Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

put forward by
Dipl.-Phys. Andreas von Manteuffel
born in Frankfurt am Main

Oral examination: 17 July 2008

Two Aspects of High Energy Physics:

Methods for Extended Higgs Models and Constraints on the Colour Dipole Picture

Referees: Prof. Dr. Otto Nachtmann
Prof. Dr. Hans-Christian Schultz-Coulon

Abstract

Theories with extended Higgs sectors such as Two-Higgs-Doublet Models (THDMs) or the Next-to-Minimal Supersymmetric Standard Model (NMSSM) allow for rich CP phenomena and involved Higgs-potential structures. Employing a gauge invariant formulation for the tree-level Higgs potential of the general THDM, we derive compact criteria for its stability, electroweak symmetry breaking, and generalised CP properties in a clear geometrical language. A new type of CP symmetry is shown to impose strong restrictions on the Lagrangian and to require at least two fermion generations for non-trivial Yukawa terms. Large regions of the NMSSM parameter space are excluded due to an instable vacuum. We present a rigorous determination of the global minimum of the tree-level potential via Gröbner bases.

In a second part, we investigate the colour dipole picture. This model of high energy photon-proton scattering permits a very successful description of available HERA data. Nevertheless, its range of applicability is limited. We derive general bounds on ratios of deep-inelastic proton structure functions within the colour dipole picture, following exclusively from its framework and photon wave function properties. Confronting these bounds with HERA data we can further restrict the range of applicability of the colour dipole picture. Finally, we calculate Ioffe times for a specific model and find them to be too small to justify the dipole picture at large photon virtualities.

Zusammenfassung

In Theorien mit erweiterten Higgs-Sektoren wie Modellen mit zwei Higgs-Doubletts (THDMs) oder dem Next-to-Minimal Supersymmetric Standard Model (NMSSM) können vielfältige CP-Phänomene und nichttriviale Strukturen des HiggsPotentials auftreten. Für das allgemeine THDM erlauben eichinvariante Freiheitsgrade eine geometrische Beschreibung des Higgs-Potentials auf Bornniveau, in der wir kompakte Kriterien für Stabilität, elektroschwache Symmetriebrechung und CP-Eigenschaften formulieren. Eine neue Art von CP-Symmetrie impliziert starke Einschränkungen an die Lagrangedichte und erfordert insbesondere mindestens zwei Fermion-Familien für nichtverschwindende Yukawa-Wechselwirkungen. Große Parameterbereiche des NMSSM können aufgrund eines instabilen Vakuums ausgeschlossen werden. Wir präsentieren eine Methode zur eindeutigen Bestimmung des globalen Minimums des Potentials auf Bornniveau mittels Gröbnerbasen.

In einem zweiten Teil dieser Arbeit untersuchen wir das Farbdipolbild (colour dipole picture). Dieses Modell der hochenergetischen Photon-Proton-Streuung erlaubt eine sehr erfolgreiche Beschreibung verfügbarer HERA-Daten. Allerdings ist sein Anwendungsbereich beschränkt. Wir leiten allgemeine Schranken an Verhältnisse tiefinelastischer Strukturfunktionen her, die sich ausschließlich aus dem konzeptuellen Rahmen und den Photon-Wellenfunktionen ergeben. Durch einen Vergleich mit HERA-Daten finden wir neue Einschränkungen für den Geltungsbereich des Dipolbildes. Schließlich berechnen wir Ioffe-Zeiten innerhalb eines speziellen Modells und zeigen, dass dort eine Standardvoraussetzung des Dipolbildes bei großen Photon-Virtualitäten verletzt ist.

Contents

Part I

Introduction

For more than three decades up to date, the Standard Model [1, 2, 3, 4, 5] of particle physics provides a theoretically consistent [6] framework to describe the measurements in high-energy physics with great success [7, 8]. Only one ingredient is still missing: the scalar Higgs boson. This particle is a crucial prediction of the Standard Model mechanism [9, 10, 11, 12] for generating particle masses and hiding the electroweak symmetry. Direct Higgs searches at LEP [13] provide a lower bound on its mass of 114.4 GeV.

Enlarging the scope of the Standard Model (SM) from collider physics to other fields, the SM can not be considered as a complete description. The observed neutrino masses [14, 15] are not contained in the SM, but they are too small to influence collider physics. Rotational speed of galaxies [16, 17], the "Bullet cluster" collision [18, 19] and large scale structure formation [20, 21, 22] suggest the existence of Cold Dark Matter for which the SM provides no candidate. In the context of collider physics, it is worth mentioning the roughly $3\,\sigma$ deviations from the SM found for the precision measurements of the muon anomalous magnetic moment [23, 24] and of B_s decays [25], which may indicate signs of beyond the SM physics.

From an aesthetical point of view, the Standard Model can not be regarded as a fundamental theory, not only because it contains a large number of free parameters. Although crucial for the stability of atoms, the quantisation of electric charge is purely coincidental within the SM. Furthermore, the Standard Model does not explain why there are three fermion generations with the observed mass and mixing patterns, except for the hint that CP violation in the Yukawa sector requires at least three families. Finally, the Standard Model provides no link to gravity. These facts suggest new physics at least at very high scales. Considering the Standard Model to be valid up to such high energies seems unnatural, since no mechanism stabilises the hierarchy between the electroweak and the high scale.

Theories able to solve at least a part of the above SM problems typically require extended Higgs sectors. In this thesis, we consider aspects of the general Two-Higgs-Doublet model (THDM) [26, 27, 28, 29] and of the Next-to-Minimal Supersymmetric Standard Model (NMSSM), see [30, 31, 32] and references therein.

The strong interactions of the Standard Model give rise to a rich phenomenology, which is theoretically challenging due to large quantum corrections and intrinsic non-perturbative features. An interesting aspect to study is the behaviour of the gluon densities at small Bjorken-x such as possible saturation effects due to non-linear interactions. HERA provides a wealth of data for deep inelastic scattering (DIS) of electrons and positrons on protons, where the region of small Bjorken-x lies in the non-perturbative region. While phenomenological models based on the colour dipole picture successfully describe the structure function measurements also in the non-perturbative regime, they can not yet be derived from the Standard Model. Further, in order to arrive at the framework used for the dipole model, a number of assumptions and approximations have to be made. In order to derive firm conclusions from colour dipole models, their range of applicability needs to be understood as precisely as possible. In this thesis we examine constraints on the colour dipole picture.

This thesis is organised as follows. In part II methods and results for different extensions of the SM Higgs sector are presented. In chapter 1 we give motivations for specific types

of extended Higgs sectors and discuss previous related work. The ends of sections 1.3 and 1.4.3 provide a detailed overview of the contents of this part of the thesis. Chapter 2 contains our discussion of THDMs. In section 2.1 we introduce the gauge invariant orbit variables as the basis for the following tree-level analysis of the general THDM. We derive the criteria for its stability and electroweak symmetry breaking (EWSB) in section 2.2 and section 2.3, respectively, followed by a discussion of masses and input parameters in our notation in section 2.4. In section 2.5 we discuss generalised CP symmetries and derive criteria for their realisation. Within the employed gauge invariant formulation, a new type of CP symmetry emerges naturally. These general methods and results are illustrated in section 2.6 for more specific models. In section 2.7 we discuss in detail the implications of the new CP symmetry for the Higgs and Yukawa sectors. An extension of these methods to the case of n doublets is discussed in section 2.8. In chapter 3 we present a method to determine the global minimum of the NMSSM. The Higgs potential and necessary EWSB conditions are given in section 3.1 and section 3.2. We introduce orbit variables for the NMSSM in section 3.3 and discuss an approach to determine the stationarity conditions via Gröbner bases in section 3.4. Section 3.5 contains our numerical results.

In part III we discuss constraints on the colour dipole picture. Chapter 4 contains an introduction and a discussion of previous work on the topic, followed by a detailed overview of the contents of this part of the thesis at the end of section 4.2. A short review of its foundations is given in chapter 5 along with a detailed discussion of the photon wave functions. In chapter 6 we derive general bounds on ratios of DIS structure functions and employ them to constrain the kinematical range of applicability of the colour dipole picture. Whether different choices for the energy variable in the dipole cross section can be considered equivalent is discussed in section 7. Finally, in chapter 8, we calculate Ioffe times for a specific colour dipole model and check whether they justify the application of the dipole picture.

In part IV we summarise our conclusions. Finally, appendix A contains a short introduction to Gröbner bases and appendix B provides the mathematical details employed in the derivations of bounds in chapter 6.

Part II

Methods for Extended Higgs Models

Chapter 1

Motivations for extended Higgs models

1.1 Open questions in the Standard Model

The detection of the SM Higgs particle is an outstandig problem for the experimental verification of the Standard Model. Direct searches at LEP give the lower bound [13] on its mass

$$m_H > 114.4\,\text{GeV} \tag{1.1}$$

when combining the measurements of all four LEP experiments. The indirect determination [8] of the SM Higgs mass from electroweak precision observables alone give $m_H = 85\,(+39/\text{-}28)\,\text{GeV}$ and $m_H < 166\,\text{GeV}$ at 95 % confidence level when Z-pole, m_t, m_W and Γ_W measurements are included. The rather large errors stem from the fact that the Higgs mass enters only logarithmically in the one-loop expressions. Before the discovery of the top quark, a similar method lead to the prediction of its mass m_t. However, the precision observables are much more sensitive on m_t, see also table 10.2 of [8], since m_t enters quadratically. The LHC may presumably discover a Higgs boson up to a mass of about 1 TeV, see e.g. [33, 34]. For these and many other aspects of Standard Model Higgs physics see also [35] of the review series [35, 36], which provides a comprehensive coverage of the state-of-the-art in this field.

Despite the fact, that not even a single scalar elementary particle has been observed in Nature up to date, it is worthwhile to consider extensions of the Standard Model Higgs sector for the following reasons.

In the Standard Model the Higgs sector is very sensitive to radiative corrections from new physics at high scales. The mass of a very heavy particle coupling to the Higgs boson, such as of an additional vector boson present in Grand Unified Theories (GUTs), enters quadratically in the radiative corrections to the Higgs boson mass. Explaining the electroweak symmetry breaking (EWSB) observed in Nature requires then a fine-tuning in the renormalisation procedure, which cancels these large terms between the bare Higgs mass and the quantum corrections to give a net renormalised Higgs mass below one TeV.

This is usually considered to be an unnatural fine-tuning, and therefore referred to as the *fine-tuning* or *naturalness problem* [37, 38, 39].

Supersymmetry [40, 41] provides a mechanism to stabilise the electroweak scale at the observed value by introducing a symmetry between bosons and fermions. Furthermore, supersymmetry provides a dark matter candidate in case of R-parity conservation. Finally, supersymmetry improves the unification of the gauge couplings [42, 43, 44, 45] and provides a link to gravity. From a mathematical point of view, supersymmetry is special in being the maximal non-trivial extension of the Poincaré symmetry of the S-matrix, which can be generated by graded Lie-algebras [46, 47]. If realised in Nature, supersymmetry must be broken. In order to give masses to both the up- and down-type quarks and to keep the theory free of chiral anomalies, supersymmetric extensions of the SM require an extended Higgs sector. The Minimal Supersymmetric Standard Model (MSSM) [48, 49, 50, 51] contains two Higgs doublets, for general reviews see e.g. [52, 53]. For the Higgs sector of the MSSM see in particular the comprehensive review [36]. In the Next-to-Minimal Supersymmetric Standard Model (NMSSM) [30, 54, 31, 32] a further Higgs singlet is added to solve the μ-problem of the MSSM as we will discuss in section 1.4.

A non-minimal Higgs sector is also favoured for cosmological reasons. CP violation is one of the three Sakharov criteria, which have to be fulfilled to explain the observed baryon-antibaryon asymmetry in our Universe through the particle dynamics [55, 56]. In the Standard Model, CP violation occurs at tree-level only in the Yukawa sector through the Kobayashi-Maskawa mechanism [57] and is therefore highly constrained. Taking furthermore the lower bound (1.1) on the SM Higgs mass into account, the electroweak phase transition is not able to provide the thermal instability needed to explain the observed baryon excess, as described in the review [58]. Extended Higgs sectors are less constrained and allow to circumvent this problem for instance by providing new sources of CP violation. In contrast to the SM, extended Higgs sectors can lead to explicit and spontaneous CP violation [59] in the Higgs sector.

Last but not least, the observed fermion masses and mixings show pronounced patterns and hierarchies. Involving several Higgs bosons in their generation might be a canonical approach. After all, there are three known generations of fermions and there is no particular reason for only one Higgs generation.

1.2 Higgs representations and the ρ parameter

In this section we discuss possible Higgs representations for a $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge theory, where we review the derivations of [60] summarised in [61]. We shall omit the gauge group $SU(3)_C$ of strong interactions since it is irrelevant for the following discussion. Let χ denote scalar fields carrying some representation of $SU(2)_L \times U(1)_Y$. In general such a representation is reducible and contains complex unitary and real orthogonal parts. We assume without loss of generality that χ carries a real orthogonal representation, since other cases may always be mapped to this case, as shown in the above citation for the general case and in appendix B of [27] for two doublets. In order to break the electroweak symmetry

spontaneously, the potential is assumed to give non-vanishing vacuum expectation values (vevs) to the components of χ,

$$\boldsymbol{v} \equiv \langle 0|\chi|0\rangle \neq 0\,, \tag{1.2}$$

such that a $U(1)$ factor of $SU(2)_L \times U(1)_Y$ remains unbroken and can be identified with the electromagnetic gauge group. Here, $\boldsymbol{v}$ is a multi-component vector, as is χ, and gives rise to particle masses. However, not all a-priori possible representations lead generically to an experimentally acceptable phenomenology.

A strong restriction stems from the non-observation of large flavour-changing neutral currents (FCNCs). While FCNCs are automatically suppressed in the SM, this is not necessarily the case for more general Higgs sectors and corresponding Yukawa couplings. One way to ensure absence of large FCNCs is to require all fermions of a given charge to receive their masses from the vev of the same Higgs field and avoid large mixings between these fields for the physical Higgs bosons [62].

The precisely measured ρ parameter gives severe constraints for the Higgs sector itself. The ρ parameter relates the masses of the W and Z boson, m_W and m_Z, to the weak mixing angle θ_w:

$$\rho \equiv \left(\frac{m_W}{\cos\theta_\mathrm{w}\, m_Z}\right)^2. \tag{1.3}$$

Precision measurements show that ρ is very close to 1 [63]. This suggests to require theoretically $\rho = 1$ at tree-level. While this is indeed the case for the SM, one finds the following for more general Higgs representations.

Extending χ to carry a unitary representation of the same dimension and decomposing it into eigenstates (t, y) of weak-isospin and hypercharge with quantum number t and y, respectively, we have

$$t = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \ldots \tag{1.4}$$

and furthermore assume y to be rational numbers, see [60]. With the normalisation chosen such that the operators of charge, hypercharge and third weak-isospin component are related by

$$Q = T_3 + Y\,, \tag{1.5}$$

the squared gauge-boson masses are given by

$$m_W^2 = \frac{1}{2}\left(\frac{e}{\sin\theta_\mathrm{w}}\right)^2 \sum_{t,y} \left[t(t+1) - y^2\right] \boldsymbol{v}^\mathrm{T}\mathbb{P}(t,y)\boldsymbol{v}, \tag{1.6}$$

$$m_Z^2 = \left(\frac{e}{\sin\theta_\mathrm{w}\cos\theta_\mathrm{w}}\right)^2 \sum_{t,y} y^2\ \boldsymbol{v}^\mathrm{T}\mathbb{P}(t,y)\boldsymbol{v}, \tag{1.7}$$

where $\mathbb{P}(t, y)$ is the projector on the subspace with representation (t, y). Here, the positron charge e, and the sine and cosine of the weak mixing angle are defined in terms of the gauge

couplings g and g' as in the SM, see for instance [64]:

$$\sin\theta_{\mathrm{w}} = \frac{g'}{\sqrt{g^2+g'^2}}, \qquad \cos\theta_{\mathrm{w}} = \frac{g}{\sqrt{g^2+g'^2}}, \qquad e = g\sin\theta_{\mathrm{w}}. \tag{1.8}$$

It turns out that

$$\boldsymbol{v}^{\mathrm{T}}\mathbb{P}(t,y)\boldsymbol{v} \neq 0 \tag{1.9}$$

is only possible if

$$y \in \{-t, -t+1, \ldots, t\}. \tag{1.10}$$

Inserting the expressions for m_W and m_Z in the definition (1.3) leads to

$$\rho = \frac{\sum_{t,y}[t(t+1)-y^2]\ \boldsymbol{v}^{\mathrm{T}}\mathbb{P}(t,y)\boldsymbol{v}}{\sum_{t,y} 2y^2\ \boldsymbol{v}^{\mathrm{T}}\mathbb{P}(t,y)\boldsymbol{v}}. \tag{1.11}$$

The value $\rho = 1$ may now either be obtained by fine-tuning the parameters of the potential in order to get the right vacuum expectation values, which seems rather unnatural and is therefore not considered here. Or one only allows those representations in (1.6) and (1.7) that *separately* lead to $\rho = 1$. There are infinitely many such representations [65], starting with the doublet with $t = 1/2$ and $y = \pm 1/2$, and the septuplet with $t = 3$ and $y = \pm 2$. From each of these representations one or more copies are allowed and one still gets $\rho = 1$. Furthermore, the singlet with $y = 0$ and all representations with

$$y \notin \{-t, -t+1, \ldots, t\} \tag{1.12}$$

are allowed since they do not contribute to the sums in (1.11).

1.3 The case of two generic doublets

A particular simple extension of the Standard Model Higgs sector while keeping $\rho = 1$ at tree-level is, according to the previous section, to include n instead of only one complex Higgs doublets. In these n-Higgs-Doublet Models models, electroweak symmetry breaking works in principle as in the Standard Model. The Lagrangian contains scalar products of the Higgs doublets without derivatives, which give rise to the tree-level Higgs potential. The potential is required to be bounded from below and possesses a minimum, which leads to non-vanishing vacuum expectation values of the Higgs fields. These vevs spontaneously break the electroweak symmetry down to the electromagnetic symmetry. They give rise to masses of W and Z bosons through the covariant derivatives of the Higgs doublets, and to masses of quarks and leptons through Yukawa interactions. As in the Standard Model three real degrees of freedom from the Higgs doublets can be identified as the Pseudo-Goldstone Bosons of EWSB, which reappear as longitudinal modes of the massive gauge bosons. Any further real degree of freedom from the Higgs doublets corresponds to a physical Higgs boson. Therefore, each additional complex Higgs doublet adds four Higgs bosons to the

physical spectrum. For more than one doublet the tree-level Higgs potential may give rise to new phenomena, such as explicit or spontaneous CP violation as well as multiple minima.

In the case of two Higgs doublets, that is, in Two-Higgs-Doublet Models (THDMs), the physical spectrum contains 5 physical Higgs bosons. Two of them are charged, $H^{\pm}$, and three of them are neutral. In case CP is conserved one can assign CP quantum numbers to the neutral Higgs bosons, such that one has one CP odd Higgs A and two CP even Higgses h and H. By conventions, h denotes the lighter of the two states h and H, which both will differ from the SM Higgs boson in the general case. For the general THDM the potential can be quite involved. We shall see that it may contain 14 parameters in contrast to the 2 parameters in the case of one Higgs doublet. The MSSM Higgs sector contains also two Higgs doublets. Its potential is a very special case of the general THDM potential, since it is highly constrained by supersymmetry. Introductory remarks specific to the MSSM will be given in the next section.

Present experiments give the following exclusion regions for various versions of the THDM. The OPAL collaboration has performed a parameter scan for the CP conserving THDM [66] and excluded at 95% C.L. large parts of the region where

$$\begin{aligned} 1\,\mathrm{GeV} < m_h < 130\,\mathrm{GeV}, \\ 3\,\mathrm{GeV} < m_A < 2\,\mathrm{TeV}, \\ 0.4 < \tan\beta < 40, \\ \alpha = -\frac{\pi}{2}, -\frac{\pi}{4}, 0, \frac{\pi}{4}, \frac{\pi}{2}. \end{aligned} \tag{1.13}$$

Here, $\tan\beta$ is the ratio of the two Higgs vevs and α is a further mixing angle, needed to diagonalise the mass matrix for the two states h and H. Furthermore, the approximate region where

$$\begin{aligned} 1\,\mathrm{GeV} < m_h < 55\,\mathrm{GeV}, \\ 3\,\mathrm{GeV} < m_A < 63\,\mathrm{GeV} \end{aligned} \tag{1.14}$$

is excluded for all $\tan\beta$ values for negative α. In a combined analysis [67] of the four LEP collaborations a lower bound on the mass of the charged Higgs in models with two Higgs doublets like the THDM or the MSSM, approximately

$$m_{H^{\pm}} > 78.6\,\mathrm{GeV} \tag{1.15}$$

is determined.

In early work [26, 27, 28, 68] on the THDM, the number of parameters were restricted by continuous or discrete symmetries or other restrictions were employed. See [29] for a review and further references. A recent proposal [69] considers a specific non-supersymmetric THDM, which postpones the naturalness problem up to a scale of several TeV. Criteria for stability and electroweak symmetry breaking in more general THDMs were derived in [70, 71]. Recently, the completely general THDM became the focus of many studies,

and different basis independent results were derived. Basis independent conditions for spontaneous CP violation are found in [72, 73]. An extensive analysis of invariants with respect to basis changes is employed in [74, 75], from which necessary and sufficient basis independent criteria for explicit and spontaneous CP violation are derived. See also [76] for another derivation of criteria for spontaneous CP violation and [77] for a general discussion of CP transformations in gauge theories. The meaning of $\tan\beta$ in a basis independent context is discussed in [78]. The coexistence of multiple minima is considered in [79, 80, 81, 82]. Criteria for stability and electroweak symmetry breaking were derived in [61] within a geometric approach in terms of gauge invariant functions, which we shall also employ in this thesis. Geometric methods are also used in [83] to study criteria for CP violation in n-Higgs-Doublet Models. A similar formulation is used in [84, 85, 86] to derive various structural statements, in particular for the coexistence of minima. There, Lorentz-boost type transformations are used for the gauge invariant functions to assume a standard basis, which corresponds to a diagonal matrix $\tilde{E}$ in our notation. Note that this modifies the kinetic terms of the Higgs fields and is not possible for all physically acceptable potentials. For a recent review of the THDM, in particular of its CP violation properties, see [87].

In chapter 2 of this thesis, we shall discuss the completely general THDM using the geometric approach of gauge invariant functions [61]. This allows us to derive compact criteria for stability, electroweak symmetry breaking, and generalised CP properties of the THDM in a clear geometrical language, see also our publications [88, 89, 90]. The scope of this analysis is the tree-level. In a more detailed study, quantum corrections should be taken into account. Our main results are formulated in a basis independent way, and may thus be applied directly to any specific model irrespective of the chosen basis.

In section 2.1 we introduce the gauge invariant functions of the Higgs field degrees of freedom. These gauge invariant functions combine to a Minkowski-type four-vector, which lies on or inside the forward light cone. Since they parameterise the gauge orbits of the Higgs fields, we shall also refer to them as orbit variables. Our discussion of stability and electroweak symmetry breaking in section 2.2, section 2.3 and physical parameters in section 2.4 rederives the results of [61] and extends them. We derive concise criteria for a theory to be phenomenologically acceptable, which restrict the parameters of the potential. The presented method gives an explicit recipe to determine all stationary points and thus the global minimum of the potential also for more involved models. We consider generalised CP transformations in section 2.5, which are shown to correspond to reflections of the space-like part of the orbit variables. CP transformations which equal the standard CP transformation in some basis are reflections on a plane, while a new type of CP symmetry corresponds to a point reflection. Within this geometric picture, we derive necessary and sufficient criteria for CP symmetries of the potential, the vacuum and the Yukawa terms. Examples to illustrate the general methods are given in section 2.6. In particular, we present an explicit analytical discussion of a model with two coexisting minima, which demonstrates the need for a global minimum check beyond considering Higgs masses. Models having the new CP symmetry mentioned before are discussed in more detail in section 2.7. There we show, that extending all CP symmetries of the Higgs sector to the Yukawa sector imposes severe restrictions on the fermion masses and couplings.

This can be used to generate mass hierarchies and ensure absence of large FCNCs by a canonical symmetry requirement. Comparisons of our results with the literature may be found in the respective sections. Similar methods as presented here may also be applied to other extended Higgs sectors. As an outlook we discuss a possible extension to the case of n Higgs doublets in section 2.8.

1.4 Supersymmetric extensions and the μ problem

1.4.1 The minimal supersymmetric extension

The MSSM is defined as the softly broken supersymmetric extension of the Standard Model with minimal additional field content and conserved R-parity. It contains soft supersymmetry breaking terms, which do not reintroduce the naturalness problem, and parameterises them in a general way. Supersymmetry requires pairings of fermions with bosons which lead effectively to a new supersymmetric partner for each of the SM fermions and gauge bosons. In order to give masses to both up- and down-type fermions, analyticity of the superpotential and absence of anomalies require two doublets for the Higgs supermultiplets. These Higgs supermultiplets contain two scalar Higgs doublets and their fermionic superpartners, the Higgsinos.

The MSSM Higgs potential is highly constrained due to supersymmetry. Quartic terms in the Higgs fields are fixed by supersymmetry in terms of gauge couplings, while quadratic terms receive supersymmetric as well as soft breaking contributions. In fact, the tree-level Higgs sector can be parameterised by only two unknown parameters, usually chosen to be the mass of the pseudoscalar boson m_A and the ratio $\tan\beta$ of the vacuum expectation values of the two Higgs doublets. In the decoupling limit $m_A \gg m_Z$, where practically $m_A \gtrsim 200\,\mathrm{GeV}$ is sufficient [91], one neutral Higgs boson h is light and has the same couplings as the SM Higgs boson whereas the other Higgs bosons H, A and $H^\pm$ are heavy and decouple.

The SM mass bound (1.1) applies thus also to the mass m_h of the lightest CP-even MSSM Higgs boson in the decoupling limit. In fact, deviations arise for the MSSM with real parameters only for small m_A and large $\tan\beta$. There, the combined LEP result [92] yields for a scan of several MSSM benchmark scenarios the bounds

$$m_A > 93\,\mathrm{GeV}\,, \tag{1.16}$$

$$m_h > 93\,\mathrm{GeV} \tag{1.17}$$

and excludes the range $0.9 < \tan\beta < 1.5$ at 95 % confidence level. Note that the MSSM predicts $m_h < m_Z = 91\,\mathrm{GeV}$ at the tree-level, see (2.260), which is in contradiction to these measurements. However, radiative corrections increase the tree-level mass bound. But a substantially higher upper bound requires contributions from rather heavy supersymmetric particles such as the stops, the superpartners of the top quark, or the gluino, the superpartner of the gluon. On the other hand, these supersymmetric particles should not be too heavy, since this would reintroduce the naturalness problem. This is called the

little fine-tuning problem of the MSSM. State-of-the-art calculations [93, 94] include the complete 1-loop corrections and many 2-loop contributions also for complex parameters. For supersymmetric masses not above 1 TeV the radiative corrections are found in [93] to change the bound on the lightest Higgs mass to

$$m_h \leq 131\,\mathrm{GeV} \tag{1.18}$$

where $m_t = 172.7\,\mathrm{GeV}$ is used. The maximum is reached for $\tan\beta \geq 10$. We also note that complex parameters of the MSSM actually still allow [92] a very light Higgs bosons to have escaped detection at LEP.

1.4.2 The μ problem of the MSSM

The superpotential of the MSSM has the form, see [95],

$$W = \hat{u^c} h_u \hat{Q} \hat{H}_u - \hat{d}^c h_d \hat{Q} \hat{H}_d - \hat{e}^c h_e \hat{L} \hat{H}_d + W_H \tag{1.19}$$

with

$$W_H^{\mathrm{MSSM}} = \mu \hat{H}_u \hat{H}_d\,, \tag{1.20}$$

where the hatted fields are superfields. The Yukawa-type terms in (1.19) contain the coupling matrices h_u, h_d, h_e. We list them only for completeness, but they are not relevant for us in the following. The Higgs superfields $\hat{H}_u$ and $\hat{H}_d$ have hypercharge $y = -1/2$ and $y = +1/2$, respectively, and $\hat{H}_u \hat{H}_d = \hat{H}_u^+ \hat{H}_d^- - \hat{H}_u^0 \hat{H}_d^0$. The parameter μ mixes $\hat{H}_u$ and $\hat{H}_d$ and thus the contained scalar Higgses as well as their fermionic superpartners, the Higgsinos. The soft breaking terms entering the Higgs potential are

$$\mathscr{L}_{\mathrm{soft}} = -m_{H_u}^2\, H_u^\dagger H_u - m_{H_d}^2\, H_d^\dagger H_d - \left(m_3^2\, H_u H_d + h.c.\right). \tag{1.21}$$

where H_u, H_d are the scalar doublets, $H_u H_d \equiv H_u^+ H_d^- - H_u^0 H_d^0$, and $m_{H_u}^2$, $m_{H_d}^2$, m_3^2 are parameters of dimension mass squared.

The μ-term enters the mass of the charginos, which are the mass eigenstates formed from the charged Higgsinos and the charged superpartners of the W bosons, the winos. Absence of signals in the search for charginos at LEP [96] put a bound on the chargino mass of $m_{\chi^\pm} > 92\,\mathrm{GeV}$ and require $|\mu| > 0$. Moreover, for a phenomenologically acceptable electroweak symmetry breaking, $|\mu|$ should have a value of roughly the electroweak scale to avoid new fine-tuning problems. On the other hand, μ is a parameter of mass dimension in the superpotential, from which one would expect its natural value to be either of the order of the Planck scale or exactly zero. A zero μ-term could be realised by a symmetry at the high scale, and the form of the renormalisation group equations [97] would then imply a zero μ-term also at the electroweak scale. The lack of a natural origin for a phenomenologically acceptable μ-term of the order of the electroweak scale v,

$$|\mu| \approx \mathcal{O}(v) \tag{1.22}$$

is called the μ-problem of the MSSM, see also [98].

Various solutions to the μ problem have been discussed in the literature. A prominent example is the Guidice-Masiero mechanism [99], where the μ-term is generated in the context of supergravity breaking. Another solution to the μ-problem has received much attention in the literature: the Next-to-Minimal Supersymmetric Standard Model (NMSSM). In this model, the μ-term is effectively generated by coupling the two doublets to an additional singlet field, which acquires a vev.

1.4.3 The NMSSM as a solution to the μ problem

The Next-to-Minimal Supersymmetric Standard Model (NMSSM) extends the MSSM by an electroweak singlet chiral superfield $\hat{S}$ in addition to the two electroweak doublet superfields $\hat{H}_u$, $\hat{H}_d$. We impose a $\mathbb{Z}_3$ symmetry on the superpotential, which is defined by multiplication of all superfields by $\exp(i2\pi/3)$. This forbids linear and mass terms in the superpotential and leaves only dimensionless couplings as supersymmetric parameters. We follow the notation of [100] and parameterise the resulting Higgs part of the NMSSM superpotential by

$$W^H_{\text{NMSSM}} = \lambda\, \hat{S}\hat{H}_u\hat{H}_d + \frac{1}{3}\kappa\, \hat{S}^3 \tag{1.23}$$

with the dimensionless real couplings $\lambda,\ \kappa$. For the parameterisation of the soft breaking terms we use

$$\mathscr{L}_{\text{soft}} = -m^2_{H_u}\,|H_u|^2 - m^2_{H_d}\,|H_d|^2 - m^2_S\,|S|^2 - \left(\lambda A_\lambda\, S H_u H_d + \frac{1}{3}\kappa A_\kappa\, S^3 + h.c.\right). \tag{1.24}$$

where $H_u,\ H_d$ are the scalar doublets and S is the complex scalar singlet. Here, $m^2_{H_u}$, $m^2_{H_d}$, m^2_S are real parameters of dimension mass squared and A_λ, A_κ are complex parameters of dimension mass.

Note, that the singlet is required to get a non-vanishing vev $\langle S\rangle$. A vanishing κ would lead to the spontaneous breaking of a continuous Peccei-Quinn $U(1)_{PQ}$ symmetry [101] and thus to a massless Goldstone boson, the axion [102, 103]. Such a massless axion is inconsistent with astrophysical and cosmological limits, see [7] for a review, except for very small λ, which in turn prohibits a natural solution of the μ problem. A non-vanishing κ explicitly breaks this continuous $U(1)_{PQ}$ down to $\mathbb{Z}_3$ such that no massless axion emerges. However, the spontaneous breaking of $\mathbb{Z}_3$ can generate domain walls in the early universe which contradict the cosmic microwave background measurements. This is called the *domain wall problem*. However, additional operators suppressed by powers of the Planck scale might break the $\mathbb{Z}_3$ symmetry at the low scale just enough to avoid the domain wall problem but still suppressed enough to leave unchanged the low-energy phenomenology considered here. The are claims in the literature that the existence of such operators introduces a new naturalness problem, see e.g. [104], while other authors propose solutions to construct them in a way they deem natural, see e.g. [105]. In the following, we just assume that the solution to the domain wall problem leaves our considerations unchanged.

Since the NMSSM is considered as a solution to the the μ-problem of the MSSM we shall explicitly define a limit to relate both models. Comparing the superpotentials of both models, (1.20) and (1.23), as well as their soft breaking terms, (1.21) and (1.24), leads to the identification of the effective μ and m_3^2 parameters within the NMSSM:

$$\mu|_{\text{eff}} = \lambda \langle S \rangle , \qquad m_3^2|_{\text{eff}} = \lambda A_\lambda \langle S \rangle \ . \tag{1.25}$$

Thus we may define the "MSSM limit" of the NMSSM as

$$\lambda, \kappa \to 0 \qquad \text{with } \ \mu|_{\text{eff}} , m_3^2|_{\text{eff}} = const \tag{1.26}$$

such that the additional singlet decouples and the standard MSSM Higgs sector with its two doublets is obtained.

Note, that the supersymmetry breaking parameters are essential for electroweak symmetry breaking in the MSSM, see section 2.6.1. Their size is roughly determined by the electroweak scale, and they may be thought of as being generated by some common mechanism. While a phenomenologically acceptable μ parameter appears at the same scale, it is a parameter of the unbroken supersymmetry in the MSSM, and therefore of different nature compared to the soft breaking terms. In contrast, the scale of the effective μ-term in the NMSSM, $\mu|_{\text{eff}}$, is set by $\langle S \rangle$. It therefore originates from soft breaking terms, such that these become the only source of scale for the electroweak sector. While the *origin* of this scale is still left unexplained in the NMSSM, it might easier be linked to one single breaking mechanism.

In the NMSSM the tree-level upper bound on the lightest Higgs boson is relaxed, which is of interest to avoid fine-tuning problems due to large masses of supersymmetric particles. The analysis [106] includes leading radiative corrections and finds a relaxation of the MSSM bound of about $12 - 16\,\text{GeV}$, where the lightest CP-even NMSSM Higgs mass reaches its maximum for $\tan\beta \approx 2$. In contrast, the lightest CP-even MSSM Higgs reaches its upper bound for $\tan\beta > 10$. More aspects of fine-tuning problems within supersymmetric theories and the NMSSM in particular may be found in [107] and the references therein. Note that even a light NMSSM Higgs boson might be difficult to detect. In fact, it might be very difficult [108] to observe just one of altogether 5 Higgs bosons at the LHC, since Higgs-to-Higgs decays may be allowed and all Higgs bosons may have reduced couplings to the gauge bosons and the top quark due to singlet admixtures.

Further investigations of the NMSSM Higgs sector can be found in [100, 109]. Computer programs to calculate spectra and check the parameters for various constraints are presented in [110, 111]. Electroweak baryogenesis and light neutralino dark matter is studied in [112]. In [113] many different constraints from collider phenomenology to cosmology are considered for the nMSSM variant of the model, whose superpotential contains a tadpole but no triple selfcouplings for the singlet.

In chapter 3 of this thesis, we shall focus on the determination of the global minimum of the Higgs potential in order to ensure a stable vacuum, see also our publication [114]. This aspect is partially omitted in the literature, or a numerical minimisation method is applied [109]. Global minimisation is generally a non-trivial problem. While for example,

combining a numerical local descent method with a grid of starting values might often succeed in finding the global minimum, it is not guaranteed by the method. From a heuristic point of view, the NMSSM Higgs potential depends on a rather high number of fields and has an involved structure in the general case, such that a safe and yet numerically affordable number of such starting points is not a priori obvious. The aim of our work here is to systematically reveal all stationary points of the potential by solving the system of equations which originates from the stationarity condition. We restrict to the tree-level and show, that this can be done for the full parameter space including CP violation. Our approach employs algebraic calculations of Gröbner bases, for which a short introduction is given in appendix A. Whether a given parameter set for the classical Higgs potential leads to the required EWSB at the global minimum or not, can be determined unambiguously with our method.

Chapter 2

The general Two-Higgs-Doublet Model

2.1 Lagrangian and real orbit variables

2.1.1 Lagrangian

We denote the two complex Higgs-doublet fields by

$$\varphi_i(x) = \begin{pmatrix} \varphi_i^+(x) \\ \varphi_i^0(x) \end{pmatrix} \tag{2.1}$$

with $i = 1, 2$. Hence we have eight real scalar degrees of freedom. The most general $SU(2)_L \times U(1)_Y$ -invariant Lagrangian for the THDM can be written as

$$\mathscr{L}_{\text{THDM}} = \mathscr{L}_\varphi + \mathscr{L}_{\text{Yuk}} + \mathscr{L}', \tag{2.2}$$

where the pure Higgs boson Lagrangian is given by

$$\mathscr{L}_\varphi = \sum_{i=1,2} \left(\mathcal{D}_\mu \varphi_i\right)^\dagger \left(\mathcal{D}^\mu \varphi_i\right) - V(\varphi_1, \varphi_2)\,. \tag{2.3}$$

This term replaces the kinetic terms of the Higgs boson and the Higgs potential in the SM Lagrangian, see e.g. [64]. Here and in the following, summation over repeated indices is understood unless explicitly noted. The covariant derivative is

$$\mathcal{D}_\mu = \partial_\mu + igW_\mu^a \mathbf{T}_a + ig' B_\mu \mathbf{Y}, \tag{2.4}$$

where $\mathbf{T}_a$ and $\mathbf{Y}$ are the generating operators of weak-isospin and weak-hypercharge transformations. For the Higgs doublets we have $\mathbf{T}_a = \sigma_a/2$, where σ_a $(a = 1, 2, 3)$ are the Pauli matrices. We assume both doublets to have weak hypercharge $y = 1/2$. Further, $\mathscr{L}_{\text{Yuk}}$ are the Yukawa-interaction terms of the Higgs fields with fermions which can be parameterised in the fully general case by

$$\mathscr{L}_{\text{Yuk}} = -\lambda_{ik}^{l\,j}\, \bar{L}_i^L \varphi_j l_k^R - \lambda_{ik}^{d'\,j}\, \bar{Q}_i^L \varphi_j d_k'^R - \lambda_{ik}^{u\,j}\, \bar{Q}_i^L \epsilon \varphi_j^* u_k^R + h.c. \tag{2.5}$$

fermion multiplet	t	y	fermion generations $i=1$	$i=2$	$i=3$
L_i^L	$1/2$	$-1/2$	$\begin{pmatrix}\nu_e\\ e_L^-\end{pmatrix}$	$\begin{pmatrix}\nu_\mu\\ \mu_L^-\end{pmatrix}$	$\begin{pmatrix}\nu_\tau\\ \tau_L^-\end{pmatrix}$
l_i^R	0	-1	e_R	μ_R	τ_R
Q_i^L	$1/2$	$1/6$	$\begin{pmatrix}u_L\\ d'_L\end{pmatrix}$	$\begin{pmatrix}c_L\\ s'_L\end{pmatrix}$	$\begin{pmatrix}t_L\\ b'_L\end{pmatrix}$
u_i^R	0	$2/3$	u_R	c_R	t_R
d_i^R	0	$-1/3$	d_R	s_R	b_R

Table 2.1: Fermion fields and their quantum numbers weak isospin t and hypercharge y. Here, the index L, R denotes the chirality of the fields. The mass eigenstates $d_i'^{L,R}$ are obtained from the fields $d_i'^{L,R}$ by the Cabbibo-Kobayashi-Maskawa basis transformation.

with the fermion electroweak multiplets Q_i^L, L_i^L, l_i^R, $d_i'^R$, u_i^R $(i=1,2,3)$ and complex Yukawa couplings $\lambda_{ik}^{l\,j}$, $\lambda_{ik}^{d'\,j}$, $\lambda_{ik}^{u\,j}$. For the usual flavour bases choice, see also [64], the fermion fields may be identified as in table 2.1, where we also give their quantum numbers. We shall specify this choice when needed in section 2.7.2. Finally, $\mathscr{L}'$ contains the terms of the Lagrangian without Higgs fields. We do not specify $\mathscr{L}'$ here since they are not relevant for our analysis. The Higgs potential V in the THDM will be specified below and discussed extensively.

We remark that in the MSSM the two Higgs doublets H_d and H_u carry hypercharges $y=-1/2$ and $y=+1/2$, respectively, whereas here we use the conventional definition of the THDM with both doublets carrying $y=+1/2$. However, our analysis can be translated to the other case, see for example (3.1) in [50], by setting

$$\begin{aligned}\varphi_1^\alpha &= -\epsilon_{\alpha\beta}(H_d^\beta)^*,\\ \varphi_2^\alpha &= H_u^\alpha,\end{aligned} \tag{2.6}$$

where ϵ is given by

$$\epsilon = \begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}. \tag{2.7}$$

The most general gauge invariant and renormalisable potential $V(\varphi_1,\varphi_2)$ for the two Higgs doublets φ_1 and φ_2 can be written, following [75], as

$$V(\varphi_1,\varphi_2) = Y_{ij}\,(\varphi_i^\dagger\varphi_j) + \frac{1}{2}Z_{ijkl}\,(\varphi_i^\dagger\varphi_j)(\varphi_k^\dagger\varphi_l) \tag{2.8}$$

where summation over repeated Higgs flavour indices $i,j,k,l\in\{1,2\}$ is understood here and in the following. The parameters Y_{ij}, Z_{ijkl} are in general complex and satisfy

$$Z_{ijkl} = Z_{klij} \tag{2.9}$$

such that their definition is unique and

$$Y_{ij} = Y_{ji}^*, \qquad\qquad Z_{ijkl} = Z_{jilk}^* \tag{2.10}$$

such that the potential is hermitian. A change of basis for the Higgs fields

$$\varphi_i \longrightarrow \varphi_i' = U_{ii'}\,\varphi_{i'}\,, \qquad (U_{ij}) \in U(2)\,, \tag{2.11}$$

leaves the kinetic terms invariant, while the potential keeps its form if its parameters are changed according to

$$Y_{ij} \longrightarrow Y_{ij}' = U_{ii'}(U^\dagger)_{jj'}\,, Y_{i'j'}\,, \tag{2.12}$$

$$Z_{ijkl} \longrightarrow Z_{ijkl}' = U_{ii'}(U^\dagger)_{jj'}U_{kk'}(U^\dagger)_{ll'}\,, Z_{i'j'k'l'}\,. \tag{2.13}$$

We note that the global phase factor of U cancels in the potential. Denoting the group of Higgs basis changes by

$$U(2)_\varphi \tag{2.14}$$

it is only the subgroup

$$SU(2)_\varphi \subset U(2)_\varphi \tag{2.15}$$

which changes the parameters of the potential. Instead of the representation (2.8) we will discuss the Higgs potential using real degrees of freedom as introduced in the following subsection.

2.1.2 Real orbit variables

It is convenient to discuss the properties of the potential such as its stability and its spontaneous symmetry breaking directly in terms of gauge invariant expressions. For this purpose we arrange the Higgs fields in a 2×2 matrix

$$\phi := \begin{pmatrix} \varphi_1^+ & \varphi_1^0 \\ \varphi_2^+ & \varphi_2^0 \end{pmatrix} \tag{2.16}$$

and define the 2×2 matrix

$$\underline{K} := \phi\,\phi^\dagger\,, \tag{2.17}$$

which contains all $SU(2)_L \times U(1)_Y$ -invariant scalar products,

$$K_{ij} = \varphi_j^\dagger \varphi_i\,, \tag{2.18}$$

with $\underline{K} = (K_{ij})$, $i,j \in \{1,2\}$. We decompose this matrix with respect to a complete set of complex matrices consisting of generators of $SU(2)_\varphi$ and the unit matrix $\mathbb{1}$,

$$\underline{K} = \frac{1}{2}\,(K_0\,\mathbb{1} + \boldsymbol{K}\,\boldsymbol{\sigma}) \tag{2.19}$$

where here and in the following boldface symbols are used to denote three component vectors. The gauge invariant functions defined by this decomposition are

$$K_0 = \mathrm{tr}(\underline{K}\mathbb{1})\,, \qquad\qquad \boldsymbol{K} = \mathrm{tr}(\underline{K}\boldsymbol{\sigma})\,. \tag{2.20}$$

Explicitly we get from (2.18) and (2.19)

$$\begin{aligned} \varphi_1^\dagger\varphi_1 &= \frac{1}{2}(K_0+K_3)\,, \qquad & \varphi_1^\dagger\varphi_2 &= \frac{1}{2}(K_1+iK_2)\,, \\ \varphi_2^\dagger\varphi_2 &= \frac{1}{2}(K_0-K_3)\,, \qquad & \varphi_2^\dagger\varphi_1 &= \frac{1}{2}(K_1-iK_2)\,, \end{aligned} \tag{2.21}$$

and from (2.20)

$$\begin{aligned} K_0 &= \varphi_1^\dagger\varphi_1+\varphi_2^\dagger\varphi_2\,, \qquad & K_1 &= 2\,\mathrm{Re}\,\varphi_1^\dagger\varphi_2\,, \\ K_3 &= \varphi_1^\dagger\varphi_1-\varphi_2^\dagger\varphi_2\,, \qquad & K_2 &= 2\,\mathrm{Im}\,\varphi_1^\dagger\varphi_2\,, \end{aligned} \tag{2.22}$$

which we arrange into the four vector

$$\tilde{\boldsymbol{K}} = \begin{pmatrix} K_0 \\ \boldsymbol{K} \end{pmatrix}. \tag{2.23}$$

We see directly from (2.17) that $\underline{K}$ is hermitian and positive semidefinite. From (2.20) we conclude that the K_0, K_a (a=1,2,3) are real due to the hermiticity of $\underline{K}$ and satisfy

$$K_0 \geq 0\,, \qquad\qquad \tilde{\boldsymbol{K}}^{\mathrm{T}}\,\tilde{g}\,\tilde{\boldsymbol{K}} \geq 0\,, \tag{2.24}$$

with

$$\tilde{g} = \mathrm{diag}(1,-1,-1,-1) \tag{2.25}$$

since the positive semi-definiteness of $\underline{K}$ gives $\mathrm{tr}\,\underline{K} \geq 0$ and $\det \underline{K} \geq 0$. These constraints on $\tilde{\boldsymbol{K}}$ are also sufficient to find corresponding fields as can be seen in an explicit calculation or in the way described in section 2.8. That is, for any real K_0, $\boldsymbol{K}$ fulfilling (2.24) it is possible to find fields φ_i such that (2.18) and (2.20) are satisfied. Two Higgs field configurations $\varphi_1(x)$, $\varphi_2(x)$ and $\varphi_1'(x)$, $\varphi_2'(x)$ which lie on the same gauge orbit, that is, which differ only by a gauge transformation, obviously give the same $\tilde{\boldsymbol{K}} = \tilde{\boldsymbol{K}}'$. On the other hand, two configurations not connected by a gauge transformation lead to different $\tilde{\boldsymbol{K}}$, $\tilde{\boldsymbol{K}}'$ as can be seen by a short calculation. Since K_0, K_a (a=1,2,3) parameterise the Higgs gauge orbits we will call them *orbit variables*. It is interesting to note that the gauge orbits of the two Higgs doublets are parameterised by Minkowski type four vectors $\tilde{\boldsymbol{K}}$ whose domain is the forward light cone, see figure 2.1.

The most general $SU(2)_L \times U(1)_Y$-invariant potential can now be written as

$$V = \tilde{\boldsymbol{\xi}}^{\mathrm{T}}\,\tilde{\boldsymbol{K}} + \tilde{\boldsymbol{K}}^{\mathrm{T}}\,\tilde{E}\,\tilde{\boldsymbol{K}} \tag{2.26}$$

where the parameters of the potential are a real four vector $\tilde{\boldsymbol{\xi}}$ and a real symmetric matrix $\tilde{E}$. More explicitly, we have

$$V = V_2 + V_4\,, \tag{2.27}$$

$$V_2 = \xi_0\,K_0 + \boldsymbol{\xi}^{\mathrm{T}}\,\boldsymbol{K}\,, \tag{2.28}$$

$$V_4 = \eta_{00}\,K_0^2 + 2\,K_0\,\boldsymbol{\eta}^{\mathrm{T}}\,\boldsymbol{K} + \boldsymbol{K}^{\mathrm{T}}\,E\,\boldsymbol{K}\,. \tag{2.29}$$

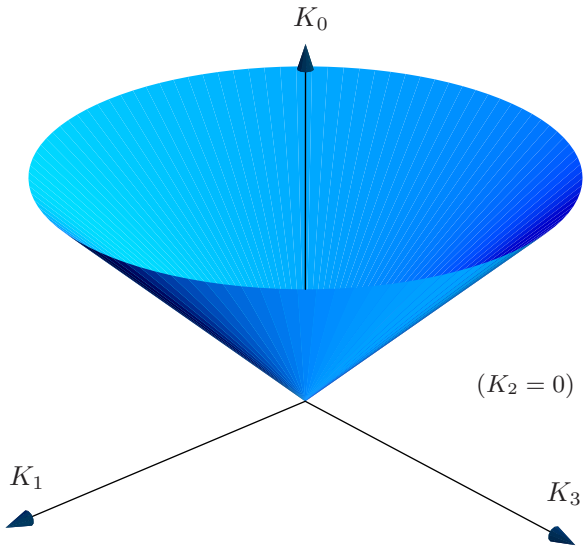

Figure 2.1: The gauge orbits of the Higgs fields in the THDM are parameterised by Minkowski type four vectors $\tilde{\boldsymbol{K}} = (K_0, \boldsymbol{K}^{\mathrm{T}})^{\mathrm{T}} = (K_0, K_1, K_2, K_3)^{\mathrm{T}}$, which lie on or inside the forward light cone.

and

$$\tilde{\boldsymbol{\xi}} = \begin{pmatrix} \xi_0 \\ \boldsymbol{\xi} \end{pmatrix}, \qquad \tilde{E} = \begin{pmatrix} \eta_{00} & \boldsymbol{\eta}^{\mathrm{T}} \\ \boldsymbol{\eta} & E \end{pmatrix},$$
$$\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}, \qquad \boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}, \qquad E = E^{\mathrm{T}} = \begin{pmatrix} \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{21} & \eta_{22} & \eta_{23} \\ \eta_{31} & \eta_{32} & \eta_{33} \end{pmatrix}. \tag{2.30}$$

Under a change of basis of the Higgs fields (2.11) we find

$$K_0' = K_0\,, \qquad \boldsymbol{K}' = R(U)\boldsymbol{K}\,, \tag{2.31}$$

where $R(U) \in SO(3)$ is defined by

$$U^{\dagger}\sigma^a U = R_{ab}(U)\,\sigma^b\,. \tag{2.32}$$

The form of the Higgs potential (2.26) remains unchanged under the replacement (2.31) if we perform an appropriate transformation of the parameters

$$\begin{aligned} \xi_0' &= \xi_0\,, & \boldsymbol{\xi}' &= R(U)\,\boldsymbol{\xi}\,, \\ \eta_{00}' &= \eta_{00}\,, & \boldsymbol{\eta}' &= R(U)\,\boldsymbol{\eta}\,, \\ E' &= R(U)\,E\,R^{\mathrm{T}}(U)\,. \end{aligned} \tag{2.33}$$

The relation (2.32) defines for any $U \in U(2)$ uniquely the corresponding real rotation R, while it specifies for an arbitrary $R \in SO(3)$ a corresponding U up to the global phase

factor. We can therefore diagonalise E, thereby reducing the number of parameters of V by three. The Higgs potential is then determined by only 11 real parameters. Considering the subgroup $SU(2)_\varphi$ of basis changes the transition from the scalar products to the orbit variables means in a group theoretical notation

$$2 \otimes 2 = 1 \oplus 3\,, \tag{2.34}$$

where K_0 is the real flavour singlet and $\boldsymbol{K}$ the real flavour triplet.

The specification of the domain of the orbit variables K_0, $\boldsymbol{K}$ enables us to discuss the potential directly in the form (2.26) with all gauge degrees of freedom eliminated, reduced powers of the occurring dynamical variables and real degrees of freedom only. In particular, this scheme allows us to determine various features of the general THDM in a clear geometrical language.

2.2 Stability of the potential

According to subsection 2.1.2 we can analyse the properties of the potential (2.26) as a function of K_0 and $\boldsymbol{K}$ on the domain determined by $K_0 \geq 0$ and $K_0^2 \geq \boldsymbol{K}^2$. For $K_0 > 0$ we define

$$\boldsymbol{k} := \boldsymbol{K}/K_0\,. \tag{2.35}$$

In fact, we have $K_0 = 0$ only for $\varphi_1 = \varphi_2 = 0$, and the potential $V = 0$ in this case. From (2.26) and (2.35) we obtain for $K_0 > 0$

$$V_2 = K_0\, J_2(\boldsymbol{k})\,, \qquad J_2(\boldsymbol{k}) := \xi_0 + \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{k}\,, \tag{2.36}$$

$$V_4 = K_0^2\, J_4(\boldsymbol{k}), \qquad J_4(\boldsymbol{k}) := \eta_{00} + 2\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{k} + \boldsymbol{k}^{\mathrm{T}} E \boldsymbol{k}\,, \tag{2.37}$$

where we introduce the functions $J_2(\boldsymbol{k})$ and $J_4(\boldsymbol{k})$ on the domain $|\boldsymbol{k}| \leq 1$.

For the potential to be stable, it must be bounded from below. This can be tested by considering its behaviour at large field configurations, that is, at large K_0. The stability is determined by the behaviour of V in the limit of large field configurations, that is, in the limit $K_0 \longrightarrow \infty$. Thus, the relevant information is given by the signs of $J_4(\boldsymbol{k})$ and $J_2(\boldsymbol{k})$ in (2.37) and (2.36). For the theory to be at least *marginally* stable

$$\left.\begin{array}{lll} J_4(\boldsymbol{k}) > 0 & \text{or} & \\ J_4(\boldsymbol{k}) = 0 & \text{and} & J_2(\boldsymbol{k}) \geq 0 \end{array}\right\} \text{ for all } |\boldsymbol{k}| \leq 1 \tag{2.38}$$

is necessary and sufficient, since this condition is equivalent to $V \geq 0$ for $K_0 \longrightarrow \infty$ in all possible directions $\boldsymbol{k}$. The more robust stability property $V \longrightarrow \infty$ for $K_0 \longrightarrow \infty$ and any $\boldsymbol{k}$ can either be guaranteed by

$$\left.\begin{array}{lll} J_4(\boldsymbol{k}) > 0 & \text{or} & \\ J_4(\boldsymbol{k}) = 0 & \text{and} & J_2(\boldsymbol{k}) > 0 \end{array}\right\} \text{ for all } |\boldsymbol{k}| \leq 1 \tag{2.39}$$

in a *weak* sense, or by

$$J_4(\boldsymbol{k}) > 0 \quad \text{for all } |\boldsymbol{k}| \leq 1 \tag{2.40}$$

in a *strong* sense, that is by the quartic terms of V solely.

To assure $J_4(\boldsymbol{k})$ being positive (semi-)definite, it is sufficient to consider its value for all stationary points of $J_4(\boldsymbol{k})$ on the domain $|\boldsymbol{k}| < 1$, and for all stationary points on the boundary $|\boldsymbol{k}| = 1$. This holds, because the global minimum of the continuous function $J_4(\boldsymbol{k})$ is reached on the *compact domain* $|\boldsymbol{k}| \leq 1$, and it is among those stationary points. This leads to bounds on η_{00}, η_a and η_{ab}, which parameterise the quartic term V_4 of the potential. Stationary points of $J_4(\boldsymbol{k})$ with $|\boldsymbol{k}| < 1$ fulfil

$$\nabla_{\boldsymbol{k}} J_4(\boldsymbol{k}) = 2(\boldsymbol{\eta} + E\boldsymbol{k}) = 0\,, \qquad 1 - \boldsymbol{k}^2 > 0\,. \tag{2.41}$$

Stationary points on the domain boundary $|\boldsymbol{k}| = 1$ satisfy

$$\nabla_{\boldsymbol{k}} \big(J_4(\boldsymbol{k}) + u(1 - \boldsymbol{k}^2)\big) = 2(\boldsymbol{\eta} + (E-u)\boldsymbol{k}) = 0\,, \qquad 1 - \boldsymbol{k}^2 = 0\,, \tag{2.42}$$

where u is a Lagrange multiplier. The conditions for both types of stationary points may be written in the combined form

$$\nabla_{\boldsymbol{k}} \big(J_4(\boldsymbol{k}) + u(1 - \boldsymbol{k}^2)\big) = 2(\boldsymbol{\eta}+(E-u)\boldsymbol{k}) = 0\,, \qquad u(1-\boldsymbol{k}^2) = 0\,, \qquad 1-\boldsymbol{k}^2 \geq 0\,, \tag{2.43}$$

where the stationary points with $|\boldsymbol{k}| < 1$ occur with the Lagrange multiplier $u = 0$. Depending on the parameters η_a and η_{ab}, there can be exceptional solutions $(\boldsymbol{k}, u)$ of (2.43) where $\det(E - u) = 0$, i.e. where u is an eigenvalue of E. For regular values of u such that $\det(E - u) \neq 0$ the stationary points fulfil

$$\boldsymbol{k}(u) = -(E - u)^{-1}\boldsymbol{\eta}, \tag{2.44}$$

where either the Lagrange multiplier u is determined from $\boldsymbol{k}(u)^2 = 1$ or $u = 0$ and $\boldsymbol{k}(0)^2 < 1$ hold. For all stationary points we can determine their existence and their J_4 values using one function only. We define the Lagrange type function

$$f(u) := \big(J_4(\boldsymbol{k}) + u(1 - \boldsymbol{k}^2)\big)_{\boldsymbol{k}=\boldsymbol{k}(u)} \tag{2.45}$$

with $\boldsymbol{k}(u)$ as in (2.44). For all regular stationary points $\boldsymbol{k}$ of $J_4(\boldsymbol{k})$ holds

$$f(u) = J_4(\boldsymbol{k})\big|_{\text{stat}}\,, \tag{2.46}$$

$$f'(u) = 1 - \boldsymbol{k}^2\big|_{\text{stat}}\,. \tag{2.47}$$

There are stationary points of $J_4(\boldsymbol{k})$ with $|\boldsymbol{k}| < 1$ and $|\boldsymbol{k}| = 1$ exactly if $f'(0) > 0$ and $f'(u) = 0$, respectively, and the value of $J_4(\boldsymbol{k})$ is then given by $f(u)$. Explicitly, we have

$$f(u) = u + \eta_{00} - \boldsymbol{\eta}^{\mathrm{T}}(E - u)^{-1}\boldsymbol{\eta}\,, \tag{2.48}$$

$$f'(u) = 1 - \boldsymbol{\eta}^{\mathrm{T}}(E - u)^{-2}\boldsymbol{\eta}\,, \tag{2.49}$$

from the definition (2.45).

In a basis where $E = \mathrm{diag}(\mu_1, \mu_2, \mu_3)$ we obtain:

$$f(u) = u + \eta_{00} - \sum_{a=1}^{3} \frac{\eta_a^2}{\mu_a - u}\,, \tag{2.50}$$

$$f'(u) = 1 - \sum_{a=1}^{3} \frac{\eta_a^2}{(\mu_a - u)^2}\,. \tag{2.51}$$

The derivative $f'(u)$ has at most 6 zeros. Notice that there are no exceptional solutions if all three components of $\boldsymbol{\eta}$ are different from zero in this basis .

The function $f(u)$ given by (2.48) allows us to discuss also the exceptional solutions of (2.41) and (2.42). Consider first $|\boldsymbol{k}| < 1$ and suppose that $\det E = 0$. In the basis where E is diagonal we have

$$\det E = \mu_1\,\mu_2\,\mu_3 = 0 \tag{2.52}$$

and (2.41) reads

$$\begin{aligned} \mu_1\,k_1 &= -\eta_1\,, \\ \mu_2\,k_2 &= -\eta_2\,, \\ \mu_3\,k_3 &= -\eta_3\,. \end{aligned} \tag{2.53}$$

Obviously, a solution of (2.53) is only possible if with $\mu_a = 0$ also $\eta_a = 0$ $(a = 1, 2, 3)$. Therefore, we see from (2.50) that exceptional solutions with $|\boldsymbol{k}| < 1$ are only possible if $f(u)$ stays finite at $u = 0$. That is, the pole which would correspond to $\mu_a = 0$ must have residue zero. Suppose now that indeed $\eta_a = 0$ for all a where $\mu_a = 0$. Take as an example $\mu_1 = \mu_2 = 0$ and $\eta_1 = \eta_2 = 0$, but $\mu_3 \neq 0$. Then we get the general solution of (2.53) as

$$k_3 = -\frac{\eta_3}{\mu_3}, \tag{2.54}$$

with k_1, k_2 arbitrary but satisfying $\boldsymbol{k}^2 < 1$. We can write this as

$$\boldsymbol{k} = \boldsymbol{k}_\parallel + \boldsymbol{k}_\perp, \tag{2.55}$$

where

$$\begin{aligned} \boldsymbol{k}_\parallel &= -\frac{1}{\mu_3}\boldsymbol{\eta}\,, \qquad E\,\boldsymbol{k}_\perp = 0\,, \\ \boldsymbol{k}_\perp^2 &< 1 - \boldsymbol{k}_\parallel^2 = 1 - \left(\frac{\eta_3}{\mu_3}\right)^2. \end{aligned} \tag{2.56}$$

Inserting the solution $\boldsymbol{k}$ from (2.55), (2.56) in $J_4(\boldsymbol{k})$ we get for the functions (2.50) and (2.51)

$$f(0) = J_4(\boldsymbol{k})\big|_{\mathrm{stat}}\,, \tag{2.57}$$

$$f'(0) = 1 - \boldsymbol{k}_\parallel^2 > \boldsymbol{k}_\perp^2 \geq 0\,. \tag{2.58}$$

Clearly these arguments work similarly, if only one of the μ_a is equal to zero or all three μ_a are zero. In all cases (2.57) holds for the exceptional points with $|\boldsymbol{k}| < 1$, which can exist only if $f(u)$ has no pole at $u = 0$. Since (2.57) involves only "scalar" quantities, it holds in any basis.

The case of exceptional solutions for $|\boldsymbol{k}| = 1$ can be treated in an analogous way. An exceptional solution of (2.42) with $u = \mu_a$ $(a = 1, 2, 3)$ can only exist if the corresponding $\eta_a = 0$. Then $f(u)$ has no pole for $u = \mu_a$ and the exceptional solutions of (2.42) fulfil

$$\boldsymbol{k} = \boldsymbol{k}_\parallel + \boldsymbol{k}_\perp \,, \tag{2.59}$$

with

$$\boldsymbol{k}_\parallel = -\,(E-u)^{-1}\boldsymbol{\eta}\big|_{u=\mu_a} \,, \quad (E-\mu_a)\boldsymbol{k}_\perp = 0$$

and

$$f(\mu_a) = J_4(\boldsymbol{k})\big|_{\text{stat}} \,, \tag{2.60}$$

$$f'(\mu_a) = 1 - \boldsymbol{k}_\parallel^2 = \boldsymbol{k}_\perp^2 \geq 0 \,. \tag{2.61}$$

Note that if a solution is possible, $\boldsymbol{k}_\perp$ may be any linear combination of the eigenvectors to the eigenvalue μ_a of E, where the norm is given by $|\boldsymbol{k}_\perp| = \sqrt{f'(\mu_a)}$. Thus we see that the function $f(u)$ is very useful for discussing the stability of the THDM potential. What we have shown so far can be formulated as follows.

Consider the functions $f(u)$ and $f'(u)$. Denote by I,

$$I = \{u_1, \ldots, u_n\}, \tag{2.62}$$

the following set of values of u. Include in I all u where $f'(u) = 0$. Add $u = 0$ to I if $f'(0) > 0$. Consider then the eigenvalues μ_a $(a = 1, 2, 3)$ of E. Add those μ_a to I where $f(\mu_a)$ is finite and $f'(\mu_a) \geq 0$. We have $n \leq 10$. The values of the function $J_4(\boldsymbol{k})$ at its stationary points are given by

$$J_4(\boldsymbol{k})\big|_{\text{stat}} = f(u_i) \tag{2.63}$$

with $u_i \in I$. The potential is stable if $f(u_i) > 0$ for all $u_i \in I$. Then the stability is given solely by the quartic terms in the potential. The potential is unstable if we have $f(u_i) < 0$ for at least one $u_i \in I$. If we have $f(u_i) \geq 0$ for all $u_i \in I$ and $f(u_i) = 0$ for at least one $u_i \in I$ we have to consider in addition $J_2(\boldsymbol{k})$ in order to decide on the stability of the potential.

We turn now to this latter case. We shall show that we have to consider in addition the function

$$g(u) := \xi_0 - \boldsymbol{\xi}^{\mathrm{T}}(E-u)^{-1}\boldsymbol{\eta} \tag{2.64}$$

for the stationary points of $J_4(\boldsymbol{k})$ with

$$J_4(\boldsymbol{k})\big|_{\text{stat}} = f(u_i) = 0 \,. \tag{2.65}$$

For the vectors $\boldsymbol{k}$ satisfying (2.65), we get

$$J_2(\boldsymbol{k})\big|_{\text{stat}} = g(u_i) \tag{2.66}$$

if $u_i \neq \mu_a$, that is u_i is not an eigenvalue of E. If u_i is an eigenvalue of E, that is $u_i = \mu_a \in I$, we have

$$\inf_{\boldsymbol{k}} J_2(\boldsymbol{k}) = g(u_i) - |\boldsymbol{\xi}_\perp(u_i)| \sqrt{f'(u_i)}\,, \tag{2.67}$$

where the infimum is taken over all exceptional solutions $\boldsymbol{k}$ to u_i and

$$\boldsymbol{\xi}_\perp(u_i) := \sum_{l=1}^{N} \frac{\boldsymbol{\xi}\boldsymbol{e}_l(u_i)}{|\boldsymbol{e}_l(u_i)|^2}\, \boldsymbol{e}_l(u_i)\,. \tag{2.68}$$

Here, $\boldsymbol{e}_l(u_i)$ $(l = 1, \ldots, N)$ are the $N \leq 3$ eigenvectors to u_i. We summarise our findings in a theorem.

Theorem 2.2.1. *The most general potential of the Two-Higgs-Doublet Model has the form* (2.26)*. Its stability is decided in the following way. If the potential has only the quadratic term V_2, it is stable for $\xi_0 > |\boldsymbol{\xi}|$, marginally stable for $\xi_0 = |\boldsymbol{\xi}|$ and unstable for $\xi_0 < |\boldsymbol{\xi}|$. Suppose now that $V_4 \not\equiv 0$. We construct then the functions $f(u)$* (2.48)*, $f'(u)$* (2.49) *and $g(u)$* (2.64)*, and the set I* (2.62) *of (at most 10) u values.*

1. *If $f(u_i) > 0$ for all $u_i \in I$ the potential is stable in the strong sense* (2.40)*.*
2. *If $f(u_i) < 0$ for at least one $u_i \in I$ the potential is unstable.*
3. *If $f(u_i) \geq 0$ for all $u_i \in I$ and $f(u_i) = 0$ for at least one $u_i \in I$ we consider also the function $g(u)$* (2.64)*. The potential is stable in the weak sense* (2.39) *if for all $u_i \in I$ where $f(u_i) = 0$ the following holds (see* (2.66) *to* (2.68)*):*

$$g(u_i) > 0 \text{ if } u_i \neq \mu_a\,, \tag{2.69}$$

$$g(u_i) - |\boldsymbol{\xi}_\perp(u_i)| \sqrt{f'(u_i)} > 0 \text{ if } u_i = \mu_a\,. \tag{2.70}$$

If in some or all of these cases we have $= 0$ instead of > 0 we have marginal stability (2.38)*. If in at least one case we have < 0 instead of > 0 the potential is unstable.*

Our theorem gives a complete characterisation of the stability properties of the general THDM potential. In the following subsection we apply the theorem to assert the strong stability condition (2.40) for a specific potential. An application for the weaker stability condition (2.39) is given in subsection 2.6.1.

2.3 Electroweak symmetry breaking

2.3.1 Stationarity conditions via orbit variables

After our stability analysis in the preceding section we now consider the stationary points of the potential, since among these points there are the local and global minima and thus

the vacuum. To determine the stationary points of the Higgs potential we consider (2.26) and take the constraints (2.24) into account. The stationary points with respect to the Higgs fields may be determined using orbit variables by employing standard techniques from constrained optimisation. We distinguish three possible types of stationary points by their location with respect to the domain boundaries.

Classification 2.3.1. *The stationary points of the THDM potential* (2.26) *are described by the following cases:*

- $K_0 = |\boldsymbol{K}| = 0$*: The tip of the forward light cone*

$$\tilde{\boldsymbol{K}} = 0 \tag{2.71}$$

  is always a stationary point as a direct consequence of the definitions.

- $K_0 = |\boldsymbol{K}| > 0$*: Stationary points on the forward light cone fulfil*

$$\nabla_{\tilde{\boldsymbol{K}}} \left(V - u\, \tilde{\boldsymbol{K}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{K}} \right) = 0 \,, \tag{2.72a}$$

$$\tilde{\boldsymbol{K}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{K}} = 0 \,, \tag{2.72b}$$

$$K_0 > 0 \,, \tag{2.72c}$$

  where u is a Lagrange multiplier.

- $K_0 > |\boldsymbol{K}|$*: Stationary points in the inner part of the forward light cone fulfil*

$$\nabla_{\tilde{\boldsymbol{K}}} V = 0 \,, \tag{2.73a}$$

$$\tilde{\boldsymbol{K}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{K}} > 0 \,, \tag{2.73b}$$

$$K_0 > 0 \,. \tag{2.73c}$$

The existence of stationary solutions as well as the value of a corresponding Lagrange multiplier is independent of the chosen Higgs flavour basis.

The vacuum will be at least a local minimum of the potential. A stable vacuum is the global minimum of a stable potential. We denote the vacuum expectation values (vevs) of the Higgs fields by

$$v_i^+ := \langle \varphi_i^+ \rangle \,, \qquad v_i^0 := \langle \varphi_i^0 \rangle \,, \tag{2.74}$$

with $i = 1, 2$. In general the v_i^+, v_i^0 are complex numbers. To exhibit the consequences for gauge invariance we consider the matrices (2.16) at the vacuum,

$$\langle \phi \rangle = \begin{pmatrix} v_1^0 & v_1^+ \\ v_2^0 & v_2^+ \end{pmatrix} . \tag{2.75}$$

The brackets denote also for the other quantities their vacuum expectation value in the following. The symmetry properties of the vacuum can be formulated using the rank

of $\langle\phi\rangle$, which is a gauge and Higgs flavour basis invariant quantity. We shall establish the connection to the orbit variables of the vacuum and discuss the implications for all possible cases. We observe that for any matrix A the relation $\operatorname{rank} A^\dagger A = \operatorname{rank} A$ holds and therefore

$$\operatorname{rank}\phi = \operatorname{rank}\underline{K}\,. \tag{2.76}$$

Furthermore we have

$$\tilde{\boldsymbol{K}}^{\mathrm{T}}\tilde{g}\tilde{\boldsymbol{K}} = 4\det\underline{K} \tag{2.77}$$

so that we find the following cases.

Classification 2.3.2. *The electroweak symmetry breaking induced by a vacuum of the THDM is characterised by the location of its orbit variables with respect to the forward-light cone.*

- $\langle K_0\rangle = \langle\boldsymbol{K}\rangle = 0$*: The trivial vacuum has*

  $$\operatorname{rank}\langle\phi\rangle = \operatorname{rank}\langle\underline{K}\rangle = 0 \tag{2.78}$$

  and vanishing vacuum expectation values for all Higgs fields. In this case, no symmetry is spontaneously broken.

- $\langle K_0\rangle = |\langle\boldsymbol{K}\rangle| > 0$*: A vacuum on the forward light cone features*

  $$\operatorname{rank}\langle\phi\rangle = \operatorname{rank}\langle\underline{K}\rangle = 1 \tag{2.79}$$

  such that after performing a $SU(2)_L \times U(1)_Y$ transformation we achieve

  $$\begin{pmatrix} v_1^+ \\ v_2^+ \end{pmatrix} = 0\,, \qquad \begin{pmatrix} v_1^0 \\ v_2^0 \end{pmatrix} = \begin{pmatrix} |v_1^0| \\ v_2^0 \end{pmatrix} \neq 0\,, \tag{2.80}$$

  and identify the unbroken $U(1)$ gauge group with the electromagnetic one.

- $\langle K_0\rangle > |\langle\boldsymbol{K}\rangle|$*: A vacuum in the interior of the forward light cone has*

  $$\operatorname{rank}\langle\phi\rangle = \operatorname{rank}\langle\underline{K}\rangle = 2\,. \tag{2.81}$$

  such that in any gauge or Higgs flavour basis at least one charged and one neutral vacuum expectation value is non-vanishing. This means that the full gauge group $SU(2)_L \times U(1)_Y$ is broken.

Clearly, the case $\langle K_0\rangle = |\langle\boldsymbol{K}\rangle| > 0$ is required for the vacuum of a phenomenologically acceptable theory. Figure 2.2 illustrates the different classes of vacua.

A scheme often used in the literature, also for other types of Higgs sectors, is to reparameterise the potential using the stationarity conditions for the Higgs fields. That is, the stationarity conditions are not solved but only one of the (in general many) stationary solutions is picked and parameterised. Typically the parameters of the terms quadratic in the Higgs fields are eliminated and the vacuum expectation values are considered as input

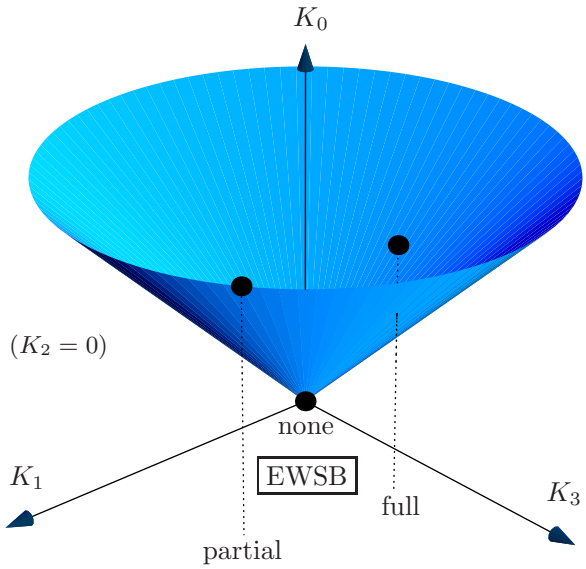

Figure 2.2: The electroweak symmetry is fully broken, partially broken or unbroken by the vacuum, depending on its position with respect to the forward light cone.

parameters instead. Physical quantities such as the electroweak scale and the Higgs masses may now be directly calculated in terms of these parameters. In a second step, parameters such as the vacuum expectation values are usually traded for the electroweak scale and maybe other physical quantities like Higgs masses. It is obvious that choosing such a set of input parameters instead of the original Lagrangian parameters has the advantage that physical observables such as the known electroweak scale may easily be adjusted to the observed values. This type of reparameterisation works actually even better when using orbit variables as we shall show in section 2.4. Note, that even without explicitly solving the stationarity equations, one achieves with this reparameterisation that the corresponding vacuum solution is a stationary point of the potential by construction. If one arranges for positive physical Higgs masses it will furthermore be a local minimum. In simple cases, where e.g. the potential is at least weakly stable and has only one local minimum, this automatically ensures a stable vacuum. In the general case however, the reparameterisation alone is not sufficient to ensure the stability of the vacuum. We will show that for the THDM indeed more than one local minimum can exist. For a stable vacuum it needs to be checked that it is indeed the global minimum of the potential. Also, to ensure sufficient stability of a meta-stable vacuum requires more global knowledge of the potential structure. Therefore, we shall develop means to calculate the stationary points and derive various statements for their classification and existence in the following.

2.3.2 Determination of stationary points

We write the non-trivial stationarity conditions (2.73a) and (2.72a) in the combined form

$$\nabla_{\tilde{\boldsymbol{K}}} \left(V - u\, \tilde{\boldsymbol{K}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{K}} \right) = \tilde{\boldsymbol{\xi}} + 2(\tilde{E} - u\tilde{g})\tilde{\boldsymbol{K}} = 0 \tag{2.82a}$$

$$u\, \tilde{\boldsymbol{K}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{K}} = 0\,, \tag{2.82b}$$

$$\tilde{\boldsymbol{K}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{K}} \geq 0\,, \tag{2.82c}$$

$$K_0 > 0\,, \tag{2.82d}$$

where a stationary point in the inner part of the forward light cone has Lagrange multiplier $u = 0$.

There may be up to 4 values $u = \tilde{\mu}_a$ with $a = 1, \ldots, 4$ for which $\det(\tilde{E} - u\tilde{g}) = 0$. Depending on the potential some or all of them may lead to exceptional solutions of (2.82a). Regular solutions of (2.82a) with $\det(\tilde{E} - u\tilde{g}) \neq 0$ are uniquely determined by

$$\tilde{\boldsymbol{K}}(u) = -\frac{1}{2}\big(\tilde{E} - u\tilde{g}\big)^{-1} \tilde{\boldsymbol{\xi}}\,. \tag{2.83}$$

We define the function

$$\tilde{f}(u) := \left(V - u\, \tilde{\boldsymbol{K}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{K}} \right) \Big|_{\tilde{\boldsymbol{K}} = \tilde{\boldsymbol{K}}(u)} \tag{2.84}$$

from which we find explicitly

$$\tilde{f}(u) = -\frac{1}{4} \tilde{\boldsymbol{\xi}}^{\mathrm{T}} \big(\tilde{E} - u\tilde{g}\big)^{-1} \tilde{\boldsymbol{\xi}}\,, \tag{2.85}$$

$$\tilde{f}'(u) = -\frac{1}{4} \tilde{\boldsymbol{\xi}}^{\mathrm{T}} \big(\tilde{E} - u\tilde{g}\big)^{-1} \tilde{g} \big(\tilde{E} - u\tilde{g}\big)^{-1} \tilde{\boldsymbol{\xi}}\,. \tag{2.86}$$

Employing this function the non-trivial regular stationary points are determined by

$$\tilde{\boldsymbol{K}} = \tilde{\boldsymbol{K}}(u) \text{ for } u \text{ with } \det(\tilde{E} - u\tilde{g}) \neq 0\,, \quad u\, f'(u) = 0\,, \quad f'(u) \leq 0\,, \quad K_0(u) > 0\,. \tag{2.87}$$

The only remaining task is now to find the zeros of $\det(\tilde{E} - u\tilde{g})$ and $\tilde{f}'(u)$, which corresponds to finding the zeros of two univariate polynomials of at most order 4 and 8 respectively. This can easily be done numerically for any potential, in some cases all solutions can be given explicitly in an analytical form, see e.g. sections 2.6.1, 2.6.3 and 2.7.

In many cases, for instance if all values $\tilde{\mu}_1, \ldots, \tilde{\mu}_4$ are different, we can diagonalise the in general non-hermitian matrix $\tilde{g}\tilde{E}$ in the following way:

$$\tilde{g}\tilde{E} = \sum_{a=1}^{4} \tilde{\mu}_a \tilde{\mathbb{P}}_a\,. \tag{2.88}$$

Here the $\tilde{\mathbb{P}}_a$ are quasi-projectors constructed from the normalised right and left eigenvectors $\chi_a, \tilde{\chi}_a$ of $\tilde{g}\tilde{E}$. We have then $\tilde{g}\tilde{E}\,\chi_a = \tilde{\mu}_a\,\chi_a$, $\tilde{\chi}_a\,\tilde{g}\tilde{E} = \tilde{\chi}_a\,\tilde{\mu}_a$, $\tilde{\chi}_a \chi_b = \delta_{ab}$ and can impose as additional normalisation condition $\chi_a^\dagger \chi_a = 1$. The $\tilde{\mathbb{P}}_a$ are given by

$$\tilde{\mathbb{P}}_a = \chi_a \tilde{\chi}_a \tag{2.89}$$

and satisfy

$$\operatorname{tr}\tilde{\mathbb{P}}_a = 1\,, \qquad \tilde{\mathbb{P}}_a\tilde{\mathbb{P}}_b = \begin{cases} \tilde{\mathbb{P}}_a & \text{for } a = b\,, \\ 0 & \text{for } a \neq b\,, \end{cases} \tag{2.90}$$

where $a, b \in \{1, \ldots, 4\}$. In terms of the $\tilde{\mathbb{P}}_a$ (2.85) and (2.86) read

$$\tilde{f}(u) = -\frac{1}{4}\sum_{a=1}^{4} \frac{\tilde{\boldsymbol{\xi}}^{\mathrm{T}}\tilde{\mathbb{P}}_a\,\tilde{g}\,\tilde{\boldsymbol{\xi}}}{\tilde{\mu}_a - u}\,, \tag{2.91}$$

$$\tilde{f}'(u) = -\frac{1}{4}\sum_{a=1}^{4} \frac{\tilde{\boldsymbol{\xi}}^{\mathrm{T}}\tilde{\mathbb{P}}_a\,\tilde{g}\,\tilde{\boldsymbol{\xi}}}{(\tilde{\mu}_a - u)^2}\,. \tag{2.92}$$

Of course, $\tilde{f}(u)$ (2.85) is always a meromorphic function of u but in general poles of higher order than one may also occur.

Alternatively to the method outlined above one may use Gröbner bases to determine the stationary points. This method also ensures at the algebraical level that all stationary points are found, which include in particular the global minimum. The advantage for the THDM with respect to the method discussed above is the full automation of the calculation. However, the Gröbner approach can also be applied to more complicated Higgs potentials, which involve non-linear multivariate stationarity conditions instead of the linear system (2.82a). We will explain the method in the context of the NMSSM in chapter 3.

2.3.3 Criteria for electroweak symmetry breaking

For *any* stationary point the potential is given by

$$V|_{stat} = \frac{1}{2}\tilde{\boldsymbol{K}}^{\mathrm{T}}\tilde{\boldsymbol{\xi}} = -\tilde{\boldsymbol{K}}^{\mathrm{T}}\tilde{E}\tilde{\boldsymbol{K}}. \tag{2.93}$$

Suppose now that the weak stability condition (2.39) holds. Then (2.93) gives for non-trivial stationary points where $\tilde{\boldsymbol{K}} \neq 0$:

$$V|_{stat} < 0\,, \tag{2.94}$$

since the cases $V_4 < 0$ and $V_4 = V_2 = 0$ are excluded by the stability condition. If the parameters fulfil $\xi_0 \geq |\boldsymbol{\xi}|$, this immediately implies $J_2(\boldsymbol{k}) \geq 0$ and hence from (2.39) $V > 0$ for all $\tilde{\boldsymbol{K}} \neq 0$, that is, $\tilde{\boldsymbol{K}} = 0$ is the only stationary point and the global minimum of the potential. For $\xi_0 < |\boldsymbol{\xi}|$ we obtain

$$\left.\frac{\partial V}{\partial K_0}\right|_{\substack{\boldsymbol{k}\text{ fixed},\\ K_0=0}} = \xi_0 + \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{k} < 0 \tag{2.95}$$

for some $\boldsymbol{k}$, i.e. $\tilde{\boldsymbol{K}} = 0$ can not be a local minimum, and the global minimum of V lies at $\tilde{\boldsymbol{K}} \neq 0$.

Whether a non-trivial stationary point is a local minimum, a local maximum or a saddle is determined by the Hessian matrix with respect to the orbit variables. For $u = 0$ the free Hessian matrix

$$H := \left(\frac{\partial^2}{\partial K_\mu \, \partial K_\nu} V \right) = 2\tilde{E} \,, \quad \text{where} \quad \mu, \nu = 0 \ldots 3 \,, \tag{2.96}$$

determines the type of the stationary point. For $u \neq 0$ the sign of u and the bordered Hessian matrix (see for instance [115])

$$\bar{H} := \begin{pmatrix} \frac{\partial^2}{(\partial u)^2} & \frac{\partial^2}{\partial u \, \partial K_\mu} \\ \frac{\partial^2}{\partial u \, \partial K_\mu} & \frac{\partial^2}{\partial K_\mu \, \partial K_\nu} \end{pmatrix} \left(V - u \, \tilde{\boldsymbol{K}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{K}} \right) \,, \quad \text{where} \quad \mu, \nu = 0 \ldots 3 \,, \tag{2.97}$$

$$= \begin{pmatrix} 0 & -2(\tilde{g}\tilde{\boldsymbol{K}})^{\mathrm{T}} \\ -2\tilde{g}\tilde{\boldsymbol{K}} & 2(\tilde{E} - u\tilde{g}) \end{pmatrix} \tag{2.98}$$

may determine the type of the stationary point. If the signs of u and the last three leading principal minors of $\bar{H}$ are $+, -, -, -$ it is a minimum, if they are $-, +, -, +$ it is a maximum, if any other combination of definite signs occur it is a saddle. The occurrence of one or more zeros as is the case e.g. in presence of a degenerate direction might prohibit the classification. Note, that a definite decision is facilitated by the fact, that no higher than second derivatives with respect to the orbit variables occur. This is in contrast to the second derivatives with respect to the Higgs fields, where for zero-directions the fourth order derivatives must be checked. However, if e.g. already the first or second leading principal minor described above is zero it will hide the information with respect to the residual directions, since the following leading principal minors will also be zero. In this case a simple reordering of the orbit variables in (2.97) may give a definite answer. See also the discussion of positive semi-definite matrices in section 2.8.

Suppose that the two points

$$\tilde{\boldsymbol{p}} = \begin{pmatrix} p_0 \\ \boldsymbol{p} \end{pmatrix} , \qquad \tilde{\boldsymbol{q}} = \begin{pmatrix} q_0 \\ \boldsymbol{q} \end{pmatrix} \tag{2.99}$$

with $p_0 \geq |\boldsymbol{p}|$ and $q_0 \geq |\boldsymbol{q}|$ are stationary points of V, that is each of them is a solution of (2.82) together with an appropriate Lagrange multiplier u_p or u_q for $\tilde{\boldsymbol{p}}$ or $\tilde{\boldsymbol{q}}$, respectively. We have

$$\begin{aligned} V(\tilde{\boldsymbol{p}}) - V(\tilde{\boldsymbol{q}}) &= \frac{1}{2} \tilde{\boldsymbol{p}}^{\mathrm{T}} \tilde{\boldsymbol{\xi}} - \frac{1}{2} \tilde{\boldsymbol{q}}^{\mathrm{T}} \tilde{\boldsymbol{\xi}} \\ &= \tilde{\boldsymbol{p}}^{\mathrm{T}} (u_q \, \tilde{g} - \tilde{E}) \tilde{\boldsymbol{q}} - \tilde{\boldsymbol{q}}^{\mathrm{T}} (u_p \, \tilde{g} - \tilde{E}) \tilde{\boldsymbol{p}} \\ &= (u_q - u_p) \, \tilde{\boldsymbol{p}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{q}} \,. \end{aligned} \tag{2.100}$$

Since $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{q}}$ are vectors on or inside the forward light cone, $\tilde{\boldsymbol{p}}^{\mathrm{T}} \tilde{g} \tilde{\boldsymbol{q}}$ is always non-negative. It becomes zero only for $\tilde{\boldsymbol{p}}$ parallel to $\tilde{\boldsymbol{q}}$ and both on the light cone. Furthermore, the case that two different $\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}$ on the forward light cone are parallel can not occur, since then

(2.100) requires $V(\tilde{\boldsymbol{p}}) = V(\tilde{\boldsymbol{q}})$, while (2.93) and (2.94) imply $V(\tilde{\boldsymbol{p}}) \neq V(\tilde{\boldsymbol{q}})$ for that case. Therefore the values of the potential at non-trivial stationary points are strictly ordered with respect to their Lagrange multipliers:

$$u_p > u_q \qquad \Longleftrightarrow \qquad V(\tilde{\boldsymbol{p}}) < V(\tilde{\boldsymbol{q}})\,. \tag{2.101}$$

We found in the previous subsection that a regular solution for some given Lagrange multiplier is uniquely determined, see (2.83). From (2.101) we see that it is the only stationary solution with its specific value of the potential. Degenerate stationary points can therefore not be regular, but need to be exceptional with a Lagrange multiplier given by one of the eigenvalues of $\tilde{g}\tilde{E}$, see previous subsection. We summarise.

Theorem 2.3.3. *For non-trivial stationary points a larger Lagrange multiplier is equivalent to a lower value of the potential, see* (2.101)*, which lies always below that of trivial point, see* (2.94)*. Degenerate stationary points share the same Lagrange multiplier given by an eigenvalue of* $\tilde{g}\tilde{E}$*.*

Assuming $p_0 = |\boldsymbol{p}|$ and $q_0 > |\boldsymbol{q}|$ we get from (2.82a):

$$V(\tilde{\boldsymbol{p}}) - V(\tilde{\boldsymbol{q}}) = (\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{q}})^{\mathrm{T}}\, \tilde{E}\, (\tilde{\boldsymbol{p}} - \tilde{\boldsymbol{q}})\,. \tag{2.102}$$

If the stationary point p on the forward light cone has a positive Lagrange multiplier, its potential value will by (2.101) have a lower potential value than any stationary point q in the interior of the forward light cone. From (2.102) follows in this case that $\tilde{E}$ has a negative eigenvalue. Since the Hessian matrix for q equals $2\tilde{E}$, see (2.96), we see that $\tilde{\boldsymbol{q}}$ cannot be a local minimum. This result and the hierarchies of the stationary points derived above agree with [80].

Suppose now that a stable potential has $\xi_0 < |\boldsymbol{\xi}|$, or equivalently, a non-trivial global minimum. The Lagrange multiplier u_0 of the global minimum is the largest of all occurring Lagrange multipliers due to (2.101), and – as necessary for any minimum – it must be non-negative

$$u_0 \geq 0\,. \tag{2.103}$$

We note that for two different stationary points in the inner part of the domain or with $u = 0$ on its boundary, any linear combination of them with $K_0 \geq |\boldsymbol{K}|$ is a stationary point as well. These points therefore belong to one connected set of degenerate stationary points. Stability requires that this set contains points with $K_0 > |\boldsymbol{K}|$ and is bounded by points with $K_0 = |\boldsymbol{K}|$. If interpreted geometrically, this degenerate set is a line segment, ellipsoidal area or volume. We summarise.

Theorem 2.3.4. *There are the following* mutually exclusive *possibilities for local minima:*

- *one or multiple local minima on the forward light cone* ($K_0 = |\boldsymbol{K}|$)
- or *a degenerate set of solutions in and on the forward light cone* ($K_0 \geq |\boldsymbol{K}|$)
- or *one local minimum in the interior of the forward light cone* ($K_0 > |\boldsymbol{K}|$)

- or *the trivial minimum* ($\tilde{\boldsymbol{K}} = 0$).

The Lagrange multiplier of any non-trivial local minimum is non-negative, and the global minimum has the largest of all occurring Lagrange multipliers.

We know from subsection 2.3.1 that in a theory with the required EWSB, the orbit variables of the vacuum lie on the forward light cone. For the particularly important case that the vacuum is stable, it must be a global minimum of the potential. Together with the discussion above we thus find the following.

Theorem 2.3.5. *The global minimum of a stable potential with the spontaneous electroweak symmetry breaking* $SU(2)_L \times U(1)_Y \longrightarrow U(1)_{em}$

(i) requires $\xi_0 < |\boldsymbol{\xi}|$,

(ii) is given and guaranteed by a stationary point on the forward light cone with the largest occurring Lagrange multiplier u_0*, which is necessarily non-negative.*

2.4 Mass matrices and reparameterisation

We assume a stable potential which leads to the desired symmetry breaking pattern as discussed in the previous sections and derive consequences for the resulting physical fields in the following. The EWSB observed in nature is achieved by a stationary solution on the forward light cone. In the gauge (2.80) we define

$$\langle\varphi_1\rangle = \begin{pmatrix} 0 \\ v_1 \end{pmatrix}, \qquad \langle\varphi_2\rangle = \begin{pmatrix} 0 \\ v_2\, e^{i\zeta} \end{pmatrix}, \qquad v_1, v_2, \zeta \in \mathbb{R} \tag{2.104}$$

and have for the orbit variables of such a partially breaking vacuum

$$\left\langle \tilde{\boldsymbol{K}} \right\rangle = \begin{pmatrix} v_1^2 + v_2^2 \\ 2v_1 v_2 \cos\zeta \\ 2v_1 v_2 \sin\zeta \\ v_1^2 - v_2^2 \end{pmatrix}. \tag{2.105}$$

Note that this decomposition into the field expectation values holds only for the gauge (2.80), while $\langle\tilde{\boldsymbol{K}}\rangle$ itself is of course gauge invariant.

In order to identify the physical content of the Lagrangian, we perform a basis change (2.11) to decouple the vacuum expectation values of the two Higgs doublets. We only note, that this basis change isolates the gauge boson mass terms (and the associated Goldstone contributions in more general gauges than considered here). We choose

$$\begin{pmatrix} \varphi_1' \\ \varphi_2' \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta\, e^{-i\zeta} \\ -\sin\beta\, e^{i\zeta} & \cos\beta \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \tag{2.106}$$

with β fulfilling $v_1 \sin\beta = v_2 \cos\beta$ ($\tan\beta = v_2/v_1$ for $v_1 \neq 0$), and thus arrange that

$$\langle \varphi_1' \rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}, \qquad \langle \varphi_2' \rangle = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.107}$$

for the vacuum expectation values in the new basis, where

$$v = \sqrt{2(v_1^2 + v_2^2)} > 0 \tag{2.108}$$

is the usual Higgs-field vacuum expectation value, $v \approx 246\,\mathrm{GeV}$ (see for instance [64]). At the level of the orbit variables the flavour basis change (2.106) means

$$\langle K_0' \rangle = \langle K_0 \rangle \tag{2.109a}$$

$$\langle \boldsymbol{K}' \rangle = \begin{pmatrix} c_{2\beta} + 2s_\beta^2 s_\zeta^2 & -s_\beta^2 s_{2\zeta} & -s_{2\beta} c_\zeta \\ -s_\beta^2 s_{2\zeta} & c_{2\zeta} + 2c_\beta^2 s_\zeta^2 & -s_{2\beta} s_\zeta \\ s_{2\beta} c_\zeta & s_{2\beta} s_\zeta & c_{2\beta} \end{pmatrix} \langle \boldsymbol{K} \rangle \tag{2.109b}$$

with $s_\beta \equiv \sin\beta$, $c_\beta \equiv \cos\beta$ etc., such that in the new basis

$$\left\langle \tilde{\boldsymbol{K}}' \right\rangle = \begin{pmatrix} \frac{1}{2}v^2 \\ 0 \\ 0 \\ \frac{1}{2}v^2 \end{pmatrix}. \tag{2.110}$$

Figure 2.2 illustrates the flavour basis change (2.109) which decouples the vacuum expectation values.

We choose a unitary gauge such that in addition to (2.107) the fields satisfy

$$\varphi_1'^{+}(x) = 0\,, \tag{2.111}$$

$$\operatorname{Im} \varphi_1'^{0}(x) = 0\,, \tag{2.112}$$

$$\operatorname{Re} \varphi_1'^{0}(x) \geq 0\,. \tag{2.113}$$

We introduce the shifted Higgs fields as the deviations of the original doublet components from their vacuum expectation values. The actual physical Higgs bosons are the mass eigenstate admixtures of these shifted fields. We define the three neutral shifted Higgs fields as

$$\rho'(x) := \sqrt{2}\, \operatorname{Re} \varphi_1'^{0}(x) - v\,, \tag{2.114}$$

$$h'(x) := \sqrt{2}\, \operatorname{Re} \varphi_2'^{0}(x)\,, \tag{2.115}$$

$$h''(x) := \sqrt{2}\, \operatorname{Im} \varphi_2'^{0}(x)\,. \tag{2.116}$$

and the charged fields as

$$H^+(x) := \varphi_2'^{+}(x)\,, \tag{2.117}$$

$$H^-(x) := \left(H^+(x)\right)^*\,. \tag{2.118}$$

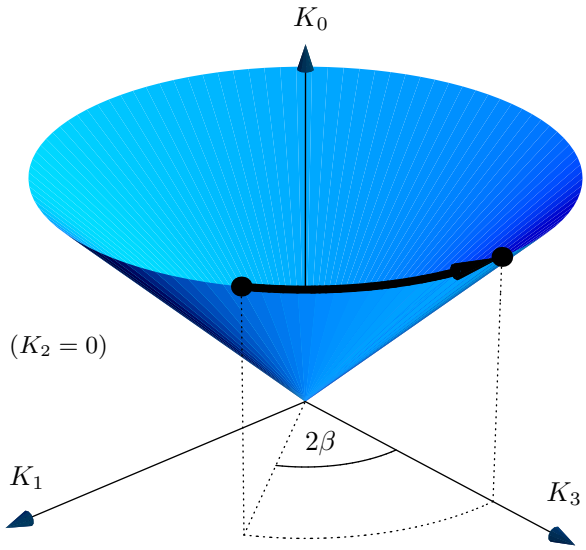

Figure 2.3: The vacuum expectation values are decoupled by a flavour basis change which is just a real rotation of the orbit variables. The angle β satisfies as usual $\tan\beta = v_2/v_1$.

This leads to

$$\varphi_1'(x) = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + \rho'(x)) \end{pmatrix} , \tag{2.119a}$$

$$\varphi_2'(x) = \begin{pmatrix} H^+(x) \\ \frac{1}{\sqrt{2}}(h'(x) + i\,h''(x)) \end{pmatrix} . \tag{2.119b}$$

It is convenient to decompose $\tilde{\boldsymbol{K}}'$ according to the power of the *physical* fields they contain

$$\tilde{\boldsymbol{K}}' = \tilde{\boldsymbol{K}}'_{\{0\}} + \tilde{\boldsymbol{K}}'_{\{1\}} + \tilde{\boldsymbol{K}}'_{\{2\}} \tag{2.120}$$

with

$$\tilde{\boldsymbol{K}}'_{\{0\}} = \frac{v^2}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} , \qquad \tilde{\boldsymbol{K}}'_{\{1\}} = v\begin{pmatrix} \rho' \\ h' \\ h'' \\ \rho' \end{pmatrix} , \qquad \tilde{\boldsymbol{K}}'_{\{2\}} = \frac{1}{2}\begin{pmatrix} \rho'^2 + h'^2 + h''^2 + 2H^+H^- \\ 2\rho' h' \\ 2\rho' h'' \\ \rho'^2 - h'^2 - h''^2 - 2H^+H^- \end{pmatrix} , \tag{2.121}$$

and, of course, $\tilde{\boldsymbol{K}}'_{\{0\}} \equiv \langle \tilde{\boldsymbol{K}}' \rangle$. By u_v we denote the Lagrange multiplier corresponding to the vacuum solution (2.110). From (2.82a) we have

$$\tilde{E}'\tilde{\boldsymbol{K}}'_{\{0\}} = u_v\,\tilde{g}\tilde{\boldsymbol{K}}'_{\{0\}} - \frac{1}{2}\tilde{\boldsymbol{\xi}}' . \tag{2.122}$$

From the explicit expressions (2.121) we further get

$$\tilde{\boldsymbol{K}}'^{\mathrm{T}}_{\{0\}}\,\tilde{g}\,\tilde{\boldsymbol{K}}'_{\{0\}} = 0 , \quad \tilde{\boldsymbol{K}}'^{\mathrm{T}}_{\{0\}}\,\tilde{g}\,\tilde{\boldsymbol{K}}'_{\{1\}} = 0 . \tag{2.123}$$

Using (2.120) to (2.123) we obtain for the potential (2.26)

$$V = V_{\{0\}} + V_{\{2\}} + V_{\{3\}} + V_{\{4\}}\,, \tag{2.124}$$

where $V_{\{k\}}$ are the terms of k^{th} order in the physical Higgs fields

$$V_{\{0\}} = (\xi'_0 + \xi'_3)\, v^2/4\,, \tag{2.125}$$

$$V_{\{2\}} = \tilde{\boldsymbol{K}}'^{\mathrm{T}}_{\{1\}}\, \tilde{E}'\, \tilde{\boldsymbol{K}}'_{\{1\}} + 2\, u_v\, \tilde{\boldsymbol{K}}'^{\mathrm{T}}_{\{0\}}\, \tilde{g}\, \tilde{\boldsymbol{K}}'_{\{2\}}\,, \tag{2.126}$$

$$V_{\{3\}} = 2\, \tilde{\boldsymbol{K}}'^{\mathrm{T}}_{\{1\}}\, \tilde{E}'\, \tilde{\boldsymbol{K}}'_{\{2\}}\,, \tag{2.127}$$

$$V_{\{4\}} = \tilde{\boldsymbol{K}}'^{\mathrm{T}}_{\{2\}}\, \tilde{E}'\, \tilde{\boldsymbol{K}}'_{\{2\}}\,. \tag{2.128}$$

The second order terms (2.126) determine the masses of the physical Higgs fields:

$$V_{\{2\}} = \frac{1}{2}(h', h'', \rho')\mathcal{M}^2_{\text{neutral}} \begin{pmatrix} h' \\ h'' \\ \rho' \end{pmatrix} + m^2_{H^\pm} H^+ H^- \tag{2.129}$$

with

$$\mathcal{M}^2_{\text{neutral}} = 2 \begin{pmatrix} v^2\,(\eta'_{11} + u_v) & v^2\,\eta'_{12} & -\xi'_1 \\ v^2\,\eta'_{12} & v^2\,(\eta'_{22} + u_v) & -\xi'_2 \\ -\xi'_1 & -\xi'_2 & -\xi'_3 - \xi'_0 \end{pmatrix}, \tag{2.130}$$

$$m^2_{H^\pm} = 2\, u_v\, v^2\,. \tag{2.131}$$

Note that the condition $u_v \geq 0$ corresponds to the non-negativity of the charged Higgs mass squared at the tree-level. Using the stationarity condition in the present flavour basis to eliminate $\tilde{\boldsymbol{\xi}}'$ we find for the mass parameters

$$\mathcal{M}^2_{\text{neutral}} = 2v^2 \begin{pmatrix} \eta'_{11} + u_v & \eta'_{12} & \eta'_{13} + \eta'_1 \\ \eta'_{12} & \eta'_{22} + u_v & \eta'_{23} + \eta'_2 \\ \eta'_{13} + \eta'_1 & \eta'_{23} + \eta'_2 & \eta'_{33} + 2\eta'_3 + \eta'_{00} \end{pmatrix}, \tag{2.132}$$

$$m^2_{H^\pm} = 2v^2\, u_v\,. \tag{2.133}$$

Note the simple structure of these mass terms. These, determining the terms quadratic in the physical fields, are proportional to simple linear combinations of the quartic parameters for the original fields and the Lagrange multiplier, with the overall scale set by v. Generically the mass terms (2.129) contain 7 real parameters. From (2.132) and (2.133) we see that all 7 parameters are in general independent in this model.

While the charged Higgs fields $H^\pm$ are already mass eigenstates, the physical neutral Higgs bosons are obtained after diagonalisation of the symmetric mass matrix $\mathcal{M}^2_{\text{neutral}}$. If CP is conserved by both the potential and the vacuum, scalar and pseudo-scalar Higgs bosons may be defined with a conserved CP quantum number. We shall discuss CP

invariances in detail in section 2.5, but already note the following. In case of standard CP conservation of both the potential and the vacuum, $\mathcal{M}^2_{\text{neutral}}$ becomes diagonal in the 2–direction. Then h'' is a CP-odd, i.e. pseudo-scalar, mass eigenstate denoted by $A := h''$, while ρ', h' are CP-even fields, which mix to two mass eigenstates denoted by h, H. In case h, H are not degenerate, h is the lighter of the two states by convention. Explicit examples for physical Higgs bosons and masses can be found in sections 2.6 and 2.7.

We close this section with a remark concerning the choice of input parameters. In general the vacuum solution (2.105) is a complicated function of the original parameters, which may not even be calculable analytically. It is therefore tempting to use a parameterisation of the vacuum to reparameterise the potential itself via the stationarity conditions, as already discussed in subsection 2.3.1. A particularly simple starting point is the parameterisation of the vacuum in the general basis (2.105). One may now replace $\tilde{\boldsymbol{\xi}}$ via (2.72a) with $\tilde{\boldsymbol{K}}$ given in (2.105) and v_1 eliminated by means of (2.108):

$$\tilde{\boldsymbol{\xi}} = \tilde{\boldsymbol{\xi}}(v, v_2, \zeta, m_{H^\pm}, \eta_{00}, \boldsymbol{\eta}, E)\,. \tag{2.134}$$

With this the potential can be reparameterised in terms of $v, v_2, \zeta, m_{H^\pm}, \eta_{00}, \boldsymbol{\eta}, E$. With this set of independent input parameters, v can be adjusted to the required value (2.108), and relations involving the vacuum solution can be evaluated directly in terms of input parameters. Note, that this parameterisation (2.134) is possible for all potentials having a non-zero stationary point $\tilde{\boldsymbol{K}}$ on the light cone. A potential not having such a point can not have the required EWSB behaviour. After the substitution (2.134) the four-vector $\tilde{\boldsymbol{K}}$ in (2.192) corresponds by construction to a stationary point of V. Thus, the parameterisation (2.134) is possible for all potentials with a stationary point at the wanted place (2.192). But for any specific values of the new parameters it remains to be checked whether $\tilde{\boldsymbol{K}}$ in (2.192) is indeed the global minimum of a stable potential V. This typically requires to perform the complete analysis presented in the previous section. Note that in the orbit variable approach this change of parameters is even possible for the cases where the phase ζ or one of v_1, v_2 vanishes. We consider this to be a slight advantage with respect to the similar change of parameters within the traditional field based approach, where case distinctions are required.

2.5 Generalised CP symmetries

2.5.1 CP transformations

Standard CP transformation

The standard CP transformation of the gauge, Higgs and fermion fields reads (see for instance [64])

$$\begin{aligned}
\mathrm{CP}_s: \qquad W^\mu(x) &\longrightarrow -W_\mu^{\mathrm{T}}(x')\,, && && \text{(2.135a)}\\
B^\mu(x) &\longrightarrow -B_\mu(x')\,, && && \text{(2.135b)}\\
\varphi_i(x) &\longrightarrow \varphi_i^*(x')\,, && i=1,2\,, && \text{(2.135c)}\\
\psi_i(x) &\longrightarrow \gamma^0\, S(C)\, \overline{\psi}_i{}^{\mathrm{T}}(x')\,, && i=1,\ldots,N_f\,. && \text{(2.135d)}
\end{aligned}$$

Here we have

$$(x^\mu) = \begin{pmatrix} x^0 \\ \boldsymbol{x} \end{pmatrix}, \quad (x'^\mu) = \begin{pmatrix} x^0 \\ -\boldsymbol{x} \end{pmatrix}, \tag{2.136}$$

γ^0 and $S(C) := i\gamma^2\gamma^0$ are the usual Dirac matrices for the parity and charge conjugation transformations, respectively (see for instance chapter 4 of [64]), and

$$W^\mu(x) = W^{\mu a}(x)\frac{\sigma_a}{2} \tag{2.137}$$

is the matrix of the W-potentials. Of course, a discussion of this CP transformation makes only sense once we have already chosen a particular basis for the two Higgs doublets since basis transformations (2.11) change (2.135). Such a particular choice of basis is, indeed, in general required when the Yukawa terms $\mathscr{L}_{\mathrm{Yuk}}$ are taken into consideration. In the MSSM, for instance, one Higgs doublet couples to the up-type fermions, one to the down type fermions. This clearly singles out a special basis. Therefore, we have denoted the CP transformations in (2.135) by CP_s for *standard* and *special*.

From the definition of our matrix $\underline{K}$ and of the four real coefficients K_0 and K_a it is obvious that the CP_s transformation (2.135) corresponds to

$$\begin{aligned}
\mathrm{CP}_s: \qquad K_0(x) &\longrightarrow K_0(x')\,,\\
\begin{pmatrix} K_1(x) \\ K_2(x) \\ K_3(x) \end{pmatrix} &\longrightarrow \begin{pmatrix} K_1(x') \\ -K_2(x') \\ K_3(x') \end{pmatrix}.
\end{aligned} \tag{2.138}$$

That is, the vector $\boldsymbol{K}(x)$ is subjected to a reflection on the $1-3$ plane and a change of argument $x \longrightarrow x'$,

$$\mathrm{CP}_s: \qquad \boldsymbol{K}(x) \longrightarrow R_2 \boldsymbol{K}(x')\,, \tag{2.139}$$

where

$$R_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.140}$$

Generalised CP transformations of the fields

We shall in this theses also consider *generalised* CP transformations of the Higgs fields [116] defined by

$$\begin{aligned}
\mathrm{CP}_g: \qquad W^\mu(x) &\longrightarrow -W_\mu^{\mathrm{T}}(x')\,, && && \text{(2.141a)}\\
B^\mu(x) &\longrightarrow -B_\mu(x')\,, && && \text{(2.141b)}\\
\varphi_i(x) &\longrightarrow U_{ij}^{\varphi}\ \varphi_j^*(x')\,, && i=1,2\,, && \text{(2.141c)}\\
\psi_i(x) &\longrightarrow U_{ij}^{\psi}\ \gamma^0\, S(C)\, \overline{\psi}_j{}^{\mathrm{T}}(x')\,, && i=1,\ldots,N_f && \text{(2.141d)}
\end{aligned}$$

with $U^\varphi = \left(U_{ij}^\varphi\right) \in U(2)$ and $U^\psi = \left(U_{ij}^\psi\right) \in U(N_f)$. That is, the usual CP transformation of the Higgs and the fermion fields are supplemented by a flavour basis change. Explicitly, we allow independent mixings for fermions with different quantum numbers:

$$\begin{aligned}
\mathrm{CP}_g: \qquad L_i^L(x) &\longrightarrow U_{ij}^L\ \gamma^0\, S(C)\, \overline{L}_j^{L\,\mathrm{T}}(x')\,, && \text{(2.142a)}\\
l_i^R(x) &\longrightarrow U_{ij}^l\ \gamma^0\, S(C)\, \bar{l}_j^{R\,\mathrm{T}}(x')\,, && \text{(2.142b)}\\
Q^L(x) &\longrightarrow U_{ij}^Q\ \gamma^0\, S(C)\, \overline{Q}_j^{L\,\mathrm{T}}(x')\,, && \text{(2.142c)}\\
d_i'^R(x) &\longrightarrow U_{ij}^{d'}\ \gamma^0\, S(C)\, \overline{d}_j'^{R\,\mathrm{T}}(x')\,, && \text{(2.142d)}\\
u_i^R(x) &\longrightarrow U_{ij}^u\ \gamma^0\, S(C)\, \overline{u}_j^{R\,\mathrm{T}}(x')\,. && \text{(2.142e)}
\end{aligned}$$

Although we presently discuss THDMs, parts of the following discussions are formulated at a rather generic level. This allows to apply the results to cases with different numbers not only of the fermions but also of the Higgs doublets.

Now we shall require that a CP_g transformation applied twice gives the original fields up to a symmetry transformation of the Lagrangian, see also [117]. We find from (2.141):

$$\begin{aligned}
\mathrm{CP}_g \circ \mathrm{CP}_g: \qquad W^\mu(x) &\longrightarrow W^\mu(x)\,, && && \text{(2.143a)}\\
B^\mu(x) &\longrightarrow B^\mu(x)\,, && && \text{(2.143b)}\\
\varphi_i(x) &\longrightarrow (U^\varphi\, U^{\varphi\,*})_{ij}\, \varphi_j(x)\,, && i=1,2\,, && \text{(2.143c)}\\
\psi_i(x) &\longrightarrow -(U^\psi\, U^{\psi\,*})_{ij}\, \psi_j(x)\,, && i=1,\ldots,N_f\,. && \text{(2.143d)}
\end{aligned}$$

Of course, the complete set of symmetries of the Lagrangian depends on the specific THDM one considers. We shall consider here completely general potentials without insisting on any further symmetries. Therefore we require that the twofold application of a CP_g transformation reproduces the original fields up to a phase. This implies that any of the mixing matrices U in (2.141) satisfy

$$U\, U^* = e^{i\varkappa} \mathbb{1} \tag{2.144}$$

with real $\varkappa$. It is interesting to note that in (2.144) only certain phases $\varkappa$ are possible depending on the number of flavours involved. Note that the global phase of the matrix U

drops out in (2.144) and a non-trivial phase $\varkappa$ can only be generated through a non-trivial matrix structure. Actually, we see from $U\,U^\dagger = \mathbb{1}$ that (2.144) implies

$$U = e^{i\varkappa} U^{\mathrm{T}} \tag{2.145}$$
$$= e^{2i\varkappa} U\,, \tag{2.146}$$

from which we read off that only the cases $U^{\mathrm{T}} = U$ with $U\,U^* = \mathbb{1}$ or $U^{\mathrm{T}} = -U$ with $U\,U^* = -\mathbb{1}$ are possible. We therefore find the following.

Theorem 2.5.1. *Generalised CP symmetries* (2.141) *may be defined which involve a unitary flavour mixing in the respective sector in addition to the standard CP transformation. Requiring a twofold CP_g symmetry to reproduce the original fields up to phase implies $U\,U^* = e^{i\varkappa}\mathbb{1}$ with $\varkappa \in \mathbb{R}$ for the flavour mixing matrices $U = U^\varphi, U^\psi$. A unitary matrix $U = U(n)$ with this property can be realised only in the following ways:*

$$U \in U(1) \qquad \Rightarrow \qquad U^{\mathrm{T}} = U \qquad \textit{with} \quad U\,U^* = \mathbb{1}, \tag{2.147}$$

$$U \in U(n)\,, \quad n \geq 2 \qquad \Rightarrow \qquad U^{\mathrm{T}} = U \qquad \textit{with} \quad U\,U^* = \mathbb{1}\,,$$
$$\textit{or} \quad U^{\mathrm{T}} = -U \qquad \textit{with} \quad U\,U^* = -\mathbb{1}\,. \tag{2.148}$$

To our knowledge, cases with $U\,U^* \neq \mathbb{1}$ have not been discussed in the literature before ref. [89]. The simplest realisation of the case $U\,U^* = -\mathbb{1}$ is possible for 2 flavours. In this case U must be of the form

$$U = e^{i\delta}\epsilon\,, \qquad \epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \delta \in \mathbb{R}\,. \tag{2.149}$$

As a preparation for the following discussion we consider a CP_g transformation in a new basis of the Higgs and the fermion fields, defined by

$$\varphi_i'(x) = W_{ij}\varphi_j(x)\,, \qquad W \in U(2)\,, \tag{2.150a}$$
$$\psi_i'(x) = V_{ij}\psi_i(x)\,, \qquad V \in U(N_f)\,. \tag{2.150b}$$

We get

$$\mathrm{CP}_g: \quad \varphi_i'(x) \longrightarrow (W\,U^\varphi\,W^{\mathrm{T}})_{ij}\varphi_j'^*(x')\,, \qquad i = 1,2\,, \tag{2.151a}$$
$$\psi_i'(x) \longrightarrow (V\,U^\psi\,V^{\mathrm{T}})_{ij}\,\gamma^0\,S(C)\,\overline{\psi}_j'^{\,\mathrm{T}}(x')\,, \quad i = 1,\ldots,N_f\,. \tag{2.151b}$$

From theorem 2.5.1 we see that a mixing U in a generalised CP transformation with $U\,U^* = \mathbb{1}$ can be realised for any number of flavours. In [75] it was shown for the Higgs sector that for any generalised CP transformation of this type there exists a Higgs flavour basis in which it is just the standard CP transformation. We briefly reproduce the proof here and extend the statement. If a unitary matrix U satisfies $U\,U^* = \mathbb{1}$ it must be symmetric, $U = U^{\mathrm{T}}$. As a consequence of Takagi's factorisation theorem, see e.g. [118], it can then be written as $U = X^{\mathrm{T}}\,X$ where X is a unitary matrix. We consider the action

of a generalised CP transformation on the Higgs fields, where the mixing matrix U^φ in (2.141c) satisfies $U^\varphi U^{\varphi *} = \mathbb{1}$. Choosing the Higgs basis $\varphi'_i = X^*_{ij}\varphi_j$ we see from (2.151a) with $W = X^*$

$$\mathrm{CP}_g: \quad \varphi'_i(x) \longrightarrow (X^* U^\varphi X^\dagger)\varphi'^*_j(x') = \varphi'^*_i(x')\,, \tag{2.152}$$

that is, the generalised CP transformation reduces to the standard CP transformation. On the other hand, let U^φ define a generalised CP transformation which is reducible to the standard CP transformation by a basis change W. From (2.151a) we see that $W\,U^\varphi\,W^{\mathrm{T}} = \mathbb{1}$ and therefore $U^\varphi\,U^{\varphi *} = W^\dagger\,W^*\,W^{\mathrm{T}}\,W^\dagger = \mathbb{1}$. For the fermion sector the arguments run along the same lines. We therefore find the following.

Theorem 2.5.2. *For a generalised CP_g transformation of the Higgs or the fermion sector,* (2.141c) *or* (2.141d)*, a flavour basis exists in which it equals the standard CP_s transformation if and only if*

$$U^{\mathrm{T}} = U \qquad \textit{with} \quad U\,U^* = \mathbb{1} \tag{2.153}$$

holds for the corresponding unitary mixing matrix $U = U^\varphi$ or $U = U^\psi$.

This theorem states in particular, that a CP_g transformation with $U\,U^* = -\mathbb{1}$, such as in (2.149), is not equivalent to standard CP transformation by a change of basis.

We would like to comment on the "concatenation" of CP_g transformations. Applying three different CP_g transformations CP^a_g, CP^b_g, CP^c_g one after the other gives

$$\mathrm{CP}^c_g \circ \mathrm{CP}^b_g \circ \mathrm{CP}^a_g\,, \tag{2.154}$$

which is again a transformation of the form (2.141) with new mixing matrices. Suppose we consider either the Higgses or a specific fermion type, and let $U_a,\ U_b,\ U_c$ be the corresponding mixing matrices in CP^a_g, CP^b_g, CP^c_g respectively. Rewriting the threefold CP_g product according to (2.141) the resulting mixing will be $U_{tot} = U_c\,U_b^*\,U_a$ and thus unitary. Even if each of $U_a,\ U_b,\ U_c$ is of type (2.144) the matrix U_{tot} must not necessarily satisfy (2.144) (take for instance $U_a = (\sigma_1 + \mathbb{1})/\sqrt{2},\ U_b = \mathbb{1},\ U_c = \epsilon$ as a counter example with 2×2 matrices). However, in other cases the $U_a,\ U_b,\ U_c$ and also the U_{tot} of all sectors are of type (2.144) and therefore interpretable in a similar way. A particular case of such a product will be constructed in section 2.7. There we show that a CP_g symmetry of the new type $U\,U^* = -\mathbb{1}$ can actually be written as a threefold product of "conventional" CP_g symmetries with $U\,U^* = \mathbb{1}$.

Concerning the normalisation of the operator $(\mathrm{CP}_g)^2$ we remark the following. The operator $(\mathrm{CP}_g)^2$ reproduces the original Higgs fields up to a possible phase factor, which can in all cases be interpreted as a hypercharge transformation:

$$\mathrm{CP}_g \circ \mathrm{CP}_g: \quad \varphi_i(x) \longrightarrow \begin{cases} \varphi_i(x) & \text{if } U^\varphi\,U^{\varphi *} = 1 \\ \exp(i6\pi Y)\varphi_i(x) & \text{if } U^\varphi\,U^{\varphi *} = -1 \end{cases} \tag{2.155}$$

If only the "standard" mixing types with $\exp(i\varkappa) = 1$ occur in a CP_g transformation, the operator $(\mathrm{CP}_g)^2$ is normalised for all fields in the same way as $(\mathrm{CP}_s)^2$, that is, for

the Higgs, the fermion and the gauge boson fields. For certain combinations involving the "non-standard" mixing types, the operator $\exp(i6\pi Y) \circ (\mathrm{CP}_g)^2$ is normalised for all fields in the usual way, that is as for $(\mathrm{CP}_s)^2$. An example for such a combination is the case $\exp(i\varkappa^\varphi) = \exp(i\varkappa^L) = \exp(i\varkappa^Q) = -1$ and $\exp(i\varkappa^l) = \exp(i\varkappa^d) = \exp(i\varkappa^u) = 1$. In the general case additional unobservable phase factors may occur for the fermions.

CP transformations of the orbit variables

We discuss now the generalised CP transformations at the level of the Higgs orbit variables and rederive some of the statements found above at the level of the Higgs fields. Using the orbit variable will allow to give a simple geometrical interpretation of the CP_g symmetries. For the standard CP transformation we get from (2.135c) for the orbit variables

$$\begin{aligned} \mathrm{CP}_g: \qquad & K_0(x) \longrightarrow K_0(x')\,, \\ & \boldsymbol{K}(x) \longrightarrow R_2 \boldsymbol{K}(x')\,. \end{aligned} \tag{2.156}$$

The generalised CP transformations of the orbit variables follow from (2.141c) and (2.11) according to

$$\begin{aligned} \mathrm{CP}_g: \qquad & K_0(x) \longrightarrow K_0(x')\,, \\ & \boldsymbol{K}(x) \longrightarrow R(U^\varphi) R_2 \boldsymbol{K}(x') \end{aligned} \tag{2.157}$$

with $R(U^\varphi) \in SO(3)$ obtained from (2.11) with U replaced by U^φ. That is, CP_g induces an *improper* rotation $\bar{R}^\varphi$ of the vector $\boldsymbol{K}$ in addition to the change of argument $x \longrightarrow x'$:

$$\mathrm{CP}_g: \qquad \boldsymbol{K}(x) \longrightarrow \bar{R}^\varphi \boldsymbol{K}(x')\,, \tag{2.158}$$

where

$$\begin{aligned} & \bar{R}^\varphi = R(U^\varphi) R_2\,, \\ & \bar{R}^\varphi \bar{R}^{\varphi\,\mathrm{T}} = \mathbb{1}_3\,, \\ & \det \bar{R}^\varphi = \det\left(R(U^\varphi) R_2\right) = -1\,. \end{aligned} \tag{2.159}$$

From the results of subsection 2.1.2 it is clear that to any improper rotation $\bar{R}^\varphi$ there is a $U^\varphi \in U(2)$ which, inserted in (2.141), gives (2.158) and (2.159).

We shall study now the effect of a basis change (2.11) on $\bar{R}_\varphi$. For this it is convenient to work with the matrix $\phi(x)$ (2.16). Let the new basis fields be $\varphi_1'(x)$, $\varphi_2'(x)$ and the corresponding matrix

$$\phi'(x) = \begin{pmatrix} \varphi_1'^+(x) & \varphi_1'^0(x) \\ \varphi_2'^+(x) & \varphi_2'^0(x) \end{pmatrix} = U\phi(x) \tag{2.160}$$

with $U \in U(2)$. The CP_g transformation (2.141) reads

$$\mathrm{CP}_g: \qquad \phi(x) \longrightarrow U^\varphi \phi^*(x')\,. \tag{2.161}$$

This implies

$$\mathrm{CP}_g : \qquad \phi'(x) \longrightarrow U^{\varphi\prime}\phi'^*(x')\,,$$

with

$$U^{\varphi\prime} = UU^{\varphi}U^{*-1}\,. \tag{2.162}$$

The transformation of $K_0'(x)$ and $\boldsymbol{K}'(x)$ in the new basis is

$$\begin{aligned} \mathrm{CP}_g : \qquad & K_0'(x) \longrightarrow K_0'(x')\,, \\ & \boldsymbol{K}'(x) \longrightarrow \bar{R}^{\varphi\prime}\boldsymbol{K}'(x')\,, \end{aligned} \tag{2.163}$$

with

$$\bar{R}^{\varphi\prime} = R(U^{\varphi\prime})R_2 = R(U)\bar{R}^{\varphi}R^{\mathrm{T}}(U)\,. \tag{2.164}$$

Here $R(U) \in SO(3)$ is the rotation matrix obtained from U according to (2.11). Thus, a basis change induces an orthogonal transformation of the improper rotation matrix $\bar{R}^{\varphi}$.

Now we shall require that a CP_g transformation applied twice gives the original fields up to a phase factor. For the orbit variables we find from (2.158):

$$\begin{aligned} \mathrm{CP}_g \circ \mathrm{CP}_g : \qquad & K_0(x) \longrightarrow K_0(x)\,, \\ & \boldsymbol{K}(x) \longrightarrow (\bar{R}^{\varphi})^2\boldsymbol{K}(x)\,. \end{aligned} \tag{2.165}$$

Requiring that $\mathrm{CP}_g \circ \mathrm{CP}_g$ gives the unit transformation for the orbit variables leads to the condition

$$\bar{R}^{\varphi}\bar{R}^{\varphi} = \mathbb{1}_3\,. \tag{2.166}$$

But we also have $\bar{R}^{\varphi}\bar{R}^{\varphi\mathrm{T}} = \mathbb{1}_3$, see (2.159). The requirement (2.166) thus means that $\bar{R}^{\varphi}$ is symmetric

$$\bar{R}^{\varphi\mathrm{T}} = \bar{R}^{\varphi}\,. \tag{2.167}$$

As a real symmetric matrix it can be diagonalised by an orthogonal matrix $R(U)$. That is, we can change the basis of the Higgs fields according to (2.160) and achieve

$$\bar{R}^{\varphi\prime} = R(U)\,\bar{R}^{\varphi}\,R^{\mathrm{T}}(U) = \text{diagonal matrix}\,. \tag{2.168}$$

Since $\bar{R}^{\varphi\prime}$ is an improper rotation it satisfies $\bar{R}^{\varphi\prime}\bar{R}^{\varphi\prime\mathrm{T}} = \mathbb{1}_3$ and $\det\bar{R}^{\varphi\prime} = -1$. Thus, we have only the possibilities $\bar{R}^{\varphi\prime} = R_1$ or R_2 or R_3 or $-\mathbb{1}_3$. Here

$$\begin{aligned} R_1 &:= \mathrm{diag}(-1,\ \ 1,\ \ 1)\,, \\ R_2 &:= \mathrm{diag}(\ \ 1,-1,\ \ 1)\,, \\ R_3 &:= \mathrm{diag}(\ \ 1,\ \ 1,-1)\,. \end{aligned} \tag{2.169}$$

The cases $\bar{R}^{\varphi\prime} = R_j$, $j = 1,2,3$ are equivalent by a basis change. In constrast, a CP_g transformation with $\bar{R}^{\varphi\prime} = -\mathbb{1}_3$ is not equivalent by a basis change to any of these cases. In fact, $\bar{R}^{\varphi\prime} = \bar{R}^{\varphi} = -\mathbb{1}_3$ is invariant under basis changes, since the latter are orthogonal transformations. Thus we find the following.

Theorem 2.5.3. *Generalised CP symmetries* (2.158) *whose twofold application reproduce the original orbit variables,* (2.166)*, are improper rotations* $\bar{R}^{\varphi}$ *in* $\boldsymbol{K}$*-space. Such an improper rotation is either*

$$(i) \quad \bar{R}^{\varphi} = -\mathbb{1}_3 \,, \tag{2.170}$$

that is, a basis independent point reflection at the origin*, or orthogonally equivalent to the standard CP transformation described by the reflection* R_2

$$(ii) \quad \bar{R}^{\varphi} = R^{\mathrm{T}}(U) R_2 R(U) \,, \tag{2.171}$$

that is, a reflection on a plane*.*

By a straight-forward calculation we find that the CP_g type (i) transformation corresponds exactly to the Higgs field mixing (2.149) with $U^{\varphi} U^{\varphi *} = -\mathbb{1}$. On the other hand we see from theorems 2.5.2 and 2.5.3 that CP_g type (ii) symmetries correspond to the CP_g symmetries with $U^{\varphi} U^{\varphi *} = \mathbb{1}$ for the Higgs fields, which are equivalent to CP_s up to a basis change. We note that in terms of the orbit variables both types of symmetries arise equally natural in terms of (2.166).

2.5.2 CP symmetries of the potential

The kinetic term in the Lagrangian (2.3) is invariant under CP_s and CP_g as defined in (2.135) and (2.141). The CP symmetries of the Lagrangian are therefore determined by the scalar Higgs potential and the Yukawa terms. We consider first the symmetries of the scalar Higgs potential. Asking if the potential V (2.26) allows a CP_g symmetry is the same as asking if it is invariant under some improper rotation (2.158) of the $\boldsymbol{K}$-vectors. That is, we have invariance under a CP_g transformation if the parameters of V (2.26) satisfy

$$\bar{R}^{\varphi} \boldsymbol{\xi} = \boldsymbol{\xi} \,, \quad \bar{R}^{\varphi} \boldsymbol{\eta} = \boldsymbol{\eta} \,, \quad \bar{R}^{\varphi} E \bar{R}^{\varphi \mathrm{T}} = E \tag{2.172}$$

for some improper rotation matrix $\bar{R}^{\varphi}$.

Standard CP symmetry

The potential V (2.26) allows CP_s as a symmetry if and only if

$$R_2 \boldsymbol{\xi} = \boldsymbol{\xi} \,, \quad R_2 \boldsymbol{\eta} = \boldsymbol{\eta} \,, \quad R_2 E R_2^{\mathrm{T}} = E \,, \tag{2.173}$$

that is, it contains no terms linear in K_2. Thus, we have the following theorem.

Theorem 2.5.4. *The Higgs Lagrangian (2.3) with the general potential (2.26) is invariant under the* CP_s *transformation* (2.135) *if and only if*

$$\xi_2 = 0 \,, \quad \eta_2 = 0 \,, \quad \eta_{12} = \eta_{23} = 0 \,. \tag{2.174}$$

CP_g type (i) symmetry

For the case (i), $\bar{R}^\varphi$ as in (2.170), the invariance conditions (2.172) for the potential parameters give us the following theorem.

Theorem 2.5.5. *The Higgs potential* (2.26) *has the* CP_g *symmetry* (2.158) *of type* (i), *where* $\bar{R}^\varphi = -\mathbb{1}_3$ *(see (2.170)), if and only if*

$$\boldsymbol{\xi} = 0 \quad \textit{and} \quad \boldsymbol{\eta} = 0\,. \tag{2.175}$$

CP_g type (ii) symmetries

For the case (ii), $\bar{R}^\varphi$ as in (2.171), we find that the original CP_g transformation (2.141) is equal to the standard CP_s transformation (2.135) for the Higgs fields after a suitable change of basis, see (2.11) and (2.160):

$$\mathrm{CP}_g: \qquad \varphi_i'(x) \longrightarrow \varphi_i'^*(x')\,, \quad i = 1,2\,. \tag{2.176}$$

Using now the results of subsection 2.5.1 we find that the THDM potential (2.26) will be invariant under a CP_g transformation of type (ii) if and only if we can find a basis transformation (2.31) eliminating all odd powers of K_2. That is, there must exist some $R(U) \in SO(3)$ such that

$$\begin{aligned}
\boldsymbol{\xi}' &= R(U)\,\boldsymbol{\xi} = \begin{pmatrix} \cdot \\ 0 \\ \cdot \end{pmatrix}, \\
\boldsymbol{\eta}' &= R(U)\,\boldsymbol{\eta} = \begin{pmatrix} \cdot \\ 0 \\ \cdot \end{pmatrix}, \\
E' &= R(U)\,E\,R^{\mathrm{T}}(U) = \begin{pmatrix} \cdot & 0 & \cdot \\ 0 & \cdot & 0 \\ \cdot & 0 & \cdot \end{pmatrix},
\end{aligned} \tag{2.177}$$

where the dots represent arbitrary entries. Note that the central entry of E', that is E'_{22}, can be different from zero since it corresponds to a quadratic term in K_2. Obviously, the first two conditions correspond to a rotation of the vector cross product $\boldsymbol{\xi} \times \boldsymbol{\eta}$ into the 2–direction which is always achievable by suitable rotations around the 1- and the 3-axis. It is advantageous to formulate the conditions (2.177) in a way independent of the chosen basis, such that no rotations of the original parameters have to be performed.

In the following we shall show that the conditions (2.177) are equivalent to a simple set of equations. We formulate this result as a theorem.

Theorem 2.5.6. *The THDM potential* V *(2.26) is invariant under a* CP_g *transformation (2.158) of type* (ii) *(see (2.171)) if and only if the following basis independent set of*

equations holds:

$$(\boldsymbol{\xi} \times \boldsymbol{\eta})^{\mathrm{T}}\, E\boldsymbol{\xi} = 0\,, \tag{2.178a}$$

$$(\boldsymbol{\xi} \times \boldsymbol{\eta})^{\mathrm{T}}\, E\boldsymbol{\eta} = 0\,, \tag{2.178b}$$

$$(\boldsymbol{\xi} \times (E\boldsymbol{\xi}))^{\mathrm{T}}\, E^2\boldsymbol{\xi} = 0\,, \tag{2.178c}$$

$$(\boldsymbol{\eta} \times (E\boldsymbol{\eta}))^{\mathrm{T}}\, E^2\boldsymbol{\eta} = 0\,. \tag{2.178d}$$

Note, that the conditions (2.178) express linear dependencies via four vanishing triple products of three-vector type quantities. These triple products are "pseudo-scalar", that is, they are invariant under rotations (basis changes) but change their sign under reflections (CP_g type (i) or (ii) transformations). The conditions (2.178c) and (2.178d) are required for the case $\boldsymbol{\xi} \times \boldsymbol{\eta} = 0$, which leads to trivial equations for (2.178a) and (2.178b) and thus gives no constraints on the matrix E.

Proof. We shall show that (2.177) is equivalent to (2.178). It is easy to see that (2.177) implies (2.178) by inserting (2.177) into (2.178). Now we want to show that from (2.178) follows (2.177) with a suitable rotation $R(U)$. First we choose a basis where

$$\boldsymbol{\xi}' = R(U)\,\boldsymbol{\xi} = \begin{pmatrix} 0 \\ 0 \\ \xi_3' \end{pmatrix}, \qquad \boldsymbol{\eta}' = R(U)\,\boldsymbol{\eta} = \begin{pmatrix} \eta_1' \\ 0 \\ \eta_3' \end{pmatrix}. \tag{2.179}$$

Note that a rotation into this basis is always possible for two vectors. It remains to be shown that in addition $\eta_{12}' = \eta_{23}'$=0 can be achieved if (2.178a)-(2.178d) hold. We remark that E is a symmetric matrix (see (2.26)) and this property is not altered by a similarity transformation (2.33). We have to consider different cases depending on whether the vector cross product

$$\boldsymbol{\xi}' \times \boldsymbol{\eta}' = \begin{pmatrix} 0 \\ \xi_3'\eta_1' \\ 0 \end{pmatrix} \tag{2.180}$$

vanishes or not. Let us first assume that the vector cross product (2.180) does not vanish, that is, we have $\xi_3'\eta_1' \neq 0$. From (2.178a) we find now

$$(\boldsymbol{\xi}' \times \boldsymbol{\eta}')^{\mathrm{T}}\, E'\,\boldsymbol{\xi}' = \eta_{23}'\, \xi_3'^2 \eta_1' = 0\,. \tag{2.181}$$

This means that $\eta_{23}' = \eta_{32}' = 0$. Then (2.178b) gives

$$(\boldsymbol{\xi}' \times \boldsymbol{\eta}')^{\mathrm{T}} E'\,\boldsymbol{\eta}' = \eta_{21}'\, \xi_3' \eta_1'^2 = 0, \tag{2.182}$$

that is, we have also $\eta_{21}' = \eta_{12}' = 0$. Thus, the explicit form (2.177) follows from (2.178a)-(2.178d) for this case.

Now we have to consider also the special case of a vanishing vector cross product (2.180). In this case (2.178a) and (2.178b) are trivially fulfilled and give no constraint for the matrix E'. We shall now use (2.178c) and (2.178d) to prove (2.177). If $\boldsymbol{\xi} \times \boldsymbol{\eta} = 0$ and $\boldsymbol{\xi} = 0$ and

$\boldsymbol{\eta} = 0$ we can achieve (2.177) trivially by diagonalising E. Thus, consider the case that $\boldsymbol{\xi} \times \boldsymbol{\eta} = 0$ and $\boldsymbol{\xi} \neq 0$. By an orthogonal transformation we can diagonalise E:

$$R(U_1)ER^{\mathrm{T}}(U_1) = E' = \mathrm{diag}(\mu_1, \mu_2, \mu_3). \tag{2.183}$$

In this basis $\eta'_{12} = \eta'_{23} = 0$ is fulfilled already. Furthermore, we have

$$\boldsymbol{\xi}' = \begin{pmatrix} \xi'_1 \\ \xi'_2 \\ \xi'_3 \end{pmatrix}, \qquad E'\boldsymbol{\xi}' = \begin{pmatrix} \mu_1\xi'_1 \\ \mu_2\xi'_2 \\ \mu_3\xi'_3 \end{pmatrix}, \qquad \boldsymbol{\xi}' \times E'\boldsymbol{\xi}' = \begin{pmatrix} (\mu_3-\mu_2)\xi'_2\xi'_3 \\ (\mu_1-\mu_3)\xi'_3\xi'_1 \\ (\mu_2-\mu_1)\xi'_1\xi'_2 \end{pmatrix}, \tag{2.184}$$

and from (2.178c),

$$\left(\boldsymbol{\xi}' \times (E'\boldsymbol{\xi}')\right)^{\mathrm{T}} E'^2\boldsymbol{\xi}' = (\mu_1-\mu_2)(\mu_2-\mu_3)(\mu_3-\mu_1)\xi'_1\xi'_2\xi'_3 = 0\,. \tag{2.185}$$

If all eigenvalues μ_a are different we find from (2.185) that at least one ξ'_a must be zero. By a change of basis which interchanges the components we can achieve $\xi'_2 = 0$ without introducing off-diagonal elements in E'. Then $\boldsymbol{\eta}'$ being parallel to $\boldsymbol{\xi}'$ implies $\eta'_2 = 0$ and we found a basis of the form (2.177). Suppose, on the other hand, that at least two eigenvalues μ_a are equal. Without loss of generality we can suppose

$$\mu_1 = \mu_2\,. \tag{2.186}$$

By a rotation around the 3-axis, leaving E' diagonal, we can then achieve

$$\boldsymbol{\xi}' = \begin{pmatrix} \xi'_1 \\ 0 \\ \xi'_3 \end{pmatrix} \tag{2.187}$$

and also $\eta'_2 = 0$ since $\boldsymbol{\eta}'$ is parallel to $\boldsymbol{\xi}'$, *q.e.d.* For the case $\boldsymbol{\xi} \times \boldsymbol{\eta} = 0$ and $\boldsymbol{\eta} \neq 0$ the argumentation runs along the same lines using (2.178d) instead of (2.178c). This completes the proof that the set of the conditions (2.178a)-(2.178d) is equivalent to the existence of a basis satisfying (2.177). □

Comparing our conditions (2.178a)-(2.178d) to the literature, we find them to be equivalent to (23)-(26) in [75] as well as to the conditions given in (A)-(B) in [83]. In [75] the conditions were found by a systematic survey of all possible complex invariants - and there is an enormous number of such invariants - within a field based formulation, which is a completely different approach. Our triple products required to vanish in (2.178a), (2.178b), and (2.178d) turn out to be equal to $-2^{-5}I_{2Y2Z}$, $2^{-7}I_{Y3Z}$, and $-2^{-13}I_{6Z}$ in their notation. Despite the fact, that the fourth invariant occurring in [75] and our condition (2.178c) are different, we can show that the full sets of conditions are equivalent. This is conveniently done by computing the *reduced Gröbner bases* for both sets which are indeed equal (for a brief introduction to the formalism of Gröbner bases see appendix A). The proof above shows how a Higgs basis is constructed for which the potential is invariant under the

standard CP transformation, provided (2.178) holds. In this basis the parameters of the potential with respect to the Higgs fields, $V(\varphi_1, \varphi_2)$, are real. Note that by construction the parameters of $V(\tilde{\boldsymbol{K}})$ are always real, independent of its CP properties.

We remark that the conditions in (2.178) guarantee that the potential has at least one CP_g invariance transformation. It is possible that a theory has more than one CP_g invariance transformation. A sufficient condition guaranteeing the uniqueness of the CP_g transformation is

$$\boldsymbol{\xi} \times \boldsymbol{\eta} \neq 0\,. \tag{2.188}$$

Then, clearly the only reflection symmetry one can have is on the plane spanned by $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$. In table 2.2, we give a classification of CP_g type (ii) invariant theories with respect to the number of independent CP_g transformations they allow.

A final remark concerns the relation of type (i) and (ii) symmetries. From theorems 2.5.5 and 2.5.6 we see that a theory having the CP_g symmetry of type (i) is also invariant under - in fact several - CP_g transformations of type (ii).

2.5.3 CP symmetries of the vacuum

If there is no CP transformation under which the potential is invariant, CP is broken *explicitly*. If the potential is invariant under a certain CP_g transformation but the vacuum expectation value does not respect this symmetry we have *spontaneous* violation of this CP_g symmetry. Note that a potential can be symmetric under several CP_g transformations where some may be conserved and some violated by the vacuum expectation value. Examples for this case are given below.

Suppose now that the potential V has a CP_g symmetry, that is, an invariance under an improper rotation $\bar{R}^{\varphi}$. The potential parameters satisfy then (2.172). This symmetry is spontaneously broken if and only if the vacuum expectation value $\boldsymbol{K}$ does not respect this symmetry, that is, it fulfils

$$\bar{R}^{\varphi}\boldsymbol{K} \neq \boldsymbol{K}. \tag{2.189}$$

Note the gauge invariance and basis independence of this condition.

Spontaneous breaking of any CP symmetry implies that the vacuum belongs to a degenerate set of stationary points. From theorem 2.3.3 we see that the Lagrange multiplier u_v of the vacuum is an eigenvalue of $\tilde{g}\tilde{E}$ in this case.

We shall now derive conditions for which the vacuum is invariant under possible CP_s, CP_g type (i) or (ii) symmetries of the potential. Note that all conditions which refer to the vacuum solution may be evaluated directly if the potential is reparameterised as described at the end of section 2.4.

Standard CP symmetry

Suppose the potential is invariant under CP_s. Spontaneous CP_s violation means that the vacuum is not invariant under this symmetry,

$$R_2\boldsymbol{K} \neq \boldsymbol{K}, \tag{2.190}$$

that is, we have

$$K_2 \neq 0. \tag{2.191}$$

For a vacuum solution with the required EWSB we write as in (2.105)

$$\tilde{\boldsymbol{K}} = \begin{pmatrix} v_1^2 + v_2^2 \\ 2v_1 v_2 \cos\zeta \\ 2v_1 v_2 \sin\zeta \\ v_1^2 - v_2^2 \end{pmatrix}. \tag{2.192}$$

From (2.191) we find the well known result that CP_s is violated spontaneously if and only if $v_1 \neq 0$, $v_2 \neq 0$, $\zeta \neq 0, \pi$. That means, the vacuum expectation values of the two Higgs fields in this special basis must be complex relative to each other. We note, however, that this statement has no basis-independent meaning. By a suitable basis transformation we can always achieve that only one Higgs doublet has a non-vanishing vacuum expectation value which, moreover, is real (see section 2.4).

Let us discuss in which cases spontaneous CP_s violation arises for a stable vacuum. For any vacuum of a CP_s conserving potential (2.174) with partial EWSB we see that (2.72a) implies

$$(\mu_2 + u_v) K_2 = 0\,, \tag{2.193}$$

where $\mu_2 = \eta_{22}$ is an eigenvalue of $\tilde{E}$ due to the latter being diagonal in the 2–direction. Since spontaneous CP_s violation means $K_2 \neq 0$, it requires for the vacuum Lagrange multiplier

$$u_v = -\mu_2\,. \tag{2.194}$$

On the other hand, u_v is non-negative and related to the charged Higgs mass squared, see theorem 2.3.4 and (2.131), such that (2.194) implies

$$\mu_2 \leq 0\,, \tag{2.195}$$

where the inequality becomes strict for absence of massless charged Higgs bosons. Taking the relation (2.131) into account we summarise as follows.

Theorem 2.5.7. *Spontaneous violation of the CP_s symmetry of the potential with the parameters* (2.174) *requires*

$$\mu_2 = -u_v = -\frac{1}{2v^2} m_{H^\pm}^2 \leq 0\,, \tag{2.196}$$

for the eigenvalue $\mu_2 = \eta_{22}$ of $\tilde{E}$ and the Lagrange multiplier u_v of the vacuum.

For a vacuum to be stable it must be the global minimum of the potential. Let us therefore consider a CP_s-symmetric potential with parameters as in (2.174) having (at least) two stationary solutions on the light cone. We suppose that the CP_s symmetry is respected by one solution $\tilde{K}^{\mathrm{CP}}$ with $K_2^{\mathrm{CP}} = 0$ and violated by the other solution $\tilde{K}^{\not{\mathrm{CP}}}$ through $K_2^{\not{\mathrm{CP}}} \neq 0$.

We denote the corresponding Lagrange multipliers by u_{CP} and $u_{\not{\mathrm{CP}}} = -\mu_2$. Perturbing the CP_g conserving point by a small amount ($0 < \varepsilon \ll 1$) within the light cone according to

$$\tilde{K}^{\mathrm{CP}} \to \tilde{K}^{\mathrm{CP}} + K_0^{\mathrm{CP}} \begin{pmatrix} \sqrt{1+\varepsilon^2} - 1 \\ 0 \\ \pm\varepsilon \\ 0 \end{pmatrix}, \tag{2.197}$$

we find for the potential value

$$V(\tilde{K}^{\mathrm{CP}}) \to V(\tilde{K}^{\mathrm{CP}}) + (u_{\mathrm{CP}} + \mu_2) \left(K_0^{\mathrm{CP}}\right)^2 \varepsilon^2 + \mathcal{O}(\varepsilon^4) \tag{2.198}$$

after employing the corresponding stationarity condition (2.72a) with $u = u_{\mathrm{CP}}$. Therefore the CP_s conserving point can only be a (local) minimum if $u_{\mathrm{CP}} + \mu_2 \geq 0$, that is, if $u_{\mathrm{CP}} \geq u_{\not{\mathrm{CP}}}$. From theorem 2.3.3 we know that a higher Lagrange multiplier means a lower potential value. Thus, we find the following.

Theorem 2.5.8. *If the THDM potential has a CP_s conserving (local) minimum, there can be no stationary points with lower values of the potential which violate this symmetry.*

This result was found before, see [80] and references therein. While the existence of a CP_g conserving light-like minimum implies that the global minimum has these properties too, there are cases with more than one CP_g conserving light-like minimum, see subsection 2.6.3. Therefore, a determination of the actual global minimum is still necessary in general.

CP_g type (i) symmetry

According to theorem 2.5.5 the potential having CP_g invariance of type (i) has the form (see (2.175))

$$V = \xi_0 K_0(x) + \eta_{00} K_0^2(x) + \boldsymbol{K}^{\mathrm{T}}(x) E \boldsymbol{K}(x). \tag{2.199}$$

From (2.24) we see that the correct EWSB requires for the vacuum $\boldsymbol{K} \neq 0$. This implies then (2.189) with $\bar{R}^{\varphi} = -\mathbb{1}_3$ to hold,

$$-\mathbb{1}_3 \boldsymbol{K} \neq \boldsymbol{K}. \tag{2.200}$$

We formulate this result as a theorem:

Theorem 2.5.9. *A theory which is invariant under the CP_g type (i) transformation has the potential (2.199). The required EWSB implies that the CP_g type (i) symmetry is spontaneously broken.*

In section 2.7 we discuss in detail the stability and EWSB properties of this class of models having the potential (2.199). There we prove the following theorem.

Theorem 2.5.10. *Consider the Higgs part of the THDM Lagrangian* (2.3) *with the potential (2.199) having CP_g invariance of type (i). Let $\mu_1 \geq \mu_2 \geq \mu_3$ be the eigenvalues of E with this ordering. The theory is stable, has the correct EWSB and no zero mass charged Higgs boson if and only if*

$$\begin{aligned} &\eta_{00} > 0, \\ &\mu_a + \eta_{00} > 0 \quad \textit{for } a = 1,2,3, \\ &\xi_0 < 0, \\ &\mu_3 < 0. \end{aligned} \tag{2.201}$$

The CP_g symmetry of type (i) is then spontaneously broken.

CP_g type (ii) symmetries

For a theory having a CP_g invariance of type (ii) the parameters of the potential V must satisfy (2.178a)-(2.178d) according to theorem 2.5.6. Such a CP_g symmetry is spontaneously broken if (2.189) holds with $\bar{R}^\varphi$ as in (2.171). Suppose now that for given parameters satisfying (2.178a)-(2.178d) it has been checked that V is a stable potential. Suppose furthermore, that the vacuum solution $\underline{K}$ has been identified. The following theorem allows then to check if CP_g is spontaneously violated or not.

Theorem 2.5.11. *Suppose that the potential is invariant under one or more CP_g type (ii) transformations, that is, its parameters respect (2.178a)-(2.178d). Let $K_0, \boldsymbol{K}$ be the vacuum solution. There is a CP_g invariance which is also respected by the vacuum if and only if the following three basis independent equations hold:*

$$(\boldsymbol{\xi} \times \boldsymbol{\eta})^{\mathrm{T}} \boldsymbol{K} = 0\,, \tag{2.202a}$$

$$(\boldsymbol{\xi} \times (E\boldsymbol{\xi}))^{\mathrm{T}} \boldsymbol{K} = 0\,, \tag{2.202b}$$

$$(\boldsymbol{\eta} \times (E\boldsymbol{\eta}))^{\mathrm{T}} \boldsymbol{K} = 0\,. \tag{2.202c}$$

We distinguish two cases.

$$(a) \quad \boldsymbol{\xi} \times \boldsymbol{\eta} \neq 0\,.$$

The theory allows then exactly one CP_g type (ii) invariance transformation which is conserved also by the vacuum if and only if (2.202a) holds. In this case (2.202b) and (2.202c) are a consequence of (2.202a).

$$(b) \quad \boldsymbol{\xi} \times \boldsymbol{\eta} = 0\,.$$

Then (2.202a) is trivial. There may be more than one CP_g type (ii) invariance transformation. At least one of these symmetries is also respected by the vacuum if (2.202b) and (2.202c) hold.

In order to prove the theorem, we need to show a basis with

$$\xi_2' = 0\,, \tag{2.203a}$$
$$\eta_2' = \eta_{12}' = \eta_{23}' = 0\,, \tag{2.203b}$$
$$K_2' = 0 \tag{2.203c}$$

exists if and only if (2.178) and (2.202) hold. As can directly be seen by insertion, (2.203) implies (2.178) and (2.202) in the primed basis and due to the basis independence of (2.178) and (2.202) also in any basis. It remains to be shown that from (2.178) and (2.202) the existence of a basis (2.203) follows. We shall give two independent proofs. The first one is short and rather formal, but reveals some more invariants which must vanish if (2.202) holds. The second is quite lengthy, but shows explicitly how to find a basis in which CP_s is a symmetry, provided (2.202) holds. Furthermore, the occurrence of more than one CP_g type (*ii*) symmetry is considered in detail.

Formal proof. We shall prove that (2.178) and (2.202) imply (2.203) for some basis choice. For the stationary point $\tilde{\boldsymbol{K}} = 0$, which leaves the electroweak symmetry unbroken, the proof is trivial. We shall now prove the statement for all other stationary points, in particular for solutions with the required EWSB. We will use the fact that any stationary point $\tilde{\boldsymbol{K}} \neq 0$ fulfils a stationarity condition (2.82a) whose three-vector part can be written as

$$\boldsymbol{\xi} = -2\big(E\boldsymbol{K} + u\boldsymbol{K} + K_0\boldsymbol{\eta}\big) \tag{2.204}$$

with a specific value of u. As a preparation we first show that certain additional invariants vanish. Replacing $\boldsymbol{\xi}$ in (2.202a) via the stationarity condition (2.204) we find

$$(\boldsymbol{\eta} \times (E\boldsymbol{K}))^{\mathrm{T}}\ \boldsymbol{K} = 0\,. \tag{2.205}$$

This implies

$$(\boldsymbol{\xi} \times (E\boldsymbol{K}))^{\mathrm{T}}\ \boldsymbol{K} = 0\,, \tag{2.206}$$

which can be seen by replacing $\boldsymbol{\xi}$ via (2.204). Next we show that

$$(\boldsymbol{\eta} \times (E\boldsymbol{\xi}))^{\mathrm{T}}\ \boldsymbol{K} = 0\,. \tag{2.207}$$

If $\boldsymbol{\eta}$ and $\boldsymbol{K}$ are linearly dependent, (2.207) follows immediately. In the other case we replace $\boldsymbol{\xi}$ in (2.207) by a linear combination of $\boldsymbol{\eta}$ and $\boldsymbol{K}$, which is possible by (2.202a). Using (2.202c) and (2.205), (2.207) follows. Similarly we find

$$(\boldsymbol{\xi} \times (E\boldsymbol{\eta}))^{\mathrm{T}}\ \boldsymbol{K} = 0\,, \tag{2.208}$$

using (2.202a), (2.202b) and (2.206). The relation

$$(E\boldsymbol{K} \times (E\boldsymbol{\xi}))^{\mathrm{T}}\ \boldsymbol{K} = 0 \tag{2.209}$$

follows after substituting $E\boldsymbol{K}$ via (2.204) from (2.202b) and (2.207). Similarly we find

$$(E\boldsymbol{K} \times (E\boldsymbol{\eta}))^{\mathrm{T}}\ \boldsymbol{K} = 0 \tag{2.210}$$

using (2.204), (2.202c) and (2.208). We find

$$(\boldsymbol{K} \times (E\boldsymbol{K}))^{\mathrm{T}}\ E^2\boldsymbol{K} = 0 \tag{2.211}$$

by replacing $E\boldsymbol{K}$ in the term $E^2\boldsymbol{K}$ via (2.204) since (2.209) and (2.210) hold.

In the case that $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are linearly independent, we may choose a basis of the form (2.177) by theorem 2.5.6. From (2.202a) follows immediately that we have $K_2 = 0$ in this basis.

In the case that $\boldsymbol{\xi}$ is a multiple of $\boldsymbol{\eta}$ we note that (2.202c), (2.205), (2.211) and (2.178d),

$$\begin{aligned} (\boldsymbol{K} \times \boldsymbol{\eta})^{\mathrm{T}}\ E\boldsymbol{K} = 0\,, \quad (\boldsymbol{K} \times (E\boldsymbol{K}))^{\mathrm{T}}\ E^2\boldsymbol{K} = 0\,, \\ (\boldsymbol{K} \times \boldsymbol{\eta})^{\mathrm{T}}\ E\boldsymbol{\eta} = 0\,, \qquad (\boldsymbol{\eta} \times (E\boldsymbol{\eta}))^{\mathrm{T}}\ E^2\boldsymbol{\eta} = 0\,, \end{aligned} \tag{2.212}$$

are equal to the explicit CP conservation conditions (2.178a)-(2.178d) if we replace $\boldsymbol{\xi}$ by $\boldsymbol{K}$ in the latter. Using the proof of theorem 2.5.6 we find that there is a basis with $\eta_2' = K_2' = \eta_{12}' = \eta_{23}' = 0$ and thus also $\xi_2' = 0$.

In the case that $\boldsymbol{\eta}$ is a multiple of $\boldsymbol{\xi}$ we use (2.202b), (2.206), (2.211) and (2.178c),

$$\begin{aligned} (\boldsymbol{K} \times \boldsymbol{\xi})^{\mathrm{T}}\ E\boldsymbol{K} = 0\,, \quad (\boldsymbol{K} \times (E\boldsymbol{K}))^{\mathrm{T}}\ E^2\boldsymbol{K} = 0\,, \\ (\boldsymbol{K} \times \boldsymbol{\xi})^{\mathrm{T}}\ E\boldsymbol{\xi} = 0\,, \qquad (\boldsymbol{\xi} \times (E\boldsymbol{\xi}))^{\mathrm{T}}\ E^2\boldsymbol{\xi} = 0\,. \end{aligned} \tag{2.213}$$

Replacing $\boldsymbol{\eta}$ by $\boldsymbol{K}$ in the proof of theorem 2.5.6 we see that we can find a basis with $\xi_2' = K_2' = \eta_{12}' = \eta_{23}' = 0$ and thus also $\eta_2' = 0$. □

Detailed proof. Here we present an alternative proof for the existence of the basis (2.203) in case that (2.178) and (2.202) hold.

We discuss first the trivial case that the potential parameters satisfy (2.178a)-(2.178d) and the vacuum expectation value is the zero four-vector $\tilde{\boldsymbol{K}} = 0$. Then (2.202a)-(2.202c) are also trivially satisfied. From theorem 2.5.6 we see that we can go to a basis where (2.203a) and (2.203b) hold. Since $\boldsymbol{K} = 0$ in our case we have also $K_2' = 0$.

Thus we can turn to the case that $\tilde{\boldsymbol{K}} \neq 0$. Then $\tilde{\boldsymbol{K}}$ fulfils according to (2.82a) the stationarity condition

$$\boldsymbol{\xi} = -2\big(E\boldsymbol{K} + u\boldsymbol{K} + K_0\boldsymbol{\eta}\big)\,. \tag{2.214}$$

Consider now a potential with parameters satisfying (2.178). We may then choose a basis with $\boldsymbol{\xi}'$, $\boldsymbol{\eta}'$ and E' of the form (2.177) by theorem 2.5.6. With a suitable rotation in the 1-3 subspace we can diagonalise E'. Then we have

$$\boldsymbol{\xi}' = \begin{pmatrix} \xi_1' \\ 0 \\ \xi_3' \end{pmatrix}, \quad \boldsymbol{\eta}' = \begin{pmatrix} \eta_1' \\ 0 \\ \eta_3' \end{pmatrix}, \quad E' = \mathrm{diag}(\mu_1, \mu_2, \mu_3)\,, \tag{2.215}$$

$$\boldsymbol{\xi}' \times \boldsymbol{\eta}' = \begin{pmatrix} 0 \\ \xi_3'\eta_1' - \xi_1'\eta_3' \\ 0 \end{pmatrix}, \tag{2.216}$$

$$\boldsymbol{\xi}' \times E'\boldsymbol{\xi}' = \begin{pmatrix} 0 \\ (\mu_1 - \mu_3)\xi_1'\xi_3' \\ 0 \end{pmatrix}, \tag{2.217}$$

$$\boldsymbol{\eta}' \times E'\boldsymbol{\eta}' = \begin{pmatrix} 0 \\ (\mu_1 - \mu_3)\eta_1'\eta_3' \\ 0 \end{pmatrix}. \tag{2.218}$$

Let us first consider the case

(a) $\boldsymbol{\xi} \times \boldsymbol{\eta} \neq 0$:

This implies, of course, $\boldsymbol{\xi}' \times \boldsymbol{\eta}' \neq 0$, that is,

$$\xi_3'\eta_1' - \xi_1'\eta_3' \neq 0\,. \tag{2.219}$$

If now (2.202a) holds we get immediately

$$\begin{aligned} &(\boldsymbol{\xi}' \times \boldsymbol{\eta}')^{\mathrm{T}}\, \boldsymbol{K}' = 0 \\ &\Longrightarrow (\xi_3'\eta_1' - \xi_1'\eta_3')K_2' = 0 \\ &\Longrightarrow K_2' = 0\,. \end{aligned} \tag{2.220}$$

Furthermore, we find from (2.217), (2.218) and (2.220) that (2.202b) and (2.202c) are automatically satisfied. We summarise this case. If $\boldsymbol{\xi}\times\boldsymbol{\eta} \neq 0$ the only possible CP_g type (ii) symmetry is the reflection on the plane spanned by $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ (see subsection 2.5.2) and this symmetry is respected by the vacuum if and only if (2.202a) holds. In this case (2.202a) implies also (2.202b) and (2.202c). This proves the case (a) of theorem 2.5.11.

Next we consider the case

(b) $\boldsymbol{\xi} \times \boldsymbol{\eta} = 0$:

Then (2.202a) is trivially fulfilled. Suppose first that $\boldsymbol{\xi} \neq 0$. Then $\boldsymbol{\eta}$ is proportional to $\boldsymbol{\xi}$,

$$\boldsymbol{\eta} = \lambda\boldsymbol{\xi}\,. \tag{2.221}$$

For the case of linearly dependent vectors $\boldsymbol{K}$ and $\boldsymbol{\xi}$ we have in particular in the basis (2.215) $K_2' = 0$, so we may assume in the following $\boldsymbol{K}$ and $\boldsymbol{\xi}$ being linearly independent. Now we distinguish various subcases.

$(b.1)$ $\boldsymbol{\xi} \times E\boldsymbol{\xi} \neq 0$:

The only reflection plane for a symmetry of the potential is spanned by $\boldsymbol{\xi}$ and $E\boldsymbol{\xi}$ in this case. We get from (2.217)

$$(\mu_1 - \mu_3)\,\xi_1'\xi_3' \neq 0 \tag{2.222}$$

and from (2.202b)

$$(\mu_1 - \mu_3)\,\xi_1'\xi_3'K_2' = 0\,. \tag{2.223}$$

This leads to K_2'=0, *q.e.d.*

(b.2) $\boldsymbol{\xi} \times E\boldsymbol{\xi} = 0$:

There we have

$$(\mu_1 - \mu_3)\,\xi_1'\xi_3' = 0\,. \tag{2.224}$$

Now we distinguish the different cases for the eigenvalues of E.

(b.2.1) μ_1, μ_2, μ_3 all different:

We get $\xi_1'\xi_3' = 0$. If, for instance, $\xi_1' = 0$ the theory has two reflection symmetries namely in this basis R_1 and R_2 (see (2.169)). From (2.214) we have

$$\begin{aligned} 0 &= -2(\mu_1 + u)K_1'\,, \\ 0 &= -2(\mu_2 + u)K_2'\,. \end{aligned} \tag{2.225}$$

Since we consider here $\mu_1 \neq \mu_2$ we must have either $K_1' = 0$ or $K_2' = 0$. That is, at least one of the reflection symmetries R_1 or R_2 is conserved by the vacuum. In case $K_1' = 0$ we can interchange the 1′- and 2′-components by a change of basis and in this way achieve $K_2' = 0$, *q.e.d.* For $\xi_3' = 0$ the argumentation is analogous, involving R_1 and R_3.

(b.2.2) $\mu_1 = \mu_2 \neq \mu_3$:

We get again from (2.224) $\xi_1'\xi_3' = 0$. For $\xi_3' = 0$ the argumentation is as in (b.2.1). For $\xi_3' \neq 0$ and $\xi_1' = 0$ we may perform a rotation around the 3′-axis such that $K_2' = 0$ *q.e.d.* Note, that E' is not affected by this rotation since $\mu_1 = \mu_2$. In this case we have reflection symmetry on every plane containing the 3′-axis, in particular on the plane spanned by $\boldsymbol{\xi}'$ and $\boldsymbol{K}'$. The reflection symmetry on this plane clearly is conserved by the vacuum.

(b.2.3) $\mu_2 = \mu_3 \neq \mu_1$:

The argumentation is analogous to the case (b.2.2).

(b.2.4) $\mu_1 = \mu_3 \neq \mu_2$:

We can, by a rotation around the 2′-axis, leaving E' diagonal, achieve $\xi_1' = \xi_2' = 0$, $\xi_3' \neq 0$. Here R_1 and R_2 are reflection symmetries. Then (2.214) gives

$$\begin{aligned} 0 &= (\mu_1 + u)K_1', \\ 0 &= (\mu_2 + u)K_2'. \end{aligned} \tag{2.226}$$

Thus, either K_1' or K_2' must be zero. In case $K_1' = 0$ we can interchange the 1′- and 2′-components by a change of basis and in this way achieve $K_2' = 0$, *q.e.d.*

(b.2.5) $\mu_1 = \mu_2 = \mu_3$:

There is reflection symmetry on all planes containing $\boldsymbol{\xi}'$, in particular on the plane spanned by $\boldsymbol{\xi}'$ and $\boldsymbol{K}'$. This reflection symmetry is obviously unbroken by the vacuum. This proves theorem 2.5.11 for the case (b) if $\boldsymbol{\xi} \neq 0$. For $\boldsymbol{\eta} \neq 0$ everything runs analogously using (2.202c) instead of (2.202b).

(b.3) $\boldsymbol{\xi} = \boldsymbol{\eta} = 0$:

In this case we have CP_g invariance of type (i). There are then at least three CP_g type (ii) invariances. Here we get from (2.214)

$$\begin{aligned} 0 &= (\mu_1 + u)K'_1\,, \\ 0 &= (\mu_2 + u)K'_2\,, \\ 0 &= (\mu_3 + u)K'_3\,. \end{aligned} \tag{2.227}$$

If not all μ_a are equal, this implies that at least one $K'_a = 0$ ($a \in \{1,2,3\}$). By a change of basis we can always achieve that $K'_2 = 0$, *q.e.d.* If $\mu_1 = \mu_2 = \mu_3$ we have reflection symmetry of the potential on any plane. The reflection symmetries on all planes containing $\boldsymbol{K}'$ are respected by the vacuum. This completes the first proof of theorem 2.5.11.

From the detailed discussion above we also found the number of independent reflections symmetries, that is, type (ii) CP_g transformations, which occur for the various cases. This is summarised in table 2.2 where it is always supposed that the potential parameters satisfy (2.178a)-(2.178d). □

We compared our conditions (2.202a)-(2.202c) for the existence of a CP symmetry of the vacuum with those of theorem 4 in [75], which were proven in [74] and found before in [72, 73]. The triple product in (2.202a) equals $-(v/2)^4 \operatorname{Im} J_1$ in the notation of [75], the other invariants in the conditions of [75] and our conditions have no one-to-one correspondence. However, we find complete agreement between our conditions (2.202) and those of [75] taking into account the respective full set of equations, that is, including the explicit CP-conservation conditions and the stationarity equations. This equivalence was obtained via *radical membership* tests using *Gröbner basis* computations (see appendix A for a short introduction to Gröbner bases). We also compared our conditions (2.202a)-(2.202c) to the corresponding conditions a)-c) in [83] and find agreement up to (2.202c), which is not contained in the latter set of criteria. We note that condition c) of [83] is no independent restriction since it is a consequence of the stationarity condition, see (2.206). Further, we do find examples where omitting (2.202c) matters, that are examples satisfying the conditions of [83] but possessing spontaneous breaking of a unique CP symmetry of the potential.

Note that the formulation *absence of spontaneous CP violation* is not quite appropriate in this context. The correct statement is given in theorem 2.5.11 above. It covers also the

	parameter conditions	number of CP_g type (ii) reflection symmetries
(a)	$\boldsymbol{\xi} \times \boldsymbol{\eta} \neq 0$	1
(b)	$\boldsymbol{\xi} \times \boldsymbol{\eta} = 0$	
$(b.1)$	$\boldsymbol{\xi} \neq 0$, $\boldsymbol{\xi} \times E\boldsymbol{\xi} \neq 0$	1
	$\boldsymbol{\eta} \neq 0$, $\boldsymbol{\eta} \times E\boldsymbol{\eta} \neq 0$	1
$(b.2)$	$\boldsymbol{\xi} \neq 0$, $\boldsymbol{\xi} \times E\boldsymbol{\xi} = 0$ or $\boldsymbol{\eta} \neq 0$, $\boldsymbol{\eta} \times E\boldsymbol{\eta} = 0$, eigenvalues of E: μ_1, μ_2, μ_3	
$(b.2.1)$	μ_1, μ_2, μ_3 all different	2
$(b.2.2)$	$\mu_1 = \mu_2 \neq \mu_3$	2 or ∞
$(b.2.3)$	$\mu_2 = \mu_3 \neq \mu_1$	2 or ∞
$(b.2.4)$	$\mu_1 = \mu_3 \neq \mu_2$	2
$(b.2.5)$	$\mu_1 = \mu_2 = \mu_3$	∞
$(b.3)$	$\boldsymbol{\xi} = 0$, $\boldsymbol{\eta} = 0$, μ_1, μ_2, μ_3:	
	all different	3
	at least 2 equal	∞

Table 2.2: The CP_g type (ii) transformations are described by reflections on planes. The table lists the number of these symmetries for a potential satisfying (2.178a)-(2.178d) depending on the different cases for the parameters. The vacuum is invariant under at least one of the symmetries if and only if (2.202a)-(2.202c) hold. The numbering of the eigenvalues μ_1, μ_2, μ_3 of E is chosen such that $\mu_2 = \eta'_{22}$ in a basis where $\boldsymbol{\xi}'$, $\boldsymbol{\eta}'$ and E' have the form (2.177).

case that the theory has more than one independent CP_g type (ii) invariance transformation, where one is respected by the vacuum and another spontaneously broken. From the discussion of the case (b) in the second proof above and from table 2.2 we see that, indeed, these *mixed* cases really occur.

The properties of a spontaneously CP_s violating vacuum in theorems 2.5.7 and 2.5.8 can be generalised to the present case of CP_g type (ii) transformations. Choosing for some given CP_g type (ii) symmetry of the potential the basis where it equals the CP_s symmetry, both theorems are directly applicable. Further, note that the conditions on the vacuum given above may be directly evaluated in terms of conditions on the potential parameters, if the reparameterisation described in subsection 2.4 is applied and the required vacuum is confirmed to be the global minimum with the methods described in section 2.3.

As discussed in the previous subsection, a type (i) symmetry is necessarily spontaneously broken in a phenomenologically acceptable theory. On the other hand, a type (i)

symmetric model has at least three type (ii) symmetries. We see from table 2.2 that the vacuum respects at least one of these symmetries.

2.5.4 CP symmetries of the Yukawa terms

We parameterised in (2.5) the Yukawa part of the Lagrangian as

$$\begin{aligned}\mathscr{L}_{\text{Yuk}}(x) = &- \lambda_{ik}^{l\,j}\ \bar{L}_i^L(x)\,\varphi_j(x)\,l_k^R(x)\\ &- \lambda_{ik}^{d'\,j}\ \bar{Q}_i^L(x)\,\varphi_j(x)\,d_k'^R(x)\\ &- \lambda_{ik}^{u\,j}\ \bar{Q}_i^L(x)\,\epsilon\,\varphi_j^*(x)\,u_k^R(x) + h.c.\end{aligned} \tag{2.228}$$

Under a generalised CP transformation (2.142) we find

$$\begin{aligned}\text{CP}_g: \qquad \mathscr{L}_{\text{Yuk}}(x) \longrightarrow &- \lambda_{ik}^{l\,j\,*}\,U_{ii'}^L\,U_{jj'}^{\varphi\,*}\,U_{kk'}^{l\,*}\ \bar{L}_{i'}^L(x')\,\varphi_{j'}(x')\,l_{k'}^R(x')\\ &- \lambda_{ik}^{d'\,j\,*}\,U_{ii'}^Q\,U_{jj'}^{\varphi\,*}\,U_{kk'}^{d'\,*}\ \bar{Q}_{i'}^L(x')\,\varphi_{j'}(x')\,d_{k'}'^R(x')\\ &- \lambda_{ik}^{u\,j\,*}\,U_{ii'}^Q\,U_{jj'}^{\varphi}\,U_{k'}^{u\,*}\ \bar{Q}_{i'}^L(x')\,\epsilon\,\varphi_{j'}^*(x')\,u_{k'}^R(x') + h.c.\end{aligned} \tag{2.229}$$

Invariance of $\mathscr{L}_{\text{Yuk}}$ under CP_g is therefore equivalent to

$$\lambda_{ik}^{l\,j} = \lambda_{i'k'}^{l\,j'\,*}\,U_{i'i}^L\,U_{j'j}^{\varphi\,*}\,U_{k'k}^{l\,*}\,, \tag{2.230a}$$

$$\lambda_{ik}^{d'\,j} = \lambda_{i'k'}^{d'\,j'\,*}\,U_{i'i}^Q\,U_{j'j}^{\varphi\,*}\,U_{k'k}^{d'\,*}\,, \tag{2.230b}$$

$$\lambda_{ik}^{u\,j} = \lambda_{i'k'}^{u\,j'\,*}\,U_{i'i}^Q\,U_{j'j}^{\varphi}\,U_{k'k}^{u\,*}\,. \tag{2.230c}$$

By recursive insertion follow the necessary conditions for CP_g invariance of the Lagrangian

$$\lambda_{ik}^{l\,j} = \lambda_{i'k'}^{l\,j'}\,(U^{L\,*}\,U^L)_{i'i}\,(U^{\varphi}\,U^{\varphi\,*})_{j'j}\,(U^l\,U^{l\,*})_{k'k}\,, \tag{2.231a}$$

$$\lambda_{ik}^{d\,j} = \lambda_{i'k'}^{d\,j'}\,(U^{Q\,*}\,U^Q)_{i'i}\,(U^{\varphi}\,U^{\varphi\,*})_{j'j}\,(U^d\,U^{d\,*})_{k'k}\,, \tag{2.231b}$$

$$\lambda_{ik}^{u\,j} = \lambda_{i'k'}^{u\,j'}\,(U^{Q\,*}\,U^Q)_{i'i}\,(U^{\varphi\,*}\,U^{\varphi})_{j'j}\,(U^u\,U^{u\,*})_{k'k}\,. \tag{2.231c}$$

We saw the requirement $\text{CP}_g \circ \text{CP}_g$ to reproduce the original fields up to a phase restricted the mixing matrices to $U^{\psi}\,U^{\psi\,*} = \pm\mathbb{1}$, where $\psi = L, l, Q, u, d'$. Together with (2.231) follows therefore that such a CP_g symmetry of the Lagrangian requires

$$U\,U^* = -\mathbb{1} \quad \text{for 0 or 2 matrices } U \in \{U^L, U^l, U^{\varphi}\} \qquad \text{or} \qquad \lambda_{ijk}^{l} \equiv 0\,, \tag{2.232a}$$

$$U\,U^* = -\mathbb{1} \quad \text{for 0 or 2 matrices } U \in \{U^Q, U^{d'}, U^{\varphi}\} \qquad \text{or} \qquad \lambda_{ijk}^{d'} \equiv 0\,, \tag{2.232b}$$

$$U\,U^* = -\mathbb{1} \quad \text{for 0 or 2 matrices } U \in \{U^Q, U^u, U^{\varphi}\} \qquad \text{or} \qquad \lambda_{ijk}^{u} \equiv 0\,. \tag{2.232c}$$

That is, a Lagrangian with this kind of CP_g symmetry requires either an appropriate matching of the mixing types in the CP_g transformation or vanishing Yukawa couplings to the Higgses and therefore zero masses for all leptons, up-type or down-type quarks respectively.

In particular, for a CP_g symmetry with anti-symmetric Higgs mixing ($U^\varphi U^{\varphi *} = -\mathbb{1}$) also one of the involved fermion mixing matrices must be anti-symmetric in order to have non-vanishing Yukawa couplings. However, this is possible only for 2 generations of fermions or more. This proves the following.

Theorem 2.5.12. *Non-vanishing CP_g type (i) symmetric couplings of fermions to Higgs fields require*

$$N_f \geq 2\,, \tag{2.233}$$

that is at least two fermion generations.

We saw in subsection 2.5.2 that a CP_g type (i) symmetric Higgs sector is automatically CP_g type (ii) invariant. We discuss this model in detail in section 2.7, where we will see that such a Higgs potential has in fact at least three different CP_g type (ii) symmetries. Extending all of these symmetries also to the Yukawa couplings requires not only 2 fermion generations or more, but also leads to severe restrictions on the Yukawa couplings for the case $N_f = 2$. We summarise the general results of section 2.7 in the following theorem.

Theorem 2.5.13. *A CP_g type (i) symmetric Higgs potential is also invariant under at least three different CP_g type (ii) transformations. Extending these three symmetries of the potential to invariances of the full Lagrangian requires at least two fermion generations for non-vanishing Yukawa terms. For the case of two coupled fermion generations different and non-vanishing masses of the two generations imply large FCNCs. It is possible to prescribe the above set of CP_g symmetries such that one fermion generation stays massless, the other acquires a mass, and no large FCNCs occur.*

2.6 Examples

Here we apply the general considerations of sections 2.2 to 2.5 to specific models.

2.6.1 Minimal Supersymmetric Standard Model

In this subsection, we consider the MSSM Higgs potential. The MSSM Higgs sector has been worked out in great detail by many authors (see e.g. the review [53] and references therein), and radiative corrections are known to be sizable [93]. The purpose of the following discussion is to illustrate the methods for the general THDM described in the previous sections. We will reproduce the well-known tree-level results for its stability, symmetry breaking and mass spectrum. The MSSM Higgs potential is

$$V = V_D + V_F + V_{\text{soft}} \tag{2.234}$$

with

$$V_D = \frac{1}{8}(g^2 + g'^2)\left(H_d^\dagger H_d - H_u^\dagger H_u\right)^2 + \frac{1}{2}g^2\left|H_d^\dagger H_u\right|^2\,, \tag{2.235}$$

$$V_F = |\mu|^2\left(H_d^\dagger H_d + H_u^\dagger H_u\right), \tag{2.236}$$

$$V_{\text{soft}} = m_{H_1}^2\, H_d^\dagger H_d + m_{H_2}^2\, H_u^\dagger H_u - \left(m_3^2\, H_d^{\mathrm{T}} \epsilon H_u + h.c.\right), \tag{2.237}$$

where we follow closely the usual notation [95]. Here, H_d and H_u are Higgs doublets with weak hypercharges $y = -1/2$ and $y = +1/2$, respectively, $m_{H_1}^2$, $m_{H_2}^2$, $|\mu|^2$ are real and m_3^2 complex parameters of dimension mass squared.

Substituting H_d and H_u by doublets φ_1, φ_2 with the same weak hypercharge $y = +1/2$ according to (2.6) and using the relations (2.21), we can put the potential in the form (2.26). The parameters are

$$\eta_{00} = \frac{1}{8}g^2, \quad \boldsymbol{\eta} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad E = \frac{1}{8}\begin{pmatrix} -g^2 & 0 & 0 \\ 0 & -g^2 & 0 \\ 0 & 0 & g'^2 \end{pmatrix} \tag{2.238}$$

for $V_4 = V_D$ and

$$\xi_0 = |\mu|^2 + \frac{1}{2}(m_{H_1}^2 + m_{H_2}^2)\,, \quad \boldsymbol{\xi} = \begin{pmatrix} -\operatorname{Re}(m_3^2) \\ \operatorname{Im}(m_3^2) \\ \frac{1}{2}(m_{H_1}^2 - m_{H_2}^2) \end{pmatrix} \tag{2.239}$$

for $V_2 = V_F + V_{soft}$.

We determine the stability of the potential by employing theorem 2.2.1. The functions $f(u)$ (2.48) and $f'(u)$ (2.49) for the MSSM are

$$f(u) = u + \frac{1}{8}g^2, \tag{2.240}$$

$$f'(u) = 1\,. \tag{2.241}$$

Here, the set I (2.62) is given by $u = 0$ and the eigenvalues of E (2.238),

$$I = \left\{ u_1 = 0,\ u_2 = -\frac{1}{8}g^2,\ u_3 = \frac{1}{8}g'^2 \right\}. \tag{2.242}$$

We find for the stationary points of J_4 with $u_i = u_1, u_3$ the values $J_4(\boldsymbol{k})|_{stat} = f(u_i) > 0$, but for those with u_2 the value $J_4(\boldsymbol{k})|_{stat} = f(u_2) = 0$. Explicitly, the stationary points of J_4 with u_2 are

$$\boldsymbol{k} = (\cos\phi, \sin\phi, 0)^{\mathrm{T}},\ \phi \in \mathbb{R},\ \text{with } J_4(\boldsymbol{k}) = 0\,. \tag{2.243}$$

They are known as the "D-flat" directions, since they have $V_D = 0$. For the MSSM, they prevent the stability assertion by the quartic terms alone. For the stability to be guaranteed by $V_2 > 0$ in these directions, theorem 2.2.1 gives as condition, see (2.70) and (2.64), the inequality

$$g(u_2) - |\boldsymbol{\xi}_\perp(u_2)|\sqrt{f'(u_2)} = \xi_0 - \sqrt{\xi_1^2 + \xi_2^2} > 0\,. \tag{2.244}$$

Inserting (2.239) we get

$$\left|m_3^2\right| < |\mu|^2 + \frac{1}{2}(m_{H_1}^2 + m_{H_2}^2) \tag{2.245}$$

as the necessary and sufficient condition for the stability of the MSSM potential in the sense of (2.39).

Checking the conditions of theorem 2.5.6 we immediately see that the potential has a CP_g type (ii) symmetry.

For the global minimum to be non-trivial, criterion (i) of theorem 2.3.5 gives $\xi_0 < \sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$, or equivalently

$$|\mu|^2 + \frac{1}{2}(m_{H_1}^2 + m_{H_2}^2) < \sqrt{|m_3^2|^2 + \frac{1}{4}\left(m_{H_1}^2 - m_{H_2}^2\right)^2} \tag{2.246}$$

as a necessary and sufficient condition. We consider the acceptable global minimum candidates on the forward light cone with the method described in subsection 2.3.2. The conditions (2.245) and (2.246) prevent exceptional solutions. The regular solutions are determined by the functions

$$\tilde{f}(u) = -\frac{1}{4}\left(\frac{\xi_0^2-\xi_1^2-\xi_2^2}{\frac{1}{8}g^2-u} - \frac{\xi_3^2}{-\frac{1}{8}g'^2-u}\right), \tag{2.247}$$

$$\tilde{f}'(u) = -\frac{1}{4}\left(\frac{\xi_0^2-\xi_1^2-\xi_2^2}{(\frac{1}{8}g^2-u)^2} - \frac{\xi_3^2}{(-\frac{1}{8}g'^2-u)^2}\right), \tag{2.248}$$

$$K_0(u) = -\frac{1}{2}\cdot\frac{\xi_0}{\frac{1}{8}g^2-u}, \tag{2.249}$$

where we omitted the insertions (2.239) for a compact notation. Employing again the conditions (2.245) and (2.246) we find the following. The function $\tilde{f}'(u)$ always has two zeros and those zeros imply values of $K_0(u)$ with opposite signs. The physical solution with $K_0(u) > 0$ has the Lagrange multiplier

$$u_0 = \frac{1}{8}\cdot\frac{|\xi_3|\,g^2 + \sqrt{\xi_0^2-\xi_1^2-\xi_2^2}\,g'^2}{|\xi_3| - \sqrt{\xi_0^2-\xi_1^2-\xi_2^2}}, \tag{2.250}$$

which is positive. Figure 2.4 shows the functions $\tilde{f}'(u), K_0(u)$ for an example set of parameters (corresponding to the SPS1a scenario [119] at the tree-level). As apparent from the graph of $\tilde{f}'(u)$ the zero-crossing at u_0 is not very pronounced compared to the other features of the function. This visualises the fact that satisfying both (2.245) and (2.246) at the same time requires some kind of fine-tuning of the original potential parameters. If (2.246) is violated, the zero-crossing of $\tilde{f}'(u)$ in the physical region with $K_0(u) > 0$ vanishes. In this context we note that the scope of the present discussion is the tree-level and these results may be modified by radiative corrections, see the introductory remark. We conclude that (2.245) and (2.246) guarantee the existence of the stationary point $\tilde{\boldsymbol{K}}(u_0)$, which fulfils criterion (ii) of theorem 2.3.5 and therefore is the global minimum with the required EWSB pattern. Moreover, there are no other local minima.

The existence of a unique minimum immediately tells us that there is no spontaneous CP_g violation, since CP_g violating minima always come in pairs. This can also be seen from table 2.2. Note from (2.238) and (2.239) that it is always possible by a rotation in the 1–2 plane of the orbit variables to choose a basis with $\xi_1 = -|m_3^2|$, $\xi_2 = 0$ without

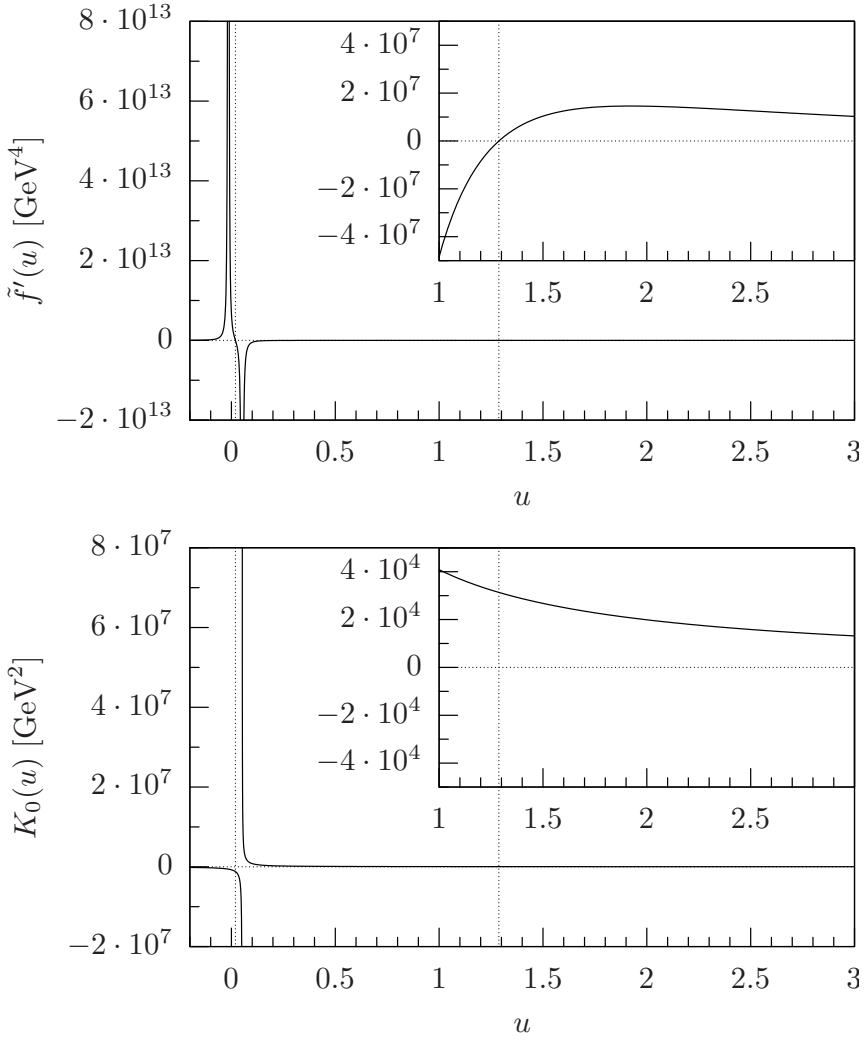

Figure 2.4: The global minimum determining functions $\tilde{f}'(u)$ and $K_0(u)$ for the MSSM, see (2.248) and (2.249), with $|\mu|^2 + m_{Hd}^2 = 157486\,\mathrm{GeV}^2$, $|\mu|^2 + m_{Hu}^2 = -2541\,\mathrm{GeV}^2$, $|m_3^2| = 15341\,\mathrm{GeV}^2$, corresponding to the SPS1a scenario [119] at the tree-level. The small boxes show the functions with enhanced ordinate resolution in the region around the physically relevant zero of $\tilde{f}'(u)$.

affecting any other parameters. We change to this basis where the potential as well as the vacuum has the standard CP$_s$ symmetry. We further choose a gauge where (2.105) holds, and perform the rotation (2.106) with

$$\tan\beta = \sqrt{\frac{\xi_0\,|\xi_3| + \sqrt{\xi_0^2 - \xi_1^2 - \xi_2^2}\,\xi_3}{\xi_0\,|\xi_3| - \sqrt{\xi_0^2 - \xi_1^2 - \xi_2^2}\,\xi_3}} \tag{2.251}$$

into a basis of the form (2.110), which has the new parameters

$$\boldsymbol{\xi}' = \begin{pmatrix} -c_{2\beta}\,|m_3^2| - s_{2\beta}\,\frac{1}{2}(m_{H_1}^2 - m_{H_2}^2) \\ 0 \\ -s_{2\beta}\,|m_3^2| + c_{2\beta}\,\frac{1}{2}(m_{H_1}^2 - m_{H_2}^2) \end{pmatrix}, \tag{2.252}$$

$$E' = \frac{1}{8}\begin{pmatrix} -g^2 + s_{2\beta}^2\,\bar{g}^2 & 0 & -\frac{1}{2}s_{4\beta}\,\bar{g}^2 \\ 0 & -g^2 & 0 \\ -\frac{1}{2}s_{4\beta}\,\bar{g}^2 & 0 & -g^2 + c_{2\beta}^2\,\bar{g}^2 \end{pmatrix} \tag{2.253}$$

with the abbreviations $\bar{g}^2 := g^2 + g'^2$ and $s_{2\beta} := \sin 2\beta$ etc. We insert the expressions into the formulae of section 2.4 and use

$$m_W^2 := \left(\frac{1}{2}gv\right)^2, \qquad m_Z^2 := \left(\frac{1}{2}\bar{g}v\right)^2 \tag{2.254}$$

with $v = \sqrt{2K_0(u_0)}$. Since we have CP_s conservation of both the potential and the vacuum we can assign conserved CP-parities to the neutral Higgs bosons. We obtain the mass squares

$$m_A^2 = 2v^2(\eta'_{22} + u_0)\,, \quad m_{H^\pm}^2 = m_A^2 + m_W^2 \tag{2.255}$$

for the pseudo-scalar boson $A := h''$ and the charged bosons $H^\pm$, which are mass eigenstates already. The non-diagonal part of the neutral mass matrix is

$$\mathcal{M}^2\big|_{\substack{\text{neutral,}\\ CP\text{ even}}} = \begin{pmatrix} c_{2\beta}^2\,m_Z^2 & -\frac{1}{2}s_{4\beta}\,m_Z^2 \\ -\frac{1}{2}s_{4\beta}\,m_Z^2 & m_A^2 + s_{2\beta}^2 m_Z^2 \end{pmatrix} \tag{2.256}$$

in the basis (ρ, h'). Its diagonalisation leads to mass squares

$$m_{h,H}^2 = \frac{1}{2}\left(m_A^2 + m_Z^2 \mp \sqrt{(m_A^2 + m_Z^2)^2 - 4\,c_{2\beta}^2\,m_A^2 m_Z^2}\right) \tag{2.257}$$

for the mass eigenstates h, H. They are obtained from

$$\begin{pmatrix} H \\ h \end{pmatrix} = \begin{pmatrix} \cos\alpha' & \sin\alpha' \\ -\sin\alpha' & \cos\alpha' \end{pmatrix}\begin{pmatrix} \rho \\ h' \end{pmatrix} \tag{2.258}$$

with the mixing angle α' determined by

$$\begin{aligned} \cos 2\alpha' &= -\,\frac{m_A^2 - c_{4\beta}\,m_Z^2}{m_H^2 - m_h^2}\,, \\ \sin 2\alpha' &= -\,\frac{s_{4\beta}\,m_Z^2}{m_H^2 - m_h^2}\,. \end{aligned} \tag{2.259}$$

Note that from tree-level formula (2.257) one may read off the *upper bound* on the lightest CP even Higgs mass

$$m_h \leq m_Z = 91\,\mathrm{GeV}\,. \tag{2.260}$$

If the tree-level bound was to hold as it stands, the MSSM would already be excluded by direct Higgs searches at LEP [92]. However, it is modified by radiative corrections. While the masses of supersymmetric particles can in general not be predicted on a rigorous basis, this bound represents a striking prediction of the MSSM. Moreover, also non-minimal scenarios predict only slightly looser versions of this bound. See the discussion in section 1.4 for more details and references.

The basis change described above corresponds to a unitary transformation (2.106) of the complete doublets. This leads to states ρ, h' with simple couplings to the gauge bosons, e.g. vanishing ZZh' and WWh' couplings at tree-level. However, usually the real parts of the neutral doublet components are excluded from that transformation (2.106), which is sufficient to disentangle the mass terms of the gauge bosons (and the associated Goldstone contributions in other gauges than considered here). Applying the inverse of the rotation (2.106) only to the neutral components (ρ, h') gives $\sqrt{2}(\mathrm{Re}\, H_d^1, \mathrm{Re}\, H_u^2)$. The mass matrix in this basis is diagonalised analogously to (2.258), where α' is replaced by the mixing angle α with

$$\begin{aligned} \cos 2\alpha &= -\cos 2\beta \, \frac{m_A^2 - m_Z^2}{m_H^2 - m_h^2}\,, \\ \sin 2\alpha &= -\sin 2\beta \, \frac{m_A^2 + m_Z^2}{m_H^2 - m_h^2}\,, \end{aligned} \tag{2.261}$$

which is the well-known result.

2.6.2 Models with softly broken $\mathbb{Z}_2$

We consider a class of THDMs with the Higgs potential

$$\begin{aligned} V(\varphi_1, \varphi_2) = {} & m_{11}^2 \varphi_1^\dagger \varphi_1 + m_{22}^2 \varphi_2^\dagger \varphi_2 - \left[m_{12}^2 \varphi_1^\dagger \varphi_2 + h.c. \right] \\ & + \frac{1}{2}\lambda_1 (\varphi_1^\dagger \varphi_1)^2 + \frac{1}{2}\lambda_2 (\varphi_2^\dagger \varphi_2)^2 + \lambda_3 (\varphi_1^\dagger \varphi_1)(\varphi_2^\dagger \varphi_2) \\ & + \lambda_4 (\varphi_1^\dagger \varphi_2)(\varphi_2^\dagger \varphi_1) + \left[\frac{1}{2}\lambda_5 (\varphi_1^\dagger \varphi_2)^2 + h.c. \right], \end{aligned} \tag{2.262}$$

written in the parameterisation of [120], where m_{12}^2 and λ_5 may be arbitrary complex and all other parameters are real. This potential breaks the $\mathbb{Z}_2$ symmetry

$$\begin{aligned} \varphi_1 &\longrightarrow -\varphi_1 \\ \varphi_2 &\longrightarrow \varphi_2 \end{aligned} \tag{2.263}$$

only softly, that is by quadratic terms in the Higgs doublet fields. This is motivated by the suppression of large FCNCs for certain choices of the Yukawa couplings. We put the

potential into the form (2.26) using the relations (2.21) and get

$$\xi_0 = \frac{1}{2}(m_{11}^2 + m_{22}^2)\,, \qquad \boldsymbol{\xi} = \begin{pmatrix} -\operatorname{Re} m_{12}^2 \\ \operatorname{Im} m_{12}^2 \\ \frac{1}{2}\,(m_{11}^2 - m_{22}^2) \end{pmatrix}, \tag{2.264a}$$

$$\eta_{00} = \frac{1}{8}(\lambda_1 + \lambda_2 + 2\lambda_3)\,, \qquad \boldsymbol{\eta} = \frac{1}{8}\begin{pmatrix} 0 \\ 0 \\ \lambda_1 - \lambda_2 \end{pmatrix},$$

$$E = \frac{1}{4}\begin{pmatrix} \lambda_4 + \operatorname{Re}\lambda_5 & -\operatorname{Im}\lambda_5 & 0 \\ -\operatorname{Im}\lambda_5 & \lambda_4 - \operatorname{Re}\lambda_5 & 0 \\ 0 & 0 & \frac{1}{2}(\lambda_1 + \lambda_2 - 2\lambda_3) \end{pmatrix}. \tag{2.264b}$$

The stability of the potential is easily investigated employing theorem 2.2.1. Stability is guaranteed by the terms quartic in the fields alone if and only if

$$\lambda_1 > 0\,, \quad \lambda_2 > 0\,, \quad \text{and } \sqrt{\lambda_1\lambda_2} + \lambda_3 > \max(0, |\lambda_5| - \lambda_4)\,. \tag{2.265}$$

In order to determine the CP properties of the potential we have to check (2.178). Two of the conditions for CP_g type (ii) invariance of the potential, (2.178b) and (2.178d), are, with (2.264), automatically fulfilled. The remaining conditions (2.178a) and (2.178c) give

$$(\lambda_1 - \lambda_2)\ \operatorname{Im}\left((m_{12}^2)^2\lambda_5^*\right) = 0\,, \tag{2.266}$$

$$\left[(\lambda_1 + \lambda_2 - 2(\lambda_3 + \lambda_4))^2 - 4\,|\lambda_5|^2\right](m_{11}^2 - m_{22}^2)\ \operatorname{Im}\left((m_{12}^2)^2\lambda_5^*\right) = 0 \tag{2.267}$$

as necessary and sufficient conditions for the existence of a CP_g invariance of type (ii) for the potential. It is obvious that for the case of real parameters m_{12}^2 and λ_5 (2.266) and (2.267) are satisfied. For $\boldsymbol{\xi} \times \boldsymbol{\eta}$ we find from (2.264)

$$\boldsymbol{\xi} \times \boldsymbol{\eta} = \frac{1}{8}(\lambda_1 - \lambda_2)\begin{pmatrix} \operatorname{Im}(m_{12}^2) \\ \operatorname{Re}(m_{12}^2) \\ 0 \end{pmatrix}. \tag{2.268}$$

From theorem 2.5.6 ff. we find, therefore, that in this model the potential allows one or more CP_g symmetries if and only if (2.266) and (2.267) hold. There is exactly one CP_g symmetry if $\lambda_1 - \lambda_2 \neq 0$ and $m_{12}^2 \neq 0$.

In the case CP_g is conserved, that is (2.266), (2.267) are fulfilled, CP_g may be violated spontaneously. We reparameterise the potential via the stationarity conditions in the general basis, see end of section 2.4, and assume that the vacuum expectation values v_1, v_2 together with the phase ζ indeed describe a sufficiently stable minimum (2.192) of the potential. The reparameterisation replaces m_{11}^2, m_{12}^2 and m_{22}^2 by

$$\begin{aligned} m_{11}^2 &= -\lambda_1 v_1^2 + (4u_v - \lambda_3)v_2^2\,, \\ m_{22}^2 &= (4u_v - \lambda_3)v_1^2 - \lambda_2 v_2^2\,, \\ m_{12}^2 &= v_1 v_2(e^{i\zeta}\lambda_5 + e^{-i\zeta}(4u_v + \lambda_4)) \end{aligned} \tag{2.269}$$

where $u_v = m_{H^\pm}^2/(2v^2)$ is the Lagrange multiplier of the vacuum. We check the conditions for spontaneous CP_g violation (2.202a)-(2.202c) and see that (2.202c) is automatically fulfilled. We find that (2.202a), (2.202b) together with (2.266), (2.267) are equivalent to the condition that either

$$v_1 v_2 \left[\cos(2\zeta)\,\mathrm{Im}\,\lambda_5 + \sin(2\zeta)\,\mathrm{Re}\,\lambda_5\right] = 0 \tag{2.270}$$

or

$$\lambda_1 = \lambda_2, \qquad (v_1^2 - v_2^2)\left[(\lambda_3 + \lambda_4 - \lambda_1)^2 - |\lambda_5|^2\right] = 0 \tag{2.271}$$

or both are fulfilled. That is, exactly if (2.270) or (2.271) or both are fulfilled, there is a CP_g symmetry of both the potential and the vacuum expectation value $\tilde{\boldsymbol{K}}$.

In the general case the potential can have more than one local minimum. The vacuum expectation values used in the parameterisation describe a stationary point by construction, for positive masses a local minimum. Whether it parameterises a stable vacuum, that is, the global minimum, must still be checked, e.g. by considering all stationary points of the potential. The generic stationary points are determined by

$$\tilde{f}(u) = \frac{\mathrm{Re}\,(\lambda_5 (m_{12}^{2\,*})^2) - (4u + \lambda_4)\,|m_{12}^2|^2}{16(u - \tilde{\mu}_1)(u - \tilde{\mu}_2)} + \frac{2(4u - \lambda_3)m_{11}^2 m_{22}^2 + \lambda_1 m_{22}^4 + \lambda_2 m_{11}^4}{32(u - \tilde{\mu}_3)(u - \tilde{\mu}_4)}\,, \tag{2.272}$$

$$\tilde{f}'(u) = \frac{|(4u + \lambda_4)m_{12}^2 - \lambda_5 m_{12}^{2\,*}|^2}{64(u - \tilde{\mu}_1)^2(u - \tilde{\mu}_2)^2} - \frac{((4u - \lambda_3)m_{11}^2 + \lambda_1 m_{22}^2)((4u - \lambda_3)m_{22}^2 + \lambda_2 m_{11}^2)}{64(u - \tilde{\mu}_3)^2(u - \tilde{\mu}_4)^2}\,, \tag{2.273}$$

$$K_0(u) = \frac{\lambda_1 m_{22}^2 + \lambda_2 m_{11}^2 + (4u - \lambda_3)(m_{11}^2 + m_{22}^2)}{16(u - \tilde{\mu}_3)(u - \tilde{\mu}_4)} \tag{2.274}$$

where we omitted the insertions (2.269) for brevity. The eigenvalues of $\tilde{E}\tilde{g}$

$$\tilde{\mu}_{1,2} = -\frac{1}{4}(\lambda_4 \pm |\lambda_5|)\,, \tag{2.275}$$

$$\tilde{\mu}_{3,4} = \frac{1}{4}(\lambda_3 \mp \sqrt{\lambda_1 \lambda_2}) \tag{2.276}$$

may lead to exceptional stationary points.

For the special case $\lambda_5 = 0$ we get the simple form

$$\tilde{f}'(u) = \frac{|m_{12}^2|^2}{4(u - \tilde{\mu}_1)^2} - \frac{((4u - \lambda_3)m_{11}^2 + \lambda_1 m_{22}^2)((4u - \lambda_3)m_{22}^2 + \lambda_2 m_{11}^2)}{64(u - \tilde{\mu}_3)^2(u - \tilde{\mu}_4)^2} \tag{2.277}$$

The determination of the generic Lagrange multipliers corresponds to finding the zeros of $\tilde{f}'(u)$, which means solving a univariate polynomial of only degree four in this case. This is possible analytically and we indeed find explicit analytical expressions. However, they are very lengthy.

From a practical point of view it is sufficient to directly determine the zeros of (2.273) (in the general basis) numerically. Actually, once the results become lengthy one might also consider to use the fully automated Gröbner basis approach which we describe in the next chapter. In the following subsection we will give explicit analytical results for a special case of the model discussed here, which has the interesting property that two minima can coexist.

2.6.3 A simple THDM with two minima

We consider the potential (2.262) with the reparameterisation (2.269) of the V_2 parameters for the case that $\lambda_1 = \lambda_2,\ \lambda_5 = 0,\ v_2 = 0$:

$$\begin{aligned} V(\varphi_1,\varphi_2) = &-\frac{1}{2}\lambda_1 v^2 \varphi_1^\dagger\varphi_1 + \frac{1}{2}(2m_{H^\pm}^2 - \lambda_3 v^2)\varphi_2^\dagger\varphi_2 \\ &+ \frac{1}{2}\lambda_1\big((\varphi_1^\dagger\varphi_1)^2 + (\varphi_2^\dagger\varphi_2)^2\big) + \lambda_3(\varphi_1^\dagger\varphi_1)(\varphi_2^\dagger\varphi_2) + \lambda_4(\varphi_1^\dagger\varphi_2)(\varphi_2^\dagger\varphi_1)\,. \end{aligned} \tag{2.278}$$

Clearly, the potential and the vacuum $\tilde{\boldsymbol{K}} = (v^2/2)(1,0,0,1)^{\mathrm{T}}$ are CP_s invariant, see (2.270). Furthermore, we replace the following V_4 parameters in favour of physical masses, see section 2.4,

$$\lambda_1 = m_{\rho'}^2/v^2\,, \tag{2.279}$$

$$\lambda_4 = 2(m_A^2 - m_{H^\pm}^2)/v^2\,. \tag{2.280}$$

All of the Higgses ρ', h' and $h'' \equiv A$ are mass eigenstates already. We denote their masses by $m_{\rho'}$, $m_{h'}$ and m_A, respectively, and assume all to be positive. We do not introduce the notation $h,\ H$ for the fields ρ' and h' since we leave open which of the latter two fields is the lighter one. One of the scalar Higgses and the pseudo-scalar Higgs are degenerate:

$$m_{h'} = m_A\,. \tag{2.281}$$

Thus we are left with the input parameters

$$m_{\rho'}^2 > 0\,,\quad m_A^2 > 0\,,\quad m_{H^\pm}^2 > 0\,,\quad \lambda_3 \tag{2.282}$$

in addition to the fixed electroweak scale $v \approx 246\,\mathrm{GeV}$. Stability in the strong sense is then equivalent to

$$\lambda_3 > \max(\,-m_{\rho'}^2\,,\,-m_{\rho'}^2 + 2(m_{H^\pm}^2 - m_A^2)\,)/v^2\,. \tag{2.283}$$

We now consider all stationary points of the potential using the methods described in subsection 2.3.2. Exceptional solutions may occur with a Lagrange multiplier u coinciding with one of the eigenvalues of $\tilde{g}\tilde{E}$:

$$\tilde{\mu}_1 = \tilde{\mu}_2 = \frac{1}{2v^2}(m_{H^\pm}^2 - m_A^2)\,, \tag{2.284}$$

$$\tilde{\mu}_{3,4} = \frac{1}{4}(\lambda_3 \mp m_{\rho'}^2/v^2)\,. \tag{2.285}$$

	u	*stationary point* $\tilde{\boldsymbol{K}}$	*conditions for existence*	*type*
none		0	$\lambda_3 v^2 \leq 2m_{H^\pm}^2$	saddle
			$\lambda_3 v^2 > 2m_{H^\pm}^2$	maximum
full (inside cone)	0	$\begin{pmatrix} \frac{v^2\left(\lambda_3 v^2 + m_{\rho'}^2 - 2m_{H^\pm}^2\right)}{2\left(\lambda_3 v^2 + m_{\rho'}^2\right)} \\ K_1 \\ K_2 \\ \frac{v^2\left(\lambda_3 v^2 - m_{\rho'}^2 - 2m_{H^\pm}^2\right)}{2\left(\lambda_3 v^2 - m_{\rho'}^2\right)} \end{pmatrix}$	$\lambda_3 v^2 > m_{H^\pm}^2 + \sqrt{m_{H^\pm}^4 + m_{\rho'}^4}$ if $m_{H^\pm}^2 \neq m_A^2$: $K_1 = K_2 = 0$ if $m_{H^\pm}^2 = m_A^2$: $K_1, K_2 : K_0 > \lvert\boldsymbol{K}\rvert$	saddle
EWSB: partial (on cone)	$\tilde{\mu}_1$	$\begin{pmatrix} \frac{v^2\left(\lambda_3 v^2 + m_{\rho'}^2 - 2m_{H^\pm}^2\right)}{2\left(\lambda_3 v^2 + m_{\rho'}^2 - 2m_{H^\pm}^2 + 2m_A^2\right)} \\ K_1 \\ K_2 \\ \frac{v^2\left(\lambda_3 v^2 - m_{\rho'}^2 - 2m_{H^\pm}^2\right)}{2\left(\lambda_3 v^2 - m_{\rho'}^2 - 2m_{H^\pm}^2 + 2m_A^2\right)} \end{pmatrix}$	$\lambda_3 v^2 > 2m_{H^\pm}^2 - m_A^2 + \sqrt{m_A^4 + m_{\rho'}^4}$, $K_1, K_2 : K_0 = \lvert\boldsymbol{K}\rvert$	saddle
	$\tilde{\mu}_5$	$\frac{1}{2}v^2 \frac{\lambda_3 v^2 - 2m_{H^\pm}^2}{m_{\rho'}^2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$	$\lambda_3 v^2 > 2m_{H^\pm}^2$ $\lambda_3 v^2 \leq \max(m_{H^\pm}^2 + \sqrt{m_{H^\pm}^4 + m_{\rho'}^4},$ $2m_{H^\pm}^2 - m_A^2 + \sqrt{m_A^4 + m_{\rho'}^4})$	saddle
			$\lambda_3 v^2 > \max(m_{H^\pm}^2 + \sqrt{m_{H^\pm}^4 + m_{\rho'}^4},$ $2m_{H^\pm}^2 - m_A^2 + \sqrt{m_A^4 + m_{\rho'}^4})$	minimum
	u_v	$\frac{1}{2}v^2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$		minimum (required vevs)

Table 2.3: Stationary points of the potential for the model (2.278). All physical masses are considered to be positive (2.282) and the stability constraint (2.283) is assumed to hold. The otherwise regular stationary points associated to $u = 0, \tilde{\mu}_5, u_v$ are exceptional if their Lagrange multipliers coincide with one of $\tilde{\mu}_1,\ \tilde{\mu}_3,\ \tilde{\mu}_4$. The minimum with the required vacuum expectation values (vevs) is the global minimum if and only if $\lambda_3 v^2 \leq 2m_{H^\pm}^2 + m_{\rho'}^2$. An illustration of the values of the potential at the stationary points is given in figure 2.5.

The Lagrange multipliers of the regular solutions are the zeros $\tilde{\mu}_5, \tilde{\mu}_6$ of $\tilde{f}'(u)$:

$$\tilde{\mu}_5 = \frac{1}{4}\left(\lambda_3 + \frac{m_{\rho'}^4}{v^2\left(2m_{H^\pm}^2 - \lambda_3 v^2\right)}\right) , \tag{2.286}$$

$$\tilde{\mu}_6 = u_v = \frac{1}{2v^2}\, m_{H^\pm}^2 . \tag{2.287}$$

Our results for all stationary points are summarised in table 2.3. Note that to determine the Lagrange multiplier of the global minimum it is sufficient to check the hierarchy of the

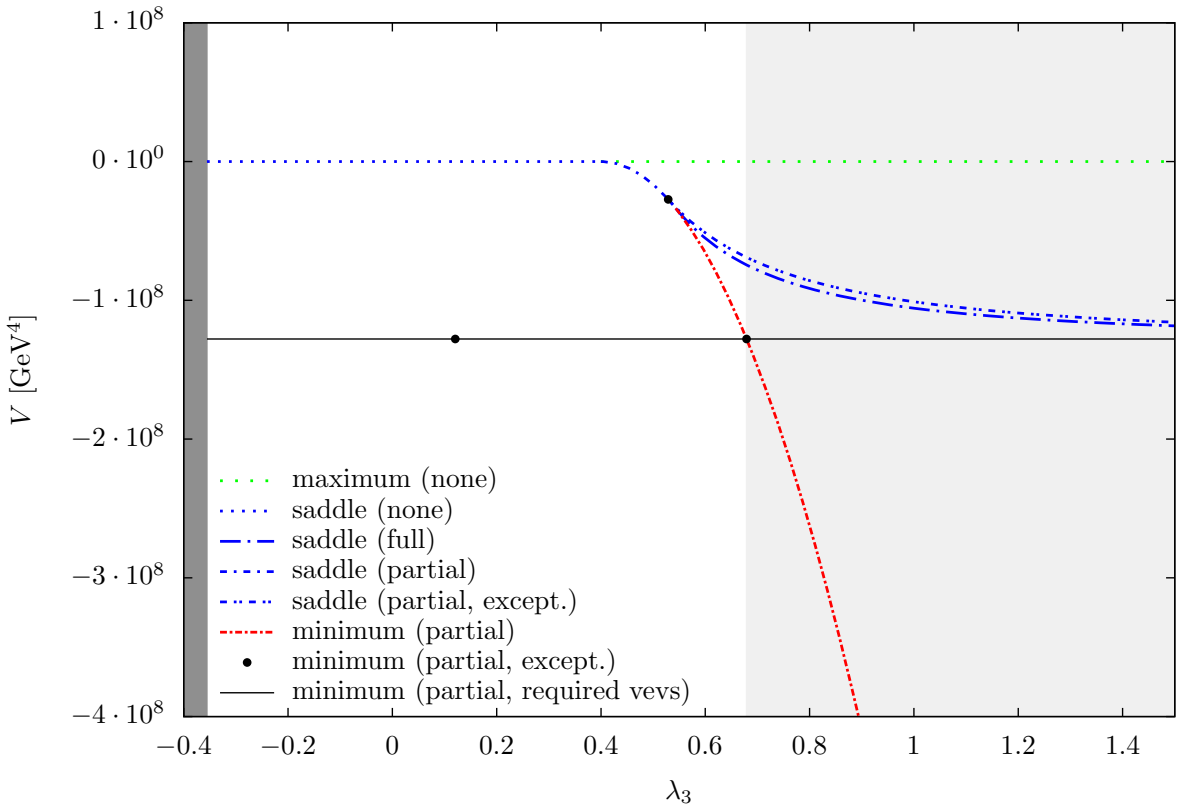

Figure 2.5: Values of the potential at its stationary points for variation of λ_3 in the model (2.278). Parameters are $m_{\rho'} = 110\,\text{GeV}$, $m_A = 120\,\text{GeV}$, $m_{H^\pm} = 130\,\text{GeV}$. The location of the orbit variables with respect to the forward light cone determines the EWSB of the respective stationary point: *none* (tip), *full* (inner part) and *partial* (on cone). Analytical expressions for all stationary points are listed in table 2.3. In the *dark grey* area the potential is unbounded from below, in the *light grey* area the required vacuum (solid line) is not the global minimum (respective lowest line). Only in the unshaded area the vacuum is stable.

Lagrange multipliers for parameters where the associated stationary solutions exist. We find that the global minimum has the Lagrange multiplier

$$u_0 = \begin{cases} u_v & \text{if and only if } \lambda_3 \leq (2m_{H^\pm}^2 + m_{\rho'}^2)/v^2\,, \\ \tilde{\mu}_5 & \text{if and only if } \lambda_3 \geq (2m_{H^\pm}^2 + m_{\rho'}^2)/v^2\,. \end{cases} \tag{2.288}$$

In other words, the required vacuum solution with $u = u_v$ is the global minimum only for $\lambda_3 \leq (2m_{H^\pm}^2 + m_{\rho'}^2)/v^2$. Figure 2.5 shows the value of the potential at all stationary points as listed in table 2.3 along with the conditions (2.283), (2.288) for a stable vacuum for fixed values of the physical masses and variation of λ_3. We stress that by construction the required vacuum with $u = u_v$ is always a local minimum, and all Higgs masses are independent of λ_3. Thus it is really impossible to recognise the instability of the vacuum for $\lambda_3 > (2m_{H^\pm}^2 + m_{\rho'}^2)/v^2$ from the leading local features of the vacuum, that is, the masses.

2.7 CP type (i) symmetric model

Here we study the theories having a CP_g invariance of type (i) in detail, see theorem 2.5.5, (2.175).

2.7.1 Higgs potential and bosons

The most general potential with type (i) invariance is given in (2.199). Note that the condition $\boldsymbol{\xi} = \boldsymbol{\eta} = 0$ is basis independent. We choose a basis where E is diagonal with ordered diagonal entries, such that the parameters of the potential are

$$\xi_0\,, \qquad \boldsymbol{\xi} = 0\,, \tag{2.289a}$$

$$\eta_{00}\,, \qquad \boldsymbol{\eta} = 0\,, \qquad E = \mathrm{diag}(\mu_1, \mu_2, \mu_3) \quad \text{with } \mu_1 \geq \mu_2 \geq \mu_3\,. \tag{2.289b}$$

Obviously this model has not only the required CP_g type (i) symmetry, but it is also automatically invariant under at least three CP_g symmetries of type (ii), namely R_1, R_2, R_3, and infinitely many if two or three eigenvalues of E coincide. Furthermore, the potential is invariant under $\varphi_1 \longrightarrow -\varphi_1$. Note that for the parameterisation used in (2.262) this corresponds to $m_{11}^2 = m_{22}^2,\ m_{12}^2 = 0,\ \lambda_1 = \lambda_2,\ \lambda_5 \in \mathbb{R}$. We choose to use the orbit space parameters (2.289) in the following discussion, the translation to the parameters of (2.262) is very simple using (2.264).

The potential is stable in the strong sense if and only if

$$\eta_{00} > 0\,, \qquad \eta_{00} + \mu_i > 0\,, \qquad i = 1, 2, 3\,. \tag{2.290}$$

The existence of a non-trivial minimum necessary for a vacuum with the correct EWSB requires

$$\xi_0 < 0\,. \tag{2.291}$$

This condition implies immediately that the potential can not be stable in the weak or marginal sense, since for increasing fields in a V_4-flat direction the potential would decrease due to (2.291) beyond any lower bound. Formulated in terms of theorem 2.2.1, we have for the function $g(u) = \xi_0 < 0$, hence no weak or marginal stability.

A global minimum with the required EWSB is given by the stationary point

$$\tilde{\boldsymbol{K}} = \frac{-\xi_0}{2(\eta_{00} + \mu_3)} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \text{with } u_v = -\mu_3 \tag{2.292}$$

if and only if

$$\mu_3 \leq 0\,. \tag{2.293}$$

Otherwise the global minimum is given by a stationary point with full EWSB. Exactly for $\mu_2 \neq \mu_3$ and $\mu_3 < 0$ the point (2.292) is a unique global minimum up to a discrete ambiguity

in the sign of K_3. If $\mu_2 = \mu_3$ the point (2.292) belongs to a continuous degenerate set of stationary points on the light cone. Any of these points may be brought into the form (2.292) by a basis change without modifying any parameter of the potential. Absence of zero-mass charged Higgses requires the strict inequality $\mu_3 < 0$. If $\mu_3 = 0$ the point (2.292) belongs to a degenerate set of stationary points extending into the inner part of the light cone. Obviously a stable vacuum (2.292) necessarily breaks a CP_g type (i) spontaneously. However, at least one CP_g symmetry of type (ii) is respected by the vacuum.

We suppose now (2.293) holds such that (2.292) is a stable vacuum and discuss the consequences for the physical Higgs boson. Comparison of (2.292) with (2.110) gives

$$v^2 = \frac{-\xi_0}{\eta_{00} + \mu_3}\,. \tag{2.294}$$

The mass squared of the charged Higgs particles is, according to (2.131),

$$m_{H^\pm}^2 = \frac{2\mu_3\xi_0}{\eta_{00} + \mu_3}\,. \tag{2.295}$$

The mass matrix squared of the neutral Higgs particles (2.130) is diagonal already with our basis choice and we find for the mass squares

$$\begin{aligned} m_{\rho'}^2 &= 2(-\xi_0)\,, \\ m_{h'}^2 &= 2\,v^2\,(\mu_1 - \mu_3)\,, \\ m_{h''}^2 &= 2\,v^2\,(\mu_2 - \mu_3)\,. \end{aligned} \tag{2.296}$$

In the following we shall require that none of the neutral physical Higgs particles is massless and that there is no mass degeneracy between h' and h''. This implies from (2.296) the condition

$$\mu_1 > \mu_2 > \mu_3 \tag{2.297}$$

which is slightly stricter than (2.289).

Our Higgs potential has five parameters ξ_0, η_{00}, μ_1, μ_2, μ_3. We now express these in terms of the five physical quantities v^2, $m_{H^\pm}^2$, $m_{\rho'}^2$, $m_{h'}^2$, $m_{h''}^2$. This gives

$$\begin{aligned} \xi_0 &= -\frac{1}{2}m_{\rho'}^2\,, \\ \eta_{00} &= \frac{1}{2v^2}\left(m_{H^\pm}^2 + m_{\rho'}^2\right)\,, \\ \mu_1 &= \frac{1}{2v^2}\left(m_{h'}^2 - m_{H^\pm}^2\right)\,, \\ \mu_2 &= \frac{1}{2v^2}\left(m_{h''}^2 - m_{H^\pm}^2\right)\,, \\ \mu_3 &= -\frac{1}{2v^2}m_{H^\pm}^2\,. \end{aligned} \tag{2.298}$$

For positive squared masses and $m^2_{h'} > m^2_{h''}$ the conditions (2.290), (2.291) and (2.293) are always satisfied.

Let us next discuss the CP symmetries of our model and the CP transformation properties of the vacuum expectation values and of the physical fields. The Higgs Lagrangian (2.3) with the potential (2.199) where we require (2.297) to hold emerged essentially from the basis independent CP_g type (i) symmetry requirement. But this Higgs Lagrangian is automatically invariant under exactly three more CP_g type (ii) symmetries. For all of these four CP_g symmetries the gauge potentials are transformed according to (2.141a)-(2.141b). But the transformation of the Higgs fields and of the orbit variables $K_0(x)$, $\boldsymbol{K}(x)$ is different. All four CP_g symmetries are unique at the level of the orbit variables. However, at the level of the Higgs fields a global phase remains arbitrary for each of the Higgs CP transformations. Such a global phase factors drop out in the Higgs potential, and for the Yukawa terms they may always be absorbed by proper redefinitions of the fermion fields, as will be explained in the next section. Therefore, without loss of generality, we may set these global phase factors to specific values, as we shall indeed do for ease of notation. Moreover, we shall give names to all four CP_g symmetries.

We choose the global phase in the type (i) CP_g transformation, see (2.149) respectively (2.170), and set

$$\begin{aligned} \mathrm{CP}_g^{(i)}: \qquad & \varphi_i(x) \longrightarrow \epsilon_{ij}\varphi_j^*(x')\,, \\ & \varphi_1(x) \longrightarrow \quad \varphi_2^*(x')\,, \\ & \varphi_2(x) \longrightarrow -\varphi_1^*(x')\,, \end{aligned} \tag{2.299}$$

$$\begin{aligned} & K_0(x) \longrightarrow K_0(x')\,, \\ & \boldsymbol{K}(x) \longrightarrow -\boldsymbol{K}(x')\,. \end{aligned} \tag{2.300}$$

Note that $\mathrm{CP}_g^{(i)}$ is basis invariant. The form of the other three CP_g symmetries, $\mathrm{CP}_{g,1}^{(ii)}$, $\mathrm{CP}_{g,2}^{(ii)}$ and $\mathrm{CP}_{g,3}^{(ii)}$, depends on the basis. Here we continue to use the convenient basis choice (2.289). For the type (ii) transformation $\mathrm{CP}_{g,1}^{(ii)}$ we set

$$\begin{aligned} \mathrm{CP}_{g,1}^{(ii)}: \qquad & \varphi_i(x) \longrightarrow \sigma^3_{ij}\varphi_j^*(x')\,, \\ & \varphi_1(x) \longrightarrow \quad \varphi_1^*(x')\,, \\ & \varphi_2(x) \longrightarrow -\varphi_2^*(x')\,, \end{aligned} \tag{2.301}$$

$$\begin{aligned} & K_0(x) \longrightarrow K_0(x')\,, \\ & \boldsymbol{K}(x) \longrightarrow R_1\,\boldsymbol{K}(x') \end{aligned} \tag{2.302}$$

with R_1 being the matrix of the reflection on the 2–3 plane, see (2.169). The type (ii) transformation $\mathrm{CP}_{g,2}^{(ii)}$ is defined as the standard CP transformation, CP_s, for the Higgs

CP_g	U^φ
$\mathrm{CP}_g^{(i)}$	ϵ
$\mathrm{CP}_{g,1}^{(ii)}$	σ^3
$\mathrm{CP}_{g,2}^{(ii)}$	$\mathbb{1}_2$
$\mathrm{CP}_{g,3}^{(ii)}$	σ^1

Table 2.4: The matrices U^φ in (2.141c) for the four CP_g transformations.

fields, where

$$\begin{aligned} \mathrm{CP}_{g,2}^{(ii)}: \qquad & \varphi_i(x) \longrightarrow \varphi_j^*(x')\,, \\ & \varphi_1(x) \longrightarrow \varphi_1^*(x')\,, \\ & \varphi_2(x) \longrightarrow \varphi_2^*(x')\,, \qquad (2.303) \\ & K_0(x) \longrightarrow K_0(x')\,, \\ & \boldsymbol{K}(x) \longrightarrow R_2\,\boldsymbol{K}(x') \qquad (2.304) \end{aligned}$$

with R_2 being the matrix of the reflection on the 1–3 plane, see (2.169). Finally, the transformation $\mathrm{CP}_{g,3}^{(ii)}$ is defined by

$$\begin{aligned} \mathrm{CP}_{g,3}^{(ii)}: \qquad & \varphi_i(x) \longrightarrow \sigma^1_{ij}\varphi_j^*(x')\,, \\ & \varphi_1(x) \longrightarrow \varphi_2^*(x')\,, \\ & \varphi_2(x) \longrightarrow \varphi_1^*(x')\,, \qquad (2.305) \\ & K_0(x) \longrightarrow K_0(x')\,, \\ & \boldsymbol{K}(x) \longrightarrow R_3\,\boldsymbol{K}(x') \qquad (2.306) \end{aligned}$$

with R_3 being the matrix of the reflection on the 1–2 plane, see (2.169). Table 2.4 summarises the four CP_g transformations by listing their Higgs flavour mixing matrices U^φ for the generic CP_g transformations defined in (2.141c). The potential with (2.289) is invariant under all of these four CP_g transformations, $\mathrm{CP}_g^{(i)}$, $\mathrm{CP}_{g,1}^{(ii)}$, $\mathrm{CP}_{g,2}^{(ii)}$ and $\mathrm{CP}_{g,3}^{(ii)}$.

Note that the symmetries $\mathrm{CP}_g^{(i)}$, $\mathrm{CP}_{g,1}^{(ii)}$, $\mathrm{CP}_{g,2}^{(ii)}$ and $\mathrm{CP}_{g,3}^{(ii)}$ are not independent since we have at the level of the transformation of the Higgs fields the relation

$$\mathrm{CP}_g^{(i)} = \mathrm{CP}_{g,1}^{(ii)} \circ \mathrm{CP}_{g,2}^{(ii)} \circ \mathrm{CP}_{g,3}^{(ii)}\,. \qquad (2.307)$$

This result was already mentioned in the general discussion of CP_g symmetries.

Any of the above CP symmetries is conserved by the vacuum if and only if the vacuum value $\boldsymbol{K}$ satisfies

$$\bar{R}\,\boldsymbol{K} = \boldsymbol{K}\,. \qquad (2.308)$$

Here we have to insert $\bar{R} = -\mathbb{1}_3$, R_1, R_2 and R_3 for the symmetries $\mathrm{CP}_g^{(i)}$, $\mathrm{CP}_{g,1}^{(ii)}$ $\mathrm{CP}_{g,2}^{(ii)}$ and $\mathrm{CP}_{g,3}^{(ii)}$, respectively. For the vacuum solution $\tilde{\boldsymbol{K}}$ as in (2.292) we immediately get

$$\begin{aligned} (-\mathbb{1}_3)\boldsymbol{K} &\neq \boldsymbol{K}\,, \\ R_1\boldsymbol{K} &= \boldsymbol{K}\,, \\ R_2\boldsymbol{K} &= \boldsymbol{K}\,, \\ R_3\boldsymbol{K} &\neq \boldsymbol{K}\,. \end{aligned} \tag{2.309}$$

Thus, the symmetries $\mathrm{CP}_g^{(i)}$ and $\mathrm{CP}_{g,3}^{(ii)}$ are spontaneously broken. On the other hand, the symmetries $\mathrm{CP}_{g,1}^{(ii)}$ and $\mathrm{CP}_{g,2}^{(ii)}$ are conserved by the vacuum.

Now we come to the CP transformation properties of the physical Higgs fields defined in (2.119). Under $\mathrm{CP}_g^{(i)}$ which transforms the Higgs doublets according to (2.299) the physical Higgs fields have no definite transformation property. This is alright, since $\mathrm{CP}_g^{(i)}$ is spontaneously broken. For the unbroken symmetry $\mathrm{CP}_{g,1}^{(ii)}$ we get from (2.119) and (2.301)

$$\begin{aligned} \mathrm{CP}_{g,1}^{(ii)}\ : \qquad \rho'(x) &\longrightarrow \rho'(x')\,, \\ h'(x) &\longrightarrow -h'(x')\,, \\ h''(x) &\longrightarrow h''(x')\,, \\ H^+(x) &\longrightarrow -H^-(x')\,. \end{aligned} \tag{2.310}$$

On the other hand, we obtain from (2.119) and (2.303) for the $\mathrm{CP}_{g,2}^{(ii)}$ symmetry

$$\begin{aligned} \mathrm{CP}_{g,2}^{(ii)}\ : \qquad \rho'(x) &\longrightarrow \rho'(x')\,, \\ h'(x) &\longrightarrow h'(x')\,, \\ h''(x) &\longrightarrow -h''(x')\,, \\ H^+(x) &\longrightarrow H^-(x')\,. \end{aligned} \tag{2.311}$$

We see that the field ρ' is $\mathrm{CP}_{g,1}^{(ii)}$ and $\mathrm{CP}_{g,2}^{(ii)}$ even, h' is $\mathrm{CP}_{g,1}^{(ii)}$ odd and h'' is $\mathrm{CP}_{g,1}^{(ii)}$ even. This role of h' and h'' is interchanged for the symmetry $\mathrm{CP}_{g,2}^{(ii)}$, see (2.311). We note, however, that this assignment of $\mathrm{CP}_g^{(ii)}$ quantum numbers is to some extent a convention since we could have inserted extra global factors of (-1) in (2.301) and (2.303). This would not change the transformation properties of the orbit variables in (2.302) and (2.304) and thus have no physical consequence.

2.7.2 Maximal CP_g invariance

We saw that a $\mathrm{CP}_g^{(i)}$-symmetric Higgs potential is automatically invariant under three more CP_g symmetries. It is interesting to study under which conditions all of these symmetries can be extended to the full Lagrangian. We shall call the simultaneous invariance of the full Lagrangian under all four CP_g transformations *maximal CP_g invariance*. We already know

from theorem 2.5.12 that non-vanishing $\mathrm{CP}_g^{(i)}$-symmetric Yukawa couplings requires at least two generations of fermions. In this section we shall treat the case of two families where, for definiteness, we consider the families 2 and 3. It is convenient to use matrix notation for the Yukawa couplings. The 2×2 matrices $\lambda^{l\,j} = (\lambda^{l\,j}_{ik})$, $\lambda^{d'\,j} = (\lambda^{d'\,j}_{ik})$, $\lambda^{u\,j} = (\lambda^{u\,j}_{ik})$ have, to start with, arbitrary complex entries.

Without changing the physical content of the theory we can make $U(2)$-rotations of the right handed fields l_i^R, $d_i'^R$ u_i^R and the left handed doublet fields L_i^L and Q_i^L. As in the SM we can use this to require, without loss of generality, for the couplings to the first Higgs doublet certain standard forms:

$$\lambda^{l\,1} = \begin{pmatrix} \lambda^{l\,1}_{22} & 0 \\ 0 & \lambda^{l\,1}_{33} \end{pmatrix}, \qquad \lambda^{l\,1}_{33} \geq \lambda^{l\,1}_{22} \geq 0\,, \tag{2.312a}$$

$$\lambda^{d\,1} = V \begin{pmatrix} \lambda^{d\,1}_{22} & 0 \\ 0 & \lambda^{d\,1}_{33} \end{pmatrix} V^\dagger\,, \qquad \tilde\lambda^{d\,1}_{33} \geq \tilde\lambda^{d\,1}_{22} \geq 0\,, \tag{2.312b}$$

$$\lambda^{u\,1} = \begin{pmatrix} \lambda^{u\,1}_{22} & 0 \\ 0 & \lambda^{u\,1}_{33} \end{pmatrix}, \qquad \lambda^{u\,1}_{33} \geq \lambda^{u\,1}_{22} \geq 0 \tag{2.312c}$$

with

$$V = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix}, \qquad 0 \leq \vartheta \leq \pi/2\,. \tag{2.312d}$$

For the derivation of the corresponding results in the SM see, for instance, chapter 22.4 of [64]. The matrix $V = (V_{ij})$ in (2.312d) will turn out to be the Cabibbo-Kobayashi-Maskawa (CKM) matrix in the 2–3 sector. As we shall see, in the basis of the fermion fields defined by (2.312) the fields l_i^R, l_i^L and u_i^R, u_i^L correspond to mass eigenstates. For the d'-fields defined in this basis the mass eigenstates will be

$$d_i^{R,L}(x) = V_{ij}^\dagger\, d_j'^{R,L}(x)\,. \tag{2.313}$$

In the following we shall always work in the fermion basis defined by (2.312) if not stated otherwise. In this basis, the usual idenfication of the fermions fields given in table 2.1 holds.

We shall construct mixing matrices matrices U^l, U^L for each of the four U^φ such that $\mathscr{L}_{\mathrm{Yuk},l}$ is invariant. We recall that each $U \in U^l, U^L$ must be either a symmetric or an antisymmetric unitary matrix. For their general parameterisation we use the scheme

$$U = e^{i\alpha} \begin{pmatrix} a & ib \\ ib & a^* \end{pmatrix} = U^{\mathrm{T}}\,, \qquad \alpha \in \mathbb{R},\ a \in \mathbb{C},\ b \geq 0,\ |a|^2 + b^2 = 1 \tag{2.314}$$

for the symmetric case and

$$U = e^{i\alpha} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -U^{\mathrm{T}}\,, \quad \alpha \in \mathbb{R} \tag{2.315}$$

for the antisymmetric case. Furthermore, we take into account that for each CP_g symmetry the total number, including U^φ, of antisymmetric mixing matrices must be 0 or 2. The necessary and sufficient conditions for CP_g invariance of the Yukawa terms (2.230) read in matrix notation

$$\lambda^{l\,j} = U^{\varphi}_{jj'}\, U^{L}\, \lambda^{l\,j'\,*}\, U^{l\,\dagger}\,, \tag{2.316}$$

$$\lambda^{d'\,j} = U^{\varphi}_{jj'}\, U^{Q}\, \lambda^{l\,j'\,*}\, U^{d'\,\dagger}\,, \tag{2.317}$$

$$\lambda^{u\,j} = U^{\varphi\,*}_{jj'}\, U^{Q}\, \lambda^{l\,j'\,*}\, U^{u\,\dagger}\,. \tag{2.318}$$

2.7.3 Invariant couplings for two lepton families

Now we consider extending all four CP_g symmetries of the Higgs sector to the Yukawa interaction (2.5). We want to find out what this implies for the coupling matrices $\lambda^{l\,j}$, $\lambda^{d'\,j}$ and $\lambda^{u\,j}$. We start by considering only the leptonic part of $\mathscr{L}_{\mathrm{Yuk}}$ in (2.5),

$$\mathscr{L}_{\mathrm{Yuk},\,l}(x) = -\lambda^{l\,j}_{ik}\; \bar{L}^{L}_{i}(x)\, \varphi_j(x)\, l^{R}_{k}(x) + h.c. \tag{2.319}$$

As explained in section 2.7.2 we can, without loss of generality, suppose (2.312a) to hold. Furthermore, the diagonal forms for $\lambda^{l\,1}$ (2.312a) still allows to redefine l^{R}_{i} and L^{L}_{i} for given index i by multiplication with an arbitrary phase factor. We use this freedom to require, without loss of generality,

$$\lambda^{l\,2}_{23} > 0 \qquad \text{if} \quad \lambda^{l\,2}_{23} \neq 0\,, \tag{2.320a}$$

$$\lambda^{l\,2}_{32} \geq 0 \qquad \text{if} \quad \lambda^{l\,2}_{23} = 0\,. \tag{2.320b}$$

Two massive leptons with different masses

Let us first consider the case

$$\lambda^{l\,1}_{22} > 0\,, \quad \lambda^{l\,1}_{33} > 0\,, \quad \lambda^{l\,1}_{22} \neq \lambda^{l\,1}_{33}\,. \tag{2.321}$$

This corresponds to non-vanishing and unequal masses for the leptons l_2 and l_3 after EWSB, since only φ_1 acquires a non-vanishing vev. By considering all possible cases under the above assumptions we find that a $\mathrm{CP}^{(ii)}_{g,2}$ symmetry is possible if and only if

$$\lambda^{l\,2} = \lambda^{l\,2\,*}\,. \tag{2.322}$$

The corresponding matrices U^L, U^l must be equal, $U^L = U^l$, symmetric and diagonal. Choosing them equal to the unit matrix up to a global phase factor, $U^L = U^l \propto \mathbb{1}$, is sufficient always, and necessary if at least one off-diagonal element of $\lambda^{l\,2}$ is non-vanishing. A $\mathrm{CP}^{(ii)}_{g,1}$ symmetry is possible in addition if and only if

$$\lambda^{l\,2} = \begin{pmatrix} 0 & \lambda^{l\,2}_{23} \\ \lambda^{l\,2}_{23} & 0 \end{pmatrix} = \lambda^{l\,2\,*}\,. \tag{2.323}$$

The matrices U^L, U^l for this symmetry must be equal, $U^L = U^l$, symmetric and diagonal again. The choice $U^L = U^l \propto \sigma^3$ is sufficient always, and necessary if at least one off-diagonal element of $\lambda^{l\,2}$ is non-vanishing. On top of that a $\mathrm{CP}_g^{(i)}$ symmetry can be constructed if and only if

$$\lambda^{l\,2} = \begin{pmatrix} 0 & \lambda^{l\,1}_{33} \\ \pm\lambda^{l\,1}_{22} & 0 \end{pmatrix}, \tag{2.324a}$$

$$\text{or} \quad \lambda^{l\,2} = \begin{pmatrix} 0 & \lambda^{l\,1}_{22} \\ \pm\lambda^{l\,1}_{33} & 0 \end{pmatrix}. \tag{2.324b}$$

For the two cases (2.324a) we have $U^L = e^{i\alpha}\epsilon$, $\alpha \in \mathbb{R}$, and $U^l = e^{i\alpha}\sigma^3$ (first case) or $U^l = -e^{i\alpha_L}\mathbb{1}$ (second case). For the two cases (2.324b) we have $U^l = e^{i\alpha}\epsilon$, $\alpha \in \mathbb{R}$, and $U^L = e^{i\alpha}\sigma^3$ (first case) or $U^L = e^{i\alpha_L}\mathbb{1}$ (second case). The remaining $\mathrm{CP}_{g,3}^{(ii)}$ transformation may then always be constructed according to

$$\mathrm{CP}_{g,3}^{(ii)} = \mathrm{CP}_{g,1}^{(ii)} \circ \mathrm{CP}_{g,2}^{(ii)} \circ \mathrm{CP}_g^{(i)}\,, \tag{2.325}$$

under which $\mathscr{L}_{\mathrm{Yuk},\,l}$ is automatically invariant, such that no additional restrictions on $\lambda^{l\,2}$ arise. We find that this is possible since the right hand side of this equation is of the form (2.141), see subsection 2.5.1, (2.325) holds for the already fixed transformations in the Higgs sector and the fermion mixings $U = U^L$, U^l defined by (2.325) are unitary matrices with $U\,U^* = \pm 1$ for the case (2.324). Note that we explicitly checked the last two of these three statements for the present case. In particular, the last of the three statements follows from the above given restrictions on the fermion mixings in $\mathrm{CP}_{g,1}^{(ii)}$, $\mathrm{CP}_{g,2}^{(ii)}$ and $\mathrm{CP}_g^{(i)}$ for non-diagonal $\lambda^{l\,2}$ as in (2.324), and not from their generic $\mathrm{CP}_g^{(i)}$, $\mathrm{CP}_g^{(ii)}$ properties already, see subsection 2.5.1.

To see the physical consequences of this result we look at $\mathscr{L}_{\mathrm{Yuk},l}$ after EWSB. Inserting for the Higgs fields the physical expressions (2.119) we get from (2.319)

$$\begin{aligned}
\mathscr{L}_{\mathrm{Yuk},l} = & -m_{l2}\,\bar{l}_2\,l_2 - m_{l3}\,\bar{l}_3\,l_3 \\
& - \tfrac{1}{\sqrt{2}}\rho'\left[\lambda^{l\,1}_{22}\,\bar{l}_2\,l_2 + \lambda^{l\,1}_{33}\,\bar{l}_3\,l_3\right] \\
& - H^+\left[\lambda^{l\,2}_{23}\,\bar{\nu}_2\omega_R l_3 + \lambda^{l\,2}_{32}\,\bar{\nu}_3\omega_R l_2\right] - H^-\left[\lambda^{l\,2}_{23}\,\bar{l}_3\omega_L\nu_2 + \lambda^{l\,2}_{32}\,\bar{l}_2\omega_L\nu_3\right] \\
& - \tfrac{1}{\sqrt{2}}h'\left[\bar{l}_3\left(\lambda^{l\,2}_{32}\omega_R + \lambda^{l\,2}_{23}\omega_L\right) l_2 + \bar{l}_2\left(\lambda^{l\,2}_{32}\omega_L + \lambda^{l\,2}_{23}\omega_R\right) l_3\right] \\
& - \tfrac{i}{\sqrt{2}}h''\left[\bar{l}_3\left(\lambda^{l\,2}_{32}\omega_R - \lambda^{l\,2}_{23}\omega_L\right) l_2 + \bar{l}_2\left(-\lambda^{l\,2}_{32}\omega_L + \lambda^{l\,2}_{23}\omega_R\right) l_3\right]
\end{aligned} \tag{2.326}$$

with the chirality projectors

$$\omega_R := \frac{1+\gamma_5}{2} \quad \text{and} \quad \omega_L := \frac{1-\gamma_5}{2} \tag{2.327}$$

and the lepton masses given by

$$\begin{aligned}
m_{l\,2} &= \lambda^{l\,1}_{22}\,\frac{v}{\sqrt{2}}\,, \\
m_{l\,3} &= \lambda^{l\,1}_{33}\,\frac{v}{\sqrt{2}}\,.
\end{aligned} \tag{2.328}$$

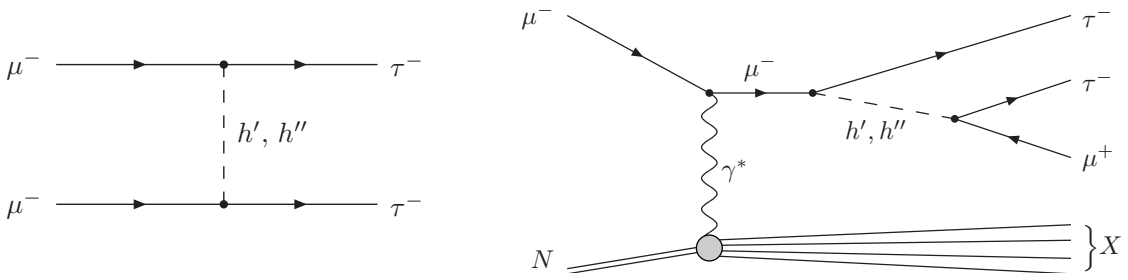

Figure 2.6: *Left*: Two Feynman diagrams for the large FCNC process $\mu^-\mu^- \longrightarrow \tau^-\tau^-$ reflecting the last two contributions in the Lagrangian (2.326) for the case l_2 and l_3 are identified with μ and τ respectively. *Right*: Two Feynman diagrams for the deep inelastic muon-nucleon scattering process which would reveal FCNCs corresponding to the same couplings as in the left diagrams.

Identifying the leptons 2 and 3 with the μ and τ leptons respectively we see that for all cases in (2.324) either $\left|\lambda^{l\,2}_{23}\right| = m_\tau\sqrt{2}/v$ or $\left|\lambda^{l\,2}_{32}\right| = m_\tau\sqrt{2}/v$. Thus (2.326) always contains large lepton flavour changing neutral currents. These would allow processes like

$$\mu^- + \mu^- \longrightarrow \tau^- + \tau^- \tag{2.329}$$

through diagrams like in figure 2.6 (left). A direct study of this process would be a topic for a muon collider which, however, is a project for the far future. But the couplings in figure 2.6 (left) would also lead to spectacular lepton-flavour-violating events in deep inelastic muon-nucleon scattering,

$$\mu^- + N \longrightarrow \mu^+ + \tau^- + \tau^- + X\,. \tag{2.330}$$

Two of the corresponding tree-level Feynman diagrams are shown in figure 2.6 (right). Here X stands for the hadronic final state. Since such FCNCs were never observed we consider them to be disfavoured phenomenologically.

Two massive leptons with equal mass

The next case to study is

$$\lambda^{l\,1}_{22} = \lambda^{l\,1}_{33} > 0\,. \tag{2.331}$$

Considering $\mathrm{CP}^{(i)}_g$ symmetries we find that a realisation requires the matrix $\lambda^{l\,2}/\lambda^{l\,1}_{22}$ to be unitary. This matrix may then be diagonalised with a single unitary matrix (for both the left and the right handed fields), such that $\lambda^{l\,1}$ is unchanged. In the new basis $\mathrm{CP}^{(i)}_g$ symmetry requires

$$\lambda^{l\,2} = e^{i\eta}\lambda^{l\,1}_{22}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \qquad \eta \in \mathbb{R}\,. \tag{2.332}$$

Moreover, this form of $\lambda^{l\,2}$ is also sufficient for a $\mathrm{CP}^{(i)}_g$ symmetry to exist. The mixings must be either $U^L = e^{i\alpha}\epsilon$ and $U^l = -e^{i(\alpha-\eta)}\sigma^1$ or $U^l = e^{i\alpha}\epsilon$ and $U^L = -e^{i(\alpha+\eta)}\sigma^1$, with

$\alpha \in \mathbb{R}$. We find an additional $\mathrm{CP}^{(ii)}_{g,2}$ symmetry to be possible if and only if

$$\lambda^{l\,2} = \pm i \lambda^{l\,1}_{22} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \text{or} \tag{2.333a}$$

$$\lambda^{l\,2} = \pm \lambda^{l\,1}_{22} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.333b}$$

For (2.333a) we must have either $U^L = U^l \propto \sigma^1$ or $U^L = U^l \propto \epsilon$. For (2.333b) U^L and U^l must be equal, $U^L = U^l$, symmetric and diagonal. On top of that a $\mathrm{CP}^{(ii)}_{g,1}$ symmetry may be constructed without further restriction on $\lambda^{l\,2}$. Here, for (2.333a) U^L and U^l must be equal, $U^L = U^l$, symmetric and diagonal. For (2.333b) U^L and U^l we must have either $U^L = U^l \propto \sigma^1$ or $U^L = U^l \propto \epsilon$. By checking all cases we find that the remaining $\mathrm{CP}^{(ii)}_{g,3}$ symmetry may be constructed according to (2.325) without further restriction on $\lambda^{2\,l}$, since the fermion mixings described above guarantee $U\,U^* = \pm \mathbb{1}$ for $U = U^L,\ U^l$.

We see from (2.312a) and (2.333) that there are no mixings between the two lepton families. Furthermore, the two Higgs doublets do not mix for our choice of basis, see subsection 2.7.1. From this we conclude that there are no FCNCs at the tree-level. However, here we have, according to (2.328), equal lepton masses, $m_{l\,2} = m_{l\,3}$, which is unacceptable phenomenologically.

One massless, one massive lepton

It remains to be studied what happens for the case of one massless and one massive lepton. Taking, by convention, l_3 to be the massive lepton we have the case

$$\lambda^{l\,1}_{22} = 0\,, \qquad \lambda^{l\,1}_{33} > 0\,. \tag{2.334}$$

In this case (2.312a) and (2.320) still leave an additional freedom in the choice of field phase factors, which we employ to require

$$\lambda^{l\,2}_{22} \geq 0 \tag{2.335}$$

in the following. This means no loss of generality, since we may first choose a basis such that (2.312a) holds, then redefine the phase of l^R_2 to achieve (2.335) and finally arrange for (2.320) in the previously described way. We find that a $\mathrm{CP}^{(i)}_{g}$ symmetry may be realised for our conventions if and only if

$$\lambda^{l\,2} = \lambda^{l\,1}_{33} \begin{pmatrix} \cos\eta & 0 \\ \sin\eta & 0 \end{pmatrix}, \qquad 0 \leq \eta < \pi/2\,, \qquad \text{or} \tag{2.336a}$$

$$\lambda^{l\,2} = \lambda^{l\,1}_{33} \begin{pmatrix} \cos\eta & \sin\eta \\ 0 & 0 \end{pmatrix}, \qquad 0 \leq \eta < \pi/2\,. \tag{2.336b}$$

An antisymmetric U^L together with a symmetric U^l are possible if (2.336a) holds and the second column of U^L equals the first column of $\lambda^{l\,2}/\lambda^{l\,1}_{33}$ up to a phase. An antisymmetric

U^l together with a symmetric U^L are possible if (2.336b) holds and the second column of $U^{l\,*}$ equals the first row of $\lambda^{l\,2}/\lambda^{l\,1}_{33}$ up to a phase. There are no other options. In addition, a $\mathrm{CP}^{(ii)}_{g,1}$ and a $\mathrm{CP}^{(ii)}_{g,2}$ symmetry can be constructed without further restrictions on $\lambda^{l\,2}$. For each of these two symmetries, the corresponding U^L, U^l must both be symmetric and diagonal. We find that a $\mathrm{CP}^{(ii)}_{g,3}$ symmetry may be defined without further restrictions on the couplings. However, here not all possible mixings in $\mathrm{CP}^{(ii)}_{g,1}$, $\mathrm{CP}^{(ii)}_{g,2}$, and $\mathrm{CP}^{(i)}_{g}$ define via (2.325) a valid $\mathrm{CP}^{(ii)}_{g,3}$ transformation with $U\,U^* = \pm 1$ for $U = U^L, U^l$. Nevertheless choosing all symmetries such that (2.325) holds for the lepton sector is possible without additional restrictions on $\lambda^{l\,2}$.

We see that (2.336) may contain an off-diagonal term which leads to FCNCs, which we consider to be phenomenologically disfavoured. Requiring absence of those FCNCs means $\eta = 0$ in (2.336) and leads to a highly symmetric form of the leptonic Yukawa terms with only one free parameter (fixed by v and $m_{l\,3}$). However, it is also clear from the above discussion, that we can *prescribe* the lepton mixings in the $\mathrm{CP}^{(i)}_{g}$ transformation in such a way, that an invariant $\lambda^{l\,2}$ is required to be diagonal and thus FCNCs are absent. Explicitly, choosing both U^L and U^l off-diagonal for that symmetry requires

$$\lambda^{l\,2} = \begin{pmatrix} \lambda^{l\,1}_{33} & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.337}$$

This phenomenologically particularly interesting case will be discussed in subsection 2.7.5 in more detail.

2.7.4 Invariant couplings for two quark families

In this section we study the quark part of the Lagrangian (2.5). Let us first look at the term which generates masses for the u-type quarks,

$$\mathscr{L}_{\mathrm{Yuk},u} = -\lambda^{u\,j}_{ik}\,\bar{Q}^L_i \epsilon \varphi^*_j u^R_k + h.c. \tag{2.338}$$

Here we can suppose without loss of generality that $\lambda^{u\,1}$ is as in (2.312c). As for the case of the leptons in subsection 2.7.3 we ask if $\mathscr{L}_{\mathrm{Yuk},u}$ in (2.338) allows to implement all four CP_g symmetries from the Higgs sector. Since all U^φ matrices are real, the invariance condition (2.318) for the present case is completely analogous to (2.316) in the lepton sector. We can immediately conclude from the results of the previous section that for the case of two massive up-type quarks with unequal masses the symmetry requirement leads to large FCNCs. Here it is important to note that these FCNCs are generated for the physical mass eigenstates u_2 and u_3. Requiring absence of these FCNCs allows then only two possibilities for a non-zero coupling $\mathscr{L}_{\mathrm{Yuk},q'}$. Either we must have non-zero equal masses for the quarks u_2 and u_3 or we must have u_2 massless and u_3 massive. Discarding the former for phenomenological reasons we are left with the case of a massless $u_2 = c$ quark and a massive $u_3 = t$ quark.

We turn next to the term in (2.5) which generates masses for d-type quarks

$$\mathscr{L}_{\mathrm{Yuk},q} = -\lambda^{d\,j}_{ik}\,\bar{Q}^L_i \varphi_j d'^R_k + h.c. \tag{2.339}$$

Here the standard form for $\lambda^{d\,1}$ is given in (2.312b) and (2.312d). Note that d'_β are – in general – not the mass eigenstates. We shall change to the basis of d_α mass according to (2.313) by setting

$$\begin{aligned} d_i^R &= V_{ii'}^\dagger \, d_{i'}'^R \,, \\ \tilde{Q}_i^L &= V_{ii'}^\dagger \, Q_{i'}^L \,. \end{aligned} \tag{2.340}$$

The coupling term (2.339) reads now

$$\mathscr{L}_{\mathrm{Yuk},d} = -\tilde{\lambda}_{ik}^{d\,j} \, \bar{\tilde{Q}}_i^L \varphi_j d_k^R + h.c. \tag{2.341}$$

where

$$\lambda^{d\,j} = V^\dagger \, \lambda^{d'\,j} \, V \,, \quad (j = 1, 2) \,. \tag{2.342}$$

From (2.312b) we see that this implies

$$\tilde{\lambda}^{d\,j} = \begin{pmatrix} \lambda_{22}^{d\,j} & 0 \\ 0 & \lambda_{33}^{d\,j} \end{pmatrix} . \tag{2.343}$$

Now we can proceed as for the lepton case. For the case of two massive down-type quarks with unequal masses, the maximal CP invariance requirement implies large FCNCs. among the physical d-quark mass eigenfields. These FCNCs can only be avoided if we require either equal masses $m_{d2} = m_{d3}$ or $m_{d2} = 0$ and $m_{d3} \neq 0$. Again we discard the former possibility for phenomenological reasons.

What we have discussed so far were only separate requirements for each of the up- and down-type quark sectors. We now take into account the connection of the latter by asking how a CP_g symmetry can actually be implemented for both sectors simultaneously. This will of course involve the CKM matrix. We consider the phenomenologically particular interesting case of one massless quark pair (u_2, d_2) and one massive pair (u_3, d_3). That is, we suppose

$$\begin{aligned} \lambda_{22}^{u\,1} &= 0 \,, & \lambda_{33}^{u\,1} &> 0 \,, \\ \tilde{\lambda}_{22}^{d\,1} &= 0 \,, & \tilde{\lambda}_{33}^{d\,1} &> 0 \,. \end{aligned} \tag{2.344}$$

Similar to the corresponding lepton case we choose, without loss of generality, a basis with

$$\begin{aligned} \lambda_{22}^{u\,2} &\geq 0 \,, \\ \tilde{\lambda}_{22}^{d\,2} &\geq 0 \,, \end{aligned} \tag{2.345}$$

and find that the requirement of maximal CP invariance together with the requirement of absence of FCNCs leads to the following structure of the quark-Higgs coupling matrices (see (2.337)):

$$\lambda^{u\,2} = \begin{pmatrix} \lambda_{33}^{u\,1} & 0 \\ 0 & 0 \end{pmatrix} , \tag{2.346}$$

$$\tilde{\lambda}^{d\,2} = \begin{pmatrix} \lambda_{33}^{d\,1} & 0 \\ 0 & 0 \end{pmatrix} . \tag{2.347}$$

Here $\tilde{\lambda}^{d\,2}$ is the CKM-rotated matrices according to (2.342). From the discussion of the lepton case we see that $\mathrm{CP}_g^{(i)}$ invariance is implementable for (2.346) and (2.347) only with certain matrices U^u, U^Q in (2.338) and certain CKM rotated matrices $\tilde{U}^d$, $\tilde{U}^Q$ for the CKM rotated fields in (2.341). Here we have according to (2.151b)

$$\tilde{U}^d = V^\dagger\, U^d\, V^* \,, \tag{2.348}$$

$$\tilde{U}^Q = V^\dagger\, U^Q\, V^* \,. \tag{2.349}$$

Now we consider matrices U^u, U^Q and $\tilde{U}^d$, $\tilde{U}^Q$ for each of the possible cases described after (2.336) and check if we can fulfil (2.349). For the case that both U^u and $\tilde{U}^d$ are antisymmetric we find the two solutions

$$U^Q = \tilde{U}^Q = e^{i\xi}\sigma^1 \qquad U^u = \tilde{U}^d = e^{i\xi}\epsilon \,, \qquad U^d = U^u \,, \qquad \vartheta = 0 \,, \qquad V = \mathbb{1} \,, \tag{2.350}$$

$$U^Q = -\tilde{U}^Q = e^{i\xi}\sigma^1 \qquad U^u = -\tilde{U}^d = e^{i\xi}\epsilon \,, \qquad U^d = -U^u \,, \qquad \vartheta = \pi/2 \,, \qquad V = \epsilon \,, \tag{2.351}$$

that is, in both cases the 2–3 sector of the CKM matrix is fixed. For the case that both U^L and $\tilde{U}^L$ are antisymmetric we find

$$U^Q = \tilde{U}^Q = e^{i\xi}\epsilon \,, \qquad U^u = \tilde{U}^d = e^{i\xi}\sigma^1 \,, \qquad U^d = e^{i\xi} \begin{pmatrix} \sin 2\vartheta & \cos 2\vartheta \\ \cos 2\vartheta & -\sin 2\vartheta \end{pmatrix} \tag{2.352}$$

without restriction on V. We find that there are no other solutions. Thus we see that, strictly speaking, the $\mathrm{CP}_g^{(i)}$ invariance plus absence of FCNCs gives no restriction on the angle ϑ in the 2–3 sector of the CKM matrix. But perhaps we can argue that also the right handed quarks $u_{\alpha\,R}$, $d'_{\alpha\,R}$ should belong to some multiplet of a bigger gauge group as would be possible in grand unified scenarios. Then a natural requirement could be $U^u = U^d$. From (2.350)-(2.352) we see that we have then only the solution $V = \mathbb{1}_2$.

2.7.5 Mass hierarchies and FCNC suppression via CP invariances

In this section we will present a model defined by a specific set of CP_g symmetries. These symmetries lead to a highly symmetric form of the Lagrangian, produce mass hierarchies and forbid FCNCs. Let us first recapitulate what we found in the previous sections. In subsection 2.7.1 we have discussed the Higgs sector of the model which is characterised by the requirement of $\mathrm{CP}_g^{(i)}$ invariance. We have seen that this leads automatically to three more CP_g invariances, $\mathrm{CP}_{g,1}^{(ii)}$, $\mathrm{CP}_{g,2}^{(ii)}$ and $\mathrm{CP}_{g,3}^{(ii)}$. Two of them, $\mathrm{CP}_g^{(i)}$ and $\mathrm{CP}_{g,3}^{(ii)}$, are broken spontaneously along with the electroweak symmetry. At tree-level, which we have discussed above, the symmetries $\mathrm{CP}_{g,1}^{(ii)}$ and $\mathrm{CP}_{g,2}^{(ii)}$ are unbroken. We know from theorem 2.5.12 that requiring a single family to have a $\mathrm{CP}_g^{(i)}$ invariant coupling leads necessarily to the coupling being identically zero. Thus, as the minimal non-trivial case, we discussed the extension of the four CP_g symmetries to non-zero Yukawa couplings for two families in the subsections 2.7.3 and 2.7.4. We chose the two families to be the second

CP_g	U^{φ}	U^{ψ_L}	U^{ψ_R}
$\mathrm{CP}_g^{(i)}$	ϵ	ϵ	$-\sigma^1$
$\mathrm{CP}_{g,1}^{(ii)}$	σ^3	σ^3	$-\mathbb{1}_2$
$\mathrm{CP}_{g,2}^{(ii)}$	$\mathbb{1}_2$	$\mathbb{1}_2$	$\mathbb{1}_2$
$\mathrm{CP}_{g,3}^{(ii)}$	σ^1	σ^1	σ^1

Table 2.5: Flavour mixing matrices for four CP_g transformations (2.141). U^{φ} describes the mixing of the two Higgs doublets, U^{ψ_L} and U^{ψ_R} the mixings of two fermion generations. Invariance under all of them with $U^l = U^d = U^u = U^{\psi_R}$ and $U^L = U^Q = U^{\psi_L}$ leads to mass hierarchies and absence of FCNCs, see (2.334).

and the third. We found that unequal non-zero masses for the leptons l_2 and l_3, the quarks u_2 and u_3, as well as d_2 and d_3 always implied large FCNCs. The absence of large FCNCs required either equal masses of corresponding fermions ($m_{l2} = m_{l3}$ etc.) or one fermion massless, the other massive. Discarding the equal mass case on phenomenological grounds we were, thus, left with the possibility

$$\begin{aligned} m_{l\,2} &= 0\,, & m_{l\,3} &\neq 0\,, \\ m_{d\,2} &= 0\,, & m_{d\,3} &\neq 0\,, \\ m_{u\,2} &= 0\,, & m_{u\,3} &\neq 0\,. \end{aligned} \tag{2.353}$$

Here we shall *prescribe* the form of the matrices $U_R^{(l)}$ and $U_L^{(l)}$ for the four CP_g transformations as shown in table 2.5 and require the Lagrangian to be invariant under them. Note that our set of CP_g transformations chosen here differs slightly from those described in [89], which will lead together with the other conventions to differences only in the signs of the resulting Yukawa couplings. In the present scheme, we chose the same mixing matrices in the CP_g transformations for all electroweak doublets, that is, the Higgs fields and the left-handed fermion fields. These transformations have the normalisation $(\mathrm{CP}_{g,1}^{(ii)})^2 = (\mathrm{CP}_{g,2}^{(ii)})^2 = (\mathrm{CP}_{g,3}^{(ii)})^2 = (\mathrm{CP}_s)^2$ and $(\mathrm{CP}_g^{(i)})^2 = \exp(i6\pi Y) \circ (\mathrm{CP}_s)^2$, see end of subsection 2.5.1. $\mathrm{CP}_{g,2}^{(ii)}$ is the standard CP transformation for all fields, $\mathrm{CP}_{g,2}^{(ii)} = \mathrm{CP}_s$. A direct calculation for the mixing matrices of each field type shows that

$$\mathrm{CP}_{g,3}^{(ii)} = \mathrm{CP}_g^{(i)} \circ \mathrm{CP}_{g,2}^{(ii)} \circ \mathrm{CP}_{g,1}^{(ii)} \tag{2.354}$$

for all fields. In this model the first family remains uncoupled to the Higgs fields. We require now invariance of the general non-trivial Yukawa interaction (2.319) under each of these four CP_g transformations, that is, we require (2.230) to hold with the corresponding mixing matrices from table 2.5. We choose a basis where (2.312) holds. We start by considering $\lambda^{d'\,1}$ and impose invariance under $\mathrm{CP}_{g,1}^{(ii)}$. This requires $V\,\lambda^{d'\,1}\,V^\dagger = \sigma^3\,V\,\lambda^{d'\,1}\,V^\dagger(-\mathbb{1})$, which can only be fulfilled with (2.312b) if the CKM matrix of the families 2 and 3 is

$$V = \mathbb{1}\,. \tag{2.355}$$

This means we are already in the mass basis for all fermions and may check the symmetry requirements for all coupling matrices in the same way. We successively evaluate the restrictions following from invariance under $\mathrm{CP}_g^{(i)}$, $\mathrm{CP}_{g,1}^{(ii)}$, $\mathrm{CP}_{g,2}^{(ii)}$. $\mathrm{CP}_{g,3}^{(ii)}$ needs not be checked, since it is a consequence of $\mathrm{CP}_g^{(i)}$, $\mathrm{CP}_{g,1}^{(ii)}$ and $\mathrm{CP}_{g,2}^{(ii)}$ symmetry (2.354). Our calculation shows that these four invariances *require* the Yukawa couplings

$$\begin{aligned}
\lambda^{l\,1} &= \begin{pmatrix} 0 & 0 \\ 0 & \lambda_{33}^{l\,1} \end{pmatrix}, \qquad \lambda^{l\,2} = \begin{pmatrix} \lambda_{33}^{l\,1} & 0 \\ 0 & 0 \end{pmatrix}, \\
\lambda^{d\,1} &= \begin{pmatrix} 0 & 0 \\ 0 & \lambda_{33}^{d\,1} \end{pmatrix}, \qquad \lambda^{d\,2} = \begin{pmatrix} \lambda_{33}^{d\,1} & 0 \\ 0 & 0 \end{pmatrix}, \\
\lambda^{u\,1} &= \begin{pmatrix} 0 & 0 \\ 0 & \lambda_{33}^{u\,1} \end{pmatrix}, \qquad \lambda^{u\,2} = \begin{pmatrix} \lambda_{33}^{u\,1} & 0 \\ 0 & 0 \end{pmatrix}
\end{aligned} \tag{2.356}$$

for the basis choice (2.312), (2.335), (2.345). This form guarantees the absence of large FCNCs and requires vanishing masses for family 2.

The full Yukawa part of the Lagrangian takes the simple form

$$\begin{aligned}
\mathscr{L}_{\mathrm{Yuk}} = &-\lambda_{33}^{l\,1}\,\left(\bar{L}_3^L\,\varphi_1 l_3^R + \bar{L}_2^L\,\varphi_2 l_2^R\right) \\
&-\lambda_{33}^{d\,1}\,\left(\bar{Q}_3^L\,\varphi_1 d_3^R + \bar{Q}_2^L\,\varphi_2 d_2^R\right) \\
&-\lambda_{33}^{u\,1}\,\left(\bar{Q}_3^L\,\epsilon\,\varphi_1^* u_3^R + \bar{Q}_2^L\,\epsilon\,\varphi_2^* u_2^R\right) + h.c.
\end{aligned} \tag{2.357}$$

Note the high degree of symmetry between the second and third family here. However, after EWSB we get, inserting (2.119) for the Higgs fields,

$$\begin{aligned}
&\mathscr{L}_{\mathrm{Yuk}} = \\
&-m_{l_3}\,\bar{l}_3 l_3 - \rho'\,\frac{m_{l_3}}{v}\,\bar{l}_3 l_3 - h'\,\frac{m_{l_3}}{v}\,\bar{l}_2 l_2 - h''\,\frac{i m_{l_3}}{v}\,\bar{l}_2\,\gamma_5\,l_2 - \left[H^+\frac{\sqrt{2}m_{l_3}}{v}\,\bar{\nu}_2\,\omega_R\,l_2 + h.c.\right] \\
&-m_{d_3}\,\bar{d}_3 d_3 - \rho'\,\frac{m_{d_3}}{v}\,\bar{d}_3 d_3 - h'\,\frac{m_{d_3}}{v}\,\bar{d}_2 d_2 - h''\,\frac{i m_{d_3}}{v}\,\bar{d}_2\,\gamma_5\,d_2 - \left[H^+\frac{\sqrt{2}m_{d_3}}{v}\,\bar{u}_2\,\omega_R\,d_2 + h.c.\right] \\
&-m_{u_3}\,\bar{u}_3 u_3 - \rho'\,\frac{m_{u_3}}{v}\,\bar{u}_3 u_3 - h'\,\frac{m_{u_3}}{v}\,\bar{u}_2 u_2 + h''\,\frac{i m_{u_3}}{v}\,\bar{u}_2\,\gamma_5\,u_2 + \left[H^+\frac{\sqrt{2}m_{u_3}}{v}\,\bar{u}_2\,\omega_L\,d_2 + h.c.\right]
\end{aligned} \tag{2.358}$$

with the fermion masses

$$\begin{aligned}
m_{l_3} &= \lambda_{33}^{l\,1}\frac{v}{\sqrt{2}} \equiv m_\tau\,, \\
m_{d_3} &= \lambda_{33}^{d\,1}\frac{v}{\sqrt{2}} \equiv m_b\,, \\
m_{u_3} &= \lambda_{33}^{u\,1}\frac{v}{\sqrt{2}} \equiv m_t\,.
\end{aligned} \tag{2.359}$$

The third generation fermions have become massive and couple to the physical ρ' Higgs. The second generation fermions stay massless but couple to h', h'' and the charged Higgses $H^\pm$.

We consider it noteworthy that our symmetry principles require more than one family. For two families we get in a natural way mass hierarchies. Choosing the simplest extension to three families we get masses unequal to zero only for τ, t and b whereas all other leptons and quarks, μ, e, c, u, s, d stay massless. In addition, the CKM matrix of the quarks equals the unit matrix, $V = \mathbb{1}$. Clearly, all this is not quite as it is observed in Nature. On the other hand, as a first approximation, it is not too bad. We have [7, 121, 122]

$$\begin{aligned} \frac{m_e}{m_\tau} &\approx 0.00029\,, & \frac{m_\mu}{m_\tau} &\approx 0.059\,, \\ \left.\frac{m_u}{m_t}\right|_v &\approx 9.9 \cdot 10^{-6}\,, & \left.\frac{m_c}{m_t}\right|_v &\approx 0.0036\,, \\ \left.\frac{m_d}{m_b}\right|_v &\approx 0.0010\,, & \left.\frac{m_s}{m_b}\right|_v &\approx 0.018\,. \end{aligned} \tag{2.360}$$

Here we have used for the quarks the $\overline{MS}$ masses at the renormalisation point $\mu = v \approx 246\,\mathrm{GeV}$ and $\alpha_s(m_Z) = 0.119$. Also, taking into account experimental measurements, the CKM matrix is not too far from unity. Indeed, one finds for the absolute values $|V_{ij}|$ (see [7])

$$\begin{pmatrix} |V_{11}| & |V_{12}| & |V_{13}| \\ |V_{21}| & |V_{22}| & |V_{23}| \\ |V_{31}| & |V_{32}| & |V_{33}| \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.227 & 0.004 \\ 0.227 & 0.973 & 0.042 \\ 0.008 & 0.042 & 0.999 \end{pmatrix}. \tag{2.361}$$

Note that in the 2–3 sector V is very close to the unit matrix. Without question, a good theory should be able to explain the experimental numbers in (2.360) and (2.361) in more detail. It remains to be seen, whether the presented model provides a starting point for this.

Summary

We have studied a Two-Higgs-Doublet Model where the scalar sector has four generalised CP symmetries. EWSB leads to the spontaneous breakdown of two of these CP symmetries. We have introduced the *principle of maximal CP invariance* requiring that these four CP symmetries can be extended to the full Lagrangian of the theory. We find that for a single fermion family this principle forbids a non-zero fermion-Higgs coupling. Thus, if we require massive fermions which arise from non-zero Yukawa couplings we need family replication. We have studied then in detail theories with two fermion families. Here, indeed, we can extend all four CP symmetries to the full Lagrangian with non-zero Yukawa couplings which are, however, highly constrained. Discarding extensions which enforce large FCNCs, we are left with the possibilities of either equal masses for each fermion type or large mass hierarchies between the families. Choosing the latter possibility we arrive at a theory with a high degree of symmetry between the two families and absence of FCNCs. We have shown that we can obtain this theory directly from a symmetry requirement. For this we have prescribed the form of the four CP transformations for the lepton and quark fields as shown in table 2.5. Our principle of maximal CP invariance leads then directly to the

Yukawa coupling (2.357) implying one massive and one massless family as well as absence of large FCNCs. The Yukawa part of this theory is given in (2.357) and after EWSB in (2.358). Through EWSB one family becomes massive the other stays massless at the tree-level, which we have discussed in this chapter. Adding a fermion family uncoupled to the Higgs particles we arrive at a model which gives a rough approximation of the fermion pattern observed in Nature. Concerning the charged fermions we have one massive family which we identify with the *third* one (τ, t, b) and two massless ones, which we identify with the *second* (μ, c, s) and the *first* (e, u, d) families. In this model the CKM matrix of the quark generations is equal to unity. As for any THDM, the spectrum of physical Higgs particles consists of three neutral scalars, ρ', h' and h'', and the charged Higgses $H^\pm$. The neutral Higgs particle ρ' – which has essentially the same properties as the SM Higgs – couples exclusively to the third family of fermions. The other Higgses h', h'' and $H^\pm$ couple exclusively to the second family of fermions. The first fermion family remains uncoupled to the Higgses. Of course, these statements are expected to be only approximately true. It remains to be seen whether such a mechanism is realised in Nature. The experiments at the LHC might be able to tell.

2.8 Orbit variables for the n-Higgs-Doublet Model

After our discussion of the general THDM in the previous chapters, we discuss here as an outlook a possible extension of the presented methods to the case of $n > 2$ Higgs doublets. We consider n complex Higgs-doublet fields

$$\varphi_i(x) = \begin{pmatrix} \varphi_i^+(x) \\ \varphi_i^0(x) \end{pmatrix}, \quad i = 1, \dots, n\,. \tag{2.362}$$

All doublets are supposed to have the same weak hypercharge $y = +1/2$. In analogy to (2.16) we introduce the matrix

$$\phi := \begin{pmatrix} \varphi_1^+ & \varphi_1^0 & 0 & \dots & 0 \\ \varphi_2^+ & \varphi_2^0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \varphi_n^+ & \varphi_n^0 & 0 & \dots & 0 \end{pmatrix}. \tag{2.363}$$

which is now a $n \times n$ matrix. As in (2.17) we define the matrix

$$\underline{K} := \phi\, \phi^\dagger \tag{2.364}$$

of all $SU(2)_L \times U(1)_Y$-invariant scalar products,

$$K_{ij} = \varphi_j^\dagger \varphi_i\,, \tag{2.365}$$

with $\underline{K} = (K_{ij})$, $i, j \in \{1, \dots, n\}$. A change of basis among the doublets means

$$\begin{aligned} \phi &\longrightarrow \phi' = U\,\phi\,, \\ \underline{K} &\longrightarrow \underline{K}' = U\,\underline{K}\,U^\dagger \end{aligned} \tag{2.366}$$

with a constant unitary matrix $U \in U(n)$. A gauge transformation from $SU(2)_L \times U(1)_Y$

$$\varphi_i'^{\alpha}(x) = (U_G(x))_{\alpha\beta}\, \varphi_i^{\beta}(x), \qquad i = 1, \ldots, n \tag{2.367}$$

with $U_G(x) \in U(2)$ means

$$\begin{aligned} \phi(x) &\longrightarrow \phi'(x) = \phi(x)\, \tilde{U}_G^{\mathrm{T}}(x)\,, \\ \underline{K}(x) &\longrightarrow \underline{K}'(x) = \underline{K}(x)\,, \end{aligned} \tag{2.368}$$

where $\tilde{U}_G(x) \in U(n)$ is block-diagonal

$$\tilde{U}_G(x) := \left(\begin{array}{c|c} U_G(x) & 0 \\ \hline 0 & \mathbb{1}_{n-2} \end{array}\right). \tag{2.369}$$

We see directly from the definitions (2.363) and (2.364) that $\underline{K}$ is hermitian, positive semidefinite and has rank ≤ 2, since

$$\operatorname{rank} \underline{K} = \operatorname{rank} \phi \tag{2.370}$$

and ϕ has rank ≤ 2. We shall show that these conditions are also sufficient for some matrix $\underline{K}$ to find corresponding Higgs fields with (2.365).

Suppose an otherwise arbitrary complex $n \times n$ matrix $\underline{K}$ is hermitian, positive semidefinite and has rank ≤ 2. We can diagonalise $\underline{K}$ and represent it as

$$\underline{K}(x) = W(x) \left(\begin{array}{cc|c} \kappa_1(x) & 0 & \\ 0 & \kappa_2(x) & 0 \\ \hline & 0 & 0 \end{array}\right) W^{\dagger}(x) \tag{2.371}$$

with $W(x) \in U(n)$ and $\kappa_1(x) \geq 0$, $\kappa_2(x) \geq 0$. We set now

$$\phi(x) = W(x) \left(\begin{array}{cc|c} \sqrt{\kappa_1(x)} & 0 & \\ 0 & \sqrt{\kappa_2(x)} & 0 \\ \hline & 0 & 0 \end{array}\right) \tag{2.372}$$

and see easily that $\phi(x)$ is of the form (2.363) and satisfies (2.365). Thus to any hermitian and positive semidefinite matrix $\underline{K}(x)$ of rank ≤ 2 there is at least one field configuration of the n Higgs doublets such that (2.364) holds. Furthermore, a direct calculation shows that any two field configurations (2.363) which give by (2.364) the same matrix $\underline{K}$ differ at most by a electroweak gauge transformation. We summarise our findings in a theorem.

Theorem 2.8.1. *For n Higgs doublet fields of the same weak hypercharge $y = +1/2$ the matrix $\underline{K}(x) = \left(\varphi_j^{\dagger}(x)\varphi_i(x)\right)$ is a positive semidefinite $n \times n$ matrix of rank ≤ 2. For any hermitian positive semidefinite $n \times n$ matrix $\underline{K}(x)$ of rank ≤ 2 there are Higgs fields such that (2.364) holds. Any two field configurations giving the same matrix $\underline{K}(x)$ are related by a $SU(2)_L \times U(1)_Y$ gauge transformation. The matrices $\underline{K}(x)$ form, therefore, the space of the gauge orbits of the n Higgs-doublet fields.*

In order to discuss a Higgs potential in terms of the gauge invariant matrix $\underline{K}$ or some decomposition of it, we need a more explicit characterisation of the domain of $\underline{K}$. In order to be incorporated in a Lagrange type minimisation of the potential, the domain should be characterised by some set of equalities and inequalities. If the constraints are polynomial in the used degrees of freedom, they might be used in the Gröbner basis approach explained in chapter 3. For example, the eigenvalues of $\underline{K}$ contain the relevant information, but their usage is complicated by the fact that no explicit formulae exist for them in the general case. Instead we shall consider the principal minors of $\underline{K}$. It is well known that a hermitian matrix is positive definite if and only if all leading principal minors are positive. Since the corresponding necessary and sufficient criterion for positive *semi*-definiteness is omitted in the text books we checked, we insert a brief intermezzo of some basic linear algebra here.

For some complex $n\times n$ matrix H one may consider $r\times s$ submatrices ($r, s \in \{1, \ldots, n\}$) obtained from H by deleting arbitrary $n-r$ rows and arbitrary $n-s$ columns. The determinants of these submatrices are called *minors*. The determinants of $r\times r$ submatrices obtained by deletion of $n-r$ rows and columns with the same indices are called *principal minors* or, when referring to a specific dimension $r \times r$ of the submatrices, principal r-minors. Finally, if a principal minor originates from a $r\times r$ submatrix of H, where the *last* $n-r$ rows and columns were deleted, it is called a *leading principal minor*. With these notations[1] fixed we state the following.

Theorem 2.8.2. *For a hermitian $n\times n$ matrix H the following is equivalent:*

(i) H is positive semi-definite,

(ii) all principal minors of H are non-negative.

Please note that in contrast to the criterion for positive definiteness here *all* principal minors are involved, not only the leading ones. We do not give a detailed proof here, but only outline the idea. For (i) $\Rightarrow$ (ii) we assume that H is positive semi-definite but has at least one negative principal minor, which is obtained from some $r\times r$ submatrix $\hat{H}$. Then $\hat{H}$ has a negative eigenvalue and there is a vector $\hat{v}$ with r components such that $\hat{v}^{\mathrm{T}}\hat{H}\hat{v} < 0$. From $\hat{v}$ a vector v with n components may be constructed by adding zero components such that $v^{\mathrm{T}}Hv < 0$, which contradicts the assumption. For (ii) $\Rightarrow$ (i) we construct a matrix $\tilde{H}$ by permutating same index rows and columns of H such that the upper left block submatrix of $\tilde{H}$ has the rank of H. A diagonalisation of the $\tilde{H}$ shows that this upper left block can be written as $B\,\hat{D}\,B^{\dagger}$, where $\hat{D}$ is a diagonal matrix with all non-zero eigenvalues of H and B is invertible. All principal minors of $B\,\hat{D}\,B^{\dagger}$ must be positive and using the proof of the well known theorem for positive matrices we conclude that $B\,\hat{D}\,B^{\dagger}$ is positive definite. Therefore also $\hat{D}$ must be positive which implies (i). Please see [123] for details omitted here, in some cases a generalisation of their proofs from symmetric to hermitian matrices is necessary.

[1] Another common nomenclature deviates from these definitions by using the name principal minor for the above definition of a leading principal minor.

Using theorem 2.8.2 we find that for some hermitian $n \times n$ matrix $\underline{K}$ there are Higgs fields with (2.364) if and only if all of the following hold:

$$\text{all principal 1-minors of } \underline{K} \text{ are non-negative, and} \tag{2.373a}$$
$$\text{all principal 2-minors of } \underline{K} \text{ are non-negative, and} \tag{2.373b}$$
$$\text{all principal 3-minors of } \underline{K} \text{ are zero.} \tag{2.373c}$$

Since the minors of $\underline{K}$ also allow to directly read off the rank of $\underline{K}$ we can characterise the possible cases as follows.

Theorem 2.8.3. *A hermitian $n \times n$ matrix $\underline{K}$ corresponds to a field configuration of n Higgs doublets with $\underline{K}(x) = \left(\varphi_j^\dagger(x)\varphi_i(x)\right)$ if and only if one of the following holds:*

- *All principal 1-minors of $\underline{K}$ are zero.*
  Then $\operatorname{rank} K = 0$ *and* $\underline{K} = 0$ *hold.*
- *All principal 2-minors of $\underline{K}$ are zero, at least one 1-minor is positive.*
  Then $\operatorname{rank} \underline{K} = 1$ *holds.*
- *All principal 3-minors of $\underline{K}$ are zero, at least one 2-minor is positive.*
  Then $\operatorname{rank} \underline{K} = 2$ *holds.*

As for the THDM, the rank of $\underline{K}$ has a direct physical meaning. We observe from (2.363) and (2.370) the following.

Classification 2.8.4. *A matrix $\langle \underline{K} \rangle$ describing the vacuum of a potential belongs to one of the following categories:*

- $\operatorname{rank} \langle \underline{K} \rangle = 0$*: A vacuum $\langle \underline{K} \rangle$ of this type is trivial, that is, it has vanishing vacuum expectation values for all Higgs fields. In this case, no symmetry is spontaneously broken.*
- $\operatorname{rank} \langle \underline{K} \rangle = 1$*: For a vacuum $\langle \underline{K} \rangle$ of this type a $SU(2)_L \times U(1)_Y$ transformation is possible such that*

$$\langle \phi \rangle = \begin{pmatrix} 0 & |\langle \varphi_1^0 \rangle| & 0 & \dots & 0 \\ 0 & \langle \varphi_2^0 \rangle & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \langle \varphi_n^0 \rangle & 0 & \dots & 0 \end{pmatrix} \tag{2.374}$$

  and we identify the unbroken $U(1)$ gauge group with the electromagnetic one. By a transformation (2.366) we can arrange that

$$\langle \phi' \rangle = \begin{pmatrix} 0 & v/\sqrt{2} & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \qquad v = \sqrt{2} \left(\sum_{i=1}^{n} \left| \varphi_i^0 \right|^2 \right)^{1/2} \tag{2.375}$$

for the vacuum expectation values in the new basis. Again, $v \approx 246\,GeV$ is the usual Higgs-field vacuum expectation value.

- $\operatorname{rank}\langle \underline{K} \rangle = 2$*: A vacuum $\underline{K}$ of this type means that in any gauge or Higgs flavour basis at least one charged and one neutral vacuum expectation value is non-vanishing. This means that the full gauge group $SU(2)_L \times U(1)_Y$ is broken.*

We now consider a decomposition of the matrix $\underline{K}(x)$ such that we arrive at real degrees of freedom. Let

$$\lambda_a/2\,, \qquad a = 1, \ldots, n^2 - 1 \tag{2.376}$$

be generators of $SU(n)$, that is hermitian and traceless $n \times n$ matrices, with the normalisation

$$\operatorname{tr} \lambda_a \lambda_b = 2\delta_{ab}\,, \qquad a, b \in \{1, \ldots, n^2 - 1\}\,. \tag{2.377}$$

We arrive at a complete set of complex $n \times n$ matrices by supplementing these generators by a multiple of the unit matrix

$$\lambda_0 := N_0 \sqrt{\frac{2}{n}}\, \mathbb{1}_n \tag{2.378}$$

with yet to be specified normalisation $N_0 \in \mathbb{R}$. The matrix $\underline{K}$ can be decomposed in the following way

$$\underline{K} = K_0 \frac{\lambda_0}{2} + \sum_{a=1}^{n^2-1} K_a \frac{\lambda_a}{2}\,. \tag{2.379}$$

The coefficients $K_0,\ K_a$ $(a = 1, \ldots, n^2 - 1)$ defined in this way are real and satisfy

$$\begin{aligned} K_0 &= \frac{1}{N_0^2} \operatorname{tr}\left(\underline{K}\lambda_0\right)\,, \\ K_a &= \operatorname{tr}\left(\underline{K}\lambda_a\right)\,, \qquad a = 1, \ldots, n^2 - 1\,. \end{aligned} \tag{2.380}$$

These are the orbit variables for the n-Higgs-Doublet Model. Although setting $N_0 = 1$ would appear most natural at this point, we keep N_0 unfixed for a reason which will become clear. From a group theoretical point of view this construction means with respect to the $SU(n)_\varphi$ subgroup of the Higgs flavour transformations $U(n)_\varphi$:

$$n \otimes n = 1 \oplus (n^2 - 1)\,, \tag{2.381}$$

where K_0 is the singlet 1 and the vector (K_a) the fundamental multiplet $(n^2 - 1)$. That is, for a Higgs flavour transformation (2.366) the doublets transform as "spinors", K_0 as an invariant "scalar" and (K_a) as a real "vector" (with the associated real rotation in $SO(n)$).

As an example we consider the case of three Higgs doublets, $n = 3$. As generators of $SU(3)$ satisfying (2.377) we have the Gell-Mann matrices

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{pmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{2.382}$$

The minor conditions (2.373) are

$$\begin{aligned} &K_{11} \geq 0\,, && K_{11}K_{22} - |K_{12}|^2 \geq 0\,, \\ &K_{22} \geq 0\,, && K_{22}K_{33} - |K_{23}|^2 \geq 0\,, \\ &K_{33} \geq 0\,, && K_{33}K_{11} - |K_{31}|^2 \geq 0\,, && \det \underline{K} = 0\,. \end{aligned} \tag{2.383}$$

Taking the sum of the principal 1-minor inequalities gives $\operatorname{tr} \underline{K} \geq 0$, which in turn is equivalent to

$$K_0 \geq 0\,. \tag{2.384}$$

From the sum of the principal 2-minor inequalities we get

$$2N_0^2\, K_0^2 - \sum_{a=1}^{8} K_a^2 \geq 0\,. \tag{2.385}$$

Here we see $N_0 = 1/\sqrt{2}$ would render (2.385) to the simplest generalisation of the "forward light cone" restriction. Note that the constraints in (2.383) are written in a basis dependent form. As basis independent constraints we find that at least (2.384), (2.385) and $\det \underline{K} = 0$ must be taken into account, where the latter is a polynomial constraint of third order in K_a ($a = 1, \ldots, 8$). This is in contrast to the THDM, where the analogues of (2.384) and (2.385) were sufficient for the specification of the domain of the orbit variables.

To summarise, we have presented with theorem 2.8.3 the domain of gauge invariant functions for the n-Higgs-Doublet Model with $n > 2$, described by a hermitian matrix. These conditions may be directly translated to the singlet plus vector type real orbit variables (2.379), where they become polynomial inequalities and equalities. We have illustrated the case $n = 3$ and found the domain of the orbit variables to be more involved than in the case of the THDM. Nevertheless, to determine the stability and EWSB of potentials for $n > 2$ doublets, Lagrange methods similar as described for the THDM may be applied.

Chapter 3

The Next-to-Minimal Supersymmetric Standard Model

3.1 Higgs potential

The Next-to-Minimal Supersymmetric Standard Model Higgs sector contains a complex scalar Higgs singlet S in addition to the two scalar Higgs doublets H_u, H_d present in the MSSM, see subsection 1.4.3. The two Higgs doublets have opposite hypercharges, $y = 1/2$ for H_u and $y = -1/2$ for H_d. We assume the soft breaking mass terms for the sleptons and squarks to be sufficiently large to prevent nonvanishing vevs for the latter. For the discussion of the electroweak symmetry breaking at the tree-level, we then need to consider only the scalar potential of the pure Higgs part. We denote the scalar Higgs fields by

$$H_u = \begin{pmatrix} H_u^+ \\ H_u^0 \end{pmatrix}, \qquad H_d = \begin{pmatrix} H_d^0 \\ H_d^- \end{pmatrix}, \qquad S\,. \tag{3.1}$$

For the scalar doublets H_u, H_d, we get contributions to the Lagrangian from the elimination of the vector supermultiplet auxiliary fields, the so-called D-terms,

$$\mathscr{L}_{\mathrm{D}}^{H} = -\frac{1}{2} g^2 \sum_{a=1}^{3} (H_u^\dagger \mathbf{T}_a H_u + H_d^\dagger \mathbf{T}_a H_d)^2 - \frac{1}{2} g'^2 (H_u^\dagger \mathbf{Y} H_u + H_d^\dagger \mathbf{Y} H_d)^2\,. \tag{3.2}$$

Contributions for all scalar Higgs fields arise from the elimination of the chiral supermultiplet auxiliary fields, the so-called F-terms,

$$\mathscr{L}_{\mathrm{F}}^{H} = -\left|\frac{\delta W^H}{\delta H_u^+}\right|^2 - \left|\frac{\delta W^H}{\delta H_u^0}\right|^2 - \left|\frac{\delta W^H}{\delta H_d^0}\right|^2 - \left|\frac{\delta W^H}{\delta H_d^-}\right|^2 - \left|\frac{\delta W^H}{\delta S}\right|^2 \tag{3.3}$$

where W^H (1.23) is taken as a functional of the scalar fields instead of the superfields. Using the identity $\sum_{a=1}^{3} \sigma_{ij}^a \sigma_{kl}^a + \delta_{ij}\delta_{kl} = 2\delta_{il}\delta_{jk}$ we find with $V_{\mathrm{D}} = -\mathscr{L}_{\mathrm{D}}^{H}$, $V_{\mathrm{F}} = -\mathscr{L}_{\mathrm{F}}^{H}$,

$V_{\rm soft} = -\mathscr{L}^H_{\rm soft}$ and (1.24) for the Higgs potential (see also [32, 100])

$$V = V_{\rm D} + V_{\rm F} + V_{\rm soft}$$

where

$$\begin{aligned} V_{\rm D} &= \frac{1}{8}\bar{g}^2 \left(|H_u|^2 - |H_d|^2\right)^2 + \frac{1}{2}g^2 \left|H_d^\dagger H_u\right|^2 \\ V_{\rm F} &= |\lambda|^2 |S|^2 \left(|H_u|^2 + |H_d|^2\right) + \left|\lambda\, H_u^{\rm T} \epsilon H_d + \kappa\, S^2\right|^2 \\ V_{\rm soft} &= m_{H_u}^2 |H_u|^2 + m_{H_d}^2 |H_d|^2 + m_S^2 |S|^2 + \left(\lambda A_\lambda\, S H_u^{\rm T} \epsilon H_d + \frac{1}{3}\kappa A_\kappa\, S^3 + h.c.\right) \end{aligned} \tag{3.4}$$

Here, λ, κ are complex dimensionless couplings, $m_{H_u}^2$, $m_{H_d}^2$, m_S^2 are real parameters of dimension mass squared and A_λ, A_κ are complex parameters of dimension mass. Further, we have $H_u^{\rm T}\epsilon H_d = H_u^+ H_d^- - H_u^0 H_d^0$ and $\bar{g} = \sqrt{g^2 + g'^2}$, where g and g' are the $SU(2)_L$ and $U(1)_Y$ gauge couplings, respectively. Thus, the parameters of the potential are given by the experimentally fixed electroweak gauge couplings and

$$\lambda,\, \kappa,\, m_{H_u}^2,\, m_{H_d}^2,\, m_S^2,\, A_\lambda,\, A_\kappa\,. \tag{3.5}$$

The quartic terms of the potential (3.4) are positive for any non-trivial field configuration, if both λ, κ are non-vanishing. The potential is therefore bounded from below for all cases considered here, and stability has not to be checked any further.

3.2 Physical parameters and necessary symmetry breaking conditions

The vacuum is required to break the electroweak symmetry $SU(2)_L \times U(1)_Y$ down to $U(1)_{em}$ as observed in Nature. In this case, we can always achieve by a $SU(2)_L \times U(1)_Y$ transformation that the vevs take the form

$$\langle H_u \rangle = e^{i\varphi_u} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} v_u \end{pmatrix}, \qquad \langle H_d \rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} v_d \\ 0 \end{pmatrix}, \qquad \langle S \rangle = \frac{1}{\sqrt{2}} e^{i\varphi_S} v_S \tag{3.6}$$

with the real parameters v_u, v_d, v_S and the real phases φ_u and φ_S. Since in the NMSSM up- and down-type fermions acquire their masses through the vacuum expectation values of H_u and H_d respectively, we require $v_u, v_d \neq 0$. Also on phenomenological grounds, we assume $v_S \neq 0$ in order to have a non-vanishing effective μ term, see subsection 1.4.3. At this point, we merely assume (3.6) to be the global minimum of the potential and postpone the question for which parameters of the potential this is indeed the case. Necessary conditions will be derived successively in this section. Our method presented in section 3.4 allows then for a definite answer.

The stationarity of the vacuum requires a vanishing gradient with respect to the Higgs fields. We define

$$\begin{aligned}
\mathcal{R} &:= \mathrm{Re}\left(\lambda\kappa^* e^{i(\varphi_u - 2\varphi_S)}\right), & \mathcal{I} &:= \mathrm{Im}\left(\lambda\kappa^* e^{i(\varphi_u - 2\varphi_S)}\right), \\
\mathcal{R}_\lambda &:= \frac{1}{\sqrt{2}}\,\mathrm{Re}\left(\lambda A_\lambda e^{i(\varphi_u + \varphi_S)}\right), & \mathcal{I}_\lambda &:= \frac{1}{\sqrt{2}}\,\mathrm{Im}\left(\lambda A_\lambda e^{i(\varphi_u + \varphi_S)}\right), \\
\mathcal{R}_\kappa &:= \frac{1}{\sqrt{2}}\,\mathrm{Re}\left(\kappa A_\kappa e^{i\,3\varphi_S}\right), & \mathcal{I}_\kappa &:= \frac{1}{\sqrt{2}}\,\mathrm{Im}\left(\kappa A_\kappa e^{i\,3\varphi_S}\right),
\end{aligned} \tag{3.7}$$

and find the stationarity conditions for (3.6) to be equivalent to

$$m_{H_u}^2 = -\frac{1}{8}\bar{g}^2(v_u^2 - v_d^2) - \frac{1}{2}|\lambda|^2(v_S^2 + v_d^2) + \frac{1}{2}\mathcal{R}\frac{v_S^2 v_d}{v_u} + \mathcal{R}_\lambda\frac{v_S v_d}{v_u}, \tag{3.8a}$$

$$m_{H_d}^2 = \frac{1}{8}\bar{g}^2(v_u^2 - v_d^2) - \frac{1}{2}|\lambda|^2(v_S^2 + v_u^2) + \frac{1}{2}\mathcal{R}\frac{v_S^2 v_u}{v_d} + \mathcal{R}_\lambda\frac{v_S v_u}{v_d}, \tag{3.8b}$$

$$m_S^2 = -|\kappa|^2 v_S^2 - \frac{1}{2}|\lambda|^2(v_u^2 + v_d^2) + \mathcal{R}v_u v_d + \mathcal{R}_\lambda\frac{v_u v_d}{v_S} - \mathcal{R}_\kappa v_S, \tag{3.8c}$$

$$\mathcal{I}_\lambda = -\frac{1}{2}v_S\mathcal{I}, \tag{3.8d}$$

$$\mathcal{I}_\kappa = -\frac{3}{2}\frac{v_u v_d}{v_S}\mathcal{I}. \tag{3.8e}$$

We shall use these equations to reparameterise the potential in the following.

In order to access the physical content of the Lagrangian, we define the shifted Higgs fields as deviations from their vacuum expectation values. Introducing new linear combinations of the shifted Higgs fields, we can isolate the gauge boson mass terms (along with the Goldstone modes). The electroweak coupling structure is complicated when working with $H_u,\ H_d$ due to their opposite hypercharges. Instead, we change to the same hypercharge convention using the substitution (2.6), where the couplings to the gauge bosons are more transparent. In this notation, the separation of the gauge boson masses becomes really simple, it reduces to a simple basis change to decouple the vacuum expectation values, see section 2.4. Inserting the rotated fields into the potential, we get for the non-interaction part of the charged Higgses

$$V_{\mathrm{mass}}^c = m_{H^\pm}^2\, H^+ H^- \tag{3.9}$$

with

$$m_{H^\pm}^2 = m_W^2 - \frac{1}{2}|\lambda|^2 v^2 + \frac{1}{s_{2\beta}}v_S(v_S\mathcal{R} + 2\mathcal{R}_\lambda), \tag{3.10}$$

showing that the charged Higgses $H^\pm$ are already mass eigenstates with mass squares $m_{H^\pm}^2$. Here, the W boson mass is identified according to (1.6) and appears as a consequence of supersymmetry. Further, $v \equiv \sqrt{v_u^2 + v_d^2} \approx 246\,\mathrm{GeV}$ is the electroweak scale, $\tan\beta \equiv v_u/v_d$ and $s_{2\beta} \equiv \sin(2\beta)$. In addition to the fields $\rho',\ h',\ h''$ there are two more real modes from the complex singlet S. These five neutral Higgses combine to five physical neutral Higgs bosons, which are for generic values of the potential parameters all massive.

Usually, stationarity conditions similar to (3.8) are used to reparameterise the potential. We employ such a reparameterisation in section 3.5. By construction it is then assured, that the required vacuum solution (3.6) is a stationary point. Further, choosing all physical masses positive would guarantee that the required vacuum is a local minimum of the potential. However, whether the required vacuum is the global minimum of the potential and thus stable, can not be decided on its local properties alone. A rigorous check is possible by considering all stationary points of the potential and test whether their values of the potential lie indeed not below that of the required vacuum, for which we get

$$\langle V\rangle = -\frac{1}{8}c_{2\beta}^2 v^2 m_Z^2 + \frac{1}{8}s_{2\beta}^2 v^2 (m_{H^\pm}^2 - m_W^2) + \frac{1}{4}v_S^2\left(-|\lambda|^2 v^2 + \frac{1}{2}s_{2\beta}\mathcal{R}v^2 - |\kappa|^2 v_S^2 - \frac{2}{3}\mathcal{R}_\kappa v_S\right). \tag{3.11}$$

We shall present a method to find all solutions of the stationarity conditions in section 3.4. Here, we want to discuss some particularly simple stationary points.

The trivial point of vanishing fields,

$$H_u = H_d = S = 0 \tag{3.12}$$

is always a stationary point with the potential value

$$V|_{triv} = 0\,. \tag{3.13}$$

Requiring the required vacuum to lie not above the trivial point in the potential gives

$$\langle V\rangle \leq V|_{triv} = 0 \tag{3.14}$$

$$\Leftrightarrow \quad m_{H^\pm}^2 \leq \cot^2(2\beta) + m_W^2 + \frac{2}{s_{2\beta}^2}v_S^2\left(|\lambda|^2 - \frac{1}{2}s_{2\beta}\mathcal{R} + |\kappa|^2\frac{v_S^2}{v^2} + \frac{2}{3}\mathcal{R}_\kappa\frac{v_S}{v^2}\right). \tag{3.15}$$

This result was also found in (39) of [109][1].

Let us next consider a specific example for the CP conserving case, where all parameters of the potential as well as the vacuum expectation values in (3.6) are real. We find a particular stationary solution, where both doublets vanish but the singlet does not,

$$H_u = H_d = 0, \quad S = \frac{\sqrt{A_\kappa^2 - 8m_S^2} - A_\kappa}{4\kappa^2} \tag{3.16}$$

where we assume $A_\kappa^2 > 8m_S^2$. The value of the potential at this stationary point is

$$V_{\rm nb} = -\frac{\left(A_\kappa - \sqrt{A_\kappa^2 - 8m_S^2}\right)^2\left(A_\kappa^2 - A_\kappa\sqrt{A_\kappa^2 - 8m_S^2} - 12m_S^2\right)}{384\kappa^2}\,. \tag{3.17}$$

A stable vacuum with the vevs (3.6) requires then

$$\langle V\rangle \leq V_{\rm nb}\,, \tag{3.18}$$

[1] Their notation differ from ours by a sign in the definition of κ.

which implies further restrictions on the parameters. Moreover, the stationary point (3.16) has an interesting property. We reparameterise the potential in terms of the vacuum expectation values as described above and consider the hierarchy between this point and the required vacuum (3.6). After reparameterisations we find as an approximation in the case $v_S \gg v, m_{H^\pm}, A_\kappa$ and κ of $\mathcal{O}(1)$ for the relative splitting of the potential values

$$\lim_{v_S\to\infty} \frac{V_{\rm nb} - \langle V\rangle}{\langle V\rangle} = -\frac{2(1+c_{4\beta})\bar{g}^2\kappa^2 + \lambda^2(-3(-1+c_{4\beta})\kappa^2 - 8\lambda^2 + 8\kappa\lambda s_{2\beta})}{32\kappa^4} \cdot \frac{v^4}{v_S^4}\,. \quad (3.19)$$

Considering the infimum and supremum of this expression in the range

$$\kappa \in [0.6, 1]\,, \quad \lambda \in]0, 1]\,, \quad \tan\beta \in [0, \infty[\,, \quad (3.20)$$

we find

$$-\,0.2 \cdot \frac{v^4}{v_S^4} \leq \lim_{v_S\to\infty} \frac{V_{\rm nb} - \langle V\rangle}{\langle V\rangle} \leq 1.7 \cdot \frac{v^4}{v_S^4}\,. \quad (3.21)$$

This means that at large v_S the value of the potential at the stationary point (3.16) comes very close to that of the required vacuum (3.6). Note that the two points are always well separated due to their doublet components, which are vanishing in the one case and given in terms of the electroweak scale in the other case. The stationary points are typically connected by a long, flat, but very narrow valley, as our numerical evaluations will show. This feature is in fact much more generic than considered here, see section 3.5.

Note concerning the stability of the vacuum, that the presented conditions are only necessary but not sufficient. Commonly encountered checks in the literature consider only stationary points, where at least one of H_u, H_d or S vanishes. In section 3.5, we shall demonstrate at the tree-level that these partial checks are insufficient.

3.3 Stationary points via orbit variables

The analysis of the electroweak symmetry breaking is simplified by using gauge invariant degrees of freedom similar to the scheme presented in chapter 2. We describe the two scalar doublets by the orbit variables (2.20) introduced for the THDM. To achieve this, we first rewrite H_u, H_d in terms of same-hypercharge doublets via (2.6), and then replace all doublet scalar products in the potential via (2.21). Further, we decompose the complex singlet field into two real fields according to $S = S_{re} + iS_{im}$. In this notation we find for

the NMSSM Higgs potential

$$\begin{aligned} V_{\mathrm{F}} = &\frac{1}{4}|\lambda|^2 \left(K_1^2 + K_2^2 + 4K_0(S_{re}^2 + S_{im}^2)\right) + |\kappa|^2(S_{re}^2 + S_{im}^2)^2 \\ &- \mathrm{Re}(\lambda\kappa^*) \left(K_1(S_{re}^2 - S_{im}^2) + 2K_2 S_{re} S_{im}\right) \\ &+ \mathrm{Im}(\lambda\kappa^*) \left(K_2(S_{re}^2 - S_{im}^2) - 2K_1 S_{re} S_{im}\right), \end{aligned} \tag{3.22}$$

$$V_{\mathrm{D}} = \frac{1}{8}\bar{g}^2 K_3^2 + \frac{1}{8}g^2 \left(K_0^2 - K_1^2 - K_2^2 - K_3^2\right), \tag{3.23}$$

$$\begin{aligned} V_{\mathrm{soft}} = &\frac{1}{2}m_{H_u}^2 (K_0 - K_3) + \frac{1}{2}m_{H_d}^2 (K_0 + K_3) + m_S^2 (S_{re}^2 + S_{im}^2) \\ &- \mathrm{Re}(\lambda A_\lambda) (K_1 S_{re} - K_2 S_{im}) + \frac{2}{3}\mathrm{Re}(\kappa A_\kappa) \left(S_{re}^3 - 3S_{re}S_{im}^2\right) \\ &+ \mathrm{Im}(\lambda A_\lambda) (K_2 S_{re} + K_1 S_{im}) + \frac{2}{3}\mathrm{Im}(\kappa A_\kappa) \left(S_{im}^3 - 3S_{re}^2 S_{im}\right), \end{aligned} \tag{3.24}$$

where all six degrees of freedom are real. As for the THDM, the domain of the orbit variables is given by the forward light cone condition

$$K_0 \geq |\boldsymbol{K}|\,, \tag{3.25}$$

see (2.24).

We shall now consider stationary points in these new degrees of freedom. With respect to a spontaneous breaking of the electroweak symmetry the singlet is irrelevant, and we may directly translate the results from section 2.3.1 to the present case.

Classification 3.3.1. *We distinguish the possible cases of stationary points by the $SU(2)_L \times U(1)_Y$ symmetry breaking behaviour which a vacuum of this type would have, see classification 2.3.1:*

- unbroken $SU(2)_L \times U(1)_Y$*: A stationary point with*

$$K_0 = K_1 = K_2 = K_3 = 0\,. \tag{3.26}$$

  A global minimum of this type implies vanishing vacuum expectation value for the doublet fields (2.1) and therefore the trivial behaviour with respect to the gauge group. The stationary points of this type are found by setting all Higgs-doublet fields (or correspondingly the K_0 as well as the K_a fields) in the potential to zero and requiring a vanishing gradient with respect to the singlet fields:

$$\nabla\, V(S_{re}, S_{im}) = 0\,. \tag{3.27}$$

- partially broken $SU(2)_L \times U(1)_Y$*: A stationary point with*

$$\begin{aligned} K_0 &> 0\,, \\ K_0^2 - K_1^2 - K_2^2 - K_3^2 &= 0\,. \end{aligned} \tag{3.28}$$

A global minimum of this type leads to the desired partial breaking of $SU(2)_L \times U(1)_Y$ *down to* $U(1)_{em}$ *. Using the Lagrange method, these stationary points are given by the real solutions of the system of equations*

$$\begin{gathered} \nabla\big[V(K_0, K_1, K_2, K_3, S_{re}, S_{im}) - u \cdot (K_0^2 - K_1^2 - K_2^2 - K_3^2)\ \big] = 0\,, \\ K_0^2 - K_1^2 - K_2^2 - K_3^2 = 0\,, \end{gathered} \tag{3.29}$$

where u *is a Lagrange multiplier. The inequality in* (3.28) *must be checked explicitly for the solutions found for* (3.29)*.*

- fully broken $SU(2)_L \times U(1)_Y$*: A stationary point with*

$$\begin{gathered} K_0 > 0\,, \\ K_0^2 - K_1^2 - K_2^2 - K_3^2 > 0\,. \end{gathered} \tag{3.30}$$

A global minimum of this type has non-vanishing vacuum expectation values for the charged components of the doublets fields in (2.1), thus leading to a fully broken $SU(2)_L \times U(1)_Y$*. The stationary points of this type are found by requiring a vanishing gradient with respect to all singlet fields and all gauge invariant functions:*

$$\nabla\, V(K_0, K_1, K_2, K_3, S_{re}, S_{im}) = 0\,. \tag{3.31}$$

The constraints (3.30) *on the gauge invariant functions must be checked explicitly for the real solutions found.*

For a potential which is bounded from below, the global minima will be among these stationary points. Solving the systems of equations (3.27), (3.31), and (3.29), and inserting the solutions in the potential, we can therefore identify the global minima as those solutions which have the lowest value of the potential. Note that in general the global minimum can be degenerate. In fact, due to the $\mathbb{Z}_3$ symmetry of the potential, stationary points generically are either threefold degenerate or non-degenerate.

3.4 Determination of stationary points via Gröbner bases

From the mathematical point of view we have to solve with (3.27), (3.31), (3.29) a non-linear, multivariate, inhomogeneous systems of polynomial equations of third order. Standard analytical approaches [124] fail due to the complexity of the problem, while a numerical approach [124] tends to be unreliable, since it sometimes misses solutions. Floating point numbers in algorithms like that should be treated with care for reason we shall shortly explain. In the following, we provide an algorithm, which reliably solves the stationarity conditions in short time despite the relatively large number of fields present in the NMSSM. The most involved case is given by (3.29), which consists of seven equations in seven indeterminates, namely six real fields and one Lagrange multiplier.

In the following we describe an algorithmic method to solve (3.27), (3.31), (3.29) for the case that the number of complex solutions is finite. The latter is indeed fulfilled for the NMSSM with generic values for the parameters, and it is automatically checked by the method. Note that the gauge invariant functions avoid spurious continuous sets of complex solutions, which we found to arise in the case of the MSSM as well as the NMSSM if the stationarity conditions are formulated with respect to the Higgs fields (2.1) in a unitary gauge. This is not surprising given the fact, that the gauge invariant functions express the contribution of the doublets to the potential by four real degrees of freedom in contrast to the five encountered for the doublet components in the unitary gauge.

The solution of multivariate polynomial systems of equations is the subject of polynomial ideal theory and can be obtained algorithmically in the Gröbner basis approach [125]. See appendix A for a brief introduction to this subject. Within this approach, the system of equations is transformed into a unique standard form with respect to a specified underlying ordering of the polynomial terms of the sum (*monomials*). This unique standard form of the system of equations is given by the corresponding *reduced Gröbner basis*. If the underlying order is the *lexicographical* ordering, the unique standard form consists of equations with a partial separation in the indeterminates. We use a variant of the F_4 algorithm [126] to compute the Gröbner bases. A Gröbner basis computation is often considerably faster if the standard form is computed with respect to *total degree* orderings and then transformed into a lexicographical Gröbner basis. The transformation of bases from total degree to lexicographical ordering is done with the help of the FGLM algorithm [127]. Finally, the system of equations represented by the lexicographical Gröbner basis has to be triangularised. The decomposition of the system of equations with a finite number of solutions into triangular sets is performed with the algorithm introduced in [128, 129]. Each triangular system consists of one univariate equation, one equation in 2 indeterminates, one equation in 3 indeterminates and so forth. This means that the solutions are found by subsequently solving just univariate equations by inserting the solutions from the previous steps.

The construction of the Gröbner basis as well as the triangularisation are done algebraically, so no approximations are needed there. However, the triangular system of equations contains in general polynomials of high order, where the zeros cannot be obtained algebraically. Here numerical methods are needed to find the in general complex roots of the univariate polynomials.

For more involved systems of equations, as is the case for the NMSSM stationarity conditions, the algorithmic solution is considerably simplified (or even made accessible), if the coefficients of the polynomials are given in form of rational numbers. Since rational numbers are arbitrarily close to real numbers and moreover the physical parameters are given only with a certain precision, this does not limit the general applicability of the method in practice. While Gröbner basis algorithms exist which use floating point number representations, their mathematical foundations are still in development. The key problem lies in the fact that the recognition of exact zeros is at the heart of the Gröbner algorithms, see appendix A, but difficult to achieve with approximate number representations. For further reading on this topic and a rigorous approach see [130, 131] and references therein.

All algorithms for the computation of the Gröbner basis with respect to a given order of the monomials, the change of the underlying order, the triangularisation, and the solution of the triangular systems are implemented in the `Singular` program package [132]. The solutions obtained can be easily checked by inserting them into the initial system of equations. Moreover, the number of complex solutions, that is the multiplicity of the system, is known at the algebraical level, so we can easily check that no stationary point is missing.

3.5 Numerical results

In order to fix experimentally known parameters like the electroweak scale, it is inappropriate to choose numerical values for the original potential parameters (3.5). Instead, we express different original parameters in terms of more physical input parameters, namely the desired vacuum expectation values (3.6) of the neutral components of the Higgs doublets and the Higgs singlet, the mass of the charged Higgs boson and a CP-violating phase. We write the complex parameters λ, κ, A_λ, and A_κ in polar coordinates with phases δ_λ, δ_κ, δ_{A_λ}, δ_{A_κ} and introduce the abbreviations

$$\delta_{\mathrm{EDM}} \equiv \delta_\lambda + \varphi_u + \varphi_S\,, \quad \delta'_\kappa \equiv \delta_\kappa + 3\varphi_S\,. \tag{3.32}$$

Together with the stationarity conditions (3.8) and the tree-level expression for the charged Higgs mass squared (3.10), the initial parameters of the potential (3.5) can be replaced by the new set of parameters

$$\lambda,\ \kappa,\ |A_\kappa|,\ \tan\beta,\ v_S,\ m_{H^\pm},\ \operatorname{sign} R_\kappa,\ \delta_{\mathrm{EDM}},\ \delta'_\kappa \tag{3.33}$$

plus the electroweak scale $v \approx 246\,\mathrm{GeV}$. In the mass matrix of the Higgs scalars, the CP violating entries which mix the "scalar" with the "pseudoscalar" fields are proportional to the imaginary part of $\exp[i(\delta_{\mathrm{EDM}} - \delta'_\kappa)]$. This new set of physical input parameters for the potential allows e.g. to adjust the electroweak scaled to the observed value. It is applicable in presence of CP violating phases as well as for their absence without the necessity of any case distinction.

As a numerical example, we choose the parameter values

$$\begin{aligned} &\lambda = 0.4\,, \quad \kappa = 0.3\,, \quad |A_\kappa| = 200\,\mathrm{GeV}, \\ &\tan\beta = 3\,, \quad v_S = 3v\,, \quad m_{H^\pm} = 2v\,, \\ &\operatorname{sign} R_\kappa = -\,, \quad \delta_{\mathrm{EDM}} = 0\,, \quad \delta'_\kappa = 0 \end{aligned} \tag{3.34}$$

and consider the variation of one parameter at a time with the values of the other parameters in (3.34) kept fixed. For a given point in parameter space, we compute all stationary points of the NMSSM potential as described above. As mentioned in section 3.4 the Gröbner basis construction is performed with numerical coefficients. Here we use a precision of 12 digits for the input parameters (3.5), which are determined from the values for the parameters (3.33). The roots of the univariate polynomials are found numerically,

where we choose a precision of 100 digits. Our approach is not limited by any fixed precision in both cases. We verify that the errors of the approximate statements described in the following are under control. For generic values of the parameters we find 52 complex solutions: 7 corresponding to the unbroken, 38 to the partially broken, and 7 to the fully broken cases. The number of real and therefore relevant solutions depends on the specific values of the parameters.

As expected from the $\mathbb{Z}_3$ symmetry of the potential, we find either 1 or 3 solutions sharing the same value of the potential within the accuracy of the numerical roots. From the computed stationary points only those may be accepted as global minima which correspond to the initial vacuum expectation values (up to the complex phases), that is which fulfil

$$\sqrt{2K_0} \approx v\,, \quad \sqrt{\frac{K_0 - K_3}{K_0 + K_3}} \approx \tan\beta\,, \quad \sqrt{2(S_{re}^2 + S_{im}^2)} \approx v_S\,. \tag{3.35}$$

Since for non-vanishing λ, κ the potential is bounded from below, the stationary point with the lowest value of the potential is the global minimum. Further, we determine for every stationary solution, whether it is a local maximum, local minimum or a saddle point. For the regular solutions, i.e. the partially- and fully-broken cases, this is achieved via the bordered Hessian matrix (see for example [115], see also (2.97)), in terms of $K_0, K_1, K_2, K_3, S_{re}, S_{im}$. This takes all powers of the doublet fields into account, which allows for a definite decision on the type of the stationary point also for frequently encountered partial breaking solutions where at least one mass squared is zero and the others have the same sign. For irregular solutions, i.e. the non-breaking solutions with $K_0 = 0$, the Lagrange formalism can not be used since the gradients of the two boundary conditions with respect to $K_0, K_1, K_2, K_3, S_{re}, S_{im}$ become linearly dependent. Instead, we resubstitute the original fields $H_u^+, H_u^0, H_d^0, H_d^-, S_{re}, S_{im}$ in this case and consider the free Hessian matrix with respect to these fields. This turns out to be sufficient in practice to judge on the type of the stationary points.

Figures 3.1-3.4 show the values of the potential at all stationary points for the parameter values (3.34) and the cases where successively one of the parameters κ, λ, $\tan\beta$, v_S, $m_{H^\pm}$, δ'_κ is varied. Each curve in the Figures represents 1- or 3-fold degenerate stationary potential values, where the gauge symmetry breaking behaviour of the solutions is denoted by different line styles. Excluded parameter regions, where the global minimum does not exhibit the required expectation values (3.35) are drawn shaded. As is illustrated by the figures, we find that substantial regions of the NMSSM parameter space are excluded. For some excluded parameter regions, the partially breaking solutions with the required vacuum expectation values (3.35) are saddle points (see for instance figure 3.1, top). This means they can be discarded as global minima without calculation of the other stationary points by checking for positive Higgs masses. However, this is not always the case. Obviously from figure 3.3, top, we find an upper bound for v_S. For the plotted v_S larger than this upper exclusion bound the solutions fulfilling (3.35) are still pronounced minima, i.e. the mass matrices have positive eigenvalues, but they are no longer the global minima. We note furthermore, that the "dangerous" solutions are often partially breaking triples, of which one is CP-invariant. That is, they are of the same type as the triple containing the

required vacuum. Even though most difficult to calculate, these cannot be omitted from the discussion. We also note that for some cases, an unwanted global minimum occurs at large field configurations. The influence of the CP-violating phase δ'_κ is shown in figure 3.4, bottom. Note that $\delta'_\kappa \to -\delta'_\kappa$ is not a symmetry of the potential. However, the potential is invariant under $(\delta'_\kappa, K_2, S_{im}) \to -(\delta'_\kappa, K_2, S_{im})$, that is $(\delta'_\kappa, H_u, H_d, S) \to (-\delta'_\kappa, H_u^*, H_d^*, S^*)$, if the residual phases are chosen as in (3.34). Therefore the stationary values of the potential in figure 3.4 depend only on $|\delta'_\kappa|$.

In all figures shown here there are non-breaking saddle points with potential values slightly above those of the *wanted* global minimum. We already discussed this feature in section 3.2 for a certain limit of the parameters. We find that this effect is not coincidental for the parameters (3.34) chosen there or here, but rather a generic feature of the NMSSM. Within the CP conserving parameter range

$$\lambda \in]0,1], \quad \kappa \in]0,1], \quad A_\kappa \in \pm\,]0,2500]\,\text{GeV}, \tag{3.36}$$

$$\tan\beta \in]0,50], \quad v_s \in]0,5000]\,\text{GeV}, \quad m_{H^\pm} \in]0,2500]\,\text{GeV} \tag{3.37}$$

we select samples producing the wanted global minimum and typically find non-breaking saddle points, where the relative separation of the potential values for the saddle points and the global minimum is below the per-mille level, in many cases even far below. We do not find fully breaking global minima for scenarios in the range (3.37) where the solutions with the required vacuum expectation values (3.35) are local minima. Eventually, we find examples, where CP conserving parameter values with the "wrong" global minimum produce the wanted global minimum if a non-vanishing phase δ'_κ is introduced.

Summary and outlook

We have demonstrated a new method to determine the global minimum of the NMSSM Higgs potential at the tree-level. We have used Gröbner basis calculations to determine all stationary points, where the approach ensures at the algebraical level that the global minimum is not missed. To our knowledge, the presented method is the first minimisation proposal for extended Higgs potentials achieving this. Our results show that the requirement of a stable vacuum excludes large regions of the NMSSM parameter space. Without fine-tuning, we have found cases which are challenging for local descent based algorithms because of strongly anisotropic potential structures or minima at large field values. Further, we have encountered non-finetuned cases, where the instability of the vacuum is neither obvious from the mass matrices nor due to often suspected minima.

Clearly, the scope of the method is not restricted to the NMSSM. We have successfully tested it also for THDMs. The run-time of existing Gröbner basis algorithms grows rapidly with the complexity of the equations and in particular the number of degrees of freedom. This limits the applicability of the method currently to cases not much more involved than the NMSSM. For the approach as it stands, it is essential to manually eliminate continuous symmetries from the potential. Here, this was achieved for the gauge degrees of freedom by employing orbit variables for the two Higgs doublets. While it was not necessary for

the case of the NMSSM, the Gröbner basis approach might also be used to ensure a Higgs potential is bounded from below. For this, the large field behaviour of the potential could be described by a function defined on a compact domain similar as shown for the THDM in section 2.2. Stability of the potential could then be achieved by checking the stationary points of such a function with the Gröbner basis method.

Since the Higgs sectors in particular of supersymmetric theories are known to receive considerably large radiative corrections, it would be very desirable to extend our method to at least the 1-loop-effective potential. This extension is not straight-forward, since the Gröbner basis notation is restricted to problems which can be formulated via polynomials. Nevertheless, the tree-level result obtained with our method might at least serve as a basic handle to access the global structure of the potential reliably.

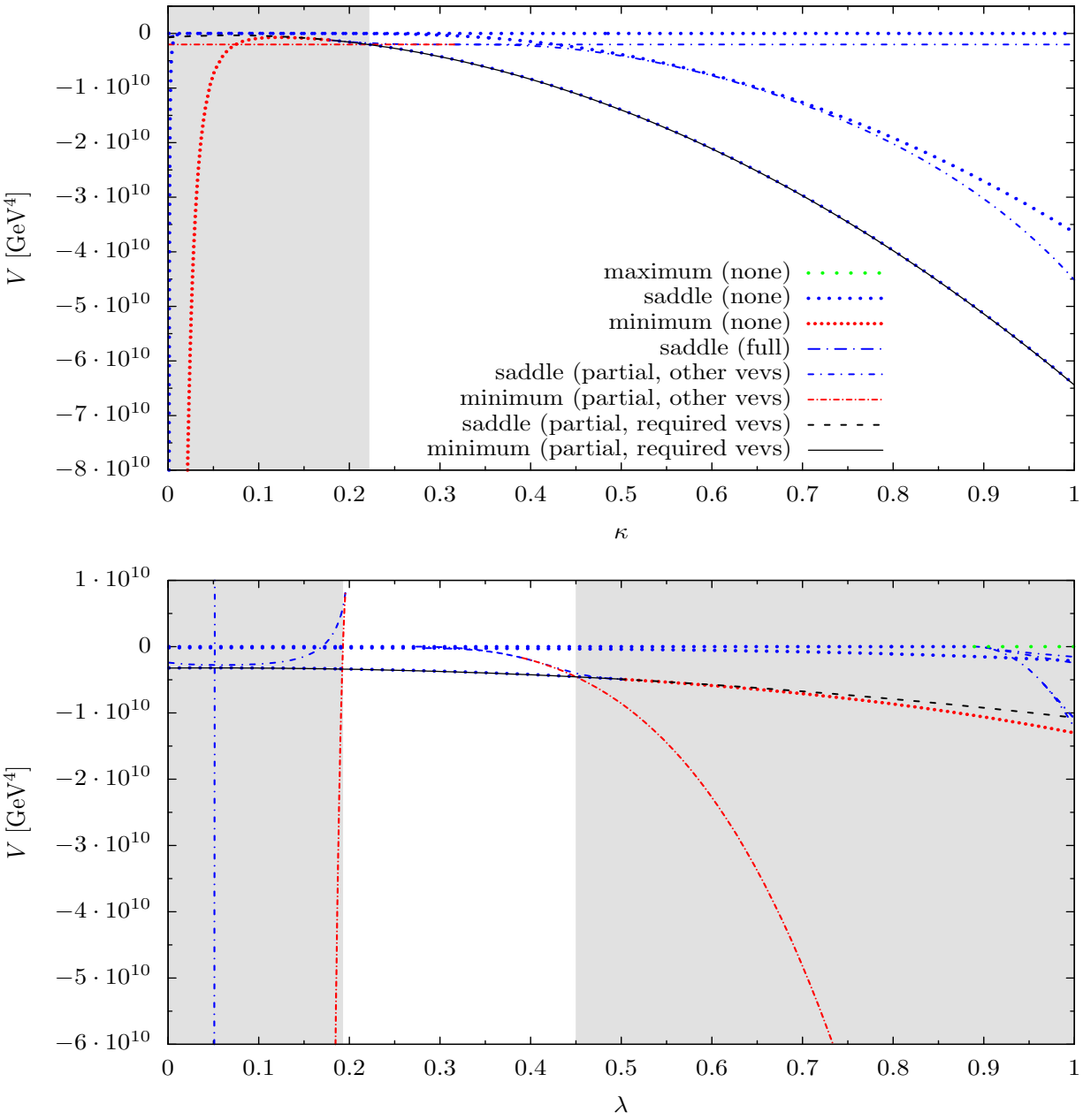

Figure 3.1: Values of the NMSSM potential at its stationary points in dependence of κ and λ, respectively. The following parameters are kept constant unless explicitly varied: $\lambda = 0.4, \kappa = 0.3, |A_\kappa| = 200\,\text{GeV}, \tan\beta = 3, v_S = 3v, m_{H^\pm} = 2v, \text{sign}\, R_\kappa = -, \delta_{\text{EDM}} = \delta'_\kappa = 0$. Each line corresponds to 1 or 3 stationary points sharing the same value of the potential. The different line styles denote saddle points, maxima, and minima. The labels 'none', 'full', and 'partial' denote solutions of the classes with unbroken (3.27), fully broken (3.31), and partially broken (3.29) $SU(2)_L \times U(1)_Y$, respectively. For solutions of the partially broken class, it is also denoted whether they correspond to the 'required vevs' v_u, v_d, v_S or to 'other vevs'. Excluded parameter values, where the global minimum does not exhibit the required vacuum expectation values, are drawn shaded.

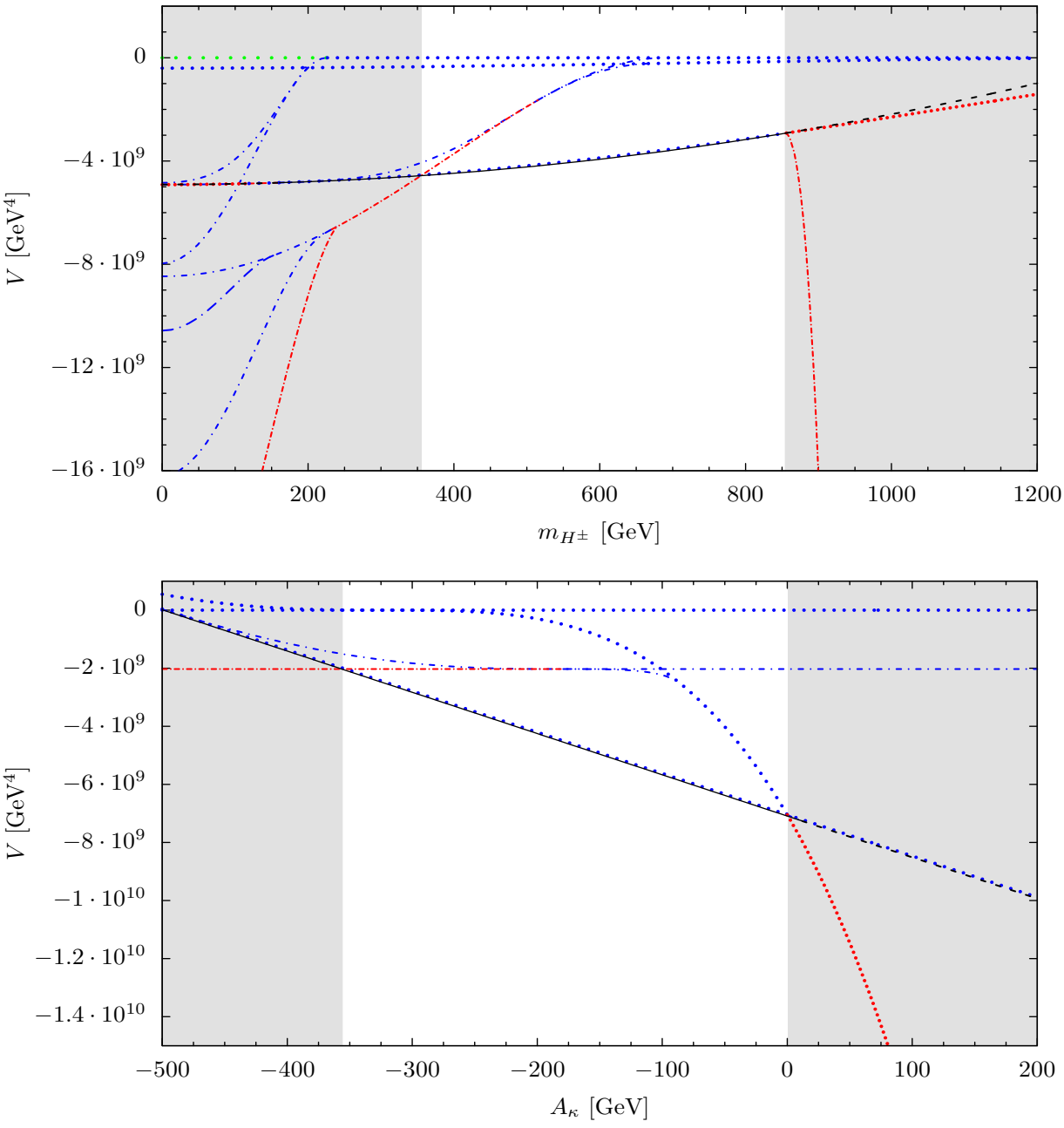

Figure 3.2: Same as in figure 3.1 but for variation of $m_{H^\pm}$ and A_κ, respectively.

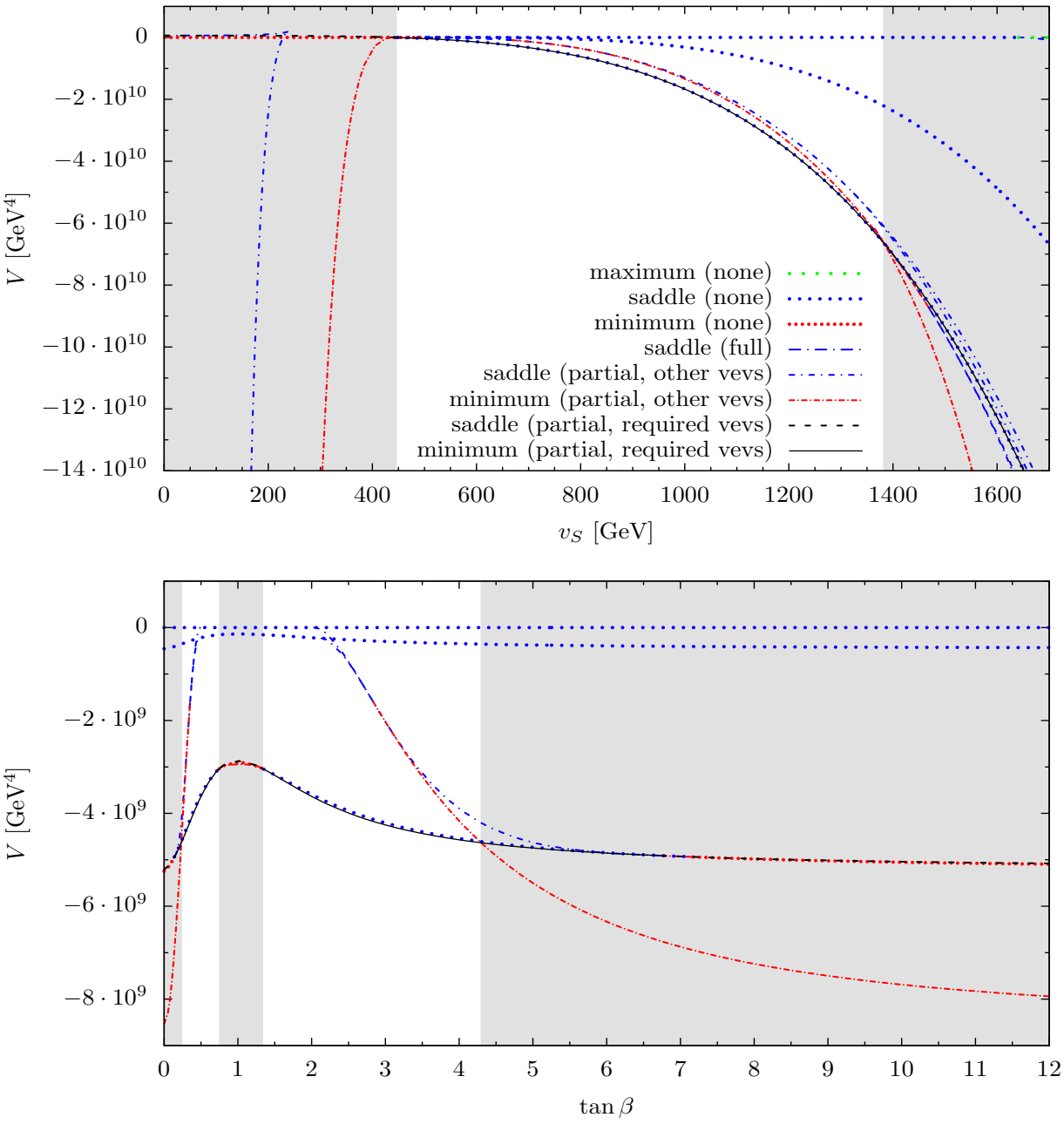

Figure 3.3: Same as in figure 3.1 but for variation of v_S and $\tan\beta$, respectively.

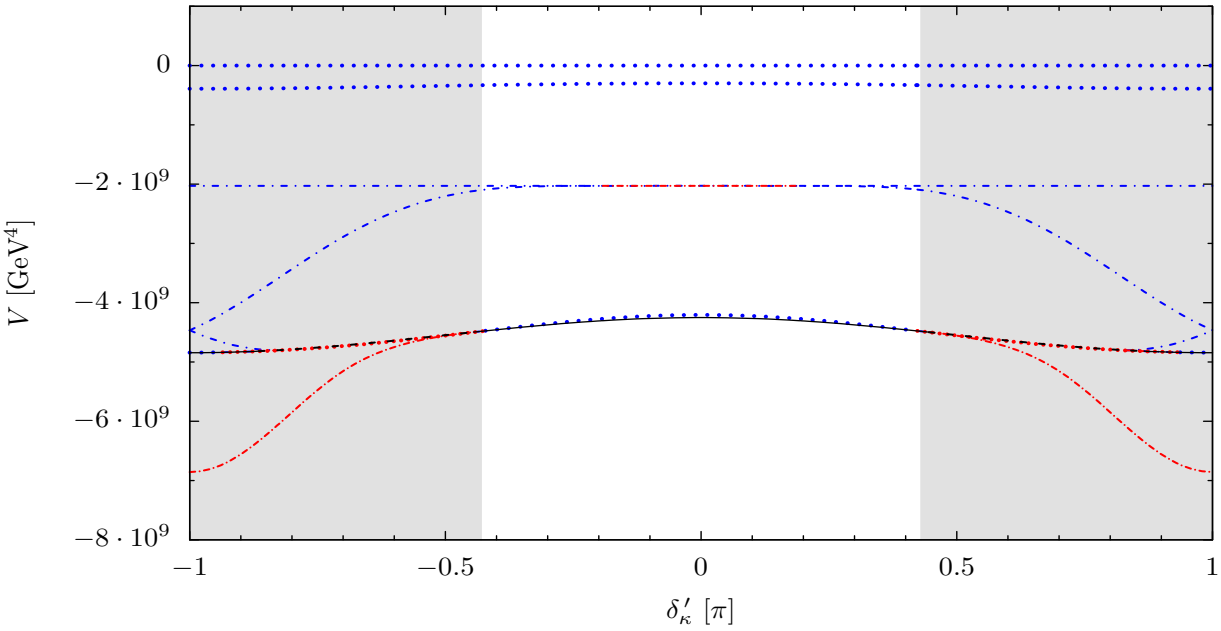

Figure 3.4: Same as in figure 3.1 but for variation of the CP-violating phase δ'_κ.

Part III

Constraints on the Colour Dipole Picture

Chapter 4

Success of the colour dipole picture

4.1 The picture

For the understanding of the proton substructure, studies of deep inelastic scattering (DIS) of electrons and positrons on protons

$$e^{\pm} + p \to e^{\pm} + X \tag{4.1}$$

play a crucial role [133]. We shall consider here momentum transfers squared $Q^2 \lesssim 1000\,\mathrm{GeV}^2$. Then it is sufficient to consider exchange of a virtual photon between the leptons and the hadrons in (4.1). Thus, the reaction which we shall study in the following is the absorption of a virtual photon γ^* on the proton,

$$\gamma^* + p \to X\,, \tag{4.2}$$

where we perform the sum over all final states X. The center-of-mass energy for this reaction is denoted by W, the virtuality of γ^* by Q^2. The proton in (4.2) is supposed to be unpolarised, but the virtual photon can have transverse or longitudinal polarisation. The corresponding total cross sections are σ_T and σ_L, respectively. The F_2 structure function takes with Hand's convention [134] for the γ^* flux factor at small Bjorken-x the simple form

$$F_2(W,Q^2) = \frac{Q^2}{4\pi^2\alpha_{\mathrm{em}}}\left(\sigma_T(W,Q^2) + \sigma_L(W,Q^2)\right), \tag{4.3}$$

where α_{em} is the fine structure constant. To obtain the standard dipole model for the cross sections $\sigma_{T,L}$, they are first related to the imaginary part of the $\gamma^* p \to \gamma^* p$ forward scattering amplitude. The latter is represented as the initial γ^* splitting in a $q\bar{q}$ pair, this pair scattering on the proton and the $q\bar{q}$ subsequently fusing to the final state γ^*, see figure 4.1. Note that this figure is to be read from right to left in order to be in complete analogy with the occurrence of the various factors in the amplitudes. In the high-energy

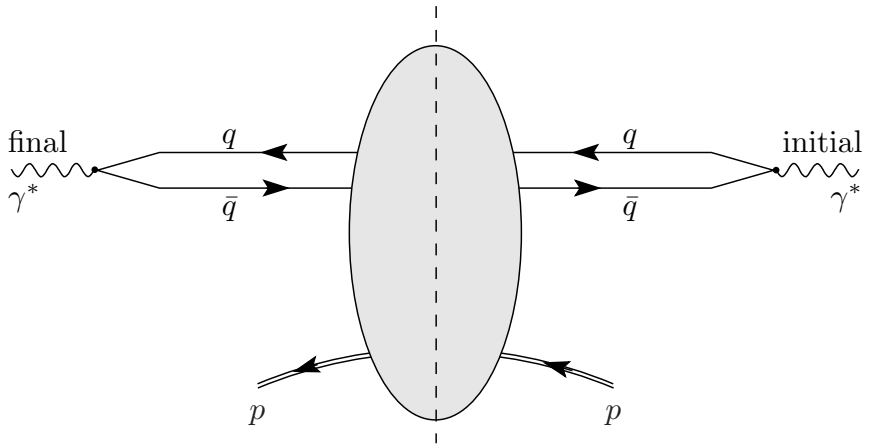

Figure 4.1: Basic diagram for the description of the cross sections $\sigma_{T,L}$ of $\gamma^* p$ scattering in the standard dipole approach.

limit, the diagram of figure 4.1 leads to

$$\sigma_T(W,Q^2) = \sum_{q,\lambda,\lambda'} \int \mathrm{d}^2 r \int_0^1 \mathrm{d}\alpha \, \left(\psi^{(q)\,\pm}_{\gamma,\,\lambda\lambda'}(\alpha,\boldsymbol{r},Q^2)\right)^* \hat{\sigma}^{(q)}(r,W,Q^2)\, \psi^{(q)\,\pm}_{\gamma,\,\lambda\lambda'}(\alpha,\boldsymbol{r},Q^2)\,, \quad (4.4)$$

$$\sigma_L(W,Q^2) = \sum_{q,\lambda,\lambda'} \int \mathrm{d}^2 r \int_0^1 \mathrm{d}\alpha \, \left(\psi^{(q)\,L}_{\gamma,\,\lambda\lambda'}(\alpha,\boldsymbol{r},Q^2)\right)^* \hat{\sigma}^{(q)}(r,W,Q^2)\, \psi^{(q)\,L}_{\gamma,\,\lambda\lambda'}(\alpha,\boldsymbol{r},Q^2)\,, \quad (4.5)$$

provided a number of assumptions and approximations hold, as worked out in detail in [135, 136][1]. In (4.4) either $\psi^{(q)\,+}_{\gamma,\,\lambda\lambda'}$ or $\psi^{(q)\,-}_{\gamma,\,\lambda\lambda'}$ may be used, both give the same result. Here, α is the longitudinal momentum fraction of the γ^* carried by the quark, $\boldsymbol{r}$ is the vector in transverse position space from the antiquark to the quark, and λ and λ' are the spin indices of q and $\bar{q}$, respectively. The total cross section for the scattering of the $q\bar{q}$ pair on the proton is denoted by $\hat{\sigma}^{(q)}$, the γ^* wave functions for transversely and longitudinally polarised γ^* by $\psi^{(q)\,\pm}_{\gamma,\,\lambda\lambda'}$ and $\psi^{(q)\,L}_{\gamma,\,\lambda\lambda'}$ respectively. Finally, a sum over all contributing quark flavours q is to be performed. The photon wave functions are calculated perturbatively. They enter the total cross sections via their absolute squares, which give to leading order in α_{em} and with the usual normalisation:

$$\sum_{\lambda,\lambda'} \left|\psi^{(q)\pm}_{\gamma,\lambda\lambda'}(\alpha,\boldsymbol{r},Q)\right|^2 = \frac{N_c \alpha_{\mathrm{em}} Q_q^2}{2\pi^2} \left[(\alpha^2+(1-\alpha)^2)\epsilon_q^2 K_1^2(\epsilon_q r) + m_q^2 K_0^2(\epsilon_q r)\right], \quad (4.6)$$

$$\sum_{\lambda,\lambda'} \left|\psi^{(q)L}_{\gamma,\lambda\lambda'}(\alpha,\boldsymbol{r},Q)\right|^2 = \frac{2N_c \alpha_{\mathrm{em}} Q_q^2}{\pi^2} Q^2 \left(\alpha(1-\alpha)\right)^2 K_0^2(\epsilon_q r)\,. \quad (4.7)$$

with the abbreviation

$$\epsilon_q := \sqrt{\alpha(1-\alpha)Q^2 + m_q^2}. \quad (4.8)$$

[1] In the next chapter, we shall give more details and point out a subtle restriction for (4.4)-(4.5) obtained in [136].

Here, $N_c = 3$ is the number of colours, Q_q denotes the quark charges in units of the proton charge, m_q the quark masses and $K_{0,1}$ are modified Bessel functions.

A prediction for the dipole cross section from first principles within QCD is not available up to now. Due to its intrinsic non-perturbative nature, it can not be calculated from perturbative QCD. While its computation via lattice QCD simulations seems to be a natural attempt, it still is an unsolved problem. A solution might possibly be found using alternative formulations of lattice QCD instead of the standard euclidean path-integral based methods. For instance, in [137] a Hamiltonian formulation of lattice QCD is proposed aiming at a well-defined continuum limit also on the light cone.

Several phenomenological models have been proposed for the dipole cross section, see e.g. [138, 139, 140, 141, 142, 143] and the review [144]. A further review and a discussion of several issues considered in this thesis can be found in [145]. The proposed phenomenological dipole cross sections involve parameters which are fitted to the data. In this thesis we derive various bounds within the colour dipole picture, which are independent on the specific form of the dipole cross section. But we shall also use particular models for $\hat{\sigma}$ for the discussion of various effects.

4.2 The Golec-Biernat-Wüsthoff model

A particular popular dipole cross section was proposed by Golec-Biernat and Wüsthoff (GBW) in [138]. It represents the prototype for models with small-x saturation. Its dipole cross section is

$$\hat{\sigma}^{(q)}_{\mathrm{GBW}}(\boldsymbol{r}, x) = \sigma_0 \left(1 - e^{-\frac{r^2}{4r_0^2(\tilde{x})}} \right), \tag{4.9}$$

where

$$r_0^2(\tilde{x}) = \left(\frac{\tilde{x}}{x_0} \right)^{\lambda} \mathrm{GeV}^{-2}, \qquad \tilde{x} = \left(1 + \frac{4m_q^2}{Q^2} \right) x \,. \tag{4.10}$$

This dipole cross section depends on Bjorken-x, $x \approx Q^2/W^2$, and therefore not only on W but also on Q^2. For small r the cross section rises with r^2 (colour transparency) and reaches a plateau at larger r (saturation). Besides the general saturation feature, the scale r_0 at which saturation sets in decreases with decreasing x. This can be interpreted as the parton structure of the proton being resolved only if testing with small dipoles. For decreasing x only decreasingly small dipoles are able to resolve the partons. In our numerical applications we use their "no charm" fit, see p.10 in [138]:

$$\hat{\sigma}_0 = 23030 \ \mu\mathrm{b}, \qquad \lambda = 0.288, \qquad x_0 = 3.04 \cdot 10^{-4}. \tag{4.11}$$

Figure 4.2 illustrates the cross section for the case of one massless quark flavour and different values of x. Despite its simple form with only three free parameters, the GBW model is actually able to describe the HERA data remarkably well. In [146] a modified version of the model was presented, which takes the DGLAP evolution [147, 148, 149, 150] of the gluon distributions into account. This modification allows improved fits to the

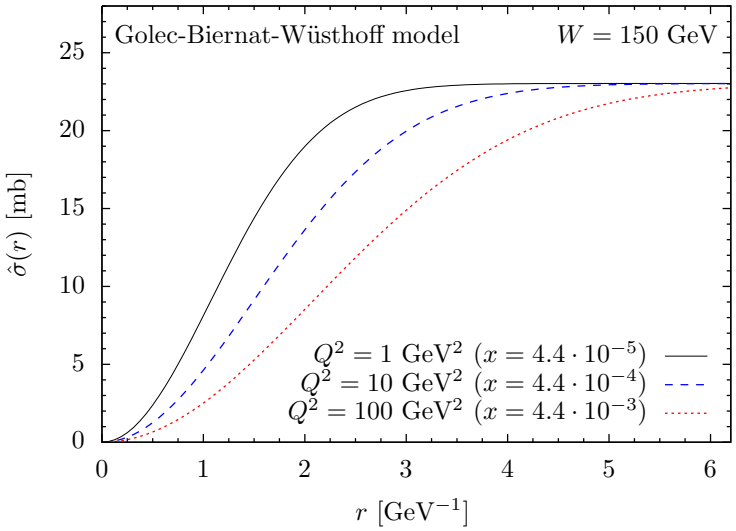

Figure 4.2: Dipole cross section $\hat{\sigma}^{(q)}$ for one massless quark flavour in dependence of the dipole size r for the Golec-Biernat-Wüsthoff model. In this model $\hat{\sigma}^{(q)}$ depends on x.

HERA data in particular for higher values of Q^2 at the cost of introducing alltogether five free parameters. Figure 4.3 shows HERA data for F_2 in dependence of x together with fits for the original GBW model and its DGLAP based modification.

Often results obtained within the dipole picture are interpreted in terms of the perturbative gluon densities $g(x, \mu^2)$, where the relation (see e.g. [151])

$$\hat{\sigma} \simeq \frac{\pi^2}{3} r^2 \alpha_s \, x g(x, \mu^2) \tag{4.12}$$

is used. There, the evolution scale μ is described in terms of $1/r$, and $1/r$ is often regarded as being effectively proportional to Q. Saturation effects at small x are then considered as non-linear phenomena due to the quick rise of the gluon densities. As will be discussed in the next chapter, the applicability of this relation might actually require some care.

In this thesis we shall consider various bounds and consistency checks for the colour dipole picture in order to deepen the understanding of its range of applicability, see also our article [152]. This is important, since the dipole picture is not exact and requires several approximations and assumptions to be made. In section 5 we review the derivation [135, 136] of the colour dipole picture and provide the building blocks for the following discussions. There, the motivations for our considerations are specified in more detail. In section 6 we derive general bounds on ratios of DIS structure functions, which follow from the general framework of the dipole picture and photon density properties alone. In section 6.1 we consider the ratio of the charm and longitudinal part of the structure function F_2. We shall derive bounds on this ratio, which are necessarily valid for any dipole cross section and might provide constraints on the dipole picture when confronted with future measurements. In section 6.2 we consider ratios of F_2 taken at the same energy W but different Q^2. We shall calculate bounds, which are valid for any dipole cross sections

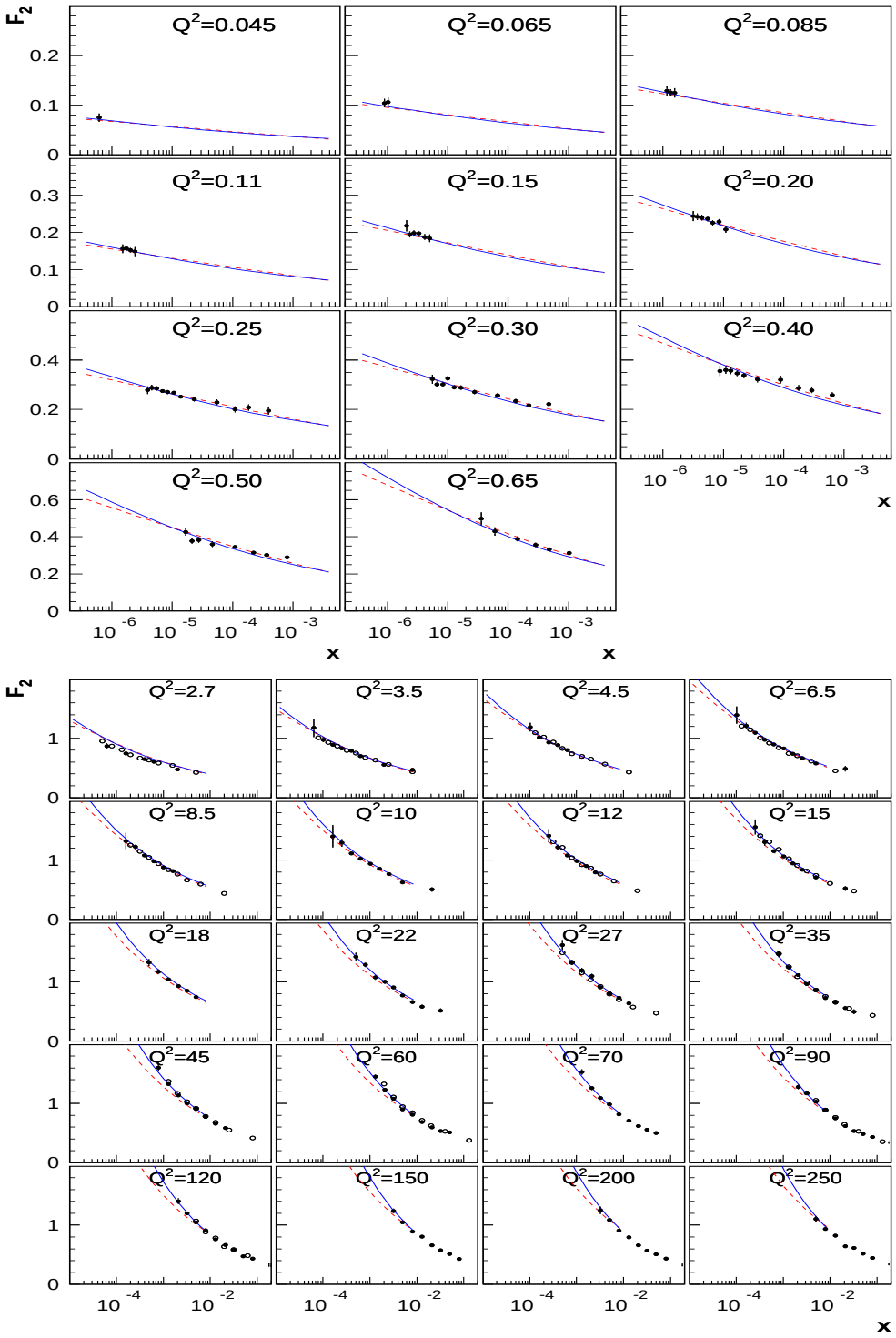

Figure 4.3: HERA data (dots) for F_2 in dependence of x together with fits for the original GBW model and its DGLAP based modification (lines). Upper graph: ZEUS BPT97 (full), GBW+DGLAP (solid), GBW (dashed). Lower graph: ZEUS (full), H1 (open), GBW+DGLAP (solid), GBW (dashed). Figures are taken from [146].

whose energy dependence is given by W alone. Confronting these bounds with data allows us to derive restrictions on the range of applicability of the framework. In section 7 we consider explicitly different choices for the energy dependence of the dipole cross section and discuss whether they may be related by effective scale arguments. Finally, in section 8 we shall calculate Ioffe times, that is, dipole lifetimes within the colour dipole picture. As a consistency check we consider, whether major contributions are associated with long Ioffe times such that the separation of dipole production and interaction with the proton used in the dipole picture is justified.

Chapter 5

Foundations and building blocks

In this chapter we review the foundations of the dipole picture as worked out in [135, 136]. Further, we shall discuss important features of the photon densities in detail as a preparation for the following chapters of this thesis.

5.1 Non-perturbative foundations

We consider deep inelastic lepton-proton scattering

$$l(l) + p(p) \to l(l') + X(p')\,, \tag{5.1}$$

where the corresponding 4-momenta of the particles are indicated in parentheses and $l = e^-, e^+$. In standard kinematics (see for instance [64]) we have

$$\begin{gathered} s = (p+l)^2\,, \qquad q = l - l' = p' - p\,, \qquad \nu = \frac{pq}{m_p}\,, \qquad y = \frac{pq}{pl} = \frac{2m_p\nu}{s - m_p^2}\,, \\ Q^2 = -q^2\,, \qquad W^2 = (p+q)^2 = 2m_p\nu - Q^2 + m_p^2\,, \qquad x = \frac{Q^2}{2m_p\nu} = \frac{Q^2}{W^2 + Q^2 - m_p^2} \end{gathered} \tag{5.2}$$

with the proton mass m_p. In the following, we consider moderate Q^2,

$$0 \leq Q^2 \lessapprox 10^3\,\mathrm{GeV}^2\,, \tag{5.3}$$

such that only photon exchange has to be taken into account. That is, we are interested in the reaction

$$\gamma^*(q) + p(p) \to X(p')\,, \tag{5.4}$$

where the proton is supposed to be unpolarised and a sum over all final states X is performed. The total cross section for (5.4) is encoded in the hadronic tensor

$$W^{\mu\nu}(p,q) = -W_1(\nu, Q^2)\left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\right) + \frac{1}{m_p^2} W_2(\nu, Q^2)\left(p^\mu - \frac{(pq)q^\mu}{q^2}\right)\left(p^\nu - \frac{(pq)q^\nu}{q^2}\right) \tag{5.5}$$

with the usual invariant functions $W_{1,2}$, see for instance [64].

In order to define the cross sections for longitudinally and transversely polarised virtual photons in (5.4) we work in the proton rest system, supposing

$$(q^\mu) = \begin{pmatrix} q^0 \\ 0 \\ 0 \\ |\boldsymbol{q}| \end{pmatrix} , \tag{5.6}$$

and define the following photon polarisation vectors:

$$\left(\varepsilon_\pm^\nu\right) = \mp\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ \pm i \\ 0 \end{pmatrix} , \tag{5.7}$$

$$(\varepsilon_L^\nu) = \frac{1}{Q} \begin{pmatrix} |\boldsymbol{q}| \\ 0 \\ 0 \\ q^0 \end{pmatrix} , \qquad (\varepsilon_L'^\nu) = (\varepsilon_L^\nu) - \frac{(q^\nu)}{Q} = \frac{1}{Q} \begin{pmatrix} |\boldsymbol{q}| - q^0 \\ 0 \\ 0 \\ q^0 - |\boldsymbol{q}| \end{pmatrix} . \tag{5.8}$$

With Hand's convention [134] the $\gamma^* p$ cross sections for transverse or longitudinal γ^* polarisation are

$$\begin{aligned} \sigma_T(W, Q^2) &= \frac{2\pi m_p}{W^2 - m_p^2}\, \varepsilon_+^{\mu *} e^2 W_{\mu\nu}\, \varepsilon_+^\nu = \frac{2\pi m_p}{W^2 - m_p^2}\, \varepsilon_-^{\mu *} e^2 W_{\mu\nu}\, \varepsilon_-^\nu \\ &= \frac{2\pi m_p}{W^2 - m_p^2}\, e^2 W_1(\nu, Q^2)\,, \end{aligned} \tag{5.9}$$

$$\begin{aligned} \sigma_L(W, Q^2) &= \frac{2\pi m_p}{W^2 - m_p^2}\, \varepsilon_L'^{\mu *} e^2 W_{\mu\nu}\, \varepsilon_L'^\nu = \frac{2\pi m_p}{W^2 - m_p^2}\, \varepsilon_L^{\mu *} e^2 W_{\mu\nu}\, \varepsilon_L^\nu \\ &= \frac{2\pi m_p}{W^2 - m_p^2} \left[e^2 W_2(\nu, Q^2)\, \frac{\nu^2 + Q^2}{Q^2} - e^2 W_1(\nu, Q^2) \right] . \end{aligned} \tag{5.10}$$

Note that due to gauge invariance the hadronic tensor $W^{\mu\nu}$ (5.5) satisfies

$$\begin{aligned} q_\mu\, W^{\mu\nu}(p, q) &= 0\,, \\ W^{\mu\nu}(p, q)\, q_\nu &= 0\,. \end{aligned} \tag{5.11}$$

Thus, choosing ε_L^ν or $\varepsilon_L'^\nu$ for the γ^* polarisation vector does not change the result of σ_L (5.10). However, as shown in [136], in applications of the dipole model it is *essential* to use $\varepsilon_L'^\nu$ and *not* ε_L^ν, in particular when one calculates the photon wave function from the Feynman rule for an incoming photon splitting into outgoing on-shell quark and antiquark. In that case the photon polarisation vector has to be chosen such that its components remain finite in the high energy limit, as is true for ε_L' but not for ε_L.

The standard structure function F_2 is defined as

$$\begin{aligned} F_2(W,Q^2) &:= \nu\, W_2(\nu,Q^2) \\ &= \frac{Q^2}{4\pi^2\alpha_{\rm em}}\left[\sigma_T(W,Q^2)+\sigma_L(W,Q^2)\right]\left(1+\frac{Q^2\left(W^2+Q^2+3m_p^2\right)}{\left(W^2-m_p^2\right)\left(W^2+Q^2-m_p^2\right)}\right)^{-1} \\ &= \frac{Q^2}{4\pi^2\alpha_{\rm em}}\left[\sigma_T(W,Q^2)+\sigma_L(W,Q^2)\right](1-x)+\mathcal{O}(m_p^2/W^2)\,. \end{aligned} \tag{5.12}$$

In the high energy limit, $W \gg Q, m_p$, this simplifies to the commonly used form

$$F_2(W,Q^2) = \frac{Q^2}{4\pi^2\alpha_{\rm em}}\left[\sigma_T(W,Q^2)+\sigma_L(W,Q^2)\right] \tag{5.13}$$

up to terms of order $\mathcal{O}(Q^2/W^2)$. Similarly, we use for the standard transverse and longitudinal structure functions F_T and F_L:

$$F_T(W,Q^2) = \frac{Q^2}{4\pi^2\alpha_{\rm em}}\sigma_T(W,Q^2)\,, \tag{5.14}$$

$$F_L(W,Q^2) = \frac{Q^2}{4\pi^2\alpha_{\rm em}}\sigma_L(W,Q^2)\,. \tag{5.15}$$

In the following we shall always use the relation (5.13) valid in the high energy limit unless explicitly noted otherwise. In section 6.2 we shall discuss how the results obtained there are modified for finite Bjorken-x if we use the exact formula (5.12) instead of (5.13). We note that one could also consider (5.13) as the defining equation for σ_T and σ_L. This would correspond to a different choice of flux factor for the virtual photons as compared to [134]. The considerations of section 6 of [136] show, however, that Hand's convention [134] is the natural one for the dipole picture; see especially (121)-(128) of [136].

In [135, 136] non-perturbative methods were employed in order to work towards a foundation of the dipole model for quasi-real and virtual photon induced reactions at high energies. The result for $W^{\mu\nu}$ (2.5) obtained there is shown diagrammatically in figure 5.1. In the high energy limit,

$$q^0 \to \infty\,, \tag{5.16}$$

taken in the proton rest frame, they find a factorisation into photon wave function and dipole-proton scattering parts. The wave function parts contain the renormalised $\gamma q\bar{q}$ vertex function plus a rescattering term. The dipole-proton scattering is built from diagrams of type (a) where the quark lines go through from right to left and type (b) where the quark lines do not go through, see figure 5.1. To get from this point to the standard dipole picture requires a number of *assumptions* and approximations as listed in [136]:

(i) Quarks of flavour q have a mass shell m_q and can be considered as asymptotic states.

(ii) The rescattering terms are dropped and the $\gamma q\bar{q}$ vertex functions are replaced by the lowest order terms in perturbation theory.

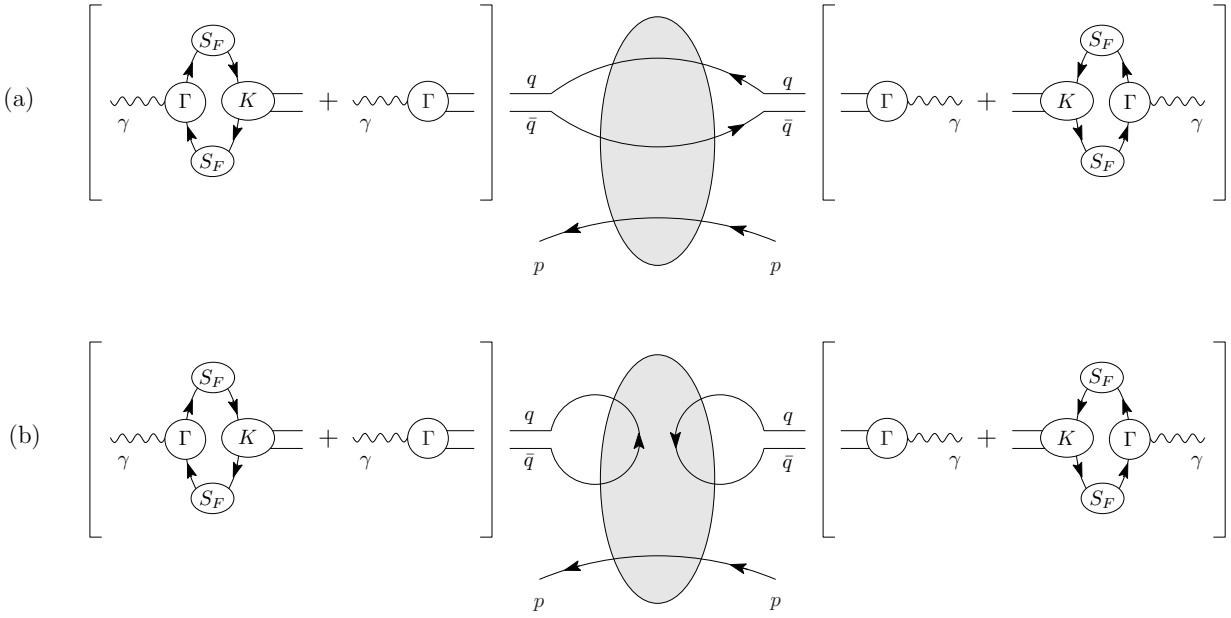

Figure 5.1: Quark skeleton diagrams for the photon-proton scattering cross section in the high energy limit. The shaded area indicates a functional integration over gluon field configurations, and Γ, K, S_F are the renormalised $\gamma q\bar{q}$ vertex, the renormalised kernel for $q'\bar{q}'$ to $q\bar{q}$ scattering, and the renormalised quark propagator, respectively. The diagrams are to be read from right to left.

(iii) The T-matrix element for the dipole-proton scattering is diagonal in the quark flavour q, in α and in $\boldsymbol{r}$. Here α is the longitudinal momentum fraction of the photon carried by the quark, and $\boldsymbol{r}$ is the two-dimensional vector from the antiquark to the quark in transverse position space. Further, the T-matrix element is proportional to the unit matrix in the space of spin orientations of the quark and antiquark in the dipole.

(iv) In the T-matrix element for the dipole-proton scattering only the contribution of type (a) is kept while that of type (b) is neglected, see figure 5.1.

(v) The proton spin averaged reduced matrix element for a given quark flavour q depends only on the dipole size $r \equiv \sqrt{\boldsymbol{r}^2}$ and on $W^2 = (p+q)^2$.

With these assumptions the authors arrive indeed at the standard dipole picture as used extensively in the literature, see the previous chapter. In detail, they find for the total cross section

$$\sigma_T(W, Q^2) = \sum_q \int \mathrm{d}^2 r \, w_T^{(q)}(r, Q^2) \, \hat{\sigma}^{(q)}(r, W) \,, \tag{5.17}$$

$$\sigma_L(W, Q^2) = \sum_q \int \mathrm{d}^2 r \, w_L^{(q)}(r, Q^2) \, \hat{\sigma}^{(q)}(r, W) \,. \tag{5.18}$$

with the photon densities $w_{T,L}^{(q)}$. They are defined by the absolute squares of the photon wave functions upon integration over α and summation over quark spins as functions of the dipole size r and of Q^2,

$$\begin{aligned} w_T^{(q)}(r,Q^2) &:= \sum_{\lambda,\lambda'} \int_0^1 \mathrm{d}\alpha \, \left| \psi_{\gamma,\lambda\lambda'}^{(q)\,\mu}(\alpha,\boldsymbol{r},Q)\, \varepsilon_{+\,\mu} \right|^2 , \\ &= \sum_{\lambda,\lambda'} \int_0^1 \mathrm{d}\alpha \, \left| \psi_{\gamma,\lambda\lambda'}^{(q)\,\mu}(\alpha,\boldsymbol{r},Q)\, \varepsilon_{-\,\mu} \right|^2 , \end{aligned} \tag{5.19}$$

$$w_L^{(q)}(r,Q^2) := \sum_{\lambda,\lambda'} \int_0^1 \mathrm{d}\alpha \, \left| \psi_{\lambda\lambda'}^{(q)\,\mu}(\alpha,\boldsymbol{r},Q)\, \varepsilon'_{L\,\mu} \right|^2 \tag{5.20}$$

and yield exactly the expressions (4.6) and (4.7), respectively. Note however, that (4.4) and (4.5) differ from (5.17) and (5.18) in the energy dependence of $\hat{\sigma}^{(q)}$, We shall return to this issue after completing further definitions.

In this thesis, we shall also consider structure functions with the production of specific quark flavours. In the dipole model the cross sections for production of a specific quark flavour are obtained as in (4.4), (4.5) but without the summation over quark flavour q,

$$\sigma_{T,L}^{(q)}(W,Q^2) = \int \mathrm{d}^2 r \, w_{T,L}^{(q)}(r,Q^2)\, \hat{\sigma}^{(q)}(r,W) \,. \tag{5.21}$$

We set as in (5.14) and (5.15) for the flavour specific structure functions

$$F_{T,L}^{(q)} = \frac{Q^2}{4\pi^2 \alpha_{\mathrm{em}}} \left[\sigma_T^{(q)}(W,Q^2) + \sigma_L^{(c)}(W,Q^2) \right] \tag{5.22}$$

and

$$F_2^{(q)} = F_T^{(q)} + F_L^{(q)} \,. \tag{5.23}$$

This assumes that all quarks of flavour q are exclusively produced directly by the inital γ^*. In particular, we consider $F_2^{(c)}$ as defined by (5.23) in section 6.1. There, we thus neglect associated charm-anticharm production in reactions initiated by other quark flavours coupling directly to the photon.

In [136] it is stressed that the dipole-proton cross section is naturally independent of Q^2, and thus its correct energy variable is given exclusively by W:

$$\hat{\sigma}^{(q)} = \hat{\sigma}^{(q)}(r,W) \,. \tag{5.24}$$

This excludes in particular the choice of Bjorken-x instead of W, since this introduces a dependence on Q^2 in addition to W, see (5.2). It was argued in [136] that using x instead of W requires additional assumptions which are difficult to assess quantitatively and which go beyond those listed above. In their derivation of the dipole picture the dipole cross section arises from a T-matrix element for scattering of a dipole state on the proton.

There, the key feature of these dipole states is that they consist of a quark and an anti-quark described by asymptotic states. The dipole states are then independent of Q^2 in the high energy limit. By a smearing in $\boldsymbol{r}$ and α they can be viewed as hadron analogues, whose normalisation is independent of continuous internal degrees of freedom. But also the mean squared invariant mass of such smeared dipole states is independent of Q^2 at large W. Nevertheless, the energy variable x – and hence a Q^2-dependence – is frequently used in popular models for the dipole cross section, such as in (4.9). Sometimes also other dependencies on Q^2 are introduced. Furthermore, the perturbative gluon density $g(x, Q^2)$ naturally depends on x and Q^2 and its often assumed connection (4.12) to the dipole cross section suggests that the latter is also Q^2 dependent. The only detailed derivation [136] of the dipole picture available uses the high energy limit

$$W \to \infty\,, \qquad Q^2 \text{ fixed}\,. \tag{5.25}$$

The formula (4.12) on the other hand is based on a comparison with the perturbative result obtained in the double leading logarithmic approximation (DLLA) for the limit

$$W \to \infty\,, \qquad Q^2 \to \infty\,, \qquad \frac{Q^2}{W^2} \text{ fixed}\,. \tag{5.26}$$

The authors of [136] point out, that a step-by-step comparison may be intrinsically difficult due to these different limites, since first taking (5.25) and then letting Q^2 become large is not necessarily equivalent to (5.26). Actually, the choice of the energy variable can be crucial at moderate values of Q^2 and W, and we shall explicitly demonstrate that different choices may in general not be considered effectively equivalent, see section 6.2 and in particular section 7.2.

For arbitrary kinematics various contributions to the forward scattering amplitude are potentially relevant which cannot be interpreted in terms of the dipole picture. In the high energy limit $q^0 \to \infty$ the two contributions shown in figure 5.1 are enhanced with respect to the others due to a pinch singularity of two poles in the integration over the quark offshellness. Let k^0 be the energy of the quark and k'^0 be the energy of the anti-quark. Defining the energy mismatch of the quark anti-quark dipole as

$$\Delta E := k^0 + k'^0 - q^0 \tag{5.27}$$

the pinch condition is

$$\Delta E \to 0\,. \tag{5.28}$$

In the high energy limit, the pinch condition (5.28) effectively restricts the range of the α integration to the finite range

$$0 < \alpha < 1\,. \tag{5.29}$$

Furthermore, the pinch condition implies a natural UV-cutoff also for the transverse momenta of the quark and the anti-quark. Thus (5.28) is a crucial element for the dipole picture to be valid. In chapter 8 we calculate the ΔE values which actually occur for a typical dipole model. Considering to what extent the condition (5.28) is fulfilled represents

an important self-consistency check and potentially allows to restrict the kinematical range of validity of the dipole model.

Discarding the rescattering effects for the quark and anti-quark forming the dipole, assumption (ii), and neglecting contributions of type (b) in figure 5.1, assumption (iv), is justified in the perturbative regime, since they are of higher order in the strong coupling constant α_s. In the non-perturbative regime of low Q^2 however it may very well be that these effects beyond the standard dipole picture are non-negligible.

5.2 Photon wave function

In this section, we provide details for the photon wave functions, since they are an important ingredient for our analysis in the following chapters. In the proton rest frame, we decompose the 4-momentum of the quark, k, and that of the antiquark, k', as follows:

$$\begin{aligned} \boldsymbol{k} &= \alpha \boldsymbol{q} + \boldsymbol{k}_T \,, & k^0 &= \sqrt{\boldsymbol{k}^2 + m_q^2} \,, \\ \boldsymbol{k}' &= (1-\alpha)\boldsymbol{q} - \boldsymbol{k}_T \,, & k'^0 &= \sqrt{\boldsymbol{k}'^2 + m_q^2} \,. \end{aligned} \tag{5.30}$$

We shall first evaluate certain expressions without any explicit restriction on $\boldsymbol{k}_T$ and α,

$$-\infty < \alpha < \infty \,. \tag{5.31}$$

This serves as a preparation for consistency checks in chapter 8. In order to arrive at the standard dipole picture formulae, we apply the restriction (5.29) at a later stage of the calculation and consider the transverse UV-regulator to be removed by taking the asymptotic limit in the final expressions. In this way we will get the standard formulae for the photon wave functions, see (4.6) and (4.7). From this we infer that in this standard scheme the corresponding dipole cross section should assure, that physical observables such as $\sigma_{T,L}$ are insensitive to the properties of the photon wave functions at large k_T, since there the dipole picture is not justified due to large ΔE.

We define the photon wave functions to leading order α_{em}, α_s in momentum space according to (40) in [136], but without the explicit transverse UV-cutoff:

$$\tilde{\psi}^{(q)\,\nu}_{\gamma,\lambda\lambda'}(\alpha, k_T, Q) = Q_q \, \frac{N}{\Delta E} \frac{|\mathbf{q}|}{2\pi 2k^0 2k'^0} \, \bar{u}_\lambda(k) \gamma^\nu v_{\lambda'}(k') \,, \tag{5.32}$$

where γ^ν are the Dirac matrices, see for instance [64]. The spinors u and v are solutions to the free Dirac equation,

$$u_\lambda(k) = \frac{1}{\sqrt{k^0 + m_q}} \begin{pmatrix} (k^0 + m_q)\, \chi_\lambda \\ \boldsymbol{\sigma} \boldsymbol{k}\, \chi_\lambda \end{pmatrix} \,, \tag{5.33}$$

$$v_{\lambda'}(k') = \frac{-1}{\sqrt{k'^0 + m_q}} \begin{pmatrix} \boldsymbol{\sigma} \boldsymbol{k}'\, \epsilon\, \chi^*_{\lambda'} \\ (k'^0 + m_q)\, \epsilon\, \chi^*_{\lambda'} \end{pmatrix} \tag{5.34}$$

with

$$\chi_{+\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \chi_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{5.35}$$

The normalisation

$$N = -2\sqrt{N_c \pi} e \sqrt{\alpha(1-\alpha)} \tag{5.36}$$

is chosen, with the proton charge $e \equiv \sqrt{4\pi\alpha_{\mathrm{em}}}$ and the number of colours $N_c = 3$. The wave functions for transversely polarised photons are defined by

$$\tilde{\psi}^{(q)\,\pm}_{\gamma,\lambda\lambda'}(\alpha, \boldsymbol{k}_T, Q) := -\varepsilon_{\pm\nu}\, \tilde{\psi}^{(q)\nu}_{\gamma,\lambda\lambda'}(\alpha, \boldsymbol{k}_T, Q)\,, \tag{5.37}$$

and for longitudinally polarised photons by

$$\tilde{\psi}^{(q)\,L}_{\gamma,\lambda\lambda'}(\alpha, \boldsymbol{k}_T, Q) := -\varepsilon'_{L\nu}\, \tilde{\psi}^{(q)\,\nu}_{\gamma,\lambda\lambda'}(\alpha, \boldsymbol{k}_T, Q)\,, \tag{5.38}$$

which give, see (52) and (58) of [136]),

$$\tilde{\psi}^{(q)\,\pm}_{\lambda\lambda'}(\alpha, \boldsymbol{k}_T, Q^2) = \mp \frac{N}{\sqrt{2}} Q_q \frac{1}{\Delta E} \frac{|\boldsymbol{q}|}{2\pi 2k^0 2k'^0}\, \bar{u}_\lambda(k) \left(\gamma^1 \pm i\gamma^2\right) v_{\lambda'}(k') \tag{5.39}$$

$$\tilde{\psi}^{(q)\,L}_{\lambda\lambda'}(\alpha, \boldsymbol{k}_T, Q^2) = -NQ_q \frac{1}{\Delta E} \frac{|\boldsymbol{q}|}{2\pi 2k^0 2k'^0} \frac{|\boldsymbol{q}| - q^0}{Q}\, \bar{u}_\lambda(k) \left(\gamma^0 + \gamma^3\right) v_{\lambda'}(k')\,. \tag{5.40}$$

Explicitly, we find by evaluating (5.39) and (5.40):

$$\begin{aligned} \tilde{\psi}^{(q)\,\pm}_{\lambda\lambda'}(\alpha, \boldsymbol{k}_T, Q) = \mp \frac{N}{\sqrt{2}} Q_q \frac{1}{\Delta E} \frac{|\boldsymbol{q}|}{2\pi k^0 2k'^0} \frac{2}{\sqrt{k^0 + m_q}\sqrt{k'^0 + m_q}} \\ \cdot \Bigg[\pm \left((k^0 + m_q)(k'^0 + m_q) - \alpha(1-\alpha)\boldsymbol{q}^2\right) \delta_{\lambda,\lambda'} \delta_{\lambda,\pm\frac{1}{2}} \\ + e^{\pm i\phi_k} k_T\, |\boldsymbol{q}|\, \delta_{\lambda,-\lambda'} \left(\alpha \delta_{\lambda,\pm\frac{1}{2}} - (1-\alpha)\delta_{\lambda,\mp\frac{1}{2}}\right) \pm e^{\pm i2\phi_k} k_T^2 \delta_{\lambda,\lambda'} \delta_{\lambda,\mp\frac{1}{2}} \Bigg], \end{aligned} \tag{5.41}$$

$$\begin{aligned} \tilde{\psi}^{(q)\,L}_{\lambda\lambda'}(\alpha, \boldsymbol{k}_T, Q) = -NQ_q \frac{1}{\Delta E} \frac{|\boldsymbol{q}|}{2\pi k^0 2k'^0} \frac{|\boldsymbol{q}| - q^0}{Q} \frac{1}{\sqrt{k^0 + m_q}\sqrt{k'^0 + m_q}} \\ \cdot \Bigg[\left(k_T^2 + (k^0 + m_q + \alpha\,|\boldsymbol{q}|)(k'^0 + m_q + (1-\alpha)\,|\boldsymbol{q}|)\right) \delta_{\lambda,-\lambda'} \\ + e^{-i(\mathrm{sign}\,\lambda)\phi_k} (\mathrm{sign}\,\lambda)(k^0 - k'^0 - (1-2\alpha)\,|\boldsymbol{q}|) k_T \delta_{\lambda,\lambda'} \Bigg], \end{aligned} \tag{5.42}$$

where $\phi_k = \arg(k_{T1} + ik_{T2})$ with k_{T1} and k_{T2} being the 1- and 2-component of the vector $\boldsymbol{k}_T$, respectively. The photon wave functions in transverse position space are obtained by

a Fourier transformation from their momentum space representation. We write as in (42) of [136]

$$\psi^{(q)\,\pm}_{\gamma,\,\lambda\lambda'}(\alpha,\boldsymbol{r},Q) = \int \frac{\mathrm{d}^2k_T}{(2\pi)^2} e^{i\boldsymbol{k}_T\boldsymbol{r}} \tilde{\psi}^{(q)\,\pm}_{\gamma,\,\lambda\lambda'}(\alpha,\boldsymbol{k}_T,Q)\,, \tag{5.43}$$

$$\psi^{(q)\,L}_{\gamma,\,\lambda\lambda'}(\alpha,\boldsymbol{r},Q) = \int \frac{\mathrm{d}^2k_T}{(2\pi)^2} e^{i\boldsymbol{k}_T\boldsymbol{r}} \tilde{\psi}^{(q)\,L}_{\gamma,\,\lambda\lambda'}(\alpha,\boldsymbol{k}_T,Q)\,. \tag{5.44}$$

From (5.41), (5.42) we find

$$\begin{aligned}\tilde{\psi}^{(q)\,\pm}_{\lambda\lambda'}(\alpha,\boldsymbol{r},Q) = {}& \frac{\mp NQ_q}{8\sqrt{2}\pi^2}\int_0^\infty \mathrm{d}k_T \frac{k_T\,|\boldsymbol{q}|}{\Delta E\,k^0k'^0\sqrt{k^0+m_q}\sqrt{k'^0+m_q}} \\ & \Big[\pm\,\delta_{\lambda,\lambda'}\left(\delta_{\lambda,\pm\frac{1}{2}}\,\left((k^0+m_q)(k'^0+m_q)-\alpha(1-\alpha)\,|\boldsymbol{q}|^2\right)\,J_0(rk_T)\right. \\ & \cdot \qquad\qquad \left. -\,\delta_{\lambda,\mp\frac{1}{2}}\,e^{\pm 2i\varphi_r}k_T^2\,J_2(rk_T)\right) \\ & +\,\delta_{\lambda,-\lambda'}\,ie^{\pm\varphi_r}k_T\,|\boldsymbol{q}|\,(\delta_{\lambda,\pm\frac{1}{2}}\,\alpha-\delta_{\lambda,\mp\frac{1}{2}}\,(1-\alpha))J_1(rk_T)\Big],\end{aligned} \tag{5.45}$$

$$\begin{aligned}\tilde{\psi}^{(q)\,L}_{\lambda\lambda'}(\alpha,\boldsymbol{r},Q) = {}& \frac{NQ_q}{16\pi^2}\int_0^\infty \mathrm{d}k_T \frac{k_T\,|\boldsymbol{q}|}{\Delta E\,k^0k'^0\sqrt{k^0+m_q}\sqrt{k'^0+m_q}}\frac{|\boldsymbol{q}|-q^0}{k_T} \\ & \cdot\Big[\delta_{\lambda,\lambda'}\,\mathrm{sign}(\lambda)ie^{-i\,\mathrm{sign}(\lambda)\varphi_r}\,k_T(k^0-k'^0-(1-2\alpha)\,|\boldsymbol{q}|)\,J_1(k_tr) \\ & +\,\delta_{\lambda,-\lambda'}\left(k_T^2+(k^0+m_q+\alpha\,|\boldsymbol{q}|)(k'^0+m_q+(1-\alpha)\,|\boldsymbol{q}|)\right)J_0(rk_T)\Big],\end{aligned} \tag{5.46}$$

where J_i $(i=0,1,2)$ are Bessel functions and $\varphi_r = \arg(r_1+ir_2)$ with r_1 and r_2 being the 1- and 2-component of the vector $\boldsymbol{r}$. In the high-energy limit $|\boldsymbol{q}|\to\infty$ we find from (5.41) and (5.42)

$$\begin{aligned}\tilde{\psi}^{(q)\,\pm}_{\lambda\lambda'}(\alpha,\boldsymbol{k}_T,Q) = \frac{\sqrt{2N_c\alpha_{\mathrm{em}}}Q_q}{\alpha(1-\alpha)Q^2+k_T^2+m_q^2}\Big[& \pm\,k_Te^{\pm i\phi_k}\delta_{\lambda,-\lambda'}\big(\alpha\delta_{\lambda,\pm\frac{1}{2}}-(1-\alpha)\delta_{\lambda,\mp\frac{1}{2}}\big) \\ & +\,m_q\delta_{\lambda,\lambda'}\delta_{\lambda,\pm\frac{1}{2}}\Big],\end{aligned} \tag{5.47}$$

$$\tilde{\psi}^{(q)\,L}_{\lambda\lambda'}(\alpha,\boldsymbol{k}_T,Q) = -\frac{2\sqrt{N_c\alpha_{\mathrm{em}}}Q_q}{\alpha(1-\alpha)Q^2+k_T^2+m_q^2}\,\alpha(1-\alpha)Q\,\delta_{\lambda,-\lambda'}\,, \tag{5.48}$$

which agrees with (53) and (60) of [136]. With (5.43) and (5.44) we obtain their position

space representation as

$$\psi^{(q)\pm}_{\gamma,\lambda\lambda'}(\alpha,\boldsymbol{r},Q) = \frac{\sqrt{N_c\alpha_{\rm em}}}{\sqrt{2}\pi}\,Q_q\left[\pm ie^{\pm i\phi_r}\epsilon_q\left(\alpha\,\delta_{\lambda,\pm\frac{1}{2}}\delta_{\lambda',-\lambda} - (1-\alpha)\,\delta_{\lambda,\mp\frac{1}{2}}\delta_{\lambda',-\lambda}\right)K_1(\epsilon_q r)\right.$$
$$\left. + m_q\,\delta_{\lambda,\pm\frac{1}{2}}\delta_{\lambda',\lambda}K_0(\epsilon_q r)\right], \tag{5.49}$$

$$\psi^{(q)L}_{\gamma,\lambda\lambda'}(\alpha,\boldsymbol{r},Q) = -\frac{\sqrt{N_c\alpha_{\rm em}}}{\pi}\,Q_q\;\alpha(1-\alpha)\,Q\,\delta_{\lambda',-\lambda}\,K_0(\epsilon_q r) \tag{5.50}$$

in agreement with (54) and (61) of [136]. These give the standard dipole picture formulae (4.6) and (4.7). If not noted otherwise, we shall always use these high-energy expressions for the photon wave functions.

In [136] a detailed derivation for this high-energy approximation is given, which in particular involves neglection of terms such as $\boldsymbol{k}_T^2/(\alpha^2\boldsymbol{q}^2)$ and $\boldsymbol{k}_T^2/((1-\alpha)^2\boldsymbol{q}^2)$ with respect to 1. However, for some given $|\boldsymbol{q}|$ those terms become non-negligible not only if k_T is large but also if α is close to 0 and 1 respectively. If relevant contributions to some observable depend on the photon wave functions in this kinematical region, the high-energy approximation could become invalid. This is potentially relevant when considering distributions in the Ioffe times, in particular for short Ioffe times. We shall calculate the distributions with and without the high-energy approximation in order to quantify this effect.

5.3 Photon densities

In this section we shall discuss the photon densities in some detail, since they are an important ingredient in the following studies.

We define the unintegrated photon densities by

$$v_T^{(q)}(\alpha,r,Q^2) = \sum_{\lambda,\lambda'}\left|\psi^{(q)\,+}_{\gamma,\lambda\lambda'}(\alpha,\boldsymbol{r},Q)\right|^2, \tag{5.51}$$

$$v_L^{(q)}(\alpha,r,Q^2) = \sum_{\lambda,\lambda'}\left|\psi^{(q)\,L}_{\lambda\lambda'}(\alpha,\boldsymbol{r},Q)\right|^2. \tag{5.52}$$

In the high energy limit we find from (5.49) and (5.50)

$$v_T^{(q)}(\alpha,r,Q^2) = \frac{N_c\alpha_{\rm em}Q_q^2}{2\pi^2}\left[(\alpha^2+(1-\alpha)^2)\epsilon_q^2K_1^2(\epsilon_q r) + m_q^2K_0^2(\epsilon_q r)\right], \tag{5.53}$$

$$v_L^{(q)}(\alpha,r,Q^2) = \frac{2N_c\alpha_{\rm em}Q_q^2}{\pi^2}Q^2\left(\alpha(1-\alpha)\right)^2K_0^2(\epsilon_q r)\,. \tag{5.54}$$

Upon integration over α we get from v_T and v_L the partially integrated photon densities

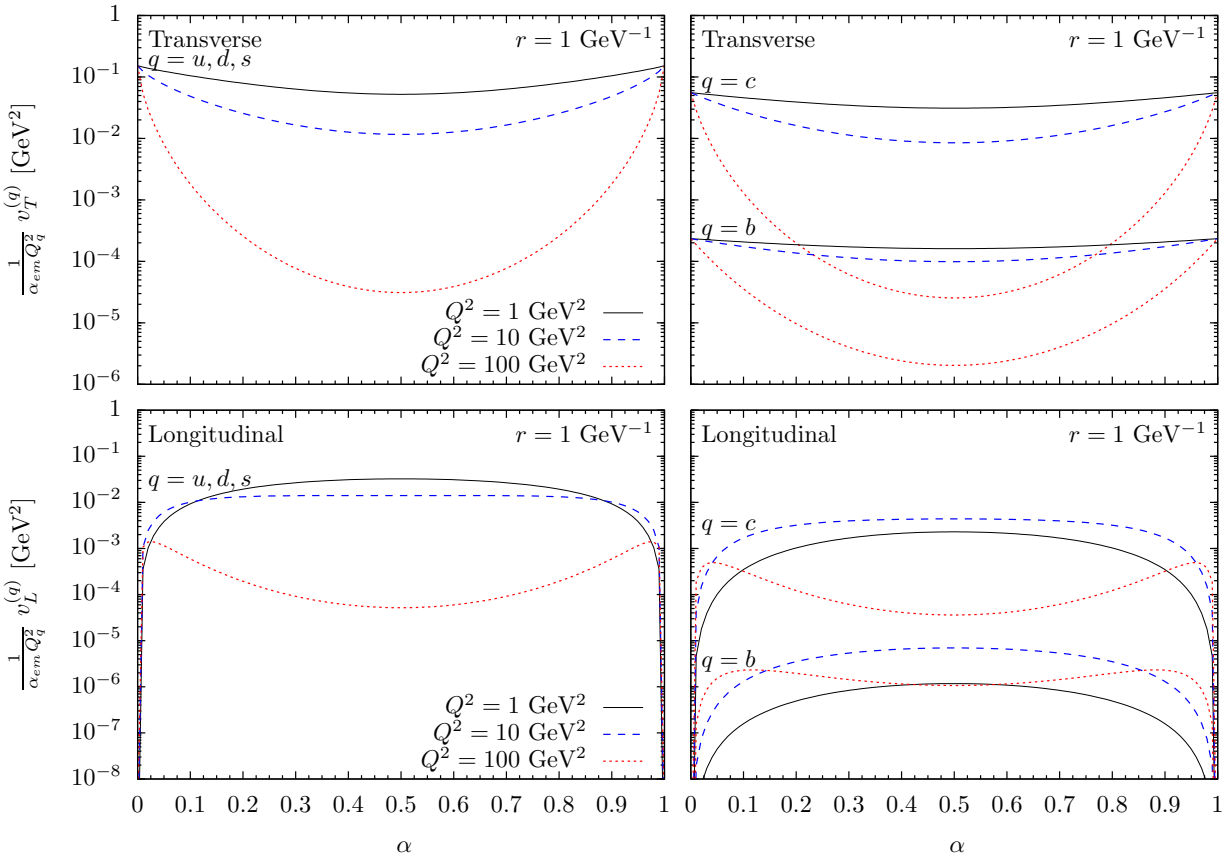

Figure 5.2: Photon densities normalised to $\alpha_{\text{em}} Q_q^2$ in dependence of the quark longitudinal momentum α.

w_T (5.19) and w_L (5.20) entering the dipole formulae (4.4) and (4.5):

$$w_T^{(q)}(r, Q^2) = \int_0^1 \mathrm{d}\alpha \, v_T^{(q)}(\alpha, r, Q^2) \,, \tag{5.55}$$

$$w_L^{(q)}(r, Q^2) = \int_0^1 \mathrm{d}\alpha \, v_L^{(q)}(\alpha, r, Q^2) \,. \tag{5.56}$$

Figure 5.2 shows the α dependencies of the unintegrated photon densities v_T, v_L. Here and in the following we use

$$m_q = 0 \qquad \text{for } u, d, s\text{-quarks}, \tag{5.57}$$

$$m_q = 1.3\,\text{GeV} \qquad \text{for } c\text{-quarks}, \tag{5.58}$$

$$m_q = 4.6\,\text{GeV} \qquad \text{for } b\text{-quarks} \tag{5.59}$$

for our numerical evaluations. The graphs for transversely polarised photons (upper plots) show a rise in v_T for approaching the longitudinal momentum endpoints $\alpha = 0, 1$ and a minimum at $\alpha = 1/2$. This means for the splitting of a transversely polarised photon into

a dipole it is most likely that either the quark or the anti-quark takes most of the photon momentum. This effect becomes increasingly pronounced with larger Q^2. For longitudinally polarised photons (lower plots) the density v_L peaks at $\alpha = 1/2$ and drops when approaching the endpoints. With increasing Q^2 this behaviour changes and v_L develops two peaks close to (but not at) the endpoints. Thus, in dipoles produced from longitudinally polarised photons with small virtuality, the quark and anti-quark typically have approximately equal momenta. For high virtualities it most probably either the quark or the anti-quark which carries most of the photon momentum. All of these effects are generally less pronounced for non-zero quark masses (right plots) than for the massless case (left plots). Furthermore, increasing quark masses lower the photon densities.

Let us consider the behaviour of the partially integrated photon densities at small and large distances r.

At small distances, $rQ,\ rm_q \ll 1$, we use

$$K_0(x) \xrightarrow{x\to 0} -\ln x\,, \tag{5.60}$$

$$K_n(x) \xrightarrow{x\to 0} \frac{\Gamma(n)}{2}\left(\frac{2}{x}\right)^n \qquad \text{for } n>0\,, \tag{5.61}$$

and find

$$w_T^{(q)}(r,Q^2) \xrightarrow{r\to 0} \frac{N_c\alpha_{\rm em}Q_q^2}{2\pi^2}Q^2\frac{2}{3}\,\frac{1}{(Qr)^2}\,, \tag{5.62}$$

$$w_L^{(q)}(r,Q^2) \xrightarrow{r\to 0} \frac{2N_c\alpha_{\rm em}Q_q^2}{\pi^2}Q^2\frac{1}{30}\,\ln^2(Qr)\,. \tag{5.63}$$

Note that these leading terms are independent of m_q.

For the limites at large distances we consider massless quarks and massive quarks separately. For massless quarks, $m_q = 0$, we find from (5.53)-(5.56) without any approximation

$$w_T^{(q)}(r,Q^2) = \frac{N_c\alpha_{\rm em}Q_q^2}{2\pi^2}Q^2\frac{2\pi}{(Qr)^6}\left(2G_{2,4}^{3,1}\left(\frac{1}{4}(Qr)^2\,\middle|\,\begin{matrix}2,7/2\\2,3,4,3/2\end{matrix}\right) - G_{2,4}^{3,1}\left(\frac{1}{4}(Qr)^2\,\middle|\,\begin{matrix}1,7/2\\2,3,4,1/2\end{matrix}\right)\right), \tag{5.64}$$

$$w_L^{(q)}(r,Q^2) = \frac{2N_c\alpha_{\rm em}Q_q^2}{\pi^2}Q^2\frac{2\pi}{(Qr)^6}G_{2,4}^{3,1}\left(\frac{1}{4}(Qr)^2\,\middle|\,\begin{matrix}-2,1/2\\0,0,0,-5/2\end{matrix}\right) \tag{5.65}$$

with Meijer's G function in the notation [124]. From this we get at large distances r, $rQ \gg 1$:

$$w_T^{(q)}(r,Q^2) \xrightarrow{r\to\infty} \frac{N_c\alpha_{\rm em}Q_q^2}{2\pi^2}Q^2\frac{8}{3}\,\frac{1}{(Qr)^4}\,, \tag{5.66}$$

$$w_L^{(q)}(r,Q^2) \xrightarrow{r\to\infty} \frac{2N_c\alpha_{\rm em}Q_q^2}{\pi^2}Q^2\frac{64}{15}\,\frac{1}{(Qr)^6}\,. \tag{5.67}$$

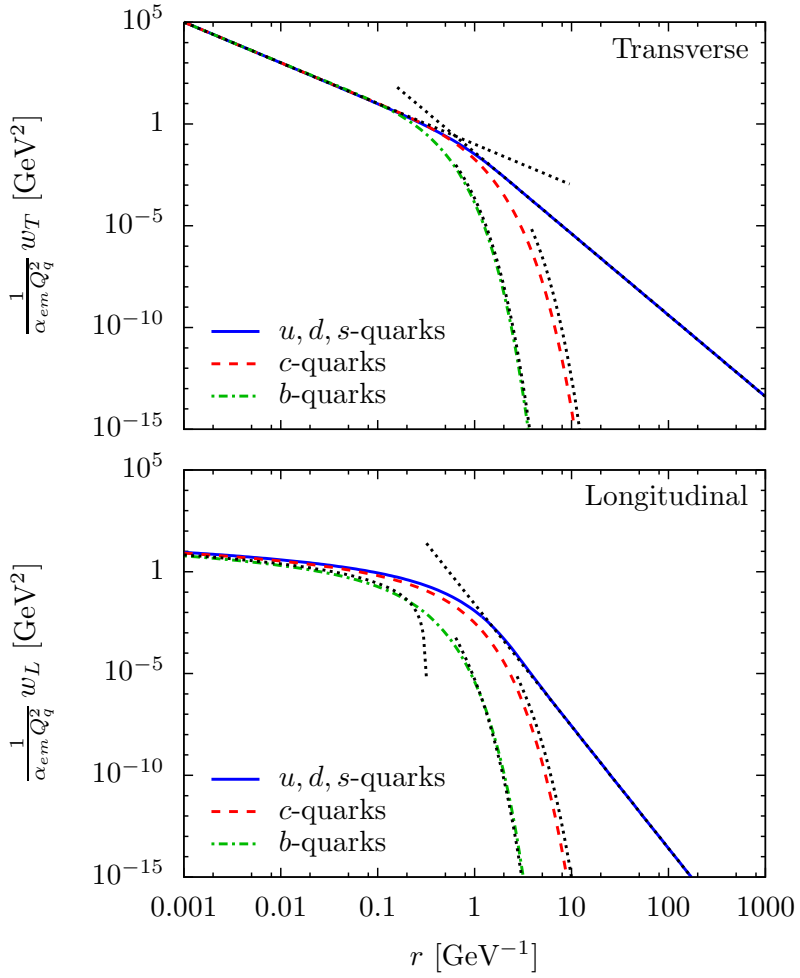

Figure 5.3: Partially integrated photon densities $w_T^{(q)}(r, Q^2)$ (upper plot) and $w_L^{(q)}(r, Q^2)$ (lower plot) both normalised to $\alpha_{\mathrm{em}} Q_q^2$ as a function of the dipole size r. In addition, the leading terms for small and large r (dotted curves) are shown. The photon virtuality is fixed to the value $Q^2 = 10\,\mathrm{GeV}^2$.

For massive quarks, $m_q > 0$, and large distances, $rQ,\ r\,m_q \gg 1$, our calculation for $w_T^{(q)}$ and our guess for $w_L^{(q)}$ give:

$$w_T^{(q)}(r, Q^2) \xrightarrow{r\to\infty} \frac{N_c \alpha_{\mathrm{em}} Q_q^2}{2\pi^2} Q^2 \frac{\pi m_q}{Q} \frac{\exp(-2m_q r)}{Qr}\,, \tag{5.68}$$

$$w_L^{(q)}(r, Q^2) \xrightarrow{r\to\infty} \frac{2N_c \alpha_{\mathrm{em}} Q_q^2}{\pi^2} Q^2 \frac{\pi Q}{2m_q} \bar{\alpha}^2(m_q/Q) \frac{\exp(-2m_q r)}{(Qr)^3} \tag{5.69}$$

with some r-independent function $\bar{\alpha}^2(m_q^2/Q^2)$, for which we find in the case $m_q < Q$ the numerical value $\bar{\alpha}^2(m_q^2/Q^2) \approx 0.25$.

Figure 5.3 shows the photon densities in dependence of the dipole size r. In addition, their asymptotic expressions discussed above are plotted. The transverse densities $w_T^{(q)}$

(upper plot) and longitudinal densities $w_L^{(q)}$ (lower plot) show a significant difference in their distributions. Compared to the longitudinal densities, the transverse densities rise faster towards small r and decrease less rapidly at large r. This means that typically a broader range of dipole sizes r will contribute to σ_T than to σ_L, see also chapter 7. How much of this effect survives depends of course on the dipole cross section. Further we see that large dipoles become strongly suppressed for increasing quark mass. The figure illustrates the photon densities reaching the discussed asymptotic expressions for the case of small r. At large r the asymptotic behaviour is visible only for the massless quarks. The aforementioned suppression of massive quarks at large distances is so strong, that their densities at actually large r are beyond the double-logarithmic plot range of figure 5.3. Our numerical tests indicate that the presented asymptotic expressions are indeed approached at large r, and the corresponding $w_{T,L}^{(q)}$ values are extremely small.

Chapter 6

Bounds on ratios of DIS observables

In this chapter, we derive bounds for ratios of structure functions from the relations (4.4), (4.5), (5.21). These bounds will rest on the explicit forms of the photon densities $w_{T,L}^{(q)}$ (5.55), (5.56) and on the non-negativity of the dipole-proton cross sections

$$\hat{\sigma}^{(q)}(r, W) \geq 0\,. \tag{6.1}$$

The bounds derived in section 6.1 remain unchanged if we assume that the dipole cross sections $\hat{\sigma}^{(q)}$ are functions of r and Bjorken-x instead of r and W. The bounds derived in section 6.2, on the other hand, depend crucially on the functional dependence indicated in (6.1).

6.1 Longitudinal and charm part of F_2

In this section we consider the structure functions F_L, $F_2^{(c)}$, and F_2 at fixed values of Q^2 and W. Arranging them into a three-vector gives according to the dipole formula

$$\begin{pmatrix} F_L(W,Q^2) \\ F_2^{(c)}(W,Q^2) \\ F_2(W,Q^2) \end{pmatrix} = \sum_q \int \mathrm{d}^2 r\, \frac{\hat{\sigma}^{(q)}(r,W)}{4\pi^2 \alpha_{\mathrm{em}}} \begin{pmatrix} f_L^{(q)}(r,Q^2) \\ \delta_{q,c}\, f^{(c)}(r,Q^2) \\ f^{(q)}(r,Q^2) \end{pmatrix}, \tag{6.2}$$

with

$$f^{(q)}(r,Q^2) = Q^2 \left[w_T^{(q)}(r,Q^2) + w_L^{(q)}(r,Q^2) \right], \tag{6.3}$$

$$f_L^{(q)}(r,Q^2) = Q^2\, w_L^{(q)}(r,Q^2)\,. \tag{6.4}$$

Note that the second entry in the vector in (6.2) receives a contribution only from the charm quark, as indicated by the Kronecker delta symbol. In the following we will make use of a geometrical interpretation of (6.2) in order to obtain correlated bounds on the structure functions involved here. Due to the Kronecker symbol the case at hand is somewhat special, which might make the geometrical interpretation slightly more difficult to conceive. An

illustration of the general argument is given in figure 6.4 in section 6.2 below where we discuss similar three-vectors of structure functions, but there without the occurrence of a Kronecker symbol.

We recall that the dipole cross sections $\hat{\sigma}^{(q)}$ are non-negative. Thus the r.h.s. of (6.2) is a sum and an integral over three-vectors multiplied by non-negative weights, or, in other words, a special linear superposition of the three-vectors appearing under the integral. We search for the set of all possible linear superpositions of this kind with non-negative coefficients. This is called a *moment problem*. In appendix B we discuss the necessary mathematical tools to solve this problem. We give there the precise definitions of the key concepts convex set, convex hull and convex cone. We also give the detailed solution of the moment problem for the case of three $F_2^{(q)}$ structure functions as discussed below in section 6.2. The solution of the moment problem in this section runs along the same lines. The analogue of the result (B.49) reads here as follows. The set of all vectors allowing a representation (6.2) is given by a convex cone. Any vector within this cone can be written as a non-negative multiple of an element within the closed convex hull (denoted by $\overline{\text{co}}$) of the three-vectors appearing in the r.h.s. of (6.2). Therefore we have

$$\begin{pmatrix} F_L(W,Q^2) \\ F_2^{(c)}(W,Q^2) \\ F_2(W,Q^2) \end{pmatrix} = \lambda \boldsymbol{u}(Q^2)$$

$$\text{with} \quad \lambda \geq 0\,, \quad \boldsymbol{u}(Q^2) \in \overline{\text{co}} \left\{ \left. \begin{pmatrix} f_L^{(q)}(r,Q^2) \\ \delta_{q,c}\, f^{(c)}(r,Q^2) \\ f^{(q)}(r,Q^2) \end{pmatrix} \right| r \in \mathbb{R}^+,\ q = u,d,\ldots \right\}. \tag{6.5}$$

Note that the three-vectors from which the convex hull is constructed involve only the functions $f^{(q)}(r,Q^2)$ and $f_L^{(q)}(r,Q^2)$ which are for any given Q^2 explicitly known for all r, see (5.53)-(5.56). Hence it is also straightforward to compute their convex hull. We further point out that these vectors are independent of the energy W, and that the condition (6.5) does not involve any model assumption about the dipole cross section $\hat{\sigma}^{(q)}$.

We can now use the condition (6.5) to derive bounds on ratios of F_L, $F_2^{(c)}$, and F_2. These bounds originate only from the photon wave functions. They will be valid for any dipole cross section $\hat{\sigma}^{(q)}$, and will be independent of the energy W. Clearly, the bounds will vary with the photon virtuality Q^2, since Q^2 explicitly enters the vectors in (6.5) via the photon wave function.

We first notice that the condition (6.5) constrains only the *directions* of the three-vectors involved, while their normalisation is irrelevant for that condition. We can therefore normalise the vector composed of the three structure functions such that its third component equals one, that is, we consider the vector $(F_L/F_2, F_2^{(c)}/F_2, 1)^{\mathrm{T}}$ instead of $(F_L, F_2^{(c)}, F_2)^{\mathrm{T}}$. That normalisation does not change the direction of the vector, and hence also the so normalised vector fulfils the condition (6.5). Similarly, we normalise the set of vectors of which the closed convex hull is formed such that its third component equals one, hence considering $(f_L^{(q)}/f^{(q)}, \delta_{q,c} f^{(c)}/f^{(q)}, 1)^{\mathrm{T}} = (f_L^{(q)}/f^{(q)}, \delta_{q,c}, 1)^{\mathrm{T}}$ in place of $(f_L^{(q)}, \delta_{q,c} f^{(c)}, f^{(q)})^{\mathrm{T}}$.

Again, that does not affect the direction of the vectors, and the condition (6.5) immediately applies with this replacement. The condition with both vectors normalised in this way contains only vectors whose third component equals one, and for this case the only possible choice for the factor λ is $\lambda = 1$. We can then eliminate the trivial third component by projecting onto the 1-2-plane and obtain from (6.5) the simpler condition

$$\begin{pmatrix} F_L(W,Q^2)/F_2(W,Q^2) \\ F_2^{(c)}(W,Q^2)/F_2(W,Q^2) \end{pmatrix} \in \overline{\text{co}} \left\{ \begin{pmatrix} f_L^{(q)}(r,Q^2)/f^{(q)}(r,Q^2) \\ \delta_{q,c} \end{pmatrix} \,\middle|\, r \in \mathbb{R}^+, q = u,d,\ldots \right\}, \tag{6.6}$$

which is in fact equivalent to the original condition (6.5) for the realistic case that F_2 and $f^{(q)}$ for $r \in \mathbb{R}^+$ are strictly positive. For a rigorous derivation of bounds on ratio vectors as in (6.6), see appendix B, where the analogous case of three F_2 structure functions is discussed in detail (see (B.52)).

The first bound that we discuss here is now obtained from (6.6) by projecting onto the 1-axis. This immediately gives

$$\inf_{r,q} \frac{f_L^{(q)}(r,Q^2)}{f^{(q)}(r,Q^2)} \leq \frac{F_L(W,Q^2)}{F_2(W,Q^2)} \leq \sup_{r,q} \frac{f_L^{(q)}(r,Q^2)}{f^{(q)}(r,Q^2)}, \tag{6.7}$$

where inf and sup denote the infimum and supremum, respectively [1]. Note that these lower and upper bounds on F_L/F_2 are given only in terms of the photon wave function. It is therefore straightforward to analyse the bounds (6.7) numerically.

In figure 6.1 we plot the ratio $f_L^{(q)}(r,Q^2)/f^{(q)}(r,Q^2)$ as a function of r for different quark flavours, choosing as an example $Q^2 = 10\,\text{GeV}^2$. Here and in the following we use vanishing masses for the light (u,d,s) quarks, $m_c = 1.3\,\text{GeV}$ for the charm quark and $m_b = 4.6\,\text{GeV}$ for the bottom quark. We find that the lower bound in (6.7) is trivial, $F_L/F_2 \geq 0$. The upper bound, on the other hand, is non-trivial. We find that the maximal value of $f_L^{(q)}(r,Q^2)/f^{(q)}(r,Q^2)$ is obtained for light quarks, as can be seen in figure 6.1 where this maximum is drawn as a dotted horizontal line. It turns out that this upper bound is independent of Q^2 and numerically leads to

$$\frac{F_L(W,Q^2)}{F_2(W,Q^2)} \leq 0.27139\,. \tag{6.8}$$

An equivalent result was already obtained in [136]. There, the authors consider the ratio

$$R := \frac{\sigma_L}{\sigma_T} \tag{6.9}$$

and derive by a direct calculation an inequality for R, which is similar to (6.7), see (144) in [136]. From this they find the limit

$$R \leq 0.37248 \tag{6.10}$$

[1] We recall that for a given subset S of $\mathbb{R}$ the infimum of the set S is the greatest number less than or equal to each element of S. Similarly, the supremum is the smallest number that is greater than or equal to each element of S. For a compact set S the minimum (maximum) coincides with the infimum (supremum).

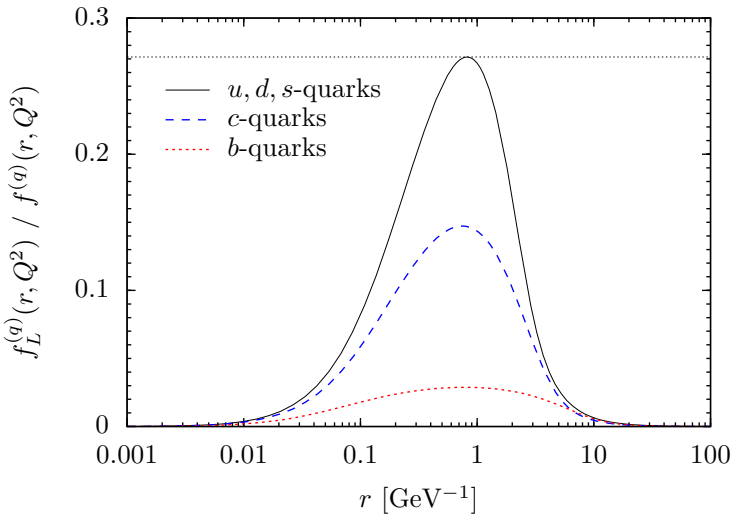

Figure 6.1: The ratio $f_L^{(q)}(r,Q^2)/f^{(q)}(r,Q^2)$ as a function of r for different quark flavours. The photon virtuality is chosen to be $Q^2 = 10\,\mathrm{GeV}^2$. The absolute maximum value (dotted line) of all curves provides an upper bound on $F_L(W,Q^2)/F_2(W,Q^2)$, see (6.7) and (6.8).

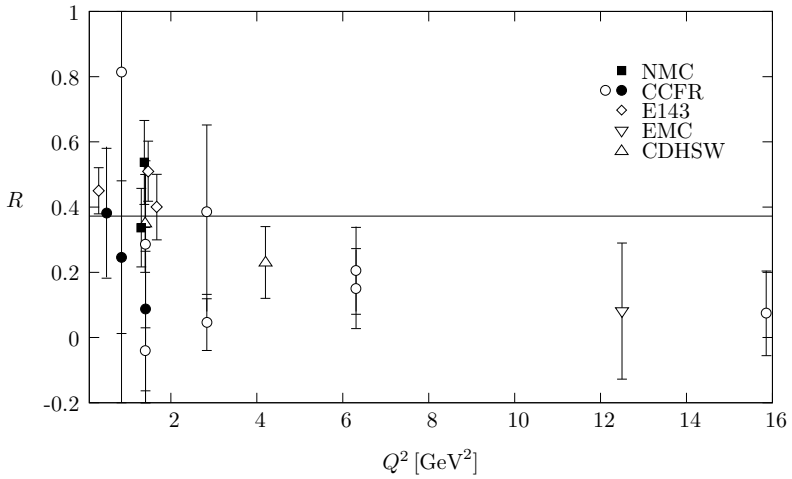

Figure 6.2: Comparison of experimental data for $R = \sigma_L/\sigma_T$ in the region $x < 0.05$ with the bound (6.10) resulting from the dipole picture. Full points correspond to data with $x < 0.01$, other points to data with $0.01 < x < 0.05$. Figure taken from [136].

originating again from the light quarks. From $F_L/F_2 = (R^{-1}+1)^{-1}$ we see that this result is equivalent to (6.8). In figure 6.2 taken from [136] the bound on R is confronted with data. It is interesting to note that at low Q^2 some data points actually lie outside the allowed range for the dipole picture. However, the errors of the low Q^2 data are still too big to judge that this really means a breakdown of the colour dipole picture in this kinematical regime.

A stronger bound can be obtained by considering the correlation of the ratios F_L/F_2 and

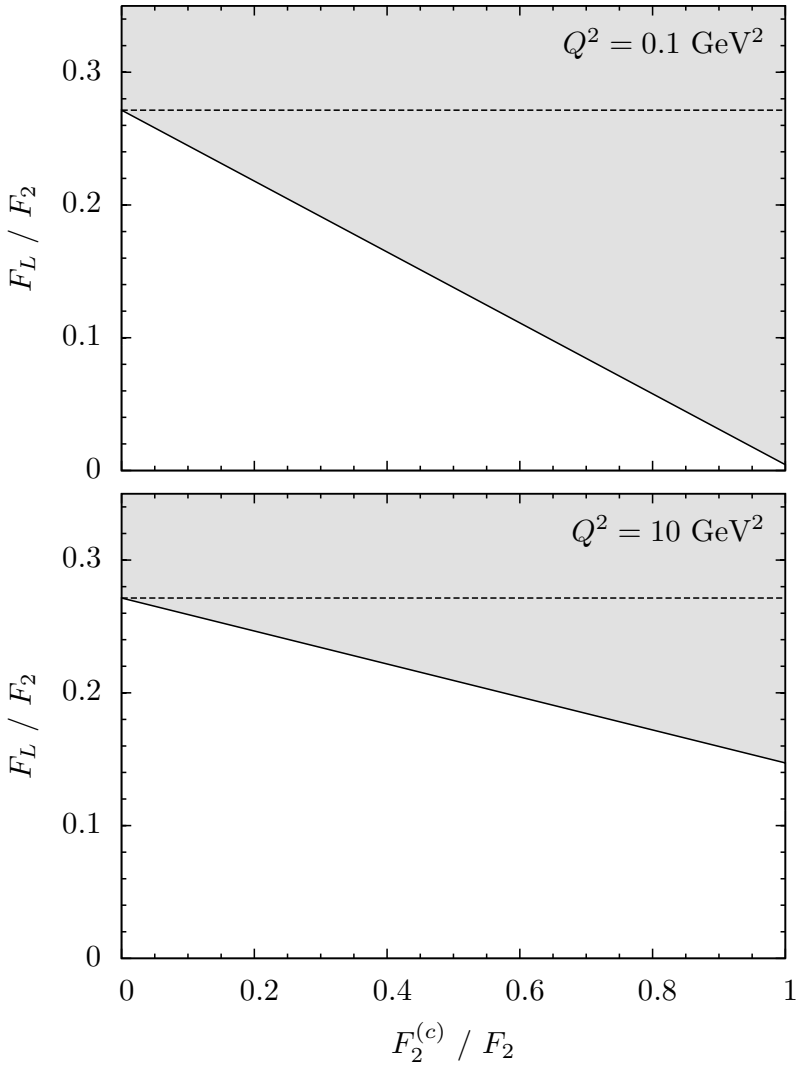

Figure 6.3: Correlated bounds on F_L/F_2 and $F_2^{(c)}/F_2$. Only the white (unshaded) area is allowed within the dipole picture. The weaker bound (6.7) is shown as a dashed line, while the stronger bound implied by (6.6) is shown as a solid line.

$F_2^{(c)}/F_2$, that is, by taking into account both components of the constraint on the vectors in (6.6). In this case the bound on the ratio F_L/F_2 will depend on the value of $F_2^{(c)}/F_2$ or vice versa. By computing the closed convex hull in (6.6) from the ratios $f_L^{(q)}(W, Q^2)/f^{(q)}(W, Q^2)$ we obtain the correlated bounds shown in figure 6.3 for the two values $Q^2 = 0.1\,\mathrm{GeV}^2$ and $Q^2 = 10\,\mathrm{GeV}^2$. The unshaded area in the two plots is the allowed region for the dipole picture. The figure is drawn for the whole range of $F_2^{(c)}/F_2$ between zero and one to illustrate origin of the bounds. Realistic values of $F_2^{(c)}/F_2$ can only range from zero to at most about 0.4. The allowed area in figure 6.3 is bounded by a straight line. This particular shape emerges due to the fact that the second component of the vectors in (6.6)

receives a contribution only from the charm quark. Due to the corresponding Kronecker symbol the upper bound on F_L/F_2 at the (unphysical) point $F_2^{(c)}/F_2 = 1$ is given by the maximum of $f_L^{(c)}(r, Q^2)/f^{(c)}(r, Q^2)$ over all r for the Q^2 under consideration. The value of this maximum for the case $Q^2 = 10\,\mathrm{GeV}^2$ can be read off from the charm quark curve in figure 6.1. For Q^2-values well below the charm mass, like for example $Q^2 = 0.1\,\mathrm{GeV}^2$, the analogous function practically vanishes and the resulting upper bound on F_L/F_2 at $F_2^{(c)}/F_2 = 1$ is practically zero. The fact that the unphysical point $F_2^{(c)}/F_2 = 1$ is relevant for the determination of the correlated bounds on F_L/F_2 and $F_2^{(c)}/F_2$ in the physical region of these ratios should not cause any worries here. It is just the consequence of omitting any assumptions about the flavour dependence of the dipole cross sections $\hat{\sigma}^{(q)}$, except their non-negativity. This general case includes for example the unphysical case that all dipole cross sections but the one for the charm quark would vanish, which would give rise to $F_2^{(c)}/F_2 = 1$. By making further assumptions about the dipole cross sections it may lead to more stringent bounds – but at the expense of introducing a dependence on those assumptions. Here, however, it is our aim to study bounds on ratios of structure functions from the dipole picture which *do not* depend on any further assumptions on the dipole cross section.

Future measurements of the structure functions F_L, $F_2^{(c)}$and F_2 at identical values of Q^2 and W might in combination with our bounds be able to constrain the range of validity of the dipole picture.

Closing this section we would like to emphasise again that the geometric argument and its implications discussed in this section remain unchanged if the dipole cross section is chosen to depend on x instead of W.

6.2 F_2 at different Q^2

In this section we use the dipole picture to derive bounds on ratios of the structure function F_2 taken at the same W but at different values of Q^2. The results found here crucially depend on choosing the functional dependence of the dipole cross section such that its arguments are r and W. In particular, the dipole cross section $\hat{\sigma}^{(q)}(r, W)$ is assumed to be independent of Q^2, see the corresponding discussion in section 5.1.

We consider the structure function F_2 at three different values of Q^2 but at the same W. Similarly to the previous section we arrange them into a three-vector, and evaluate it according to the dipole formula,

$$\begin{pmatrix} F_2(W, Q_1^2) \\ F_2(W, Q_2^2) \\ F_2(W, Q_3^2) \end{pmatrix} = \sum_q \int \mathrm{d}^2 r \, \frac{\hat{\sigma}^{(q)}(r, W)}{4\pi^2 \alpha_{\mathrm{em}}} \begin{pmatrix} f^{(q)}(r, Q_1^2) \\ f^{(q)}(r, Q_2^2) \\ f^{(q)}(r, Q_3^2) \end{pmatrix} , \tag{6.11}$$

where the $f^{(q)}(r, Q_i^2)$ are defined in (6.3). We can now derive bounds on ratios of such structure functions following the same procedure as in the preceding section. To find all vectors allowing a representation (6.11) is again a moment problem. In appendix B we

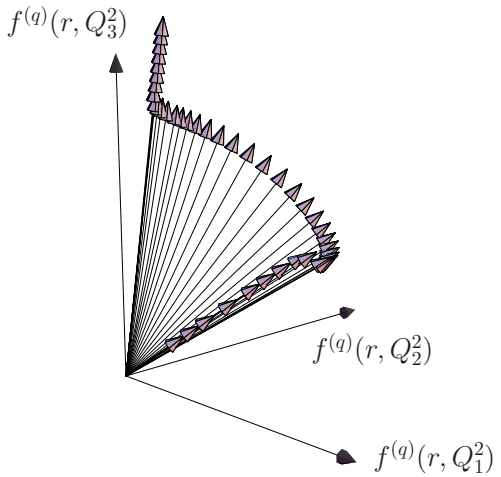

Figure 6.4: The vectors $(f^{(q)}(r, Q_1^2), f^{(q)}(r, Q_2^2), f^{(q)}(r, Q_3^2))^{\mathrm{T}}$ for different values of r, shown here for a massless quark flavour q and for one particular choice of the triple (Q_1^2, Q_2^2, Q_3^2).

discuss the solution of this problem for the case at hand in a mathematically rigorous way. A simple argument, leaving out some subtleties, is as follows.

The vector on the l.h.s. of (6.11) is a linear superposition of the three-vectors $(f^{(q)}(r, Q_1^2)$, $f^{(q)}(r, Q_2^2)$, $f^{(q)}(r, Q_3^2))^{\mathrm{T}}$ which appear under the integral. For a given flavour q and given values of the Q_i^2 that vector follows a trajectory as $r \in \mathbb{R}^+$ is varied. Figure 6.4 illustrates a number of vectors along such a trajectory for the case of massless quarks and for one particular choice of Q_1^2, Q_2^2, and Q_3^2. We recall again that the dipole cross sections $\hat{\sigma}^{(q)}$ are non-negative. Accordingly, the r.h.s. of (6.11) is a linear superposition with non-negative weights of the vectors that appear under the integral. Therefore the resulting vector on the l.h.s. must lie in the closed convex cone formed by all possible linear superpositions with non-negative weights of those vectors and their boundary. Any vector within such a cone is a non-negative multiple of a vector that lies in the closed convex hull (denoted by $\overline{\mathrm{co}}$) of the vectors appearing under the integral in (6.11), see (B.49) of appendix B. Hence we obtain the condition

$$\begin{pmatrix} F_2(W, Q_1^2) \\ F_2(W, Q_2^2) \\ F_2(W, Q_3^2) \end{pmatrix} = \mu \boldsymbol{v}(Q_1^2, Q_2^2, Q_3^2)$$

$$\text{with} \quad \mu \geq 0\,, \quad \boldsymbol{v}(Q_1^2, Q_2^2, Q_3^2) \in \overline{\mathrm{co}} \left\{ \begin{pmatrix} f^{(q)}(r, Q_1^2) \\ f^{(q)}(r, Q_2^2) \\ f^{(q)}(r, Q_3^2) \end{pmatrix} \,\middle|\, r \in \mathbb{R}^+,\ q = u, d, \ldots \right\}. \tag{6.12}$$

As in the case of (6.6) in the previous section this condition constrains only the directions of the vectors involved, but not their length. This applies both to the vector with components $F_2(W, Q_i^2)$ and to the vector with components $f^{(q)}(r, Q_i^2)$. Accordingly, we can normalise these vectors such that their third component equals one. Performing this for both vectors

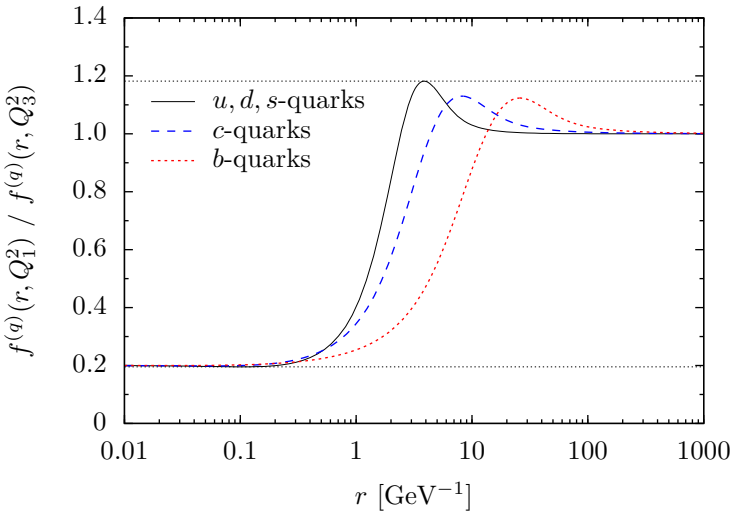

Figure 6.5: The ratio $f^{(q)}(r, Q_1^2)/f^{(q)}(r, Q_3^2)$ as a function of r for different quark masses, here for the choice $(Q_1^2, Q_3^2) = (2, 10)\,\mathrm{GeV}^2$. The minimum and maximum values of all curves (shown as dotted lines) provide a lower and an upper bound on $F_2(W, Q_1^2)/F_2(W, Q_3^2)$, respectively, see (6.14).

that appear in (6.12) we obtain an equivalent condition. In this condition the only possible value for μ is obviously $\mu = 1$. Since the third component of the condition is now trivial we discard it by projecting onto the 1-2-plane to obtain

$$\begin{pmatrix} F_2(W, Q_1^2)/F_2(W, Q_3^2) \\ F_2(W, Q_2^2)/F_2(W, Q_3^2) \end{pmatrix} \in \overline{\mathrm{co}} \left\{ \begin{pmatrix} f^{(q)}(r, Q_1^2)/f^{(q)}(r, Q_3^2) \\ f^{(q)}(r, Q_2^2)/f^{(q)}(r, Q_3^2) \end{pmatrix} \,\middle|\, r \in \mathbb{R}^+, q = u, d, \ldots \right\}, \tag{6.13}$$

which is fully equivalent to (6.12) because $F_2(W, Q_3^2)$ and $f^{(q)}(r, Q_3^2)$ are strictly positive for the relevant range of their arguments. For a rigorous derivation of (6.13) see (B.50)-(B.52) of appendix B.

Let us first consider the two components of the condition (6.13) separately. Projecting it onto the 1-axis and onto the 2-axis immediately gives the conditions

$$\inf_{r,q} \frac{f^{(q)}(r, Q_1^2)}{f^{(q)}(r, Q_3^2)} \leq \frac{F_2(W, Q_1^2)}{F_2(W, Q_3^2)} \leq \sup_{r,q} \frac{f^{(q)}(r, Q_1^2)}{f^{(q)}(r, Q_3^2)}, \tag{6.14}$$

$$\inf_{r,q} \frac{f^{(q)}(r, Q_2^2)}{f^{(q)}(r, Q_3^2)} \leq \frac{F_2(W, Q_2^2)}{F_2(W, Q_3^2)} \leq \sup_{r,q} \frac{f^{(q)}(r, Q_2^2)}{f^{(q)}(r, Q_3^2)}, \tag{6.15}$$

respectively. The condition (6.15) equals (6.14) if we replace Q_2^2 by Q_1^2. Thus, for two given values of Q^2 we actually obtain one condition here which contains an upper and a lower bound. The same result was already presented in [153], where these bounds were derived in a different way.

Here we discuss briefly the bounds of (6.14), for a more detailed discussion we refer the reader to [153]. We first note that the upper and lower bound (6.14) depend only on

the values of Q_1^2 and Q_3^2, but do not involve the energy W. Figure 6.5 shows the ratio $f^{(q)}(r, Q_1^2)/f^{(q)}(r, Q_3^2)$ as a function of r for different quark masses along with the resulting bounds on $F_2(W, Q_1^2)/F_2(W, Q_3^2)$ for the specific choice $(Q_1^2, Q_3^2) = (2, 10)\,\mathrm{GeV}^2$.

In the following we confront the bounds that we obtain in this section from the colour dipole picture with HERA data. Before we proceed a remark concerning the comparison with data is in order. Data on F_2 and measurements of the reduced cross section are available for a large range of Q^2 values with (x, Q^2)-binning. However, throughout this section we deal with bounds involving values of F_2 (or of the reduced cross section, see below) at the same W but at different Q_i^2. Hence a comparison with our bounds requires different values of Q^2 at the same value of W, and data with (W, Q^2)-binning are published only for comparatively small kinematical ranges. We therefore use a fit to the F_2 data that can, to a good approximation, be considered as a substitute of actual data. We do this in most of the following comparisons, except for two illustrations where HERA data are used directly (see figure 6.9 and the corresponding discussion below). Concretely, we use the ALLM97 fit to F_2 [154, 155] which represents the measured data points of [156, 157] within their errors, except maybe for the region of very low Q^2 where the fit appears to be slightly worse. We emphasise that we use the fit only *inside* the kinematical range in which actual HERA data are available. No extrapolation beyond that range is done here.

Figure 6.6 confronts the bound (6.14) with the ALLM97 fit to F_2 for a fixed value of Q_3^2 and variation of Q_1^2, as presented in [153] before. It is apparent from the figure that there is a value of Q_1^2 beyond which the dipole picture fails to be compatible with the ALLM97 fit. This maximal Q_1^2 value depends on the value of W, as can be seen in the figure from the three curves for different W, and it also depends on the value chosen for Q_3^2. With the choice $Q_3^2 = 10\,\mathrm{GeV}^2$ made for the figure, this maximal Q_1^2 is in the range of about 150-300 GeV^2, depending on W.

Let us discuss, to what extent the bound (6.14) is affected by different normalizations of the $\gamma^* p$ cross sections σ_T and σ_L relative to F_2. Above, we used the simple relation (5.13), as often encountered in the context of high energy scattering, in particular for applications of the dipole picture. The relation (5.12) derived from Hand's convention reduces to that simple expression if the high energy limit is taken for fixed Q^2. For finite x the normalization of $\sigma_{T,L}$ relative to F_2 differs by a factor $(1-x)$ between the two choices (neglecting terms of $\mathcal{O}(m_p^2/W^2)$), as is indicated in (5.12). The bounds discussed here are all based on the simpler formula (5.13). It is straightforward, however, to derive similar bounds based on the relation (5.12). The additional factor $(1-x)$ depends both on Q^2 and W. Therefore this factor does not cancel if ratios of structure functions are taken at different Q_i^2. Furthermore, the bounds on ratios of F_2 inherit a dependence on W from this factor. The bounds on ratios of F_L, $F_2^{(c)}$, and F_2 discussed in subsection 6.1, on the other hand, are not affected. There, the structure functions are evaluated at the same W and Q^2 and the additional factor $(1-x)$ cancels in the ratios. In figure 6.7 we show for one energy $W = 60\,\mathrm{GeV}$ how the bound (6.14) is changed if one uses (5.12) instead of (5.13). A sizable deviation from the original bound occurs only at relatively large Q^2 where the new bound is closer to the data than the original bound. However, both bounds are violated by the

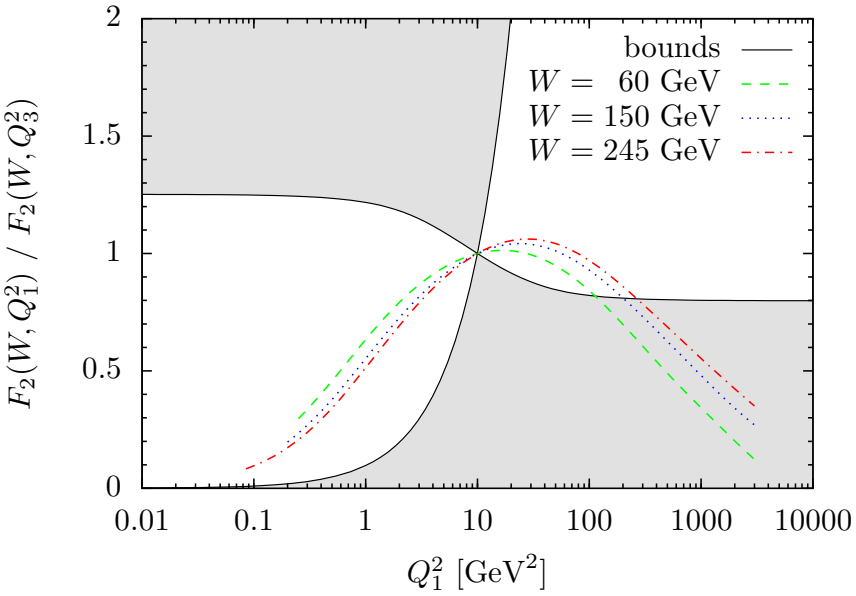

Figure 6.6: The bounds (6.14) on $F_2(W,Q_1^2)/F_2(W,Q_3^2)$ resulting from the dipole picture (solid lines) confronted with the corresponding ratios obtained using the ALLM97 fit to F_2 for three different values of W. Here Q_1^2 is varied while the value $Q_3^2 = 10\,\mathrm{GeV}^2$ is kept fixed. The shaded region is excluded by the bounds.

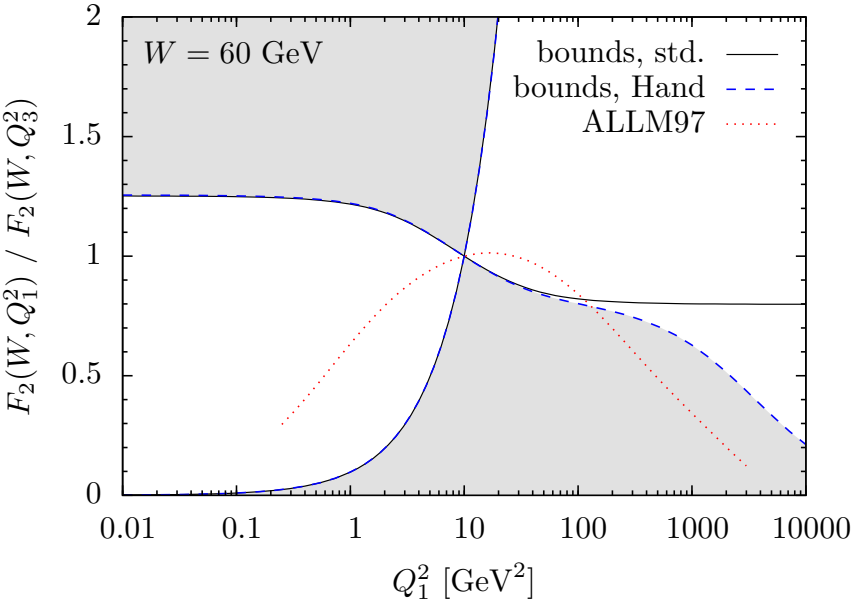

Figure 6.7: Change of the bound (6.14) on $F_2(W,Q_1^2)/F_2(W,Q_3^2)$ due to using Hand's convention (5.12) instead of the simpler (5.13), shown here for $W = 60\,\mathrm{GeV}$ and the choice $Q_3^2 = 10\,\mathrm{GeV}^2$. The solid lines are the original bounds (6.14), while the dashed lines represent the modified bounds. The dotted line is the ALLM97 fit to $F_2(W,Q_1^2)/F_2(W,Q_3^2)$.

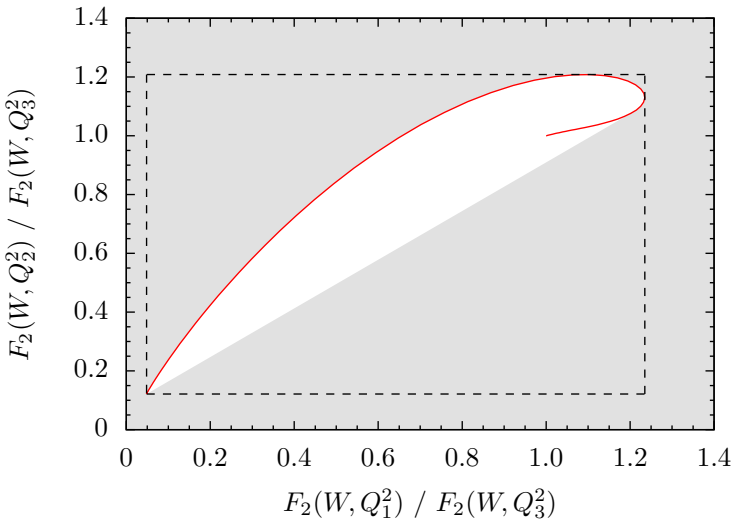

Figure 6.8: The trajectory of the vectors $(f^{(q)}(r,Q_1^2)/f^{(q)}(r,Q_3^2), f^{(q)}(r,Q_2^2)/f^{(q)}(r,Q_3^2))^{\mathrm{T}}$ for variation of r (solid curve), here for a massless quark flavour q and for the choice $(Q_1^2,Q_2^2,Q_3^2) = (4,10,80)\,\mathrm{GeV}^2$. The unshaded area is the closed convex hull of the vectors that form the trajectory. According to the weaker bounds (6.14) and (6.15) the vectors $(F_2(W,Q_1^2)/F_2(W,Q_3^2), F_2(W,Q_2^2)/F_2(W,Q_3^2))^{\mathrm{T}}$ must lie within the dashed rectangle, while the stronger bound (6.13) requires them to lie within the convex hull of the unshaded area and the corresponding areas obtained for massive quarks. The curves for massive quarks have a similar shape and are not shown here for simplicity.

data at about the same Q^2 and the difference between the original and the modified bound grows only at larger Q^2. Therefore the normalization of the $\gamma^* p$ cross sections according to Hand's convention would not significantly alter our results concerning the range of validity of the dipole picture. We restrict to the simpler relation (5.13) in the following.

So far we have discussed in some detail the bounds (6.14) and (6.15) which resulted from considering the two components of (6.13) separately. We can improve these bounds by taking into account the correlation of those two components, that is the correlation of the two ratios $F_2(W,Q_1^2)/F_2(W,Q_3^2)$ and $F_2(W,Q_2^2)/F_2(W,Q_3^2)$. According to (6.13) the 2-vector constructed from these two ratios for a given set of Q_i^2 lies in the closed convex hull of the vectors $(f^{(q)}(r,Q_1^2)/f^{(q)}(r,Q_3^2), f^{(q)}(r,Q_2^2)/f^{(q)}(r,Q_3^2))^{\mathrm{T}}$. As r is varied the latter vector (for each quark flavour q) follows a trajectory in 2-dimensional space. For the case of a massless quark flavour q that trajectory is shown as the solid curve in figure 6.8, where we have chosen the values $(Q_1^2,Q_2^2,Q_3^2) = (4,10,80)\,\mathrm{GeV}^2$ for this example. The white (unshaded) area is the closed convex hull of the vectors that form the trajectory. Similar but slightly different trajectories are obtained for massive quark flavours, which are not shown here in order to keep the figure simple. As a consequence of the dipole picture the 2-vectors $(F_2(W,Q_1^2)/F_2(W,Q_3^2), F_2(W,Q_2^2)/F_2(W,Q_3^2))^{\mathrm{T}}$ must lie within the closed convex hull of those trajectories, independently of the energy W, see (6.13). The dashed lines in figure 6.8 represent the two bounds (6.14) and (6.15). Clearly, the correlated bound (6.13)

is much stronger than the separate bounds on the ratios.

Next we want to compare the stronger bound (6.13) with experimental data. For this purpose we need data points of F_2 at three different Q_i^2 but at the same W. As we have mentioned before, most of the available data are not published in (W, Q^2)-binning. We have found only few points which are suitable for a direct comparison with our bound, that is with the same W and three different Q_i^2 that are not too close to each other. We will now present two of these examples. Further below we will then again use the ALLM97 fit for a more comprehensive analysis of the kinematical range in which the bound (6.13) is respected. For the comparison with actual HERA data we choose as the observable the reduced cross section instead of F_2, since the former is the one which was directly measured. The reduced cross section is defined as

$$\sigma_{\mathrm{r}}(W, Q^2) = \frac{Q^2}{4\pi^2 \alpha_{\mathrm{em}}} \left(\sigma_T + \frac{2(1-y)}{1+(1-y)^2} \sigma_L \right), \tag{6.16}$$

with $y \approx (W^2 + Q^2)/s$, see (5.2), where $\sqrt{s} \approx 300\,\mathrm{GeV}$ is the lepton-proton center-of-mass energy for the available HERA data. It is straightforward to derive correlated bounds for ratios of reduced cross sections instead of F_2 structure functions from the dipole picture. The derivation is completely analogous to the one described above. We just have to replace $f^{(q)}(r, Q^2)$ by

$$f_{\mathrm{r}}^{(q)}(r, Q^2) = Q^2 \left[w_T^{(q)}(r, Q^2) + \frac{2(1-y)}{1+(1-y)^2} w_L^{(q)}(r, Q^2) \right], \tag{6.17}$$

as can be seen from (5.13) and (6.16) together with (6.2), (6.3). The resulting bound is then as given by (6.13) but with F_2 replaced by σ_{r} and $f^{(q)}$ replaced by $f_{\mathrm{r}}^{(q)}$. Due to this modification the bound for the reduced cross section now depends on W (which enters via y), which was not the case for the original bound for F_2. Figure 6.9 confronts the bound on the quantity $(\sigma_{\mathrm{r}}(W, Q_1^2)/\sigma_{\mathrm{r}}(W, Q_3^2), \sigma_{\mathrm{r}}(W, Q_2^2)/\sigma_{\mathrm{r}}(W, Q_3^2))^{\mathrm{T}}$ with its measured values from ZEUS [158] and H1 [159] for two different choices of W and of the triple of Q_i^2. The depicted errors on the ratios are the combination of the experimental errors on σ_{r} in quadrature. The curves in figure 6.9 show the correlated bounds for contributions to $(\sigma_{\mathrm{r}}(W, Q_1^2)/\sigma_{\mathrm{r}}(W, Q_3^2), \sigma_{\mathrm{r}}(W, Q_2^2)/\sigma_{\mathrm{r}}(W, Q_3^2))^{\mathrm{T}}$ from different quark flavours as given by the analogue of (6.13) for σ_{r}. Only if the point obtained from the data lies within the convex hull of all these curves it can possibly be described in the framework of the dipole picture. We see that this condition is fulfilled for the high W, moderate Q^2 sample (upper graph), while it is violated by approximately two standard deviations for the lower W, higher Q^2 sample (lower graph).

The above discussion refers to the applicability of the dipole picture at a given value of W for one particular triple of Q^2-values. For a determination of the range of applicability of the dipole picture it is more desirable to determine for a given W a maximal *range* in Q^2 in which the three Q_i^2 can be chosen without giving rise to a violation of the bound. For this purpose we now consider again the structure function F_2 (and no longer the reduced cross section). Using the ALLM97 fit to the measured F_2 data we can then perform a

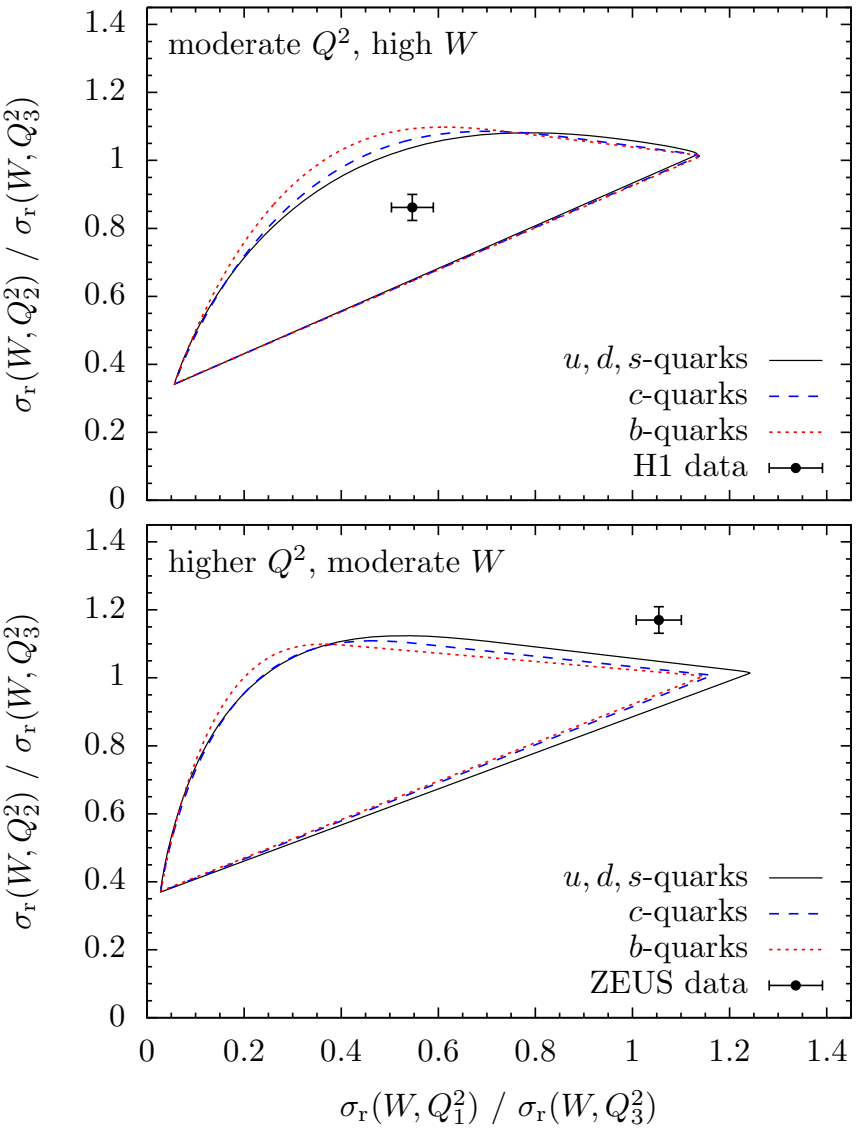

Figure 6.9: Correlated bounds on ratios of the reduced cross section (6.16) at different values of Q^2 obtained from the dipole picture, confronted with HERA data. The inner parts of the different curves show the allowed regions if only contributions from specific quark flavours are considered, while the convex hull of these regions gives the net bound if no further assumptions are made. The kinematical values are $W = 247\,\text{GeV}$, $(Q_1^2, Q_2^2, Q_3^2) = (2, 12, 35)\,\text{GeV}^2$ for the upper plot and $W = 75\,\text{GeV}$, $(Q_1^2, Q_2^2, Q_3^2) = (3.5, 45, 120)\,\text{GeV}^2$ for the lower plot.

continuous scan in Q^2 and determine precisely the kinematical range in which the bounds are respected. We first consider the correlated bounds obtained from (6.13), and later compare the allowed Q^2-range with the one resulting from the weaker bounds (6.14) and (6.15).

Let us first fix the energy W at some value. We will in the following call a violation of

the dipole-picture bound a 'significant' violation if the ALLM97 F_2 ratios give a relative deviation of more than 10% from the bound. This accounts for a kind of error band which should be associated with the ALLM97 fit or with the corresponding ratios of F_2. If for a certain triple (Q_1^2, Q_2^2, Q_3^2) the ratios obtained from the ALLM97 fit violate the bounds by a significant amount (in the above sense) any Q^2-range containing the values Q_1^2, Q_2^2, Q_3^2 is excluded for a successful description within the dipole picture. In contrast, agreement with the bounds for a triple does not necessarily imply agreement for the full range $[\min_i(Q_i^2), \max_i(Q_i^2)]$ of that triple since the bounds depend on all three Q_i^2. We therefore systematically search for the maximal Q^2-range that contains no Q^2-triple for which the bounds are violated significantly. Technically, we do this by searching for the minimal Q^2-range in which we can find at least one Q^2-triple for which the bounds are violated significantly. The lower bound of a given Q^2-interval turns out to have only mild influence on whether a significant violation of the bounds can be found within that interval – provided it is not much larger than $1\,\mathrm{GeV}^2$. We therefore keep the lower end of the considered Q^2-range fixed at $1\,\mathrm{GeV}^2$ and determine the upper end $Q^2_{\max}$ of the Q^2-range within which the bounds are not significantly violated. We can then repeat this procedure for each energy W and determine $Q^2_{\max}$ as a function of W.

The solid line in figure 6.10 shows the result of such a calculation based on the correlated bounds obtained from (6.13). The allowed Q^2-range slowly grows with increasing energy, as can be expected on general grounds. $Q^2_{\max}$ ranges from about $100\,\mathrm{GeV}^2$ for $W = 60\,\mathrm{GeV}$ to about $200\,\mathrm{GeV}^2$ for $W = 245\,\mathrm{GeV}$. The dashed line in figure 6.10 represents the analogous curve obtained from the uncorrelated bounds (6.14) and (6.15). Here we have varied both Q_1^2 and Q_3^2 in (6.14) in order to determine the maximal virtuality, $Q^2_{\max}$, below which both Q_1^2 and Q_3^2 can be chosen arbitrarily without giving rise to a significant violation of the bound. We see that the correlated bounds resulting from (6.13) indeed give stronger restrictions on the range of validity of the dipole picture than the uncorrelated bounds (6.14), (6.15).

Note that the violation of the correlated bound does not take place at a constant value of x. In the contrary, the value of x changes along the solid line in figure 6.10. For $Q^2 = 100\,\mathrm{GeV}^2$ we find that $x < 0.03$ is required for the bound not to be violated, while for $Q^2 = 200\,\mathrm{GeV}^2$ the bound is only respected for $x < 0.003$. A similar observation applies to the uncorrelated bounds (6.14), (6.15) (see the dashed line in figure 6.10) as already observed in [153].

Obviously, a violation of the above bounds indicates that some contributions to the cross section become important which are not contained in the dipole picture. We would like to emphasise that such corrections to the dipole picture might become sizable already before the bounds are actually violated. One should therefore expect that corrections to the standard dipole picture are important already if the data come close to the bounds.

The upper limit on the kinematical range of validity of the dipole picture that we find here appears to be rather low in view of the fact that phenomenological fits to F_2 data based on the dipole picture often work quite well up to rather high Q^2, see for example [146]. However, the good quality of those fits at large Q^2 is not in contradiction with our result. We recall that the bounds derived in this section crucially depend on the

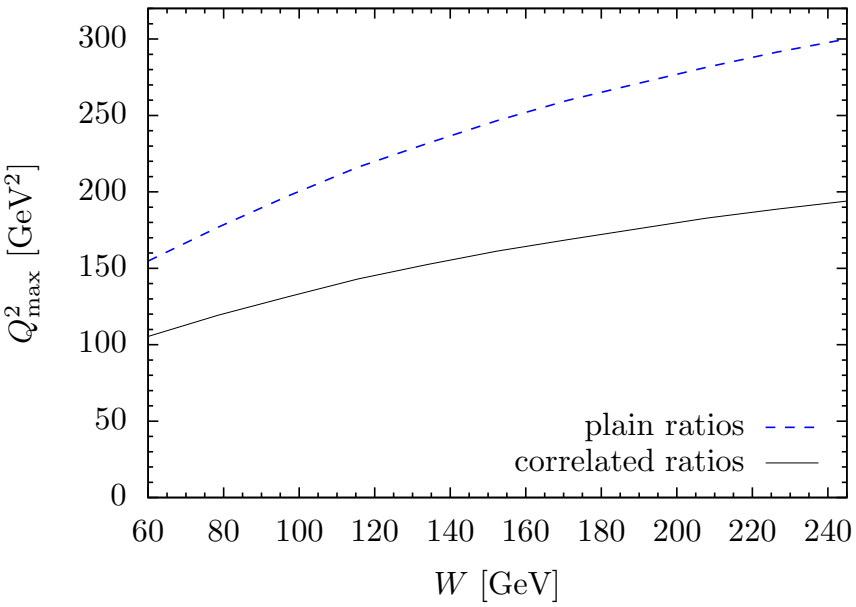

Figure 6.10: Upper limit $Q^2_{\max}$ of the Q^2-range in which the ALLM97 fit to F_2 is consistent with the bounds obtained from the dipole picture within a 10% relative deviation of the F_2 ratios, plotted as a function of the energy W. The curves are for consistency with the correlated bounds (6.13) (solid line), and with the uncorrelated bounds (6.14) and (6.15) (dashed line), respectively.

functional dependence of the dipole cross section $\hat{\sigma}$ on r and W, as obtained naturally from the derivation of the dipole picture presented in [135, 136]. In particular, $\hat{\sigma}$ needs to be independent of Q^2 for our bounds to be valid. But almost all recent models for $\hat{\sigma}$ assume it to depend on x, and hence on Q^2. The transition from the energy variable W to the energy variable x in the dipole cross section requires additional assumptions the justification and the physical significance of which appears difficult to assess. In practice, they might capture – at least partly – some corrections that are left out in the usual dipole picture (see the discussion in chapter 5). It would be very desirable to obtain a better understanding of this situation. An important step would be to check whether it is also possible to describe the presently available HERA data by models for the dipole cross section based on the more natural variables r and W.

We finally note that our bounds are modified if one uses, as suggested by Hand's convention, the relation (5.12) between F_2 and the $\gamma^* p$ cross sections σ_T and σ_L instead of the simpler relation (5.13) that has been used here. The modification becomes relevant for not so small x, and hence for large Q^2. The natural kinematical region for the application of the dipole picture is the region of small x, so that this modification is only of minor relevance for the dipole picture. Nevertheless, we illustrated the effect of using (5.12) on our bounds for an example. Our conclusions about the range of validity of the dipole picture would not to be significantly affected.

Summary

We have derived correlated bounds on F_L/F_2 and $F_2^{(c)}/F_2$ within the dipole picture. These are valid for any dipole model, independent of the dipole cross section and its energy dependence. It will be interesting to compare these bounds with the very recent [160] and future HERA structure function measurements.

Further, we have derived correlated bounds on ratios of F_2 taken at three different values of Q^2 but the same energy W. Here, we have taken the energy dependence of the dipole cross section to be given by W alone. This is crucial for the derived bounds, such that these do not apply for an x-dependent dipole cross section. Similarly, modifications of the standard dipole formuae might alter these bounds as illustrated for one example. We find, that the correlated bounds are significantly more restrictive than their uncorrelated counter-parts. We have confronted the bounds with experimental data. More precisely, we have used the ALLM97 fit to the measured F_2 data except for two examples, where we used actual data points. A more direct comparison would require a (W, Q^2) instead of the usual (x, Q^2) binning of the data. Our results show that these bounds give no relevant limitation at low Q^2. But we find the dipole picture to fail in describing high Q^2 data. Depending on W, this results in an upper bound on Q^2 of roughly $100 - 200\,\mathrm{GeV}^2$ for the range of applicability of any such dipole model.

Chapter 7

Energy dependence of the dipole cross section

7.1 Typical dipole sizes

We discussed in chapter 5 that in a natural scheme the dipole cross section is independent of Q^2 and depends on the energy W only. We also noted that in contrast to this, prominent dipole models discussed in the literature introduce a dependence of $\hat{\sigma}$ on Q^2, typically through the Bjorken-x variable. Often it is argued that in the dipole model one has a relation of the kind

$$r \propto \frac{1}{Q}\,, \tag{7.1}$$

that is, the most relevant dipole sizes r are given by the scale Q. The reasoning behind this is the fact, that due to the interplay of the Q^2 dependence of the photon wave function and the r dependence of the dipole cross section a typical size is generated. Neglecting the W dependency, the masses and all non-perturbative intrinsic scales, this typical size must by dimensional reasons obviously given by (7.1). Taking (7.1) seriously one might rewrite energy dependencies such that the choice of the energy dependence in the dipole cross section would be a mere convention without physical consequences. In this chapter we shall show that this is not true, that is, a strict identification of the type (7.1) fails.

Let us first see how the relation (7.1) for typical dipole sizes arises in the generic case. We consider a dipole cross section

$$\hat{\sigma}(r) = c_0\, r^2 \tag{7.2}$$

where $c_0 \in \mathbb{R}$ is constant. This dipole cross section implements colour transparency as present in many proposed models at small r. The non-perturbative region of large r is of course unphysical for this cross section. However, the region of large r is not relevant for our aim here. A physical model might deviate from (7.2) also in the intermediate r region. These effects could alter the present discussion but are neglected here, since we are interested in a generic discussion with minimal model assumptions here. With (7.2) the

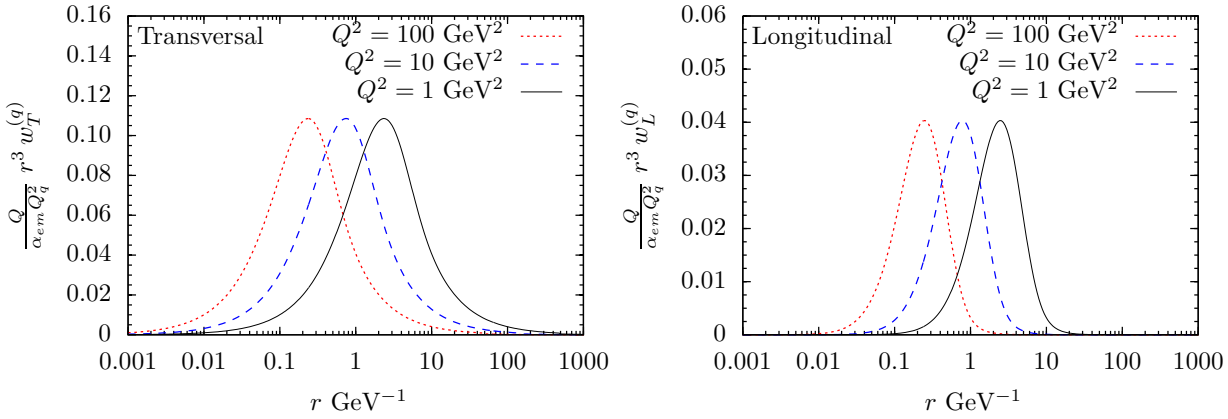

Figure 7.1: Dependence of the dimensionless quantities $Qr^3 w_{T,L}^{(q)}(Q^2)/(\alpha_{\mathrm{em}} Q_q^2)$ on r for the case of massless quarks. Left graph for transverse, right graph for longitudinal photon polarisation.

$\gamma^* p$ cross sections are formally

$$\sigma_T(Q^2) = \sum_q \int_0^\infty \mathrm{d}r\, 2\pi c_0\, r^3\, w_T^{(q)}(r, Q^2)\,, \tag{7.3}$$

$$\sigma_L(Q^2) = \sum_q \int_0^\infty \mathrm{d}r\, 2\pi c_0\, r^3\, w_L^{(q)}(r, Q^2)\,. \tag{7.4}$$

Figure 7.1 shows the behaviour of the integrands in (7.3) (left graph) and (7.4) (right graph) for the case of massless quarks and different values of Q^2. It is apparent from the curves that the integrands feature a maximum, and the position of the maximum depends on Q^2. We note however, that in particular for transverse photon polarisation the distribution of the integrand is rather broad.

For the *massless case* we find the rescaling property $r^2 w_{T,L}^{(q)}(r, Q^2) = (cr)^2 w_{T,L}^{(q)}(cr, Q^2/c^2)$ for any $c \in \mathbb{R}$. We numerically maximise $r^3(w_T^{(q)}(r, Q^2) + w_L^{(q)}(r, Q^2))$ with respect to r for some constant value of Q^2 and conclude that for arbitrary fixed Q^2 the maximum position is given by

$$r|_{\max} = \frac{C}{Q} \qquad \text{with } C = 2.40\,. \tag{7.5}$$

Note that in the massless case, $\sigma_T^{(q)}$ is not well defined due to logarithmic endpoint singularities in the α integration:

$$\sigma_T^{(q)}(Q^2) = 2\pi c_0\, \alpha_{\mathrm{em}}\, N_c \int_0^1 \mathrm{d}\alpha\, \frac{1 + (1-2\alpha)^2}{6\pi^2 Q^2 \alpha(1-\alpha)}\,, \tag{7.6}$$

$$\sigma_L^{(q)}(Q^2) = 2\pi c_0\, \alpha_{\mathrm{em}}\, N_c \int_0^1 \mathrm{d}\alpha\, \frac{2}{3\pi^2 Q^2}\,. \tag{7.7}$$

These endpoint singularities arise from the large r behaviour of the integrand. For dipole cross sections which rise for large r only as $\propto r$ the singularities vanish. This applies in

particular to the physical dipole models we found in the literature. However, also in the case of convergence, for increasing Q^2 the integral becomes dominated by integrand peaks in narrow regions with α near the endpoints and relatively large r.

The aforementioned endpoint singularities with $\hat{\sigma}$ as in (7.2) are regularised by the introduction of a *non-zero quark mass* in the photon densities:

$$\sigma_T^{(q)}(Q^2) = 2\pi c_0 \alpha_{\mathrm{em}} N_c \int_0^1 \mathrm{d}\alpha \, \frac{2(1-2\alpha(1-\alpha))(\alpha(1-\alpha)Q^2+m_q^2)+m_q^2}{6\pi^2(\alpha(1-\alpha)Q^2+m_q^2)^2}\,, \tag{7.8}$$

$$\sigma_L^{(q)}(Q^2) = 2\pi c_0 \alpha_{\mathrm{em}} N_c \int_0^1 \mathrm{d}\alpha \, \frac{2(\alpha(1-\alpha)Q^2)^2}{3\pi^2 Q^2(\alpha(1-\alpha)Q^2+m_q^2)^2}\,, \tag{7.9}$$

$$\begin{aligned}\sigma_{tot}^{(q)}(Q^2) &= \sigma_T^{(q)}(Q^2) + \sigma_L^{(q)}(Q^2) \\ &= 2\pi c_0 \, \alpha_{\mathrm{em}} \, N_c \, \frac{Q\sqrt{Q^2+4m_q^2} + 4(Q^2-m_q^2)\,\mathrm{arctanh}\sqrt{\frac{Q^2}{Q^2+4m_q^2}}}{3\pi^2 Q^3 \sqrt{Q^2+4m_q^2}}\,.\end{aligned} \tag{7.10}$$

For the massive case the simple rescaling property does not hold. Fitting the ansatz $r_{\max} = \sqrt{C^2/(Q^2+4m_q^2)}$ to samples with $Q^2 = 1, 10, 100\,\mathrm{GeV}^2$ and $m_q = 0.14, 1.3\,\mathrm{GeV}$ gives $C = 2.28$ (or $C = 2.34$ if massless samples are included). This fit works well for small masses but produces errors up to the order 25 % for the maximum position in the charm mass case. For the ansatz $r|_{\max} = \sqrt{C^2/(Q^2+Dm_q^2)}$ a simultaneous fit of C, D automatically reproduces the value of C needed for the massless case, even if no massless samples are included in the fit. We perform another fit with the aforementioned samples where the value of C is fixed to the massless value and D is varied. The result is

$$r|_{\max} = \sqrt{\frac{C^2}{Q^2+Dm_q^2}} \qquad \text{with } C = 2.40\,, \quad D = 7.12 \tag{7.11}$$

with maximal relative errors for the maximum position of the order 1 %.

We thus see that for the simple ansatz (7.2) there is a dipole size for which the contribution to the γp cross section is maximal, and this size is roughly of type (7.1), see (7.5) and (7.11). On the other hand we keep in mind, that the relevant distributions are quite broad, see figure 7.1. For a better estimate of the dominant dipole size one could consider the centre of the integrand instead of the integrand's maximum position. This goes beyond the (unintegrable) ansatz (7.2) and depends on more details of the specific dipole cross section. We shall demonstrate in the following section that a type (7.1) relation may not be taken in a straight-forward way to substitute energy dependencies. A mere change of the numerical constants in a type (7.1) relation by considering the integrand centre instead of its maximum will turn out to be rather irrelevant to this overall conclusion.

7.2 Substitution of scales via typical dipole sizes

Let us consider the Golec-Biernat-Wüsthoff model described in section 4.2. Its dipole cross section $\hat{\sigma}_{GBW}(r,x)$ depends on Bjorken-x, see (4.9), and therefore not only on W but also

on Q^2. We may ask now, whether this corresponds effectively to a Q^2-independent dipole cross section $\hat{\sigma}^{mod}_{GBW}(r,W)$ by taking typical dipole sizes via (7.1) into account. Let us restrict to the case of massless quarks and define an effective Q^2 via (7.5):

$$\bar{Q}^2(r) := \frac{C^2}{r^2}\,, \qquad C = 2.40\,. \tag{7.12}$$

We consider the GBW dipole cross section as a function of W and Q^2,

$$\hat{\sigma}_{GBW}(r,x) = \hat{\sigma}_{GBW}\left(r, \frac{Q^2}{Q^2+W^2}\right)\,. \tag{7.13}$$

The simplest effective $\hat{\sigma}^{mod}_{GBW}(r,W)$ is obtained now by keeping W fixed and eliminating all occurrences of Q^2 in $\hat{\sigma}_{GBW}$ via identification with $\bar{Q}^2(r)$:

$$\to \hat{\sigma}^{mod}_{GBW}(r,W) := \hat{\sigma}_{GBW}\left(r, \frac{\bar{Q}^2(r)}{\bar{Q}^2(r)+W^2}\right) = \hat{\sigma}_{GBW}\left(r, \frac{C^2}{C^2+r^2W^2}\right)\,. \tag{7.14}$$

Figure 7.2 (left) shows the effect of the replacement (7.14) on F_2. There, contributions from massless u, d, s-quarks are considered. The curves for F_2 of the original GBW model (solid lines) are significantly modified by the substitution (7.14) (dashed lines), in particular, the dependence on Q^2 and x is altered. This is also the case when using $C = 4.8$ in (7.14) (dotted lines), which corresponds to choosing higher typical dipole sizes. Note that a similar procedure for non-vanishing quark masses by "inversion" of (7.11) instead of (7.5) would lead to negative $\bar{Q}^2$ which is obviously questionable. The reason is that for $Q^2 \to 0$ the maximum position given by the massive formula (7.11) converges to a finite upper bound. In other words, sizes r larger than this limit do not correspond to any maximum position.

We shall now consider the reversed case, that is, introducing a Q^2 dependence via the typical dipole size into a dipole cross section, which depends only on r and W but not on Q^2. Actually there are only few examples for dipole models of the latter kind. We choose here a model proposed by Donnachie and Dosch in [140] based on previous work by them and Rueter [161, 162, 163], which in turn is based on work by Nachtmann [164]. This model is based on Regge theory with a two-Pomeron exchange, that is, the exchange of a soft and a hard Pomeron. The intercepts of the Pomeron trajectories are denoted by $(1+\epsilon_{\rm soft})$ for the soft and $(1+\epsilon_{\rm hard})$ for the hard Pomeron. The dipole cross section in this model is:

$$\hat{\sigma}_{\rm DDR}(r,W^2) = A_0\, r\, \left(1-e^{-r/R_0}\right) \times \begin{cases} 0 & \text{if } r < R_c \\ (W^2/W_0^2)^{\epsilon_{\rm hard}} & \text{if } R_c \le r \le a \\ (W^2/W_0^2)^{\epsilon_{\rm soft}} & \text{if } r > a \end{cases} \tag{7.15}$$

with the parameter values, see p.3-4 of [140],

$$\epsilon_{\rm hard} = 0.42, \qquad A_0 = 10954\,\mu{\rm b\,GeV}, \qquad R_c = 0.81/{\rm GeV}, \qquad R_0 = 3.1\,a, \tag{7.16}$$

$$\epsilon_{\rm soft} = 0.09, \qquad W_0^2 = (20\,{\rm GeV})^2, \qquad a = 1.75/{\rm GeV}. \tag{7.17}$$

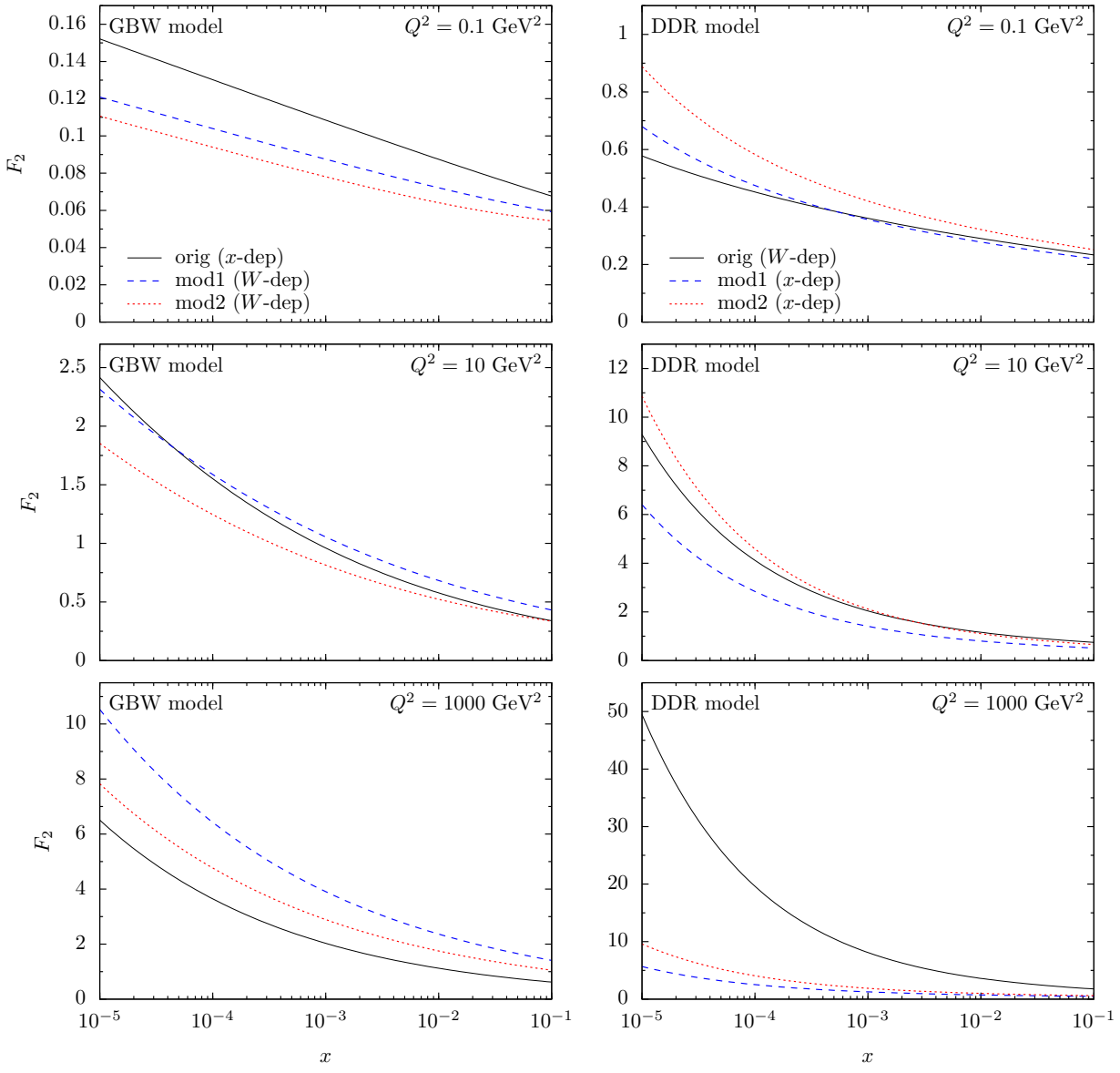

Figure 7.2: F_2 for two dipole models and modified versions with altered energy dependencies via "effective scale" substitutions. Here, x is varied and Q^2 is kept fixed for each plot. Left column for the Golec-Biernat-Wüsthoff model (solid curves) and modification (7.14) with $C = 2.4$ (dashed curves) and $C = 4.8$ (dotted curves). Right column for the Donnachie-Dosch-Rueter model (solid curves) and modification (7.20) with $C = 2.4$ (dashed curves) and $C = 4.8$ (dotted curves). Quark masses are set to zero in all cases.

The short-distance cut-off R_c is introduced to suppress the otherwise overestimated contributions for extrapolation of this non-perturbative model to higher values of Q^2.

We consider now the introduction of a Q^2 dependency into the dipole cross section based on typical dipole sizes according to (7.1), again for massless quarks only. For this we define the typical dipole size according to (7.5) as

$$\bar{r}(Q^2) = \frac{C}{Q}\,. \tag{7.18}$$

We consider the dipole cross section as a function f of r and the dimensionless quantity Wr,

$$\hat{\sigma}_{DDR}(r, W) =: f(r, Wr) \tag{7.19}$$

and introduce a Q^2 dependency into $\hat{\sigma}_{DDR}$ by replacing r by $\bar{r}(Q^2)$ only in the dimensionless argument Wr

$$\rightarrow \hat{\sigma}^{mod}_{DDR}(r, W, Q^2) := f(r, W\bar{r}(Q^2)) = \hat{\sigma}_{DDR}(r, WC/(rQ)). \tag{7.20}$$

That is, the replacement changes the cross section only via terms which are sensitive to the external scale W. Figure 7.2 (right column) shows the effect of this substitution on F_2. Only contributions from massless u, d, s-quarks are considered. There are significant deviations in F_2 between using $\hat{\sigma}_{DDR}(r, W)$ (solid curves) and using $\hat{\sigma}^{mod}_{DDR}(r, W, Q^2)$ with $C = 2.4$ (dashed curves) or $C = 4.8$ (dotted curves). In particular, the x and Q^2 dependence is heavily altered. We note that the solid curves are well off the data as a consequence of using three massless quarks in contrast to the model presented in [140]. Nevertheless, setting the quark masses to zero avoids additional complications for the effect we are demonstrating here.

Actually, we could infer already from the previous chapters that the choice of the energy variable in $\hat{\sigma}$ is crucial at high Q^2. We saw that *any* dipole cross section of the form $\hat{\sigma}(r, W)$ fails to describe the data at high Q^2. In contrast, the GBW model with $\hat{\sigma}_{GBW}(r, x)$ provides a good fit to the data also at high Q^2. From this we see that a dependence of $\hat{\sigma}$ on Q^2 in addition to W can certainly not be eliminated or introduced by an effective scale argument of the type (7.1) in the regime of high Q^2. Let us compare ratios of F_2 for different values of Q^2 for the models and replacements discussed in this section above. An illustration is given in figure 7.3, where in addition the general bound (6.14) valid for any $\hat{\sigma}(r, W)$ are shown. We see as expected, that $\hat{\sigma}^{mod}_{GBW}(r, W)$ as well as $\hat{\sigma}_{DDR}(r, W)$ respect the general bounds. In contrast, the bounds are violated for both Q^2-dependent cross sections $\hat{\sigma}_{GBW}(r, x)$ and $\hat{\sigma}^{mod}_{DDR}(r, W, Q^2)$.

We conclude that different choices of the energy variables are in general not effectively equivalent at high Q^2. Furthermore, also at medium or low Q^2, the effective dipole size argument (7.1) does not allow to consider different choices of the energy variable to be equivalent in the straight-forward way tested here. Since the involved distributions are rather broad in r, a reduction to one effective dipole size tends to oversimplify the picture.

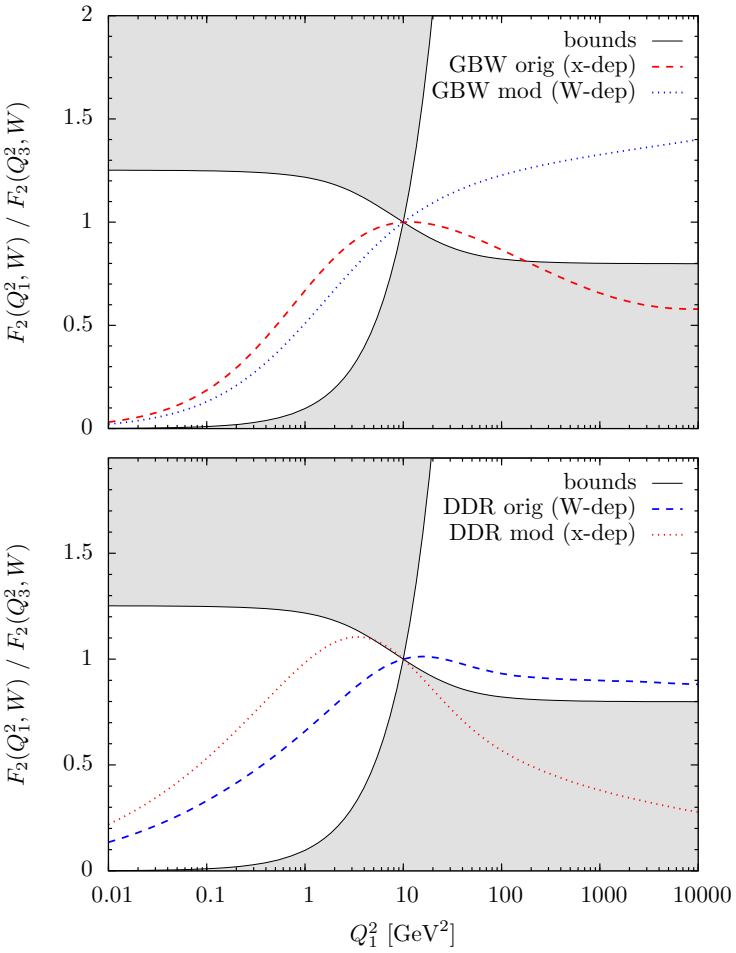

Figure 7.3: Ratio $F_2(W, Q_1^2)/F_2(W, Q_3^2)$ for two dipole models and modified versions with altered energy dependencies via "effective scale" substitutions. Here, Q_1^2 is varied and $Q_3^2 = 10\,\text{GeV}$ is kept fixed. Upper plot for the Golec-Biernat-Wüsthoff model (dashed curves) and modification (7.14) with $C = 2.4$ (dotted curves). Lower plot for the Donnachie-Dosch-Rueter model (dashed curves) and modification (7.20) with $C = 2.4$ (dotted curves). Quark masses are set to zero in all cases. The shaded area is excluded by the bound (6.14) for any Q^2 independent dipole cross section $\hat{\sigma}^{(q)}(r, W)$.

The dipole size argument (7.1) furthermore neglects the influence of W and further details entering through the dipole cross section. Therefore we suggest that the effective dipole size (7.1) can not be regarded as a strict equality without further qualification.

Chapter 8

Ioffe times

8.1 Ioffe times in DIS

In this chapter we study the Ioffe times [165, 166] for scattering of real and virtual photons on hadrons. In particular, we shall calculate such Ioffe times for $\gamma^* p$ reactions from the DIS structure functions within the framework of the colour dipole model. Please see [167] for a discussion of Ioffe time distributions in the context of the Operator Product Expansion formalism, and [168] for the Ioffe time structure of the gluon distribution function in the double logarithmic approximation.

In the colour dipole picture, the virtual photon fluctuates into a quark-antiquark pair where the three-momenta are conserved but the energy is not, see chapter 5. The energy mismatch

$$\Delta E = k^0 + k'^0 - q^0 \,, \tag{8.1}$$

defined in the proton rest frame, implies via the uncertainty relation that this colour dipole is unstable and lives no longer than the Ioffe time

$$\tau = \frac{1}{\Delta E} \,. \tag{8.2}$$

The key idea of the dipole model is the splitting of the γp scattering into a two-step process: First the photon splits into a dipole, then the dipole interacts with proton. This separation however is expected to be valid only as long as the dipole lifetime is well above the timescale of the hadronic interactions. In applications of the dipole model one frequently finds simple estimates for Ioffe times and if they turn out to be of order of several femtometer or larger this is used as justification of the dipole model. Furthermore, we saw in section 6.2 that for $W = 60 - 245$ GeV the standard dipole model necessarily fails to describe the HERA data for Q^2 larger than about 100 to 200 GeV^2, see figure 6.10. It is interesting to see whether this could be related to Ioffe times becoming too short.

Here, we shall really calculate the Ioffe times and not estimate them based on effective quantities, since the latter may lead to oversimplifications, see section 7.2. It turns out that even for a fixed kinematical point (W, Q^2) for the $\gamma^* p$ reaction one has a whole distribution

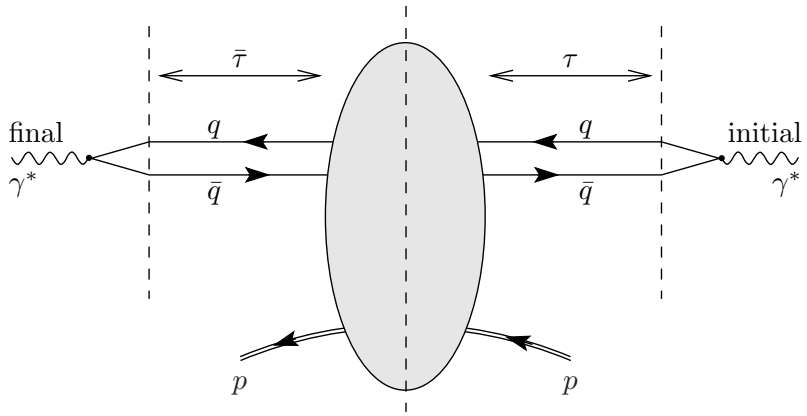

Figure 8.1: Basic diagram for the description of the cross sections $\sigma_{T,L}$ of $\gamma^* p$ scattering in the standard dipole approach.

of Ioffe times. In fact, the $\gamma^* p$ total absorption cross sections must be considered as imaginary part of the forward scattering amplitude $\gamma^* p \to \gamma^* p$, see figure 8.1. Thus, we shall have to deal with two Ioffe times, one for the initial state γ^* and one for the final state γ^*. Both times have distributions which also depend on the polarisation, transverse or longitudinal, of the γ^*. We shall calculate such distributions in the following for the GBW model, which describes the HERA data rather well. It is clear, that such an analysis of the occurring Ioffe times serves only as a consistency check. That is, the analysis can not prove the applicability of the dipole picture, but it may point out kinematical ranges where the interpretation of the picture becomes questionable.

Let us first note some basic features for the offshellness ΔE in (8.1) which follow directly from the kinematics of the $\gamma \to q\bar{q}$ splitting. Explicitly we have from (8.1)

$$\Delta E(\boldsymbol{k}_T, \alpha) = \sqrt{\alpha^2 \boldsymbol{q}^2 + k_T^2 + m_q^2} + \sqrt{(1-\alpha)^2 \boldsymbol{q}^2 + k_T^2 + m_q^2} - q^0 \,. \tag{8.3}$$

Figure 8.2 shows the dependencies of ΔE on the quark transverse momentum k_T (left graph) and longitudinal momentum fraction α for the case of massless quarks. We note that ΔE strongly peaks at the longitudinal momentum endpoints and rises monotonously with the transverse momentum, see (8.3) and the figure. The minimal possible energy mismatch as a function of α is reached at vanishing transverse momentum. The minimal possible energy mismatch as a function of α is

$$\Delta E_{min}(\alpha) := \Delta E(0, \alpha) \quad = \sqrt{\alpha^2 \boldsymbol{q}^2 + m_q^2} + \sqrt{(1-\alpha)^2 \boldsymbol{q}^2 + m_q^2} - q^0 \tag{8.4}$$

reaching the absolute minimum of ΔE at $\alpha = 1/2$,

$$\Delta E_{min,tot} := \Delta E(0, 1/2) = \sqrt{\boldsymbol{q}^2 + 4m_q^2} - q^0 \,. \tag{8.5}$$

That is, we have

$$\Delta E(\boldsymbol{k}_T, \alpha) \geq \Delta E_{min}(\alpha) \geq \Delta E_{min,tot} \geq 0 \,, \tag{8.6}$$

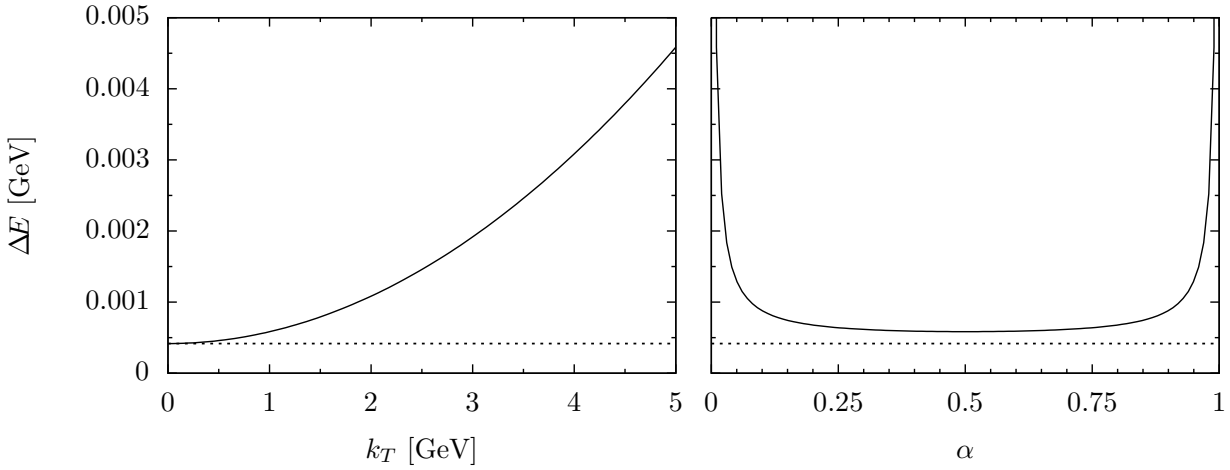

Figure 8.2: Offshellness ΔE of the dipole in dependence of quark transverse momentum k_T (left graph) and longitudinal momentum fraction α (right graph). Parameters are $W = 150$ GeV, $Q^2 = 10$ GeV^2, $m_q = 0$. For the left graph $\alpha = 1/2$ and for the right graph $k_T = 1$ GeV are used. The dotted line shows the absolute minimum ΔE_{min} for the given external kinematics, that is ΔE at $k_T = 0$, $\alpha = 1/2$.

where the last inequality is strict for any $Q^2 > 0$[1]. Thus we see that there is an a priori minimal value for the energy mismatch ΔE. Figure 8.3 shows the minimal energy mismatch $\Delta E_{min,tot}$ in dependence of Q^2 for massless quarks (left graph) and massive quarks (right graph). We see that ΔE becomes non-negligible for small W, large m_q or large Q^2. This means the corresponding Ioffe times become short and the applicability of the dipole picture becomes questionable. The restrictions following directly from $\Delta E_{min,tot}$ are numerically not very tight, see figure 8.3.

We considered so far only the *a priori* lower bound on ΔE. Let us now study which values for ΔE are actually relevant for the cross section. From this we expect stronger bounds on the range of applicability of the dipole model. The expressions (4.4) and (4.5) for σ_T and σ_L, respectively, involve the γ^* wave functions in transverse position space. This is appropriate since the formulae are simple there due to the assumption that the $q\bar{q}$ pair does not change its transverse size $\boldsymbol{r}$ nor the longitudinal momentum fraction α of the quark in the scattering on the proton. These forms are not appropriate for a study of the energy mismatches ΔE occurring in the initial splitting process $\gamma^* \to q\bar{q}$ and $\Delta\bar{E}$ occurring in the final fusion process $q\bar{q} \to \gamma^*$. For this we have to employ the transverse momentum representation of the initial and final photon wave functions. We insert the representations (5.43) and (5.44) for both the initial and final state photon wave functions in (4.4) and

[1] The inequality is also strict for $Q^2 = 0$ and $m_q > 0$. This is the reason why a real photon can not decay into an on-shell e^+e^- pair in the vacuum.

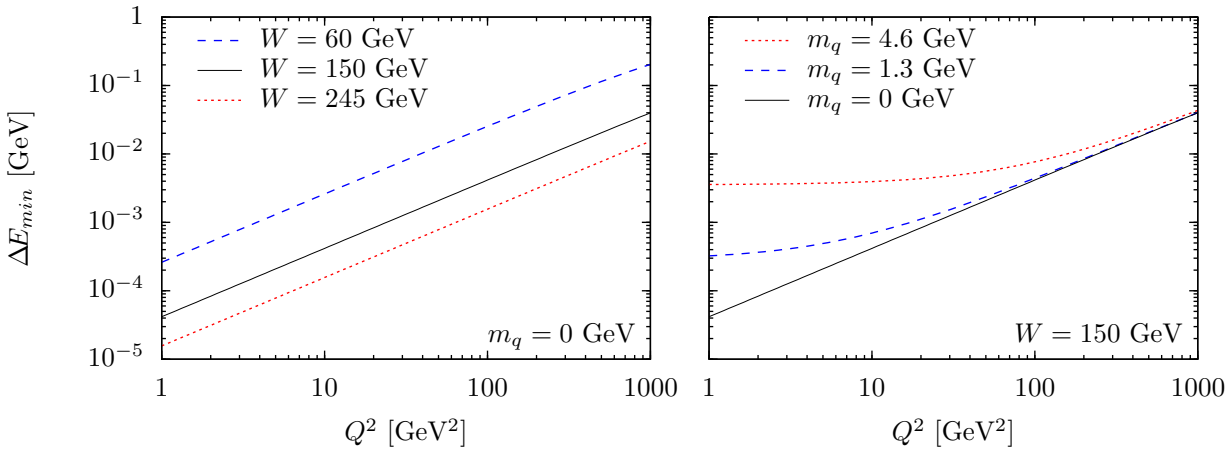

Figure 8.3: Minimal offshellness ΔE_{min} of dipole in dependence of photon virtuality Q^2 for different values of the energy W (left graph) and quark masses m_q (right graph).

(4.5), respectively, and obtain

$$\sigma_T = \sum_q \sum_{\lambda,\lambda'} \int \mathrm{d}\alpha \int \frac{\mathrm{d}^2\bar{k}_T}{(2\pi)^2} \int \frac{\mathrm{d}^2 k_T}{(2\pi)^2} \left(\tilde{\psi}^{(q)\,\pm}_{\gamma,\,\lambda\lambda'}(\alpha, \bar{\boldsymbol{k}}_T)\right)^* \tilde{\hat{\sigma}}^{(q)}(\boldsymbol{k}_T - \bar{\boldsymbol{k}}_T) \tilde{\psi}^{(q)\,\pm}_{\gamma,\,\lambda\lambda'}(\alpha, \boldsymbol{k}_T)\,, \tag{8.7}$$

$$\sigma_L = \sum_q \sum_{\lambda,\lambda'} \int \mathrm{d}\alpha \int \frac{\mathrm{d}^2\bar{k}_T}{(2\pi)^2} \int \frac{\mathrm{d}^2 k_T}{(2\pi)^2} \left(\tilde{\psi}^{(q)\,L}_{\gamma,\,\lambda\lambda'}(\alpha, \bar{\boldsymbol{k}}_T)\right)^* \tilde{\hat{\sigma}}^{(q)}(\boldsymbol{k}_T - \bar{\boldsymbol{k}}_T) \tilde{\psi}^{(q)\,L}_{\gamma,\,\lambda\lambda'}(\alpha, \boldsymbol{k}_T) \tag{8.8}$$

with the Fourier transform of the dipole cross section

$$\tilde{\hat{\sigma}}^{(q)}(\boldsymbol{k}_T, W) = \int \mathrm{d}^2 r\, e^{i\boldsymbol{k}_T \boldsymbol{r}} \hat{\sigma}^{(q)}(r, W)\,, \tag{8.9}$$

where we omit here and in the following dependencies e.g. on Q^2 in the notation for the sake of brevity. Note, that we omitted the explicit restriction $0 < \alpha < 1$ on the α integration range for another reason. Also, we shall consider the expressions (5.41) and (5.41) for the photon wave functions instead of their high-energy approximations (5.47) and (5.48). We omit both of these high-energy approximations to begin with, since it is not a priori obvious, that they are irrelevant for the calculation of Ioffe times, see the discussion in section 5.2. At a later stage, we shall then consider both approximations in order to quantify their influence on our results. We shall see, that the α integration is automatically restricted to finite integration intervals for the calculation of Ioffe times.

The energy mismatches ΔE for the initial γ^* splitting to $q\bar{q}$ and $\Delta\bar{E}$ for the final $q\bar{q}$ fusion to γ^* are (see chapter 5)

$$\Delta E := k^0 + k'^0 - q^0\,, \tag{8.10}$$

$$\Delta\bar{E} := \bar{k}^0 + \bar{k}'^0 - q^0\,. \tag{8.11}$$

The corresponding Ioffe times are

$$\tau = \frac{1}{\Delta E}\,, \tag{8.12}$$

$$\bar{\tau} = \frac{1}{\Delta \bar{E}}\,. \tag{8.13}$$

The expressions (8.10) and (8.11) are not very suitable for direct numerical evaluation due to large cancellations between the terms of the sum. We use the identity

$$\sqrt{1+\epsilon} - 1 = \frac{(\sqrt{1+\epsilon}-1)(\sqrt{1+\epsilon}+1)}{\sqrt{1+\epsilon}+1} = \frac{\epsilon}{\sqrt{1+\epsilon}+1}, \qquad \epsilon \in \mathbb{R} \tag{8.14}$$

to rewrite (8.10) and (8.11) as sums of non-negative terms:

$$\Delta E = \frac{\alpha^2 Q^2 + m_q^2}{k^0 + |\alpha|\, q^0} + \frac{(1-\alpha)^2 Q^2 + m_q^2}{k^0 + |1-\alpha|\, q^0} + \begin{cases} 2\,|\alpha|\, q^0 & \text{if } \alpha < 0 \\ 0 & \text{if } 0 \le \alpha \le 1 \\ 2\,|1-\alpha|\, q^0 & \text{if } \alpha > 1 \end{cases}, \tag{8.15}$$

$$\Delta \bar{E} = \frac{\alpha^2 Q^2 + m_q^2}{\bar{k}^0 + |\alpha|\, q^0} + \frac{(1-\alpha)^2 Q^2 + m_q^2}{\bar{k}^0 + |1-\alpha|\, q^0} + \begin{cases} 2\,|\alpha|\, q^0 & \text{if } \alpha < 0 \\ 0 & \text{if } 0 \le \alpha \le 1 \\ 2\,|1-\alpha|\, q^0 & \text{if } \alpha > 1 \end{cases}. \tag{8.16}$$

We see that ΔE depends on $\boldsymbol{k}_T$, $\Delta \bar{E}$ on $\bar{\boldsymbol{k}}_T$. Thus, the cross sections (8.7) and (8.8) involve the superpositions of amplitudes of various ΔE in the initial and various $\Delta \bar{E}$ in the final state. We define now the joint distribution of $\eta = \Delta E$ and $\bar{\eta} = \Delta \bar{E}$ by

$$\begin{aligned} \frac{\partial^2 \sigma_T}{\partial \eta\, \partial \bar{\eta}} = \sum_{q,\lambda,\lambda'} \int \mathrm{d}\alpha \frac{\mathrm{d}^2 \bar{k}_T}{(2\pi)^2} \frac{\mathrm{d}^2 k_T}{(2\pi)^2} \left(\tilde{\psi}^{(q)\,\pm}_{\gamma\,\lambda\lambda'}(\alpha, \bar{\boldsymbol{k}}_T) \right)^* \delta(\bar{\eta} - \Delta\bar{E}(\bar{\boldsymbol{k}}_T, \alpha)) \\ \cdot\, \tilde{\tilde{\sigma}}^{(q)}(\boldsymbol{k}_T - \bar{\boldsymbol{k}}_T)\, \delta(\eta - \Delta E(\boldsymbol{k}_T, \alpha))\, \tilde{\psi}^{(q)\,\pm}_{\gamma\,\lambda\lambda'}(\alpha, \boldsymbol{k}_T)\,, \end{aligned} \tag{8.17}$$

$$\begin{aligned} \frac{\partial^2 \sigma_L}{\partial \eta\, \partial \bar{\eta}} = \sum_{q,\lambda,\lambda'} \int \mathrm{d}\alpha \frac{\mathrm{d}^2 \bar{k}_T}{(2\pi)^2} \frac{\mathrm{d}^2 k_T}{(2\pi)^2} \left(\tilde{\psi}^{(q)\,L}_{\gamma\,\lambda\lambda'}(\alpha, \bar{\boldsymbol{k}}_T) \right)^* \delta(\bar{\eta} - \Delta\bar{E}(\bar{\boldsymbol{k}}_T, \alpha)) \\ \cdot\, \tilde{\tilde{\sigma}}^{(q)}(\boldsymbol{k}_T - \bar{\boldsymbol{k}}_T)\, \delta(\eta - \Delta E(\boldsymbol{k}_T, \alpha))\, \tilde{\psi}^{(q)\,L}_{\gamma\,\lambda\lambda'}(\alpha, \boldsymbol{k}_T)\,. \end{aligned} \tag{8.18}$$

Note that for $\eta \neq \bar{\eta}$ these distributions need not be positive. In the next section we shall study in particular the distribution of the sum of ΔE and $\Delta \bar{E}$

$$\Delta E_+ = \Delta E + \Delta \bar{E}\,. \tag{8.19}$$

We also define the corresponding Ioffe time

$$\tau_+ = \frac{1}{\Delta E_+}\,. \tag{8.20}$$

Due to $\Delta E \geq 0$, $\Delta\bar{E} \geq 0$ we find that at a given ΔE_+ only ΔE and $\Delta\bar{E}$ less or equal to ΔE_+ can contribute. Moreover, for some given ΔE_+ at least one of ΔE, $\Delta\bar{E}$ must be greater or equal to $\Delta E_+/2$. Therefore the distributions

$$\frac{\mathrm{d}\sigma_{T,L}}{\mathrm{d}\Delta E_+} = \int_0^\infty \mathrm{d}\eta_1 \int_0^\infty \mathrm{d}\eta_2 \, \frac{\partial^2 \sigma_{T,L}}{\partial\eta_1 \, \partial\eta_2} \, \delta(\Delta E_+ - \eta_1 - \eta_2) \tag{8.21}$$

allow to test for large energy mismatches, i.e. short Ioffe time, which are relevant for the cross section. For the typical interaction time τ_{had} of the dipole with the proton we follow [145] and use the inverse of the QCD scale, $\tau_{had} \approx 1/\lambda_{QCD}$, corresponding to the confinement scale of the light quarks. A necessary condition for both dipole lifetimes τ and $\bar{\tau}$ being much larger than τ_{had} is thus $\Delta E_+/2 \ll \lambda_{QCD}$ or, equivalently,

$$2\tau_+ \gg \tau_{had} \approx \frac{1}{\lambda_{QCD}} \approx 2 \text{ fm}\,. \tag{8.22}$$

If this condition is violated for a relevant portion of the contributions to the total $\gamma^* p$ cross section, the applicability of the dipole model is questionable.

8.2 Results for the Golec-Biernat-Wüsthoff model

In this section we shall evaluate $\mathrm{d}\sigma_{T,L}/\mathrm{d}\Delta E_+$ defined in (8.21) using the Golec-Biernat-Wüstoff model [138]. We described this model in section 4.2 and recall that it matches the F_2 data from HERA quite well. If it also describes the longitudinal structure function F_L must be left open since F_L measurements from HERA have been published just recently [160]. We note again, that its dipole cross section (4.9) depends not only on W but also on Q^2. The Fourier transformation of (4.9) gives

$$\tilde{\sigma}^{(q)}_{\mathrm{GBW}}(\boldsymbol{k}_T) = \sigma_0 \left[(2\pi)^2 \delta^{(2)}(\boldsymbol{k}_T) - 4\pi r_0^2 \exp(-r_0^2 k_T^2)\right]. \tag{8.23}$$

In order to integrate out the azimuthal angles we decompose the photon wave functions

$$\tilde{\psi}^{(q)\,\pm,L}_{\lambda\lambda'}(\alpha, \boldsymbol{k}_T) = \sum_{n=-2}^{2} e^{in\phi_k} \tilde{\psi}^{(q)\,\pm,L}_{\lambda\lambda' n}(\alpha, k_T) \tag{8.24}$$

with $\phi_k = \arg(k_{T,1} + i\, k_{T,2})$. We use the substitution

$$\int_0^\infty \mathrm{d}k_T \, f(k_T) = \int_{\Delta E_{min}(\alpha)}^\infty \mathrm{d}(\Delta E) \, \frac{k^0 k'^0}{k_T(k^0 + k'^0)} f(k_T(\Delta E, \alpha)) \tag{8.25}$$

and find for the distribution in ΔE_+

$$\frac{\mathrm{d}\sigma_{T,L}^{(q)}}{\mathrm{d}\Delta E_+} = \frac{\mathrm{d}\sigma_{T,L}^{(q),const}}{\mathrm{d}\Delta E_+} + \frac{\mathrm{d}\sigma_{T,L}^{(q),exp}}{\mathrm{d}\Delta E_+}\,, \tag{8.26}$$

$$\frac{\mathrm{d}\sigma_{T,L}^{(q),const}}{\mathrm{d}\Delta E_+} = \sigma_0 \int_{\alpha_{min}(\Delta E_+)}^{\alpha_{max}(\Delta E_+)} \mathrm{d}\alpha \; \frac{k^0 k'^0}{2\pi(k^0+k'^0)} \sum_{\lambda,\lambda',n} \left|\tilde{\psi}_{\lambda\lambda' n}^{(q)\,\pm,L}(\alpha, k_T)\right|^2\,, \tag{8.27}$$

$$\begin{aligned}\frac{\mathrm{d}\sigma_{T,L}^{(q),exp}}{\mathrm{d}\Delta E_+} = \sigma_0 \int_{\alpha_{min}(\Delta E_+)}^{\alpha_{max}(\Delta E_+)} \mathrm{d}\alpha \int_{-\Delta E_{min}(\alpha)}^{\Delta E_{min}(\alpha)} \mathrm{d}\Delta E_- \; \frac{k^0 k'^0 \bar{k}^0 \bar{k}'^0 r_0^2}{2\pi(k^0+k'^0)(\bar{k}^0+\bar{k}'^0)} \\ \cdot\, e^{-r_0^2(k_T-\bar{k}_T)^2} \sum_{\lambda,\lambda',n} \frac{I_n(2r_0^2 k_T \bar{k}_T)}{\exp(2r_0^2 k_T \bar{k}_T)}\, \tilde{\psi}_{\lambda\lambda' n}^{(q)\,\pm,L}(\alpha, k_T) \left(\tilde{\psi}_{\lambda\lambda' n}^{(q)\,\pm,L}(\alpha, \bar{k}_T)\right)^*\,.\end{aligned} \tag{8.28}$$

Here, the energies and momenta of quark and anti-quark (5.30) are understood as functions of α, ΔE, $\Delta\bar{E}$ via

$$k_T = \sqrt{\frac{((\Delta E+q^0)^2-\boldsymbol{q}^2)\,((\Delta E+q^0)^2-(1-2\alpha)^2\boldsymbol{q}^2)}{4(\Delta E+q^0)^2} - m_q^2}\,, \tag{8.29}$$

$$\bar{k}_T = \sqrt{\frac{\left((\Delta\bar{E}+q^0)^2-\boldsymbol{q}^2\right)\left((\Delta\bar{E}+q^0)^2-(1-2\alpha)^2\boldsymbol{q}^2\right)}{4(\Delta\bar{E}+q^0)^2} - m_q^2}\,. \tag{8.30}$$

The integration area in (8.26) is finite for fixed ΔE_+ and its boundaries correspond to vanishing transverse momenta. The extremal values of α are given by the two solutions of the equation $\Delta E_+/2 = \Delta E|_{k_T=0}$ with respect to α:

$$\alpha_{\substack{max\\min}} = \frac{1}{2} \pm \frac{\Delta E_+/2+q^0}{2\,|\boldsymbol{q}|}\sqrt{1-\frac{4m^2}{(\Delta E_+/2+q^0)^2-\boldsymbol{q}^2}}\,. \tag{8.31}$$

In all cases considered in the following $\alpha_{min} \approx 0$ and $\alpha_{max} \approx 1$.

We perform the residual integrations in (8.26) numerically with a `C++` program and take care of various numerical issues. We avoid loss of precision due to large cancellations in the evaluation of sums by rewritings similar to (8.14). Non-negligible contributions occur at large arguments of the Bessel functions I_n, thus it is important to avoid finite range problems due to the exponential factor in I_n in that region. We already wrote (8.26) in a form which shows that this exponential rise is effectively cancelled. We achieve this cancellation at the algebraical level by using the the expansion of [169] for the I_n at large arguments. Further, we find that the implementation of the I_n in [170], also used by other general purpose libraries such as [171], gives wrong results at large arguments already before the expected finite range problem arises. We make sure the I_n evaluations in our program are reliable by alternatively using the implementations of [172] and cross-checking with various versions of [124]. Last but not least, we perform the same numerical

integrations employing different algorithms and cross-check the results. Specifically, we use Monte-Carlo as well as deterministic quadratures implemented in [173, 172] and modify them slightly to overcome deficiencies due to significant underestimation of the integration errors. By these means we make sure the integration error is under control.

Figure 8.4 shows our results for the ΔE_+ distributions of the total $\gamma^* p$ cross section in the GBW model. Here, the full expressions (5.41) and (5.42) are used for the photon wave functions rather than using their high energy approximation. Also we integrate over the full range in α which is slightly bigger than $[0,1]$, see (8.31). Contributions to the total cross section may be directly read off the plots in terms of areas under the curves, since the normalisation includes a ΔE_+ factor which compensates for the logarithmic abscissa scale. The onset at the respective kinematical lower bound $\Delta E_+ = 2\Delta E_{min,tot}$, see (8.19) and (8.5), is clearly visible for all curves and sharp for longitudinal γ^* polarisation. The ΔE_+ distributions for transverse γ polarisation are significantly broader than those for longitudinal γ polarisation. For both polarisation types the distributions become narrow for increasing values of Q^2. In all cases the cross section is dominated by ΔE_+-values not more than one order of magnitude above the absolute minimum $2\Delta E_{min,tot}$ given by pure kinematical restrictions as shown in figure 8.3. We note that the typical Ioffe times τ_+ are well above 1 fm for all $Q^2 \leq 100$ GeV^2. There, the necessary separation condition (8.22) is satisfied. However, for $Q^2 = 1000$ GeV^2 we see that at least for transversely polarised photons the separation condition (8.22) is clearly violated – non-negligible contributions come from Ioffe times τ_+ which are actually smaller than 1 fm.

We study the effect of the high energy approximation (5.49) and (5.50) for the photon wave functions as well as of restricting the α integration range to $[0,1]$. The Ioffe-time distributions in figure 8.4 are essentially unaltered when performing the α integration only over $[0,1]$. The α range specified by (8.31) is only slightly larger than $[0,1]$ and no sufficiently strong enhancement of the integrand compensates for this fact in the kinematical ranges considered here. As expected, deviations in the Ioffe-time distributions by the high energy approximation for the photon wave functions become larger for increasing ΔE_+. However, these deviations are numerically negligible for the peak regions of the distributions considered here. They would be visible in figure 8.4 only at the large ΔE_+ tails of the curves for high Q^2 and may therefore be omitted in the present discussion. We stress that this result is not obvious, since the interplay of the high energy approximation and the dipole energy mismatch is non-trivial in the longitudinal endpoint regions, as discussed above.

Conclusions

We have calculated Ioffe times in the Golec-Biernat-Wüsthoff dipole model, where the energy $W = 150$ GeV was chosen. We find that typical Ioffe times are large with respect to hadronical scales for $Q^2 \leq 100$ GeV^2, such that no inconsistency arises. However, at large photon virtuality, $Q^2 = 1000$ GeV^2, typical Ioffe times are of similar order as hadronical scales, which violates a standard assumption for the validity of the dipole picture. In section 6.2 we have seen that Q^2-independent dipole cross sections fail to describe the data

at $Q^2 \approx 150$ GeV^2 for $W = 150$ GeV, see figure 6.10. In this context it is interesting to note, that a model with Q^2-dependent dipole cross section starts to lack justification at Q^2 not much higher than this value. This puts the inability of Q^2-independent dipole cross sections to describe the data at high Q^2 into perspective again, since the dipole picture itself might generally lack justification at Q^2 values in a similar region. Although a breakdown of the dipole picture is generally expected at high Q^2, our analysis helps to quantify the scale at which this should be expected.

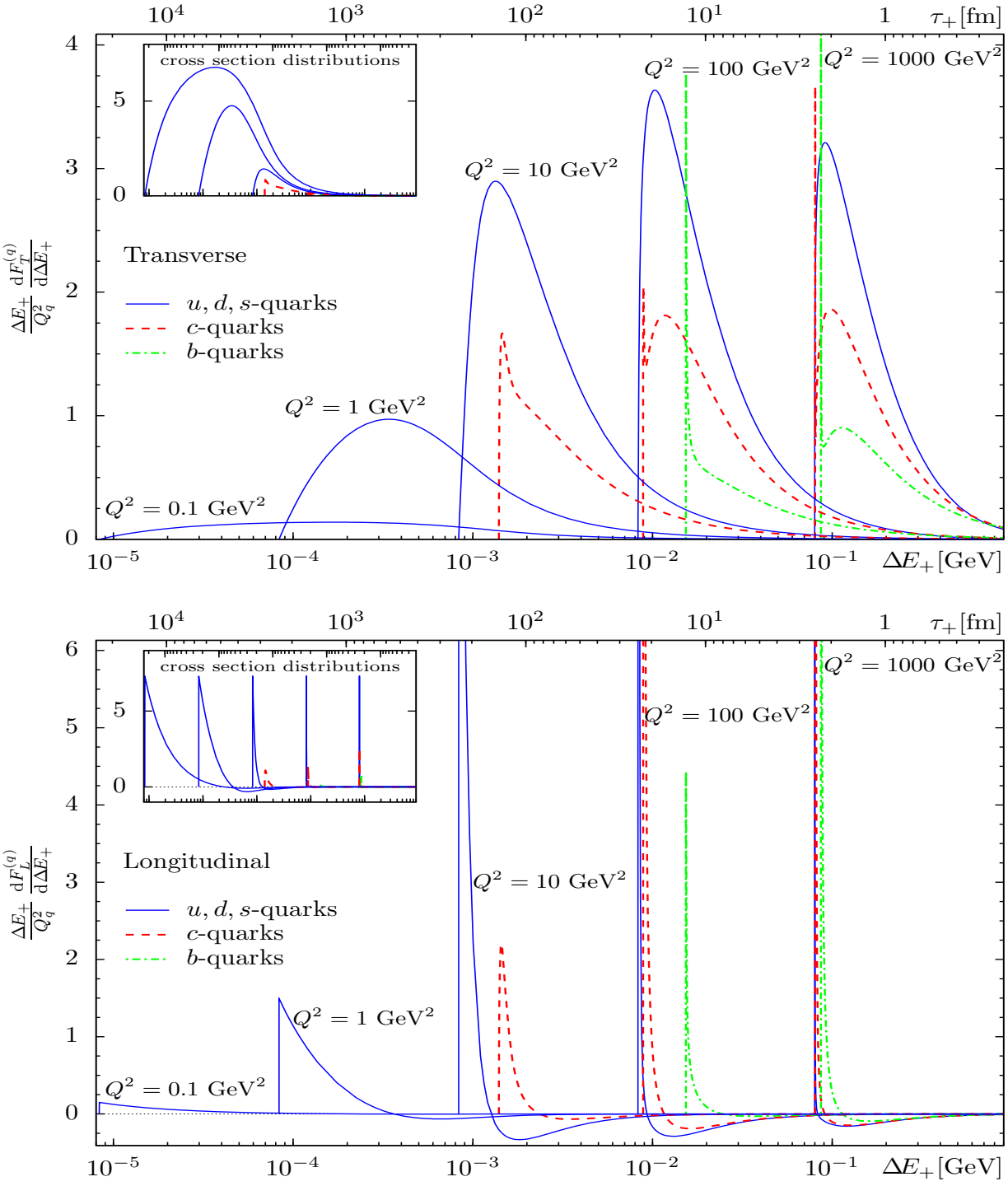

Figure 8.4: F_2 contributions in dependence of the joint dipole energy mismatch ΔE_+ respectively Ioffe time τ_+ for the GBW model (large plots). Upper graph for transverse, lower graph for longitudinal γ^* polarisation. Curves are for different values of Q^2 and quark flavours, $W = 150$ GeV is kept fixed. The small plots show correspondingly $(\Delta E_+ \mathrm{d}\sigma^{(q)}/\mathrm{d}\Delta E_+)/(\alpha_{\mathrm{em}} Q_q^2)$ [mb].

Part IV

Conclusions

In this thesis, we employed gauge invariant functions to analyse extended Higgs potentials in a transparent geometrical picture. In the case of two Higgs doublets these gauge invariant functions form a Minkowski-type four-vector, which has to lie on or inside the forward light cone. For the general Two-Higgs-Doublet Model (THDM), we confirmed and extended previous results by giving concise criteria for the stability and electroweak symmetry breaking properties of its tree-level potential. These criteria cover in particular the case of multiple minima, show how to find the global minimum of the potential and constrain the parameters of a phenomenologically acceptable theory. We discussed generalised CP transformations, which are space-like reflections at the level of the gauge invariant four-vectors. Reflections on a plane were shown to correspond to the standard CP transformation up to a basis change. We derive basis independent necessary and sufficient conditions for the potential and the vacuum to be invariant under such a type of CP transformation. We find that the point-reflection at the origin represents a new basis independent type of CP transformation. A CP symmetry of this type is necessarily spontaneously broken for a phenomenlogically acceptable Higgs potential. Invariance of non-vanishing Yukawa terms under the new CP symmetry requires at least two fermion families. Starting from this new symmetry, a set of CP invariance conditions was prescribed, which produces large mass hierachies and absence of large flavour changing neutral currents for the considered case of two fermion families. The gauge invariant function approach may be combined with algebraic Gröbner basis techniques, which allows for a fully automated unambigous determination of the global minimum of a tree-level Higgs potential. We applied this method to the Next-to-Minimal Supersymmetric Standard Model (NMSSM), where we find that the study of the minima structure is indeed necessary to avoid a destabilisation of the required vacuum. Requiring the vacuum to be not only a local but a global minimum of the potential excludes further regions in the NMSSM parameter space.

In a second part of this thesis, we presented bounds for the colour dipole picture, which help to further clarify its range of application. We derived general bounds on the ratio of the charm and the longitudinal part of the structure function F_2, which apply for any choice of the dipole cross section. Future measurements of these ratios might exclude kinematical ranges, where this bound is violated. We derived bounds on ratios of F_2 at different photon virtualities Q^2, assuming the energy dependence of the dipole cross section to be given only by the center-of-mass energy. Confronting these bounds with a fit to the data, we find that any such dipole model necessarily fails to describe the data for Q^2 above 100-200 GeV^2, depending on the center-of-mass energy. Furthermore, we explicitly considered different choices of energy variables for the dipole cross section and find, that they may not be related through effective scale arguments in a straight-forward way. Finally, we calculated Ioffe times, that is, dipole lifetimes, for a popular dipole model. This serves as a consistency check, since the Ioffe times are supposed to be large with respect to hadronical sizes in order to justify a separate treatment of dipole production and scattering on the proton. We find this justification to fail at large Q^2.

At the time of this writing, final preparations are taken at the Large Hadron Collider (LHC) and its experiments to start operation. Whatever the origin of electroweak symmetry breaking may be, new phenomena are expected below 1-2 TeV, and thus the LHC

may shed light on this long outstanding question of physics within the nearer future.

Part V

Appendix

Appendix A

Gröbner Bases

In this appendix we want to sketch the construction of the *Buchberger algorithm* which transforms a given set of polynomials F into a *Gröbner basis* G. The Gröbner basis G has exactly the same simultaneous zeros as the initial set of polynomials F, but allows better access to the actual calculation of these zeros. The general idea is to *complete* the set F by adjoining differences of polynomials. Before we can present the algorithms themselves we have to introduce the two basic ingredients, that is *Reduction* and the *S-polynomial*. For a more detailed discussion we refer the reader to the literature [125, 174, 175, 130, 131]. Here we repeat the summary given in [114], which follows closely [175]. First of all we recall some definitions.

Definition A.1. *Polynomial Ring*
A Polynomial Ring $K[x_1, \ldots, x_n] \equiv K[\mathbf{x}]$ *is the set of all n-variate polynomials with variables* $x_1, \ldots, x_n$ *and coefficients in the field* K*.*

Definition A.2. *Generated Ideal*
Let $F = \{f_1, \ldots, f_n\} \subset K[\mathbf{x}]$ *be finite,* F *generates an ideal defined by*
$$I(F) \equiv \left\{ \sum_{f_i \in F} r_i \cdot f_i \;\middle|\; r_i \in K[\mathbf{x}],\ f_i \in F,\ i = 1, \ldots, n \right\}.$$

In the following we want to consider an explicit example, that is a set $F = \{f_1, f_2, f_3\} \subset \mathbb{Q}[x, y]$ of polynomials with rational coefficients:

$$\begin{aligned} f_1 &= 3x^2y + 2xy + y + 9x^2 + 5x - 3, \\ f_2 &= 2x^3y - xy - y + 6x^3 - 2x^2 - 3x + 3, \\ f_3 &= x^3y + x^2y + 3x^3 + 2x^2. \end{aligned} \tag{A.1}$$

The set F generates an ideal $I(F)$, which is given by the set of sums of f_1, f_2, and f_3, where each polynomial is multiplied with another arbitrary polynomial from the ring $\mathbb{Q}[x, y]$. The summands of the polynomial are denoted as *monomials* and each monomial is the product of a coefficient and a *power product*.

Further, we introduce an *ordering* ($\succ$) of the monomials. In the *lexicographical ordering* ($\succ_{\text{lex}}$) the monomials are ordered with respect to the power of each variable subsequently. The ring notation $\mathbb{Q}[x, y]$ defines $y \succ_{\text{lex}} x$, that is for the lexicographical ordering of monomials powers of y are considered first, then powers of x. Explicitly, this means $2x^2y^3 \succ_{\text{lex}} 5xy^2$ because the power of y is larger in the first monomial and $2xy^2 \succ_{\text{lex}} 5y^2$, because both monomials have the same power of y, but the first monomial has a larger power of x. The monomials of the polynomials (A.1) from the ring $\mathbb{Q}[x, y]$ are ordered with respect to lexicographical ordering. In *total degree ordering* ($\succ_{\text{deg}}$) the monomials are ordered with respect to the sum of powers in each monomial. If two monomials have the same sum of powers, they are ordered with respect to another ordering, for instance lexicographical. For polynomials in $\mathbb{Q}[x, y]$ we have $x^2y \succ_{\text{deg}} 4xy$ since the sum of powers of the left power product is 3 compared to 2 for the right power product.

The largest power product with respect to the underlying ordering ($\succ$) of a polynomial f is denoted as the *leading power product*, $\text{LP}(f)$, the corresponding coefficient as *leading coefficient*, $\text{LC}(f)$. With help of these preparations we can define the two essential parts of the Buchberger algorithm, that is *Reduction* and the *S-polynomial*.

Definition A.3. *Reduction*
Let $f, p \in K[\mathbf{x}]$. We call f reducible modulo p, if for a power product t of f there exists a power product u with $\text{LP}(p) \cdot u = t$. Then we say, f reduces to h modulo p, where $h = f - \frac{Coefficient(f,t)}{\text{LC}(p)} \cdot u \cdot p$.

In the example (A.1) the polynomial f_3 is reducible modulo f_1, since for example the second monomial of f_3, that is x^2y, is a multiple of the $\text{LP}(f_1)$, and $h = f_3 - 1/3f_1 = x^3y - 2/3xy - 1/3y + 3x^3 - x^2 - 5/3x + 1$.

Reduction of a polynomial modulo a set $P \subset K[\mathbf{x}]$ is accordingly defined if there is a $p \in P$ such that f is reducible modulo p. Further, we say, a polynomial h is in *reduced form* or *normal form* modulo F, in short $\text{normf}(h, F)$, if there is no h' such that h reduces to h' modulo F. A set $P \subset K[\mathbf{x}]$ is called reduced, if each $p \in P$ is in reduced form modulo $P \backslash \{p\}$. Note that reduction is defined with respect to the underlying ordering of the monomials, since the leading power product is defined with respect to the ordering. In general, a normal form is not unique, neither for a polynomial nor for a set.

Now we can present an algorithm, to compute a normal form $Q \subset K[\mathbf{x}]$ of a finite $F \subset K[\mathbf{x}]$.

Algorithm A.4. *Normal form*
For a given finite set $F \subset K[\mathbf{x}]$ determine a normal form $Q \subset K[\mathbf{x}]$.

$Q := F$
while *exists* $p \in Q$ *which is reducible modulo* $Q \backslash \{p\}$ **do**
    $Q := Q \backslash \{p\}$
    $h := \mathrm{normf}(p, Q)$
    if $h \neq 0$ **then**
        $Q := Q \cup \{h\}$
return Q

Clearly, the simultaneous zeros of all $f_i \in F$ are also simultaneous zeros of all $q_i \in Q$ and vice versa.

Definition A.5. *S-polynomial*
For $g_1, g_2 \in K[\mathbf{x}]$ *the S-polynomial of* g_1 *and* g_2 *is defined as*

$$\mathrm{spol}(g_1, g_2) \equiv \frac{\mathrm{lcm}\left(\mathrm{LP}(g_1), \mathrm{LP}(g_2)\right)}{\mathrm{LP}(g_1)} g_1 - \frac{\mathrm{LC}(g_1)}{\mathrm{LC}(g_2)} \frac{\mathrm{lcm}\left(\mathrm{LP}(g_1), \mathrm{LP}(g_2)\right)}{\mathrm{LP}(g_2)} g_2,$$

where lcm *denotes the least common multiple.*

Clearly, a simultaneous zero of g_1 and g_2 is also a zero of $\mathrm{spol}(g_1, g_2)$. In the example (A.1) we can build the S-polynomial for any two polynomials, for instance

$$\begin{aligned} \mathrm{spol}(f_1, f_2) &= \frac{x^3 y}{x^2 y}\, f_1 - \frac{3}{2}\, \frac{x^3 y}{x^3 y}\, f_2 = x\, f_1 - 3/2\, f_2 \\ &= 2x^2 y + 5/2xy + 3/2y + 8x^2 + 3/2x - 9/2. \end{aligned} \tag{A.2}$$

Finally we define the Gröbner basis.

Definition A.6. *Gröbner basis*
$G \subset K[\mathbf{x}]$ *is called Gröbner Basis, if for all* $f_1, f_2 \in G$ $\mathrm{normf}(\mathrm{spol}(f_1, f_2), G) = 0$.

Now everything is prepared to present the Buchberger algorithm.

Algorithm A.7. *Buchberger*
For a given finite set $F \subset K[\mathbf{x}]$ *determine the Gröbner basis* $G \subset K[\mathbf{x}]$ *with* $I(F) = I(G)$.

$G := F$
$B := \{\{g_1, g_2\} | g_1, g_2 \in G \text{ with } g_1 \neq g_2\}$
while $B \neq \emptyset$ **do**
  choose $\{g_1, g_2\}$ *from* B
  $B := B \backslash \{\{g_1, g_2\}\}$
  $h := \mathrm{spol}(g_1, g_2)$
  $h' := \mathrm{normf}(h, G)$
  if $h' \neq 0$ **then**
    $B := B \cup \{\{g, h'\} | g \in G\}$
    $G := G \cup \{h'\}$
return G

Note, that since G just follows by adjoining reduced S-polynomials to F both sets generate the same ideal, especially, both sets have exactly the same simultaneous zeros. It can be proven, that the Buchberger algorithm terminates. The final step is to construct the reduced Gröbner basis by applying the normal form algorithm defined above to the Gröbner basis G. It can be shown that the reduced Gröbner basis is unique [174]. If we apply the Buchberger algorithm to the set (A.1) with subsequent reduction we end up with the reduced Gröbner basis (with underlying lexicographical ordering)

$$\begin{aligned} g_1 &= y + x^2 - 3/2x - 3, \\ g_2 &= x^3 - 5/2x^2 - 5/2x. \end{aligned} \tag{A.3}$$

The system of equations $g_1 = g_2 = 0$ is equivalent to $f_1 = f_2 = f_3 = 0$, but the former allows to directly calculate the solutions: Since $g_2 = 0$ is univariate it can be solved immediately and subsequently $g_1 = 0$ for each partial solution inserted.

Despite the correctness of the Buchberger algorithm tractability of practical examples requires to improve this algorithm. In particular, the number of iterations in the algorithm drastically grows with an increasing number of polynomials and with higher degrees of the polynomials. In this respect much progress has been made with the improvement of this original Buchberger algorithm from 1965 (see [174, 175, 126]).

The scope of Gröbner bases is not limited to system solving. In this thesis it is also used to unambiguously check equivalences of polynomial systems of equations via *radical membership* tests. For this and many more topics we refer the interested reader to the recent text books [130, 131].

Appendix B

Convex hulls, convex cones and moment problems

In this appendix we discuss the notions of convex hull and convex cone as well as some further mathematical relations, as given in [152]. The precise mathematical definitions can be found in [176], for our notation see also [177].

Let us consider the n-dimensional Euclidean space $\mathbb{R}^n$ with elements $\boldsymbol{x}$, $\boldsymbol{y}$ etc. A nonempty subset X of $\mathbb{R}^n$ is called a *convex set* if for any elements $\boldsymbol{x}$, $\boldsymbol{y}$ in X and any real number a with $0 \leq a \leq 1$ the element $a\boldsymbol{x} + (1-a)\boldsymbol{y}$ is also contained in X. That is, with any two points of X the complete straight line connecting them is also in X.

Let now Y be an arbitrary nonempty subset of $\mathbb{R}^n$. The minimum convex set containing Y exists [176] and is called the *convex hull* of Y and denoted by $\mathrm{co}(Y)$. Its closure is denoted by $\overline{\mathrm{co}}(Y)$. To illustrate this concept we give a physical example. Let $Y = \{\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(N)}\}$ be a set of N points in $\mathbb{R}^n$. Consider arbitrary distributions of masses $m_i \geq 0$ $(i = 1, \ldots, N)$ on these points. The center-of-mass is then

$$\boldsymbol{x} = \frac{\sum_{i=1}^{N} m_i \boldsymbol{y}^{(i)}}{\sum_{i=1}^{N} m_i} \,. \tag{B.1}$$

The convex hull of Y, $\mathrm{co}(Y)$, is the set of all possible center-of-mass points of such mass distributions.

Next we discuss the notion of *convex cone*. A nonempty subset X of $\mathbb{R}^n$ is called a convex cone if for any elements $\boldsymbol{x}$, $\boldsymbol{y}$ of X and any real number $a \geq 0$ the elements $a\boldsymbol{x}$ and $\boldsymbol{x} + \boldsymbol{y}$ are also contained in X. Let Y be an arbitrary non-empty subset of $\mathbb{R}^n$, then the minimal convex cone containing Y exists and is denoted by $\mathrm{K}(Y)$. Its closure is denoted by $\overline{\mathrm{K}}(Y)$.

We illustrate these notions with a two-dimensional example. Let Y consist of three points in $\mathbb{R}^2$

$$Y = \left\{\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)}, \boldsymbol{y}^{(3)}\right\} \tag{B.2}$$

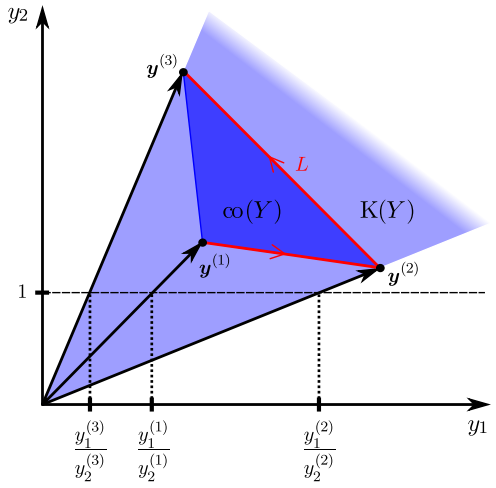

Figure B.1: Illustration of the convex hull $\mathrm{co}(Y)$ and the convex cone $\mathrm{K}(Y)$ for the set Y of (B.2). Here both sets are closed, that is $\overline{\mathrm{co}}(Y) = \mathrm{co}(Y)$ and $\overline{\mathrm{K}}(Y) = \mathrm{K}(Y)$.

as shown in figure B.1. Here

$$\boldsymbol{y}^{(i)} = \begin{pmatrix} y_1^{(i)} \\ y_2^{(i)} \end{pmatrix} \tag{B.3}$$

and we suppose

$$y_2^{(i)} > 0 \qquad \text{for } i = 1, 2, 3\,. \tag{B.4}$$

The convex hull of Y, $\mathrm{co}(Y)$, is given by the dark grey triangle bounded by the polygon from $\boldsymbol{y}^{(1)}$ to $\boldsymbol{y}^{(2)}$, $\boldsymbol{y}^{(3)}$ and back to $\boldsymbol{y}^{(1)}$. The cone $\mathrm{K}(Y)$ is indicated by the light grey area bounded by the rays $\lambda\,\boldsymbol{y}^{(2)}$ with $\lambda \geq 0$ and $\mu\,\boldsymbol{y}^{(3)}$ with $\mu \geq 0$.

Let Y be a non-empty subset of $\mathbb{R}^n$ and let $\mathrm{co}(Y)$ be the convex hull of Y and $\mathrm{K}(Y)$ the minimal convex cone containing Y. We define the set

$$\mathrm{K}'(Y) = \{\, \boldsymbol{x} \mid\ \boldsymbol{x} = \lambda \boldsymbol{y},\ \lambda \geq 0,\ \boldsymbol{y} \in \mathrm{co}(Y) \,\} \tag{B.5}$$

and assert that

$$\mathrm{K}'(Y) = \mathrm{K}(Y)\,. \tag{B.6}$$

The proof of (B.6) goes as follows. It is easy to see that $\mathrm{K}'(Y)$ is a convex cone containing Y. Thus, since $\mathrm{K}(Y)$ is the minimal such cone we have

$$\mathrm{K}'(Y) \supset \mathrm{K}(Y)\,. \tag{B.7}$$

On the other hand, $\mathrm{K}(Y)$ is a convex set containing Y and $\mathrm{co}(Y)$ is the minimal such set. Thus

$$\mathrm{K}(Y) \supset \mathrm{co}(Y)\,. \tag{B.8}$$

Since $\mathrm{K}(Y)$ is a convex cone this implies that for any element $\boldsymbol{x} \in \mathrm{co}(Y)$ and any $\lambda \geq 0$ we have $\lambda \boldsymbol{x} \in \mathrm{K}(Y)$. That is, we have

$$\mathrm{K}'(Y) \subset \mathrm{K}(Y)\,. \tag{B.9}$$

Therefore, we have shown that $\mathrm{K}'(Y) = \mathrm{K}(Y)$. That is, every element of $\mathrm{K}(Y)$ can be written in the form given in (B.5). For the closures we find in a similar way

$$\overline{\mathrm{K}}(Y) = \{\, \boldsymbol{x} \mid \ \boldsymbol{x} = \lambda \boldsymbol{y},\ \lambda \geq 0,\ \boldsymbol{y} \in \overline{\mathrm{co}}(Y) \,\} \,. \tag{B.10}$$

Next we come to the moment problem which is at the heart of our derivations of bounds. Suppose we have a continuous vector function on a closed interval $[t_0, t_1] \subset \mathbb{R}$ defining a curve L in $\mathbb{R}^n$:

$$\begin{aligned} L: \quad & [t_0, t_1] \to \mathbb{R}^n \,, \\ & t \mapsto \boldsymbol{y}(t) \,. \end{aligned} \tag{B.11}$$

We also suppose that there is at least one constant vector $\boldsymbol{a}$ such that

$$\boldsymbol{a}^{\mathrm{T}} \boldsymbol{y}(t) > 0 \qquad \text{for all } t \in [t_0, t_1] \,. \tag{B.12}$$

We are interested in the set $\tilde{K}$ of all points $\boldsymbol{x}$ of $\mathbb{R}^n$ which can be represented as

$$\boldsymbol{x} = \int_{t_0}^{t_1} \boldsymbol{y}(t) \, \mathrm{d}\Sigma(t) \tag{B.13}$$

where $\Sigma(t)$ is some non-decreasing function on $[t_0, t_1]$. Note that such a function is bounded from below and above since

$$\Sigma(t_0) \leq \Sigma(t) \leq \Sigma(t_1) \,. \tag{B.14}$$

Before we discuss the solution of this problem as given in [178] we note that in (B.13) we are dealing with so-called Stieltjes integrals, see for example [179, 180]. The reader not familiar with these integrals may always set

$$\mathrm{d}\Sigma(t) = \sigma(t) \, \mathrm{d}t \tag{B.15}$$

where $\sigma(t)$ is some non-negative distribution. That is, $\sigma(t)$ can be an ordinary non-negative function but can also contain non-negative δ-distributions.

The solution of the problem posed above is as follows, see [178]. The set $\tilde{K}$ of points which can be represented in the form (B.13) is given by $\overline{\mathrm{K}}(L)$, that is, by the smallest closed convex cone containing the curve L:

$$\tilde{K} = \overline{\mathrm{K}}(L) \,. \tag{B.16}$$

Consider for illustration the two-dimensional example as in figure B.1 and the following curve L defined for $t \in [0, 1]$,

$$L: \ t \mapsto \boldsymbol{y}(t) = \begin{cases} \boldsymbol{y}^{(1)} + 2t(\boldsymbol{y}^{(2)} - \boldsymbol{y}^{(1)}) & : 0 \leq t \leq \frac{1}{2} \,, \\ \boldsymbol{y}^{(2)} + (2t-1)(\boldsymbol{y}^{(3)} - \boldsymbol{y}^{(2)}) & : \frac{1}{2} < t \leq 1 \,. \end{cases} \tag{B.17}$$

We are interested in the points $\boldsymbol{x}$ allowing a representation

$$\boldsymbol{x} = \int_0^1 \boldsymbol{y}(t)\,\mathrm{d}\Sigma(t) \tag{B.18}$$

with some non-decreasing function $\Sigma(t)$. According to the theorem quoted above $\boldsymbol{x}$ has to be in the closed convex cone $\overline{\mathrm{K}}(L)$ as shown in figure B.1. We ask now for the allowed range for x_1 given some x_2. The possible x_1 values are obtained by cutting the cone $\overline{\mathrm{K}}(L)$ at $y_2 = x_2 = \textit{const.}$ and reading off the corresponding y_1 values. Similarly, the allowed range of the ratio x_1/x_2 is obtained if we choose $y_2 = 1$ for cutting the cone. Clearly, to get the extremal values of x_1/x_2 we just have to consider the generating rays $\lambda\boldsymbol{y}^{(i)}$, $\lambda \geq 0$, for $i = 1, 2, 3$. Cutting them at $y_2 = 1$ gives $y_1^{(i)}/y_2^{(i)}$, $i = 1, 2, 3$. Among these ratios there are the extremal points of x_1/x_2. In our example we get

$$\frac{y_1^{(3)}}{y_2^{(3)}} \leq \frac{x_1}{x_2} \leq \frac{y_1^{(2)}}{y_2^{(2)}}\,. \tag{B.19}$$

Note that the interval $[y_1^{(3)}/y_2^{(3)}\,,\ y_1^{(2)}/y_2^{(2)}]$ is the convex hull of the set $\{y_1^{(i)}/y_2^{(i)} \mid i = 1, 2, 3\}$.

For the general case, in $\mathbb{R}^n$, the situation is completely analogous. Consider (B.13) in $\mathbb{R}^n$ ($n \geq 2$) and let us write in components

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \int_{t_0}^{t_1} \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} \mathrm{d}\Sigma(t)\,. \tag{B.20}$$

Let us suppose that

$$0 < c_0 \leq y_n(t) \leq c_1 \qquad \text{for } t_0 \leq t \leq t_1\,. \tag{B.21}$$

Thereby (B.12) is satisfied with $\boldsymbol{a}^{\mathrm{T}} = (0, \ldots, 0, 1)$. To get the bounds for the ratio vector

$$\boldsymbol{x}' = \begin{pmatrix} x_1/x_n \\ \vdots \\ x_{n-1}/x_n \\ 1 \end{pmatrix} \tag{B.22}$$

we just have to cut the cone $\overline{\mathrm{K}}(L)$ with the hyperplane $x_n = 1$. The corresponding set in $\mathbb{R}^n$ obtained by this cutting is given by the closed convex hull of the ratio vectors of the curve L generating the cone $\overline{\mathrm{K}}(L)$. That is, we denote by L' the following curve in $\mathbb{R}^n$

$$L': \quad t \mapsto \begin{pmatrix} y_1(t)/y_n(t) \\ \vdots \\ y_{n-1}(t)/y_n(t) \\ 1 \end{pmatrix}, \qquad t \in [t_0, t_1]. \tag{B.23}$$

Let $\overline{\mathrm{co}}(L')$ be the closed convex hull of L'. The intersetion of $\overline{\mathrm{K}}(L)$ with the hyperplane $x_n = 1$ is given by $\overline{\mathrm{co}}(L')$. Clearly, the extremal points of the above cone–hyperplane intersection must be given by the intersections of the rays generating the cone $\overline{\mathrm{K}}(L)$, that is by the rays through the curve L. But this gives just L'.

We give now a formal proof of the above statements. For this consider the minimal closed convex cone $\overline{\mathrm{K}}(L)$ containing L and analogously $\overline{\mathrm{K}}(L')$ containing L'. We assert that

$$\overline{\mathrm{K}}(L) = \overline{\mathrm{K}}(L')\,. \tag{B.24}$$

To prove (B.24) we note that according to (B.13) and (B.23) all vectors $\boldsymbol{x}' \in \overline{\mathrm{K}}(L')$ are of the form

$$\boldsymbol{x}' = \int_{t_0}^{t_1} \frac{\boldsymbol{y}(t)}{y_n(t)}\,\mathrm{d}\Sigma'(t) \tag{B.25}$$

with some non-decreasing function $\Sigma'(t)$. Due to (B.21) the division by $y_n(t)$ in (B.25) is harmless and we can define a non-decreasing function $\Sigma(t)$ on $[t_0, t_1]$ by

$$\Sigma(t) = \int_{t_0}^{t} \frac{1}{y_n(t')}\,\mathrm{d}\Sigma'(t')\,. \tag{B.26}$$

We get then for $\boldsymbol{x}'$ of (B.25)

$$\boldsymbol{x}' = \int_{t_0}^{t_1} \boldsymbol{y}(t)\,\mathrm{d}\Sigma(t)\,. \tag{B.27}$$

That is, $\boldsymbol{x}' \in \overline{\mathrm{K}}(L)$ according to (B.13) and we have shown

$$\overline{\mathrm{K}}(L') \subset \overline{\mathrm{K}}(L)\,. \tag{B.28}$$

Now we consider an arbitrary element $\boldsymbol{x} \in \overline{\mathrm{K}}(L)$ which according to (B.13) has the form

$$\boldsymbol{x} = \int_{t_0}^{t_1} \boldsymbol{y}(t)\,\mathrm{d}\Sigma(t) \tag{B.29}$$

with some non-decreasing function $\Sigma(t)$. We define a non-decreasing function $\Sigma'(t)$ on $[t_0, t_1]$ by

$$\Sigma'(t) = \int_{t_0}^{t} y_n(t')\,\mathrm{d}\Sigma(t')\,. \tag{B.30}$$

Again, we use here (B.21). We get then

$$\boldsymbol{x} = \int_{t_0}^{t_1} \boldsymbol{y}(t)\frac{1}{y_n(t)}\,\mathrm{d}\Sigma'(t)\,. \tag{B.31}$$

That is, $\boldsymbol{x} \in \overline{\mathrm{K}}(L')$ and therefore

$$\overline{\mathrm{K}}(L) \subset \overline{\mathrm{K}}(L')\,. \tag{B.32}$$

From (B.28) and (B.32) follows (B.24), *q.e.d.*

From (B.31) we can now draw the following conclusion for any non-zero element $\boldsymbol{x} \in \overline{\mathrm{K}}(L)$. Such an $\boldsymbol{x}$ is of the form (B.29) with $\Sigma(t) \neq const.$ We have then with $\Sigma'(t)$ from (B.30),

$$x_n = \Sigma'(t_1) = \int_{t_0}^{t_1} \mathrm{d}\Sigma'(t) > 0 \tag{B.33}$$

where we use (B.21). From (B.31) we can represent $\boldsymbol{x}$ as

$$\boldsymbol{x} = x_n \, \boldsymbol{x}' \tag{B.34}$$

where

$$\boldsymbol{x}' = \int_{t_0}^{t_1} \begin{pmatrix} y_1(t)/y_n(t) \\ \vdots \\ y_{n-1}(t)/y_n(t) \\ 1 \end{pmatrix} \mathrm{d}\Sigma''(t)\,,$$

$$\mathrm{d}\Sigma''(t) = \frac{1}{x_n}\mathrm{d}\Sigma'(t)\,, \qquad \int_{t_0}^{t_1} \mathrm{d}\Sigma''(t) = 1\,. \tag{B.35}$$

Clearly $\boldsymbol{x}'$ is in the intersection of $\overline{\mathrm{K}}(L')$ with the hyperplane $x'_n = 1$. Since $\overline{\mathrm{K}}(L')$ is the minimal closed convex cone containing L' this intersection is the minimal closed convex set containing L', that is, the closed convex hull $\overline{\mathrm{co}}(L')$:

$$\boldsymbol{x}' \in \overline{\mathrm{co}}(L')\,. \tag{B.36}$$

Thus we have shown that every non-zero vector $\boldsymbol{x} \in \overline{\mathrm{K}}(L)$, that is of the form (B.29), can be represented as $x_n \, \boldsymbol{x}'$ where $\boldsymbol{x}' \in \overline{\mathrm{co}}(L')$.

Now we come to the application of the above mathematical theorems to our problems. Consider three structure functions as in (6.11) but – for simplicity – only for fixed massless flavour q

$$\begin{pmatrix} F_2^{(q)}(W, Q_1^2) \\ F_2^{(q)}(W, Q_2^2) \\ F_2^{(q)}(W, Q_3^2) \end{pmatrix} = \int_0^\infty \mathrm{d}r\, r \frac{\hat{\sigma}^{(q)}(r, W)}{2\pi\alpha_{\mathrm{em}}} \begin{pmatrix} f^{(q)}(r, Q_1^2) \\ f^{(q)}(r, Q_2^2) \\ f^{(q)}(r, Q_3^2) \end{pmatrix}. \tag{B.37}$$

To bring this into the form of the problem (B.11), (B.13), we change variables and set

$$r = r_0 \frac{t}{1-t}\,, \qquad r_0 = 1 \text{ fm}\,, \qquad 0 < t < 1\,. \tag{B.38}$$

Furthermore, we shall split off from $f^{(q)}(r, Q^2)$ the asymptotic terms for $r \to 0$ and $r \to \infty$ derived in subsection 5.3. In the present case of a massless flavour q, we see from (5.62),(5.63),(5.66) and (5.67) that $f^{(q)}(r, Q^2)$ increases for $r \to 0$ as $1/r^2$. and decreases for $r \to \infty$ as $1/r^4$. Thus, we define a function

$$g(r) = \frac{r_0^4}{r^2(r_0 + r)^2} \tag{B.39}$$

which is independent of Q^2. This allows us to define the function

$$\hat{f}^{(q)}(t,Q^2) = \begin{cases} \lim_{t'\to 0} \left.\frac{f^{(q)}(r,Q^2)}{g(r)}\right|_{r=r_0\frac{t'}{1-t'}} & \text{if } t=0\,, \\ \left.\frac{f^{(q)}(r,Q^2)}{g(r)}\right|_{r=r_0\frac{t}{1-t}} & \text{if } 0<t<1\,, \\ \lim_{t'\to 1} \left.\frac{f^{(q)}(r,Q^2)}{g(r)}\right|_{r=r_0\frac{t'}{1-t'}} & \text{if } t=1 \end{cases} \tag{B.40}$$

for all t in the closed interval $[0,1]$, since the limites in (B.40) exist. It is easy to show that $\hat{f}^{(q)}(t,Q^2)$ is continuous as function of t. Moreover, we find

$$0 < c_0(Q^2) \le \hat{f}^{(q)}(t,Q^2) \le c_1(Q^2) \tag{B.41}$$

for all $t \in [0,1]$. Here $c_j(Q^2)$ $(j=0,1)$ are fixed positive constants for fixed Q^2. Next we note that the dipole model makes only sense if the dipole cross section $\hat{\sigma}^{(q)}(r,W)$ can be integrated with $g(r)$, that is, if

$$\int_0^\infty \mathrm{d}r\, r\, g(r)\, \frac{\hat{\sigma}^{(q)}(r,W)}{2\pi\alpha_{\mathrm{em}}} \;<\; \infty\,. \tag{B.42}$$

Further, we assume for $0<r<\infty$:

$$\hat{\sigma}^{(q)}(r,W) \ge 0\,. \tag{B.43}$$

This allows us to define a function $\Sigma^{(q)}(t,W)$ which is non-decreasing in t for fixed W:

$$\Sigma^{(q)}(t,W) = \int_0^t \mathrm{d}t' \left[\frac{\mathrm{d}r'}{\mathrm{d}t'} r'\, g(r') \frac{\hat{\sigma}^{(q)}(r',W)}{2\pi\alpha_{\mathrm{em}}}\right]_{r'=r_0\frac{t'}{1-t'}} \tag{B.44}$$

for $0 \le t \le 1$. Conversely, every non-decreasing function $\Sigma^{(q)}(t,W)$ gives, via (B.44), an acceptable dipole cross section $\hat{\sigma}^{(q)}(r,W)$. Furthermore, we define the curves L and L' as follows

$$L: \quad t \mapsto \boldsymbol{y}(t) := \begin{pmatrix} \hat{f}^{(q)}(t,Q_1^2) \\ \hat{f}^{(q)}(t,Q_2^2) \\ \hat{f}^{(q)}(t,Q_3^2) \end{pmatrix}, \qquad 0 \le t \le 1\,, \tag{B.45}$$

$$L': \quad t \mapsto \boldsymbol{y}'(t) := \begin{pmatrix} y_1(t)/y_3(t) \\ y_2(t)/y_3(t) \\ 1 \end{pmatrix}, \qquad 0 \le t \le 1\,. \tag{B.46}$$

Our original integrals (B.37) take now exactly the form of (B.13):

$$\boldsymbol{x} \equiv \begin{pmatrix} F_2^{(q)}(W,Q_1^2) \\ F_2^{(q)}(W,Q_2^2) \\ F_2^{(q)}(W,Q_3^2) \end{pmatrix} = \int_0^1 \boldsymbol{y}(t)\, \mathrm{d}\Sigma^{(q)}(t,W)\,. \tag{B.47}$$

The vector function $\boldsymbol{y}(t)$ is continuous for $t \in [0,1]$ and (B.12) is satisfied with $\boldsymbol{a}^{\mathrm{T}} = (0,0,1)$ due to (B.41) which also guarantees (B.21). From (B.16) we conclude that $\boldsymbol{x}$ must be in the smallest closed convex cone containing L, that is, in $\overline{\mathrm{K}}(L)$. It is easy to see that this cone coincides with the cone defined as in (6.12) but for fixed flavour q. Indeed, we have from (B.10) that $\overline{\mathrm{K}}(L)$ can be represented as

$$\overline{\mathrm{K}}(L) = \{\, \boldsymbol{x} \mid \boldsymbol{x} = \lambda \boldsymbol{z},\ \lambda \geq 0,\ \boldsymbol{z} \in \overline{\mathrm{co}}(L) \,\}\,. \tag{B.48}$$

With a simple rescaling by $g(r)$ from (B.39) and using that the closure takes care of the limiting points $t = 0$ and $t = 1$ which correspond to $r \to 0$ and $r \to \infty$ we get for every $\boldsymbol{x} \in \overline{\mathrm{K}}(L)$ the representation

$$\begin{aligned} &\boldsymbol{x} = \mu\, \boldsymbol{v}, \\ &\text{with} \quad \mu \geq 0, \quad \boldsymbol{v} \in \overline{\mathrm{co}} \left\{ \left. \begin{pmatrix} f^{(q)}(r, Q_1^2) \\ f^{(q)}(r, Q_2^2) \\ f^{(q)}(r, Q_3^2) \end{pmatrix} \right| r \in \mathbb{R}^+ \right\}. \end{aligned} \tag{B.49}$$

The extension of these arguments to more than one flavour is straightforward. With this we have given a rigorous proof of (6.12).

Consider next the ratio vector

$$\boldsymbol{x}' = \frac{1}{x_3}\boldsymbol{x} = \begin{pmatrix} F_2^{(q)}(W, Q_1^2)/F_2^{(q)}(W, Q_3^2) \\ F_2^{(q)}(W, Q_2^2)/F_2^{(q)}(W, Q_3^2) \\ 1 \end{pmatrix}. \tag{B.50}$$

From (B.29), (B.34) and (B.36) we see that $\boldsymbol{x}'$ must be in the closed convex hull $\overline{\mathrm{co}}(L')$:

$$\boldsymbol{x}' \in \overline{\mathrm{co}}(L')\,. \tag{B.51}$$

We have from (B.46), (B.45) and (B.40)

$$\begin{aligned} \overline{\mathrm{co}}(L') &= \overline{\mathrm{co}} \left\{ \left. \begin{pmatrix} \hat{f}^{(q)}(t, Q_1^2)/\hat{f}^{(q)}(t, Q_3^2) \\ \hat{f}^{(q)}(t, Q_2^2)/\hat{f}^{(q)}(t, Q_3^2) \\ 1 \end{pmatrix} \right| 0 \leq t \leq 1 \right\} \\ &= \overline{\mathrm{co}} \left\{ \left. \begin{pmatrix} f^{(q)}(r, Q_1^2)/f^{(q)}(r, Q_3^2) \\ f^{(q)}(r, Q_2^2)/f^{(q)}(r, Q_3^2) \\ 1 \end{pmatrix} \right| 0 < r < \infty \right\}. \end{aligned} \tag{B.52}$$

Taking the closure eliminates differences which could otherwise exist between the two convex hulls of the sets in (B.52) originating from the fact that $\hat{f}^{(q)}(t, Q^2)$ is defined on a closed t interval whereas $f^{(q)}(r, Q^2)$ is defined on an open r interval. The straightforward extension of (B.51) and (B.52) to the case of several flavours q proves (6.13).

With this we have illustrated for one particular case how our bounds are derived in a mathematically rigorous way. For all other cases analogous arguments can be applied.

Bibliography

[1] S. L. Glashow, *"Partial symmetries of weak interactions"*, *Nucl. Phys.* **22** (1961) 579–588.

[2] S. Weinberg, *"A model of leptons"*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[3] A. Salam, *Elementary Particle Theory*, p. 367. N. Svartholm *ed.*, Almquist and Wiksell, Stockholm, 1968.

[4] S. L. Glashow, J. Iliopoulos, and L. Maiani, *"Weak interactions with lepton-hadron symmetry"*, *Phys. Rev.* **D2** (1970) 1285–1292.

[5] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *"Advantages of the color octet gluon picture"*, *Phys. Lett.* **B47** (1973) 365–368.

[6] G. 't Hooft and M. J. G. Veltman, *"Regularization and renormalization of gauge fields"*, *Nucl. Phys.* **B44** (1972) 189–213.

[7] Particle Data Group, W. M. Yao *et al.*, *"Review of particle physics"*, *J. Phys.* **G33** (2006) 1–1232. `http://pdg.lbl.gov`. and 2007 partial update for the 2008 edition available on the PDG WWW pages.

[8] ALEPH, DELPHI, L3, OPAL Collaborations and LEP Electroweak Working Group, J. Alcaraz *et al.*, *"A combination of preliminary electroweak measurements and constraints on the standard model"*, `arXiv:hep-ex/0612034`.

[9] P. W. Higgs, *"Broken symmetries, massless particles and gauge fields"*, *Phys. Lett.* **12** (1964) 132–133.

[10] P. W. Higgs, *"Spontaneous symmetry breakdown without massless bosons"*, *Phys. Rev.* **145** (1966) 1156–1163.

[11] F. Englert and R. Brout, *"Broken symmetry and the mass of gauge vector mesons"*, *Phys. Rev. Lett.* **13** (1964) 321–322.

[12] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *"Global conservation laws and massless particles"*, *Phys. Rev. Lett.* **13** (1964) 585–587.

[13] LEP Higgs Working Group, R. Barate *et al.*, *"Search for the standard model Higgs boson at LEP"*, *Phys. Lett.* **B565** (2003) 61–75, arXiv:hep-ex/0306033.

[14] G. L. Fogli, E. Lisi, A. Marrone, and A. Palazzo, *"Global analysis of three-flavor neutrino masses and mixings"*, *Prog. Part. Nucl. Phys.* **57** (2006) 742–795, arXiv:hep-ph/0506083.

[15] B. Kayser, *"Neutrino mass, mixing, and flavor change"*, arXiv:0804.1497 [hep-ph].

[16] M. Kamionkowski and A. Kinkhabwala, *"Galactic halo models and particle dark matter detection"*, *Phys. Rev.* **D57** (1998) 3256–3263, arXiv:hep-ph/9710337.

[17] G. Bertone, D. Hooper, and J. Silk, *"Particle dark matter: Evidence, candidates and constraints"*, *Phys. Rept.* **405** (2005) 279–390, arXiv:hep-ph/0404175.

[18] D. Clowe *et al.*, *"A direct empirical proof of the existence of dark matter"*, *Astrophys. J.* **648** (2006) L109–L113, arXiv:astro-ph/0608407.

[19] G. W. Angus, H. Shan, H. Zhao, and B. Famaey, *"On the law of gravity, the mass of neutrinos and the proof of dark matter"*, *Astrophys. J.* **654** (2007) L13–L16, arXiv:astro-ph/0609125.

[20] G. R. Blumenthal, S. M. Faber, J. R. Primack, and M. J. Rees, *"Formation of galaxies and large-scale structure with cold dark matter"*, *Nature* **311** (1984) 517–525.

[21] S. M. Faber and J. S. Gallagher, *"Masses and mass-to-light ratios of galaxies"*, *Ann. Rev. Astron. Astrophys.* **17** (1979) 135–183.

[22] SDSS Collaboration, M. Tegmark *et al.*, *"Cosmological constraints from the SDSS luminous red galaxies"*, *Phys. Rev.* **D74** (2006) 123507, arXiv:astro-ph/0608632.

[23] Muon G-2 Collaboration, G. W. Bennett *et al.*, *"Final report of the muon E821 anomalous magnetic moment measurement at BNL"*, *Phys. Rev.* **D73** (2006) 072003, arXiv:hep-ex/0602035.

[24] F. Jegerlehner, *"Essentials of the muon $g-2$"*, *Acta Phys. Polon.* **B38** (2007) 3021, arXiv:hep-ph/0703125.

[25] UTfit Collaboration, M. Bona *et al.*, *"First evidence of new physics in $b \leftrightarrow s$ transitions"*, arXiv:0803.0659 [hep-ph].

[26] N. G. Deshpande and E. Ma, *"Pattern of symmetry breaking with two Higgs doublets"*, *Phys. Rev.* **D18** (1978) 2574.

[27] H. E. Haber, G. L. Kane, and T. Sterling, *"The fermion mass scale and possible effects of Higgs bosons on experimental observables", Nucl. Phys.* **B161** (1979) 493.

[28] J. F. Donoghue and L. F. Li, *"Properties of charged Higgs bosons", Phys. Rev.* **D19** (1979) 945.

[29] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, *"The Higgs Hunter's Guide",* . SCIPP-89/13.

[30] P. Fayet, *"Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino", Nucl. Phys.* **B90** (1975) 104–124.

[31] M. Drees, *"Supersymmetric models with extended Higgs sector", Int. J. Mod. Phys.* **A4** (1989) 3635.

[32] J. R. Ellis, J. F. Gunion, H. E. Haber, L. Roszkowski, and F. Zwirner, *"Higgs bosons in a nonminimal supersymmetric model", Phys. Rev.* **D39** (1989) 844.

[33] S. Asai *et al.*, *"Prospects for the search for a standard model Higgs boson in ATLAS using vector boson fusion", Eur. Phys. J.* **C32S2** (2004) 19–54, `arXiv:hep-ph/0402254`.

[34] S. Abdullin *et al.*, *"Summary of the CMS Potential for the Higgs Boson Discovery",* . `http://cmsdoc.cern.ch/documents/03/note03_033.pdf`. CMS Note 2003/033.

[35] A. Djouadi, *"The anatomy of electro-weak symmetry breaking (I). The Higgs boson in the standard model", Phys. Rept.* **457** (2008) 1–216, `arXiv:hep-ph/0503172`.

[36] A. Djouadi, *"The anatomy of electro-weak symmetry breaking (II). The Higgs bosons in the minimal supersymmetric model", Phys. Rept.* **459** (2008) 1–241, `arXiv:hep-ph/0503173`.

[37] S. Weinberg, *"Gauge hierarchies", Phys. Lett.* **B82** (1979) 387.

[38] M. J. G. Veltman, *"The infrared – ultraviolet connection", Acta Phys. Polon.* **B12** (1981) 437.

[39] C. H. Llewellyn Smith and G. G. Ross, *"The real gauge hierarchy problem", Phys. Lett.* **B105** (1981) 38.

[40] J. Wess and B. Zumino, *"Supergauge transformations in four dimensions", Nucl. Phys.* **B70** (1974) 39–50.

[41] D. V. Volkov and V. P. Akulov, *"Is the neutrino a Goldstone particle?", Phys. Lett.* **B46** (1973) 109–110.

[42] S. Dimopoulos, S. Raby, and F. Wilczek, *"Supersymmetry and the scale of unification"*, *Phys. Rev.* **D24** (1981) 1681–1683.

[43] L. E. Ibanez and G. G. Ross, *"Low-energy predictions in supersymmetric Grand Unified Theories"*, *Phys. Lett.* **B105** (1981) 439.

[44] S. Dimopoulos and H. Georgi, *"Softly broken supersymmetry and $SU(5)$"*, *Nucl. Phys.* **B193** (1981) 150.

[45] U. Amaldi, W. de Boer, and H. Furstenau, *"Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP"*, *Phys. Lett.* **B260** (1991) 447–455.

[46] R. Haag, J. T. Lopuszanski, and M. Sohnius, *"All possible generators of supersymmetries of the S matrix"*, *Nucl. Phys.* **B88** (1975) 257.

[47] S. R. Coleman and J. Mandula, *"All possible symmetries of the S matrix"*, *Phys. Rev.* **159** (1967) 1251–1256.

[48] H. P. Nilles, *"Supersymmetry, supergravity and particle physics"*, *Phys. Rept.* **110** (1984) 1–162.

[49] H. E. Haber and G. L. Kane, *"The search for supersymmetry: Probing physics beyond the Standard Model"*, *Phys. Rept.* **117** (1985) 75–263.

[50] J. F. Gunion and H. E. Haber, *"Higgs bosons in supersymmetric models (I)."*, *Nucl. Phys.* **B272** (1986) 1, Erratum-ibid. **B402** (1993) 567-569.

[51] J. F. Gunion and H. E. Haber, *"Higgs bosons in supersymmetric models (II). Implications for phenomenology"*, *Nucl. Phys.* **B278** (1986) 449.

[52] J. Wess and J. Bagger, *"Supersymmetry and supergravity"*, .

[53] S. P. Martin, *"A supersymmetry primer"*, `arXiv:hep-ph/9709356`.

[54] H. P. Nilles, M. Srednicki, and D. Wyler, *"Weak interaction breakdown induced by supergravity"*, *Phys. Lett.* **B120** (1983) 346.

[55] A. D. Sakharov, *"Violation of CP invariance, c Asymmetry, and baryon asymmetry of the universe"*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32–35.

[56] L. Fromme, S. J. Huber, and M. Seniuch, *"Baryogenesis in the two-Higgs doublet model"*, *JHEP* **11** (2006) 038, `arXiv:hep-ph/0605242`.

[57] M. Kobayashi and T. Maskawa, *"CP violation in the renormalizable theory of weak interaction"*, *Prog. Theor. Phys.* **49** (1973) 652–657.

[58] W. Bernreuther, *"CP violation and baryogenesis", Lect. Notes Phys.* **591** (2002) 237–293, arXiv:hep-ph/0205279.

[59] T. D. Lee, *"A theory of spontaneous T violation", Phys. Rev.* **D8** (1973) 1226–1239.

[60] W. Bernreuther and O. Nachtmann, *"Flavour dynamics with general scalar fields", Eur. Phys. J.* **C9** (1999) 319–333, arXiv:hep-ph/9812259.

[61] F. Nagel, *New aspects of gauge-boson couplings and the Higgs sector.* PhD thesis, University Heidelberg, 2004.

[62] S. L. Glashow and S. Weinberg, *"Natural conservation laws for neutral currents", Phys. Rev.* **D15** (1977) 1958.

[63] ALEPH, DELPHI, L3, OPAL, SLD Collaborations and LEP Electroweak Group, SLD Working Group, SLD Heavy Flavour Group, *"Precision electroweak measurements on the Z resonance", Phys. Rept.* **427** (2006) 257, arXiv:hep-ex/0509008.

[64] O. Nachtmann, *Elementary Particle Physics: Concepts and Phenomena.* Springer Verlag, Berlin, Heidelberg, 1990.

[65] H.-S. Tsao, *"Higgs boson quantum numbers and the Pell equation", .* COO-2232B-199, in: *Proc. of Guangzhou Conf. on Theoretical Particle Physics, Guangzhou, PRC, Jan 5-12, 1980.*

[66] OPAL Collaboration, G. Abbiendi *et al.*, *"Flavour independent h^0 A^0 search and two Higgs doublet model interpretation of neutral Higgs boson searches at LEP", Eur. Phys. J.* **C40** (2005) 317–332, arXiv:hep-ex/0408097.

[67] LEP Higgs Working Group, *"Search for charged Higgs bosons: Preliminary combined results using LEP data collected at energies up to 209 GeV",* arXiv:hep-ex/0107031.

[68] L. J. Hall and M. B. Wise, *"Flavour changing Higgs-boson couplings", Nucl. Phys.* **B187** (1981) 397.

[69] R. Barbieri and L. J. Hall, *"Improved naturalness and the two Higgs doublet model",* arXiv:hep-ph/0510243.

[70] B. M. Kastening, *"Bounds from stability and symmetry breaking on parameters in the two Higgs doublet potential",* arXiv:hep-ph/9307224.

[71] J. Velhinho, R. Santos, and A. Barroso, *"Tree level vacuum stability in two-Higss doublet models", Phys. Lett.* **B322** (1994) 213–218.

[72] L. Lavoura and J. P. Silva, *"Fundamental CP violating quantities in a $SU(2) \times U(1)$ model with many Higgs doublets"*, *Phys. Rev.* **D50** (1994) 4619–4624, arXiv:hep-ph/9404276.

[73] F. J. Botella and J. P. Silva, *"Jarlskog - like invariants for theories with scalars and fermions"*, *Phys. Rev.* **D51** (1995) 3870–3875, arXiv:hep-ph/9411288.

[74] S. Davidson and H. E. Haber, *"Basis-independent methods for the two-Higgs-doublet model"*, *Phys. Rev.* **D72** (2005) 035004, Erratum-ibid. **D72** (2005) 099902, arXiv:hep-ph/0504050.

[75] J. F. Gunion and H. E. Haber, *"Conditions for CP-violation in the general two-Higgs- doublet model"*, *Phys. Rev.* **D72** (2005) 095002, arXiv:hep-ph/0506227v2.

[76] G. C. Branco, M. N. Rebelo, and J. I. Silva-Marcos, *"CP-odd invariants in models with several Higgs doublets"*, *Phys. Lett.* **B614** (2005) 187–194, arXiv:hep-ph/0502118.

[77] W. Grimus and M. N. Rebelo, *"Automorphisms in gauge theories and the definition of CP and P"*, *Phys. Rept.* **281** (1997) 239–308, arXiv:hep-ph/9506272.

[78] H. E. Haber and D. O'Neil, *"Basis-independent methods for the two-Higgs-doublet model (II). The significance of* $\tan\beta$*"*, *Phys. Rev.* **D74** (2006) 015018, Erratum-ibid. **D74** (2006) 059905, arXiv:hep-ph/0602242.

[79] P. M. Ferreira, R. Santos, and A. Barroso, *"Stability of the tree-level vacuum in two Higgs doublet models against charge or CP spontaneous violation"*, *Phys. Lett.* **B603** (2004) 219–229, arXiv:hep-ph/0406231.

[80] A. Barroso, P. M. Ferreira, and R. Santos, *"Charge and CP symmetry breaking in two Higgs doublet models"*, *Phys. Lett.* **B632** (2006) 684–687, arXiv:hep-ph/0507224.

[81] I. F. Ginzburg, *"Symmetries of 2HDM, different vacua, CP violation and possible relation to a history of time"*, *Acta Phys. Polon.* **B37** (2006) 1161–1172, arXiv:hep-ph/0512102.

[82] A. Barroso, P. M. Ferreira, and R. Santos, *"Neutral minima in two-Higgs doublet models"*, *Phys. Lett.* **B652** (2007) 181–193, arXiv:hep-ph/0702098.

[83] C. C. Nishi, *"CP violation conditions in N-Higgs-doublet potentials"*, *Phys. Rev.* **D74** (2006) 036003, Erratum-ibid. **D76** (2007) 119901, arXiv:hep-ph/0605153v3.

[84] I. P. Ivanov, *"Two-Higgs-doublet model from the group-theoretic perspective"*, *Phys. Lett.* **B632** (2006) 360–365, arXiv:hep-ph/0507132.

[85] I. P. Ivanov, *"Minkowski space structure of the Higgs potential in 2HDM"*, *Phys. Rev.* **D75** (2007) 035001, Erratum-ibid. **D76** (2007) 039902, arXiv:hep-ph/0609018.

[86] I. P. Ivanov, *"Minkowski space structure of the Higgs potential in 2HDM (II). Minima, symmetries, and topology"*, *Phys. Rev.* **D77** (2008) 015017, arXiv:0710.3490 [hep-ph].

[87] S. Kraml *et al.*, eds., *Workshop on CP studies and non-standard Higgs physics*. 2006. arXiv:hep-ph/0608079.

[88] M. Maniatis, A. von Manteuffel, O. Nachtmann, and F. Nagel, *"Stability and symmetry breaking in the general two-Higgs- doublet model"*, *Eur. Phys. J.* **C48** (2006) 805–823, arXiv:hep-ph/0605184.

[89] M. Maniatis, A. von Manteuffel, and O. Nachtmann, *"A new type of CP symmetry, family replication and fermion mass hierarchies"*, arXiv:0711.3760 [hep-ph].

[90] M. Maniatis, A. von Manteuffel, and O. Nachtmann, *"CP violation in the general two-Higgs-doublet model: A geometric view"*, arXiv:0707.3344 [hep-ph].

[91] M. S. Carena and H. E. Haber, *"Higgs boson theory and phenomenology"*, *Prog. Part. Nucl. Phys.* **50** (2003) 63–152, arXiv:hep-ph/0208209.

[92] LEP Higgs Working Group, S. Schael *et al.*, *"Search for neutral MSSM Higgs bosons at LEP"*, *Eur. Phys. J.* **C47** (2006) 547–587, arXiv:hep-ex/0602042.

[93] S. Heinemeyer, W. Hollik, and G. Weiglein, *"Electroweak precision observables in the minimal supersymmetric standard model"*, *Phys. Rept.* **425** (2006) 265–368.

[94] M. Frank, T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, and G. Weiglein, *"The Higgs boson masses and mixings of the complex MSSM in the Feynman-diagrammatic approach"*, *JHEP* **02** (2007) 047, arXiv:hep-ph/0611326.

[95] P. Skands *et al.*, *"SUSY Les Houches accord: Interfacing SUSY spectrum calculators, decay packages, and event generators"*, *JHEP* **07** (2004) 036, arXiv:hep-ph/0311123.

[96] ALEPH, DELPHI, L3, OPAL Collaborations and LEP SUSY Working Group. http://lepsusy.web.cern.ch/lepsusy/Welcome.html. note LEPSUSYWG/04-01.1.

[97] N. K. Falck, *"Renormalization group equations for softly broken supersymmetry: The most general case"*, *Z. Phys.* **C30** (1986) 247.

[98] J. E. Kim and H. P. Nilles, *"The μ problem and the strong CP problem"*, *Phys. Lett.* **B138** (1984) 150.

[99] G. F. Giudice and A. Masiero, *"A natural solution to the μ problem in supergravity theories"*, *Phys. Lett.* **B206** (1988) 480–484.

[100] D. J. Miller, R. Nevzorov, and P. M. Zerwas, *"The Higgs sector of the Next-to-Minimal Supersymmetric Standard Model"*, *Nucl. Phys.* **B681** (2004) 3–30, arXiv:hep-ph/0304049.

[101] R. D. Peccei and H. R. Quinn, *"CP conservation in the presence of instantons"*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.

[102] S. Weinberg, *"A new light boson?"*, *Phys. Rev. Lett.* **40** (1978) 223–226.

[103] F. Wilczek, *"Problem of strong P and T invariance in the presence of instantons"*, *Phys. Rev. Lett.* **40** (1978) 279–282.

[104] S. A. Abel, S. Sarkar, and P. L. White, *"On the cosmological domain wall problem for the minimally extended supersymmetric standard model"*, *Nucl. Phys.* **B454** (1995) 663–684, arXiv:hep-ph/9506359.

[105] C. Panagiotakopoulos and K. Tamvakis, *"Stabilized NMSSM without domain walls"*, *Phys. Lett.* **B446** (1999) 224–227, arXiv:hep-ph/9809475.

[106] U. Ellwanger and C. Hugonie, *"The upper bound on the lightest Higgs mass in the NMSSM revisited"*, *Mod. Phys. Lett.* **A22** (2007) 1581–1590, arXiv:hep-ph/0612133v4.

[107] P. C. Schuster and N. Toro, *"Persistent fine-tuning in supersymmetry and the NMSSM"*, arXiv:hep-ph/0512189.

[108] U. Ellwanger, J. F. Gunion, and C. Hugonie, *"Difficult scenarios for NMSSM Higgs discovery at the LHC"*, *JHEP* **07** (2005) 041, arXiv:hep-ph/0503203.

[109] K. Funakubo and S. Tao, *"The Higgs sector in the Next-to-MSSM"*, *Prog. Theor. Phys.* **113** (2005) 821–842, arXiv:hep-ph/0409294v2.

[110] U. Ellwanger, J. F. Gunion, and C. Hugonie, *"NMHDECAY: A Fortran code for the Higgs masses, couplings and decay widths in the NMSSM"*, *JHEP* **02** (2005) 066, arXiv:hep-ph/0406215.

[111] U. Ellwanger and C. Hugonie, *"NMHDECAY 2.0: An updated program for sparticle masses, Higgs masses, couplings and decay widths in the NMSSM"*, hep-ph/0508022.

[112] A. Menon, D. E. Morrissey, and C. E. M. Wagner, *"Electroweak baryogenesis and dark matter in the nMSSM"*, *Phys. Rev.* **D70** (2004) 035005, arXiv:hep-ph/0404184.

[113] C. Balazs, M. S. Carena, A. Freitas, and C. E. M. Wagner, *"Phenomenology of the nMSSM from colliders to cosmology"*, *JHEP* **06** (2007) 066, arXiv:0705.0431 [hep-ph].

[114] M. Maniatis, A. von Manteuffel, and O. Nachtmann, *"Determining the global minimum of Higgs potentials via Groebner bases – applied to the NMSSM"*, *Eur. Phys. J.* **C49** (2007) 1067–1076, arXiv:hep-ph/0608314.

[115] R. Fletcher, *Practical Methods of Optimization*. Wiley, Chichester & New York, 2nd edition, 1987.

[116] G. Ecker, W. Grimus, and H. Neufeld, *"A standard form for generalized CP transformations"*, *J. Phys.* **A20** (1987) L807.

[117] S. Weinberg, *"The Quantum Theory of Fields. Vol. I: Foundations"*, .

[118] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, Cambridge, UK, 1985.

[119] B. C. Allanach *et al.*, *"The Snowmass points and slopes: Benchmarks for SUSY searches"*, arXiv:hep-ph/0202233.

[120] H. E. Haber and R. Hempfling, *"The renormalization group improved Higgs sector of the minimal supersymmetric model"*, *Phys. Rev.* **D48** (1993) 4280–4309, arXiv:hep-ph/9307201.

[121] M. Jamin, J. A. Oller, and A. Pich, *"Scalar K π form factor and light quark masses"*, *Phys. Rev.* **D74** (2006) 074009, arXiv:hep-ph/0605095.

[122] Quarkonium Working Group, N. Brambilla *et al.*, *"Heavy quarkonium physics"*, arXiv:hep-ph/0412158.

[123] M. Koecher, *Lineare Algebra und analytische Geometrie*. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 2nd edition, 1985.

[124] I. Wolfram Research, Mathematica, versions 5.0, 5.1, 5.2 and 6.0, 2003-2007.

[125] B. Buchberger, *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*. PhD thesis, University Innsbruck, 1965.

[126] J. C. Faugère, *"A bew efficient algorithm for computing Gröbner bases (F_4)"*, *J. of Pure and Applied Algebra* **139, Issue 1-3** (1999) 61.

[127] J. C. Faugère, P. Gianni, D. Lazard, and T. Mora, *"Efficient computation of zero-dimensional Gröbner basis by change of ordering", J. Symb. Comput.* **16, Number 4** (1993) 329.

[128] H. M. Moeller, *"On decomposing systems of polynomial equations with finitely many solutions", Appl. Algebra Eng. Commun. Comput.* **4** (1993) 217.

[129] D. Hillebrand, *Triangulierung nulldimensionaler Ideale - Implementierung und Vergleich zweier Algorithmen.* PhD thesis, University Dortmund, 1999.

[130] M. Kreuzer and L. Robbiano, *Computational Commutative Algebra 1.* Springer, Berlin, Heidelberg, New-York, 2000.

[131] M. Kreuzer and L. Robbiano, *Computational Commutative Algebra 2.* Springer, Berlin, Heidelberg, New-York, 2005.

[132] G. M. Greuel, G. Pfister, and H. Schoenemann, *"*`Singular` *- A computer algebra system for polynomial computations",* in *Symbolic Computation and Automated Reasoning, The Calculemus-2000 Symposium*, M. Kerber and M. Kohlhase, eds., p. 227. 2001. `http://www.singular.uni-kl.de`.

[133] B. Foster, A. D. Martin, and M. G. Vincter, *"Structure functions", Phys. Rev.* **D66** (2002) 010001.

[134] L. N. Hand, *"Experimental investigation of pion electroproduction", Phys. Rev.* **129** (1963) 1834–1846.

[135] C. Ewerz and O. Nachtmann, *"Towards a nonperturbative foundation of the dipole picture (I). Functional methods", Annals Phys.* **322** (2007) 1635–1669, `arXiv:hep-ph/0404254`.

[136] C. Ewerz and O. Nachtmann, *"Towards a nonperturbative foundation of the dipole picture (II). High energy limit", Annals Phys.* **322** (2007) 1670–1726, `arXiv:hep-ph/0604087`.

[137] D. Grunewald, E. M. Ilgenfritz, E. V. Prokhvatilov, and H. J. Pirner, *"Formulating light cone QCD on the lattice", Phys. Rev.* **D77** (2008) 014512, `arXiv:0711.0620 [hep-lat]`.

[138] K. J. Golec-Biernat and M. Wusthoff, *"Saturation effects in deep inelastic scattering at low Q^2 and its implications on diffraction", Phys. Rev.* **D59** (1998) 014017, `arXiv:hep-ph/9807513`.

[139] J. R. Forshaw, G. Kerley, and G. Shaw, *"Extracting the dipole cross-section from photo- and electro-production total cross-section data", Phys. Rev.* **D60** (1999) 074012, `hep-ph/9903341`.

[140] A. Donnachie and H. G. Dosch, *"Diffractive exclusive photon production in deep inelastic scattering"*, *Phys. Lett.* **B502** (2001) 74–78, hep-ph/0010227.

[141] A. Donnachie and H. G. Dosch, *"A comprehensive approach to structure functions"*, *Phys. Rev.* **D65** (2001) 014019, arXiv:hep-ph/0106169.

[142] E. Iancu, K. Itakura, and S. Munier, *"Saturation and BFKL dynamics in the HERA data at small x"*, *Phys. Lett.* **B590** (2004) 199–208, arXiv:hep-ph/0310338.

[143] J. R. Forshaw and G. Shaw, *"Gluon saturation in the colour dipole model?"*, *JHEP* **12** (2004) 052, arXiv:hep-ph/0411337.

[144] J. R. Forshaw, R. Sandapen, and G. Shaw, *"Further success of the colour dipole model"*, *JHEP* **11** (2006) 025, arXiv:hep-ph/0608161.

[145] S. Donnachie, H. G. Dosch, O. Nachtmann, and P. Landshoff, *Pomeron Physics and QCD*. Cambridge University Press, 2002.

[146] J. Bartels, K. J. Golec-Biernat, and H. Kowalski, *"A modification of the saturation model: DGLAP evolution"*, *Phys. Rev.* **D66** (2002) 014001, arXiv:hep-ph/0203258.

[147] V. N. Gribov and L. N. Lipatov, *"Deep inelastic ep scattering in perturbation theory"*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[148] L. N. Lipatov, *"The parton model and perturbation theory"*, *Sov. J. Nucl. Phys.* **20** (1975) 94–102.

[149] Y. L. Dokshitzer, *"Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in Quantum Chromodynamics. (In Russian)"*, *Sov. Phys. JETP* **46** (1977) 641–653.

[150] G. Altarelli and G. Parisi, *"Asymptotic freedom in parton language"*, *Nucl. Phys.* **B126** (1977) 298.

[151] L. Frankfurt, A. Radyushkin, and M. Strikman, *"Interaction of small size wave packet with hadron target"*, *Phys. Rev.* **D55** (1997) 98–104, arXiv:hep-ph/9610274.

[152] C. Ewerz, A. von Manteuffel, and O. Nachtmann, *"On the range of validity of the dipole picture"*, *Phys. Rev.* **D77** (2008) 074022, arXiv:0708.3455 [hep-ph].

[153] C. Ewerz and O. Nachtmann, *"Bounds on ratios of DIS structure functions from the color dipole picture"*, *Phys. Lett.* **B648** (2007) 279–283, arXiv:hep-ph/0611076.

[154] H. Abramowicz, E. M. Levin, A. Levy, and U. Maor, *"A parametrization of $\sigma_T(\gamma^* p)$ above the resonance region $Q^2 \geq 0$"*, *Phys. Lett.* **B269** (1991) 465–476.

[155] H. Abramowicz and A. Levy, *"The ALLM parameterization of $\sigma_{tot}(\gamma^* p)$: An update"*, arXiv:hep-ph/9712415.

[156] ZEUS Collaboration, J. Breitweg *et al.*, *"Measurement of the proton structure function F_2 at very low Q^2 at HERA"*, *Phys. Lett.* **B487** (2000) 53–73, arXiv:hep-ex/0005018.

[157] ZEUS Collaboration, S. Chekanov *et al.*, *"High-Q**2 neutral current cross sections in e^+p deep inelastic scattering at $\sqrt{s} = 318\,GeV$"*, *Phys. Rev.* **D70** (2004) 052001, arXiv:hep-ex/0401003.

[158] ZEUS Collaboration, S. Chekanov *et al.*, *"Measurement of the neutral current cross section and F_2 structure function for deep inelastic e^+p scattering at HERA"*, *Eur. Phys. J.* **C21** (2001) 443–471, arXiv:hep-ex/0105090.

[159] H1 Collaboration, C. Adloff *et al.*, *"Deep-inelastic inclusive ep scattering at low x and a determination of α_s"*, *Eur. Phys. J.* **C21** (2001) 33–61, hep-ex/0012053.

[160] H1 Collaboration, *"Measurement of the proton structure function F_L at low x"*, arXiv:0805.2809 [hep-ex].

[161] M. Rueter, *"Energy and Q^2 dependence of elastic vector meson production and the proton structure function F_2"*, *Eur. Phys. J.* **C7** (1999) 233–249, hep-ph/9807448.

[162] A. Donnachie, H. G. Dosch, and M. Rueter, *"Two photon reactions at high energies"*, *Phys. Rev.* **D59** (1999) 074011, hep-ph/9810206.

[163] A. Donnachie, H. G. Dosch, and M. Rueter, *"$\gamma^*\gamma^*$ reactions at high energies"*, *Eur. Phys. J.* **C13** (2000) 141–150, hep-ph/9908413.

[164] O. Nachtmann, *"Considerations concerning diffraction scattering in Quantum Chromodynamics"*, *Annals Phys.* **209** (1991) 436–478.

[165] V. N. Gribov, B. L. Ioffe, and I. Y. Pomeranchuk, *"What is the range of interactions at high-energies"*, *Sov. J. Nucl. Phys.* **2** (1966) 549. [Yad. Fiz. **2** (1965) 768].

[166] B. L. Ioffe, *"Space-time picture of photon and neutrino scattering and electroproduction cross-section asymptotics"*, *Phys. Lett.* **B30** (1969) 123–125.

[167] V. Braun, P. Gornicki, and L. Mankiewicz, *"Ioffe - time distributions instead of parton momentum distributions in description of deep inelastic scattering"*, *Phys. Rev.* **D51** (1995) 6036–6051, arXiv:hep-ph/9410318.

[168] Y. V. Kovchegov and M. Strikman, *"Ioffe time in double logarithmic approximation"*, *Phys. Lett.* **B516** (2001) 314–320, arXiv:hep-ph/0107015.

[169] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. National Bureau of Standards, USA, 10th printing, 1972. `http://www.math.sfu.ca/~cbm/aands`.

[170] S. L. Moshier, *Methods and Programs for Mathematical Functions*. Prentice Hall Europe, 1989. `http://www.netlib.org/cephes`. Cephes Math Library, Release 2.8, 2000.

[171] `IT++`. *An open-source* `C++` *library of mathematical, signal processing, speech processing, and communications classes and functions*, versions 3.10.9 and 4.0.0. `http://itpp.sourceforge.net`.

[172] M. Galassi *et al.*, *GNU Scientific Library Reference Manual*, 2nd edition, 2006. `http://www.gnu.org/software/gsl/`. GSL library, versions 1.7 and 1.8.

[173] T. Hahn, *"*`CUBA`*: A library for multidimensional numerical integration", Comput. Phys. Commun.* **168** (2005) 78–95, `arXiv:hep-ph/0404043`.

[174] T. Becker and V. Weispfenning, *Gröbner Bases*. Springer-Verlag New York, 1993.

[175] N. K. Bose, J. P. Guiver, E. W. Kamen, H. M. Valenzuela, and B. Buchberger, *Multidimensional Systems Theory, Progress, Directions and Open Problems in Multidimensional Systems*. D. Reidel Publishing Company, 1985.

[176] S. Iyanaga and Y. Kawada, eds., *Encyclopedic Dictionary of Mathematics*, vol. I, p. 304 ff. MIT Press, Cambridge, Massachusetts, and London, England, 1980.

[177] B. Schölkopf and A. Smola, *Learning with Kernels*, p. 150 ff. MIT Press, Cambridge, MA, 2002.

[178] N. I. Ahiezer and M. Krein, *Some Questions in the Theory of Moments*, p. 128 ff. American Mathematical Society, Providence, Rhode Island, 1962.

[179] W. I. Smirnov, *Lehrgang der höheren Mathematik*, p. 15 ff. VEB Deutscher Verlag der Wissenschaften, Berlin, 1988.

[180] G. James and R. C. James, eds., *Mathematics Dictionary*. D. Van Nostrand Comp., Inc., Toronto, New York, London, 1949.

[181] `Gnuplot`. *A freeware command-line driven plotting utility.* `http://www.gnuplot.info`.

[182] `Inkscape`. *An open-source vector graphics editor.* `http://inkscape.org`.

[183] `Xfig`. *An open-source vector graphics editor.* `http://xfig.org`.

[184] L. Lamport, *LaTeX: A Document Preparation System*. Addison-Wesley, Reading, Massachusetts, 2nd edition, 1994.

Acknowledgments

First I would like to express my deep gratitude to my supervisor Professor Nachtmann for the proposal of this thesis, his enduring encouragement and permanent extraordinary support. I thank him for many inspiring discussions and very much appreciate what I could learn from him.

Professor Schultz-Coulon took over the burden of the second referee. I thank him very much for this work and his thorough questions.

For the dipole picture part of this work, I thank Carlo Ewerz for the good collaboration with many interesting discussions. The Higgs part of this work was developed in close collaboration with Markos Maniatis, who also initiated my contact to Professor Nachtmann. I thank him for both as well as for many suggestions during the work and comments on the manuskript.

For this thesis I employed various free software packages [132, 170, 171, 172, 173, 181, 182, 183, 184]. I thank the authors for their generosity in providing these packages, in particular when this includes also open source code.

I enjoyed very much the nice atmosphere at the institute, to which many people contributed. I thank Eduard Thommes for his valuable and kind help with various organisational issues. For their contribution to a pleasant working atmosphere with interesting discussions, I thank the former members of Professor Nachtmann's group Juliane Behrend, Timo Bergmann and Dieter Greiner as well as Andreas Braun and Matthias Kronenwett. I thank Tania Robens for proof-reading parts of the manuskript. I am very indebted to Birgit Eberle for her general support during my work on this thesis, many comments at different stages of this work and proof-reading of the manuskript.

Last but not least, I am especially gratefull to my parents. Their support laid the foundation on which this work became possible.