpha = -10.5°$, $\beta = 35°$, $\beta_P = 38°$, $\beta_L = 37.77°$, $\gamma = 35°$, $h = 59.52$ m, $n = 103.51$ m, $r_1 = 93$ m, $r_2 = 105$ m, $r_L = 356.5$ m, $l_1 = 18.67$ m, $l_2 = 13.90$ m, $t = 6.7$ m, $s = 3$ m. For all jumping hills approved by the FIS the parameters can be found in the FIS Certificates of Jumping Hills. The trajectories and velocities for three different masses (55, 65 and 75 kg) are shown. The approach velocity v_0 was 26.27 m s^{-1}, the air density $\rho = 1.0$ kg m^{-3} and v_{p0} was 2.5 m s^{-1}. The gust velocity v_g was set to 0.
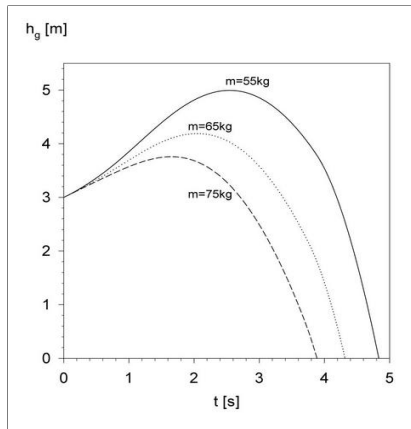
(b) Height above ground h_g for three different masses as a function of the flight time t. The solid line shows a jump simulation using a mass of $m = 55$ kg, the dotted line for $m = 65$ kg, and the broken line for $m = 75$ kg.
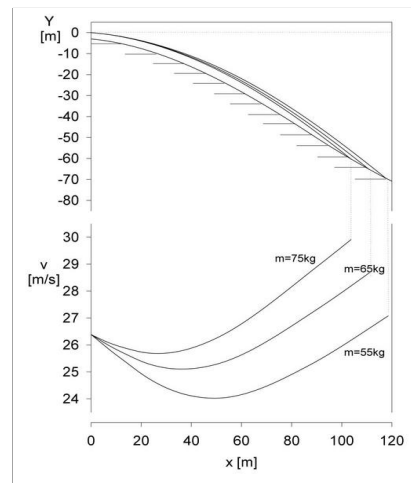
(c) Analogous to Fig. (a), however, in this case a gust blowing constantly with $v_g = 3$ m s^{-1} ($\zeta = 135°$) during the whole flight was used.

Fig. 8: Simulated jumps using the L and D tables from the reference jump for model A and the profile of the jumping hill in Park City

Organizing Committee

Egil Lillestøl (CERN and University of Bergen), CERN Schools Director
Emmerich Kneringer (University of Innsbruck), School Director
Laurenz Widhalm (HEPHY, Vienna), Co-Chair
Manfred Jeitler (HEPHY, Vienna and CERN), Co-ordinator
Tatyana Donskova (JINR), Organizing Secretary
Nick Ellis (CERN)
Robert Fleischer (CERN)
Danielle Métral (CERN), School Administrator
Alexander Olchevsky (JINR)
Andreas Salzburger (University of Innsbruck)

Lecturers

Wilfried Buchmüller (DESY)
Gerhard Ecker (University of Vienna)
John Ellis (CERN)
Robert Fleischer (CERN)
Manfred Jeitler (HEPHY, Vienna and CERN)
Rocky Kolb (Fermilab)
Manfred Lindner (Technical University, Munich)
Larry McLerran (BNL)
Laurenz Widhalm (HEPHY, Vienna)
Wolfgang Mueller (University of Graz)
Herbert Pietschmann (University of Vienna)

Discussion Leaders

Helmut Eberl (HEPHY, Vienna)
Alexei Gladyshev (JINR)
Stephan Huber (CERN)
Alejandro Ibarra (IFT, University of Madrid)
Harald Markum (Technical University, Vienna)
Sergey Troitsky (INR, Moscow)

Other Attendees

Alexei Sissakian, Vice-Director, JINR
Jos Engelen, Representing CERN Director-General, CERN

Students

Cyril Adamuscin
Cristina Adorisio
Paolo Adragna
Philip Allfrey
Thomas Atkinson
Giuseppe Avolio
Kanstantsin Babich
Oxana Barsukova
Arkadiy Bel'kov
Mark Bell
Ivan Belotelov
Michele Bianco
Marc-Oliver Boenig
Richard Brauer
Nicholas Brett
Philippe Calfayan
Nuno Filipe Castro
Teh Lee Cheng
Yaw Ming Chia
Claudio Chiri
Christian Holm Christensen
Thijs Cornelissen
Jonathan Cox
Sergei Demidov
Daniel Dobos
Uladzimir Druhakou
Michael Duehrssen
Viacheslav Duk
Christina Edgar
Saöa Fratina
Esteban Fullana Torregrosa
Vincent Giangiobbe
Carlos Gonzalez
Elena Guardincerri
Hayk Hakobyan
Per Hansson
Sam Harper
Zdenek Hubacek
Vincenzo Izzo
Christian Jacoby
Christopher Jung
Vojtech Juranek
Patrick Jussel
Nicolas Kerschen
Graham Kilvington
Thomas Kittelmann
Esben Bryndt Klinkby
Jens Konrath

Truls Martin Larsen
Alfio Lazzaro
Sabina Lehocka
Debora Leone
Phillip Litchfield
Christoph Luedeling
Jaroslaw Lukasik
Rasmus Mackeprang
Tuula Maki
Evelina Marinova
Marine Michaut
Thomas Millet
Jovan Mitrevski
Ciprian Mitu
Jukka Nysten
Fadmar Osmic
Siarhei Padolski
Sung Jin Park
Shabnaz Pashapour
Sergey Perov
Peicho Petkov
Anton Poluektov
Paul Prideaux
Sebastien Procureur
Tobias Raufer
Stefan Rieke
Belen Salvachua
Pavel Sazhin
Gerolf Schlager
Sezen Sekmen
Serhiy Senyukov
Anna Sfyrla
Vincent Siccardi
Alexey Stadnik
Bernd Stelzer
Oliver Stelzer-Chilton
Tadeusz Szymocha
Fabien Tarrade
Giovanni Francesco Tassielli
Barbara Toczek
Niels van Eldik
Filipe Veloso
Sara Vigano
Ian Vollrath
Dmytro Volyanskyy
Markus Warsinsky
Pasquale Federico Zema

Listeners

Harald Duer   Michael Rudolf Lerchster

Posters

Author	Poster title
Philip Allfrey	Measurement of charm production with the ZEUS microvertex detector
Thomas Atkinson	Track fitting with the Gaussian sum filter
Kanstantsin Babich	$B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ perspectives for the CMS experiment at LHC
Mark Bell	Semi-leptonic decays of charm to electrons at ZEUS
Ivan Belotelov	Physics with di-muons at CMS. High-p_t muons simulation and reconstruction
Michele Bianco	ATLAS RPC tests and results at INFN in Lecce
Marc-Oliver Boenig	The H1 fast track trigger
Richard Brauer	The end caps of the CMS silicon microstrip tracker
Nicholas Brett	Micro black holes — the end of short distance physics?
Philippe Calfayan	Search for second-generation leptoquarks in the $q\bar{q} \rightarrow LQ\overline{LQ} \rightarrow (\mu\,\text{jet}, \nu\,\text{jet})$ channel
Nuno Filipe Castro	Study of the ATLAS sensitivity to FCNC decays in single top events (with F. Veloso)
Teh Lee Cheng	Can we see flavour changing $t \rightarrow c$ or $t \rightarrow u$ via neutral-current-like interactions at LHC?
Yaw Ming Chia	BaBar L1 drift chamber trigger
Claudio Chiri	The MEG experiment: a possible strategy for pattern recognition in the spectrometer (with G. F. Tassielli)
Christian Holm Christensen	QBD Quantum beer dynamics (non physics) (with C. Holm, T. Kittelmann, E. Klinkby, T. M. Larsen and R. Mackeprang)
Thijs Cornelissen	Track reconstruction in the ATLAS combined test beam
Jonathan Cox	ATLAS FSI retroreflectors
Sergei Demidov	Electroweak baryogenesis in split SUSY
Daniel Dobos	ATLAS pixel module test and characterization
Michael Duehrssen	A study of $gg \rightarrow H \rightarrow WW$ in NLO and fast shower — a new parametrization for the ATLAS fast simulation
Viacheslav Duk	ISTRA+ experiment
Christina Edgar	Electromagnetic calorimeter background studies at BaBar
Saöa Fratina	Belle silicon vertex detector intrinsic resolution
Esteban Fullana Torregrosa	Digital signal reconstruction in the ATLAS hadronic tile calorimeter (with B. Salvachua)
Vincent Giangiobbe	Analysis of the 2004 combined test beam in the ATLAS experiment
Hayk Hakobyan	(Non physics) 1) A glimpse on a day of graduate students 2) Expanding science in the flat world (with Z. Hubacek)
Per Hansson	Measuring the top quark charge with the D0 experiment
Christopher Jung	Grid computing for LHC physics
Vojtech Juranek	A reconstruction of the $H \rightarrow b\bar{b}$ channel
Graham Kilvington	Backgrounds to SUSY at ATLAS

Authors	Poster title
Jens Konrath	Top quark spin correlation at the Tevatron
Sabina Lehocka	Ring-like substructures in heavy-ion interactions
Debora Leone	Hadronic cross-section at KLOE
Phillip Litchfield	NuMI-MINOS (with T. Raufer)
Christoph Luedeling	Physics in higher dimensions
Tuula Maki	Top quark mass measurement in dileptonic channel at CDF
Thomas Millet	Search for SUSY at the Tevatron: gluino decaying in b-quark and missing E_T
Ciprian Mitu	A possibility to detect upsilons in forward spectrometer at ALICE
Siarhei Padolski	The E391a experiment for rare $K_L^0 \to \pi^0 \nu \bar{\nu}$ decay searching (with S. Perov)
Shabnaz Pashapour	An investigation of top quark pair production
Peicho Petkov	Resistive plate chambers for CMS barrel — production and tests
Anton Poluektov	Measurement of CP-violating parameter $\phi_3(\gamma)$ at the Belle detector
Paul Prideaux	Multi-jet physics at HERA
Sebastien Procureur	Recent measurement of $\Delta G/G$ at COMPASS
Stefan Rieke	ATLAS LVL1 trigger, jet/energy processor and combined test beam
Sezen Sekmen	Susy d'Arc at LC
Anna Sfyrla	The ATLAS SCT assembly
Bernd Stelzer	Search for electroweak single-top-quark production at the Tevatron with CDFII
Oliver Stelzer-Chilton	Measuring the W boson mass with CDF in RunII (with I. Vollrath)
Fabien Tarrade	ATLAS: signal reconstruction and calibration for test beam, tau identification and search for standard Higgs boson using vector boson fusion
Dmytro Volyanskyy	The $B_s \to J/\psi \, \eta'$ decay at LHCb
Markus Warsinsky	Study of the process $pp \to (h/A/H \to \mu\mu) + b + X$ at ATLAS

Photos

More photos

More photos