Thesis for the degree of DOCTOR OF PHILOSOPHY in Fundamental Physics at
CHALMERS UNIVERSITY OF TECHNOLOGY
AND
DOCTOR OF SCIENCE in Theoretical Physics and Mathematics at
UNIVERSITE LIBRE DE BRUXELLES

Arithmetic and Hyperbolic Structures in String Theory

Daniel Persson

This doctoral thesis is the result of joint supervision between Chalmers University of Technology and Université Libre de Bruxelles under the framework of the double doctoral degree agreement signed by both universities on the 8th of September 2008.

Fundamental Physics
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden
and
Physique Théorique et Mathémathique
UNIVERSITE LIBRE DE BRUXELLES & INTERNATIONAL SOLVAY INSTITUTES
Bruxelles, Belgium

2009

Arithmetic and Hyperbolic Structures in String Theory
Daniel Persson
ISBN 978-91-7385-301-9

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 2982
ISSN 0346-718X

Fundamental Physics
Chalmers University of Technology
SE-412 96 Göteborg, Sweden

Physique Théorique et Mathématique
Université Libre de Bruxelles & International Solvay Institutes
Campus Plaine CP 231
1050 Bruxelles, Belgium

The research presented in this thesis has been funded by The International Solvay Institutes, by a grant from PAI, Belgium, by IISN-Belgium (conventions 4.4505.86, 4.4514.08, 4.4511.06 and 4.4514.08), by the Belgian National Lottery, by the European Commission FP6 RTN programme MRTN-CT-2004-005104 and by the Belgian Federal Science Policy Office through the Interuniversity Attraction Poles P5/27 and P6/11.
I am very grateful for all the generous support.

Chalmers Reproservice
Göteborg, Sweden 2009

Arithmetic and Hyperbolic Structures in String Theory

Daniel Persson

Fundamental Physics
Chalmers University of Technology
and
Physique Théorique et Mathématique
Université Libre de Bruxelles & International Solvay Institutes

Abstract

This thesis consists of an introductory text followed by two separate parts which may be read independently of each other. In **Part I** we analyze certain hyperbolic structures arising when studying gravity in the vicinity of spacelike singularities (the BKL-limit). In this limit, spatial points decouple and the dynamics exhibits ultralocal behaviour which may be mapped to an auxiliary problem given in terms of a (possibly chaotic) hyperbolic billiard. In all supergravities arising as low-energy limits of string theory or M-theory, the billiard dynamics takes place within the fundamental Weyl chambers of certain hyperbolic Kac-Moody algebras, suggesting that these algebras generate hidden infinite-dimensional symmetries of gravity. We investigate the modification of the billiard dynamics when the original gravitational theory is formulated on a compact spatial manifold of arbitrary topology, revealing fascinating mathematical structures known as *galleries*. We further use the conjectured hyperbolic symmetry $\mathcal{E}_{10}$ to generate and classify certain cosmological (S-brane) solutions in eleven-dimensional supergravity. Finally, we show in detail that eleven-dimensional supergravity and massive type IIA supergravity are dynamically unified within the framework of a geodesic sigma model for a particle moving on the infinite-dimensional coset space $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$.

Part II of the thesis is devoted to a study of how (U-)dualities in string theory provide powerful constraints on perturbative and non-perturbative quantum corrections. These dualities are typically given by certain arithmetic groups $G(\mathbb{Z})$ which are conjectured to be preserved in the effective action. The exact couplings are given by moduli-dependent functions which are manifestly invariant under $G(\mathbb{Z})$, known as *automorphic forms.* We discuss in detail various methods of constructing automorphic forms, with particular emphasis on a special class of functions known as (non-holomorphic) *Eisenstein series.* We provide detailed examples for the physically relevant cases of $SL(2,\mathbb{Z})$ and $SL(3,\mathbb{Z})$, for which we construct their respective Eisenstein series and compute their (non-abelian) Fourier expansions. We also discuss the possibility that certain generalized Eisenstein series, which are covariant under the maximal compact subgroup $K(G)$, could play a role in determining the exact effective action for toroidally compactified higher derivative corrections. Finally, we propose that in the case of rigid Calabi-Yau compactifications in type IIA string theory, the exact universal hypermultiplet moduli space exhibits a quantum duality group given by the *Picard modular group* $SU(2,1;\mathbb{Z}[i])$. To verify this proposal we construct an $SU(2,1;\mathbb{Z}[i])$-invariant Eisenstein series, and we present preliminary results for its Fourier expansion which reveals the expected contributions from D2-brane and NS5-brane instantons.

This thesis is based on the following seven papers, henceforth referred to as **Paper I–VII**:

I M. Henneaux, M. Leston, D. Persson and Ph. Spindel,
Geometric Configurations, Regular Subalgebras of E_{10} and M-Theory Cosmology,
JHEP **0610** (2006) 021 [arXiv:hep-th/0606123].

II M. Henneaux, M. Leston, D. Persson and Ph. Spindel,
A Special Class of Rank 10 and 11 Coxeter Groups,
J. Math. Phys. **48** (2007) 053512 [arXiv:hep-th/0610278].

III M. Henneaux, D. Persson and Ph. Spindel,
Spacelike Singularities and Hidden Symmetries of Gravity,
Living Rev. Rel. **1** (2008) 1 [arXiv:0710.1818].

IV L. Bao, J. Bielecki, M. Cederwall, B.E.W. Nilsson and D. Persson,
U-Duality and the Compactified Gauss-Bonnet Term,
JHEP **0807** (2008) 048 [arXiv:0710.4907].

V M. Henneaux, D. Persson and D.H. Wesley,
Coxeter Group Structure of Cosmological Billiards on Compact Spatial Manifolds,
JHEP **0809** (2008) 052 [arXiv:0805.3793].

VI M. Henneaux, E. Jamsin, A. Kleinschmidt and D. Persson,
On the E_{10}/Massive Type IIA Supergravity Correspondence,
Phys. Rev. **D79** (2009) 045008 [arXiv:0811.4358].

VII B. Pioline and D. Persson,
The Automorphic NS5-Brane (2.0),
[arXiv:0902.3274v2].

In addition, Chapter 12 is based on the following work in progress, referred to as **Paper VIII**:

VIII L. Bao, C. Colonnello, A. Kleinschmidt, B.E.W. Nilsson and D. Persson,
Instanton Corrections to the Universal Hypermultiplet and Automorphic Forms,
Work in progress.

Acknowledgments

I have been given the great privilege of being guided through my PhD with the aid of two amazing advisors. I am very grateful to Marc Henneaux and Bengt E. W. Nilsson for your unwavering support and encouragement these last four years. Thank you so much Marc for the many enjoyable and rewarding collaborations, helpful discussions and for always making sure that I am busy. Thank you so much Bengt for very close collaboration during the course of six years now, and for our numerous discussions on theoretical physics in general and string theory in particular, during which I have learned so many things.

In the second half of my PhD, I have also been fortunate enough to work on several projects in close collaboration with Axel Kleinschmidt, who has become much like an extra advisor. I am very grateful for our extensive discussions on everything from E_{10} to number theory, and for many helpful remarks which greatly improved the contents of this thesis. Thank you so much Axel!

Part II of this thesis had probably not been very coherent if it had not been for the generous help I have received from Boris Pioline during this last year. I am enormously grateful to Boris for all your explanations and clarifications regarding automorphic forms and their role in string theory, as well as for a very enjoyable collaboration during the last few months, and for numerous useful comments and suggestions on a previous draft of this thesis.

I want to extend a very special thanks to my two companions Christoffer Petersson and Jakob Palmkvist for your close friendship and support during all these years. To Christoffer, I am particularly grateful for our endless discussions on string theory, and in particular on the topics presented in Chapters 8 and 12 of this thesis. I also want to thank Jakob especially for being the only one who has actually read through this entire manuscript and given numerous helpful remarks.

During the last months, I have also benefitted greatly from discussions and correspondence with Pierre Vanhove regarding the contents of Chapter 12, and for this I am very grateful. In addition, I want to thank Stefan Vandoren for very useful discussions on instanton corrections in string theory.

I want to thank the members of my defense jury, Riccardo Argurio, Marc Henneaux, Laurent Houart, Bengt E.W. Nilsson, Boris Pioline and Michel Tytgat, for kindly agreeing to put me on "trial" for the work presented in this thesis. Bring it on!

The time of my PhD has been divided between two different research groups, an arrangement from which I have benefitted enormously. I want to begin by thanking the entire PTM-group at ULB; I have enjoyed my time in Brussels enormously, and I am very grateful to all the people in the group for creating a wonderful atmosphere, both scientific and social. In addition, I want to thank everyone in the Fundamental Physics group at Chalmers for "adopting" me, and for always making me feel at home on the sixth floor. Special thanks is due to Ling Bao and Martin Cederwall for very fruitful collaborations – some of which already led to publications, and others based on ideas which are still simmering – and for many enjoyable discussions during the last four years.

I want to thank Daniel H. Wesley for extensive collaboration in the course of almost two years, during which time I benefitted a lot from our discussions about Kac-Moody algebras, BKL-chaos, Big Bang-cosmology and other related topics. I am also grateful to Philippe Spindel for fruitful collaborations during the first two years of my PhD. I was very fortunate to have you as an unofficial advisor when I arrived in Brussels, and I learned so much from our discussions. In addition, I am grateful to Mauricio Leston for six months of intense collaboration, and really good times. You definitely deserve extra credit for all the stories you left behind: "In order to be efficient"... :-)

At Chalmers, I want to thank Ulf Gran and Niklas Wyllard especially for contributing to the very nice atmosphere on the sixth floor. I am grateful to Ulf for many useful discussions, and in particular for invaluable help with GAMMA. I am also grateful to Niklas for helpful discussions on automorphic forms and instantons. A special thanks is due also to Gabriele Ferretti, to whom I am very grateful for constant encouragement and support throughout the course of my PhD. An additional big thanks goes out to all the other members of the Fundamental Physics group for making it a great place to do research: Lars Brink, Måns Henningson, Robert Marnelius, Fredrik Ohlsson and Per Salomonson.

At ULB, I want to begin by thanking all the “old-school” students, Nazim Bouatta, Sophie de Buyl, Geoffrey Compère and Stephane Detournay for welcoming a lost Swede with open arms, and for all the fun times we had in Modave! I am especially grateful to Jarah Evslin and Carlo Maccaferri for the first years when I didn’t just work non-stop, but actually took the time to enjoy the remarkable assortment of beers that Brussels has to offer, during which time I also learned all about the Atiyah-Hirzebruch spectral sequence, Weinberg Chapter 3, string field theory, bubbling AdS_3... Jarah also deserves a special acknowledgment for being the ultimate landlord; may “Casa Jarah” prosper for a long time after I’m gone! I wish to thank my favourite Russian, Stanislav Kuperstein – the world’s leading expert on AdS/CFT and Wikipedia – for the great times we shared on the sixth floor, before I abandoned you for the Kac-Moody group down below. Sorry! I am grateful for our endless discussions on physics, Russian jokes, Karlsson på taket, Pomperipossa, Carrollian peculiarities... A very special acknowledgment goes out to the multicultural team Ella Jamsin and Alex “Fiber Bundle” Wijns for your close friendship during these years. Thanks Ella for being a great office-mate, for making amazing lemon chicken, for very nice collaboration and for always taking care of the Swedish guy when he is lost in the bureaucratic jungle of Brussels. Thanks Alex for sharing my interest in music, for very close (almost) collaboration, and for all the papers we will write together in the future!

I am very grateful to Riccardo Argurio for taking care of me from day one in Brussels, and for many nice discussions which almost led us (and Stanislav) to uncover BKL-chaos and Kac-Moody algebras in AdS/CFT; I still think those duality walls are important! I thank Laurent Houart, Josef Lindman-Hörnlund and Nassiba Tabti for very nice “non-split” collaborations during the second half of my PhD, and I’m really sorry for being a hopeless co-author these last few months. Now that this monster is finished, I’ll get better! Chethan Krishnan deserves a special thanks for being a great friend and a really good physicist who is always ready to tackle new things! To my companions at “Casah Jarah”, Alice Bernamonti and Cyril Closset, I want to extend an big thanks for your friendship and for being the best flatmates! A big thanks also goes out to all the remaining members of the groups at ULB/VUB in Brussels; in particular Glenn Barnich, Francesco Bigazzi, Ben Craps, Francois Dehouck, Frank Ferrari, Emiliano Imeroni, Sung-Soo Kim, Alberto Marriotti, Sergio Montanez Naz, Cedric Troessaert, Amitabh Virmani and Vincent Wens. I want to thank Stephane Detournay, Gaston Giribet and Monica Guica for nice collaboration in the organization of the (very successful!) working group on 3d Gravity at the 2009 CERN winter school. I thank Luca Forte for the good times we had during your year in Brussels, and especially for teaching me all about the remarkable lemons of Amalfi!

I am very grateful to the physics department at École Normale Supérieure in Paris for kind hospitality during last fall, and in particular to Costas Bachas, Bernard Julia and Jan Troost for welcoming me and for many enjoyable discussions. In addition, I am especially grateful to Raphael Benichou for your friendship, and for past and future collaborations. Oh, and congratulations!

Let us at last also move out into the real world. Here, I would like to extend my deepest gratitude to my parents, Nina and Tommy, for your constant encouragement and for always supporting whatever decisions I make in life. To my dearest sister, **KATJA**, I want to say thank you so much for being there for me at all times; our close friendship means everything to me. And keep the songs coming!

I want to send a big thanks to my “bonus family” of farmers: Karin, Micke, David, Rebecca and Bina for taking me in and making me a part of your lives. For this, I am forever grateful. In addition, I want to say to all my friends out there: You mean so much to me, even if I may seem a bit disconnected at times. I want to thank especially Calle, Chadster, Fiolen, Mr. Milton and Spiiken for all the years of friendship, and for many more to come. From now on you will have to call me “Doktor Daniel”!

Finally, and most importantly, I want to extend the greatest acknowledgment of all to my wonderful Jojo, without whom this thesis would not have existed. It is not easy to live with a globetrotter of a physicist, and I am so grateful that you have always supported me and stood by me in my choices.

This thesis is dedicated to you.

"Man cannot survive except by gaining knowledge,
and reason is his only means to gain it.
Reason is the faculty that perceives, identifies and integrates
the material provided by his senses.
The task of his senses is to give him the evidence of existence,
but the task of identifying it belongs to his reason,
his senses tell him only that something is,
but what it is must be learned by his mind."

-Ayn Rand

Till Johanna

Contents

1

Introduction – A Tour Through the Dualities of String Theory

1.1 The Principle of Beauty

To unravel the fundamental laws of nature is the ultimate goal of theoretical physics. In order to advance in this ambitious undertaking, it is desirable to have a set of underlying principles to guide us through the mist. A striking example is provided by Einstein's *principle of equivalence*, which states that the concepts of inertial and gravitational mass are merely two sides of the same coin. With this principle as a guiding light, Einstein was led to the discovery of the general theory of relativity.

It was argued by Dirac that the most basic principle that should serve as a lamppost for the pursuit of truth in fundamental laws, is the *principle of beauty.* Beauty is of course a subjective notion, but, nonetheless, physicists and mathematicians alike tend to agree on whether a certain theory, or a mathematical equation, is beautiful. This principle of beauty in fundamental laws should *not* be confused with *simplicity.* The mathematical formulation of the laws of nature exhibits a remarkably rich and complicated structure, for which "simple" is not a fitting description. Dirac therefore restated the basic principle by saying that we should pursue *mathematical beauty*, since this has often, perhaps surprisingly, proven to go hand in hand with an accurate description of nature.

In addition to the general theory of relativity, another great achievement of the twentieth century was the discovery of quantum mechanics. This relied heavily on what may in retrospect be referred to as a *principle of symmetry*, stating that the fundamental laws should look the same irrespective of whether one derives them standing upright or hanging upside down. More accurately, the principle of symmetry demands that the mathematical laws of a theory should remain invariant under a set of *transformations.* In mathematical terminology, such a set of transformations form a *group*, the properties of which are described in terms of *group theory.* These mathematical techniques were vital in the development of the standard model of particle physics, the framework in which the fundamental constituents of nature are described. This theory is essentially determined uniquely by specifying that its mathematical laws should remain invariant under symmetry transformations described by the group $SU(3) \times SU(2) \times U(1)$. This shows that the principle of symmetry is extremely powerful in constraining the possible descriptions of nature. Of course, to uncover which particular symmetry group is relevant for describing a certain property of nature is in general a difficult problem; but it is a *well defined* problem, and therefore sets the physicist onto a clear path forward.

The underlying theme of this thesis is the quest for finding the fundamental symmetry of *string theory*, which is a leading candidate for a theory unifying the two descriptions of nature mentioned above: general relativity - describing the large-scale structure of the universe - and quantum mechanics - the framework in which we describe interactions between the fundamental building blocks of nature, the *elementary particles*. The basic idea of string theory is that these elementary particles correspond to different vibrational modes of a single fundamental object, a *string*. Originally developed in the 1970's, string theory has grown into a vast subject of research, by now encompassing large areas of both physics and mathematics. String theory is a well defined and consistent theory in a mathematical sense; whether or not it describes our universe is yet to be verified. Despite the frustrating lack of experimental evidence for string theory, it exhibits a truly beautiful and intricate structure which may potentially answer some of the most difficult questions we can ask: How did the universe begin? What is the fundamental structure of spacetime? What is inside a black hole? Whether or not the answers are true remains to be revealed.

One of the outstanding problems in string theory is to find a satisfactory fundamental formulation of the theory from first principles, analogous to the way Einstein formulated general relativity based on the principle of equivalence. At present, such a formulation is not known, but rather we have variety of different descriptions which are valid in certain limits and regimes. The remarkable property of the theory, as realized by Witten, is that all of these different descriptions are related in various ways to each other, creating an intricate web of *dualities* which tie the whole structure together into a robust and consistent framework. These dualities reveal an elaborate generalization of the principle of symmetry: although the various descriptions of the theory change under *duality transformations*, the complete structure is left invariant. It is the purpose of this thesis to further explore the notion of duality in string theory from several different viewpoints, with the hope that this line of research will eventually lead us to uncover its basic underlying symmetry.

As a motivational prequel to the remainder of this thesis, we shall now embark on a brief guided tour through the duality web of string theory. The bulk of the thesis is then divided into two main parts – referred to as **Part I** and **Part II** – and the following introduction provides a general overview of the main concepts and ideas which underlie the analysis. It should be noted that the two main parts of the thesis may be read independently. However, it is also the purpose of the present introduction to give the reader a notion of how **Part I** and **Part II** may ultimately be related. Before we begin, it is in order to extend an apology, and a warning, to the layman reader because, despite the gentle words of motivation above, the following text is written at a fairly technical level, and is primarily aimed towards a theoretical physicist who is well versed in the framework of string theory.

1.2 The Duality Web of Type II String Theory

String theory lives in ten spacetime dimensions, with its "mother", M-theory, hovering above in eleven dimensions. There are five different ten-dimensional string theories – known under the esoteric names type I, type IIA and IIB, Heterotic $E_8 \times E_8$ and Heterotic $SO(32)$ – which are all related by various non-trivial dualities, often via M-theory [1]. We will in our current treatment restrict to the type II theories and their relation with M-theory. The other theories are also interesting in their own right, but do not play an immediate role for the analysis in this thesis. We refer the reader to [2–4] for more information.

1.2.1 Duality and Maximal Supersymmetry

The main feature that enables us to relate various string theories to each other is through *compactification*, namely formulating the theory on a product space $\mathbb{R}^{1,D-n} \times X$, where X is an n-dimensional compact internal manifold. One of the prime examples is the relation between ten-dimensional type IIA string theory and M-theory on $\mathbb{R}^{1,9} \times S^1$ [1]. It has been known for a long time that the corresponding low-energy limits, eleven-dimensional supergravity and type IIA supergravity, are related by dimensional reduction on a circle [5–7]. Subsequently, Witten realized that this correspondence in fact extends to the full quantum theories, with one of the key features being the interpretation of the type IIA string coupling $g_s^{(\mathrm{A})} = e^{\phi_{(\mathrm{A})}}$ in terms of the radius of the M-theory circle. This implies that at weak coupling, $g_s^{(\mathrm{A})} \to 0$, there is a nice description of type IIA string theory in terms of a perturbative expansion in $g_s^{(\mathrm{A})}$, while in the opposite, strong-coupling limit, $g_s^{(\mathrm{A})} \to \infty$, the compact dimension of the circle appears, and the theory becomes eleven-dimensional. This implies that *strongly-coupled* type IIA string theory is accurately described in terms of *weakly-coupled* eleven-dimensional supergravity, which in turn is the low-energy limit of M-theory.

By taking yet another direction to be compact, we end up with M-theory compactified on a torus T^2, giving rise to a nine-dimensional theory which is dual to type IIA on S^1 in the sense described above. Furthermore, in nine dimensions, type IIA string theory is related to type IIB string theory via T-duality on the circle. More precisely, T-duality relates type IIA on a circle S^1 of radius R to type IIB on a different circle $\tilde{S}^1$ of radius $1/R$.

We have seen above that the strongly coupled type IIA theory is eleven-dimensional supergravity. What is the corresponding statement for strongly coupled type IIB string theory? The relation between type IIB and M-theory turns out to be slightly more subtle: via T-duality in $D = 9$ one obtains an equivalence between type IIB string theory on S^1 and M-theory on T^2. Under this map, the complex structure $\Omega = \Omega_1 + i\Omega_2$ of the torus is identified with the "axio-dilaton" $\tau = C_{(0)} + ie^{-\phi_{(\mathrm{B})}}$, which contains the type IIB string coupling $g_s^{(\mathrm{B})} = e^{\phi_{(\mathrm{B})}}$. The M-theory formulation exhibits a natural $SL(2,\mathbb{Z})$-invariance associated with modular transformations of the torus, which maps to an $SL(2,\mathbb{Z})$-action on the axio-dilaton τ of type IIB in nine dimensions. However, since the axio-dilaton exists already in $D = 10$, this suggests that ten-dimensional type IIB string theory should in fact be invariant under $SL(2,\mathbb{Z})$ [1,8]. In particular, the action of the modular group on τ inverts the type IIB string coupling, $e^{\phi_{(\mathrm{B})}} \to e^{-\phi_{(\mathrm{B})}}$, implying that type IIB is *self-dual* under S-duality. In other words, strongly coupled type IIB string theory is *equivalent* to weakly coupled type IIB string theory.

So far we have restricted our attention to dualities appearing in nine and ten dimensions. However, more interesting structures appear as we take a larger internal manifold. It has been known since the work of Cremmer and Julia [9, 10] that maximal supergravity on a torus T^n exhibits a chain of hidden symmetries described by Lie groups $G_{D-n}(\mathbb{R})$ in their split real form. The moduli space of scalars in lower dimensions is then given by the coset space $\mathcal{M}_{D-n} = G_{D-n}(\mathbb{R})/K$, where $K = K(G_{D-n})$ is the maximal compact subgroup. For example, in eight dimensions one finds $G_8(\mathbb{R}) = SL(3,\mathbb{R}) \times SL(2,\mathbb{R})$, while in six dimensions the global symmetry group is $SO(5,5)$. These groups are referred to as "hidden symmetries" because they are always larger than the naive isometry $GL(n,\mathbb{R})$ of the internal torus. The extra structure which is responsible for the enhancement of the symmetry arises from the dualisation of Ramond-Ramond and NS-NS p-form fields as we descend in dimension. Remarkably, starting in $D = 5$ these global symmetries are given by the exceptional Lie groups $\mathcal{E}_n(\mathbb{R})$ [10]. In particular, for dimensional reduction to $D = 3$, all vector fields may be dualized into scalars and the theory is globally invariant under the largest exceptional Lie group $\mathcal{E}_8(\mathbb{R})$.

The groups $G_{D-n}(\mathbb{R})$ are exact symmetries only of the *classical* effective action in $D-n$ dimensions, while the embedding into string/M-theory breaks the symmetry. This fact can be understood quite generally by the following argument [1, 8]. In addition to the moduli space of scalars $G_{D-n}(\mathbb{R})/K$, the reduced theories contain abelian vector fields arising from the reduction of the D-dimensional metric and the various p-form fields. The charges associated with these vector fields transform in a representation of G_{D-n} (see, e.g. [11]). However, through the embedding in string theory, these charges arise from the couplings of the vector fields to the basic objects in the theory, i.e. strings and D-branes, and are therefore subject to Dirac-Zwanziger quantization conditions. One then deduces that a continuous symmetry group $G_{D-n}(\mathbb{R})$ can never be an exact symmetry of string theory since it does not preserve the lattice of charges [8].

Then what is the fate of the classical symmetry groups $G_{D-n}(\mathbb{R})$ in string theory? It has been conjectured by Hull and Townsend that these are broken to discrete subgroups $G_{D-n}(\mathbb{Z}) \subset G_{D-n}(\mathbb{R})$, and that the full string/M-theory remains invariant under $G_{D-n}(\mathbb{Z})$ [8]. This is known as *U-duality*, and provides one of the cornerstones for the consistency of the string/M-theory duality web [1, 8] (see also [11]). In the quantum theory, the exact moduli space is then given by

$$\mathcal{M}^{\text{exact}}_{D-n} = G_{D-n}(\mathbb{Z})\backslash G_{D-n}(\mathbb{R})/K. \tag{1.2.1}$$

Let us discuss this in some more detail for the case of compactifications to $D = 4$, where the classical symmetry is given by the exceptional group $\mathcal{E}_7(\mathbb{R})$, and the conjectured U-duality group is $\mathcal{E}_7(\mathbb{Z})$. We want to understand the structure of this group from the general arguments above. In $D = 4$, all p-form fields can be dualized to axionic scalars or to abelian vector fields. For the case at hand, the bosonic sector of the four-dimensional effective action contains, in addition to the metric, 70 scalar fields parametrizing the coset space $\mathcal{E}_7(\mathbb{R})/(SU(8)/\mathbb{Z}_2)$, together with 28 Maxwell fields $\mathcal{A}^I_\mu$. The electric and magnetic charges (q^I, p_I) associated with these vector fields transform in a 56-dimensional representation of the Lie algebra E_7. However, as argued above, in the quantum theory these charges are subject to Dirac-Zwanziger charge quantization. More specifically, for two particles with electric-magnetic charges (q^I, p_I) and $(\tilde{q}^I, \tilde{p}_I)$ the quantization condition enforces the constraint

$$q^I \tilde{p}_I - \tilde{q}^I p_I \in \mathbb{Z}. \tag{1.2.2}$$

This implies that in the quantum theory the charges (q^I, p_I) span a 56-dimensional integral lattice which is invariant under the *electric-magnetic duality group* $Sp(56;\mathbb{Z})$. Based on these arguments, it was conjectured in [8] that the largest possible subgroup of $\mathcal{E}_7(\mathbb{R})$ that can be preserved in the quantum theory is defined by the subset of transformations which leave the electric-magnetic charge lattice invariant. In other words, the U-duality group $\mathcal{E}_7(\mathbb{Z})$ is *defined* by [8]

$$\mathcal{E}_7(\mathbb{Z}) := \mathcal{E}_7(\mathbb{R}) \cap Sp(56;\mathbb{Z}). \tag{1.2.3}$$

The U-duality conjecture implies that all couplings in the quantum effective action should be functions on $\mathcal{E}_7(\mathbb{R})/(SU(8)/\mathbb{Z}_2)$, which in addition are completely invariant under the discrete group $\mathcal{E}_7(\mathbb{Z})$ [12,13]. Such functions are known as *automorphic forms* and will play a leading role in **Part II** of this thesis. All quantum corrections to the effective action must respect the U-duality symmetry and this enforces powerful constraints which in principle are sufficient to determine the action exactly (see for instance [12–15]). Moreover, this statement is expected to hold in any dimension $4 \leq D \leq 11$, with the respective duality groups $G_{D-n}(\mathbb{Z})$ encoding the relevant quantum corrections in each dimension.

1.2.2 Duality in $\mathcal{N} = 2$ Theories

The fascinating structures that we have seen emerging above can be viewed as a consequence of supersymmetry. In ten dimensions, the type II theories exhibit $\mathcal{N} = 2$ supersymmetry, corresponding to 32 supercharges. Upon dimensional reduction on a torus T^n, all of these supercharges are preserved. For example, reduction on T^6 gives rise to $\mathcal{N} = 8$ supergravity in $D = 4$ [9,10]. It is this large amount of supersymmetry which constrains the classical moduli spaces of the reduced theories to be given by symmetric spaces $G_{D-n}(\mathbb{R})/K$. When taking more complicated internal manifolds X, many of the 32 supercharges are broken, giving rise to theories exhibiting a lesser amount of supersymmetry.

Of particular relevance for the topics treated in **Part II** of this thesis is the case when X is a Calabi-Yau threefold. This compactification breaks a quarter of the original 32 supercharges and thus preserves $\mathcal{N} = 2$ supersymmetry in four dimensions. Originally, Calabi-Yau compactifications were introduced in order to obtain four-dimensional theories containing chiral fermions [16,17]. These are also interesting examples for the purpose of probing the intricate duality web of string theory, because they provide an intermediate step between the simpler compactifications preserving $\mathcal{N} \geq 4$ supersymmetry in $D = 4$ – for which the moduli space is constrained to be a symmetric space G/K – and the difficult case of (more realistic) $\mathcal{N} = 1$ theories, for which quantum corrections are essentially unconstrained.

The classical limit of type II theories compactified on a Calabi-Yau threefold X is described by pure $\mathcal{N} = 2$ supergravity in four dimensions coupled to vector multiplets and hypermultiplets. One of the complications in describing this theory is the fact that the scalar fields no longer parametrize a symmetric space, but is rather given by a product [18,19]

$$\mathcal{M}_4 = \mathcal{M}_{\text{SK}} \times \mathcal{M}_{\text{QK}}, \tag{1.2.4}$$

where $\mathcal{M}_{\text{SK}}$ is a *special Kähler manifold*, parametrized by scalars belonging to vector multiplets, and $\mathcal{M}_{\text{QK}}$ is a *quaternionic-Kähler manifold*, parametrized by hypermultiplet scalars. The same structure appears both for the type IIA and type IIB theories, while the particular details of the effective actions are quite different. In type IIA compactifications, the moduli space $\mathcal{M}_{\text{SK}}^{(\text{A})}$ is $2h_{1,1}$-dimensional and encodes the (complexified) Kähler structure deformations of X, while the quaternionic manifold $\mathcal{M}_{\text{QK}}^{(\text{A})}$ is $4(h_{2,1}+1)$-dimensional and encodes the complex structure deformations of X. On the type IIB side the situation is reversed: the vector multiplet moduli space $\mathcal{M}_{\text{SK}}^{(\text{B})}$ is $2h_{2,1}$-dimensional, while the hypermultiplet moduli space $\mathcal{M}_{\text{QK}}^{(\text{B})}$ is $4(h_{1,1}+1)$-dimensional.

Although $\mathcal{N}=2$ supersymmetry does constrain the moduli space $\mathcal{M}_4$ to always split into a special Kähler and a quaternionic-Kähler manifold, this constraint is still not very strong: there is a zoo of possible quantum corrections to the four-dimensional effective action which deform the metric on the moduli space $\mathcal{M}_4$. Nevertheless, even when all quantum corrections are turned on, the decoupling between hypermultiplets and vector multiplets is preserved. This fact provides the basis for yet another powerful string duality: *mirror symmetry*. Consider again type IIA string theory on X. Then the mirror symmetry conjecture states that there exists a different *mirror* Calabi-Yau manifold $\tilde{X}$, such that the moduli spaces of type IIA and type IIB are related as follows [20, 21]

$$\mathcal{M}_{\text{SK}}^{(\text{A})}(X) \equiv \mathcal{M}_{\text{SK}}^{(\text{B})}(\tilde{X}), \qquad \mathcal{M}_{\text{QK}}^{(\text{A})}(X) \equiv \mathcal{M}_{\text{QK}}^{(\text{B})}(\tilde{X}), \tag{1.2.5}$$

which is a consequence of the fact that the complex structures and Kähler structures for a mirror pair $(X, \tilde{X})$ are interchanged: $h_{1,1}(X) = h_{2,1}(\tilde{X})$ and $h_{2,1}(X) = h_{1,1}(\tilde{X})$. We should emphasize that this relation between type IIA on X and type IIB on $\tilde{X}$ is not restricted to a relation between the moduli spaces: the statement of (generalized) mirror symmetry is that the complete theories, including quantum corrections, should be equivalent [20].

Recall now that type IIA and type IB are simply related by T-duality in $D=9$. Could there exist a version of this in lower dimensions, in addition to mirror symmetry? Indeed, there is such a duality which at the level of the moduli space is known as the *c-map* [22]. To reveal this, we start from the type IIA point of view and consider the further reduction on a circle to $D=3$.[1] Under this reduction, the moduli space $\mathcal{M}_{\text{QK}}^{(\text{A})}$ simply goes along unchanged, while the vector multiplet moduli space $\mathcal{M}_{\text{SK}}^{(\text{A})}$ gets enhanced due to the additional scalars arising from the reduction process. More precisely, in addition to the $2h_{1,1}$ scalar fields, the vector multiplets in $D=4$ also contain $h_{1,1}$ abelian vector fields $\mathcal{A}_\mu^I$, $I=1,\ldots,h_{1,1}$. Upon reduction on S^1 to $D=3$ these vector fields give rise to $h_{1,1}$ new scalars $\zeta^I \equiv \mathcal{A}_4^I$, while the three-dimensional vector fields $\mathcal{A}_\alpha^I$ can in turn be dualized into an additional set of $h_{1,1}$ scalars $\tilde{\zeta}_I$, giving rise to a total of $2h_{1,1}$ new scalar fields $(\zeta^I, \tilde{\zeta}_I)$. These scalars parametrize a $2h_{1,1}$-dimensional torus $T^{2h_{1,1}}$ which is fibered over the four-dimensional moduli space $\mathcal{M}_{\text{SK}}^{(\text{A})}$. Had we been considering the reduction of $D=4$, $\mathcal{N}=2$ supersymmetric *gauge theory*, this would have been the end of the story (see [24]).

However, in the present analysis, we are analyzing $\mathcal{N}=2$ supergravity, and therefore we must also take into account the reduction of the *gravity multiplet* in $D=4$. This corresponds simply to the four-dimensional metric $\text{g}_{\mu\nu}$ together with an additional abelian vector A_μ, known as the graviphoton. Similarly to the other vector fields, the reduction of A_μ gives rise

[1] For a nice discussion of the *c*-map, see [23].

to two axionic scalars ζ^0 and $\tilde{\zeta}_0$, while the reduction of the metric contributes with the radius of the circle e^U as well as a Kaluza-Klein vector $A_\alpha \equiv \mathrm{g}_{\alpha 4}$ which can be dualized into yet another axionic scalar σ. Together with the $2h_{1,1}$ scalars of $\mathcal{M}_{\mathrm{SK}}^{(\mathrm{A})}$, we thus conclude that the theory in $D = 3$ contains a total of $4(h_{1,1} + 1)$ scalar fields. Recall that this is precisely the dimension of the hypermultiplet moduli space $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{B})}$ of type IIB string theory compactified on the *same* Calabi-Yau threefold X. In fact, the new moduli space in $D = 3$ is also a quaternionic manifold which precisely coincides with $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{B})}$. Microscopically, this relation is actually not unexpected: it arises from T-duality along the circle S^1, in particular mapping the (inverse) radius e^U in type IIA on $X \times S^1$ to the string coupling $e^{\phi_{(\mathrm{B})}}$ in type IIB on $X \times \tilde{S}^1$. In a similar way, one may deduce that the vector multiplet moduli space $\mathcal{M}_{\mathrm{SK}}^{(\mathrm{B})}$ is mapped, via the c-map, to the hypermultiplet moduli space $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{A})}$ of type IIA on X. This is therefore in strong contrast to the case of mirror symmetry, which, as we have seen above, interchanges the vector- and hypermultiplet moduli spaces between type IIA and type IIB on *different* Calabi-Yau threefolds.

The analysis above has revealed that there is a fascinating duality structure within the type II sector of string theory also for compactifications on more complicated manifolds than tori. But there appears to be one piece of the puzzle missing: Where is the relation with M-theory? Previously we have learned that boosting the type IIA string coupling $g_s^{(\mathrm{A})}$ to infinity takes us to eleven-dimensional supergravity. So what happens if we do the same in $D = 4$? To answer this question, we must first note an important point. After the compactification on X the dilaton modulus $e^{\phi_{(\mathrm{A})}}$ belongs to a hypermultiplet, and therefore parametrizes the quaternionic moduli space $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{A})}$. This is, however, not quite the whole truth because for the decoupling between the moduli spaces (1.2.4) to work out, one must rescale the ten-dimensional dilaton by the volume of the internal manifold, and define a new four-dimensional dilaton by $e^{-2\phi_4} \equiv \mathrm{Vol}(X) e^{-2\phi_{(\mathrm{A})}}$. The string coupling constant in $D = 4$ is then defined with respect to the rescaled dilaton $g_s^{(4)} = e^{\phi_4}$, and the strong-coupling limit corresponds to $g_s^{(4)} \to \infty$.

Now consider M-theory compactified on the same Calabi-Yau manifold X. This gives rise to $\mathcal{N} = 1$ supergravity in *five* dimensions, with a hypermultiplet moduli space taking the same form as $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{A})}$ in (1.2.4), although with a different interpretation of the scalar fields. Upon further reduction on S^1 to $D = 4$ this moduli space goes along unchanged and the type IIA dilaton $e^{\phi_{(\mathrm{A})}}$ is again interpreted as the radius of the circle. Hence the statement is that the strong-coupling limit $g_s^{(4)} \to \infty$ of type IIA on X should be interpreted as blowing up the M-theory circle $e^{\phi_{(\mathrm{A})}} \to \infty$ while keeping the volume of X fixed. To conclude, in the strong-coupling limit, type IIA on X is equivalent to M-theory on $X \times S^1$, in analogy with the 10/11-dimensional duality.

1.2.3 Instanton Effects and String Dualities

Because of the lesser amount of supersymmetry preserved, quantum corrections to the four-dimensional effective action in Calabi-Yau compactifications are considerably less constrained compared to the toroidal case. As we have seen, the string coupling g_s belongs to the quaternionic hypermultiplet moduli space $\mathcal{M}_{\mathrm{QK}}$ both in type IIA and type IIB. This implies that the hypermultiplet moduli space is sensitive to quantum corrections associated with the perturbative power expansion in g_s corresponding worldsheets of higher genus. Since we have seen that supersymmetry enforces a complete decoupling between the two moduli spaces (1.2.4) these quantum corrections will only affect the hypermultiplet moduli space $\mathcal{M}_{\mathrm{QK}}$. For the vector multiplet moduli space, on the other hand, the situation is different in type IIA and type IIB. In type IIA, the vector multiplet scalars are associated with Kähler structure deformations of X, and the associated moduli space $\mathcal{M}_{\mathrm{SK}}^{(\mathrm{A})}$ only receives quantum corrections with respect to the perturbative expansion in the worldsheet coupling α', while being completely insensitive to the corrections in the string coupling g_s. In type IIB, the Kähler structure deformations are encoded in the hypermultiplet moduli space $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{B})}$, which therefore also receives α'-corrections in addition to the corrections in g_s. Thus, because all quantum corrections on the IIB side affect the quaternionic manifold $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{B})}$, the vector multiplet moduli space $\mathcal{M}_{\mathrm{SK}}^{(\mathrm{B})}$ is tree-level exact both in α' and in g_s (see [25] for a nice review).

In addition, the metric on $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{A/B})}$ receives non-perturbative corrections which are exponentially suppressed of order e^{-1/g_s} in the weak-coupling limit. These non-perturbative effects arise from Euclidean D-branes whose worldvolume wraps cycles in the Calabi-Yau manifold [26]. On the type IIA side, such effects are attributed to Euclidean D2-branes wrapping special Lagrangian submanifolds in X, while in type IIB they arise from all Euclidean Dp-branes, with $p = -1, 1, 3, 5$, wrapping even cycles in X. In addition, both in type IIA and type IIB there are effects from Euclidean NS5-branes wrapping the entire Calabi-Yau threefold. Such NS5-brane instantons behave differently compared to the D-brane instantons since the tension of the NS5-brane scales as $T_{\mathrm{NS5}} \sim g_s^{-2}$ while for a D-brane we have $T_{\mathrm{D}p} \sim g_s^{-1}$. This implies that NS5-brane instanton effects are exponentially suppressed of order e^{-1/g_s^2} and are therefore subdominant in the weak-coupling limit compared to D-brane instantons.

To compute the complete quantum corrected metric on the hypermultiplet moduli spaces $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{A/B})}$ has long been an outstanding problem. There is generically no analogue in $\mathcal{N} = 2$ theories of the U-duality groups $G(\mathbb{Z})$ which proved to be so powerful for toroidal compactifications. Nonetheless, considerable progress has been made recently by assuming that the $SL(2,\mathbb{Z})$-invariance of type IIB string theory should be unaffected by the compactification. Enforcing this constraint directly in the four-dimensional effective action made it possible to sum up all instanton corrections to the moduli space $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{B})}$ due to D(-1) and D1-instantons [27]. Using the mirror map (1.2.5) between the IIA and IIB moduli spaces, it was also possible to sum up the contributions to the moduli space $\mathcal{M}_{\mathrm{QK}}^{(\mathrm{A})}$ due to a subset of the D2-brane instantons [28]. Additional progress has also been made recently [29–31] using twistor techniques, and this will be discussed in more detail in Chapter 12. In Chapter 12 we will also see that for certain special cases of Calabi-Yau threefolds, known as *rigid*, there is an appealing candidate for a larger discrete group than $SL(2,\mathbb{Z})$ that might potentially constrain completely also the NS5-brane instantons.

1.3 Hyperbolic Billiards and Infinite-Dimensional Dualities

1.3.1 Infinite-Dimensional U-Duality?

The observant reader may have noticed an apparent lack of continuity in the previous section. If the various compactifications of the type II string theories reveal such interesting structures and dualities, why stop the reduction process in $D = 3$? The naive answer is: No reason! However, below $D = 3$ things start to become slightly more subtle. As we have seen in the previous section, the dimension $D = 3$ is special since it is the first dimension where all p-form fields, originating from ten or eleven dimensions, may be dualized into scalars. In the context of Calabi-Yau compactifications, this feature was responsible for the enhancement of the $2n$-dimensional special Kähler moduli space $\mathcal{M}_{\text{SK}}$ to a quaternionic $4(n+1)$-dimensional manifold $\mathcal{M}_{\text{QK}}$.[2]

For the case of toroidal compactifications to $D = 4$ we learned that the classical moduli space is given by $\mathcal{E}_7(\mathbb{R})/(SU(8)/\mathbb{Z}_2)$. Recall that this four-dimensional theory also contained 28 Maxwell fields. Upon further reduction to $D = 3$ all of these 28 abelian vector fields may be dualized and give rise to 56 new axionic scalars. Moreover, as we have seen, the $D = 4$ metric yields a dilaton e^U and a Kaluza-Klein scalar σ in $D = 3$. These scalars always appear during the reduction from $D = 4$ to $D = 3$ and are independent of which four-dimensional theory we started with. Adding up all the scalars we find a total of $2 + 56 + 70 = 128$ scalar fields in $D = 3$. These parametrize the coset space $\mathcal{E}_8(\mathbb{R})/(Spin(16)/\mathbb{Z}_2)$, with the global symmetry described by the largest of the exceptional Lie groups $\mathcal{E}_8(\mathbb{R})$.

So what *does* happen if take the compactification below three-dimensions? It was in fact conjectured long ago by Julia [32,33] that the chain of hidden symmetries should persist, with the groups $\mathcal{E}_9(\mathbb{R}), \mathcal{E}_{10}(\mathbb{R})$ and $\mathcal{E}_{11}(\mathbb{R})$ appearing upon toroidal reduction to $D = 2, D = 1$ and $D = 0$, respectively. This is not as obvious as it might sound, because these groups are no longer finite-dimensional Lie groups, but rather infinite-dimensional generalisations of these, known as *Kac-Moody groups* (see [34]). This makes it very tricky to understand the statement that these should become symmetries in dimensions below $D = 3$. How can a theory with a finite number of degrees of freedom suddenly reveal an infinite-dimensional symmetry? This is indeed a crucial point which is not at all understood for the case of $D = 1$ and $D = 0$, while for reductions to two dimensions it is possible to make sense of the enhanced symmetry.

In $D = 3$ all p-form fields can be dualized to scalars, enhancing the symmetry from $\mathcal{E}_7(\mathbb{R})$ to $\mathcal{E}_8(\mathbb{R})$. It turns out that $D = 2$ also has a very special property, namely that *scalars are dual to scalars*, in the sense that any axionic scalar χ can be dualized into another axionic scalar $\tilde{\chi}$ through the relation $d\chi = \star d\tilde{\chi}$. The symmetry group $\mathcal{E}_8(\mathbb{R})$, which acts on the original scalar χ, has an induced non-linear action on the dual scalar $\tilde{\chi}$ through the duality relation $d\chi = \star d\tilde{\chi}$. Moreover, the duality can be applied an infinite number of times giving rise to an infinite number of dual axions. It is the infinitely many new axionic scalars which are responsible for enhancing the symmetry to $\mathcal{E}_9(\mathbb{R})$, with the total moduli space in $D = 2$ being the infinite-dimensional coset space $\mathcal{E}_9(\mathbb{R})/K(\mathcal{E}_9)$ [35]. This phenomenon was originally uncovered in the simpler setting of pure gravity in $D = 4$ by Geroch [36], in which case the relevant infinite-dimensional group was later identified as $SL(2,\mathbb{R})^+$ [37], corresponding to the *affine extension* of the three-dimensional duality group $G_3 = SL(2,\mathbb{R})$. For this reason, the infinite exceptional group $\mathcal{E}_9$ is sometimes called the *Geroch group*.

Similarly to $SL(2,\mathbb{R})^+$, the group $\mathcal{E}_9$ is an affine Kac-Moody group which can be obtained through an affine extension of $\mathcal{E}_8$, and is therefore often denoted $\mathcal{E}_8^+$. Concretely, the extension is performed by adjoining an extra node to the Dynkin diagram of $\mathcal{E}_8$ in a certain prescribed way which will be explained in detail in Chapter 2. The groups $\mathcal{E}_{10}$ and $\mathcal{E}_{10}$, which are conjectured to appear upon further reduction below $D = 2$, may in turn be constructed by extending the Dynkin diagram of $\mathcal{E}_8$ with one or two extra nodes, respectively [38]. For this

[2]Recall $n = h_{1,1}, h_{2,1}$ for type IIA, IIB, respectively.

reason, they are also denoted by $\mathcal{E}_{10} = \mathcal{E}_8^{++}$ and $\mathcal{E}_{11} = \mathcal{E}_8^{+++}$.

The affine $\mathcal{E}_9$-symmetry of type II supergravities in $D = 2$ is by now well-established, and may be attributed to the fact that any gravitational theory becomes completely integrable in two dimensions [35, 37]. However, it is still not understood *how*, or even *if*, the infinite Kac-Moody groups $\mathcal{E}_{10}$ and $\mathcal{E}_{11}$ are supposed to appear upon further reduction. Nevertheless, these groups encode interesting algebraic structures that might turn out to play important roles in the ultimate formulation of string/M-theory. An important difference compared to the finite Lie groups discussed in the previous section is that the invariant metric on the Cartan subgalgebra $\mathfrak{h}$ of the associated Lie algebras E_{10} and E_{11} is of Lorentzian signature. In contrast, all finite-dimensional Lie algebras have Killing forms which are Euclidean when restricted to the Cartan subalgebra. For this reason, the Kac-Moody algebras E_{10} and E_{11} belong to the class of *Lorentzian Kac-Moody algebras* [38].

For the purposes of this thesis, it is the Kac-Moody algebra E_{10} which will play a leading role. This belongs to the small subclass of Lorentzian Kac-Moody algebras known as *hyperbolic*. The nomenclature here refers to the fact that the Weyl group $\mathcal{W}(E_{10})$ leaves invariant the unit hyperboloid inside the Cartan subalgebra $\mathfrak{h} \subset E_{10}$, analogously to the Weyl groups of finite Lie algebras which leave invariant a sphere at infinity, and are a therefore known as *spherical*. If the conjecture that E_{10} appears as a symmetry of eleven-dimensional supergravity on T^{10} is true, then it begs the question of what the fate of this symmetry is when the theory is embedded into M-theory. Following the general arguments of Hull and Townsend [8], discussed extensively in the previous section, one would expect that the continuous symmetry $\mathcal{E}_{10}(\mathbb{R})$ be broken to some discrete subgroup $\mathcal{E}_{10}(\mathbb{Z})$, and that M-theory on T^{10} (or type II string theory on T^9) are invariant under this infinite U-duality group. In fact, an analysis of the moduli space of M-theory on T^{10} was initiated in [39], where it was verified that the moduli space has Lorentzian signature, signaling the underlying structure of E_{10}.[3] If these speculations are true, then we would expect that $\mathcal{E}_{10}(\mathbb{Z})$ encodes information about novel non-perturbative effects in string theory, such as Euclidean D8-branes wrapping the internal torus T^9.

1.3.2 Hyperbolic Weyl Groups and Cosmological Billiards

The discussion above fits well into the philosophy of Section 1.1 regarding *the principle of beauty*: the hyperbolic Kac-Moody algebra E_{10} is without a doubt a beautiful and intriguing mathematical object, and its suggestive appearance opens up new and largely unexplored territories which lie on the borderline between mathematics and physics. However, despite these suggestive speculations the facts remain clear: we do not yet understand the underlying role of $\mathcal{E}_{10}$ within string theory and M-theory.[4] Nevertheless, we shall now discuss a very different physical setting in which the hyperbolic Kac-Moody algebra E_{10} also makes a surprising appearance, giving yet another indication of its possible importance. More precisely, it turns out that the Weyl group of E_{10} controls the dynamics of eleven-dimensional supergravity close to a cosmological singularity [44]. To understand this statement, we shall begin by discussing some of the underlying ideas which form the basis of **Part I** of this thesis.

[3] Extensions of these ideas have also been pursued in [40].

[4] Other appearances of E_{10} in string theory are discussed in [41–43].

In the 1970's, Belinskii, Khalatnikov and Lifshitz (BKL) analyzed the generic behaviour of four-dimensional pure gravity close to a cosmological (spacelike) singularity [45–47]. Their conclusion was that in this limit the temporal dependence of the dynamical degrees of freedom dominate over the spatial dependence, creating an effective *decoupling of spatial points*. The approach to the singularity is described by a sequence of *Kasner epochs*, where the dynamics in each epoch is associated to a certain Kasner solution. It was also found that this oscillating behaviour is *chaotic*, implying that there is an infinite sequence of oscillations before one reaches the singularity. Subsequently, Misner showed that the BKL-dynamics at a generic spatial point can be recast in terms of the dynamics of an auxiliary particle moving in a bounded region of hyperbolic space, called *mixmaster behaviour* [48].

These results were later extended to higher-dimensional gravitational theories, and it was found that pure gravity exhibits chaotic BKL-oscillations only in dimensions $4 \leq D \leq 10$, while pure gravity in eleven dimensions ceases to be chaotic [49]. However, it it turns out that all supergravities arising as low-energy limits of string theory and M-theory do reveal chaotic dynamics in the "BKL-limit" [50]. The surprising appearance of chaos in eleven-dimensional supergravity is attributed to the additional 3-form field $C_{(3)}$ which is not present in pure gravity.

Following the original ideas of Misner, it has been shown that the dynamics in the BKL-limit for any gravitational theory can be recast into geodesic motion in a region of hyperbolic space, bounded by hyperplanes [51]. The auxiliary particle undergoes geometric reflections against these hyperplanes, thus giving rise to a billiard-type behaviour. From the point if view of the original BKL-analysis, free-flight motion of the particle represents a single Kasner solution, while a geometric reflection flips between two distinct Kasner solutions. The behaviour close to a spacelilke singularity has therefore been dubbed *cosmological billiards* [51]. Whether a certain theory exhibits chaotic behaviour is then equivalent to determining whether the region in which the billiard dynamics takes place is of finite volume or not. It is well-known that random motion in a finite-volume hyperbolic billiard is chaotic.

It turns out that all of these discoveries have a very elegant algebraic explanation [44]. The auxiliary hyperbolic space hosting the billiard may be identified with the Cartan subalgebra of a Lorentzian Kac-Moody algebra $\mathfrak{g}$, while the bounded region in which billiard ball is confined corresponds to the *fundamental Weyl chamber*. In this new interpretation, it becomes clear that the geometric reflections responsible for the sequence of Kasner oscillations in fact generate the Weyl group $\mathcal{W}$ of $\mathfrak{g}$. Hence, the dynamics in the BKL-limit is completely controlled by the Weyl group of a Lorentzian Kac-Moody algebra. This gives a powerful way of determining whether a given theory exhibits chaotic dynamics or not: if the Lorentzian Kac-Moody algebra $\mathfrak{g}$ is of *hyperbolic type*, then the fundamental Weyl chamber is of finite volume, and the theory is chaotic [52].

Remarkably, when performing this analysis for eleven-dimensional supergravity one finds that the algebra whose Weyl group controls the dynamics in the BKL-limit is precisely the hyperbolic Kac-Moody algebra E_{10}, thus explaining the appearance of chaos [44]. Moreover, it turns out that the Weyl group of E_{10} also appears for the type IIA and type IB supergravities, revealing that these three theories display identical dynamical behaviour in the BKL-limit. This is in perfect accordance with the fact that type IIA, type IIB as well as eleven-dimensional supergravity become equivalent upon reduction on a torus T^n as discussed above. We should also emphasize that the Weyl group $\mathcal{W}(E_{10})$ is in fact a subgroup of the conjectured U-duality group $\mathcal{E}_{10}(\mathbb{Z})$, thus suggesting that the BKL limit is an alternative way of "unveiling" a possible underlying symmetry of string theory. **Part I** of this thesis is to a large extent devoted to a careful study of the dynamics in the BKL-limit, with particular emphasis on the underlying algebraic structures.

1.3.3 Geodesic Motion on Infinite-Dimensional Coset Spaces

The particle dynamics in a finite-volume billiard is only sensitive to the hyperplanes which are "closest" to the particle geodesic. These are known as the *dominant walls*. From the algebraic point of view, this translates to the statement that the relevant Weyl-reflections governing the BKL-oscillations are only those with respect to the *simple roots* of the Kac-Moody algebra $\mathfrak{g}$, implying that there is a correspondence between dominant billiard walls and simple roots of $\mathfrak{g}$. However, generically all theories give rise to many more walls than just the dominant ones. Depending on the structure of the theory, there can be a number of different *subdominant walls* which play no role in the strict BKL-limit [51]. However, when going away from the strict BKL-limit it is expected that the dominant walls become less sharp and subdominant walls come into play. Another way to say this is that the complete decoupling of spatial points is no longer valid, and spatial dependence gradually reappears. From the algebraic point of view, the subdominant walls are associated with non-simple positive roots of the Kac-Moody algebra, and hence probes the possible hidden Kac-Moody symmetry $\mathfrak{g}$ beyond its Weyl group $\mathcal{W}(\mathfrak{g})$.

These ideas were put on a more concrete footing in the context of eleven-dimesional supergravity by Damour, Henneaux and Nicolai [53]. Inspired by the BKL-analysis discussed above, they constructed a manifestly $\mathcal{E}_{10}$-invariant non-linear sigma model for a particle moving on the infinite-dimensional coset space $\mathcal{E}_{10}/K(\mathcal{E}_{10})$.[5] To make sense of the coset space $\mathcal{E}_{10}/K(\mathcal{E}_{10})$, the Kac-Moody algebra E_{10} was sliced up into an infinite gradation, called a *level decomposition*, with finite-dimensional subspaces falling into representations of a distinguished $\mathfrak{sl}(10,\mathbb{R})$ subalgebra. At low levels this decomposition reveals tensorial representations matching the field content of eleven-dimensional supergravity, and it was shown that there is a dynamical equivalence between the geodesic motion on $\mathcal{E}_{10}/K(\mathcal{E}_{10})$ and a certain truncation of eleven-dimensional supergravity. The match works up to *height 30* on the algebraic side, and therefore probes the Kac-Moody algebra E_{10} far beyond its simple roots, which by definition all have height one.[6]

This correspondence between $\mathcal{E}_{10}$ and eleven-dimensional supergravity was later extended also to encompass type IIA [61] and type IIB supergravity [62].[7] In the second half of **Part I** of this thesis we will discuss the original example in great detail, and we will extend the

[5]It has also been conjectured by West that the Lorentzian Kac-Moody algebra E_{11} should correspond to a non-linearly realized hidden symmetry of eleven-dimensional supergravity, or possibly of M-theory [54, 55]. A lot of work has been done within the framework of this alternative proposal (see for example [56–60]). In the present work, we are mainly motivated by the cosmological billiard picture, and we will therefore not treat this point of view in more detail.

[6]The height of a root is defined as the sum of all the coefficients when this root is written as a linear combination of the simple roots. This is explained in more detail in Chapter 2.

[7]This type of construction has also been extended by Englert and Houart to $E_{11} = E_8^{+++}$, as well as to other Lorentzian Kac-Moody algebras of triple-extended type [63–65]. This framework is particularly powerful for analyzing the algebraic structure of BPS solutions in the context of maximal supergravities [66–68].

correspondence to the case of massive type IIA supergravity.

It has also been understood how to incorporate fermions into the correspondence between supergravity and infinite-dimensional sigma models. The fermionic degrees of freedom are associated with certain spinorial representations of the maximal compact subalgebra $K(E_{10})$ [69–72]. For exampe, in the case of eleven-dimensional supergravity, the gravitino arises as a 320-dimensional vector-spinor representation of $K(E_{10})$. This is an unfaithful representation of the infinite-dimensional algebra $K(E_{10})$, and it is an outstanding problem to understand the construction of non-trivial faithful representations which could match the infinite gradation involving the bosonic degrees of freedom. Because of this mismatch it has not yet been possible to construct a supersymmetric version of the non-linear sigma model on $\mathcal{E}_{10}/K(\mathcal{E}_{10})$.

With these words we end our tour through the duality web of type II string theory and supergravity. The topics discussed in this introduction will all play important roles in the analysis of **Part I** and **Part II** of this thesis.

Part I

Spacelike Singularities and Hyperbolic Structures in (Super-)Gravity

2

Kac-Moody Algebras

In this chapter we present the basic theory of Kac-Moody algebras, with emphasis on the affine and Lorentzian cases.[1] Our treatment is aimed towards physicists, and to this end we do not give formal definitions, theorems or proofs, but rather we introduce the reader to a "toolbox", whose constituents can, in principle, be mastered relatively quickly. We presuppose a working knowledge of the theory of finite simple Lie algebras, and explain in detail how these structures generalize to arbitrary infinite-dimensional Kac-Moody algebras. The techniques presented here play a central role in the remainder of this thesis, in particular in **Part I** but also to some extent in **Part II**. Recommended references for this chapter are [34, 73].

2.1 Preliminary Example: A_1 – The Fundamental Building Block

In this section we consider a simple "warm-up" example which nevertheless contains many of the important features encountered later on. Even though being the smallest finite-dimensional simple Lie algebra, A_1 plays an important role in the general theory of Kac-Moody algebras. In particular, one may view any simple rank r Kac-Moody algebra $\mathfrak{g}$ as a set of r distinct A_1-subalgebras which are intertwined in a non-trivial way. This fact is a cornerstone in the representation theory of Kac-Moody algebras [34]. For this reason, A_1 may be described as the fundamental building block of all Lie algebras, finite- as well as infinite-dimensional. All terminology introduced in this example will be properly defined in subsequent sections.

$A_1 \simeq \mathfrak{sl}(2, \mathbb{C})$ is the algebra of 2×2 complex traceless matrices. This algebra is 3-dimensional and we take as a basis the set of generators $\{T_i \mid i = 1, 2, 3\}$, subject to the commutation relations

$$[T_i, T_j] = -\varepsilon_{ijk} T_k, \tag{2.1.1}$$

[1] This chapter is based on lectures given by the author at the *Third Modave International Summer School on Mathematical Physics*, held in Modave, Belgium, August 2008.

with $\varepsilon_{123} = 1$. Writing this out we find the following relations between the generators:

$$[T_1, T_2] = -T_3, \qquad [T_1, T_3] = T_2, \qquad [T_2, T_3] = -T_1. \tag{2.1.2}$$

In the fundamental representation we have a matrix realization of the form

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad T_2 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad T_3 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \tag{2.1.3}$$

which can easily be seen to satisfy (2.1.2). For the purposes of this thesis, it is useful to switch to another basis, where the three basis elements take the form

$$\begin{aligned} e &\equiv & T_2 - iT_1 &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ f &\equiv & -(T_2 + iT_1) &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ h &\equiv & -2iT_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{2.1.4}$$

The new basis $\{e, f, h\}$ is similar to the familiar basis $\{J^+, J^-, J^3\}$ of $\mathfrak{su}(2)$. The commutation relations now become

$$[e, f] = h, \qquad [h, e] = 2e, \qquad [h, f] = -2f, \tag{2.1.5}$$

implying that the generators e and f can be thought of a "step operators", taking us, respectively, "up" and "down" between the lowest and highest weights of the representation. The basis $\{e, f, h\}$ is called the Chevalley basis, and (2.1.5) corresponds to a Chevalley presentation of A_1. There are two main reasons for working in this basis. Firstly, it is the starting point for the generalization to arbitrary Kac-Moody algebras which we shall consider in the next section. Secondly, in the Chevalley basis, the matrix realization of the generators only involves *real* traceless matrices. This ensures that simply by restricting all linear combinations of generators to the real numbers, we obtain a *real* Lie algebra, namely the split real form $\mathfrak{sl}(2, \mathbb{R})$, consisting of 2×2 real traceless matrices.

The Chevalley presentation reveals a natural decomposition of $\mathfrak{sl}(2, \mathbb{R})$ as the following direct sum of vector spaces

$$\mathfrak{sl}(2, \mathbb{R}) = \mathbb{R}f \oplus \mathbb{R}h \oplus \mathbb{R}e. \tag{2.1.6}$$

This is the triangular decomposition of the Lie algebra, which in this case simply means that each matrix can be decomposed as a sum of a lower triangular, a diagonal and an upper triangular matrix. The two subspaces $\mathbb{R}f$ and $\mathbb{R}e$ form nilpotent subalgebras and $\mathbb{R}h$ form the abelian Cartan subalgebra. Of course, these concepts are all somewhat trivial in this example, but it is nevertheless useful to employ this terminology for later use.

An important concept which we shall encounter numerous times is that of a maximal compact subalgebra. It is well known that the maximal compact subalgebra of $\mathfrak{sl}(2, \mathbb{R})$ is $\mathfrak{so}(2)$, the algebra of 2×2 (traceless) antisymmetric matrices. Let us now try to understand this from a more abstract point of view. The algebra $\mathfrak{so}(2) \subset \mathfrak{sl}(2, \mathbb{R})$ is a so-called involutory subalgebra, meaning that there exists an involution ω of $\mathfrak{sl}(2, \mathbb{R})$ such that $\mathfrak{so}(2)$ coincides with the set $\{x \in \mathfrak{sl}(2,\mathbb{R}) \mid \omega(x) = x\}$, which is pointwise fixed by ω. The involution ω is called the Chevalley involution and is defined on the Chevalley generators as follows

$$\omega(e) = -f, \qquad \omega(f) = -e, \qquad \omega(h) = -h. \tag{2.1.7}$$

This is obviously an involution, $\omega^2 = 1$, and it leaves the commutation relations invariant, so ω is an automorphism of $\mathfrak{sl}(2,\mathbb{R})$.

Now note that the combination $e - f$ is fixed by ω and so spans the one-dimensional maximal compact subalgebra of $\mathfrak{sl}(2,\mathbb{R})$. Indeed we have

$$e - f = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathfrak{so}(2), \tag{2.1.8}$$

so that we may write $\mathfrak{so}(2) = \mathbb{R}(e-f)$. The Chevalley involution thus induces another decomposition of $\mathfrak{sl}(2,\mathbb{R})$ into pointwise invariant and anti-invariant subsets under the action of ω. Explicitly, this yields the Cartan decomposition

$$\mathfrak{sl}(2,\mathbb{R}) = \mathbb{R}(e-f) \oplus \Big(\mathbb{R}h \oplus \mathbb{R}(e+f)\Big). \tag{2.1.9}$$

It is important to note that this is a direct sum of vector spaces, and, in particular, that the anti-invariant part $\mathfrak{p} \equiv \mathbb{R}h \oplus \mathbb{R}(e+f)$ is *not* a subalgebra. This is in contrast to the triangular decomposition above, for which each subspace is a subalgebra in itself. It is easy to see that $\mathfrak{p}$ does not form a subalgebra by noticing that the commutation relations do not close,

$$[h, e+f] = 2(e-f) \in \mathfrak{so}(2). \tag{2.1.10}$$

Moreover we have

$$[h, e-f] = 2(e+f) \in \mathfrak{p}, \qquad [e-f, e+f] = 2h \in \mathfrak{p}, \tag{2.1.11}$$

revealing that the algebraic structure of the decomposition is

$$[\mathfrak{p},\mathfrak{p}] \subset \mathfrak{so}(2), \qquad [\mathfrak{so}(2),\mathfrak{p}] \subset \mathfrak{p}, \qquad [\mathfrak{so}(2),\mathfrak{so}(2)] \subset \mathfrak{so}(2), \tag{2.1.12}$$

indicating that at the "group level" $\mathfrak{p}$ corresponds to a symmetric space, which in this case coincides with the coset space $SL(2,\mathbb{R})/SO(2)$. We shall come back to this in Section 2.2.3.

Finally, we consider an additional useful decomposition, known as the Iwasawa decomposition, which will play an important role in what follows. It takes the form (direct sum of vector spaces)

$$\mathfrak{sl}(2,\mathbb{R}) = \mathfrak{so}(2) \oplus \mathbb{R}h \oplus \mathbb{R}e, \tag{2.1.13}$$

where each subspace is now a subalgebra. Note that in this decomposition, even though $\mathfrak{so}(2)$ is defined as before through the Chevalley involution, the second part, $\mathbb{R}h \oplus \mathbb{R}e$ (the Borel subalgebra) is not anti-invariant under ω.

2.2 Basic Definitions

Kac-Moody algebras are infinite-dimensional generalizations of the finite simple Lie algebras, as classified by Cartan and Killing. A standard treatment of an l-dimensional Lie algebra is in terms of a set of generators $\{T_m \mid m = 1, \ldots, l\}$, subject to the commutation relations

$$[T_m, T_n] = f_{mn}{}^p T_p, \tag{2.2.1}$$

where the structure constants $f_{mn}{}^p$ contain all the information of the algebra. This construction is not very convenient if we want to generalize it to cases when $l \to \infty$. It is for this reason that we in the previous section emphasized the importance of the Chevalley-Serre presentation, a construction which is amenable for generalization. We turn now to discuss the basic properties of Kac-Moody algebras, using the Chevalley-Serre presentation.

2.2.1 The Chevalley-Serre Presentation

Let (e_i, f_i, h_i) be a triple of generators, satisfying the commutation relations of $\mathfrak{sl}(2, \mathbb{R})$:

$$[e_i, f_i] = h_i, \qquad [h_i, e_i] = 2e_i, \qquad [h_i, f_i] = -2f_i. \tag{2.2.2}$$

We can now construct an algebra $\tilde{\mathfrak{g}}$ by letting the index i run from 1 to r and "intertwining" the r copies of $\mathfrak{sl}(2, \mathbb{R})$ through the following relations

$$\begin{aligned}
[e_i, f_j] &= \delta_{ij} h_j, \\
[h_i, e_j] &= A_{ij} e_j, \\
[h_i, f_j] &= -A_{ij} f_j, \\
[h_i, h_j] &= 0.
\end{aligned} \tag{2.2.3}$$

It is important to note that here r is always finite, even if the algebra itself might be infinite-dimensional. The structure of the algebra $\tilde{\mathfrak{g}}$ is encoded in the *Cartan matrix* A, whose entries, A_{ij}, determine the commutation relations between the generators of the different $\mathfrak{sl}(2, \mathbb{R})$-subalgebras. We shall discuss the Cartan matrix in more detail below, and for now we just impose that it be non-degenerate, $\det A \neq 0$. This condition will be lifted later on. We shall often emphasize the dependence of $\tilde{\mathfrak{g}}$ on the Cartan matrix, and write $\tilde{\mathfrak{g}}(A)$. From now on we also fix the base field of $\tilde{\mathfrak{g}}$ to $\mathbb{R}$.

Further elements of $\tilde{\mathfrak{g}}$ are obtained by taking multiple commutators as follows

$$[e_{i_1}, [e_{i_2}, \ldots, [e_{i_{k-1}}, e_{i_k}] \cdots]], \qquad [f_{i_1}, [f_{i_2}, \ldots, [f_{i_{k-1}}, f_{i_k}] \cdots]]. \tag{2.2.4}$$

Note that at this point the algebra $\tilde{\mathfrak{g}}$ is infinite-dimensional because there are no relations between the e_i or f_i that restrict these multicommutators.

Through the use of the Chevalley relations, (2.2.3), any commutator involving the h_i may be reduced to one of the form of (2.2.4). This gives rise to the so-called *triangular decomposition* of $\tilde{\mathfrak{g}}$, which takes the form (direct sum of vector spaces)

$$\tilde{\mathfrak{g}} = \tilde{\mathfrak{n}}_- \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}_+. \tag{2.2.5}$$

Here, $\mathfrak{h}$ is a real vector space spanned by the h_i,

$$\mathfrak{h} = \sum_{i=1}^{r} \mathbb{R} h_i, \tag{2.2.6}$$

which forms an abelian subalgebra of $\tilde{\mathfrak{g}}$, called the *Cartan subalgebra*. The subspaces $\tilde{\mathfrak{n}}_+$ and $\tilde{\mathfrak{n}}_-$ are freely generated by the e_i's and the f_i's, respectively. The algebra $\tilde{\mathfrak{g}}$ is not simple, but has a maximal ideal $\mathfrak{i}$, which decomposes as a direct sum of ideals,

$$\mathfrak{i} = (\mathfrak{i} \cap \tilde{\mathfrak{n}}_-) \oplus (\mathfrak{i} \cap \tilde{\mathfrak{n}}_+) \equiv \mathfrak{i}_- \oplus \mathfrak{i}_+, \tag{2.2.7}$$

where the two subspaces $\mathfrak{i}_+$ and $\mathfrak{i}_-$ are ideals in $\tilde{\mathfrak{n}}_+$ and $\tilde{\mathfrak{n}}_-$, respectively. It follows that we have

$$\mathfrak{i}_\pm \cap \mathfrak{h} = 0, \qquad \mathfrak{i}_+ \cap \mathfrak{i}_- = 0. \tag{2.2.8}$$

The two ideals $\mathfrak{i}_\pm$ are generated by the subsets of elements $S_\pm$ given by

$$\begin{aligned} S_+ &= \left\{ \mathrm{ad}_{e_i}^{1-A_{ij}}(e_j) \;\middle|\; i \neq j,\ i,j = 1, \dots, r \right\}, \\ S_- &= \left\{ \mathrm{ad}_{f_i}^{1-A_{ij}}(f_j) \;\middle|\; i \neq j,\ i,j = 1, \dots, r \right\}, \end{aligned} \tag{2.2.9}$$

where "ad" denotes the adjoint action, i.e., for $x, y \in \tilde{\mathfrak{g}}$, $\mathrm{ad}_x(y) = [x, y]$. We shall now take the quotient of $\tilde{\mathfrak{g}}(A)$ by the ideal $\mathfrak{i}$. This gives rise to relations among the generators of $\tilde{\mathfrak{n}}_+$ and $\tilde{\mathfrak{n}}_-$. We thereby define the *Kac-Moody algebra* $\mathfrak{g}(A)$, associated with the Cartan matrix A, as follows

$$\mathfrak{g}(A) = \tilde{\mathfrak{g}}(A) \;/\; \mathfrak{i}. \tag{2.2.10}$$

Since we have chosen the base field to be $\mathbb{R}$, this defines the split real form of the corresponding Kac-Moody algebra over $\mathbb{C}$.

The rank r Kac-Moody algebra $\mathfrak{g}(A)$ is now generated by the $3r$ generators e_i, f_i, h_i subject to the Chevalley relations, (2.2.3), and the *Serre relations*,

$$\begin{aligned} \mathrm{ad}_{e_i}^{1-A_{ij}}(e_j) &= [e_i, [e_i, \dots, [e_i, e_j] \cdots]] = 0, \\ \mathrm{ad}_{f_i}^{1-A_{ij}}(f_j) &= [f_i, [f_i, \dots, [f_i, f_j] \cdots]] = 0, \end{aligned} \tag{2.2.11}$$

with each relation containing $1 - A_{ij}$ commutators. The triangular decomposition of $\mathfrak{g}(A)$ then reads

$$\mathfrak{g}(A) = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+, \tag{2.2.12}$$

where

$$\mathfrak{n}_\pm = \tilde{\mathfrak{n}}_\pm \;/\; \mathfrak{i}_\pm. \tag{2.2.13}$$

The Serre relations thus impose restrictions on $\mathfrak{n}_\pm$ which cut the chains of multiple commutators involving e_i and f_i. These restrictions might, or might not, render the algebra $\mathfrak{g}(A)$ finite-dimensional. We shall see in the next section how this depends on the properties of the Cartan matrix.

2.2.2 The Cartan Matrix and Dynkin Diagrams

So far we have discussed how the algebra $\mathfrak{g}(A)$ is constructed by imposing relations between the Chevalley generators e_i, f_i, h_i. These relations are completely determined by the entries of the matrix A, an important object which we shall now discuss more closely.

An $r \times r$ matrix $A = (A_{ij})_{i,j=1,\ldots,r}$ is called a *generalized Cartan matrix* if it satisfies the following properties:

$$\begin{aligned} A_{ii} &= 2, \quad i = 1, \ldots, r, \\ A_{ij} &= 0 \Leftrightarrow A_{ji} = 0, \\ A_{ij} &\in \mathbb{Z}_- \quad (i \neq j). \end{aligned} \tag{2.2.14}$$

For brevity, we shall in the following refer to A simply as a Cartan matrix. The Cartan matrix is called *indecomposable* if the index set $\mathcal{S} = \{1, \ldots, r\}$ can not be divided into two non-empty subsets $\mathcal{I}$ and $\mathcal{J}$ such that $A_{ij} = 0$ for $i \in \mathcal{I}$ and $j \in \mathcal{J}$. An important statement is then the following: *when the Cartan matrix A is non-degenerate, det $A \neq 0$, and indecomposable, the Kac-Moody algebra $\mathfrak{g}(A)$ is simple* [34]. In the following we shall always assume that A is indecomposable.

It is now possible to provide a (partial) classification of the various types of algebras $\mathfrak{g}(A)$ that can be constructed from a Cartan matrix. There exist three main classes:

- If A is positive definite, the algebra $\mathfrak{g}(A)$ is finite-dimensional and falls under the Cartan-Killing classification, i.e., it is one of the finite simple Lie algebras $A_n, B_n, C_n, D_n, G_2, F_4, E_6, E_7$ or E_8.

- If A is positive-semidefinite, i.e., det $A = 0$ with one zero eigenvalue, the algebra is infinite-dimensional and is said to be an *affine* Kac-Moody algebra.[2] All affine Kac-Moody algebras are classified [34].

- If A is not part of the two classes above, the algebra $\mathfrak{g}(A)$ is infinite-dimensional and is generally called an *indefinite* Kac-Moody algebra, by virtue of the fact that A is of indefinite signature.

For the third class above, no general classification exists. We shall, however, mainly be interested in a subclass of the indefinite Kac-Moody algebras, corresponding to the case when the matrix A has one negative eigenvalue and $r-1$ positive eigenvalues. The associated Kac-Moody algebras are called *Lorentzian*, because of the signature $(- + + \cdots + +)$ of A. A special subclass of the Lorentzian algebras, known as *hyperbolic* Kac-Moody algebras, have in fact been classified. We shall define hyperbolic Kac-Moody algebras in Section 2.4.2.

Since most of the entries of the Cartan matrix are zero, it is convenient to encode the non-vanishing entries in a diagram, $\Gamma = \Gamma(A)$, called a *Dynkin diagram*. To this end, we associate a node $\circ$ in Γ to each Chevalley triple (e_i, f_i, h_i), and if $A_{ij} \neq 0$ for $i \neq j$ the nodes i and j are connected by $\max(|A_{ij}|, |A_{ji}|)$ lines. In addition, when $|A_{ij}| > |A_{ji}|$ we draw an arrow from node j to node i. Indecomposability of the Cartan matrix A is equivalent to the statement that the Dynkin diagram $\Gamma(A)$ is connected.

[2]Strictly speaking, in the case of det $A = 0$ the algebra constructed from the Chevalley-Serre relations only corresponds to the *derived* algebra $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$. We shall come back to this issue in Section 2.3.2.

Let us now discuss some simple examples in order to illustrate the relation between the Kac-Moody algebra $\mathfrak{g}(A)$, its Cartan matrix A and the associated Dynkin diagram $\Gamma(A)$. We begin with the simplest possible case, namely the Lie algebra $A_1 = \mathfrak{sl}(2, \mathbb{R})$, discussed at length in Section 2.1. Here there is only one Chevalley triple, (e, f, h), and consequently the Cartan matrix is just the number (2), with Dynkin diagram consisting of one node $\circ$. The Lie algebra $A_2 = \mathfrak{sl}(3, \mathbb{R})$, in turn, is described by the Cartan matrix

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

and Dynkin diagram $\circ\!\!-\!\!-\!\!\circ$. This corresponds to two copies of $\mathfrak{sl}(2, \mathbb{R})$ which are intertwined through the non-vanishing off-diagonal components of the Cartan matrix. In contrast, if $A_{12} = A_{21} = 0$ we have a direct sum of Lie algebras $A_1 \oplus A_1$ corresponding to the decomposable Cartan matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

with Dynkin diagram $\circ\ \ \circ$. This algebra is not simple, since the two A_1's constitute two non-trivial ideals. Later on we shall discuss these examples, and more involved ones, in more detail.

2.2.3 The Root System and the Root Lattice

A very important notion in the theory of Kac-Moody algebras is that of a *root*. In this section we shall develop the basic theory of roots and examine the vector space which they span, hopefully convincing the reader that these issues are extremely useful for a deeper understanding of Kac-Moody algebras.

Let us begin by noting that, by virtue of the Chevalley relations, (2.2.3), the adjoint action of $\mathfrak{h}$ on $\mathfrak{n}_{\pm}$ is diagonal,

$$\mathrm{ad}_h(e_i) = [h, e_i] = \alpha_i(h) e_i, \qquad h \in \mathfrak{h}, \tag{2.2.15}$$

and similarly for the action on f_i. The eigenvalue $\alpha_i(h)$ represents the value of a linear map from $\mathfrak{h}$ to the real numbers,

$$\alpha_i \ : \ \mathfrak{h} \ \ni \ h \ \longmapsto \ \alpha_i(h) \ \in \ \mathbb{R}. \tag{2.2.16}$$

The linear maps α_i are called *simple roots* and belong to the dual space $\mathfrak{h}^\star$. We shall sometimes employ the notation $\langle \alpha, h \rangle = \alpha(h)$ for the pairing between a form $\alpha \in \mathfrak{h}^\star$ and a vector $h \in \mathfrak{h}$. If the eigenvalue $\alpha(h)$ vanishes, then α is not a root. It is also common to refer to the Cartan generators $h_i$as *simple coroots* to emphasize that they belong to the dual of the space of roots. In this case one also writes $\alpha_i^\vee \equiv h_i$. The same analysis can be performed for multiple commutators, e.g,

$$\begin{aligned} [h, [e_i, e_j]] &= -[e_i, [e_j, h]] - [e_j, [h, e_i]] \\ &= (\alpha_i + \alpha_j)(h)[e_i, e_j], \end{aligned} \tag{2.2.17}$$

where in the first line we made use of the Jacobi identity. If the generator $e_\alpha \equiv [e_i, e_j]$ is non-vanishing, i.e., is not killed by the Serre relations, then $\alpha \equiv \alpha_i + \alpha_j$ is the root associated with e_α. We denote by $\Pi = \{\alpha_1, \ldots, \alpha_r\}$ the basis of simple roots and by $\Pi^\vee = \{\alpha_1^\vee, \ldots, \alpha_r^\vee\}$ the basis of simple coroots. Any root can be expressed as an integer linear combination of the simple roots. We denote by Φ the complete set of roots. This is called the *root system*. In analogy with (2.2.6) we also have

$$\mathfrak{h}^\star = \sum_{i=1}^{r} \mathbb{R}\alpha_i. \tag{2.2.18}$$

A root is called *positive* (*negative*) if it can be written as a linear combination of the simple roots Π with only non-negative (non-positive) coefficients. From the triangular decomposition it follows that all roots are either positive or negative. Thus, the root system, Φ, splits into a disjoint union of positive and negative roots,

$$\Phi = \Phi_+ \cup \Phi_-. \tag{2.2.19}$$

For $\mathfrak{g} \ni x_\alpha \neq 0$ the associated root

$$\alpha = \sum_{i=1}^{r} m_i \alpha_i \tag{2.2.20}$$

belongs to Φ_+ if all $m_i \in \mathbb{Z}_{\geq 0}$ and to Φ_- if all $m_i \in \mathbb{Z}_{\leq 0}$. Let us for definiteness take α to be a positive root. Then $-\alpha$ is necessarily a negative root, and we write

$$x_\alpha \equiv e_\alpha \in \mathfrak{n}_+, \tag{2.2.21}$$
$$x_{-\alpha} \equiv f_\alpha \in \mathfrak{n}_-. \tag{2.2.22}$$

In the Chevalley basis the eigenvalues $\alpha_i(h)$ are always integers, called the *Cartan integers*, revealing that the set of roots Φ lie on an r-dimensional lattice Q spanned by the simple roots,

$$Q = \sum_{i=1}^{r} \mathbb{Z}\alpha_i \subset \mathfrak{h}^\star. \tag{2.2.23}$$

All elements of the root system thus belong to Q but the converse is not true, hence

$$\Phi \subset Q. \tag{2.2.24}$$

In the Cartan subalgebra $\mathfrak{h}$ we similarly have the dual notion of a *coroot lattice*:

$$Q^\vee = \sum_{i=1}^{r} \mathbb{Z}\alpha_i^\vee \subset \mathfrak{h}. \tag{2.2.25}$$

We may now decompose the algebra $\mathfrak{g}$ into disjoint subsets $\mathfrak{g}_\alpha \subset \mathfrak{g}$, where each subset is spanned by those generators $x \in \mathfrak{g}$, whose eigenvalue under the action of $h \in \mathfrak{h}$ is given by $\alpha(h)$. This decomposition is called the *root space decomposition* and reads

$$\mathfrak{g} = \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha, \tag{2.2.26}$$

where the subspace $\mathfrak{g}_\alpha$ is the *root space* associated to the root α. Explicitly these are given by

$$\mathfrak{g}_\alpha = \big\{ x \in \mathfrak{g} \,\big|\, \forall\ h \in \mathfrak{h}\ :\ \mathrm{ad}_h(x) = \alpha(h)x \big\}. \tag{2.2.27}$$

Note that the zeroth subspace $\mathfrak{g}_0$ coincides with the Cartan subalgebra, $\mathfrak{g}_0 = \mathfrak{h}$. Because of the disjoint split $\Phi_+ \cup \Phi_-$ of the root system we can write the root space decomposition as follows

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Phi_+} \mathfrak{g}_\alpha \oplus \bigoplus_{\alpha\in\Phi_+} \mathfrak{g}_{-\alpha}. \tag{2.2.28}$$

The dimension of each subspace $\mathfrak{g}_\alpha$ is called the *multiplicity*, $\mathrm{mult}(\alpha)$, of the root α,

$$\mathrm{mult}(\alpha) \equiv \dim\, \mathfrak{g}_\alpha. \tag{2.2.29}$$

Thus, for a given root $\gamma \in \Phi$, with root space

$$\mathfrak{g}_\gamma = \mathbb{R}x_\gamma^{(1)} \oplus \mathbb{R}x_\gamma^{(2)} \oplus \cdots \oplus \mathbb{R}x_\gamma^{(k-1)} \oplus \mathbb{R}x_\gamma^{(k)}, \tag{2.2.30}$$

we have

$$\mathrm{mult}(\gamma) = k\ \in\ \mathbb{Z}_+\backslash\{0\}. \tag{2.2.31}$$

The root spaces corresponding to the simple roots are one-dimensional

$$\mathfrak{g}_{\alpha_i} = \mathbb{R}e_i, \qquad \mathfrak{g}_{-\alpha_i} = \mathbb{R}f_i, \tag{2.2.32}$$

and, consequently, the multiplicities of the simple roots are one,

$$\mathrm{mult}(\alpha_i) = 1. \tag{2.2.33}$$

For finite-dimensional Lie algebras the root multiplicities are always one. This does not carry over to infinite-dimensional Kac-Moody algebras, for which roots can have arbitrarily large multiplicity. We shall come back to the issue of root multiplicities in Section 2.2.6 when we discuss the Weyl group. We can now write the full root system as follows

$$\Phi = \{\alpha \in \mathfrak{h}^\star \mid \alpha \neq 0,\ \mathfrak{g}_\alpha \neq 0\}. \tag{2.2.34}$$

A useful notion is that of the *height* of a root. This is a linear integral map

$$\mathrm{ht} : \alpha\ \longmapsto\ \mathrm{ht}(\alpha) \in \mathbb{Z} \tag{2.2.35}$$

defined as the sum of the coefficients of α in the basis of simple roots (see (2.2.20)),

$$\mathrm{ht}(\alpha) = \sum_{i=1}^{r} m_i. \tag{2.2.36}$$

It follows that for $\alpha \in \Phi_+$ we have $\mathrm{ht}(\alpha) > 0$, and vice versa for the negative roots. An important object will turn out to be half the sum of all positive roots,

$$\rho := \frac{1}{2}\sum_{\alpha\in\Phi_+} \alpha. \tag{2.2.37}$$

We now have a better understanding of the appearance of the Cartan matrix in (2.2.3). It simply corresponds to the values of the simple roots $\alpha_j \in \mathfrak{h}^\star$ acting on the simple coroots $\alpha_i^\vee \in \mathfrak{h}$, i.e.,

$$A_{ij} = \alpha_j(\alpha_i^\vee) = \langle \alpha_j, \alpha_i^\vee \rangle . \tag{2.2.38}$$

Finally we shall here develop a more geometric description of the root system, which is very useful for our understanding of the Kac-Moody algebra $\mathfrak{g}(A)$. An arbitrary root $\gamma \in \Phi$ may be seen as a "vector" in $\mathfrak{h}^\star$ with components given by

$$\gamma_i \equiv \gamma(h_i), \tag{2.2.39}$$

i.e., the components of γ correspond to the different values of the root γ acting on the simple coroots $h_i = \alpha_i^\vee$. We shall sometimes write the *root vector* $\vec{\gamma}$ as

$$\vec{\gamma} = (\gamma_1, \ldots, \gamma_r) \ \in \ \mathfrak{h}^\star. \tag{2.2.40}$$

From this point of view the entries A_{ij} of the Cartan matrix correspond to the components of the root vectors $\vec{\alpha}_i$ associated with the simple roots:

$$\begin{aligned} \vec{\alpha}_i &= (\alpha_{i(1)}, \alpha_{i(2)}, \ldots, \alpha_{i(r)}) \\ &= (A_{1i}, A_{2i}, \ldots, A_{ri}), \end{aligned} \tag{2.2.41}$$

where we have indicated the component index (i) of the simple roots within parenthesis to distinguish it from the index i labeling the different simple roots.

We conclude this section by defining an involution ω on the Kac-Moody algebra $\mathfrak{g}(A)$, known as the *Chevalley involution*. This is defined as follows on the Chevalley generators:

$$\omega(e_i) = -f_i, \qquad \omega(f_i) = e_i, \qquad \omega(h_i) = -h_i. \tag{2.2.42}$$

This involution leaves the Chevalley relations, (2.2.3), invariant and therefore corresponds to an automorphism of $\mathfrak{g}(A)$. The involution ω acts as on multiple commutators in the standard way, e.g, on $e_3 \equiv [e_1, e_2] \in \mathfrak{n}_+ \subset \mathfrak{g}(A)$ one has

$$\omega(e_3) = \omega([e_1, e_2]) = [\omega(e_1), \omega(e_2)] = [f_1, f_2] = f_3 \in \mathfrak{n}_- \subset \mathfrak{g}(A). \tag{2.2.43}$$

The subset of $\mathfrak{g}(A)$ which is pointwise fixed under ω defines the *maximal compact subalgebra*

$$K(\mathfrak{g}) = \big\{ x \in \mathfrak{g}(A) \ \big| \ \omega(x) = x \big\} \ \subset \ \mathfrak{g}. \tag{2.2.44}$$

The maximal compact subalgebra is generated by the combinations $e_i - f_i, \ i = 1, \ldots, r$, of Chevalley generators. We have the induced *Cartan decomposition* of $\mathfrak{g}(A)$ (direct sum of vector spaces):

$$\mathfrak{g} = K(\mathfrak{g}) \oplus \mathfrak{p}, \tag{2.2.45}$$

where the complement $\mathfrak{p}$ is the subset of $\mathfrak{g}$ which is pointwise anti-invariant under ω,

$$\mathfrak{p} = \big\{ x \in \mathfrak{g}(A) \ \big| \ \omega(x) = -x \big\}. \tag{2.2.46}$$

This is not a subalgebra of $\mathfrak{g}$, but elements of $\mathfrak{p}$ transforms in some representation of $K(\mathfrak{g})$. The Cartan decomposition yields the following characteristic properties of a *symmetric space*:

$$[\mathfrak{p},\mathfrak{p}] \subset K(\mathfrak{g}), \qquad [K(\mathfrak{g}),\mathfrak{p}] \subset \mathfrak{p}, \qquad [K(\mathfrak{g}),K(\mathfrak{g})] \subset K(\mathfrak{g}). \tag{2.2.47}$$

Let us also note here an additional important decomposition of $\mathfrak{g}(A)$. This is the *Iwasawa decomposition* which reads

$$\mathfrak{g} = K(\mathfrak{g}) \oplus \mathfrak{h} \oplus \mathfrak{n}_+. \tag{2.2.48}$$

In the finite-dimensional case this decomposition reduces to the familiar fact that any matrix can be decomposed into an orthogonal part, a diagonal part and an upper triangular part. The subset

$$\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}_+ \tag{2.2.49}$$

is known as the *Borel subalgebra*. There is an alternative Iwasawa decomposition which instead utilizes the negative nilpotent subspace $\mathfrak{n}_-$:

$$\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus K(\mathfrak{g}), \tag{2.2.50}$$

with an associated negative Borel subalgebra

$$\mathfrak{b}_- = \mathfrak{n}_- \oplus \mathfrak{h}. \tag{2.2.51}$$

The first of version of the Iwasawa decomposition will be used frequently in **Part I** of this thesis, while in **Part II** we will make use of the second version. This is purely a matter of convention.

2.2.4 The Invariant Bilinear Form

To proceed with the analysis of the roots of a Kac-Moody algebra it is useful to first define a "metric" $(\cdot|\cdot)$ on the space $\mathfrak{h}^\star$. This will then be extended to an invariant bilinear form on the entire Kac-Moody algebra $\mathfrak{g}(A)$, and will thereby also play an important role in many of the subsequent developments.

We shall assume, as before, that the Cartan matrix, is non-degenerate, and, in addition, we shall take it to be *symmetrizable*. The first condition will be lifted later on, while the second condition will be kept throughout the remainder of these lectures. Symmetrizability of A implies that there exists a diagonal matrix $D = \mathrm{diag}(\epsilon_1, \ldots, \epsilon_r)$, with all $\epsilon_i > 0$, such that the Cartan matrix decomposes according to

$$A = DS, \tag{2.2.52}$$

where S is a symmetric $r \times r$ matrix. The matrix $S = (S_{ij})$ now defines a symmetric invertible bilinear form $(\cdot|\cdot)$ on $\mathfrak{h}^\star$ as follows

$$S_{ij} \equiv (\alpha_i|\alpha_j), \tag{2.2.53}$$

for $\alpha_i, \alpha_j \in \Pi$. Moreover, by imposing the defining relation $A_{ii} = 2$ we find

$$\epsilon_i = \frac{2}{(\alpha_i|\alpha_i)}. \tag{2.2.54}$$

We have now defined a bilinear form on the space $\mathfrak{h}^\star$, which in turn induces a bilinear form on the root lattice Q. An important consequence of this is the following:

- For finite-dimensional Lie algebras the bilinear form $(\cdot|\cdot)$ is of Euclidean signature and, consequently, Q is a Euclidean lattice. In this case the bilinear form coincides with the standard Killing form.
- For Lorentzian Kac-Moody algebras the bilinear form $(\cdot|\cdot)$ is a flat metric with signature $(- + \cdots +)$ and, consequently, Q is a Lorentzian lattice.

The bilinear form can now be extended to the full Kac-Moody algebra. Since $(\cdot|\cdot)$ is non-degenerate it defines an isomorphism $\mu : \mathfrak{h}^\star \to \mathfrak{h}$ as follows:

$$\langle \alpha, \mu(\beta)\rangle \equiv (\alpha|\beta), \qquad \beta, \alpha \in \mathfrak{h}^\star,\ \mu(\beta) \in \mathfrak{h}, \tag{2.2.55}$$

with the inverse map $\mu^{-1} : \mathfrak{h} \to \mathfrak{h}^\star$ then defines a bilinear form $(\cdot|\cdot)$ on the Cartan subalgebra $\mathfrak{h}$ through

$$\left\langle \mu^{-1}(\alpha^\vee), \beta^\vee \right\rangle \equiv (\alpha^\vee|\beta^\vee), \qquad \alpha^\vee, \beta^\vee \in \mathfrak{h},\ \mu(\alpha^\vee) \in \mathfrak{h}^\star. \tag{2.2.56}$$

Then, from the definition of $(\cdot|\cdot)$ in (2.2.53), we find

$$(\alpha_i|\beta) = \frac{1}{\epsilon_i} \left\langle \beta, \alpha_i^\vee \right\rangle. \tag{2.2.57}$$

In addition, by virtue of (2.2.55), we have the relation

$$(\alpha_i|\beta) = \langle \beta, \mu(\alpha_i)\rangle, \tag{2.2.58}$$

and by equating (2.2.57) with (2.2.58) we arrive at the explicit expressions

$$\mu(\alpha_i) = \frac{1}{\epsilon_i}\alpha_i^\vee \qquad \text{or} \qquad \mu^{-1}(\alpha_i^\vee) = \epsilon_i \alpha_i. \tag{2.2.59}$$

We can use this result to find a relation between the bilinear forms on $\mathfrak{h}$ and $\mathfrak{h}^\star$:

$$(\alpha_i^\vee|\alpha_j^\vee) = \epsilon_i\epsilon_j(\alpha_i|\alpha_j), \tag{2.2.60}$$

and in the special case $i = j$ we thus have

$$\frac{(\alpha_i^\vee|\alpha_i^\vee)}{2} = \frac{2}{(\alpha_i|\alpha_i)}. \tag{2.2.61}$$

Let us further note that (2.2.57) ensures that the Cartan matrix can be expressed solely in terms of the bilinear form

$$A_{ij} = \frac{2(\alpha_i|\alpha_j)}{(\alpha_i|\alpha_i)}, \tag{2.2.62}$$

an expression which is very useful for practical purposes.

At this point we have a non-degenerate symmetric bilinear form on the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}(A)$. To extend this to the entire algebra, one exploits the *invariance* of $(\cdot|\cdot)$, i.e., the property

$$([x, y], z) = (x, [y, z]), \qquad x, y, z \in \mathfrak{g}(A). \tag{2.2.63}$$

For example, by computing $(\alpha_i^\vee|[\alpha_k^\vee, e_j]) = A_{kj}(\alpha_i^\vee|e_j)$ and using the invariance on the left hand side we find

$$(\alpha_i^\vee|e_j) = 0, \tag{2.2.64}$$

because of the fact that $[\alpha_i^\vee, \alpha_k^\vee] = 0$. A similar argument gives $(\alpha_i^\vee | f_j) = 0$. Moreover, by computing $([\alpha_k^\vee, e_i] | f_j)$ one finds

$$(e_i | f_j) = \epsilon_i \delta_{ij}, \tag{2.2.65}$$

or, more generally, for two arbitrary generators $x_\alpha, x_\beta \in \mathfrak{g}(A)$ one has

$$(x_\alpha | x_\beta) \sim \delta_{\alpha,-\beta}, \tag{2.2.66}$$

where the proportionality constant depends on the normalization of the Chevalley generators.

Before we proceed with some examples, let us discuss some additional features of the root system Φ of a Kac-Moody algebra $\mathfrak{g}(A)$. In the special case when $\mathfrak{g}(A)$ is a rank r finite-dimensional Lie algebra we have seen that the root lattice is an r-dimensional Euclidean lattice, thus implying that all roots have positive norm, $\alpha^2 > 0$, $\forall \alpha \in \Phi$. In the general case however, the root lattice can have arbitrary signature, and thus roots can in general have positive, zero or negative norm. We shall adopt the standard terminology and call roots of positive norm *real* roots, and those of zero or negative norm, *imaginary* roots. In this way the root system of a Kac-Moody algebra decomposes into two disjoint sets $\Phi_{\Re}$ and $\Phi_{\Im}$ of real and imaginary roots, respectively. The largest norm squared of the real roots is, by analogy with the finite-dimensional case, restricted to 2, while the imaginary roots can come with arbitrarily large negative norm squared. We can thus describe these two types of roots as follows

$$\begin{aligned} \Phi_{\Re} &= \big\{ \alpha \in \Phi \, \big| \, 0 < (\alpha|\alpha) \leq 2 \big\}, \\ \Phi_{\Im} &= \big\{ \beta \in \Phi \, \big| \, (\beta|\beta) \leq 0 \big\}, \end{aligned} \tag{2.2.67}$$

and we have

$$\Phi = \Phi_{\Im} \cup \Phi_{\Re}. \tag{2.2.68}$$

The multiplicity of the real roots is always one, $\text{mult}(\alpha) = 1$, $\forall \alpha \in \Phi_{\Re}$, while the imaginary roots generally come with a non-trivial multiplicity, $\text{mult}(\beta) > 1$, $\forall \beta \in \Phi_{\Im}$. In particular, for the indefinite Kac-Moody algebras the multiplicity of the imaginary roots grows exponentially with increasing height, thus rendering these algebras very difficult to control. We shall come back to the issue of real and imaginary roots in Section 2.2.6 after we have learned some of the basic properties of the Weyl group.

2.2.5 Example: A_2 versus A_1^+

Let us now try to make all this a bit more concrete, by introducing and comparing two examples in detail. We shall consider the familiar finite-dimensional Lie algebra $A_2 = \mathfrak{sl}(3, \mathbb{R})$ and the infinite-dimensional affine Kac-Moody algebra A_1^+ (the notation will be explained in Section 2.4.1). Our goal is firstly to understand what it is that makes the first one finite and the second one infinite-dimensional. Secondly, we shall investigate and compare the two different root systems.

Serre Relations

The rank 2 Lie algebras A_2 and A_1^+ are described by the Cartan matrices

$$A[A_2] = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \qquad A[A_1^+] = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}, \tag{2.2.69}$$

with the associated Dynkin diagrams displayed in Figure 2.1. For simplicity of notation, we shall refer to the two different Cartan matrices simply as $A = A[A_2]$ and $\bar{A} = A[A_1^+]$.

Figure 2.1: On the left the Dynkin diagram of the Lie algebra $A_2 = \mathfrak{sl}(3,\mathbb{R})$ and on the right the Dynkin diagram of the affine Kac-Moody algebra A_1^+.

The Chevalley generators for A_2 are $\{e_1, e_2, f_1, f_2, h_1, h_2\}$ and the ones for A_1^+ are $\{\bar{e}_1, \bar{e}_2, \bar{f}_1, \bar{f}_2, \bar{h}_1, \bar{h}_2\}$. The commutation relations follow from the general form of the Chevalley relations in (2.2.3), with the insertion of the individual Cartan matrix components. For example, we have

$$\begin{aligned} A_2 \quad &: \quad [h_1,h_2]=0, \quad [h_1,e_2]=-e_2, \quad [h_2,e_1]=-e_1, \quad [e_1,f_1]=h_1, \\ A_1^+ \quad &: \quad [\bar{h}_1,\bar{h}_2]=0, \quad [\bar{h}_1,\bar{e}_2]=-2\bar{e}_2, \quad [\bar{h}_2,\bar{e}_1]=-2\bar{e}_1, \quad [\bar{e}_1,\bar{f}_1]=\bar{h}_1. \end{aligned} \tag{2.2.70}$$

It is clear that the relations in A_2 are remarkably similar to those in A_1^+ with the only difference arising in the relations involving the off-diagonal entries of the Cartan matrices, which, of course, is the only obvious distinction between the algebras at this point. We now want to understand how this seemingly trivial change in the Cartan matrix can render an algebra infinite-dimensional. The answer lies in the Serre relations.

Let us now proceed to check the Serre relations involving the positive nilpotent generators of "e-type". The analysis is analogous for the negative ones. For A_2 we then have

$$\mathrm{ad}_{e_1}^{1-A_{12}}(e_2) = [e_1,[e_1,e_2]] = 0, \tag{2.2.71}$$

implying that the generator $[e_1,[e_1,e_2]]$ does not exist in the algebra. The commutators $e_3 \equiv [e_1,e_2]$, on the other hand, is not killed by (2.2.71) and so corresponds to a new generator. On the negative side, we similarly find the new generator $f_3 \equiv -[f_1,f_2]$. No other non-vanishing generators exist in A_2 and therefore the algebra is eight-dimensional. We may take as a basis of A_2 the eight elements $\{e_1,e_2,e_3,f_1,f_2,f_3,h_1,h_2\}$. This corresponds to the adjoint representation $\mathbf{8}$ of $\mathfrak{sl}(3,\mathbb{R})$.

We now turn to A_1^+. The Serre relation for $\bar{e}_1$ and $\bar{e}_2$ reads

$$\mathrm{ad}_{\bar{e}_1}^{1-\bar{A}_{12}}(\bar{e}_2) = [\bar{e}_1,[\bar{e}_1,[\bar{e}_1,\bar{e}_2]]] = 0. \tag{2.2.72}$$

This condition therefore kills the generator $[\bar{e}_1,[\bar{e}_1,[\bar{e}_1,\bar{e}_2]]]$ in A_1^+, while there are no restrictions on the following two generators:

$$\bar{e}_3 \equiv [\bar{e}_1,\bar{e}_2], \qquad \bar{e}_4 = [\bar{e}_1,[\bar{e}_1,\bar{e}_2]]. \tag{2.2.73}$$

In addition, we have the Serre relation for $\bar{e}_2$ acting on $\bar{e}_1$ which yields yet another non-vanishing generator

$$\bar{e}_5 \equiv [\bar{e}_2, [\bar{e}_2, \bar{e}_1]]. \tag{2.2.74}$$

It is the existence of $\bar{e}_4$ and $\bar{e}_5$ which renders A_1^+ infinite-dimensional. For example, consider the following multicommutator, alternating between $\bar{e}_1$ and $\bar{e}_2$,

$$[\bar{e}_1, [\bar{e}_2, [\bar{e}_1, \bar{e}_2]]] = -[\bar{e}_1, [\bar{e}_2, [\bar{e}_2, \bar{e}_1]] \neq 0. \tag{2.2.75}$$

In A_2 this commutator would have been zero because the Serre relations impose

$$[e_2, [e_2, e_1]] = 0,$$

while in A_1^+ it corresponds to $-[\bar{e}_1, \bar{e}_5]$ which is unrestricted. It is possible to continue in this way, and any alternating multicommutator is non-vanishing, e.g,

$$[\bar{e}_1, [\bar{e}_2, [\bar{e}_1, [\bar{e}_2, \ldots, [\bar{e}_1, \bar{e}_2] \cdots]]]] \; \in \; A_1^+. \tag{2.2.76}$$

Root Systems

We shall now proceed to compare the two algebras A_2 and A_1^+ at the level of their respective systems of roots. To A_2 we associate the simple roots $\Pi = \{\alpha_1, \alpha_2\}$, and to A_1^+, $\bar{\Pi} = \{\bar{\alpha}_1, \bar{\alpha}_2\}$. Since the Cartan matrices are symmetric, they give directly the bilinear forms on the space of roots. We have

$$\begin{aligned} &(\alpha_1|\alpha_1) = 2, \quad (\alpha_2|\alpha_2) = 2, \quad (\alpha_1|\alpha_2) = -1, \\ &(\bar{\alpha}_1|\bar{\alpha}_1) = 2, \quad (\bar{\alpha}_2|\bar{\alpha}_2) = 2, \quad (\bar{\alpha}_1|\bar{\alpha}_2) = -2. \end{aligned} \tag{2.2.77}$$

All roots can be described as integral non-negative or non-positive linear combinations of the simple roots. For A_2, we find that $\alpha_1 + \alpha_2$ is a root, because the corresponding generator $e_3 \equiv [e_1, e_2]$ survives the Serre relations. Thus we define $\alpha_3 \equiv \alpha_1 + \alpha_2$. Let is now try to add yet another simple root and take, say, $\alpha_3 + \alpha_2$. This corresponds to the generator $[e_3, e_2] = [[e_1, e_2], e_2]$ which is zero in A_2 because of the Serre relations. In this way we find that the root system $\Phi = \Phi(A)$ of A_2 is given by

$$\Phi = \Phi_+ \cup \Phi_- = \big\{\alpha_1, \alpha_2, \alpha_3\big\} \cup \big\{ -\alpha_1, -\alpha_2, -\alpha_3\big\}, \tag{2.2.78}$$

revealing that indeed the root system of A_2 is finite. Of course, any vector $m\alpha_1 + n\alpha_2,\ m, n \in \mathbb{Z}$, lies on the root lattice Q of A_2, even though it is not a root. Because the root α_3 is the root with largest height, $\mathrm{ht}(\alpha_3) = 2$, of Φ, it is the highest root of the algebra. This is also the highest weight of the adjoint representation.

It is illuminating to draw the root system in a root diagram, which makes it easier to visualize the structure of the algebra. To this end we define, as described in the previous section, the simple root vectors $\vec{\alpha}_1$ and $\vec{\alpha}_2$, with components

$$\begin{aligned} \vec{\alpha}_1 &= (A_{11}, A_{12}) = (2, -1), \\ \vec{\alpha}_2 &= (A_{21}, A_{22}) = (-1, 2). \end{aligned} \tag{2.2.79}$$

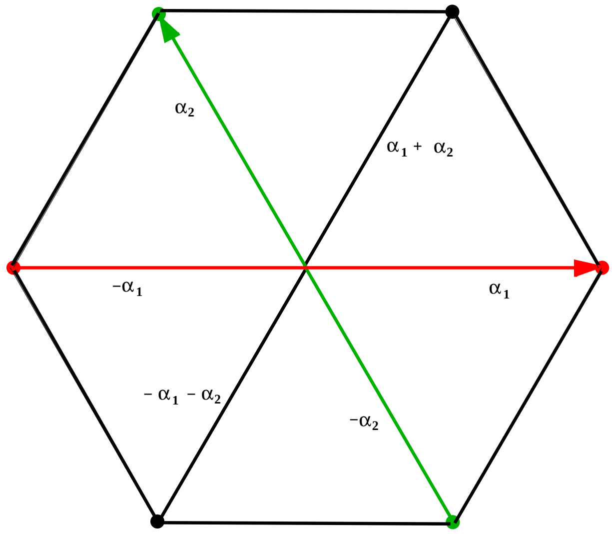


Figure 2.2: The root diagram of A_2, representing the adjoint representation **8**. The root $\alpha_1 + \alpha_2$ is the highest root corresponding to the highest weight of the representation.

The two vectors $\vec{\alpha}_1$ and $\vec{\alpha}_2$ span a two-dimensional Euclidean lattice, with separating angle of $2\pi/3$. We have indicated the root diagram of A_2 in Figure 2.2. Let us now analyze the root system $\bar{\Phi} = \Phi(\bar{A})$ of A_1^+. We begin by noting that the determinant of the Cartan matrix vanishes

$$\det \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} = 0, \tag{2.2.80}$$

as we have seen is the distinguishing feature of affine Kac-Moody algebras. This implies that the bilinear form of the algebra constructed from A is degenerate. We shall discuss how to deal with this feature in Section 2.3.2. For our present purposes, however, we just note that (2.2.80) implies that there exists a root $\bar{\delta} \in \bar{\Phi}$, which has zero norm,

$$(\bar{\delta}|\bar{\delta}) = 0. \tag{2.2.81}$$

In terms of the simple roots, we have

$$\bar{\delta} = \bar{\alpha}_1 + \bar{\alpha}_2, \tag{2.2.82}$$

as follows from (2.2.77). That $\bar{\delta}$ is indeed a root of $\bar{\Phi}$ can be seen by noting that the associated generator $[\bar{e}_1, \bar{e}_2]$ is non-vanishing. The existence of a null root is indicative of the fact that the algebra, as well as the associated root system, is infinite-dimensional.

In order to understand the root system it will prove convenient to write $\bar{\alpha}_2 = \bar{\delta} - \bar{\alpha}_1$, and treat the root system as a two-dimensional *Lorentzian* space, with basis vectors $\bar{\alpha}_1$ and $\bar{\delta}$. Is $\bar{\delta} - 2\bar{\alpha}_1$ a root of the algebra? The answer is no, because the associated generator vanishes,

$$\bar{e}_{\bar{\delta}-2\bar{\alpha}_1} \equiv [\bar{e}_{\bar{\delta}-\bar{\alpha}_1}, \bar{f}_1] = [\bar{e}_2, \bar{f}_1] = 0, \tag{2.2.83}$$

as follows from the Chevalley relations. However, $2\bar{\delta} - \bar{\alpha}_1 = \bar{\alpha}_1 + 2\bar{\alpha}_2$ is a root since it corresponds to the generator $[\bar{e}_2[\bar{e}_2, \bar{e}_1]] \neq 0$. One can iterate this procedure and find a complete description of the root system. All null roots are multiples of $\bar{\delta}$, while the real roots are combinations of $\pm\bar{\alpha}_1$ with $\bar{\delta}$. As discussed in the previous section, the root system thereby splits into disjoint sets corresponding to the *real* (i.e., spacelike) and the *imaginary* (i.e., lightlike) roots. Explicitly, the root system of A_1^+ then reads

$$\bar{\Phi} = \bar{\Phi}_{\Re} \cup \bar{\Phi}_{\Im} = \big\{ \pm\bar{\alpha}_1 + n\bar{\delta} \bigm| n \in \mathbb{Z} \big\} \cup \big\{ k\bar{\delta} \bigm| k \in \mathbb{Z} \setminus \{0\} \big\}. \tag{2.2.84}$$

In addition we have, of course, the usual split of $\bar{\Phi}$ into positive and negative roots.

2.2.6 The Weyl Group

A very important concept, which we shall use extensively in subsequent sections, is that of the *Weyl group* $\mathcal{W} = \mathcal{W}(A)$ of the Kac-Moody algebra $\mathfrak{g}(A)$. We begin by constructing the group $\mathcal{W}(A)$ abstractly and then we show how it is related to a Kac-Moody algebra. Fix a set of generators $\mathcal{S} = \{s_1, \ldots, s_r\}$ and let $\tilde{\mathcal{W}}$ be the free group generated by $\mathcal{S}$. Let $m = (m_{ij})_{i,j=1,\ldots,r}$ be an $r \times r$ matrix satisfying: (i) $m_{ii} = 1$, (ii) $m_{ij} \in \mathbb{Z}_{\geq 1}$, $i \neq j$, and (iii) $m_{ij} = m_{ji}$. The group $\tilde{\mathcal{W}}$ then has a normal subgroup $\mathcal{N}$ generated by the particular combinations [73]

$$(s_i s_j)^{m_{ij}}. \tag{2.2.85}$$

The Weyl group $\mathcal{W}$, associated to the set $\mathcal{S}$, is the quotient group

$$\mathcal{W} \equiv \tilde{\mathcal{W}} \,/\, \mathcal{N}. \tag{2.2.86}$$

This is a particular instance of a *Coxeter group*, and the entries of the matrix (m_{ij}) are called *Coxeter exponents*. Our construction implies that the Weyl group, $\mathcal{W}$, is the group generated by the set $\mathcal{S}$ modulo the relations

$$(s_i s_j)^{m_{ij}} = 1, \quad i, j = 1, \ldots, r. \tag{2.2.87}$$

The elements $s_i \in \mathcal{S}$ are called the *fundamental reflections*, by virtue of the property

$$s_i^2 = 1, \tag{2.2.88}$$

as follows from the first condition on (m_{ij}) above.

We now show how the group $\mathcal{W}(A)$ enters the story of Kac-Moody algebras. The Weyl group is a group of automorphisms of the root lattice Q,

$$\mathcal{W} \,:\, Q \longrightarrow Q, \tag{2.2.89}$$

with the fundamental reflections s_i being geometrically realized as reflections in the hyperplanes orthogonal to the simple roots α_i. More specifically, we associate a fundamental reflection s_i to each simple root α_i, such that the action on $\gamma \in Q$ is given by

$$s_i \,:\, \gamma \;\longmapsto\; s_i(\gamma) = \gamma - \big\langle \gamma, \alpha_i^\vee \big\rangle\, \alpha_i. \tag{2.2.90}$$

It is clear from this definition that s_i reverses the sign of α_i and pointwise fixes the hyperplane

$$T_i = \big\{ \beta \in Q \bigm| \big\langle \beta, \alpha_i^\vee \big\rangle = 0 \big\}. \tag{2.2.91}$$

$A_{ij}A_{ji}$	0	1	2	3	≥ 4
m_{ij}	2	3	4	6	∞

Table 2.1: The relation between the entries A_{ij} of the Cartan matrix and the Coxeter exponents m_{ij}.

Let us also check the condition $s_i^2 = 1$ explicitly for the geometric realization. Applying (2.2.90) twice on $\gamma \in Q$ yields

$$\begin{aligned} s_i \cdot s_i(\gamma) &= s_i\big(\gamma - \left\langle \gamma, \alpha_i^\vee \right\rangle \alpha_i\big) \\ &= \gamma - 2\left\langle \gamma, \alpha_i^\vee \right\rangle \alpha_i + \left\langle \gamma, \alpha_i^\vee \right\rangle \left\langle \alpha_i, \alpha_i^\vee \right\rangle \alpha_i \\ &= \gamma, \end{aligned} \tag{2.2.92}$$

where, in the last step, we made use of the fact that $\langle \alpha_i, \alpha_i^\vee \rangle = A_{ii} = 2$.

When realized geometrically in this way, the fundamental reflections are commonly called *Weyl reflections.* When acting on the simple roots themselves, the Weyl reflections become

$$s_i(\alpha_j) = \alpha_j - \frac{2(\alpha_i|\alpha_j)}{(\alpha_i|\alpha_i)}\alpha_i = \alpha_j - A_{ij}\alpha_i, \tag{2.2.93}$$

where (A_{ij}) is the Cartan matrix. There is a simple relation between the entries of the Cartan matrix and the associated Coxeter exponents. This is displayed in Table 2.1. When $A_{ij}A_{ji} \geq 4$ the corresponding Coxeter exponents are infinite, implying that there are no relations between the generators s_i and s_j for these particular values of i and j.

The bilinear form $(\cdot|\cdot)$ is $\mathcal{W}$-invariant,

$$\big(\omega(\beta)\big|\omega(\beta')\big) = \big(\beta\big|\beta'\big), \tag{2.2.94}$$

which follows by direct calculation using (2.2.90). This implies that the Weyl group is "orthogonal" with respect to the bilinear form $(\cdot|\cdot)$ and hence is a discrete subgroup of the isometry group $O(\mathfrak{h}^\star)$ of $\mathfrak{h}^\star$,

$$\mathcal{W} \subset O(\mathfrak{h}^\star). \tag{2.2.95}$$

We can now make use of the Weyl group to get a better handle on the structure of the root system. The first important fact is that the root system $\Phi(A)$ of a Kac-Moody algebra $\mathfrak{g}(A)$ is $\mathcal{W}(A)$-invariant,

$$\mathcal{W} \cdot \Phi = \Phi. \tag{2.2.96}$$

We can associate a general Weyl reflection ω_α to any root $\alpha \in \Phi$. This will be described by a finite product of the fundamental reflections,

$$\omega_\alpha = s_{i_1}s_{i_2}\cdots s_{i_{k-1}}s_{i_k}, \tag{2.2.97}$$

where the minimal number k of fundamental reflections needed to describe ω_α is called the *length* of ω_α, and is denoted by $\ell(\alpha)$. By definition, the fundamental reflections have length one, $\ell(s_i) = 1$.

The reflection $\omega_\alpha \in \mathcal{W}$ fixes the hyperplane $T_\alpha \subset \mathfrak{h}^\star$,

$$T_\alpha = \{\gamma \in \mathfrak{h}^\star \mid (\gamma|\alpha) = 0,\ \alpha \in \Phi\}, \tag{2.2.98}$$

orthogonal to α. By removing all such hyperplanes we may decompose $\mathfrak{h}^\star$ into connected subsets, called *chambers*. We choose one such subset $\mathcal{C} \subset \mathfrak{h}^\star$ and give it a distinguished status as the *fundamental chamber*. The fundamental chamber is conventionally chosen as the region enclosed by the hyperplanes T_i orthogonal to the simple roots. This implies that $\mathcal{C}$ contains all vectors $\gamma \in \mathfrak{h}^\star$ such that $(\gamma|\alpha_i)$ is positive,

$$\mathcal{C} = \{\gamma \in \mathfrak{h}^\star \mid (\gamma|\alpha_i) > 0, \forall\ \alpha_i \in \Pi\}, \tag{2.2.99}$$

The chambers in $\mathfrak{h}^\star$ correspond to the images $\omega(\mathcal{C})$ for $\omega \in \mathcal{W}$. The union $\mathcal{X}$ of all such images is called the *Tits cone*,

$$\mathcal{X} = \bigcup_{\omega \in \mathcal{W}} \omega(\mathcal{C}). \tag{2.2.100}$$

For finite-dimensional Lie algebras, the Tits cone coincides with the space $\mathfrak{h}^\star$, while in general, when $\Phi_{\Im} \neq \emptyset$, one has $\mathcal{X} \subset \mathfrak{h}^\star$.

We are now at a stage where we can describe the root system Φ with more precision. A first important fact is that the sets of real and imaginary roots are separately invariant under the Weyl group,

$$\begin{aligned} \mathcal{W} \cdot \Phi_{\Re} &= \Phi_{\Re}, \\ \mathcal{W} \cdot \Phi_{\Im} &= \Phi_{\Im}. \end{aligned} \tag{2.2.101}$$

The sets of real and imaginary roots have a decomposition into disjoint sets of positive and negative roots, and we write

$$\begin{aligned} \Phi_{\Re} &= \Phi_{\Re+} \cup \Phi_{\Re-}, \\ \Phi_{\Im} &= \Phi_{\Im+} \cup \Phi_{\Im-}, \end{aligned} \tag{2.2.102}$$

with

$$\Phi_{\Re+} \cap \Phi_{\Im+} = 0, \qquad \Phi_{\Re-} \cap \Phi_{\Im-} = 0. \tag{2.2.103}$$

The only root $\alpha \in \Phi_{\Re+}$ for which $s_i(\alpha) \in \Phi_{\Re-}$ is $\alpha = \alpha_i$, implying that

$$s_i \cdot \Phi_{\Re+}/\{\alpha_i\} = \Phi_{\Re+}/\{\alpha_i\}, \tag{2.2.104}$$

and similarly for the negative real roots. A consequence of this is that, since $\alpha_i \notin \Phi_{\Im}$, the positive and negative imaginary roots are separately invariant under the Weyl group

$$\mathcal{W} \cdot \Phi_{\Im+} = \Phi_{\Im+}, \qquad \mathcal{W} \cdot \Phi_{\Im-} = \Phi_{\Im-}. \tag{2.2.105}$$

We can now state which elements of the root lattice Q are actually roots. First, all real roots lie in Weyl orbits of the simple roots, and thus we have

$$\Phi_{\Re} = \bigcup_{\omega \in \mathcal{W}} \omega(\Pi). \tag{2.2.106}$$

We now want to find a similar description for the set of imaginary roots. To this end it is useful to first introduce the notion of the *support*, supp(α), of an element $\alpha \in Q$. Let

$$\alpha = \sum_{i=1}^{r} k_i \alpha_i \ \in \ Q, \tag{2.2.107}$$

and introduce the subdiagram $\Xi_\alpha(A) \subset \Gamma(A)$, of the Dynkin diagram $\Gamma(A)$, as the diagram consisting only of those vertices i for which $k_i \neq 0$ and of all lines joining these vertices. We then have

$$\text{supp}(\alpha) \equiv \Xi_\alpha(A). \tag{2.2.108}$$

Next, we define a region $\mathcal{K} \subset \Phi_{\Im+}$ as follows

$$\mathcal{K} = \big\{ \alpha \in Q_+ \ \big| \ \langle \alpha, \alpha_i^\vee \rangle \leq 0, \ \forall \ i \ : \ \Xi_\alpha(A) \text{ is connected} \big\}. \tag{2.2.109}$$

The set of positive imaginary roots can now be elegantly described as the union of all images of $\mathcal{K}$ under the Weyl group [34],

$$\Phi_{\Im+} = \bigcup_{\omega \in \mathcal{W}} \omega(\mathcal{K}). \tag{2.2.110}$$

with a similar description of the negative imaginary roots.

Through the aid of the Weyl group $\mathcal{W}(A)$ we have now obtained a complete description of the root system $\Phi(A)$ of any Kac-Moody algebra $\mathfrak{g}(A)$. In subsequent sections we shall discuss in more detail the root systems for the classes of affine and hyperbolic Kac-Moody algebras, for which some simplifications arise.

We have defined the Weyl group as the group of reflections with respect to the simple roots. Through a natural generalization of (2.2.90) one can define reflections $s_\alpha \in \mathcal{W}$ through any real root $\alpha \in \Phi_\Re$. These reflections act as follows on $\beta \in \mathfrak{h}^\star$:

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha^\vee \rangle \alpha = \beta - \frac{2(\beta|\alpha)}{\alpha|\alpha)}\alpha. \tag{2.2.111}$$

How are these reflections related to the fundamental reflections $s_i \equiv s_{\alpha_i}$? To answer this question we first note that since α is real we must have $\alpha = \omega(\alpha_i)$ for some $\omega \in \mathcal{W}$ and some $\alpha_i \in \Pi$. Inserting $\alpha = \omega(\alpha_i)$ into (2.2.111) then yields

$$\begin{aligned} s_\alpha(\beta) &= \beta - \frac{2(\beta|\omega(\alpha_i))}{(\omega(\alpha_i)|\omega(\alpha_i))}\omega(\alpha_i) \\ &= \beta - \frac{2(\omega^{-1}(\beta)|\alpha_i)}{(\alpha_i|\alpha_i)}\omega(\alpha_i) \\ &= \beta - \langle \omega^{-1}(\beta), \alpha_i^\vee \rangle \omega(\alpha_i). \end{aligned} \tag{2.2.112}$$

where we made use of the invariance of $(\cdot|\cdot)$ under the Weyl group. We can rewrite (2.2.112) as

$$\begin{aligned} s_\alpha(\beta) &= \omega\big(\omega^{-1}(\beta) - \langle \omega^{-1}(\beta), \alpha_i^\vee \rangle \alpha_i\big), \\ &= \omega \cdot s_i\big(\omega^{-1}(\beta)\big), \end{aligned} \tag{2.2.113}$$

revealing that the generalized reflection s_α corresponds to a conjugation of the fundamental reflection s_i by some element $\omega \in \mathcal{W}$:

$$s_\alpha = \omega s_i \omega^{-1} \ \in \ \mathcal{W}. \tag{2.2.114}$$

2.3 Affine Kac-Moody Algebras

In this section we shall explore affine Kac-Moody algebras in more detail. This class of algebras corresponds to the first step away from the finite-dimensional Lie algebras, and is the only class of infinite-dimensional Kac-Moody algebras which are well understood. We recall that a Kac-Moody algebra $\mathfrak{g}(A)$ is said to be of affine type if the associated Cartan matrix A is positive semi-definite, det $A = 0$, with one zero eigenvalue. Because of the degeneracy of A, the bilinear form as constructed in Section 2.2.4 is degenerate. We shall explain how this problem is circumvented through the inclusion of an additional generator, called the *derivation* d, in the Cartan subalgebra. This new generator ensures that the invariant bilinear form on the full algebra is non-degenerate.

2.3.1 The Center of a Kac-Moody Algebra

The center Z of a Kac-Moody algebra $\mathfrak{g}(A)$ is defined as follows:

$$Z = \big\{x \in \mathfrak{g}(A) \;\big|\; \forall\, y \in \mathfrak{g}(A)\ :\ [x, y] = 0\big\}. \tag{2.3.1}$$

It is a general result that $Z \neq 0$ if and only if det $A = 0$ [34]. This is related to the rank of the matrix A. In previous sections we have treated the Cartan matrix as an $r \times r$ matrix of matrix rank equal to the rank r of the associated Kac-Moody algebra $\mathfrak{g}(A)$. Now we shall be more general and let $A = (A_{ij})_{i,j=1,\dots,r}$ be an $r \times r$ matrix of *matrix rank* n. In this case, det $A = 0$ and the rank r Kac-Moody algebra $\mathfrak{g}(A)$ has a non-trivial center $Z \neq 0$ of dimension

$$\dim Z = \text{corank } A = r - n. \tag{2.3.2}$$

For affine Kac-Moody algebras the Cartan matrix has only one zero eigenvalue and hence the corank of A is one, implying that the center is one-dimensional and is spanned by the *central element* c,

$$Z = \mathbb{R}c. \tag{2.3.3}$$

A consequence of this is that affine Kac-Moody algebras are *not* simple, since the center forms a non-trivial ideal in $\mathfrak{g}(A)$. The center Z is always contained in the Cartan subalgebra

$$Z \subset \mathfrak{h}, \tag{2.3.4}$$

implying that the central element c must be expressible as a linear combination $\sum_{i=1}^{r} c_i\alpha_i^\vee$ of the simple corrots $\Pi^\vee = \{\alpha_1^\vee, \dots, \alpha_r^\vee\}$. To see this, let $v = (v_1, \dots, v_r)^T$ be the non-trivial element of the kernel of the transposed matrix A^T, i.e.,

$$A^T \cdot v = \sum_{j=1}^{r} (A^T)_{ij} v_j = 0. \tag{2.3.5}$$

We then have $c_i = v_i$, i.e.,

$$c = \sum_{i=1}^{r} v_i \alpha_i^\vee \ \in\ Z \subset \mathfrak{h}. \tag{2.3.6}$$

This result follows from the fact that c must commute with all the Chevalley generators e_i, $i=1,\ldots,r$, and hence

$$0=[c,e_i]=\sum_{j=1}^{r}c_j[\alpha_j^\vee,e_i]=\sum_{j=1}^{r}c_j\left\langle\alpha_i,\alpha_j^\vee\right\rangle e_i=\sum_{j=1}^{r}c_jA_{ji}e_i=\Big[\sum_{j=1}^{r}(A^T)_{ij}c_j\Big]e_i, \tag{2.3.7}$$

which is only satisfied when $\sum_{j=1}^{r}(A^T)_{ij}c_j=0$, and hence $c_j=v_j$ as announced.

2.3.2 The Derived Algebra and the Derivation

It is now time to define what we mean by an affine Kac-Moody algebra. In fact, the algebra constructed from an "affine" Cartan matrix A using the Chevalley-Serre relations is only the *derived* Kac-Moody algebra

$$\mathfrak{g}'(A)=[\mathfrak{g}(A),\mathfrak{g}(A)]. \tag{2.3.8}$$

When $\det A\neq 0$ the derived algebra $\mathfrak{g}'$ coincides with the full algebra $\mathfrak{g}$, i.e., $\mathfrak{g}=[\mathfrak{g},\mathfrak{g}]$. To understand these statements we must introduce the notion of a *derivation* d. The motivation for this is that the complete Kac-Moody algebra must have a well-defined non-degenerate bilinear form, an object which does not exist for the derived algebra $\mathfrak{g}'$ because of the degeneracy of the Cartan matrix. For the following discussion it will be convenient to make a slight relabelling of the simple roots and the simple coroots. This is motivated by the fact that for any affine Kac-Moody algebra $\mathfrak{g}$ of rank r one may identify a maximal rank $r-1$ finite-dimensional subalgebra $\bar{\mathfrak{g}}\subset\mathfrak{g}$, and view $\mathfrak{g}$ as an *extension* $\mathfrak{g}\equiv\bar{\mathfrak{g}}^+$ of $\bar{\mathfrak{g}}$, where the superscript "+" indicates that the affine Kac-Moody algebra $\mathfrak{g}$ is obtained by adding a single node to the Dynkin diagram of $\bar{\mathfrak{g}}$ in a prescribed way. We shall discuss the general theory of extensions of Lie algebras in Section 2.4.1 but for now this will suffice. To this end we take the $r=k+1$ simple roots of $\mathfrak{g}(A)$ to be $\Pi=\{\alpha_0,\alpha_1,\ldots,\alpha_k\}$, with $\bar{\Pi}=\{\alpha_1,\ldots,\alpha_k\}$ representing the k simple roots of the finite subalgebra $\bar{\mathfrak{g}}$. The root α_0 is called the *affine root*. It is always of the form

$$\alpha_0=\delta-\theta, \tag{2.3.9}$$

where δ is a null root, $(\delta|\delta)=0$, and θ is the highest root of the finite subalgebra $\bar{\mathfrak{g}}\subset\mathfrak{g}$.

We now follow Kac [34] and add by hand a generator $d\in\mathfrak{h}$ to the Cartan subalgebra, with the property

$$\langle\alpha_i,d\rangle=\delta_{i0},\qquad i=0,1,\ldots,k. \tag{2.3.10}$$

The basis of the Cartan subalgebra $\mathfrak{h}$ is then taken to be $\Pi^\vee=\{\alpha_0^\vee,\alpha_1^\vee,\ldots,\alpha_k^\vee,d\}$, on which a non-degenerate bilinear form now exists with the following properties:

$$(c|\alpha_i^\vee)=0,\quad (c|c)=0,\quad (c|d)=1, \tag{2.3.11}$$

where the non-degeneracy of $(\cdot|\cdot)$ on $\mathfrak{h}$ follows from the non-vanishing scalar product between the derivation d and the central element c. An *affine* Kac-Moody algebra $\mathfrak{g}(A)$ is then defined as the derived algebra $\mathfrak{g}'$ augmented with the derivation d,

$$\mathfrak{g}=\mathfrak{g}'+\mathbb{R}d. \tag{2.3.12}$$

Note that this implies that an affine Kac-Moody algebra of rank r has a Cartan subalgebra of dimension $r+1$.

Let us now make an effort towards understanding the structure of (2.3.12). Because of its definition, (2.3.10), the derivation $d \in \mathfrak{h}$ will never appear on the right hand side of any commutator in the algebra. For example, the commutation relations with the positive Chevalley generators are

$$\begin{aligned} [d, e_a] &= 0, \quad a = 1, \ldots, k, \\ [d, e_0] &= e_0, \end{aligned} \tag{2.3.13}$$

and similarly for the f_i's. This implies that the derived algebra $\mathfrak{g}'$ does not contain the derivation d, thus explaining the structure of (2.3.12). In fact, the derivation can be viewed as a "counting operator" which counts the number of times the affine generator e_0, corresponding to the affine root α_0, appears in any commutator.

2.3.3 The Affine Root System

In the case of affine Kac-Moody algebras, the appearance of imaginary roots does imply a drastic complication. As we have alluded to before, the only "independent" imaginary root is the null root δ, of which all other imaginary roots are multiples. The complete root system Φ is therefore determined by the finite root system $\bar{\Phi}$ of the maximal finite subalgebra $\bar{\mathfrak{g}} \subset \mathfrak{g}$ and the null root δ. As mentioned in the previous section, the Dynkin diagram of $\bar{\mathfrak{g}}$ is obtained by deleting the zeroth node in the Dynkin diagram of the affine algebra $\mathfrak{g}$.

Let A be the Cartan matrix of an affine Kac-Moody algebra $\mathfrak{g}$, with associated root system $\Phi = \Phi(A)$. We begin by splitting Φ into its real and imaginary parts,

$$\Phi = \Phi_{\Re} \cup \Phi_{\Im}. \tag{2.3.14}$$

In the example in Section 2.2.5 we saw that the real roots of A_1^+ were given by all roots of the form $\alpha_1 + n\delta$, $n\mathbb{Z}$, with α_1 being the simple root (and, in fact, the only positive root) of the underlying finite algebra A_1. This fact generalizes to any affine Kac-Moody algebra $\mathfrak{g}(A)$ in the following way

$$\Phi_{\Re} = \big\{\alpha + n\delta \;\big|\; \forall\ \alpha \in \bar{\Phi};\ n \in \mathbb{Z}\big\}, \tag{2.3.15}$$

with $\bar{\Phi}$ being, as usual, the root system of the underlying finite-dimensional Lie algebra $\bar{\mathfrak{g}}$. The positive part of the real roots can then be described as follows

$$\Phi_{\Re+} = \big\{\alpha + n\delta \;\big|\; \forall\ \alpha \in \bar{\Phi};\ n \in \mathbb{Z}_{\geq 0}\big\} \cup \bar{\Phi}_+. \tag{2.3.16}$$

It is well known in Lie theory that if α is a root of a finite-dimensional Lie algebra, then the only multiples of α which are also roots are $\pm\alpha$. This feature carries over to the real part of the root system of general Kac-Moody algebras, while it is no longer true for the imaginary roots. For affine Kac-Moody algebras there exists only one independent imaginary root, and this is the root δ which appears in the construction of the zeroth simple root $\alpha_0 = \delta - \theta$. Any multiple of δ is also an imaginary root, and hence the imaginary part of the root system of any affine Kac-Moody algebra is very easy to describe:

$$\Phi_{\Im} = \Phi_{\Im+} \cup \Phi_{\Im-} = \big\{n\delta \;\big|\; n \in \mathbb{Z}_{\geq 0}\big\} \cup \big\{n\delta \;\big|\; n \in \mathbb{Z}_{\leq 0}\big\}. \tag{2.3.17}$$

2.3.4 The Affine Weyl Group

We now want to perform a closer analysis of the Weyl group of an affine Kac-Moody algebra. During our study we will find a natural explanation for where the name *affine* has its origin.

The Weyl group $\mathcal{W}(A)$ associated with an affine Kac-Moody algebra $\mathfrak{g}(A)$ is defined through the geometric action of the fundamental reflections $\mathcal{S} = \{s_0, s_1, \ldots, s_k\}$ on the dual space $\mathfrak{h}^\star$. Moreover we have seen that one can associate a reflection s_α with respect to any real root $\alpha \in \Phi_{\Re}$ as $s_\alpha = \omega s_i \omega^{-1}$ for $\omega \in \mathcal{W}$ and $\alpha = \omega(\alpha_i)$. On the other hand, no such construction exists for the imaginary roots since for $\beta \in \Phi_{\Im}$ the pairing $\langle \alpha, \beta^\vee \rangle$ is not defined. Therefore, although they act *on* $\Phi_{\Im}$, all Weyl reflections are defined *with respect to* real roots only.

An important new feature owing to the existence of null roots is that, since $(\delta|\alpha) = 0$ for all $\alpha \in \Phi_{\Re}$, we have

$$s_\alpha(\delta) = \delta - \frac{2(\delta|\alpha)}{(\alpha|\alpha)}\alpha = \delta. \tag{2.3.18}$$

This implies that the entire Weyl group acts as the identity on the set of imaginary roots:

$$\omega(\beta) = \beta, \quad \forall\, \beta \in \Phi_{\Im}. \tag{2.3.19}$$

Note that this is true only in the affine case for which all roots in $\Phi_{\Im}$ are lightlike, but not in the general case when $\Phi_{\Im}$ also contains timelike roots.

Let $\bar{\mathcal{W}} \subset \mathcal{W}$ be the finite Weyl group of $\bar{\mathfrak{g}} \subset \mathfrak{g}$. A particular feature of affine Weyl groups is the fact that they decompose into a semidirect product of the form

$$\mathcal{W} = \bar{\mathcal{W}} \ltimes \bar{T}^\vee, \tag{2.3.20}$$

where $\bar{T}^\vee$ denotes the abelian group of translations of the coroot lattice $\bar{Q}^\vee$ of $\bar{\mathfrak{g}}$.[3] We shall explain this phenomenon in detail below in the context of a simple example.

Example: The Weyl Group of A_1^+

Let $\Pi = \{\alpha_0, \alpha_1\}$ be the simple roots of $\mathfrak{g} = A_1^+$ and $\bar{\Pi} = \{\alpha_1\}$ the simple root of the underlying finite subalgebra $\bar{\mathfrak{g}} = A_1$. The Weyl group $\bar{\mathcal{W}}$ of A_1^+ is generated by a single reflection s_1 with respect to α_1:

$$s_1(\alpha_1) = -\alpha_1, \tag{2.3.21}$$

implying that $\bar{\mathcal{W}} = \mathbb{Z}_2$. The Cartan matrix of A_1^+ is

$$\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

which has a kernel spanned by the column vector $\begin{pmatrix} 1 & 1 \end{pmatrix}^T$. Thus, from our discussion of the center of Kac-Moody algebras in Section 2.3.1, we find that A_1^+ has a central element given by

$$c = \alpha_0^\vee + \alpha_1^\vee. \tag{2.3.22}$$

[3] Here we refer to the coroot lattice $Q^\vee$ as a lattice in $\mathfrak{h}^\star$, in the sense that we can use the non-degenerate bilinear form $(\cdot|\cdot)$ on $\mathfrak{g}(A)$ to identify $\mathfrak{h}$ with $\mathfrak{h}^\star$. Then the coroot lattice is spanned by the simple coroots $\alpha_i^\vee \equiv 2\alpha_i/(\alpha_i|\alpha_i) \in \mathfrak{h}^\star$ and we have $Q \subset Q^\vee$.

It is easy to check that c indeed commutes with all generators of the algebra. For example, we have

$$[c, e_1] = [\alpha_0^\vee, e_1] + [\alpha_1^\vee, e_1] = (-2+2)e_1 = 0. \qquad (2.3.23)$$

Recall now from Section 2.3.2 that so far we have been dealing only with the derived algebra

$$\mathfrak{g}' = [A_1^+, A_1^+] \qquad (2.3.24)$$

whose Cartan subalgebra is

$$\mathfrak{h}' = \mathbb{R}\alpha_0^\vee + \mathbb{R}\alpha_1^\vee. \qquad (2.3.25)$$

However, to understand the Weyl group $\bar{\mathcal{W}}$ it is crucial that we treat the full algebra

$$A_1^+ = [A_1^+, A_1^+] + \mathbb{R}d. \qquad (2.3.26)$$

Thus, we add by hand the derivation d to the algebra, with the properties

$$(c|d) = 1, \qquad (d|\alpha_0^\vee) = 0, \qquad (d|d) = 0. \qquad (2.3.27)$$

In the following we shall also view the central element c as a basis element of the Cartan subalgebra, instead of the generator $\alpha_0^\vee$. This will prove convenient later on. The full Cartan subalgebra now takes the form

$$\mathfrak{h} = \bar{\mathfrak{h}} \oplus \big(\mathbb{R}c + \mathbb{R}d\big), \qquad (2.3.28)$$

where $\bar{\mathfrak{h}} = \mathbb{R}\alpha_1^\vee$ is the Cartan subalgebra of A_1.

We shall now proceed to write the simple roots as the root vectors $\vec{\alpha}_0$ and $\vec{\alpha}_1$ with components given by the eigenvalues under the adjoint action of $\mathfrak{h}$ on the Chevalley generators e_0 and e_1. For e_0 we find

$$\begin{aligned} &[\alpha_1^\vee, e_0] = \langle \alpha_0, c\rangle\, e_0 = -2e_0, \\ &[c, e_0] = 0, \\ &[d, e_0] = \langle \alpha_0, d\rangle\, e_0 = e_0, \end{aligned} \qquad (2.3.29)$$

and for e_1 we have

$$\begin{aligned} &[\alpha_1^\vee, e_1] = \big\langle \alpha_1, \alpha_1^\vee\big\rangle\, e_1 = 2e_1, \\ &[c, e_1] = 0, \\ &[d, e_1] = \langle \alpha_1, d\rangle\, e_1 = 0, \end{aligned} \qquad (2.3.30)$$

where we made use of the defining relation, (2.3.10), for the derivation. Consequently, the component forms of the simple root vectors, in the basis determined by $\alpha_1^\vee, c$ and d, are

$$\begin{aligned} \vec{\alpha}_0 &= (-2, 0, 1), \\ \vec{\alpha}_1 &= (2, 0, 0). \end{aligned} \qquad (2.3.31)$$

We now have all the ingredients to understand the Weyl group of A_1^+. The group $\mathcal{W}$ is generated by the two fundamental reflections s_0 and s_1. We shall compute the action of these generators on an arbitrary vector

$$\vec{\lambda} = (\bar{\lambda}, k, m) \;\in\; \mathfrak{h}^\star. \qquad (2.3.32)$$

We begin by computing the action of s_1:

$$\begin{aligned} s_1(\vec{\lambda}) &= \vec{\lambda} - \frac{2(\vec{\lambda}|\vec{\alpha}_1)}{(\vec{\alpha}_1|\vec{\alpha}_1)}\vec{\alpha}_1 \\ &= (\bar{\lambda}, k, m) - (2\bar{\lambda}, 0, 0) \\ &= (-\bar{\lambda}, k, m). \end{aligned} \tag{2.3.33}$$

This result was expected, namely that the action of s_1 on the root lattice $\bar{Q} = \mathbb{Z}\alpha_1$ is the same as for the Weyl group of A_1, i.e., in component form we have

$$s_1(\bar{\lambda}) = -\bar{\lambda}, \qquad s_1(k) = k, \qquad s_1(m) = m. \tag{2.3.34}$$

Now, let us move on to the more interesting case of the s_0-generator. Its action on $\vec{\lambda}$ becomes

$$\begin{aligned} s_0(\vec{\lambda}) &= \vec{\lambda} - \frac{2(\vec{\lambda}|\vec{\alpha}_0)}{(\vec{\alpha}_0|\vec{\alpha}_0)}\vec{\alpha}_0 \\ &= (\bar{\lambda}, k, m) - \big[-\bar{\lambda} + k\big](-2, 0, 1) \\ &= (-\bar{\lambda} + 2k, k, m + \bar{\lambda} - k). \end{aligned} \tag{2.3.35}$$

We shall focus on the "projected" action of s_0 on the root lattice $\bar{Q}$, which reads

$$s_0(\bar{\lambda}) = -\bar{\lambda} + 2k. \tag{2.3.36}$$

This corresponds to a reflection of $\bar{\lambda}$, not with respect to the origin, but rather with respect to the point displaced by k away from the origin in $\bar{Q}$. We can make things more clear by noting that the combined reflection $t \equiv s_0 \circ s_1$ acts on $\bar{Q}$ as a pure translation:

$$t(\bar{\lambda}) = s_0 \circ s_1(\bar{\lambda}) = \bar{\lambda} + 2k. \tag{2.3.37}$$

Thus the Weyl group of A_1^+, acting on the Euclidean root lattice $\bar{Q}$ of $\bar{\mathfrak{g}}$, contains translations, i.e., *affine transformations*, thus explaining the origin of the name *affine* Kac-Moody algebras.

The action of s_0 can now be written as

$$s_0(\bar{\lambda}) = t \circ s_1(\bar{\lambda}), \tag{2.3.38}$$

which corresponds to the combination of an element $s_1 \in \bar{\mathcal{W}}$ and an element $t \in \bar{T}$, with $\bar{T}$ being the abelian group of translations of the root lattice $\bar{Q}$. This group is generated by t with a general element $t^n \in \bar{T}$ acting as follows

$$t^n(\bar{\lambda}) = \bar{\lambda} + 2kn. \tag{2.3.39}$$

The complete Weyl group of A_1^+ is therefore

$$\mathcal{W} = \big\{t^n,\ t^n \circ s_1 \ \big|\ n \in \mathbb{Z};\ s_1 \in \bar{\mathcal{W}}\big\}, \tag{2.3.40}$$

which is equivalent to the semidirect product

$$\mathcal{W} = \bar{\mathcal{W}} \ltimes \bar{T}, \tag{2.3.41}$$

as announced in the beginning of this section. The reason why it is the root lattice $\bar{Q}$ which appears here, and not the coroot lattice $\bar{Q}^\vee$, is that for A_1^+ they coincide, since all roots have the same length. However, in the general case it is the coroot lattice which is relevant.

2.4 Lorentzian Kac-Moody Algebras

Kac-Moody algebras of indefinite type constitute a vast wasteland of unexplored territory. Not much is known in general about these algebras; most notably there is no indefinite Kac-Moody algebra for which the root multiplicities are known in closed form to arbitrary height. We shall here focus on a particular subclass of indefinite Kac-Moody algebras, namely those for which the invariant bilinear form is of Lorentzian signature. These Lorentzian Kac-Moody algebras similarly constitute an infinite class of unclassified algebras, but they nevertheless has a subclass which is, in a certain sense, under control. We refer here to the Lorentzian algebras which can be obtained through prescribed extensions of finite-dimensional simple Lie algebras [38]. Extensions of finite Lie algebras can in this way give rise to Lorentzian Kac-Moody algebras, with properties which have proven to be of interest in string and M-theory. The precise extension procedure is described in detail in Section 2.4.1. Then, in Section 2.4.2, analyze a very interesting subclass of the Lorentzian Kac-Moody algebras, known as *hyperbolic*. As was the case with affine Kac-Moody algebras, we shall see that the hyperbolic class also draws its name from properties of the associated Weyl groups. The hyperbolic Kac-Moody algebras are the only examples of indefinite Kac-Moody algebras which have been completely classified.

2.4.1 Extensions of Lie Algebras

Let A be a symmetrizable generalized Cartan matrix and S its symmetric part. If the bilinear form $(\cdot|\cdot)$ defined by S is of Lorentzian signature then the Kac-Moody algebra $\mathfrak{g}(A)$ is of Lorentzian type. We shall in this section describe how algebras of Lorentzian type can be obtained from finite simple Lie algebras $\bar{\mathfrak{g}} = \mathfrak{g}(\bar{A})$ through double and triple extensions, denoted $\bar{\mathfrak{g}}^{++}$ and $\bar{\mathfrak{g}}^{+++}$, respectively, of the Dynkin diagram $\Gamma(\bar{A})$ associated with $\bar{\mathfrak{g}}$.

As in Section 2.3 we let $\bar{\Pi} = \{\alpha_1, \dots, \alpha_k\}$ be a basis of simple roots for the finite rank k Lie algebra $\bar{\mathfrak{g}}$. The root lattice $\bar{Q} = \sum_{i=1}^{k} \mathbb{Z}\alpha_i$ is Euclidean. Let also $\mathbb{Z}^{1,1}$ be the (unique) even two-dimensional unimodular Lorentzian lattice, spanned by the vectors u_1 and u_2. We define a non-degenerate bilinear form $(\cdot|\cdot)$ on $\mathbb{Z}^{1,1}$, of signature $(-+)$, by

$$(u_1|u_2) = 1, \qquad (u_1|u_1) = 0, \qquad (u_2|u_2) = 0. \tag{2.4.1}$$

This scalar product is induced from the standard Minkowski metric on $\mathbb{R}^{1,1}$ by taking

$$u_1 = (1, 0), \qquad u_2 = (0, -1), \tag{2.4.2}$$

and we write

$$\mathbb{R}^{1,1} = \mathbb{R}u_1 + \mathbb{R}u_2, \qquad \mathbb{Z}^{1,1} = \mathbb{Z}u_1 + \mathbb{Z}u_2. \tag{2.4.3}$$

We now want to extend the Dynkin diagram $\bar{\Gamma} = \Gamma(\bar{A})$ with one node in such a way that the new diagram $\bar{\Gamma}^{+}$ corresponds to the Dynkin diagram of an affine Kac-Moody algebra $\bar{\mathfrak{g}}^{+}$. This step was actually already discussed in Section 2.3, and we recall that an affine algebra can be obtained from any finite Lie algebra $\bar{\mathfrak{g}}$ by augmenting the set of simple roots with the *affine root*

$$\alpha_0 \equiv u_1 - \theta, \tag{2.4.4}$$

where θ is the highest root of $\bar{\mathfrak{g}}$ and $u_1 \in \mathbb{Z}^{1,1}$ corresponds to the null root δ. We now have

$$(\alpha_i|u_1) = 0, \quad \text{for} \quad i = 1, \ldots, k, \tag{2.4.5}$$

implying that the new root lattice

$$\bar{Q}^+ = \mathbb{Z}\alpha_0 + \bar{Q} \tag{2.4.6}$$

is contained in the direct sum of $\bar{Q}$ and $\mathbb{Z}^{1,1}$,

$$\bar{Q}^+ \subset \bar{Q} \oplus \mathbb{Z}^{1,1}. \tag{2.4.7}$$

We define new indices $A, B = (0, i)$, such that we can write the matrix of scalar products between the new simple roots $\bar{\Pi}^+ = \{\alpha_0, \alpha_1, \ldots, \alpha_k\}$ as follows

$$\bar{A}^+_{AB} = \frac{2(\alpha_A|\alpha_B)}{(\alpha_A|\alpha_A)}, \tag{2.4.8}$$

which then corresponds to the entries of the Cartan matrix $\bar{A}^+$ of the affine Kac-Moody algebra $\bar{\mathfrak{g}}^+ = \mathfrak{g}(\bar{A}^+)$.

Let us now proceed to include also the second basis vector $u_2 \in \mathbb{Z}^{1,1}$. This is done by adding the simple root

$$\alpha_{-1} \equiv -u_1 - u_2, \tag{2.4.9}$$

which has non-vanishing scalar product only with α_0:

$$(\alpha_{-1}|\alpha_0) = -1, \qquad (\alpha_{-1}|\alpha_i) = 0, \quad \forall\, i = 1, \ldots, k. \tag{2.4.10}$$

Since we also have $(\alpha_{-1}|\alpha_{-1}) = (\alpha_0|\alpha_0) = 2$ this implies that the node associated with α_{-1} attaches by a single link to the zeroth node of the Dynkin diagram $\bar{\Gamma}^+$ of $\bar{\mathfrak{g}}^+$. As before, we define new collective indices $I, J = (-1, 0, i)$, and the matrix of scalar products

$$\bar{A}^{++}_{IJ} = \frac{2(\alpha_I|\alpha_J)}{(\alpha_I|\alpha_I)} \tag{2.4.11}$$

is the Cartan matrix of the *Lorentzian* Kac-Moody algebra $\bar{\mathfrak{g}}^{++}$. The root lattice $\bar{Q}^{++}$ is then of Lorentzian signature and is equivalent to the direct sum of $\bar{Q}$ with $\mathbb{Z}^{1,1}$,

$$\bar{Q}^{++} = \bar{Q} \oplus \mathbb{Z}^{1,1}. \tag{2.4.12}$$

In order to obtain a triple extension, while still keeping the Lorentzian signature of the root lattice, we introduce yet another two-dimensional Lorentzian lattice $\tilde{\mathbb{Z}}^{1,1}$, spanned by the basis vectors v_1 and v_2. The scalar product on $\tilde{\mathbb{Z}}^{1,1}$ is of the same form as the one on $\mathbb{Z}^{1,1}$,

$$(v_1|v_2) = 1, \qquad (v_1|v_1) = 0, \qquad (v_2|v_2) = 0. \tag{2.4.13}$$

We now note that the vector $v_1 + v_2 \in \tilde{\mathbb{Z}}^{1,1}$ is spacelike, $(v_1 + v_2|v_1 + v_2) = 2$. Thus, by including this vector into the new root lattice we ensure that the Lorentzian signature is preserved, i.e., we do not introduce zero eigenvalues in the bilinear form. We augment the set of simple roots with the "triple-extended" root

$$\alpha_{-2} \equiv u_1 - (v_1 + v_2). \tag{2.4.14}$$

Again, this root is spacelike, $(\alpha_{-2}|\alpha_{-2}) = 2$, and the associated node in the Dynkin diagram connects with a single link to the node corresponding to α_{-1},

$$(\alpha_{-2}|\alpha_{-1}) = -1, \qquad (\alpha_{-2}|\alpha_A) = 0, \quad \forall\ A = 0, 1, \ldots, k. \tag{2.4.15}$$

The new Lorentzian root lattice is given by

$$\bar{Q}^{+++} = \mathbb{Z}\alpha_{-2} + \mathbb{Z}\alpha_{-1} + \mathbb{Z}\alpha_0 + \bar{Q}, \tag{2.4.16}$$

and we have

$$\bar{Q}^{+++} \ \subset\ \tilde{\mathbb{Z}}^{1,1} \oplus \mathbb{Z}^{1,1} \oplus \bar{Q}. \tag{2.4.17}$$

Introducing new indices $M, N = (-2, -1, 0, i)$ we can once again organize the scalar products as

$$\bar{A}^{+++}_{MN} = \frac{2(\alpha_M|\alpha_N)}{(\alpha_M|\alpha_M)}, \tag{2.4.18}$$

corresponding to the Cartan matrix $\bar{A}^{+++}$ of a Lorentzian Kac-Moody algebra $\bar{\mathfrak{g}}^{+++} = \mathfrak{g}(\bar{A}^{+++})$.

Nice examples of extended Lie algebras, which shall be discussed extensively later on, are the Kac-Moody algebras obtained by extending the largest exceptional Lie algebra E_8. This gives rise to the following chain of embeddings

$$E_8 \ \subset E_9 = E_8^{+} \ \subset\ E_{10} = E_8^{++} \ \subset\ E_{11} = E_8^{+++}, \tag{2.4.19}$$

of which E_9 is affine, E_{10} is hyperbolic and E_{11} is Lorentzian (but not hyperbolic). In the next section we shall focus on the subclass of hyperbolic Kac-Moody algebras. In particular, the hyperbolic algebra E_{10} will play an important role in **Part I** of this thesis, and is discussed in more detail in Section 2.5.4 below.

2.4.2 Hyperbolic Kac-Moody Algebras

A hyperbolic Kac-Moody algebra is defined as a Lorentzian algebra which upon removal of any node in the Dynkin diagram yields only finite or (at most one) affine subdiagrams. By this criterion it is easy to understand why E_{11} is not hyperbolic; removal of the node associated with the triple-extended root α_{-2} gives the Dynkin diagram of E_{10} which is neither finite nor affine. In this section we shall see that the class of hyperbolic algebras exhibit some very intriguing features which are unique among all indefinite Kac-Moody algebras.

In the rest of this section we shall take $\mathfrak{g}(A)$ to be a rank r hyperbolic Kac-Moody algebra, unless otherwise specified. By virtue of its Lorentzian signature the space $\mathfrak{h}^\star$ is isomorphic to r-dimensional Minkowski space:

$$\mathfrak{h}^\star \simeq \mathbb{R}^{1,r-1}. \tag{2.4.20}$$

An important consequence of this is that there exists a lightcone in $\mathfrak{h}^\star$, defined as

$$\mathcal{O} = \{x \in \mathfrak{h}^\star \mid (x|x) \leq 0\}. \tag{2.4.21}$$

The lightcone clearly separates real and imaginary roots

$$\Phi_{\Im} = \Phi \cap \mathcal{O}. \tag{2.4.22}$$

These properties are of course shared among all Lorentzian algebras. However, a unique feature of hyperbolic Kac-Moody algebras is that its root system is in principle known. By this we mean that any element of the root lattice $Q(A)$ is also an element of the root system $\Pi(A)$ if its norm is less than or equal to 2. In this way we find the following description of the root system:

$$\Phi = \{\alpha \in Q \mid (\alpha|\alpha) \leq 2\}. \tag{2.4.23}$$

Furthermore, we recall that the fundamental Weyl chamber $\mathcal{C}$ is defined as the region of $\mathfrak{h}^\star$ which is bounded by the hyperplanes T_i which are orthogonal to the simple roots α_i. As a consequence of their definition, hyperbolic Kac-Moody algebras have the special property that all these hyperplanes intersect *inside or on the lightcone.* Thus, for any hyperbolic Kac-Moody algebra the fundamental Weyl chamber is contained inside the lighcone

$$\mathcal{C} \ \subset \ \mathcal{O}. \tag{2.4.24}$$

Because of the Lorentzian signature, the lightcone decomposes into future and past components, $\mathcal{O}_+$ and $\mathcal{O}_-$, respectively. We shall employ the convention that the simple roots have future temporal directions, implying that the fundamental Weyl chamber, defined as $\{\beta \in \mathfrak{h}^\star \mid (\beta|\alpha_i) > 0,\ i = 1, \ldots, r\}$, is actually contained in the *past* lightcone. Since no simple roots are inside the lightcone, the union of all the images of the Weyl group, $\mathcal{W}$, acting on $\mathcal{C}$, will not extend outside of $\mathcal{C}$, and, in fact, we have that the Tits cone, $\mathcal{X}$, coincides with the past lighcone:

$$\mathcal{O} \equiv \mathcal{X} = \bigcup_{\omega \in \mathcal{W}} \omega(\mathcal{C}). \tag{2.4.25}$$

Because of this, the Weyl chamber $\mathcal{C}$ is not a fundamental domain for the action of $\mathcal{W}$ on all of $\mathfrak{h}^\star$, as is the case for finite Lie algebras, but rather is the fundamental domain for the action of $\mathcal{W}$ on the Tits cone, $\mathcal{X}$.

The Weyl group $\mathcal{W}(A)$ of a hyperbolic Kac-Moody algebra $\mathfrak{g}(A)$ is a discrete subgroup of the isometry group of $\mathfrak{h}^\star = \mathbb{R}^{1,r-1}$. Moreover, since all of the hyperplanes T_i are either timelike or lightlike the Weyl group preserves the temporal direction of any $x \in \mathfrak{h}^\star$. This implies that the Weyl group of a hyperbolic Kac-Moody algebra is a subgroup of the ortochronous Lorentz group, $O^\dagger(1, r-1)$, i.e., the time-preserving part of the isometry group of $\mathbb{R}^{1,r-1}$,

$$\mathcal{W} \ \subset \ O^\dagger(1, r-1). \tag{2.4.26}$$

Because of this fact, the Weyl group preserves spaces of constant negative curvature in $\mathcal{O}$, i.e., the r-dimensional hyperbolic space $\mathcal{H}_r$. The hyperplanes T_i project onto hyperplanes in $\mathcal{H}_r$ and because we then have $r+1$ hyperplanes bounding a region in an r-dimensional space, the Weyl chamber $\mathcal{C}$ projects onto a simplex of finite volume in $\mathcal{H}_r$. Geometric reflections in the faces of a finite volume simplex in hyperbolic space are elements of a *hyperbolic Coxeter group.* The associated Weyl groups therefore correspond to such hyperbolic Coxeter groups, and it is this fact which is the origin of the name hyperbolic Kac-Moody algebras.

2.4.3 Extended Example: The Weyl Group of A_1^{++}

We have previously discussed the Weyl groups of $\bar{\mathfrak{g}} = A_1$ and $\bar{\mathfrak{g}}^+ = A_1^+$ in some detail. Recall that we found:

$$\mathcal{W}(\bar{A}) = \mathbb{Z}_2, \qquad \mathcal{W}(\bar{A}^+) = \mathbb{Z}_2 \ltimes \bar{T}, \tag{2.4.27}$$

where $\bar{T}$ is the abelian group of translations of the Euclidean root lattice $\bar{Q}$. The first group $\mathbb{Z}_2$ is of course of finite order, while the second one, $\mathbb{Z}_2 \ltimes \bar{T}$, is of infinite order due to the presence of the translation group. Finite Coxeter groups, such as $\mathbb{Z}_2$, are often called *spherical* because they are reflections about the origin in a Euclidean space and so leaves invariant a sphere at infinity. Similarly the affine Coxeter groups leave the entire Euclidean spaces themselves invariant. Finally, on the other side we have the *hyperbolic* Coxeter groups which leave the hyperbolic space invariant. In this section we shall check this more explicitly by investigating the Weyl group of the hyperbolic Kac-Moody algebra $\mathfrak{g} = A_1^{++}$, i.e., the double extension of A_1. The Cartan matrix of this algebra is

$$A(A_1^{++}) = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -2 \\ 0 & -2 & 2 \end{pmatrix}, \tag{2.4.28}$$

and the associated Dynkin diagram $\Gamma(A_1^{++})$ is displayed in Figure 2.3. We associate a funda-

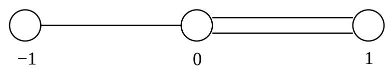


Figure 2.3: The Dynkin diagram of the hyperbolic Kac-Moody algebra A_1^{++}.

mental reflection s_I with each of the simple roots α_I, $I = -1, 0, 1$. By making use of Table 2.1 and (2.2.87) we find that these generators obey

$$(s_1 s_0)^\infty = 1, \qquad (s_0 s_{-1})^3 = 1, \qquad (s_1 s_{-1})^2 = 1. \tag{2.4.29}$$

The root space of A_1^{++} is three-dimensional and thus the (projected) Weyl chamber $\mathcal{C}$ is a simplex in the two-dimensional hyperbolic plane $\mathcal{H}_2 = \{\beta \in \mathfrak{h}^\star \mid (\beta|\beta) = -1\}$, bounded by the three hyperplanes T_{-1}, T_0 and T_1. The angle ϑ_{IJ} between two adjacent faces I and J are related to the Coxeter exponents, m_{IJ}, as follows (see, e.g., [73])

$$\vartheta_{IJ} = \frac{\pi}{m_{IJ}}. \tag{2.4.30}$$

In our case we thus find that the angle between T_1 and T_0 is zero, the one between T_1 and T_{-1} is $\pi/2$ and the one between T_0 and T_{-1} is $\pi/3$. This implies that the hyperplanes T_1 and T_0 intersect on the border of the lighcone, while the other walls intersect inside the lightcone. This verifies that A_1^{++} is indeed hyperbolic.

We now proceed to show that the Weyl group of A_1^{++} is isomorphic to the extended modular group $PGL(2,\mathbb{Z})$ (see, e.g, [73]). We define the group $PGL(2,\mathbb{Z})$ as the group of 2×2 matrices

$$PGL(2,\mathbb{Z}) \ni \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad a,b,c,d \in \mathbb{Z}, \tag{2.4.31}$$

with determinant $ad - bc = 1$ and the identification $\{a, b, c, d\} \sim \{-a, -b, -c, -d\}$. [The determinant of any element $X \in GL(2, \mathbb{Z})$ is restricted to $+1$ or -1 in order for the inverse X^{-1} to be contained in the group. In the projected group $PGL(2, \mathbb{Z})$ elements with determinant -1 are projected out.]

Now, we recall that the hyperbolic plane $\mathcal{H}_2$ has a realization as the complex upper half plane $\mathbb{H}$ (see, e.g, [74])

$$\mathbb{H} = \{\tau \in \mathbb{C} \mid \Im(\tau) > 0\}. \tag{2.4.32}$$

The group $PGL(2, \mathbb{Z})$ acts on $\tau \in \mathcal{H}_2$ as

$$\tau \longrightarrow \tau' = \frac{au + b}{cu + d}, \tag{2.4.33}$$

where we take

$$u = \begin{cases} \tau & ; \quad ad - bc = 1 \\ \bar{\tau} & ; \quad ad - bc = -1. \end{cases} \tag{2.4.34}$$

The reason for taking $u = \bar{\tau}$ when $ad - bc = -1$ is to ensure that the upper half plane is preserved, i.e.,

$$\Im(\tau) > 0 \quad \Longrightarrow \quad \Im(\tau') > 0. \tag{2.4.35}$$

We can think of the transformation (2.4.33) as the ordinary action of the modular group $PSL(2, \mathbb{Z}) \subset PGL(2, \mathbb{Z})$ together with the action of complex conjugation. Recall that $PSL(2, \mathbb{Z})$ is generated by the translation $\mathcal{T} : \tau \to \tau + 1$ and the inversion $\mathcal{I} : \tau \to -1/\tau$, with the realization

$$\mathcal{T} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \mathcal{I} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{2.4.36}$$

The generators of $PGL(2, \mathbb{Z})$ can now be obtained simply by adding the complex conjugation transformation $\tau \longrightarrow -\bar{\tau}$, i.e., we obtain the three generators:

$$\begin{aligned} r_1 &: \tau \longrightarrow -\bar{\tau} \\ r_0 \equiv r_1 \circ \mathcal{T} &: \tau \longrightarrow 1 - \bar{\tau} \\ r_{-1} \equiv r_1 \circ \mathcal{I} &: \tau \longrightarrow \frac{1}{\bar{\tau}}. \end{aligned} \tag{2.4.37}$$

These have the matrix realization:

$$r_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad r_0 = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, \qquad r_{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.4.38}$$

Using the explicit action of $PGL(2, \mathbb{Z})$ in (2.4.33) one may verify, e.g., that

$$(r_0 r_{-1}) \circ (r_0 r_{-1}) \circ (r_0 r_{-1}) \; : \; \tau \longrightarrow \tau \tag{2.4.39}$$

and similarly that $(r_1 r_{-1})^2 = 1$. We also have that no product of $(r_1 r_0)$ gives the identity, and so we have $(r_1 r_0)^\infty$. The group generated by r_1, r_0 and r_{-1} therefore coincides with the Weyl group of A_1^{++} and we may conclude that

$$\mathcal{W}(A_1^{++}) \simeq PGL(2, \mathbb{Z}). \tag{2.4.40}$$

Let us finally note that one can see that the groups are the same by comparing the geometric properties of the Weyl chamber with the fundamental domain for the action of $PGL(2,\mathbb{Z})$ on the upper half plane. To this end we write $\tau = x + iy \in \mathfrak{U}$, for $x, y \in \mathbb{R}$, and check that r_1 acts as

$$r_1 \ : \ x + iy \quad \longrightarrow \quad -x + iy, \tag{2.4.41}$$

which implies that this is a reflection in the "hyperplane" $W_1 = \{\tau \in \mathbb{C} \mid \Re(\tau) = 0\}$, i.e., a reflection in the line $x = 0$. By similar arguments one finds that the s_0 transformation is a reflection in the line $x = 1/2$ ($W_0 = \{\tau \in \mathbb{C} \mid \Re(\tau) = 1/2\}$), and that s_{-1} is a reflection in the unit circle $|\tau| = 1$ ($W_{-1} = \{\tau \in \mathbb{C} \mid |\tau| = 1\}$). The angle between W_1 and W_0 is therefore zero, the angle between W_1 and W_{-1} is $\pi/2$ and the angle between W_0 and W_{-1} is $\pi/3$. Hence the $PGL(2,\mathbb{Z})$-elements r_1, r_0 and r_{-1} generate a Coxeter group with Coxeter exponents $m_{10} = \infty, m_{1(-1)} = 2$ and $m_{0(-1)} = 3$, which, again, is the same as for the Weyl group of A_1^{++}.

Subgroups of $\mathcal{W}(A_1^{++})$

There exists a powerful method to obtain regular subalgebras of a Kac-Moody algebra $\mathfrak{g}$ by studying subgroups of its Weyl group $\mathcal{W}(\mathfrak{g})$. In this section we shall illustrate this procedure for the example of $\mathfrak{g} = A_1^{++}$.

Recall that since A_1^{++} is hyperbolic, all elements of the lightcone $\mathcal{C} \subset \mathfrak{h}$ are timelike or lightlike. This implies that reflections in these walls preserve the time-orientiation, and hence the Weyl group $\mathcal{W}(A_1^{++})$ is a subgroup of the ortochronous Lorentz group $O^+(2,1)$. As we saw above, the Weyl chamber $\mathcal{C}$ is contained in the future lightcone $\mathcal{O}_+ \subset \mathfrak{h}$, and hence the unit hyperboloid $\mathcal{H}_2 = \{\beta \in \mathcal{O}_+ \mid (\beta|\beta) = -1\}$ is invariant under $\mathcal{W}(A_1^{++})$. We may therefore project the Weyl chamber onto the hyperbolic space $\mathcal{H}_2$. This can be viewed as a projection from the interior of the future lightcone $\mathcal{O}_+$ onto the coset space $SL(2,\mathbb{R})/SO(2)$. The Weyl group of A_1^{++} is then realized as the group of reflections in $\mathcal{H}_2$, and the walls of the Weyl chamber correspond to the intersections of the hyperplanes W_1, W_0 and W_{-1}, defined above, with the unit hyperboloid. We have seen above that this Weyl group is isomorphic to the extended modular group $PGL(2,\mathbb{Z}) = GL(2,\mathbb{Z})/\mathbb{Z}_2$.

The action of $PGL(2,\mathbb{Z})$ on $\mathcal{C}$ gives rise to a tiling of the hyperbolic space into an infinite union of subregions, which are all images of the fundamental region $\mathcal{C}$ under $\mathcal{W}(A_1^{++})$. Each image $\omega(\mathcal{C}) \in \mathfrak{h}$, for $\omega \in \mathcal{W}(A_1^{++})$, is a simplex in $\mathcal{H}_2$, i.e., is bounded by three walls. This is indicated in Figure 2.4.

We shall here use the Weyl group $\mathcal{W}(A_1^{++})$ to exhibit two explicit embeddings of regular subalgebras of A_1^{++}. We consider the rank 3 hyperbolic Kac-Moody algebras $\mathfrak{g}(A_{1,I})$ and $\mathfrak{g}(A_{1,II})$ discussed by Gritsenko and Nikulin in the context of Siegel modular forms [75]. The second of these, $\mathfrak{g}(A_{1,II})$, has also played an important role recently in understanding the wall-crossing behaviour of BPS states in $D = 4$, $\mathcal{N} = 4$ string theory [76–79].

The two hyperbolic Kac-Moody algebras $\mathfrak{g}(A_{1,I})$ and $\mathfrak{g}(A_{1,II})$ are constructed from the Cartan matrices

$$A_{1,I} = \begin{pmatrix} 2 & -2 & -1 \\ -2 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \qquad A_{1,II} = \begin{pmatrix} 2 & -2 & -2 \\ -2 & 2 & -2 \\ -2 & -2 & 2 \end{pmatrix}. \tag{2.4.42}$$

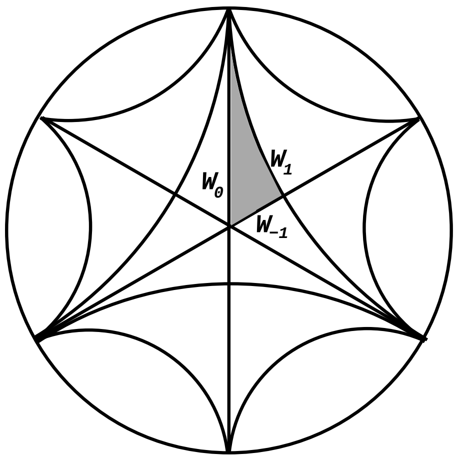


Figure 2.4: Tiling of the hyperbolic plane as defined by the action of the Weyl group $\mathcal{W}(A_{1,0})$ on the fundamental chamber $\mathcal{C} = \{\beta \in \mathfrak{h} \mid \alpha_i(\beta) \geq 0, i = 1, 0, -1\}$, with W_1, W_0, W_{-1} being the hyperplanes orthogonal to the simple roots $\alpha_1, \alpha_0, \alpha_{-1}$ of A_1^{++}.

The associated Dynkin diagrams are displayed in Figures 2.5 and 2.6.

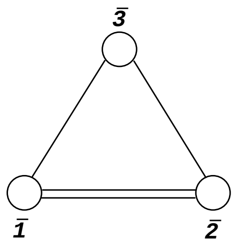


Figure 2.5: The Dynkin diagram of the rank 3 hyperbolic Kac-Moody algebra $\mathfrak{g}(A_{1,I})$. Its embedding into A_1^{++} is defined by $\bar{\alpha}_1 = 2\alpha_1 + \alpha_0, \bar{\alpha}_2 = \alpha_0, \bar{\alpha}_3 = \alpha_{-1}$, where $\alpha_1, \alpha_0, \alpha_{-1}$ are the simple roots of A_1^{++}.

The Weyl groups $\mathcal{W}(A_{1,I})$ and $\mathcal{W}(A_{1,II})$ are hyperbolic Coxeter groups with the following presentations,

$$\begin{aligned} \mathcal{W}(A_{1,I}) &= \left\langle s_i \mid s_i^2 = 1,\ (s_1 s_2)^\infty = 1,\ (s_1 s_3)^3 = 1,\ (s_2 s_3)^3 = 1, \right\rangle, \\ \mathcal{W}(A_{1,II}) &= \left\langle r_i \mid r_i^2 = 1,\ (r_1 r_2)^\infty = 1,\ (r_1 r_3)^\infty = 1,\ (r_2 r_3)^\infty = 1 \right\rangle. \end{aligned} \tag{2.4.43}$$

We begin by considering $\mathfrak{g}(A_{1,I})$. Define a new basis of simple roots $\bar{\Pi} \subset \Delta$ as follows

$$\bar{\alpha}_1 \equiv 2\alpha_1 + \alpha_0, \qquad \bar{\alpha}_2 \equiv \alpha_0, \qquad \bar{\alpha}_3 \equiv \alpha_{-1}. \tag{2.4.44}$$

Clearly, $\bar{\Delta} \subset \Delta$. The scalar products between the new simple roots give rise to the Cartan

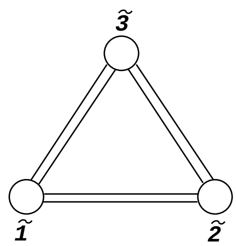


Figure 2.6: The Dynkin diagram of the rank 3 hyperbolic Kac-Moody algebra $\mathfrak{g}(A_{1,II})$. Its embedding into A_1^{++} is defined by $\tilde{\alpha}_1 = \alpha_1, \tilde{\alpha}_2 = \alpha_1 + 2\alpha_0, \tilde{\alpha}_3 = \alpha_1 + 2\alpha_0 + 2\alpha_{-1}$, where $\alpha_1, \alpha_0, \alpha_{-1}$ are the simple roots of A_1^{++}.

matrix $A_{1,I}$,

$$\frac{2(\bar{\alpha}_i|\bar{\alpha}_j)}{(\bar{\alpha}_i|\bar{\alpha}_i)} = \begin{pmatrix} 2 & -2 & -1 \\ -2 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} = A_{1,I}, \tag{2.4.45}$$

implying that the algebra $\mathfrak{g}(A_{1,I})$ is a regular subalgebra of $\mathfrak{g}(A_{1,0}) = A_1^{++}$, with the embedding defined by (2.4.44). We further note that the set of simple roots $\bar{\Pi}$ of $\mathfrak{g}(A_{1,I})$ is obtained by replacing α_1 by the image of α_0 under the fundamental reflection s_1,

$$s_1(\alpha_0) = 2\alpha_1 + \alpha_0, \tag{2.4.46}$$

implying that the fundamental chamber $\bar{\mathcal{C}}$ of $\mathcal{W}(A_{1,I})$ is twice as large as the original fundamental domain $\mathcal{C}$, and it corresponds to the union

$$\bar{\mathcal{C}} = \mathcal{C} \cup s_1(\mathcal{C}), \tag{2.4.47}$$

where

$$s_1(\mathcal{C}) = \{h \in \mathfrak{h} \mid \alpha_1(h) \leq 0, (2\alpha_1 + \alpha_0)(h) \geq 0, \alpha_{-1}(h) \geq 0\}. \tag{2.4.48}$$

Thus, the wall $W_1 = \{h \in \mathfrak{h} \mid \alpha_1(h) = 0\}$ "slices" the region $\bar{\mathcal{C}}$ into two subregions of equal size. This is indicated in Figure 2.7, where the new wall $\bar{W}_1$ is explicitly given by

$$\bar{W}_1 = \{h \in \mathfrak{h} \mid \bar{\alpha}_1(h) = (2\alpha_1 + \alpha_0)(h) = 0\}. \tag{2.4.49}$$

We thus deduce

$$\text{Area } \bar{\mathcal{C}} = 2 \cdot \text{Area } \mathcal{C} = \frac{\pi}{3}. \tag{2.4.50}$$

One way to see that the precise numerical factor is $\pi/3$ is to note that the fundamental domain of $\mathcal{C}$ is half that of the quotient $PSL(2,\mathbb{Z})\backslash\mathcal{H}_2$, which is the familiar fundamental domain for the action of $PSL(2,\mathbb{Z})$ on the upper half plane $\mathcal{H}_2$, which is well known to have area $\pi/3$ (see, e.g., [80, 81]).

We proceed to the case of $\mathfrak{g}(A_{1,II})$. To this end we define yet another set of simple roots $\tilde{\Pi}$ by

$$\tilde{\alpha}_1 = \alpha_1, \qquad \tilde{\alpha}_2 = \alpha_1 + 2\alpha_0, \qquad \tilde{\alpha}_3 = \alpha_1 + 2\alpha_0 + 2\alpha_{-1}. \tag{2.4.51}$$

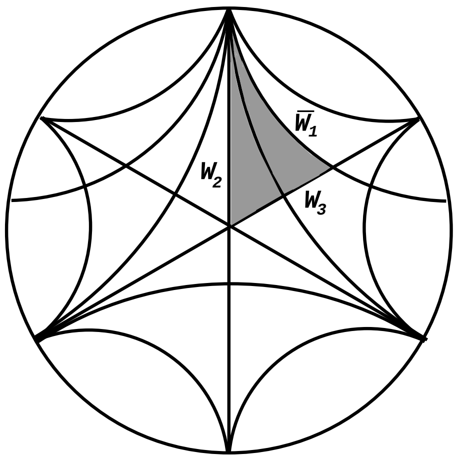


Figure 2.7: The fundamental Weyl chamber $\bar{\mathcal{C}}$ of $\mathcal{W}(A_{1,I})$ is shown as the union of $\mathcal{C}$ and $s_1(\mathcal{C})$, where $\mathcal{C}$ is the fundamental Weyl chamber of $\mathcal{W}(A_{1,0})$.

The normalized scalar products between these simple roots give rise to the Cartan matrix $A_{1,II}$,

$$\frac{2(\tilde{\alpha}_i|\tilde{\alpha}_j)}{(\tilde{\alpha}_i|\tilde{\alpha}_i)} = \begin{pmatrix} 2 & -2 & -2 \\ -2 & 2 & -2 \\ -2 & -2 & 2 \end{pmatrix} = A_{1,II}, \tag{2.4.52}$$

and we thereby deduce that $\mathfrak{g}(A_{1,II})$ is a regular subalgebra of A_1^{++} with the embedding defined by (2.4.51). The structure of the fundamental domain $\tilde{\mathcal{C}}$ of $\mathfrak{g}(A_{1,II})$ can be analyzed by considering its embedding inside the lightcone of A_1^{++}, as above. First note that $\tilde{\mathcal{C}}$ obviously contains the original fundamental domain $\mathcal{C}$ of $\mathcal{W}(A_{1,0})$ as a subregion. In fact, the fundamental Weyl chamber $\tilde{\mathcal{C}}$ can be embedded into the Cartan subalgebra $\mathfrak{h} \subset A_1^{++}$ as the following union of chambers,

$$\tilde{\mathcal{C}} = \bigcup_{i=1}^{6} \mathcal{C}_i, \tag{2.4.53}$$

where each chamber $\mathcal{C}_i$ is a copy of the fundamental chamber. Explicitly we have

$$\begin{aligned} \mathcal{C}_1 &= \mathcal{C}, \\ \mathcal{C}_2 &= s_2(\mathcal{C}_1), \\ \mathcal{C}_3 &= s_3 s_2 s_3(\mathcal{C}_2), \\ \mathcal{C}_4 &= s_3(\mathcal{C}_3), \\ \mathcal{C}_5 &= s_2(\mathcal{C}_4), \\ \mathcal{C}_6 &= s_3 s_2 s_3(\mathcal{C}_5). \end{aligned} \tag{2.4.54}$$

It is easy to see that the sequence stops at $\mathcal{C}_6$, since applying s_3 on $\mathcal{C}_6$ takes us back to $\mathcal{C}$, and any other reflection on $\mathcal{C}_6$ yields either a chamber outside of $\tilde{\mathcal{C}}$ or another of the six chambers of $\tilde{\mathcal{C}}$. We thus find that the size of $\tilde{\mathcal{C}}$ is

$$\text{Area}\,\tilde{\mathcal{C}} = 6 \cdot \text{Area}\,\mathcal{C} = \pi. \tag{2.4.55}$$

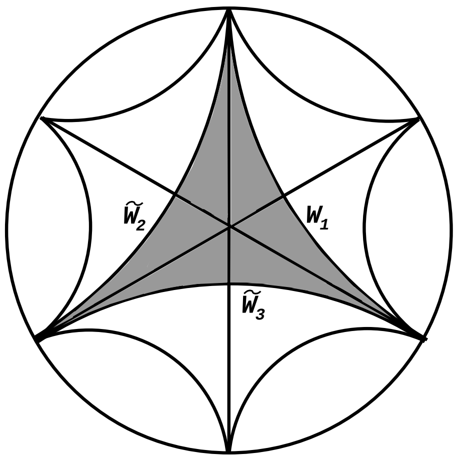


Figure 2.8: The fundamental Weyl chamber $\tilde{\mathcal{C}}$ of $\mathcal{W}(A_{1,II})$ is displayed as the ideal triangle inside the Cartan subalgebra of A_1^{++}.

The embedding $\tilde{\mathcal{C}} \subset \mathfrak{h}$ is displayed in Figure 2.8, and it corresponds to the so called "ideal triangle". In Chapter 4 we will learn that the union of chambers (2.4.53) has a nice description in the theory of buildings, and is known as a *gallery*.

2.5 Level Decomposition in Terms of Finite Regular Subalgebras

So far we have discussed general aspects of Kac-Moody algebras, with special emphasis on the Lorentzian subclass. When dealing with these infinite-dimensional structures in a physical context in Chapters 5, 6 and 7 it will prove to be very convenient to slice up the algebras into finite-dimensional subspaces $\mathfrak{g}_\ell$. More precisely, following [53], we will in this section define a so-called *level decomposition* of the adjoint representation of $\mathfrak{g}$ such that each level ℓ corresponds to a finite number of representations of a finite regular subalgebra $\mathfrak{r}$ of $\mathfrak{g}$. Generically the decomposition will take the form of the adjoint representation of $\mathfrak{r}$ plus a (possibly infinite) number of additional representations of $\mathfrak{r}$. This type of expansion of $\mathfrak{g}$ will prove to be very useful when considering sigma models invariant under $\mathfrak{g}$ for which we may use the level expansion to consistently truncate the theory to any finite level ℓ (see Chapter 5). We begin by illustrating these ideas for the finite-dimensional Lie algebra $\mathfrak{sl}(3,\mathbb{R})$ after which we generalize the procedure to the indefinite case in Sections 2.5.2, 2.5.3 and 2.5.4. This section is based on **Paper III**.

2.5.1 A finite-dimensional example: $\mathfrak{sl}(3,\mathbb{R})$

The rank 2 Lie algebra $\mathfrak{g} = \mathfrak{sl}(3,\mathbb{R})$ is characterized by the Cartan matrix

$$A[\mathfrak{sl}(3,\mathbb{R})] = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \tag{2.5.1}$$

whose Dynkin diagram is displayed in Figure 2.9.

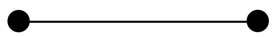


Figure 2.9: The Dynkin diagram of $\mathfrak{sl}(3,\mathbb{R})$.

We recall that $\mathfrak{sl}(3,\mathbb{R})$ is the split real form of $\mathfrak{sl}(3,\mathbb{C}) \equiv A_2$, and is thus defined through the same Chevalley–Serre presentation as for $\mathfrak{sl}(3,\mathbb{C})$, but with all coefficients restricted to the real numbers.

The Cartan generators $\{h_1, h_2\}$ will indifferently be denoted by $\{\alpha_1^\vee, \alpha_2^\vee\}$. As we have seen, they form a basis of the Cartan subalgebra $\mathfrak{h}$, while the simple roots $\{\alpha_1, \alpha_2\}$, associated with the raising operators e_1 and e_2, form a basis of the dual root space $\mathfrak{h}^\star$. Any root $\gamma \in \mathfrak{h}^\star$ can thus be decomposed in terms of the simple roots as follows,

$$\gamma = m\alpha_1 + \ell\alpha_2, \tag{2.5.2}$$

and the only values of (m,n) are $(1,0)$, $(0,1)$, $(1,1)$ for the positive roots and minus these values for the negative ones.

The algebra $\mathfrak{sl}(3,\mathbb{R})$ defines through the adjoint action a representation of $\mathfrak{sl}(3,\mathbb{R})$ itself, called the adjoint representation, which is eight-dimensional and denoted $\mathbf{8}$. The weights of the adjoint representation are the roots, plus the weight $(0,0)$ which is doubly degenerate. The lowest weight of the adjoint representation is

$$\Lambda_{\mathfrak{g}} = -\alpha_1 - \alpha_2, \tag{2.5.3}$$

corresponding to the generator $[f_1, f_2]$. We display the weights of the adjoint representation in Figure 2.10.

The idea of the level decomposition is to decompose the adjoint representation into representations of one of the regular $\mathfrak{sl}(2,\mathbb{R})$-subalgebras associated with one of the two simple roots α_1 or α_2, i.e., either $\{e_1, \alpha_1^\vee, f_1\}$ or $\{e_2, \alpha_2^\vee, f_2\}$. For definiteness we choose the level to count the number ℓ of times the root α_2 occurs, as was anticipated by the notation in Equation (2.5.2). Consider the subspace of the adjoint representation spanned by the vectors with a fixed value of ℓ. This subspace is invariant under the action of the subalgebra $\{e_1, \alpha_1^\vee, f_1\}$, which only changes the value of m. Vectors at a definite level transform accordingly in a representation of the regular $\mathfrak{sl}(2,\mathbb{R})$-subalgebra

$$\mathfrak{r} \equiv \mathbb{R}e_1 \oplus \mathbb{R}\alpha_1^\vee \oplus \mathbb{R}f_1. \tag{2.5.4}$$

Let us begin by analyzing states at level $\ell = 0$, i.e., with weights of the form $\gamma = m\alpha_1$. We see from Figure 2.10 that we are restricted to move along the horizontal axis in the root diagram. By the defining Lie algebra relations we know that $\mathrm{ad}_{f_1}(f_1) = 0$, implying that $\Lambda_{\mathrm{ad}}^{(0)} = -\alpha_1$ is a lowest weight of the $\mathfrak{sl}(2,\mathbb{R})$-representation. Here, the superscript 0 indicates that this is a level $\ell = 0$ representation. The corresponding complete irreducible module is found by acting on f_1 with e_1, yielding

$$[e_1, f_1] = \alpha_1^\vee, \qquad [e_1, \alpha_1^\vee] = -2e_1, \qquad [e_1, e_1] = 0. \tag{2.5.5}$$

We can then conclude that $\Lambda_{\mathrm{ad}}^{(0)} = -\alpha_1$ is the lowest weight of the three-dimensional adjoint representation $\mathbf{3}_0$ of $\mathfrak{sl}(2,\mathbb{R})$ with weights $\{\Lambda_{\mathrm{ad}}^{(0)}, 0, -\Lambda_{\mathrm{ad}}^{(0)}\}$, where the subscript on $\mathbf{3}_0$ again

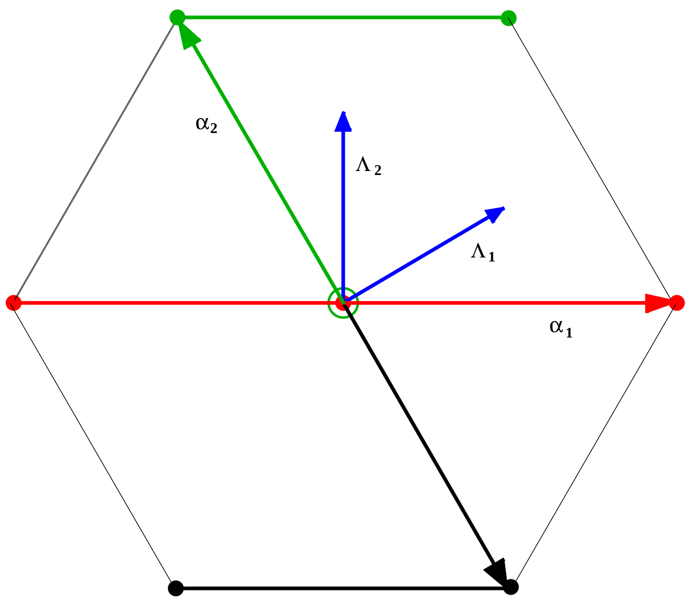


Figure 2.10: Level decomposition of the adjoint representation $\mathcal{R}_{\text{ad}} = \mathbf{8}$ of $\mathfrak{sl}(3,\mathbb{R})$ into representations of the subalgebra $\mathfrak{sl}(2,\mathbb{R})$. The labels 1 and 2 indicate the simple roots α_1 and α_2. Level zero corresponds to the horizontal axis where we find the adjoint representation $\mathcal{R}_{\text{ad}}^{(0)} = \mathbf{3}_0$ of $\mathfrak{sl}(2,\mathbb{R})$ (red nodes) and the singlet representation $\mathcal{R}_s^{(0)} = \mathbf{1}_0$ (green circle about the origin). At level one we find the two-dimensional representation $\mathcal{R}^{(1)} = \mathbf{2}_1$ (green nodes). The black arrow denotes the negative level root $-\alpha_2$ and so gives rise to the level $\ell = -1$ representation $\mathcal{R}^{(-1)} = \mathbf{2}_{(-1)}$. The blue arrows represent the fundamental weights Λ_1 and Λ_2.

indicates that this representation is located at level $\ell = 0$ in the decomposition. The module for this representation is $\mathcal{L}(\Lambda^{(0)}_{\text{ad}}) = \text{span}\{f_1, \alpha_1^\vee, e_1\}$.

This is, however, not the complete content at level zero since we must also take into account the Cartan generator $\alpha_2^\vee$ which remains at the origin of the root diagram. We can combine $\alpha_2^\vee$ with $\alpha_1^\vee$ into the vector

$$h = \alpha_1^\vee + 2\alpha_2^\vee, \tag{2.5.6}$$

which constitutes the one-dimensional singlet representation $\mathbf{1}_0$ of $\mathfrak{r}$ since it is left invariant under all generators of $\mathfrak{r}$,

$$[e_1, h] = [f_1, h] = [\alpha_1^\vee, h] = 0, \tag{2.5.7}$$

as follows trivially from the Chevalley relations. Thus level zero contains the representations $\mathbf{3}_0$ and $\mathbf{1}_0$.

Note that the vectors at level 0 not only transform in a (reducible) representation of $\mathfrak{sl}(2, \mathbb{R})$, but also form a subalgebra since the level is additive under taking commutators. The algebra in question is $\mathfrak{gl}(2, \mathbb{R}) = \mathfrak{sl}(2, \mathbb{R}) \oplus \mathbb{R}$. Accordingly, if the generator $\alpha_2^\vee$ is added to the subalgebra $\mathfrak{r}$, through the combination in Equation (2.5.6), so as to take the entire $\ell = 0$ subspace, $\mathfrak{r}$ is enlarged from $\mathfrak{sl}(2, \mathbb{R})$ to $\mathfrak{gl}(2, \mathbb{R})$, the generator h being somehow the "trace" part of $\mathfrak{gl}(2, \mathbb{R})$. This fact will prove to be important in subsequent sections.

Let us now ascend to the next level, $\ell = 1$. The weights of $\mathfrak{r}$ at level 1 take the general form $\gamma = m\alpha_1 + \alpha_2$ and the lowest weight is $\Lambda^{(1)} = \alpha_2$, which follows from the vanishing of the commutator

$$[f_1, e_2] = 0. \tag{2.5.8}$$

Note that $m \geq 0$ whenever $\ell > 0$ since $m\alpha_1 + \ell\alpha_2$ is then a positive root. The complete representation is found by acting on the lowest weight $\Lambda^{(1)}$ with e_1 and we get that the commutator $[e_1, e_2]$ is allowed by the Serre relations, while $[e_1, [e_1, e_2]]$ is killed, i.e.,

$$\begin{aligned} [e_1, e_2] &\neq 0, \\ [e_1, [e_1, e_2]]] &= 0. \end{aligned} \tag{2.5.9}$$

The non-vanishing commutator $e_\theta \equiv [e_1, e_2]$ is the vector associated with the highest root θ of $\mathfrak{sl}(3, \mathbb{R})$ given by

$$\theta = \alpha_1 + \alpha_2. \tag{2.5.10}$$

This is just the negative of the lowest weight $\Lambda_{\mathfrak{g}}$. The only representation at level one is thus the two-dimensional representation $\mathbf{2}_1$ of $\mathfrak{r}$ with weights $\{\Lambda^{(1)}, \theta\}$. The decomposition stops at level one for $\mathfrak{sl}(3, \mathbb{R})$ because any commutator with two e_2's vanishes by the Serre relations. The negative level representations may be found simply by applying the Chevalley involution and the result is the same as for level one.

Hence, the total level decomposition of $\mathfrak{sl}(3, \mathbb{R})$ in terms of the subalgebra $\mathfrak{sl}(2, \mathbb{R})$ is given by

$$\mathbf{8} = \mathbf{3}_0 \oplus \mathbf{1}_0 \oplus \mathbf{2}_1 \oplus \mathbf{2}_{(-1)}. \tag{2.5.11}$$

Although extremely simple (and familiar), this example illustrates well the situation encountered with more involved cases below. In the following analysis we will not mention the negative levels any longer because these can always be obtained simply through a reflection with respect to the $\ell = 0$ "hyperplane", using the Chevalley involution.

2.5.2 Some Formal Considerations

Before we proceed with a more involved example, let us formalize the procedure outlined above. We mainly follow the excellent treatment given in [82], although we restrict ourselves to the cases where $\mathfrak{r}$ is a *finite* regular subalgebra of $\mathfrak{g}$.

In the previous example, we performed the decomposition of the roots (and the ensuing decomposition of the algebra) with respect to one of the simple roots which then defined the level. In general, one may consider a similar decomposition of the roots of a rank r Kac–Moody algebra with respect to an arbitrary number $s < r$ of the simple roots and then the level ℓ is generalized to the "multilevel" $\ell = (\ell_1, \cdots, \ell_s)$.

Gradation

We consider a Kac–Moody algebra $\mathfrak{g}$ of rank r and we let $\mathfrak{r} \subset \mathfrak{g}$ be a finite regular rank $m < r$ subalgebra of $\mathfrak{g}$ whose Dynkin diagram is obtained by deleting a set of nodes $\mathcal{N} = \{n_1, \cdots, n_s\}$ $(s = r - m)$ from the Dynkin diagram of $\mathfrak{g}$.

Let γ be a root of $\mathfrak{g}$,

$$\gamma = \sum_{i \notin \mathcal{N}} m_i \alpha_i + \sum_{a \in \mathcal{N}} \ell_a \alpha_a. \tag{2.5.12}$$

To this decomposition of the roots corresponds a decomposition of the algebra, which is called a *gradation* of $\mathfrak{g}$ and which can be written formally as

$$\mathfrak{g} = \bigoplus_{\ell \in \mathbb{Z}^s} \mathfrak{g}_\ell, \tag{2.5.13}$$

where for a given ℓ, $\mathfrak{g}_\ell$ is the subspace spanned by all the vectors e_γ with that definite value ℓ of the multilevel,

$$[h, e_\gamma] = \gamma(h) e_\gamma, \qquad l_a(\gamma) = \ell_a. \tag{2.5.14}$$

Of course, if $\mathfrak{g}$ is finite-dimensional this sum terminates for some finite level, as in Equation (2.5.11) for $\mathfrak{sl}(3, \mathbb{R})$. However, in the following we shall mainly be interested in cases where Equation (2.5.13) is an infinite sum.

We note for further reference that the following structure is inherited from the gradation:

$$[\mathfrak{g}_\ell, \mathfrak{g}_{\ell'}] \subseteq \mathfrak{g}_{\ell + \ell'}. \tag{2.5.15}$$

This implies that for $\ell = 0$ we have

$$[\mathfrak{g}_0, \mathfrak{g}_{\ell'}] \subseteq \mathfrak{g}_{\ell'}, \tag{2.5.16}$$

which means that $\mathfrak{g}_{\ell'}$ is a representation of $\mathfrak{g}_0$ under the adjoint action. Furthermore, $\mathfrak{g}_0$ is a subalgebra. Now, the algebra $\mathfrak{r}$ is a subalgebra of $\mathfrak{g}_0$ and hence we also have

$$[\mathfrak{r}, \mathfrak{g}_{\ell'}] \subseteq \mathfrak{g}_{\ell'}, \tag{2.5.17}$$

so that *the subspaces $\mathfrak{g}_\ell$ at definite values of the multilevel are invariant subspaces under the adjoint action of* $\mathfrak{r}$. In other words, the action of $\mathfrak{r}$ on $\mathfrak{g}_\ell$ does not change the coefficients ℓ_a.

At level zero, $\ell = (0, \cdots, 0)$, the representation of the subalgebra $\mathfrak{r}$ in the subspace $\mathfrak{g}_0$ contains the adjoint representation of $\mathfrak{r}$, just as in the case of $\mathfrak{sl}(3, \mathbb{R})$ discussed in Section 2.5.1.

All positive and negative roots of $\mathfrak{r}$ are relevant. Level zero contains in addition s singlets for each of the Cartan generator associated to the set $\mathcal{N}$.

Whenever one of the ℓ_a's is positive, all the other ones must be non-negative for the subspace $\mathfrak{g}_\ell$ to be nontrivial and only positive roots appear at that value of the multilevel.

Weights of $\mathfrak{g}$ and Weights of $\mathfrak{r}$

Let V be the module of a representation $\mathcal{R}(\mathfrak{g})$ of $\mathfrak{g}$ and $\Lambda \in \mathfrak{h}^\star_\mathfrak{g}$ be one of the weights occurring in the representation. We define the action of $h \in \mathfrak{h}_\mathfrak{g}$ in the representation $\mathcal{R}(\mathfrak{g})$ on $x \in V$ as

$$h \cdot x = \Lambda(h)x \tag{2.5.18}$$

(we consider representations of $\mathfrak{g}$ for which one can speak of "weights" [34]). Any representation of $\mathfrak{g}$ is also a representation of $\mathfrak{r}$. When restricted to the Cartan subalgebra $\mathfrak{h}_\mathfrak{r}$ of $\mathfrak{r}$, Λ defines a weight $\bar{\Lambda} \in \mathfrak{h}^\star_\mathfrak{r}$, which one can realize geometrically as follows.

The dual space $\mathfrak{h}^\star_\mathfrak{r}$ may be viewed as the m-dimensional subspace Π of $\mathfrak{h}^\star_\mathfrak{g}$ spanned by the simple roots α_i, $i \notin \mathcal{N}$. The metric induced on that subspace is positive definite since $\mathfrak{r}$ is finite-dimensional. This implies, since we assume that the metric on $\mathfrak{h}^\star_\mathfrak{g}$ is non-degenerate, that $\mathfrak{h}^\star_\mathfrak{g}$ can be decomposed as the direct sum

$$\mathfrak{h}^\star_\mathfrak{g} = \mathfrak{h}^\star_\mathfrak{r} \oplus \Pi^\perp. \tag{2.5.19}$$

To that decomposition corresponds the decomposition

$$\Lambda = \Lambda^\parallel + \Lambda^\perp \tag{2.5.20}$$

of any weight, where $\Lambda^\parallel \in \mathfrak{h}^\star_\mathfrak{r} \equiv \Pi$ and $\Lambda^\perp \in \Pi^\perp$. Now, let $h = \sum k_i \alpha_i^\vee \in \mathfrak{h}_\mathfrak{r}$ $(i \notin \mathcal{N})$. One has $\Lambda(h) = \Lambda^\parallel(h) + \Lambda^\perp(h) = \Lambda^\parallel(h)$ because $\Lambda^\perp(h) = 0$: The component perpendicular to $\mathfrak{h}^\star_\mathfrak{r}$ drops out. Indeed, $\Lambda^\perp(\alpha_i^\vee) = \frac{2(\Lambda^\perp|\alpha_i)}{(\alpha_i|\alpha_i)} = 0$ for $i \notin \mathcal{N}$.

It follows that one can identify the weight $\bar{\Lambda} \in \mathfrak{h}^\star_\mathfrak{r}$ with the orthogonal projection $\Lambda^\parallel \in \mathfrak{h}^\star_\mathfrak{r}$ of $\Lambda \in \mathfrak{h}^\star_\mathfrak{g}$ on $\mathfrak{h}^\star_\mathfrak{r}$. This is true, in particular, for the fundamental weights Λ_i. The fundamental weights Λ_i project on 0 for $i \in \mathcal{N}$ and project on the fundamental weights $\bar{\Lambda}_i$ of the subalgebra $\mathfrak{r}$ for $i \notin \mathcal{N}$. These are also denoted λ_i. For a general weight, one has

$$\Lambda = \sum_{i \notin \mathcal{N}} p_i \Lambda_i + \sum_{a \in \mathcal{N}} k_a \Lambda_a \tag{2.5.21}$$

and

$$\bar{\Lambda} = \Lambda^\parallel = \sum_{i \notin \mathcal{N}} p_i \lambda_i. \tag{2.5.22}$$

The coefficients p_i can easily be extracted by taking the scalar product with the simple roots,

$$p_i = \frac{2}{(\alpha_i|\alpha_i)}(\alpha_i|\Lambda), \tag{2.5.23}$$

a formula that reduces to

$$p_i = (\alpha_i|\Lambda) \tag{2.5.24}$$

in the simply-laced case. Note that $(\Lambda^\parallel|\Lambda^\parallel) > 0$ even when Λ is non-spacelike.

Outer Multiplicity

There is an interesting relationship between root multiplicities in the Kac–Moody algebra $\mathfrak{g}$ and weight multiplicites of the corresponding $\mathfrak{r}$-weights, which we will explore here.

For finite Lie algebras, the roots always come with multiplicity one. This is in fact true also for the real roots of indefinite Kac–Moody algebras. However, recall from Sections 2.2.3 and 2.2.6 that imaginary roots can have arbitrarily large multiplicity. This must therefore be taken into account in the sum (2.5.13).

Let $\gamma \in \mathfrak{h}_{\mathfrak{g}}^{\star}$ be a root of $\mathfrak{g}$. There are two important ingredients:

- The multiplicity $\mathrm{mult}(\gamma)$ of each $\gamma \in \mathfrak{h}_{\mathfrak{g}}^{\star}$ at level ℓ as a *root* of $\mathfrak{g}$.
- The multiplicity $\mathrm{mult}_{\mathcal{R}_{\gamma}^{(\ell)}}(\gamma)$ of the corresponding weight $\bar{\gamma} \in \mathfrak{h}_{\mathfrak{r}}^{\star}$ at level ℓ as a *weight* in the representation $\mathcal{R}_{\gamma}^{(\ell)}$ of $\mathfrak{r}$. (Note that two distinct roots at the same level project on two distinct $\mathfrak{r}$-weights, so that given the $\mathfrak{r}$-weight and the level, one can reconstruct the root.)

It follows that the root multiplicity of γ is given as a sum over its multiplicities as a weight in the various representations $\{\mathcal{R}_q^{(\ell)} \,|\, q = 1, \cdots, N_\ell\}$ at level ℓ. Some representations can appear more than once at each level, and it is therefore convenient to introduce a new measure of multiplicity, called the *outer multiplicity* $\mu(\mathcal{R}_q^{(\ell)})$, which counts the number of times each representation $\mathcal{R}_q^{(\ell)}$ appears at level ℓ. So, for each representation at level ℓ we must count the individual weight multiplicities in that representation and also the number of times this representation occurs. The total multiplicity of γ can then be written as

$$\mathrm{mult}(\gamma) = \sum_{q=1}^{N_\ell} \mu(\mathcal{R}_q^{(\ell)})\, \mathrm{mult}_{\mathcal{R}_q^{(\ell)}}(\gamma). \tag{2.5.25}$$

This simple formula might provide useful information on which representations of $\mathfrak{r}$ are allowed within $\mathfrak{g}$ at a given level. For example, if γ is a real root of $\mathfrak{g}$, then it has multiplicity one. This means that in the formula (2.5.25), only the representations of $\mathfrak{r}$ for which γ has weight multiplicity equal to one are permitted. The others have $\mu(\mathcal{R}_q^{(\ell)}) = 0$. Furthermore, only one of the permitted representations does actually occur and it has necessarily outer multiplicity equal to one, $\mu(\mathcal{R}_q^{(\ell)}) = 1$.

The subspaces $\mathfrak{g}_\ell$ can now be written explicitly as

$$\mathfrak{g}_\ell = \bigoplus_{q=1}^{N_\ell} \left[\bigoplus_{k=1}^{\mu(\mathcal{R}_q^{(\ell)})} \mathcal{L}^{[k]}(\Lambda_q^{(\ell)}) \right], \tag{2.5.26}$$

where $\mathcal{L}(\Lambda_q^{(\ell)})$ denotes the module of the representation $\mathcal{R}_q^{(\ell)}$ and N_ℓ is the number of inequivalent representations at level ℓ. It is understood that if $\mu(\mathcal{R}_q^{(\ell)}) = 0$ for some ℓ and q, then $\mathcal{L}(\Lambda_q^{(\ell)})$ is absent from the sum. Note that the superscript $[k]$ labels multiple modules associated to the same representation, e.g., if $\mu(\mathcal{R}_q^{(\ell)}) = 3$ this contributes to the sum with a term

$$\mathcal{L}^{[1]}(\Lambda_q^{(\ell)}) \oplus \mathcal{L}^{[2]}(\Lambda_q^{(\ell)}) \oplus \mathcal{L}^{[3]}(\Lambda_q^{(\ell)}). \tag{2.5.27}$$

Finally, we mention that the multiplicity $\text{mult}(\alpha)$ of a root $\alpha \in \mathfrak{h}^\star$ can be computed recursively using the *Peterson recursion relation*, defined as [34]

$$(\alpha|\alpha - 2\,\rho)c_\alpha = \sum_{\substack{\gamma + \gamma' = \alpha \\ \gamma, \gamma' \in Q_+}} (\gamma|\gamma')c_\gamma c_{\gamma'}, \tag{2.5.28}$$

where Q_+ denotes the set of all positive integer linear combinations of the simple roots, i.e., the positive part of the root lattice, and ρ is the Weyl vector (see Eq. (2.2.37). The coefficients c_γ are defined as

$$c_\gamma = \sum_{k \geq 1} \frac{1}{k} \,\text{mult}\left(\frac{\gamma}{k}\right), \tag{2.5.29}$$

and, following [83], we call this the *co-multiplicity.* Note that if γ/k is not a root, this gives no contribution to the co-multiplicity. Another feature of the co-multiplicity is that even if the multiplicity of some root γ is zero, the associated co-multiplicity c_γ does not necessarily vanish. Taking advantage of the fact that all real roots have multiplicity one it is possible, in principle, to compute recursively the multiplicity of any imaginary root. Since no closed formula exists for the outer multiplicity μ, one must take a detour via the Peterson relation and Equation (2.5.25) in order to find the outer multiplicity of each representation at a given level.

2.5.3 Level Decomposition of A_1^{++}

The Kac–Moody algebra A_1^{++} is one of the simplest hyperbolic algebras and so provides a nice testing ground for investigating general properties of hyperbolic Kac–Moody algebras. From a physical point of view, it is the Weyl group of A_1^{++} which governs the chaotic behavior of pure four-dimensional gravity close to a spacelike singularity [52], as we have explained. Moreover, as we saw in Section 2.4.3, the Weyl group of A_1^{++} is isomorphic to the well-known arithmetic group $PGL(2, \mathbb{Z})$.

The level decomposition of $\mathfrak{g} = A_1^{++}$ follows a similar route as for $\mathfrak{sl}(3, \mathbb{R})$ above, but the result is much more complicated due to the fact that A_1^{++} is infinite-dimensional. This decomposition has been treated before in [51]. Recall that the Cartan matrix for A_1^{++} is given by

$$\begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \tag{2.5.30}$$

and the associated Dynkin diagram is given in Figure 2.11.

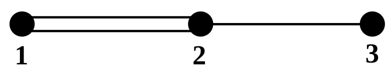


Figure 2.11: The Dynkin diagram of the hyperbolic Kac–Moody algebra $A_1^{++} \equiv A_1^{++}$. The labels indicate the simple roots α_1, α_2 and α_3. The nodes “2” and “3” correspond to the subalgebra $\mathfrak{r} = \mathfrak{sl}(3, \mathbb{R})$ with respect to which we perform the level decomposition.

We see that there exist three rank 2 regular subalgebras that we can use for the decomposition: $A_2, A_1 \oplus A_1$ or A_1^+. We will here focus on the decomposition into representations of $\mathfrak{r} = A_2 = \mathfrak{sl}(3, \mathbb{R})$ because this is the one relevant for pure gravity in four dimensions [52][4]. The level ℓ is then the coefficient in front of the simple root α_1 in an expansion of an arbitrary root $\gamma \in \mathfrak{h}_{\mathfrak{g}}^\star$, i.e.,

$$\gamma = \ell\alpha_1 + m_2\alpha_2 + m_3\alpha_3. \tag{2.5.31}$$

We restrict henceforth our analysis to positive levels only, $\ell \geq 0$. Before we begin, let us develop an intuitive idea of what to expect. We know that at each level we will have a set of finite-dimensional representations of the subalgebra $\mathfrak{r}$. The corresponding weight diagrams will then be represented in a Euclidean two-dimensional lattice in exactly the same way as in Figure 2.10 above. The level ℓ can be understood as parametrizing a third direction that takes us into the full three-dimensional root space of A_1^{++}. We display the level decomposition up to positive level two in Figure 2.12.

From previous sections we recall that A_1^{++} is hyperbolic so its root space is of Lorentzian signature. This implies that there is a lightcone in $\mathfrak{h}_{\mathfrak{g}}^\star$ whose origin lies at the origin of the root diagram for the adjoint representation of $\mathfrak{r}$ at level $\ell = 0$. The lightcone separates real roots from imaginary roots and so it is clear that if a representation at some level ℓ intersects the walls of the lightcone, this means that some weights in the representation will correspond to imaginary roots of $\mathfrak{h}_{\mathfrak{g}}^\star$ but will be real as weights of $\mathfrak{h}_{\mathfrak{r}}^\star$. On the other hand if a weight lies outside of the lightcone it will be real both as a root of $\mathfrak{h}_{\mathfrak{g}}^\star$ and as a weight of $\mathfrak{h}_{\mathfrak{r}}^\star$.

Level $\ell = 0$

Consider first the representation content at level zero. Given our previous analysis we expect to find the adjoint representation of $\mathfrak{r}$ with the additional singlet representation from the Cartan generator $\alpha_1^\vee$. The Chevalley generators of $\mathfrak{r}$ are $\{e_2, f_2, e_3, f_3, \alpha_2^\vee, \alpha_3^\vee\}$ and the generators associated to the root defining the level are $\{e_1, f_1, \alpha_1^\vee\}$. As discussed previously, the additional Cartan generator $\alpha_1^\vee$ that sits at the origin of the root space enlarges the subalgebra from $\mathfrak{sl}(3, \mathbb{R})$ to $\mathfrak{gl}(3, \mathbb{R})$. A canonical realisation of $\mathfrak{gl}(3, \mathbb{R})$ is obtained by defining the Chevalley generators in terms of the matrices ${K^i}_j$ $(i, j = 1, 2, 3)$ whose commutation relations are

$$[{K^i}_j, {K^k}_l] = \delta_j^k {K^i}_l - \delta_l^i {K^k}_j. \tag{2.5.32}$$

All the defining Lie algebra relations of $\mathfrak{gl}(3, \mathbb{R})$ are then satisfied if we make the identifications

$$\begin{array}{rclrclrcl} & & & & & & \alpha_1^\vee &=& {K^1}_1 - {K^2}_2 - {K^3}_3, \\ e_2 &=& {K^2}_1, & f_2 &=& {K^1}_2, & \alpha_2^\vee &=& {K^2}_2 - {K^1}_1, \\ e_3 &=& {K^3}_2, & f_3 &=& {K^2}_3, & \alpha_3^\vee &=& {K^3}_3 - {K^2}_2. \end{array} \tag{2.5.33}$$

Note that the trace ${K^1}_1 + {K^2}_2 + {K^3}_3$ is equal to $-4\alpha_2^\vee - 2\alpha_3^\vee - 3\alpha_1^\vee$. The generators e_1 and f_1 can of course not be realized in terms of the matrices ${K^i}_j$ since they do not belong to level zero. The invariant bilinear form $(\ |\ )$ at level zero reads

$$({K^i}_j | {K^k}_l) = \delta_l^i \delta_j^k - \delta_j^i \delta_l^k, \tag{2.5.34}$$

[4] The decomposition of A_1^{++} into representations of A_1^+ was done in [84].

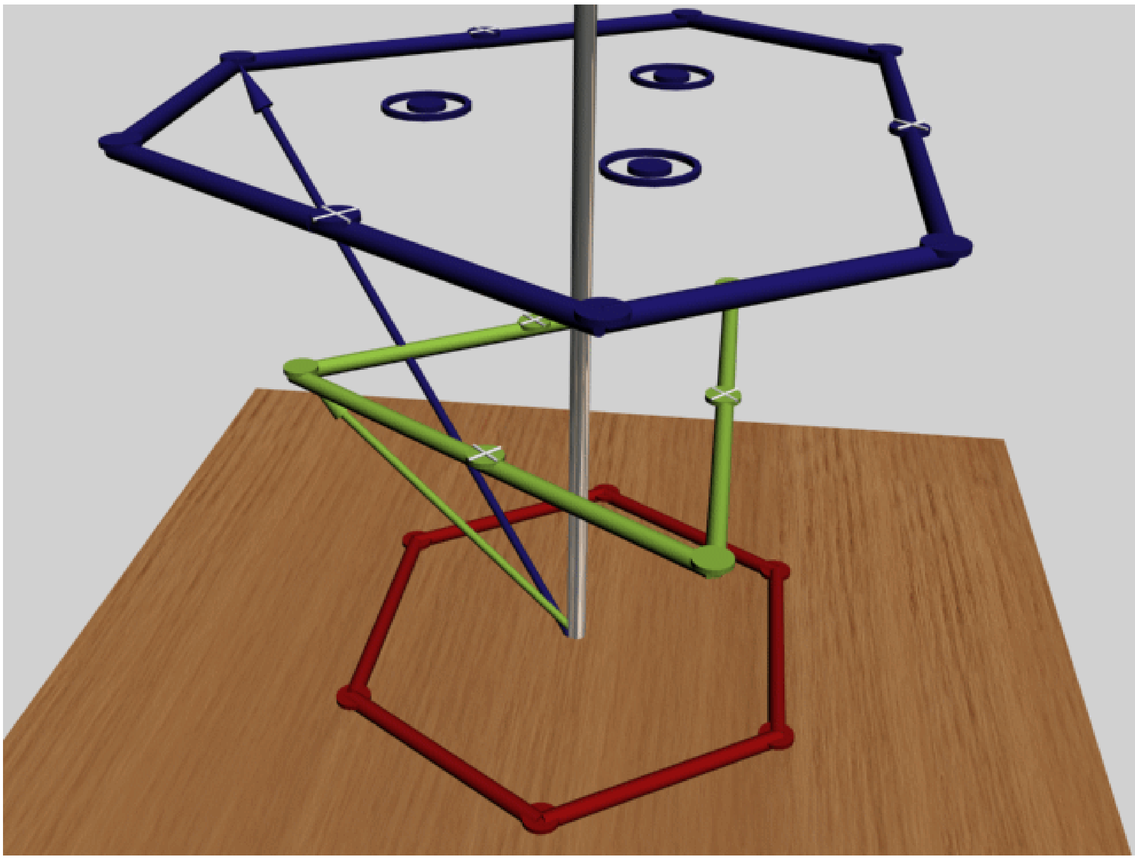


Figure 2.12: Level decomposition of the adjoint representation of A_1^{++}. We have displayed the decomposition up to positive level $\ell = 2$. At level zero we have the adjoint representation $\mathcal{R}_1^{(0)} = \mathbf{8}_0$ of $\mathfrak{sl}(3,\mathbb{R})$ and the singlet representation $\mathcal{R}_2^{(0)} = \mathbf{1}_0$ defined by the simple Cartan generator $\alpha_1^\vee$. Ascending to level one with the root α_1 (green vector) gives the lowest weight $\Lambda^{(1)}$ of the representation $\mathcal{R}^{(1)} = \mathbf{6}_1$. The weights of $\mathcal{R}^{(1)}$ labelled by white crosses are on the lightcone and so their norm squared is zero. At level two we find the lowest weight $\Lambda^{(2)}$ (blue vector) of the 15-dimensional representation $\mathcal{R}^{(2)} = \mathbf{15}_2$. Again, the white crosses label weights that are on the lightcone. The three innermost weights are inside of the lightcone and the rings indicate that these all have multiplicity 2 as weights of $\mathcal{R}^{(2)}$. Since these also have multiplicity 2 as *roots* of $\mathfrak{h}_{\mathfrak{g}}^\star$ we find that the outer multiplicity of this representation is one, $\mu(\mathcal{R}^{(2)}) = 1$.

where the coefficient in front of the second term on the right hand side is fixed to -1 through the embedding of $\mathfrak{gl}(3,\mathbb{R})$ in A_1^{++}.

The commutation relations in Equation (2.5.32) characterize the adjoint representation of $\mathfrak{gl}(3,\mathbb{R})$ as was expected at level zero, which decomposes as the representation $\mathcal{R}_{\text{ad}}^{(0)} \oplus \mathcal{R}_s^{(0)}$ of $\mathfrak{sl}(3,\mathbb{R})$ with $\mathcal{R}_{\text{ad}}^{(0)} = \mathbf{8}_0$ and $\mathcal{R}_s^{(0)} = \mathbf{1}_0$.

Dynkin Labels

It turns out that at each positive level ℓ, the weight that is easiest to identify is the lowest weight. For example, at level one, the lowest weight is simply α_1 from which one builds all the other weights by adding appropriate positive combinations of the roots α_2 and α_3. It will therefore turn out to be convenient to characterize the representations at each level by their (conjugate) Dynkin labels p_2 and p_3 defined as the coefficients of minus the (projected) lowest weight $-\bar{\Lambda}_{\text{lw}}^{(\ell)}$ expanded in terms of the fundamental weights λ_2 and λ_3 of $\mathfrak{sl}(3,\mathbb{R})$ (blue arrows in Figure 2.13),

$$-\bar{\Lambda}_{\text{lw}}^{(\ell)} = p_2\lambda_2 + p_3\lambda_3. \tag{2.5.35}$$

Note that for any weight Λ we have the inequality

$$(\Lambda|\Lambda) \leq (\bar{\Lambda}|\bar{\Lambda}) \tag{2.5.36}$$

since $(\Lambda|\Lambda) = (\bar{\Lambda}|\bar{\Lambda}) - |(\Lambda^{\perp}|\Lambda^{\perp})|$.

The Dynkin labels can be computed using the scalar product $(\ |\ )$ in $\mathfrak{h}_{\mathfrak{g}}^{\star}$ in the following way:

$$p_2 = -(\alpha_2|\Lambda_{\text{lw}}^{(\ell)}), \qquad p_3 = -(\alpha_3|\Lambda_{\text{lw}}^{(\ell)}). \tag{2.5.37}$$

For the level zero sector we therefore have

$$\begin{aligned} \mathbf{8}_0 \ &: \ [p_2,p_3] = [1,1], \\ \mathbf{1}_0 \ &: \ [p_2,p_3] = [0,0]. \end{aligned} \tag{2.5.38}$$

The module for the representation $\mathbf{8}_0$ is realized by the eight traceless generators $K^i{}_j$ of $\mathfrak{sl}(3,\mathbb{R})$ and the module for the representation $\mathbf{1}_0$ corresponds to the "trace" $\alpha_1^{\vee}$.

Note that the highest weight Λ_{hw} of a given representation of $\mathfrak{r}$ is not in general equal to minus the lowest weight Λ of the same representation. In fact, $-\Lambda_{\text{hw}}$ is equal to the lowest weight of the *conjugate* representation. This is the reason our Dynkin labels are really the conjugate Dynkin labels in standard conventions. It is only if the representation is self-conjugate that we have $\Lambda_{\text{hw}} = -\Lambda$. This is the case for example in the adjoint representation $\mathbf{8}_0$.

It is interesting to note that since the weights of a representation at level ℓ are related by Weyl reflections to weights of a representation at level $-\ell$, it follows that the negative of a lowest weight $\Lambda^{(\ell)}$ at level ℓ is actually equal to the *highest* weight $\Lambda_{\text{hw}}^{(-\ell)}$ of the conjugate representation at level $-\ell$. Therefore, the Dynkin labels at level ℓ as defined here are the standard Dynkin labels of the representations at level $-\ell$.

Level $\ell = 1$

We now want to exhibit the representation content at the next level $\ell = 1$. A generic level one commutator is of the form $[e_1, [\cdots[\cdots]]]$, where the ellipses denote (positive) level zero generators. Hence, including the generator e_1 implies that we step upwards in root space, i.e., in the direction of the forward lightcone. The root vector e_1 corresponds to a lowest weight of $\mathfrak{r}$ since it is annihilated by f_2 and f_3,

$$\begin{aligned} \mathrm{ad}_{f_2}(e_1) &= [f_2, e_1] = 0, \\ \mathrm{ad}_{f_3}(e_1) &= [f_3, e_1] = 0, \end{aligned} \tag{2.5.39}$$

which follows from the defining relations of A_1^{++}.

Explicitly, the root associated to e_1 is simply the root α_1 that defines the level expansion. Therefore the lowest weight of this level one representation is

$$\Lambda^{(1)}_{\mathrm{lw}} = \bar{\alpha}_1, \tag{2.5.40}$$

Although α_1 is a real *positive* root of $\mathfrak{h}^\star_{\mathfrak{g}}$, its projection $\bar{\alpha}_{(1)}$ is a *negative* weight of $\mathfrak{h}^\star_{\mathfrak{r}}$. Note that since the lowest weight $\Lambda^{(1)}_1$ is real, the representation $\mathcal{R}^{(1)}$ has outer multiplicity one, $\mu(\mathcal{R}^{(1)}) = 1$.

Acting on the lowest weight state with the raising operators of $\mathfrak{r}$ yields the six-dimensional representation $\mathcal{R}^{(1)} = \mathbf{6}_1$ of $\mathfrak{sl}(3, \mathbb{R})$. The root α_1 is displayed as the green vector in Figure 2.12, taking us from the origin at level zero to the lowest weight of $\mathcal{R}^{(1)}$. The Dynkin labels of this representation are

$$\begin{aligned} p_2(\mathcal{R}^{(1)}) &= -(\alpha_2|\alpha_1) = 2, \\ p_3(\mathcal{R}^{(1)}) &= -(\alpha_3|\alpha_1) = 0, \end{aligned} \tag{2.5.41}$$

which follows directly from the Cartan matrix of A_1^{++}. Three of the weights in $\mathcal{R}^{(1)}$ correspond to roots that are located on the lightcone in root space and so are null roots of $\mathfrak{h}^\star_{\mathfrak{g}}$. These are $\alpha_1+\alpha_2$, $\alpha_1+\alpha_2+\alpha_3$ and $\alpha_1+2\alpha_2+\alpha_3$ and are labelled with white crosses in Figure 2.12. The other roots present in the representation, in addition to α_1, are $\alpha_1+2\alpha_2$ and $\alpha_1+2\alpha_2+2\alpha_3$, which are real. This representation therefore contains no weights inside the lightcone.

The $\mathfrak{gl}(3, \mathbb{R})$-generator encoding this representation is realized as a symmetric 2-index tensor E^{ij} which indeed carries six independent components. In general we can easily compute the dimensionality of a representation given its Dynkin labels using the *Weyl dimension formula* which for $\mathfrak{sl}(3, \mathbb{R})$ takes the form [85]

$$d_{\Lambda_{\mathrm{hw}}}\left(\mathfrak{sl}(3, \mathbb{R})\right) = (p_2+1)(p_3+1)\left(\frac{1}{2}(p_2+p_3)+1\right). \tag{2.5.42}$$

In particuar, for $(p_2, p_3) = (2, 0)$ this gives indeed $d_{\Lambda^{(1)}_{\mathrm{hw},1}} = 6$.

It is convenient to encode the Dynkin labels, and, consequently, the index structure of a given representation module, in a Young tableau. We follow conventions where the first Dynkin label gives the number of columns with 1 box and the second Dynkin label gives the number of columns with 2 boxes[5]. For the representation $\mathbf{6}_1$ the first Dynkin label is 2 and the second is 0, hence the associated Young tableau is

$$\mathbf{6}_1 \quad \Longleftrightarrow \quad \square\!\square\,. \tag{2.5.43}$$

At level $\ell = -1$ there is a corresponding negative generator F_{ij}. The generators E^{ij} and F_{ij} transform contravariantly and covariantly, respectively, under the level zero generators, i.e.,

$$\begin{aligned} [K^i{}_j, E^{kl}] &= \delta^k_j E^{il} + \delta^l_j E^{ki}, \\ [K^i{}_j, F_{kl}] &= -\delta^i_k F_{jl} - \delta^i_l F_{kj}. \end{aligned} \tag{2.5.44}$$

The internal commutator on level one can be obtained by first identifying

$$e_1 \equiv E^{11}, \qquad f_1 \equiv F_{11}, \tag{2.5.45}$$

and then by demanding $[e_1, f_1] = \alpha_1^\vee$ we find

$$[E^{ij}, F_{kl}] = 2\delta^{(i}_{(k} K^{j)}{}_{l)} - \delta^{(i}_k \delta^{k)}_l (K^1{}_1 + K^2{}_2 + K^3{}_3), \tag{2.5.46}$$

which is indeed compatible with the realisation of $\alpha_1^\vee$ given in Equation (2.5.33). The Killing form at level 1 takes the form

$$\left(F_{ij} | E^{kl}\right) = \delta_i^{(k} \delta_j^{l)}. \tag{2.5.47}$$

Constraints on Dynkin Labels

As we go to higher and higher levels it is useful to employ a systematic method to investigate the representation content. It turns out that it is possible to derive a set of equations whose solutions give the Dynkin labels for the representations at each level [53].

We begin by relating the Dynkin labels to the expansion coefficients ℓ, m_2 and m_3 of a root $\gamma \in \mathfrak{h}^\star_{\mathfrak{g}}$, whose projection $\bar{\gamma}$ onto $\mathfrak{h}^\star_{\mathfrak{r}}$ is a lowest weight vector for some representation of $\mathfrak{r}$ at level ℓ. We let $a = 2, 3$ denote indices in the root space of the subalgebra $\mathfrak{sl}(3, \mathbb{R})$ and we let $i = 1, 2, 3$ denote indices in the full root space of A_1^{++}. The formula for the Dynkin labels then gives

$$p_a = -(\alpha_a | \gamma) = -\ell A_{a1} - m_2 A_{a2} - m_3 A_{a3}, \tag{2.5.48}$$

where A_{ij} is the Cartan matrix for A_1^{++}, given in Equation (2.5.30). Explicitly, we find the following relations between the coefficients m_2, m_3 and the Dynkin labels:

$$\begin{aligned} p_2 &= 2\ell - 2m_2 + m_3, \\ p_3 &= m_2 - 2m_3. \end{aligned} \tag{2.5.49}$$

These formulae restrict the possible Dynkin labels for each ℓ since the coefficients m_2 and m_3 must necessarily be non-negative integers. Therefore, by inverting Equation (2.5.49) we obtain two Diophantine equations that restrict the possible Dynkin labels,

$$\begin{aligned} m_2 &= \frac{4}{3}\ell - \frac{2}{3}p_2 - \frac{1}{3}p_3 \geq 0, \\ m_3 &= \frac{2}{3}\ell - \frac{1}{3}p_2 - \frac{2}{3}p_3 \geq 0. \end{aligned} \tag{2.5.50}$$

[5] Since we are, in fact, using conjugate Dynkin labels, these conventions are equivalent to the standard ones if one replaces covariant indices by contravariant ones, and vice-versa.

In addition to these constraints we can also make use of the fact that we are decomposing the adjoint representation of A_1^{++}. Since the weights of the adjoint representation are the roots of the algebra we know that the lowest weight vector Λ must satisfy

$$(\Lambda|\Lambda) \leq 2. \tag{2.5.51}$$

Taking $\Lambda = \ell\alpha_1 + m_2\alpha_2 + m_3\alpha_3$ then gives the following constraint on the coefficients ℓ, m_2 and m_3:

$$(\Lambda|\Lambda) = 2\ell^2 + 2m_2^2 + 2m_3^2 - 4\ell m_2 - 2m_2m_3 \leq 2. \tag{2.5.52}$$

We are interested in finding an equation for the Dynkin labels, so we insert Equation (2.5.50) into Equation (2.5.52) to obtain the constraint

$$p_2^2 + p_3^2 + p_2p_3 - \ell^2 \leq 3. \tag{2.5.53}$$

The inequalities in Equation (2.5.50) and Equation (2.5.53) are sufficient to determine the representation content at each level ℓ. However, this analysis does not take into account the outer multiplicities, which must be analyzed separately by comparing with the known root multiplicities of A_1^{++} as given in Table H_3 on page 215 of [34]. We shall return to this issue later.

Level $\ell = 2$

Let us now use these results to analyze the case for which $\ell = 2$. The following equations must then be satisfied:

$$\begin{array}{rcl} a8 - 2p_2 - p_3 & \geq & 0, \\ 4 - p_2 - 2p_3 & \geq & 0, \\ p_2^2 + p_3^2 + p_2p_3 & \leq & 7. \end{array} \tag{2.5.54}$$

The only admissible solution is $p_2 = 2$ and $p_3 = 1$. This corresponds to a 15-dimensional representation $\mathbf{15}_2$ with the following Young tableau

$$\mathbf{15}_2 \quad \Longleftrightarrow \quad \ydiagram{3,1}\,. \tag{2.5.55}$$

Note that $p_2 = p_3 = 0$ is also a solution to Equation (2.5.54) but this violates the constraint that m_2 and m_3 be integers and so is not allowed.

Moreover, the representation $[p_2, p_3] = [0, 2]$ is also a solution to Equation (2.5.54) but has not been taken into account because it has vanishing outer multiplicity. This can be understood by examining Figure 2.13 a little closer. The representation $[0, 2]$ is six-dimensional and has highest weight $2\lambda_3$, corresponding to the middle node of the top horizontal line in Figure 2.13. This weight lies outside of the lightcone and so is a real root of A_1^{++}. Therefore we know that it has root multiplicity one and may therefore only occur once in the level decomposition. Since the weight $2\lambda_3$ already appears in the larger representation $\mathbf{15}_2$ it cannot be a highest weight in another representation at this level. Hence, the representation $[0, 2]$ is not allowed within A_1^{++}. A similar analysis reveals that also the representation $[p_2, p_3] = [1, 0]$, although allowed by Equation (2.5.54), has vanishing outer multiplicity.

The level two module is realized by the tensor $E_i{}^{jk}$ whose index structure matches the Young tableau above. Here we have used the $\mathfrak{sl}(3,\mathbb{R})$-invariant antisymmetric tensor ϵ^{abc} to lower the two upper antisymmetric indices leading to a tensor $E_i{}^{jk}$ with the properties

$$E_i{}^{jk} = E_i{}^{(jk)}, \qquad E_i{}^{ik} = 0. \tag{2.5.56}$$

This corresponds to a positive root generator and by the Chevalley involution we have an associated negative root generator $F^i{}_{jk}$ at level $\ell = -2$. Because the level decomposition gives a gradation of A_1^{++} we know that all higher level generators can be obtained through commutators of the level one generators. More specifically, the level two tensor $E_i{}^{jk}$ corresponds to the commutator

$$[E^{ij}, E^{kl}] = \epsilon^{mk(i}E_m{}^{j)l} + \epsilon^{ml(i}E_m{}^{j)k}, \tag{2.5.57}$$

where ϵ^{ijk} is the totally antisymmetric tensor in three dimensions. Inserting the result $p_2 = 2$ and $p_3 = 1$ into Equation (2.5.50) gives $m_2 = 1$ and $m_3 = 0$, thus providing the explicit form of the root taking us from the origin of the root diagram in Figure 2.12 to the lowest weight of $\mathbf{15}_2$ at level two:

$$\Lambda^{(2)} = 2\alpha_1 + \alpha_2. \tag{2.5.58}$$

This is a real root of A_1^{++}, $(\gamma|\gamma) = 2$, and hence the representation $\mathbf{15}_2$ has outer multiplicity one. We display the representation $\mathbf{15}_2$ of $\mathfrak{sl}(3,\mathbb{R})$ in Figure 2.13. The lower leftmost weight is the lowest weight $\Lambda^{(2)}$. The expansion of the lowest weight $\Lambda^{(2)}_{\text{lw}}$ in terms of the fundamental weights λ_2 and λ_3 is given by the (conjugate) Dynkin labels

$$-\Lambda^{(2)}_{\text{hw}} = p_2\lambda_2 + p_3\lambda_3 = 2\lambda_2 + \lambda_3. \tag{2.5.59}$$

The three innermost weights all have multiplicity 2 as weights of $\mathfrak{sl}(3,\mathbb{R})$, as indicated by the black circles. These lie inside the lightcone of $\mathfrak{h}^\star_{\mathfrak{g}}$ and so are timelike roots of A_1^{++}.

Level $\ell = 3$

We proceed quickly past level three since the analysis does not involve any new ingredients. Solving Equation (2.5.50) and Equation (2.5.53) for $\ell = 3$ yields two admissible $\mathfrak{sl}(3,\mathbb{R})$ representations, $\mathbf{27}_3$ and $\mathbf{8}_3$, represented by the following Dynkin labels and Young tableaux:

$$\begin{aligned} \mathbf{27}_3 \quad &: \quad [p_2, p_3] = [2,2] \quad \Longleftrightarrow \quad \ydiagram{4,2}\,, \\ \mathbf{8}_3 \quad &: \quad [p_2, p_3] = [1,1] \quad \Longleftrightarrow \quad \ydiagram{2,1}\,. \end{aligned} \tag{2.5.60}$$

The lowest weight vectors for these representations are

$$\begin{aligned} \Lambda^{(3)}_{\mathbf{15}} &= 3\alpha_1 + 2\alpha_2, \\ \Lambda^{(3)}_{\mathbf{8}} &= 3\alpha_1 + 3\alpha_2 + \alpha_3. \end{aligned} \tag{2.5.61}$$

The lowest weight vector for $\mathbf{27}_3$ is a real root of A_1^{++}, $(\Lambda^{(3)}_{\mathbf{27}}|\Lambda^{(3)}_{\mathbf{27}}) = 2$, while the lowest weight vectors for $\mathbf{8}_3$ is timelike, $(\Lambda^{(3)}_{\mathbf{8}}|\Lambda^{(3)}_{\mathbf{8}}) = -4$. This implies that the entire representation $\mathbf{8}_3$ lies inside the lightcone of $\mathfrak{h}^\star_{\mathfrak{g}}$. Both representations have outer multiplicity one.

Note that $[0,3]$ and $[3,0]$ are also admissible solutions but have vanishing outer multiplicities by the same arguments as for the representation $[0,2]$ at level 2.

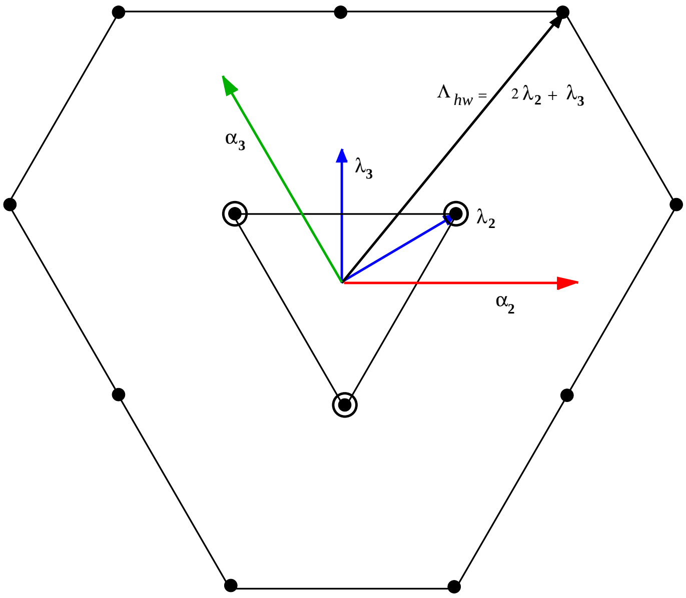


Figure 2.13: The representation $\mathbf{15}_2$ of $\mathfrak{sl}(3,\mathbb{R})$ appearing at level two in the decomposition of the adjoint representation of A_1^{++} into representations of $\mathfrak{sl}(3,\mathbb{R})$. The lowest leftmost node is the lowest weight of the representation, corresponding to the real root $\Lambda^{(2)} = 2\alpha_1 + \alpha_2$ of A_1^{++}. This representation has outer multiplicity one.

Level $\ell = 4$

At this level we encounter for the first time a representation with non-trivial outer multiplicity. It is a 15-dimensional representation with the following Young tableau structure:

$$\bar{\mathbf{15}}_4 : [p_2, p_3] = [1, 2] \quad \Longleftrightarrow \quad \ydiagram{3,2}\,. \tag{2.5.62}$$

The lowest weight vector is

$$\Lambda^{(4)}_{\bar{\mathbf{15}}} = 4\alpha_1 + 4\alpha_2 + \alpha_3, \tag{2.5.63}$$

which is an imaginary root of A_1^{++},

$$(\Lambda^{(4)}_{\bar{\mathbf{15}}}|\Lambda^{(4)}_{\bar{\mathbf{15}}}) = -6. \tag{2.5.64}$$

Table H_3 of [34] we find that this root has multiplicity 5 as a root of A_1^{++},

$$\mathrm{mult}(\Lambda^{(4)}_{\bar{\mathbf{15}}}) = 5. \tag{2.5.65}$$

In order for Equation (2.5.26) to make sense, this multiplicity must be matched by the total multiplicity of $\Lambda^{(4)}_{\bar{\mathbf{15}}}$ as a weight of $\mathfrak{sl}(3,\mathbb{R})$ representations at level four. The remaining representations at this level are

$$\begin{aligned}
\mathbf{24}_4 &: [3,1] \iff \ydiagram{4,1}\,, \\
\bar{\mathbf{3}}_4 &: [0,1] \iff \ydiagram{1,1}\,, \\
\mathbf{6}_4 &: [2,0] \iff \ydiagram{2}\,, \\
\mathbf{42}_4 &: [2,3] \iff \ydiagram{6,3}\,.
\end{aligned} \tag{2.5.66}$$

By drawing these representations explicitly, one sees that the root $4\alpha_1+4\alpha_2+\alpha_3$, representing the weight $\Lambda^{(4)}_{\bar{\mathbf{15}}}$, also appears as a weight (but not as a lowest weight) in the representations $\mathbf{42}_4$ and $\mathbf{24}_4$. It occurs with weight multiplicity 1 in the $\mathbf{24}_4$ but with weight multiplicity 2 in the $\mathbf{42}_4$. Taking also into account the representation $\bar{\mathbf{15}}_4$ in which it is the lowest weight we find a total weight multiplicity of 4. This implies that, since in A_1^{++}

$$\text{mult}(4\alpha_1 + 4\alpha_2 + \alpha_3) = 5, \tag{2.5.67}$$

the outer multiplicity of $\bar{\mathbf{15}}_4$ must be 2, i.e.,

$$\mu\left(\Lambda^{(2)}_{\bar{\mathbf{15}}}\right) = 2. \tag{2.5.68}$$

When we go to higher and higher levels, the outer multiplicities of the representations located entirely inside the lightcone in $\mathfrak{h}_{\mathfrak{g}}$ increase exponentially.

2.5.4 Level decomposition of E_{10}

In this section we shall describe in detail the level decomposition of the Kac–Moody algebra E_{10} which is one of the four hyperbolic algebras of maximal rank [34]; the others being BE_{10}, DE_{10} and CE_{10} (see, e.g., **Paper III** for more information). As already mentioned in Section 2.4.1, E_{10} can be constructed as an overextension of E_8 and is therefore often denoted by E_8^{++}. Similarly to E_8 in the rank 8 case, E_{10} is the unique indefinite rank 10 algebra with an even self-dual root lattice, namely the Lorentzian lattice $\Pi_{1,9}$.

Our first encounter with E_{10} in a physical application will be in Chapter 3, where we will show that the Weyl group of E_{10} describes the chaotic dynamics that emerges when studying eleven-dimensional supergravity close to a spacelike singularity [44].

In Chapter 5, we will also discuss how to construct a Lagrangian manifestly invariant under global $\mathcal{E}_{10}$-transformations, and compare its dynamics to that of eleven-dimensional supergravity. The level decomposition associated with the removal of the "exceptional node" labelled "10" in Figure 2.14 will be central to the analysis. It turns out that the low-level structure in this decomposition precisely reproduces the bosonic field content of eleven-dimensional supergravity [53].

Moreover, decomposing E_{10} with respect to different regular subalgebras reproduces also the bosonic field contents of the Type IIA and Type IIB supergravities. The fields of the IIA theory are obtained by decomposition in terms of representations of the $D_9 = \mathfrak{so}(9,9,\mathbb{R})$ subalgebra obtained by removing the first simple root α_1 [86]. An alternative way of revealing the degrees of freedom of type IIA supergravity is to decompose E_{10} with respect to and $\mathfrak{sl}(9,\mathbb{R})$-subalgebra. This viewpoint has a natural interpretation as describing the emergence of type IIA supergavity from a dimensional reduction of eleven-dimensional supergravity. This was analyzed in detail in **Paper VI** and is discussed in Chapter 7.

Similarly the degrees of freedom of type IIB supergavity appear at low levels in a decomposition of E_{10} with respect to the $A_9 \oplus A_1 = \mathfrak{sl}(9,\mathbb{R}) \oplus \mathfrak{sl}(2,\mathbb{R})$ subalgebra found upon removal of the second simple root α_2 [87]. The extra A_1-factor in this decomposition ensures that the global $SL(2,\mathbb{R})$-symmetry of IIB supergravity is recovered. This $SL(2,\mathbb{R})$-symmetry, and its arithmetic subgroup $SL(2,\mathbb{Z}) \subset SL(2,\mathbb{R})$ will also play an important role in **Part II** of this thesis.

Decomposition with Respect to $\mathfrak{sl}(10,\mathbb{R})$

Let $\alpha_1, \cdots, \alpha_{10}$ denote the simple roots of E_{10} and $\alpha_1^\vee, \cdots, \alpha_{10}^\vee$ the Cartan generators. These span the root space $\mathfrak{h}^\star$ and the Cartan subalgebra $\mathfrak{h}$, respectively. Since E_{10} is simply laced the Cartan matrix is given by the scalar products between the simple roots:

$$A_{ij}[E_{10}] = (\alpha_i|\alpha_j) = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}. \tag{2.5.69}$$

The associated Dynkin diagram is displayed in Figure 2.14. We will perform the decomposition with respect to the $\mathfrak{sl}(10,\mathbb{R})$ subalgebra represented by the horizontal line in the Dynkin diagram so the level ℓ of an arbitrary root $\alpha \in \mathfrak{h}^\star$ is given by the coefficient in front of the exceptional simple root, i.e.,

$$\gamma = \sum_{i=1}^{9} m^i \alpha_i + \ell \alpha_{10}. \tag{2.5.70}$$

As before, the weight that is easiest to identify for each representation $\mathcal{R}(\Lambda^{(\ell)})$ at positive level ℓ is the lowest weight $\Lambda^{(\ell)}_{\text{lw}}$. We denote by $\bar{\Lambda}^{(\ell)}_{\text{lw}}$ the projection onto the spacelike slice of the root lattice defined by the level ℓ. The (conjugate) Dynkin labels $p_1, \cdots, p_9$ characterizing the representation $\mathcal{R}(\Lambda^{(\ell)})$ are defined as before as minus the coefficients in the expansion of $\bar{\Lambda}^{(\ell)}_{\text{lw}}$ in terms of the fundamental weights λ^i of $\mathfrak{sl}(10,\mathbb{R})$:

$$-\bar{\Lambda}^{(\ell)}_{\text{lw}} = \sum_{i=1}^{9} p_i \lambda^i. \tag{2.5.71}$$

The Killing form on each such slice is positive definite so the projected weight $\bar{\Lambda}^{(\ell)}_{\text{hw}}$ is of course real. The fundamental weights of $\mathfrak{sl}(10,\mathbb{R})$ can be computed explicitly from their

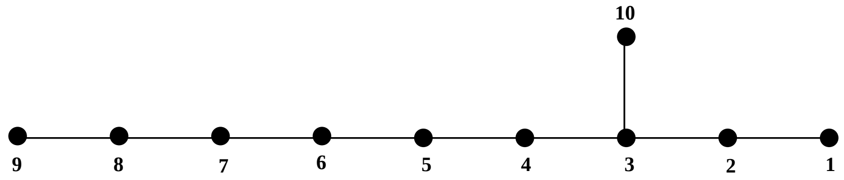


Figure 2.14: The Dynkin diagram of E_{10}. Labels $i = 1, \cdots, 9$ enumerate the nodes corresponding to simple roots α_i of the $\mathfrak{sl}(10, \mathbb{R})$ subalgebra and "10" labels the exceptional node.

definition as the duals of the simple roots:

$$\lambda^i = \sum_{j=1}^{9} B^{ij} \alpha_j, \tag{2.5.72}$$

where B^{ij} is the inverse of the Cartan matrix of A_9,

$$(B_{ij}[A_9])^{-1} = \frac{1}{10} \begin{pmatrix} 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 8 & 16 & 14 & 12 & 10 & 8 & 6 & 4 & 2 \\ 7 & 14 & 21 & 18 & 15 & 12 & 9 & 6 & 3 \\ 6 & 12 & 18 & 24 & 20 & 16 & 12 & 8 & 4 \\ 5 & 10 & 15 & 20 & 25 & 20 & 15 & 10 & 5 \\ 4 & 8 & 12 & 16 & 20 & 24 & 18 & 12 & 6 \\ 3 & 6 & 9 & 12 & 15 & 18 & 21 & 14 & 7 \\ 2 & 4 & 6 & 8 & 10 & 12 & 14 & 16 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{pmatrix}. \tag{2.5.73}$$

Note that all the entries of B^{ij} are positive which will prove to be important later on. As we saw for the A_1^{++} case we want to find the possible allowed values for $(m_1, \cdots, m_9)$, or, equivalently, the possible Dynkin labels $[p_1, \cdots, p_9]$ for each level ℓ.

The corresponding diophantine equation, Equation (2.5.50), for E_{10} was found in [53] and reads

$$m^i = B^{i3}\ell - \sum_{j=1}^{9} B^{ij} p_j \geq 0. \tag{2.5.74}$$

Since the two sets $\{p_i\}$ and $\{m^i\}$ both consist of non-negative integers and all entries of B^{ij} are positive, these equations put strong constraints on the possible representations that can occur at each level. Moreover, each lowest weight vector $\Lambda^{(\ell)} = \gamma$ must be a root of E_{10}, so we have the additional requirement

$$(\Lambda^{(\ell)}|\Lambda^{(\ell)}) = \sum_{i,j=1}^{9} B^{ij} p_i p_j - \frac{1}{10}\ell^2 \leq 2. \tag{2.5.75}$$

The representation content at each level is represented by $\mathfrak{sl}(10, \mathbb{R})$-tensors whose index structure are encoded in the Dynkin labels $[p_1, \cdots, p_9]$. At level $\ell = 0$ we have the adjoint

Table 2.2: The low-level representations in a decomposition of the adjoint representation of E_{10} into representations of its A_9 subalgebra obtained by removing the exceptional node in the Dynkin diagram in Figure 2.14.

ℓ	$\Lambda^{(\ell)} = [p_1, \cdots, p_9]$	$\Lambda^{(\ell)} = (m_1, \cdots, m_{10})$	A_9-representation	E_{10}-generator
1	$[0,0,1,0,0,0,0,0,0]$	$(0,0,0,0,0,0,0,0,0,1)$	$\mathbf{120}_1$	E^{abc}
2	$[0,0,0,0,0,1,0,0,0]$	$(1,2,3,2,1,0,0,0,0,2)$	$\mathbf{210}_2$	$E^{a_1\cdots a_6}$
3	$[1,0,0,0,0,0,0,1,0]$	$(1,3,5,4,3,2,1,0,0,3)$	$\mathbf{440}_3$	$E^{a\|b_1\cdots b_8}$

representation of $\mathfrak{sl}(10,\mathbb{R})$ represented by the generators ${K^a}_b$ obeying the same commutation relations as in Equation (2.5.32) but now with $\mathfrak{sl}(10,\mathbb{R})$-indices.

All higher (lower) level representations will then be tensors transforming contravariantly (covariantly) under the level $\ell = 0$ generators. The resulting representations are displayed up to level 3 in Table 2.2. We see that the level 1 and 2 representations have the index structures of a 3-form and a 6-form respectively. In the E_{10}-invariant sigma model, to be constructed in Chapter 5, these generators will become associated with the time-dependent physical "fields" $A_{abc}(t)$ and $A_{a_1\cdots a_6}(t)$ which are related to the electric and magnetic component of the 3-form in eleven-dimensional supergravity. Similarly, the level 3 generator $E^{a|b_1\cdots b_9}$ with mixed Young symmetry will be associated to the dual of the spatial part of the eleven-dimensional vielbein. This field is therefore sometimes referred to as the "dual graviton".

Algebraic Structure at Low Levels

Let us now describe in a little more detail the commutation relations between the low-level generators in the level decomposition of E_{10} (see Table 2.2). We recover the Chevalley generators of A_9 through the following realisation:

$$e_i = {K^{i+1}}_i, \qquad f_i = {K^i}_{i+1}, \qquad h_i = {K^{i+1}}_{i+1} - {K^i}_i \qquad (i = 1, \cdots, 9), \tag{2.5.76}$$

where, as before, the ${K^i}_j$'s obey the commutation relations

$$[{K^i}_j, {K^k}_l] = \delta^k_j {K^i}_l - \delta^i_l {K^k}_j. \tag{2.5.77}$$

At levels ± 1 we have the positive root generators E^{abc} and their negative counterparts $F_{abc} = -\tau(E^{abc})$, where τ denotes the Chevalley involution as defined in Section 2.2.3. Their transformation properties under the $\mathfrak{sl}(10,\mathbb{R})$-generators ${K^a}_b$ follow from the index structure and reads explicitly

$$\begin{aligned} [{K^a}_b, E^{cde}] &= 3\delta^{[c}_b E^{de]a}, \\ [{K^a}_b, F_{cde}] &= -3{\delta^a}_{[c} F_{de]b}, \\ [E^{abc}, F_{def}] &= 18\delta^{[ab}_{[de} {K^{c]}}_{f]} - 2\delta^{abc}_{def} \sum_{a=1}^{10} {K^a}_a, \end{aligned} \tag{2.5.78}$$

where we defined

$$\begin{aligned}\delta^{ab}_{cd} &= \frac{1}{2}(\delta^a_c\delta^b_d - \delta^b_c\delta^a_d)\\ \delta^{abc}_{def} &= \frac{1}{3!}(\delta^a_d\delta^b_e\delta^c_f \pm 5 \text{ permutations}).\end{aligned} \tag{2.5.79}$$

The "exceptional" generators e_{10} and f_{10} are fixed by Equation (2.5.76) to have the following realisation:

$$e_{10} = E^{123}, \qquad f_{10} = F_{123}. \tag{2.5.80}$$

The corresponding Cartan generator is obtained by requiring $[e_{10}, f_{10}] = h_{10}$ and upon inspection of the last equation in Equation (2.5.78) we find

$$h_{10} = -\frac{1}{3}\sum_{i\neq 1,2,3} K^a{}_a + \frac{2}{3}(K^1{}_1 + K^2{}_2 + K^3{}_3), \tag{2.5.81}$$

enlarging $\mathfrak{sl}(10,\mathbb{R})$ to $\mathfrak{gl}(10,\mathbb{R})$.

The bilinear form at level zero is

$$(K^i{}_j|K^k{}_l) = \delta^i_l\delta^k_j - \delta^i_j\delta^k_l \tag{2.5.82}$$

and can be extended level by level to the full algebra by using its invariance, $([x,y]|z) = (x|[y,z])$ for $x,y,z \in E_{10}$ (see Section 2.2.4). For level 1 this yields

$$\left(E^{abc}|F_{def}\right) = 3!\delta^{abc}_{def}, \tag{2.5.83}$$

where the normalization was chosen such that

$$(e_{10}|f_{10}) = \left(E^{123}|F_{123}\right) = 1. \tag{2.5.84}$$

Now, by using the graded structure of the level decomposition we can infer that the level 2 generators can be obtained by commuting the level 1 generators

$$[\mathfrak{g}_1, \mathfrak{g}_1] \subseteq \mathfrak{g}_2. \tag{2.5.85}$$

Concretely, this means that the level 2 content should be found from the commutator

$$[E^{a_1a_2a_3}, E^{a_4a_5a_6}]. \tag{2.5.86}$$

We already know that the only representation at this level is $\mathbf{210}_2$, realized by an antisymmetric 6-form. Since the normalization of this generator is arbitrary we can choose it to have weight one and hence we find

$$E^{a_1\cdots a_6} = [E^{a_1a_2a_3}, E^{a_4a_5a_6}]. \tag{2.5.87}$$

The bilinear form is lifted to level 2 in a similar way as before with the result

$$(E^{a_1\cdots a_6}|F_{b_1\cdots b_6}) = 6!\delta^{a_1\cdots a_6}_{b_1\cdots b_6}. \tag{2.5.88}$$

Continuing these arguments, the level 3-generators can be obtained from

$$[[\mathfrak{g}_1, \mathfrak{g}_1], \mathfrak{g}_1] \subseteq \mathfrak{g}_3. \tag{2.5.89}$$

From the index structure one would expect to find a 9-form generator $E^{a_1\cdots a_9}$ corresponding to the Dynkin labels $[0,0,0,0,0,0,0,0,1]$. However, we see from Table 2.2 that only the representation $[1,0,0,0,0,0,0,1,0]$ appears at level 3. The reason for the disappearance of the representation $[0,0,0,0,0,0,0,0,1]$ is because the generator $E^{a_1\cdots a_9}$ is not allowed by the Jacobi identity. A detailed explanation for this can be found in [88]. The right hand side of Equation (2.5.89) therefore only contains the index structure compatible with the generators $E^{a|b_1\cdots b_8}$,

$$[[E^{ab_1b_2}, E^{b_3b_4b_5}], E^{b_6b_7b_8}] = -E^{[a|b_1b_2]b_3\cdots b_8}, \tag{2.5.90}$$

where the minus sign is purely conventional.

For later reference, we list here some additional commutators that are useful [89]:

$$\begin{aligned}
[E^{a_1\cdots a_6}, F_{b_1b_2b_3}] &= -5!\delta^{[a_1a_2a_3}_{b_1b_2b_3}E^{a_1a_2a_3]}, \\
[E^{a_1\cdots a_6}, F_{b_1\cdots b_6}] &= 6\cdot 6!\delta^{[a_1\cdots a_5}_{[b_1\cdots b_5}K^{a_6]}{}_{b_6]} - \frac{2}{3}\cdot 6!\delta^{a_1\cdots a_6}_{b_1\cdots b_6}\sum_{a=1}^{10}K^a{}_a, \\
[E^{a_1|a_2\cdots a_9}, F_{b_1b_2b_3}] &= -7\cdot 48\left(\delta^{a_1[a_2a_3}_{b_1b_2b_3}E^{a_4\cdots a_9]} - \delta^{[a_2a_3a_4}_{b_1b_2b_3}E^{a_5\cdots a_9]a_1}\right), \\
[E^{a_1|a_2\cdots a_9}, F_{b_1\cdots b_6}] &= -8!\left(\delta^{a_1[a_2\cdots a_6}_{b_1\cdots b_6}E^{a_7a_8a_9]} - \delta^{[a_2\cdots a_7}_{b_1\cdots b_6}E^{a_8a_9]a_1}\right).
\end{aligned} \tag{2.5.91}$$

3

Chaotic Cosmology and Hyperbolic Coxeter Groups

In this chapter we will see explicitly how the mathematical structures presented in Chapter 2 appears in a physical context. To this end we begin by analyzing gravity close to a spacelike singularity and explain in detail the reformulation of the dynamics in terms of a hyperbolic billiard. The key steps in this analysis are to work within a Hamiltonian framework, and to perform an Iwasawa decomposition of the spatial metric. This separates diagonal and off-diagonal degrees of freedom and leads in a natural way to the billiard description. Next, we make the connection with Kac-Moody algebras by showing that the billiard dynamics takes place within a bounded region of the Cartan subalgebra of a Lorentzian Kac-Moody algebra. This reveals that the approach to the singularity is controlled by the Weyl group of this Kac-Moody algebra, and provides a natural explanation for the (non-)existence of chaos. This chapter is based on **Paper III**, written in collaboration with Marc Henneaux and Philippe Spindel.

3.1 Cosmological Billiards

Here we present a lightning review of the billiard interpretation of the dynamics in the BKL limit. We emphasize features which are relevant for understanding the connection between the billiards and the underlying Kac-Moody algebraic structure, a concept which is discussed in more detail in Section 3.2. For more details, we refer the reader to **Paper III** or the review [51].

Our starting point is the conjecture of Belinskii, Khalatnikov and Lifshitz (BKL) that the dynamics of gravity close to a spacelike (or "big bang type") singularity becomes ultralocal due to a complete decoupling of spatial points [45–47]. We will here make use of Misner's reformulation of the dynamics in terms of a a billiard-type particle motion in an auxiliary hyperbolic space, originally called "mixmaster behaviour" [48, 90].

This conjecture was originally framed in the context of pure four-dimensional gravity, but has subsequently been generalised to any dimension and to gravity coupled to any number of p-forms. For our purposes it is useful to divide the conjecture into two parts:

- The dynamics in the vicinity of the singularity is inherently "local". This means that the partial differential equations describing gravity reduce to a collection of ordinary differential equations, at each spatial point, with respect to proper time. This feature is sometimes referred to as a "decoupling of spatial points".
- Each collection of ordinary differential equations can be equivalently described by a billiard motion in a region of hyperbolic space. If this region is of finite volume then the billiard dynamics is chaotic, and if the region is of infinite volume then the dynamics is non-chaotic.

For what follows we will assume these conjectures to be true.

3.1.1 General Considerations

Our main interest in this thesis will be to analyze the BKL-limit in the context of gravitational theories which arise as low-energy limits of string theory and M-theory. However, the methods applied in this section are completely general and can be applied to any gravitational theory coupled to bosonic matter fields. The only known bosonic matter fields that consistently couple to gravity are p-form fields and "dilatonic" scalars, so we begin by considering a general action in $D = d + 1$ dimensions of the form

$$S\left[g_{\mu\nu}, \phi, A^{(p)}\right] = \int d^D x \sqrt{-G}\left[R - \partial_\mu\phi\,\partial^\mu\phi - \frac{1}{2}\sum_p \frac{e^{\lambda^{(p)}\phi}}{(p+1)!}F^{(p)}_{\mu_1\cdots\mu_{p+1}}F^{(p)\,\mu_1\cdots\mu_{p+1}}\right], \tag{3.1.1}$$

where we have chosen units such that $16\pi G = 1$, and we have chosen the Einstein metric $G_{\mu\nu}$ to have Lorentzian signature $(-, +, \cdots, +)$. We further assume that among the scalars, there is only one dilaton[1], denoted ϕ, whose kinetic term is normalized with weight 1 with respect to the Ricci scalar. The real parameter $\lambda^{(p)}$ measures the strength of the coupling to the dilaton. The other scalar fields, collectively referred to as *axions*, are denoted $A^{(0)}$ and have dilaton coupling $\lambda^{(0)} \neq 0$. The integer $p \geq 0$ labels the various p-forms $A^{(p)}$ present in the theory, with field strengths $F^{(p)} = dA^{(p)}$. We assume that the degree p of $A^{(p)}$ is strictly smaller than $D-1$, since a $(D-1)$-form in D dimensions carries no local degree of freedom. Furthermore, if $p = D-2$ the p-form is dual to a scalar and we impose also $\lambda^{(D-2)} \neq 0$.

The field strength, $F^{(p)} = dA^{(p)}$, of each p-form could be modified by additional coupling terms of Yang–Mills or Chapline–Manton type [91, 92][2], but we neglect such terms in the action since they are of no consequence in the BKL-limit. For the same reason, we also neglect Chern-Simons terms in what follows.

We shall at this stage consider arbitrary dilaton couplings and menus of p-forms. The following discussion remains valid no matter what these are; all theories described by the

[1]This is done mostly for notational convenience. If there were other dilatons among the 0-forms, these should be separated off from the p-forms because they play a distinct role. They would appear as additional scale factors and would increase the dimensions of the relevant hyperbolic billiard (they define additional spacelike directions in the space of scale factors).

[2]For example, $F_C = dC^{(2)} - C^{(0)}dB^{(2)}$ for two 2-forms $C^{(2)}$ and $B^{(2)}$ and a 0-form $C^{(0)}$, as it occurs in ten-dimensional type IIB supergravity.

general action Equation (3.1.1) lead to the billiard picture. However, it is only for particular p-form menus, spacetime dimensions and dilaton couplings that the billiard region is regular and associated with a Kac–Moody algebra. This will be discussed in Section 3.2. Note that the action, Equation (3.1.1), contains as particular cases the bosonic sectors of all known supergravity theories.

3.1.2 Hamiltonian Treatment

For the purposes of analyzing the action (3.1.1) in the BKL-limit, it is convenient to first reformulate it in Hamiltonian form. To this end, we assume that there is a spacelike singularity at a finite distance in proper time. We adopt a spacetime slicing adapted to the singularity, which "occurs" on a slice of constant time. The slicing is then built up starting from the singularity by taking pseudo-Gaussian coordinates defined by $N = \sqrt{\mathrm{g}}$ and $N^i = 0$, where N is the lapse and N^i is the shift [51]. Here, $\mathrm{g} \equiv \det(\mathrm{g}_{ij})$, with g_{ij} being the spatial metric. Thus, in some spacetime patch, our metric ansatz reads[3]

$$ds^2 = -\mathrm{g}(dx^0)^2 + \mathrm{g}_{ij}(x^0, x^i)\, dx^i\, dx^j, \tag{3.1.2}$$

where the local volume g collapses at each spatial point as $x^0 \to +\infty$, in such a way that the proper time $dT = -\sqrt{\mathrm{g}}\, dx^0$ remains finite (and tends conventionally to 0^+). Here we have assumed the singularity to occur in the past, as in the original BKL analysis, but a similar discussion holds for future spacelike singularities.

In the Hamiltonian description of the dynamics, the canonical variables are the spatial metric components g_{ij}, the dilaton ϕ, the spatial p-form components $A^{(p)}_{m_1\cdots m_p}$ and their respective conjugate momenta π^{ij}, π_ϕ and $\pi^{m_1\cdots m_p}_{(p)}$. The Hamiltonian action in the pseudo-Gaussian gauge is given by

$$S\left[\mathrm{g}_{ij}, \pi^{ij}, \phi, \pi_\phi, A^{(p)}, \pi_{(p)}\right] = \int dx^0 \left[\int d^dx \left(\pi^{ij}\dot{\mathrm{g}}_{ij} + \pi_\phi\dot{\phi} + \sum_p \pi^{m_1\cdots m_p}_{(p)} \dot{A}^{(p)}_{m_1\cdots m_p}\right) - H\right], \tag{3.1.3}$$

where the Hamiltonian is

$$\begin{aligned} H &= \int d^dx\, \mathcal{H} = \int d^dx \big(K' + V'\big), \\ K' &= \pi^{ij}\pi_{ij} - \frac{1}{d-1}(\pi^i_i)^2 + \frac{1}{4}(\pi_\phi)^2 + \sum_p \frac{(p!)e^{-\lambda^{(p)}\phi}}{2}\, \pi^{m_1\cdots m_p}_{(p)}\pi_{(p)\,m_1\cdots m_p}, \\ V' &= -R\mathrm{g} + \mathrm{g}^{ij}\mathrm{g}\,\partial_i\phi\,\partial_j\phi + \sum_p \frac{e^{\lambda^{(p)}\phi}}{2\,(p+1)!}\,\mathrm{g}\, F^{(p)}_{m_1\cdots m_{p+1}} F^{(p)\,m_1\cdots m_{p+1}}. \end{aligned} \tag{3.1.4}$$

In addition to imposing the coordinate conditions $N = \sqrt{\mathrm{g}}$ and $N^i = 0$, we have also set the temporal components of the p-forms equal to zero ("temporal gauge").

[3] Note that we have for convenience chosen to work with a coordinate coframe dx^i, with the imposed constraint $N = \sqrt{\mathrm{g}}$. In general, one may of course use an arbitrary spatial coframe, say $\theta^i(x)$, for which the associated gauge choice reads $N = w(x)\sqrt{\mathrm{g}}$, with $w(x)$ being a density of weight -1. This general kind of spatial coframe was also used extensively in the recent work [93].

The dynamical equations of motion are obtained by varying the above action w.r.t. the canonical variables. Moreover, there are additional constraints on the dynamical variables, which are preserved by the dynamical evolution and need only be imposed at some "initial" time, say $x^0 = 0$. The detailed form of these constraints can be found in [51, 94], and here we shall only be concerned with the Hamiltonian constraint,

$$\mathcal{H} = 0, \tag{3.1.5}$$

obtained by varying the action with respect to the lapse function N, before imposing the gauge condition $N = \sqrt{\mathrm{g}}$.

3.1.3 Iwasawa Decomposition of the Metric

In order to study the dynamical behavior of the fields as $x^0 \to \infty$ ($\mathrm{g} \to 0$) and to exhibit the billiard picture, it is particularly convenient to perform separate diagonal from off-diagonal components of the spatial metric. Effectively, this corresponds to performing an Iwasawa decomposition of the spatial vielbein. Let $\mathrm{e}(x^0, x^i)$ be the spatial vielbein, such that the spatial metric can be written as $\mathrm{g}(x^0, x^i) = \mathrm{e}(x^0, x^i)^T \mathrm{e}(x^0, x^i)$. We then perform the following Iwasawa decomposition of the vielbein[4]

$$\mathrm{e} = \mathcal{KAN}, \qquad \mathcal{K} \in SO(d, \mathbb{R}), \tag{3.1.6}$$

where $\mathcal{N} = \mathcal{N}(x^0, x^i)$ is an upper triangular matrix with 1's on the diagonal ($\mathcal{N}_{ii} = 1$, $\mathcal{N}_{ij} = 0$ for $i > j$) and $\mathcal{A} = \mathcal{A}(x^0, x^i)$ is a diagonal matrix with positive elements, which we parametrize as

$$\mathcal{A} = \exp(-\beta), \qquad \beta = \mathrm{diag}(\beta^1, \beta^2, \cdots, \beta^d). \tag{3.1.7}$$

The spatial metric then reads

$$\mathrm{g} = \mathrm{e}^T \mathrm{e} = \mathcal{N}^T \mathcal{A}^2 \mathcal{N}, \tag{3.1.8}$$

or, in components,

$$d\sigma^2 = \mathrm{g}_{ij}\, dx^i\, dx^j = \sum_{k=1}^{d} e^{(-2\beta^k)} (\omega^k)^2 \tag{3.1.9}$$

with

$$\omega^k = \sum_i \mathcal{N}_{k\,i}\, dx^i. \tag{3.1.10}$$

The variables β^i of the Iwasawa decomposition give the (logarithmic) scale factors in the new, orthogonal, basis. The variables $\mathcal{N}_{ij}$ characterize the change of basis that diagonalizes the metric and hence they parametrize the off-diagonal components of the original g_{ij}.

We extend the transformation Equation (3.1.8) in configuration space to a canonical transformation in phase space through the formula

$$\pi^{ij} d\mathrm{g}_{ij} = \pi^i d\beta_i + \sum_{i<j} P_{ij}\, d\mathcal{N}_{ij}. \tag{3.1.11}$$

[4] See Chapter 2 for a definition of the Iwasawa decomposition.

Since the scale factors and the off-diagonal variables play very distinct roles in the asymptotic behavior, we split off the Hamiltonian as a sum of a kinetic term for the scale factors (including the dilaton),

$$K = \frac{1}{4}\left[\sum_{i=1}^{d} \pi_i^2 - \frac{1}{d-1}\left(\sum_{i=1}^{d} \pi_i\right)^2 + \pi_\phi^2\right], \tag{3.1.12}$$

plus the rest, denoted by V, which will act as a potential for the scale factors. The Hamiltonian then becomes

$$\mathcal{H} = K + V, \tag{3.1.13}$$

with

$$\begin{aligned}
V &= V_S + V_G + \sum_p V_p + V_\phi, \\
V_S &= \frac{1}{2}\sum_{i<j} e^{-2(\beta^j-\beta^i)}\Big(\sum_m P_{im}\mathcal{N}_{jm}\Big)^2, \\
V_G &= -R\mathrm{g}, \\
V_{(p)} &= V_{(p)}^{\mathrm{el}} + V_{(p)}^{\mathrm{magn}}, \\
V_{(p)}^{\mathrm{el}} &= \frac{p!e^{-\lambda^{(p)}\phi}}{2}\,\pi_{(p)}^{m_1\cdots m_p}\pi_{(p)\,m_1\cdots m_p}, \\
V_{(p)}^{\mathrm{magn}} &= \frac{e^{\lambda^{(p)}\phi}}{2\,(p+1)!}\,\mathrm{g}\,F^{(p)}_{m_1\cdots m_{p+1}}F^{(p)\,m_1\cdots m_{p+1}}, \\
V_\phi &= \mathrm{g}^{ij}\mathrm{g}\,\partial_i\phi\,\partial_j\phi.
\end{aligned} \tag{3.1.14}$$

The kinetic term K is quadratic in the momenta conjugate to the scale factors and defines the inverse of a metric in the space of the scale factors. The space of scale factors will play a crucial role in what follows and we will denote it by $\mathcal{M}_\beta$. The metric $\gamma_{\mu\nu}$ on $\mathcal{M}_\beta$ can be extracted by inverting the metric in the kinetic term K and one finds

$$ds^2_{\mathcal{M}_\beta} = \gamma_{\mu\nu}d\beta^\mu d\beta^\nu = \sum_i (d\beta^i)^2 - \Big(\sum d\beta^i\Big)^2 + (d\phi)^2. \tag{3.1.15}$$

Since the metric coefficients do not depend on the scale factors, this metric is always flat and, moreover, is of Lorentzian signature. A conformal transformation where all scale factors are scaled by the same number ($\beta^i \to \beta^i + \epsilon$) defines a timelike direction. It will be convenient in the following to collectively denote all the scale factors (the β^i's and the dilaton ϕ) as β^μ, i.e., $(\beta^\mu) = (\beta^i, \phi)$. We will discuss these important properties of $\mathcal{M}_\beta$ later on.

The analysis is further simplified if we take for new p-form variables the components of the p-forms in the Iwasawa basis of the ω^k's,

$$\mathcal{A}^{(p)}_{i_1\cdots i_p} = \sum_{m_1,\cdots,m_p} (\mathcal{N}^{-1})_{m_1 i_1}\cdots(\mathcal{N}^{-1})_{m_p i_p} A_{(p)m_1\cdots m_p}, \tag{3.1.16}$$

and again extend this configuration space transformation to a point canonical transformation in phase space,

$$\left(\mathcal{N}_{ij}, P_{ij}, A^{(p)}_{m_1\cdots m_p}, \pi^{m_1\cdots m_p}_{(p)}\right) \quad \to \quad \left(\mathcal{N}_{ij}, P'_{ij}, \mathcal{A}^{(p)}_{m_1\cdots m_p}, \mathcal{E}^{i_1\cdots i_p}_{(p)}\right), \tag{3.1.17}$$

using the formula $\sum p\, dq = \sum p'\, dq'$, which reads

$$\sum_{i<j} P_{ij}\dot{\mathcal{N}}_{ij} + \sum_p \pi^{m_1\cdots m_p}_{(p)} \dot{A}^{(p)}_{m_1\cdots m_p} = \sum_{i<j} P'_{ij}\dot{\mathcal{N}}_{ij} + \sum_p \mathcal{E}^{i_1\cdots i_p}_{(p)} \dot{\mathcal{A}}^{(p)}_{m_1\cdots m_p}. \tag{3.1.18}$$

Note that the scale factor variables are unaffected, while the momenta P_{ij} conjugate to $\mathcal{N}_{ij}$ get redefined by terms involving $\mathcal{E}$, $\mathcal{N}$ and $\mathcal{A}$ since the components $\mathcal{A}^{(p)}_{m_1\cdots m_p}$ of the p-forms in the Iwasawa basis involve the $\mathcal{N}$'s. On the other hand, the new p-form momenta, i.e., the components of the electric field $\pi^{m_1\cdots m_p}_{(p)}$ in the basis $\{\omega^k\}$ are simply given by

$$\mathcal{E}^{i_1\cdots i_p}_{(p)} = \sum_{m_1,\cdots,m_p} \mathcal{N}_{i_1 m_1}\, \mathcal{N}_{i_2 m_2} \cdots \mathcal{N}_{i_p m_p} \pi^{m_1\cdots m_p}_{(p)}. \tag{3.1.19}$$

In terms of the new variables, the electromagnetic potentials become

$$\begin{aligned} V^{\text{el}}_{(p)} &= \frac{p!}{2} \sum_{i_1,i_2,\cdots,i_p} e^{-2e_{i_1\cdots i_p}(\beta)} (\mathcal{E}^{i_1\cdots i_p}_{(p)})^2, \\ V^{\text{magn}}_{(p)} &= \frac{1}{2\,(p+1)!} \sum_{i_1,i_2,\cdots,i_{p+1}} e^{-2m_{i_1\cdots i_{p+1}}(\beta)} (\mathcal{F}_{(p)\,i_1\cdots i_{p+1}})^2. \end{aligned} \tag{3.1.20}$$

Here, $e_{i_1\cdots i_p}(\beta)$ are the electric linear forms

$$e_{i_1\cdots i_p}(\beta) = \beta^{i_1} + \cdots + \beta^{i_p} + \frac{\lambda^{(p)}}{2}\phi \tag{3.1.21}$$

(the indices i_j are all distinct because $\mathcal{E}^{i_1\cdots i_p}_{(p)}$ is completely antisymmetric) while $\mathcal{F}_{(p)\,i_1\cdots i_{p+1}}$ are the components of the magnetic field $F_{(p)m_1\cdots m_{p+1}}$ in the basis $\{\omega^k\}$,

$$\mathcal{F}_{(p)\,i_1\cdots i_{p+1}} = \sum_{m_1,\cdots,m_{p+1}} (\mathcal{N}^{-1})_{m_1 i_1} \cdots (\mathcal{N}^{-1})_{m_{p+1} i_{p+1}} F_{(p)m_1\cdots m_{p+1}}, \tag{3.1.22}$$

and $m_{i_1\cdots i_{p+1}}(\beta)$ are the magnetic linear forms

$$m_{i_1\cdots i_{p+1}}(\beta) = \sum_{j\notin\{i_1,i_2,\cdots i_{p+1}\}} \beta^j - \frac{\lambda^{(p)}}{2}\phi. \tag{3.1.23}$$

One sometimes rewrites $m_{i_1\cdots i_{p+1}}(\beta)$ as $b_{i_{p+2}\cdots i_d}(\beta)$, where $\{i_{p+2}, i_{p+3}, \cdots, i_d\}$ is the set complementary to $\{i_1, i_2, \cdots i_{p+1}\}$, e.g.,

$$b_{1\,2\cdots\,d-p-1}(\beta) = \beta^1 + \cdots + \beta^{d-p-1} - \frac{\lambda^{(p)}}{2}\phi = m_{d-p\cdots\,d}. \tag{3.1.24}$$

The exterior derivative $\mathcal{F}$ of $\mathcal{A}$ in the non-holonomic frame $\{\omega^k\}$ involves of course the structure coefficients $C^i{}_{jk}$ in that frame, i.e.,

$$\mathcal{F}_{(p)\, i_1 \cdots i_{p+1}} = \partial_{[i_1} \mathcal{A}_{i_2 \cdots i_{p+1}]} + \text{“}C\mathcal{A}\text{”-terms}, \tag{3.1.25}$$

where

$$\partial_{i_1} \equiv \sum_{m_1} (\mathcal{N}^{-1})_{m_1 i_1} (\partial / \partial x^{m_1}) \tag{3.1.26}$$

is here the frame derivative. Similarly, the potential V_ϕ reads

$$V_\phi = \sum_i e^{-2\bar{m}_i(\beta)} (\mathcal{F}_i)^2, \tag{3.1.27}$$

where $\mathcal{F}_i$ is

$$\mathcal{F}_i = (\mathcal{N}^{-1})_{ji} \partial_j \phi \tag{3.1.28}$$

and

$$\bar{m}_i(\beta) = \sum_{j \neq i} \beta^j. \tag{3.1.29}$$

3.1.4 Decoupling of Spatial Points

So far we have only redefined the variables without making any approximation. We now start the discussion of the BKL-limit, which investigates the leading behavior of the degrees of freedom as $x^0 \to \infty$ ($\mathrm{g} \to 0$). As already discussed in the introduction to Section 3.1, it is illuminating to separate two aspects of the BKL conjecture.[5]

The first aspect is that the spatial points decouple in the limit $x^0 \to \infty$, in the sense that one can replace the Hamiltonian by an effective “ultralocal” Hamiltonian H^{UL} involving no spatial gradients, and hence leading at each spatial point to a set of dynamical equations that are ordinary differential equations with respect to time. The effective ultralocal Hamiltonian has a form similar to that of the Hamiltonian governing certain spatially homogeneous cosmological models, as will be explained below.

The second aspect of the BKL-limit is to take the sharp wall limit of the ultralocal Hamiltonian. This leads then directly to the billiard description of the dynamics.

Spatially Homogeneous Models

In spatially homogeneous models, the fields depend only on time in invariant frames, e.g., for the metric

$$ds^2 = \mathrm{g}_{ij}(x^0) \psi^i \psi^j, \tag{3.1.30}$$

where the invariant forms fulfill

$$d\psi^i = -\frac{1}{2} f^i{}_{jk} \psi^j \wedge \psi^k.$$

[5]The heuristic derivation of [51] in the Hamiltonian framework shares many features in common with the work of [95–98], extended to some higher-dimensional models in [99, 100]. The central feature of [51] is the Iwasawa decomposition which enables one to clearly see the role of off-diagonal variables.

Here, the ${f^i}_{jk}$ are the structure constants of the spatial homogeneity group. Similarly, for a 1-form and a 2-form,

$$A^{(1)} = A_i(x^0)\psi^i, \qquad A^{(2)} = \frac{1}{2}A_{ij}(x^0)\psi^i \wedge \psi^j, \qquad \text{etc.} \tag{3.1.31}$$

The Hamiltonian constraint yielding the field equations in the spatially homogeneous context[6] is obtained by substituting the form of the fields in the general Hamiltonian constraint and contains, of course, no explicit spatial gradients since the fields are homogeneous. Note, however, that the structure constants ${f^i}_{ik}$ contain implicit spatial gradients. The Hamiltonian can now be decomposed as before and reads

$$\begin{aligned} \mathcal{H}^{\mathrm{UL}} &= K + V^{\mathrm{UL}}, \\ V^{\mathrm{UL}} &= V_S + V_G^{\mathrm{UL}} + \sum_p \left(V_{(p)}^{\mathrm{el}} + V_{(p)}^{\mathrm{UL,magn}} \right), \end{aligned} \tag{3.1.32}$$

where K, V_S and $V_{(p)}^{\mathrm{el}}$, which do not involve spatial gradients, are unchanged and where V_ϕ disappears since $\partial_i \phi = 0$. The potential V_G is given by [49]

$$V_G \equiv -\mathrm{g}R = \frac{1}{4} \sum_{i\neq j, i\neq k, j\neq k} e^{-2G_{ijk}(\beta)} ({C^i}_{jk})^2 + \frac{1}{2} \sum_j e^{-2\bar{m}_j(\beta)} \left({C^i}_{jk}\, {C^k}_{ji} + \cdots \right), \tag{3.1.33}$$

where the linear forms $G_{ijk}(\beta)$ (with i, j, k distinct) read

$$G_{ijk}(\beta) = 2\beta^i + \sum_{m\,:\, m\neq i, m\neq j, m\neq k} \beta^m, \tag{3.1.34}$$

and where the ellipsis represent the terms in the first sum that arise upon taking $i = j$ or $i = k$. The structure constants in the Iwasawa frame (with respect to the coframe in Equation (3.1.30)) are related to the structure constants ${f^i}_{jk}$ through

$${C^i}_{jk} = \sum_{i', j', k'} {f^{i'}}_{j'k'} \mathcal{N}^{-1}_{ii'} \mathcal{N}_{jj'} \mathcal{N}_{kk'} \tag{3.1.35}$$

and depend therefore on the dynamical variables. Similarly, the potential $V_{(p)}^{\mathrm{magn}}$ becomes

$$V_{(p)}^{\mathrm{magn}} = \frac{1}{2\,(p+1)!} \sum_{i_1, i_2, \cdots, i_{p+1}} e^{-2m_{i_1\cdots i_{p+1}}(\beta)} (\mathcal{F}^h_{(p)\, i_1 \cdots i_{p+1}})^2, \tag{3.1.36}$$

where the field strengths $\mathcal{F}^h_{(p)\, i_1 \cdots i_{p+1}}$ reduce to the "AC" terms in dA and depend on the potentials and the off-diagonal Iwasawa variables.

[6]This Hamiltonian exists if ${f^i}_{ik} = 0$, as we shall assume from now on.

The Ultralocal Hamiltonian

Let us now come back to the general, inhomogeneous case and express the dynamics in the frame $\{dx^0, \psi^i\}$ where the ψ^i's form a "generic" non-holonomic frame in space,

$$d\psi^i = -\frac{1}{2} f^i{}_{jk}(x^m)\, \psi^j \wedge \psi^k. \tag{3.1.37}$$

Here the $f^i{}_{jk}$'s are in general space-dependent. In the non-holonomic frame, the exact Hamiltonian takes the form

$$\mathcal{H} = \mathcal{H}^{\mathrm{UL}} + \mathcal{H}^{\mathrm{gradient}}, \tag{3.1.38}$$

where the ultralocal part $\mathcal{H}^{\mathrm{UL}}$ is given by Equations (3.1.32) and (3.1.33) with the relevant $f^i{}_{jk}$'s, and where $\mathcal{H}^{\mathrm{gradient}}$ involves the spatial gradients of $f^i{}_{jk}$, β^m, ϕ and $\mathcal{N}_{ij}$.

The first part of the BKL conjecture states that one can drop $\mathcal{H}^{\mathrm{gradient}}$ asymptotically; namely, the dynamics of a generic solution of the Einstein–p-form-dilaton equations (not necessarily spatially homogeneous) is asymptotically determined, as one goes to the spatial singularity, by the ultralocal Hamiltonian

$$H^{\mathrm{UL}} = \int d^d x\, \mathcal{H}^{\mathrm{UL}}, \tag{3.1.39}$$

provided that the phase space constants $f^i{}_{jk}(x^m) = -f^i{}_{kj}(x^m)$ are such that all exponentials in the above potentials do appear. In other words, the f's must be chosen such that none of the coefficients of the exponentials, which involve f and the fields, identically vanishes – as would be the case, for example, if $f^i{}_{jk} = 0$ since then the potentials V_G and $V^{\mathrm{magn}}_{(p)}$ are equal to zero. This is always possible because the $f^i{}_{jk}$, even though independent of the dynamical variables, may in fact depend on x and so are not required to fulfill relations "$ff = 0$" analogous to the Bianchi identity since one has instead "$\partial f + ff = 0$".

3.1.5 The Sharp Wall Limit

The second step in the BKL-limit is to take the sharp wall limit of the potentials.[7] This leads to the billiard picture. It is crucial here that the coefficients in front of the dominant walls are all positive. Again, just as for the first step, this limit has not been rigorously justified.

To summarize, we first recall that the ultralocal Hamiltonian reads explicitly

$$\mathcal{H}^{\mathrm{UL}} = K + V_S + V_G^{\mathrm{UL}} + \sum_p \left(V^{\mathrm{el}}_{(p)} + V^{\mathrm{UL,magn}}_{(p)} \right) \tag{3.1.40}$$

where the potentials are completely determined by exponentials containing linear forms $\omega(\beta)$ in the scale factors β. We can then write the ultralocal Hamiltonian as

$$\begin{aligned} \mathcal{H}^{\mathrm{UL}} &= K + \sum_{i<j} c_S\, e^{-2s_{ji}(\beta)} + \sum_{i\neq j, i\neq k, j\neq k} c_G\, e^{-2G_{ijk}(\beta)} \\ &\quad + \sum_{i_1<i_2<\cdots<i_p} c_E\, e^{-2e_{i_1\cdots i_p}(\beta)} + \sum_{i_1<i_2<\cdots<i_{p+1}} c_M\, e^{-2m_{i_1\cdots i_{p+1}}(\beta)}, \end{aligned} \tag{3.1.41}$$

[7] In this article we will exclusively restrict ourselves to considerations involving the sharp wall limit. However, in recent work [93] it was argued that in order to have a rigorous treatment of the dynamics close to the singularity also in the chaotic case, it is necessary to go beyond the sharp wall limit. This implies that one should retain the exponential structure of the dominant walls.

where the coefficients are irrelevant for the dynamics.

The idea is now that as one goes to the singularity, the exponential potentials get sharper and sharper, and in the strict limit $x^0 \to \infty$ they can be replaced by corresponding Θ-functions, defined by

$$\Theta(x) := \begin{cases} 0 & x < 0 \\ \infty & x > 0. \end{cases} \tag{3.1.42}$$

Since $a\Theta(x) = \Theta(x)$ for all $a > 0$, the coefficients in (3.1.41) may be neglected and that the "sharp wall" Hamiltonian becomes

$$\begin{aligned} \mathcal{H}^{\text{sharp}} &= K + \sum_{i<j} \Theta\left(-2s_{ji}(\beta)\right) + \sum_{i\neq j, i\neq k, j\neq k} \Theta(-2\alpha_{ijk}(\beta)) \\ &\quad + \sum_{i_1<i_2<\cdots<i_p} \Theta(-2e_{i_1\cdots i_p}(\beta)) + \sum_{i_1<i_2<\cdots<i_{p+1}} \Theta(-2m_{i_1\cdots i_{p+1}}(\beta)). \end{aligned} \tag{3.1.43}$$

For the readers convenience we list here all of the linear forms appearing in the arguments of the step functions:

- *Symmetry walls*, originating from off-diagonal components of the spatial metric,

$$s_{ji}(\beta) = \beta^j - \beta^i, \qquad j > i. \tag{3.1.44}$$

- *Gravity walls*, originating from the spatial curvature,

$$G_{ijk}(\beta) = 2\beta^i + \sum_{m\,:\,m\neq i, m\neq j, m\neq k} \beta^m. \tag{3.1.45}$$

- *Electric walls*, originating from the electric components of each p-form,

$$e_{i_1\cdots i_p}(\beta) = \beta^{i_1} + \cdots + \beta^{i_p} + \frac{\lambda^{(p)}}{2}\phi. \tag{3.1.46}$$

- *Magnetic walls*, originating from the magnetic components of each p-form,

$$m_{i_1\cdots i_{p+1}}(\beta) = \sum_{j\notin\{i_1,i_2,\cdots i_{p+1}\}} \beta^j - \frac{\lambda^{(p)}}{2}\phi. \tag{3.1.47}$$

We shall refer to these linear forms as *wall forms* and denote them collectively by $\omega(\beta) = \omega_\mu \beta^\mu$. Each wall form defines a co-dimension one subspace W in the space of scale factors $\mathcal{M}_\beta$, which is defined by the locus where the associated wall form vanishes, i.e.

$$W[\omega] := \big\{\beta \in \mathcal{M}_\beta \;\big|\; \omega(\beta) = 0\big\}. \tag{3.1.48}$$

It is now clear that the complete dynamics of any gravitational theory, described by the action (3.1.1), reduces in the limit $x^0 \to \infty$ to the dynamics of the scale factors $\beta(x^0, \mathbf{x})$ at each spatial point $\mathbf{x}$. Because the potential walls are infinite (and positive), the motion of the scale factors is constrained to the region of $\mathcal{M}_\beta$ where the arguments of all Θ-functions are negative, i.e., to

$$s_{ji}(\beta) \geq 0 \; (i < j), \qquad G_{ijk}(\beta) \geq 0, \qquad e_{i_1 \cdots i_p}(\beta) \geq 0, \qquad m_{i_1 \cdots i_{p+1}}(\beta) \geq 0. \tag{3.1.49}$$

This corresponds to the "positive side" of all the walls $W[\omega]$. In this region, the dynamics is governed by the kinetic term K, implying that the motion is a geodesic for the metric in the space of scale factors $\mathcal{M}_\beta$. Since that metric is flat, this is a straight line. In addition, the Hamiltonian constraint $\mathcal{H} = 0$ reduces to $K = 0$ away from the potential walls and therefore forces the straight line to be null. In other words, the dynamics of the particle $\beta^\mu(x^0)$ is given by *lightlike* motion in $\mathcal{M}_\beta$. We shall furthermore assume that the time orientation in the space of the scale factors is such that this lightlike motion is future oriented ($\mathrm{g} \to 0$ in the future).

3.1.6 Dynamics as a Billiard in Hyperbolic Space

Our analysis so far has led us to the realization that the original gravity-p-form action in (3.1.1) has a very simple description in the limit $x^0 \to \infty$ (proper time $T \to 0^+$) in terms of the lightlike linear motion of the scale factors $\beta^\mu(x^0, \mathbf{x})$ in a region of the space $\mathcal{M}_\beta$. This implies that in the BKL-limit we may replace the action (3.1.1) by an effective geodesic action at each spatial point $\mathbf{x}$ given by

$$S_{\mathrm{BKL}}(\beta) := \int dx^0 \Big[\gamma_{\mu\nu} \frac{\partial \beta^\mu}{\partial x^0} \frac{\partial \beta^\nu}{\partial x^0} - V(\beta)\Big], \tag{3.1.50}$$

which is a non-linear sigma model describing maps from the worldline parametrized by x^0 into the target space $\mathcal{M}_\beta$. In defining (3.1.50) we have set $N = \sqrt{\mathrm{g}}$ as before. All the remaining non-trivial information about the original action (3.1.1) is now stored in the potential $V(\beta)$, which takes the general form

$$V(\beta) = \sum_{\omega \in \{\text{wall forms}\}} e^{-2\omega(\beta)}, \tag{3.1.51}$$

where we sum over all possible wall forms.

Now let us analyze this action in more detail. Firstly, one may check that all the walls $W[\omega]$, defined by Equations (3.1.47) and (3.1.48), are timelike hyperplanes in $\mathcal{M}_\beta$. This follows from the fact that the squared norms of all the wall forms ω in (3.1.47) are positive:

$$\begin{aligned} (\omega|\omega) &:= \gamma^{\mu\nu}\omega_\mu \omega_\nu \\ &= \sum_i \omega_i \omega_i - \frac{1}{d-1}\Big(\sum_i \omega_i\Big)\Big(\sum_j \omega_j\Big) + \omega_\phi \omega_\phi \;>\; 0, \end{aligned} \tag{3.1.52}$$

where we have used the inverse metric $\gamma^{\mu\nu}$ in $\mathcal{M}_\beta$. Explicitly, for the wall forms defined in

(3.1.47) one finds

$$\begin{aligned}(s_{ji}|s_{ji}) &= 2,\\ (G_{ijk}|G_{ijk}) &= 2,\\ (e_{i_1\cdots i_p}|e_{i_1\cdots i_p}) &= \frac{p(d-p-1)}{d-1}+\frac{\left(\lambda^{(p)}\right)^2}{4},\\ (m_{i_1\cdots i_{p+1}}|m_{i_1\cdots i_{p+1}}) &= \frac{p(d-p-1)}{d-1}+\frac{\left(\lambda^{(p)}\right)^2}{4}.\end{aligned} \tag{3.1.53}$$

Because these potential walls are timelike, they have a non-empty intersection with the forward light cone in the space of the scale factors. The motion thus corresponds to lightlike free flight motion in $\mathcal{M}_\beta$, occasionally interrupted by collisions with the timelike hyperplanes $W[\omega] \subset \mathcal{M}_\beta$. When the null straight line representing the evolution of the scale factors hits one of the walls, it gets reflected according to the rule [50]

$$v^\mu \quad \to \quad v^\mu - 2\frac{v^\nu\omega_\nu}{(\omega|\omega)}\omega^\mu, \tag{3.1.54}$$

where v^μ is the velocity vector associated with particle $\beta^\mu(x^0)$ (see Section 3.1.7). Since the hyperplanes are timelike, this reflection preserves the time orientation of the motion and therefore belong to the orthochronous Lorentz group $O^\uparrow(d,1)$ which is a subgroup of the isometry group of $\mathcal{M}_\beta$.[8]

The motion is thus a succession of future-oriented null straight line segments interrupted by reflections against the walls, where the motion undergoes a reflection belonging to $O^\uparrow(k,1)$. We can simplify the description of the dynamics further by noting that not all of the scale-factors are independent. This follows from the Hamiltonian constraint, which enforces

$$\sum_i \frac{\partial\beta^i}{\partial x^0}\frac{\partial\beta^i}{\partial x^0} - \Big(\sum_i \frac{\partial\beta^i}{\partial x^0}\Big)\Big(\sum_j \frac{\partial\beta^j}{\partial x^0}\Big) + \frac{\partial\phi}{\partial x^0}\frac{\partial\phi}{\partial x^0} = 0. \tag{3.1.55}$$

Since the collisions against the hyperplance $W[\omega]$ preserves the time-orientation, it is consistent to use the Hamiltonian constraint to restrict to a physical subset of scale factors. This is conveniently done by projecting the motion onto the unit hyperboloid $\mathbb{H}_d$, defined by

$$\mathbb{H}_d := \big\{\beta \in \mathcal{M}_\beta \;\big|\; (\beta|\beta) = -1\big\}. \tag{3.1.56}$$

Since we have chosen the time-orientation such that $\sum_i \beta^i > 0$ the motion is now confined to the positive sheet $\mathbb{H}_d^+$ of the unit hyperboloid $\mathbb{H}_d$. We then recall that the positive sheet of the unit hyperboloid provides a model for *hyperbolic space* (see, e.g., [74]). The straight line motion on $\mathcal{M}_\beta$ projects to lightlike geodesic motion on $\mathbb{H}_d^+$, while the walls $W[\omega]$ project onto hyperplanes in hyperbolic space. In this new picture the dynamics corresponds to that of a particle bouncing around in a bounded region of hyperbolic space, very much like a billiard ball bouncing around on a billiard table. This is the origin of the name *cosmological billiards.*

The description of the dynamics in terms of a hyperbolic billiard is very useful for determining the nature of the approach to the singularity. In the original work of BKL, it was

[8] If there are k dilatons, this changes to $O^\uparrow(d-1+k,1)$.

found that the approach to the singularity may exhibit chaotic behaviour. This can be most easily understood in our present framework.

As mentioned above, the intersection of a timelike hyperplane $W[\omega]$ with the unit hyperboloid $\mathbb{H}_d$ defines a hyperplane in hyperbolic space $\mathbb{H}_d^+$. The region in $\mathbb{H}_d^+$ on the positive side of all hyperplanes is the allowed dynamical region and is called the "billiard table", denoted $\mathcal{B}_\beta \subset \mathbb{H}_d^+$. It is never compact in the cases relevant to gravity, but it may or may not have finite volume. In this picture it is also clear that not all of the walls $W[\omega]$ will have dynamical relevance, since some of the walls will be "behind" others, and will thus never be seen by the geodesic motion. The billiard table is therefore uniquely defined by a subset $\{W'\} \subset \{W\}$ of *dominant walls*, and can be described as follows

$$\mathcal{B}_\beta := \big\{\beta \in \mathcal{M}_\beta \;\big|\; \omega_A(\beta) \geq 0,\ A = 1, \ldots, d+1\big\}, \tag{3.1.57}$$

where ω_A denotes the $d+1$ dominant walls. The billiard table thus defines a *simplex* in hyperbolic space, i.e. a region in a d-dimensional space bounded by $d+1$ walls. If the intersection between two dominant hyperplanes W and W' in $\mathcal{M}_\beta$ lies outside of the lightcone, then the projected hyperplanes in $\mathbb{H}_d^+$ will never intersect. As a consequence, the projected region $\mathcal{B}_\beta$ will be of infinite volume. In contrast, when all of the dominant walls intersect inside or on the lightcone, then the billiard table $\mathcal{B}_\beta$ will be of finite volume.

When the volume of the billiard table is finite, the collisions with the potential walls never end (for generic initial data) and the motion is chaotic. Onn the other hand, when the volume is infinite, generic initial data lead to a motion that ultimately freely runs away to infinity since after a few collisions there will be no wall to interrupt the motion. The dynamics is then non-chaotic. In Section 3.2 we will describe a powerful mathematical technique which makes it straightforward to determine whether or not the volume of $\mathcal{B}_\beta$ is finite, and hence to determine the type of dynamics.

Let us now for convenience list which of the possible wall forms in (3.1.47) are dominant. These are

- dominant symmetry walls ($i = 1, 2, \cdots, d-1$):

$$\beta^{i+1} - \beta^i = 0, \tag{3.1.58}$$

- dominant gravity walls:

$$2\beta^1 + \beta^2 + \cdots + \beta^{d-2} = 0, \tag{3.1.59}$$

- dominant electric wall:

$$\beta^1 + \cdots + \beta^p + \frac{\lambda^{(p)}}{2}\phi = 0, \tag{3.1.60}$$

- dominant magnetic wall:

$$\beta^1 + \cdots + \beta^{d-p-1} - \frac{\lambda^{(p)}}{2}\phi = 0. \tag{3.1.61}$$

3.1.7 Recovering the Kasner Solution

We conclude this section by discussing the relation between the billiard description of the BKL-limit and the more conventional description in terms of a (possibly infinite) sequence of different Kasner solutions. The free motion between two bounces in the billiard is a straight line in the space of the scale factors. In terms of the original metric components, this simply corresponds to a Kasner solution with dilaton. To see this, we note that a straight line in $\mathcal{M}_\beta$ is described by

$$\beta^\mu = v^\mu x^0 + \beta_0^\mu,$$

where the "velocities" v^μ are subject to

$$\sum_i (v^i)^2 - \Big(\sum_i v^i\Big)\Big(\sum_j v^j\Big) + v_\phi^2 = 0,$$

since the motion is lightlike by the Hamiltonian constraint. The proper time $dT = -\sqrt{\mathrm{g}}\, dx^0$ is then $T = B\exp(-Kx^0)$, with $K = \sum_i v^i$ for some constant B (we assume, as before, that the singularity is at $T = 0^+$). By the following redefinition

$$p^\mu := \frac{v^\mu}{\sum_i v^i}$$

we then obtain the standard Kasner solution

$$ds^2 = -dT^2 + \sum_i T^{2p^i} \left(dx^i\right)^2, \tag{3.1.62}$$

$$\phi = -p_\phi \ln T + A, \tag{3.1.63}$$

subject to the constraints

$$\sum_i p^i = 1, \qquad \sum_i (p^i)^2 + p_\phi^2 = 1, \tag{3.1.64}$$

where A is a constant of integration and where the coordinates x^i have been suitably rescaled (if necessary). We may therefore conclude that each segment of free flight motion of the billiard ball is in one-to-one correspondence with a certain Kasner solution defined by (3.1.63). Moreover, each reflection against a wall has the effect of modifying the dynamics to a different Kasner solution. Hence, a finite volume billiard indeed leads to an infinite sequence of Kasner solutions as we approach the singularity, while a finite volume billiard leads to dynamics which eventually settle down in one specific Kasner regime all the way to $T = 0^+$.

3.2 Kac-Moody Billiards

We shall now reinterpret the results of previous sections within the mathematical framework introduced in Chapter 2. In particular, we explore the correspondence between Weyl groups of Lorentzian Kac-Moody algebras and the limiting behavior of the dynamics of gravitational theories close to a spacelike singularity.

We have seen in Section 3.1.6 that in the BKL-limit, the dynamics of gravitational theories is equivalent to a billiard dynamics in a region of hyperbolic space. In the generic case, the billiard region has no particular feature. However, as shown in [44], in the special cases of the supergravities arising as low-energy limits of string theory or M-theory, the billiard exhibits nice regularity properties which are linked to underlying hyperbolic Kac-Moody algebras.

In fact, this feature arises for all gravitational theories whose toroidal dimensional reduction to three dimensions exhibits hidden symmetries, in the sense that the reduced theory can be reformulated as three-dimensional gravity coupled to a nonlinear sigma-model based on a coset space $K_3\backslash G_3$, where K_3 is the maximal compact subgroup of G_3. The "hidden" symmetry group G_3 is also called, by a generalization of language, "the U-duality group" [11]. This situation covers the cases of pure gravity in any spacetime dimension, as well as all known supergravity models. In all these cases, the billiard region is the fundamental domain of a Lorentzian Coxeter group ("Coxeter billiard"). Furthermore, the Coxeter group in question is crystallographic and turns out to be the Weyl group of a Lorentzian Kac–Moody algebra. The billiard table is then the fundamental Weyl chamber of a Lorentzian Kac–Moody algebra [44, 52] and the billiard is also called a "Kac–Moody billiard". This enables one to reformulate the dynamics as a motion in the Cartan subalgebra of the Lorentzian Kac–Moody algebra, hinting at the potential – and still conjectural at this stage – existence of a deeper, infinite-dimensional symmetry of the theory.

3.2.1 The Kac-Moody Billiard of Pure Gravity

We start by providing some examples of theories leading to regular billiards, focusing first on pure gravity in any number of D (> 3) spacetime dimensions. In this case, there are $d = D - 1$ scale factors β^i and the relevant walls are the symmetry walls, Equation (3.1.58),

$$s_i(\beta) \equiv \beta^{i+1} - \beta^i = 0 \qquad (i = 1, 2, \cdots, d-1), \tag{3.2.1}$$

and the curvature wall, Equation (3.1.59),

$$r(\beta) \equiv 2\beta^1 + \beta^2 + \cdots + \beta^{d-2} = 0. \tag{3.2.2}$$

There are thus d relevant walls, which define a simplex in $(d-1)$-dimensional hyperbolic space $\mathcal{H}_{d-1}$. The scalar products of the linear forms defining these walls are easily computed. One finds as non-vanishing products

$$\begin{aligned}
(s_i|s_i) &= 2 \qquad (i = 1, \cdots, d-1),\\
(r|r) &= 2,\\
(s_{i+1}|s_i) &= -1 \qquad (i = 2, \cdots, d-1)\\
(r|s_1) &= -1,\\
(r|s_{d-2}) &= -1.
\end{aligned} \tag{3.2.3}$$

The matrix of the scalar products of the wall forms is thus the Cartan matrix of the (simply-laced) Lorentzian Kac–Moody algebra A^{++}_{d-2} with Dynkin diagram as in Figure 3.1. The roots of the underlying finite-dimensional algebra A_{d-2} are given by s_i $(i = 1, \cdots, d-3)$ and r. The affine root is s_{d-2} and the overextended root is s_{d-1}.

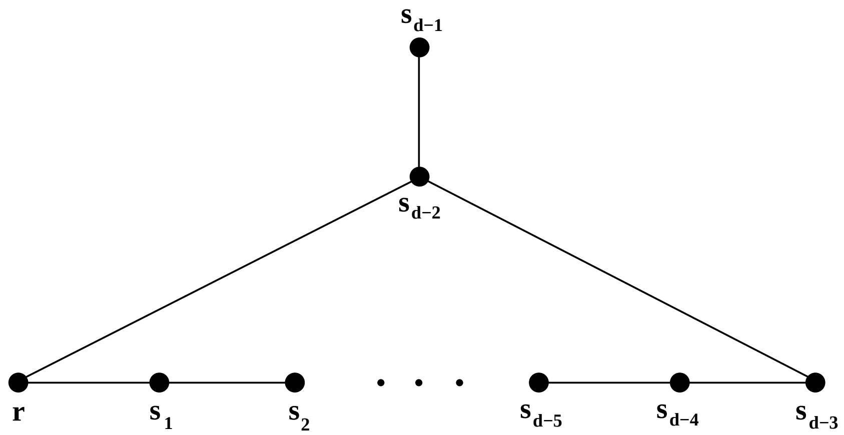


Figure 3.1: The Dynkin diagram of the hyperbolic Kac–Moody algebra A_{d-2}^{++} which controls the billiard dynamics of pure gravity in $D = d + 1$ dimensions. The nodes $s_1, \cdots, s_{d-1}$ represent the "symmetry walls" arising from the off-diagonal components of the spatial metric, and the node r corresponds to a "curvature wall" coming from the spatial curvature. The horizontal line is the Dynkin diagram of the underlying A_{d-2}-subalgebra and the two topmost nodes, s_{d-2} and s_{d-1}, give the affine- and overextension, respectively.

Accordingly, in the case of pure gravity in any number of spacetime dimensions, one finds also that the billiard region is regular. This provides new examples of Coxeter billiards, with Coxeter groups A_{d-2}^{++}, which are also Kac–Moody billiards since the Coxeter groups are the Weyl groups of the Kac–Moody algebras A_{d-2}^{++}.

3.2.2 The Kac-Moody Billiard for the Coupled Gravity-3-Form System

Let us review the conditions that must be fulfilled in order to get a Kac–Moody billiard and let us emphasize how restrictive these conditions are. The billiard region computed from any theory coupled to gravity with n dilatons in $D = d + 1$ dimensions always defines a convex polyhedron in a $(d+n-1)$-dimensional hyperbolic space $\mathcal{H}_{d+n-1}$. In the general case, the dihedral angles between adjacent faces of $\mathcal{H}_{d+n-1}$ can take arbitrary continuous values, which depend on the dilaton couplings, the spacetime dimensions and the ranks of the p-forms involved. However, only if the dihedral angles are integer submultiples of π (meaning of the form π/k for $k \in \mathbb{Z}_{\geq 2}$) do the reflections in the faces of $\mathcal{H}_{d+n-1}$ define a Coxeter group. In this special case the polyhedron is called a *Coxeter polyhedron*. This Coxeter group is then a (discrete) subgroup of the isometry group of $\mathcal{H}_{d+n-1}$.

In order for the billiard region to be identifiable with the fundamental Weyl chamber of a Kac–Moody algebra, the Coxeter polyhedron should be a *simplex*, i.e., bounded by $d + n$ walls in a $d + n - 1$-dimensional space. In general, the Coxeter polyhedron need not be a simplex.

There is one additional condition. The angle ϑ between two adjacent faces i and j is given, in terms of the Coxeter exponents, by

$$\vartheta = \frac{\pi}{m_{ij}}. \tag{3.2.4}$$

Coxeter groups that correspond to Weyl groups of Kac–Moody algebras are the *crystallographic* Coxeter groups for which $m_{ij} \in \{2, 3, 4, 6, \infty\}$. So, the requirement for a gravitational theory to have a Kac–Moody algebraic description is not just that the billiard region is a Coxeter simplex but also that the angles between adjacent walls are such that the group of reflections in these walls is crystallographic.

These conditions are very restrictive and hence gravitational theories which can be mapped to a Kac–Moody algebra in the BKL-limit are rare.

The Kac-Moody Billiard of Eleven-Dimensional Supergravity

Consider for instance the action (3.1.1) for gravity coupled to a single three-form in $D = d+1$ spacetime dimensions. We assume $D \geq 6$ since in lower dimensions the 3-form is equivalent to a scalar ($D = 5$) or has no degree of freedom ($D < 5$).

Theorem: Whenever a p-form ($p \geq 1$) is present, the curvature wall is subdominant as it can be expressed as a linear combination with positive coefficients of the electric and magnetic walls of the p-forms. (These walls are all listed in Section 3.1.6.)

Proof: The dominant electric wall is (assuming the presence of a dilaton)

$$e_{1\cdots p}(\beta) \equiv \beta^1 + \beta^2 + \cdots + \beta^p - \frac{\lambda_p}{2}\phi = 0, \tag{3.2.5}$$

while one of the magnetic wall reads

$$m_{1,p+1,\cdots,d-2}(\beta) \equiv \beta^1 + \beta^{p+1} + \cdots + \beta^{d-2} + \frac{\lambda_p}{2}\phi = 0, \tag{3.2.6}$$

so that the dominant curvature wall is just the sum $e_{1\cdots p}(\beta) + m_{1,p+1,\cdots,d-2}(\beta)$.

It follows that in the case of gravity coupled to a single three-form in $D = d+1$ spacetime dimensions, the relevant walls are the symmetry walls, Equation (3.1.58),

$$s_i(\beta) \equiv \beta^{i+1} - \beta^i = 0, \qquad i = 1, 2, \cdots, d-1 \tag{3.2.7}$$

(as always) and the electric wall

$$e_{123}(\beta) \equiv \beta^1 + \beta^2 + \beta^3 = 0 \tag{3.2.8}$$

($D \geq 8$) or the magnetic wall

$$m_{1\cdots D-5}(\beta) \equiv \beta^1 + \beta^2 + \cdots \beta^{D-5} = 0 \tag{3.2.9}$$

($D \leq 8$). Indeed, one can express the magnetic walls as linear combinations with (in general non-integer) positive coefficients of the electric walls for $D \geq 8$ and vice versa for $D \leq 8$.

Hence the billiard table is always a simplex (this would not be true had one a dilaton and various forms with different dilaton couplings).

However, it is only for $D = 11$ that the billiard is a Coxeter billiard. In all the other spacetime dimensions, the angle between the relevant p-form wall and the symmetry wall that does not intersect it orthogonally is not an integer submultiple of π. More precisely, the angle between

- the magnetic wall β^1 and the symmetry wall $\beta^2 - \beta^1$ ($D = 6$),
- the magnetic wall $\beta^1 + \beta^2$ and the symmetry wall $\beta^3 - \beta^2$ ($D = 7$), and
- the electric wall $\beta^1 + \beta^2 + \beta^3$ and the symmetry wall $\beta^4 - \beta^3$ ($D \geq 8$),

is easily verified to be an integer submultiple of π only for $D = 11$, for which it is equal to $\pi/3$.

From the point of view of the regularity of the billiard, the spacetime dimension $D = 11$ is thus privileged. This is of course also the dimension privileged by supersymmetry. It is quite intriguing that considerations *a priori* quite different (billiard regularity on the one hand, supersymmetry on the other hand) lead to the same conclusion that the gravity-3-form system is quite special in $D = 11$ spacetime dimensions.

For completeness, we here present the wall system relevant for the special case of $D = 11$. We obtain ten dominant wall forms, which we rename $\alpha_1, \cdots, \alpha_{10}$,

$$\begin{aligned} \alpha_m(\beta) &= \beta^{m+1} - \beta^m \qquad (m = 1, \cdots, 10), \\ \alpha_{10}(\beta) &= \beta^1 + \beta^2 + \beta^3. \end{aligned} \tag{3.2.10}$$

Then, defining a new collective index $i = (m, 10)$, we see that the scalar products between these wall forms can be organized into the matrix

$$A_{ij} = 2\frac{(\alpha_i|\alpha_j)}{(\alpha_i|\alpha_i)} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}, \tag{3.2.11}$$

which can be identified with the Cartan matrix of the hyperbolic Kac–Moody algebra E_{10} that we have encountered in Section 2.5.4. We again display the corresponding Dynkin diagram in Figure 7.3, where we point out the explicit relation between the simple roots and the walls of the Einstein–3-form theory. It is clear that the nine dominant symmetry wall forms correspond to the simple roots α_m of the subalgebra $\mathfrak{sl}(10, \mathbb{R})$. The enlargement to E_{10} is due to the tenth exceptional root realized here through the dominant electric wall form e_{123}.

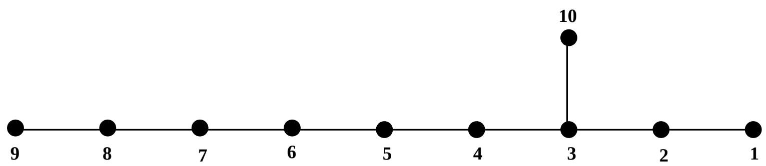


Figure 3.2: The Dynkin diagram of E_{10}. Labels $m = 1, \cdots, 9$ enumerate the nodes corresponding to simple roots, α_m, of the $\mathfrak{sl}(10, \mathbb{R})$ subalgebra and the exceptional node, labeled "10", is associated to the electric wall $\alpha_{10} = e_{123}$.

3.2.3 Dynamics in the Cartan Subalgebra

We have just learned that, in some cases, the group of reflections that describe the (possibly chaotic) dynamics in the BKL-limit is a Lorentzian Coxeter group. In this section we fully exploit this algebraic fact and show that whenever this Coxeter group is *crystallographic*, the dynamics takes place in the Cartan subalgebra $\mathfrak{h}$ of the Lorentzian Kac–Moody algebra $\mathfrak{g}$, for which the relevant Coxeter group is identified with the Weyl group $\mathcal{W}(\mathfrak{g})$. Moreover, we show that the "billiard table" can be identified with the fundamental Weyl chamber in $\mathfrak{h}$.

Scale Factor Space and the Wall System

Let us first briefly review some of the salient features encountered so far in the analysis. In the following we denote by $\mathcal{M}_\beta$ the Lorentzian "scale factor"-space (or β-space) in which the billiard dynamics takes place. Recall that the metric in $\mathcal{M}_\beta$, induced by the Einstein–Hilbert action, is a flat Lorentzian metric, whose explicit form in terms of the (logarithmic) scale factors reads

$$G_{\mu\nu}\, d\beta^\mu\, d\beta^\nu = \sum_{i=1}^{d} d\beta^i\, d\beta^i - \left(\sum_{i=1}^{d} d\beta^i\right)\left(\sum_{j=1}^{d} d\beta^j\right) + d\phi\, d\phi, \tag{3.2.12}$$

where d counts the number of physical spatial dimensions (see Section 3.1.6). The role of all other "off-diagonal" variables in the theory is to interrupt the free-flight motion of the particle, by adding walls in $\mathcal{M}_\beta$ that confine the motion to a limited region of scale factor space, namely a convex cone bounded by timelike hyperplanes. When projected onto the unit hyperboloid, this region defines a simplex in hyperbolic space which we refer to as the "billiard table".

One has, in fact, more than just the walls. The theory provides these walls with a specific normalization through the Lagrangian, which is crucial for the connection to Kac–Moody algebras. Let us therefore discuss in somewhat more detail the geometric properties of the wall system. The metric, Equation (3.2.12), in scale factor space can be seen as an extension of a flat Euclidean metric in Cartesian coordinates, and reflects the Lorentzian nature of the vector space $\mathcal{M}_\beta$. In this space we may identify a pair of coordinates (β^i, ϕ) with the components of a vector $\beta \in \mathcal{M}_\beta$, with respect to a basis $\{\bar{u}_\mu\}$ of $\mathcal{M}_\beta$, such that

$$\bar{u}_\mu \cdot \bar{u}_\nu = G_{\mu\nu}. \tag{3.2.13}$$

The walls themselves are then defined by hyperplanes in this linear space, i.e., as linear forms $\omega = \omega_\mu \underline{\sigma}^\mu$, for which $\omega = 0$, where $\{\underline{\sigma}^\mu\}$ is the basis dual to $\{\bar{u}^\mu\}$. The pairing $\omega(\beta)$ between a vector $\beta \in \mathcal{M}_\beta$ and a form $\omega \in \mathcal{M}_\beta^\star$ is sometimes also denoted by $\langle \omega, \beta \rangle$, and for the two dual bases we have, of course,

$$\langle \underline{\sigma}^\mu, \bar{u}_\nu \rangle = \delta^\mu_\nu. \tag{3.2.14}$$

We therefore find that the walls can be written as linear forms in the scale factors:

$$\omega(\beta) = \sum_{\mu,\nu} \omega_\mu \beta^\nu \langle \underline{\sigma}^\mu, \bar{u}_\nu \rangle = \sum_\mu \omega_\mu \beta^\mu = \sum_{i=1}^d \omega_i \beta^i + \omega_\phi \phi. \tag{3.2.15}$$

We call $\omega(\beta)$ *wall forms*. With this interpretation they belong to the dual space $\mathcal{M}_\beta^\star$, i.e.,

$$\begin{aligned} \mathcal{M}_\beta^\star \ni \omega : \mathcal{M}_\beta \quad &\longrightarrow \quad \mathbb{R}, \\ \beta \quad &\longmapsto \quad \omega(\beta). \end{aligned} \tag{3.2.16}$$

From Equation (3.2.16) we may conclude that the walls bounding the billiard are the hyperplanes $\omega = 0$ through the origin in $\mathcal{M}_\beta$ which are orthogonal to the vector with components $\omega^\mu = G^{\mu\nu}\omega_\nu$.

It is important to note that it is the wall forms that the theory provides, as arguments of the exponentials in the potential, and not just the hyperplanes on which these forms ω vanish. The scalar products between the wall forms are computed using the metric in the dual space $\mathcal{M}_\beta^\star$, whose explicit form was given in Section 3.1.6,

$$(\omega|\omega') \equiv G^{\mu\nu}\omega_\mu\omega_\nu = \sum_{i=1}^d \omega_i\omega_i' - \frac{1}{d-1}\left(\sum_{i=1}^d \omega_i\right)\left(\sum_{j=1}^d \omega_j'\right) + \omega_\phi\omega_\phi', \qquad \omega, \omega' \in \mathcal{M}_\beta. \tag{3.2.17}$$

Scale Factor Space and the Cartan Subalgebra

The crucial additional observation is that (for the "interesting" theories) the matrix A associated with the relevant walls ω_A,

$$A_{AB} = 2\frac{(\omega_A|\omega_B)}{(\omega_A|\omega_A)} \tag{3.2.18}$$

is a Cartan matrix, i.e., besides having 2's on its diagonal, which is rather obvious, it has as off-diagonal entries non-positive integers (with the property $A_{AB} \neq 0 \Rightarrow A_{BA} \neq 0$). This Cartan matrix is of course symmetrizable since it derives from a scalar product.

For this reason, one can usefully identify the space of the scale factors with the Cartan subalgebra $\mathfrak{h}$ of the Kac–Moody algebra $\mathfrak{g}(A)$ defined by A. In that identification, the wall forms become the simple roots, which span the vector space $\mathfrak{h}^\star = \mathrm{span}\{\alpha_1, \cdots, \alpha_r\}$ dual to the Cartan subalgebra. The rank r of the algebra is equal to the number of scale factors β^μ, including the dilaton(s) if any ($(\beta^\mu) \equiv (\beta^i, \phi)$). This number is also equal to the number of walls since we assume the billiard to be a simplex. So, both A and μ run from 1 to r. The metric in $\mathcal{M}_\beta$, Equation (3.2.12), can be identified with the invariant bilinear form of $\mathfrak{g}$, restricted to the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$. The scale factors β^μ of $\mathcal{M}_\beta$ become then coordinates h^μ on the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}(A)$.

The Weyl group of a Kac–Moody algebra has been defined first in the space $\mathfrak{h}^\star$ as the group of reflections in the walls orthogonal to the simple roots. Since the metric is non-degenerate, one can equivalently define by duality the Weyl group in the Cartan algebra $\mathfrak{h}$ itself. For each reflection r_i on $\mathfrak{h}^\star$ we associate a dual reflection $r_i^\vee$ as follows,

$$r_i^\vee(\beta) = \beta - \langle \alpha_i, \beta \rangle \, \alpha_i^\vee, \qquad \beta, \alpha_i^\vee \in \mathfrak{h}, \tag{3.2.19}$$

which is the reflection relative to the hyperplane $\alpha_i(\beta) = \langle \alpha_i, \beta \rangle = 0$. This expression can be rewritten,

$$r_i^\vee(\beta) = \beta - \frac{2(\beta|\alpha_i^\vee)}{(\alpha_i^\vee|\alpha_i^\vee)}\alpha_i^\vee, \tag{3.2.20}$$

or, in terms of the scale factor coordinates β^μ,

$$\beta^\mu \longrightarrow \beta^{\mu\prime} = \beta^\mu - \frac{2(\beta|\omega^\vee)}{(\omega^\vee|\omega^\vee)}\omega^{\vee\mu}. \tag{3.2.21}$$

This is precisely the billiard reflection Equation (3.1.54) found in Section 3.1.5.

Thus, we have the following correspondence:

$$\begin{aligned} \mathcal{M}_\beta &\equiv \mathfrak{h}, \\ \mathcal{M}_\beta^\star &\equiv \mathfrak{h}^\star, \\ \omega_A(\beta) &\equiv \alpha_A(h), \\ \text{billiard wall reflections} &\equiv \text{fundamental Weyl reflections}, \end{aligned} \tag{3.2.22}$$

As we have also seen, the Kac–Moody algebra $\mathfrak{g}(A)$ is Lorentzian since the signature of the metric Equation (3.2.12) is Lorentzian. This fact will be crucial in the analysis of subsequent sections and is due to the presence of gravity, where conformal rescalings of the metric define timelike directions in scale factor space.

We thereby arrive at the following important result [44, 51, 52]:

The dynamics of certain gravity-p-form systems can in the BKL-limit be mapped to a billiard motion in the Cartan subalgebra $\mathfrak{h}$ of a Lorentzian Kac–Moody algebra $\mathfrak{g}$.

3.2.4 Hyperbolicity Implies Chaos

Let us now point out the main consequence of the identifications in (3.2.22). As in Section 3.1.6, let $\mathcal{B}_\beta$ denote the region in scale factor space to which the billiard motion is confined, i.e. the "billiard table",

$$\mathcal{B}_\beta = \{\beta \in \mathcal{M}_\beta \,|\, \omega_A(\beta) \geq 0\}, \tag{3.2.23}$$

where the index A runs over all dominant walls. On the algebraic side, recall from Section 2.4.2 that the fundamental Weyl chamber $\mathcal{C}_\mathfrak{h} \subset \mathfrak{h}$ is the closed convex (half) cone given by

$$\mathcal{C}_\mathfrak{h} = \{h \in \mathfrak{h} \,|\, \alpha_A(h) \geq 0;\ A = 1, \cdots, \operatorname{rank}\mathfrak{g}\}, \tag{3.2.24}$$

where we have put a subscript on $\mathcal{C}_{\mathfrak{h}}$ to emphasize that this is the fundamental chamber in the *Cartan subalgebra* $\mathfrak{h}$, in contrast to the fundamental chamber $\mathcal{C}$ of the dual root space $\mathfrak{h}^\star$. We see that the conditions $\alpha_A(h) \geq 0$ defining $\mathcal{C}_{\mathfrak{h}}$ are equivalent, upon examination of Equation (3.2.22), to the conditions $\omega_A(\beta) \geq 0$ defining the billiard table $\mathcal{B}_{\mathcal{M}_\beta}$. We may therefore make the crucial identification

$$\mathcal{C}_{\mathfrak{h}} \equiv \mathcal{B}_{\mathcal{M}_\beta}, \tag{3.2.25}$$

which means that the particle geodesic is confined to move within the fundamental Weyl chamber of $\mathfrak{h}$. From the billiard analysis in Section 3.1.6 we know that the piecewise motion in scale-factor space is controlled by geometric reflections with respect to the walls $\omega_A(\beta) = 0$. By comparing with the dominant wall forms and using the correspondence in Equation (3.2.22) we may then conclude that the geometric reflections of the coordinates $\beta^\mu(\tau)$ are controlled by the Weyl group in the Cartan subalgebra of $\mathfrak{g}(A)$.

Now recall further that the BKL dynamics is chaotic if and only if the billiard table is of finite volume when projected onto the unit hyperboloid. From our discussion of hyperbolic Coxeter groups in Section 2.4.2, and from the identification (3.2.25), we deduce that this feature is equivalent to hyperbolicity of the corresponding Kac–Moody algebra. This leads to the crucial statement [44, 51, 52]:

If the billiard region of a gravity-p-form system can be identified with the fundamental Weyl chamber of a hyperbolic Kac–Moody algebra, then the dynamics is chaotic.

As we have also discussed above, hyperbolicity can be rephrased in terms of the fundamental weights Λ_i defined as

$$\left\langle \Lambda_j, \alpha_i^\vee \right\rangle = \frac{2(\Lambda_j|\alpha_i)}{(\alpha_i|\alpha_i)} \equiv \delta_{ij}, \qquad \alpha_i^\vee \in \mathfrak{h},\ \Lambda_i \in \mathfrak{h}^\star. \tag{3.2.26}$$

Just as the fundamental Weyl chamber in $\mathfrak{h}^\star$ can be expressed in terms of the fundamental weights, the fundamental Weyl chamber in $\mathfrak{h}$ can be expressed in a similar fashion in terms of the fundamental coweights:

$$\mathcal{C}_{\mathfrak{h}} = \{\beta \in \mathfrak{h} \,|\, \beta = \sum_i a_i \Lambda_i^\vee,\ a_i \in \mathbb{R}_{\geq 0}\}. \tag{3.2.27}$$

Thus, hyperbolicity holds if and only if none of the fundamental weights are spacelike,

$$(\Lambda_i|\Lambda_i) \leq 0, \tag{3.2.28}$$

for all $i \in \{1, \cdots, \operatorname{rank} \mathfrak{g}\}$.

4

Compactification, Cohomology and Coxeter Groups

In the previous chapter we have seen that many gravitational theories of physical interest exhibit chaotic behavour in the BKL-limit, which in turn is related to an underlying hyperbolic Kac-Moody algebraic structure. In this chapter we shall analyze this in more detail, and in particular study the modification of the billiard dynamics when there is a spacetime splitting of the form $\Sigma = \mathbb{R} \times X$, with X being a smooth compact internal manifold of arbitrary topology. This question is of particular interest for clarifying the role of BKL-chaos within Big Bang/Big Crunch scenarios which typically relies on a smooth collapsing phase in the approach towards the singularity, in apparent contradiction with the chaotic BKL-behaviour exhibited by, for example, string theory related supergravities. We will learn that for certain special internal manifolds, chaos is removed by the compactification, thus providing a possible reconciliation with the aforementioned analysis. Along the way we will also discover fascinating new algebraic structures of the modified billiards which are described using the theory of *buildings*. This chapter is based on **Paper V**, written in collaboration with Marc Henneaux and Daniel H. Wesley. For explicit examples of the results presented in this chapter, as well as a complete classification, the reader is referred to **Paper V**.

4.1 Intermediate Asymptotics

The original BKL analysis is classical and has been pushed all the way to the singularity [45,47]. As such it is valid for any spatial topology. In the approach to a spacelike singularity there is an asymptotic decoupling of spatial points, and the dynamics becomes “ultralocal”. These results (decoupling of spatial points and chaotic oscillations) are by now well supported by extensive analytical and numerical evidence; a non-exhaustive list of references include [93,101–108].

The classical analysis has however obvious limitations and it is not clear what becomes of the BKL results for energy scales above the Planck scale, where quantum gravity effects cannot be ignored. In the standard BKL analysis, which ignores quantum effects, no walls are removed as the big crunch is approached. But, when some spatial dimensions are compact, there is a wide range of initial conditions for which walls corresponding to massive modes are always subdominant until the universe enters the quantum regime, and these walls are not relevant for the billiard analysis while the universe is described by classical physics [109,110]. For this broad set of initial conditions, there is no epoch in which the usual classical BKL analysis, with the full set of walls, applies.

For this reason, it is of interest to consider the regime of *intermediate asymptotics* where the curvature is much smaller than the Planck curvature but where the billiard analysis applies (see [111]). In that pre-Planck regime, it is not true that the topology of spacetime is irrelevant [109]. A modification might arise in the presence of p-forms because the massless spectrum of p-forms in the lower-dimensional theory depends on the de Rham cohomology of the internal manifold, and since massless degrees of freedom dominate in the BKL-limit before reaching the Planck scale [109], non-trivial compactification eliminates billiard walls corresponding to degrees of freedom which are rendered massive in the compactification. Depending on which walls are removed by the compactification, a chaotic theory can be rendered non-chaotic.

There are a number of reasons to better understand the interplay between BKL dynamics and compactification. On the physics side, the BKL-limit with compact internal spaces is relevant for certain types of cyclic or "pre-big bang" cosmological models built from string or M-theory (see, e.g., [112–115] and references therein). Cosmological models with a big crunch/big bang transition rely on a smooth collapsing phase as they approach the big crunch singularity, hence chaotic BKL oscillations close to the singularity are a potential problem for these models. If chaos can be removed by interpreting our four-dimensional world as an effective description of a more fundamental higher-dimensional theory where the "troublesome" billiard walls are eliminated through the cohomology of the internal space, the problems with BKL chaos in these models may be circumvented. On the mathematical side, the billiard regions and the reflection groups that emerge after compactification possess a rich structure which deserves investigation.

4.2 The "Uncompactified Billiard"

We shall here recall some useful features of the billiards which were discussed in Chapter 3. This corresponds to the case when all the walls are switched on, as it occurs when no internal dimension is compact [109]. This is also the case relevant when the BKL analysis is pushed all the way to the singularity. The billiard region is then the smallest possible one in the sense that the billiard region of all the other cases will contain the billiard region of the uncompactified theory. We shall denote by $\omega_{A'}$ the dominant walls of the uncompactified theory. While the number of dominant walls relevant to the compact cases might not be equal to the dimension M of $\mathcal{M}_\beta$, it turns out that for all theories whose dimensional reduction to three dimensions is described by a symmetric space, the number of dominant walls is equal to M [51].

In this case the dominant walls confine the billiard motion to the region $\mathcal{B}_\beta \subset \mathcal{M}_\beta$ defined by (see Section 3.1.6)

$$\mathcal{B}_\beta = \big\{\beta \in \mathcal{M}_\beta \,\big|\, \omega_{A'}(\beta) \geq 0,\ A' = 1, \cdots, M\big\}, \tag{4.2.1}$$

The billiard table is a simplex in $\mathbb{H}_m$. We shall call somewhat improperly the region (4.2.1) the "uncompactified billiard region" and its projection on $\mathbb{H}_m$ the "uncompactified billiard table".

The scalar products $(\omega_{A'}|\omega_{B'})$ between the dominant walls can be organized into a matrix,

$$A_{A'B'} = \frac{2(\omega_{A'}|\omega_{B'})}{(\omega_{A'}|\omega_{A'})}. \tag{4.2.2}$$

In the noncompact case, the matrix A turns out to possess the properties of a Lorentzian Cartan matrix [44,51], thereby identifying the dominant wall forms $\omega_{A'}$ with the simple roots of the Kac-Moody algebra $\mathfrak{g}(A)$ constructed from A [34]. This Kac-Moody algebra is the "overextension" $\mathfrak{u}^{++}$ of the U-duality algebra $\mathfrak{u}$. The group generated by the reflections in the billiard walls of the uncompactified theory is a Coxeter group, which is the Weyl group of the corresponding Kac-Moody algebra [44,51]. We shall denote it by $\mathcal{W}$.

Recall from Section 2.4.2) that the action of the Weyl group $\mathcal{W}$ on the (future) lightcone $\mathcal{O}^+$ splits up into a disjoint union of chambers, called *Weyl chambers*. One of these chambers is defined by the inequalities $\omega_{A'} \geq 0$ and is called the *fundamental chamber* $\mathcal{F}$.[1] Then, all other chambers in $\mathcal{O}^+$ correspond to images of $\mathcal{F}$ under $\mathcal{W}$. The action of $\mathcal{W}$ on the Weyl chambers is simply transitive. When projected onto $\mathbb{H}_m$, these chambers become m-simplices of finite volume. The fundamental chamber $\mathcal{F}$ is the uncompactified billiard region in which the chaotic dynamics takes place. The $m+1$ hyperplanes (or dominant walls) which bound the fundamental chamber, correspond to the codimension-one faces of $\mathcal{F}$ when projected onto $\mathbb{H}_n$ (see also Section 4.5.3).

4.3 Compactification and Cohomology

Compactification can modify the billiard, as was shown in [109, 110]. This occurs because the billiard dynamics in the intermediate regime considered here depends not only on the p–form menu, but also on the topology of the space upon which the theory is formulated, specifically on the de Rham cohomology $\mathcal{H}^p(X)$ of the compactification manifold X. The rules for constructing the noncompact billiard system are given above, and here we focus on the "selection rule" that describes how this billiard is modified after compactification.

4.3.1 Selection Rule

We study situations in which all spatial dimensions are compact, and thus spacetime Σ has topology

$$\Sigma = \mathbb{R} \times X, \tag{4.3.1}$$

where X is closed, compact, and orientable. Electric and magnetic walls, $e(\beta)$ and $m(\beta)$, arise from the electric and magnetic components, F_E and F_M, of a given p-form F. On a compact manifold X the p-form fields which remain massless during the compactification

[1] Note that in Chapters 2 and 3 the fundamental Weyl chamber was denoted $\mathcal{C}$. In this chapter we let $\mathcal{C}$ denote an arbitrary chamber and reserve the notation $\mathcal{F}$ for the specific choice of fundamental chamber. We trust that the reader will not confuse the two.

correspond to solutions of the equations of motion of the form

$$F_E = f_E(t)\,\omega_p \wedge dt, \qquad F_M = f_B(t)\,\omega_{p+1}, \tag{4.3.2}$$

where ω_p and ω_{p+1} are representatives of the de Rham cohomology classes $\mathcal{H}^p(X)$ and $\mathcal{H}^{p+1}(X)$, respectively. When X is compact, solutions that yield electric and magnetic billiard walls can therefore only be found when the de Rham cohomology classes are nontrivial.

We may now state the influence of the topology of X on the billiard structure simply in terms of a *selection rule*. This rule makes use of the Betti numbers $b_j(X)$, which are the dimensions of the de Rham cohomology classes $\mathcal{H}^p(X)$. The selection rule reads as follows [109, 110]:

- **Selection Rule**: *When $b_s(X) = 0$ for some s, we remove all billiard walls corresponding to electric s–forms, or magnetic $(s-1)$–forms.*

The selection rule is established using the same assumptions as the noncompact BKL analysis, namely that we are in a regime where classical gravity is valid, and studying a sufficiently "generic" spacetime solution.

It has been known for some time that the algebraic structure of the billiard is invariant under Kaluza–Klein reduction on tori [116]. The selection rule is compatible with these results, since tori have no vanishing Betti numbers. Note also that none of the symmetry walls (or gravity walls) is eliminated by the compactification. Hence, among the dominant walls of the compactified theory we always have the $(d-1)$ dominant symmetry walls $\beta^2 - \beta^1$, ..., $\beta^d - \beta^{d-1}$.

4.4 Gram Matrices and Coxeter Groups

In order to understand the structure of the reflection group that emerges when some p-form walls are switched off as well as the features of the corresponding billiard domain, it is useful to recall a few facts about polyhedra and reflection groups in hyperbolic space. The main reference for this section is [117].

4.4.1 Convex Polyhedra

We shall consider convex polyhedra P of hyperbolic space, i.e., regions of the form

$$P = \bigcap_{s=1}^{N} H_s^+, \tag{4.4.1}$$

where H_s^+ is a half-space bounded by the hyperplane H_s, and N is the number of such bounding hyperplanes. In our case, H_s is one of the walls of the relevant dominant wall system,

$$H_s = \{\beta \in \mathcal{M}_\beta \mid \omega_s(\beta) = 0\}, \tag{4.4.2}$$

and H_s^+ is defined by

$$H_s^+ = \{\beta \in \mathcal{M}_\beta \mid \omega_s(\beta) \geq 0\}. \tag{4.4.3}$$

The polyhedra P therefore contain the fundamental domain (4.2.1), defined by the dominant walls of the uncompactified theory. Hence it is clear that P has non-vanishing volume.

4.4.2 Relative Positions of Walls in Hyperbolic Space

It is customary to associate with the convex polyhedron P a matrix $G(P)$, the so-called *Gram matrix*, which differs from the matrix A by normalization. The construction proceeds as follows. To each hyperplane H_s we associate a unit spacelike vector e_s pointing inside P, i.e., pointing towards the billiard region (which is thus defined by $(\beta, e_s) \geq 0$). We then construct the $N \times N$ matrix $G(P)$ of scalar products $(e_s|e_{s'})$. Four cases can occur for the scalar product $(e_s|e_{s'})$ between a given pair of distinct vectors e_s and $e_{s'}$ [117][2]:

i. $-1 \leq (e_s|e_{s'}) \leq 0$. In this case, the hyperplanes H_s and $H_{s'}$ intersect and form an acute angle. The limiting case $(e_s|e_{s'}) = -1$ means that the hyperplanes intersect at infinity, i.e., are parallel. The other limiting case $(e_s|e_{s'}) = 0$ means that the hyperplanes form a right angle, which is both acute and obtuse.

ii. $0 \leq (e_s|e_{s'}) \leq 1$. In this case, the hyperplanes also intersect, but form an obtuse angle. The limiting case $(e_s|e_{s'}) = 1$ corresponds again to parallel hyperplanes meeting at infinity.

iii. $(e_s|e_{s'}) < -1$.

iv. $(e_s|e_{s'}) > \quad 1$.

In the latter two cases, the hyperplanes diverge, i.e., do not meet even at infinity. The difference between these two cases is that while the condition $H_s^+ \cap H_{s'}^+$ defines a non-empty region when $(e_s|e_{s'}) \leq -1$, one has $H_s^+ \cap H_{s'}^+ = \emptyset$ whenever $(e_s|e_{s'}) \geq 1$. The fourth case is thus excluded from our analysis, as is the limiting case $(e_s|e_{s'}) = 1$ $(s \neq s')$. We denote also the dihedral angle between the hyperplanes H_s and $H_{s'}$ by $H_s^+ \cap H_{s'}^+$. When the hyperplanes intersect, the value of the dihedral angle $H_s^+ \cap H_{s'}^+$ can be found from the relation

$$\cos(H_s^+ \cap H_{s'}^+) = -(e_s|e_{s'}). \tag{4.4.4}$$

4.4.3 Acute-Angled Polyhedra

If the number of dominant walls is strictly smaller than the dimension M of $\mathcal{M}_\beta$, the billiard table in $\mathbb{H}_m$ has infinite volume and the motion is non-chaotic. After a finite number of collisions, the billiard ball escapes to infinity [49, 118]. We shall therefore assume that the number of dominant walls is greater than or equal to the dimension M of $\mathcal{M}_\beta$, a case that needs a more detailed analysis. The Gram matrix is then of rank M because among the S dominant wall forms, one can find a subset of M of them that defines a basis, namely the $(d-1)$ symmetry walls and one of the other dominant walls if there is no dilaton (or two linearly independent dominant non-symmetry walls if there is a dilaton etc.). The convex polyhedron defined by the dominant walls is therefore non-degenerate (see [117]). We shall also assume that the Gram matrix is indecomposable (cannot be written as a direct sum upon reordering of the e_i's), as the decomposable case can be analysed in terms of the indecomposable one.

[2]Note that one cannot have $e_s = -e_{s'}$ since then the region $H_s^+ \cap H_{s'}^+ = H_s$ has vanishing volume, which is excluded as we observed above.

If the number of dominant walls is exactly equal to M, the billiard table is a simplex in hyperbolic space. Otherwise, one has a non-simplex billiard table, with the number of faces exceeding $\dim \mathbb{H}_n + 1 = n + 1$.

A crucial notion in the study of reflection groups is that of *acute-angled polyhedra*. A convex polyhedron is said to be *acute-angled* if for any pair of distinct hyperplanes defining it, either the hyperplanes do not intersect, or, if they do, the dihedral angle $H_s^+ \cap H_{s'}^+$ does not exceed $\frac{\pi}{2}$. The Gram matrix, which has 1's on the diagonal, has then negative entries off the diagonal. While non-degenerate, indecomposable acute-angled polyhedra on the sphere or on the plane are necessarily simplices, this is not the case on hyperbolic space.

4.4.4 Coxeter Polyhedra and the Billiard Group

We have seen that the dynamics of gravity is described in all cases (uncompactified or compactified) by the motion of a billiard ball in a region of hyperbolic space. The reflections s_s $(s = 1, \ldots, N)$ with respect to the billiard walls generate a discrete reflection group which we want to characterize. This group, which we shall call the *billiard group* and denote by $\mathfrak{B}$, is a subgroup of the Coxeter group relevant to the uncompactified case, where the total number of walls is maximum and the billiard region the smallest (and contained in all other billiard regions). The billiard group is a crystallographic Coxeter group since it is generated by reflections [117] and since it preserves the root lattice of the Kac-Moody algebra of the uncompactified case. Its presentation in terms of the billiard walls might, however, be non-standard. The billiard group $\mathfrak{B}$ will be examined more carefully in Sections 4.5.2 and 4.5.2.

The billiard table has an important property which it inherits from the complete wall system of the theory. Consider the dihedral angle $H_i^+ \cap H_j^+$ between two different walls H_i and H_j that intersect. If this angle is acute, then it is an integer submultiple of π, i.e., of the form π/m_{ij} where m_{ij} is an integer greater than or equal to 2. If this angle is obtuse, then $\pi - H_i^+ \cap H_j^+$ is an integer submultiple of π, i.e., $\pi - H_i^+ \cap H_j^+ = \pi/m_{ij}$, where $m_{ij} \in \mathbb{Z}_{\geq 2}$. If the walls do not intersect, they are parallel and one has $m_{ij} = \infty$. Thus, given any pair of distinct walls, one can associate to it an integer $m_{ij} = m_{ji} \geq 2$.

In the case when all the angles are acute, and hence integer submultiples of π, the polyhedron is called a *Coxeter polyhedron*. Coxeter polyhedra are thus acute-angled. In hyperbolic space, they may or may not be simplices.

4.5 General Results

In this section we describe our new results concerning general features of the billiard structures after compactification. In Section 4.5.1, we use features of the wall system to explain why the billiard table need not be a Coxeter polyhedron after compactification. In Section 4.5.3 we describe the billiard region after compactification in terms of galleries, which we explain. Finally, we describe in Section 4.5.4 our methods for determining the chaotic properties for all possible compactifications.

4.5.1 Rules of the Game

We show why the billiard region need not be a Coxeter polyhedron after compactification with the aid of two facts about billiard wall systems:

- **Fact 1**: *If both an electric and a magnetic wall are present for a given p–form, then the gravitational walls are subdominant.* This was proven in [105,116], by noticing that for any p we have
$$G_{ijk} = e^{[p]}_{r_1\cdots r_p} + m^{[p]}_{s_1\cdots s_{d-p-1}}, \tag{4.5.1}$$
where one of the r_q and one of the s_q are equal to i, and neither j nor k appears in either the r_q or s_q.

- **Fact 2**: *The inner product between a gravitational wall and a p–form wall is unity for*

  – *electric walls with* $p \leq D-3$,
  – *magnetic walls with* $p \geq 1$,

  and the inner product vanishes for

  – *electric walls with* $p = D-2$,
  – *magnetic walls with* $p = 0$.

  Typically, p forms with $p > \lfloor D/2 \rfloor - 1$ are dualised so that we may safely assume that $p \leq \lfloor D/2 \rfloor - 1$. (This is a stronger condition than $p \leq D-3$ for $D \geq 4$). Then, we have that the inner product between gravitational and p–form walls is always positive, except when the p–form is a scalar (axion) and the inner product vanishes. This fact is proven by computing the relevant inner products (which are independent of the dilaton coupling(s) of the p–form) using the metric between wall forms.

  This fact is significant in combination with the requirement that a system of dominant walls define an acute-angled polyhedron if the inner products between each pair of walls are either zero or negative. Therefore we can conclude:

  If the dominant wall set includes a gravitational wall and any non-scalar p–form wall, then the dominant wall system does not define a Coxeter polyhedron.

For compactifications on tori, only **Fact 1** is relevant since all components of the same p-forms are preserved in the compactification [116]. Hence, electric and magnetic walls are always present in pairs, ensuring that the gravity wall is always subdominant. More general compactifications can eliminate one of the electric or magnetic walls of a given p–form, while leaving the other intact. Unlike the noncompact case, it is possible for both gravitational and p–form walls to appear simultaneously in the set of dominant walls. **Fact 2** tells us that when this happens, we no longer have a Coxeter polyhedron.

In our analysis, we never eliminate gravitational walls (although there are some partial results regarding their selection rules that were given in [109]). This means that after compactification there is always a gravitational wall in the root system, though it is not necessarily dominant. Therefore all compactifications we study fall into one of three classes:

- A pair of electric and magnetic walls from the noncompact theory has not been eliminated by compactification, so the gravitational wall is subdominant. The resulting billiard table may be a Coxeter polyhedron, depending on the details of the p-form menu and couplings in the theory.

- One (or both) members of each pair of electric/magnetic walls are eliminated, so the gravitational wall is exposed. If there are any other p-form walls left, because of **Fact 2**, the billiard table cannot be a Coxeter polyhedron. However, it may occur that "coincidentally" two walls from different p-forms can combine and make the gravitational wall subdominant.

- All of the p-form walls are eliminated by compactification, and so only the gravitational wall is left. In this case one always obtains A_n^{++}, the billiard corresponding to pure gravity. This billiard sits at the bottom of every list when all possible Betti numbers are set to zero. Occasionally there are also direct summands corresponding to scalar fields that are never eliminated by compactification.

For examples of these three cases we refer the reader to **Paper V**. Only the first of the cases described above arises in the noncompact theory, where gravitational walls are never dominant, except when they are the only non-symmetry walls.

4.5.2 Describing the Billiard Group After Compactification

We shall now begin to assemble the various structures described so far in order to analyze the group-theoretical properties of the billiard dynamics after compactification. This involves understanding the relation between the billiard table and the fundamental domain of the associated reflection group. In this context it is important to distinguish between the formal Coxeter group and the billiard group. We consider these concepts in turn in the following two sections, and explain how they are related in Section 4.5.2.

The Formal Coxeter Group

Recall from Section 4.4.4 that the reflections s_i with respect to the dominant walls H_i generate a Coxeter group. To describe this group we must characterize the relations among the reflections s_i. Being reflections, they clearly satisfy

$$s_i^2 = 1. \tag{4.5.2}$$

Consider next the reflections s_i and s_j with respect to two different hyperplanes H_i and H_j. Then, the product $s_i s_j$ is a rotation by the angle $\frac{2\pi}{m_{ij}}$ (where the integers m_{ij} were introduced in Section 4.4.4) and so

$$(s_i s_j)^{m_{ij}} = 1. \tag{4.5.3}$$

These relations alone define a Coxeter group, which we shall call the *formal Coxeter group* $\mathfrak{C}$ associated with the billiard. One can describe $\mathfrak{C}$ more precisely as follows. Let $\tilde{\mathfrak{C}}$ be the group freely generated by the elements of the set $\mathcal{S} = \{r_1, \ldots, r_N\}$, and let $\mathfrak{N}$ be the normal subgroup generated by $(r_i r_j)^{m_{ij}}$, where the *Coxeter exponents* m_{ij} satisfy [34,73] (see also [94])

$$\begin{aligned} m_{ii} &= 1, \\ m_{ij} &\in \mathbb{Z}_{\geq 2},\ i \neq j, \\ m_{ij} &= m_{ji}. \end{aligned} \tag{4.5.4}$$

Then $\mathfrak{C}$ is defined as the quotient group $\tilde{\mathfrak{C}}/\mathfrak{N}$ and has the following standard presentation:

$$\mathfrak{C} = \langle r_1, \ldots, r_n \mid (r_i r_j)^{m_{ij}} = 1,\ i, j = 1, \ldots, N \rangle\,. \tag{4.5.5}$$

One can associate a Coxeter graph with the formal Coxeter group, i.e., with the set of m_{ij}'s. Each r_i defines a node of the Coxeter graph, and two different nodes i and j are connected by a line whenever $m_{ij} > 2$, with m_{ij} explicitly written over the line whenever $m_{ij} > 3$ [73]. To a Coxeter graph, one can further associate a symmetric matrix defined by

$$B_{ij} = -\cos\left(\frac{\pi}{m_{ij}}\right) \tag{4.5.6}$$

with 1's on the diagonal and non-positive elements otherwise.

The Billiard Group and its Fundamental Domain

In Section 4.4.4 we introduced the concept of the *billiard group* $\mathfrak{B}$. Here we shall elaborate on this object and elucidate the structure of its fundamental domain.

The billiard group $\mathfrak{B}$ is defined as the group generated by the reflections s_i with respect to the dominant walls of the billiard table. This group coincides with the formal Coxeter group $\mathfrak{C}$ (with $r_i \equiv s_i$) if and only if there are no additional relations among the s_i's. Two cases must be considered.

i. **Acute-angled billiard tables.** There is no additional relations among the s_i's if and only if the billiard table is acute-angled, i.e., is a Coxeter polyhedron [117]. In that case, the matrix $B = (B_{ij})$ coincides with the Gram matrix $G(P)$. Furthermore, the billiard table is a fundamental domain [117].

   In the hyperbolic case relevant here, the billiard table need not be a simplex. When it is a simplex, however, there is further structure. The Gram matrix is non-degenerate and the action of the Coxeter group on the β-space $\mathcal{M}_\beta$ coincides with the standard geometric realization considered in [73]. In addition, the matrix

$$A_{ss'} = \frac{2(\omega_s|\omega_{s'})}{(\omega_s|\omega_s)}, \tag{4.5.7}$$

   has non-positive integers off the diagonal and hence is a non-degenerate Cartan matrix. It is obvious that the off-diagonal entries are integers since the walls correspond to some roots of the Kac-Moody algebra of the uncompactified case. Moreover, it follows from the fact that the billiard table is acute-angled that they are negative. The billiard group is then the Weyl group of a simple Kac-Moody algebra.[3] When the matrix A is a Cartan matrix, one can associate to it a Dynkin diagram.

ii. **Non acute-angled billiard tables.** If the billiard table is not acute-angled (and hence not a Coxeter polyhedron), there are further relations among the s_i's and the billiard group is therefore the quotient of the formal Coxeter group $\mathfrak{C}$ by these additional relations. Moreover, the billiard table is not a fundamental domain of the billiard group. One may understand this feature as follows. Consider a dihedral angle $H_i^+ \cap H_j^+$ of the polyhedron that is obtuse. The rotation $s_i s_j$ by the angle $2\pi/m_{ij}$ is an element of the group and hence maps reflection hyperplanes to reflection hyperplanes. The image by this rotation of H_j is the hyperplane $s_i H_j$ that intersects the interior of the billiard table.[4] The reflection $s_i s_j s_i$ with respect to this hyperplane belongs to the group, and hence the billiard table cannot be a fundamental region of the billiard group since the orbit of a point sufficiently close to $s_i H_j$ intersects the billiard table at least twice. Fundamental regions are obtained by considering all the mirrors (reflection hyperplanes) associated with the group (most of which are not billiard walls), which decompose the space into equivalent chambers that are permuted by the group (homogeneous decomposition). Each of these chambers is a fundamental domain. The billiard group $\mathfrak{B}$ is generated by the reflections in the mirrors of the fundamental domain, which provide a standard presentation of the group, and the billiard table is a union of chambers [117]. Examples of the occurrence of this phenomenon will be discussed below.

   Although the billiard table is not a fundamental domain, it can be naturally described as a gallery defined by the Coxeter group of the uncompactified theory. This is described in Section 4.5.3.

Non-Standard Presentations of Coxeter Groups

We have seen how one can associate a formal Coxeter group $\mathfrak{C}$ to the billiard region using the Coxeter presentation $\mathfrak{C} = \tilde{\mathfrak{C}}/\mathfrak{N}$. The billiard group $\mathfrak{B}$ – which is also a Coxeter group – differs from $\mathfrak{C}$ when the billiard table possesses obtuse angles because the reflections in the walls of the billiard then fulfill additional relations. This yields a non-standard presentation of the billiard group, which can be formally described as follows.

Let $\mathcal{B}_\beta$ be the billiard region after compactification, let the elements of the set $\mathcal{S} = \{s_i \mid i = 1, \ldots, N\}$ be the reflections in the walls W_i bounding $\mathcal{B}_\beta$, and let $\tilde{\mathfrak{C}}$ be the formal group freely generated by $\mathcal{S}$. The dihedral angles between the W_i give rise to a set of Coxeter exponents m_{ij}, with associated normal subgroup $\mathfrak{N} \subset \tilde{\mathfrak{C}}$. Suppose now that the region $\mathcal{B}_\beta$ is not a fundamental domain of $\mathfrak{B}$, and denote by $\mathfrak{J}$ the normal subgroup of $\tilde{\mathfrak{C}}$ generated by $(s_i s_j)^{m_{ij}}$ and any other non-standard relations between the elements of $\mathcal{S}$. Note that we have $\mathfrak{N} \subset \mathfrak{J}$. The billiard group $\mathfrak{B}$ is then the quotient

$$\mathfrak{B} = \tilde{\mathfrak{C}}/\mathfrak{J}. \tag{4.5.8}$$

[3] In the case of a non-simplex, acute-angled billiard table, the matrix A is also a valid Cartan matrix, but it is degenerate and the corresponding Kac-Moody algebra is not simple. Furthermore, the standard geometric realization is defined in a space of dimension larger than M, while the billiard realization is defined in the M-dimensional β-space.

[4] When the Coxeter group is crystallographic, the converse is also true: if the angle between H_i and H_j is acute, then the image $s_i H_j$ does *not* intersect the interior of the billiard table.

Equivalently, if we denote by $\mathfrak{F}$ the normal subgroup of $\tilde{\mathfrak{C}}$ generated only by the non-standard relations between the elements of $\mathcal{S}$, we may describe the billiard group as

$$\mathfrak{B} = \mathfrak{C}/\mathfrak{F}. \tag{4.5.9}$$

Neither of these presentations is a Coxeter presentation.

In all cases we consider in this analysis, the uncompactified billiard is described by the Weyl group $\mathcal{W}[\mathfrak{u}^{++}]$ of a Lorentzian Kac-Moody algebra $\mathfrak{u}^{++}$. For general compactifications, the billiard group $\mathfrak{B}$ is therefore a Coxeter subgroup of $\mathcal{W}[\mathfrak{u}^{++}]$. However, in **Paper V** examples were found of cases when the formal Coxeter group $\mathfrak{C}$ after compactification differs from $\mathfrak{u}^{++}$, while the billiard group $\mathfrak{B}$ is actually isomorphic $\mathfrak{u}^{++}$, with a non-standard presentation. We refer to **Paper V** for more details.

4.5.3 Fundamental Domains, Chamber Complexes and Galleries

We have seen that the billiard table need not be a Coxeter polyhedron upon compactification. When it is not a Coxeter polyhedron, it no longer corresponds a fundamental domain of the billiard group $\mathfrak{B}$. Moreover, $\mathfrak{B}$ is the quotient by nontrivial extra relations of the formal Coxeter group associated with the billiard table. In this section, we describe how the billiard region relevant to the compactified case can then be built as a union of images of the uncompactified billiard region. This is achieved using the theory of buildings, in terms of chambers and galleries.

Chambers and Galleries

The analysis in this section makes use of the treatment of Coxeter groups as a theory of *buildings*, a formalism mainly developed by J. Tits. Introductions and references may be found in [119, 120].

The basic idea is to study Coxeter groups in geometric language by defining them in terms of the objects on which they act nicely. The *buildings* are the fundamental objects which then defines the associated Coxeter group. For example, finite Coxeter groups act on so-called *spherical buildings*, because these groups preserve the unit sphere. We are interested in Coxeter groups which act on *hyperbolic buildings*, i.e., hyperbolic Coxeter groups, which preserve the hyperbolic space.

An *n-simplex* $\mathcal{X}$ in hyperbolic space is determined by its $n+1$ vertices. A 1-simplex is determined by its two endpoints, a 2-simplex (a triangle) is determined by its three vertices etc. For this reason, it is convenient to identify an n-simplex $\mathcal{X}$ with the set of its $n+1$ vertices, $\mathcal{X} = \{\text{set of } n+1 \text{ vertices}\}$. A *face* f of $\mathcal{X}$ is a simplex corresponding to a non-empty subset $f \subset \mathcal{X}$. The *codimension* of f with respect to $\mathcal{X}$ is given by $\dim \mathcal{X} - \dim f$. For example, there are three codimension-one faces in a 2-simplex, which are the three edges of the triangle.

Next we define the notion of a simplicial complex. Let V be a set of vertices, and $\mathcal{K}$ a set of finite subsets $\mathcal{X}_k \subset V$. We assume that the subsets containing a single vertex of V are all elements of $\mathcal{K}$. Then $\mathcal{K}$ is called a *simplicial complex* if it is such that given $\mathcal{X} \in \mathcal{K}$ and a face f of $\mathcal{X}$, then $f \in \mathcal{K}$. The elements $\mathcal{X}_k$ of $\mathcal{K}$ are the simplices in the simplicial complex. Two simplices $\mathcal{X}_1$ and $\mathcal{X}_2$ of the same dimension in $\mathcal{K}$ are called *adjacent* if they share a codimension-one face, i.e., if they are separated by a common wall. Figure 4.1 illustrates a simplicial complex of 2-simplices.

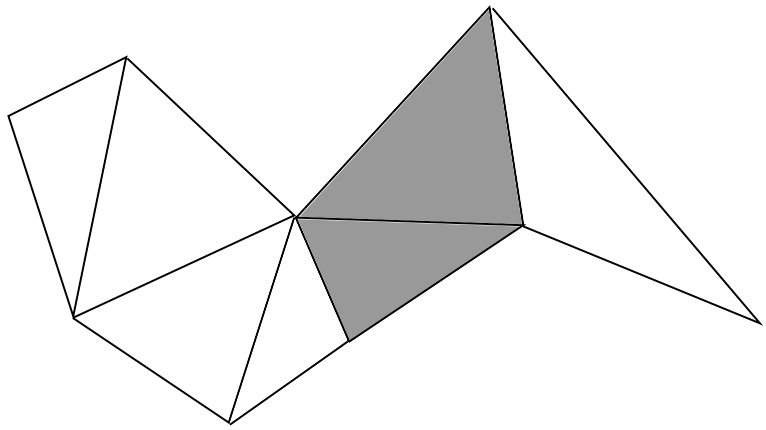


Figure 4.1: A simplicial complex of 2-simplices (triangles). The two shaded regions represent adjacent (maximal) simplices.

A *maximal simplex* $\mathcal{C}$ in $\mathcal{K}$ is such that $\mathcal{C}$ does not correspond to the face of another simplex in $\mathcal{K}$. Maximal simplices in a simplicial complex are called *chambers* and they shall be our main objects of study. A sequence of chambers $\mathcal{C}_1, \dots, \mathcal{C}_k$, such that any two consecutive chambers $\mathcal{C}_i$ and $\mathcal{C}_{i+1}$ are adjacent is called a *gallery* Γ. Thus, a gallery corresponds to a connected path between two chambers $\mathcal{C}_1$ and $\mathcal{C}_k$ in $\mathcal{K}$, and we write

$$\Gamma \ : \ \mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_{k-1}, \mathcal{C}_k. \tag{4.5.10}$$

The *length* of Γ is k, and the *distance* between $\mathcal{C}_1$ and $\mathcal{C}_k$ is the length of the shortest gallery connecting them. If any two chambers in $\mathcal{K}$ are connected by a gallery, then the simplicial complex is called a *chamber complex.* A simple example of a gallery inside a chamber complex is displayed in Figure 4.2.

The Billiard Region as Gallery

We describe the billiard regions of compactified gravity-dilaton-p-form theories in the language presented above. This is achieved by expressing them in terms of the billiard region of the uncompactified theory (or compactified on a torus) which we recall is the fundamental Weyl chamber $\mathcal{F}$ of the Weyl group $\mathcal{W}[\mathfrak{u}^{++}]$ of the Kac-Moody algebra $\mathfrak{u}^{++}$, whose Cartan matrix is defined through the scalar products between the dominant wall forms.

Compactification amounts to a process of removing dominant walls, and so the billiard table is enlarged. The inequalities $\omega_{A'} \geq 0$ associated with the simple roots of the underlying Kac-Moody algebra are indeed replaced by weaker inequalities. The bigger region so defined can be written as a union of Weyl chambers of $\mathcal{W}[\mathfrak{u}^{++}]$. We shall illustrate this phenomenon on the example of the familiar Lie algebra A_3, whose Weyl group is a spherical Coxeter group. The fundamental Weyl chamber $\mathcal{F}$ is defined by

$$\omega_1(\beta) \geq 0, \ \ \omega_2(\beta) \geq 0, \ \ \omega_3(\beta) \geq 0, \tag{4.5.11}$$

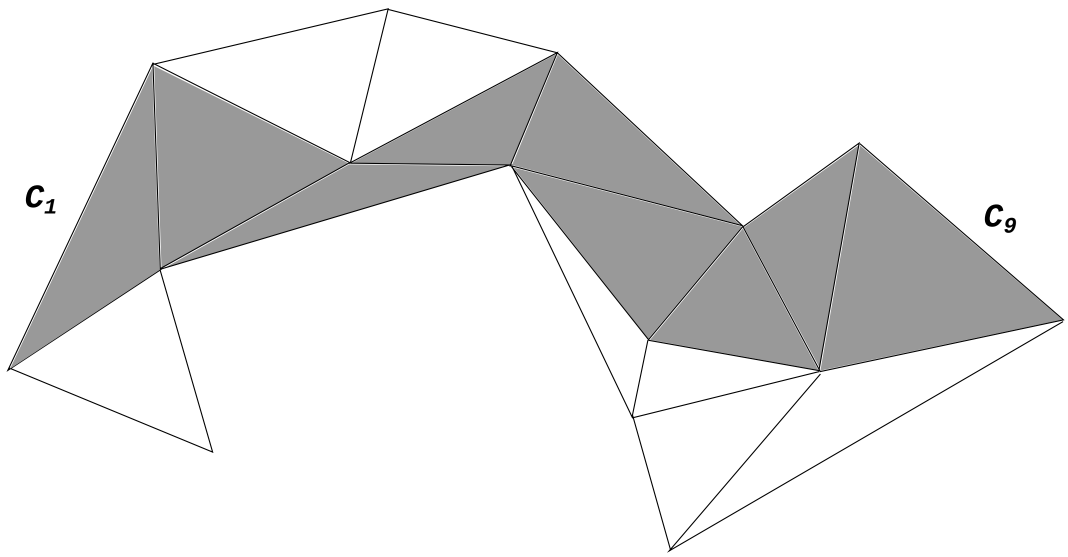


Figure 4.2: A chamber complex with a gallery $\Gamma : \mathcal{C}_1, \ldots, \mathcal{C}_9$, represented by the shaded region. The length of the gallery is $k = 9$, which is also the distance between $\mathcal{C}_1$ and $\mathcal{C}_9$ since Γ is the shortest gallery connecting $\mathcal{C}_1$ and $\mathcal{C}_9$.

corresponding to the three simple roots. The non simple roots are $\omega_1 + \omega_2$, $\omega_2 + \omega_3$ and $\omega_1 + \omega_2 + \omega_3$. Suppose that the single dominant wall W_1 defined by $\omega_1(\beta) = 0$ is suppressed. Effectively, this implies that the particle geodesic may cross the wall W_1. Thus it moves from the region where $\omega_1(\beta) \geq 0$ into the region where

$$\omega_1(\beta) \leq 0. \tag{4.5.12}$$

We shall consider two cases. (i) The new billiard region is defined by

$$\omega_1(\beta) + \omega_2(\beta) \geq 0, \;\; \omega_2(\beta) \geq 0, \;\; \omega_3(\beta) \geq 0, \tag{4.5.13}$$

i.e., the wall W_1 is replaced by the wall $\omega_1(\beta) + \omega_2(\beta) = 0$; (ii) In the second case, we suppose that also the wall defined by $\omega_1(\beta) + \omega_2(\beta) = 0$ is suppressed. Then the new billiard region is defined by the inequalities

$$\omega_1(\beta) + \omega_2(\beta) + \omega_3(\beta) \geq 0, \;\; \omega_2(\beta) \geq 0, \;\; \omega_3(\beta) \geq 0 \tag{4.5.14}$$

i.e., the wall W_1 is replaced by the wall $\omega_1(\beta) + \omega_2(\beta) + \omega_3(\beta) = 0$.

In the first case, one can write the new billiard region as the union $\mathcal{F} \cup \mathcal{A}_2$ where $\mathcal{A}_2$ is defined by the inequalities

$$\omega_1(\beta) \leq 0, \;\; \omega_1(\beta) + \omega_2(\beta) \geq 0, \;\; \omega_3(\beta) \geq 0 \tag{4.5.15}$$

(which imply $\omega_2(\beta) \geq 0$). The region $\mathcal{A}_2$ is the Weyl chamber obtained by reflecting the fundamental Weyl chamber across the wall W_1 since

$$\begin{aligned} s_1(\omega_1) &= -\omega_1, \\ s_1(\omega_2) &= \omega_1 + \omega_2, \\ s_1(\omega_3) &= \omega_3. \end{aligned} \tag{4.5.16}$$

Hence,

$$s_1 \cdot \mathcal{F} = \{\beta \in \mathfrak{h} \mid \omega_1(\beta) \leq 0, (\omega_1 + \omega_2)(\beta) \geq 0, \omega_3(\beta) \geq 0\} = \mathcal{A}_2. \tag{4.5.17}$$

A reflection of this type which maps a chamber $\mathcal{C}$ to an adjacent chamber $\mathcal{C}'$ is known as an *adjacency transformation* [117]. By removing the single dominant W_1, we therefore get in the first case a new region which is precisely twice as large as the original fundamental region.

In the second case, although we also remove a single dominant wall of the original billiard, we get a larger region. This is because we also remove the subdominant wall $\omega_1(\beta)+\omega_2(\beta) = 0$ which is exposed once W_1 is removed. Indeed, the new billiard region can now be written as the union $\mathcal{F} \cup \mathcal{A}_2 \cup \mathcal{A}_3$ where $\mathcal{A}_3$ is defined by the inequalities

$$\omega_2(\beta) \geq 0, \;\; \omega_1(\beta) + \omega_2(\beta) \leq 0, \;\; \omega_1(\beta) + \omega_2(\beta) + \omega_3(\beta) \geq 0 \tag{4.5.18}$$

(which imply $\omega_1(\beta) \leq 0$ and $\omega_3(\beta) \geq 0$). The region $\mathcal{A}_3$ is again a Weyl chamber, obtained from $\mathcal{A}_2$ by acting with the reflection s with respect to $\omega_1 + \omega_2$ since

$$\begin{aligned} s(-\omega_1) &= \omega_2 \\ s(\omega_1 + \omega_2) &= -(\omega_1 + \omega_2), \\ s(\omega_3) &= \omega_1 + \omega_2 + \omega_3. \end{aligned} \tag{4.5.19}$$

Hence,

$$s \cdot \mathcal{A}_2 = \{\beta \in \mathfrak{h} \mid \omega_2(\beta) \geq 0, (\omega_1 + \omega_2)(\beta) \leq 0, \omega_1(\beta) + \omega_2(\beta) + \omega_3(\beta) \geq 0\} = \mathcal{A}_3. \tag{4.5.20}$$

Note that $\mathcal{A}_2$ and $\mathcal{A}_3$ are adjacent. Thus, while in the second case we also remove a single dominant wall, we now obtain a region three times larger than the fundamental Weyl chamber. This new region can be described as a union of three pairwise adjacent Weyl chambers. In Figure 4.3 we describe pictorially a similar example for the case of the hyperbolic Coxeter group A_1^{++}.

By extrapolating this analysis to the general case where we remove an arbitrary number $r \leq n+1$ of dominant walls, we may conclude that the new billiard region corresponds to a union of images of the fundamental Weyl chamber. This naturally has the structure of a simplicial complex, and moreover, by a suitable ordering of the chambers, one sees that it corresponds to a gallery covering the whole complex. In conclusion, we have found the following: *the billiard region $\mathcal{B}$ obtained by compactification on a manifold of non-trivial topology is described by a gallery Γ inside the Cartan subalgebra $\mathfrak{h}$ of the original hyperbolic Kac-Moody algebra $\mathfrak{g}$.*

The dynamics after compactification is chaotic if the new billiard region is a finite union of images of the fundamental chamber, i.e., if the gallery Γ has finite length, while if this union is infinite the particle motion will eventually settle down in a single asymptotic Kasner solution, and chaos is removed. Since the Coxeter reflections preserve the volume, the volume of $\mathcal{B}$ is

$$\text{vol}\ \mathcal{B} = k \cdot \text{vol}\ \mathcal{F}, \tag{4.5.21}$$

where k is the length of the gallery Γ associated with $\mathcal{B}$.

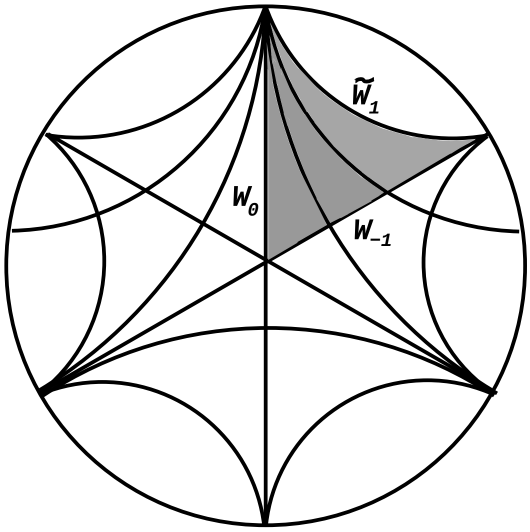


Figure 4.3: Here we illustrate a gallery $\Gamma\ :\ \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ of length 3 for the case of the hyperbolic Kac-Moody algebra A_1^{++}. The two walls W_0 and W_{-1} are associated with the affine and overextended simple roots α_0 and α_{-1}, respectively. The original fundamental Weyl chamber $\mathcal{C}_1 = \{h \in \mathfrak{h} \mid \alpha_1(h) \geq 0,\ \alpha_0(h) \geq 0,\ \alpha_{-1}(h) \geq 0\}$ corresponds to the leftmost shaded region. We have removed the wall $W_1 = \{h \in \mathfrak{h} \mid \alpha_1(h) = 0\}$ as well as the wall $(\alpha_0 + \alpha_1)(h) = 0$. The far end of the billiard region is now bounded by the new wall $\tilde{W}_1 = \{h \in \mathfrak{h} \mid (2\alpha_0 + 3\alpha_1)(h) = 0\}$. Each of the three chambers is clearly a copy of the fundamental chamber, and the total region is of finite volume. See, e.g., [94, 121] for more detailed discussions of the Weyl group of A_1^{++}.

4.5.4 Determining the Chaotic Properties After Compactification

The selection rules described in Section 4.5.1 provide a straightforward means to determine the billiard system after compactification. Determining whether or not this billiard system is chaotic, i.e., computing the biliard table volume, is somewhat more involved because finding explicitly the corresponding galleries might be intricate. In most cases it is possible to answer this question purely analytically without working out the gallery, although there are several different techniques that work for different types of billiard system. In this section we describe the various methods we employ.

The simplest case is when the billiard table is a Coxeter simplex. The matrix $\bar{A}_{ab}$ is then a Cartan matrix. The associated Kac-Moody algebra $\mathfrak{g}(\bar{A})$ is by construction a regular subalgebra[5] of the Kac-Moody algebra $\mathfrak{g}(A)$, whose Weyl group controlled the uncompactified billiard. The dynamics of the compactified billiard is described by the Weyl group $\bar{\mathcal{W}}$ of $\mathfrak{g}(\bar{A})$, and the billiard region $\bar{\mathcal{B}}$ coincides with the fundamental domain $\bar{\mathcal{F}}$ of $\bar{\mathcal{W}}$. Thus, if $\mathfrak{g}(\bar{A})$ is

[5]A subalgebra $\bar{\mathfrak{g}} \subset \mathfrak{g}$ is *regularly embedded* in $\mathfrak{g}$ if and only if two conditions are fulfilled: (i) the root vectors of $\bar{\mathfrak{g}}$ are root vectors of $\mathfrak{g}$; and (ii) the simple roots of $\bar{\mathfrak{g}}$ are real roots of $\mathfrak{g}$. Moreover, the embedding is *positive regular* if the positive root vectors of $\bar{\mathfrak{g}}$ are positive root vectors of $\mathfrak{g}$. See, e.g., [121, 122] for more detailed discussions on regular subalgebras of Kac-Moody algebras.

hyperbolic, then $\bar{\mathcal{B}}$ is of finite volume, yielding chaotic dynamics. If the Kac-Moody algebra $\mathfrak{g}(\bar{A})$ is Lorentzian but not hyperbolic, then the billiard is non-chaotic. This is illustrated in Figure 4.4.

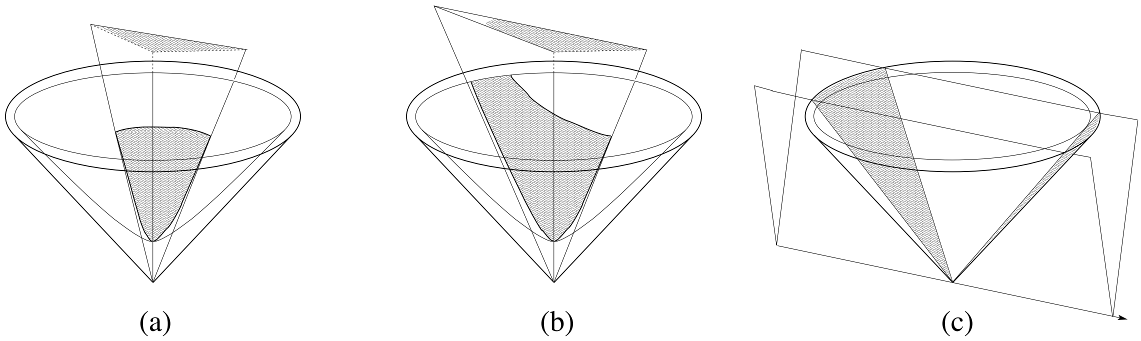


Figure 4.4: Some examples of the wall systems and their chaotic properties: (**a**) the wall system corresponding to a hyperbolic Kac-Moody algebra, (**b**) the wall system of a non-hyperbolic Kac-Moody algebra, or one for which the coweight construction is possible, (**c**) a wall system with fewer walls than the dimension of the β-space.

In many cases the billiard table is a simplex, but some dihedral angles are obtuse (positive inner product between two different walls) and the matrix A_{ab} is not a proper Cartan matrix. It is however non-degenerate, so that it is possible to define a set of dominant "coweights" $\Lambda^{A'\mu}$ such that

$$\omega_{A'\mu}\Lambda^{B'\mu} = \delta^{B'}{}_{A'}, \tag{4.5.22}$$

where ω is a dominant wall labelled by A'. As in the standard Kac-Moody algebra case, these "coweights" span the space of rays that lie within the wall cone, provided we only combine "coweights" using non-negative coefficients. A non-chaotic solution to the equations of motion, which corresponds to a null ray within the wall cone, exists if and only if there is at least one timelike and one spacelike "coweight". This condition is readily checked once the "coweights" are in hand. This technique was employed in [110].

When the billiard table is not a simplex and the number of walls is less than the dimension of $\mathcal{M}_\beta$, then the theory is not chaotic, essentially because there are too few walls to prevent a ray from reaching infinity.

When the billiard table is not a simplex and the number of walls is greater than the dimension of $\mathcal{M}_\beta$, the analysis becomes more complex. This situation is illustrated in Figure 4.5. One method to determine whether chaos is present is to successively remove dominant walls until the billiard region is again a simplex. If there is (at least) one way to do this such that the resulting structure is hyperbolic, then we can conclude that the full region is of finite volume, since reinserting the walls that were removed can never render the volume infinite. In a small number of cases, all wall removals lead to non-chaotic billiards and one cannot conclude immediately whether or not the volume of the billiard table is finite. Another method is then to search numerically for whether a spacelike direction in the wall cone exists. We do this by maximising the Lorentzian norm of points on the unit sphere that lie within the wall cone. If the maximal norm is negative, then no spacelike direction exists and the system is chaotic.

Also, as we have described in Section 4.5.3 it is sometimes possible to compute the volume of the billiard region exactly by making use of certain properties of the Weyl group $\mathcal{W}[\mathfrak{g}]$, associated with the uncompactified theory and construct the associated gallery.

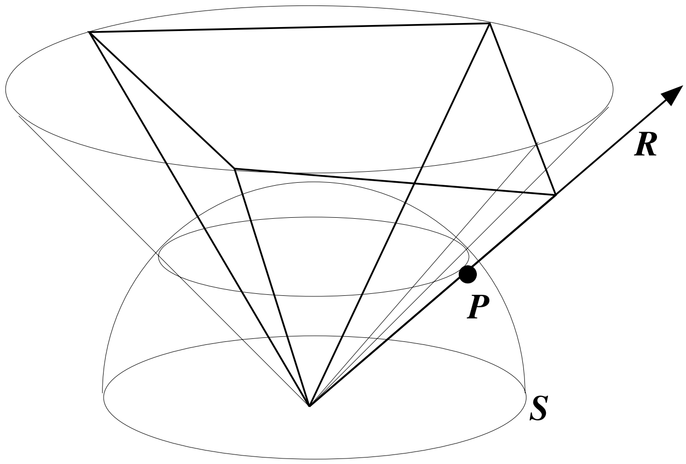


Figure 4.5: The figure illustrates a non-simplex billiard table. To determine whether the theory is chaotic, it is sufficient to locate a spacelike ray. This is equivalent to maximising the Lorentzian norm on the unit sphere surrounding the origin.

As a byproduct of the "coweight construction" mentioned above, we can easily prove the following useful fact:

- **Fact 3:** *Whenever the billiard is described by a direct product $\mathfrak{B}_{fin} \times \mathfrak{B}_{hyp}$ of a finite and a hyperbolic Coxeter group, then the dynamics is non-chaotic.*

This follows by noting that the metric is a direct sum of the metric of $\mathfrak{B}_{fin}$ and that of $\mathfrak{B}_{hyp}$ so the coweights associated with each factor define orthogonal subspaces. Since the coweights associated with the finite factor are spacelike, there will always exist at least one spacelike intersection in the region inside the dominant walls. A different intuitive way to see this is to recall that the volume of the fundamental Weyl chamber of the first factor is finite after projection on the sphere and that of the second factor after projection on the hyperboloid. Two projections are needed to have a finite volume but there is only one here (on a hyperboloid living in the product space).

5

Beyond the Weyl Group – Infinite Symmetries Made Manifest

We have learned that there is an intimate connection between gravitational theories and hyperbolic Kac-Moody algebras, which is revealed when studying the theory close to a space-like singularity. It is then natural to speculate whether this indicates the existence of a huge hidden hyperbolic symmetry of gravity. The aim of this chapter is to describe one line of research intended to put these speculations to the test. This approach is directly inspired by the results obtained through toroidal dimensional reduction of gravitational theories, where the scalar fields form coset manifolds exhibiting explicitly larger and larger symmetries as one goes down in dimensions. In the case of eleven-dimensional supergravity, reduction on an n-torus T^n reveals a chain of exceptional U-duality symmetries $\mathcal{E}_n$ [9, 10], culminating with $\mathcal{E}_8$ in three dimensions [123]. This has lead to the conjecture [32] that the chain of enhanced symmetries should in fact remain unbroken and give rise to the infinite-dimensional duality groups $\mathcal{E}_9, \mathcal{E}_{10}$ and $\mathcal{E}_{11}$, as one reduces the theory to two, one and zero dimensions, respectively.

The connection between the symmetry groups controlling the billiards in the BKL-limit, and the symmetry groups appearing in toroidal dimensional reduction to three dimensions has led to the attempt to reformulate eleven-dimensional supergravity as a one-dimensional nonlinear sigma model with target space given by the infinite-dimensional coset space $\mathcal{K}(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ [53]. This sigma model describes the geodesic flow of a particle moving on $\mathcal{K}(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$, whose dynamics can be seen to match the dynamics of the associated (suitably truncated) supergravity theory.

We begin by describing some general aspects of nonlinear sigma models for finite-dimensional coset spaces. We then explain how to generalize the construction to the infinite-dimensional case. We finally apply the construction in detail to the case of eleven-dimensional supergravity where the conjectured symmetry group is $\mathcal{E}_{10}$.

This chapter is based on **Paper III**, written in collaboration with Marc Henneaux and Philippe Spindel.

5.1 Nonlinear Sigma Models on Finite-Dimensional Coset Spaces

A nonlinear sigma model describes maps ξ from one Riemannian space X, equipped with a metric γ, to another Riemannian space, the "target space" M, with metric g. Let x^m $(m = 1, \cdots, p = \dim X)$ be coordinates on X and ξ^α $(\alpha = 1, \cdots, q = \dim M)$ be coordinates on M. Then the standard action for this sigma model is

$$S = \int_X d^p x \sqrt{\gamma}\, \gamma^{mn}(x)\, \partial_m \xi^\alpha(x)\, \partial_n \xi^\beta(x)\, g_{\alpha\beta}\left(\xi(x)\right). \tag{5.1.1}$$

Solutions to the equations of motion resulting from this action will describe the maps ξ^α as functions of x^m.

A familiar example, of direct interest to the analysis below, is the case where X is one-dimensional, parametrized by the coordinate t. Then the action for the sigma model reduces to

$$S_{\text{geodesic}} = \int dt\, A\, \frac{d\xi^\alpha(t)}{dt} \frac{d\xi^\beta(t)}{dt} g_{\alpha\beta}\left(\xi(t)\right), \tag{5.1.2}$$

where A is $\gamma^{11}\sqrt{\gamma}$ and ensures reparametrization invariance in the variable t. Extremization with respect to A enforces the constraint

$$\frac{d\xi^\alpha(t)}{dt} \frac{d\xi^\beta(t)}{dt} g_{\alpha\beta}\left(\xi(t)\right) = 0, \tag{5.1.3}$$

ensuring that solutions to this model are null geodesics on M. We have already encountered such a sigma model before, namely as describing the free lightlike motion of the billiard ball in the $(\dim M - 1)$-dimensional scale-factor space. In that case A corresponds to the inverse "lapse-function" N^{-1} and the metric $g_{\alpha\beta}$ is a constant Lorentzian metric.

5.1.1 The Cartan Involution and Symmetric Spaces

In what follows, we shall be concerned with sigma models on symmetric spaces $K\backslash G$ where G is a Lie group with semi-simple real Lie algebra $\mathfrak{g}$ in its split real form and $K = K(G)$ its maximal compact subgroup with real Lie algebra $\mathfrak{k}$ (sometimes also denoted $K(\mathfrak{g})$) corresponding to the maximal compact subalgebra of $\mathfrak{g}$. Since elements of the coset are obtained by factoring out K, this subgroup is referred to as the "local gauge symmetry group" (see below). Our aim is to provide an algebraic construction of the metric on the coset space $K\backslash G$ and explain how to obtain the associated sigma model Lagrangian.

We recall from Section 7.2.1 that for any maximally split real Lie algebra $\mathfrak{g}$, the Chevalley involution ω induces a Cartan decomposition of $\mathfrak{g}$ into even and odd eigenspaces:

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p} \tag{5.1.4}$$

(direct sum of vector spaces), where

$$\begin{aligned} \mathfrak{k} &= \{x \in \mathfrak{g} \,|\, \omega(x) = x\}, \\ \mathfrak{p} &= \{y \in \mathfrak{g} \,|\, \omega(y) = -y\} \end{aligned} \tag{5.1.5}$$

play central roles. The decomposition (5.1.4) is orthogonal, in the sense that $\mathfrak{p}$ is the orthogonal complement of $\mathfrak{k}$ with respect to the invariant bilinear form $(\cdot|\cdot) \equiv B(\cdot,\cdot)$,

$$\mathfrak{p} = \{y \in \mathfrak{g} \,|\, \forall x \in \mathfrak{k} : (y|x) = 0\}. \tag{5.1.6}$$

The commutator relations split in a way characteristic for symmetric spaces,

$$[\mathfrak{k},\mathfrak{k}] \subset \mathfrak{k}, \qquad [\mathfrak{k},\mathfrak{p}] \subset \mathfrak{p}, \qquad [\mathfrak{p},\mathfrak{p}] \subset \mathfrak{k}. \tag{5.1.7}$$

The subspace $\mathfrak{p}$ is not a subalgebra. Elements of $\mathfrak{p}$ transform in some representation of $\mathfrak{k}$, which depends on the Lie algebra $\mathfrak{g}$. We stress that if the commutator $[\mathfrak{p},\mathfrak{p}]$ had also contained elements in $\mathfrak{p}$ itself, this would not have been a symmetric space.

The left coset space $K\backslash G$ is defined as the set of equivalence classes $[g]$ of G defined by the equivalence relation

$$g \sim g' \qquad \text{iff } g{g'}^{-1} \in K \text{ and } g, g' \in G, \tag{5.1.8}$$

i.e.,

$$[g] = \{kg \,|\, \forall k \in K\}. \tag{5.1.9}$$

5.1.2 Nonlinear Realisations

The group G naturally acts through right multiplication on the quotient space $K\backslash G$ as

$$[h] \mapsto [hg]. \tag{5.1.10}$$

This definition makes sense because if $h \sim h'$, i.e., $h' = kh$ for some $k \in K$, then $h'g \sim hg$ since $h'g = (kh)g = k(hg)$ (left and right multiplications commute).

In order to describe a dynamical theory on the quotient space $K\backslash G$, it is convenient to introduce as dynamical variable the group element $\mathrm{V}(x) \in G$ and to construct the action for $\mathrm{V}(x)$ in such a way that the equivalence relation

$$\forall k(x) \in K : \mathrm{V}(x) \sim k(x)\mathrm{V}(x) \tag{5.1.11}$$

corresponds to a gauge symmetry. The physical (gauge invariant) degrees of freedom are then parametrized indeed by points of the coset space. We also want to impose Equation (5.1.10) as a rigid symmetry. Thus, the action should be invariant under

$$\mathrm{V}(x) \quad \longmapsto \quad k(x)\mathrm{V}(x)g, \qquad k(x) \in K,\, g \in G. \tag{5.1.12}$$

One may develop the formalism without fixing the K-gauge symmetry, or one may instead fix the gauge symmetry by choosing a specific coset representative $\mathrm{V}(x) \in K\backslash G$. When K is a maximal compact subgroup of G there are no topological obstructions, and a standard choice, which is always available, is to take $\mathrm{V}(x)$ to be of upper triangular form as allowed by the Iwasawa decomposition. This is usually called the *Borel gauge* and will be discussed in more detail later. In this case an arbitrary global transformation,

$$\mathrm{V}(x) \quad \longmapsto \quad \mathrm{V}(x)' = \mathrm{V}(x)g, \qquad g \in G, \tag{5.1.13}$$

will destroy the gauge choice because $\mathrm{V}'(x)$ will generically not be of upper triangular form. Then, a compensating local K-transformation is needed that restores the original gauge choice. The total transformation is thus

$$\mathrm{V}(x) \quad \longmapsto \quad \mathrm{V}(x)'' = k\left(\mathrm{V}(x), g\right) \mathrm{V}(x) g, \qquad k\left(\mathrm{V}(x), g\right) \in K,\, g \in G, \tag{5.1.14}$$

where $\mathrm{V}''(x)$ is again in the upper triangular gauge. Because now $k\left(\mathrm{V}(x), g\right)$ depends non-linearly on $\mathrm{V}(x)$, this is called a *nonlinear realisation* of G.

5.1.3 Three Ways of Writing the Quadratic $K \times G$-Invariant Action

Given the field $\mathrm{V}(x)$, we can form the Lie algebra valued one-form (Maurer–Cartan form)

$$d\mathrm{V}(x)\, \mathrm{V}(x)^{-1} = dx^{\mu}\, \partial_{\mu}\mathrm{V}(x)\mathrm{V}(x)^{-1}. \tag{5.1.15}$$

Under the Cartan decomposition, this element splits according to Equation (5.1.4),

$$\partial_{\mu}\mathrm{V}(x)\, \mathrm{V}(x)^{-1} = \mathrm{Q}_{\mu}(x) + \mathrm{P}_{\mu}(x), \tag{5.1.16}$$

where $\mathrm{Q}_{\mu}(x) \in \mathfrak{k}$ and $\mathrm{P}_{\mu}(x) \in \mathfrak{p}$. We can use the Cartan involution θ to write these explicitly as projections onto the odd and even eigenspaces as follows:

$$\begin{aligned} \mathrm{Q}_{\mu}(x) &= \frac{1}{2}\left[\partial_{\mu}\mathrm{V}(x)\mathrm{V}(x)^{-1} + \theta\left(\partial_{\mu}\mathrm{V}(x)\mathrm{V}(x)^{-1}\right)\right] \in \mathfrak{k}, \\ \mathrm{P}_{\mu}(x) &= \frac{1}{2}\left[\partial_{\mu}\mathrm{V}(x)\mathrm{V}(x)^{-1} - \theta\left(\partial_{\mu}\mathrm{V}(x)\mathrm{V}(x)^{-1}\right)\right] \in \mathfrak{p}. \end{aligned} \tag{5.1.17}$$

If we define a *generalized transpose* $\mathcal{T}$ by

$$(\,)^{\mathcal{T}} \equiv -\theta(\,), \tag{5.1.18}$$

then $\mathrm{P}_{\mu}(x)$ and $\mathrm{Q}_{\mu}(x)$ correspond to symmetric and antisymmetric elements, respectively,

$$\mathrm{P}_{\mu}(x)^{\mathcal{T}} = \mathrm{P}_{\mu}(x), \qquad \mathrm{Q}_{\mu}(x)^{\mathcal{T}} = -\mathrm{Q}_{\mu}(x). \tag{5.1.19}$$

Of course, in the special case when $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{R})$ and $\mathfrak{k} = \mathfrak{so}(n)$, the generalized transpose $(\,)^{\mathcal{T}}$ coincides with the ordinary matrix transpose $(\,)^{T}$. The Lie algebra valued one-forms with components $\partial_{\mu}\mathrm{V}(x)\mathrm{V}(x)^{-1}$, $\mathrm{Q}_{\mu}(x)$ and $\mathrm{P}_{\mu}(x)$ are invariant under rigid right multiplication, $\mathrm{V}(x) \mapsto \mathrm{V}(x)g$.

Being an element of the Lie algebra of the gauge group, $\mathrm{Q}_{\mu}(x)$ can be interpreted as a gauge connection for the local symmetry K. Under a local transformation $k(x) \in K$, $\mathrm{Q}_{\mu}(x)$ transforms as

$$K : \mathrm{Q}_{\mu}(x) \quad \longmapsto \quad k(x)\mathrm{Q}_{\mu}(x)k(x)^{-1} + \partial_{\mu}k(x)k(x)^{-1}, \tag{5.1.20}$$

which indeed is the characteristic transformation property of a gauge connection. On the other hand, $\mathrm{P}_{\mu}(x)$ transforms covariantly,

$$K : \mathrm{P}_{\mu}(x) \quad \longmapsto \quad k(x)\mathrm{P}_{\mu}(x)k(x)^{-1}, \tag{5.1.21}$$

because the element $\partial_{\mu}k(x)k(x)^{-1}$ is projected out due to the negative sign in the construction of $\mathrm{P}_{\mu}(x)$ in Equation (5.1.17).

Using the bilinear form $(\cdot|\cdot)$ we can now form a manifestly $K \times G$-invariant expression by simply "squaring" $\mathrm{P}_\mu(x)$, i.e., the p-dimensional action takes the form (cf. Equation (5.1.1))

$$S_{G/K} = \int_X d^p x \,\sqrt{\gamma}\gamma^{\mu\nu} \left(\mathrm{P}_\mu(x)\middle|\mathrm{P}_\nu(x)\right). \tag{5.1.22}$$

We can rewrite this action in a number of ways. First, we note that since $\mathrm{Q}_\mu(x)$ can be interpreted as a gauge connection we can form a "covariant derivative" D_μ in a standard way as

$$\mathrm{D}_\mu \mathrm{V}(x) \equiv \partial_\mu \mathrm{V}(x) - \mathrm{Q}_\mu(x)\mathrm{V}(x), \tag{5.1.23}$$

which, by virtue of Equation (5.1.17), can alternatively be written as

$$\mathrm{D}_\mu \mathrm{V}(x) = \mathrm{P}_\mu(x)\mathrm{V}(x). \tag{5.1.24}$$

We see now that the action can indeed be interpreted as a gauged nonlinear sigma model, in the sense that the local invariance is obtained by minimally coupling the globally G-invariant expression $(\partial_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}|\partial^\mu \mathrm{V}(x)\mathrm{V}(x)^{-1})$ to the gauge field $\mathrm{Q}_\mu(x)$ through the "covariantization" $\partial_\mu \to \mathrm{D}_\mu$,

$$\left(\partial_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}\middle|\partial^\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}\right) \quad \longrightarrow \quad \left(\mathrm{D}_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}\middle|\mathrm{D}^\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}\right) = \left(\mathrm{P}_\mu(x)\middle|\mathrm{P}_\nu(x)\right). \tag{5.1.25}$$

Thus, the action then takes the form

$$S_{G/K} = \int_X d^p x \,\sqrt{\gamma}\gamma^{\mu\nu} \left(\mathrm{D}_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}\middle|\mathrm{D}_\nu \mathrm{V}(x)\mathrm{V}(x)^{-1}\right). \tag{5.1.26}$$

We can also form a generalized "metric" $\mathrm{M}(x)$ that does not transform at all under the local symmetry, but only transforms under rigid G-transformations. This is done, using the generalized transpose (extended from the algebra to the group through the exponential map [124]), in the following way,

$$\mathrm{M}(x) \equiv \mathrm{V}(x)^{\mathcal{T}}\mathrm{V}(x), \tag{5.1.27}$$

which is clearly invariant under local transformations

$$K : \mathrm{M}(x) \quad \longmapsto \quad \left(k(x)\mathrm{V}(x)\right)^{\mathcal{T}} \left(k(x)\mathrm{V}(x)\right) = \mathrm{V}(x)^{\mathcal{T}} \left(k(x)^{\mathcal{T}} k(x)\right) \mathrm{V}(x) = \mathrm{M}(x) \tag{5.1.28}$$

for $k(x) \in K$, and transforms as follows under global transformations on $\mathrm{V}(x)$ from the right,

$$G : \mathrm{M}(x) \quad \longmapsto \quad g^{\mathcal{T}}\mathrm{M}(x)g, \qquad g \in G. \tag{5.1.29}$$

A short calculation shows that the relation between $\mathrm{M}(x) \in G$ and $\mathrm{P}(x) \in \mathfrak{p}$ is given by

$$\begin{aligned}
\frac{1}{2}\mathrm{M}(x)^{-1}\partial_\mu \mathrm{M}(x) &= \frac{1}{2}\left(\mathrm{V}(x)^{\mathcal{T}}\mathrm{V}(x)\right)^{-1}\partial_\mu \mathrm{V}(x)^{\mathcal{T}}\mathrm{V}(x) + \left(\mathrm{V}(x)^{\mathcal{T}}\mathrm{V}(x)\right)^{-1}\mathrm{V}(x)^{\mathcal{T}}\partial_\mu \mathrm{V}(x) \\
&= \frac{1}{2}\mathrm{V}(x)^{-1}\left[\left(\partial_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}\right)^{\mathcal{T}} + \partial_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}\right]\mathrm{V}(x) \\
&= \mathrm{V}(x)^{-1}\mathrm{P}_\mu(x)\mathrm{V}(x).
\end{aligned} \tag{5.1.30}$$

Since the factors of $\mathrm{V}(x)$ drop out in the squared expression,

$$\left(\mathrm{V}(x)^{-1}\mathrm{P}_\mu(x)\mathrm{V}(x)\middle|\mathrm{V}(x)^{-1}\mathrm{P}^\mu(x)\mathrm{V}(x)\right) = \left(\mathrm{P}_\mu(x)\middle|\mathrm{P}^\mu(x)\right), \tag{5.1.31}$$

Equation (5.1.30) provides a third way to write the $K \times G$-invariant action, completely in terms of the generalized metric $\mathrm{M}(x)$,

$$S_{G/K} = \frac{1}{4}\int_X d^p x\,\sqrt{\gamma}\gamma^{\mu\nu}\left(\mathrm{M}(x)^{-1}\partial_\mu\mathrm{M}(x)\middle|\mathrm{M}(x)^{-1}\partial_\nu\mathrm{M}(x)\right). \tag{5.1.32}$$

We call M a "generalized metric" because in the $SO(n)\backslash GL(n,\mathbb{R})$-case, it does correspond to the metric, the field V being the "vielbein"; see Section 5.3.2.

All three forms of the action are manifestly gauge invariant under K. If desired, one can fix the gauge, and thereby eliminating the redundant degrees of freedom.

5.1.4 Equations of Motion and Conserved Currents

Let us now take a closer look at the equations of motion resulting from an arbitrary variation $\delta\mathrm{V}(x)$ of the action in Equation (5.1.22). The Lie algebra element $\delta\mathrm{V}(x)\mathrm{V}(x)^{-1} \in \mathfrak{g}$ can be decomposed according to the Cartan decomposition,

$$\delta\mathrm{V}(x)\mathrm{V}(x)^{-1} = \Sigma(x) + \Lambda(x), \qquad \Sigma(x) \in \mathfrak{k},\ \Lambda(x) \in \mathfrak{p}. \tag{5.1.33}$$

The variation $\Sigma(x)$ along the gauge orbit will be automatically projected out by gauge invariance of the action. Thus we can set $\Sigma(x) = 0$ for simplicity. Let us then compute $\delta\mathrm{P}_\mu(x)$. One easily gets

$$\delta\mathrm{P}_\mu(x) = \partial_\mu\Lambda(x) - [\mathrm{Q}_\mu(x), \Lambda(x)]. \tag{5.1.34}$$

Since $\Lambda(x)$ is a Lie algebra valued scalar we can freely set $\partial_\mu\Lambda(x) \to \nabla_\mu\Lambda(x)$ in the variation of the action below, where ∇^μ is a covariant derivative on X compatible with the Levi–Civita connection. Using the symmetry and the invariance of the bilinear form one then finds

$$\delta S_{K\backslash G} = \int_X d^p x\,\sqrt{\gamma}\gamma^{\mu\nu}2\left[\left(-\nabla_\nu\mathrm{P}_\mu(x) + [\mathrm{Q}_\nu(x), \mathrm{P}_\mu(x)]\middle|\Lambda(x)\right)\right]. \tag{5.1.35}$$

The equations of motion are therefore equivalent to

$$\mathrm{D}^\mu\mathrm{P}\mu(x) = 0, \tag{5.1.36}$$

with

$$\mathrm{D}_\mu\mathrm{P}_\nu(x) = \nabla_\mu\mathrm{P}_\nu(x) - [\mathrm{Q}_\mu(x), \mathrm{P}_\nu(x)], \tag{5.1.37}$$

and simply express the covariant conservation of $\mathrm{P}_\mu(x)$.

It is also interesting to examine the dynamics in terms of the generalized metric $\mathrm{M}(x)$. The equations of motion for $\mathrm{M}(x)$ are

$$\frac{1}{2}\nabla^\mu\left(\mathrm{M}(x)^{-1}\partial_\mu\mathrm{M}(x)\right) = 0. \tag{5.1.38}$$

These equations ensure the conservation of the current

$$\mathcal{J}_\mu \equiv \frac{1}{2}\mathrm{M}(x)^{-1}\partial_\mu\mathrm{M}(x) = \mathrm{V}(x)^{-1}\mathrm{P}_\mu(x)\mathrm{V}(x), \tag{5.1.39}$$

i.e.,

$$\nabla^\mu\mathcal{J}_\mu = 0. \tag{5.1.40}$$

This is the conserved Noether current associated with the rigid G-invariance of the action.

5.1.5 Parametrization of $K\backslash G$

The Borel gauge choice is always accessible when the group K is the maximal compact subgroup of G. In the noncompact case this is no longer true since one cannot invoke the Iwasawa decomposition (see, e.g. [125] for a discussion of the subtleties involved when K is noncompact). This point will, however, not be of concern to us in this thesis. We shall now proceed to write down the sigma model action in the Borel gauge for the coset space $K\backslash G$, with K being the maximal compact subgroup. Let $\Pi = \{\alpha_1^\vee, \cdots, \alpha_n^\vee\}$ be a basis of the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$, and let $\Delta_+ \subset \mathfrak{h}^\star$ denote the set of positive roots. The Borel subalgebra of $\mathfrak{g}$ can then be written as

$$\mathfrak{b} = \sum_{i=1}^{n} \mathbb{R}\alpha_i^\vee + \sum_{\alpha\in\Delta_+} \mathbb{R}E_\alpha, \tag{5.1.41}$$

where E_α is the positive root generator spanning the one-dimensional root space $\mathfrak{g}_\alpha$ associated to the root α. The coset representative is then chosen to be

$$\mathrm{V}(x) = \mathrm{V}_1(x)\mathrm{V}_2(x) = \mathrm{Exp}\left[\sum_{i=1}^{n} \phi_i(x)\alpha_i^\vee\right] \mathrm{Exp}\left[\sum_{\alpha\in\Delta_+} \chi_\alpha(x)E_\alpha\right] \in K\backslash G. \tag{5.1.42}$$

Because $\mathfrak{g}$ is a finite Lie algebra, the sum over positive roots is finite and so this is a well-defined construction.

From Equation (5.1.42) we may compute the Lie algebra valued one-form $\partial_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}$ explicitly. Let us do this in some detail. First, we write the general expression in terms of $\mathrm{V}_1(x)$ and $\mathrm{V}_2(x)$,

$$\partial_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1} = \partial_\mu \mathrm{V}_1(x)\mathrm{V}_1(x)^{-1} + \mathrm{V}_1(x)\left(\partial_\mu \mathrm{V}_2(x)\mathrm{V}_2(x)^{-1}\right)\mathrm{V}_1(x)^{-1}. \tag{5.1.43}$$

To compute the individual terms in this expression we need to make use of the Baker–Hausdorff formulas:

$$\begin{aligned} \partial_\mu e^A e^{-A} &= \partial_\mu A + \frac{1}{2!}[A, \partial_\mu A] + \frac{1}{3!}[A,[A,\partial_\mu A]] + \cdots, \\ e^A B e^{-A} &= B + [A,B] + \frac{1}{2!}[A,[A,B]] + \cdots. \end{aligned} \tag{5.1.44}$$

The first term in Equation (5.1.43) is easy to compute since all generators in the exponential commute. We find

$$\partial_\mu \mathrm{V}_1(x)\mathrm{V}_1(x)^{-1} = \sum_{i=1}^{n} \partial_\mu\phi_i(x)\alpha_i^\vee \in \mathfrak{h}. \tag{5.1.45}$$

Secondly, we compute the corresponding expression for $\mathrm{V}_2(x)$. Here we must take into account all commutators between the positive root generators $E_\alpha \in \mathfrak{n}_+$. Using the first of the Baker–

Hausdorff formulas above, the first terms in the series become

$$\begin{aligned}\partial_\mu \mathrm{V}_2(x)\mathrm{V}_2(x)^{-1} &= \partial_\mu \operatorname{Exp}\left[\sum_{\alpha\in\Delta_+}\chi_\alpha(x)E_\alpha\right]\operatorname{Exp}\left[-\sum_{\alpha'\in\Delta_+}\chi_{\alpha'}(x)E_{\alpha'}\right]\\ &= \sum_{\alpha\in\Delta_+}\partial_\mu\chi_\alpha(x)E_\alpha+\frac{1}{2!}\sum_{\alpha,\alpha'\in\Delta_+}\chi_\alpha(x)\partial_\mu\chi_{\alpha'}(x)[E_\alpha,E_{\alpha'}]\\ &\quad+\frac{1}{3!}\sum_{\alpha,\alpha',\alpha''\in\Delta_+}\chi_\alpha(x)\chi_{\alpha'}(x)\partial_\mu\chi_{\alpha''}(x)[E_\alpha,[E_{\alpha'},E_{\alpha''}]]+\cdots.\end{aligned} \tag{5.1.46}$$

Each multi-commutator $[E_\alpha,[E_{\alpha'},\cdots]\cdots,E_{\alpha'''}]$ corresponds to some new positive root generator, say $E_\gamma\in\mathfrak{n}_+$. However, since each term in the expansion (5.1.46) is a sum over all positive roots, the specific generator E_γ will get a contribution from all terms. We can therefore write the sum in "closed form" with the coefficient in front of an arbitrary generator E_γ taking the form

$$\mathcal{R}_{\gamma,\mu}(x)\equiv\partial_\mu\chi_\gamma(x)+\frac{1}{2!}\underbrace{\chi_\zeta(x)\partial_\mu\chi_{\zeta'}(x)}_{\zeta+\zeta'=\gamma}+\cdots+\frac{1}{k_\gamma!}\underbrace{\chi_\eta(x)\chi_{\eta'}(x)\cdots\chi_{\eta''}(x)\partial_\mu\chi_{\eta'''}(x)}_{\eta+\eta'+\cdots+\eta''+\eta'''=\gamma}, \tag{5.1.47}$$

where k_γ denotes the number corresponding to the last term in the Baker–Hausdorff expansion in which the generator E_γ appears. The explicit form of $\mathcal{R}_{\gamma,\mu}(x)$ must be computed individually for each root $\gamma\in\Delta_+$.

The sum in Equation (5.1.46) can now be written as

$$\partial_\mu\mathrm{V}_2(x)\mathrm{V}_2(x)^{-1}=\sum_{\alpha\in\Delta_+}\mathcal{R}_{\alpha,\mu}(x)E_\alpha. \tag{5.1.48}$$

To proceed, we must conjugate this expression with $\mathrm{V}_1(x)$ in order to compute the full form of Equation (5.1.43). This involves the use of the second Baker–Hausdorff formula in Equation (5.1.44) for each term in the sum, Equation (5.1.48). Let h denote an arbitrary element of the Cartan subalgebra,

$$h=\sum_{i=1}^{n}\phi_i(x)\alpha_i^\vee\in\mathfrak{h}. \tag{5.1.49}$$

Then the commutators we need are of the form

$$[h,E_\alpha]=\alpha(h)E_\alpha, \tag{5.1.50}$$

where $\alpha(h)$ denotes the value of the root $\alpha\in\mathfrak{h}^\star$ acting on the Cartan element $h\in\mathfrak{h}$,

$$\alpha(h)=\sum_{i=1}^{n}\phi_i(x)\alpha(\alpha_i^\vee)=\sum_{i=1}^{n}\phi_i(x)\left\langle\alpha,\alpha_i^\vee\right\rangle\equiv\sum_{i=1}^{n}\phi_i(x)\alpha_i. \tag{5.1.51}$$

So, for each term in the sum in Equation (5.1.48) we obtain

$$
\begin{aligned}
\mathrm{V}_1(x)E_\alpha\mathrm{V}_1(x)^{-1} &= E_\alpha + \sum_i \phi_i(x)\alpha_i E_\alpha + \frac{1}{2}\sum_{i,j}\phi_i(x)\phi_j(x)\alpha_i\alpha_j E_\alpha + \cdots \\
&= \mathrm{Exp}\left[\sum_i \phi_i(x)\alpha_i\right] E_\alpha \\
&= e^{\alpha(h)}E_\alpha. \qquad (5.1.52)
\end{aligned}
$$

We can now write down the complete expression for the element $\partial_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1}$,

$$
\partial_\mu \mathrm{V}(x)\mathrm{V}(x)^{-1} = \sum_{i=1}^{n}\partial_\mu\phi_i(x)\alpha_i^\vee + \sum_{\alpha\in\Delta_+} e^{\alpha(h)}\mathcal{R}_{\alpha,\mu}(x)E_\alpha. \qquad (5.1.53)
$$

Projection onto the coset $\mathfrak{p}$ gives (see Equation (5.1.17))

$$
\mathrm{P}_\mu(x) = \sum_{i=1}^{n}\partial_\mu\phi_i(x)\alpha_i^\vee + \frac{1}{2}\sum_{\alpha\in\Delta_+} e^{\alpha(h)}\mathcal{R}_{\alpha,\mu}(x)\left(E_\alpha + E_{-\alpha}\right), \qquad (5.1.54)
$$

where we have used that $E_\alpha^{\mathcal{T}} = E_{-\alpha}$ and $(\alpha_i^\vee)^{\mathcal{T}} = \alpha_i^\vee$.

Next we want to compute the explicit form of the action in Equation (5.1.22). Choosing the following normalization for the root generators,

$$
(E_\alpha|E_{\alpha'}) = \delta_{\alpha,-\alpha'}, \qquad (\alpha_i^\vee|\alpha_j^\vee) = \delta_{ij}, \qquad (5.1.55)
$$

which implies

$$
(E_\alpha|E_{\alpha'}^{\mathcal{T}}) = (E_\alpha|E_{-\alpha'}) = \delta_{\alpha,\alpha'} \qquad (5.1.56)
$$

one finds the form of the $K\times G$-invariant action in the parametrization of Equation (5.1.42),

$$
S_{K\backslash G} = \int_X d^px\,\sqrt{\gamma}\gamma^{\mu\nu}\left[\sum_{i=1}^{n}\partial_\mu\phi_i(x)\partial_\nu\phi_i(x) + \frac{1}{2}\sum_{\alpha\in\Delta_+} e^{2\alpha(h)}\mathcal{R}_{\alpha,\mu}(x)\mathcal{R}_{\alpha,\nu}(x)\right]. \qquad (5.1.57)
$$

5.2 Geodesic Sigma Models on Infinite-Dimensional Coset Spaces

In the following we shall both "generalize and specialize" the construction from Section 5.1. The generalization amounts to relaxing the restriction that the algebra $\mathfrak{g}$ be finite-dimensional. Although in principle we could consider $\mathfrak{g}$ to be any indefinite Kac–Moody algebra, we shall be focusing on the case where it is of Lorentzian type. The analysis will also be a specialization, in the sense that we consider only *geodesic* sigma models, meaning that the Riemannian space X is the one-dimensional worldline of a particle, parametrized by one variable t. This restriction is of course motivated by the billiard description of gravity close to a spacelike singularity, where the dynamics at each spatial point is effectively described by a particle geodesic in the fundamental Weyl chamber of a Lorentzian Kac–Moody algebra.

The motivation is that the construction of a geodesic sigma model that exhibits this Kac–Moody symmetry in a manifest way, would provide a link to understanding the role of the full algebra $\mathfrak{g}$ beyond the BKL-limit.

5.2.1 Formal Construction

For definiteness, we consider only the case when the Lorentzian algebra $\mathfrak{g}$ is a split real form, although this is not really necessary as the Iwasawa decomposition holds also in the non-split case.

A very important difference from the finite-dimensional case is that we now have nontrivial *multiplicities* of the imaginary roots (see Section 2.2.3). Recall that if a root $\alpha \in \Delta$ has multiplicity m_α, then the associated root space $\mathfrak{g}_\alpha$ is m_α-dimensional. Thus, it is spanned by m_α generators $E_\alpha^{[s]}$ $(s = 1, \cdots, m_\alpha)$,

$$\mathfrak{g}_\alpha = \mathbb{R}E_\alpha^{[1]} + \cdots + \mathbb{R}E_\alpha^{[m_\alpha]}. \tag{5.2.1}$$

The root multiplicities are not known in closed form for any indefinite Kac–Moody algebra, but must be computed recursively as described in Section 2.5.

Our main object of study is the coset representative $\mathcal{V}(t) \in K\backslash G$, which must now be seen as "formal" exponentiation of the infinite number of generators in $\mathfrak{p}$. We can then proceed as before and choose $\mathcal{V}(t)$ to be in the Borel gauge, i.e., of the form

$$\mathcal{V}(t) = \mathrm{Exp}\left[\sum_{\mu=1}^{\dim\mathfrak{h}} \beta^\mu(t)\alpha_\mu^\vee\right] \mathrm{Exp}\left[\sum_{\alpha\in\Delta_+}\sum_{s=1}^{m_\alpha} \xi_\alpha^{[s]}(t)E_\alpha^{[s]}\right] \in K\backslash G. \tag{5.2.2}$$

Here, the index μ does not correspond to "spacetime" but instead is an index in the Cartan subalgebra $\mathfrak{h}$, or, equivalently, in "scale-factor space" (see Section 3.1.6). In the following we shall dispose of writing the sum over μ explicitly. The second exponent in Equation (5.2.2) contains a formal infinite sum over all positive roots Δ_+. We will describe in detail in subsequent sections how it can be suitably truncated. The coset representative $\mathcal{V}(t)$ corresponds to a nonlinear realisation of G and transforms as

$$G : \mathcal{V}(t) \quad\longmapsto\quad k\left(\mathcal{V}(t), g\right)\mathcal{V}(t)g, \qquad k\left(\mathcal{V}(t), g\right) \in K,\ g \in G. \tag{5.2.3}$$

A $\mathfrak{g}$-valued "one-form" can be constructed analogously to the finite-dimensional case,

$$\partial\mathcal{V}(t)\mathcal{V}(t)^{-1} = \mathcal{Q}(t) + \mathcal{P}(t), \tag{5.2.4}$$

where $\partial \equiv \partial_t$. The first term on the right hand side represents a $\mathfrak{k}$-connection that is fixed under the Chevalley involution,

$$\tau(\mathcal{Q}) = \mathcal{Q}, \tag{5.2.5}$$

while $\mathcal{P}(t)$ lies in the orthogonal complement $\mathfrak{p}$ and so is anti-invariant,

$$\tau(\mathcal{P}) = -\mathcal{P} \tag{5.2.6}$$

(for the split form, the Cartan involution coincides with the Chevalley involution). Using the projections onto the coset $\mathfrak{p}$ and the compact subalgebra $\mathfrak{k}$, as defined in Equation (5.1.17), we can compute the forms of $\mathcal{P}(t)$ and $\mathcal{Q}(t)$ in the Borel gauge, and we find

$$\begin{aligned} \mathcal{P}(t) &= \partial\beta^\mu(t)\alpha_\mu^\vee + \frac{1}{2}\sum_{\alpha\in\Delta_+}\sum_{s=1}^{m_\alpha} e^{\alpha(\beta)}\mathfrak{R}_\alpha^{[s]}(t)\left(E_\alpha^{[s]} + E_{-\alpha}^{[s]}\right), \\ \mathcal{Q}(t) &= \frac{1}{2}\sum_{\alpha\in\Delta_+}\sum_{s=1}^{m_\alpha} e^{\alpha(\beta)}\mathfrak{R}_\alpha^{[s]}(t)\left(E_\alpha^{[s]} - E_{-\alpha}^{[s]}\right), \end{aligned} \tag{5.2.7}$$

where $\mathfrak{R}^{[s]}_{\alpha}(t)$ is the analogue of $\mathcal{R}_{\alpha}(x)$ in the finite-dimensional case, i.e., it takes the form

$$\mathfrak{R}^{[s]}_{\alpha}(t) = \partial\xi^{[s]}_{\alpha}(t) + \frac{1}{2}\underbrace{\xi^{[s]}_{\zeta}(t)\partial\xi^{[s]}_{\zeta'}(t)}_{\zeta+\zeta'=\alpha} + \cdots, \tag{5.2.8}$$

which still contains a finite number of terms for each positive root α. The value of the root $\alpha \in \mathfrak{h}^{\star}$ acting on $\beta = \beta^{\mu}(t)\alpha^{\vee}_{\mu} \in \mathfrak{h}$ is

$$\alpha(\beta) = \alpha_{\mu}\beta^{\mu}. \tag{5.2.9}$$

Note that here the notation α_{μ} does not correspond to a simple root, but denotes the components of an arbitrary root vector $\alpha \in \mathfrak{h}^{\star}$.

The action for a particle moving on the infinite-dimensional coset space $K\backslash G$ can now be constructed using the invariant bilinear form $(\cdot|\cdot)$ on $\mathfrak{g}$,

$$S_{K\backslash G} = \int dt\, n(t)^{-1}\left(\mathcal{P}(t)|\mathcal{P}(t)\right), \tag{5.2.10}$$

where $n(t)$ ensures invariance under reparametrizations of t. The variation of the action with respect to $n(t)$ constrains the motion to be a *null geodesic* on $K\backslash G$,

$$(\mathcal{P}(t)|\mathcal{P}(t)) = 0. \tag{5.2.11}$$

Defining, as before, a covariant derivative $\mathfrak{D}$ with respect to the local symmetry K as

$$\mathfrak{D}\mathcal{P}(t) \equiv \partial\mathcal{P}(t) - [\mathcal{Q}(t), \mathcal{P}(t)], \tag{5.2.12}$$

the equations of motion read simply

$$\mathfrak{D}\left(n(t)^{-1}\mathcal{P}(t)\right) = 0. \tag{5.2.13}$$

The explicit form of the action in the parametrization of Equation (5.2.2) becomes

$$S_{K\backslash G} = \int dt\, n(t)^{-1}\left[G_{\mu\nu}\,\partial\beta^{\mu}(t)\,\partial\beta^{\nu}(t) + \frac{1}{2}\sum_{\alpha\in\Delta_+}\sum_{s=1}^{m_{\alpha}} e^{2\alpha(\beta)}\,\mathfrak{R}^{[s]}_{\alpha}(t)\,\mathfrak{R}^{[s]}_{\alpha}(t)\right], \tag{5.2.14}$$

where $G_{\mu\nu}$ is the flat Lorentzian metric, defined by the restriction of the bilinear form $(\cdot|\cdot)$ to the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$. The metric $G_{\mu\nu}$ is exactly the same as the metric in scale-factor space (see Section 3.1.6), and the kinetic term for the Cartan parameters $\beta^{\mu}(t)$ reads explicitly

$$G_{\mu\nu}\,\partial\beta^{\mu}(t)\,\partial\beta^{\nu}(t) = \sum_{i=1}^{\dim\mathfrak{h}-1}\partial\beta^{i}(t)\,\partial\beta^{i}(t) - \left(\sum_{i=1}^{\dim\mathfrak{h}-1}\partial\beta^{i}(t)\right)\left(\sum_{j=1}^{\dim\mathfrak{h}-1}\partial\beta^{j}(t)\right) + \partial\phi(t)\,\partial\phi(t). \tag{5.2.15}$$

Although $\mathfrak{g}$ is infinite-dimensional we still have the notion of "formal integrability", owing to the existence of an infinite number of conserved charges, defined by the equations of motion in Equation (5.2.13). We can define the generalized metric for any $\mathfrak{g}$ as

$$\mathcal{M}(t) \equiv \mathcal{V}(t)^{\mathcal{T}}\mathcal{V}(t), \tag{5.2.16}$$

where the transpose $(\,)^{\mathcal{T}}$ is defined as before in terms of the Chevalley involution,

$$(\,)^{\mathcal{T}} = -\tau(\,). \tag{5.2.17}$$

Then the equations of motion $\mathfrak{D}\mathcal{P}(t) = 0$ are equivalent to the conservation $\partial\mathfrak{J} = 0$ of the current

$$\mathfrak{J} \equiv \frac{1}{2}\mathcal{M}(t)^{-1}\partial\mathcal{M}(t). \tag{5.2.18}$$

This can be formally solved in closed form

$$\mathcal{M}(t) = e^{t\mathfrak{J}^{\mathcal{T}}}\mathcal{M}(0)e^{t\mathfrak{J}}, \tag{5.2.19}$$

and so an arbitrary group element $g \in G$ evolves according to

$$g(t) = k(t)g(0)e^{t\mathfrak{J}}, \qquad k(t) \in K. \tag{5.2.20}$$

Although the explicit form of $\mathcal{P}(t)$ contains infinitely many terms, we have seen that each coefficient $\mathfrak{R}_{\alpha}^{[s]}(t)$ can, in principle, be computed exactly for each root α. This, however, is not the case for the current $\mathfrak{J}$. To find the form of $\mathfrak{J}$ one must conjugate $\mathcal{P}(t)$ with the coset representative $\mathcal{V}(t)$ and this requires an infinite number of commutators to get the correct coefficient in front of any generator in $\mathfrak{J}$.

5.2.2 Consistent Truncations

One method for dealing with infinite expressions like Equation (5.2.7) consists in considering successive finite expansions allowing more and more terms, while still respecting the dynamics of the sigma model.

This leads us to the concept of a *consistent truncation* of the sigma model for $K\backslash G$. We take as definition of such a truncation any sub-model S' of $S_{K\backslash G}$ whose solutions are also solutions of the original model.

There are two main approaches to finding suitable truncations that fulfill this latter criterion. These are the so-called *subgroup truncations* and the *level truncations*, which will both prove to be useful for our purposes, and we consider them in turn below.

Subgroup Truncation

The first consistent truncation we shall treat is the case when the dynamics of a sigma model for some global group G is restricted to that of an appropriately chosen subgroup $\bar{G} \subset G$. We consider here only subgroups $\bar{G}$ which are obtained by exponentiation of regular subalgebras $\bar{\mathfrak{g}}$ of $\mathfrak{g}$. The concept of regular embeddings of Lorentzian Kac–Moody algebras is discussed in detail in **Paper I**.

To restrict the dynamics to that of a sigma model based on the coset space $K(\bar{G})\backslash\bar{G}$, we first assume that the initial conditions $g(t)\big|_{t=0} = g(0)$ and $\partial g(t)\big|_{t=0}$ are such that the following two conditions are satisfied:

i. The group element $g(0)$ belongs to $\bar{G}$.

ii. The conserved current $\mathfrak{J}$ belongs to $\bar{\mathfrak{g}}$.

When these conditions hold, then $g(0)e^{t\mathfrak{J}}$ belongs to $\bar{G}$ for all t. Moreover, there always exists $\bar{k}(t) \in K(\bar{G})$ such that

$$\bar{g}(t) \equiv \bar{k}(t)g(0)e^{t\mathfrak{J}} \in K(\bar{G})\backslash\bar{G}, \tag{5.2.21}$$

i.e, $\bar{g}(t)$ belongs to the Borel subgroup of $\bar{G}$. Because the embedding is regular, $\bar{k}(t)$ belongs to K and we thus have that $\bar{g}(t)$ also belongs to the Borel subgroup of the full group G.

Now recall that from Equation (5.2.20), we know that $\bar{g}(t) = \bar{k}(t)g(0)e^{t\mathfrak{J}}$ is a solution to the equations of motion for the sigma model on $K(\bar{G})\backslash\bar{G}$. But since we have found that $\bar{g}(t)$ preserves the Borel gauge for $K\backslash G$, it follows that $\bar{k}(t)g(0)e^{t\mathfrak{J}}$ is a solution to the equations of motion for the full sigma model. Thus, the dynamical evolution of the subsystem $S' = S_{K(\bar{G})\backslash\bar{G}}$ preserves the Borel gauge of G. These arguments show that initial conditions in $\bar{G}$ remain in $\bar{G}$, and hence the dynamics of a sigma model on $K\backslash G$ can be consistently truncated to a sigma model on $K(\bar{G})\backslash\bar{G}$.

Finally, we recall that because the embedding $\bar{\mathfrak{g}} \subset \mathfrak{g}$ is regular, the restriction of the bilinear form on $\mathfrak{g}$ coincides with the bilinear form on $\bar{\mathfrak{g}}$. This implies that the Hamiltonian constraints for the two models, arising from time reparametrization invariance of the action, also coincide.

We shall make use of subgroup truncations in Chapter 6.

Level Truncation and Height Truncation

Alternative ways of consistently truncating the infinite-dimensional sigma model rest on the use of *gradations* of $\mathfrak{g}$,

$$\mathfrak{g} = \cdots + \mathfrak{g}_{-2} + \mathfrak{g}_{-1} + \mathfrak{g}_0 + \mathfrak{g}_1 + \mathfrak{g}_2 + \cdots , \tag{5.2.22}$$

where the sum is infinite but each subspace is finite-dimensional. One also has

$$[\mathfrak{g}_{\ell'}, \mathfrak{g}_{\ell''}] \subseteq \mathfrak{g}_{\ell'+\ell''}. \tag{5.2.23}$$

Such a gradation was for instance constructed in Section 2.5 and was based on a so-called *level decomposition* of the adjoint representation of $\mathfrak{g}$ into representations of a finite regular subalgebra $\mathfrak{r} \subset \mathfrak{g}$. We will now use this construction to truncate the sigma model based of $K\backslash G$, by "terminating" the gradation of $\mathfrak{g}$ at some finite level $\bar{\ell}$. More specifically, the truncation will involve setting to zero all coefficients $\mathfrak{R}_\alpha^{[s]}(t)$, in the expansion of $\mathcal{P}(t)$, corresponding to roots α whose generators $E_\alpha^{[s]}$ belong to subspaces $\mathfrak{g}_\ell$ with $\ell > \bar{\ell}$. Part of this section draws inspiration from the treatment in [51, 53, 82].

The level ℓ might be the height, or it might count the number of times a specified single simple root appears. In that latter case, the actual form of the level decomposition must of course be worked out separately for each choice of algebra $\mathfrak{g}$ and each choice of decomposition. We will do this in detail in Section 5.3 for a specific level decomposition of the hyperbolic algebra E_{10}. Here, we shall display the general construction in the case of the *height truncation*, which exists for any algebra.

Let α be a positive root, $\alpha \in \Delta_+$. It has the following expansion in terms of the simple roots

$$\alpha = \sum_i m_i\alpha_i \qquad (m_i \geq 0). \tag{5.2.24}$$

Then the *height* of α is defined as (see Section 2.2.3)

$$\mathrm{ht}(\alpha) = \sum_i m_i. \tag{5.2.25}$$

The height can thus be seen as a linear integral map $\mathrm{ht} : \Delta \to \mathbb{Z}$, and we shall sometimes use the notation $\mathrm{ht}(\alpha) = h_\alpha$ to denote the value of the map ht acting on a root $\alpha \in \Delta$.

To achieve the height truncation, we assume that the sum over all roots in the expansion of $\mathcal{P}(t)$, Equation (5.2.7), is ordered in terms of increasing height. Then we can consistently set to zero all coefficients $\mathfrak{R}_\alpha^{[s]}(t)$ corresponding to roots with greater height than some, suitably chosen, finite height $\bar{h}$. We thus find that the finitely truncated coset element $\mathcal{P}_0(t)$ is

$$\mathcal{P}_0(t) \equiv \mathcal{P}(t)\big|_{\mathrm{ht}\leq\bar{h}} = \partial\beta^\mu(t)\alpha_\mu^\vee + \frac{1}{2}\sum_{\substack{\alpha\in\Delta_+\\ \mathrm{ht}(\alpha)\leq\bar{h}}}\sum_{s=1}^{m_\alpha} e^{\alpha(\beta)}\mathfrak{R}_\alpha^{[s]}(t)\left(E_\alpha^{[s]} + E_{-\alpha}^{[s]}\right), \tag{5.2.26}$$

which is equivalent to the statement

$$\mathfrak{R}_\gamma^{[s]}(t) = 0 \qquad \forall\gamma\in\Delta_+,\ \mathrm{ht}(\gamma) > \bar{h}. \tag{5.2.27}$$

For further use, we note here some properties of the coefficients $\mathfrak{R}_\alpha^{[s]}(t)$. By examining the structure of Equation (5.2.8), we see that $\mathfrak{R}_\alpha^{[s]}(t)$ takes the form of a temporal derivative acting on $\xi_\alpha^{[s]}(t)$, followed by a sequence of terms whose individual components, for example $\xi_\zeta^{[s]}(t)$, are all associated with roots of *lower* height than α, $\mathrm{ht}(\zeta) < \mathrm{ht}(\alpha)$. It will prove useful to think of $\mathfrak{R}_\alpha^{[s]}(t)$ as representing a kind of "generalized" derivative operator acting on the field $\xi_\alpha^{[s]}$. Thus we define the operator $\mathcal{D}$ by

$$\mathcal{D}\xi_\alpha^{[s]}(t) \equiv \partial\xi_\alpha^{[s]}(t) + \mathcal{F}_\alpha^{[s]}\left(\xi\partial\xi, \xi^2\partial\xi, \cdots\right), \tag{5.2.28}$$

where $\mathcal{F}_\alpha^{[s]}(t)$ is a polynomial function of the coordinates $\xi(t)$, whose explicit structure follows from Equation (5.2.8). It is common in the literature to refer to the level truncation as "setting all higher level covariant derivatives to zero", by which one simply means that all $\mathcal{D}\xi_\gamma^{[s]}(t)$ corresponding to roots γ above a given finite level $\bar{\ell}$ should vanish. Following [53] we shall call the operators $\mathcal{D}$ "covariant derivatives".

It is clear from the equations of motion $\mathfrak{D}\mathcal{P}(t) = 0$, that if all covariant derivatives $\mathcal{D}\xi_\gamma^{[s]}(t)$ above a given height are set to zero, this choice is preserved by the dynamical evolution. Hence, the height (and any level) truncation is indeed a consistent truncation. Let us here emphasize that it is *not* consistent by itself to merely put all fields $\xi_\gamma^{[s]}(t)$ above a certain level to zero, but one must take into account the fact that combinations of lower level fields may parametrize a higher level generator in the expansion of $\mathcal{P}(t)$, and therefore it is crucial to define the truncation using the derivative operator $\mathcal{D}\xi_\gamma^{[s]}(t)$.

5.3 Eleven-Dimensional Supergravity and $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$

We shall now illustrate the results of the previous sections by explicitly constructing an action for the coset space $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$. We employ the level decomposition of $E_{10} = \mathrm{Lie}\,\mathcal{E}_{10}$ in terms of its regular $\mathfrak{sl}(10,\mathbb{R})$-subalgebra (see Section 2.5), to write the coordinates on the coset space as (time-dependent) $\mathfrak{sl}(10,\mathbb{R})$-tensors. It is then shown that for a truncation of the sigma model at level $\ell = 3$, these fields can be interpreted as the physical fields of eleven-dimensional supergravity. This "dictionary" ensures that the equations of motion arising from the sigma model on $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ are equivalent to the (suitably truncated) equations of motion of eleven-dimensional supergravity [53].

5.3.1 Low-Level Fields

We perform the level decomposition of E_{10} with respect to the $\mathfrak{sl}(10,\mathbb{R})$-subalgebra obtained by removing the exceptional node in the Dynkin diagram in Figure 2.14. This procedure was described in Section 2.5. When using this decomposition, a sum over (positive) roots becomes a sum over all $\mathfrak{sl}(10,\mathbb{R})$-indices in each (positive) representation appearing in the decomposition.

We recall that up to level three the following representations appear

$$\begin{aligned} \ell = 0: &\qquad K^a{}_b, \\ \ell = 1: &\qquad E^{abc} = E^{[abc]}, \\ \ell = 2: &\qquad E^{a_1\cdots a_6} = E^{[a_1\cdots a_6]}, \\ \ell = 3: &\qquad E^{a|b_1\cdots b_8} = E^{a|[b_1\cdots b_8]}, \end{aligned} \tag{5.3.1}$$

where all indices are $\mathfrak{sl}(10,\mathbb{R})$-indices and so run from 1 to 10. The level zero generators $K^a{}_b$ correspond to the adjoint representation of $\mathfrak{sl}(10,\mathbb{R})$ and the higher level generators correspond to an infinite tower of raising operators of E_{10}. As indicated by the square brackets, the level one and two representations are completely antisymmetric in all indices, while the level three representation has a mixed Young tableau symmetry: It is antisymmetric in the eight indices $b_1\cdots b_8$ and is also subject to the constraint

$$E^{[a|b_1\cdots b_8]} = 0. \tag{5.3.2}$$

In the scale factor space (β-basis), the roots of E_{10} corresponding to the above generators act as follows on $\beta \in \mathfrak{h}$:

$$\begin{aligned} K^a{}_b &\iff \alpha_{ab}(\beta) = \beta^a - \beta^b \qquad (a > b), \\ E^{abc} &\iff \alpha_{abc}(\beta) = \beta^a + \beta^b + \beta^c, \\ E^{a_1\cdots a_6} &\iff \alpha_{a_1\cdots a_6}(\beta) = \beta^{a_1} + \cdots + \beta^{a_6}, \\ E^{a|ab_1\cdots b_7} &\iff \alpha_{ab_1\cdots b_7}(\beta) = 2\beta^a + \beta^{b_1} + \cdots + \beta^{b_7}, \\ E^{a_1|a_2\cdots a_9} &\iff \alpha_{a_1\cdots a_9}(\beta) = \beta^{a_1} + \cdots + \beta^{a_9}. \end{aligned} \tag{5.3.3}$$

We can use the scalar product in root space, $\mathfrak{h}^\star$, to compute the norms of these roots. The metric on $\mathfrak{h}^\star$ is the inverse of the metric in Equation (5.2.15), and for E_{10} it takes the form

$$(\omega|\omega) = G^{ij}\omega_i\omega_j = \sum_{i=1}^{10}\omega_i\omega_i - \frac{1}{9}\left(\sum_{i=1}^{10}\omega_i\right)\left(\sum_{j=1}^{10}\omega_j\right), \qquad \omega \in \mathfrak{h}^\star. \tag{5.3.4}$$

The level zero, one and two generators correspond to real roots of E_{10},

$$(\alpha_{ab}|\alpha_{cd}) = 2, \qquad (\alpha_{abc}|\alpha_{def}) = 2, \qquad (\alpha_{a_1\cdots a_6}|\alpha_{b_1\cdots b_6}) = 2. \tag{5.3.5}$$

We have split the roots corresponding to the level three generators into two parts, depending on whether or not the special index a takes the same value as one of the other indices. The resulting two types of roots correspond to real and null roots, respectively,

$$(\alpha_{ab_1\cdots b_7}|\alpha_{cd_1\cdots d_7}) = 2, \quad (\alpha_{a_1\cdots a_9}|\alpha_{b_1\cdots b_9}) = 0. \tag{5.3.6}$$

Thus, the first time that generators corresponding to imaginary roots appear in the level decomposition is at level three. This will prove to be important later on in our analysis.

5.3.2 The $SO(10)\backslash GL(10,\mathbb{R})$-Sigma Model

Because of the importance and geometric significance of level zero, we shall first develop the formalism for the $SO(10)\backslash GL(10,\mathbb{R})$-sigma model. A general group element H in the subgroup $GL(10,\mathbb{R})$ reads

$$H = \text{Exp}\left[h_a{}^b K^a{}_b\right] \tag{5.3.7}$$

where $h_a{}^b$ is a 10×10 matrix (with a being the row index and b the column index). Although the $K^a{}_b$'s are generators of E_{10} and can, within this framework, at best be viewed as infinite matrices, it will prove convenient – for streamlining the calculations – to view them in the present section also as 10×10 matrices, since we confine our attention to the finite-dimensional subgroup $GL(10,\mathbb{R})$. Namely, $K^a{}_b$ is treated as a 10×10 matrix with 0's everywhere except 1 in position (a,b). The final formulation in terms of the variables $h_a{}^b(t)$ – which are 10×10 matrices irrespectively as to whether one deals with $GL(10,\mathbb{R})$ *per se* or as a subgroup of E_{10} – does not depend on this interpretation.

It is also useful to describe $GL(10,\mathbb{R})$ as the set of linear combinations $m_i{}^j K^i{}_j$ where the 10×10 matrix $m_i{}^j$ is invertible. The product of the $K^i{}_j$'s is given by

$$K^i{}_j\, K^k{}_m = \delta^k{}_j\, K^i{}_m. \tag{5.3.8}$$

One easily verifies that if $M = m_i{}^j K^i{}_j$ and $N = n_i{}^j K^i{}_j$ belong to $GL(10,\mathbb{R})$, then $MN = (mn)_i{}^j K^i{}_j$ where mn is the standard product of the 10×10 matrices m and n. Furthermore, $\text{Exp}\left(h_i{}^j K^i{}_j\right) = \left(e^h\right)_i{}^j K^i{}_j$ where e^h is the standard matrix exponential.

Under a general transformation, the representative $H(t)$ is multiplied from the left by a time-dependent $SO(10)$ group element R and from the right by a constant linear $GL(10,\mathbb{R})$-group element L. Explicitly, the transformation takes the form (suppressing the time-dependence for notational convenience)

$$H \to H' = RHL. \tag{5.3.9}$$

In terms of components, with $H = e_a{}^b K^a{}_b$, $e_a{}^b = (e^h)_a{}^b$, $R = R_a{}^b K^a{}_b$ and $L = L_a{}^b K^a{}_b$, one finds

$$e'^{\,b}_a = R_a{}^c e_c{}^d L_d{}^b, \tag{5.3.10}$$

where we have set $H' = e'^{\,b}_a K^a{}_b$. The indices on the coset representative have different covariance properties. To emphasize this fact, we shall write a bar over the first index, $e_a{}^b \to e_{\bar{a}}{}^b$. Thus, barred indices transform under the local $SO(10)$ gauge group and are called “local”, or also “flat”, indices, while unbarred indices transform under the global

$GL(10,\mathbb{R})$ and are called "world", or also "curved", indices. The gauge invariant matrix product $M = H^T H$ is equal to

$$M = g^{ab} K_{ab}, \tag{5.3.11}$$

with $K_{ab} \equiv K^c{}_b \delta_{ac}$ and

$$g^{ab} = \sum_{\bar{c}} e_{\bar{c}}{}^a e_{\bar{c}}{}^b. \tag{5.3.12}$$

The g^{ab} do not transform under local $SO(10)$-transformations and transform as a (symmetric) contravariant tensor under rigid $GL(10,\mathbb{R})$-transformations,

$$g'^{ab} = g^{cd} L_c{}^a L_d{}^b. \tag{5.3.13}$$

They are components of a nondegenerate symmetric matrix that can be identified with an inverse Euclidean metric.

Indeed, the coset space $SO(10)\backslash GL(10,\mathbb{R})$ can be identified with the space of symmetric tensors of Euclidean signature, i.e., the space of metrics. This is because two symmetric tensors of Euclidean signature are equivalent under a change of frame, and the isotropy subgroup, say at the identity, is evidently $SO(10)$. In that view, the coset representative $e_a{}^b$ is the spatial vielbein.

The action for the coset space $SO(10)\backslash GL(10,\mathbb{R})$ with the metric of Equation (2.5.82) is easily found to be

$$\mathcal{L}_0 = \frac{1}{4}\Big(g^{ac}(t)g^{bd}(t) - g^{ab}(t)g^{cd}(t)\Big)\,\partial g_{ab}(t)\,\partial g_{cd}(t). \tag{5.3.14}$$

Note that the quadratic form multiplying the time derivatives is just the "De Witt supermetric" [126]. Note also for future reference that the invariant form $\partial H\, H^{-1}$ reads explicitly

$$\partial H\, H^{-1} = \partial e_{\bar{a}}{}^{\;b}\, e_b{}^{\bar{c}}\, K^a{}_c, \tag{5.3.15}$$

where $e_b{}^{\bar{n}}$ is the inverse vielbein.

5.3.3 Sigma Model Fields and $SO(10)_{\text{local}} \times GL(10,\mathbb{R})_{\text{rigid}}$-Covariance

We now turn to the full nonlinear sigma model for $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$. Rather than exponentiating the Cartan subalgebra separately as in Equation (5.2.2), it will here prove convenient to instead single out the level zero subspace $\mathfrak{g}_0 = \mathfrak{gl}(10,\mathbb{R})$. This permits one to control easily $SO(10)_{\text{local}} \times GL(10,\mathbb{R})_{\text{rigid}}$-covariance. To make this level zero covariance manifest, we shall furthermore assume that the Borel gauge has been fixed only for the non-zero levels, and we keep all level zero fields present. The residual gauge freedom is then just multiplication by an $SO(10)$ rotation from the left.

Thus, we take a coset representative of the form

$$\mathcal{V}(t) = H(t)\ \mathrm{Exp}\left[\frac{1}{3!}\mathcal{A}_{abc}(t)E^{abc} + \frac{1}{6!}\mathcal{A}_{a_1\cdots a_6}(t)E^{a_1\cdots a_6} + \frac{1}{9!}\mathcal{A}_{a|b_1\cdots b_8}(t)E^{a|b_1\cdots b_8} + \cdots\right], \tag{5.3.16}$$

where the sum in the first exponent would be restricted to all $a \geq b$ if we had taken a full Borel gauge also at level zero. The parameters $\mathcal{A}_{abc}(t), \mathcal{A}_{a_1\cdots a_6}(t)$ and $\mathcal{A}_{a|b_1\cdots b_8}(t)$ are coordinates on the coset space $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ and will eventually be interpreted as physical time-dependent fields of eleven-dimensional supergravity.

How do the fields transform under $SO(10)_{\text{local}} \times GL(10,\mathbb{R})_{\text{rigid}}$? Let $R \in SO(10)$, $L \in GL(10,\mathbb{R})$ and decompose $\mathcal{V}$ according to Equation (5.3.16) as the product

$$\mathcal{V} = HT, \tag{5.3.17}$$

with

$$\begin{aligned} H &= \text{Exp}\left[h_a{}^b(t)K^a{}_b\right] \in GL(10,\mathbb{R}), \\ T &= \text{Exp}\left[\frac{1}{3!}\mathcal{A}_{abc}(t)E^{abc} + \frac{1}{6!}\mathcal{A}_{a_1\cdots a_6}(t)E^{a_1\cdots a_6} + \frac{1}{9!}\mathcal{A}_{a|b_1\cdots b_8}(t)E^{a|b_1\cdots b_8} + \cdots\right]. \end{aligned} \tag{5.3.18}$$

One has

$$\mathcal{V} \to \mathcal{V}' = R(HT)L = (RHL)(L^{-1}TL). \tag{5.3.19}$$

Now, the first matrix $H' = RHL$ clearly belongs to $GL(10,\mathbb{R})$, since it is the product of a rotation matrix by two $GL(10,\mathbb{R})$-matrices. It has exactly the same transformation as in Equation (5.3.9) above in the context of the nonlinear sigma model for $SO(10)\backslash GL(10,\mathbb{R})$. Hence, the geometric interpretation of $e_{\bar{a}}{}^b = (e^h)_{\bar{a}}{}^b$ as the vielbein remains.

Similarly, the matrix $T' \equiv L^{-1}TL$ has exactly the same form as T,

$$\begin{aligned} T' &= \text{Exp}\left(L^{-1}\left[\frac{1}{3!}\mathcal{A}_{abc}(t)E^{abc} + \frac{1}{6!}\mathcal{A}_{a_1\cdots a_6}(t)E^{a_1\cdots a_6} + \frac{1}{9!}\mathcal{A}_{a|b_1\cdots b_8}(t)E^{a|b_1\cdots b_8} + \cdots\right]L\right) \\ &= \text{Exp}\left[\frac{1}{3!}\mathcal{A}'_{abc}(t)E^{abc} + \frac{1}{6!}\mathcal{A}'_{a_1\cdots a_6}(t)E^{a_1\cdots a_6} + \frac{1}{9!}\mathcal{A}'_{a|b_1\cdots b_8}(t)E^{a|b_1\cdots b_8} + \cdots\right], \end{aligned} \tag{5.3.20}$$

where the variables $\mathcal{A}'_{abc}$, $\mathcal{A}'_{a_1\cdots a_6}$, ..., are obtained from the variables $\mathcal{A}_{abc}$, $\mathcal{A}_{a_1\cdots a_6}$, ..., by computing $L^{-1}E^{abc}L$, $L^{-1}E^{a_1\cdots a_6}L$, ..., using the commutation relations with $K^a{}_b$. Explicitly, one gets

$$\mathcal{A}'_{abc} = (L^{-1})_a{}^e(L^{-1})_b{}^f(L^{-1})_c{}^g\mathcal{A}_{efg}, \qquad \mathcal{A}'_{a_1\cdots a_6} = (L^{-1})_{a_1}{}^{b_1}\cdots(L^{-1})_{a_6}{}^{b_6}\mathcal{A}_{b_1\cdots b_6}, \qquad \text{etc.} \tag{5.3.21}$$

Hence, the fields $\mathcal{A}_{abc}$, $\mathcal{A}_{a_1\cdots a_6}$, ... do not transform under local $SO(10)$ transformations. However, they do transform under rigid $GL(10,\mathbb{R})$-transformations as tensors of the type indicated by their indices. Their indices are world indices and not flat indices.

5.3.4 "Covariant Derivatives"

The invariant form $\partial\mathcal{V}\,\mathcal{V}^{-1}$ reads

$$\partial\mathcal{V}\,\mathcal{V}^{-1} = \partial H\,H^{-1} + H(\partial T\,T^{-1})H^{-1}. \tag{5.3.22}$$

The first term is the invariant form encountered above in the discussion of the level zero nonlinear sigma model for $SO(10)\backslash GL(10,\mathbb{R})$. So let us focus on the second term. It is clear that $\partial T T^{-1}$ will contain only positive generators at level ≥ 1. So we set, in a manner similar to Equation (5.1.48),

$$\partial T T^{-1} = \sum_{\alpha\in\Delta_+} \sum_s \mathcal{D}\mathcal{A}_{\alpha,s}\, E_{\alpha,s}, \tag{5.3.23}$$

where the sum is over positive roots at levels one and higher and takes into account multiplicities (through the extra index s). The expressions $\mathcal{D}\mathcal{A}_{\alpha,s}$ are linear in the time derivatives $\partial\mathcal{A}$. As before, we call them "covariant derivatives". They are computed by making use of the Baker–Hausdorff formula, as in Section 5.2.1. Explicitly, up to level 3, one finds

$$\dot{T}T^{-1} = \frac{1}{3!}\mathcal{D}\mathcal{A}_{abc}(t)E^{abc} + \frac{1}{6!}\mathcal{D}\mathcal{A}_{a_1\cdots a_6}(t)E^{a_1\cdots a_6} + \frac{1}{9!}\mathcal{D}\mathcal{A}_{a|b_1\cdots b_8}(t)E^{a|b_1\cdots b_8} + \cdots, \tag{5.3.24}$$

with

$$\begin{aligned}
\mathcal{D}\mathcal{A}_{abc}(t) &= \partial\mathcal{A}_{abc}(t), \\
\mathcal{D}\mathcal{A}_{a_1\cdots a_6}(t) &= \partial\mathcal{A}_{a_1\cdots a_6}(t) + 10\mathcal{A}_{[a_1a_2a_3}(t)\partial\mathcal{A}_{a_4a_5a_6]}(t), \\
\mathcal{D}\mathcal{A}_{a|b_1\cdots b_8}(t) &= \partial\mathcal{A}_{a|b_1\cdots b_8}(t) + 42\mathcal{A}_{\langle ab_1b_2}(t)\partial\mathcal{A}_{b_3\cdots b_8\rangle}(t) - 42\partial\mathcal{A}_{\langle ab_1b_2}(t)\mathcal{A}_{b_3\cdots b_8\rangle}(t), \\
&\quad +280\mathcal{A}_{\langle ab_1b_2}(t)\mathcal{A}_{b_3b_4b_5}(t)\partial\mathcal{A}_{b_6b_7b_8\rangle}(t),
\end{aligned} \tag{5.3.25}$$

as computed in [53]. The notation $\langle a_1\cdots a_k\rangle$ denotes projection onto the Young tableaux symmetry carried by the field upon which the covariant derivative acts[1]. It should be stressed that the covariant derivatives $\mathcal{D}\mathcal{A}$ have the same transformation properties under $SO(10)$ (under which they are inert) and $GL(10,\mathbb{R})$ as $\mathcal{A}$ since the $GL(10,\mathbb{R})$ transformations do not depend on time.

5.3.5 The $K(\mathcal{E}_{10})\times\mathcal{E}_{10}$-Invariant Action at Low Levels

The action can now be computed using the bilinear form $(\cdot|\cdot)$ on E_{10},

$$S_{K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}} = \int dt\, n(t)^{-1}\left(\mathcal{P}(t)\middle|\mathcal{P}(t)\right), \tag{5.3.26}$$

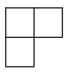
where $\mathcal{P}$ is obtained by projecting orthogonally onto the subalgebra $\mathfrak{k}_{E_{10}}$ by using the generalized transpose,

$$(K^a{}_b)^{\mathcal{T}} = K^b{}_a, \qquad (E^{abc})^{\mathcal{T}} = F_{abc}, \cdots \text{etc.}, \tag{5.3.27}$$

where as above $(\ )^{\mathcal{T}} = -\omega(\ )$ (with ω being the Chevalley involution). We shall compute the action up to, and including, level 3,

$$S_{K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}} = \int dt\, n(t)^{-1}\left(\mathcal{L}_0 + \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \cdots\right). \tag{5.3.28}$$

[1] As an example, consider the projection $P_{\alpha\beta\gamma} \equiv T_{\langle\alpha\beta\gamma\rangle}$ of a three index tensor $T_{\alpha\beta\gamma}$ onto the Young tableaux



.

This projection is given by

$$P_{\alpha\beta\gamma} = \frac{1}{3}(T_{\alpha\beta\gamma} + T_{\beta\alpha\gamma} - T_{\gamma\beta\alpha} - T_{\beta\gamma\alpha}),$$

which clearly satisfies

$$P_{\alpha\beta\gamma} = -P_{\gamma\beta\alpha}, \qquad P_{[\alpha\beta\gamma]} = 0.$$

Note also that $P_{\alpha\beta\gamma} \neq P_{\beta\alpha\gamma}$.

From Equation (5.3.22) and the fact that generators at level zero are orthogonal to generators at levels $\neq 0$, we see that $\mathcal{L}_0$ will be constructed from the level zero part $\dot{H}\,H^{-1}$ and will coincide with the Lagrangian (5.3.14) for the nonlinear sigma model $SO(10)\backslash GL(10,\mathbb{R})$,

$$\mathcal{L}_0 = \frac{1}{4}\left(g^{ac}(t)g^{bd}(t) - g^{ab}(t)g^{cd}(t)\right)\partial g_{ab}(t)\,\partial g_{cd}(t). \tag{5.3.29}$$

To compute the other terms, we use the following trick. The Lagrangian must be a $GL(10,\mathbb{R})$ scalar. One can easily compute it in the frame where $H = 1$, i.e., where the metric g_{ab} is equal to δ_{ab}. One can then covariantize the resulting expression by replacing everywhere δ_{ab} by g_{ab}. To illustrate the procedure consider the level 1 term. One has, for $H = 1$ and at level 1, $\partial\mathcal{V}\,\mathcal{V}^{-1} = \frac{1}{3!}\mathcal{D}\mathcal{A}_{abc}(t)E^{abc}$ and thus, with the same gauge conditions, $\mathcal{P}(t) = \frac{1}{2\cdot 3!}\mathcal{D}\mathcal{A}_{abc}(t)\left(E^{abc} + F^{abc}\right)$ (where we have raised the indices of F_{abc} with δ^{ab}, $F_{123} \equiv F^{123}$ etc). Using $(E^{a_1a_2a_3}|F^{b_1b_2b_3}) = \delta^{a_1b_1}\delta^{a_2b_2}\delta^{a_3b_3}\pm$permutations that make the expression antisymmetric (3! terms; see Section 2.5.4), one then gets $\mathcal{L}_1 = \frac{1}{2\cdot 3!}\mathcal{D}\mathcal{A}_{abc}(t)\,\mathcal{D}\mathcal{A}_{def}(t)\,\delta^{ad}\delta^{be}\delta^{cf}$ in the frame where $g_{ab} = \delta_{ab}$. This yields the level 1 Lagrangian in a general frame,

$$\begin{aligned}\mathcal{L}_1 &= \frac{1}{2\cdot 3!}g^{a_1c_1}g^{a_2c_2}g^{a_3c_3}\mathcal{D}\mathcal{A}_{a_1a_2a_3}(t)\,\mathcal{D}\mathcal{A}_{c_1c_2c_3}(t)\\ &= \frac{1}{2\cdot 3!}\mathcal{D}\mathcal{A}_{a_1a_2a_3}(t)\,\mathcal{D}\mathcal{A}^{a_1a_2a_3}(t).\end{aligned} \tag{5.3.30}$$

By a similar analysis, the level 2 and 3 contributions are

$$\begin{aligned}\mathcal{L}_2 &= \frac{1}{2\cdot 6!}\,\mathcal{D}\mathcal{A}_{a_1\cdots a_6}(t)\,\mathcal{D}\mathcal{A}^{a_1\cdots a_6}(t),\\ \mathcal{L}_3 &= \frac{1}{2\cdot 9!}\,\mathcal{D}\mathcal{A}_{a|b_1\cdots b_8}(t)\,\mathcal{D}\mathcal{A}^{a|b_1\cdots b_8}(t).\end{aligned} \tag{5.3.31}$$

Collecting all terms, the final form of the action for $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ up to and including level $\ell = 3$ is

$$\begin{aligned}S_{K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}} = \int dt\,n(t)^{-1}\Bigg[&\frac{1}{4}\left(g^{ac}(t)g^{bd}(t) - g^{ab}(t)g^{cd}(t)\right)\partial g_{ab}(t)\,\partial g_{cd}(t)\\ &+\frac{1}{2\cdot 3!}\mathcal{D}\mathcal{A}_{a_1a_2a_3}(t)\,\mathcal{D}\mathcal{A}^{a_1a_2a_3}(t) + \frac{1}{2\cdot 6!}\mathcal{D}\mathcal{A}_{a_1\cdots a_6}(t)\,\mathcal{D}\mathcal{A}^{a_1\cdots a_6}(t)\\ &+\frac{1}{2\cdot 9!}\mathcal{D}\mathcal{A}_{a|b_1\cdots b_8}(t)\,\mathcal{D}\mathcal{A}^{a|b_1\cdots b_8}(t) + \cdots\Bigg],\end{aligned} \tag{5.3.32}$$

which agrees with the action found in [53].

5.3.6 The Correspondence

We shall now relate the equations of motion for the $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ sigma model to the equations of motion of eleven-dimensional supergravity. As the precise correspondence is not yet known, we shall here only sketch the main ideas. These work remarkably well at low levels but need unknown ingredients at higher levels.

We have seen that the sigma model for $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ can be consistently truncated level by level. More precisely, one can consistently set equal to zero all covariant derivatives of the fields above a given level and get a reduced system whose solutions are solutions of the full system. We shall show here that the consistent truncations of $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ at levels 0, 1 and 2 yields equations of motion that coincide with the equations of motion of appropriate consistent truncations of eleven-dimensional supergravity, using a prescribed dictionary presented below. We will also show that the correspondence extends up to parts of level 3.

We recall that in the gauge $N^i = 0$ (vanishing shift) and $A_{0bc} = 0$ (temporal gauge), the bosonic fields of eleven-dimensional supergravity are the spatial metric $\mathrm{g}_{ab}(x^0, x^i)$, the lapse $N(x^0, x^i)$ and the spatial components $A_{abc}(x^0, x^i)$ of the vector potential 3-form. The physical field is $F = dA$ and its electric and magnetic components are, respectively, denoted F_{0abc} and F_{abcd}. The electric field involves only time derivatives of $A_{abc}(x^0, x^i)$, while the magnetic field involves spatial gradients.

Levels 0 and 1

If one keeps only levels zero and one, the sigma model action (5.3.32) reduces to

$$\begin{aligned} S[g_{ab}(t), \mathcal{A}_{abc}(t), n(t)] \;=\; & \int dt\, n(t)^{-1} \left[\frac{1}{4}\left(g^{ac}(t)\, g^{bd}(t) - g^{ab}(t)\, g^{cd}(t)\right) \partial g_{ab}(t)\, \partial g_{cd}(t) \right. \\ & \left. + \frac{1}{2\cdot 3!} \partial \mathcal{A}_{a_1a_2a_3}(t)\, \partial \mathcal{A}^{a_1a_2a_3}(t)\right]. \end{aligned} \tag{5.3.33}$$

Consider now the consistent homogeneous truncation of eleven-dimensional supergravity in which the spatial metric, the lapse and the vector potential depend only on time (no spatial gradient). Then the reduced action for this truncation is precisely Equation (5.3.33) provided one makes the identification $t = x^0$ and

$$g_{ab}(t) = \mathrm{g}_{ab}(t), \tag{5.3.34}$$

$$\mathcal{A}_{abc}(t) = A_{abc}(t), \tag{5.3.35}$$

$$n(t) = \frac{N(t)}{\sqrt{\mathrm{g}(t)}} \tag{5.3.36}$$

(see, for instance, [49]). Also the Hamiltonian constraints (the only one left) coincide. Thus, there is a perfect match between the sigma model truncated at level one and supergravity "reduced on a 10-torus". If one were to drop level one, one would find perfect agreement with pure gravity. In the following, we shall make the gauge choice $N = \sqrt{\mathrm{g}}$, equivalent to $n = 1$.

Level 2

At levels 0 and 1, the supergravity fields g_{ab} and A_{abc} depend only on time. When going beyond this truncation, one needs to introduce some spatial gradients. Level 2 introduces spatial gradients of a very special type, namely allows for a homogeneous magnetic field. This means that A_{abc} acquires a space dependence, more precisely, a linear one (so that its gradient does not depend on x). However, because there is no room for x-dependence on the sigma model side, where the only independent variable is t, we shall use the trick to describe the magnetic field in terms of a dual potential $A_{a_1\cdots a_6}$. Thus, there is a close interplay between duality, the sigma model formulation, and the introduction of spatial gradients.

There is no tractable, fully satisfactory variational formulation of eleven-dimensional supergravity where both the 3-form potential and its dual appear as independent variables in the action, with a quadratic dependence on the time derivatives (this would be double-counting, unless an appropriate self-duality condition is imposed [127,128]). This means that from now on, we shall not compare the actions of the sigma model and of supergravity but, rather, only their respective equations of motion. As these involve the electromagnetic field and not the potential, we rewrite the correspondence found above at levels 0 and 1 in terms of the metric and the electromagnetic field as

$$\begin{aligned} g_{ab}(t) &= \mathrm{g}_{ab}(t), \\ \mathcal{D}\mathcal{A}_{abc}(t) &= F_{0abc}(t). \end{aligned} \tag{5.3.37}$$

The equations of motion for the nonlinear sigma model, obtained from the variation of the Lagrangian Equation (5.3.32), truncated at level two, read explicitly

$$\begin{aligned} \frac{1}{2}\partial\left(n^{-1}g^{ac}\,\partial g_{cb}\right) &= \frac{n^{-1}}{4}\left(\mathcal{D}\mathcal{A}^{ac_1c_2}\,\mathcal{D}\mathcal{A}_{bc_1c_2} - \frac{1}{9}\delta^a{}_b\mathcal{D}\mathcal{A}^{c_1c_2c_3}\,\mathcal{D}\mathcal{A}_{c_1c_2c_3}\right) \\ &\quad + \frac{n^{-1}}{2\cdot 5!}\left(\mathcal{D}\mathcal{A}^{ac_1\cdots c_5}\,\mathcal{D}\mathcal{A}_{bc_1\cdots c_5} - \frac{1}{9}\delta^a{}_b\mathcal{D}\mathcal{A}^{c_1\cdots c_6}\,\mathcal{D}\mathcal{A}_{c_1\cdots c_6}\right), \\ \partial\left(n^{-1}\mathcal{D}\mathcal{A}^{a_1a_2a_3}\right) &= -\frac{1}{3!}\,n^{-1}\mathcal{D}\mathcal{A}^{a_1\cdots a_6}\,\mathcal{D}\mathcal{A}_{a_4a_5a_6}, \\ \partial\left(n^{-1}\mathcal{D}\mathcal{A}^{a_1\cdots a_6}\right) &= 0. \end{aligned} \tag{5.3.38}$$

In addition, we have the constraint obtained by varying n,

$$\begin{aligned} (\mathcal{P}|\mathcal{P}) &= \frac{1}{4}\left(g^{ac}\,g^{bd} - g^{ab}\,g^{cd}\right)\partial g_{ab}\,\partial g_{cd} \\ &\quad + \frac{1}{2\cdot 3!}\mathcal{D}\mathcal{A}^{a_1a_2a_3}\,\mathcal{D}\mathcal{A}_{a_1a_2a_3} + \frac{1}{2\cdot 6!}\mathcal{D}\mathcal{A}^{a_1\cdots a_6}\,\mathcal{D}\mathcal{A}_{a_1\cdots a_6} \\ &= 0. \end{aligned} \tag{5.3.39}$$

On the supergravity side, we truncate the equations to metrics $\mathrm{g}_{ab}(t)$ and electromagnetic fields $F_{0abc}(t)$, $F_{abcd}(t)$ that depend only on time. We take, as in Section 3.1.6), the spacetime metric to be of the form

$$ds^2 = -N^2(t)\,dt^2 + \mathrm{g}_{ab}(t)\,dx^a\,dx^b, \tag{5.3.40}$$

but now with $x^0 \equiv t$. In the following we use Greek letters $\lambda, \sigma, \rho, \cdots$ to denote eleven-dimensional spacetime indices, and Latin letters $a, b, c, \cdots$ to denote ten-dimensional spatial indices.

The equations of motion and the Hamiltonian constraint for eleven-dimensional supergravity have been explicitly written in [49], so they can be expediently compared with the equations of motion of the sigma model. The dynamical equations for the metric read

$$\begin{aligned} \frac{1}{2}\partial\left(\sqrt{\mathrm{g}}N^{-1}\mathrm{g}^{ac}\partial\mathrm{g}_{cb}\right) &= \frac{1}{12}N\sqrt{\mathrm{g}}F^{a\rho\sigma\tau}F_{b\rho\sigma\tau} - \frac{1}{144}N\sqrt{\mathrm{g}}\,\delta^a{}_b\,F^{\lambda\rho\sigma\tau}F_{\lambda\rho\sigma\tau} \\ &= \frac{1}{4}N^{-1}\sqrt{\mathrm{g}}F^{0ac_1c_2}F_{0bc_1c_2} - \frac{1}{36}N^{-1}\sqrt{\mathrm{g}}\,\delta^a{}_b\,F^{0c_1c_2c_3}F_{0c_1c_2c_3} \\ &\quad + \frac{1}{12}N\sqrt{\mathrm{g}}F^{ac_1c_2c_3}F_{bc_1c_2c_3} - \frac{1}{144}N\sqrt{\mathrm{g}}\,\delta^a{}_b\,F^{c_1c_2c_3c_4}F_{c_1c_2c_3c_4}, \end{aligned} \tag{5.3.41}$$

and for the electric and magnetic fields we have, respectively, the equations of motion and the Bianchi identity,

$$\begin{aligned} \partial\left(F^{0abc}N\sqrt{\mathrm{g}}\right) &= \frac{1}{144}\varepsilon^{0abcd_1d_2d_3e_1e_2e_3e_4}F_{0d_1d_2d_3}F_{e_1e_2e_3e_4}, \\ \partial F_{a_1a_2a_3a_4} &= 0. \end{aligned} \tag{5.3.42}$$

Furthermore we have the Hamiltonian constraint

$$\frac{1}{4}\left(\mathrm{g}^{ac}\mathrm{g}^{bd}-\mathrm{g}^{ab}\mathrm{g}^{cd}\right)\partial\mathrm{g}_{ab}\,\partial\mathrm{g}_{cd}+\frac{1}{12}F^{0abc}F_{0abc}+\frac{1}{48}N^2F^{abcd}F_{abcd}=0. \tag{5.3.43}$$

(We shall not consider the other constraints here; see remarks as at the end of this section.)

One finds again perfect agreement between the sigma model equations, Equation (5.3.38) and (5.3.39), and the equations of eleven-dimensional supergravity, (5.3.41) and (5.3.43), provided one extends the above dictionary through [53]

$$\mathcal{D}\mathcal{A}^{a_1\cdots a_6}(t)=-\frac{1}{4!}\varepsilon^{a_1\cdots a_6b_1b_2b_3b_4}F_{b_1b_2b_3b_4}(t). \tag{5.3.44}$$

This result appears to be quite remarkable, because the Chern–Simons term in (5.3.42) is in particular reproduced with the correct coefficient, which in eleven-dimensional supergravity is fixed by invoking supersymmetry.

Level 3

Level 3 should correspond to the introduction of further controlled spatial gradients, this time for the metric. Because there is no room for spatial derivatives as such on the sigma model side, the trick is again to introduce a dual graviton field. When this dual graviton field is non-zero, the metric does depend on the spatial coordinates.

Satisfactory dual formulations of non-linearized gravity do not exist. At the linearized level, however, the problem is well understood since the pioneering work by Curtright [129]. In eleven spacetime dimensions, the dual graviton field is *described precisely by a tensor* $\mathcal{A}_{a|b_1\cdots b_8}$ *with the mixed symmetry of the Young tableau* $[1,0,0,0,0,0,0,1,0]$ *appearing at level 3 in the sigma model description.* Exciting this field, i.e., assuming $\mathcal{D}\mathcal{A}_{a|b_1\cdots b_8}\neq 0$ amounts to introducing spatial gradients for the metric – and, for that matter, for the other fields as well – as follows. Instead of considering fields that are homogeneous on a torus, one considers fields that are homogeneous on non-Abelian group manifolds. This introduces spatial gradients (in coordinate frames) in a well controlled manner.

Let θ^a be the group invariant one-forms, with structure equations

$$d\theta^a=\frac{1}{2}C^a{}_{bc}d\theta^b\wedge d\theta^c. \tag{5.3.45}$$

We shall assume that $C^a{}_{ac}=0$ ("Bianchi class A"). Truncation at level 3 assumes that the metric and the electric and magnetic fields depend only on time in this frame and that the $C^a{}_{bc}$ are constant (corresponding to a group). The supergravity equations have been written in that case in [49] and can be compared with the sigma model equations. There is almost a complete match between both sets of equations provided one extends the dictionary at level 3 through

$$\mathcal{D}\mathcal{A}^{a|b_1\cdots b_8}(t) = \frac{3}{2}\varepsilon^{b_1\cdots b_8 cd}C^a{}_{cd} \tag{5.3.46}$$

(with the equations of motion of the sigma model implying indeed that $\mathcal{D}\mathcal{A}^{a|b_1\cdots b_8}$ does not depend on time). Note that to define an invertible mapping between the level three fields and the $C^a{}_{bc}$, it is important that $C^a{}_{bc}$ be traceless; there is no "room" on level three on the sigma model side to incorporate the trace of $C^a{}_{bc}$.

With this correspondence, the match works perfectly for real roots up to, and including, level three. However, it fails for fields associated with imaginary roots (level 3 is the first time imaginary roots appear, at height 30) [53]. In fact, the terms that match correspond to "$SL(10,\mathbb{R})$-covariantized E_8", i.e., to fields associated with roots of E_8 and their images under the Weyl group of $SL(10,\mathbb{R})$.

Since the match between the sigma model equations and supergravity fails at level 3 under the present line of investigation, we shall not provide the details but refer instead to [53] for more information. The correspondence up to level 3 was also checked in [89] through a slightly different approach, making use of a formulation with local frames, i.e., using local flat indices rather than global indices as in the present treatment.

Let us note here that higher level fields of E_{10}, corresponding to imaginary roots, have been considered from a different point of view in [40], where they were associated with certain brane configurations (see also [130, 131]).

The Dictionary

One may view the above failure at level 3 as a serious flaw to the sigma model approach to exhibiting the E_{10} symmetry[2]. Let us, however, be optimistic for a moment and assume that these problems will somehow get resolved, perhaps by changing the dictionary or by including higher order terms. So, let us proceed.

What would be the meaning of the higher level fields? As discussed in Section 5.3.7, there are indications that fields at higher levels contain higher order spatial gradients and therefore enable us to reconstruct completely, through something similar to a Taylor expansion, the most general field configuration from the fields at a given spatial point.

From this point of view, the relation between the supergravity degrees of freedom $\mathrm{g}_{ij}(t,x)$ and $F_{(4)}(t,x) = dA_{(3)}(t,x)$ would be given, at a specific spatial point $x = \mathbf{x}_0$ and in a suitable spatial frame $\theta^a(x)$ (that would also depend on x), by the following "dictionary":

$$\begin{aligned} g_{ab}(t) &= \mathrm{g}_{ab}(t,\mathbf{x}_0), \\ \mathcal{D}\mathcal{A}_{abc}(t) &= F_{tabc}(t,\mathbf{x}_0), \\ \mathcal{D}\mathcal{A}^{a_1\cdots a_6}(t) &= -\frac{1}{4!}\varepsilon^{a_1\cdots a_6 bcde}F_{bcde}(t,\mathbf{x}_0), \\ \mathcal{D}\mathcal{A}^{a|b_1\cdots b_8}(t) &= \frac{3}{2}\varepsilon^{b_1\cdots b_8 cd}C^a{}_{cd}(\mathbf{x}_0), \end{aligned} \tag{5.3.47}$$

[2] This does not exclude that other approaches would be successful. That E_{10}, or perhaps E_{11}, does encode a lot of information about M-theory is a fact, but that this should be translated into a sigma model reformulation of the theory appears to be questionable.

which reproduces in the homogeneous case what we have seen up to level 3.

This correspondence goes far beyond that of the algebraic description of the BKL-limit in terms of Weyl reflections in the simple roots of a Kac–Moody algebra. Indeed, the dynamics of the billiard is controlled entirely by the walls associated with *simple* roots and thus does not transcend height one. Here, we go to a much higher height and successfully extend (unfortunately incompletely) the intriguing connection between eleven-dimensional supergravity and E_{10}.

5.3.7 Higher Levels and Spatial Gradients

We have seen that the correspondence between the $\mathcal{E}_{10}$-invariant sigma model and eleven-dimensional supergravity fails when we include spatial gradients beyond first order. It is nevertheless believed that the information about spatial gradients is somehow encoded within the algebraic description: One idea is that space is "smeared out" among the infinite number of fields contained in $\mathcal{E}_{10}$ and it is for this reason that a direct dictionary for the inclusion of spatial gradients is difficult to find. If true, this would imply that we can view the level expansion on the algebraic side as reflecting a kind of "Taylor expansion" in spatial gradients on the supergravity side. Below we discuss some speculative ideas about how such a correspondence could be realized in practice.

The "Gradient Conjecture"

One intriguing suggestion put forward in [53] was that fields associated to certain "affine representations" of E_{10} could be interpreted as spatial derivatives acting on the level one, two and three fields, thus providing a direct conjecture for how space "emerges" through the level decomposition of E_{10}. The representations in question are those for which the Dynkin label associated with the overextended root of E_{10} vanishes, and hence these representations are realized also in a level decomposition of the regular E_9-subalgebra obtained by removing the overextended node in the Dynkin diagram of E_{10}.

The affine representations were discussed in Section 2.5 and we recall that they are given in terms of three infinite towers of generators, with the following $\mathfrak{sl}(10,\mathbb{R})$-tensor structures,

$$E^{a_1a_2a_3}{}_{b_1\cdots b_k}, \qquad E^{a_1\cdots a_6}{}_{b_1\cdots b_k}, \qquad E^{a_1|a_2\cdots a_9}{}_{b_1\cdots b_k}, \tag{5.3.48}$$

where the upper indices have the same Young tableau symmetries as the $\ell = 1, 2$ and 3 representations, while the lower indices are all completely symmetric. In the sigma model these generators of $\mathcal{E}_{10}$ are parametrized by fields exhibiting the same index structure, i.e., $\mathcal{A}_{a_1a_2a_2}{}^{b_1\cdots b_k}(t)$, $\mathcal{A}_{a_1\cdots a_6}{}^{b_1\cdots b_k}(t)$ and $\mathcal{A}_{a_1|a_2\cdots a_9}{}^{b_1\cdots b_k}(t)$.

The idea is now that the three towers of fields have precisely the right index structure to be interpreted as spatial gradients of the low level fields

$$\begin{aligned}
\mathcal{A}_{a_1a_2a_2}{}^{b_1\cdots b_k}(t) &= \partial^{b_1}\cdots\partial^{b_k}\mathcal{A}_{a_1a_2a_3}(t),\\
\mathcal{A}_{a_1\cdots a_6}{}^{b_1\cdots b_k}(t) &= \partial^{b_1}\cdots\partial^{b_k}\mathcal{A}_{a_1\cdots a_6}(t),\\
\mathcal{A}_{a_1|a_2\cdots a_9}{}^{b_1\cdots b_k}(t) &= \partial^{b_1}\cdots\partial^{b_k}\mathcal{A}_{a_1|a_2\cdots a_9}(t).
\end{aligned} \tag{5.3.49}$$

Although appealing and intuitive as it is, this conjecture is difficult to prove or to check explicitly, and not much progress in this direction has been made since the original proposal.

Let us point out, however, that recently [68] this problem was attacked from a rather different point of view with some very interesting results, indicating that the gradient conjecture may need modification.

6

Geometric Configurations and Cosmological Solutions from $\mathcal{E}_{10}$

In this chapter we will show that the low level equivalence between the $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ sigma model and eleven-dimensional supergravity can be put to practical use for finding exact solutions. We will restrict our analysis to a cosmological sector where it is assumed that all spatial gradients can be neglected so that all fields depend only on time. Moreover, we impose diagonality of the spatial metric. These conditions must of course be compatible with the equations of motion; if the conditions are imposed initially, they should be preserved by the time evolution.

A large class of solutions to eleven-dimensional supergravity preserving these conditions were found in [49]. These solutions have zero magnetic field but have a restricted number of electric field components turned on. Surprisingly, it was found that such solutions have an elegant interpretation in terms of so called *geometric configurations*, denoted (n_m, g_3), of n points and g lines (with $n \leq 10$) drawn on a plane with certain pre-determined rules. That is, for each geometric configuration (whose definition is recalled below) one can associate a diagonal solution with some non-zero electric field components F_{tijk}, determined by the configuration. In this section we re-examine this result from the point of view of the sigma model based on $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$.

We show, following [122], that each configuration (n_m, g_3) encodes information about a (regular) subalgebra $\bar{\mathfrak{g}}$ of E_{10}, and the supergravity solution associated to the configuration $(n_m g_3)$ can be obtained by restricting the $\mathcal{E}_{10}$-sigma model to the subgroup $\bar{G}$ whose Lie algebra is $\bar{\mathfrak{g}}$. Therefore, we will here make use of both the level truncation and the subgroup truncation simultaneously; first by truncating to a certain level and then by restricting to the relevant $\bar{\mathfrak{g}}$-algebra generated by a subset of the representations at this level.

This chapter is based on **Paper I**, written in collaboration with Marc Henneaux, Mauricio Leston and Philippe Spindel, as well as on **Paper III**, written in collaboration with Marc Henneaux and Philippe Spindel.

6.1 Bianchi I Models and Eleven-Dimensional Supergravity

On the supergravity side, we will restrict the metric and the electromagnetic field to depend on time only,

$$\begin{aligned} ds^2 &= -N^2(t)\,dt^2 + \mathrm{g}_{ab}(t)\,dx^a\,dx^b, \\ F_{\lambda\rho\sigma\tau} &= F_{\lambda\rho\sigma\tau}(t). \end{aligned} \tag{6.1.1}$$

Recall from Section 5.3 that with these ansätze the dynamical equations of motion of eleven-dimensional supergravity reduce to [49]

$$\frac{1}{2}\partial\left(\sqrt{\mathrm{g}}N^{-1}\mathrm{g}^{ac}\partial\mathrm{g}_{cb}\right) = \frac{1}{12}N\sqrt{\mathrm{g}}F^{a\rho\sigma\tau}F_{b\rho\sigma\tau} - \frac{1}{144}N\sqrt{\mathrm{g}}\,\delta^a{}_b\,F^{\lambda\rho\sigma\tau}F_{\lambda\rho\sigma\tau}, \tag{6.1.2}$$

$$\partial\left(F^{tabc}N\sqrt{\mathrm{g}}\right) = \frac{1}{144}\varepsilon^{tabcd_1d_2d_3e_1e_2e_3e_4}F_{td_1d_2d_3}F_{e_1e_2e_3e_4}, \tag{6.1.3}$$

$$\partial F_{a_1a_2a_3a_4} = 0. \tag{6.1.4}$$

This corresponds to the truncation of the sigma model at level 2 which, as we have seen, completely matches the supergravity side. We also defined $\partial \equiv \partial_t$ as in Section 5.3. Furthermore we have the following constraints,

$$\frac{1}{4}\left(\mathrm{g}^{ac}\mathrm{g}^{bd} - \mathrm{g}^{ab}\mathrm{g}^{cd}\right)\partial\mathrm{g}_{ab}\,\partial\mathrm{g}_{cd} + \frac{1}{12}F^{tabc}F_{tabc} + \frac{1}{48}N^2F^{abcd}F_{abcd} = 0, \tag{6.1.5}$$

$$\frac{1}{6}NF^{tbcd}F_{abcd} = 0, \tag{6.1.6}$$

$$\varepsilon^{tabc_1c_2c_3c_4d_1d_2d_3d_4}F_{c_1c_2c_3c_4}F_{d_1d_2d_3d_4} = 0, \tag{6.1.7}$$

which are, respectively, the Hamiltonian constraint, momentum constraint and Gauss' law. Note that Greek indices $\alpha, \beta, \gamma, \cdots$ correspond to the full eleven-dimensional spacetime, while Latin indices $a, b, c, \cdots$ correspond to the ten-dimensional spatial part.

We will further take the metric to be purely time-dependent and diagonal,

$$ds^2 = -N^2(t)\,dt^2 + \sum_{i=1}^{10} a_i^2(t)(dx^i)^2. \tag{6.1.8}$$

This form of the metric has manifest invariance under the ten distinct spatial reflections

$$\begin{aligned} x^j &\to -x^j, \\ x^{i\neq j} &\to x^{i\neq j}, \end{aligned} \tag{6.1.9}$$

and in order to ensure compatibility with the Einstein equations, the energy-momentum tensor of the 4-form field strength must also be diagonal.

6.1.1 Diagonal Metrics and Geometric Configurations

Assuming zero magnetic field (this restriction will be lifted below), one way to achieve diagonality of the energy-momentum tensor is to assume that the non-vanishing components of the electric field $F^{\perp abc} = N^{-1}F_{tabc}$ are determined by *geometric configurations* (n_m, g_3) with $n \leq 10$ [49].

A geometric configuration (n_m, g_3) is a set of n points and g lines with the following incidence rules [132–134]:

i. Each line contains three points.

ii. Each point is on m lines.

iii. Two points determine at most one line.

It follows that two lines have at most one point in common. It is an easy exercise to verify that $mn = 3g$. An interesting question is whether the lines can actually be realized as straight lines in the (real) plane, but, for our purposes, it is not necessary that it should be so; the lines can be bent.

Let (n_m, g_3) be a geometric configuration with $n \leq 10$ points. We number the points of the configuration $1, \cdots, n$. We associate to this geometric configuration a pattern of electric field components $F^{\perp abc}$ with the following property: $F^{\perp abc}$ can be non-zero only if the triple (a, b, c) is a line of the geometric configuration. If it is not, we take $F^{\perp abc} = 0$. It is clear that this property is preserved in time by the equations of motion (in the absence of magnetic field). Furthermore, because of Rule iii above, the products $F^{\perp abc} F^{\perp a'b'c'} g_{bb'} g_{cc'}$ vanish when $a \neq a'$ so that the energy-momentum tensor is diagonal.

6.2 Geometric Configurations and Regular Subalgebras of E_{10}

We prove here that the conditions on the electric field embodied in the geometric configurations (n_m, g_3) have a direct Kac–Moody algebraic interpretation. They simply correspond to a consistent truncation of the E_{10} nonlinear sigma model to a $\bar{\mathfrak{g}}$ nonlinear sigma model, where $\bar{\mathfrak{g}}$ is a rank g Kac–Moody subalgebra of E_{10} (or a quotient of such a Kac–Moody subalgebra by an appropriate ideal when the relevant Cartan matrix has vanishing determinant), with three crucial properties: (i) It is regularly embedded in E_{10} (see Section 2.2.3 for the definition of regular subalgebras), (ii) it is generated by electric roots only, and (iii) every node P in its Dynkin diagram $\mathbb{D}_{\bar{\mathfrak{g}}}$ is linked to a number k of nodes that is independent of P (but depend on the algebra). We find that the Dynkin diagram $\mathbb{D}_{\bar{\mathfrak{g}}}$ of $\bar{\mathfrak{g}}$ is the *line incidence diagram* of the geometric configuration (n_m, g_3), in the sense that (i) each line of (n_m, g_3) defines a node of $\mathbb{D}_{\bar{\mathfrak{g}}}$, and (ii) two nodes of $\mathbb{D}_{\bar{\mathfrak{g}}}$ are connected by a single bond iff the corresponding lines of (n_m, g_3) have no point in common. This defines a geometric duality between a configuration (n_m, g_3) and its associated Dynkin diagram $\mathbb{D}_{\bar{\mathfrak{g}}}$. In the following we shall therefore refer to configurations and Dynkin diagrams related in this way as *dual*.

None of the algebras $\bar{\mathfrak{g}}$ relevant to the truncated models turn out to be hyperbolic: They can be finite, affine, or Lorentzian with infinite-volume Weyl chamber. Because of this, the solutions are non-chaotic. After a finite number of collisions, they settle asymptotically into a definite Kasner regime (both in the future and in the past).

6.2.1 General Considerations

In order to match diagonal Bianchi I cosmologies with the sigma model, one must truncate the $\mathcal{E}_{10}/\mathcal{K}(\mathcal{E}_{10})$ action in such a way that the sigma model metric g_{ab} is diagonal. This will be the case if the subalgebra $\bar{\mathfrak{g}}$ to which one truncates has no generator $K^i{}_j$ with $i \neq j$. The off-diagonal components of the metric are precisely the exponentials of the associated sigma model fields. The set of simple roots of $\bar{\mathfrak{g}}$ should therefore not contain any root at level zero.

Consider "electric" regular subalgebras of E_{10}, for which the simple roots are all at level one, where the 3-form electric field variables live. These roots can be parametrized by three indices corresponding to the indices of the electric field, with $i_1 < i_2 < i_3$. We denote them $\alpha_{i_1 i_2 i_3}$. For instance, $\alpha_{123} \equiv \alpha_{10}$. In terms of the β-parametrization of [44, 51], one has $\alpha_{i_1 i_2 i_3} = \beta^{i_1} + \beta^{i_2} + \beta^{i_3}$.

Now, for $\bar{\mathfrak{g}}$ to be a regular subalgebra, it must fulfill, as we have seen, the condition that the difference between any two of its simple roots is not a root of E_{10}: $\alpha_{i_1 i_2 i_3} - \alpha_{i'_1 i'_2 i'_3} \notin \Phi_{E_{10}}$ for any pair $\alpha_{i_1 i_2 i_3}$ and $\alpha_{i'_1 i'_2 i'_3}$ of simple roots of $\bar{\mathfrak{g}}$. But one sees by inspection of the commutator of $E^{i_1 i_2 i_3}$ with $F_{i'_1 i'_2 i'_3}$ in Equation (2.5.78) that $\alpha_{i_1 i_2 i_3} - \alpha_{i'_1 i'_2 i'_3}$ is a root of E_{10} if and only if the sets $\{i_1, i_2, i_3\}$ and $\{i'_1, i'_2, i'_3\}$ have exactly two points in common. For instance, if $i_1 = i'_1$, $i_2 = i'_2$ and $i_3 \neq i'_3$, the commutator of $E^{i_1 i_2 i_3}$ with $F_{i'_1 i'_2 i'_3}$ produces the off-diagonal generator ${K^{i_3}}_{i'_3}$ corresponding to a level zero root of E_{10}. In order to fulfill the required condition, one must avoid this case, i.e., one must choose the set of simple roots of the electric regular subalgebra $\bar{\mathfrak{g}}$ in such a way that given a pair of indices (i_1, i_2), there is at most one i_3 such that the root α_{ijk} is a simple root of $\bar{\mathfrak{g}}$, with (i, j, k) being the re-ordering of (i_1, i_2, i_3) such that $i < j < k$.

To each of the simple roots $\alpha_{i_1 i_2 i_3}$ of $\bar{\mathfrak{g}}$, one can associate the line (i_1, i_2, i_3) connecting the three points i_1, i_2 and i_3. If one does this, one sees that the above condition is equivalent to the following statement: *The set of points and lines associated with the simple roots of* $\bar{\mathfrak{g}}$ *must fulfill the third rule defining a geometric configuration, namely, that two points determine at most one line.* Thus, this geometric condition has a nice algebraic interpretation in terms of regular subalgebras of E_{10}.

The first rule, which states that each line contains 3 points, is a consequence of the fact that the E_{10}-generators at level one are the components of a 3-index antisymmetric tensor. The second rule, that each point is on m lines, is less fundamental from the algebraic point of view since it is not required to hold for $\bar{\mathfrak{g}}$ to be a regular subalgebra. It was imposed in [49] in order to allow for solutions isotropic in the directions that support the electric field. We keep it here as it yields interesting structure.

6.2.2 Incidence Diagrams and Dynkin Diagrams

We have just shown that each geometric configuration (n_m, g_3) with $n \leq 10$ defines a regular subalgebra $\bar{\mathfrak{g}}$ of E_{10}. In order to determine what this subalgebra $\bar{\mathfrak{g}}$ is, one needs, according to the theorem recalled in Section 2.2.3, to compute the Cartan matrix

$$C = [C_{i_1 i_2 i_3, i'_1 i'_2 i'_3}] = \left[\left(\alpha_{i_1 i_2 i_3} | \alpha_{i'_1 i'_2 i'_3}\right)\right] \tag{6.2.1}$$

(the real roots of E_{10} have length squared equal to 2). According to that same theorem, the algebra $\bar{\mathfrak{g}}$ is then just the rank g Kac–Moody algebra with Cartan matrix C, unless C has zero determinant, in which case $\bar{\mathfrak{g}}$ might be the quotient of that algebra by a nontrivial ideal.

Using for instance the root parametrization of [44, 51] and the expression of the scalar product in terms of this parametrization, one easily verifies that the scalar product is equal to

$$\left(\alpha_{i_1 i_2 i_3} | \alpha_{i'_1 i'_2 i'_3}\right) = \begin{cases} 2 & \text{if all three indices coincide,} \\ 1 & \text{if two and only two indices coincide,} \\ 0 & \text{if one and only one index coincides,} \\ -1 & \text{if no indices coincide.} \end{cases} \quad (6.2.2)$$

The second possibility does not arise in our case since we deal with geometric configurations. For completeness, we also list the scalar products of the electric roots α_{ijk} $(i < j < k)$ with the symmetry roots $\alpha_{\ell m}$ $(\ell < m)$ associated with the raising operators $K^m{}_\ell$:

$$(\alpha_{ijk} | \alpha_{\ell m}) = \begin{cases} -1 & \text{if } \ell \in \{i,j,k\} \text{ and } m \notin \{i,j,k\}, \\ 0 & \text{if } \{\ell, m\} \subset \{i,j,k\} \text{ or } \{\ell, m\} \cap \{i,j,k\} = \emptyset, \\ 1 & \text{if } \ell \notin \{i,j,k\} \text{ and } m \in \{i,j,k\}, \end{cases} \quad (6.2.3)$$

as well as the scalar products of the symmetry roots among themselves,

$$(\alpha_{ij} | \alpha_{\ell m}) = \begin{cases} -1 & \text{if } j = \ell \text{ or } i = m, \\ 0 & \text{if } \{\ell, m\} \cap \{i,j\} = \emptyset, \\ 1 & \text{if } i = \ell \text{ or } j \neq m, \\ 2 & \text{if } \{\ell, m\} = \{i,j\}. \end{cases} \quad (6.2.4)$$

Given a geometric configuration (n_m, g_3), one can associate with it a "line incidence diagram" that encodes the incidence relations between its lines. To each line of (n_m, g_3) corresponds a node in the incidence diagram. Two nodes are connected by a single bond if and only if they correspond to lines with no common point ("parallel lines"). Otherwise, they are not connected[1]. By inspection of the above scalar products, we come to the important conclusion that *the Dynkin diagram of the regular, rank g, Kac–Moody subalgebra $\bar{\mathfrak{g}}$ associated with the geometric configuration (n_m, g_3) is just its line incidence diagram.* We shall call the Kac–Moody algebra $\bar{\mathfrak{g}}$ the algebra "dual" to the geometric configuration (n_m, g_3).

Because the geometric configurations have the property that the number of lines through any point is equal to a constant m, the number of lines parallel to any given line is equal to a number k that depends only on the configuration and not on the line. This is in fact true in general and not only for $n \leq 10$ as can be seen from the following argument. For a configuration with n points, g lines and m lines through each point, any given line Δ admits $3(m-1)$ true secants, namely, $(m-1)$ through each of its points[2]. By definition, these secants are all distinct since none of the lines that Δ intersects at one of its points, say P, can coincide with a line that it intersects at another of its points, say P', since the only line joining P to P' is Δ itself. It follows that the total number of lines that Δ intersects is the number of true secants plus Δ itself, i.e., $3(m-1)+1$. As a consequence, each line in the configuration admits $k = g - [3(m-1)+1]$ parallel lines, which is then reflected by the fact that each node in the associated Dynkin diagram has the same number k of adjacent nodes.

[1] One may also consider a point incidence diagram defined as follows: The nodes of the point incidence diagram are the points of the geometric configuration. Two nodes are joined by a single bond if and only if there is no straight line connecting the corresponding points. The point incidence diagrams of the configurations $(9_3, 9_3)$ are given in [133]. For these configurations, projective duality between lines and points lead to identical line and point incidence diagrams. Unless otherwise stated, the expression "incidence diagram" will mean "line incidence diagram".

[2] A true secant is here defined as a line, say Δ', distinct from Δ and with a non-empty intersection with Δ.

6.3 Cosmological Solutions With Electric Flux

Let us now make use of these considerations to construct some explicit supergravity solutions. We begin by analyzing the simplest configuration $(3_1, 1_3)$, of three points and one line. It is displayed in Figure 6.1. This case is the only possible configuration for $n = 3$.

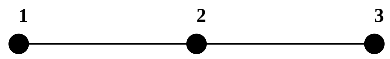


Figure 6.1: $(3_1, 1_3)$: The only allowed configuration for $n = 3$.

This example also exhibits some subtleties associated with the Hamiltonian constraint and the ensuing need to extend $\bar{\mathfrak{g}}$ when the algebra dual to the geometric configuration is finite-dimensional. We will come back to this issue below.

6.3.1 General Discussion

In light of our discussion, considering the geometric configuration $(3_1, 1_3)$ is equivalent to turning on only the component $\mathcal{A}_{123}(t)$ of the 3-form that parametrizes the generator E^{123} in the coset representative $\mathcal{V}(t) \in K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$. Moreover, in order to have the full coset description, we must also turn on the diagonal metric components corresponding to the Cartan generator $h = [E^{123}, F_{123}]$. The algebra has thus basis $\{e, f, h\}$ with

$$e \equiv E^{123}, \qquad f \equiv F_{123}, \qquad h = [e, f] = -\frac{1}{3} \sum_{a \neq 1,2,3} K^a{}_a + \frac{2}{3}(K^1{}_1 + K^2{}_2 + K^3{}_3), \tag{6.3.1}$$

where the form of h followed directly from the general commutator between E^{abc} and F_{def} in Section 2.5. The Cartan matrix is just (2) and is nondegenerate. It defines an $A_1 = \mathfrak{sl}(2, \mathbb{R})$ regular subalgebra. The Chevalley–Serre relations, which are guaranteed to hold according to the general argument, are easily verified. The configuration $(3_1, 1_3)$ is thus dual to A_1,

$$\mathfrak{g}_{(3_1,3_1)} = A_1. \tag{6.3.2}$$

This A_1 algebra is simply the $\mathfrak{sl}(2, \mathbb{R})$-algebra associated with the simple root α_{10}. Because the Killing form of A_1 restricted to the Cartan subalgebra $\mathfrak{h}_{A_1} = \mathbb{R}h$ is positive definite, one cannot find a solution of the Hamiltonian constraint if one turns on only the fields corresponding to A_1. One needs to enlarge A_1 (at least) by a one-dimensional subalgebra $\mathbb{R}l$ of $\mathfrak{h}_{E_{10}}$ that is timelike. As will be discussed further below, we take for l the Cartan element $K^4{}_4 + K^5{}_5 + K^6{}_6 + K^7{}_7 + K^8{}_8 + K^9{}_9 + K^{10}{}_{10}$, which ensures isotropy in the directions not supporting the electric field. Thus, the appropriate regular subalgebra of E_{10} in this case is $A_1 \oplus \mathbb{R}l$.

The need to enlarge the algebra A_1 was discussed in the paper [135] where a group theoretical interpretation of some cosmological solutions of eleven-dimensional supergravity was given. In that paper, it was also observed that $\mathbb{R}l$ can be viewed as the Cartan subalgebra of the (non-regularly embedded) subalgebra A_1 associated with an imaginary root at level 21, but since the corresponding field is not excited, the relevant subalgebra is really $\mathbb{R}l$.

6.3.2 The Solution

In order to make the above discussion a little less abstract we now show how to obtain the relevant supergravity solution by solving the $\mathcal{E}_{10}$-sigma model equations of motion and then translating these, using the dictionary from Section 5.3.6, to supergravity solutions. For this particular example the analysis was done in [135].

In order to better understand the role of the timelike generator $l \in \mathfrak{h}$ we begin the analysis by omitting it. The truncation then amounts to considering the coset representative

$$\mathcal{V}(t) = e^{\phi(t)h}\, e^{\mathcal{A}_{123}(t)E^{123}} \in K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}. \tag{6.3.3}$$

The projection $\mathcal{P}(t)$ onto the coset becomes

$$\begin{aligned}\mathcal{P}(t) &= \frac{1}{2}\Big[\partial\mathcal{V}(t)\,\mathcal{V}(t)^{-1} + \big(\partial\mathcal{V}(t)\,\mathcal{V}(t)^{-1}\big)^{\mathcal{T}}\Big] \\ &= \partial\phi(t)h + \frac{1}{2}e^{2\phi(t)}\partial\mathcal{A}_{123}(t)\left(E^{123} + F_{123}\right),\end{aligned} \tag{6.3.4}$$

where the exponent is the linear form $\alpha(\phi) = 2\phi$ representing the exceptional simple root α_{123} of E_{10}. More precisely, it is the linear form α acting on the Cartan generator $\phi(t)h$, as follows:

$$\alpha(\phi h) = \phi\,\langle\alpha, h\rangle = \phi\,\langle\alpha, \alpha^{\vee}\rangle = \alpha^2\phi = 2\phi. \tag{6.3.5}$$

The Lagrangian becomes

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\left(\mathcal{P}(t)|\mathcal{P}(t)\right) \\ &= \partial\phi(t)\partial\phi(t) + \frac{1}{4}e^{4\phi(t)}\,\partial\mathcal{A}_{123}(t)\,\partial\mathcal{A}_{123}(t).\end{aligned} \tag{6.3.6}$$

For convenience we have chosen the gauge $n = 1$ of the free parameter in the $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$-Lagrangian. Recall that for the level one fields we have $\mathcal{D}\mathcal{A}_{abc}(t) = \partial\mathcal{A}_{abc}(t)$, which is why only the partial derivative of $\mathcal{A}_{123}(t)$ appears in the Lagrangian.

The reason why this simple looking model contains information about eleven-dimensional supergravity is that the A_1 subalgebra represented by (e, f, h) is embedded in E_{10} through the level 1-generator E^{123}, and hence this Lagrangian corresponds to a consistent subgroup truncation of the $\mathcal{E}_{10}$- sigma model.

Let us now study the dynamics of the Lagrangian in Equation (6.3.6). The equations of motion for $\mathcal{A}_{123}(t)$ are

$$\partial\left(\frac{1}{2}e^{4\phi(t)}\partial\mathcal{A}_{123}(t)\right) = 0 \quad\Longrightarrow\quad \frac{1}{2}e^{4\phi(t)}\partial\mathcal{A}_{123}(t) = a, \tag{6.3.7}$$

where a is a constant. The equations for the $\ell = 0$ field ϕ may then be written as

$$\partial^2\phi(t) = 2a^2 e^{-4\phi(t)}. \tag{6.3.8}$$

Integrating once yields

$$\partial\phi(t)\,\partial\phi(t) + a^2 e^{-4\phi(t)} = E, \tag{6.3.9}$$

where E plays the role of the energy for the dynamics of $\phi(t)$. This equation can be solved exactly with the result [135]

$$\phi(t) = \frac{1}{2} \ln \left[\frac{2a}{\sqrt{E}} \cosh \sqrt{E} t \right] \equiv \frac{1}{2} \ln H(t). \tag{6.3.10}$$

We must also take into account the Hamiltonian constraint

$$\mathcal{H} = (\mathcal{P}|\mathcal{P}) = 0, \tag{6.3.11}$$

arising from the variation of $n(t)$ in the $\mathcal{E}_{10}$-sigma model. The Hamiltonian becomes

$$\begin{aligned} \mathcal{H} &= 2\partial\phi(t)\,\partial\phi(t) + \frac{1}{2} e^{4\phi(t)} \partial\mathcal{A}_{123}(t)\,\partial\mathcal{A}_{123}(t) \\ &= 2\left(\partial\phi(t)\partial\phi(t) + a^2 e^{-4\phi(t)}\right) \\ &= 2E. \end{aligned} \tag{6.3.12}$$

It is therefore impossible to satisfy the Hamiltonian constraint unless $E = 0$. This is the problem which was discussed above, and the reason why we need to enlarge the choice of coset representative to include the timelike generator $l \in \mathfrak{h}$. We choose l such that it commutes with h and E^{123},

$$[l, h] = [l, E^{123}] = 0, \tag{6.3.13}$$

and such that isotropy in the directions not supported by the electric field is ensured. Most importantly, in order to solve the problem of the Hamiltonian constraint, l must be timelike,

$$l^2 = (l|l) < 0, \tag{6.3.14}$$

where $(\cdot|\cdot)$ is the scalar product in the Cartan subalgebra of E_{10}. The subalgebra to which we truncate the sigma model is thus given by

$$\bar{\mathfrak{g}} = A_1 \oplus \mathbb{R}l \subset E_{10}, \tag{6.3.15}$$

and the corresponding coset representative is

$$\tilde{\mathcal{V}}(t) = e^{\phi(t)h + \tilde{\phi}(t)l} e^{\mathcal{A}_{123}(t)E^{123}}. \tag{6.3.16}$$

The Lagrangian now splits into two disconnected parts, corresponding to the direct product $SO(2)\backslash SL(2, \mathbb{R}) \times \mathbb{R}$,

$$\tilde{\mathcal{L}} = \left(\partial\phi(t)\,\partial\phi(t) + \frac{1}{4} e^{4\phi(t)} \partial\mathcal{A}_{123}(t)\,\partial\mathcal{A}_{123}(t) \right) + \frac{l^2}{2} \partial\tilde{\phi}(t)\,\partial\tilde{\phi}(t). \tag{6.3.17}$$

The solution for $\tilde{\phi}$ is therefore simply linear in time,

$$\tilde{\phi} = |l^2| \sqrt{\tilde{E}}\, t. \tag{6.3.18}$$

The new Hamiltonian now gets a contribution also from the Cartan generator l,

$$\tilde{\mathcal{H}} = 2E - |l^2| \tilde{E}. \tag{6.3.19}$$

This contribution depends on the norm of l and since $l^2 < 0$, it is possible to satisfy the Hamiltonian constraint, provided that we set

$$\tilde{E} = \frac{2}{|l^2|}E. \tag{6.3.20}$$

We have now found a consistent truncation of the $K(\mathcal{E}_{10}) \times \mathcal{E}_{10}$-invariant sigma model which exhibits $SL(2,\mathbb{R}) \times SO(2) \times \mathbb{R}$-invariance. We want to translate the solution to this model, Equation (6.3.10), to a solution of eleven-dimensional supergravity. The embedding of $\mathfrak{sl}(2,\mathbb{R}) \subset E_{10}$ in Equation (6.3.1) induces a natural "Freund–Rubin" type $(1+3+7)$ split of the coordinates in the physical metric, where the 3-form is supported in the three spatial directions x^1, x^2, x^3. We must also choose an embedding of the timelike generator l. In order to ensure isotropy in the directions $x^4, \cdots, x^{10}$, where the electric field has no support, it is natural to let l be extended only in the "transverse" directions and we take [135]

$$l = K^4{}_4 + \cdots + K^{10}{}_{10}, \tag{6.3.21}$$

which has norm

$$(l|l) = \big(K^4{}_4 + \cdots + K^{10}{}_{10}|K^4{}_4 + \cdots + K^{10}{}_{10}\big) = -42. \tag{6.3.22}$$

To find the metric solution corresponding to our sigma model, we first analyze the coset representative at $\ell = 0$,

$$\tilde{\mathcal{V}}(t)\big|_{\ell=0} = \mathrm{Exp}\left[\phi(t)h + \tilde{\phi}(t)l\right]. \tag{6.3.23}$$

In order to make use of the dictionary from Section 5.3.6 it is necessary to rewrite this in a way more suitable for comparison, i.e., to express the Cartan generators h and l in terms of the $\mathfrak{gl}(10,\mathbb{R})$-generators $K^a{}_b$. We thus introduce parameters $\xi^a{}_b(t)$ and $\tilde{\xi}^a{}_b(t)$ representing, respectively, ϕ and $\tilde{\phi}$ in the $\mathfrak{gl}(10,\mathbb{R})$-basis. The level zero coset representative may then be written as

$$\begin{aligned}
\tilde{\mathcal{V}}(t)\big|_{\ell=0} &= \mathrm{Exp}\left[\sum_{a=1}^{10}\left(\xi^a{}_a(t) + \tilde{\xi}^a{}_a(t)\right)K^a{}_a\right] \\
&= \mathrm{Exp}\left[\sum_{a=4}^{10}\left(\xi^a{}_a(t) + \tilde{\xi}^a{}_a(t)\right)K^a{}_a + \left(\xi^1{}_1(t)K^1{}_1 + \xi^2{}_2(t)K^2{}_2 + \xi^3{}_3(t)K^3{}_3\right)\right],
\end{aligned} \tag{6.3.24}$$

where in the second line we have split the sum in order to highlight the underlying spacetime structure, i.e., to emphasize that $\tilde{\xi}^a{}_b$ has no non-vanishing components in the directions x^1, x^2, x^3. Comparing this to Equation (6.3.1) and Equation (6.3.21) gives the diagonal components of $\xi^a{}_b$ and $\tilde{\xi}^a{}_b$,

$$\xi^1{}_1 = \xi^2{}_2 = \xi^3{}_3 = 2\phi/3, \qquad \xi^4{}_4 = \cdots = \xi^{10}{}_{10} = -\phi/3, \qquad \tilde{\xi}^4{}_4 = \cdots = \tilde{\xi}^{10}{}_{10} = \tilde{\phi}. \tag{6.3.25}$$

Now, the dictionary from Section 5.3.6 identifies the physical spatial metric as follows:

$$\mathrm{g}_{ab}(t) = e_a{}^{\bar{a}}(t)e_b{}^{\bar{b}}(t)\delta_{\bar{a}\bar{b}} = (e^{\xi(t)+\tilde{\xi}(t)})_a{}^{\bar{a}}(e^{\xi(t)+\tilde{\xi}(t)})_b{}^{\bar{b}}\delta_{\bar{a}\bar{b}} \tag{6.3.26}$$

By observation of Equation (6.3.25) we find the components of the metric to be

$$\begin{aligned} \mathrm{g}_{11} &= \mathrm{g}_{22} = \mathrm{g}_{33} = e^{4\phi/3}, \\ \mathrm{g}_{44} &= \cdots = \mathrm{g}_{(10)(10)} = e^{-2\phi/3+2\tilde{\phi}}. \end{aligned} \tag{6.3.27}$$

This result shows clearly how the embedding of h and l into E_{10} is reflected in the coordinate split of the metric. The gauge fixing $N = \sqrt{\mathrm{g}}$ (or $n = 1$) gives the g_{tt}-component of the metric,

$$\mathrm{g}_{tt} = N^2 = e^{14\tilde{\phi}-2\phi/3}. \tag{6.3.28}$$

Next we consider the generator E^{123}. The dictionary tells us that the field strength of the 3-form in eleven-dimensional supergravity at some fixed spatial point $\mathbf{x}_0$ should be identified as

$$F_{t123}(t, \mathbf{x}_0) = \mathcal{D}\mathcal{A}_{123}(t) = \partial \mathcal{A}_{123}(t). \tag{6.3.29}$$

It is possible to eliminate the $\mathcal{A}_{123}(t)$ in favor of the Cartan field $\phi(t)$ using the first integral of its equations of motion, Equation (6.3.7),

$$\frac{1}{2} e^{-4\phi(t)} \partial \mathcal{A}_{123}(t) = a. \tag{6.3.30}$$

In this way we may write the field strength in terms of a and the solution for ϕ,

$$F_{t123}(t, \mathbf{x}_0) = 2a e^{4\phi(t)} = 2a H^{-2}(t). \tag{6.3.31}$$

Finally, we write down the solution for the spacetime metric explicitly:

$$\begin{aligned} ds^2 &= -e^{14\tilde{\phi}+2\phi/3}\, dt^2 + e^{4\phi/3} \left[(dx^1)^2 + (dx^2)^2 + (dx^3)^2\right] + e^{2\tilde{\phi}-2\phi/3} \sum_{\bar{a}=4}^{10} (dx^{\bar{a}})^2 \\ &= -H^{1/3}(t) e^{\frac{1}{3}\sqrt{\tilde{E}}t}\, dt^2 + H^{-2/3}(t) \left[(dx^1)^2 + (dx^2)^2 + (dx^3)^2\right] + H^{1/3}(t) e^{\frac{\sqrt{\tilde{E}}}{21}t} \sum_{\bar{a}=4}^{10} (dx^{\bar{a}})^2, \end{aligned} \tag{6.3.32}$$

where

$$H(t) = \frac{2a}{\sqrt{E}} \cosh \sqrt{E} t. \tag{6.3.33}$$

This solution coincides with the cosmological solution first found in [49] for the geometric configuration $(3_1, 3_1)$, and it is intriguing that it can be exactly reproduced from a manifestly $\mathcal{E}_{10} \times K(\mathcal{E}_{10})$-invariant action, a priori unrelated to any physical model.

Note that in modern terminology, this solution is an SM2-brane solution (see, e.g., [136] for a review) since it can be interpreted as a spacelike (i.e., time-dependent) version of the M2-brane solution. From this point of view the world volume of the SM2-brane is extended in the directions x^1, x^2 and x^3, and so is Euclidean.

In the BKL-limit this solution describes two asymptotic Kasner regimes, at $t \to \infty$ and at $t \to -\infty$. These are separated by a collision against an electric wall, corresponding to the blow-up of the electric field $F_{t123}(t) \sim H^{-2}(t)$ at $t = 0$. In the billiard picture the dynamics in the BKL-limit is thus given by free-flight motion interrupted by one geometric reflection against the electric wall,

$$e_{123}(\beta) = \beta^1 + \beta^2 + \beta^3, \tag{6.3.34}$$

which is the exceptional simple root of E_{10}. This indicates that in the strict BKL-limit, electric walls and SM2-branes are actually equivalent.

6.3.3 Intersecting Spacelike Branes From Geometric Configurations

Let us now examine a slightly more complicated example. We consider the configuration $(6_2, 4_3)$, shown in Figure 6.2. This configuration has four lines and six points. As such the associated supergravity model describes a cosmological solution with four components of the electric field turned on, or, equivalently, it describes a set of four intersecting SM2-branes [122].

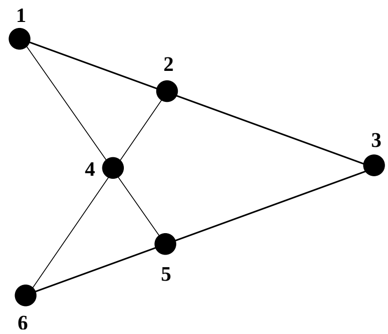


Figure 6.2: The configuration $(6_2, 4_3)$, dual to the Lie algebra $A_1 \oplus A_1 \oplus A_1 \oplus A_1$.

From the configuration we read off the Chevalley–Serre generators associated to the simple roots of the dual algebra:

$$e_1 = E^{123}, \qquad e_2 = E^{145}, \qquad e_3 = E^{246}, \qquad e_4 = E^{356}. \tag{6.3.35}$$

The first thing to note is that all generators have one index in common since in the graph any two lines share one node. This implies that the four lines in $(6_2, 4_3)$ define four *commuting* A_1 subalgebras,

$$(6_2, 4_3) \quad \Longleftrightarrow \quad \mathfrak{g}_{(6_2,4_3)} = A_1 \oplus A_1 \oplus A_1 \oplus A_1. \tag{6.3.36}$$

One can make sure that the Chevalley–Serre relations are indeed fulfilled for this embedding. For instance, the Cartan element $h = [E^{b_1 b_2 b_3}, F_{b_1 b_2 b_3}]$ (no summation on the fixed, distinct indices b_1, b_2, b_3) reads

$$h = -\frac{1}{3} \sum_{a \neq b_1, b_2, b_3} K^a{}_a + \frac{2}{3}(K^{b_1}{}_{b_1} + K^{b_2}{}_{b_2} + K^{b_3}{}_{b_3}). \tag{6.3.37}$$

Hence, the commutator $[h, E^{b_i cd}]$ vanishes whenever $E^{b_i cd}$ has only one b-index,

$$\begin{aligned} [h, E^{b_i cd}] &= -\frac{1}{3}[(K^c{}_c + K^d{}_d), E^{b_i cd}] + \frac{2}{3}[(K^{b_1}{}_{b_1} + K^{b_2}{}_{b_2} + K^{b_3}{}_{b_3}), E^{b_i cd}] \\ &= \left(-\frac{1}{3} - \frac{1}{3} + \frac{2}{3}\right) E^{b_i cd} = 0 \qquad (i = 1, 2, 3). \end{aligned} \tag{6.3.38}$$

Furthermore, multiple commutators of the step operators are immediately killed at level 2 whenever they have one index or more in common, e.g.,

$$[E^{123}, E^{145}] = E^{123145} = 0. \tag{6.3.39}$$

To fulfill the Hamiltonian constraint, one must extend the algebra by taking a direct sum with $\mathbb{R}l$, $l = K^7{}_7 + K^8{}_8 + K^9{}_9 + K^{10}{}_{10}$. Accordingly, the final algebra is $A_1 \oplus A_1 \oplus A_1 \oplus A_1 \oplus \mathbb{R}l$. Because there is no magnetic field, the momentum constraint and Gauss' law are identically satisfied.

By investigating the sigma model solution corresponding to the algebra $\mathfrak{g}_{(6_2,4_3)}$, augmented with the timelike generator l,

$$\bar{\mathfrak{g}} = A_1 \oplus A_1 \oplus A_1 \oplus A_1 \oplus \mathbb{R}l, \tag{6.3.40}$$

we find a supergravity solution which generalizes the one found in [49]. The solution describes a set of four intersecting SM2-branes, with a five-dimensional transverse spacetime in the directions t, x^7, x^8, x^9, x^{10}.

Let us write down also this solution explicitly. The full set of generators for $\mathfrak{g}_{(6_2,4_3)}$ is

$$\begin{aligned}
e_1 &= E^{123}, \quad e_2 = E^{145}, \quad e_3 = E^{246}, \quad e_4 = E^{356} \\
f_1 &= F_{123}, \quad f_2 = F_{145}, \quad f_3 = F_{246}, \quad f_4 = F_{356} \\
h_1 &= -\frac{1}{3}\sum_{a\neq 1,2,3} K^a{}_a + \frac{2}{3}(K^1{}_1 + K^2{}_2 + K^3{}_3), \\
h_2 &= -\frac{1}{3}\sum_{a\neq 1,4,5} K^a{}_a + \frac{2}{3}(K^1{}_1 + K^4{}_4 + K^5{}_5), \\
h_3 &= -\frac{1}{3}\sum_{a\neq 2,4,6} K^a{}_a + \frac{2}{3}(K^2{}_2 + K^4{}_4 + K^6{}_6), \\
h_4 &= -\frac{1}{3}\sum_{a\neq 3,5,6} K^a{}_a + \frac{2}{3}(K^3{}_3 + K^5{}_5 + K^6{}_6).
\end{aligned} \tag{6.3.41}$$

The coset element for this configuration then becomes

$$\mathcal{V}(t) = e^{\phi_1(t)h_1+\phi_2(t)h_2+\phi_3(t)h_3+\phi_4(t)h_4+\tilde{\phi}(t)l}\, e^{\mathcal{A}_{123}(t)E^{123}+\mathcal{A}_{145}(t)E^{145}+\mathcal{A}_{246}(t)E^{246}+\mathcal{A}_{356}(t)E^{356}}. \tag{6.3.42}$$

We must further choose the timelike Cartan generator, $l \in \mathfrak{h}$, appropriately. Examination of Equation (6.3.41) reveals that the four electric fields are supported only in the spatial directions $x^1, \cdots, x^6$ so, again, in order to ensure isotropy in the directions transverse to the S-branes, we choose the timelike Cartan generator as follows:

$$l = K^7{}_7 + K^8{}_8 + K^9{}_9 + K^{10}{}_{10}, \tag{6.3.43}$$

which implies

$$l^2 = (l|l) = \big(K^7{}_7 + K^8{}_8 + K^9{}_9 + K^{10}{}_{10} | K^7{}_7 + K^8{}_8 + K^9{}_9 + K^{10}{}_{10}\big) = -12. \tag{6.3.44}$$

The Lagrangian for this system becomes

$$\mathcal{L}_{(6_2,4_3)} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \frac{l^2}{2}\partial\tilde{\phi}(t)\,\partial\tilde{\phi}(t), \tag{6.3.45}$$

where $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3$ and $\mathcal{L}_4$ represent the $SL(2,\mathbb{R}) \times SO(2)$-invariant Lagrangians corresponding to the four A_1-algebras. The solutions for $\phi_1(t), \cdots, \phi_4(t)$ and $\tilde{\phi}(t)$ are separately identical to the ones for $\phi(t)$ and $\tilde{\phi}(t)$, respectively, in Section 6.3.2. From the embedding into E_{10}, provided in Equation (6.3.41), we may read off the solution for the spacetime metric,

$$\begin{aligned} ds^2_{(6_2,4_3)} &= -(H_1H_2H_3H_4)^{1/3} e^{\frac{2}{3}\sqrt{E_-}t}\, dt^2 + (H_1H_4)^{-2/3}(H_2H_3)^{1/3}(dx^1)^2 \\ &\quad +(H_1H_3)^{-2/3}(H_2H_4)^{1/3}(dx^2)^2 + (H_1H_2)^{-2/3}(H_3H_4)^{1/3}(dx^3)^2 \\ &\quad +(H_3H_4)^{-2/3}(H_1H_2)^{1/3}(dx^4)^2 + (H_2H_4)^{-2/3}(H_1H_3)^{1/3}(dx^5)^2 \\ &\quad +(H_2H_3)^{-2/3}(H_1H_4)^{1/3}(dx^6)^2 + (H_1H_2H_3H_4)^{1/3} e^{\frac{1}{6}\sqrt{E_-}t} \sum_{\bar{a}=7}^{10}(dx^{\bar{a}})^2. \end{aligned} \tag{6.3.46}$$

As announced, this describes four intersecting SM2-branes with a 1+4-dimensional transverse spacetime. For example the brane that couples to the field associated with the first Cartan generator is extended in the directions x^1, x^2, x^3. By restricting to the case $\phi_1 = \phi_2 = \phi_3 = \phi_4 \equiv \phi$ the metric simplifies to

$$\begin{aligned} ds^2_{(6_2,4_3)} &= -\left(\frac{2a}{\sqrt{E}}\right)^{4/3} \cosh^{4/3}\sqrt{E}t\, e^{\frac{2}{3}\sqrt{\tilde{E}}t} dt^2 + \left(\frac{2a}{\sqrt{E}}\right)^{-2/3} \cosh^{-2/3}\sqrt{E}t \sum_{a'=1}^{6}(dx^{a'})^2 \\ &\quad + \left(\frac{2a}{\sqrt{E}}\right)^{4/3} \cosh^{4/3}\sqrt{E}t\, e^{\frac{1}{6}\sqrt{\tilde{E}}t} \sum_{\bar{a}=7}^{10}(dx^{\bar{a}})^2, \end{aligned} \tag{6.3.47}$$

which coincides with the cosmological solution found in [49] for the configuration $(6_2, 4_3)$. We can therefore conclude that the algebraic interpretation of the geometric configurations found in this chapter generalizes the solutions given in the aforementioned reference.

In a more general setting where we excite more roots of E_{10}, the solutions of course become more complex. However, as long as we consider *commuting* subalgebras there will naturally be no coupling in the Lagrangian between fields parametrizing different subalgebras. This implies that if we excite a direct sum of m A_1-algebras the total Lagrangian will split according to

$$\mathcal{L} = \sum_{k=1}^{m} \mathcal{L}_k + \tilde{\mathcal{L}}, \tag{6.3.48}$$

where $\mathcal{L}_k$ is of the same form as Equation (6.3.6), and $\tilde{\mathcal{L}}$ is the Lagrangian for the timelike Cartan element, needed in order to satisfy the Hamiltonian constraint. It follows that the associated solutions are

$$\begin{aligned} \phi_k(t) &= \frac{1}{2}\ln\left[\frac{a_k}{E_k}\cosh\sqrt{E_k}t\right] \qquad (k = 1, \cdots, m), \\ \tilde{\phi}(t) &= |l^2|\sqrt{\tilde{E}}t. \end{aligned} \tag{6.3.49}$$

Furthermore, the resulting structure of the metric depends on the embedding of the A_1-algebras into E_{10}, i.e., which level 1-generators we choose to realize the step-operators and hence which Cartan elements that are associated to the ϕ_k's. Each excited A_1-subalgebra will turn on an electric 3-form that couples to an SM2-brane and hence the solution for the metric will describe a set of m intersecting SM2-branes.

As an additional nice example, we mention here the configuration $(7_3, 7_3)$, also known as the *Fano plane*, which consists of 7 lines and 7 points (see Figure 6.3). This configuration is well known for its relation to the octonionic multiplication table [137]. For our purposes, it is interesting because none of the lines in the configuration are parallell. Thus, the algebra dual to the Fano plane is a direct sum of seven A_1-algebras and the supergravity solution derived from the sigma model describes a set of seven intersecting SM2-branes.

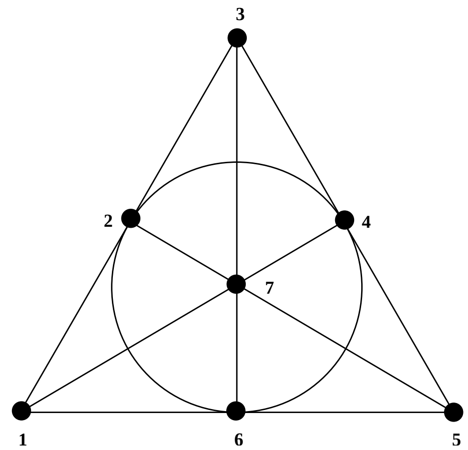


Figure 6.3: The Fano Plane, $(7_3, 7_3)$, dual to the Lie algebra $A_1 \oplus A_1 \oplus A_1 \oplus A_1 \oplus A_1 \oplus A_1 \oplus A_1$.

6.3.4 Intersection Rules for Spacelike Branes

For multiple brane solutions, there are rules for how these branes may intersect in order to describe allowed BPS-solutions [138]. These intersection rules also apply to spacelike branes [139] and hence they apply to the solutions considered here. In this section we will show that the intersection rules for multiple S-brane solutions are encoded in the associated geometric configurations [122].

For two spacelike q-branes, A and B, in M-theory the rules are

$$SMq_A \cap SMq_B = \frac{(q_A+1)(q_B+1)}{9} - 1. \tag{6.3.50}$$

So, for example, if we have two SM2-branes the result is

$$SM2 \cap SM2 = 0, \tag{6.3.51}$$

which means that they are allowed to intersect on a 0-brane. Note that since we are dealing with spacelike branes, a zero-brane is extended in one spatial direction, so the two SM2-branes may therefore intersect in one spatial direction only. We see from Equation (6.3.46) that these rules are indeed fulfilled for the configuration $(6_2, 4_3)$.

In [67] it was found in the context of $\mathfrak{g}^{+++}$-algebras that the intersection rules for extremal branes are encoded in orthogonality conditions between the various roots from which the branes arise. This is equivalent to saying that the subalgebras that we excite are commuting, and hence the same result applies to $\mathfrak{g}^{++}$-algebras in the cosmological context[3]. From this point of view, the intersection rules can also be read off from the geometric configurations in the sense that the configurations encode information about whether or not the algebras commute.

The next case of interest is the Fano plane, $(7_3, 7_3)$. As mentioned above, this configuration corresponds to the direct sum of 7 commuting A_1 algebras and so the gravitational solution describes a set of 7 intersecting SM2-branes. The intersection rules are guaranteed to be satisfied for the same reason as before.

6.4 Cosmological Solutions With Magnetic Flux

We will now briefly sketch how one can also obtain the SM5-brane solutions from geometric configurations and regular subalgebras of E_{10}. In order to do this we consider "magnetic" subalgebras of E_{10}, constructed only from simple root generators at level two in the level decomposition of E_{10}. To the best of our knowldege, there is no theory of geometric configurations developed for the case of having 6 points on each line, which would be needed here. However, we may nevertheless continue to investigate the simplest example of such a configuration, namely $(6_1, 1_6)$, displayed in Figure 6.4.

Figure 6.4: The simplest "magnetic configuration" $(6_1, 1_6)$, dual to the algebra A_1. The associated supergravity solution describes an SM5-brane, whose world volume is extended in the directions $x^1, \cdots, x^6$.

The algebra dual to this configuration is an A_1-subalgebra of E_{10} with the following generators:

$$\begin{aligned} e &= E^{123456} = F_{123456}, \\ h &\equiv [E^{123456}, F_{123456}] = -\frac{1}{6}\sum_{a\neq 1,\cdots,6} K^a{}_a + \frac{1}{3}(K^1{}_1 + \cdots + K^6{}_6). \end{aligned} \tag{6.4.1}$$

Although the embedding of this algebra is different from the electric cases considered previously, the sigma model solution is still associated to an $SL(2,\mathbb{R})/SO(2)$ coset space and therefore the solutions for $\phi(t)$ and $\tilde{\phi}(t)$ are the same as before. Because of the embedding, however, the sigma model translates to a different type of supergravity solution, namely a spacelike five-brane whose world volume is extended in the directions $x^1, \cdots, x^6$. The metric is given by

$$ds^2 = -H^{-4/3}(t)\, e^{\frac{2}{3}\sqrt{E_-}t}\, dt^2 + H^{-1/3}(t)\sum_{a'=1}^{6}(dx^{a'})^2 + H^{1/6}(t)e^{\frac{1}{6}\sqrt{E_-}t}\sum_{\bar{a}=7}^{10}(dx^{\bar{a}})^2. \tag{6.4.2}$$

[3]This was also pointed out in [135].

This solution coincides with the SM5-brane found by Strominger and Gutperle in [140][4]. Note that the correct power of $H(t)$ for the five-brane arises here entirely due to the embedding of h into E_{10} through Equation (6.4.1).

Because of the existence of electric-magnetic duality on the supergravity side, it is suggestive to expect the existence of a duality between the two types of configurations (n_m, g_3) and (n_m, g_6), of which we have here seen the simplest realisation for the configurations $(3_1, 1_3)$ and $(6_1, 1_6)$.

6.5 The Petersen Algebra and the Desargues Configuration

We want to end this section by considering an example which is more complicated, but very interesting from the algebraic point of view. There exist ten geometric configurations of the form $(10_3, 10_3)$, i.e., with exactly ten points and ten lines. In [49], these were associated to supergravity solutions with ten components of the electric field turned on. This result was re-analyzed by some of the present authors in [122] where it was found that many of these configurations have a dual description in terms of Dynkin diagrams of rank 10 Lorentzian Kac–Moody subalgebras of E_{10}. One would therefore expect that solutions of the sigma models for these algebras should correspond to new solutions of eleven-dimensional supergravity. However, since these algebras are infinite-dimensional, the corresponding sigma models are difficult to solve without further truncation. Nevertheless, one may argue that explicit solutions should exist, since the algebras in question are all non-hyperbolic, so we know that the supergravity dynamics is non-chaotic.

We shall here consider one of the $(10_3, 10_3)$-configurations in some detail, referring the reader to [122] for a discussion of the other cases. The configuration we will treat is the well known *Desargues configuration*, displayed in Figure 6.5. The Desargues configuration is associated with the 17th century French mathematician *Gérard Desargues* to illustrate the following "Desargues theorem" (adapted from [134]):

> *Let the three lines defined by* $\{4,1\}, \{5,2\}$ *and* $\{6,3\}$ *be concurrent, i.e., be intersecting at one point, say* $\{7\}$*. Then the three intersection points* $8 \equiv \{1,2\} \cap \{4,5\}, 9 \equiv \{2,3\} \cap \{5,6\}$ *and* $10 \equiv \{1,3\} \cap \{4,6\}$ *are colinear.*

Another way to say this is that the two triangles $\{1,2,3\}$ and $\{4,5,6\}$ in Figure 6.5 are in perspective from the point $\{7\}$ and in perspective from the line $\{8,10,9\}$.

As we will see, a new fascinating feature emerges for this case, namely that the Dynkin diagram dual to this configuration *also* corresponds in itself to a geometric configuration. In fact, the Dynkin diagram dual to the Desargues configuration turns out to be the famous *Petersen graph*, denoted $(10_3, 15_2)$, which is displayed in Figure 6.6.

To construct the Dynkin diagram we first observe that each line in the configuration is disconnected from three other lines, e.g., $\{4,1,7\}$ have no nodes in common with the lines $\{2,3,9\}$, $\{5,6,9\}$, $\{8,10,9\}$. This implies that all nodes in the Dynkin diagram will be connected to three other nodes. Proceeding as in Section 6.2.2 leads to the Dynkin diagram

[4] In [140] they were dealing with a hyperbolic internal space so there was an additional sinh-function in the transverse spacetime.

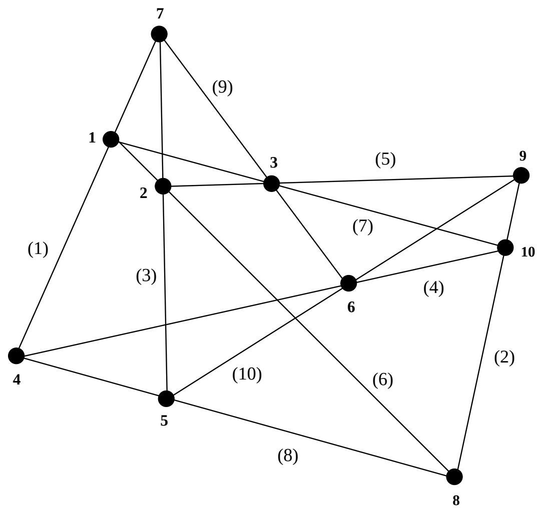


Figure 6.5: $(10_3, 10_3)_3$: The Desargues configuration, dual to the Petersen graph.

in Figure 6.6, which we identify as the Petersen graph. The corresponding Cartan matrix is

$$A(\mathfrak{g}_{\text{Petersen}}) = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ -1 & 2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ -1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 2 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 2 \end{pmatrix}, \tag{6.5.1}$$

which is of Lorentzian signature with

$$\det A(\mathfrak{g}_{\text{Petersen}}) = -256. \tag{6.5.2}$$

The Petersen graph was invented by the Danish mathematician *Julius Petersen* in the end of the 19th century. It has several embeddings on the plane, but perhaps the most famous one is as a star inside a pentagon as depicted in Figure 6.6. One of its distinguishing features from the point of view of graph theory is that it contains a *Hamiltonian path* but no *Hamiltonian cycle*[5].

Because the algebra is Lorentzian (with a metric that coincides with the metric induced from the embedding in E_{10}), it does not need to be enlarged by any further generator to be compatible with the Hamiltonian constraint. Recently, cosmological solutions associated with the Petersen algebra were further analyzed in [141].

[5]We recall that a Hamiltonian path is defined as a path in an undirected graph which intersects each node once and only once. A Hamiltonian cycle is then a Hamiltonian path which also returns to its initial node.

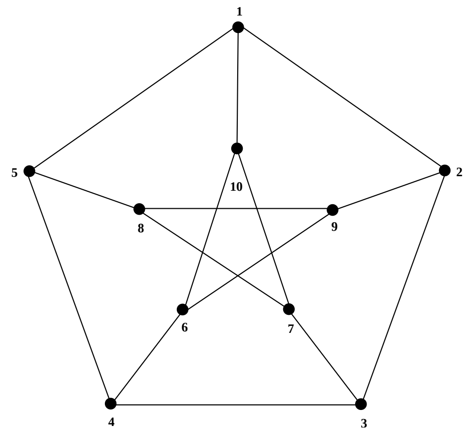


Figure 6.6: This is the so-called Petersen graph. It is the Dynkin diagram dual to the Desargues configuration, and is in fact a geometric configuration itself, denoted $(10_3, 15_2)$.

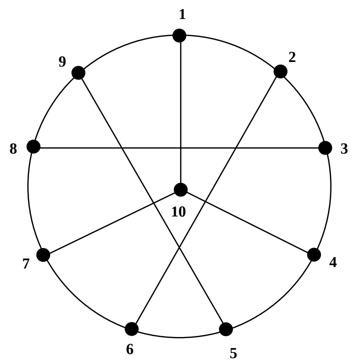


Figure 6.7: An alternative drawing of the Petersen graph in the plane. This embedding reveals an S_3 permutation symmetry about the central point.

7

$\mathcal{E}_{10}$ and Massive Type IIA Supergravity

In the previous chapter, we discussed in detail the example of eleven-dimensional supergravity, and how its associated dynamics in a cosmological regime is reproduced from a geodesic sigma model with target space $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$. This correspondence relied heavily on the level deomposition of the Lie algebra E_{10} with respect to the finite-dimensional subalgebra $\mathfrak{sl}(10,\mathbb{R})$, which at low levels reveals representations matching the tensor structure of the p-form fields of eleven-dimensional supergravity. As discussed in some detail in Section 2.5.4, slicing up E_{10} with respect to different finite subalgebras gives rise to the field contents also of the type IIA and IIB supergravities in ten dimensions. In this chapter we shall investigate this correspondence in the context of massive type IIA supergravity, originally discovered by Romans [142], whose field content arises from a decomposition of E_{10} with respect to a specific $\mathfrak{sl}(9,\mathbb{R})$-subalgebra. Here it is particularly interesting to note that from the E_{10} point of view, the mass deformation arises from a representation at $\mathfrak{sl}(10,\mathbb{R})$-level 4 in the terminology of Sections 2.5.4 and 5.3. Although this level 4 representation has no interpretation in the eleven-dimensional theory, we shall find that it matches precisely with the mass-parameter in the type IIA context, thus revealing an algebraic and dynamic unification of massive IIA supergravity and eleven-dimensional supergravity with the framework of the $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$-sigma model at low levels. In this chapter we will spell out this relation in detail, including both the bosonic and fermionic sectors of massive IIA supergravity. To include fermions in the correspondence, we must extend the $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$-sigma model of Section 5.3 to incorporate spinorial representations of $K(\mathcal{E}_{10})$. This is developed in Section 7.2 following earlier work [69–72]. To improve readability, many of the details of the calculations are relegated to Appendix A. This chapter is based on **Paper VI**, written in collaboration with Marc Henneaux, Ella Jamsin and Axel Kleinschmidt.

7.1 Massive IIA supergravity

Massive type IIA supergravity was first constructed by Romans [142] by deforming the standard type IIA supergravity through a Stückelberg mechanism, giving a mass to the two-form potential through the replacement $F_{\mu\nu} \to F_{\mu\nu} + mA_{\mu\nu}$, where $F_{\mu\nu} = 2\partial_{[\mu}A_{\nu]}$. The potential

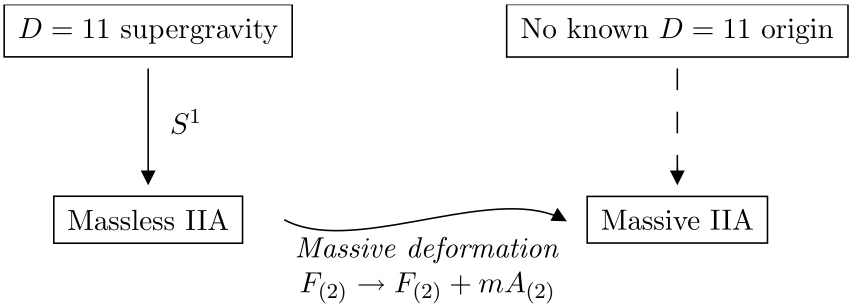


Figure 7.1: *Massive IIA supergravity from $D = 11$ supergravity: Massive type IIA supergravity is obtained as a deformation of the standard type IIA supergravity, but unlike the latter, it does not possess any known eleven-dimensional origin. See Figure 7.2 for a pictorial description of how $D = 11$ supergravity and massive IIA supergravity are unified inside $\mathcal{E}_{10}$.*

A_μ can then be gauged away by a gauge transformation of $A_{\mu\nu}$. This process unfortunately obscures the massless limit to recover the standard IIA theory [5–7] as some of the supersymmetry variations involve a coefficient m^{-1}. This is remedied by a field redefinition presented in [143, 144], that we will make use of in this chapter.

Moreover, a more democratic version of massive type IIA is given in [144, 145], in which every form field comes with its dual. In particular, it involves a nine-form as a (potential) dual to the mass (seen as a field strength). This democratic formulation makes explicit the striking feature of massive IIA supergravity that it allows the existence of a D8-brane, that is known to exist in type IIA string theory. Indeed, this feature requires a nine-form potential that couples to the D8-brane, and accordingly does not appear in massless IIA supergravity. Although in our analysis we will not use directly the democratic formulation, we will employ the field redefinitions used in [143] in order to clarify the massless limit.

In addition, as a motivation for the present work, one also finds a nine-form in a certain decomposition of $\mathcal{E}_{10}$, as will be developed in the next section, and this nine-form appears in the $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ Lagrangian in the same way as the mass term in the massive IIA Lagrangian. Massive IIA supergravity has in common with many other deformed maximal supergravities that it does not possess any known higher dimensional origin, as illustrated in Figure 7.1. A consequence of the present work is to show that, although they are not related by dimensional reduction, eleven dimensional supergravity and massive IIA supergravity have the same $\mathcal{E}_{10}$ origin as displayed in Figure 7.2, see also [58, 116, 146].[1]

7.1.1 Romans' Theory

In this section, and throughout the chapter, the curved ten-dimensional space-time indices are denoted by μ and decompose into time t and nine spatial indices m. The flat indices are similarly denoted by $\alpha = (0, a)$. Our space-time signature is $(- + \ldots +)$. More details on the

[1] We note that the search for an eleven-dimensional origin of the Romans mass parameter, and hence the D8-brane, has led to studies of an M-theory M9-brane which is meant to exist in the presence of one Killing direction and then reduces to the IIA D8-brane, see e.g. [147, 148].

$$\begin{tikzcd}[column sep=small]
& \boxed{E_{10}} \arrow[dl, "\ell_1"'] \arrow[dr, "(\ell_1,\ell_2)"] & \\
\boxed{\mathcal{L}_{\mathrm{SUGRA}_{11}} \subset \mathcal{L}_{E_{10}/K(E_{10})}} \arrow[rr, "\text{`Dimensional Reduction'}", "\ell_2"'] & & \boxed{\mathcal{L}_{\mathrm{mIIA}} \subset \mathcal{L}_{E_{10}/K(E_{10})}}
\end{tikzcd}$$

Figure 7.2: *This picture describes the common $\mathcal{E}_{10}$ origin of eleven-dimensional supergravity and massive type IIA supergravity. First, if one considers a level ℓ_1 decomposition of $\mathcal{E}_{10}$ with respect to A_9 (cf. Figure 7.3), one sees that the first levels ($\ell_1 = 0$ to $\ell_1 = 3$) of an $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ sigma-model correspond to a truncated version of eleven-dimensional supergravity, with Lagrangian $\mathcal{L}_{\mathrm{SUGRA}_{11}}$ [149]. Taking this as a starting point, we can perform an additional level ℓ_2 decomposition on the sigma model. On the lower ℓ_1 levels ($\ell_1 = 0$ to $\ell_1 = 3$) this is equivalent to a dimensional reduction of eleven-dimensional supergravity, which gives massless IIA supergravity (cf. Figure 7.1). However, if one includes one of the generators appearing at $\ell_1 = 4$, this leads to a theory that coincides with a truncated version of massive IIA supergravity, with Lagrangian $\mathcal{L}_{\mathrm{mIIA}}$. This procedure is equivalent to a multi-level (ℓ_1, ℓ_2) decomposition of $\mathcal{E}_{10}$ with respect to A_8. More details on this are given in Section 7.2.2. By similar arguments, one could add IIB supergravity to this picture, in the sense that it has the same $\mathcal{E}_{10}$ origin in a level decomposition with respect to $A_8 \times A_1$, i.e. with respect to node 8 in the Dynkin diagram in Figure 7.3 [62].*

conventions can be found in Appendix A.1.

The complete supersymmetric Lagrangian (up to second order in fermions) in Einstein frame is given by the sum of a bosonic and a fermionic part, $\mathcal{L} = \mathcal{L}^{[B]} + \mathcal{L}^{[F]}$. The bosonic sector contains a metric $G_{\mu\nu}$, a dilaton ϕ, a one-form $A_{(1)}$ with field strength $F_{(2)}$, a two-form $A_{(2)}$ with field strength $F_{(3)}$, a three-form $A_{(3)}$ with field strength $F_{(4)}$, and a real mass parameter m. The bosonic part of the Lagrangian reads

$$\begin{aligned}
\mathcal{L}^{[B]} \;=\;& \sqrt{-G}\left[R - \frac{1}{2}|\partial\phi|^2 - \frac{1}{4}e^{3\phi/2}|F_{(2)}|^2 - \frac{1}{12}e^{-\phi}|F_{(3)}|^2 - \frac{1}{48}e^{\phi/2}|F_{(4)}|^2 - \frac{1}{2}m^2 e^{5\phi/2}\right] \\
&+ \tilde{\varepsilon}^{\mu_1\ldots\mu_{10}}\left[\frac{1}{144}\partial_{\mu_1}A_{\mu_2\mu_3\mu_4}\partial_{\mu_5}A_{\mu_6\mu_7\mu_8}A_{\mu_9\mu_{10}} + \frac{m}{288}\partial_{\mu_1}A_{\mu_2\mu_3\mu_4}A_{\mu_5\mu_6}A_{\mu_7\mu_8}A_{\mu_9\mu_{10}}\right. \\
&\left. + \frac{m^2}{1280}A_{\mu_1\mu_2}A_{\mu_3\mu_4}A_{\mu_5\mu_6}A_{\mu_7\mu_8}A_{\mu_9\mu_{10}}\right]. \qquad (7.1.1a)
\end{aligned}$$

Here, we already point out that the massive deformation induces a positive definite potential on the scalar sector. This is in marked contrast to what happens for gauged supergravity in lower dimensions where the scalar potential is generically indefinite [150, 151], causing also problems in the connection to $\mathcal{E}_{10}$ [152].

On the fermionic side, we have two gravitini, combined in a single 10×32 component vector-spinor ψ_μ, and two dilatini, combined in a single 32 component Dirac-spinor λ, which decompose into two fields of opposite chirality under $SO(1,9)$. For this sector the Lagrangian takes the form

$$
\begin{aligned}
\mathcal{L}^{[F]} &= i\sqrt{-G}\Bigg[-2\bar{\psi}_{\mu_1}\Gamma^{\mu_1\ldots\mu_3}D_{\mu_2}\psi_{\mu_3} - \bar{\lambda}\Gamma^{\mu}D_\mu\lambda + \partial_{\mu_1}\phi\bar{\lambda}\Gamma^{\mu_2}\Gamma^{\mu_1}\psi_{\mu_2} \\
&+ \frac{1}{4}e^{3\phi/4}F_{\mu_1\mu_2}\Big(\bar{\psi}_{[\nu_1}\Gamma^{\nu_1}\Gamma^{\mu_1\mu_2}\Gamma^{\nu_2}\Gamma_{10}\psi_{\nu_2]} - \frac{3}{2}\bar{\lambda}\Gamma^{\nu}\Gamma^{\mu_1\mu_2}\Gamma_{10}\psi_\nu + \frac{5}{8}\bar{\lambda}\Gamma^{\mu_1\mu_2}\Gamma_{10}\lambda\Big) \\
&+ \frac{1}{12}e^{-\phi/2}F_{\mu_1\ldots\mu_3}\Big(\bar{\psi}_{[\nu_1}\Gamma^{\nu_1}\Gamma^{\mu_1\ldots\mu_3}\Gamma^{\nu_2}\Gamma_{10}\psi_{\nu_2]} - \bar{\lambda}\Gamma^{\nu}\Gamma^{\mu_1\ldots\mu_3}\Gamma_{10}\psi_\nu\Big) \\
&- \frac{1}{48}e^{\phi/4}F_{\mu_1\ldots\mu_4}\Big(\bar{\psi}_{[\nu_1}\Gamma^{\nu_1}\Gamma^{\mu_1\ldots\mu_4}\Gamma^{\nu_2}\psi_{\nu_2]} - \frac{1}{2}\bar{\lambda}\Gamma^{\nu}\Gamma^{\mu_1\ldots\mu_4}\psi_\nu + \frac{3}{8}\bar{\lambda}\Gamma^{\mu_1\ldots\mu_4}\lambda\Big) \\
&+ \frac{21}{8}me^{5\phi/4}\bar{\lambda}\lambda - \frac{1}{2}me^{5\phi/4}\bar{\psi}_{\mu_1}\Gamma^{\mu_1\mu_2}\psi_{\mu_2} + \frac{5}{4}me^{5\phi/4}\bar{\lambda}\Gamma^{\mu}\psi_\mu\Bigg].
\end{aligned}
\tag{7.1.1b}
$$

The last line contains the explicit mass terms for the fermions. There are also implicit mass deformations in the definitions of the field strength in terms of the gauge potentials

$$
\begin{aligned}
F_{\mu_1\mu_2} &= 2\partial_{[\mu_1}A_{\mu_2]} + mA_{\mu_1\mu_2}\,, \\
F_{\mu_1\mu_2\mu_3} &= 3\partial_{[\mu_1}A_{\mu_2\mu_3]}\,, \\
F_{\mu_1\mu_2\mu_3\mu_4} &= 4\partial_{[\mu_1}A_{\mu_2\mu_3\mu_4]} + 4A_{[\mu_1}F_{\mu_2\mu_3\mu_4]} + 3mA_{[\mu_1\mu_2}A_{\mu_3\mu_4]}\,.
\end{aligned}
\tag{7.1.2}
$$

In the Lagrangian, they are contracted without additional factors, for example

$$|F_{(2)}|^2 = F_{\mu_1\mu_2}F^{\mu_1\mu_2}. \tag{7.1.3}$$

In (7.1.2) we see the characteristic feature of a deformed theory that the usual tensor hierarchy of gauge fields is broken: there are forms coupling to potentials of higher degree but only through terms proportional to the deformation parameter. The tangent space field strengths are defined as usual by conversion with the (inverse) vielbein $e_\alpha{}^\mu$, for example $F_{\alpha_1\alpha_2} = e_{\alpha_1}{}^{\mu_1}e_{\alpha_2}{}^{\mu_2}F_{\mu_1\mu_2}$. Flat indices are raised and lowered with the Minkowski metric and we will often write contracted flat spatial indices on the same level.

7.1.2 Supersymmetry Variations

The supersymmetry variations leaving (7.1.1) invariant, up to total derivatives and higher order fermion terms, are listed in this section. For the fermions, they read

$$
\begin{aligned}
\delta_\varepsilon\psi_\mu &= D_\mu\varepsilon - \frac{1}{32}me^{5\phi/4}\Gamma_\mu\varepsilon - \frac{1}{64}e^{3\phi/4}F_{\nu\rho}\big(\Gamma_\mu{}^{\nu\rho} - 14\delta_\mu^{[\nu}\Gamma^{\rho]}\big)\Gamma_{10}\varepsilon \\
&\quad+\frac{1}{96}e^{-\phi/2}F_{\nu\rho\sigma}\big(\Gamma_\mu{}^{\nu\rho\sigma} - 9\delta_\mu^{[\nu}\Gamma^{\rho\sigma]}\big)\Gamma_{10}\varepsilon \\
&\quad+\frac{1}{256}e^{\phi/4}F_{\nu\rho\sigma\gamma}\big(\Gamma_\mu{}^{\nu\rho\sigma\gamma} - \frac{20}{3}\delta_\mu^{[\nu}\Gamma^{\rho\sigma\gamma]}\big)\varepsilon,
\end{aligned}
\tag{7.1.4}
$$

and

$$\begin{aligned}\delta_\varepsilon\lambda &= \frac{1}{2}\partial_\mu\phi\Gamma^\mu\varepsilon + \frac{5}{8}e^{5\phi/4}m\,\varepsilon - \frac{3}{16}e^{3\phi/4}F_{\mu\nu}\Gamma^{\mu\nu}\Gamma_{10}\varepsilon \\ &\quad -\frac{1}{24}e^{-\phi/2}F_{\mu\nu\rho}\Gamma^{\mu\nu\rho}\Gamma_{10}\varepsilon + \frac{1}{192}e^{\phi/4}F_{\mu\nu\rho\sigma}\Gamma^{\mu\nu\rho\sigma}\varepsilon,\end{aligned} \tag{7.1.5}$$

while for the bosons we have

$$\delta_\varepsilon e_\mu{}^\alpha = i\bar{\varepsilon}\Gamma^\alpha\psi_\mu, \qquad \delta_\varepsilon\phi = i\bar{\lambda}\varepsilon, \tag{7.1.6}$$

and

$$\delta_\varepsilon A_\mu = \theta_\mu, \qquad \delta_\varepsilon A_{\mu\nu} = \theta_{\mu\nu}, \qquad \delta_\varepsilon A_{\mu\nu\rho} = \theta_{\mu\nu\rho} + 6A_{[\mu}\theta_{\nu\rho]}, \tag{7.1.7}$$

where

$$\begin{aligned}\theta_\mu &:= ie^{-3\phi/4}\left(-\bar{\psi}_\mu - \frac{3}{4}\bar{\lambda}\Gamma_\mu\right)\Gamma_{10}\varepsilon, \\ \theta_{\mu\nu} &:= ie^{\phi/2}\left(2\,\bar{\psi}_{[\mu}\Gamma_{\nu]} - \frac{1}{2}\bar{\lambda}\Gamma_{\mu\nu}\right)\Gamma_{10}\varepsilon, \\ \theta_{\mu\nu\rho} &:= ie^{-\phi/4}\left(3\,\bar{\psi}_{[\mu}\Gamma_{\nu\rho]} + \frac{1}{4}\bar{\lambda}\Gamma_{\mu\nu\rho}\right)\varepsilon.\end{aligned} \tag{7.1.8}$$

Note that the mass only enters in the supersymmetry variations of the fermions.

7.2 On $\mathcal{E}_{10}$ and the Geodesic Sigma Model for $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$

In this section we give some basic properties of the Kac–Moody algebra E_{10} and explain how to construct a non-linear sigma model for geodesic motion on the infinite-dimensional coset space $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$. To this end we shall slice up the adjoint representation of E_{10} in a multi-level decomposition, suitable to reveal the field content of massive IIA supergravity [146, 153]. In order to incorporate also the fermionic sector in the sigma model, we will analyse the relevant (unfaithful) Dirac-spinor and vector-spinor representations of $K(E_{10})$ up to the desired level.[2] For a more detailed discussion of the general E_{10} methods employed here we refer to the review papers [51, 94], for more information about $K(E_{10})$ see [69–72], and for a mathematical introduction to Kac-Moody algebras the canonical reference is [34].

7.2.1 Generalities of the Kac-Moody Algebra E_{10}

Here we discuss the salient features of the hyperbolic Kac–Moody algebra E_{10}, the group of which we shall denote by $\mathcal{E}_{10}$. We will furthermore only be concerned with the split real form

[2] The vector-spinor representation Ψ that we construct (following [69–72]) is unfaithful in the sense that it is a finite-dimensional representation of the infinite-dimensional algebra $K(E_{10})$. However, we stress that in the spirit of the original 'gradient conjecture' of [149] it would be more natural to regard Ψ as the first component in a faithful (infinite-dimensional) representation $\hat{\Psi} := (\Psi, \partial_a\Psi, \dots)$ of $K(E_{10})$, where the remaining components encode spatial gradients of the gravitino Ψ [70, 71]. As the employed **320** representation is a fully consistent representation of $K(\mathcal{E}_{10})$, its transformations (that will be shown to be in good agreement with supergravity below) will never leave this 320-dimensional (invariant) representation space. A reconciliation of this with the gradient conjecture is beyond the scope of the present analysis.

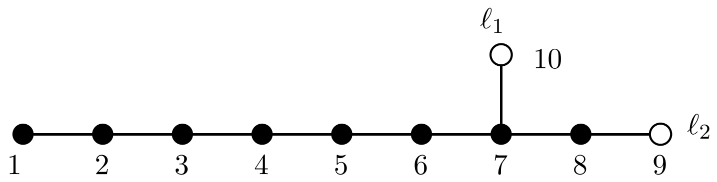


Figure 7.3: *The Dynkin diagram of E_{10} with the nodes associated with the level decomposition indicated in white.*

$E_{10} := \mathfrak{e}_{10(10)}(\mathbb{R})$ of the complex Lie algebra $E_{10}(\mathbb{C})$. The split real form is generated by ten triples (e_i, f_i, h_i), $i = 1, \ldots, 10$, of Chevalley generators, each triple making up a distinguished subalgebra,

$$\mathfrak{sl}_i(2, \mathbb{R}) = \mathbb{R}f_i \oplus \mathbb{R}h_i \oplus \mathbb{R}e_i \subset E_{10}. \tag{7.2.1}$$

These subalgebras are intertwined inside E_{10} according to the stucture of the Dynkin diagram in Figure 7.3. The full structure of the algebra follows from multiple commutators of the form $[e_{i_1}, [e_{i_2}, \cdots [e_{i_{k-1}}, e_{i_k}] \cdots]]$ (and similarly for the f_i's) modulo the so-called Serre relations. We have the standard triangular decomposition

$$E_{10} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+, \tag{7.2.2}$$

where $\mathfrak{h} = \sum_i \mathbb{R}h_i$ is the Cartan subalgebra and the nilpotent parts $\mathfrak{n}_\pm$ are generated by the e_i's and f_i's, respectively. In other words, the subspace $\mathfrak{n}_+$ contains the positive step operators, while $\mathfrak{n}_-$ contains the negative step operators.

The maximal compact subalgebra $K(E_{10}) \subset E_{10}$ is defined as the subalgebra which is pointwise fixed by the Chevalley involution ω,

$$K(E_{10}) := \{x \in E_{10} \mid \omega(x) = x\}, \tag{7.2.3}$$

where ω is defined through its action on each triple (e_i, f_i, h_i):

$$\omega(e_i) = -f_i, \qquad \omega(f_i) = -e_i, \qquad \omega(h_i) = -h_i. \tag{7.2.4}$$

By virtue of the existence of $K(E_{10})$ we have the standard Iwasawa and Cartan decompositions (direct sums of vector spaces),

$$\begin{aligned} E_{10} &= K(E_{10}) \oplus \mathfrak{h} \oplus \mathfrak{n}_+, \qquad &\text{(Iwasawa)} \\ E_{10} &= K(E_{10}) \oplus \mathfrak{p}, \qquad &\text{(Cartan)}, \end{aligned} \tag{7.2.5}$$

which will both be of importance in subsequent sections. The generators belonging to $\mathfrak{p}$ are those which are anti-invariant under ω. Note that the subspace $\mathfrak{p}$ does not close under the Lie bracket, but rather transforms in a representation of $\mathfrak{k}$. On the other hand, the subspace $\mathfrak{b} := \mathfrak{h} \oplus \mathfrak{n}_+$ is a subalgebra of E_{10}, known as a *Borel subalgebra*. Using the Chevalley involution we project an arbitrary generator $x \in E_{10}$ onto the different subspaces of the Cartan decomposition according to

$$\begin{aligned} \mathcal{Q} &:= \frac{1}{2}\big[x+\omega(x)\big] \in K(E_{10}), \\ \mathcal{P} &:= \frac{1}{2}\big[x-\omega(x)\big] \in \mathfrak{p}. \end{aligned} \tag{7.2.6}$$

The Cartan matrix A of E_{10}, deduced from the Dynkin diagram in Figure 7.3, induces an indefinite and non-degenerate bilinear form $\langle\cdot|\cdot\rangle$ on $\mathfrak{h}$ as follows

$$\langle h_i|h_j\rangle := A_{ij}. \tag{7.2.7}$$

By invariance, this bilinear form can be extended to all of E_{10}, and in particular for the Chevalley generators we have

$$\langle e_i|f_j\rangle = \delta_{ij}, \qquad \langle h_i|e_j\rangle = 0, \qquad \langle h_i|f_j\rangle = 0. \tag{7.2.8}$$

This bilinear form will be used in subsequent sections to construct a manifestly E_{10}-invariant Lagrangian.

7.2.2 The IIA Level Decomposition of E_{10} and $K(E_{10})$

To elucidate the relation between the infinite-dimensional algebra E_{10} and the field content of massive type IIA supergravity, we shall perform a decomposition of the adjoint representation of E_{10} into representations of the finite-dimensional subalgebra $A_8 \cong \mathfrak{sl}(9,\mathbb{R})$, defined by nodes $1,\dots,8$ in the Dynkin diagram in Figure 7.3 (see also [146]). Each generator of E_{10} is then represented as an $\mathfrak{sl}(9,\mathbb{R})$-tensor, say $X_{a_1\cdots a_k} \in E_{10}$, where the indices are interpreted as the flat spatial indices of the ten-dimensional supergravity theory and transform in a given irreducible representation described by a set of Dynkin indices, or equivalently by a Young tableau. Since E_{10} is of rank 10 and A_8 is of rank 8, this decomposition is a multilevel $\ell := (\ell_1,\ell_2)$-decomposition, with ℓ_1 being associated with the exceptional node in the Dynkin diagram, while ℓ_2 corresponds to the rightmost node.

A useful alternative point of view on this decomposition is to first perform only the ℓ_1 decomposition with respect to the horizontal $A_9 \cong \mathfrak{sl}(10,\mathbb{R})$ subalgebra consisting of nodes 1 through to 9 in Figure 7.3. As is well-known, at low ℓ_1 levels, this decomposition gives rise to the field content of eleven-dimensional supergravity [89, 149]. We can then view the additional decomposition with respect to ℓ_2 as a 'dimensional reduction' from $D = 11$ to $D = 10$. It is this point of view which enables us to relate the mass term in massive type IIA supergravity to a generator of E_{10} at level $\ell_1 = 4$ in the standard 'M-theory decomposition' of [58,89,149]. This is intriguing since in $D = 11$ the matching between supergravity and E_{10} has only been successful up to $\ell_1 = 3$. Thus, the mass term in $D = 10$ provides a non-trivial check of E_{10} beyond its '$\mathfrak{sl}(10,\mathbb{R})$-covariantized $\mathfrak{e}_8$' subset, i.e. the generators of $\mathfrak{e}_8$ and their images under (the Weyl group of) $\mathfrak{sl}(10,\mathbb{R})$.

The relation between E_{10} and $\mathfrak{e}_{11}$ and the mass deformation parameter of type IIA supergravity has been pointed out in various earlier references. In [153] the massive type IIA theory was reformulated as a non-linear realization where the mass term was associated with a certain nine-form generator of $\mathfrak{e}_{11}$. Shortly after, in [116], it was observed that the mass term in

(ℓ_1, ℓ_2)	$\mathfrak{sl}(9,\mathbb{R})$ Dynkin labels	Generator of E_{10}
$(0,0)$	$[1,0,0,0,0,0,0,1] \oplus [0,0,0,0,0,0,0,0]$	$K^a{}_b$
$(0,0)$	$[0,0,0,0,0,0,0,0]$	T
$(0,1)$	$[0,0,0,0,0,0,0,1]$	E^a
$(1,0)$	$[0,0,0,0,0,0,1,0]$	$E^{a_1a_2}$
$(1,1)$	$[0,0,0,0,0,1,0,0]$	$E^{a_1a_2a_3}$
$(2,1)$	$[0,0,0,1,0,0,0,0]$	$E^{a_1\dots a_5}$
$(2,2)$	$[0,0,1,0,0,0,0,0]$	$E^{a_1\dots a_6}$
$(3,1)$	$[0,1,0,0,0,0,0,0]$	$E^{a_1\dots a_7}$
$(3,2)$	$[1,0,0,0,0,0,0,0]$	$E^{a_1\dots a_8}$
$(3,2)$	$[0,1,0,0,0,0,0,1]$	$E^{a_0\vert a_1\dots a_7}$
$(4,1)$	$[0,0,0,0,0,0,0,0]$	$E^{a_1\dots a_9}$

Table 7.1: *IIA Level decomposition of E_{10} under $\mathfrak{sl}(9,\mathbb{R})$.*

Romans' theory corresponds to a positive real root of E_{10}. This was further elaborated upon in [61] where a truncated version of massive type IIA supergravity in an $SO(9,9)$-covariant formulation was shown to be equivalent to the sigma model for E_{10} in a level decomposition with respect to a $D_9 \cong \mathfrak{so}(9,9)$-subalgebra. This analysis mainly focused on the bosonic sector, but a preliminary analysis of the fermionic sector was initiated by restricting to the zeroth level of the E_{10}-decomposition. It was also pointed out in [58, 94, 146] that the mass term has a natural interpretation as a generator at level four in a decomposition of E_{10} with respect to $A_8 \cong \mathfrak{sl}(9,\mathbb{R})$. From this point of view, the kinetic term in the sigma model associated with this generator naturally gives rise to the mass term of type IIA upon dimensional reduction. This is the viewpoint which we extend in the present work.

Generators of E_{10}

The level decomposition of E_{10} under the $A_8 \cong \mathfrak{sl}(9,\mathbb{R})$ subalgebra up to level $(\ell_1,\ell_2) = (4,1)$ is shown in Table 7.1. At level $(\ell_1,\ell_2) = (0,0)$ there is a copy of $\mathfrak{gl}(9,\mathbb{R}) = \mathfrak{sl}(9,\mathbb{R}) \oplus \mathbb{R}$, as well as a scalar generator associated with the dilaton. The commutation relations at this level are $(a, b = 1, \dots, 9)$

$$\begin{aligned} [K^a{}_b, K^c{}_d] &= \delta^c_b K^a{}_d - \delta^a_d K^c{}_b, \\ [T, K^a{}_b] &= 0, \end{aligned} \tag{7.2.9}$$

and the bilinear form reads

$$\langle K^a{}_b | K^c{}_d \rangle = \delta^a_d \delta^c_b - \delta^a_b \delta^c_d, \qquad \langle T|T\rangle = \frac{1}{2}, \qquad \langle T|K^a{}_b\rangle = 0. \tag{7.2.10}$$

All objects transform as $\mathfrak{gl}(9,\mathbb{R})$ tensors in the obvious way. The positive level generators are obtained through multiple commutators between the (fundamental) generators E^a and E^{ab} on levels $(0,1)$ and $(1,0)$, respectively. For example, the generator on level $(1,1)$ is obtained simply as the commutator

$$[E^{ab}, E^c] = E^{abc}. \tag{7.2.11}$$

All the remaining relevant commutators up to level $(4,1)$ are given in Appendix A.2.1. No generators of E_{10} appear on mixed positive and negative levels, meaning that for any $X_\ell \in E_{10}$ the levels ℓ_1 and ℓ_2 are either both non-positive or both non-negative. In terms of the multilevel $\ell = (\ell_1, \ell_2)$ we shall denote this by $\ell \leq 0$ or $\ell \geq 0$, respectively. This implies that the level decomposition induces a grading of E_{10} into an infinite set of finite-dimensional subspaces $\mathfrak{g}_\ell$ with respect to the multilevel ℓ. Let E_ℓ and $E_{\ell'}$ be arbitrary root vectors in the subspaces $\mathfrak{g}_\ell$ and $\mathfrak{g}_{\ell'}$ of E_{10}. Then a generic commutator, generalizing (7.2.11), takes the form

$$[E_\ell, E_{\ell'}] = E_{\ell+\ell'} \ \in \ \mathfrak{g}_{\ell+\ell'}. \tag{7.2.12}$$

The negative level generators are simply obtained using the Chevalley involution, and for an arbitrary positive level generator $E_\ell \in \mathfrak{g}_\ell$ we define the associated negative generator as

$$F_\ell := -\omega(E_\ell) \in \mathfrak{g}_{-\ell}, \qquad \ell \in \mathbb{Z}^2_{\geq 0}. \tag{7.2.13}$$

Because of the graded structure, commutators between generators in $\mathfrak{g}_\ell$ and $\mathfrak{g}_{-\ell}$ belong to the zeroth subspace $\mathfrak{g}_0$. For example, for the fundamental generator $E^{ab} \in \mathfrak{g}_{(0,1)}$ we have

$$[E^{a_1a_2}, F_{b_1b_2}] = -\frac{1}{2}\delta^{a_1a_2}_{b_1b_2}K + 4\delta^{[a_1}_{[b_1}K^{a_2]}{}_{b_2]} - 2\delta^{a_1a_2}_{b_1b_2}T. \tag{7.2.14}$$

The explicit form of the remaining commutators can be found in Appendix A.2.1.

The "Mass Generator"

Let us now discuss in more detail how the generator associated with the mass term appears in the level decomposition of E_{10}. First consider the decomposition with respect to ℓ_1. At level $\ell_1 = 4$ one finds a generator corresponding to an $\mathfrak{sl}(10,\mathbb{R})$-tensor with 12 indices of the form $E^{\dot{a}|\dot{b}|\dot{c}_1\cdots\dot{c}_{10}}$ (with $\dot{a} = (10, a)$, etc.), which has mixed Young symmetry, i.e. it is antisymmetric in the block of indices $\dot{c}_1, \ldots, \dot{c}_{10}$ and symmetric in $\dot{a}, \dot{b}$. This generator has no physical interpretation in $D = 11$ supergravity. However, consider now the further level decomposition with respect to ℓ_2. This can be realized by doing a dimensional reduction on $E^{\dot{a}|\dot{b}|\dot{c}_1\cdots\dot{c}_{10}}$ along the direction '10'. We are interested in the generator obtained in this way by fixing the first three indices to $\dot{a} = \dot{b} = \dot{c}_1 = 10$, which yields [58, 94]

$$E^{a_1\cdots a_9} := \frac{1}{8}E^{10|10|10a_1\cdots a_9}. \tag{7.2.15}$$

The resulting tensor corresponds to the $(\ell_1, \ell_2) = (4,1)$-generator $E^{a_1\cdots a_9}$ in Table 7.1. This $\mathfrak{sl}(9,\mathbb{R})$-tensor is a nine-form, and is therefore associated with a supergravity potential $A_{a_1\cdots a_9}$ whose field-strength is a top-form in $D = 10$. This is the generator associated with the mass term. From this analysis it is clear that the mass deformation of type IIA supergravity probes E_{10} beyond the realm of what has been successfully verified previously in the context of $D = 11$ supergravity. This is especially interesting due to the fact that the $\ell_1 = 4$-generator $E^{\dot{a}|\dot{b}|\dot{c}_1\cdots\dot{c}_{10}}$ is a 'genuine' E_{10} element, with no components contained in $\mathfrak{e}_8$ nor in $\mathfrak{e}_9$.

Generators of $K(E_{10})$

The level decomposition of E_{10} induces a decomposition of its maximal compact subalgebra $K(E_{10})$ which was defined in (7.2.3). Using the Cartan decomposition, (7.2.5) and (7.2.6), we define compact and noncompact generators, $J_\ell \in K(E_{10})$ and $S_\ell \in \mathfrak{p}$, as follows

$$J_\ell := E_\ell - F_\ell, \qquad S_\ell := E_\ell + F_\ell, \qquad \ell \in \mathbb{Z}^2_{\geq 0} \backslash \{(0,0)\}. \tag{7.2.16}$$

Furthermore, at level $(0,0)$ we have the $\mathfrak{so}(9)$ Lorentz generators

$$M^{ab} := K^a{}_b - K^b{}_a. \tag{7.2.17}$$

This decomposition of $K(E_{10})$ is not a gradation, but rather corresponds to a *filtration* [72], in the sense that an arbitrary commutator between two 'positive level' generators exhibits a graded structure modulo lower-level generators only,

$$[J_\ell, J_{\ell'}] = J_{|\ell-\ell'|} + J_{\ell+\ell'}, \tag{7.2.18}$$

where it is understood that $J_{|\ell-\ell'|} \neq 0$ if and only if $(\ell-\ell') \geq 0$ or $(\ell-\ell') \leq 0$. In other words, $J_{|\ell-\ell'|}$ is non-zero only when the difference $(\ell - \ell')$ involves no mixing between negative and positive levels.

Using (7.2.16), together with the E_{10} commutators in Appendix A.2.1, we may deduce the abstract $K(E_{10})$-relations at each 'level'. Let us consider a few examples to illustrate the procedure. We begin by defining the $K(E_{10})$-generators associated with level $(0,0)$ and the fundamental generators at level $(0,1)$ and $(1,0)$:

$$J^a := E^a - F_a, \qquad J^{ab} := E^{ab} - F_{ab}. \tag{7.2.19}$$

Since there are no generators of mixed level $(1,-1)$ the commutator $[E^a, F_b]$ vanishes, and the commutator between J^{ab} and J^c simply gives

$$[J^{ab}, J^c] = J^{abc}, \tag{7.2.20}$$

with $J^{abc} := E^{abc} - F_{abc}$. Proceeding in the same way for $J^{a_1a_2}$ and $J^{a_1a_2a_3}$, we obtain $J^{a_1\cdots a_5}$ modulo lower level terms,

$$[J^{a_1a_2}, J^{a_3a_4a_5}] = J^{a_1\cdots a_5} - 6\delta^{[a_3a_4}_{a_1a_2} J^{a_5]}. \tag{7.2.21}$$

Note that one may project onto $J^{a_1\cdots a_5}$ as follows,

$$J^{a_1\cdots a_5} = \left[J^{[a_1a_2}, J^{a_3a_4a_5]}\right] . \tag{7.2.22}$$

Here, all indices are $\mathfrak{so}(9)$ vector indices and are raised and lowered with the Euclidean δ_{ab} so that the position does not matter. The other relevant $K(E_{10})$-commutators are listed in Appendix A.2.2.

7.2.3 Spinorial Representations

The fermionic degrees of freedom in the sigma model for $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ transform in spinorial representations of $K(E_{10})$. The two relevant $K(E_{10})$ representations are finite-dimensional (unfaithful) and of dimensions 32 and 320, respectively. In the decomposition of $K(E_{10})$ associated with eleven-dimensional supergravity these representations transform as a 32-dimensional Dirac spinor representation ϵ of $\mathfrak{so}(10) \subset K(E_{10})$ and a 320-dimensional vector-spinor representation $\Psi_{\dot{a}}$, $\dot{a} = (10, a)$, of $\mathfrak{so}(10) \subset K(E_{10})$, identified with the gravitino [69–72]. Upon reduction to the IIA theory (through the additional level decomposition with respect to ℓ_2) the gravitino decomposes into a 32-dimensional spinor Ψ_{10} and a 288-dimensional vector spinor Ψ_a of $\mathfrak{so}(9)$. However, because they both descend from $\Psi_{\dot{a}}$, these two representations will mix under $K(E_{10})$ [154]. No such complication arises for the supersymmetry parameter ϵ, for which we keep the same notation in the IIA picture. The spinor Ψ_{10} will be associated with the ten-dimensional dilatino, while the vector-spinor Ψ_a is related to the gravitino.

For the first three levels, the transformation properties of the spinor ϵ are

$$\begin{aligned} M^{ab} \cdot \epsilon &= \frac{1}{2}\Gamma^{ab}\epsilon, \\ J^{a} \cdot \epsilon &= \frac{1}{2}\Gamma_{10}\Gamma^{a}\epsilon, \\ J^{ab} \cdot \epsilon &= \frac{1}{2}\Gamma_{10}\Gamma^{ab}\epsilon, \end{aligned} \tag{7.2.23}$$

where Γ^a and Γ^{ab} are $\mathfrak{so}(9)$ Γ-matrices (our Γ-matrix conventions are given in Appendix A.1.1). The higher level transformations are now defined through the abstract $K(E_{10})$-relations; for example:

$$\begin{aligned} J^{a_1a_2a_3} \cdot \epsilon &:= \left[J^{a_1a_2}, J^{a_3}\right] \cdot \epsilon = \frac{1}{2}\Gamma^{a_1a_2a_3}\epsilon, \\ J^{a_1\cdots a_5} \cdot \epsilon &:= \left[J^{[a_1a_2}, J^{a_3a_4a_5]}\right] \cdot \epsilon = \frac{1}{2}\Gamma_{10}\Gamma^{a_1\cdots a_5}\epsilon. \end{aligned} \tag{7.2.24}$$

Similarly, we have for the spinor component Ψ_{10} the following low-level transformations:

$$\begin{aligned} M^{a_1a_2} \cdot \Psi_{10} &= \frac{1}{2}\Gamma^{a_1a_2}\Psi_{10} \\ J^{a} \cdot \Psi_{10} &= \frac{1}{2}\Gamma_{10}\Gamma^{a}\Psi_{10} + \Psi^{a}, \\ J^{a_1a_2} \cdot \Psi_{10} &= \frac{1}{6}\Gamma_{10}\Gamma^{a_1a_2}\Psi_{10} + \frac{4}{3}\Gamma^{[a_1}\Psi^{a_2]}, \end{aligned} \tag{7.2.25}$$

showing explicitly the mixing between Ψ_{10} and Ψ_a under $K(E_{10})$. Finally, we also have for Ψ_a:

$$\begin{aligned} M^{a_1a_2} \cdot \Psi_b &= \frac{1}{2}\Gamma^{a_1a_2}\Psi_b + 2\delta_b^{[a_1}\Psi^{a_2]}, \\ J^{a} \cdot \Psi_b &= \frac{1}{2}\Gamma_{10}\Gamma^{a}\Psi_b - \delta_b^{a}\Psi_{10}, \end{aligned}$$

$$\begin{aligned} J^{a_1a_2}\cdot\Psi_b &= \frac{1}{2}\Gamma_{10}\Gamma^{a_1a_2}\Psi_b - \frac{4}{3}\Gamma_{10}\delta_b^{[a_1}\Psi^{a_2]} + \frac{2}{3}\Gamma_{10}\Gamma_b{}^{[a_1}\Psi^{a_2]} \\ &\quad + \frac{4}{3}\delta_b^{[a_1}\Gamma^{a_2]}\Psi_{10} - \frac{1}{3}\Gamma_b{}^{a_1a_2}\Psi_{10}. \end{aligned} \tag{7.2.26}$$

More details on these $K(E_{10})$-representations can be found in Appendix A.2.2. We note that we could also have redefined the $SO(9)$ spinor Ψ_a by a shift with $\Gamma_a\Gamma^{10}\Psi_{10}$ as one does in Kaluza–Klein reduction (cf. (A.1.18)) but refrain from doing so here.

7.2.4 The Non-Linear Sigma Model for $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$

A non-linear sigma model with rigid $\mathcal{E}_{10}$-invariance and local $K(\mathcal{E}_{10})$-invariance may now be constructed using the properties of E_{10} described in previous sections. By virtue of the Iwasawa decomposition we can always choose a coset representative in the partial 'Borel gauge' by taking

$$\mathcal{V}(t) := \mathcal{V} = \mathcal{V}_0 e^{\phi T} e^{A\star E} \in K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}, \tag{7.2.27}$$

where $\mathcal{V}_0 = \exp\, h^a{}_b K^b{}_a$ represents the $GL(9,\mathbb{R})$ inverse vielbein $e_a{}^m$, while $e^{A\star E}$ contains the positive step operators of E_{10}. We call this a partial Borel gauge since $\mathcal{V}_0$ is not constrained to contain only positive step operators of $\mathfrak{gl}(9,\mathbb{R})$. In other words, $\mathcal{V}_0$ is an arbitrary $GL(9,\mathbb{R})$-matrix and not a representative of the coset $SO(9)\backslash GL(9,\mathbb{R})$. With some abuse of terminology, we shall sometimes refer to the part of $\mathcal{E}_{10}$ parametrized by $\mathcal{V}$ as the 'Borel subgroup', and denote it by $\mathcal{E}_{10}^+$ [3]. The positive level part $e^{A\star E}$ of $\mathcal{V}$ is defined as

$$e^{A\star E} := \exp\big[A_m(t)E^m\big]\exp\big[\frac{1}{2}A_{m_1m_2}(t)E^{m_1m_2}\big]\exp\big[\frac{1}{3!}A_{m_1m_2m_3}(t)E^{m_1m_2m_3}\big]\cdots \tag{7.2.28}$$

with similar exponentials occurring for the higher levels. Note that in this expression the indices $m_1, m_2, \ldots$ are $\mathfrak{sl}(9,\mathbb{R})$-indices, and hence correspond to curved spatial indices from a supergravity point of view.

The coset representative $\mathcal{V}$ transforms under global $g \in \mathcal{E}_{10}$-transformations from the right and local $k \in K(\mathcal{E}_{10})$-transformations from the left:

$$\mathcal{V} \longmapsto k\mathcal{V}g. \tag{7.2.29}$$

From $\mathcal{V}$ we construct the Lie algebra-valued Maurer–Cartan form

$$\partial_t\mathcal{V}\mathcal{V}^{-1} = \mathcal{P} + \mathcal{Q}, \tag{7.2.30}$$

where we also employed the Cartan decomposition, according to (7.2.6). The transformation property of $\mathcal{V}$ in (7.2.29) implies that the coset part $\mathcal{P} \in \mathfrak{p}$ of the Maurer-Cartan form is globally $\mathcal{E}_{10}$-invariant while transforms covariantly under $K(\mathcal{E}_{10})$:

$$K(\mathcal{E}_{10}) \;:\; \mathcal{P} \longmapsto k\mathcal{P}k^{-1}. \tag{7.2.31}$$

On the other hand, $\mathcal{Q} \in K(E_{10})$ properly transforms as a connection,

$$K(\mathcal{E}_{10}) \;:\; \mathcal{Q} \longmapsto k\mathcal{Q}k^{-1} + \partial_t k k^{-1}. \tag{7.2.32}$$

[3] This differs from the 'true' Borel subgroup $B := \exp\,\mathfrak{b} \subset \mathcal{E}_{10}$ through the negative root generators in $\mathcal{V}_0$ (see Section 7.2.1 for the definition of $\mathfrak{b}$). $\mathcal{E}_{10}^+$ is only a parabolic subgroup of $\mathcal{E}_{10}$.

Using the bilinear form on E_{10}, we can now construct a manifestly $\mathcal{E}_{10} \times K(\mathcal{E}_{10})_{\text{local}}$-invariant Lagrangian as follows [89, 149]

$$\mathcal{L}^{[B]}_{K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}} = \frac{1}{4} n(t)^{-1} \big\langle \mathcal{P} \big| \mathcal{P} \big\rangle , \tag{7.2.33}$$

where the lapse function $n(t)$ ensures invariance under reparametrizations of the geodesic parameter t. We have also included a superscript $[B]$ to emphasize that this Lagrangian is only the bosonic part of a sigma model which also includes the fermionic degrees of freedom in the **320** of $K(\mathcal{E}_{10})$ to be introduced shortly. The equations of motion for $n(t)$ enforces a lightlike (Hamiltonian) constraint on the dynamics,

$$\big\langle \mathcal{P} \big| \mathcal{P} \big\rangle = 0, \tag{7.2.34}$$

while the equations of motion for $\mathcal{P}$ (in the gauge $n = 1$) read

$$\mathcal{D}\mathcal{P} := \partial_t \mathcal{P} - [\mathcal{Q}, \mathcal{P}] = 0, \tag{7.2.35}$$

where we defined the $K(\mathcal{E}_{10})$-covariant derivative $\mathcal{D}$. Equation (7.2.35) encodes the dynamics of the bosonic sector of the sigma model and is written out in detail in Appendix A.3.1.

The fermionic degrees of freedom are included in the Lagrangian through the spinor representation Ψ as follows [70–72]

$$\mathcal{L}^{[F]}_{K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}} = -\frac{i}{2} \big\langle \Psi \big| \mathcal{D}\Psi \big\rangle , \tag{7.2.36}$$

where the bracket now denotes an invariant inner product on the representation space. The associated 'Dirac equation' reads

$$\mathcal{D}\Psi := \partial_t \Psi - \mathcal{Q} \cdot \Psi = 0, \tag{7.2.37}$$

where it is understood that the connection $\mathcal{Q}$ acts on Ψ in the vector-spinor representation constructed in the previous section. Equation (7.2.37) is written in terms of the full 320-dimensional spinor Ψ, which encodes both the 'gravitino' Ψ_a and the 'dilatino' Ψ_{10}. The separate equations of motion for these fields can be written out as follows

$$\begin{aligned} \mathcal{D}\Psi_a &:= \partial_t \Psi_a - \mathcal{Q} \cdot \Psi_a = 0, \\ \mathcal{D}\Psi_{10} &:= \partial_t \Psi_{10} - \mathcal{Q} \cdot \Psi_{10} = 0. \end{aligned} \tag{7.2.38}$$

Recall that the $K(E_{10})$-action on Ψ_a contains Ψ_{10}-terms, and vice versa. This is important since the same mixing between the gravitino and the dilatino occurs in the corresponding supergravity equations of motion (see Appendix A.1.4). The fermionic equations of motion in (7.2.38) are written out in more detail in Appendix A.3.2.

The bosonic equations of motion (7.2.35) were written for the gauge choice $n = 1$. The lapse function n has a superpartner Ψ_t, which is a Dirac spinor under $K(E_{10})$, and the associated supersymmetry transformations are

$$\begin{aligned} \delta_\epsilon n &= i\epsilon^T \Psi_t, \\ \delta_\epsilon \Psi_t &= \mathcal{D}\epsilon. \end{aligned} \tag{7.2.39}$$

The fermionic equations of motion are then valid in the 'supersymmetric gauge' $\Psi_t = 0$. The associated constraint (analogous to the Hamiltonian constraint) that should be imposed is the supersymmetry constraint which states that the spin Ψ is orthogonal to the velocity $\mathcal{P}$. The full form of this constraint is unknown due to the fact that it is not known how to supersymmetrize the $\mathcal{E}_{10}$ model correctly. Nevertheless, low level expressions have been obtained in [70, 72], which the reader should also consult for further discussions on this point.

Let us now analyse these properties of the geodesic sigma model in more detail by utilizing the level decomposition of E_{10} with respect to $\mathfrak{sl}(9,\mathbb{R})$. Expanding the coset element $\mathcal{P}$ up to level $(3,2)$ we obtain

$$\mathcal{P} = \partial_t \phi T + \frac{1}{2} p_{ab}(K^a{}_b + K^b{}_a) + \sum_{\ell>0} P_\ell \star S_\ell \tag{7.2.40}$$

and the associated expansion for $\mathcal{Q}$ is given by

$$\mathcal{Q} = \frac{1}{2} q_{ab} M^{ab} + \sum_{\ell>0} Q_\ell \star J_\ell, \tag{7.2.41}$$

where $\star$ schematically denotes the coupling between the dynamical fields P_ℓ and the associated generators S_ℓ. The explicit form of this expansion will be given below. In Borel gauge, the fields at non-zero levels are the same in $\mathcal{Q}$ and $\mathcal{P}$:

$$Q_\ell \equiv P_\ell, \qquad \ell \in \mathbb{Z}^2_{\geq 0} \backslash \{(0,0)\}\,. \tag{7.2.42}$$

The explicit form of the higher level terms in the expansion of $\mathcal{P}$ reads

$$\begin{aligned}\sum_{\ell>0} P_\ell \star S_\ell \quad := \quad & e^{3\phi/4} P_{a_1} S^{a_1} + \frac{1}{2} e^{-\phi/2} P_{a_1a_2} S^{a_1a_2} + \frac{1}{3!} e^{\phi/4} P_{a_1a_2a_3} S^{a_1a_2a_3} \\ & + \frac{1}{5!} e^{-\phi/4} P_{a_1\ldots a_5} S^{a_1\ldots a_5} + \frac{1}{6!} e^{\phi/2} P_{a_1\ldots a_6} S^{a_1\ldots a_6} + \frac{1}{7!} e^{-3\phi/4} P_{a_1\ldots a_7} S^{a_1\ldots a_7} \\ & + \frac{1}{8!} P_{a_1\ldots a_8} S^{a_1\ldots a_8} + \frac{1}{8!} P_{a_0|a_1\ldots a_7} S^{a_0|a_1\ldots a_7} + \frac{1}{9!} e^{-5\phi/4} P_{a_1\ldots a_9} S^{a_1\ldots a_9} + \ldots\end{aligned} \tag{7.2.43}$$

with a similar expression for $\mathcal{Q}$ with the S_ℓ's replaced by the corresponding J_ℓ's.

From a given parametrization of $\mathcal{V}$ as in (7.2.28) one can work out explicit expressions for $\mathcal{P}$ in terms of the 'potentials' A_ℓ (which *a fortiori* carries flat indices as it transforms under $K(E_{10})$) which appear in the construction of $\mathcal{V}$. For example, one finds

$$P_a = \frac{1}{2} e_a{}^m \partial_t A_m = \frac{1}{2} e_a{}^m D A_m \tag{7.2.44}$$

in terms of the 'covariant derivatives' of [149]. Here, we have written $\mathcal{V}_0 = e_a{}^m$ as an inverse $GL(9,\mathbb{R})$ vielbein. We do not require the exact expressions for the higher level components of $\mathcal{P}$ for establishing the correspondence and their expression will be more complicated due to the appearance of additional terms, arising when expanding the Maurer-Cartan form $\partial_t \mathcal{V}\mathcal{V}^{-1}$ using the Baker–Campbell–Hausdorff formula $de^X e^{-X} = dX + \frac{1}{2}[X, dX] + \ldots$.

The geodesic equations $\partial_t \mathcal{P} = [\mathcal{Q}, \mathcal{P}]$ can be written conveniently by treating the $\mathcal{V}_0$ contribution separately in a partially covariant derivative $D^{(0)}$, see [89]. For example, for the fundamental generators at level $(0,1)$ and $(1,0)$ the equations of motion become (for $n=1$)

$$\begin{aligned} D^{(0)}(e^{3\phi/2}P_a) &= -e^{\phi/2}P_{ac_1c_2}P_{c_1c_2} + \frac{2}{5!}e^{\phi}P_{ac_1\ldots c_5}P_{c_1\ldots c_5} + \frac{12}{8!}P_{ac_1\ldots c_7}P_{c_1\ldots c_7} \\ &\quad + \frac{1}{4\cdot 7!}P_{a|c_1\ldots c_7}P_{c_1\ldots c_7} \qquad (7.2.45) \\ D^{(0)}(e^{-\phi}P_{a_1a_2}) &= 2e^{\phi/2}P_{a_1a_2c}P_c + \frac{1}{3}e^{-\phi/2}P_{a_1a_2c_1c_2c_3}P_{c_1c_2c_3} \\ &\quad + \frac{2}{5!}e^{-3\phi/2}P_{a_1a_2c_1\ldots c_5}P_{c_1\ldots c_5} + \frac{2}{7!}e^{-5\phi/2}P_{a_1a_2c_1\ldots c_7}P_{c_1\ldots c_7} \\ &\quad + \frac{1}{6!}P_{a_1a_2c_1\ldots c_6}P_{c_1\ldots c_6} + \frac{1}{4\cdot 5!}P_{c_1|c_2\ldots c_6a_1a_2}P_{c_1\ldots c_6}. \qquad (7.2.46) \end{aligned}$$

In Section 7.3 we will show that these equations are equivalent to the equations of motion for the electric fields F_{ta} and F_{tab}, respectively. The equations for all the remaining levels are given in Appendix A.3.1. The fermionic equations of motion can be found in A.3.2 where the connection term $\mathcal{Q}$ is evaluated in the vector-spinor representation. The supersymmetry variation (7.2.39) requires the evaluation of $\mathcal{Q}$ in the Dirac-spinor representation, the resulting expression can be found in A.3.3.

7.3 The Correspondence

In this section we make the $\mathcal{E}_{10}$/massive IIA correspondence explicit by giving a dictionary between the dynamical variables of the $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$-sigma model and the fields of massive IIA supergravity. The comparison cannot be done at the level of the respective Lagrangians since the $\mathcal{E}_{10}$ sigma model naturally incorporates kinetic terms for all fields as well as their duals, which is not the case for the IIA Lagrangian. The matching is rather done at the level of the equations of motion, where we will see that bosonic equations of motion and Bianchi identities on the supergravity side all become associated with geodesic equations of motion on the $\mathcal{E}_{10}$ side. Similarly, the fermionic supergravity equations will be associated with the Dirac equation of the spinning coset particle.

7.3.1 Bosonic Equations of Motion and Truncation

To be able to compare the equations of motion on the supergravity side (A.1.5) and (A.1.6) to the ones on the sigma model side (7.2.35), spelt out in Appendix A.3.1, we need to rewrite the former. As is customary in the correspondence between $\mathcal{E}_{10}$ and supergravity we split the indices $\alpha = (0, a)$ into temporal and spatial indices and also adopt a pseudo-Gaussian gauge for the ten-dimensional vielbein:

$$e_\mu{}^\alpha = \begin{pmatrix} N & 0 \\ 0 & e_m{}^a \end{pmatrix}. \qquad (7.3.1)$$

In addition we demand that the spatial trace of the spin connection vanishes (see Appendix A.1.1 for our conventions)

$$\omega_{a\,ab} = 0 \quad \Rightarrow \quad \Omega_{ba\,a} = 0 \qquad (7.3.2)$$

and write the tracefree spin connection as $\tilde{\omega}_{a\,bc}$ as a reminder. We also choose temporal gauges for all supergravity gauge potentials,

$$A_t = 0\,, \quad A_{tm} = 0\,, \quad A_{tm_1m_2} = 0\,. \tag{7.3.3}$$

Moreover, we can only expect that a truncated version of the supergravity equation corresponds to the coset model equations. This truncation was originally devised in the context of eleven-dimensional supergravity, where it was strongly motivated by the billiard analysis of the theory close to a spacelike singularity (the 'BKL-limit') [89, 149]. Recall from Chapter 3 that in this limit spatial points decouple and the dynamics becomes effectively time-dependent, ensuring that the truncation is a valid one in this regime. In this chapter, we analyse the same question in the context of massive IIA supergravity, and an identical procedure requires the truncation of a set of spatial gradients. These can be obtained from a BKL-type analysis of massive IIA and their full list is presented in Appendix A.1.5. Notice that except for the expression involving the mass, all the spatial gradients to be truncated away can be obtained by dimensional reduction of the eleven-dimensional truncation.

Let us illustrate the implications of this truncation on the supergravity equations in some detail for an explicit example. The following truncation of massive IIA supergravity can be deduced from the billiard analysis:

$$\partial_a \left(N e^{3\phi/4} F_{b_1b_2} \right) = 0. \tag{7.3.4}$$

The effect of this truncation appears in the equation of motion for the two-form field strength $F_{\alpha\beta}$ (see (A.1.5a) in Appendix A.1.3). After splitting time and space indices, the space component of this equation reads[4]

$$\begin{aligned} D^{(0)}(e^{3\phi/2}F_{tb}) \;\; = \;\; & -\frac{1}{2}e^{\phi/2}F_{ta_1a_2b}F_{ta_1a_2} + \frac{1}{3!}e^{\phi/2}N^2F_{a_1\ldots a_3b}F_{a_1\ldots a_3} \\ & +\frac{3}{4}e^{3\phi/2}N^2\partial_a\phi F_{ab} - \frac{1}{2}e^{3\phi/2}N^2\Omega_{a_1a_2b}F_{a_1a_2} \\ & +Ne^{3\phi/4}\partial_a\left(Ne^{3\phi/4}F_{ab}\right)\,. \end{aligned} \tag{7.3.5}$$

Using the truncation (7.3.4), the last term on the right hand side vanishes, and consequently the equation to compare with the $\mathcal{E}_{10}$ equation of motion (7.2.45) is (7.3.5) without the bottom line. Let us emphasize that the truncation (7.3.4) also follows from dimensional reduction of the associated truncation in the equation of motion for the electric field $F_{0\dot{a}_1\dot{a}_2\dot{a}_3}$ in eleven dimensions [89, 149]. Moreover, this example makes clear the important comment that the truncation we impose is *not* equivalent to discarding all spatial gradients on the supergravity side, since it is clear that spatial derivatives of the one-form potential A_a are implicitly contained in $F_{a_1a_2}$.

From comparing in detail the supergravity equations (Appendix A.1) with the truncations applied with the bosonic equations of the coset model (Appendix A.3.1) one can derive a correspondence between the components of the coset velocity $\mathcal{P}$ and the fields of supergravity. The dictionary is given in Table 7.2 where, in the last line, $e = \det\, e_m{}^a = \sqrt{\det\, g_{mn}} = \sqrt{g}$. It

[4] $D^{(0)}$ here contains the contributions from the time derivatives of the spatial vielbein $e_m{}^a$ and will be identified with the corresponding $\mathcal{E}_{10}$ coset derivative operator below.

(ℓ_1, ℓ_2)	$\mathcal{E}_{10}$ fields	Bosonic fields of supergravity
$(0,0)$	p_{ab}	$-N\omega_{a\,b0}$
$(0,0)$	$q_{a_1a_2}$	$-N\omega_{0\,a_1a_2}$
$(0,1)$	P_a	$\frac{1}{2}NF_{0a}$
$(1,0)$	$P_{a_1a_2}$	$\frac{1}{2}NF_{0a_1a_2}$
$(1,1)$	$P_{a_1a_2a_3}$	$\frac{1}{2}NF_{0a_1a_2a_3}$
$(2,1)$	$P_{a_1\cdots a_5}$	$\frac{1}{2\cdot 4!}Ne^{\phi/2}\epsilon_{a_1\cdots a_5}{}^{b_1b_2b_3b_4}F_{b_1b_2b_3b_4}$
$(2,2)$	$P_{a_1\cdots a_6}$	$\frac{1}{2\cdot 3!}Ne^{-\phi}\epsilon_{a_1\cdots a_6}{}^{b_1b_2b_3}F_{b_1b_2b_3}$
$(3,1)$	$P_{a_1\cdots a_7}$	$-\frac{1}{2\cdot 2!}Ne^{3\phi/2}\epsilon_{a_1\cdots a_7}{}^{b_1b_2}F_{b_1b_2}$
$(3,2)$	$P_{a_1\cdots a_8}$	$-\frac{1}{2}N\epsilon_{a_1\cdots a_8}{}^{b}\partial_b\phi$
$(3,2)$	$P_{a_0\|a_1\cdots a_7}$	$2N\epsilon_{a_1\cdots a_7}{}^{b_1b_2}\tilde{\Omega}_{b_1b_2\,a_0}$
$(4,1)$	$P_{a_1\cdots a_9}$	$\frac{1}{2}Ne^{5\phi/2}\epsilon_{a_1\ldots a_9}m$
-	n	Ne^{-1}

Table 7.2: *Bosonic dictionary: This table shows the correspondence between $\mathcal{E}_{10}$ fields and bosonic fields of massive IIA supergravity that one obtains by considering the equations of motion and the Bianchi equations.*

is perfectly consistent with identifying the form equations of motion (A.1.5) with the sigma model expressions (A.3.2a) to (A.3.2c), the Bianchi identities (A.1.4) with the equations (A.3.2d) to (A.3.2g) and the dilaton equation of motion (A.1.6a) with (A.3.1a).

The remaining equation, the Einstein equation (A.1.6b), does not fit perfectly in this picture. More precisely, two terms do not match completely with (A.3.1b). One is similar to the mismatch in the A_9 decomposition relevant to $D = 11$ supergravity and is a contribution to the Ricci tensor R_{ab} of the form $\Omega_{a\,cd}\Omega_{b\,dc}$ [89]. The other term is a mismatch of the coefficient in the energy momentum tensor of the dilaton $T_{ab} \sim \partial_a\phi\partial_b\phi$; the coefficient is off by a factor two. This can be traced back to $D = 11$ where both mismatches were part of the $D = 11$ Ricci tensor. In this sense this is not a new discrepancy but a known one. It is to be noted that all the terms involved in the mismatch are related to contributions to the Lagrangian which would give rise to walls corresponding to imaginary roots in the cosmological billiards picture [149]. There is no mismatch in the equation of motion of the dilaton ϕ since in the reduction the missing term in $D = 11$ does not contribute to this equation.

Let us study the effect of the mass term more closely. In the bosonic equations, the mass appears in five places: the dilaton equation (A.1.6a), the Einstein equation (A.1.6b), the equation of motion for $F_{ta_1a_2}$ (A.1.5b), the Bianchi equation for $F_{a_1a_2}$ (A.1.4a) and of course in its own equation of motion $\partial_t m = 0$. It is remarkable that the contribution of the real root corresponding to the mass deformation enters all equivalent sigma-model equations correctly, even though it is beyond the realm of E_8 generators and above height 30. In particular, the $\mathcal{E}_{10}$-invariant sigma model produces the right potential for the scalar ϕ. This is possible since the supergravity potential is positive definite in agreement with positive definiteness of the $\mathcal{E}_{10}$-invariant Lagrangian (7.2.33) away from the Cartan subalgebra. By contrast, for gauged deformations in lower dimensions where the supergravity potential is indefinite and not reproduced fully by $\mathcal{E}_{10}$ [152]. From this point of view the massive Romans theory seems to be special since the $\mathcal{E}_{10}$ model reproduces correctly all the effects of the deformation.

7.3.2 The Truncation Revisited

The truncation we applied to supergravity, using billiard arguments, also proves useful for ensuring the consistency of the dictionary. Indeed, notice that all the sigma model variables depend on (sigma model) 'time' only, and hence we must demand that their spatial gradients vanish:

$$\partial_a \mathcal{P}(t) = 0. \tag{7.3.6}$$

Applied to the Maurer–Cartan expansion, this equation translates to constraints on the supergravity variables upon using the dictionary in Table 7.2. Of course, (7.3.6) does not really make sense on the sigma model side, where spatial gradients have no meaning, but must rather be understood as a convenient way of encoding the relevant truncations on the supergravity side. Nevertheless, the truncations encoded in (7.3.6) correspond precisely to the truncations we have just imposed. Let us illustrate this for the example of the truncation on the magnetic field in (7.3.5). The level $(3,1)$ part of the expansion of the Maurer-Cartan form contains the following term (see (7.2.43))

$$\mathcal{P}(t)\big|_{(3,1)} = \frac{1}{7!} e^{-3\phi/4} P_{a_1\cdots a_7}(t) S^{a_1\cdots a_7}. \tag{7.3.7}$$

The coefficient of the generator $S^{a_1\cdots a_7} \in E_{10}$ is a dynamical quantity which is purely time-dependent. This implies that on the supergravity side we must make sure that the corresponding quantity is also purely time-dependent, i.e. we must demand

$$\partial_a (e^{-3\phi/4} P_{a_1\cdots a_7}) = 0, \tag{7.3.8}$$

which, after using the dictionary in Table 7.2, precisely yields the truncation in (7.3.4). Similarly, one sees that requiring (7.3.6) for each term of the Maurer-Cartan expansion corresponds to the truncations of Appendix A.1.5.

7.3.3 Fermionic Equations of Motion

To make contact between the fermionic equations of Romans' theory, (equations of motion (A.1.7) and (A.1.11) and supersymmetry variation (7.1.4)), and the Kac-Moody side of the story, (equations of motion (A.3.3) and (A.3.4) and supersymmetry variation (A.3.5)), we must make some further field redefinitions. We redefine the gravitino components, the dilatino and the supersymmetry parameter as

$$\begin{aligned}
\tilde{\psi}_0 &\equiv g^{1/4}\Big(\psi_0 - \Gamma_0 \Gamma^a \psi_a\Big), \\
\tilde{\psi}_a &\equiv g^{1/4}\Big(\psi_a - \frac{1}{12}\Gamma_a \lambda\Big),
\end{aligned}$$

$K(\mathcal{E}_{10})$ representations	Fermionic fields of supergravity
Ψ_a	$\tilde{\psi}_a = g^{1/4}\Big(\psi_a - \frac{1}{12}\Gamma_a\lambda\Big)$
$\Gamma^{10}\Psi_{10}$	$\tilde{\lambda} = \frac{2}{3}g^{1/4}\lambda$
ϵ	$\tilde{\varepsilon} = g^{-1/4}\varepsilon$
Ψ_t	$\tilde{\psi}_t = ng^{1/4}\Big(\psi_0 - \Gamma_0\Gamma^a\psi_a\Big)$

Table 7.3: *Fermionic dictionary: This table presents the relation between the spinor and vector spinor unfaithful representations of $K(E_{10})$ and the fermionic fields of massive type IIA supergravity. The first half is obtained by requiring the equations of motion for the fermions to match with the $K(E_{10})$ equation and the second half comes from the supersymmetry variation of ψ_t.*

$$\begin{aligned} \tilde{\lambda} &\equiv \frac{2}{3}g^{1/4}\lambda, \\ \tilde{\varepsilon} &\equiv g^{-1/4}\varepsilon. \end{aligned} \tag{7.3.9}$$

With these redefinitions, using the bosonic dictionary obtained in the previous section and in the gauge

$$\tilde{\psi}_0 = 0, \tag{7.3.10}$$

we can show that the equations of motion of the fermions of massive IIA supergravity (A.1.7) and (A.1.11) are equivalent to the Kac-Moody fermionic spinor equations (A.3.3) and (A.3.4) if we assume a correspondence between the unfaithful representation of $K(E_{10})$ and the redefined fermionic fields displayed in the first half of Table 7.3.

To illustrate how the correspondence is proved, let us consider the mass terms of the equation of motion for the gravitino. On the supergravity side, one looks at the spatial components of the equation of motion (A.1.11). We only keep the terms involving the time derivative of ψ_a and the mass:

$$\partial_0\psi_a + \frac{1}{16}e^{5\phi/4}m\Gamma^0\Gamma_{ab}\psi^b + \frac{5}{16}e^{5\phi/4}m\Gamma^0\psi_a + \frac{5}{64}e^{5\phi/4}m\Gamma^0\Gamma_a\lambda + \cdots = 0. \tag{7.3.11}$$

One the sigma model side, we must consider the terms at level $(4,1)$ of the explicit version of the equation of motion (A.3.4), that give

$$\partial_t\Psi_a - \frac{1}{2\cdot 9!}e^{-5\phi/4}P_{b_1\cdots b_9}\Gamma_{10}\Gamma^{b_1\cdots b_9}\Psi_a + \frac{12}{9!}e^{-5\phi/4}P_{b_1\cdots b_9}\Gamma_{10}\Gamma_a{}^{b_1\cdots b_8}\Psi^{b_9} + \cdots = 0. \tag{7.3.12}$$

Using the bosonic and fermionic dictionaries given in Tables 7.2 and 7.3, it is now a purely algebraic exercise to see that (7.3.11) and (7.3.12) are equivalent.

7.3.4 Supersymmetry Variations of Fermions

Considering the supersymmetry variation of the gravitino provides us with a consistency check of the previously obtained bosonic and fermionic dictionaries. To this end it is natural to define

$$\tilde{\psi}_t \equiv n\tilde{\psi}_0, \tag{7.3.13}$$

where t is considered as a 'vector index' along the world line with respect to the einbein n.

Using the previous redefinitions yields the following supersymmetry transformation on the redefined gravitino

$$\begin{aligned}
\delta_{\vec{e}}\tilde{\psi}_t \;=\;& \partial_t\vec{e} + \frac{1}{4}g^{-1}\partial_t g\vec{e} + \frac{1}{4}N\omega_{0ab}\Gamma^{ab}\vec{e} + \frac{1}{4}Ne^{5\phi/4}m\Gamma_0\vec{e} \\
&+\frac{1}{4\cdot 4!}e^{\phi/4}N\Gamma^{abcd}\Gamma_0 F_{abcd}\vec{e} - \frac{1}{4\cdot 3!}e^{\phi/4}N\Gamma^{abc}F_{0abc}\vec{e} \\
&-\frac{1}{8}e^{3\phi/4}N\Gamma^{ab}\Gamma_0\Gamma_{10}F_{ab}\vec{e} + \frac{1}{4}e^{3\phi/4}N\Gamma^{a}\Gamma_{10}F_{0a}\vec{e} \\
&-\frac{1}{24}e^{-\phi/2}N\Gamma^{abc}\Gamma_0\Gamma_{10}F_{abc}\vec{e} - \frac{1}{8}e^{-\phi/2}N\Gamma^{ab}\Gamma_{10}F_{0ab}\vec{e} \\
&+\frac{1}{4}N\omega_{abc}\Gamma^{abc}\Gamma_0\vec{e} - N\Gamma_0\Gamma^a\Big[\partial_a\vec{e} + \frac{1}{4}g^{-1}\partial_a g\vec{e}\Big]. \qquad (7.3.14)
\end{aligned}$$

We can now identify the right hand side of (7.3.14) with the $K(\mathcal{E}_{10})$-covariant derivative acting on the 32-dimensional spinor representation ϵ (7.2.39), given explicitly in (A.3.5), using the bosonic and fermionic dictionaries already computed (Tables 7.2 and 7.3). We then obtain the second half of Table 7.3, that is, the relations for the supersymmetry parameter and the time component of the gravitino.

In conclusion, we see that all the fermionic equations of motion as well as the supersymmetry variation of ψ_t match with the $K(\mathcal{E}_{10})\backslash\mathcal{E}_{10}$ fermionic theory. In particular, we notice that the mass enters these equations correctly everywhere.

Part II

U-Duality and Arithmetic Structures in String Theory

8

Aspects of String Duality and Quantum Corrections

This chapter is intended to serve as a motivation for the topics treated in **Part II** of this thesis.[1] The main philosophy that we shall advocate is that quantum corrections in string theory are strongly constrained by invariance under discrete (U-)duality groups $G(\mathbb{Z})$, and in many instances may be completely summed up in terms of certain $G(\mathbb{Z})$-invariant automorphic forms. The purpose of **Part II** is to understand the construction of such automorphic forms, and how to extract the physical information which they encode. In this first chapter, we shall discuss general aspects of these techniques, as well as consider some specific examples in more detail. The main example which will be used as a source of inspiration throughout the remainder of this thesis is the analysis of [14], which revealed that an infinite series of instanton corrections to type IIB string theory can be completely summed up in terms of certain $SL(2,\mathbb{Z})$-invariant Eisenstein series. This example is discussed in Section 8.2, while a more careful treatment of Eisenstein series is given in Chapters 9 and 10.

8.1 Perturbative and Non-Perturbative Effects in String Theory

We begin here with a general discussion of some important aspects of string perturbation theory, emphasizing in particular the interplay between the worldsheet and spacetime viewpoints. We explain how this analysis naturally leads us to also consider non-perturbative effects, and we discuss the origin of such effects in terms of D-brane instantons.

8.1.1 General Structure of String Perturbation Theory

String theory is defined perturbatively in terms of an asymptotic expansion in two coupling constants: the worldsheet coupling $\alpha' \equiv \ell_s^2$, with ℓ_s being the fundamental string length, and

[1] I am very grateful to Axel Kleinschmidt and Christoffer Petersson for extensive discussions on the topics presented in this chapter, and for helpful remarks on an early version.

the string coupling $g_s \equiv e^{\langle\phi\rangle}$, where $e^{\langle\phi\rangle}$ denotes the expectation value of the dilaton. Any (closed string) scattering amplitude may then be written schematically as follows

$$\mathcal{A} = \sum_{n=0}^{\infty}\sum_{g=0}^{\infty} \left(\alpha'\right)^{n-4} g_s^{2(g-1)} \mathcal{A}_{(n,g)}, \tag{8.1.1}$$

where g denotes the genus of the string worldsheet and $\mathcal{A}_{(n,g)}$ is the genus g amplitude of order $n-4$ in α'. We emphasize that α' is dimensionful, a property of the worldsheet expansion which will play an important role in what follows.

The theory is under good control only in either of the two limits $\alpha' \to 0$ or $g_s \to 0$. In the limit of small string coupling, $g_s \to 0$, string theory is well described (on-shell) in terms of the two-dimensional worldsheet conformal field theory. On the other hand, when the worldsheet coupling is small, $\alpha' \to 0$, the strings themselves are well approximated by point particles, and the theory is effectively described by spacetime (super-)gravity. Even though morally true, this picture is somewhat oversimplified. In reality there is an interesting interplay between worldsheet effects, governed by α', and effects associated with the genus expansion in g_s.

To illustrate the discussion, let us consider in more detail the spacetime point of view of string perturbation theory. To leading order in α' and g_s, we have an effective description in terms of ten-dimensional Einstein-Hilbert gravity coupled to p-forms and fermionic matter fields, with an action of the schematic form

$$S_{(0,0)} = (\alpha')^{-4} \int d^{10}x \sqrt{G^{(s)}} e^{-2\phi}\Big[R - \star d\phi \wedge d\phi - \sum_p e^{\lambda_p \phi} \star F_{(p+1)} \wedge F_{(p+1)} + \cdots\Big], \tag{8.1.2}$$

where $F_{(p+1)}$ denotes the field strengths of whichever p-form potentials $A_{(p)}$ the theory contains, and λ_p represents the associated dilaton couplings.[2] The ellipsis indicate possible Chern-Simons couplings that are not important for the present discussion, and for the same reason we have also suppressed fermionic terms. The action (8.1.2) is written in *string frame* as is indicative of the dilaton prefactor in front of the bracket. By comparison to the schematic amplitude in Equation (8.1.1), we note that $e^{-2\phi}$ is the correct power of the string coupling corresponding to the genus zero, $g = 0$, part of the sum. This is the leading order effect and corresponds to a tree-level contribution in g_s. Further note the overall factor of $(\alpha')^{-4}$, which has the correct dimension to compensate for the dimension of the measure.[3] This is then what is referred to as tree-level in α', and hence the action (8.1.2) is a tree-level contribution with respect to the worldsheet coupling as well as the string coupling. This is indicated by the subscript on $S_{(0,0)}$ in (8.1.2).

In general, the spacetime effective action receives perturbative quantum corrections both in α' and in g_s. Since α' has dimension length squared, amplitudes which are higher order in α' naturally come with higher powers of momentum insertions. From the point of view of the effective action, these higher powers of momenta are associated with higher derivatives acting on the fundamental fields. Hence, the worldsheet α'-expansion is tantamount to a *derivative expansion* in spacetime.

Let us now consider an explicit example, namely type IIB string theory in $D = 10$. The tree-level effective action then corresponds to the action of type IIB supergravity. The next

[2] In the Ramond-Ramond sector all of the dilaton couplings are zero, $\lambda_p = 0$.

[3] This is related to Newton's constant in $D = 10$ as $16\pi G_N^{(10)} = (2\pi)^7(\alpha')^4 g_s^2$.

to leading order correction in α' arises from the $n = 1$ contribution in (8.1.1). This is of order $(\alpha')^{-3}$ and is therefore associated with a term in the effective action which is quartic in derivatives. One possible such term contains four derivatives of the metric $G^{(s)}_{\mu\nu}$, and therefore corresponds to a quadratic curvature correction in the effective action. However, it turns out that because of the large amount of supersymmetry ($\mathcal{N} = 2$) in the theory, the corresponding amplitude vanishes, $\mathcal{A}_{(1,g)} = 0$. In fact, the amplitude $\mathcal{A}_{(2,g)}$ also vanishes and the first non-trivial α'-correction arises from the $n = 3$ term in (8.1.1) and is of order $(\alpha')^{-1}$. This induces a term in the action with eight derivatives on the metric $G^{(s)}_{\mu\nu}$. This correction enters the effective action in the following way

$$S_{(3,0)} = (\alpha')^{-4} \int d^{10}x \sqrt{G^{(s)}} e^{-2\phi} \Big[R + (\alpha')^3 \mathcal{R}^4 + \cdots \Big], \tag{8.1.3}$$

where $\mathcal{R}^4$ denotes a specific combination of curvature scalars constructed out of the Ricci scalar R, the Ricci tensor $R_{\mu\nu}$ and the Riemann tensor $R_{\mu\nu\rho\sigma}$. The precise kinematical structure will not concern us here, and we refer the interested reader to [155–157] for more details. There will also be additional terms in (8.1.3) of the same order in α' corresponding to combinations of curvature terms and p-forms (e.g. R^2F^4), as well as pure p-form terms (e.g. F^8). These terms also play an important role in the analysis but will for simplicity of argument be neglected in the present discussion.

The effective action (8.1.3) now displays the first non-trivial α'-correction, but still only contains tree-level contributions in the string coupling g_s. In principle, there might be an additional infinite series of perturbative corrections in g_s at each order in α'. However, it turns out that the lowest order terms in α' (two-derivative terms) do not receive any corrections in g_s. We will see in Section 8.2 that this is also compatible with S-duality of type IIB string theory. On the other hand, the $\mathcal{R}^4$-term does receive g_s-corrections, and the action (8.1.3) must therefore be completed to include these additional contributions:

$$S_{(3,g)} = (\alpha')^{-4} \int d^{10}x \sqrt{G^{(s)}} \Big[e^{-2\phi} R + (\alpha')^3 \sum_{g=0}^{\infty} c_g \; e^{2(g-1)\phi} \mathcal{R}^4 + \cdots \Big], \tag{8.1.4}$$

where c_g are some coefficients which in general depend on the moduli of the theory. To calculate the precise values of these coefficients is a difficult problem, in particular since there might in principle be an infinite series of terms with non-vanishing coefficients c_g. In reality, however, it turns out that there are powerful duality arguments which constrain the coefficients such that the moduli-dependent coefficient in front of the $\mathcal{R}^4$-term may be determined exactly [14]. This will be discussed in detail in Section 8.2.

8.1.2 Including Non-Perturbative Effects

So far we have discussed only perturbative aspects of string theory. In addition, there exist *non-perturbative* effects which are crucial for the consistency of the theory. As argued long ago by Shenker [158], to understand this it is important to note that the genus-expansion in (8.1.1) does not converge, but rather should be interpreted as an *asymptotic series*.[4] This

[4]See, e.g., [159] for a discussion to asymptotic series in perturbation theory.

follows from the fact that for any fixed order k in α' the large genus, $g \to \infty$, behaviour of $\mathcal{A}_k$ is [158]

$$\lim_{g\to\infty} \mathcal{A}_k \equiv \lim_{g\to\infty} \sum_{g=0}^{\infty} e^{2(g-1)\langle\phi\rangle} \mathcal{A}_{(k,g)} \sim \sum_{g=0}^{\infty} e^{2(g-1)\langle\phi\rangle}\ a^{-2g}(2g)!\ , \tag{8.1.5}$$

where a is some constant. The factorial growth of $\mathcal{A}_{(k,g)} \sim (2g)!$ for large genus is due to the growth of the volume of the moduli space of Riemann surfaces $\mathcal{M}_g$ for large g [158,160,161]. In [158] it was suggested that in order to cure this divergence, one must include additional effects in (8.1.1) which are suppressed by a factor e^{-1/g_s} in the limit $g_s \to 0$. These would therefore correspond to non-perturbative effects from the point of view of string perturbation theory. This prediction of instanton effects in string theory was later verified with the discovery of D-branes [162], which are solitonic objects whose tension scales as $T \sim g_s^{-1}$, ensuring that they indeed contribute by exponentially suppressed terms of order e^{-1/g_s}, in accordance with the arguments of [158]. More specifically, consider a spacetime splitting of the form $\mathbb{R}^{1,9-r} \times X$, where X is some r-dimensional compact internal manifold. Then any Euclidean Dp-brane wrapping a $p+1$-dimensional submanifold $\mathcal{C} \subset X$ is completely localized in the external spacetime $\mathbb{R}^{1,10-r}$, and may therefore be interpreted as a *D-brane instanton.* In general, such instanton effects are suppressed by factors of $e^{-\mathrm{Vol}(\mathcal{C})/g_s}$, where $\mathrm{Vol}(\mathcal{C})$ denotes the volume of the submanifold $\mathcal{C}$. In addition, there exist non-perturbative effects from the worldsheet point of view. These are exponentially suppressed by factors $e^{-\mathrm{Vol}(\mathcal{B})/\alpha'}$ in the limit $\alpha' \to 0$, and arises from Euclidean fundamental strings wrapping holomorphic two-cycles $\mathcal{B}$ in X [163,164].

These results indicate that the action (8.1.4) is not yet the full story, but rather should be further completed to include possible non-perturbative effects at each order in α'. More precisely, we should add to the action an infinite series of such contributions, schematically indicated as follows

$$S^{\mathrm{np}}_{(3,\infty)} = (\alpha')^{-4} \int d^{10}x \sqrt{G^{(s)}} \left[g_s^{-2} R + (\alpha')^3 \left(\sum_{g=0}^{\infty} c_g\ g_s^{2(g-1)} + \sum_{N\neq 0} b_N e^{-|N|/g_s} \right) \mathcal{R}^4 + \cdots \right], \tag{8.1.6}$$

where N denotes the "instanton charge" and b_N is some moduli-dependent coefficient.[5] Since we are in $D = 10$ there is no volume dependence in the suppression factor $e^{-|N|/g_s}$. These kinds of effects indeed do appear in type IIB string theory, and may be attributed to D(-1)-instantons, which are 0-dimensional objects and hence localized in ten-dimensional spacetime. In the next section, we will see that the coefficients b_N are highly non-trivial and in fact contain an infinite series of additional perturbative excitations for each instanton number N.

Let us finally note that recently [165] the arguments of [158], regarding the large-order behaviour of string perturbation theory, were generalized to the large-order behaviour of certain infinite series of D-instanton corrections in type II string theory on a Calabi-Yau threefold X. From this analysis it was concluded that the D-instanton series should itself be treated asymptotically and, in order for the series to converge, additional non-perturbative effects of order e^{-1/g_s^2} must be included. Such effects are naturally attributed to Euclidean NS5-branes wrapping the entire internal manifold X and it was conjectured that in order to

[5] In this expression we have been deliberately sloppy and replaced the dynamical dilaton field e^{ϕ} directly with the string coupling g_s to illustrate the point. It is of course understood that g_s really corresponds to expectation value of the dilaton, as in (8.1.1).

ensure a convergent instanton series, the contributions from NS5-brane instantons must be included. We will explicitly see the effects of NS5-brane instantons arising in the analysis of Chapter 12, which is based on results from **Paper VIII**.

8.2 Summing Up Instantons With Automorphic Forms

The discussion in the previous section made it clear that there is an intricate interplay between worldsheet effects governed by α' and effects from the genus expansion governed by the string coupling g_s. While we have seen that perturbative α'-corrections are to a certain extent constrained by spacetime supersymmetry, no such constraint was put on the g_s-corrections. In this section we will see that there exist extra constraints, loosely referred to as "string dualities", which severly constrain the form of perturbative and non-perturbative corrections in the string coupling at each order in α'. These dualities are generally known as S-, T, or U-duality depending on the context. Typically, they are described by discrete groups $G(\mathbb{Z})$ which should leave the full quantum effective action invariant. Enforcing this invariance on the lowest order terms in g_s may then potentially fix completely the infinite series of perturbative and non-perturbative corrections in (8.1.6). In this section we will describe in detail the simplest example where this can be done, namely in type IIB string theory which exhibits invariance under the S-duality group $G(\mathbb{Z}) = SL(2,\mathbb{Z})$. Towards the end of the section we will also discuss extensions of these techniques to lower dimensions where larger duality groups appear.

8.2.1 The Exact $\mathcal{R}^4$-Correction in Type IIB String Theory

In Section 8.1.1 we discussed the existence of an α'-correction corresponding to a quartic curvature correction to Einstein gravity in $D = 10$. This occurs as the first non-trivial α'-correction to type IIB supergravity. We will now take a closer look at the possible perturbative and non-perturbative g_s-corrections to the $\mathcal{R}^4$-term.

Let us begin by reminding the reader about the bosonic field content of the theory. In addition to the metric, there is a dilaton modulus e^{ϕ}, an NS-NS 2-form $B_{(2)}$, and a set of Ramond-Ramond p-form fields $C_{(p)}$ with $p = 0, 2, 4, 6, 8$. The Ramond-Ramond forms couple to the D-branes of the theory, namely D(-1), D1, D3, D5, and D7-branes, respectively. In the NS-NS sector, $B_{(2)}$ couples to the fundamental string F1, while its dual $\tilde{B}_{(6)}$ couples to the NS5-brane. These are the basic objects in the theory, and upon compactification they might all give rise to various non-perturbative effects in the lower-dimensional theory.

For our present purposes, however, we shall focus on the D(-1)-branes. These exhibit the special property of having no extension in any of the ten spacetime directions, and therefore simply correspond to points. The D(-1)-branes thus have an interpretation as spacetime instantons in $D = 10$. This is in marked contrast to the other objects in the theory which have no such interpretation in ten dimensions. The D(-1)-instantons are sourced by the zero form $C_{(0)}$, which we shall refer to as the *axion* and denote by $C_{(0)} \equiv \chi$. The axion and the dilaton further combine into a complex scalar

$$\tau = \tau_1 + i\tau_2 := \chi + ie^{-\phi}, \tag{8.2.1}$$

which parametrizes the moduli space $\mathcal{M}_{10} = SL(2,\mathbb{R})/SO(2)$. The group $SL(2,\mathbb{R})$ is a global symmetry of the classical type IIB action. This invariance is however not manifest when the action is written in string frame as in (8.1.2). This is apparent by noting that the Einstein-Hilbert term is multiplied by an overall factor $e^{-2\phi}$ which transforms non-trivially under $SL(2,\mathbb{R})$. To exhibit the $SL(2,\mathbb{R})$-invariance, we perform a Weyl-rescaling of the string frame metric, $G^{(s)}_{\mu\nu} \equiv e^{\phi/2}G_{\mu\nu}$, which then reduces the action to the standard Einstein-Hilbert form

$$S_{(0,0)} = (\alpha')^{-4}\int d^{10}x\sqrt{G}\Big[R+\cdots\Big]. \tag{8.2.2}$$

This is referred to as *Einstein frame* and the rescaled metric $G_{\mu\nu}$ does not transform under $SL(2,\mathbb{R})$, ensuring that the Einstein-Hilbert term is now manifestly invariant under $SL(2,\mathbb{R})$. The other terms in the action also organize themselves in such a way that the full action in Einstein frame is $SL(2,\mathbb{R})$-invariant. It is generally expected in string theory that such classical continuous global symmetries $G(\mathbb{R})$ should be broken by quantum effects. However, one expects that there will in fact be some remnant of this classical symmetry preserved in the full quantum theory in the form of a discrete subgroup $G(\mathbb{Z}) \subset G(\mathbb{R})$ [8].

Let us then investigate this question within the present framework of type IIB supergravity. To this end we turn on the quartic correction term $\mathcal{R}^4$ in the effective action, as discussed in Section 8.1.1. We have seen in (8.1.4) that in string frame this term enters with the same power $e^{-2\phi}$ of the dilaton as the Einstein-Hilbert term, signifying a tree-level effect. To make contact with the discussion of $SL(2,\mathbb{R})$-invariance above, we convert also the $\mathcal{R}^4$-term to Einstein frame, with the result

$$S_{(3,0)} = (\alpha')^{-4}\int d^{10}x\sqrt{G}\Big[R+(\alpha')^3 c_0\tau_2^{3/2}\mathcal{R}^4+\cdots\Big], \tag{8.2.3}$$

where we recall that $\tau_2 = e^{-\phi}$, and we have also reinstalled the genus zero coefficient from (8.1.4). By explicit string theory calculations [166–168], it has been verified that c_0 is non-vanishing and takes the precise value $c_0 = 2\zeta(3)$, where $\zeta(3)$ denotes the value of the Riemann zeta function $\zeta(z)$ for $z = 3$.

In addition, in [157] the next to leading order term, corresponding to genus $g = 1$ in (8.1.1), was calculated and it was found that also this coefficient is non-vanishing, with $c_1 = 4\zeta(2)$. Hence, the $\mathcal{R}^4$-term receives at least a 1-loop correction, and up to genus one we then have

$$S_{(3,1)} = (\alpha')^{-4}\int d^{10}x\sqrt{G}\Big[R+(\alpha')^3\Big(2\zeta(3)\tau_2^{3/2}+4\zeta(2)\tau_2^{-1/2}\Big)\mathcal{R}^4+\cdots\Big]. \tag{8.2.4}$$

From this expression it is clear that the perturbative contributions to the $\mathcal{R}^4$-correction explicitly break the $SL(2,\mathbb{R})$-invariance of the effective action. This follows since in Einstein frame the $\mathcal{R}^4$-term is itself invariant, while the dilaton-dependent coefficients transform non-trivially. From the generic form of (8.1.4) we may in fact deduce that each term in the entire infinite series of possible perturbative corrections will separately break the invariance under $SL(2,\mathbb{R})$. However, there is by now a large body of evidence [1,8] that type IIB string theory is self-dual under the S-duality group $SL(2,\mathbb{Z}) \subset SL(2,\mathbb{R})$, i.e. the modular group of integer-valued 2×2 matrices with unit determinant. It is then expected that after all perturbative and non-perturbative quantum corrections have been taken into account, the coefficient in front of $\mathcal{R}^4$ should in fact be invariant under $SL(2,\mathbb{Z})$.

This philosophy was first utilized in the seminal work of Green and Gutperle [14], where they conjectured that the action (8.2.4) should be completed to an exact expression of the form:

$$S_{(3,\text{exact})} = (\alpha')^{-4} \int d^{10}x\sqrt{G}\Big[R + (\alpha')^3 f(\tau)\mathcal{R}^4 + \cdots\Big], \tag{8.2.5}$$

where $f(\tau)$ is some $SL(2,\mathbb{Z})$-invariant function of the modulus $\tau \in SL(2,\mathbb{R})/SO(2)$, i.e. an *automorphic form*.[6] This function is further constrained so that in a weak-coupling expansion, $\tau_2 = g_s^{-1} \to \infty$, it should reproduce the tree-level and one-loop terms in (8.2.4),

$$f(\tau) \sim 2\zeta(3)\tau_2^{3/2} + 4\zeta(2)\tau_2^{-1/2} + \cdots, \tag{8.2.6}$$

where the ellipsis indicate possible additional perturbative and non-perturbative contributions to the $\mathcal{R}^4$-term. It turns out that there is a unique[7] candidate for $f(\tau)$ given by the non-holomorphic Eisenstein series [14]

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau; s) = \sum_{(m,n)\neq(0,0)} \frac{\tau_2^s}{|m+n\tau|^{2s}}, \tag{8.2.7}$$

where the parameter s is known as the *order* of the Eisenstein series. For the special value $s = 3/2$ this Eisenstein series has a Fourier expansion of the form

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau; 3/2) = 2\zeta(3)\tau_2^{3/2} + 4\zeta(2)\tau_2^{-1/2} + 4\pi\sqrt{\tau_2}\sum_{N\neq 0}\mu_{3/2}(N)NK_1\big(2\pi|N|\tau_2\big)e^{-2\pi i N\tau_1}, \tag{8.2.8}$$

where $K_1(x)$ is the modified Bessel function, and $\mu_{3/2}(N)$ is a combinatorial coefficient which plays an important role, and will be discussed in more detail at the end of this section. Remarkably, the expansion (8.2.8) reproduces correctly the two perturbative terms in (8.2.4), including the exact genus zero and genus one coefficients $c_0 = 2\zeta(3)$ and $c_1 = 4\zeta(2)$.

The next important thing to note about the expansion (8.2.8) is the absence of perturbative contributions beyond one-loop, implying that $SL(2,\mathbb{Z})$-invariance enforces $c_g = 0$ for all coefficients in (8.1.4) with $g \geq 2$. This is a strong prediction of S-duality in type IIB string theory, and lies at the heart of powerful non-renormalisation theorems beyond one-loop for the $\mathcal{R}^4$-term also in dimensions $D \leq 10$ [14, 172–179].

It remains to discuss the last part of (8.2.8), apparently encoding an infinite series of additional non-perturbative corrections. The non-perturbative effects are best revealed by

[6] We shall generally refer to functions invariant under some discrete group $G(\mathbb{Z})$ as *automorphic forms*, while in the special case of $G(\mathbb{Z}) = SL(2,\mathbb{Z})$ these are sometimes called *modular forms*. Let us also note that in the literature on holomorphic modular forms for $SL(2,\mathbb{Z})$, it is common to distinguish between modular *forms* and *functions*, the latter being completely $SL(2,\mathbb{Z})$-invariant while the former transform with some overall (modular) weight. In this thesis we shall not make this distinction.

[7] A priori, it is possible that so called *cusp forms* (roughly, automorphic forms for which the perturbative part vanishes) could also contribute, as they would not spoil the perfect agreement with the tree-level and one-loop terms. However, this possibility has been ruled out by constraints from supersymmetry [15, 169–171].

noting that for large argument the modified Bessel function $K_1(x)$ in (8.2.8) has an expansion of the form

$$K_1(x) = \sqrt{\frac{\pi}{2x}} e^{-x} \Big[1 + \mathcal{O}(1/x) + \cdots \Big]. \tag{8.2.9}$$

This expansion is justified in the weak-coupling limit $\tau_2 \to \infty$ and yields

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau; 3/2) = 2\zeta(3)\tau_2^{3/2} + 4\zeta(2)\tau_2^{-1/2} + 2\pi \sum_{N\neq 0} \mu_{3/2}(N)\sqrt{N} e^{-S_{\text{inst}}(\tau)} \Big[1 + \mathcal{O}(\tau_2) + \cdots \Big]. \tag{8.2.10}$$

This shows that for each N we have infinite series of perturbative excitations which are exponentially suppressed by a factor $e^{-S_{\text{inst}}(\tau)}$, where we have defined

$$S_{\text{inst}}(\tau) := 2\pi |N| \tau_2 + 2\pi i N \tau_1 = 2\pi |N| e^{-\phi} + 2\pi i N \chi. \tag{8.2.11}$$

This is precisely the Euclidean action for D(-1)-instantons [14, 180]. The imaginary part of $S_{\text{inst}}(\tau)$ corresponds to the analogue of the "theta-angle" in Yang-Mills theory, and encodes the coupling to the Ramond-Ramond axion χ. The infinite series of corrections in (8.2.10), induced by the expansion of the Bessel function, may then be attributed to perturbative excitations around the instanton background [14]. The existence of these perturbative corrections for each N is another prediction of type IIB $SL(2,\mathbb{Z})$-duality, albeit one which is very difficult to verify since explicit methods for instanton calculus are not well developed in string theory. We note that the expansion (8.2.10) is indeed of the expected form indicated in (8.1.6).

The success of this analysis led to the conjecture that the exact $\mathcal{R}^4$-term be given by [14]

$$S_{(3,\text{exact})} = (\alpha')^{-4} \int d^{10}x \sqrt{G} \Big[R + (\alpha')^3 \mathcal{E}^{SL(2,\mathbb{Z})}(\tau; 3/2) \mathcal{R}^4 + \cdots \Big]. \tag{8.2.12}$$

This result may be seen as a direct realization of the arguments of Shenker [158], discussed in Section 8.1.2, in the sense that the perturbative series in (8.1.4) was "completed" by non-perturbative effects into the form (8.2.12). In fact, the Eisenstein series $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau; s)$ is absolutely convergent for $\Re(s) > 1$, and hence the series encoded in the coefficient in (8.2.12) indeed converges.

We conclude this section with a discussion of the important coefficient $\mu_{3/2}(N)$, appearing in the expansion (8.2.8). This is a number theoretic quantity, known in the physics literature as the *instanton measure*, which is given explicitly by a sum over divisors of the instanton number N [14, 181–185]

$$\mu_{3/2}(N) := \sum_{n|N} n^{-2}. \tag{8.2.13}$$

Roughly speaking, this counts the degeneracies of charge N D(-1)-instantons. More accurately, the non-perturbative effects in (8.2.8) are not to be interpreted as multi-instanton contributions, but rather as contributions from a single instanton with arbitrary charge N. This is due to the fact that processes involving multiple instantons only contribute to terms which are higher order in α', because of the saturation of zero modes required for a non-vanishing contribution [14]. Hence, the proper statement is that $\mu_{3/2}(N)$ should correspond to the measure on the moduli space of a single D(-1)-instanton of charge N [183].

It is illuminating to consider the further reduction of type IIB string theory on a circle $S^1(R)$ to $D = 9$. By T-duality, the D(-1)-instantons are then mapped to Euclidean D0-particles wrapping the circle in type IIA string theory on $S^1(R^{-1})$. More precisely, each D(-1)-instanton with (real) action $\Re(S_{\text{inst}}) = 2\pi|N|e^{-\phi}$ is mapped to a D0-particle with mass $2\pi R|N|e^{-\phi} = 2\pi R|\tilde{m}n|e^{-\phi}$, which has n units of momentum along the circle and winding number $\tilde{m}$. Under this duality, there is therefore a simple way to understand the sum over divisors of N in (8.2.13): it counts the number of partitions of N into two integers $\tilde{m}$ and n, or, in other words, the number of ways in which a single charge N D(-1)-instanton may be constructed from a bound state of D0-particles with momentum n and winding $\tilde{m}$ along the Euclidean circle S^1 [14].

8.2.2 U-Duality and the Exact $\mathcal{R}^4$-Correction in $D < 10$

The analysis so far has been restricted to type IIA and type IIB string theory in $D = 10$, with the exception of a short detour to $D = 9$. We have seen that type IIB already in maximal dimension exhibits non-perturbative effects due to D(-1)-instantons, while in type IIA similar effects only appear in nine dimensions after reduction on S^1. In fact, as briefly mentioned in Section 8.1.2, this is a general feature of non-perturbative effects in string theory; as we compactify the theory to lower dimensions new instanton effects appear. These arise because when spacetime splits into an external part $\mathbb{R}^{1,10-r}$ and an r-dimensional compact internal part X, then Euclidean D-branes, or NS5-branes, may wrap submanifolds of X, hence becoming completely localized with respect to the external spacetime $\mathbb{R}^{1,10-r}$. This also implies that in lower dimensions it becomes increasingly difficult to sum up all instanton corrections to the effective action.

One generic problem that arises when studying the theory on $\mathbb{R}^{1,10-r} \times X$ is that the lower-dimensional theory exhibits a complicated moduli space $\mathcal{M}$ parametrized by the scalar fields of the theory. The geometry of this moduli space is in turn related to the topology of the compact manifold X. This often means that the same techniques used in the previous sections for D(-1)-instantons may no longer work since it is not known in general what the analogue of the S-duality group $SL(2,\mathbb{Z})$ is in lower dimensions. In other words, the power of automorphy is lost.

The situation improves considerably if we restrict to compact manifolds which preserve a large amount of supersymmetry. In particular, the simplest case is to take $X = T^r$ which preserves all of the 32 supercharges of type II string theory. For such toroidal compactifications it turns out that the classical moduli space $\mathcal{M}$ is always a symmetric space of the form G/K, where G is a global symmetry of the classical action and K is the "R-symmetry". For example, consider type IIB string theory reduced on T^2. Under this reduction, the classical moduli space is enhanced from $\mathcal{M}_{10} = SL(2,\mathbb{R})/SO(2)$ to

$$\mathcal{M}_8 = \frac{SL(2,\mathbb{R}) \times SL(3,\mathbb{R})}{SO(2) \times SO(3)} \tag{8.2.14}$$

in $D = 8$. It is generally expected that the global symmetry group will again be broken by quantum effects to a discrete subgroup

$$G(\mathbb{Z}) = SL(2,\mathbb{Z}) \times SL(3,\mathbb{Z}), \tag{8.2.15}$$

known as the *U-duality group* [8] (see [11] for a review). Hence, in this case we expect that the power of automorphy is retained, since the exact effective action in $D = 8$ should exhibit manifest invariance under $G(\mathbb{Z})$. This was indeed shown to be true in [12] (see also preceeding work [15]), where a proposal for the exact coefficient of the $\mathcal{R}^4$-term was given in terms of an $SL(3,\mathbb{Z})$-invariant Eisenstein series of the form

$$\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s) = \sum_{m\in\mathbb{Z}^3\backslash\{(0,0,0)\}} \Big[\vec{m}^T\cdot\mathcal{K}\cdot\vec{m}\Big]^{-s}, \tag{8.2.16}$$

where the relevant order is again $s = 3/2$. The matrix $\mathcal{K}$ parametrizes the coset space $SL(3,\mathbb{R})/SO(3)$. This Eisenstein series is a direct generalization of (10.2.1), and some of its properties will be discussed in detail in Sections 9.3 and 10.3.

In accordance with our general discussion above, new instanton effects appear in $D = 8$ from Euclidean fundamental strings and D1-branes wrapping the torus T^2. It was shown in [12] that these effects are indeed correctly reproduced in the Fourier expansion of $\mathcal{E}^{SL(3,\mathbb{Z})}$, which also encodes the D(-1)-instanton sum in (10.2.1). Another interesting feature of this example is that the Eisenstein series (8.2.16) is only absolutely convergent for $\Re(s) > 3/2$ and in fact exhibits a pole for $s = 3/2$. However, after regularization, this pole is replaced by a logarithmic divergence which precisely reproduces the physical infrared logarithmic divergence of the one-loop correction to the $\mathcal{R}^4$-term in $D = 8$ [12].[8]

As also discussed in Chapter 1, when taking a larger and larger internal torus, the duality groups $G(\mathbb{R})$ also become larger, culminating in $D = 4$ with the group $G(\mathbb{R}) = \mathcal{E}_7(\mathbb{R})$ [186], and with classical moduli space given by

$$\mathcal{M}_4 = \frac{\mathcal{E}_7(\mathbb{R})}{SU(8)/\mathbb{Z}_2}. \tag{8.2.17}$$

In addition to the metric and the scalar fields parametrizing the coset space $\mathcal{M}_4$, the bosonic sector in $D = 4$ also contains 28 abelian vector fields, whose electric and magnetic charges together span a symplectic lattice invariant under $Sp(56;\mathbb{Z})$. Based on these arguments, it was conjectured in [8] that the U-duality group $G(\mathbb{Z})$ in $D = 4$ should be

$$\mathcal{E}_7(\mathbb{Z}) := \mathcal{E}_7(\mathbb{R}) \cap Sp(56;\mathbb{Z}). \tag{8.2.18}$$

In $D = 4$ yet more non-perturbative effects are expected to arise from Euclidean Dp-branes ($p = -1, 1, 3, 5$) as well as NS5-branes wrapping even cycles in T^6. Similarly, from the type IIA perspective these effects arise from Euclidean Dp-branes ($p = 0, 2, 4$) wrapping odd cycles in T^6, as well as NS5-branes wrapping the entire six-torus. It was speculated in [13] that such effects could possibly be summed up with a certain $\mathcal{E}_7(\mathbb{Z})$-invariant Eisenstein series. However, it is not clear at present whether or not it is possible to make this proposal explicit; most notably to extract the Fourier coefficients of such an Eisenstein series appears to be a very hard challenge which has not yet been achieved.

Another problem is related to the perturbative contributions in the Fourier expansion. As mentioned above, there is no reason to believe that the one-loop non-renormalization theorems that are valid for the $\mathcal{R}^4$-correction in $D = 10$ should be violated in lower dimensions. In other

[8] I am grateful to Pierre Vanhove for emphasizing this point to me.

words, it is expected that the $\mathcal{R}^4$-term in $D = 4$ should also be perturbatively finite beyond one-loop. However, in general the number of perturbative terms in the Fourier expansion of a $G(\mathbb{Z})$-invariant automorphic form is equal to the order of the Weyl group $\mathcal{W}(\mathfrak{g})$ of the Lie algebra $\mathfrak{g} = \text{Lie } G$ [187].[9] For example, in the case of $SL(2,\mathbb{R})$ the Weyl group is of order 2, $\mathcal{W}(\mathfrak{sl}(2,\mathbb{R})) = \mathbb{Z}_2$, and this is the reason for the existence of only two perturbative terms in the expansion (8.2.8). In the case of $\mathcal{E}_7(\mathbb{Z})$-invariant Eisenstein series this presents a problem, since the order of the Weyl group of $\mathcal{E}_7$ is very large. A possible way out is that the number of constant terms decreases for "residues" of the general Eisenstein series, i.e. corresponding to a specific choice of representation of the duality group G. It is therefore plausible that the correct coefficient in front of the $\mathcal{R}^4$-term in $D = 4$ should be given by some high-order residue of the general $\mathcal{E}_7(\mathbb{Z})$-invariant Eisenstein series proposed in [13].[10] In Chapter 10, we will discuss some examples of this phenomenon in the simpler context of $SL(3,\mathbb{Z})$, revealing that the Eisenstein series (8.2.16) in fact arises from a more general Eisenstein series $\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K}; s_1, s_2)$ in the limit $s_2 \to 0$.

To conclude the discussion, let us provide some words of motivation for subsequent chapters. We have learned that (U-)duality $G(\mathbb{Z})$ in any dimension $D \leq 10$ provides powerful constraints on the structure of perturbative and non-perturbative effects in string theory. In particular, $G(\mathbb{Z})$-invariant automorphic forms may potentially be used to sum up all quantum corrections to the effective action of string theory on $\mathbb{R}^{1,10-r} \times X$, at least in cases where X preserves a sufficient amount of supersymmetry. In addition to the toroidal case discussed in this section, another example is provided by type IIA string theory on $X = K3 \times T^2$ which is known to be dual to heterotic string theory on T^6 [188, 189]. In this case the U-duality group is given by $SL(2,\mathbb{Z}) \times SO(6,22;\mathbb{Z})$ and automorphic techniques have been extensively used to constrain the effective action [190–192]. All this motivates us to study the general construction of automorphic forms on coset space G/K, and in particular to understand how to compute their Fourier expansions. This is the purpose of the remaining chapters of **Part II** of this thesis. We will also see in Chapter 12 that in certain cases automorphic techniques may be used also for compact manifolds X preserving a smaller amount of supersymmetry, e.g. for Calabi-Yau threefolds.

[9] These issues will be explained in more detail in Chapter 10.

[10] I am grateful to Boris Pioline for emphasizing this to me.

9

Constructing Automorphic Forms

In the previous chapter, we learned that discrete subgroups $G(\mathbb{Z})$ of certain Lie groups G play an important role as underlying symmetry groups in string theory. We have seen how invariance under $G(\mathbb{Z})$ enforces strong constraints on perturbative and non-perturbative quantum corrections. These contributions to the low-energy effective action are encoded in $G(\mathbb{Z})$-invariant functions of the moduli fields, i.e *automorphic forms.* We briefly touched upon one of the simplest examples, namely the well-known non-holomorphic Eisenstein series $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau; s)$, which encodes D$(-1)$-instanton corrections in type IIB string theory. In this section we shall formalize these techniques and give a more general treatment of automorphic forms on arithmetic quotients $G(\mathbb{Z})\backslash G/K$. We will also consider in detail the Fourier expansion of these automorphic forms, which is the framework in which physical effects are revealed. In this context, it is important to distinguish between two different cases: *abelian* and *non-abelian* Fourier expansions. This refers to whether the nilpotent subgroup $N \subset G$ is abelian or non-abelian. Representative examples are given by Eisenstein series on $SL(2,\mathbb{R})$ and $SL(3,\mathbb{R})$, which are treated in detail in Sections 10.2 and 10.3, respectively.

Before we begin, let us give a brief guide to some of the relevant literature on automorphic forms. A canonical mathematical reference for Eisenstein series are the lecture notes by Langlands [187]. Perhaps more accessible are the books by Terras [80, 81], or in the special case of $G = SL(2,\mathbb{R})$ a highly recommended reference is the book by Borel [193], while for $G = SL(3,\mathbb{R})$ we also recommend the thesis by Miller [194]. For nice physicist's accounts, upon which the present treatment is mainly based, we recommend the paper [13] by Obers and Pioline, and the review [195] by Pioline and Waldron.

9.1 Constructing Automorphic Forms

We shall here describe several different methods for constructing automorphic forms on coset spaces G/K. We first give a general treatment of these methods, comparing and contrasting the mathematician's viewpoint with more physics-inspired constructions. For the applications considered in this thesis we will mainly be interested in a special class of automorphic forms known as *Eisenstein series.* The general construction of Eisenstein series is therefore treated in detail in Sections 9.1.1 and 9.1.2. Then in Section 9.1.3 we discuss a more general framework which in principle allows for constructing any automorphic form on a coset space G/K, of which interesting examples include Eisenstein series and theta series.

9.1.1 Eisenstein Series - Mathematicians' Approach

Let $G(\mathbb{R})$ be a Lie group in its split real form[1], and let $\mathfrak{g}$ be the associated Lie algebra of rank r. We know from Section 2.2.3 that G exhibits an Iwasawa decomposition of the form

$$G = NAK, \tag{9.1.1}$$

where N denotes the nilpotent subgroup, A is the Cartan torus and K is the maximal compact subgroup. Let furthermore h_i $(i = 1, \ldots, r)$ be the generators of the Cartan subalgebra $\mathfrak{h} = \text{Lie } A$ and choose a parametrization of the Cartan torus of the following form

$$a(\phi) = \exp\Big[-\sum_{i=1}^{r} \phi_i h_i\Big] := \prod_{i=1}^{r} \Phi_i^{h_i} \in A, \tag{9.1.2}$$

where we defined the variables

$$\Phi_i := \exp\big[-\phi_i\big]. \tag{9.1.3}$$

Now note that by virtue of the Iwasawa decomposition (9.1.1) the left action action of any element $x \in N$ on $g = na(\phi)k \in G$ leaves invariant the variables Φ_i parametrizing the Cartan torus. A function $E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}})$ which is invariant under a discrete subgroup $G(\mathbb{Z}) \subset G(\mathbb{R})$ may now be constructed as follows

$$E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}}) := \sum_{\gamma \in N(\mathbb{Z})\backslash G(\mathbb{Z})} \prod_{i=1}^{r} \Big(\gamma \cdot \Phi_i\Big)^{s_i}, \tag{9.1.4}$$

where $\lambda_{\mathcal{R}} = (s_1, \ldots, s_r)$ is a highest weight vector[2] associated with the representation $\mathcal{R}$ of G. Since the product of all Φ_i is invariant under N, we must in order to ensure convergence sum only over the orbit $N(\mathbb{Z})\backslash G(\mathbb{Z}) := \big(G(\mathbb{Z}) \cap N\big)\backslash G(\mathbb{Z})$. The function $E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}})$ is known as an *Eisenstein series*, and a choice of the parameters s_i is called the *order* of the Eisenstein series. The general method of constructing $G(\mathbb{Z})$-invariant functions by summing over images of a coset $\Gamma\backslash G(\mathbb{Z})$, for Γ some discrete subgroup of G, is known as *Poincaré series*, and may also be used to construct other types of automorphic forms.

An important remark is that the Eisenstein series $E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}})$ is an eigenfunction of all G-invariant differential operators on G/K. In particular, it is an eigenfunction of the Laplacian on G/K, with eigenvalue proportional to the value of the quadratic Casimir in the representation $\mathcal{R}$:

$$\Delta_{G/K} E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}}) = c\big(\lambda_{\mathcal{R}}, \lambda_{\mathcal{R}} + 2\rho\big) E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}}), \tag{9.1.5}$$

[1] As we shall see in Chapter 12 the construction can be generalized to the case of non-split real forms, but for simplicity we here restrict to the split case.

[2] In general, it is not necessary for $\lambda_{\mathcal{R}}$ to be a highest weight vector, but only some arbitrary vector in weight space. However, for our purposes, it is sufficient to restrict to the case when $\lambda_{\mathcal{R}}$ is a highest weight, in which case the present construction relates nicely to the one in the following section. I thank Boris Pioline for pointing this out to me.

where ρ is the Weyl vector (see Eq. (2.2.37)), and the proportionality constant $c \neq 0$ depends on the conventions. We shall see explicit examples of Eq. (9.1.5) in Sections 9.2 and 9.3, as well as in Chapter 12. The Eisenstein series $E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}})$ has poles for those values of $\lambda_{\mathcal{R}}$ where the quadratic Casimir vanishes [187].[3]

9.1.2 Eisenstein Series - Physicists' Approach

In this section we shall describe a more "physics-oriented" method for constructing non-holomorphic Eisenstein series on arithmetic quotients,

$$G(\mathbb{Z})\backslash G(\mathbb{R})/K(G), \tag{9.1.6}$$

where $G(\mathbb{Z})$ denotes a discrete subgroup of the continuous Lie group $G(\mathbb{R})$. In other words, we are interested in $G(\mathbb{Z})$-invariant functions whose argument takes values in the coset G/K. The method we will discuss was developed in [13] by Obers and Pioline.

We will be able to use some of the technology discussed in **Part I** of this thesis. First we choose a coset representative $\mathcal{V} \in G/K$ in the Borel gauge (see Section 5.1.5). Note that this is a representative for the right coset G/K, rather than the left coset $K\backslash G$ which was used extensively in **Part I**. This choice is more convenient for our present purposes. The difference lies in the choice of Iwasawa decomposition, which here reads $G = NAK$, as in (9.1.1), in contrast to $G = KAN$ used in **Part I**. The main building block of the Eisenstein series will then be the "generalized metric" $\mathcal{K}$, constructed from $\mathcal{V}$ as follows:

$$\mathcal{K} := \mathcal{V}\mathcal{V}^{\mathcal{T}}, \tag{9.1.7}$$

where $(\ )^{\mathcal{T}}$ denotes the "generalized transpose" defined in Section 5.1.3, which for the special case of $G = SL(n, \mathbb{R})$ reduces to the ordinary matrix transpose. The coset representative $\mathcal{V}$ transforms by $k^{-1} \in K$ from the right and by $g \in G$ from the left,

$$\mathcal{V} \quad \longmapsto \quad g\mathcal{V}k^{-1}, \tag{9.1.8}$$

which in turn implies that $\mathcal{K}$ is manifestly invariant under K, while transforms covariantly under G:

$$\mathcal{K} \quad \longmapsto \quad g\mathcal{K}g^{\mathcal{T}}. \tag{9.1.9}$$

Any function of $\mathcal{K}$ will then by construction live on the coset G/K. To proceed, we must specify what is meant by the discrete subgroup $G(\mathbb{Z}) \subset G(\mathbb{R})$. For the purposes of this thesis the group $G(\mathbb{Z})$ simply corresponds to the subset of $G(\mathbb{R})$ that leaves invariant some discrete lattice $\Lambda_{\mathbb{Z}}$. In physical applications, this lattice can for example correspond to the lattice of electric and magnetic charges (see, eg., [8]).

The desired Eisenstein series may then be defined as follows

$$\mathcal{E}^{G(\mathbb{Z})}(\mathcal{K}; s) := \sum_{m \in \Lambda_{\mathbb{Z}}}{}' \delta\big(\vec{m} \wedge m\big) \Big[m^{\mathcal{T}} \cdot \mathcal{K} \cdot m\Big]^{-s}, \tag{9.1.10}$$

where the prime on the sum indicates that the term with $m = (m_1, \ldots, m_n) = (0, \ldots, 0)$ should be omitted. This is indeed a function on G/K because it is constructed from $\mathcal{K}$, and

[3] I am grateful to Pierre Vanhove for stressing this point to me.

it is also invariant under $G(\mathbb{Z})$ since the action of $\gamma \in G(\mathbb{Z})$ on $\mathcal{K}$ will hit the lattice vectors $m \in \Lambda_{\mathbb{Z}}$ and simply results in a change in summation variables, i.e.

$$\mathcal{E}^{G(\mathbb{Z})}(\gamma\cdot\mathcal{K};s) = \sum_{m\in\Lambda_{\mathbb{Z}}} \delta\big(m\wedge m\big)\Big[m^{\mathcal{T}}\cdot\gamma\cdot\mathcal{K}\cdot\gamma^{\mathcal{T}}\cdot m\Big]^{-s} = \sum_{m'\in\Lambda_{\mathbb{Z}}} \delta\big(m'\wedge m'\big)\Big[m'^{\mathcal{T}}\cdot\mathcal{K}\cdot m'\Big]^{-s}, \quad (9.1.11)$$

where $m' := \gamma^{\mathcal{T}} \cdot m \in \Lambda_{\mathbb{Z}}$. We have also inserted a delta function $\delta(m \wedge m)$ which enforces a certain quadratic constraint

$$m \wedge m = 0 \quad (9.1.12)$$

on the summation variables. This is required in general to make sure that $\mathcal{E}^{G(\mathbb{Z})}(\mathcal{K};s)$ is an eigenfunction of the Laplacian on G/K, as is required for non-holomorphic Eisenstein series. For the case of $G = SL(n,\mathbb{R})$ this constraint is not needed, but it will appear explicitly in Chapter 12 when we consider the example of $G = SU(2,1)$. For a more detailed discussion of the constraint see [13]. We note that for an n-dimensional integral lattice $\Lambda_{\mathbb{Z}}$, the Eisenstein series (9.1.10) converges absolutely for $\Re(s) > n/2$ [81].

Let us briefly conclude with a comparison to the construction of the previous section. The Eisenstein series $\mathcal{E}^{G(\mathbb{Z})}(\mathcal{K};s)$ may be considered as a special case of the Poincaré series $E^{G(\mathbb{Z})}(g;\lambda_{\mathcal{R}})$ in (9.1.4). More precisely, $\mathcal{E}^{G(\mathbb{Z})}(\mathcal{K};s)$ arises as a certain limiting choice of the parameters s_i. This will become more clear in Sections 9.2 and 9.3 where we consider explicit examples.

9.1.3 Constructing Automorphic Forms Using Spherical Vectors

There exists a general method for constructing automorphic forms on $G(\mathbb{Z})\backslash G/K$ as advocated in [195–197]. The fundamental object in this construction is the so called *spherical vector* f_K which belongs to a Hilbert space of square-integrable functions. The spherical vector is defined by its invariance under $K(G)$, acting on $\mathcal{H}$ through a linear representation ρ of $G(\mathbb{R})$. To understand this construction we begin by recalling a few facts about the theory of *induced representations*.

A Note on Parabolic Induction

We want to describe representations ρ of a Lie group G acting on functions belonging to a Hilbert space $\mathcal{H}$, which in our case will be identified with the space $L^2(P\backslash G)$ of square integrable, real-valued functions on the coset space $P\backslash G$, where P is the Borel subgroup of G. We will see how P is defined in explicit examples later on. Elements of $L^2(P\backslash G)$ are then functions of the form

$$f \,:\, n \longmapsto f(n) \in \mathbb{R}, \qquad n \in P\backslash G. \quad (9.1.13)$$

The group G acts on $n \in P\backslash G$ from the right and a compensating transformation of $p \in P$ from the left is required to obtain a new representative n' for the coset space $P\backslash G$,

$$n \longmapsto png = n' \in P\backslash G. \quad (9.1.14)$$

We will then consider representations ρ of G acting on functions $f \in L^2(P\backslash G)$ with the property that

$$\rho(g)\cdot f(n) = f(ng) = f(pn') = \chi(p)f(n'), \quad (9.1.15)$$

where

$$\chi \,:\, p \longmapsto \chi(p) \in \mathbb{R} \tag{9.1.16}$$

is the so called *infinitesimal character* defining the representation ρ. Representations of G obtained in this way by induction from the Heisenberg parabolic P belong to the *principal continuous series* of G-representations.

Automorphic Forms and Spherical Vectors

We shall now apply these techniques to construct an automorphic form Ψ on G/K, invariant under $G(\mathbb{Z})$. As mentioned above, to this end we need to find a *spherical vector*, i.e. a K-invariant function $f_K \in \mathcal{H} = L^2(P\backslash G)$ belonging to the representation ρ discussed above. In addition we need two more ingredients, namely a $G(\mathbb{Z})$-invariant *distribution* $f_{\mathbb{Z}}$ which belongs to the dual space $\mathcal{H}^\star$, as well as an inner product $\langle\ ,\ \rangle$, provided by the canonical pairing between $\mathcal{H}$ and $\mathcal{H}^\star$. Putting things together, we may now quite generally define Ψ as [195–197]

$$\Psi(g) := \langle f_{\mathbb{Z}}, \rho(g) \cdot f_K \rangle\,, \tag{9.1.17}$$

The calculation of Ψ can be simplified by invoking the Iwasawa decomposition of G,

$$G = NAK, \tag{9.1.18}$$

with N representing the nilpotent part, A the abelian part, and K the maximal compact subgroup. We thus have the corresponding decomposition of $g \in G$,

$$g = nak, \qquad n \in N,\ a \in A,\ k \in K. \tag{9.1.19}$$

Now we may use the property that f_K is invariant under K,

$$\rho(k) \cdot f_K = f_K, \tag{9.1.20}$$

implying that

$$\rho(g) \cdot f_K = \rho(nak) \cdot f_K = \rho(na) \cdot f_K. \tag{9.1.21}$$

Defining the standard coset representative in the Borel gauge as follows

$$\mathcal{V} = na \ \in\ G/K, \tag{9.1.22}$$

we may write the automorphic form as

$$\Psi(\mathcal{V}) = \langle f_{\mathbb{Z}}, \rho(\mathcal{V}) \cdot f_K \rangle\,. \tag{9.1.23}$$

This minor simplification will turn out to be very useful in the construction of Ψ. The coset representative $\mathcal{V} \in G/K$ transforms by $k^{-1} \in K$ on the right and $\gamma \in G(\mathbb{Z})$ on the left,

$$\mathcal{V} \longmapsto \gamma \mathcal{V} k^{-1}. \tag{9.1.24}$$

On Ψ the left action by k becomes a left action on f_K, which is invariant, and the right action of γ becomes a left action on $f_{\mathbb{Z}}$, which is also invariant. Hence, $\Psi(\mathcal{V})$ is by construction a function on the double quotient $G(\mathbb{Z})\backslash G/K$.

p-Adic Automorphic Forms

Although very appealing, the method described above is complicated by the fact that the distribution $f_{\mathbb{Z}}$ is in general difficult to obtain. There is however a powerful mathematical technique, developed in [195–197], to compute $f_{\mathbb{Z}}$ using p-adic number theory. In this approach, the distribution $f_{\mathbb{Z}}$ is reinterpreted in terms of a *p-adic spherical vector* f_p that can be straightforwardly constructed from its real counterpart f_K in a way that will be explained shortly. The automorphic form Ψ can then be written in the following way

$$\Psi(\mathcal{V}) = \sum_{\vec{x}\in\mathbb{Q}^n} \Big[\prod_{p<\infty} f_p(\vec{x})\Big]\rho(\mathcal{V}) \cdot f_K(\vec{x}), \tag{9.1.25}$$

where $\vec{x}$ is a vector of rational numbers in $\mathbb{Q}^n$, and the product is over all prime numbers p. This construction is based on the fact that for the field of p-adic numbers the analogue of the maximal compact subgroup K is $G(\mathbb{Z}_p)$, where $\mathbb{Z}_p$ denotes the field of p-adic integers which is a compact subset of the p-adic numbers $\mathbb{Q}_p$ (see Eq. 9.1.28). Hence, in the p-adic framework, the $G(\mathbb{Z})$-invariant vector $f_{\mathbb{Z}}$ simply corresponds to the $G(\mathbb{Z}_p)$-invariant "spherical" vector f_p. Moreover, it was found in [195–197] that once the K-invariant spherical vector f_K is constructed, there is a simple prescription to obtain the p-adic spherical vector f_p. This prescription will be explained in explicit examples below. We will then also see how to evaluate the somewhat intimidating infinite product over primes in (9.1.25), as well as how to recover the lattice summation appearing in the Eisenstein series in (9.1.10) from the sum over rational numbers in $\Psi(\mathcal{V})$. This will reveal that for a certain choice of spherical vector f_K the construction in (9.1.25) is equivalent to the construction of $\mathcal{E}^{G(\mathbb{Z})}(\mathcal{K}; s)$ in (9.1.10). In the next section will introduce some of the basic terminology, and in Sections 9.2 and 9.3 will discuss some detailed examples of how this works.

A Lightning Review of p-Adic Numbers

In order to use the formula (9.1.25) later on, we need to introduce a few basic facts about p-adic number theory. This is a vast subject in mathematics, and the following treatment is merely intended to provide the minimum set of tools required for subsequent sections. For a very nice introduction to p-adic numbers for physicists we recommend [198]. A relatively accessible mathematical treatment is also given in [199].

Perhaps the most intuitive definition of a p-adic number x_p is to consider it as a Laurent series in the prime number p:

$$x_p = c_k p^k + c_{k+1} p^{k+1} + \cdots, \qquad k \in \mathbb{Z}, \tag{9.1.26}$$

with the coefficients c_k being strictly positive integral, and bounded by p. The field of p-adic numbers is denoted by $\mathbb{Q}_p$, and may be considered as a "completion" of the real numbers $\mathbb{R}$ under the p-adic norm $|\cdot|_p$. This norm is defined as

$$|x|_p := p^{-k}, \qquad x \in \mathbb{Q}, \tag{9.1.27}$$

where k is the largest integer such that $x/p^k \in \mathbb{Z}$.

We also have the subset of p-adic integers $\mathbb{Z}_p$ which similary may be considered as the completion of the standard integers $\mathbb{Z}$ under the p-adic norm. They have the following definition:

$$\mathbb{Z}_p = \{x \in \mathbb{Q}_p \mid |x|_p \leq 1\} \subset \mathbb{Q}_p. \tag{9.1.28}$$

We can also extend the p-adic norm to vectors of rational numbers. Recall that the Euclidean norm $|\vec{v}|_\infty = ||\vec{v}||$ of a vector $\vec{v} = (v_1, \ldots, v_n) \in \mathbb{Q}^n$ is defined as

$$||\vec{v}|| := \sqrt{v_1^2 + \cdots v_n^2}. \tag{9.1.29}$$

The p-adic counterpart of this is then given by

$$|\vec{v}|_p := \max(|v_1|_p, \ldots, |v_n|_p). \tag{9.1.30}$$

Note that, in contrast to the Euclidean norm, there is no square root in the definition of $|\vec{v}|_p$.

9.2 Eisenstein series on $SL(2,\mathbb{R})/SO(2)$

In this section we shall illustrate the various constructions discussed above in the context of the standard non-holomorphic $SL(2,\mathbb{Z})$-Eisenstein series on $SL(2,\mathbb{R})/SO(2)$. In Section 8.2 we saw that this Eisenstein series takes the form

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau; s) = \sum_{(m,n)\neq(0,0)} \frac{\tau_2^s}{|m+n\tau|^{2s}}, \tag{9.2.1}$$

where the complex parameter τ parametrizes the upper half plane

$$\mathbb{H} = \{\tau \in \mathbb{C} \mid \tau_2 \geq 0\}. \tag{9.2.2}$$

We will now see how Eq. (9.2.1) may be reproduced using either of the three methods described in Sections 9.1.1, 9.1.2 and 9.1.3.

9.2.1 Poincaré Series

Recall from Section 2.1 that the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ is generated by the Chevalley triple (f, h, e), subject to the relations

$$[h, e] = 2e, \qquad [h, f] = -2f, \qquad [e, f] = h. \tag{9.2.3}$$

Following the logic of Section 9.1.1 we then perform an Iwasawa decomposition of $g \in SL(2,\mathbb{R})$:

$$g = e^{\chi e} e^{-\tilde{\phi} h} k = \begin{pmatrix} 1 & \chi \\ & 1 \end{pmatrix} \begin{pmatrix} e^{-\tilde{\phi}} & \\ & e^{\tilde{\phi}} \end{pmatrix} k, \qquad k \in SO(2), \tag{9.2.4}$$

where

$$a(\tilde{\phi}) = \begin{pmatrix} e^{-\tilde{\phi}} & \\ & e^{\tilde{\phi}} \end{pmatrix} := \Phi^H. \tag{9.2.5}$$

Implementing this in the general formula (9.1.4) yields the Eisenstein series

$$E^{SL(2,\mathbb{Z})}(g;\tilde{s}) = \sum_{\gamma\in N(\mathbb{Z})\backslash SL(2,\mathbb{Z})} \big(\gamma\cdot\Phi\big)^{\tilde{s}}. \tag{9.2.6}$$

To make contact with Eq. 9.2.1 we note that the isomorphism between the upper half plane $\mathbb{H}$ and the coset space $SL(2,\mathbb{R})$ is given by

$$\tau = \tau_1 + i\tau_2 := \chi + ie^{-\phi}, \tag{9.2.7}$$

where $\phi \equiv 2\tilde{\phi}$, and hence $\Phi \equiv \tau_2^2$. Moreover, the action of $SL(2,\mathbb{Z})$ on the upper half plane is given by the standard fractional transformation

$$\gamma\cdot\tau = \frac{a\tau+b}{c\tau+d}, \qquad a,b,c,d\in\mathbb{Z}, \quad ad-bc=1, \tag{9.2.8}$$

which in particular implies

$$\gamma\cdot\Phi = \big(\gamma\cdot\tau_2\big)^2 = \left(\frac{\tau_2}{|c\tau+d|^2}\right)^2. \tag{9.2.9}$$

To understand the sum in Eq. (9.2.6), we must also analyze the structure of the coset $N(\mathbb{Z})\backslash SL(2,\mathbb{Z})$. This corresponds to the equivalence relation

$$SL(2,\mathbb{Z}) \ni \begin{pmatrix} a & b \\ c & d \end{pmatrix} \sim \begin{pmatrix} 1 & \ell \\ & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+\ell c & b+\ell d \\ c & d \end{pmatrix}, \qquad \ell\in\mathbb{Z}, \tag{9.2.10}$$

implying that, in the coset $N(\mathbb{Z})\backslash SL(2,\mathbb{Z})$, a and b are determined by c and d. Hence, the coset $N(\mathbb{Z})\backslash SL(2,\mathbb{Z})$ is completely characterized by the two integers c and d. In addition, the determinant condition $ad-bc=1$ further constrains c and d to be coprime: $(c,d)=1$. Finally, redefining $\tilde{s}\equiv s/2$ and using Eq. 9.2.9 we find that the Eisenstein series (9.2.6) becomes

$$E^{SL(2,\mathbb{Z})}(\tau;s) = {\sum_{(c,d)=1}}' \frac{\tau_2^s}{|c\tau+d|^{2s}}, \tag{9.2.11}$$

where the prime on the sum indicates that the term with $c=d=0$ should be removed. The additional coprime condition compared to Eq. (9.2.1) can be understood as follows. Consider the sum over $m,n\in\mathbb{Z}$ in (9.2.1) and make the change of variables $m\equiv kd$ and $n\equiv kc$, with $k=\gcd(m,n)$ and $(c,d)=1$. Implementing this in (9.2.1) yields

$$\sum_{(m,n)\neq(0,0)} \frac{\tau_2^s}{|m+n\tau|^{2s}} = {\sum_{(c,d)=1}}' \left[\sum_{k\neq 0} k^{-2s}\right] \frac{\tau_2^s}{|c\tau+d|^{2s}} = 2\zeta(2s) {\sum_{(c,d)=1}}' \frac{\tau_2^s}{|c\tau+d|^{2s}}, \tag{9.2.12}$$

where we recall that the Riemann zeta function is given by

$$\zeta(2s) = \sum_{k=1}^{\infty} k^{-2s}. \tag{9.2.13}$$

We may therefore conclude that the two different Eisenstein series are related in the following simple way

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = 2\zeta(2s)E^{SL(2,\mathbb{Z})}(\tau;s). \tag{9.2.14}$$

9.2.2 Lattice Construction

Let us now see how we can construct $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)$ using the method discussed in Section 9.1.2. Again, invoking the Iwasawa decomposition, we may choose a coset representative $\mathcal{V}$ in the Borel gauge as follows

$$\mathcal{V} = e^{\chi e} e^{-\frac{\phi}{2}h} = \begin{pmatrix} e^{-\phi/2} & \\ & e^{\phi/2} \end{pmatrix} \begin{pmatrix} 1 & \chi \\ & 1 \end{pmatrix} \in SL(2,\mathbb{R})/SO(2). \tag{9.2.15}$$

The generalized metric $\mathcal{K}$ then takes the form

$$\mathcal{K} := \mathcal{V}\mathcal{V}^T = \begin{pmatrix} e^{-\phi} + e^{\phi}\chi^2 & e^{\phi}\chi \\ e^{\phi}\chi & e^{\phi} \end{pmatrix}. \tag{9.2.16}$$

According to the general prescription described in Section 9.1.2 the Eisenstein series is obtained by taking a lattice vector $\vec{m} = (n,m) \in \mathbb{Z}^2$, corresponding to the fundamental representation of $SL(2,\mathbb{Z})$, and then summing over the integers m_1, m_2 as follows

$$\begin{aligned} \mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) &= {\sum_{\vec{m}\in\mathbb{Z}^2}}' \Big[\vec{m}^T \cdot \mathcal{K} \cdot \vec{m}\Big]^{-s} \\ &= \sum_{(m,n)\neq(0,0)} \Big[\frac{e^{\phi}}{n^2 + e^{2\phi}(m+n\chi)^2}\Big]^{s}. \end{aligned} \tag{9.2.17}$$

This result agrees with Eq. (9.2.1) if we identify

$$\tau := \chi + ie^{-\phi}, \tag{9.2.18}$$

which is the standard map between the upper half plane $\mathbb{H}$ and the coset space $SL(2,\mathbb{R})/SO(2)$.

9.2.3 Spherical Vector

Finally, we will show how to construct the Eisenstein series $\mathcal{E}_s^{SL(2,\mathbb{Z})}(\tau,\bar{\tau})$ using the more abstract framework of Section 9.1.3. This will be done using induction from the parabolic subgroup

$$P = \left\{ \begin{pmatrix} t_1 & \\ \eta & t_2 \end{pmatrix} \Big| \, \eta, t_1, t_2 \in \mathbb{R};\ t_1 t_2 = 1 \right\}. \tag{9.2.19}$$

Note that P coincides with the Borel subgroup $B_- \subset SL(2,\mathbb{R})$ generated by $\{f, h\}$. A representative for the parabolic coset $P\backslash SL(2,\mathbb{R})$ can now be chosen as

$$n = \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \in P\backslash SL(2,\mathbb{R}). \tag{9.2.20}$$

Functions in $L^2\big(P\backslash SL(2,\mathbb{R})\big)$ corresponding to the principal continuous series then obey

$$\rho(g)\cdot f(n) = f(ng) = f(pn') = \chi_s(p) f(n'), \tag{9.2.21}$$

where the character $\chi_s(p)$ is defined as

$$\chi_s(p) := t_1^{-2s}, \qquad p = \begin{pmatrix} t_1 & \\ \eta & t_2 \end{pmatrix} \in P. \tag{9.2.22}$$

We want to find a suitable spherical vector, i.e. an $SO(2)$-invariant function f_K belonging to $L^2\big(P\backslash SL(2,\mathbb{R})\big)$. The maximal compact subgroup $SO(2)$ acts on $n \in P\backslash SL(2,\mathbb{R})$ from the right, and a compensating transformation by $p \in P$ from the left is required to restore the upper-triangular form of (9.2.20). Since $k \in SO(2)$ acts on n as a rotation, it will by definition leave invariant the Euclidean norms $||\vec{r}_i||$ of the two rows $\vec{r}_1$ and $\vec{r}_2$ in n. Note furthermore that the compensating transformation of $p \in P$ acts simply by multiplication on the top row:

$$pn = \begin{pmatrix} t_1 & t_1 x \\ \eta & t_2 + \eta x \end{pmatrix}. \tag{9.2.23}$$

We may therefore take the spherical vector f_K to be defined as the Euclidean norm of the top row of n:

$$f_K(x) := ||\vec{r}_1||^{-2s} = (1+x^2)^{-s}. \tag{9.2.24}$$

Although the compensating action of p from the left modifies f_K by an overall factor t_1^{2s}, we must also include the character $\chi_s(p) = t_1^{-2s}$ since f_K belongs to the principal continuous series. Hence, we find that $f_K(x)$ is indeed invariant, and corresponds to the desired spherical vector.

To construct the automorphic form $\Psi(\mathcal{V})$, we must also compute the action of $\rho(\mathcal{V})$ on $f_K(x)$, where $\mathcal{V} \in SL(2,\mathbb{R})/SO(2)$ is the standard coset representative, Eq. (9.2.15), in the Borel gauge. This is straightforward, following the prescription in (9.2.21): the action of $\rho(\mathcal{V})$ on any function $f(n)$ is given by

$$\rho(\mathcal{V}) \cdot f(n) = f(n\mathcal{V}) = f(p_0 n') = \chi_s(p_0) f(n') = e^{s\phi} f(n'), \tag{9.2.25}$$

where

$$p_0 = \begin{pmatrix} e^{-\phi/2} & \\ & e^{\phi/2} \end{pmatrix}, \qquad n' = \begin{pmatrix} 1 & e^{\phi}(x+\chi) \\ & 1 \end{pmatrix}. \tag{9.2.26}$$

The spherical vector thus transforms as follows

$$\rho(\mathcal{V}) \cdot f_K(x) = e^{s\phi}\big(1 + e^{2\phi}(x+\chi)^2\big)^{-s}. \tag{9.2.27}$$

The last missing piece in the construction of $\Psi(\mathcal{V})$ is the p-adic spherical vector f_p. It was found in [195–197] that f_p can be constructed according to the following simple prescription: replace the Euclidean norm $||\cdot||$ in f_K by the p-adic version $|\cdot|_p$.[4] Hence, we take the following definition of the p-adic spherical vector

$$f_p(x) := |\vec{r}_1|_p^{-2s} = \max\big(1, |x|_p\big)^{-2s}, \tag{9.2.28}$$

[4]This procedure works because the action of the group is (projective) linear in this representation. I am grateful to Boris Pioline for pointing this out to me.

where the p-adic norm $|\cdot|_p$ was defined in Section 9.1.3. Inserting Eq. (9.2.27) and Eq. (9.2.28) into Eq. (9.1.25), we find that the automorphic form $\Psi(\mathcal{V})$ takes the form

$$\Psi(\mathcal{V}) = \sum_{x\in\mathbb{Q}} \Big[\prod_{p<\infty} \max\big(1, |x|_p\big)^{-2s}\Big] e^{s\phi}\big(1+e^{2\phi}(x+\chi)^2\big)^{-s}. \tag{9.2.29}$$

In order to make sense of this expression, we must evaluate the infinite sum over prime numbers. To this end we first make a change of variables in the summation, and write the rational number $x \in \mathbb{Q}$ as

$$x = \frac{m}{n}, \qquad m, n \in \mathbb{Z},\ n \neq 0. \tag{9.2.30}$$

This yields

$$\Psi(\mathcal{V}) = \sum_{m\in\mathbb{Z}}\sum_{n\neq 0} \Big[\prod_{p<\infty} \max\Big(1, \Big|\frac{m}{n}\Big|_p\Big)^{-2s}\Big] e^{s\phi}\Big(1+e^{2\phi}\big(\frac{m}{n}+\chi\big)^2\Big)^{-s}, \tag{9.2.31}$$

Now we can evaluate the product over primes. Suppose first that $m/n \in \mathbb{Z}$. Then we know by definition of the p-adic norm that $\left|\frac{m}{n}\right|_p < 1$, and hence the max-function will simply pick out the first component:

$$\prod_{p<\infty} \max\Big[1, \Big|\frac{m}{n}\Big|_p\Big] = 1, \qquad \forall\ m/n \in \mathbb{Z}. \tag{9.2.32}$$

On the other hand, when m and n are coprime, the maximum value of f_p for each p will be when $p = n$, for which we have

$$\prod_{p<\infty} \max\Big[1, \Big|\frac{m}{n}\Big|_p\Big] = n. \tag{9.2.33}$$

All other cases, i.e. when $p \neq n$, will again give unity. Hence, the result is

$$\begin{aligned}\Psi(\mathcal{V}) &= \sum_{(m,n)=1, n\neq 0} n^{-2s} e^{s\phi}\left[1+e^{2\phi}\Big(\frac{m}{n}+\chi\Big)^2\right]^{-s} \\ &= \sum_{(m,n)=1, n\neq 0} \frac{\tau_2^s}{|m+n\tau|^{2s}},\end{aligned} \tag{9.2.34}$$

where $\tau = \chi + ie^{-\phi}$ as usual. We conclude that indeed for f_K in the principal continuous series, this final form of $\Psi(\mathcal{V})$ coincides with the non-holomorphic Eisenstein series $\mathcal{E}_s^{SL(2,\mathbb{Z})}(\tau,\bar{\tau})$ in Eq. (9.2.1), modulo the missing term corresponding to $n = 0$.

9.3 Eisenstein Series on $SL(3,\mathbb{R})/SO(3)$

We will now consider in detail also the example of $SL(3,\mathbb{Z})$-invariant Eisenstein series in the principal continuous series of $SL(3,\mathbb{R})$-representations.

9.3.1 Poincaré Series

The rank 2 Lie algebra $\mathfrak{sl}(3,\mathbb{R})$ is generated by the two Chevalley triples (e_1, f_1, h_1) and (e_2, f_2, h_2) represented in the fundamental representation by the following matrices:

$$e_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad h_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad h_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{9.3.1}$$

together with $f_i = (f_i)^T$, $i = 1, 2$. An arbitrary element $g \in SL(3,\mathbb{R})$ may then be represented in Iwasawa form as

$$\begin{aligned} g &= \exp\left[ve_1 + \frac{1}{2}\mathcal{A}_2 e_2 + \left(\mathcal{A}_1 - \frac{1}{2}v\mathcal{A}_1\right)e_3\right] \exp\left[-\tilde{\phi}_1 h_1 - \tilde{\phi}_2 h_2\right] k \\ &= \begin{pmatrix} 1 & v & \mathcal{A}_1 \\ & 1 & \mathcal{A}_2 \\ & & 1 \end{pmatrix} \begin{pmatrix} e^{\tilde{\phi}_1} & & \\ & e^{-\tilde{\phi}_1 + \tilde{\phi}_2} & \\ & & e^{-\tilde{\phi}_2} \end{pmatrix} k, \end{aligned} \tag{9.3.2}$$

where $e_3 = [e_1, e_2]$. The Cartan torus is thus parametrized by

$$a(\tilde{\phi}) = \begin{pmatrix} e^{\tilde{\phi}_1} & & \\ & e^{-\tilde{\phi}_1 + \tilde{\phi}_2} & \\ & & e^{-\tilde{\phi}_2} \end{pmatrix} := \left(\Phi_1\right)^{h_1} \left(\Phi_2\right)^{h_2}. \tag{9.3.3}$$

From Eq. (9.1.4) we then find that an $SL(3,\mathbb{Z})$-invariant Eisenstein series in the principal series of $SL(3,\mathbb{R})$ is constructed as follows

$$E^{SL(3,\mathbb{Z})}(g; \tilde{s}_1, \tilde{s}_2) = \sum_{\gamma \in N(\mathbb{Z}) \backslash SL(3,\mathbb{Z})} \left(\gamma \cdot \Phi_1\right)^{\tilde{s}_1} \left(\gamma \cdot \Phi_2\right)^{\tilde{s}_2}. \tag{9.3.4}$$

For the purpose of simplifying the analysis in Section 10.3, it will be useful to make a slight change of variables which is in better accordance with the conventions of [200]:

$$\Phi_1 = y^{-1/6} u^{-1/2}, \qquad \Phi_2 = y^{-1/3}, \tag{9.3.5}$$

and we also redefine the parameters $\tilde{s}_1$ and $\tilde{s}_2$ as follows

$$\tilde{s}_1 = -2s_2, \qquad \tilde{s}_2 = -2s_1. \tag{9.3.6}$$

Implementing these changes in (9.3.4) yields

$$E^{SL(3,\mathbb{Z})}(g; s_1, s_2) = \sum_{\gamma \in N(\mathbb{Z}) \backslash SL(3,\mathbb{Z})} \left(\gamma \cdot y\right)^{\frac{2s_1 + s_2}{3}} \left(\gamma \cdot u\right)^{s_2}. \tag{9.3.7}$$

To write this out explicitly, we consider an arbitrary $SL(3,\mathbb{Z})$-element of the form

$$\gamma = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \in SL(3,\mathbb{Z}), \tag{9.3.8}$$

and the Eisenstein series may then be written as

$$\begin{aligned} E^{SL(3,\mathbb{Z})}(g;s_1,s_2) &= y^{\frac{s_2-s_1}{3}} \sum_{(c,b)\in\mathbb{Z}^6}{}' \bigg\{ \Big[c_1^2 u + \frac{1}{u}\big(c_2+c_1 v\big)^2 + \frac{1}{y}\big(c_1\mathcal{A}_1+c_2\mathcal{A}_2+c_3\big)^2\Big]^{-s_1} \\ &\qquad \times\Big[yD_3^2 + u\big(D_2-\mathcal{A}_2 D_3\big)^2 + \frac{1}{u}\big(D_1 - vD_2 - (\mathcal{A}_1 - v\mathcal{A}_2)D_3\big)^2\Big]^{-s_2}\bigg\}, \end{aligned} \tag{9.3.9}$$

where we defined

$$D_1 := b_2c_3 - b_3c_2, \qquad D_2 := b_3c_1 - b_1c_3, \qquad D_3 := b_1c_2 - b_2c_1. \tag{9.3.10}$$

The summation variables in (9.3.9) are further constrained by the structure of the coset $N(\mathbb{Z})\backslash SL(3,\mathbb{Z})$, and it was shown in [200] that this implies the following relations:

$$\begin{aligned} (c_1,c_2,c_3) = (D_1,D_2,D_3) &= 1, \\ c_1D_1+c_2D_2+c_3D_3 &= 0, \\ c_1, D_1 &\geq 0, \end{aligned} \tag{9.3.11}$$

where we used the notation $(x,y,z)=\gcd(x,y,z)$.

We will see in Section 10.3 that these variables are natural when computing the Fourier expansion. In particular, it will be important that the complex parameter $z=v+iu$ transforms in the standard fractional way under a certain $SL(2,\mathbb{Z})$ subgroup of $SL(3,\mathbb{Z})$.

9.3.2 Lattice Construction

Let us now proceed to the method of Section 9.1.2. To this end we begin by choosing the coset representative

$$\mathcal{V} = \begin{pmatrix} 1 & \mathcal{A} & \mathcal{C} \\ & 1 & \mathcal{B} \\ & & 1 \end{pmatrix} \begin{pmatrix} 1/\eta_1 & & \\ & \eta_1/\eta_2 & \\ & & \eta_2 \end{pmatrix} \in SL(3,\mathbb{R})/SO(3). \tag{9.3.12}$$

We are here using slightly different variables compared to the previous subsection in order to make the comparison with the analysis in Section 9.3.3 straightforward. Below we will also explain the relation with the variables used in Eq. (9.3.9).

We proceed to construct the $SO(3)$-invariant generalized metric:

$$\mathcal{K} = \mathcal{V}\mathcal{V}^T = \begin{pmatrix} \frac{1}{\eta_1^2} + \frac{\eta_1^2}{\eta_2^2}\mathcal{A} + \eta_2^2\mathcal{C}^2 & \frac{\eta_1^2}{\eta_2^2}\mathcal{A} + \eta_2^2\mathcal{B}\mathcal{C} & \eta_2^2\mathcal{C} \\ \frac{\eta_1^2}{\eta_2^2}\mathcal{A} + \eta_2^2\mathcal{B}\mathcal{C} & \frac{\eta_1^2}{\eta_2^2} + \eta_2^2\mathcal{B}^2 & \eta_2^2\mathcal{B} \\ \eta_2^2\mathcal{C} & \eta_2^2\mathcal{B} & \eta_2^2 \end{pmatrix}. \tag{9.3.13}$$

We furthermore choose a lattice vector $m=(m_3,m_2,m_1)\in\mathbb{Z}^3$ in the fundamental representation of $SL(3,\mathbb{Z})$. Using Eq. (9.1.10) we then obtain the following Eisenstein series:

$$\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s) = \sum_{m\in\mathbb{Z}^3}{}' \Big[m^T\cdot\mathcal{M}\cdot m\Big]^{-s}$$

$$= \sum_{(m_1,m_2,m_3)\neq(0,0,0)} \left[\frac{m_3^2}{\eta_1^2} + \frac{\eta_1^2}{\eta_2^2}\big(m_2+m_3\mathcal{A}\big)^2 + \eta_2^2\big(m_1+m_2\mathcal{B}+m_3\mathcal{C}\big)^2\right]^{-s}. \tag{9.3.14}$$

To compare this result with Eq. (9.3.9) we change variables to

$$\mathcal{A} = v, \quad \mathcal{B} = \mathcal{A}_2, \quad \mathcal{C} = \mathcal{A}_1, \quad \eta_1 = y^{-1/6}u^{-1/2}, \quad \eta_2 = y^{-1/3}, \tag{9.3.15}$$

after which (9.3.14) becomes

$$\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s) = y^{-s/3} \sum_{m\in\mathbb{Z}^3}{}' \left[m_1^2 u + \frac{1}{u}\big(m_2+m_1 v\big)^2 + \frac{1}{y}\big(m_1\mathcal{A}_1+m_2\mathcal{A}_2+m_3\big)^2\right]^{-s_1}. \tag{9.3.16}$$

Here we recognize the first summand of Eq. (9.3.9), while the range of summation is different because of the extra coprime condition $(c_1,c_2,c_3)=1$ in (9.3.9). This can be understood in the same way as in the $SL(2,\mathbb{R})$-example in Section 9.2, namely redefine $m_i \equiv kc_i,\ i=1,2,3$, with $(c_1,c_2,c_3)=1$ and $k=(m_1,m_2,m_3)$. After this change of summation variables, Eq. (9.3.16) becomes

$$\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s) = 2\zeta(2s)y^{-s/3} \sum_{(c_1,c_2,c_3)=1}{}' \left[c_1^2 u+\frac{1}{u}\big(c_2+c_1 v\big)^2+\frac{1}{y}\big(c_1\mathcal{A}_1+c_2\mathcal{A}_2+c_3\big)^2\right]^{-s}. \tag{9.3.17}$$

From this result we conclude that the minimal Eisenstein series arises as the limiting case $s_2 \to 0$ of the general principal series in (9.3.9), i.e.

$$\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s_1) = \lim_{s_2\to 0}\left[2\zeta(s_1)E^{SL(3,\mathbb{Z})}(g;s_1,s_2)\right]. \tag{9.3.18}$$

It is in fact possible to write the full principal series in terms of the method of Section 9.1.2. In this case, one introduces two lattice vectors $m\in\mathbb{Z}^3$ and $n\in\mathbb{Z}^3$ and consider the following series [201]

$$\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s_1,s_2) = \sum_{(m,n)\in\mathbb{Z}^6}{}' \left[m^T\cdot\mathcal{K}\cdot m\right]^{-s_2}\left[\big(m^T\cdot\mathcal{K}\cdot m\big)\big(n^T\cdot\mathcal{K}\cdot n\big) - \big(m^T\cdot\mathcal{K}\cdot n\big)^2\right]^{-s_2}, \tag{9.3.19}$$

where the summation excludes $(m_1,m_2,m_3)=(0,0,0)$ and $(n_1,n_2,n_3)=(0,0,0)$, modulo the equivalence relation $n\sim n+m$. The relation with the principal series in Eq. (9.3.7) again becomes apparent after eliminating the coprime condition, which then yields a generalized version of Eq. (9.2.14) [201]:

$$\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s_1,s_2) = 4\zeta(2s_1+2s_2)\zeta(2s_2)E^{SL(3,\mathbb{Z})}(g;s_1,s_2). \tag{9.3.20}$$

9.3.3 Spherical Vector

In this section we will finally reproduce the minimal Eisenstein series in Eq. (9.3.14) using the general approach of Section 9.1.3, based on the spherical vector f_K. As in the case of

$SL(2,\mathbb{R})$ above, we shall consider representations ρ induced from the parabolic subgroup consisting of lower-triangular matrices

$$P = \left\{ \begin{pmatrix} t_1 & & \\ * & t_2 & \\ * & * & t_3 \end{pmatrix} \,\Big|\, t_1 t_2 t_3 = 1 \right\} \subset SL(3,\mathbb{R}). \tag{9.3.21}$$

A natural representative for $P\backslash SL(3,\mathbb{R})$ is given by

$$n = \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} \in P\backslash SL(3,\mathbb{R}). \tag{9.3.22}$$

We consider the representation ρ of $g \in SL(3,\mathbb{R})$ induced from P which acts on functions $f(n)$ on $P\backslash SL(3,\mathbb{R})$ by

$$\rho(g) \cdot f(n) = f(ng) = f(pn') = \chi_s(p) f(n'), \tag{9.3.23}$$

where $\chi_s(p)$ is the character

$$\chi_s(p) := t_1^{-2s}. \tag{9.3.24}$$

We have here chosen the simplest possible character for the purpose of reproducing the minimal Eisenstein series $\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s)$. The general principal series $\mathcal{E}^{SL(3,\mathbb{Z})}(\mathcal{K};s_1,s_2)$ may also be constructed by induction from the parabolic subgroup P in (9.3.21). However, carrying out the full program for this case is much more complicated and we will therefore refrain from doing so here. See however [195, 201] for a discussion.

The element $ng \in SL(3,\mathbb{R})$ is generically not in the form of (9.3.22), but by a compensating transformation of $p \in P$ from the right we obtained $ng = pn'$, with $n' \in P\backslash SL(3,\mathbb{R})$. The specific choice of character in (9.3.24) corresponds to a certain limit of the principal continuous series of $SL(3,\mathbb{R})$ representation (see [201]). In fact, it restricts from the reducible representation acting on functions $f(x,y,z)$ of three variables to an irreducible representation acting on functions of two variables, $f(x,z) := f(x,1,z)$, i.e. to the minimal representation. The spherical vector $f_K(x,z)$ is defined by the condition

$$\rho(k) \cdot f_K(x,z) = f_K(x,z), \qquad k \in SO(3). \tag{9.3.25}$$

A simple way of determining $f_K(x,z)$ is to consider the action of $SO(3)$ on $n \in P\backslash SL(3,\mathbb{R})$. $SO(3)$ acts from the right on n and a compensating P-transformation from the left is needed to restore the upper-triangular form. The action of $p \in P$ on n is

$$pn = \begin{pmatrix} t_1 & & \\ p_{21} & t_2 & \\ p_{31} & p_{32} & t_3 \end{pmatrix} \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} = \begin{pmatrix} t_1 & t_1 x & t_1 z \\ p_{21} & t_2 + p_{21} x & t_2 y + p_{21} z \\ p_{31} & p_{32} + p_{31} x & t_3 + p_{32} y + p_{31} z \end{pmatrix}. \tag{9.3.26}$$

Since the action of $K = SO(3)$ on n is just a rotation, it preserves the Euclidean norm of each row $\vec{r}_i$ of n. If we act by $k \in K$ on n from the right we will however destroy the upper triangular form of n, and a compensating transformation by $p \in P$ from the left is required. Moreover, we see that the action of P on the top row $\vec{r}_1 = (1,x,z)$ is very simple: $p \cdot \vec{r}_1 = t_3 \vec{r}_1$.

The spherical vector $f_K(x,z)$, in the minimal representation, can therefore be taken as the Euclidean norm of the vector $\vec{r}_1$ raised to the power $-2s$:

$$f_K(x,z) := ||\vec{r}_1||^{-2s} = (1+x^2+z^2)^{-s}. \tag{9.3.27}$$

The action of $k \in SO(3)$ will transform f_K by an overall factor t_1^{2s} because of the compensating P-transformation in 9.3.26, but since f_K belongs to the principal continuous series we must also include the character $\chi_s(p) = t_1^{-2s}$ from (9.3.24), implying that $f_K(x,z)$ is indeed invariant.

To find the action of $\mathcal{V} \in SL(3,\mathbb{R})/SO(3)$ on the spherical vector in this representation, we simply apply the formula (9.3.23) on f_K, which yields

$$\rho(\mathcal{V}) \cdot f_K(n) = f_K(n\mathcal{V}) = f_K(p_0 n') = \chi_s(p_0) f_K(n') = \eta_1^{2s} ||\vec{r}_1{}'||^{-2s}, \tag{9.3.28}$$

where

$$\begin{aligned} n' &= \begin{pmatrix} 1 & \frac{\eta_1^2}{\eta_2}\big(x+\mathcal{A}\big) & \eta_1\eta_2\big(z+x\mathcal{B}+\mathcal{C}\big) \\ & 1 & \frac{\eta_2^2}{\eta_1}\big(y+\mathcal{B}\big) \\ & & 1 \end{pmatrix}, \\ p &= \begin{pmatrix} 1/\eta_1 & & \\ & \eta_1/\eta_2 & \\ & & \eta_2 \end{pmatrix}. \end{aligned} \tag{9.3.29}$$

The transformed spherical vector thus corresponds to the norm squared $||\vec{r}_1{}'||^{-2s}$ of the top row $\vec{r}_1{}'$ of the transformed matrix $n' \in P\backslash SL(3,\mathbb{R})$. In addition, we must include the overall character η_1^{2s} and the final result is

$$\rho(\mathcal{V}) \cdot f_K(x,z) = \eta_1^{2s}\Big[1 + \frac{\eta_1^4}{\eta_2^2}\big(x+\mathcal{A}\big)^2 + \eta_1^2\eta_2^2\big(z+x\mathcal{B}+\mathcal{C}\big)^2\Big]^{-s}. \tag{9.3.30}$$

The p-adic spherical vector is again obtained by replacing the Euclidean norm $||\cdot||$ in $f_K(x,z)$ by its p-adic version, which yields

$$f_p(x,z) := |1,x,z|_p^{-2s} = \max\big(1, |x|_p, |z|_p\big)^{-2s}. \tag{9.3.31}$$

At this stage, the automorphic form $\Psi(\mathcal{V})$ in Eq. (9.1.25) for the minimal representation of $SL(3,\mathbb{R})$ thus reads

$$\Psi(\mathcal{V}) = \sum_{(x,z)\in\mathbb{Q}^2} \Big[\prod_{p<\infty} \max\big(1,|x|_p,|z|_p\big)^{-2s}\Big]\Big[\frac{1}{\eta_1^2} + \frac{\eta_1^2}{\eta_2^2}\big(x+\mathcal{A}\big)^2 + \eta_2^2(z+x\mathcal{B}+\mathcal{C}\big)^2\Big]^{-s}. \tag{9.3.32}$$

To proceed, we must again evaluate the infinite product over primes. This can be done by the same method as for the $SL(2,\mathbb{R})$-example in the previous subsection. We begin by extracting the greatest common divisor between the rational summation variables x and z:

$$x = \frac{m_2}{m_3}, \qquad z = \frac{m_1}{m_3}, \tag{9.3.33}$$

where $m_i \in \mathbb{Z}$ ($m_3 \neq 0$), and $(m_1, m_2) = 1$. By the same reasoning as before the infinite product over the p-adic spherical vector becomes

$$\prod_{p<\infty} \max\Big(1, \Big|\frac{m_1}{m_3}\Big|_p, \Big|\frac{m_2}{m_3}\Big|_p\Big)^{-2s} = m_3^{-2s}. \tag{9.3.34}$$

Inserting this result into Eq. (9.3.32) then yields

$$\Psi(\mathcal{V}) = \sum_{(m_1,m_2)=1} \sum_{m_3 \neq 0} \left[\frac{m_3^2}{\eta_1^2} + \frac{\eta_1^2}{\eta_2^2}\big(m_2 + m_3\mathcal{A}\big)^2 + \eta_2^2\big(m_1 + m_2\mathcal{B} + m_3\mathcal{C}\big)^2\right]^{-s}. \tag{9.3.35}$$

This result indeed agrees with the Eisenstein series in Eq. (9.3.14), up to the missing term in the sum with $m_3 = 0$, and the extra coprime condition $(m_1, m_2) = 1$ on the summation.

10

Fourier Expansions of Automorphic Forms

We have now learned how to construct $G(\mathbb{Z})$-invariant automorphic forms on a symmetric space G/K in a variety of different ways. The remaining task is then to extract the physical information that they contain. For the simplest case of $G(\mathbb{Z}) = SL(2, \mathbb{Z})$ we have seen in Chapter 8 that this is done through Fourier expansion, which recasts the automorphic form into an infinite sum, revealing perturbative and non-perturbative quantum effects. In this chapter we shall discuss in more detail how to compute such Fourier expansions, again with particular emphasis on the special examples of $SL(2, \mathbb{Z})$ and $SL(3, \mathbb{Z})$. We begin in Section 10.1 by explaining the main philosophy, stressing certain key points, while in Section 10.2 and 10.3 we consider explicit examples. Section 10.3 is based on **Paper VII**, written in collaboration with Boris Pioline.

10.1 General Considerations

In this section[1] we shall discuss some general features of the Fourier expansion of Eisenstein series $E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}})$, constructed as in Section 12.4.2. We will be deliberately schematic in order to emphasize the key issues. Details are relegated to the examples in the following sections. We begin by noting the important fact that by virtue of the Iwasawa decomposition $g = nak \in G$, the function $E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}})$ is by construction invariant under the nilpotent subgroup $N(\mathbb{Z}) \subset G(\mathbb{Z})$,

$$E^{G(\mathbb{Z})}(xg; \lambda_{\mathcal{R}}) = E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}}), \qquad x \in N(\mathbb{Z}). \tag{10.1.1}$$

[1] I thank Axel Kleinschmidt for helpful remarks which considerably improved the logic of the presentation in this section.

To understand the implications of this fact, we parametrize the element $n \in N$ in the standard way (see Chapter 5 where this was called the "Borel gauge")

$$n = \exp\Big[\sum_{\alpha\in\Phi_+} \chi_\alpha E_\alpha\Big] \in N, \tag{10.1.2}$$

where Φ_+ denotes the positive root system of the Lie algebra $\mathfrak{g} = \text{Lie } G$ (see Section 2.2.3), and E_α are the positive root generators. We shall refer to the parameters χ_α as "axions", by analogy with the axion χ in the type IIB example of Chapter 8. In this parametrization, it is clear that the left action of $x \in N(\mathbb{Z})$ on $g = nak$ will leave the "dilatonic" parameters ϕ of $a \in A$ invariant, while acting as integer-valued shifts on the axions:

$$N(\mathbb{Z}) \ni x : \chi_\alpha \longmapsto \chi_\alpha + \ell_\alpha, \tag{10.1.3}$$

where $\ell_\alpha \in \mathbb{Z}$ are the corresponding parameters of $x \in N(\mathbb{Z})$. This expression is somewhat schematic but nevertheless illustrates the main point. By comparing (10.1.3) with (10.1.2), we then deduce that $E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}})$ exhibits a *periodicity* in the set of axionic moduli $\{\chi_\alpha\}$, schematically indicated as follows

$$E^{G(\mathbb{Z})}\Big(\phi, \{\chi_\alpha\}; \lambda_{\mathcal{R}}\Big) = E^{G(\mathbb{Z})}\Big(\phi, \{\chi_\alpha + \ell_\alpha\}; \lambda_{\mathcal{R}}\Big), \tag{10.1.4}$$

where we collectively denoted the Cartan parameters by ϕ. This periodicity ensures that $E^{G(\mathbb{Z})}(xg; \lambda_{\mathcal{R}})$ has a Fourier expansion of the general form

$$E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}}) = \sum_{\ell_\alpha\in\mathbb{Z}^D} \mathfrak{C}_{\ell_\alpha}(\phi) \exp\Big[2\pi i \sum_{\alpha\in\Phi_+} \ell_\alpha \chi_\alpha\Big], \tag{10.1.5}$$

where $D = \dim \text{Lie } N$. From this expression it is clear that the Fourier expansion essentially corresponds to a diagonalization of the action of the nilpotent subgroup $N(\mathbb{Z}) \subset G(\mathbb{Z})$.

Although morally true, in the discussion above we have actually suppressed the main subtlety: the nilpotent group N is generically not abelian and therefore can not be diagonalized according to the general form indicated in (10.1.5). This implies that in the general case when N is non-abelian, Eq. (10.1.3) is actually *wrong*. To remedy this, one must decompose the Fourier expansion into an *abelian* part and a *non-abelian* part:

$$E^{G(\mathbb{Z})}(g; \lambda_{\mathcal{R}}) = E_A(g) + E_{NA}(g), \tag{10.1.6}$$

where the abelian term corresponds to the Fourier expansion with respect to the "abelianized" nilpotent group $\tilde{N} \equiv N/Z$. Here Z denotes the center of N. The non-abelian part $E_{NA}(g)$ then represents the remaining decomposition with respect to the center Z. We will see in Section 10.3 that for $G = SL(3, \mathbb{R})$, N is isomorphic to a three-dimensional Heisenberg group, and in this case the center Z is one-dimensional, corresponding to the commutator subgroup $[N, N]$.

Let us proceed to write the correct form of (10.1.3) when N is non-abelian. To this end, let $\tilde{\chi}_i$, $i = 1, \ldots, \tilde{D} = \dim \text{Lie } \tilde{N}$ denote the parameters of the abelianized group $\tilde{N}$, and let z_a, $a = 1, \ldots, d = \dim \text{Lie } Z$ denote the parameters of the center Z. The abelianized discrete group $\tilde{N}(\mathbb{Z})$ then acts on the scalars $\tilde{\chi}_i$ in the same way as in (10.1.3),

$$\tilde{N}(\mathbb{Z}) \ni \tilde{x} : \tilde{\chi}_i \longmapsto \tilde{\chi}_i + \tilde{\ell}_i, \tag{10.1.7}$$

while the center Z leaves all $\tilde{\chi}_i$ invariant. On the other hand, the full group $N(\mathbb{Z}) = \tilde{N}(\mathbb{Z}) \times Z(\mathbb{Z})$ acts non-trivially on the scalars z_a:

$$\tilde{N}(\mathbb{Z}) \times Z(\mathbb{Z}) \ni y : z_a \longmapsto z_a + k_a + \sum_{i=1}^{\tilde{D}} c_{ij}\tilde{\ell}_i\tilde{\chi}_j, \qquad k_a \in \mathbb{Z}, \tag{10.1.8}$$

for some numerical coefficients c_{ij} which depend on the group N. As a consequence of (10.1.8), the Fourier expansion will no longer display a complete separation of variables between the Cartan parameters ϕ and the axions χ_α, in contrast to the abelian case (10.1.5). The correct form of the non-abelian Fourier expansion then reads

$$\begin{aligned} E^{G(\mathbb{Z})}(g;\lambda_{\mathcal{R}}) \quad = \quad & \sum_{\tilde{\ell}\in\mathbb{Z}^{\tilde{D}}} \mathfrak{C}_{\tilde{\ell}}(\phi) \exp\left[2\pi i \sum_{i=1}^{\tilde{D}} \tilde{\ell}_i\tilde{\chi}_i\right] \\ & + \sum_{k\in\mathbb{Z}^d} \mathfrak{C}_\ell\big(\phi,\tilde{\chi}_1,\ldots,\tilde{\chi}_{\tilde{D}}\big) \exp\left[2\pi i \sum_{a=1}^{d} k_a z_a\right]. \end{aligned} \tag{10.1.9}$$

We emphasize that in the second (non-abelian) term the coefficients $\mathfrak{C}_\ell\big(\phi,\tilde{\chi}_i\big)$ are periodic functions of the axionic scalars $\tilde{\chi}_i$ with respect to the action of the abelianized group $\tilde{N}(\mathbb{Z})$.

Let us conclude this section with a brief discussion of the "perturbative" part of the Fourier expansion corresponding to the zeroth Fourier coefficients $\mathfrak{C}_0$ in (10.1.9). In mathematical terminology, these are known as the *constant terms* and effectively correspond to integrating out all of the axionic scalars:

$$\begin{aligned} \mathfrak{C}_0\big(\phi\big) \quad &= \quad \int_{N(\mathbb{Z})\backslash N} E^{G(\mathbb{Z})}(g;\lambda_{\mathcal{R}})dn \\ &= \quad \int_0^1 d\tilde{\chi}_1\cdots\int_0^1 \tilde{\chi}_{\tilde{D}}\int_0^1 dz_1\cdots\int_0^1 dz_d\ E^{G(\mathbb{Z})}\big(\phi,\tilde{\chi}_i,z_a;\lambda_{\mathcal{R}}\big). \end{aligned} \tag{10.1.10}$$

Being completely independent of the axionic scalars, these are precisely the terms which encode the perturbative contributions to the Fourier expansion, analogously to the tree-level and one-loop term in Eq. (8.2.8).

We turn now to explicit examples of Fourier expansions, starting with simplest case of $G(\mathbb{Z}) = SL(2,\mathbb{Z})$, in which case N is abelian and the complications discussed above are absent. However, in the example of $SL(3,\mathbb{Z})$ considered in Section 10.2 we will explicitly see the structure of (10.1.9) emerging.

10.2 Fourier Expansion of $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)$

As our first example, we shall compute in detail the Fourier expansion of the non-holomorphic Eisenstein series $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)$ discussed in Section 8.2. We begin by analyzing the general structure of the expansion by utilizing the fact that $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)$ is an eigenfunction of the Laplacian on $SL(2,\mathbb{R})$. In Section 10.2.2 we then use Poisson resummation to explicitly evaluate the Fourier coefficients.

10.2.1 General Structure

Recall that the Eisenstein series $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)$ takes the form

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = \sum_{(m,n)\neq(0,0)} \frac{\tau_2^s}{|m+n\tau|^{2s}}, \tag{10.2.1}$$

where $\tau = \tau_1 + i\tau_2$ parametrizes the coset space space $SL(2,\mathbb{R})/SO(2)$ via the standard map discussed in Section 9.2. We have also seen in Section 9.2 that $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)$ is by construction invariant under the discrete subgroup $SL(2,\mathbb{Z})$ which acts in the following way on the coordinate τ:

$$\tau \longmapsto \frac{a\tau+b}{c\tau+d}, \qquad a,b,c,d\in\mathbb{Z}, \quad ad-bc=1. \tag{10.2.2}$$

This group is generated by two transformations S and T (see, e.g. [202], and Section 2.4.3), acting on τ as

$$S\,:\,\tau \longmapsto -\frac{1}{\tau}, \qquad T\,:\,\tau \longmapsto \tau+1. \tag{10.2.3}$$

Let us for a moment focus on the shift transformation T. Notice that this leaves invariant the imaginary part of τ, while shifts the real part according to $\tau_1 \longmapsto \tau_1+1$. Since the Eisenstein series is invariant under the entire group $SL(2,\mathbb{Z})$ it will in particular satisfy

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau_1,\tau_2;s) = \mathcal{E}^{SL(2,\mathbb{Z})}(\tau_1+1,\tau_2;s), \tag{10.2.4}$$

implying that it is periodic in the variable τ_1 with period 1, in accordance with the general discussion of the previous section. This ensures that $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)$ has a Fourier expansion of the form

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = \sum_{N\in\mathbb{Z}} \mathfrak{C}_N(\tau_2)e^{2\pi i N\tau_1}, \tag{10.2.5}$$

where the Fourier coefficients $\mathfrak{C}_N(\tau_2)$ do not depend on the real part of τ. It is useful to separate out the zeroth Fourier coefficient $\mathfrak{C}_0$ and write

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = \mathfrak{C}_0(\tau_2) + \sum_{N\neq 0} \mathfrak{C}_N(\tau_2)e^{2\pi i N\tau_1}. \tag{10.2.6}$$

The zeroth coefficient corresponds to the *constant term*

$$\mathfrak{C}_N(\tau_2) = \int_{N(\mathbb{Z})\backslash N} \mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)dn = \int_0^1 \mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)d\tau_1, \tag{10.2.7}$$

as discussed in general in Section 10.1. Now recall further that the Eisenstein series is an eigenfunction of the Laplacian on $SL(2,\mathbb{R})/SO(2)$:

$$\Delta\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = s(s-1)\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s), \tag{10.2.8}$$

with

$$\Delta = \tau_2^2\Big(\frac{\partial^2}{\partial\tau_1^2} + \frac{\partial^2}{\partial\tau_2^2}\Big). \tag{10.2.9}$$

This Laplacian condition on $\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s)$ induces differential equations for the coefficients $\mathfrak{C}_N(\tau_2)$ which fixes their dependence on τ_2. Let us consider the constant term first. Since this term is independent of τ_1, it must obey the equation

$$\Delta_{\mathcal{R}}\mathfrak{C}_0(\tau_2) = s(s-1)\mathfrak{C}_0(\tau_2), \tag{10.2.10}$$

where we define the "radial part" part of the Laplacian

$$\Delta_{\mathcal{R}} := \tau_2^2\frac{\partial^2}{\partial\tau_2^2}. \tag{10.2.11}$$

The terminology refers to the fact that this is the Laplacian only along the directions of the Cartan subalgebra of $SL(2,\mathbb{R})$, corresponding roughly to the "radius" of the hyperbolic space $SL(2,\mathbb{R})/SO(2)$. From Equation (10.2.10) we deduce that the constant term is of the form

$$\mathfrak{C}_0(\tau_2) = A(s)\tau_2^s + B(s)\tau_2^{1-s}, \tag{10.2.12}$$

where $A(s)$ and $B(s)$ are unknown functions that remains to be determined. These are fixed by $SL(2,\mathbb{Z})$-invariance of the Fourier expansion, as we shall see explicitly below. Let us now proceed to analyze the remaining Fourier coefficients $\mathfrak{C}_N(\tau_2)$ for $N\neq 0$. Inserting the bulk part of the general Fourier expansion (10.2.6) into the Laplacian equation (10.2.8) yields the following equation for the coefficients:

$$\left[\tau_2^2\frac{\partial^2}{\partial\tau_2^2} - 4\pi^2N^2\tau_2^2 - s(s-1)\right]\mathfrak{C}_N(\tau_2) = 0. \tag{10.2.13}$$

To solve this equation, it is convenient to redefine $\mathfrak{C}_N(\tau_2)\equiv\sqrt{\tau_2}\,\tilde{\mathfrak{C}}_N(\tau_2)$, after which the equation for $\tilde{\mathfrak{C}}_N$ is identified with a Bessel equation, with solution given by the modified Bessel function:

$$\tilde{\mathfrak{C}}_N(\tau_2) = a(N)K_{s-1/2}\big(2\pi|N|\tau_2\big), \tag{10.2.14}$$

where $a(N)$ is some τ_2-independent prefactor. Recalling that $\mathfrak{C}_N(\tau_2)=\sqrt{\tau_2}\tilde{\mathfrak{C}}_N(\tau_2)$, we then find that the Fourier expansion (10.2.6) becomes

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = A(s)\tau_2^s + B(s)\tau_2^{1-s} + \sqrt{\tau_2}\sum_{N\neq 0}a(N)K_{s-1/2}\big(2\pi|N|\tau_2\big)e^{2\pi iN\tau_1}. \tag{10.2.15}$$

The remaining coefficients $A(s), B(s)$ and $a(N)$ contain the missing information required to make this expression $SL(2,\mathbb{Z})$-invariant. To determine these unknowns, we shall now consider the manifestly $SL(2,\mathbb{Z})$-invariant expression (10.2.1) and explicitly compute its Fourier expansion through Poisson resummation techniques. This will cast (10.2.1) into the form (10.2.15), allowing us to deduce the missing pieces.

10.2.2 Explicit Fourier Coefficients from Poisson Resummation

We follow the general method described in [13]. Starting from (10.2.1), the first step is to extract the term with $n=0$ from the sum, which yields

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = 2\zeta(2s)\tau_2^s + \sum_{m\in\mathbb{Z}}\sum_{n\neq 0}\frac{\tau_2^s}{|m+n\tau|^{2s}}, \tag{10.2.16}$$

where the Riemann zeta function is defined as follows

$$\zeta(t) := \sum_{m=1}^{\infty} \frac{1}{m^t}. \tag{10.2.17}$$

By comparing this with (10.2.15) we may already deduce that $A(s) = 2\zeta(2s)$. Note further that since $n \neq 0$ in (10.2.16) the sum over m is unrestricted. To proceed, it is convenient to employ the following integral representation of the summand

$$\frac{1}{|m+n\tau|^{2s}} = \frac{\pi^s}{\Gamma(s)} \int_0^\infty \frac{dt}{t^{s+1}}\ e^{-\frac{\pi}{t}|m+n\tau|^2}, \tag{10.2.18}$$

and we expand the argument of the exponential as follows

$$|m+n\tau|^2 = (m+n\tau)(m+n\bar{\tau}) = (m+n\tau_1)^2 + n^2\tau_2^2. \tag{10.2.19}$$

Inserting this into (10.2.16) then yields

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = 2\zeta(2s)\tau_2^s + \frac{\pi^s\tau_2^s}{\Gamma(s)} \sum_m \sum_{n\neq 0} \int_0^\infty \frac{dt}{t^{s+1}} e^{-\frac{\pi}{t}\left[(m+n\tau_1)^2 - n^2\tau_2^2\right]}. \tag{10.2.20}$$

Since the sum over m is unrestricted we may perform the following Poisson resummation:

$$\sum_m e^{-\frac{\pi}{t}(m+n\tau_1)^2} = \sqrt{t} \sum_{\tilde{m}} e^{-\pi t\tilde{m}^2 - 2\pi i\tilde{m}n\tau_1}. \tag{10.2.21}$$

Inserting this into (10.2.20) ensures that the term involving τ_1 is independent of the integration variable t, and we obtain

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = 2\zeta(2s)\tau_2^s + \frac{\pi^s\tau_2^s}{\Gamma(s)} \sum_{\tilde{m}} \sum_{n\neq 0} e^{-2\pi i\tilde{m}n\tau_1} \int_0^\infty \frac{dt}{t^{s+1/2}} e^{-\pi t\tilde{m}^2 - \frac{\pi}{t}n^2\tau_2^2}. \tag{10.2.22}$$

We now extract the $\tilde{m} = 0$ part of the sum, which allows us to determine the coefficient $B(s)$ in (10.2.15),

$$\begin{aligned}\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) &= 2\zeta(2s)\tau_2^s + 2\sqrt{\pi}\frac{\Gamma(s-\frac{1}{2})}{\Gamma(s)}\zeta(2s-1)\tau_2^{1-s} \\ &+ \frac{\pi^s\tau_2^s}{\Gamma(s)} \sum_{\tilde{m}\neq 0} \sum_{n\neq 0} e^{-2\pi i\tilde{m}n\tau_1} \int_0^\infty \frac{dt}{t^{s+1/2}} e^{-\pi t\tilde{m}^2 - \frac{\pi}{t}n^2\tau_2^2},\end{aligned} \tag{10.2.23}$$

where we again made use of (10.2.17) and (10.2.18). The remaining integral is of Bessel type, and may be evaluated as follows

$$\int_0^\infty \frac{dt}{t^{s+1/2}} e^{-\pi t\tilde{m}^2 - \frac{\pi}{t}n^2\tau_2^2} = 2\left|\frac{\tilde{m}}{n}\right|^{s-1/2} \tau_2^{1/2-s} K_{s-1/2}(2\pi|\tilde{m}n|\tau_2), \tag{10.2.24}$$

which yields

$$\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) = 2\zeta(2s)\tau_2^s + 2\sqrt{\pi}\frac{\Gamma(s-\frac{1}{2})}{\Gamma(s)}\zeta(2s-1)\tau_2^{1-s}$$

$$+\frac{2\pi^s\sqrt{\tau_2}}{\Gamma(s)}\sum_{\tilde{m}\neq 0}\sum_{n\neq 0}\left|\frac{\tilde{m}}{n}\right|^{s-1/2}K_{s-1/2}(2\pi|\tilde{m}n|\tau_2)e^{-2\pi i\tilde{m}n\tau_1}. \tag{10.2.25}$$

To make contact with (10.2.15) we further change variables in the sum to $N \equiv -\tilde{m}n$. Replacing the sum over $\tilde{m}$ by a sum over $N \in \mathbb{Z}\backslash\{0\}$ puts a restriction on the sum over n which enforces the constraint $\tilde{m} = N/n \in \mathbb{Z}\backslash\{0\}$. Effectively this implies that the sum over n becomes a sum over divisors of N, written as $\sum_{n|N}$. We thus have

$$\sum_{\tilde{m}\neq 0}\sum_{n\neq 0}\left|\frac{\tilde{m}}{n}\right|^{s-1/2} = \sum_{N\neq 0}\sum_{n|N}N^{s-1/2}n^{1-2s} \equiv \sum_{N\neq 0}N^{s-1/2}\mu_{1-2s}(N), \tag{10.2.26}$$

where we defined the quantity

$$\mu_{1-2s}(N) := \sum_{n|N}n^{1-2s}, \tag{10.2.27}$$

which is known as the *instanton measure* [14, 181–183] (see Section 8.2). The final form of the Fourier expansion is therefore

$$\begin{aligned}\mathcal{E}^{SL(2,\mathbb{Z})}(\tau;s) \;=\;& 2\zeta(2s)\tau_2^s + 2\sqrt{\pi}\frac{\Gamma(s-\frac{1}{2})}{\Gamma(s)}\zeta(2s-1)\tau_2^{1-s}\\ &+\frac{2\pi^s\sqrt{\tau_2}}{\Gamma(s)}\sum_{N\neq 0}\mu_{1-2s}(N)N^{s-1/2}K_{s-1/2}(2\pi|N|\tau_2)e^{2\pi iN\tau_1}.\end{aligned} \tag{10.2.28}$$

By comparing this expression with (10.2.15) we finally deduce that the remaining Fourier coefficient $a(N)$ is given by

$$a(N) = \frac{2\pi^s}{\Gamma(s)}N^{s-1/2}\mu_{1-2s}(N). \tag{10.2.29}$$

Let us finally note that we may rewrite this in a slightly more compact way in terms of the so called *completed zeta function*:

$$\xi(s) := \pi^{-s/2}\Gamma(s/2)\zeta(s), \tag{10.2.30}$$

which obeys the nice functional relation

$$\xi(s) = \xi(1-s). \tag{10.2.31}$$

Extracting an overall factor of $2\zeta(2s)$, we may then write the Fourier expansion of the Poincaré series discussed in Section (9.2) in the following way:

$$E^{SL(2,\mathbb{Z})}(\tau;s) \;=\; \sum_{\gamma\in N(\mathbb{Z})\backslash SL(2,\mathbb{Z})}\big(\gamma\cdot\tau_2\big)^s$$

$$= \tau_2^s + \frac{\xi(2s-1)}{\xi(2s)}\tau_2^{1-s} + \frac{\sqrt{\tau_2}}{\xi(2s)}\sum_{N\neq 0}\mu_{1-2s}(N)N^{s-1/2}K_{s-1/2}(2\pi|N|\tau_2)e^{2\pi i N\tau_1}. \tag{10.2.32}$$

Let us finally mention that the functional relation (10.2.31) of the completed zeta function $\xi(s)$ can be extended to the following nice functional relation between Eisenstein series of different order [187]:

$$\xi(s)E^{SL(2,\mathbb{Z})}(\tau;s) = \xi(-s)E^{SL(2,\mathbb{Z})}(\tau;-s). \tag{10.2.33}$$

10.2.3 Spherical Vector Construction of the Fourier Expansion

From the general formula (9.1.25) it is clear that the two important building blocks of any automorphic form are the spherical vector f_K and the p-adic counterpart f_p. For the case of $SL(2,\mathbb{Z})$, we have seen in Section 9.2 that the infinite product over prime numbers in (9.1.25) can be evaluated and gives rise to the correct form of the Eisenstein series $E^{SL(2,\mathbb{Z})}(\tau;s)$. It turns out that also the bulk part of the Fourier expansion of $E^{SL(2,\mathbb{Z})}(\tau;s)$ may be recast into this form, by identifying a new spherical vector corresponding to the Fourier transform of $f_K(x) = (1+x^2)^{-s}$. Unfortunately the constant terms do not follow from this construction, but may in principle be obtained by enforcing invariance under the Weyl group of $SL(2,\mathbb{R})$ [195, 196].

We begin by noting that, as a consequence of the p-adic norm, the sum over primes in (9.1.25) has only support on the integers. Thus, Eq. (9.1.25) can be rewritten as follows

$$E^{SL(2,\mathbb{Z})}(\tau;s) = \sum_{x\in\mathbb{Z}}{}' f_{\mathbb{Z}}(x)\rho(\mathcal{V})\cdot f_K(x), \tag{10.2.34}$$

where $f_{\mathbb{Z}}$ is the $SL(2,\mathbb{Z})$-invariant distribution discussed in Section 9.1.3, which arises after evaluating the product over primes. To identify the individual building blocks in (10.2.34) from the Fourier expansion of $E^{SL(2,\mathbb{Z})}(\tau,s)$ in (10.2.32), we begin by restricting to the origin of the moduli space: $\tau_1 = 0,\ \tau_2 = 1$, in which case (10.2.34) may be written as

$$E^{SL(2,\mathbb{Z})}(1;s) = \sum_{x\in\mathbb{Z}}{}' f_{\mathbb{Z}}(x)f_K(x). \tag{10.2.35}$$

Similarly, at the origin of the moduli space, the bulk piece of the Fourier expansion (10.2.32) becomes

$$E^{SL(2,\mathbb{Z})}(1;s) = \frac{1}{\xi(2s)}\sum_{x\in\mathbb{Z}}{}' \mu_s(x)x^{s-1/2}K_{s-1/2}(2\pi|x|). \tag{10.2.36}$$

Comparing (10.2.36) to (10.2.35) we may extract the new spherical vector:

$$f_K(x) := x^{s-1/2}K_{s-1/2}(2\pi|x|), \tag{10.2.37}$$

together with the arithmetic coefficient

$$f_{\mathbb{Z}}(x) := \frac{1}{\xi(2s)}\mu_s(x), \tag{10.2.38}$$

where $\mu_{1-2s}(x)$ is the sum over divisors defined in 10.2.27. We finally note that (10.2.37) is indeed the Fourier transform of the old spherical vector $f_K(x) = (1+x^2)^{-s}$.

10.3 Fourier Expansion of $E^{SL(3,\mathbb{Z})}(g;s_1,s_2)$

In this section we shall give the complete Fourier expansion of the general Eisenstein series $E^{SL(3,\mathbb{Z})}(g;s_1,s_2)$ in the principal series. This was computed in detail by Vinogradov and Takhtadzhyan in [200], and we will here reproduce their results in a slightly adapted form. The following section is based on **Paper VII**, which is written in collaboration with Boris Pioline. The Eisenstein series $E(g;s_1,s_2)$ is conjectured to capture D(-1), D5 and NS5-brane instanton corrections to the hypermultiplet moduli space in type IIB string theory compactified on a Calabi-Yau threefold X. The physical interpretation of our results is briefly commented on in the text below, while a more detailed account can be found in **Paper VII**.

10.3.1 General Structure

For the readers convenience we begin by recalling the explicit form (9.3.9) of the principal Eisenstein series:

$$\begin{aligned} E^{SL(3,\mathbb{Z})}(g;s_1,s_2) &= y^{\frac{s_2-s_1}{3}} \sum_{(c,b)\in\mathbb{Z}^6}{}' \bigg\{\Big[c_1^2 u+\frac{1}{u}\big(c_2+c_1 v\big)^2+\frac{1}{y}\big(c_1\mathcal{A}_1+c_2\mathcal{A}_2+c_3\big)^2\Big]^{-s_1} \\ &\qquad\times\Big[yD_3^2+u\big(D_2-\mathcal{A}_2 D_3\big)^2+\frac{1}{u}\big(D_1-vD_2-(\mathcal{A}_1-v\mathcal{A}_2)D_3\big)^2\Big]^{-s_2}\bigg\}, \end{aligned} \tag{10.3.1}$$

where

$$D_1 := b_2c_3-b_3c_2, \qquad D_2 := b_3c_1-b_1c_3, \qquad D_3 := b_1c_2-b_2c_1. \tag{10.3.2}$$

The general structure of the Fourier expansion is given by [200]

$$E^{SL(3,\mathbb{Z})}(g;s_1,s_2) = \sum_{(m_1,m_2,m_3)\in\mathbb{Z}^3} \mathfrak{C}_{m_1,m_2,m_3}(y,u_{1,2};s_1,s_2)e^{2\pi i(m_1\mathcal{A}_1+m_2\mathcal{A}_2+m_3 v_{1,2})}, \tag{10.3.3}$$

where $v_{1,2}\equiv v_{m_1,m_2}$ and $u_{1,2}\equiv u_{m_1,m_2}$ denote the real and imaginary parts of the image of the complex scalar $z=v+iu$ under an $SL(2,\mathbb{Z})$ subgroup of $SL(3,\mathbb{Z})$ acting on z as follows:

$$\delta\cdot z = \frac{\alpha z+\beta}{m_1' z+m_2'} \equiv v_{m_1',m_2'}+iu_{m_1',m_2'}, \tag{10.3.4}$$

where $m_1'=m_1/d$ and $m_2'=m_2/d$ with $d=\gcd(m_1,m_2)$, and α,β are two integers such that $\alpha m_2'-\beta m_1'=1$. When $m_1'\neq 0$ one may rewrite this in the following useful form

$$v_{m_1',m_2'} = \frac{\alpha}{m_1'}-\frac{m_1' v+m_2'}{m_1'|m_1' z+m_2'|^2}\,, \qquad u_{m_1',m_2'} = \frac{u}{|m_1' z+m_2'|^2}\,, \tag{10.3.5}$$

where we used $\beta=(m_2'\alpha-1)/m_1'$ to derive the first relation. In terms of the group element $g\in SL(3,\mathbb{R})$ (see Eq. (9.3.2)) this $SL(2,\mathbb{Z})$ corresponds to the left action on g by the $SL(3,\mathbb{Z})$ element

$$\gamma = \begin{pmatrix} \alpha & \beta & 0 \\ m_1' & m_2' & 0 \\ 0 & 0 & 1 \end{pmatrix} \in SL(3,\mathbb{Z}), \tag{10.3.6}$$

under which y is invariant while $(\mathcal{A}_1, \mathcal{A}_2)$ transform linearly. The appearance of the somewhat complicated expressions for $v_{1,2}$ and $u_{1,2}$ in the Fourier expansion is a consequence of the non-abelian structure of the Heisenberg group N. In order to make contact with the conventions of **Paper VII** we make the following minor change of moduli variables:

$$y = \nu^{-1}\ , \qquad u = \tau_2\ , \qquad \mathcal{A}_1 = -\psi\ , \qquad \mathcal{A}_2 = -c_0\ , \qquad v = -\tau_1\ , \tag{10.3.7}$$

and for the summation variables we define

$$k = -m_1, \qquad p = m_2, \qquad q = m_3. \tag{10.3.8}$$

The nilpotent group $N \subset SL(3,\mathbb{R})$ is isomorphic to the three-dimensional Heisenberg group, and is therefore non-abelian. From (9.3.2) and the change of variables (10.3.7) we deduce that ψ parametrizes the central generator of this Heisenberg group. As explained in Section 10.1, the Fourier expansion then splits into an abelian part and a non-abelian part,

$$E^{SL(3,\mathbb{Z})}(g; s_1, s_2) = E_A(g; s_1, s_2) + E_{NA}(g; s_1, s_2), \tag{10.3.9}$$

where

$$\begin{aligned} E_A(g; s_1, s_2) &= \sum_{(p,q)\in\mathbb{Z}^2} \Psi_{p,q}(\nu, \tau_2)\, e^{2\pi i(pc_0 - q\tau_1)}, \\ E_{NA}(g; s_1, s_2) &= \sum_{k\neq 0} \sum_{(p,q)\in\mathbb{Z}^2} \Psi_{q,(k,p)}(\nu, [\tau_2]_{-k,p})\, e^{-2\pi i\left(q[\tau_1]_{-k,p} - pc_0 - k\psi\right)}, \end{aligned} \tag{10.3.10}$$

where $(k,p) = \gcd(k,p)$. Recall from Section 10.1 that the abelian term represents the Fourier expansion with respect to the abelianized group $\tilde{N} = N/Z$, where Z is the center of N, corresponding to the commutator subgroup:

$$Z = [N, N] = \left\{ \begin{pmatrix} 1 & & * \\ & 1 & \\ & & 1 \end{pmatrix} \right\}. \tag{10.3.11}$$

The non-abelian term is then the Fourier expansion with respect to the center Z. Let us begin to consider the zeroth Fourier coefficient $\Psi_{0,0}$. This is given by the constant term

$$\Psi_{0,0}(\nu, \tau_2; s_1, s_2) = \int_0^1 d\tau_1 \int_0^1 dc_0 \int_0^1 d\psi\ E(\nu, \tau_2, \tau_2, c_0, \psi; s_1, s_2), \tag{10.3.12}$$

which was evaluated in [200] with the result

$$\begin{aligned} \Psi_{0,0}(\nu, \tau_2; s_1, s_2) &= \nu^{-\frac{2s_1+s_2}{3}} \tau_2^{s_2} + c(s_1), \nu^{-\frac{1}{2}+\frac{s_1-s_2}{3}} \tau_2^{s_3} + c(s_2)\nu^{-\frac{2s_1+s_2}{3}} \tau_2^{1-s_2} \\ &\quad + c(s_1)\, c(s_3)\, \nu^{-\frac{1}{2}+\frac{s_1-s_2}{3}} \tau_2^{1-s_3} + c(s_2)\, c(s_3)\, \nu^{\frac{2s_1+s_2}{3}-1} \tau_2^{s_1} \\ &\quad + c(s_1)\, c(s_2)\, c(s_3)\, \nu^{\frac{2s_1+s_2}{3}-1} \tau_2^{1-s_1}\ , \end{aligned} \tag{10.3.13}$$

where $s_3 = s_1 + s_2 - 1/2$ and we defined

$$c(s) := \frac{\xi(2s-1)}{\xi(2s)}. \tag{10.3.14}$$

Recall also that $\xi(s)$ is the completed Riemann zeta function (10.2.30).

10.3.2 Abelian Fourier Coefficients

Let us now proceed to analyze the Abelian Fourier coefficients $\Psi_{p,q}$ in (10.3.10). Here it is convenient to begin by studying the simplest cases $\Psi_{0,q}$ and $\Psi_{p,0}$. As we shall see, the general abelian coefficient is considerably more involved. We find:

$$\begin{aligned}\Psi_{0,q}(\nu,\tau_2) = &\frac{2(2\pi)^{1/2-s_1}c(s_2)c(s_3)}{\xi(2s_1)}\nu^{\frac{s_1+2s_2}{3}-1}\tau_2^{1-s_1}\,\mu_{1-2s_1}(|q|)\,\mathcal{K}_{1/2-s_1}\big(2\pi|q|\tau_2\big)\\ &+\frac{2(2\pi)^{1/2-s_2}}{\xi(2s_2)}\nu^{-\frac{2s_1+s_2}{3}}\tau_2^{1-s_2}\,\mu_{1-2s_2}(|q|)\,\mathcal{K}_{1/2-s_2}\big(2\pi|q|\tau_2\big)\\ &+\frac{2(2\pi)^{1/2-s_3}c(s_1)}{\xi(2s_3)}\nu^{\frac{s_1-s_2-1}{2}}\tau_2^{1-s_3}\,\mu_{1-2s_3}(|q|)\,\mathcal{K}_{1/2-s_3}\big(2\pi|q|\tau_2\big)\ ,\end{aligned} \tag{10.3.15}$$

and

$$\begin{aligned}\Psi_{p,0}(\nu,\tau_2) = &\frac{2(2\pi)^{1/2-s_1}}{\xi(2s_1)}\nu^{-\frac{2s_1+s_2}{3}}\tau_2^{s_2}\mu_{1-2s_1}(|p|)\,\mathcal{K}_{1/2-s_1}\big(2\pi|p|/\sqrt{\nu\tau_2}\big)\\ &+\frac{2(2\pi)^{1/2-s_2}c(s_1)c(s_3)}{\xi(2s_2)}\nu^{\frac{s_1-s_2}{3}-\frac{1}{2}}\tau_2^{\frac{3}{2}-s_1-s_2}\,\mu_{1-2s_2}(|p|)\,\mathcal{K}_{1/2-s_2}\big(2\pi|p|/\sqrt{\nu\tau_2}\big)\\ &+\frac{2(2\pi)^{1/2-s_3}c(s_2)}{\xi(2s_3)}\nu^{-\frac{s_2+2s_1}{3}}\tau_2^{1-s_2}\,\mu_{1-2s_3}(|p|)\,\mathcal{K}_{1/2-s_3}\big(2\pi|p|/\sqrt{\nu\tau_2}\big)\ ,\end{aligned} \tag{10.3.16}$$

where we have defined the rescaled modified Bessel function

$$\mathcal{K}_s(x) := x^{-s}K_s(x), \tag{10.3.17}$$

and we defined the sum over divisors,

$$\mu_s(r) := \sum_{n|r} n^s. \tag{10.3.18}$$

Following the general logic of Chapter 8, we expect that these Fourier coefficients contain non-perturbative information. As mentioned above, in the context of **Paper VII** this non-perturbative information is due to D(−1), D5 and NS5-brane instantons. Recall further from Chapter 8 that this information is extracted by studying the Fourier coefficients in the semiclassical limit where the argument of the Bessel functions are large. To this end we use the expansion of the (rescaled) Bessel function given by

$$\mathcal{K}_s(x) \sim \sqrt{\frac{\pi}{2}}x^{-(s+1/2)}e^{-x}\Big[1+\mathcal{O}(1/x)\Big]\,. \tag{10.3.19}$$

Using this expansion in (10.3.15) we deduce that in the weak-coupling limit $\tau_2\to\infty$, the coefficients $\Psi_{0,q}$ contribute to the expansion (10.3.10) by exponentially suppressed contributions of the form

$$\sum_{q\neq 0}\Psi_{0,q}(\nu,\tau_2)\,e^{-2\pi i q\tau_1} \sim \sum_{q\neq 0}\mu(0,q)e^{-2\pi S_{0,q}(\tau)}\Big[1+\mathcal{O}(1/\tau_2)+\cdots\Big], \tag{10.3.20}$$

where we defined

$$S_{0,q}(\tau) = |q|\,\tau_2 + i\,q\,\tau_1. \tag{10.3.21}$$

This is precisely the instanton action for D(−1) instantons [14,180]. The remaining coefficient $\mu(0,q)$ corresponds to the instanton measure and is given by

$$\begin{aligned}\mu(0,q) \;\; &:= \;\; \frac{1}{\sqrt{\pi}}|q|^{-1}\Bigg[\frac{c(s_2)c(s_3)}{\xi(2s_1)}|q|^{s_1}\mu_{1-2s_1}(|q|)\\ &+\frac{1}{\xi(2s_2)}|q|^{s_2}\mu_{1-2s_2}(|q|)+\frac{c(s_1)}{\xi(2s_3)}|q|^{s_3}\mu_{1-2s_3}(|q|)\Bigg].\end{aligned} \tag{10.3.22}$$

Repeating this analysis for the coefficients $\Psi_{p,0}$ we find again exponentially suppressed contributions of order $e^{-2\pi S_{p,0}}$ where

$$S_{p,0}(\nu,\tau_2,c_0) = |p|(\nu\tau_2)^{-1/2} - i\,p\,c_0 = |p|\,\tau_2 V \; - i\,p\,c_0 \;, \tag{10.3.23}$$

which is then the instanton action for p D5-branes wrapping the internal manifold X. In the second step we have rewritten the instanton action in terms of the volume of X (see **Paper VII** for details). From the semiclassical expansion we may also extract the D5-instanton measure:

$$\begin{aligned}\mu(p,0) \;\; &:= \;\; \frac{1}{\sqrt{\pi}}|p|^{-1}\Bigg[\frac{1}{\xi(2s_1)}|p|^{s_1}\mu_{1-2s_1}(|p|)\\ &+\frac{c(s_1)c(s_3)}{\xi(2s_2)}|p|^{s_2}\mu_{1-2s_2}(|p|)+\frac{c(s_2)}{\xi(2s_3)}|p|^{s_3}\mu_{1-2s_3}(|p|)\Bigg].\end{aligned} \tag{10.3.24}$$

We now proceed to analyze the coefficients $\Psi_{p,q}$ for $pq \neq 0$. These coefficients may be written as

$$\begin{aligned}\Psi_{p,q}(\nu,\tau_2) = \frac{4\nu^{\frac{s_2-s_1}{6}-\frac{1}{2}}\tau_2^{\frac{s_2-s_1}{2}+\frac{1}{2}}}{\xi(2s_1)\xi(2s_2)\xi(2s_3)}\sum_{d_1|p}\sum_{d_2|\frac{p}{d_1}} d_1^{1-2s_3}d_2^{1-2s_2}\\ \times\,\sigma_{1-2s_1,1-2s_3}\Big(\frac{p}{d_1d_2},|q|\Big)(pq)^{s_3-\frac{1}{2}}\,\mathcal{I}\left(R_{p,q},\mathbf{x}_{p,q}\right)\;,\end{aligned} \tag{10.3.25}$$

where

$$R_{p,q} \equiv \left(\frac{p^2q}{\nu}\right)^{1/3}, \qquad \mathbf{x}_{p,q} \equiv \tau_2^{-1}\left(\frac{p^2}{\nu q^2}\right)^{1/3}. \tag{10.3.26}$$

Here, we have also defined the "double divisor sum"

$$\sigma_{\alpha\beta}(n,m) := \sum_{\substack{m=d_1d_2d_3,\\ d_1,d_2,d_3>0,(d_3,n)=1}} d_2^{\alpha}d_3^{\beta}\;, \tag{10.3.27}$$

and the integral

$$\mathcal{I}(R,\mathbf{x}) := \int_0^\infty K_{s_3-1/2}\big(2\pi R\,\mathbf{x}^{-1}\sqrt{1+x}\big)\,K_{s_3-1/2}\Big(2\pi R\,\mathbf{x}^{1/2}\sqrt{1+1/x}\Big)\,x^{\frac{s_2-s_1}{2}}\frac{dx}{x}\;. \tag{10.3.28}$$

We see that these coefficients cannot be written simply in terms of a single Bessel function, but rather involves a rather complicated integral over two Bessel functions. This reflects the non-abelian nature of the Heisenberg group N. To deduce the instanton action for these combined D(−1)-D5 contributions, we must find the semi-classical expansion $\tau_2 \to \infty$ of the integral $\mathcal{I}(R, \mathbf{x})$.

To this end we shall now evaluate the saddle point approximation of (10.3.28), corresponding to the limit $R \to \infty$, while keeping $\mathbf{x}$ fixed. We begin by expanding the Bessel functions for large argument as above

$$K_s(x) \sim \sqrt{\frac{\pi}{2x}} e^{-x} \left[1 + \frac{4s^2 - 1}{8x} + \mathcal{O}(1/x^2) \right] \ . \tag{10.3.29}$$

To leading order, the integral then simplifies to

$$\mathcal{I}(R, \mathbf{x}) \sim \frac{\mathbf{x}^{1/4}}{4R} \int_0^\infty \frac{x^{\frac{s_2 - s_1}{2} - \frac{3}{4}} \ e^{-2\pi S(x)}}{\sqrt{1+x}} dx \ , \tag{10.3.30}$$

where $S(x)$ is given by

$$S(x) = R \left(\mathbf{x}^{-1} \sqrt{1+x} + \mathbf{x}^{1/2} \sqrt{1 + \frac{1}{x}} \right) \ . \tag{10.3.31}$$

This is extremized at $x = \mathbf{x}$, with

$$S(\mathbf{x}) = \frac{R(1+\mathbf{x})^{3/2}}{\mathbf{x}} \ , \qquad \partial_x^2 S(\mathbf{x}) = \frac{3R}{4\mathbf{x}^2 \sqrt{1+\mathbf{x}}} \ . \tag{10.3.32}$$

The saddle point approximation is then

$$\mathcal{I}(R, \mathbf{x}) \sim \frac{\mathbf{x}^{\frac{s_2 - s_1 - 1}{2}}}{4R\sqrt{1+\mathbf{x}}} \left[\frac{1}{2} \partial_x^2 S(\mathbf{x}) \right]^{-1/2} \exp\left(-2\pi S(\mathbf{x})\right) \left[1 + \mathcal{O}(1/R) \right], \tag{10.3.33}$$

which may be evaluated to

$$\mathcal{I}(R, \mathbf{x}) \sim \frac{\mathbf{x}^{\frac{s_2 - s_1 + 1}{2}}}{\sqrt{6}\, R^{3/2} (1+\mathbf{x})^{1/4}} \exp\left[-\frac{2\pi R (1+\mathbf{x})^{3/2}}{\mathbf{x}} \right] \left[1 + \mathcal{O}(1/R) \right] \ . \tag{10.3.34}$$

Plugging in the values of $R_{p,q}$ and $\mathbf{x}_{p,q}$ given in (10.3.26), we find that such terms give contributions of order $e^{-2\pi S_{p,q}}$ where

$$S_{p,q}(\nu, \tau, c_0) = \left[p^{2/3} (\tau_2 V)^{2/3} + q^{2/3} \tau_2^{2/3} \right]^{3/2} + i(q\tau_1 - p c_0) \ . \tag{10.3.35}$$

We note that the real part of this action is proportional to the mass formula for bound states of D0-D6-branes found in [203–205]. Moreover, in the limit $q = 0$ or $p = 0$, (10.3.35) reduces to (10.3.21) and (10.3.23). We conclude that general Abelian terms with $pq \neq 0$ correspond to bound states of D(-1) and D5-brane instantons. The general abelian summation measure may also be extracted with the result

$$\mu(p, q) := \frac{4}{\xi(2s_1)\xi(2s_2)\xi(2s_3)} \sum_{d_1 | p} \sum_{d_2 | \frac{p}{d_1}} d^{2 - s_1 - s_2} d_1^{1 - 2s_3} d_2^{1 - 2s_2} \sigma_{1-2s_1, 1-2s_3}\left(\frac{p}{d_1 d_2}, |q| \right). \tag{10.3.36}$$

10.3.3 Non-Abelian Fourier Coefficients

We now proceed to the non-abelian terms in (10.3.10) corresponding to $k \neq 0$. These may be written as

$$\Psi_{q,(k,p)} = \frac{4\nu^{\frac{s_2-s_1}{6}-\frac{1}{2}}}{\xi(2s_1)\xi(2s_2)\xi(2s_3)} \sum_{q\in\mathbb{Z}} \sum_{d_1|d} \sum_{d_2|\frac{d}{d_1}} d_1^{1-2s_3} d_2^{1-2s_2} \sigma_{1-2s_1,1-2s_3}\Big(\frac{d}{d_1 d_2}, |q|\Big) \\ \times [\tau_2]_{-k,p}^{\frac{s_2-s_1}{2}+\frac{1}{2}} (d\,q)^{s_3-\frac{1}{2}} \mathcal{I}\left(R_{p,q,k}, \mathbf{x}_{p,q,k}\right) \; e^{-2\pi i q[\tau_1]_{-k,p}}, \tag{10.3.37}$$

where

$$R_{p,q,k} \equiv \left(\frac{d^2|q|}{\nu}\right)^{1/3}, \qquad \mathbf{x}_{p,q,k} \equiv [\tau_2]_{-k,p}^{-1} \left(\frac{d^2}{\nu q^2}\right)^{1/3}, \tag{10.3.38}$$

and $d \equiv \gcd(p,k) > 0$ and the integral $\mathcal{I}(R,\mathbf{x})$ is the same as in (10.3.28).

As before, using the saddle point approximation of the integral $\mathcal{I}(R,\mathbf{x})$, we find that $\Psi_{q,(p,k)}$ contributes to (10.3.10) with terms which are exponentially suppressed by $e^{-2\pi S_{Q,p,k}}$, where

$$S_{Q,p,k} = \frac{\left[(V\tau_2)^{2/3}|p-k\tau|^2 + \tau_2^{2/3}Q^{2/3}\right]^{3/2}}{|p-k\tau|^2} - i(p\,c_0 + k\,\psi) + i\,Q\frac{p-k\tau_1}{k|p-k\tau|^2} \tag{10.3.39}$$

and we denoted $Q \equiv d^2 q$ in comparison to (10.3.37). Note that the last term proportional to Q originates from the non-abelian coefficient $\Psi_{q,(p,k)}$ (10.3.37), and is in addition to the overall phase factors already appearing in the Fourier expansion (10.3.10). From the saddle point approximation, we may also extract the summation measure

$$\mu(p,q,k) \equiv \frac{4\,e^{\frac{2\pi i d q \alpha}{k}}}{\xi(2s_1)\xi(2s_2)\xi(2s_3)} \sum_{d_1|d} \sum_{d_2|\frac{d}{d_1}} d^{2-s_1-s_2} d_1^{1-2s_3} d_2^{1-2s_2} \sigma_{1-2s_1,1-2s_3}\Big(\frac{d}{d_1 d_2}, |q|\Big)\,. \tag{10.3.40}$$

For more details on this Fourier expansion, and in particular its physical interpretation, we refer the reader to **Paper VII**.

10.3.4 The Exact Spherical Vector in the Principal Series

As was done in Section 10.2.3 for the Fourier expansion of $E^{SL(2,\mathbb{Z})}(\tau;s)$, we can now cast the non-abelian Fourier expansion of $E^{SL(3,\mathbb{Z})}(g;s_1,s_2)$ into the general framework of Section 9.1.3. This will allow us to extract the exact spherical vector f_K for the full princial series of $SL(3,\mathbb{R})$, whose semi-classical limit was obtained previously in [201]. As above, we follow **Paper VII** closely.

To make the link with the analysis in [201], we begin by making the following change of variables

$$y = -k\,, \qquad x_0 = p\,, \qquad x_1 = (d^2 q)^{1/3}\,, \tag{10.3.41}$$

and, as in Section 10.2.3, we work at the origin of moduli space where $\tau_1 = c_0 = \psi = 0$ and $\tau_2 = \nu = 1$, such that

$$[\tau_2]_{1,2}^{(0)} = \frac{d^2}{y^2+x_0^2}\,, \qquad R = x_1\,, \qquad \mathbf{x} = \frac{y^2+x_0^2}{x_1^2}\,, \tag{10.3.42}$$

where we have used (10.3.5) and (10.3.7). Moreover, at the origin of moduli space, the phase factor in the non-abelian term of (10.3.10) becomes

$$\exp\left[-2\pi iq[\tau_1]^{(0)}_{1,2}\right]=\exp\left[\frac{2\pi iqd\alpha}{k}\right]\,\exp\left[\frac{2\pi i\;pq}{k(k^2+p^2)}\right]=\exp\left[\frac{2\pi ix_1^3\alpha}{dk}\right]\,\exp\left[-\frac{2\pi ix_0x_1^3}{y(y^2+x_0^2)}\right].\tag{10.3.43}$$

Following the reasoning of Section 10.2.3, we may further write the non-abelian part of the expansion at the origin of moduli space as

$$E_{NA}(1;s_1,s_2)=\sum_{(y,x_0,x_1)\in\mathbb{Z}}{}^{\prime}\, f_{\mathbb{Z}}(y,x_0,x_1)\,f_K(y,x_0,x_1).\tag{10.3.44}$$

Comparing this expression with (10.3.37), we deduce that the exact spherical vector is given by

$$\begin{aligned}f_K(y,x_0,x_1)=\;&(y^2+x_0^2)^{\frac{1}{2}(s_1-s_2-1)}\,x_1^{3(s_1+s_2-1)}\,e^{-\frac{2\pi ix_0x_1^3}{y(y^2+x_0^2)}}\\ &\times\int_0^\infty K_{s_3-\frac{1}{2}}\Big(\frac{2\pi\,x_1^3}{y^2+x_0^2}\sqrt{1+x}\Big)K_{s_3-\frac{1}{2}}\Big(2\pi\sqrt{(y^2+x_0^2)(1+1/x)}\Big)x^{\frac{s_2-s_1}{2}}\frac{dx}{x}\,,\end{aligned}\tag{10.3.45}$$

while the “summation measure” is

$$f_{\mathbb{Z}}(y,x_0,x_1)\equiv\frac{4\,e^{\frac{2\pi ix_1^3\alpha}{dk}}}{\xi(2s_1)\xi(2s_2)\xi(2s_3)}\sum_{d_1|d}\sum_{d_2|\frac{d}{d_1}}d^{2-s_1-s_2}d_1^{1-2s_3}d_2^{1-2s_2}\sigma_{1-2s_1,1-2s_3}\Big(\frac{d}{d_1d_2},\frac{x_1^3}{d^2}\Big)\,,\tag{10.3.46}$$

where we recall that $d=\gcd(y,x_0)$.

The spherical vector simplifies considerably in the limit where y,x_0,x_1 are scaled to infinity with fixed ratio: in this case the saddle point approximation (10.3.33) becomes

$$\mathcal{I}(y,x_0,x_1)\sim\frac{(y^2+x_0^2)^{\frac{s_2-s_1+1}{2}}x_1^{s_1-s_2-2}}{(y^2+x_0^2+x_1^2)^{1/4}}\exp\left[-\frac{2\pi(y^2+x_0^2+x_1^2)^{3/2}}{y^2+x_0^2}\right]\,,\tag{10.3.47}$$

and the spherical vector simplifies to

$$f_K(y,x_0,x_1)\sim\frac{x_1^{4s_1+2s_2-5}}{(y^2+x_0^2+x_1^2)^{1/4}}\exp\left[-\frac{2\pi(y^2+x_0^2+x_1^2)^{3/2}}{y^2+x_0^2}-\frac{2\pi ix_0x_1^3}{y(y^2+x_0^2)}\right]\,.\tag{10.3.48}$$

As a consistency check, we note that in the special case $(s_1,s_2)=(0,0)$[2], this result agrees with the semi-classical spherical vector of the principal series representation of $SL(3,\mathbb{R})$ obtained in [201]. We finally note that upon restoring the moduli dependence in the semi-classical spherical vector (10.3.48) by acting with $\rho(g)$, we obtain exponentially suppressed terms with instanton action (10.3.39) as expected. For more information, see **Paper VII**.

[2]In the conventions of [201] this corresponds to$(\lambda_{23},\lambda_{21})=(1,1)$.

11

Compactification and the Automorphic Lift

The purpose of this chapter is to analyze the fate the continuous "U-duality" groups $G_3(\mathbb{R})$, appearing in toroidal compactifications of supergravity theories when higher derivative interactions are turned on. Recently, several authors [206–210] have considered this problem from various perspectives, with the consensus that toroidal compactifications of quadratic and higher order corrections give rise to terms in $D = 3$ which are not G_3-invariant.[1] We propose that the result of the compactification should then be viewed as the leading term in a large-volume expansion of some $G_3(\mathbb{Z})$-invariant expression, where $G_3(\mathbb{Z}) \subset G_3(\mathbb{R})$ is the underlying discrete duality group which is expected to be preserved in the full quantum theory [8]. This completion under $G_3(\mathbb{Z})$ is performed through the introduction of $G_3(\mathbb{Z})$-invariant automorphic forms, and is generally referred to as the *automorphic lift*. The analysis suggests that generalized versions of Eisenstein series which transform in some representation of the maximal compact subgroup $K(G_3)$ are required.

This chapter is based on **Paper IV**, written in collaboration with Ling Bao, Johan Bielecki, Martin Cederwall and Bengt E. W. Nilsson.

11.1 Automorphic Lift - General Philosophy

Keeping in mind the discussion from Chapter 8 regarding the $\mathcal{R}^4$-correction in type IIB supergravity, a natural interpretation of the compactified higher-derivative Lagrangian presents itself: the result of the compactification should correspond to the leading contribution in a large volume expansion of some manifestly $G_3(\mathbb{Z})$-invariant higher-derivative effective action. Inspired by the philosophy of Chapter 8, we would therefore like to "complete" the Lagrangian in $D = 3$ using a certain $G_3(\mathbb{Z})$-invariant automorphic form, whose Fourier expansion repro-

[1] One exception being ref. [209] in which the authors considered quadratic curvature corrections to pure gravity in four dimensions. In that special case, the most general correction can be related, through suitable field redefinitions, to the Gauss-Bonnet term which is topological in four dimensions and does not contribute to the dynamics. Hence, the $SL(2, \mathbb{R})$-symmetry of the compactified Lagrangian is trivially preserved.

duces the result of the compactification to leading order. Let us now consider this appealing idea in some more detail.

In the in the analysis of Chapter 8 the completion to an S-duality invariant $\mathcal{R}^4$-term was achieved through the use of an Eisenstein series which was completely $SL(2,\mathbb{Z})$-invariant. More generally, one might find terms in the effective action whose non-perturbative completion requires automorphic forms transforming under the maximal compact subgroup $K(G_3)$. For example, this was found to be the case in [211], where interaction terms of sixteen fermions were analyzed. These terms transform under the maximal compact subgroup $U(1) \subset SL(2,\mathbb{R})$ and so the U-duality invariant completion requires in this case an automorphic form which transforms with a $U(1)$ weight that compensates for the transformation of the fermionic term, and thus renders the effective action invariant.

The need for automorphic forms which transform under the maximal compact subgroup $K(G_3)$ was also emphasized in [207], based on the observation that the dilaton exponents in compactified higher curvature corrections correspond to weights of the global symmetry group G_3, implying that these terms transform non-trivially in some representation of $K(G_3)$. An explicit realisation of these arguments was found in [210] for the case of compactification on S^1 of the four-dimensional coupled Einstein-Liouville system, supplemented by a four-derivative curvature correction. The resulting effective action was shown to explicitly break the Ehlers $SL(2,\mathbb{R})$-symmetry; however, an $SL(2,\mathbb{Z})_{\text{global}} \times U(1)_{\text{local}}$-invariant effective action was obtained by "lifting" the scalar coefficients to automorphic forms transforming with compensating $U(1)$ weights. The non-perturbative completion implied by this lifting is in this case attributed to gravitational Taub-NUT instantons [210].

Similar conclusions were drawn in [208], in which compactifications of derivative corrections of second, third and fourth powers of the Riemann tensor were analyzed. Again, it was concluded that the G_3-symmetry is explicitly broken by the correction terms. It was argued, in accordance with the type IIB analysis discussed above, that the result of the compactification – being inherently perturbative in nature – should be considered as the large volume expansion of a $G_3(\mathbb{Z})$-invariant effective action. It was shown on general grounds that any term resulting from such a compactification can always be lifted to a U-duality invariant expression through the use of automorphic forms transforming in some representation of $K(G_3)$.

In chapter we extend some aspects of the analysis of [208]. In [208] only parts of the compactification of the Riemann tensor squared, $\hat{R}_{ABCD}\hat{R}^{ABCD}$, were presented. The terms which were analyzed were sufficient to show that the continuous symmetry was broken, and to argue for the necessity of introducing transforming automorphic forms to restore the U-duality symmetry $G_3(\mathbb{Z})$. Moreover, the overall volume factor of the internal torus was neglected in the analysis.

Here we shall restrict our attention to corrections quadratic in the Riemann tensor in order for a complete compactification to be a feasible task. More precisely, we shall focus on a four-derivative correction to the Einstein-Hilbert action in the form of the Gauss-Bonnet term $\hat{R}_{ABCD}\hat{R}^{ABCD} - 4\hat{R}_{AB}\hat{R}^{AB} + \hat{R}^2$. Modulo field equations, this is the only independent invariant quadratic in the Riemann tensor. We extend the investigations of [208] by giving the complete compactification on T^n of the Gauss-Bonnet term from D dimensions to $D-n$ dimensions. In the special case of compactifications to $D-n=3$ dimensions the resulting expression simplifies, making it amenable for a more careful analysis. In particular, one of the main points of this chapter is to study the full structure of the dilaton exponents, with the purpose of determining the $\mathfrak{sl}(n+1,\mathbb{R})$-representation structure associated with quadratic curvature corrections.

11.2 Compactification of the Gauss-Bonnet Term

In this section we outline the derivation of the toroidal compactification of the Gauss-Bonnet term from D dimensions to $D-n$ dimensions. The full result for the compactification to arbitrary dimensions is given in **Paper IV**. Here we focus on the special case of $D-n=3$, which is the most relevant case for the questions we pursue.

11.2.1 The General Procedure

The Gauss-Bonnet Lagrangian density is quadratic in the Riemann tensor and takes the explicit form

$$\mathcal{L}_{GB} = \hat{e}\big[\hat{R}_{ABCD}\hat{R}^{ABCD} - 4\hat{R}_{AB}\hat{R}^{AB} + \hat{R}^2\big]. \tag{11.2.1}$$

The compactification of the D-dimensional Riemann tensor $\hat{R}^{A}_{BCD}$ on an n-torus, T^n, is done in three steps: first we perform a Weyl-rescaling of the internal vielbein, followed by a splitting of the external and internal indices, and finally we perform yet another rescaling of the vielbein. In the following we shall always assume that the torsion vanishes.

Our index conventions are as follows. $M, N, \ldots$ denote D dimensional curved indices, and $A, B, \ldots$ denote D dimensional flat indices. Upon compactification we split the indices according to $M = (\mu, m)$, where $\mu, \nu, \ldots$ and $m, n, \ldots$ are curved external and internal indices, respectively. Similarly, the flat indices split into external and internal parts according to $A = (\alpha, a)$.

Our reduction Ansatz for the vielbein is

$$\hat{e}_M{}^A = e^{\varphi}\tilde{e}_M{}^A = e^{\varphi}\begin{pmatrix} e_\mu{}^\alpha & \mathcal{A}^m_\mu\tilde{e}_m{}^a \\ 0 & \tilde{e}_m{}^a \end{pmatrix}, \tag{11.2.2}$$

where the internal vielbein $\tilde{e}_m{}^a$ is an element of the isometry group $GL(n,\mathbb{R})$ of the n-torus. Later on we shall parametrise $\tilde{e}_m{}^a$ in various ways. With this Ansatz, the line element becomes

$$ds^2_D = e^{2\varphi}\Big\{ds^2_{D-n} + \big[(dx^m + \mathcal{A}^m_{(1)})\tilde{e}_m{}^a\big]^2\Big\}. \tag{11.2.3}$$

In order to obtain a Lagrangian in Einstein frame after dimensional reduction, we perform a Weyl-rescaling of the D-dimensional vielbein,

$$\hat{e}_M{}^A \longrightarrow \tilde{e}_M{}^A = e^{-\varphi}\hat{e}_M{}^A. \tag{11.2.4}$$

Note that all D-dimensional objects before rescaling are denoted $\hat{X}$, the Weyl-rescaled objects are denoted $\tilde{X}$, while the $d=(D-n)$-dimensional objects are written without any diacritics. After the Weyl-rescaling the Gauss-Bonnet Lagrangian, including the volume measure $\hat{e} = e^{D\varphi}\tilde{e}$, can be conveniently organized in terms of equations of motion and total derivatives. This is achieved using integration by parts, where $\tilde{\nabla}_{(A}\tilde{\partial}_{B)}\varphi$ does not appear explicitly. The resulting Lagrangian is (see the appendix of **Paper IV** for the derivation):

$$\begin{aligned}\mathcal{L}_{\rm GB} &= \tilde{e}e^{(D-4)\varphi}\Big\{\tilde{R}^2_{\rm GB} - (D-3)(D-4)\big[2(D-2)(\tilde{\partial}\varphi)^2\tilde{\Box}\varphi + (D-2)(D-3)(\tilde{\partial}\varphi)^4 \\ &\quad +4\big(\tilde{R}_{AB} - \frac{1}{2}\eta_{AB}\tilde{R}\big)(\tilde{\partial}^A\varphi)(\tilde{\partial}^B\varphi)\big]\big\} \\ &\quad +2(D-3)\tilde{e}\tilde{\nabla}_A\big\{e^{(D-4)\varphi}\big[(D-2)^2(\tilde{\partial}\varphi)^2\tilde{\partial}^A\varphi + 2(D-2)(\tilde{\Box}\varphi)\tilde{\partial}^A\varphi \\ &\quad -(D-2)\tilde{\partial}^A(\tilde{\partial}\varphi)^2 + 4\big(\tilde{R}^{AB} - \frac{1}{2}\eta^{AB}\tilde{R}\big)\tilde{\partial}_B\varphi\big]\big]\Big\},\end{aligned} \tag{11.2.5}$$

where $\tilde{R}^2_{GB}$ represents the rescaled Gauss-Bonnet combination. In $D=4$ the Lagrangian is only altered by a total derivative, while in $D=3$ the Lagrangian it is merely rescaled by a factor of $e^{-\varphi}$. The total derivative terms here will remain total derivatives even after the compactification. Along with the volume factor the Weyl-rescaling will determine the overall exponential dilaton factor, which shall play an important role in the analysis that follows.

11.2.2 Tree-Level Scalar Coset Symmetries

The internal vielbein $\hat{e}_m{}^a$ can be used to construct the internal metric $\hat{g}_{mn} = \hat{e}_m{}^a\hat{e}_n{}^b\delta_{ab}$, which is manifestly invariant under local $SO(n)$ rotations in the reduced directions. Thus we are free to fix a gauge for the internal vielbein using the $SO(n)$-invariance. After compactification the volume measure becomes $\tilde{e} = e\tilde{e}_{\rm int}$, where e is the determinant of the spacetime vielbein and $\tilde{e}_{\rm int}$ is the determinant of the internal vielbein. Defining the Weyl-rescaling coefficient as $e^{-(D-2)\varphi} \equiv \tilde{e}_{\rm int}$ ensures that the reduced Lagrangian is in Einstein frame.

The $GL(n,\mathbb{R})$ group element $\tilde{e}_m{}^a$ can now be parameterized in several ways, and we will discuss the two most natural choices here. The first choice is included for completeness, while it is the second choice which we shall subsequently employ in the compactification of the Gauss-Bonnet term.

First Parametrisation - Making the Symmetry Manifest

First, there is the possibility of separating out the determinant of the internal vielbein according to $\tilde{e}_m{}^a = (\tilde{e}_{\rm int})^{1/n}\varepsilon_m{}^a = e^{-\frac{(D-2)}{n}\varphi}\varepsilon_m{}^a$, where $\varepsilon_m{}^a$ is an element of $SL(n,\mathbb{R})$ in any preferred gauge. The line element takes the form

$$ds^2_D = e^{2\varphi}\left\{ds^2_{D-n} + e^{-2\frac{(D-2)}{n}\varphi}\left[(dx^m + \mathcal{A}^m_{(1)})\varepsilon_m{}^a\right]^2\right\}. \tag{11.2.6}$$

This Ansatz is nice for investigating the symmetry properties of the reduced Lagrangian because the $GL(n,\mathbb{R})$-symmetry of the internal torus is manifestly built into the formalism. More precisely, the reduction of the tree-level Einstein-Hilbert Lagrangian, $\hat{e}\hat{R}$, to $d = D-n$ dimensions becomes,

$$\mathcal{L}^{[d]}_{\rm EH} = e\left[R - \frac{1}{4}e^{-2\frac{(D-2)}{n}\xi\rho}F_{c\alpha\beta}F^{c\alpha\beta} - \frac{1}{2}(\partial\rho)^2 - {\rm tr}(P_\alpha P^\alpha) - 2\xi\Box\rho\right], \tag{11.2.7}$$

where $F^c{}_{\alpha\beta} \equiv \varepsilon_m{}^a F^m{}_{\alpha\beta}$ and

$$P_\alpha{}^{bc} \equiv \varepsilon^{m(b}\partial_\alpha \varepsilon_m{}^{c)} = \tilde{P}_\alpha{}^{bc} + \frac{(D-2)}{n}\xi\partial_\alpha\rho\delta^{bc}. \tag{11.2.8}$$

Notice that $P_\alpha{}^{bc}$ is $\mathfrak{sl}(n,\mathbb{R})$ valued and hence fulfills $\mathrm{tr}(P_\alpha) = 0$. To obtain Eq. (11.2.7) we also performed a scaling $\varphi = \xi\rho$ with $\xi = \sqrt{\frac{n}{2(D-2)(D-n-2)}}$, so as to ensure that the scalar field ρ appears canonically normalized in the Lagrangian.

The $SL(n,\mathbb{R})$-symmetry is manifest in this Lagrangian because the term $\mathrm{tr}(P_\alpha P^\alpha)$ is constructed using the invariant Killing form on $\mathfrak{sl}(n,\mathbb{R})$. By dualising the two-form field strength $F_{(2)}$, the symmetry is enhanced to $SL(n+1,\mathbb{R})$. With a slight abuse of terminology we call this the (classical) "U-duality" group. Since we are only investigating the pure gravity sector, this is of course only a subgroup of the full continuous U-duality group.

It was already shown in [208], that the tree-level symmetry $SL(n+1,\mathbb{R})$ is not realized in the compactified Gauss-Bonnet Lagrangian. It was argued, however, that the quantum symmetry $SL(n+1,\mathbb{Z})$ could be reinstated by "lifting" the result of the compactification through the use of automorphic forms. In the present analysis we take the same point of view, but since we now have access to the complete expression of the compactified Gauss-Bonnet Lagrangian we can here extend the analysis of [208] in some aspects. In order to do this we shall make use of a different parametrisation than the one displayed above, which illuminates the structure of the dilaton exponents in the Lagrangian. The dilaton exponents reveals the weight structure of the global symmetry group and so can give information regarding which representation of the U-duality group we are dealing with. Because we here have access to a complete expression this analysis is more exhaustive than the one presented in [206, 207].

Second Parametrisation - Revealing the Root Structure

The second natural choice of the internal vielbein is to parameterize it in triangular form by using dimension by dimension compactification [127, 128]. Instead of extracting only the determinant of the vielbein, one dilaton scalar is pulled out for each compactified dimension according to $\tilde{e}_m{}^a = e^{-\frac{1}{2}\vec{f}_a\cdot\vec{\phi}} u_m{}^a$, where $\vec{\phi} = (\phi_1,\ldots,\phi_n)$ and

$$\vec{f}_a = 2(\alpha_1,\ldots,\alpha_{a-1},(D-n-2+a)\alpha_a,\underbrace{0,\ldots,0}_{n-a}), \tag{11.2.9}$$

with

$$\alpha_a = \frac{1}{\sqrt{2(D-n-2+a)(D-n-3+a)}}. \tag{11.2.10}$$

The internal vielbein is now the Borel representative of the coset $GL(n,\mathbb{R})/SO(n)$, with the diagonal degrees of freedom $e^{-\frac{1}{2}\vec{f}_a\cdot\vec{\phi}}$ corresponding to the Cartan generators and the upper triangular degrees of freedom

$$u_m{}^a = [(1-\mathcal{A}_{(0)})^{-1}]_m{}^a = [1+\mathcal{A}_{(0)} + (\mathcal{A}_{(0)})^2 + \ldots]_m{}^a \tag{11.2.11}$$

corresponding to the positive root generators. The form of Eq. (11.2.11) follows naturally from a step by step compactification, where the scalar potentials $(\mathcal{A}_{(0)})^i_j$, arising from the compactification of the graviphotons, are nonzero only when $i > j$. The sum of the vectors $\vec{f}_a$ can be shown to be

$$\sum_{a=1}^{n} \vec{f}_a = \frac{D-2}{3}\vec{g}, \tag{11.2.12}$$

$\vec{g} \equiv 6(\alpha_1, \alpha_2 \ldots, \alpha_n)$. In addition, $\vec{g}$ and $\vec{f}_a$ obey

$$\begin{aligned} \vec{g}\cdot\vec{g} &= \frac{18n}{(D-2)(D-n-2)}, \\ \vec{g}\cdot\vec{f}_a &= \frac{6}{D-n-2}, \\ \vec{f}_a\cdot\vec{f}_b &= 2\delta_{ab} + \frac{2}{D-n-2}, \end{aligned} \tag{11.2.13}$$

and

$$\sum_{a=1}^{n}(\vec{f}_a\cdot\vec{x})(\vec{f}_a\cdot\vec{y}) = 2(\vec{x}\cdot\vec{y}) + \frac{D-2}{9}(\vec{g}\cdot\vec{x})(\vec{g}\cdot\vec{y}). \tag{11.2.14}$$

These scalar products can naturally be used to define the Cartan matrix, once a set of simple root vectors are found. The line element becomes

$$ds_D^2 = e^{\frac{1}{3}\vec{g}\cdot\vec{\phi}}\Big\{ds_{D-n}^2 + \sum_{a=1}^{n} e^{-\vec{f}_a\cdot\vec{\phi}}\left[(dx^m + \mathcal{A}_{(1)}^m)u_m{}^a\right]^2\Big\}, \tag{11.2.15}$$

yielding the corresponding Einstein-Hilbert Lagrangian in d dimensions

$$\mathcal{L}_{\mathrm{EH}}^{[d]} = e\left[R - \frac{1}{2}(\partial\vec{\phi})^2 - \frac{1}{4}\sum_{a=1}^{n} e^{-\vec{f}_a\cdot\vec{\phi}}F_{a\beta\gamma}F^{a\beta\gamma} - \frac{1}{2}\sum_{\substack{b,c=1\\ b<c}}^{n} e^{(\vec{f}_b-\vec{f}_c)\cdot\vec{\phi}}G_{\alpha bc}G^{\alpha bc} - \frac{1}{3}\vec{g}\cdot\Box\vec{\phi}\right], \tag{11.2.16}$$

with $F^c{}_{\alpha\beta} \equiv u_m{}^a F^m{}_{\alpha\beta}$ and

$$G_\alpha{}^{bc} = u^{mb}\partial_\alpha u_m{}^c = e^{-\frac{1}{2}(\vec{f}_b-\vec{f}_c)\cdot\vec{\phi}}\left[(\tilde{P}_\alpha{}^{bc} + \frac{1}{2}\vec{f}_b\cdot\partial_\alpha\vec{\phi}\delta^{bc}) + Q_\alpha{}^{bc}\right]. \tag{11.2.17}$$

Here, no Einstein's summation rule is assumed for the flat internal indices.

We shall refer to the various exponents of the form $e^{\vec{x}\cdot\vec{\phi}}$ ($\vec{x}$ being some vector in $\mathbb{R}^n$) collectively as "dilaton exponents". If relevant, this also includes the contribution from the overall volume factor.

All the results obtained in this parametrisation can be converted to the first parametrisation simply by using the following identifications

$$\begin{aligned} \frac{1}{3}(\vec{g}\cdot\vec{\phi}) &= 2\xi\rho, \\ \vec{f}_a\cdot\vec{\phi} &= 2\frac{(D-2)}{n}\xi\rho, \qquad \forall a, \\ \vec{\phi}\cdot\vec{\phi} &= \rho^2. \end{aligned} \tag{11.2.18}$$

Notice also that our compactification procedure breaks down at $D-n=2$, in which case the scalar products in Eq. (11.2.13) become ill-defined.

Even though proving the symmetry contained in the Lagrangian is somewhat more cumbersome compared to the first choice of parametrisation, since all the group actions have to be carried out adjointly in a formal manner, the second choice comes to its power when dealing with the exceptional symmetry groups of the supergravities for which no matrix representations exist. This parametrisation is particularly suitable for reading off the root vectors of the underlying symmetry algebra; they appear as exponential factors in front of each term in the Lagrangian. Identifying a complete set of root vectors in this way gives a necessary but not sufficient constraint on the underlying symmetry.

11.2.3 The Gauss-Bonnet Lagrangian Reduced to Three Dimensions

When reducing to $D-n=3$ dimensions, we can dualise the two-form field strength $\tilde{F}^{a}{}_{\alpha\beta} \equiv \tilde{e}_{m}{}^{a} F^{m}{}_{\alpha\beta}$ of the graviphoton $\mathcal{A}_{(1)}$ into the one-form $\tilde{H}_{a\alpha}$. More explicitly, we employ the standard dualisation so that

$$\delta_{ab}\tilde{F}^{b}{}_{\alpha\beta} = \epsilon_{\alpha\beta\gamma}\tilde{e}^{m}{}_{a}\partial^{\gamma}\chi_{m} \equiv \epsilon_{\alpha\beta\gamma}\tilde{H}_{a}{}^{\gamma}. \tag{11.2.19}$$

When we go to Einstein frame, the appearance of the inverse vielbein $\tilde{e}^{m}{}_{a}$ in the definition of the one-form $\tilde{H}_{a\alpha}$ implies there is a sign flip in the Lagrangian after dualisation. The dualisation presented here follows from the tree-level Lagrangian, but in general receives higher order α'-corrections. However, these lead to terms of higher derivative order than quartic and so can be neglected in the present analysis [206, 208].

The final result for the compactification is written in such way that the only explicit derivative terms appearing are divergences, total derivatives and first derivatives on the dilatons φ. The end result reads explicitly

$$\begin{aligned}
\mathcal{L}_{\mathrm{GB}}^{[3]} \;=\; & \sqrt{|g|}e^{-2\varphi}\Big\{ -\frac{1}{4}\tilde{H}_{a\gamma}\tilde{H}_{b}{}^{\gamma}\tilde{H}^{a}{}_{\delta}\tilde{H}^{b\delta} + \frac{1}{4}\tilde{H}^{2}\tilde{H}^{2} - 4\tilde{H}^{2}(\partial\varphi)^{2} + 2\tilde{H}^{c\alpha}\tilde{P}_{\alpha cd}\tilde{P}^{\beta de}\tilde{H}_{e\beta} \\
& -2\tilde{H}^{c\alpha}\tilde{P}_{\beta cd}\tilde{P}^{\beta de}\tilde{H}_{e\alpha} + 4\tilde{H}_{c\alpha}\tilde{P}^{\alpha cd}\tilde{H}_{d}{}^{\beta}\partial_{\beta}\varphi - 6\tilde{H}_{c\alpha}\tilde{P}^{\beta cd}\tilde{H}_{d}{}^{\alpha}\partial_{\beta}\varphi \\
& +2\mathrm{tr}(\tilde{P}_{\alpha}\tilde{P}_{\beta}\tilde{P}^{\alpha}\tilde{P}^{\beta}) + 2\mathrm{tr}(\tilde{P}_{\alpha}\tilde{P}_{\beta})\mathrm{tr}(\tilde{P}^{\alpha}\tilde{P}^{\beta}) - (\tilde{P}^{2})^{2} + 8\mathrm{tr}(\tilde{P}_{\alpha}\tilde{P}_{\beta}\tilde{P}^{\beta})\partial^{\alpha}\varphi \\
& -4(D-2)\mathrm{tr}(\tilde{P}_{\alpha}\tilde{P}_{\beta})\partial^{\alpha}\varphi\partial^{\beta}\varphi + 2(D+2)\tilde{P}^{2}(\partial\varphi)^{2} + (D-2)(D-4)(\partial\varphi)^{2}(\partial\varphi)^{2}\Big\},
\end{aligned} \tag{11.2.20}$$

where $\tilde{H}^{2} \equiv \tilde{H}_{a\beta}\tilde{H}^{a\beta}$ and $\tilde{P}^{2} \equiv \tilde{P}_{\alpha bc}\tilde{P}^{\alpha bc}$. Note that contributions from the boundary terms and terms proportional to the equations of motion have been ignored. The one-form $\tilde{P}_{\alpha}$ is the Maurer-Cartan form associated with the internal vielbein $\tilde{e}_{m}^{\ a}$, and so takes values in the Lie algebra $\mathfrak{gl}(n,\mathbb{R}) = \mathfrak{sl}(n,\mathbb{R}) \oplus \mathbb{R}$. Here, the abelian summand $\mathbb{R}$ corresponds to the "trace-part" of $\tilde{P}_{\alpha}$. Explicitly, we have $\mathrm{tr}(\tilde{P}_{\alpha}) = -(D-2)\partial_{\alpha}\varphi$. We shall discuss various properties of $\tilde{P}_{\alpha}$ in more detail below.

Finally, we note that the three-dimensional Gauss-Bonnet term is absent from the reduced Lagrangian because it vanishes identically in three dimensions:

$$R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} - 4R_{\alpha\beta}R^{\alpha\beta} + R^{2} = 0, \qquad (\alpha,\beta,\gamma,\delta = 1,2,3). \tag{11.2.21}$$

The remainder of this chapter is devoted to a detailed analysis of the symmetry properties of Eq. (11.2.20).

11.3 Algebraic Structure of the Compactified Gauss-Bonnet Term

We have seen that the Ansatz presented in Eq. (11.2.15) is particularly suitable for identifying the roots of the relevant symmetry algebra from the dilaton exponents associated with the diagonal components of the internal vielbein. Through this analysis one may deduce that for the lowest order effective action, the terms in the action are organized according to the adjoint representation of $\mathfrak{sl}(n+1,\mathbb{R})$, for which the weights are the roots. The aim of this section is to extend the analysis to the Gauss-Bonnet Lagrangian. By general arguments [206, 207], it has been shown that the exponents no longer correspond to roots of the symmetry algebra but rather they now lie on the weight lattice. Here, however, we have access to the complete compactified Lagrangian and we may therefore present an explicit counting of the weights in the dilaton exponents and identify the relevant $\mathfrak{sl}(n+1,\mathbb{R})$-representation.

An exhaustive analysis of the $\mathfrak{sl}(4,\mathbb{R})$-representation structure of the Gauss-Bonnet term compactified from 6 to 3 dimensions on T^3 is performed. We do this in two alternative ways.

First, we neglect the contribution from the overall dilaton factor $e^{-2\varphi}$ in the representation structure. This is consistent before dualisation because this factor is $SL(3,\mathbb{R})$-invariant. However, after dualisation this is no longer true and one must understand what role this factor plays in the algebraic structure. If one continues to neglect this factor then all the weights fit into the **84**-representation of $\mathfrak{sl}(4,\mathbb{R})$ with Dynkin labels $[2,0,2]$.

On the other hand, including the overall exponential dilaton factor in the weight structure induces a shift on the weights which "destroys" the **84** of $\mathfrak{sl}(4,\mathbb{R})$. Instead one finds that the highest weight is now associated with the **36**-representation of $\mathfrak{sl}(4,\mathbb{R})$, with Dynkin labels $[2,0,1]$. However, this representation is not "big enough" to incorporate all the weights in the Lagrangian. It turns out that there are additional weights outside of the **36** that fit into a **27** of $\mathfrak{sl}(3,\mathbb{R})$. Unfortunately, we are unable to determine which $\mathfrak{sl}(4,\mathbb{R})$-representation this belongs to, since the associated highest weight does not seem to be represented in the Lagrangian.

These results indicate that the general analysis performed in [206] is only partly correct, or, rather, that the interpretation given there might be incorrect. The highest weight of the representation $[2,0,1]$ does indeed appear in the reduced Lagrangian but it does not incorporate the full representation structure of the compactified quadratic curvature correction.

11.3.1 Kaluza-Klein Reduction and $\mathfrak{sl}(n,\mathbb{R})$-Representations

We shall begin by rewriting the reduction Ansatz in a way which has a more firm Lie algebraic interpretation. Recall from Eq. (11.2.15) that the standard Kaluza-Klein Ansatz for the metric is

$$ds_D^2 = e^{\frac{1}{3}\vec{g}\cdot\vec{\phi}}ds_d^2 + e^{\frac{1}{3}\vec{g}\cdot\vec{\phi}}\sum_{i=1}^{n} e^{-\vec{f}_i\cdot\vec{\phi}}\left[\left(dx^m + \mathcal{A}_{(1)}^m\right)u_m{}^a\right]^2. \tag{11.3.1}$$

The exponents in this Ansatz are linear forms on the space of dilatons. Let $\vec{e}_i$, $i = 1,\ldots,n$, constitute an n-dimensional orthogonal basis of $\mathbb{R}^n$,

$$\vec{e}_i\cdot\vec{e}_j = \delta_{ij}. \tag{11.3.2}$$

Since there is a non-degenerate metric on the space of dilatons (the Cartan subalgebra $\mathfrak{h}\subset\mathfrak{sl}(n+1,\mathbb{R})$) we can use this to identify this space with its dual space of linear forms. Thus, we may express all exponents in the orthogonal basis $\vec{e}_i$ and the vectors $\vec{f}_i, \vec{g}$ and $\vec{\phi}$ may then be written as

$$\begin{aligned}\vec{f}_i &= \sqrt{2}\vec{e}_i + \alpha\vec{g},\\ \vec{g} &= \beta\sum_{i=1}^{n}\vec{e}_i,\\ \vec{\phi} &= \phi_i\vec{e}_i,\end{aligned} \tag{11.3.3}$$

where the constants α and β are defined as

$$\begin{aligned}\alpha &= \frac{1}{3n}\Big(D-2-\sqrt{(D-n-2)(D-2)}\Big),\\ \beta &= \sqrt{\frac{18}{(D-n-2)(D-2)}}.\end{aligned} \tag{11.3.4}$$

Note here that the constant α is not the same as the α_a of Eq. (11.2.9).

The combinations

$$\vec{f}_i - \vec{f}_j = \sqrt{2}\vec{e}_i - \sqrt{2}\vec{e}_j \tag{11.3.5}$$

span an $(n-1)$-dimensional lattice which can be identified with the root lattice of $A_{n-1} = \mathfrak{sl}(n,\mathbb{R})$. For compactification of the pure Einstein-Hilbert action to three dimensions, the dilaton exponents precisely organize into the complete set of positive roots of $\mathfrak{sl}(n,\mathbb{R})$, revealing that it is the adjoint representation which is the relevant one for the U-duality symmetries of the lowest order (two-derivative) action. After dualisation of the Kaluza-Klein one forms $\mathcal{A}_{(1)}$ the symmetry is lifted to the full adjoint representation of $\mathfrak{sl}(n+1,\mathbb{R})$.

When we compactify higher derivative corrections to the Einstein-Hilbert action it is natural to expect that other representations of $\mathfrak{sl}(n,\mathbb{R})$ and $\mathfrak{sl}(n+1,\mathbb{R})$ become relevant. In order to pursue this question for the Gauss-Bonnet Lagrangian, we shall need some features of the representation theory of $A_n = \mathfrak{sl}(n+1,\mathbb{R})$.

Representation Theory of $A_n = \mathfrak{sl}(n+1,\mathbb{R})$

For the infinite class of simple Lie algebras A_n, it is possible to choose an embedding of the weight space $\mathfrak{h}^\star$ in $\mathbb{R}^{n+1}$ such that $\mathfrak{h}^\star$ is isomorphic to the subspace of $\mathbb{R}^{n+1}$ which is orthogonal to the vector $\sum_{i=1}^{n+1}\vec{e}_i$. We can use this fact to construct an embedding of the $(n-1)$-dimensional weight space of $A_{n-1} = \mathfrak{sl}(n,\mathbb{R})$ into the n-dimensional weight space of $A_n = \mathfrak{sl}(n+1,\mathbb{R})$, in terms of the n basis vectors $\vec{e}_i$ of $\mathbb{R}^n$.

To this end we define the new vectors

$$\begin{aligned}\vec{\omega}_i &= \vec{f}_i - (\alpha + \frac{\sqrt{2}}{n\beta})\vec{g}\\ &= \sqrt{2}\vec{e}_i - \frac{\sqrt{2}}{n}\sum_{j=1}^{n}\vec{e}_j,\end{aligned} \tag{11.3.6}$$

which have the property that

$$\vec{\omega}_i \cdot \vec{g} = \sqrt{2}\beta - \sqrt{2}\beta = 0. \tag{11.3.7}$$

This implies that the vectors $\vec{\omega}_i$ form a (non-orthogonal) basis of the $(n-1)$-dimensional subspace $U \subset \mathbb{R}^n$, orthogonal to $\vec{g}$. The space U is then isomorphic to the weight space $\mathfrak{h}^\star$ of $A_{n-1} = \mathfrak{sl}(n, \mathbb{R})$. Since there are n vectors $\vec{\omega}_i$, this basis is overcomplete. However, it is easy to see that not all $\vec{\omega}_i$ are independent, but are subject to the relation

$$\sum_{i=1}^{n} \vec{\omega}_i = 0. \tag{11.3.8}$$

A basis of simple roots of $\mathfrak{h}^\star$ can now be written in three alternative ways

$$\vec{\alpha}_i = \vec{f}_i - \vec{f}_{i+1} = \vec{\omega}_i - \vec{\omega}_{i+1} = \sqrt{2}(\vec{e}_i - \vec{e}_{i+1}), \qquad (i = 1, \ldots, n-1). \tag{11.3.9}$$

What is the algebraic interpretation of the vectors $\vec{\omega}_i$? It turns out that they may be identified with the weights of the n-dimensional fundamental representation of $\mathfrak{sl}(n, \mathbb{R})$. The condition $\sum_{i=1}^{n} \vec{\omega}_i = 0$ then reflects the fact that the generators of the fundamental representation are traceless.

In addition, we can use the weights of the fundamental representation to construct the fundamental weights $\vec{\Lambda}_i$, defined by

$$\vec{\alpha}_i \cdot \vec{\Lambda}_j = 2\delta_{ij}. \tag{11.3.10}$$

One finds

$$\vec{\Lambda}_i = \sum_{j=1}^{i} \vec{\omega}_j, \qquad (i = 1, \ldots, n-1), \tag{11.3.11}$$

which can be seen to satisfy Eq. (11.3.10).

The relation, Eq. (11.3.11), between the fundamental weights $\vec{\Lambda}_i$ and the weights of the fundamental representation $\vec{\omega}_i$ can be inverted to

$$\vec{\omega}_i = \vec{\Lambda}_i - \vec{\Lambda}_{i-1}, \qquad (i = 1, \ldots, n-1). \tag{11.3.12}$$

In addition, the n:th weight is

$$\vec{\omega}_n = -\vec{\Lambda}_{n-1}, \tag{11.3.13}$$

corresponding to the lowest weight of the fundamental representation.

We may now rewrite the Kaluza-Klein Ansatz in a way such that the weights $\vec{\omega}_i$ appear explicitly in the metric[2]

$$ds_D^2 = e^{\frac{1}{3}\vec{g}\cdot\vec{\phi}} ds_d^2 + e^{\gamma \vec{g}\cdot\vec{\phi}} \sum_{i=1}^{n} e^{-\vec{\omega}_i\cdot\vec{\phi}} \Big[\big(dx^m + \mathcal{A}^m_{(1)}\big) u_m{}^a\Big]^2, \tag{11.3.14}$$

with

$$\gamma = \frac{1}{3} - \alpha - \frac{\sqrt{2}}{n\beta}. \tag{11.3.15}$$

[2]A similar construction was given in [212].

11.3.2 The Algebraic Structure of Gauss-Bonnet in Three Dimensions

We are interested in the dilaton exponents in the scalar part of the three-dimensional Lagrangian. For the Einstein-Hilbert action we know that these are of the forms

$$\vec{f}_a - \vec{f}_b \quad (b > a), \qquad \text{and} \qquad \vec{f}_a. \tag{11.3.16}$$

The first set of exponents $\vec{f}_a - \vec{f}_b$ correspond to the positive roots of $\mathfrak{sl}(n,\mathbb{R})$ and the second set $\vec{f}_a$, which contributes to the scalar sector after dualisation, extends the algebraic structure to include all positive roots of $\mathfrak{sl}(n+1,\mathbb{R})$. The highest weight $\vec{\lambda}^{\text{hw}}_{\text{ad},n}$ of the adjoint representation of $A_n = \mathfrak{sl}(n+1,\mathbb{R})$ can be expressed in terms of the fundamental weights as

$$\vec{\lambda}^{\text{hw}}_{\text{ad},n} = \vec{\Lambda}_1 + \vec{\Lambda}_n, \tag{11.3.17}$$

corresponding to the Dynkin labels

$$[1, 0, \ldots, 0, 1].$$

We see that before dualisation the highest weight of the adjoint representation of $\mathfrak{sl}(n,\mathbb{R})$ arises in the dilaton exponents in the form $\vec{f}_1 - \vec{f}_n = \vec{\omega}_1 - \vec{\omega}_n = \vec{\Lambda}_1 + \vec{\Lambda}_{n-1} = \vec{\lambda}^{\text{hw}}_{\text{ad},n-1}$.

We proceed now to analyze the various dilaton exponents arising from the Gauss-Bonnet term after compactification to three dimensions. These can be extracted from each term in the Lagrangian Eq. (11.2.20) by factoring out the diagonal components of the internal vielbein according to $\tilde{e}_m{}^a = e^{-\frac{1}{2}\vec{f}_a\cdot\vec{\phi}} u_m{}^a$. For example, before dualisation we have the manifestly $SL(n,\mathbb{R})$-invariant term $\operatorname{tr}(\tilde{P}_\alpha \tilde{P}_\beta \tilde{P}^\alpha \tilde{P}^\beta)$. Expanding this gives (among others) the following types of terms

$$\begin{aligned} \operatorname{tr}\big(\tilde{P}_\alpha \tilde{P}_\beta \tilde{P}^\alpha \tilde{P}^\beta\big) \quad \sim \quad & \sum_{\substack{c\,<\,a,\,b\\ d\,<\,a,\,b}} e^{-2(\vec{f}_c+\vec{f}_d-\vec{f}_a-\vec{f}_b)\cdot\vec{\phi}} G_{\alpha ca} G^{\alpha c}{}_b G_{\beta d}{}^a G^{\beta db} + \cdots \\ & + \sum_{\substack{a\,>\,c\,>\,b\\ a\,>\,d\,>\,b}} e^{-2(\vec{f}_b-\vec{f}_a)\cdot\vec{\phi}} G_{\alpha ca} G^\alpha{}_b{}^c G^{\beta da} G_\beta{}^b{}_d + \cdots . \end{aligned} \tag{11.3.18}$$

After dualisation, we need to take into account also terms containing $\tilde{H}^\alpha{}_a$. We have then, for example, the term

$$\tilde{H}^4 \sim \sum_{a,b} e^{-2(\vec{f}_a+\vec{f}_b)\cdot\vec{\phi}} H^4. \tag{11.3.19}$$

Many different terms in the Lagrangian might in this way give rise to the same dilaton exponents. As can be seen from Eq. (11.3.18), the internal index contractions yield constraints on the various exponents. We list below all the "independent" exponents, i.e., those which are the least constrained. All other exponents follow as special cases of these. Before dualisation we then find the following exponents:

$$\begin{aligned} \vec{f}_a - \vec{f}_b \qquad & (b > a), \\ \vec{f}_c + \vec{f}_d - \vec{f}_a - \vec{f}_b \qquad & (c < a,\ c < b,\ d < a,\ d < b), \\ \vec{f}_a + \vec{f}_b - \vec{f}_c - \vec{f}_d \qquad & (b < c,\ a < d), \end{aligned} \tag{11.3.20}$$

and after dualisation we also get a contribution from

$$\begin{aligned}&\vec{f}_a,\\ &\vec{f}_a+\vec{f}_b,\\ &\vec{f}_a+\vec{f}_b-\vec{f}_c, \qquad (b<c).\end{aligned} \tag{11.3.21}$$

Let us investigate the general weight structure of the dilaton exponents before dualisation. The highest weight arises from the terms of the form $\vec{f}_c+\vec{f}_d-\vec{f}_a-\vec{f}_b$ when $c=d=1$ and $a=b=n$, i.e., for the dilaton vector $2\vec{f}_1-2\vec{f}_n$. This can be written in terms of the fundamental weights as follows

$$2\vec{f}_1-2\vec{f}_n=2\vec{\omega}_1-2\vec{\omega}_n=2\vec{\Lambda}_1+2\vec{\Lambda}_{n-1}, \tag{11.3.22}$$

which is the highest weight of the $[2,0,\ldots,0,2]$-representation of $\mathfrak{sl}(n,\mathbb{R})$.

11.3.3 Special Case: Compactification from $D=6$ on T^3

In order to determine if this is indeed the correct representation for the Gauss-Bonnet term, we shall now restrict to the case of $n=3$, i.e., compactification from $D=6$ on T^3. We do this so that a complete counting of the weights in the Lagrangian is a tractable task. Before dualisation we then expect to find the representation **27** of $\mathfrak{sl}(3,\mathbb{R})$, with Dynkin labels $[2,2]$. We wish to investigate if, after dualisation, this representation lifts to the representation **84** of $\mathfrak{sl}(4,\mathbb{R})$, with Dynkin labels $[2,0,2]$.

It is important to realize that of course the Lagrangian will not display the complete set of weights in these representations, but only the *positive* weights, i.e., the ones that can be obtained by summing positive roots only. Let us begin by analyzing the weight structure before dualisation. From Eq. (11.3.20) we find the weights

$$\begin{aligned}&\vec{f}_1-\vec{f}_2, \qquad \vec{f}_2-\vec{f}_3, \qquad \vec{f}_1-\vec{f}_3,\\ &2(\vec{f}_1-\vec{f}_2), \qquad 2(\vec{f}_2-\vec{f}_3), \qquad 2(\vec{f}_1-\vec{f}_3),\\ &2\vec{f}_1-\vec{f}_2-\vec{f}_3, \qquad \vec{f}_1+\vec{f}_2-2\vec{f}_3.\end{aligned} \tag{11.3.23}$$

The first three may be identified with the positive roots of $\mathfrak{sl}(3,\mathbb{R})$, $\vec{\alpha}_1=\vec{f}_1-\vec{f}_2$, $\vec{\alpha}_2=\vec{f}_2-\vec{f}_3$ and $\vec{\alpha}_\theta=\vec{f}_1-\vec{f}_3$. The second line then corresponds to $2\vec{\alpha}_2$, $2\vec{\alpha}_2$ and $2\vec{\alpha}_\theta$. The remaining weights are

$$\begin{aligned}\vec{f}_1+\vec{f}_2-2\vec{f}_3 &= \vec{\alpha}_1+2\vec{\alpha}_2,\\ 2\vec{f}_1-\vec{f}_2-\vec{f}_3 &= 2\vec{\alpha}_1+\vec{\alpha}_2.\end{aligned} \tag{11.3.24}$$

These weights are precisely the eight positive weights of the **27** representation of $\mathfrak{sl}(3,\mathbb{R})$.

We now wish to see if this representation lifts to the corresponding representation of $\mathfrak{sl}(4,\mathbb{R})$, upon inclusion of the weights in Eq. (11.3.21). As mentioned above, the representation of $\mathfrak{sl}(4,\mathbb{R})$ with Dynkin labels $[2,0,2]$ is 84-dimensional. It is illuminating to first decompose this in terms of representations of $\mathfrak{sl}(3,\mathbb{R})$,

$$\mathbf{84}=\mathbf{27}\oplus\mathbf{15}\oplus\overline{\mathbf{15}}\oplus\mathbf{6}\oplus\overline{\mathbf{6}}\oplus\mathbf{8}\oplus\mathbf{3}\oplus\overline{\mathbf{3}}\oplus\mathbf{1}, \tag{11.3.25}$$

or, in terms of Dynkin labels,

$$[2,0,2] = [2,2] + [2,1] + [1,2] + [2,0] + [0,2] + [1,1] + [1,0] + [0,1] + [0,0]. \tag{11.3.26}$$

We may view this decomposition as a *level decomposition* of the representation **84**, with the level ℓ being represented by the number of times the third simple root $\vec{\alpha}_3$ appears in each representation. From this point of view, and as we shall see in more detail shortly, the representations $\mathbf{27}, \mathbf{8}$ and $\mathbf{1}$ reside at $\ell = 0$, the representations **15** and **3** at $\ell = 1$, and the representation **6** at $\ell = 2$. The "barred" representations then reside at the associated negative levels. We know that we can only expect to find the strictly positive weights in these representations. Let us therefore begin to count these.

Firstly, we may neglect all representations at negative levels since these do not contain any positive weights. However, not all weights for $\ell \geq 0$ are positive. If we had decomposed the adjoint representation of $\mathfrak{sl}(4,\mathbb{R})$ this problem would not have been present since all roots are either positive or negative, and hence all weights at positive level are positive and vice versa. In our case this is not true because for representations larger than the adjoint many weights are neither positive nor negative. It is furthermore important to realize that after dualisation it is the positive weights of $\mathfrak{sl}(4,\mathbb{R})$ that we will obtain and not of $\mathfrak{sl}(3,\mathbb{R})$. As can be seen in Figure 11.1 the decomposition indeed includes weights which are negative weights of $\mathfrak{sl}(3,\mathbb{R})$ but nevertheless positive weights of $\mathfrak{sl}(4,\mathbb{R})$. An explicit counting reveals the following number of positive weights at each level (not counting weight multiplicities):

$$\begin{aligned} \ell = 0 &\ : \ 8, \\ \ell = 1 &\ : \ 8, \\ \ell = 2 &\ : \ 6. \end{aligned} \tag{11.3.27}$$

The eight weights at level zero are of course the positive weights of the **27** representation of $\mathfrak{sl}(3,\mathbb{R})$ that we had before dualisation. In order to verify that we find all positive weights of **84** we must now check explicitly that after dualisation we get 8+6 additional positive weights. The total number of distinct weights of $\mathfrak{sl}(4,\mathbb{R})$ that should appear in the Lagrangian after compactification and dualisation is thus 22.

The lifting from $\mathfrak{sl}(3,\mathbb{R})$ to $\mathfrak{sl}(4,\mathbb{R})$ is done by adding the third simple root $\vec{\alpha}_3 \equiv \vec{f}_3$, from Eq. (11.3.21). The complete set of new weights arising from Eq. (11.3.21) is then

$$\begin{aligned}
\ell = 1 \quad : \quad & \vec{f}_1 = \vec{\alpha}_1 + \vec{\alpha}_2 + \vec{\alpha}_3, & \vec{f}_2 &= \vec{\alpha}_2 + \vec{\alpha}_3, \\
& 2\vec{f}_1 - \vec{f}_2 = 2\vec{\alpha}_1 + \vec{\alpha}_2 + \vec{\alpha}_3, & 2\vec{f}_2 - \vec{f}_3 &= 2\vec{\alpha}_2 + \vec{\alpha}_3, \\
& 2\vec{f}_1 - \vec{f}_3 = 2\vec{\alpha}_1 + 2\vec{\alpha}_2 + \vec{\alpha}_3 & \vec{f}_1 + \vec{f}_2 - \vec{f}_3 &= \vec{\alpha}_1 + 2\vec{\alpha}_2 + \vec{\alpha}_3, \\
& \vec{f}_1 + \vec{f}_3 - \vec{f}_2 = \vec{\alpha}_1 + \vec{\alpha}_3, & \vec{f}_3 &= \vec{\alpha}_3, \\
\\
\ell = 2 \quad : \quad & 2\vec{f}_1 = 2\vec{\alpha}_1 + 2\vec{\alpha}_2 + 2\vec{\alpha}_3, & 2\vec{f}_2 &= 2\vec{\alpha}_2 + 2\vec{\alpha}_3, \\
& 2\vec{f}_3 = 2\vec{\alpha}_3, & \vec{f}_1 + \vec{f}_2 &= \vec{\alpha}_1 + 2\vec{\alpha}_2 + 2\vec{\alpha}_3, \\
& \vec{f}_1 + \vec{f}_3 = \vec{\alpha}_1 + \vec{\alpha}_2 + 2\vec{\alpha}_3, & \vec{f}_2 + \vec{f}_2 &= \vec{\alpha}_2 + 2\vec{\alpha}_3.
\end{aligned} \tag{11.3.28}$$

In Table 11.1 we indicate which representations these weights belong to and in Figure 11.1 we give a graphical presentation of the level decomposition. These results show that the

Reps	ℓ	Positive Weights of $\mathfrak{sl}(4,\mathbb{R})$
3	1	$\vec{\alpha}_3,\ \vec{\alpha}_2+\vec{\alpha}_3,\ \vec{\alpha}_1+\vec{\alpha}_2+\vec{\alpha}_3$
15	1	$2\vec{\alpha}_2+\vec{\alpha}_3,\ \vec{\alpha}_1+2\vec{\alpha}_2+\vec{\alpha}_3,\ 2\vec{\alpha}_1+2\vec{\alpha}_2+\vec{\alpha}_3,$ $2\vec{\alpha}_1+\vec{\alpha}_2+\vec{\alpha}_3,\ \vec{\alpha}_1+\vec{\alpha}_3$
6	2	$2\vec{\alpha}_2+\vec{\alpha}_3,\ \vec{\alpha}_1+2\vec{\alpha}_2+\vec{\alpha}_3,\ 2\vec{\alpha}_1+2\vec{\alpha}_2+2\vec{\alpha}_3,$ $\vec{\alpha}_1+\vec{\alpha}_2+2\vec{\alpha}_3,\ 2\vec{\alpha}_3,\ \vec{\alpha}_2+2\vec{\alpha}_3$

Table 11.1: Positive weights at levels one and two.

Gauss-Bonnet term in $D=6$ compactified on T^3 to three dimensions indeed gives rise to all strictly positive weights of the **84**-representation of $\mathfrak{sl}(4,\mathbb{R})$.

Weight Multiplicities

We have shown that the six-dimensional Gauss-Bonnet term compactified to three dimensions gives rise to all positive weights of the **84**-representation of $\mathfrak{sl}(4,\mathbb{R})$. However, we have not yet addressed the issue of weight multiplicities. It is not clear how approach this problem. Naively, one might argue that if k distinct terms in the Lagrangian are multiplied by the same dilaton exponential, corresponding to some weight $\vec{\lambda}$, then this weight has multiplicity k. Unfortunately, this type of counting does not seem to work.

Consider, for instance, the representations at $\ell=1$. Both representations **15** and **3** contain the weights $\vec{f}_1$, $\vec{f}_2$ and $\vec{f}_3$. In **15** these have all multiplicity 2, while in **3** they have multiplicity 1. Thus, in total these weights have multiplicity 3 as weights of $\mathfrak{sl}(3,\mathbb{R})$. Now, investigation of the dilaton exponents in Eq. (11.3.21) reveals that these weights arise from the two types of exponents $\vec{f}_a$ and $\vec{f}_a+\vec{f}_b-\vec{f}_c$, $(b<c)$. The first type gives rise to all three weights $\vec{f}_1$, $\vec{f}_2$ and $\vec{f}_3$, while the second type only contributes $\vec{f}_2=\vec{f}_3+\vec{f}_2-\vec{f}_3$ and $\vec{f}_1=\vec{f}_3+\vec{f}_1-\vec{f}_3$. Thus, by the arguments above the weights $\vec{f}_1$ and $\vec{f}_2$ has multiplicity 2 in the Lagrangian, and $\vec{f}_3$ has multiplicity 1. We therefore deduce that for all these weights there appears to be a mismatch in the multiplicity.

We suggest that the correct way to interpret this discrepancy in the weight multiplicities is as an indicative of the need to introduce transforming automorphic forms in order to restore the $SL(4,\mathbb{Z})$-invariance. This will be discussed more closely in Section 11.4.

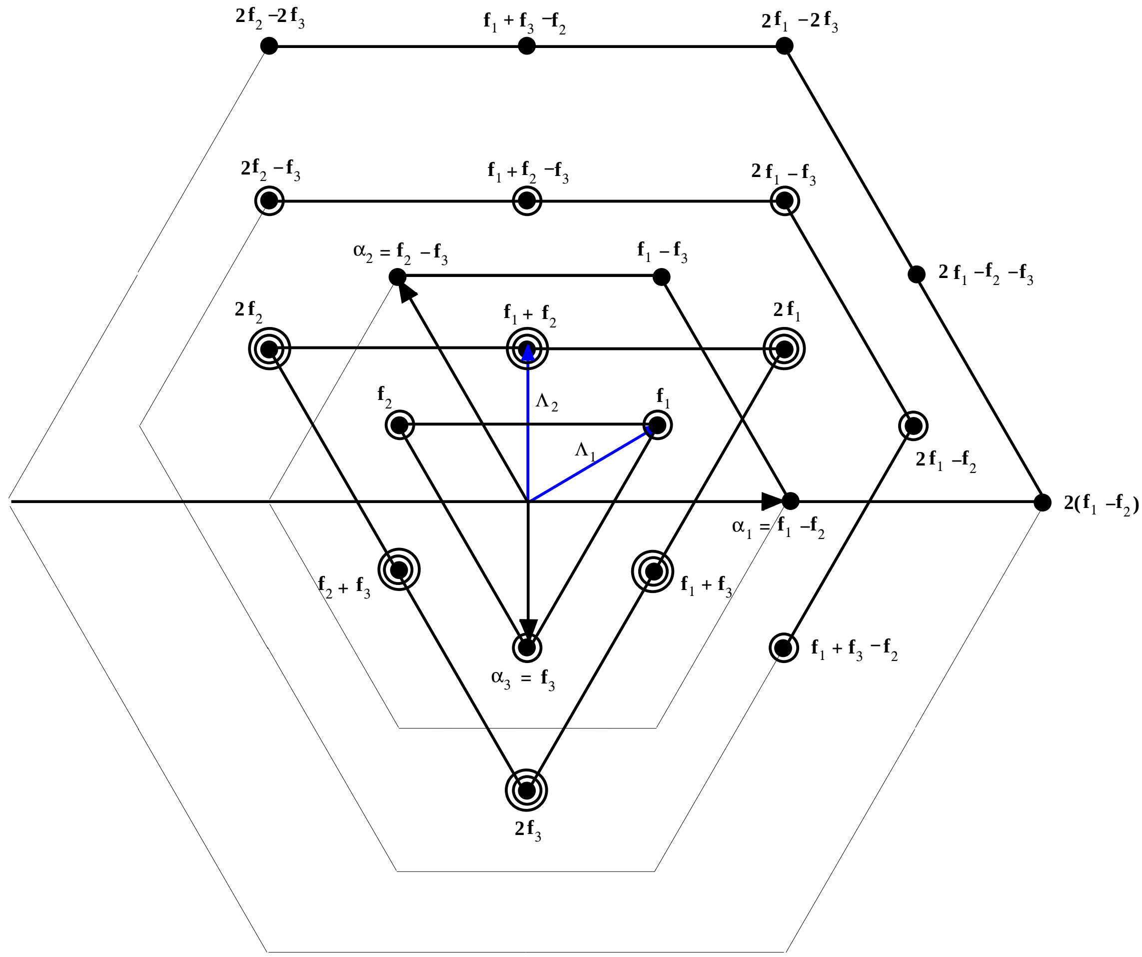


Figure 11.1: Graphical presentation of the representation structure of the compactified Gauss-Bonnet term. The black nodes arise from distinct dilaton exponents in the three-dimensional Lagrangian. The figure displays the level decomposition of the **84**-representation of $\mathfrak{sl}(4,\mathbb{R})$ into representations of $\mathfrak{sl}(3,\mathbb{R})$. Only positive levels are displayed. The black nodes correspond to positive weights of **84** of $\mathfrak{sl}(4,\mathbb{R})$. Nodes with no rings represent the positive weights of the level zero representation **27**, nodes with one ring represent the positive weights of the level one representations **15** and **3**, while nodes with two rings represent the positive weights of the level two representation **6**. The shaded lines complete the representations with non-positive weights which are not displayed explicitly.

Including the Dilaton Prefactor

We will now revisit the analysis from Section 11.3.3, but here we include the contribution from the overall exponential factor $e^{-2\varphi}$ in the Lagrangian Eq. (11.2.20). This factor arises as follows. The determinant of the D-dimensional vielbein is given by $\hat{e} = e^{D\varphi}\tilde{e}$, because of the Weyl-rescaling. Moreover, upon compactification the determinant of the rescaled vielbein splits according to $\tilde{e} = e\tilde{e}_{\text{int}}$, where e represents the external vielbein and $\tilde{e}_{\text{int}}$ the internal vielbein. By the definition of the Weyl-rescaling we have $\tilde{e}_{\text{int}} = e^{-(D-2)\varphi}$. This represents the volume of the n-torus, upon which we perform the reduction. Thus, the overall scaling contribution from the measure is $e^{D\varphi}e^{-(D-2)\varphi} = e^{2\varphi}$. In addition, we have a factor of $e^{-4\varphi}$ from the Gauss-Bonnet term. This gives a total overall dilaton prefactor of $e^{-2\varphi}$, which, after inserting $\varphi = \frac{1}{6}\vec{\phi}\cdot\vec{g}$, becomes $e^{-\frac{1}{3}\vec{\phi}\cdot\vec{g}}$.

The importance of the volume factor for compactified higher derivative terms was emphasized in [206], using the argument that after dualisation this factor is no longer invariant under the extended symmetry group $SL(n+1,\mathbb{R})$ and so must be included in the weight structure. We shall see that the inclusion of this factor drastically modifies the previous structure and destroys the representation **84** of $\mathfrak{sl}(4,\mathbb{R})$.

In order to perform this analysis, it is useful to first rewrite the simple roots and fundamental weights in a way which makes a comparison with [206] possible. We define arbitrary 3-vectors in $\mathbb{R}^3$ as follows

$$\hat{\vec{v}} = v_1\vec{\Lambda}_1 + v_2\vec{\Lambda}_2 + v_g\vec{g} = \big(\vec{v}, v_g\big) = \big(v_1, v_2, v_g\big), \tag{11.3.29}$$

where $\vec{\Lambda}_1$ and $\vec{\Lambda}_2$ are the fundamental weights of $\mathfrak{sl}(3,\mathbb{R})$ and $\vec{g}$ is the basis vector taking us from the weight space $\mathbb{R}^2$ of $\mathfrak{sl}(3,\mathbb{R})$ to the weight space $\mathbb{R}^3$ of $\mathfrak{sl}(4,\mathbb{R})$. Note that

$$\vec{\Lambda}_1\cdot\vec{g} = \vec{\Lambda}_2\cdot\vec{g} = 0, \tag{11.3.30}$$

by virtue of Eq. (11.3.7) and Eq. (11.3.11), which implies

$$\hat{\vec{v}}\cdot\hat{\vec{u}} = \vec{v}\cdot\vec{u} + v_g v_g \vec{g}\cdot\vec{g}. \tag{11.3.31}$$

The scalar products may all be deduced from the orthonormal basis $\vec{e}_i$ of $\mathbb{R}^3$. Restricting to $D=6$ and $n=3$ gives

$$\vec{f}_a = \sqrt{2}\vec{e}_a + \frac{2}{9}\vec{g}, \tag{11.3.32}$$

and thus

$$\vec{\omega}_a = \vec{f}_a - \frac{4}{9}\vec{g}. \tag{11.3.33}$$

The relevant scalar products become

$$\begin{aligned}
\vec{g}\cdot\vec{g} &= \frac{54}{4}, \\
\vec{g}\cdot\vec{f}_a &= 6, \\
\vec{f}_a\cdot\vec{f}_b &= 2\delta_{ab} + 2, \\
\vec{\omega}_a\cdot\vec{\omega}_b &= 2\delta_{ab} - \frac{2}{3}.
\end{aligned} \tag{11.3.34}$$

The simple roots of $\mathfrak{sl}(3,\mathbb{R})$ may now be written as

$$\begin{aligned}\hat{\vec{\alpha}}_1 &= \big(\vec{\alpha}_1, 0\big) = \big(2,-1,0\big),\\ \hat{\vec{\alpha}}_2 &= \big(\vec{\alpha}_2, 0\big) = \big(-1,2,0\big),\end{aligned} \tag{11.3.35}$$

and the third simple root becomes

$$\hat{\vec{\alpha}}_3 = \vec{f}_3 = \vec{\omega}_3 + \frac{4}{9}\vec{g} = -\vec{\Lambda}_2 + \frac{4}{9}\vec{g} = \big(0,-1,\frac{4}{9}\big). \tag{11.3.36}$$

In addition, the associated fundamental weights $\hat{\vec{\Lambda}}_i$, $i = 1,2,3$, of $\mathfrak{sl}(4,\mathbb{R})$, defined by

$$\hat{\vec{\alpha}}_i \cdot \hat{\vec{\Lambda}}_j = 2\delta_{ij}, \tag{11.3.37}$$

become

$$\hat{\vec{\Lambda}}_1 = \big(1,0,\frac{1}{9}\big), \quad \hat{\vec{\Lambda}}_2 = \big(0,1,\frac{2}{9}\big), \quad \hat{\vec{\Lambda}}_3 = \big(0,0,\frac{1}{3}\big). \tag{11.3.38}$$

Let us check that these indeed correspond to the fundamental weights of $\mathfrak{sl}(4,\mathbb{R})$, by computing the highest weight $2\hat{\vec{\Lambda}}_1 + 2\hat{\vec{\Lambda}}_3$ explicitly,

$$\begin{aligned}2\hat{\vec{\Lambda}}_1 + 2\hat{\vec{\Lambda}}_3 &= 2\vec{\Lambda}_1 + \frac{2}{9}\vec{g} + \frac{2}{3}\vec{g}\\ &= 2(\vec{\omega}_1 + \frac{4}{9}\vec{g})\\ &= 2\vec{f}_1\\ &= 2\hat{\vec{\alpha}}_1 + 2\hat{\vec{\alpha}}_2 + 2\hat{\vec{\alpha}}_3.\end{aligned} \tag{11.3.39}$$

This result is consistent with being the highest weight of the **84** representation of $\mathfrak{sl}(4,\mathbb{R})$ as can be seen in Figure 11.1.

Let us now include the dilaton prefactor in the analysis. In terms of $\mathfrak{sl}(4,\mathbb{R})$-vectors the volume factor can be identified with a negative shift in $\hat{\vec{\Lambda}}_3$, i.e.,

$$e^{-\frac{1}{3}\vec{g}\cdot\vec{\phi}} = e^{-\hat{\vec{\Lambda}}_3\cdot\vec{\phi}}. \tag{11.3.40}$$

As already mentioned above, this factor is irrelevant before dualisation because $\vec{g}\cdot\vec{\phi}$ is invariant under $SL(3,\mathbb{R})$. Thus, before dualisation the manifest $SL(3,\mathbb{R})$-symmetry of the compactified Gauss-Bonnet term is associated with the **27**-representation of $\mathfrak{sl}(3,\mathbb{R})$.

After dualisation, all the dilaton exponents in Eq. (11.3.20) and Eq. (11.3.21) become shifted by a factor of $-\hat{\vec{\Lambda}}_3$. In particular, the new highest weight is

$$\big(2\hat{\vec{\Lambda}}_1 + 2\hat{\vec{\Lambda}}_3\big) - \hat{\vec{\Lambda}}_3 = 2\hat{\vec{\Lambda}}_1 + \hat{\vec{\Lambda}}_3, \tag{11.3.41}$$

corresponding to the **36** representation of $\mathfrak{sl}(4,\mathbb{R})$, with Dynkin labels $[2,0,1]$. This is consistent with the general result of [206] that a generic curvature correction to pure Einstein gravity of order $l/2$ should be associated with an $\mathfrak{sl}(n+1,\mathbb{R})$-representation with highest weight $\frac{l}{2}\hat{\vec{\Lambda}}_1 + \hat{\vec{\Lambda}}_n$.

However, this is not the full story. A more careful examination in fact reveals that the **36** representation cannot incorporate all the dilaton exponents appearing in the Lagrangian, in contrast to the **84**-representation of Figure 11.1. To see this, let us decompose **36** in terms of representations of $\mathfrak{sl}(3,\mathbb{R})$. The result is:

$$\begin{aligned} \mathbf{36} &= \mathbf{15} \oplus \mathbf{8} \oplus \mathbf{6} \oplus \mathbf{3} \oplus \bar{\mathbf{3}} \oplus \mathbf{1}, \\ [2,0,1] &= [2,1]+[1,1]+[2,0]+[1,0]+[0,1]+[0,0]. \end{aligned} \tag{11.3.42}$$

Comparing this with Eq. (11.3.25), we see that the representations $\mathbf{27}, \bar{\mathbf{15}}$ and $\bar{\mathbf{6}}$ are no longer present. For the latter two this is not a problem since they were never present in the previous analysis. What happens is that the **6** of **84** gets shifted "downwards" and becomes the **6** of **36**. Similarly, the **15** and **3** of **84** become the **15** and **3** of **36**. This takes into account all the shifted dilaton exponents arising from the dualisation process. However, since there is not enough "room" for the **27** of $\mathfrak{sl}(3,\mathbb{R})$ in Eq. (11.3.42), some of the dilaton exponents (the ones corresponding to $2\vec{f}_2-2\vec{f}_3$, $\vec{f}_1+\vec{f}_3-\vec{f}_2$, $2\vec{f}_1-2\vec{f}_3, 2\vec{f}_1-\vec{f}_2-\vec{f}_3$ and $2\vec{f}_1-2\vec{f}_2$) arising from pure $\tilde{P}$-terms, i.e., before dualisation, remain outside of **36**. In fact, due to the shift of $-\hat{\vec{\Lambda}}_3$ these have now become *negative* weights of $\mathfrak{sl}(4,\mathbb{R})$, because they are below the hyperplane defined by $\vec{g}\cdot\vec{\phi}=0$. Although we know that these weights still correspond to positive weights of the **27** of $\mathfrak{sl}(3,\mathbb{R})$, we cannot determine which representation of $\mathfrak{sl}(4,\mathbb{R})$ they belong to.

By a straightforward generalisation of this analysis to compactifications of quadratic curvature corrections from arbitrary dimensions D, we may conclude that the highest weight $2\hat{\vec{\Lambda}}_1+\hat{\vec{\Lambda}}_n$, can never incorporate the dilaton exponents associated with the $[2,0,\ldots,0,2]$-representation of $\mathfrak{sl}(n,\mathbb{R})$ before dualisation.

11.4 An $SL(n+1,\mathbb{Z})$-Invariant Quartic Effective Action?

It is clear from the analysis in the previous section that the overall dilaton factor $e^{-\hat{\vec{\Lambda}}_3\cdot\vec{\phi}}$ (or, more generally, $e^{-\hat{\vec{\Lambda}}_n\cdot\vec{\phi}}$) complicates the interpretation of the dilaton exponents in terms of $\mathfrak{sl}(n+1,\mathbb{R})$-representations. To analyze this question, let us discuss what kind of information is encoded in the weight structure. Apart from the overall dilaton factor, the reduction of any higher derivative term $\sim \mathcal{R}^p$ will give rise to terms with $\mathcal{P}^{2p}$ (and terms with more derivatives and fewer $\mathcal{P}$'s), where $\mathcal{P}$ represents any of the "building blocks" P, H and $\partial\phi$ (we suppress all 3-dimensional indices). The appearance of weights of $\mathfrak{sl}(n+1,\mathbb{R})$ (without the uniform shift from the overall dilaton factor) reflects the fact that we use fields which are components of the symmetric part of the left-invariant Maurer–Cartan form $\mathcal{P}$ of $\mathfrak{sl}(n+1,\mathbb{R})$. Moreover, the dilaton factor contains information about the number of such fields. A term $\mathcal{R}^{l/2}$ will generically give weights in the weight space of the representation $[l/2,0,\ldots,0,l/2]$ of $\mathfrak{sl}(n+1,\mathbb{R})$, and fill out the positive part of this weight space.[3] This much is clear from the observation that the overall dilaton factor really is "overall".

The presence of the overall dilaton factor shifts this weight space uniformly in a negative direction. This shift happens to be by a vector in the weight lattice of $\mathfrak{sl}(n+1,\mathbb{R})$ for any

[3]We note that the representation structure encountered here is of the same type as for the lattice of BPS charges in string theory on T^n [13].

value of p. However, we emphasize that the dilaton exponents still lie in the weight space of the representation $[l/2, 0, \ldots, 0, l/2]$, albeit shifted "downwards". From this point of view, the weight space of the representation with the shifted highest weight of $[l/2, 0, \ldots, 0, l/2]$ as highest weight – for example, the representation $[2, 0, 1]$ in the case discussed above – does not contain all the weights that appear in the reduced Lagrangian, and is thus not relevant.

11.4.1 Completion Under $SL(n+1,\mathbb{Z})$

Consider now the fact that it is really the discrete "U-duality" group $SL(n+1,\mathbb{Z}) \subset SL(n+1,\mathbb{R})$ which is expected to be a symmetry of the complete effective action. Therefore, the compactified action should be seen as a remnant of the full U-duality invariant action, arising from a "large volume expansion" of certain automorphic forms.

Schematically, a generic, quartic, scalar term in the action after compactification of the Gauss-Bonnet term is of the form

$$\int d^3x \sqrt{|g|} e^{-\hat{\vec{\Lambda}}_n \cdot \vec{\phi}} F(\mathcal{P}), \tag{11.4.1}$$

where $F(\mathcal{P})$ is a quartic polynomial in the components of the Maurer–Cartan form mentioned above. F will be invariant under $SO(n)$ by construction, but generically not under $SO(n+1)$.

To obtain an action which is a scalar under $SO(n+1)$ we must first "lift" the result of the compactification to a globally $SL(n+1,\mathbb{Z})$-invariant expression. This can be done by replacing $e^{-\hat{\vec{\Lambda}}_n \cdot \vec{\phi}} F(\mathcal{P})$ by a suitable automorphic form contracted with four $\mathcal{P}$'s:

$$\Psi_{I_1 \ldots I_8}(X) \mathcal{P}^{I_1 I_2} \mathcal{P}^{I_3 I_4} \mathcal{P}^{I_5 I_6} \mathcal{P}^{I_7 I_8}, \tag{11.4.2}$$

where the I's are vector indices of $SO(n+1)$. Here, $\Psi(X)$ is an automorphic form transforming in some representation of $SO(n+1)$, and is constructed as an Eisenstein series, following, e.g., refs. [207, 208]. We must demand that when the large volume limit, $\hat{\vec{\Lambda}}_n \cdot \vec{\phi} \to -\infty$, is imposed, the leading behaviour is

$$\Psi_{I_1 \ldots I_8}(X) \mathcal{P}^{I_1 I_2} \mathcal{P}^{I_3 I_4} \mathcal{P}^{I_5 I_6} \mathcal{P}^{I_7 I_8} \quad \longrightarrow \quad e^{-\hat{\vec{\Lambda}}_n \cdot \vec{\phi}} F(\mathcal{P}). \tag{11.4.3}$$

This limit was taken explicitly in [207, 208]. This gives conditions on which irreducible $SO(n+1)$ representations the automorphic forms transform under (from the tensor structure), as well as a single condition on the "weights" of the automorphic forms (from the matching of the overall dilaton factor). Automorphic forms exist for continuous values of the weight (unlike holomorphic Eisenstein series) above some minimal value derived from convergence of the Eisenstein series. It was proven in [208] that any $SO(n)$-covariant tensor structure can be reproduced as the large volume limit of some automorphic form, and that the weight dictated by the overall dilaton factor is consistent with the convergence criterion.

Under the assumption that these arguments are valid, we may conclude that the representation theoretic structure of the dilaton exponents in the polynomial F should be analyzed without inclusion of the volume factor $e^{-\hat{\vec{\Lambda}}_n \cdot \vec{\phi}}$, and hence, for the Gauss-Bonnet term ($l = 4$), it is the $[2, 0, \ldots, 0, 2]$-representation which is the relevant one (in the sense above, that we are dealing with products of four Maurer–Cartan forms), and *not* the representation $[2, 0, \ldots, 0, 1]$ which was advocated in [206]. Another indication for why the representation with highest weight $2\hat{\vec{\Lambda}}_1 + \hat{\vec{\Lambda}}_n$ cannot be the relevant one is that it is not contained in the tensor product of the adjoint representation $[1, 0, \ldots, 0, 1]$ of $\mathfrak{sl}(n+1, \mathbb{R})$ with itself.

The present point of view also suggest a possible explanation for the discrepancy of the weight multiplicities observed in the previous section. In the complete $SL(n+1, \mathbb{Z})$-invariant four-derivative effective action the multiplicities of the weights in the $[2, 0, \ldots, 0, 2]$-representation necessarily match because the action is constructed directly from the $\mathfrak{sl}(n+1, \mathbb{R})$-valued building block $\mathcal{P}$. When taking the large volume limit, Eq. (11.4.3), a lot of information is lost (see, e.g., [208]) and it is therefore natural that the result of the compactification does not display the correct weight multiplicities. Thus, it is only after taking the non-perturbative completion, Eq. (11.4.2), that we can expect to reproduce correctly the weight multiplicities of the representation $[2, 0, \ldots, 0, 2]$.

12

Instanton Corrections to the Universal Hypermultiplet

The analysis in **Part II** of this thesis has so far been restricted to string compactifications on tori, in which case maximal supersymmetry is preserved, giving rise to moduli spaces described by symmetric spaces $\mathcal{M} = G/K$, where G is a global continuous symmetry of the classical effective action. These moduli spaces are known as *rigid*, since they allow no global deformations that preserve the holonomy K (i.e. is compatible with supersymmetry) except for an overall additional quotient by a discrete group $G(\mathbb{Z}) \subset G(\mathbb{R})$ [25]. Hence, for these cases the exact quantum moduli space must be

$$\mathcal{M}_{\text{exact}} = G(\mathbb{Z})\backslash G(\mathbb{R})/K, \tag{12.0.1}$$

where $G(\mathbb{Z})$ is the U-duality group [8]. This implies that all quantum corrections to the effective action are encoded in the duality group $G(\mathbb{Z})$, leading to powerful techniques for summing up perturbative and non-perturbative corrections in terms of $G(\mathbb{Z})$-invariant automorphic forms, as discussed in Chapter 8. The purpose of the present chapter is to analyze cases where the internal manifold is more complicated that a torus, for which the moduli space is no longer given by a symmetric space. In particular, we shall focus on the special case of type IIA compactifications on *rigid* Calabi-Yau threefolds in which case we propose that a natural candidate for the underlying duality group $G(\mathbb{Z})$ is the *Picard modular group* $SU(2,1;\mathbb{Z}[i])$. We further construct an $SU(2,1;\mathbb{Z}[i])$-invariant Eisenstein series, and show that its Fourier expansion potentially reproduces the expected contributions from D2- and NS5-brane instantons. This chapter is based on **Paper VIII**, which is work in progress with Ling Bao, Claudia Colonnello, Axel Kleinschmidt and Bengt E. W. Nilsson.

Let us also mention that related results were recently obtained in **Paper VII**, which was written in collaboration with Boris Pioline. There it was proposed that for generic (non-rigid) Calabi-Yau compactifications, a candidate for the underlying duality group in $D = 4$ is $G(\mathbb{Z}) = SL(3,\mathbb{Z})$. This suggestion is complementary to the discussion in the present chapter since for generic Calabi-Yau manifolds, the Picard group is not expected to play any role, while on the other hand for rigid compactifications $SL(3,\mathbb{Z})$ is not relevant. A small part of the analysis of **Paper VII** was presented in Section 10.3.

In the following section we will give a general overview of the relevant techniques, after which we begin a more detailed analysis in Section 12.2.

12.1 Instanton Corrections to Quaternionic Moduli Spaces

As was briefly mentioned above, and discussed in Chapter 1, for compactifications on more complicated internal manifolds than tori, the moduli space $\mathcal{M}$ is typically not described in terms of a simple symmetric space, rendering it difficult to find the duality group $G(\mathbb{Z})$, if it exists. Particularly interesting examples are type II Calabi-Yau threefold compactifications, which give rise to moduli spaces corresponding to a product of a special Kähler manifold $\mathcal{M}_{\text{SK}}$ and a quaternionic-Kähler manifold $\mathcal{M}_{\text{QK}}$. Due to the small amount of supersymmetry preserved, these moduli space are subject to quantum corrections which deform the geometry. Recall from the discussion in Chapter 1 that the special Kähler manifold $\mathcal{M}_{\text{SK}}$ receives perturbative quantum corrections associated with the worldsheet coupling α', as well as non-perturbative corrections from worldsheet instantons, which are exponentially suppressed of order $e^{-1/\alpha'}$ in the weak-coupling limit $\alpha' \to \infty$ [163, 164]. Recall further that the four-dimensional dilaton e^{ϕ} belongs to the hypermultiplet moduli space $\mathcal{M}_{\text{QK}}$, implying that this moduli space is sensitive to perturbative corrections from the genus expansion in $g_s = e^{\phi}$. In addition, $\mathcal{M}_{\text{QK}}$ receives non-perturbative corrections from D-brane instantons [26]. In the type IIA picture, these instantons aries from Euclidean D2-branes wrapping special Lagrangian 3-cycles in the Calabi-Yau manifold X, while in type IIB the contributions are due to Euclidean Dp-branes, $p = -1, 1, 3, 5$, wrapping even cycles in X. In both type IIA and type IIB there are also instanton effects due to Euclidean NS5-branes wrapping all of X.

Special Kähler manifolds are described by a holomorphic prepotential F, and it is therefore possible to additively encode quantum corrections to the metric on $\mathcal{M}_{\text{SK}}$ in terms of F [20,21]. For the quaternionic-Kähler metric on $\mathcal{M}_{\text{QK}}$, however, no such natural description exists and one must resort to more sophisticated techniques to encode deformations of $\mathcal{M}_{\text{QK}}$. It has long been an outstanding problem to compute the instanton-corrected metric on $\mathcal{M}_{\text{QK}}$, in particular the contributions from NS5-brane instantons. Recently, however, considerable progress has been made in this endeavour by utilizing techniques from twistor theory. These developments will play an important role in this chapter and we shall here give a brief introductory overview. More details are given in Section 12.5.

The basic idea, developed in a series of papers [27–31, 213], is that deformations of the quaternionic moduli space $\mathcal{M}_{\text{QK}}$ can be uplifted to deformations of the associated *twistor space* $\mathcal{Z}$, a $\mathbb{C}P^1$-bundle over $\mathcal{M}_{\text{QK}}$. Quantum corrections are then encoded in local complex (Darboux) coordinates on $\mathcal{Z}$ together with their associated transition functions gluing together local patches in $\mathcal{Z}$. These complex coordinates are called *twistor lines*, and play a similar role for $\mathcal{M}_{\text{QK}}$ as the holomorphic prepotential for special Kähler manifolds. A crucial object in this story is the *contact potential* e^{Φ} which determines the Kähler potential on the twistor space $\mathcal{Z}$. As shown in [27, 28, 31], perturbative and non-perturbative quantum corrections can then be elegantly encoded in an $SL(2,\mathbb{Z})$-invariant completion of the contact potential e^{Φ}, in a very similar way as the corrections to the $\mathcal{R}^4$-term discussed in Chapter 8.

Despite these new techniques, an unsolved problem has been to incorporate the corrections from NS5-brane instantons which start to contribute in $D = 4$, in which case $SL(2,\mathbb{Z})$-invariance has proven insufficient. It would therefore be desirable to have access to a larger duality group $G(\mathbb{Z})$, such that the $G(\mathbb{Z})$-invariant completion of e^{Φ} also sums up all NS5-brane instanton corrections. For generic Calabi-Yau threefolds it is not known what the group $G(\mathbb{Z})$ might be, but in this chapter we shall see that for the simpler case of compactifications of *rigid* Calabi-Yau threefolds there is a very appealing candidate for the underlying duality group. As mentioned above, we argue that for rigid Calabi-Yau compactifications in type IIA string theory, the underlying duality group is given by the Picard modular group $SU(2,1;\mathbb{Z})$, which is discrete subgroup of $SU(2,1)$ defined over the Gaussian integers $\mathbb{Z}[i]$ rather than the standard integers $\mathbb{Z}$ (see, e.g., [214]).

To verify this claim, we construct an $SU(2,1;\mathbb{Z})$-invariant Eisenstein series $\mathcal{E}_s$ in the principal continuous representation of $SU(2,1)$, and we propose that Eisenstein series provides the desired completion of the contact potential $e^{\Phi(x^{\mu},z)}$, restricted to the north pole $z = 0$ of the twistor space $\mathcal{Z}$. The restriction to the north pole is needed since we do not yet have access to the associated completion of the transition functions required to reconstruct the global deformations of $\mathcal{Z}$. Nonetheless, as a first step to finding the exact moduli space, we show that for the special value $s = 3/2$ the Fourier expansion of $\mathcal{E}_s$ correctly reproduces the perturbative tree-level and one-loop corrections to the moduli space, as well as an infinite series of exponentially suppressed contributions, attributed to D2-brane and NS5-brane instantons. Despite these successes, we should also mention that our analysis presents some puzzles which have not yet been resolved. These are discussed in detail in their appropriate context in subsequent sections.

12.2 The Universal Sector of Type IIA on Calabi-Yau Threefolds

Let us now present the physical setting in which our analysis applies. We consider the low-energy effective theory arising from the compactification of type IIA string theory on a rigid Calabi-Yau threefold $\mathcal{X}$. The bosonic sector of this compactification correponds to $D = 4$ Maxwell-Einstein gravity coupled to the so called *universal hypermultiplet*, with moduli space $\mathcal{M}_{\text{UH}}$. In the analysis we will disregard the vector multiplets, which decouple and do not play any role in what follows.

We describe the further compactification of this theory on S^1, in which case the gravity multiplet gives rise to an additional moduli space $\mathcal{M}_{\text{GM}}$, which turns out to be identical to $\mathcal{M}_{\text{UH}}$. In this context we discuss the *c*-map which relates the degrees of freedom of the hypermultiplet sector to the degrees of freedom in the gravity sector.

12.2.1 Type IIA on Calabi-Yau Threefolds

Type IIA compactifications on a generic Calabi-Yau threefold X give rise to $\mathcal{N} = 2$ supergravity in four dimensions. This theory splits into three separate pieces: (1) the *gravity multiplet*, consisting of the $D = 4$ metric $g_{\mu\nu}$ and an abelian vector $\mathcal{A}_{\mu}$ known as the graviphoton; (2) the *vector multiplets*, consisting of $n_V = h_{1,1}(X)$ abelian vectors A^I_{μ} and $n_V = h_{1,1}(X)$ complex scalars Z^I; (3) and, finally, the *hypermultiplets*, consisting of $4n_H = 4(h_{2,1}(X) + 1)$ real scalars φ^i. The fermionic degrees of freedom of these multiplets will not be needed in the present work.

The vector multiplet moduli Z^I parametrize a complex $2n_V$-dimensional *special Kähler manifold* $\mathcal{M}_{\text{V}}$, while the hypermultiplet moduli φ^i parametrize a real $4n_H$-dimensional *quaternionic-Kähler manifold* $\mathcal{M}_{\text{H}}$. The total moduli space $\mathcal{M}(X)$ splits locally into a direct product

$$\mathcal{M}(X) = \mathcal{M}_{\text{V}} \times \mathcal{M}_{\text{H}}. \tag{12.2.1}$$

The four-dimensional dilaton e^ϕ belongs to the hypermultiplet moduli space, implying that $\mathcal{M}_{\text{H}}$ is sensitive to quantum corrections originating from the string genus expansion. At the same time, the vector multiplet moduli space $\mathcal{M}_{\text{V}}$ receives corrections which are higher order in α', as well as non-perturbative worldsheet instanton corrections which scale as $e^{-1/\alpha'}$. The quantum corrections to $\mathcal{M}_{\text{V}}$ can be conveniently encoded in corrections to the prepotential $F(Z)$, whose exact form is known [20].

The corrections to the hypermultiplet moduli space has, however, proven to be much more complicated to understand. This is partly due to the fact that there is no analogue of the prepotential for describing the geometry of quaternionic-Kähler manifolds. Moreover, since the geometry of $\mathcal{M}_{\text{H}}$ is sensitive to corrections associated with the string coupling $g_s = e^{-\phi}$, the moduli space of the hypermultiplets is expected to also receive contributions which are non-perturbative as $g_s \to 0$. More precisely, $\mathcal{M}_{\text{H}}$ receives instanton corrections from Euclidean D2-branes wrapping special Lagrangian submanifolds in X, as well as Euclidean NS5-branes wrapping all of X [26]. The former scales as e^{-1/g_s}, as is characteristic for D-brane instantons, while the latter contributions scale as e^{-1/g_s^2}, in accordance with the tension of NS5-branes. These quantum effects will be further discussed in Section 12.5.

12.2.2 Restricting to the Rigid Case

We shall now restrict to the case of interest for our analysis, namely when X is a rigid Calabi-Yau threefold, i.e. $h_{2,1}(X) = 0$. In the following, we denote by $\mathcal{X}$ a rigid Calabi-Yau threefold, while reserving the notation X for generic Calabi-Yau threefolds. In this case the cohomology group $H^3(\mathcal{X})$ simplifies considerably:

$$H^3(\mathcal{X}) = H^{3,0}(\mathcal{X}) + H^{0,3}(\mathcal{X}). \tag{12.2.2}$$

Since $H^{3,0}(\mathcal{X})$ and $H^{0,3}(\mathcal{X})$ are both one-dimensional, we have only two "universal" 3-cycles, $\mathcal{A}$ and $\mathcal{B}$, corresponding to the Hodge numbers $h_{3,0}(\mathcal{X}) = h_{0,3}(\mathcal{X}) = 1$. No restriction is imposed on the Kähler structure $h_{1,1}(\mathcal{X})$. Such Calabi-Yau manifolds have been discussed both in the mathematical (see, e.g., [215]) and the physics literature (see, e.g., [16, 17, 216, 217]).

A peculiarity of compactification on a rigid Calabi-Yau manifold $\mathcal{X}$ is that there exists no mirror manifold $\mathcal{Y}$, such that type IIA on $\mathcal{X}$ would be dual to type IIB on $\mathcal{Y}$. This is due to the fact that $\mathcal{Y}$ would necessarily have $h_{1,1}(\mathcal{Y}) = h_{2,1}(\mathcal{X}) = 0$, which is not possible for a Kähler manifold. Nevertheless, there exist dual non-geometric type IIB compactifications, which are purely described in terms of the conformal field theory (certain "Landau-Ginzburg" models) of the internal manifold [216]. Moreover, for generic Calabi-Yau manifolds X which admit K3-fibrations, the type IIA theory on X is dual to heterotic compactifications on $K3 \times T^2$ [218] (see also [189]). A curious feature of rigid Calabi-Yau compactifications is that no such dual heterotic description exists, since rigid Calabi-Yau threefolds do not admit

K3-fibrations.[1] For these reasons we shall in the present analysis solely restrict to the type IIA picture, where our analysis has a clear geometric interpretation in terms of the underlying rigid Calabi-Yau manifold $\mathcal{X}$.

The moduli space of type IIA on $\mathcal{X}$ is now given by

$$\mathcal{M}(\mathcal{X}) = \mathcal{M}_{\mathrm{V}} \times \mathcal{M}_{\mathrm{UH}}, \tag{12.2.3}$$

where the vector multiplet moduli space $\mathcal{M}_{\mathrm{V}}$ is unchanged, while the total hypermultiplet moduli space $\mathcal{M}_{\mathrm{UH}}$ is a real 4-dimensional quaternionic-Kähler manifold corresponding to the *universal hypermultiplet* [22]. $\mathcal{M}_{\mathrm{UH}}$ is universal in the sense that it does not depend on the details of the compactification manifold $\mathcal{X}$. A first important note is that $\mathcal{M}_{\mathrm{UH}}$ contains the four-dimensional dilaton e^{ϕ}, and so is still sensitive to stringy quantum effects. In addition, the ten-dimensional Ramond-Ramond 3-form $C_{(3)}$ gives rise to a universal complex scalar $C = \chi + i\tilde{\chi}$ in $D = 4$, associated with the reduction of $C_{(3)}$ along the universal holomorphic and anti-holomorphic 3-cycles $\mathcal{A}$ and $\mathcal{B}$, respectively. Finally, the reduction of the NS-NS 2-form $B_{(2)}$ gives rise to a 2-form $B_{\mu\nu}$ in $D = 4$ which can further be dualized to a third axionic scalar ψ.

It was shown in [22] that the four real scalars $\{e^{\phi}, \chi, \tilde{\chi}, \psi\}$ parametrize the coset space

$$\mathcal{M}_{\mathrm{UH}} = \frac{SU(2,1)}{SU(2) \times U(1)}, \tag{12.2.4}$$

which is a quaternionic manifold with holonomy given by $SU(2) \times U(1)$. Moreover, this coset space has the unusual feature of being quaternionic-Kähler as well as Kähler, a property not shared by the hypermultiplet moduli space $\mathcal{M}_{\mathrm{H}}$ appearing for type IIA compactifications on generic Calabi-Yau threefolds X. While the quaternionic-Kähler condition is preserved by supersymmetry, the Kähler property of $\mathcal{M}_{\mathrm{UH}}$ will generically be broken by quantum effects [219]. Hence, for this reason quantum corrections to $\mathcal{M}_{\mathrm{UH}}$ cannot be encoded directly in terms of corrections to the classical Kähler potential.

In the following we shall ignore the vector multiplets completely since they decouple from the hypermultiplet sector. However, the universal sector does not only contain the hypermultiplet degrees of freedom, but also the four-dimensional metric $g_{\mu\nu}$ and the graviphoton $\mathcal{A}_{\mu}$ of the gravity multiplet. Therefore, the bosonic part of the full universal sector is described by the standard Maxwell-Einstein action coupled to a non-linear sigma model with target space $\mathcal{M}_{\mathrm{UH}}$:

$$\mathcal{S}_4 = \int d^4x\sqrt{g}\Big[R - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \gamma_{ij}\partial_{\mu}\varphi^i\partial^{\mu}\varphi^j\Big], \tag{12.2.5}$$

where $F_{\mu\nu} = 2\partial_{[\mu}\mathcal{A}_{\nu]}$, and the metric on $\mathcal{M}_{\mathrm{UH}}$ is given by

$$ds^2_{\mathcal{M}_{\mathrm{UH}}} = \gamma_{ij}d\varphi^i d\varphi^j = d\phi^2 + e^{2\phi}\big(d\chi^2 + d\tilde{\chi}^2\big) + e^{4\phi}\big(d\psi + \chi d\tilde{\chi} - \tilde{\chi}d\chi\big)^2. \tag{12.2.6}$$

In Section 12.3.1 we show in detail how this metric is associated with the invariant metric on the Lie algebra $\mathfrak{su}(2,1)$.

[1] I am grateful to Nick Halmagyi and Boris Pioline for discussions on this point.

12.2.3 Reduction to $D = 3$ and the c-Map

Let us now consider the further reduction of this theory on a spacelike circle S^1 to $D = 3$, i.e. we want to analyze type IIA on $\mathcal{X} \times S^1$, where S^1 is the spacelike circle associated with the compact coordinate x^3. The resulting theory is also dual to M-theory on $\mathcal{X} \times T^2$, but in this chapter we will restrict to the type IIA picture.

Under the reduction on $\mathcal{X} \times S^1$, the moduli space $\mathcal{M}_{\text{UH}}$ of the universal hypermultiplet goes through the reduction unchanged, with the scalars parametrizing $\mathcal{M}_{\text{UH}}$ being taken to be independent of the S^1-direction, $\varphi^i(x^M) = \varphi^i(x^\mu, x^3) \equiv \varphi^i(x^\mu)$, where $x^\mu = (t, x^1, x^2)$ are the spacetime coordinates in $D = 3$.

On the other hand, the reduction of the Maxwell-Einstein sector gives rise to additional scalar fields in three dimensions. First of all, we have the new "dilatonic" scalar $e^{-U} \equiv R(S^1)$, parametrizing the radius of the S^1. From the Maxwell field $\mathcal{A}_\mu$ we obtain an axionic scalar $\zeta \equiv \mathcal{A}_3$, as well as a three-dimensional abelian vector $\mathcal{A}_\alpha$. This vector can in turn be dualized, and yields another axionic scalar $\tilde{\zeta}$. Finally, the four-dimensional metric gives rise to the Kaluza-Klein vector field $A_\alpha = g_{3\alpha}$ in $D = 3$. This vector can also be dualized into a scalar σ of axionic type. The four new scalars $\{e^U, \zeta, \tilde{\zeta}, \sigma\}$, arising from the reduction of the gravity multiplet, parametrize a moduli space $\mathcal{M}_{\text{GM}}$ which turns out to be identical to $\mathcal{M}_{\text{UH}}$:

$$\mathcal{M}_{\text{GM}} = SU(2,1)/\big(SU(2) \times U(1)\big). \tag{12.2.7}$$

Hence, the theory in $D = 3$ is described by the action

$$\mathcal{S}_3 = \int d^3x \sqrt{\mathrm{g}} \Big[R_3 - \tilde{\gamma}_{ij} \partial_\mu \tilde{\varphi}^i \partial^\mu \tilde{\varphi}^j - \gamma_{ij} \partial_\mu \varphi^i \partial^\mu \varphi^j \Big], \tag{12.2.8}$$

where $\mathrm{g}_{\mu\nu}$ is the $D = 3$ metric, and the scalars $\tilde{\varphi}^i$ parametrize the moduli space $\mathcal{M}_{\text{GM}}$:

$$ds^2_{\mathcal{M}_{\text{GM}}} = \tilde{\gamma}_{ij} d\tilde{\varphi}^i d\tilde{\varphi}^j = dU^2 + e^{2U}\big(d\zeta^2 + d\tilde{\zeta}^2\big) + e^{4U}\big(d\sigma + \zeta d\tilde{\zeta} - \tilde{\zeta} d\zeta\big)^2. \tag{12.2.9}$$

What we have seen here is a special case of the c-map [22], which is a classical relation between special Kähler manifolds and quaternionic-Kähler manifolds. More specifically, upon compactification of type IIA on $X \times S^1$, the complex $h_{1,1}$-dimensional special Kähler manifold $\mathcal{M}_{\text{V}}$, corresponding to the vector multiplet moduli space, gets enhanced to a real $4(h_{1,1}+4)$-dimensional quaternionic-Kähler manifold $\mathcal{M}_{\text{V}}^{\text{QK}}$. The total moduli space in $D = 3$ is therefore

$$\mathcal{M}_3(X) = \mathcal{M}_{\text{V}}^{\text{QK}} \times \mathcal{M}_{\text{H}}. \tag{12.2.10}$$

The new quaternionic moduli space $\mathcal{M}_{\text{V}}^{\text{QK}}$ is known as the *c-map* of the special Kähler manifold $\mathcal{M}_{\text{V}}$. Microscopically, this can be understood as T-duality on the S^1, where the moduli space $\mathcal{M}_{\text{V}}^{\text{QK}}$ is interpreted as the hypermultiplet moduli space $\mathcal{M}_{\text{H}}^{\text{IIB}}$ of type IIB on $X \times \tilde{S}^1$, where $\tilde{S}^1$ is the T-dual circle. The radius $R(S^1) = e^{-U}$ of S^1 gets mapped to the (inverse) dilaton $e^{-\phi_{\text{IIB}}}$ on the IIB side. Hence, the large radius limit $e^{-U} \to \infty$ in type IIA on $X \times S^1$ corresponds to the weak-coupling limit $e^{\phi_{\text{IIB}}} \to 0$ in type IIB on $X \times \tilde{S}^1$. Similarly, the hypermultiplet moduli space $\mathcal{M}_{\text{H}}$ on the IIA side lies in the image of the c-map of the vector multiplet moduli space $\mathcal{M}_{\text{V}}^{\text{IIB}}$ of type IIB compactified on the same Calabi-Yau X, in which case the type IIA dilaton e^{ϕ} gets mapped to the (inverse) radius $R^{-1}(\tilde{S}^1)$ of the T-dual circle

$\tilde{S}^1$. Also here we see that after the c-map the low-energy limit $g_s = e^\phi \to 0$ in type IIA corresponds to the large radius limit $R(\tilde{S}^1) \to \infty$ in type IIB on $X \times \tilde{S}^1$.

For the special case of interest in our present analysis, namely the universal sector arising from compactification on $\mathcal{X} \times S^1$, the low-energy effective action is self-dual under the c-map [22,220]. This is evident by inspection of the three-dimensional action $\mathcal{S}_3$ in Eq. (12.2.8): the action is manifestly invariant under the simultaneous exchange of the degrees of freedom in $\mathcal{M}_{\text{GM}}$ and $\mathcal{M}_{\text{UH}}$ according to:

$$c - \text{map} \; : \; \{e^U, \zeta, \tilde{\zeta}, \sigma\} \;\longleftrightarrow\; \{e^\phi, \chi, \tilde{\chi}, \psi\}, \tag{12.2.11}$$

Note that under this map, the dilaton e^ϕ is mapped to the inverse radius e^U such that the weak-coupling limit $e^\phi \to 0$ indeed corresponds to the large radius limit $e^{-U} \to \infty$ as discussed above.

Effectively, the c-map can be used as a kind of solution-generating technique [220,221]. For example, a BPS black hole solution to the gravity multiplet sector in $D = 4$ is mapped under the c-map to a certain D-brane instanton solution to the hypermultiplet sector. For example, the extremal Reissner-Nordström solution is dual under the c-map to a D2-brane instanton arising from a Euclidean D2-brane wrapping the holomorphic 3-cycle $\mathcal{A}$ in $\mathcal{X}$ (or, equivalently, the $\mathcal{B}$-cycle). Similarly, a Taub-NUT solution on the gravity side is dual to an NS5-brane instanton in the hypermultiplet sector. In the three-dimensional picture, Euclidean BPS solutions in the gravity sector will wrap the S^1 and appear as gravitational instanton effects in $D = 3$. Under the c-map these instantons are then identified with the instanton effects in the universal hypermultiplet.

12.3 On the Coset Space $SU(2,1)/(SU(2) \times U(1)$

This section introduces our conventions for the group $SU(2,1)$ and a convenient parametrization of the coset space $SU(2,1)/(SU(2) \times U(1))$. We will also discuss the isomorphism between $SU(2,1)/(SU(2) \times U(1))$ and complex hyperbolic space $\mathbb{CH}^2$. Finally, we introduce the Picard modular group $SU(2,1;\mathbb{Z}[i])$ which acts as a modular group on $\mathbb{CH}^2$.

12.3.1 The Group $SU(2,1)$ and its Lie Algebra $\mathfrak{su}(2,1)$

The Lie group $SU(2,1)$ is defined as a subgroup of the group $GL(3,\mathbb{C})$ of invertible (3×3) complex matrices via

$$SU(2,1) = \Big\{ g \in GL(3,\mathbb{C}) \; : \; g^\dagger \eta g = \eta \; \text{ and } \det(g) = 1 \Big\} \,. \tag{12.3.1}$$

Here, the defining metric η is given by

$$\eta = \begin{pmatrix} 0 & 0 & -i \\ 0 & 1 & 0 \\ i & 0 & 0 \end{pmatrix} \tag{12.3.2}$$

and has signature $(++-)$. We note that the condition $g^\dagger \eta g = \eta$ already implies $|\det(g)| = 1$ and so we can also think of $SU(2,1)$ as the set of unitary matrices $U(2,1)$ modulo a pure phase, $SU(2,1) \cong PU(2,1)$, with the projectivization P referring to the equivalence relation $g \sim ge^{i\alpha}$ for $\alpha \in [0,2\pi)$. The diagonal matrices $e^{i\alpha}\mathrm{diag}(1,1,1)$ form the center of the group $U(2,1)$.

The Lie group $SU(2,1)$ as defined in (12.3.1) has as Lie algebra

$$\mathfrak{su}(2,1) = \left\{ X \in \mathfrak{gl}(3,\mathbb{C}) \,:\, X^\dagger \eta + \eta X = 0 \text{ and } \mathrm{tr}(X) = 0 \right\} . \tag{12.3.3}$$

It is a Lie algebra of real dimension 8 and a particular real form of $\mathfrak{sl}(3,\mathbb{C})$. It consists of four compact and four non-compact generators, the maximal real torus is one-dimensional. Since we will have ample opportunity to refer to specific generators we define the non-compact and compact Cartan generators

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} , \quad J = \begin{pmatrix} i & 0 & 0 \\ 0 & -2i & 0 \\ 0 & 0 & i \end{pmatrix} , \tag{12.3.4}$$

the positive step operators

$$X_1 = \begin{pmatrix} 0 & -1+i & 0 \\ 0 & 0 & 1-i \\ 0 & 0 & 0 \end{pmatrix} , \quad \tilde{X}_1 = \begin{pmatrix} 0 & 1+i & 0 \\ 0 & 0 & 1+i \\ 0 & 0 & 0 \end{pmatrix} , \quad X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \tag{12.3.5}$$

and the negative step operators

$$Y_{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 1+i & 0 & 0 \\ 0 & -1-i & 0 \end{pmatrix} , \quad \tilde{Y}_{-1} = \begin{pmatrix} 0 & 0 & 0 \\ -1+i & 0 & 0 \\ 0 & -1+i & 0 \end{pmatrix} , \quad Y_{-2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} . \tag{12.3.6}$$

The subscript refers to the eigenvalue under the adjoint action of the non-compact Cartan generator H, e.g. $[H, X_1] = X_1$ — the adjoint action of the compact Cartan generator J is not diagonalisable over the real numbers. Furthermore, the generators satisfy

$$X_2 = -\frac{1}{4}\left[X_1, \tilde{X}_1\right] , \tag{12.3.7}$$

such that the positive step operators form a Heisenberg algebra. Furthermore, the negative step operators Y are minus the Hermitian conjugate of the positive step operator X.

The Lie algebra $\mathfrak{su}(2,1)$ has a natural five grading by the generator H as a direct sum of vector spaces

$$\mathfrak{su}(2,1) = \mathfrak{g}_{-2} \oplus \mathfrak{g}_{-1} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2 , \tag{12.3.8}$$

with

$$\mathfrak{g}_{-2} = \mathbb{R}Y_{-2} ,\, \mathfrak{g}_{-1} = \mathbb{R}Y_{-1} \oplus \mathbb{R}\tilde{Y}_{-1} ,\, \mathfrak{g}_0 = \mathbb{R}H \oplus \mathbb{R}J ,\, \mathfrak{g}_1 = \mathbb{R}X_1 \oplus \mathbb{R}\tilde{X}_1 ,\, \mathfrak{g}_2 = \mathbb{R}X_2 . \tag{12.3.9}$$

One sees that the H-eigenspaces with eigenvalue ± 1 are degenerate. This is a characteristic feature of the reduced root system BC_1 underlying the real form $\mathfrak{su}(2,1)$ of $\mathfrak{sl}(3,\mathbb{C})$. There is a single root α since the real rank of $\mathfrak{su}(2,1)$ is one, and there are non-trivial root spaces

$\mathfrak{g}_1$ and $\mathfrak{g}_2$ corresponding to α and 2α, respectively.[2] The $\mathfrak{sl}(2,\mathbb{R})$ subalgebra associated with the 2α root space is canonically normalised and can be given a standard basis for example with H, $E=-X_2$ and $F=Y_{-2}$, so that $[E,F]=H$.

The Iwasawa decomposition of the Lie algebra $\mathfrak{su}(2,1)$ reads

$$\mathfrak{su}(2,1)=\mathfrak{n}_+\oplus\mathfrak{a}\oplus\mathfrak{k}, \tag{12.3.10}$$

where the non-compact (abelian) Cartan subalgebra $\mathfrak{a}=\mathbb{R}H$ while the nilpotent subspace $\mathfrak{n}_+=\mathbb{R}X_1\oplus\mathbb{R}\tilde{X}_1\oplus X_2$ is spanned by the positive step operators. The compact subalgebra of $\mathfrak{su}(2,1)$ is $\mathfrak{k}=\mathfrak{su}(2)\oplus\mathfrak{u}(1)$ as a direct sum of Lie algebras.[3] The generators of $\mathfrak{su}(2)$ and $\mathfrak{u}(1)$ are given explicitly by the anti-Hermitian matrices

$$\begin{aligned}\hat{K}_1 &= \frac{1}{4}\left(X_1+Y_{-1}\right), \quad \hat{K}_2=\frac{1}{4}\left(\tilde{X}_1+\tilde{Y}_{-1}\right), \quad \hat{K}_3=\frac{1}{4}\left(X_2+Y_{-2}+J\right),\\ \hat{J} &= \frac{3}{4}\left(X_2+Y_{-2}\right)-\frac{1}{4}J\,. \end{aligned} \tag{12.3.11}$$

These satisfy $\left[\hat{J},\hat{K}_i\right]=0$ and $\left[\hat{K}_i,\hat{K}_j\right]=-\epsilon_{ijk}\hat{K}_k$.

12.3.2 Complex Hyperbolic Space

The group $SU(2,1)$ has a natural action on complex hyperbolic two-space which we model here using the unbounded hyperquadric model [214] of complex dimension two

$$\mathbb{CH}^2=\left\{\mathcal{Z}=(z_1,z_2)\in\mathbb{C}^2\ :\ \mathcal{F}(\mathcal{Z})>0\right\}, \tag{12.3.12}$$

where the function $\mathcal{F}:\mathbb{C}^2\to\mathbb{R}$ is defined by

$$\mathcal{F}(\mathcal{Z}):=\Im(z_1)-\frac{1}{2}|z_2|^2>0\,, \tag{12.3.13}$$

and is sometimes referred to as height function. The condition $\mathcal{F}(\mathcal{Z})>0$ is the generalization of the usual condition $\Im(\tau)>0$ for a complex number τ to lie on the upper half plane of complex dimension one, and we see that for $z_2=0$ the space $\mathbb{CH}^2$ contains the usual upper half plane. We will refer to the space $\mathbb{CH}^2$ defined in (12.3.12) as the *complex hyperbolic space*, or the *complex upper half plane*. The link to the upper half plane will be made more explicitly below in Section 12.3.3.

The action of $SU(2,1)$ on $\mathcal{Z}\in\mathbb{CH}^2$ can now be written in the compact fractional linear form

$$g\cdot\mathcal{Z}=\frac{A\mathcal{Z}+B}{C\mathcal{Z}+D}\qquad\text{for}\quad g=\begin{pmatrix}A & B\\ C & D\end{pmatrix}, \tag{12.3.14}$$

where the blocks A, B, C and D have the sizes (2×2), (2×1), (1×2) and (1×1), respectively, so that the denominator is a complex number. In order to verify that (12.3.14) defines an

[2] A discussion of the restricted root system can for example be found in [94].

[3] By contrast, the Iwasawa decomposition (12.3.10) is only a direct sum of vector spaces and not of Lie algebras.

action of $SU(2,1)$ on complex hyperbolic space one needs to check that it preserves the defining condition in (12.3.12). This can be seen from computing

$$\mathcal{F}(g\cdot\mathcal{Z}) = \frac{\mathcal{F}(\mathcal{Z})}{|C\mathcal{Z}+D|^2}\,, \tag{12.3.15}$$

which again is a generalization of the usual $SL(2,\mathbb{Z})$-transformation of $\Im(\tau)$ on the (real) upper half plane. Furthermore, one needs to check the group property of the action which is straight-forward. In fact, when verifying (12.3.15) one only requires the condition $g^\dagger\eta g = \eta$ so that (12.3.14) defines an action of all of $U(2,1)$ on complex hyperbolic two-space. Since elements from the center act trivially, one can restrict to $PU(2,1) \cong SU(2,1)$ to obtain a simply transitive action.

The Kähler metric on $\mathbb{CH}^2$ can be written in terms of the following Kähler potential

$$K(\mathcal{Z}) = -\log\mathcal{F}(\mathcal{Z})\,. \tag{12.3.16}$$

Written out in terms of the two complex coordinates $\mathcal{Z} = (z_1, z_2)$ this gives the Euclidean metric

$$ds^2 = \frac{1}{4}\mathcal{F}^{-2}\left[dz_1d\bar{z}_1 + iz_2dz_1d\bar{z}_2 - i\bar{z}_2dz_2d\bar{z}_1 + 2\Im(z_1)dz_2d\bar{z}_2\right]\,. \tag{12.3.17}$$

The group $SU(2,1)$ acts isometrically on $\mathbb{CH}^2$.

12.3.3 Relation to the Scalar Coset Manifold $SU(2,1)/(SU(2)\times U(1))$

The complex hyperbolic upper half plane is also isomorphic to (a connected component) of the Hermitian symmetric space

$$\mathbb{CH}^2 \cong SU(2,1)/(SU(2)\times U(1))\,, \tag{12.3.18}$$

where the right hand side should properly be restricted to the connected component of the identity. The Hermitian symmetric space is of real dimension four and can be parametrized by four real variables $\{\phi,\chi,\tilde{\chi},\psi\}$ in triangular gauge, using the Iwasawa decomposition (12.3.10), as

$$\mathcal{V} = e^{\chi X_1+\tilde{\chi}\tilde{X}_1+2\psi X_2}e^{-\phi H} = \begin{pmatrix} e^{-\phi} & \tilde{\chi}-\chi+i(\chi+\tilde{\chi}) & e^{\phi}\left(2\psi+i(\chi^2+\tilde{\chi}^2)\right) \\ 0 & 1 & e^{\phi}\left(\chi+\tilde{\chi}+i(\tilde{\chi}-\chi)\right) \\ 0 & 0 & e^{\phi} \end{pmatrix}. \tag{12.3.19}$$

The symmetric space is a right coset in our conventions and the element $\mathcal{V}$ transforms as $\mathcal{V}\to g\mathcal{V}k^{-1}$ with $g\in SU(2,1)$ and $k\in SU(2)\times U(1)$. The four scalar fields can take arbitrary real values.

It is convenient to define the Hermitian matrix

$$\mathcal{K} = \mathcal{V}\mathcal{V}^\dagger \tag{12.3.20}$$

that transforms as $\mathcal{K}\to g\mathcal{K}g^\dagger$ under the action of $g\in SU(2,1)$. Explicitly, this matrix reads

$$\mathcal{K} = \begin{pmatrix} e^{-2\phi}+|\lambda|^2+e^{2\phi}|\gamma|^2 & i\bar{\lambda}+e^{2\phi}\bar{\lambda}\gamma & e^{2\phi}\gamma \\ -i\lambda+e^{2\phi}\lambda\bar{\gamma} & 1+e^{2\phi}|\lambda|^2 & e^{2\phi}\lambda \\ e^{2\phi}\bar{\gamma} & e^{2\phi}\bar{\lambda} & e^{2\phi} \end{pmatrix}, \tag{12.3.21}$$

where, for later convenience, we defined the complex variables

$$\lambda := \chi + \tilde{\chi} + i(\tilde{\chi} - \chi), \qquad \gamma := 2\psi + \frac{i}{2}|\lambda|^2. \tag{12.3.22}$$

From $\mathcal{K}$ one can define the metric on the symmetric space via

$$ds^2 = -\frac{1}{8}\mathrm{tr}\left(d\mathcal{K}\, d(\mathcal{K}^{-1})\right) = \frac{1}{8}\mathrm{tr}\left(\mathcal{V}^{-1}d\mathcal{V} + (\mathcal{V}^{-1}d\mathcal{V})^\dagger\right)^2 . \tag{12.3.23}$$

Working this out for the coset element (12.3.19) one finds the following $SU(2,1)$ invariant metric

$$ds^2 = d\phi^2 + e^{2\phi}(d\chi^2 + d\tilde{\chi}^2) + e^{4\phi}(d\psi + \chi d\tilde{\chi} - \tilde{\chi}d\chi)^2 . \tag{12.3.24}$$

Comparing (12.3.24) to (12.3.17) leads to the identification

$$\begin{aligned} z_1 &= 2\psi + i\left(e^{-2\phi} + \frac{1}{2}|z_2|^2\right) = 2\psi + i\left(e^{-2\phi} + \chi^2 + \tilde{\chi}^2\right) , \\ z_2 &= \chi + \tilde{\chi} + i(\tilde{\chi} - \chi) . \end{aligned} \tag{12.3.25}$$

Note that $z_1 = \gamma + ie^{-2\phi}$ and $z_2 = \lambda$. In Section 12.4 it will prove to be convenient to use the complex variable γ rather than z_1.

We see that the condition $\mathcal{F}(\mathcal{Z}) > 0$ is satisfied in the parametrization (12.3.25) of $\mathcal{Z}$ since $\mathcal{F}(\mathcal{Z}) = e^{-2\phi}$. The advantage of using the complex hyperbolic upper half plane rather than the scalar coset in the form (12.3.19) is that the action of $SU(2,1)$ is simpler to evaluate. The action of $g \in SU(2,1)$ on $\mathcal{V}$ requires a compensating transformation $k \in SU(2) \times U(1)$ to restore the triangular gauge chosen in (12.3.19) whenever g is itself not of triangular form. Finding this compensating transformation in general can be involved. Since it has been implicitly carried out in the non-linear action (12.3.14) one does not require the precise form of the compensator. We note that setting $z_2 = 0$ gives back the real upper half plane but the parametrization (12.3.25) appears to have additional factors of two compared to the usual expressions at first sight. However, these factors have a physical significance as will be discussed further below. Besides being a quaternionic Kähler manifold, $SU(2,1)/(SU(2) \times U(1))$ also has the stronger property of being Kähler, as is clear from (12.3.16).

In the variables $\mathcal{Z} = (z_1, z_2)$ given by (12.3.25), the matrix $\mathcal{K}$ of (12.3.20) takes the simple form

$$\mathcal{K} = \tilde{\mathcal{K}} + \eta , \tag{12.3.26}$$

where η is the defining matrix of $SU(2,1)$ given in (12.3.2) and

$$\tilde{\mathcal{K}} = e^{2\phi}\begin{pmatrix} |z_1|^2 & z_1\bar{z}_2 & z_1 \\ \bar{z}_1 z_2 & |z_2|^2 & z_2 \\ \bar{z}_1 & \bar{z}_2 & 1 \end{pmatrix} , \tag{12.3.27}$$

bearing in mind that $e^{2\phi} = \mathcal{F}(\mathcal{Z})^{-1}$.

12.3.4 Coset Transformations and Subgroups of $SU(2,1)$

We now study the effect of some particular elements of $SU(2,1)$ on complex hyperbolic two-space. The specific transformations we investigate are the ones with an immediate physical interpretation.

Heisenberg Translations

Let N denote the exponential of the nilpotent algebra of positive step operators $\mathfrak{n}_+$. We define the following elements of N

$$T_1 = \begin{pmatrix} 1 & -1+i & i \\ 0 & 1 & 1-i \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{T}_1 = \begin{pmatrix} 1 & 1+i & i \\ 0 & 1 & 1+i \\ 0 & 0 & 1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{12.3.28}$$

These are such that $T_1 = \exp(X_1)$ etc. Any element $n \in N$ then can be written as

$$\begin{aligned} n &= (T_1)^a(\tilde{T}_1)^b(T_2)^{c+2ab} = e^{aX_1+b\tilde{X}_1+cX_2} \\ &= \begin{pmatrix} 1 & a(-1+i)+b(1+i) & c+i(a^2+b^2) \\ 0 & 1 & a(1-i)+b(1+i) \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{12.3.29}$$

for $a, b, c \in \mathbb{R}$. The effect of this transformation on $\mathcal{Z} = (z_1, z_2)$ is

$$\begin{aligned} z_1 &\longmapsto z_1 + \big[a(-1+i)+b(1+i)\big]z_2 + c + i(a^2+b^2)\,, \\ z_2 &\longmapsto z_2 + a(1-i) + b(1+i)\,, \end{aligned} \tag{12.3.30}$$

or in terms of the four scalars fields of (12.3.19)

$$\begin{aligned} \phi &\longmapsto \phi\,, \\ \chi &\longmapsto \chi + a\,, \\ \tilde{\chi} &\longmapsto \tilde{\chi} + b\,, \\ \psi &\longmapsto \psi + \frac{1}{2}c - a\tilde{\chi} + b\chi\,. \end{aligned} \tag{12.3.31}$$

The appearance of the shift parameters a and b in the transformation of ψ is due to the non-Abelian structure of $\mathfrak{n}_+$ given by the Heisenberg algebra (12.3.7). This effect is also evident in the first line of the expression (12.3.29) for the general element of N. From the point of view of the coset, the Heisenberg translations do not require any compensating transformation as they preserve the Borel gauge.

Rotations

Rotations are generated by the compact Cartan element J of $\mathfrak{su}(2,1)$ given in (12.3.4). Let

$$R = \exp(\pi J/2) = \begin{pmatrix} i & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & i \end{pmatrix}, \tag{12.3.32}$$

then the most general transformation of this type is given by R^m, for $m \in [0, 4)$, and acts on $\mathcal{Z} = (z_1, z_2)$ via

$$z_1 \to z_1\,, \quad z_2 \to e^{i\pi\sigma/2}z_2\,. \tag{12.3.33}$$

In terms of the four scalar fields this transformation reads

$$\begin{aligned}
\phi &\longmapsto \phi\,,\\
\chi &\longmapsto \cos(\pi\sigma/2)\chi - \sin(\pi\sigma/2)\tilde{\chi}\,,\\
\tilde{\chi} &\longmapsto \sin(\pi\sigma/2)\chi + \cos(\pi\sigma/2)\tilde{\chi}\,,\\
\psi &\longmapsto \psi
\end{aligned} \tag{12.3.34}$$

and so rotates the two scalars χ and $\tilde{\chi}$ among each other while leaving the other two invariant. The compensating transformation to restore the Borel gauge for the coset element (12.3.19) is $k = R^m$.

Involution

The last transformation of interest is the following involution

$$S = \begin{pmatrix} 0 & 0 & i \\ 0 & -1 & 0 \\ -i & 0 & 0 \end{pmatrix}, \tag{12.3.35}$$

which acts on $\mathcal{Z} = (z_1, z_2)$ according to

$$z_1 \mapsto -\frac{1}{z_1}\,, \quad z_2 \mapsto -i\frac{z_2}{z_1}\,, \tag{12.3.36}$$

making it apparent that this is an inversion of z_1. For the real scalars themselves we find the following transformation

$$\begin{aligned}
\phi &\longmapsto -\frac{1}{2}\ln\left[\frac{e^{-2\phi}}{4\psi^2 + [e^{-2\phi} + \chi^2 + \tilde{\chi}^2]^2}\right],\\
\chi &\longmapsto \frac{2\psi\tilde{\chi} - (e^{-2\phi} + \chi^2 + \tilde{\chi}^2)\chi}{4\psi^2 + [e^{-2\phi} + \chi^2 + \tilde{\chi}^2]^2}\,,\\
\tilde{\chi} &\longmapsto \frac{2\psi\chi + (e^{-2\phi} + \chi^2 + \tilde{\chi}^2)\tilde{\chi}}{4\psi^2 + [e^{-2\phi} + \chi^2 + \tilde{\chi}^2]^2}\,,\\
\psi &\longmapsto -\frac{\psi}{4\psi^2 + [e^{-2\phi} + \chi^2 + \tilde{\chi}^2]^2}\,.
\end{aligned} \tag{12.3.37}$$

It is straightforward to check that the required compensating transformation in this case indeed belongs to the maximal compact subgroup $SU(2)\times U(1)$.

We note that the involution (12.3.35) can also be written as

$$S = e^{Y_{-2}} e^{X_2} e^{Y_{-2}} R\,, \tag{12.3.38}$$

using the rotation matrix (12.3.32). The expression (12.3.38) is almost the expression of Kac [34] for Weyl reflections (see Lemma 3.8 in [34]). There is a difference in our expression for S compared to [34] since we are dealing with a non-split real form. Firstly, one should take the generators corresponding to the canonically normalised, split $\mathfrak{sl}(2,\mathbb{R}) \subset \mathfrak{su}(2,1)$, that is, use the generators Y_{-2} and $-X_2$ of the graded decomposition (12.3.8). Secondly, one has to verify the action of the resulting transformation on the degree one subspaces in (12.3.8) that are degenerate. Demanding that no additional transformation on their basis is introduced requires the inclusion of R in (12.3.38) since then $SX_1S^{-1} = Y_{-1}$ (and not $\tilde{Y}_{-1}$). We also note that without the inclusion of R the transformation (12.3.38) would be of order four and not two as required for a reflection. The reduced root system BC_1 of $\mathfrak{su}(2,1)$ is of (real) rank one and therefore its Weyl group is generated by a single reflection, namely the one displayed in (12.3.35). Therefore one should call the Weyl group

$$\mathcal{W}(\mathfrak{su}(2,1)) = \mathcal{W}(BC_1) \cong \mathbb{Z}_2\,. \tag{12.3.39}$$

In the context of U-duality symmetries for torus compactifications it is often the Weyl group that is retained as a minimal discrete symmetry group acting on BPS states [11,222]. In the present case this would correspond to studying the action of the involution (12.3.35) on BPS states.

Let us also briefly discuss various subgroups of $SU(2,1)$. The canonically normalised $SL(2,\mathbb{R})$ subgroup mentioned in mentioned in Section 12.3.1 consists of matrices of the type

$$SL(2,\mathbb{R}) = \left\{ \begin{pmatrix} a & 0 & b \\ 0 & 1 & 0 \\ c & 0 & d \end{pmatrix} \,:\, a,b,c,d \in \mathbb{R} \text{ and } ad-bc=1 \right\} \subset SU(2,1)\,. \tag{12.3.40}$$

Restricting complex hyperbolic space to the slice $z_2 = 0$, one sees that this condition is left invariant by this $SL(2,\mathbb{R})$ subgroup which furthermore acts in the standard non-linear fashion on the variable z_1.

Another $SL(2,\mathbb{R})$ subgroup of $SU(2,1)$ consists of the following matrices

$$SL(2,\mathbb{R}) = \left\{ \begin{pmatrix} a^2 & (-1+i)ab & ib^2 \\ (-1-i)ac & ad+bc & (1-i)bd \\ -ic^2 & (1+i)cd & d^2 \end{pmatrix} \,:\, a,b,c,d \in \mathbb{R} \text{ and } ad-bc=1 \right\}. \tag{12.3.41}$$

Unlike (12.3.40), this group is not regularly embedded in the sense that the restriction of the canonical bilinear form of $\mathfrak{su}(2,1)$ to this subgroup does not coincide with the canonical bilinear form on $\mathfrak{sl}(2;\mathbb{R})$ but rather differs by a factor four. A consequence of the non-regularity of the embedding is that all entries in (12.3.41) are quadratic functions of the $SL(2,\mathbb{R})$ matrix in the fundamental representation.

The relevance of the subgroup (12.3.41) is the following. The slice $\tilde{\chi} = \psi = 0$ can be parametrized by a complex variable $\tau := \chi + ie^{-\phi}$ that lives on the standard upper half plane. On this parameter τ the group (12.3.41) can be seen to act in the usual fractional linear fashion. Note also that the standard S transformation of $SL(2,\mathbb{Z})$ expressed in terms of the parametrization above corresponds to the involution (12.3.35). Turning on the transverse coordinates again, the action on these is more complicated. The parameter τ relevant for this subgroup is composed out of the dilaton and a RR scalar as one would expect in analogy with the IIB case. It is this parameter that appears for example in the D2-instanton corrections to type IIA studied in [28], obtained via a mirror map of the D(-1)-instanton corrrections in type IIB. The two subgroups (12.3.40) and (12.3.41) together generate the whole of $SU(2,1)$.

12.3.5 The Picard Modular Group

We finally discuss the Picard modular group $SU(2,1;\mathbb{Z}[i])$. This group can be defined as the intersection [214]

$$SU(2,1;\mathbb{Z}[i]) := SU(2,1)\cap SL(3,\mathbb{Z}[i]), \tag{12.3.42}$$

where $\mathbb{Z}[i]$ denotes the Gaussian integers

$$\mathbb{Z}[i]=\{z\in\mathbb{C}\,:\,\Re(z),\Im(z)\in\mathbb{Z}\}. \tag{12.3.43}$$

This definition implies that any element $g\in SU(2,1)$ which has only Gaussian integer matrix entries belongs to $SU(2,1;\mathbb{Z}[i])$. In view of the discussion of $PU(2,1)\cong SU(2,1)$ the Picard modular group can also be called $PU(2,1;\mathbb{Z}[i])$.[4]

Let us now examine the particular $SU(2,1)$-transformations of the previous subsection to check whether they belong to the Picard group. The Heisenberg group $N\subset SU(2,1)$ contains a subgroup $N(\mathbb{Z}):=N\cap SU(2,1;\mathbb{Z}[i])$. By inspection of Eq. (12.3.29) we see that $N(\mathbb{Z})$ must be of the form

$$N(\mathbb{Z})=\left\{e^{aX_1+b\tilde{X}_1+cX_2}\;:\;a,b,c\in\mathbb{Z}\right\}. \tag{12.3.44}$$

In view of (12.3.29), a natural set of generators for $N(\mathbb{Z})$ is given by the three matrices in (12.3.28) T_1, $\tilde{T}_1$ and T_2. The action of these discrete shifts are then as given in (12.3.31) with parameters $a,b,c\in\mathbb{Z}$. The translations (12.3.28) are of infinite order in the Picard modular group.

The rotation R^m defined in (12.3.32) only is an element for the discrete values of the exponent $m=0,1,2,3$, and R is an element of order 4 in the Picard modular group. The action of R on the scalar fields is

$$R:\quad(\chi,\tilde{\chi})\longmapsto(-\tilde{\chi},\chi). \tag{12.3.45}$$

Physically speaking, this corresponds to electric-magnetic duality, which is expected to be preserved in the quantum theory [223].

Finally, we will examine the involution S in Eq. (12.3.35). Clearly, the involution is an element of the Picard modular group. The involution S is of order 2 in the Picard modular group. As already noted above, the involution (12.3.35) corresponds to the Weyl reflection of the restricted root system BC_1 of the non-split real form $\mathfrak{su}(2,1)$. The Weyl reflection is associated with the (long) root 2α. We can also give an interpretation to the rotation R. This is a transformation that rotates within the degenerate, two-dimensional α root space, spanned by the generators X_1 and $\tilde{X}_1$.

The Picard modular group acts discontinuously on the complex hyperbolic space $\mathbb{CH}^2$. A fundamental domain for its action has been given by Francsics and Lax in [214]. Recently, they have also proven the following [224][5]:

[4] The nomenclature "Picard group" is not unique, in fact our Picard group is a member of a family of similar groups $PSU(1,n+1;\mathbb{Z}[i])$ of which the case $n=0$, corresponding to $PSL(2,\mathbb{Z}[i])$ is also often called *the* Picard group. In this chapter we will always mean $SU(2,1;\mathbb{Z}[i])$ when speaking of the Picard group.

[5] I am very grateful to Gabor Francsics and Peter Lax for communicating this result prior to publication.

Theorem (Francsics & Lax): *The Picard modular group $SU(2,1;\mathbb{Z}[i])$ is generated by the translations T_1 and T_2, together with the rotation R and the involution S.*

Since the two translations T_1 and $\tilde{T}_1$ are related through "electric-magnetic duality" by $\tilde{T}_1 = RT_1R^{-1}$, one may equivalently choose either of the translations T_1 or $\tilde{T}_1$ associated with the α root space in the theorem. Since all three translations T_1, $\tilde{T}_1$ and T_2 will turn out to have a clear physical interpretation we present the Picard modular group as generated (non-minimally) by the following five elements:

$$T_1 = \begin{pmatrix} 1 & -1+i & i \\ 0 & 1 & 1-i \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{T}_1 = \begin{pmatrix} 1 & 1+i & i \\ 0 & 1 & 1+i \\ 0 & 0 & 1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$R = \begin{pmatrix} i & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & i \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 0 & i \\ 0 & -1 & 0 \\ -i & 0 & 0 \end{pmatrix}. \tag{12.3.46}$$

In accordance with the $SL(2,\mathbb{R})$ subgroup identified in (12.3.40) we note that there is an $SL(2,\mathbb{Z}) \subset SU(2,1;\mathbb{Z}[i])$ that acts on the slice $z_2 = 0$ of complex hyperbolic space as the usual modular group on the remaining variable z_1.

12.4 Eisenstein Series for the Picard Modular Group

In this section we shall construct an Eisenstein series for the Picard modular group in the principal continuous series representation of $SU(2,1)$. We show that this can be done in three different ways, which all mutually enlighten each other.

12.4.1 Lattice Construction and Quadratic Constraint

Our first method will be the one developed in [13]. A non-holomorphic Eisenstein series, of order s, on the quotient

$$SU(2,1;\mathbb{Z}[i])\backslash SU(2,1)/(SU(2)\times U(1)), \tag{12.4.1}$$

can then be constructed in the following way

$$\mathcal{E}_s(\mathcal{K}) := \sum_{\vec{\omega}\in\mathbb{Z}[i]^3} \delta\big(\vec{\omega}^\dagger \wedge \vec{\omega}\big)\Big[\vec{\omega}^\dagger \cdot \mathcal{K} \cdot \vec{\omega}\Big]^{-s}, \tag{12.4.2}$$

where $\mathcal{K} = \mathcal{V}\mathcal{V}^\dagger$ is the "generalized metric" which was constructed explicitly in Eq. (12.3.21), and $\delta(\vec{\omega}^\dagger \wedge \vec{\omega})$ is a quadratic constraint on the Gaussian integers $\vec{\omega} \in \mathbb{Z}[i]^3$ which ensures that the Eisenstein series is an eigenfunction of the Laplacian on the coset space $SU(2,1)/(SU(2)\times U(1))$ [13]. We will discuss this constraint in detail below. Moreover, it is always understood that the summation is restricted such that the vector $\vec{\omega} = (0,0,0)$ is excluded. Writing out

Eq. (12.4.2) explicitly yields

$$\mathcal{E}_s(\phi,\lambda,\gamma)=\sum_{\vec{\omega}\in\mathbb{Z}[i]^3}\delta\big(\vec{\omega}^\dagger\wedge\vec{\omega}\big)e^{-2s\phi}\Big[|\bar{\omega}_1+\bar{\omega}_2\bar{\lambda}+\bar{\omega}_3\bar{\gamma}|^2+e^{-2\phi}|\bar{\omega}_2-i\omega_3\lambda|^2+e^{-4\phi}|\omega_3|^2\Big]^{-s}. \tag{12.4.3}$$

The variables λ and γ were defined as functions of $\mathcal{Z}=(z_1,z_2)$ in (12.3.22). To better understand the quadratic constraint $\delta(\vec{\omega}^\dagger\wedge\vec{\omega})$, it is illuminating to utilize the isomorphism between the coset space $SU(2,1)/(SU(2)\times U(1))$ and the complex hyperbolic space $\mathbb{CH}^2$, as discussed in Section 12.3.3. We recall from (12.3.27) that in terms of the variable $\mathcal{Z}=(z_1,z_2)\in\mathbb{CH}^2$, the matrix $\mathcal{K}$ reads

$$\mathcal{K}=\tilde{\mathcal{K}}+\eta, \tag{12.4.4}$$

where η is the $SU(2,1)$-invariant metric, Eq. (12.3.2), and the matrix $\tilde{\mathcal{K}}$ is given by

$$\tilde{\mathcal{K}}=e^{2\phi}\begin{pmatrix}|z_1|^2 & z_1\bar{z}_2 & z_1\\ \bar{z}_1z_2 & |z_2|^2 & z_2\\ \bar{z}_1 & \bar{z}_2 & 1\end{pmatrix}. \tag{12.4.5}$$

In this new parametrization, the Eisenstein series becomes

$$\begin{aligned}\mathcal{E}_s(\mathcal{Z}) &= \sum_{\vec{\omega}\in\mathbb{Z}[i]^3}\delta\big(\vec{\omega}^\dagger\wedge\vec{\omega}\big)\Big[\vec{\omega}^\dagger\cdot\tilde{\mathcal{K}}\cdot\vec{\omega}+\vec{\omega}^\dagger\cdot\eta\cdot\vec{\omega}\Big]^{-s}\\ &= \sum_{\vec{\omega}\in\mathbb{Z}[i]^3}\delta\big(\vec{\omega}^\dagger\wedge\vec{\omega}\big)\Big[e^{2\phi}|\omega_1+\omega_2z_2+\omega_3z_1|^2+\vec{\omega}^\dagger\cdot\eta\cdot\vec{\omega}\Big]^{-s}.\end{aligned} \tag{12.4.6}$$

This decomposition of the Eisenstein series can be understood as follows [13]. By virtue of Eq. (12.3.19) the coset representative $\mathcal{V}\in SU(2,1)/(SU(2)\times U(1))$ is constructed in the fundamental representation $\mathcal{R}$ of $SU(2,1)$. This implies that the generalized metric $\mathcal{K}=\mathcal{V}\mathcal{V}^\dagger$ transforms in the symmetric tensor product $\mathcal{R}\otimes_s\mathcal{R}$. The contraction of $\mathcal{K}$ with the lattice vectors $\vec{\omega}$ is only an eigenfunction of the Laplacian on $SU(2,1)/(SU(2)\times U(1))$ if this contraction is irreducible. The split of the summand in Eq. (12.4.6) precisely shows that this is not the case; $\vec{\omega}^\dagger\cdot\mathcal{K}\cdot\vec{\omega}$ is reducible and its irreducible components are the two terms in Eq. (12.4.6). The second term is the trivial (moduli-independent) component of the tensor product, and must be projected out in order to obtain an eigenfunction of the Laplacian. The Laplacian will be discussed in detail in Section 12.4.2 below. By the reasoning above we find that the quadratic constraint $\delta(\vec{\omega}^\dagger\wedge\vec{\omega})$ must be

$$\vec{\omega}^\dagger\wedge\vec{\omega}:=\vec{\omega}^\dagger\cdot\eta\cdot\vec{\omega}=|\omega_2|^2-2\Im(\omega_1\bar{\omega}_3)=0. \tag{12.4.7}$$

The Eisenstein series then finally becomes

$$\mathcal{E}_s(\mathcal{Z})=\sum_{\vec{\omega}\in\mathbb{Z}[i]^3}\delta\big(\vec{\omega}^\dagger\cdot\eta\cdot\vec{\omega}\big)e^{-2s\phi}|\omega_1+\omega_2z_2+\omega_3z_1|^{-2s}. \tag{12.4.8}$$

It will also be convenient to have the equivalent form of the Eisenstein series in the original variables ϕ,λ and γ. This is of course just Eq. (12.4.3) with the explicit constraint inserted. Imposing the (real) quadratic constraint $|\omega_2|^2-2\Im(\omega_1\bar{\omega}_3)=0$ eliminates one of the summation variables, so that we are effectively summing over only five integers.

12.4.2 Poincaré Series on the Complex Upper Half Plane

As we have seen in Chapter 9, a standard way of constructing non-holomorphic Eisenstein series on a symmetric space G/K is in terms of *Poincaré series*. For example, for the case of $SL(2,\mathbb{R})/SO(2)$, parametrized by a complex coordinate τ, we have seen in Section 9.2 that such a Poincaré series is obtained by summing the function $\Im(\gamma\cdot\tau)$ over the orbit $\gamma\in\Gamma_\infty\backslash SL(2,\mathbb{Z})$, where Γ_∞ is generated by $T:\tau\mapsto\tau+1$. This indeed produces a non-holomorphic Eisenstein series on the double quotient $SL(2,\mathbb{Z})\backslash SL(2,\mathbb{R})/SO(2)$.

Here we would like to generalize this construction to a Poincaré series on the complex upper half plane $\mathbb{CH}^2$, parametrized by the variable $\mathcal{Z}=(z_1,z_2)$. The generalization of $\Im(\tau)$ is then given by the $N(\mathbb{Z})$-invariant function $\mathcal{F}(\mathcal{Z})$, constructed in (12.3.13) [225].[6] The invariance of $\mathcal{F}(\mathcal{Z})$ under $N(\mathbb{Z})$ can be checked by direct substitution of the Heisenberg translations in Eq. (12.3.30). As we have seen in Section 12.3, the Picard modular group $SU(2,1;\mathbb{Z}[i])$ acts by fractional transformations on $\mathcal{Z}\in\mathbb{CH}^2$ such that the function $\mathcal{F}(\mathcal{Z})$ transforms as

$$\mathcal{F}(\gamma\cdot\mathcal{Z})=\frac{\mathcal{F}(\mathcal{Z})}{|C\mathcal{Z}+D|^2},\qquad \gamma=\begin{pmatrix} A & B\\ C & D\end{pmatrix}\in SU(2,1;\mathbb{Z}[i]). \tag{12.4.9}$$

A Poincaré series for the Picard group may now be constructed as follows

$$\mathcal{P}_s(\mathcal{Z}):=\sum_{\gamma\in N(\mathbb{Z})\backslash SU(2,1;\mathbb{Z}[i])}\mathcal{F}(\gamma\cdot\mathcal{Z})^s=\sum_{\gamma\in N(\mathbb{Z})\backslash SU(2,1;\mathbb{Z}[i])}\Big(\frac{\mathcal{F}(\mathcal{Z})}{|C\mathcal{Z}+D|^2}\Big)^s. \tag{12.4.10}$$

Taking $C\equiv(\omega_3,\omega_2)\in\mathbb{Z}[i]^2$ and $D\equiv\omega_1\in\mathbb{Z}[i]$, and recalling $\mathcal{F}(\mathcal{Z})=e^{-2\phi}$, then reproduces the same form of the Eisenstein series as in Eq. (12.4.8), i.e.

$$\mathcal{P}_s(\mathcal{Z})=\sum_{\gamma\in N(\mathbb{Z})\backslash SU(2,1;\mathbb{Z}[i])}e^{-2s\phi}|\omega_1+\omega_2z_2+\omega_3z_1|^{-2s}, \tag{12.4.11}$$

which agrees with (12.4.8) modulo a possible discrepancy in the range of summation.

Let us now also discuss the Laplacian condition on $\mathcal{P}_s$. The Laplacian on the coset space $\mathbb{CH}^2=SU(2,1)/(SU(2)\times U(1))$ is most conveniently written in terms of the real variables $\{y=e^{-2\phi},\chi,\tilde{\chi},\psi\}$ and reads explicitly [226]

$$\Delta_{\mathbb{CH}^2}=\frac{1}{4}y(\partial_\chi^2+\partial_{\tilde{\chi}}^2)+\frac{1}{4}(y^2+y(\chi^2+\tilde{\chi}^2))\partial_\psi^2+\frac{1}{2}y(\tilde{\chi}\partial_\chi-\chi\partial_{\tilde{\chi}})\partial_\psi+y^2\partial_y^2-y\partial_y. \tag{12.4.12}$$

We furthermore recall that in terms of these variables the height function $\mathcal{F}(\mathcal{Z})$ takes a particularly simple form

$$\mathcal{F}(\mathcal{Z})=y. \tag{12.4.13}$$

Hence, the only part of the Laplacian which is non-vanishing when acting on $\mathcal{F}(\mathcal{Z})$ is the "radial part" $y^2\partial_y^2-y\partial_y$.[7] We then find $\Delta_{\mathbb{CH}^2}\mathcal{F}(\mathcal{Z})^s=s(s-2)\mathcal{F}(\mathcal{Z})$, and consequently

$$\Delta_{\mathbb{CH}^2}\mathcal{P}_s(\mathcal{Z})=s(s-2)\mathcal{P}_s(\mathcal{Z}). \tag{12.4.14}$$

[6] We are grateful to Genkai Zhang for helpful discussions on this construction.

[7] Note the extra linear term compared to the "radial part" of the Laplacian on $SL(2,\mathbb{R})/SO(2)$. This is what gives rise to the difference in the eigenvalues.

Moreover, since the Eisenstein series $\mathcal{E}_s$ in Eq. (12.4.3) has the same summand as $\mathcal{P}_s$ when the constraint $\vec{\omega}^\dagger \cdot \eta \cdot \vec{\omega} = 0$ is imposed, we may deduce that $\mathcal{E}_s$ is indeed an eigenfunction,

$$\Delta_{\mathbb{CH}^2}\mathcal{E}_s(\phi,\lambda,\gamma) = s(s-2)\mathcal{E}_s(\phi,\lambda,\gamma)\,. \tag{12.4.15}$$

12.4.3 Spherical Vector and p-Adic Eisenstein Series

As we have discussed in detail in Chapter 9, there exists a more general method for constructing automorphic forms on a coset space G/K, as developed in [195–197]. In this subsection, we show that this method may be used to construct the Eisenstein series $\mathcal{E}_s(\phi,\lambda,\gamma)$ for the Picard modular group.[8] This alternative approach also sheds light on the relation between the quadratic constraint $\vec{\omega}^\dagger \wedge \vec{\omega} = 0$ and the representation theoretic structure of the Eisenstein series. Recall from Section 9.1.3 that an automorphic form Ψ can be rewritten in the following way

$$\Psi(\mathcal{V}) = \sum_{\vec{x}\in\mathbb{Q}^n}\Big[\prod_{p<\infty} f_p(\vec{x})\Big]\rho(\mathcal{V})\cdot f_K(\vec{x}), \tag{12.4.16}$$

where $\vec{x}$ is a vector of rational numbers in $\mathbb{Q}^n$, and the product is over all prime numbers p.

To apply this method to obtain an Eisenstein series for the Picard modular group we shall begin by explicitly constructing a spherical vector f_K, invariant under $SU(2)\times U(1)$, which belongs to the principal continuous series of $SU(2,1)$-representations. Recall from Section 9.1.3 that this means that f_K belongs to the Hilbert space $\mathcal{H} = L^2(P\backslash SU(2,1))$ of real-valued, square-integrable functions on the coset space $P\backslash SU(2,1)$, where P is the parabolic subgroup of $SU(2,1)$ corresponding to the Lie algebra

$$\mathfrak{p} = \mathfrak{g}_{-2}\oplus\mathfrak{g}_{-1}\oplus\mathfrak{g}_0 \subset \mathfrak{su}(2,1), \tag{12.4.17}$$

associated with the 5-grading of $\mathfrak{su}(2,1)$ defined in (12.3.8). The parabolic group P thus corresponds to the subgroup of lower-triangular matrices,

$$P = \left\{\begin{pmatrix} t_1 & & \\ * & t_2 & \\ * & * & t_3\end{pmatrix}\in SU(2,1)\ :\ t_1t_2t_3 = 1\right\}. \tag{12.4.18}$$

The coset space $P\backslash SU(2,1)$ can be identified with the Heisenberg group N, which is parameterized as follows

$$n = e^{xX_1+\tilde{x}\tilde{X}_1+yX_2} = \begin{pmatrix} 1 & i\bar{C}_2 & C_1 \\ & 1 & C_2 \\ & & 1\end{pmatrix} \equiv \begin{pmatrix}\vec{r}_1\\ \vec{r}_2\\ \vec{r}_3\end{pmatrix}\in N, \tag{12.4.19}$$

where

$$C_1 := 2y + \frac{i}{2}|C_2|^2\,,\quad C_2 := x+\tilde{x}+i(\tilde{x}-x) \tag{12.4.20}$$

and the last equality in (12.4.19) defines the row vectors of the Heisenberg group element. It is important to note that the two complex variables are not independent, but obey the relation

$$|C_2|^2 - 2\Im(C_1) = 0\,. \tag{12.4.21}$$

[8] I am very grateful to Boris Pioline for helpful explanations of this method.

This relation is of course obvious in the present context, but we will see that this is exactly what gives rise to the quadratic constraint in the lattice construction of Section 12.4.1.

Any function $f \in L^2(P\backslash SU(2,1))$ obeys

$$\rho(g)\cdot f(n) = f(ng) = f(pn') = \chi(p)f(n'), \qquad p\in P,\ n'\in P\backslash SU(2,1), \tag{12.4.22}$$

where $\chi(p)$ is an infinitesimal character. For ρ in the principal continuous series we may choose the character:

$$\chi_s(p) := t_1^{-2s}, \qquad p = \begin{pmatrix} t_1 & & \\ * & t_2 & \\ * & * & t_3 \end{pmatrix} \in P, \tag{12.4.23}$$

where we added a subscript on the character to indicate that the principal series depends on a single real parameter s. We are therefore interested in functions $f(x,\tilde{x},y)\in L^2(P\backslash SU(2,1))$ of three real variables which transform by the overall character t_1^{-2s}.

As was apparent from Eq. (12.4.22), a general group element $g\in SU(2,1)$ acts on $n\in P\backslash SU(2,1)$ from the right, and since this will destroy the upper triangular structure, a compensating transformation of $p\in P$ from the left is needed to restore the upper triangular form of n. This left-action of P on the second and third rows, $\vec{r}_2$ and $\vec{r}_3$, of n is quite complicated, while the action on the first row $\vec{r}_1$ is very simple: $p\in P$ simply modifies $\vec{r}_1$ by an overall factor of t_1. Moreover, the action of $k\in SU(2)\times U(1)$ leaves invariant the (complex) norms of the rows $\vec{r}_i$. We may therefore construct a spherical vector $f_K\in L^2(P\backslash SU(2,1))$ as the norm of the first row $\vec{r}_1$ of n, raised to the appropriate power of s [227][9]:

$$f_K(x,\tilde{x},y) := |\vec{r}_1|^{-2s} = \left(1+|C_1|^2+|C_2|^2\right)^{-s} = \Big(1+2(x^2+\tilde{x}^2)+4y^2+(x^2+\tilde{x}^2)^2\Big)^{-s}.$$

This object is indeed invariant under $SU(2)\times U(1)$, since the right action of k on n is a "rotation" that preserves the norm, while the compensating left action of p merely modifies f_K by an overall factor t_1^{2s}, which in turn is canceled against the character $\chi_s(p)=t_1^{-2s}$ which is present since f_K is in the principal series. We have thus found our desired spherical vector.

The next step is to compute the action of $\rho(\mathcal{V})$ on f_K. Following the prescription above, this can be done by first computing $n\cdot\mathcal{V} = p_0\cdot n'$, with

$$p_0 = \begin{pmatrix} e^{-\phi} & & \\ & 1 & \\ & & e^{\phi} \end{pmatrix} \in P, \qquad n' = \begin{pmatrix} 1 & ie^{\phi}(\bar{\lambda}+\bar{C}_2) & e^{2\phi}(\gamma+i\bar{C}_2\lambda+C_1) \\ & 1 & e^{\phi}(\lambda+C_2) \\ & & 1 \end{pmatrix} \in P\backslash SU(2,1). \tag{12.4.24}$$

Applying this to the spherical vector $f_K(x,\tilde{x},y) = f_K(n)$ yields

$$\rho(\mathcal{V})\cdot f_K(n) = f_K(n\mathcal{V}) = f_K(p_0 n') = \chi_s(p_0) f_K(n') = e^{2s\phi}|\vec{r}_1{}'|^{-2s}, \tag{12.4.25}$$

which may be written explicitly in the form

$$\rho(\mathcal{V})\cdot f_K(C_1,C_2) = e^{-2s\phi}\Big(|\bar{C}_1 - iC_2\bar{\lambda}+\bar{\gamma}|^2 + e^{-2\phi}|C_2+\lambda|^2 + e^{-4\phi}\Big)^{-s}. \tag{12.4.26}$$

[9] See also [201] for a similar construction in the context of $SL(3,\mathbb{R})$.

The p-adic spherical vector $f_p(C_1, C_2)$ can now be found by the following method [195–197]: replace the complex norm $|\cdot|$ in the real spherical vector f_K by the p-adic counterpart $|\cdot|_p^{\mathbb{Q}[i]}$. Actually, this differs slighthly from the analysis in [195–197] (which dealt with the real Euclidean norm $||\cdot||$) in the sense that we must here consider the p-adic norm associated with the quadratic extension $\mathbb{Q}[i]$ of the rational numbers $\mathbb{Q}$. In this case the complex p-adic norm is defined as follows [198]:

$$|z|_p^{\mathbb{Q}[i]} := \sqrt{\big|z\bar{z}\big|_p}, \qquad z \in \mathbb{Q}[i], \tag{12.4.27}$$

where the right hand side is evaluated using the standard p-adic norm $|\cdot|_p$ since $z\bar{z} \in \mathbb{Q}$. For more information on p-adic numbers we refer the reader to [198]. The p-adic spherical vector is then given by

$$f_p(C_1, C_2) := \Big[\big|v_1\big|_p^{\mathbb{Q}[i]}\Big]^{-2s} = \max\Big(1, \sqrt{\big|C_1\bar{C}_1\big|_p}, \sqrt{\big|C_2\bar{C}_2\big|_p}\Big)^{-2s}. \tag{12.4.28}$$

The automorphic form $\Psi(\mathcal{V})$ in this representation thus reads

$$\Psi(\mathcal{V}) = \sum_{(x,\tilde{x},y)\in\mathbb{Q}^3} \Bigg[\prod_{p<\infty} \max\Big(1, \sqrt{\big|C_1\bar{C}_1\big|_p}, \sqrt{\big|C_2\bar{C}_2\big|_p}\Big)^{-2s}\Bigg]\rho(\mathcal{V}) \cdot f_K(x, \tilde{x}, y). \tag{12.4.29}$$

We can also write the summation over the complex rational variables $(C_1, C_2) \in \mathbb{Q}[i]^2$ instead of the rational variables $(x, \tilde{x}, y) \in \mathbb{Q}^3$, if we incorporate the constraint from Eq. (12.4.21) as follows

$$\Psi(\mathcal{V}) = \sum_{(C_1,C_2)\in\mathbb{Q}[i]^2} \delta\Big(|C_2|^2 - \Im(C_1)\Big)\Big[\prod_{p<\infty} f_p(C_1, C_2)\Big]\rho(\mathcal{V}) \cdot f_K(C_1, C_2). \tag{12.4.30}$$

Next we must evaluate the infinite product over prime numbers. To this end we extract the greatest common divisor between the variables C_1 and C_2 in the following way:[10]

$$C_1 := \frac{\omega_1}{\omega_3}, \quad C_2 := \frac{i\bar{\omega}_2}{\bar{\omega}_3}, \tag{12.4.31}$$

with $\omega_j \in \mathbb{Z}[i]$, for $j = 1, 2, 3$. We can now evaluate the infinite product over primes with the simple result

$$\prod_{p<\infty} \max\Big(1, \sqrt{\Big|\frac{\omega_1\bar{\omega}_1}{\omega_3\bar{\omega}_3}\Big|_p}, \sqrt{\Big|\frac{\omega_2\bar{\omega}_2}{\omega_3\bar{\omega}_3}\Big|_p}\Big)^{-2s} = |\omega_3|^{-2s}. \tag{12.4.32}$$

We also multiply the constraint $|C_2|^2 - \Im(C_1)$ by a factor of $|\omega_3|^2$ which yields

$$|\omega_3|^2\Big(|C_2|^2 - \Im(C_1)\Big) = |\omega_2|^2 - 2\Im(\omega_1\bar{\omega}_3) = \vec{\omega}^\dagger \cdot \eta \cdot \vec{\omega} = 0. \tag{12.4.33}$$

Combining Eqs. (12.4.30), (12.4.32) and (12.4.33) then yields the final form of $\Psi(\mathcal{V})$:

$$\Psi(\mathcal{V}) = \sum_{\vec{\omega}\in\mathbb{Z}[i]^3} \delta(\vec{\omega}^\dagger \cdot \eta \cdot \vec{\omega})e^{-2s\phi}\Big[|\bar{\omega}_1 + \bar{\omega}_2\bar{\lambda} + \bar{\omega}_3\bar{\gamma}|^2 + e^{-2\phi}|\bar{\omega}_2 - i\bar{\omega}_3\lambda|^2 + e^{-4\phi}|\omega_3|^2\Big]^{-s}, \tag{12.4.34}$$

[10] We note that the greatest common divisor in $\mathbb{Q}[i]$ is defined up to Gaussian units which are a subgroup of order 4 in the Gaussian integers $\mathbb{Z}[i]$.

which we recognize as the Eisenstein series $\mathcal{E}_s(\phi, \lambda, \gamma)$ constructed in Section 12.4.1. More correctly, this result is equal to $\mathcal{E}_s(\phi, \lambda, \gamma)$ modulo the term in the sum with $\omega_3 = 0$, since this term is not allowed in Eq. (12.4.34) by virtue of Eq. (12.4.31). The same phenomenon also happens when constructing non-holomorphic $SL(2, \mathbb{Z})$-invariant Eisenstein series using the p-adic approach [195, 196].

12.5 Instanton Corrections to the Universal Hypermultiplet

In this section we discuss in detail the physical interpretation of the Eisenstein series $\mathcal{E}_s(\phi, \lambda, \gamma)$ and its Fourier expansion. To this end we begin by recalling what is known about quantum corrections to hypermultiplet moduli spaces in type II Calabi-Yau compactifications, with particular emphasis on recent developments involving twistor theory. We then discuss the restriction to rigid Calabi-Yau compactifications, and show that the Eisenstein series $\mathcal{E}_s(\phi, \lambda, \gamma)$ encodes information about the non-perturbative corrections to the moduli space $\mathcal{M}_{\text{UH}}$ of the universal hypermultiplet.

12.5.1 Quantum Corrected Hypermultiplet Moduli Spaces

Perturbative corrections to hypermultiplet moduli spaces are well understood, and it has been established that the metric on $\mathcal{M}_{\text{H}}$ receives only tree-level and one-loop corrections, but no perturbative corrections beyond one loop [140,219,228–231]. For the universal hypermultiplet this was rigorously proven in [230]. The general form of the perturbative corrections can be inferred from compactifications of higher derivative couplings in ten dimensions [219, 228]. Note, however, that due to field redefinition ambiguities for higher derivative terms, the precise coefficients must be inferred from explicit string theory calculations in $D = 4$ [219].[11] In the string frame, the tree-level correction is of the form $\zeta(3)\chi_X g_s^{-2}$, where χ_X is the Euler number of the Calabi-Yau threefold X. while the one-loop correction to the metric on $\mathcal{M}_{\text{H}}$ is of the form $\zeta(2)\chi_X$. The complete perturbatively corrected metric can be found in [30, 231].

Since the work of Becker, Becker and Strominger [26] it is also known that $\mathcal{M}_{\text{H}}$ should receive non-perturbative corrections due to D2-brane and NS5-brane instantons. The contributions to $\mathcal{M}_{\text{H}}$ from these effects have, however, remained very difficult to determine, mainly due to the fact that quaternionic-Kähler geometry is much more complicated than special Kähler geometry, and because techniques for instanton calculus in string theory are not at all well developed.

Nevertheless, in a series of recent papers [27–31, 213], part of the non-perturbative corrections to $\mathcal{M}_{\text{H}}$ have been gradually understood. The key idea, following advances in the mathematics [232–234] and physics literature [235–237] on quaternionic-Kähler spaces, is that linear deformations of the hypermultiplet moduli space $\mathcal{M}_{\text{H}}$ can be lifted to its *twistor space* $\mathcal{Z}_{\text{H}}$, which is a $\mathbb{C}P^1$ bundle over $\mathcal{M}_{\text{H}}$. In contrast to $\mathcal{M}_{\text{H}}$, the $4h_{2,1} + 6$-dimensional twistor space $\mathcal{Z}_{\text{H}}$ is Kähler, and quantum corrections to $\mathcal{M}_{\text{H}}$ can thereby, in principle, be encoded in the Kähler potential of its twistor space $\mathcal{Z}_{\text{H}}$. Practically, however, the authors of [31] emphasize that a more convenient way of describing corrections to $\mathcal{M}_{\text{H}}$ is by analyzing deformations of the transition functions which glue together local patches on $\mathcal{Z}_{\text{H}}$. This can be done since

[11] I thank Pierre Vanhove for emphasizing this.

the twistor space is identified with a complex $2h_{2,1}+3$-dimensional *contact manifold* (see, e.g., [238] for an introduction to contact geometry), which, by Darboux's theorem, can be completely described by generic local complex coordinates $\left(\xi^\Lambda_{[i]}, \tilde{\xi}^{[i]}_\Lambda, \alpha_{[i]}\right)$, $\Lambda = 1, \dots, h_{2,1}+1$, in each patch $U_i \subset \mathcal{Z}_{\mathrm{H}}$, and *contact transformations* S_{ij}, relating the coordinates $\left(\xi^\Lambda_{[i]}, \tilde{\xi}^{[i]}_\Lambda, \alpha_{[i]}\right)$ to the coordinates $\left(\xi^\Lambda_{[j]}, \tilde{\xi}^{[j]}_\Lambda, \alpha_{[j]}\right)$ on the overlap $U_i \cap U_j$. The geometry of the hypermultiplet moduli space can then be extracted by determining the *contact twistor lines*, i.e. expressing the generic coordinates $\left(\xi_\Lambda, \tilde{\xi}^\Lambda, \alpha\right)$, in some patch U, in terms of the coordinates $x^\mu \in \mathcal{M}_{\mathrm{H}}$ on the base manifold, and the complex coordinate $z \in \mathbb{C}P^1$ on the fiber. Deformations of the contact transformations S_{ij} then determine the corrections to the contact twistor lines, from which the corrected geometry of $\mathcal{M}_{\mathrm{H}}$ may be extracted in terms of the *contact potential* $e^{\Phi_{[i]}(x^\mu, z)}$, which determines the Kähler potential on $\mathcal{Z}_{\mathrm{H}}$ through

$$\mathcal{K}^{[i]}_{\mathcal{Z}_{\mathrm{H}}} = \log \frac{1+|z|^2}{|z|} + \Re\big[\Phi_{[i]}(x^\mu, z)\big]. \tag{12.5.1}$$

For details on this construction we refer the reader to [30,31]. In Section 12.5.3 we will discuss some aspects of the twistor space of the universal hypermultiplet moduli space.

$SL(2,\mathbb{Z})$-Invariance and D-Brane Instantons

This elaborate scheme has been successfully carried out for all instanton corrections arising from Euclidean D2-branes wrapping both A- and B-cycles in the Calabi-Yau threefold X [27,28,31]. One of the crucial features of this analysis was to utilize the twistor techniques discussed above, while at the same time imposing $SL(2,\mathbb{Z})$-invariance of the effective action [27]. This $SL(2,\mathbb{Z})$-invariance descends from the familiar $SL(2,\mathbb{Z})$-invariance of type IIB string theory, and carries over to the IIA side via mirror symmetry in four dimensions [28]. In type IIB, the vector multiplet moduli space $\mathcal{M}^{\mathrm{IIB}}_{\mathrm{VM}}$ is classically exact to all orders in α', as well as in g_s. On the other hand, the hypermultiplet moduli space $\mathcal{M}^{\mathrm{IIB}}_{\mathrm{H}}$ receives perturbative α'- and g_s-corrections, as well as non-perturbative worldsheet instanton and brane instanton corrections. The hypermultiplet moduli in type IIB encode the Kähler structure of the Calabi-Yau threefold, and therefore receives corrections from Euclidean Dp-branes ($p=-1,1,3,5$) wrapping even cycles in X, as well as Euclidean NS5-branes wrapping the entire Calabi-Yau.

We have learned in Chapter 8 that the effects of D(-1)-instantons in ten dimensions are automatically captured by demanding invariance under $SL(2,\mathbb{Z})$ [14]. By assuming that $SL(2,\mathbb{Z})$ remains unbroken upon compactification on X, the authors of [27] showed that the D(-1)-instanton corrections to the metric on the hypermultiplet moduli space $\mathcal{M}^{\mathrm{IIB}}_{\mathrm{H}}$ are also captured by the same Eisenstein series $\mathcal{E}^{SL(2,\mathbb{Z})}_{3/2}(\tau,\bar{\tau})$, which in this context is interpreted as a contribution to the contact potential $e^{\Phi(x^\mu,z)}$ on the twistor space $\mathcal{Z}_{\mathcal{M}^{\mathrm{IIB}}_{\mathrm{H}}}$ [31].[12] Moreover,

[12] To be more specific, references [27,28] were analyzing this problem within the framework of the so-called *hyper-Kähler cone* (or *Swann bundle*) $\mathcal{S}$, which is a $\mathbb{C}^2/\mathbb{Z}_2$ bundle over the quaternionic-Kähler space $\mathcal{M}_{\mathrm{H}}$. The space $\mathcal{S}$ is then a hyperkähler manifold of real dimension $4h_{2,1}+8$. Deformations of $\mathcal{M}_{\mathrm{H}}$ can in this approach be described in terms of the twistor space $\mathcal{Z}_{\mathcal{S}} = \mathbb{C}P^1 \times \mathcal{S}$, a trivial $\mathbb{C}P^1$ bundle over $\mathcal{S}$. This construction is geared towards the so-called *projective superspace* description of the hypermultiplet moduli space, which relies on the existence of $h_{2,1}+2$ commuting isometries (in the IIA picture). This is indeed the case in the presence of only D2-brane instantons on the IIA side, or D(-1), D1 and F1-instantons in type IIB. The Eisenstein

it is known that the (D1, F1)-system transforms as a doublet under $SL(2,\mathbb{R})$, and hence by imposing $SL(2,\mathbb{Z})$-invariance in four dimensions also these instanton corrections to $\mathcal{M}_{\rm H}^{\rm IIB}$ were determined [27]. However, $SL(2,\mathbb{Z})$-invariance is not sufficient to determine the contributions to $\mathcal{M}_{\rm H}^{\rm IIB}$ from Euclidean D3, D5- and NS5-brane instantons.

In [28], these results were translated into the type IIA picture using mirror symmetry. After the mirror map, the contact potential $e^{\Phi(x^\mu,z)}$ carries information about the corrected geometry of $\mathcal{M}_{\rm H} := \mathcal{M}_{\rm H}^{\rm IIA}$ due to Euclidean D2-branes wrapping A-cycles in X. However, $SL(2,\mathbb{Z})$-invariance is insufficient also in the IIA picture for summing up all possible non-perturbative corrections, and indeed misses the contributions from B-type D2-brane instantons, and NS5-brane instantons. The insight of the recent paper [31], was that the contributions from D2-branes wrapping B-cycles may nevertheless be obtained by imposing *electric-magnetic duality* (or, more generally, *symplectic invariance*) between A- and B-cycles. This led to a proposal for the contact potential $e^{\Phi(x^\mu,z)}$ encoding all D2-brane instanton corrections to the hypermultiplet moduli space $\mathcal{M}_{\rm H}$ in type IIA on a Calabi-Yau threefold X.

NS5-Brane Instantons Corrections

The NS5-brane contributions were unfortunately still out of reach in the analysis of [28, 31], due to the fact that deformations of the transition functions S_{ij} on the twistor space $\mathcal{Z}_{\rm H}$ become considerably more complicated if the NS5-brane contribution is taken into account. Nevertheless, a "roadmap" for how to incorporate NS5-brane effects was proposed in [28]. The main idea is to use mirror symmetry to transform the moduli space $\mathcal{M}_{\rm H}$ back to the IIB side, after electric-magnetic duality between A- and B-cycles has been imposed. In type IIB this will then automatically include the previously unknown effects of Euclidean D3- and D5-branes. The NS5-branes should then in principle emerge after once more imposing $SL(2,\mathbb{Z})$-invariance, since the (D5, NS5)-system transforms as a doublet under $SL(2,\mathbb{R})$. The final step would be to use mirror symmetry to map the result back to the IIA side, where now the NS5-brane effects are included. Although quite appealing, this program is technically very challenging due to the complicated action of $SL(2,\mathbb{Z})$ on the moduli after mapping the results back to the IIB-side [31]. In the next subsection, we give a proposal for how to circumvent this problem in the special case of type IIA compacifications on a rigid Calabi-Yau threefold $\mathcal{X}$.

12.5.2 Instanton Corrections to the Universal Hypermultiplet

As discussed in detail in Section 12.2, rigid Calabi-Yau compactifications give rise to pure $\mathcal{N} = 2$ supergravity in four dimensions (ignoring the vector multiplets), coupled to the *universal hypermultiplet*, whose moduli space $\mathcal{M}_{\rm UH}$ coincides with the coset space $SU(2,1)/(SU(2)\times U(1))$, parametrized by the four-dimensional dilaton e^{ϕ}, the Ramond-Ramond axions χ and $\tilde{\chi}$, and the axionic scalar ψ, dual to the NS 2-form. This may be viewed as a "toy example" of the more general discussion of Section 12.5.1, but one which still contains the essential

series $\mathcal{E}_{3/2}^{SL(2,\mathbb{Z})}(\tau,\bar{\tau})$, encoding the D(-1)-instantons, then gives a contribution to the hyperkähler potential χ on $\mathcal{S}$ [27]. This formalism fails, however, when B-type D2-brane instantons as well as NS5-brane instantons in IIA (or D3, D5, and NS5-brane instantons in IIB) are included. For such generic instanton configurations no isometries remain and therefore the projective superspace approach breaks down. However, the general method of [31], based on the twistor space $\mathcal{Z}_{\rm H}$ of $\mathcal{M}_{\rm H}$ itself, still applies.

features: namely D2-brane instantons as well as NS5-brane instantons [223]. This follows from the fact that, even though there is no complex structure, Euclidean D2-branes may still wrap the universal holomorphic and anti-holomorphic 3-cycles $\mathcal{A}$ and $\mathcal{B}$, and Euclidean NS5-branes may wrap the entire rigid Calabi-Yau manifold $\mathcal{X}$. Perturbative and non-perturbative corrections to $\mathcal{M}_{\text{UH}}$ have been discussed extensively in the literature [140, 213, 219, 223, 230], but no proposal has been put forward for the exact quantum corrected geometry including the effects of NS5-branes.

On the Physical Relevance of the Picard Modular Group

As discussed extensively in Chapters 1 and 8, it is generally expected that the full quantum effective action in string compactifications should be invariant under a global discrete symmetry group $G(\mathbb{Z})$ (see, e.g., [8, 11–14]). However, for Calabi-Yau threefold compactifications little is known about the structure of $G(\mathbb{Z})$. For the case of rigid Calabi-Yau compactifications the situation is different. In this case we know that the classical effective action exhibits a global $SU(2,1)$-symmetry, which is broken to a discrete subgroup $G(\mathbb{Z}) \subset SU(2,1)$ by quantum corrections [223]. The problem is then reduced to finding which discrete subgroup $G(\mathbb{Z})$ is preserved. The following transformations are expected to be preserved in the quantum theory:

- A discrete subgroup $N(\mathbb{Z}) \subset N$ of the Heisenberg group, acting by discrete (Peccei-Quinn) shift symmetries on the axions $\chi, \tilde{\chi}$ and ψ [223]:

$$\begin{aligned} \chi &\longmapsto \chi + a\,, \\ \tilde{\chi} &\longmapsto \tilde{\chi} + b\,, \\ \psi &\longmapsto \psi + \frac{1}{2}c - a\tilde{\chi} + b\chi, \end{aligned} \qquad (12.5.2)$$

  where $a, b, c \in \mathbb{Z}$, while leaving the dilaton invariant. The breaking of the continuous shifts of χ and $\tilde{\chi}$ are due to D2-brane instantons, while the breaking of the shift of ψ is due to NS5-brane instantons. The factor $1/2$ appearing in front of c is in agreement with the quantisation condition on the NS5-brane instantons derived in [223].

- A transformation R which interchanges the R-R scalars χ and $\tilde{\chi}$ [223]:

$$R\ :\ (\chi, \tilde{\chi}) \longmapsto (-\tilde{\chi}, \chi). \qquad (12.5.3)$$

  Microscopically, this corresponds to a phase shift on the holomorphic 3-form $\Omega_3 \sim dz^1 \wedge dz^2 \wedge dz^3$ of $\mathcal{X}$, and can be interpreted as "electric-magnetic duality" between A- and B-cycles in the Calabi-Yau manifold.

- Finally, a transformation S which inverts the string coupling g_s, and is interpreted as a kind of S*-duality*. This should act in the standard non-linear way on the complex parameter $\tau = \chi + ie^{-\phi}$.

  Let us emphasize, however, that it is debatable whether type IIA in $D = 4$ really exhibits S-duality in the same sense as on the type IIB side.[13] Since the strong-coupling limit of

[13] I am grateful to Boris Pioline, Stefan Vandoren and Pierre Vanhove for helpful discussions on this crucial point of the analysis.

type IIA is expected to be M-theory, it does not appear sensible to talk about type IIA as self-dual under S-duality. Nevertheless, the hypermultiplet moduli space is the same for M-theory on $X \times S^1$ and type IIA on X, and one might therefore argue that the existence of S-duality is a consequence of the M-theory picture of the moduli space. In particular, the universal hypermultiplet appears also for compactifications of M-theory on $\mathcal{X}$, where it receives non-perturbative corrections from M2-brane as well as M5-brane instantons [26]. It is also argued in [239] that the mirror image of the type IIB $SL(2,\mathbb{Z})$ in $D = 4$ is broken by worldsheet instanton effects, although that it might be restored after including spacetime instanton effects in type IIA.

In the following analysis, we will assume that the universal hypermultiplet does exhibit S-duality, an assumption which is crucial for the main conclusions of our analysis.

In view of the theorem cited in Section 12.3.5, we recognize that these transformations together generate precisely the *Picard modular group* $SU(2,1;\mathbb{Z}[i]) \subset SU(2,1)$. Our proposal is thus that for type IIA compactifications on rigid Calabi-Yau manifolds, the discrete symmetry group $G(\mathbb{Z}) = SU(2,1;\mathbb{Z}[i])$ should be preserved in the quantum theory, and the exact form of the hypermultiplet moduli space is

$$\mathcal{M}_{\text{UH}}^{\text{exact}} = SU(2,1;\mathbb{Z}[i])\backslash SU(2,1)/\big(SU(2)\times U(1)\big)\,. \tag{12.5.4}$$

We would now like to test this proposal by analyzing whether the expected instanton corrections to the universal hypermultiplet can be determined by imposing invariance under $SU(2,1;\mathbb{Z}[i])$. More specifically, we want to understand if the Eisenstein series $\mathcal{E}_s(\phi,\lambda,\gamma)$, constructed in Section 12.4, may be related to a generalized version of the contact potential $e^{\Phi(x^\mu,z)}$ given in [31], but here restricted to the special case of rigid Calabi-Yau compactifications. To this end we shall in the following section investigate in detail the Fourier expansion of $\mathcal{E}_s(\phi,\lambda,\gamma)$ to determine if it encodes the expected instanton contributions, in the spirit of [14, 27, 28, 31].

12.5.3 The Non-Abelian Fourier Expansion of $\mathcal{E}_s(\phi,\lambda,\gamma)$

In Chapters 8 and 10 it was emphasized that the method of revealing which physical contributions are contained in a certain Eisenstein series is given by Fourier expansion. We would therefore like to compute the Fourier expansion of the $SU(2,1;\mathbb{Z}[i])$-invariant Eisenstein series $\mathcal{E}_s(\phi,\lambda,\gamma)$ to determine whether or not it reproduces the expected perturbative and non-perturbative corrections. The calculation of the Fourier expansion of $\mathcal{E}_s(\phi,\lambda,\gamma)$ is currently in progress, and will appear in **Paper VIII**.[14] In this section we shall provide some preliminary remarks on the general structure, which appears to exhibit the desired features. This general structure of the Fourier expansion of automorphic forms for the Picard modular group is discussed from a mathematical point of view in [240, 241], to which we refer the interested reader. A similar discussion may also be found in the mathematics [242, 243] and physics [244] (**Paper VII**) literature for the case of automorphic forms on $SL(3,\mathbb{R})/SO(3)$.

We learned in Section 10.1 that when the nilpotent subgroup $N \subset G$ is non-abelian, the Fourier expansion must split up into an abelian and a non-abelian part. The abelian

[14] I am very grateful to Boris Pioline for collaboration in understanding this Fourier expansion.

part represents the Fourier expansion with respect to the abelianized group $\tilde{N} = N/Z$, and the non-abelian part corresponds to the Fourier expansion with respect to the center Z. For the case of $G = SU(2,1)$ we have seen in Eq. (12.3.7) that the nilpotent subgroup N is isomorphic to a three-dimensional Heisenberg group, and is therefore non-abelian. The center of N coincides with the commutator subgroup $[N, N]$:

$$Z = [N, N] = \left\{ \begin{pmatrix} 1 & & * \\ & 1 & \\ & & 1 \end{pmatrix} \right\}, \tag{12.5.5}$$

while the abelianized group $\tilde{N}$ is given by

$$\tilde{N} = N/Z = \left\{ \begin{pmatrix} 1 & * & \\ & 1 & * \\ & & 1 \end{pmatrix} \right\}. \tag{12.5.6}$$

The action of $\tilde{N}(\mathbb{Z})$ on the axionic scalars χ and $\tilde{\chi}$ then reflects the general structure discussed in Section 10.1:

$$\begin{aligned} \tilde{N}(\mathbb{Z}) \quad : \quad \chi &\longmapsto \chi + a, \\ \tilde{\chi} &\longmapsto \tilde{\chi} + b, \end{aligned} \tag{12.5.7}$$

with $a, b \in \mathbb{Z}$. Similarly, the full action of $N(\mathbb{Z}) = \tilde{N}(\mathbb{Z}) \times Z(\mathbb{Z})$ on the third axionic scalar ψ is given by

$$\tilde{N}(\mathbb{Z}) \times Z(\mathbb{Z}) \; : \; \psi \; \longmapsto \psi + \frac{1}{2}c - a\tilde{\chi} + b\chi, \tag{12.5.8}$$

where $c \in \mathbb{Z}$. This is also in accordance with the general structure of Eq. (10.1.8). Note that the additional factor of 1/2 arises because of our choice of parametrization in (12.3.19).

With these preparations out of the way, we can immediately write the general form of the Fourier expansion, following (10.1.9),

$$\mathcal{E}_s(\phi, \chi, \tilde{\chi}) = \mathcal{E}_s^{(0)}(\phi) + \mathcal{E}_s^{(\mathrm{A})}(\phi, \chi, \tilde{\chi}) + \mathcal{E}_s^{(\mathrm{NA})}(\phi, \chi, \tilde{\chi}, \psi), \tag{12.5.9}$$

where $\mathcal{E}^{(0)}(\phi)$ is the constant term and

$$\begin{aligned} \mathcal{E}_s^{(\mathrm{A})}(\phi, \chi, \tilde{\chi}) \quad &= \quad \sum_{(p,q)\neq(0,0)} \mathfrak{C}_{p,q}(\phi) e^{2\pi i (q\chi - p\tilde{\chi})}, \\ \mathcal{E}_s^{(\mathrm{NA})}(\phi, \chi, \tilde{\chi}, \psi) \quad &= \quad \sum_{k\neq 0} \mathfrak{C}_k(\phi, \chi, \tilde{\chi}) e^{4\pi i k \psi}. \end{aligned} \tag{12.5.10}$$

We shall now discuss the three pieces of the Fourier expansion in turn, starting with the constant term $\mathcal{E}_s^{(0)}$. By matching this term with the known perturbative contributions to the universal hypermultiplet will allow us to determine which order s is physically relevant.

Constant Terms and Perturbative Contributions

As we have seen in Section 10.1, the constant term is defined generally as

$$\mathcal{E}_s^{(0)}(\phi) = \int_0^1 d\chi \int_0^1 \tilde{\chi} \int_0^{1/2} d\psi \; \mathcal{E}_s(\phi, \chi, \tilde{\chi}, \psi), \tag{12.5.11}$$

where the integral over the NS-scalar ψ runs from 0 to 1/2 because of the extra factor of 2 in front of ψ in our parametrization of N in Eq. (12.3.19). Since the Cartan subgroup A appearing in the Iwasawa decomposition of $SU(2,1)$ is one-dimensional, the constant term only depends on the dilatonic scalar ϕ. Moreover, recall from the discussion in Section 12.3.4 that the Weyl group of $\mathfrak{su}(2,1)$ is the Weyl group of the restricted root system BC_1, which is isomorphic with $\mathbb{Z}_2$. Hence, the constant term $\mathcal{E}_s^{(0)}(\phi)$ consists of two contributions, which are permuted by $\mathbb{Z}_2$.[15]

Similarly to the analysis in Section 10.2 the powers of e^{ϕ} in $\mathcal{E}_s^{(0)}(\phi)$ may be determined by the Laplacian condition on $\mathcal{E}_s$. In Section 12.4.2 we have seen that the Eisenstein series is an eigenfunction of the Laplacian $\Delta_{\mathbb{CH}^2}$ with eigenvalue $s(s-2)$. This implies that all the constant terms must individually be eigenfunctions of $\Delta_{\mathbb{CH}^2}$ with the same eigenvalue. It turns out that there is a unique solution to this, and we find that $\mathcal{E}_s^{(0)}$ must be of the form

$$\mathcal{E}_s^{(0)}(\phi) = A(s)e^{-2s\phi} + B(s)e^{2(s-2)\phi}. \tag{12.5.12}$$

We see that for the special value $s = 3/2$ the dilaton exponents become

$$\mathcal{E}_{3/2}^{(0)}(\phi) = e^{-\phi}\Big(A(3/2)e^{-2\phi} + B(3/2)\Big), \tag{12.5.13}$$

where the terms within the bracket have precisely the correct dilaton dependence to correspond to tree-level and one-loop effects in the string coupling $g_s = e^{\phi}$. It is known from explicit string theory calculations [140, 219] that the tree-level and one-loop coefficients read $A(3/2) = 2\zeta(3)$ and $B(3/2) = 4\zeta(2)$, respectively. We expect that these values will be reproduced by the constant term calculation in Eq. (12.5.11). This is work in progress which will be reported in **Paper VIII**.

General Structure of the Fourier Coefficients

The precise Fourier coefficients in (12.5.10) will be given in **Paper VIII**, while we shall here discuss some preliminary features. The semiclassical limit $g_s = e^{\phi} \to 0$ of the abelian Fourier coefficients is given by

$$\mathfrak{C}_{p,q}(\phi) \sim e^{-\frac{2\pi}{g_s}\sqrt{p^2+q^2}}, \tag{12.5.14}$$

which exhibits the correct scaling to be attributed to D2-brane instantons. Combining this with the overall phase factor in (12.5.10) we find the abelian term contributes with exponentially suppressed corrections of the form

$$\mathcal{E}_s^{(\mathrm{A})}(\phi, \chi, \tilde{\chi}) \sim \sum_{(p,q)\neq(0,0)} \mu_s(p,q) e^{-S_{p,q}(\phi,\chi,\tilde{\chi})}\Big[1 + \mathcal{O}(g_s) + \cdots\Big], \tag{12.5.15}$$

[15] I am grateful to Pierre Vanhove for helpful discussions on the constant terms.

where we defined

$$S_{p,q}(\phi,\chi,\tilde{\chi}) = \frac{2\pi}{g_s}\sqrt{q^2+p^2} - 2\pi i q\chi + 2\pi i p\tilde{\chi}. \tag{12.5.16}$$

This may be identified with the instanton action for Euclidean D2-branes wrapping the homology class $q\mathcal{A}+p\mathcal{B}\in H_3(\mathcal{X})$. The instanton measure $\mu_s(p,q)$ in (12.5.15) will be given in **Paper VIII**.

Let us now proceed to the non-abelian term $\mathcal{E}_s^{(\mathrm{NA})}(\phi,\chi,\tilde{\chi},\psi)$. In the semi-classical limit the non-abelian coefficients exhibit the following exponential suppression

$$\mathfrak{C}_k(\phi,\chi,\tilde{\chi},\psi) \sim \sum_{(p,q)\in\mathbb{Z}^2} \mu_s(p,q,k) e^{-\frac{2\pi}{g_s^2}\sqrt{g_s^2(q^2+p^2)+k}+2\pi i(q\chi-p\tilde{\chi})}. \tag{12.5.17}$$

Combining this result with the phase factor in (12.5.10) we find a total contribution of the form

$$\mathcal{E}_s^{(\mathrm{NA})}(\phi,\chi,\tilde{\chi},\psi) \sim \sum_{k\neq 0}\sum_{(p,q)\in\mathbb{Z}^2} \mu_s(p,q,k) e^{-S_{p,q,k}(\phi,\chi,\tilde{\chi},\psi)}\Big[1+\mathcal{O}(g_s)+\cdots\Big], \tag{12.5.18}$$

where we have identified the instanton action for combined D2-brane and NS5-brane instanton effects:

$$S_{p,q,k}(\phi,\chi,\tilde{\chi},\psi) = \frac{2\pi}{g_s^2}\sqrt{k^2+g_s^2(q^2+p^2)} + 2\pi i(-q\chi+p\tilde{\chi}-2k\psi). \tag{12.5.19}$$

In particular, setting $p=q=0$ in (12.5.18) we recover the instanton action for k NS5-branes wrapping $\mathcal{X}$:

$$S_k(\phi,\psi) = \frac{2\pi|k|}{g_s^2} - 4\pi i k\psi. \tag{12.5.20}$$

The non-abelian instanton measure $\mu_s(p,q,k)$ will appear in **Paper VIII**.

It is satisfactory to see from this preliminary analysis that the Fourier expansion of $\mathcal{E}_s(\phi,\lambda,\gamma)$ appears to exhibit the expected quantum corrections to the metric on the universal hypermultiplet moduli space. Of course, in order to verify this claim we must also extract the perturbative coefficients $A(s)$ and $B(s)$ in the constant term, as well as compute the abelian and non-abelian instanton measures. These results will appear in **Paper VIII**.

12.5.4 Twistorial Interpretation of the Eisenstein Series

The analysis carried out in the previous subsection appears to capture parts of the expected non-perturbative effects that correct the geometry of the moduli space $\mathcal{M}_{\mathrm{UH}}$. But this also begs the question: How exactly is the Eisenstein series $\mathcal{E}_{3/2}(\phi,\lambda,\gamma)$ related to the metric on $\mathcal{M}_{\mathrm{UH}}$? We do not have a clear answer to this question, but in this subsection we shall offer some natural speculations.

As was briefly discussed in Section 12.5.1, the quantum corrected geometry of $\mathcal{M}_{\mathrm{H}}$ is preferably not described directly by its quaternionic-Kähler metric, but rather through certain properties of its twistor space $\mathcal{Z}_{\mathrm{H}}$, namely the contact twistor lines $(\xi_\Lambda,\tilde{\xi}^\Lambda,\alpha)$, in some local patch $U\subset\mathcal{Z}_{\mathrm{H}}$, and the contact potential $e^{\Phi(x^\mu,z)}$ which in turn determines the Kähler potential on $\mathcal{Z}_{\mathrm{H}}$. We have been deliberately vague in our discussion of this construction because the direct relation to our analysis is not perfectly clear. Let us, however, try to be a little more specific in the context of the universal hypermultiplet.

The twistor space $\mathcal{Z}_{\rm UH}$ of the moduli space $\mathcal{M}_{\rm UH}$ can be nicely described group-theoretically as follows. The twistor fiber $\mathbb{C}P^1$ can be taken to be the two-sphere $S^2 = SU(2)/U(1)$, and the fibration over $\mathcal{M}_{\rm UH}$ is then described as a direct product of coset spaces [227, 245]:

$$\mathcal{Z}_{\rm UH} = \frac{SU(2)}{U(1)} \times \frac{SU(2,1)}{SU(2)\times U(1)} = \frac{SU(2,1)}{U(1)\times U(1)}. \tag{12.5.21}$$

The twistor space $\mathcal{Z}_{\rm UH}$ is a complex 3-dimensional contact manifold, described by the local coordinates $(\xi, \tilde{\xi}, \alpha)$. These coordinates also have a nice group-theoretic interpretation: they parametrize the complexified Heisenberg group $N_{\mathbb{C}}$, or, equivalently, they are coordinates on the complex coset space $P_{\mathbb{C}}\backslash SU(2,1;\mathbb{C})$, where $P_{\mathbb{C}}$ is the complexification of the parabolic subgroup $P \subset SU(2,1)$ discussed in Section 12.4.3. In terms of the coordinates $(\xi, \tilde{\xi}, \alpha)$ on $P_{\mathbb{C}}\backslash SU(2,1;\mathbb{C})$ the Kähler potential of $\mathcal{Z}_{\rm UH}$ takes the following form [227]

$$K_{\mathcal{Z}_{\rm UH}} = \frac{1}{2}\log\left[\Big((\xi-\bar{\xi})^2 + (\tilde{\xi}-\bar{\tilde{\xi}})^2\Big)^2 + 4\Big(\alpha - \bar{\alpha} + \bar{\xi}\tilde{\xi} - \xi\bar{\tilde{\xi}}\Big)^2\right]. \tag{12.5.22}$$

The contact twistor lines for the unperturbed twistor space correspond to the change of variables that relate the coordinates $(\xi, \tilde{\xi}, \alpha)$ on $\mathcal{Z}_{\rm UH}$ to the coordinates $x^\mu = \{e^\phi, \chi, \tilde{\chi}, \psi\}$ on the base $\mathcal{M}_{\rm UH}$ and the coordinate z on the fiber $\mathbb{C}P^1 = SU(2)/U(1)$. These twistor lines were obtained in [227], and in the present conventions they read

$$\begin{aligned}
\xi &= -\sqrt{2}\chi + \frac{1}{\sqrt{2}}e^{-\phi}\big(z - z^{-1}\big),\\
\tilde{\xi} &= -\sqrt{2}\tilde{\chi} - \frac{i}{\sqrt{2}}e^{-\phi}\big(z + z^{-1}\big),\\
\alpha &= 2\psi - e^{-\phi}\Big[z(\tilde{\chi} + i\chi) - z^{-1}(\tilde{\chi} - i\chi)\Big].
\end{aligned} \tag{12.5.23}$$

Plugging these into (12.5.22) we find that the Kähler potential for the twistor space of the classical universal hypermultiplet, in the coordinates $x^\mu \in \mathcal{M}_{\rm UH}$ and $z \in \mathbb{C}P^1$, reads

$$K_{\mathcal{Z}_{\rm UH}} = \log\frac{1+|z|^2}{|z|} - 2\phi. \tag{12.5.24}$$

which indeed agrees with the general form of the Kähler potential in Eq. (12.5.1) upon identifying the classical contact potential as follows

$$e^{\Phi_{\rm classical}(x^\mu, z)} = e^{-2\phi}. \tag{12.5.25}$$

The approach of [31] was to compute the most general deformations of the twistor lines compatible with D2-brane instanton effects in the linear approximation. These were then used to compute the instanton-corrected contact potential $e^{\Phi(x^\mu, z)}$. By restricting the expression for $e^{\Phi_{A/B}}$ given in Eq. (4.17) of [31] to a rigid Calabi-Yau compactification[16] one obtains

[16] For example, after translation into our conventions, the expression for $m\Theta_\gamma$ in that equation becomes, in the rigid case,

$$m\Theta_\gamma = m(k_\Lambda\zeta^\Lambda - l^\Lambda\tilde{\zeta}_\Lambda) \longrightarrow \ell_1\chi + \ell_2\tilde{\chi}.$$

an expression very similar to the abelian part $\mathcal{E}^{(\mathrm{A})}_{3/2}$ of the Fourier expansion of $\mathcal{E}_{3/2}(\phi,\lambda,\gamma)$. It would therefore be natural to conjecture that the Eisenstein series $\mathcal{E}_{3/2}(\phi,\lambda,\gamma)$ should be seen as a generalized version, $e^{\Phi_{A/B/NS5}}$, of the contact potential $e^{\Phi_{A/B}}$ for the universal hypermultiplet in the presence of NS5-brane instantons.

However, this naive suggestion is most likely not quite correct. For instance, as emphasized repeatedly in [31], in the absence of NS5-brane corrections, the perturbed twistor lines simplify considerably in such a way that the contact potential $e^{\Phi_{A/B}}$ is globally defined on the twistor space $\mathcal{Z}_{\mathrm{UH}}$, and, moreover, its dependence on the twistor variable $z\in\mathbb{C}P^1$ drops out, $\Phi_{A/B}(x^\mu,z)\equiv\Phi_{A/B}(x^\mu)$. When NS5-brance corrections are incorporated this is no longer true, and one must consider $\Phi^{[i]}(x^\mu,z)$ separately in each local patch $U_i\subset\mathcal{Z}_{\mathrm{UH}}$, and the dependence on the twistor fiber is no longer absent. By construction, the Eisenstein series $\mathcal{E}_{3/2}(\phi,\lambda,\gamma)$ has of course no dependence on the fiber coordinate z and therefore cannot be directly associated with the contact potential $e^{\Phi^{[i]}(x^\mu,z)}$.

One possibility is that the Eisenstein series $\mathcal{E}_{3/2}(\phi,\lambda,\gamma)$ should coincide with the contact potential $e^{\Phi(x^\mu,z)}$ only at the north pole $z=0$ of the fiber $\mathbb{C}P^1$. Similar arguments are given in **Paper VII** in a related context. In order to have a complete understanding of the global structure of the deformed twistor space $\mathcal{Z}_{\mathrm{UH}}$ one must also analyze the quantum corrections to the twistor lines $(\xi,\tilde{\xi},\alpha)$ in Eq. (12.5.23). It is not clear at this stage how to properly incorporate $SU(2,1;\mathbb{Z}[i])$-invariance in this context, and we leave a more careful investigation of this issue for future work.

12.5.5 Relation with the Przanowski Framework

In a series of works [246,247], Przanowski developed a method to describe any four-dimensional quaternionic metric determined by a single real-valued function $h(\omega_1,\omega_2)$, depending on local complex coordinates ω_1 and ω_2. This function h satisfies the so-called *master equation*

$$\Big[\frac{\partial^2}{\partial\omega^1\partial\omega^{\bar{1}}}h\Big]\Big[\frac{\partial^2}{\partial\omega^2\partial\omega^{\bar{2}}}h\Big]-\Big[\frac{\partial^2}{\partial\omega^1\partial\omega^{\bar{2}}}h\Big]\Big[\frac{\partial^2}{\partial\omega^{\bar{1}}\partial\omega^2}h\Big]+\Big(2\frac{\partial^2}{\partial\omega^1\partial\omega^{\bar{1}}}h-\Big[\frac{\partial}{\partial\omega^1}h\Big]\Big[\frac{\partial}{\partial\omega^{\bar{1}}}h\Big]\Big)e^h=0. \tag{12.5.26}$$

The map between the local complex coordinates ω_1,ω_2 and the physical variables of the universal hypermultiplet was found in [213], where it was shown that for the classical metric the function h is simply given by $h_0=2\phi$. Hence, by comparing with (12.5.25), we deduce that the (inverse) function e^{-h_0} coincides with the classical contact potential, i.e.

$$e^{-h_0}=e^{\Phi_{\mathrm{classical}}(x^\mu,z)}. \tag{12.5.27}$$

More generally, one expects that the function e^{-h} corresponds to the contact potential $e^{\Phi(x^\mu,z)}$ restricted to the north pole $z=0$ of the twistor space.[17] It is then tempting to conjecture that the non-perturbative completion of e^{-h_0} should be given in terms of the Eisenstein series $\mathcal{E}_{3/2}(\phi,\lambda,\gamma)$. However, due to the overall dilaton factor $e^{-\phi}$ appearing in the constant terms of $\mathcal{E}_{3/2}$ (see Eq. (12.5.13)), in order to obtain the correct powers of the dilaton in the constant terms we have to make a rescaling: $\mathcal{E}_{3/2}\longrightarrow e^{\phi}\mathcal{E}_{3/2}$. Then we may propose a

[17] I am grateful to Boris Pioline for pointing this out to me.

non-perturbative completion of e^{-h_0}, or, equivalently, of the contact potential $e^{\Phi_{\rm classical}}$, given by

$$e^{-h_{\rm exact}} = e^{\Phi_{\rm exact}(x^\mu,0)} = e^{-h_0} + e^{\phi}\mathcal{E}_{3/2}(\phi,\lambda,\gamma) = e^{-2\phi}\big(1+2\zeta(3)\big) + c(3/2) + \cdots\,, \quad (12.5.28)$$

where we recall from (12.5.13) that $c(3/2)$ is the one-loop coefficient. The rescaling of $\mathcal{E}_{3/2}$ cannot be attributed to a Weyl rescaling, and at this stage we do not have a clear understanding of its meaning. The same problem occurs if one tries to interpret the two constant terms in (12.5.13) within the Calderbank-Pedersen framework (see, e.g., [219]).

A

Details on the E_{10}/Massive IIA Correspondence

A.1 Details For Massive IIA Supergravity

In this appendix we give all the relevant details of massive IIA supergravity which are required to establish the correspondence with the E_{10}-sigma model. The complete Lagrangian and supersymmetry variations were already given in Section 2, and will not be repeated here. Here we complement this with information regarding our conventions, as well as explicit expressions for the bosonic and fermionic equations of motion, and the Bianchi identities. Moreover, we discuss in detail the truncations that we need to impose on the supergravity side to ensure the matching with the geodesic sigma model. Finally, we also discuss how our conventions for massless type IIA supergravity matches with a reduction from eleven-dimensional supergravity.

A.1.1 Conventions

We use the signature $(-+\ldots+)$ for space-time. The indices $\mu=(t,m)$ are $(1+9)$-dimensional curved indices, whereas $\alpha=(0,a)$ are the corresponding flat indices. Partial derivatives with flat indices are defined via conversion with the inverse vielbein: $\partial_\alpha = e_\alpha{}^\mu\partial_\mu$. The Lorentz covariant derivative D_α acts on (co-)vectors via $D_\alpha V_\beta = \partial_\alpha V_\beta + \omega_{\alpha\,\beta}{}^\gamma V_\gamma$ in terms of the Lorentz connection, which is defined in turn as

$$\omega_{\alpha\,\beta\gamma} = \frac{1}{2}\left(\Omega_{\alpha\beta\,\gamma}+\Omega_{\gamma\alpha\,\beta}-\Omega_{\beta\gamma\,\alpha}\right) \tag{A.1.1}$$

in terms of the anholonomy of the orthonormal frame $e_\mu{}^\alpha$ defined by $\Omega_{\mu\nu}{}^\alpha = 2\partial_{[\mu}e_{\nu]}{}^\alpha$.

Our fermions are Majorana–Weyl spinors of $SO(1,9)$ of real dimension 16. As the theory is type II non-chiral we can combine two spinors into a 32-dimensional Majorana representation on which the Γ-matrices of $SO(1,10)$ act. These are the eleven real 32×32 matrices $(\Gamma^0,\Gamma^a,\Gamma^{10})$ which are symmetric except for Γ^0 which is antisymmetric. We choose the representation such that Γ^{10} is block diagonal and projects on the two 16 component spinors of opposite chirality. Γ^0 is the charge conjugation matrix such that our conventions are identical to those of [61].

A useful identity for our Γ-matrices is

$$\Gamma^{a_1\dots a_k} = \frac{(-1)^{(k+1)(k+2)/2}}{(9-k)!}\epsilon^{a_1\dots a_k b_1\dots b_{9-k}}\Gamma_{b_1\dots b_{9-k}}\Gamma^0\Gamma^{10}\,. \tag{A.1.2}$$

The various ϵ tensors we use are such that

$$\epsilon^{0\,1\dots 10} = +1\,, \quad \epsilon^{0\,1\dots 9} = +1\,, \quad \epsilon^{1\,\dots 9} = +1\,. \tag{A.1.3}$$

A.1.2 Bianchi Identities

For the comparison with E_{10} it is useful to write all supergravity equations in a non-coordinate orthonormal frame described by the vielbein $e_\mu{}^\alpha$ and use only Lorentz covariant objects. In these flat indices the Bianchi identities following from (7.1.2) are

$$3D_{[\alpha_1}F_{\alpha_2\alpha_3]} = mF_{\alpha_1\alpha_2\alpha_3}\,, \tag{A.1.4a}$$
$$4D_{[\alpha_1}F_{\alpha_2\alpha_3\alpha_4]} = 0\,, \tag{A.1.4b}$$
$$5D_{[\alpha_1}F_{\alpha_2\alpha_3\alpha_4\alpha_5]} = 10F_{[\alpha_1\alpha_2}F_{\alpha_3\alpha_4\alpha_5]}\,. \tag{A.1.4c}$$

A.1.3 Bosonic Equations of Motion

The form equations of motion can be rewritten in 10-dimensional flat indices as

$$D_\alpha(e^{3\phi/2}F^{\alpha\beta}) = -\frac{1}{3!}e^{\phi/2}F^{\alpha_1\alpha_2\alpha_3\beta}F_{\alpha_1\alpha_2\alpha_3}\,, \tag{A.1.5a}$$

$$\begin{aligned} D_\alpha(e^{-\phi}F^{\alpha\beta_1\beta_2}) &= me^{3\phi/2}F^{\beta_1\beta_2} + \frac{1}{2!}e^{\phi/2}F^{\beta_1\beta_2\alpha_1\alpha_2}F_{\alpha_1\alpha_2} \\ &\quad -\frac{1}{1152}F_{\alpha_1\dots\alpha_4}F_{\alpha_5\dots\alpha_8}\epsilon^{\alpha_1\dots\alpha_8\beta_1\beta_2}\,, \end{aligned} \tag{A.1.5b}$$

$$D_\alpha(e^{\phi/2}F^{\alpha\beta_1\dots\beta_3}) = \frac{1}{144}F_{\alpha_1\dots\alpha_4}F_{\alpha_5\dots\alpha_7}\epsilon^{\alpha_1\dots\alpha_7\beta_1\dots\beta_3}, \tag{A.1.5c}$$

while the dilaton and gravity equations are

$$D^\alpha\partial_\alpha\phi = \frac{3}{8}e^{3\phi/2}|F_{(2)}|^2 - \frac{1}{12}e^{-\phi}|F_{(3)}|^2 + \frac{1}{96}e^{\phi/2}|F_{(4)}|^2 + \frac{5}{4}m^2e^{5\phi/2}\,, \tag{A.1.6a}$$

$$\begin{aligned} R_{\alpha\beta} &= \frac{1}{2}\partial_\alpha\phi\partial_\beta\phi + \frac{m^2}{16}\eta_{\alpha\beta}e^{5\phi/2} + \frac{1}{2}e^{3\phi/2}F_{\alpha\gamma}F_\beta{}^\gamma - \frac{1}{32}\eta_{\alpha\beta}e^{3\phi/2}F_{\gamma_1\gamma_2}F^{\gamma_1\gamma_2} \\ &\quad +\frac{1}{4}e^{-\phi}F_{\alpha\gamma_1\gamma_2}F_\beta{}^{\gamma_1\gamma_2} - \frac{1}{48}\eta_{\alpha\beta}e^{-\phi}F_{\gamma_1\gamma_2\gamma_3}F^{\gamma_1\gamma_2\gamma_3} \\ &\quad +\frac{1}{12}e^{\phi/2}F_{\alpha\gamma_1\gamma_2\gamma_3}F_\beta{}^{\gamma_1\gamma_2\gamma_3} - \frac{1}{128}\eta_{\alpha\beta}e^{\phi/2}F_{\gamma_1\gamma_2\gamma_3\gamma_4}F^{\gamma_1\gamma_2\gamma_3\gamma_4}. \end{aligned} \tag{A.1.6b}$$

A.1.4 Fermionic Equations of Motion

Besides the bosonic equations one also deduces the fermionic equations of motion from the Lagrangian (7.1.1) which we write out in flat indices. For the dilatino this gives

$$\begin{aligned}
\Gamma^{\alpha} D_{\alpha}\lambda \quad & - \quad \frac{5}{32} e^{3\phi/4} F_{\alpha_1\alpha_2}\Gamma^{\alpha_1\alpha_2}\Gamma_{10}\lambda + \frac{3}{16} e^{3\phi/4} F_{\alpha_1\alpha_2}\Gamma^{\beta}\Gamma^{\alpha_1\alpha_2}\Gamma_{10}\psi_{\beta} \\
& + \quad \frac{1}{24} e^{-\phi/2} F_{\alpha_1\cdots\alpha_3}\Gamma^{\beta}\Gamma^{\alpha_1\cdots\alpha_3}\Gamma_{10}\psi_{\beta} \\
& + \quad \frac{1}{128} e^{\phi/4} F_{\alpha_1\cdots\alpha_4}\Gamma^{\alpha_1\cdots\alpha_4}\lambda - \frac{1}{192} e^{\phi/4} F_{\alpha_1\cdots\alpha_4}\Gamma^{\beta}\Gamma^{\alpha_1\cdots\alpha_4}\psi_{\beta} \\
& - \quad \frac{1}{2}\partial_{\alpha}\phi\Gamma^{\beta}\Gamma^{\alpha}\psi_{\beta} \\
& - \quad \frac{21}{16} m e^{5\phi/4}\lambda - \frac{5}{8} m e^{5\phi/4}\Gamma^{\alpha}\psi_{\alpha} = 0,
\end{aligned} \tag{A.1.7}$$

The gravitino equation obtained directly from the variation of (7.1.1) is of the form

$$\mathcal{E}^{\alpha} = \Gamma^{\alpha\beta\gamma} D_{\beta}\psi_{\gamma} + R^{\alpha} = 0. \tag{A.1.8}$$

After multiplication with two gamma matrices it can be rewritten as

$$E_{\alpha} = \Gamma^{\beta}(D_{\alpha}\psi_{\beta} - D_{\beta}\psi_{\alpha}) + L_{\alpha} = 0, \tag{A.1.9}$$

where

$$L_{\alpha} = \frac{1}{8}(\Gamma_{\alpha\beta}R^{\beta} - 7R_{\alpha}), \tag{A.1.10}$$

Although $\mathcal{E}^{\alpha} = 0$ is equivalent to $E^{\alpha} = 0$, the spatial components E^{a} and $\mathcal{E}^{a}$ are only equivalent when the supersymmetry constraint $\mathcal{E}^{0}$ is also taken into account. It turns out that the dynamical equation that corresponds directly to the $K(E_{10})$ Dirac equation is $E^{a} = 0$, not too surprisingly since in this form one obtains directly a Dirac equation for ψ_{a}. The

$SO(1,9)$ covariant equation $E_\alpha = 0$ reads explicitly as follows:

$$
\begin{aligned}
\Gamma^\beta\left(D_\alpha\psi_\beta - D_\beta\psi_\alpha\right) \quad &+ \quad \frac{21}{64}e^{3\phi/4}F_{\alpha\beta}\Gamma^\beta\Gamma_{10}\lambda - \frac{3}{128}e^{3\phi/4}F_{\beta_1\beta_2}\Gamma_\alpha{}^{\beta_1\beta_2}\Gamma_{10}\lambda \\
&+ \quad \frac{1}{64}e^{3\phi/4}F_{\beta_1\beta_2}\Gamma_\alpha{}^{\beta_1\beta_2\gamma}\Gamma_{10}\psi^\gamma - \frac{1}{32}e^{3\phi/4}F_{\beta_1\beta_2}\Gamma_\alpha{}^{\beta_1}\Gamma_{10}\psi^{\beta_2} \\
&- \quad \frac{7}{32}e^{3\phi/4}F_{\alpha\beta}\Gamma^{\beta\gamma}\Gamma_{10}\psi_\gamma - \frac{7}{64}e^{3\phi/4}F_{\beta_1\beta_2}\Gamma^{\beta_1\beta_2}\Gamma_{10}\psi_\alpha \\
&+ \quad \frac{7}{32}e^{3\phi/4}F_{\alpha\beta}\Gamma_{10}\psi^\beta \\
&+ \quad \frac{1}{96}e^{-\phi/2}F_{\beta_1\beta_2\beta_3}\Gamma_\alpha{}^{\beta_1\beta_2\beta_3}\Gamma_{10}\lambda - \frac{3}{32}e^{-\phi/2}F_{\alpha\beta_1\beta_2}\Gamma^{\beta_1\beta_2}\Gamma_{10}\lambda \\
&+ \quad \frac{1}{96}e^{-\phi/2}F_{\beta_1\beta_2\beta_3}\Gamma_\alpha{}^{\beta_1\beta_2\beta_3\gamma}\Gamma_{10}\psi^\gamma + \frac{1}{32}e^{-\phi/2}F_{\beta_1\beta_2\beta_3}\Gamma^{\beta_1\beta_2\beta_3}\Gamma_{10}\psi_\alpha \\
&- \quad \frac{1}{32}e^{-\phi/2}F_{\beta_1\beta_2\beta_3}\Gamma_\alpha{}^{\beta_1\beta_2}\Gamma_{10}\psi^{\beta_3} - \frac{3}{32}e^{-\phi/2}F_{\alpha\beta_1\beta_2}\Gamma^{\beta_1\beta_2\gamma}\Gamma_{10}\psi_\gamma \\
&+ \quad \frac{3}{16}e^{-\phi/2}F_{\alpha\beta_1\beta_2}\Gamma^{\beta_1}\Gamma_{10}\psi^{\beta_2} \\
&- \quad \frac{1}{512}e^{\phi/4}F_{\beta_1\ldots\beta_4}\Gamma_\alpha{}^{\beta_1\ldots\beta_4}\lambda + \frac{5}{384}e^{\phi/4}F_{\alpha\beta_1\ldots\beta_3}\Gamma^{\beta_1\ldots\beta_3}\lambda \\
&- \quad \frac{1}{256}e^{\phi/4}F_{\beta_1\ldots\beta_4}\Gamma_\alpha{}^{\beta_1\ldots\beta_4\gamma}\psi^\gamma + \frac{5}{768}e^{\phi/4}F_{\beta_1\ldots\beta_4}\Gamma^{\beta_1\ldots\beta_4}\psi_\alpha \\
&+ \quad \frac{1}{64}e^{\phi/4}F_{\beta_1\ldots\beta_4}\Gamma_\alpha{}^{\beta_1\ldots\beta_3}\psi^{\beta_4} + \frac{5}{192}e^{\phi/4}F_{\alpha\beta_1\ldots\beta_3}\Gamma^{\beta_1\ldots\beta_3\gamma}\psi_\gamma \\
&- \quad \frac{5}{64}e^{\phi/4}F_{\alpha\beta_1\ldots\beta_3}\Gamma^{\beta_1\beta_2}\psi^{\beta_3} + \frac{1}{2}\partial_\alpha\phi\lambda \\
&+ \quad \frac{1}{32}e^{5\phi/4}m\Gamma_{\alpha\beta}\psi^\beta + \frac{9}{32}e^{5\phi/4}m\psi_\alpha + \frac{5}{64}e^{5\phi/4}m\Gamma_\alpha\lambda = 0\,. \qquad \text{(A.1.11)}
\end{aligned}
$$

A.1.5 Truncation on the Supergravity Side

As explained in Section 7.3.1 of the main text, the correspondence between the dynamics of the E_{10}-invariant sigma model and the dynamics of type IIA supergravity only works if a certain truncation is applied.

As dictated from the BKL analysis, the following truncations must be imposed on the equations of motion and Bianchi identities:

$$\partial_a(N\partial_0\phi) = \partial_a(N\omega_{0\,ab}) = \partial_a(N\omega_{a\,b0}) = \partial_a(N\partial_b\phi) = \partial_a(N\omega_{b\,cd}) = 0 \qquad \text{(A.1.12a)}$$

and

$$
\begin{aligned}
\partial_a\left(Ne^{3\phi/4}F_{0b}\right) = \partial_a\left(Ne^{-\phi/2}F_{0b_1b_2}\right) = \partial_a\left(Ne^{\phi/4}F_{0b_1b_2b_3}\right) &= 0\,, \\
\partial_a\left(Ne^{3\phi/4}F_{b_1b_2}\right) = \partial_a\left(Ne^{-\phi/2}F_{b_1b_2b_3}\right) = \partial_a\left(Ne^{\phi/4}F_{b_1b_2b_3b_4}\right) &= 0\,, \\
\partial_a\left(Ne^{5\phi/4}m\right) &= 0\,. \qquad \text{(A.1.12b)}
\end{aligned}
$$

Furthermore, as already indicated in (7.3.2), the spatial trace of the spin connection has to be set to zero

$$\omega_{b\,ba} = 0\,. \qquad \text{(A.1.12c)}$$

Equations (A.1.12) exhaust all truncations of the bosonic variables. As explained in the text, with these truncations the bosonic geodesic equations agree with the supergravity up to one term in the Einstein equation coming from the contribution to the Ricci tensor R_{ab} which is proportional to $\Omega_{ac\,d}\Omega_{bd\,c}$. We emphasize that there are no mismatches associated with the mass parameter m.

For the fermionic variables one also needs to apply appropriate truncations of spatial gradients. These turn out to be

$$N^{-1}\partial_a(N\lambda) = N^{-1/2}\partial_a(N^{1/2}\psi_b) = 0\,. \tag{A.1.13}$$

With this choice of truncation the Dirac equation of the coset match exactly the fermionic equations of motion of supergravity if in addition the supersymmetric gauge

$$\psi_0 - \Gamma_0\Gamma^a\psi_a = 0 \tag{A.1.14}$$

of (7.3.10) is adopted.

A.1.6 Reduction from $D = 11$

Our conventions for massless type IIA supergravity are consistent with the reduction of eleven-dimensional supergravity through the reduction ansatz and field redefinitions given in this section. This construction of massless type IIA was first carried out in [5–7]. In this section only, we will denote the $D = 11$ gravitino by Ψ_M.

The supersymmetry variation of the gravitino in eleven-dimensional supergravity is

$$\delta_{\varepsilon^{(11)}}\Psi_M^{(11)} = D_M^{(11)}\varepsilon^{(11)} + \frac{1}{288}\Big(\Gamma_M{}^{N_1\cdots N_4} - 8\delta_M^{[N_1}\Gamma^{N_2N_3N_4]}\Big)\varepsilon^{(11)}F^{(11)}_{N_1\cdots N_4}. \tag{A.1.15}$$

We reduce this expression along x^{10} with the following ansatz for the eleven-dimensional vielbein:

$$E_M{}^A = \begin{pmatrix} e^{-\frac{1}{12}\phi}e_\mu{}^\alpha & e^{\frac{2}{3}\phi}A_\mu \\ 0 & e^{\frac{2}{3}\phi} \end{pmatrix}. \tag{A.1.16}$$

The four-form field strength is reduced as follows in curved indices

$$\begin{aligned} F^{(10)}_{\mu\nu\rho} &\equiv F^{(11)}_{\mu\nu\rho\tilde{10}}, \\ F^{(10)}_{\mu\nu\rho\sigma} &\equiv F^{(11)}_{\mu\nu\rho\sigma} + 4A_{[\mu}F^{(10)}_{\nu\rho\sigma]}, \end{aligned} \tag{A.1.17}$$

where $\tilde{10}$ denotes a curved index. The eleven-dimensional gravitino $\Psi_M^{(11)}$ splits into the ten-dimensional gravitino ψ_μ and the dilatino λ, according to the following field redefinitions

$$\begin{aligned} \Psi_\mu^{(11)} &= e^{-\frac{1}{24}\phi}\big(\psi_\mu - \frac{1}{12}\Gamma_\mu\lambda\big) + \frac{2}{3}e^{\frac{1}{24}\phi}\Gamma_{\tilde{10}}A_\mu\lambda, \\ \Gamma^{\tilde{10}}\Psi_{\tilde{10}}^{(11)} &= \frac{2}{3}e^{\frac{1}{24}\phi}\lambda. \end{aligned} \tag{A.1.18}$$

We also rescale the supersymmetry parameter according to

$$\varepsilon \equiv e^{\frac{1}{24}\phi}\varepsilon^{(11)}. \tag{A.1.19}$$

(ℓ_1, ℓ_2)	$\mathfrak{sl}(9,\mathbb{R})$ Dynkin labels	$\mathfrak{e}_{10}$ root α of lowest weight
$(0,0)$	$[1,0,0,0,0,0,0,1] \oplus [0,0,0,0,0,0,0,0]$	$(-1,-1,-1,-1,-1,-1,-1,-1,0,0)$
$(0,0)$	$[0,0,0,0,0,0,0,0]$	$(0,0,0,0,0,0,0,0,0,0)$
$(0,1)$	$[0,0,0,0,0,0,0,1]$	$(0,0,0,0,0,0,0,0,1,0)$
$(1,0)$	$[0,0,0,0,0,0,1,0]$	$(0,0,0,0,0,0,0,0,0,1)$
$(1,1)$	$[0,0,0,0,0,1,0,0]$	$(0,0,0,0,0,0,1,1,1,1)$
$(2,1)$	$[0,0,0,1,0,0,0,0]$	$(0,0,0,0,1,2,3,2,1,2)$
$(2,2)$	$[0,0,1,0,0,0,0,0]$	$(0,0,0,1,2,3,4,3,2,2)$
$(3,1)$	$[0,1,0,0,0,0,0,0]$	$(0,0,1,2,3,4,5,3,1,3)$
$(3,2)$	$[1,0,0,0,0,0,0,0]$	$(0,1,2,3,4,5,6,4,2,3)$
$(3,2)$	$[0,1,0,0,0,0,0,1]$	$(0,0,1,2,3,4,5,3,2,3)$
$(4,1)$	$[0,0,0,0,0,0,0,0]$	$(1,2,3,4,5,6,7,4,1,4)$
$(3,3)$	$[1,0,0,0,0,0,0,1]$	$(0,1,2,3,4,5,6,4,3,3)$

Table A.1: $\mathfrak{sl}(9,\mathbb{R})$ *level decomposition of* $\mathfrak{e}_{10}$ *with root vectors. All shown levels are complete. The very last entry is that of a mixed symmetry generator not studied for the dictionary of Table 7.2. Its possible relation to trombone gauging is discussed in* **Paper VI**.

A.2 Details on the IIA Level Decomposition of $\mathfrak{e}_{10}$ and $\mathfrak{k}(\mathfrak{e}_{10})$

In this appendix we give all the details of the level deomposition of $\mathfrak{e}_{10}$ with respect to $\mathfrak{sl}(9,\mathbb{R})$, up to level $(\ell_1,\ell_2) = (4,1)$. In particular we give all the relevant $\mathfrak{e}_{10}$ commutators which are needed to compute the explicit expressions of the bosonic and fermionic equations of motion in Appendix A.3. Moreover, we give details on the spinor and and vector-spinor representations of $\mathfrak{k}(\mathfrak{e}_{10})$.

A.2.1 Commutation Relations for Fields Appearing in the Dictionary

At level $(\ell_1,\ell_2) = (0,0)$ there is a copy of $\mathfrak{gl}(9,\mathbb{R}) = \mathfrak{sl}(9,\mathbb{R}) \oplus \mathbb{R}$, as well as a scalar generator associated with the dilaton. Their relations are $(a, b = 1, \dots, 9)$

$$
\begin{aligned}
&[K^a{}_b, K^c{}_d] = \delta^c_b K^a{}_d - \delta^a_d K^c{}_b\,, \qquad \langle K^a{}_b | K^c{}_d\rangle = \delta^a_d \delta^c_b - \delta^a_b \delta^c_d\,,\\
&[T, K^a{}_b] = 0\,, \qquad \langle T|T\rangle = \frac{1}{2}\,, \qquad \langle T|K^a{}_b\rangle = 0\,.
\end{aligned} \tag{A.2.1}
$$

Here, $\langle\cdot|\cdot\rangle$ is the invariant bilinear form. We define also the trace $K = \sum_{a=1}^{9} K^a{}_a$. For completeness

$$
\begin{aligned}
K &= 8h_1 + 16h_2 + 24h_3 + 32h_4 + 40h_5 + 48h_6 + 56h_7 + 37h_8 + 18h_9 + 27h_{10}\,,\\
T &= \frac{1}{2}h_1 + h_2 + \frac{3}{2}h_3 + 2h_4 + \frac{5}{2}h_5 + 3h_6 + \frac{7}{2}h_7 + \frac{25}{12}h_8 + \frac{2}{3}h_9 + \frac{23}{12}h_{10}\,.
\end{aligned} \tag{A.2.2}
$$

All objects transform as $\mathfrak{gl}(9,\mathbb{R})$ tensors in the obvious way. The T commutator relations are

$$\begin{aligned}
[T, E^{a_1}] &= \frac{3}{4} E^{a_1}\,, & [T, E^{a_1 a_2}] &= -\frac{1}{2} E^{a_1 a_2}\,,\\
[T, E^{a_1 a_2 a_3}] &= \frac{1}{4} E^{a_1 a_2 a_3}\,, & [T, E^{a_1 \dots a_5}] &= -\frac{1}{4} E^{a_1 \dots a_5}\,,\\
[T, E^{a_1 \dots a_6}] &= \frac{1}{2} E^{a_1 \dots a_6}\,, & [T, E^{a_1 \dots a_7}] &= -\frac{3}{4} E^{a_1 \dots a_7}\,,\\
[T, E^{a_1 \dots a_9}] &= -\frac{5}{4} E^{a_1 \dots a_9}\,, & &\\
\left[T, E^{a_0|a_1 \dots a_7}\right] &= 0\,, & [T, E^{a_1 \dots a_8}] &= 0\,.
\end{aligned} \tag{A.2.3}$$

The positive level generators are generated by the simple (fundamental) generators on levels $(0,1)$ and $(1,0)$ by

$$\begin{aligned}
[E^{a_1}, E^{a_2}] &= 0\,, & [E^{a_1 a_2}, E^{a_3 a_4}] &= 0\,,\\
[E^{a_1 a_2}, E^{a_3}] &= E^{a_1 a_2 a_3}\,, & [E^{a_1 a_2}, E^{a_3 \dots a_5}] &= E^{a_1 \dots a_5}\,,\\
[E^{a_1 a_2}, E^{a_3 \dots a_7}] &= E^{a_1 \dots a_7}\,, & [E^{a_1 a_2}, E^{a_3 \dots a_9}] &= E^{a_1 \dots a_9}\,,\\
[E^{a_1}, E^{a_2 \dots a_6}] &= E^{a_1 \dots a_6}\,, & [E^{a_0}, E^{a_1 \dots a_7}] &= E^{a_0|a_1 \dots a_7} + \frac{3}{2} E^{a_0 a_1 \dots a_7}\,.
\end{aligned} \tag{A.2.4}$$

These defining relations imply for example

$$\begin{aligned}
[E^{a_1 a_2 a_3}, E^{a_4 a_5 a_6}] &= -E^{a_1 \dots a_6}\,,\\
[E^{a_1 a_2}, E^{a_3 \dots a_8}] &= -2E^{[a_1|a_2]a_3 \dots a_8} + E^{a_1 \dots a_8}\,,\\
[E^{a_1 a_2 a_3}, E^{a_4 \dots a_8}] &= -3E^{[a_1|a_2 a_3]a_4 \dots a_8} - \frac{1}{2} E^{a_1 \dots a_8}\,,\\
[E^{a_1 \dots a_5}, E^{a_6 a_7 a_8}] &= 5E^{[a_1|a_2 \dots a_5]a_6 a_7 a_8} - \frac{1}{2} E^{a_1 \dots a_8}\,,\\
[E^{a_1 \dots a_6}, E^{a_7 a_8}] &= -6E^{[a_1|a_2 \dots a_6]a_7 a_8} - E^{a_1 \dots a_8}\,,\\
[E^{a_1 \dots a_7}, E^{a_8}] &= -7E^{[a_1|a_2 \dots a_7]a_8} + \frac{3}{2} E^{a_1 \dots a_8}\,.
\end{aligned} \tag{A.2.5}$$

The Young symmetry on the dual graviton implies $E^{a_0|a_1 \dots a_7} = 7E^{[a_1|a_2 \dots a_7]a_0}$. The two irreducible representations on $(3,2)$ are projected onto via

$$\begin{aligned}
E^{a_1 \dots a_8} &= \frac{2}{3}\left[E^{[a_1}, E^{a_2 \dots a_8]}\right]\,, \quad E^{a_0|a_1 \dots a_7} = \frac{7}{8}\left([E^{a_0}, E^{a_1 \dots a_7}] + \left[E^{[a_1}, E^{a_2 \dots a_7]a_0}\right]\right)\,,\\
E^{a_0|a_1 \dots a_7} &= -\frac{7}{8}\left(\left[E^{a_0[a_1}, E^{a_2 \dots a_7]}\right] + \left[E^{[a_1 a_2}, E^{a_3 \dots a_7]a_0}\right]\right)\,.
\end{aligned} \tag{A.2.6}$$

The definitions (A.2.4) are such that the normalisations are

$$\begin{aligned}
\langle E^{a_1 \dots a_p} | F_{b_1 \dots b_p} \rangle &= p!\, \delta^{a_1 \dots a_p}_{b_1 \dots b_p} & &\Rightarrow \quad \langle E^{1 \dots p} | F_{1 \dots p} \rangle = 1 \quad (p \neq 8)\,,\\
\langle E^{a_1 \dots a_8} | F_{b_1 \dots b_8} \rangle &= \frac{1}{2} \cdot 8! \delta^{a_1 \dots a_8}_{b_1 \dots b_8} & &\Rightarrow \quad \langle E^{1 \dots 8} | F_{1 \dots 8} \rangle = \frac{1}{2}\,,\\
\langle E^{a_0|a_1 \dots a_7} | F_{b_0|b_1 \dots b_7} \rangle &= \frac{7}{8} \cdot 7! \left(\delta^{a_0}_{b_0} \delta^{a_1 \dots a_7}_{b_1 \dots b_7} + \delta^{[a_1}_{b_0} \delta^{a_2 \dots a_7]a_0}_{b_1 \dots\dots b_7}\right)\,. & &
\end{aligned} \tag{A.2.7}$$

The additional factor of 2 in the normalisation of the 8-form is chosen such that all structure constants remain rational.

Defining the transposed generators via $F \equiv E^T$ as usual gives the following commutations relation between the form generators and their transposes:

$$\begin{aligned}
[E^a, F_b] &= -\frac{1}{8}\delta^a_b K + K^a{}_b + \frac{3}{2}\delta^a_b T\,,\\
[E^{a_1a_2}, F_{b_1b_2}] &= -\frac{1}{2}\delta^{a_1a_2}_{b_1b_2} K + 4\delta^{[a_1}_{[b_1} K^{a_2]}{}_{b_2]} - 2\delta^{a_1a_2}_{b_1b_2} T\,,\\
[E^{a_1a_2a_3}, F_{b_1b_2b_3}] &= -\frac{3}{8}\cdot 3!\,\delta^{a_1a_2a_3}_{b_1b_2b_3} K + 3\cdot 3!\,\delta^{[a_1a_2}_{[b_1b_2} K^{a_3]}{}_{b_3]} + 3\delta^{a_1a_2a_3}_{b_1b_2b_3} T\,,\\
[E^{a_1\dots a_5}, F_{b_1\dots b_5}] &= -\frac{5}{8}\cdot 5!\,\delta^{a_1\dots a_5}_{b_1\dots b_5} K + 5\cdot 5!\,\delta^{[a_1\dots a_4}_{[b_1\dots b_4} K^{a_5]}{}_{b_5]} - \frac{1}{2}\cdot 5!\,\delta^{a_1\dots a_5}_{b_1\dots b_5} T\,,\\
[E^{a_1\dots a_6}, F_{b_1\dots b_6}] &= -\frac{3}{4}\cdot 6!\,\delta^{a_1\dots a_6}_{b_1\dots b_6} K + 6\cdot 6!\,\delta^{[a_1\dots a_5}_{[b_1\dots b_5} K^{a_6]}{}_{b_6]} + 6!\,\delta^{a_1\dots a_6}_{b_1\dots b_6} T\,,\\
[E^{a_1\dots a_7}, F_{b_1\dots b_7}] &= -\frac{7}{8}\cdot 7!\,\delta^{a_1\dots a_7}_{b_1\dots b_7} K + 7\cdot 7!\,\delta^{[a_1\dots a_6}_{[b_1\dots b_6} K^{a_7]}{}_{b_7]} - \frac{3}{2}\cdot 7!\,\delta^{a_1\dots a_7}_{b_1\dots b_7} T\,,\\
[E^{a_1\dots a_9}, F_{b_1\dots b_9}] &= -\frac{9}{8}\cdot 9!\,\delta^{a_1\dots a_9}_{b_1\dots b_9} K + 9\cdot 9!\,\delta^{[a_1\dots a_8}_{[b_1\dots b_8} K^{a_9]}{}_{b_9]} - \frac{5}{2}\cdot 9!\,\delta^{a_1\dots a_9}_{b_1\dots b_9} T\,,\\
[E^{a_1\dots a_8}, F_{b_1\dots b_8}] &= -\frac{1}{2}\cdot 8!\,\delta^{a_1\dots a_8}_{b_1\dots b_8} K + 4\cdot 8!\,\delta^{[a_1\dots a_7}_{[b_1\dots b_7} K^{a_8]}{}_{b_8]}\,.
\end{aligned} \tag{A.2.8}$$

The commutator of the dual graviton generator $E^{a_0|a_1\dots a_7}$ can be most conveniently written using a dummy tensor $X_{a_0|a_1\dots a_7}$ as

$$\Big[F_{b_0|b_1\dots b_7}, X_{a_0|a_1\dots a_7}E^{a_0|a_1\dots a_7}\Big] = 7!\left(X_{b_0|b_1\dots b_7}K - X_{c|b_1\dots b_7}K^c{}_{b_0} - 7X_{b_0|c[b_1\dots b_6}K^c{}_{b_7]}\right). \tag{A.2.9}$$

The generators of different rank commute in the following non-trivial way:

$$\begin{aligned}
[E^a, F_{b_1b_2b_3}] &= 3\delta^a_{[b_1}F_{b_2b_3]}\,, & [E^{a_1a_2}, F_{b_1b_2b_3}] &= -6\delta^{a_1a_2}_{[b_1b_2}F_{b_3]}\,,\\
[E^{a_1a_2}, F_{b_1\dots b_5}] &= -20\delta^{a_1a_2}_{[b_1b_2}F_{b_3b_4b_5]}\,, & [E^{a_1a_2a_3}, F_{b_1\dots b_5}] &= 60\delta^{a_1a_2a_3}_{[b_1b_2b_3}F_{b_4b_5]}\,,\\
[E^a, F_{b_1\dots b_6}] &= -6\delta^a_{[b_1}F_{b_2\dots b_6]}\,, & [E^{a_1a_2a_3}, F_{b_1\dots b_6}] &= 120\delta^{a_1a_2a_3}_{[b_1b_2b_3}F_{b_4b_5b_6]}\,,\\
[E^{a_1\dots a_5}, F_{b_1\dots b_6}] &= -6!\,\delta^{a_1\dots a_5}_{[b_1\dots b_5}F_{b_6]}\,, & [E^{a_1a_2}, F_{b_1\dots b_7}] &= -7\cdot 6\delta^{a_1a_2}_{[b_1b_2}F_{b_3\dots b_7]}\,,\\
[E^{a_1\dots a_5}, F_{b_1\dots b_7}] &= \frac{1}{2}\cdot 7!\,\delta^{a_1\dots a_5}_{[b_1\dots b_5}F_{b_6b_7]}\,, & [E^{a_1a_2}, F_{b_1\dots b_9}] &= -9\cdot 8\delta^{a_1a_2}_{[b_1b_2}F_{b_3\dots b_9]}\,,\\
[E^{a_1\dots a_7}, F_{b_1\dots b_9}] &= \frac{1}{2}\cdot 9!\,\delta^{a_1\dots a_7}_{[b_1\dots b_7}F_{b_8b_9]}\,.
\end{aligned} \tag{A.2.10}$$

Anticipating the geodesic equation we know that (A.2.10) describes all the couplings between the different forms occurring in the matter equations, so that for example the 9-form (=mass term) occurs only in the Bianchi identity for $F_{(2)}$ and in the eom of $F_{(3)}$, consistent with (A.1.4) and (A.1.5). The dilaton and Einstein equation are described by the couplings of equations (A.2.8) and (A.2.9).

The commutators with the dual dilaton are

$$
\begin{aligned}
{[E^{a}, F_{b_1\dots b_8}]} &= -6\delta^{a}_{[b_1}F_{b_2\dots b_8]}\,, & {[E^{a_1a_2}, F_{b_1\dots b_8}]} &= -4\cdot 7\delta^{a_1a_2}_{[b_1b_2}F_{b_3\dots b_8]}\,,\\
{[E^{a_1a_2a_3}, F_{b_1\dots b_8}]} &= 2\cdot 7\cdot 6\delta^{a_1a_2a_3}_{[b_1b_2b_3}F_{b_4\dots b_8]}\,, & {[E^{a_1\dots a_5}, F_{b_1\dots b_8}]} &= 2\cdot 7\cdot 5!\,\delta^{a_1\dots a_5}_{[b_1\dots b_5}F_{b_6b_7b_8]}\,,\\
{[E^{a_1\dots a_6}, F_{b_1\dots b_8}]} &= 2\cdot 7!\,\delta^{a_1\dots a_6}_{[b_1\dots b_6}F_{b_7b_8]}\,, & {[E^{a_1\dots a_7}, F_{b_1\dots b_8}]} &= -6\cdot 7!\,\delta^{a_1\dots a_7}_{[b_1\dots b_7}F_{b_8]}\,,
\end{aligned}
\tag{A.2.11}
$$

whereas for the dual graviton one finds

$$
\begin{aligned}
\left[E^{a}, F_{b_0|b_1\dots b_7}\right] &= -\frac{7}{8}\left(\delta^{a}_{b_0}F_{b_1\dots b_7} + \delta^{a}_{[b_1}F_{b_2\dots b_7]b_0}\right)\,,\\
\left[E^{a_1a_2}, F_{b_0|b_1\dots b_7}\right] &= \frac{21}{2}\left(\delta^{a_1a_2}_{b_0[b_1}F_{b_2\dots b_7]} + \delta^{a_1a_2}_{[b_1b_2}F_{b_3\dots b_7]b_0}\right)\,,\\
\left[E^{a_1a_2a_3}, F_{b_0|b_1\dots b_7}\right] &= \frac{45\cdot 7}{4}\left(\delta^{a_1a_2a_3}_{b_0[b_1b_2}F_{b_3\dots b_7]} + \delta^{a_1a_2a_3}_{[b_1b_2b_3}F_{b_4\dots b_7]b_0}\right)\,,\\
\left[E^{a_1\dots a_5}, F_{b_0|b_1\dots b_7}\right] &= -\frac{5\cdot 7!}{16}\left(\delta^{a_1\dots a_5}_{b_0[b_1\dots b_4}F_{b_5\dots b_7]} + \delta^{a_1\dots a_5}_{[b_1\dots b_5}F_{b_6b_7]b_0}\right)\,,\\
\left[E^{a_1\dots a_6}, F_{b_0|b_1\dots b_7}\right] &= \frac{3\cdot 7!}{4}\left(\delta^{a_1\dots a_6}_{b_0[b_1\dots b_5}F_{b_6b_7]} + \delta^{a_1\dots a_6}_{[b_1\dots b_6}F_{b_7]b_0}\right)\,,\\
\left[E^{a_1\dots a_7}, F_{b_0|b_1\dots b_7}\right] &= \frac{7\cdot 7!}{8}\left(\delta^{a_1\dots a_7}_{b_0[b_1\dots b_6}F_{b_7]} + \delta^{a_1\dots a_7}_{b_1\dots b_7}F_{b_0}\right)\,.
\end{aligned}
\tag{A.2.12}
$$

A.2.2 Spinor Representations of $\mathfrak{k}(\mathfrak{e}_{10})$

For the E_{10} model to incorporate all low energy limits of M-theory in a single model, the fermionic representations used in (7.2.36) should not depend on the particular supergravity one wishes to study. Rather the unfaithful **320** and **32** representations of $\mathfrak{k}(\mathfrak{e}_{10})$ should be decomposed under a suitable subalgebra. Here, this subalgebra is $\mathfrak{so}(9) \subset \mathfrak{gl}(9,\mathbb{R})$ and this appendix provides the details of the action of the $\mathfrak{k}(\mathfrak{e}_{10})$ generators in this basis. When doing the following calculations we found the computer package GAMMA [248] useful.[1]

Dirac Spinor

The result of writing the $K(E_{10})$ action on the Dirac spinor is (we recall the notation $M^{a_1a_2}$ for the level $(0,0)$ generator of $K(E_{10})$ from (7.2.17))

$$
\begin{aligned}
M^{a_1a_2}\cdot\epsilon &= \frac{1}{2}\Gamma^{a_1a_2}\epsilon, & J^{a}_{(0,1)}\cdot\epsilon &= \frac{1}{2}\Gamma_{10}\Gamma^{a}\epsilon,\\
J^{a_1a_2}_{(1,0)}\cdot\epsilon &= \frac{1}{2}\Gamma_{10}\Gamma^{a_1a_2}\epsilon, & J^{a_1a_2a_3}_{(1,1)}\cdot\epsilon &= \frac{1}{2}\Gamma^{a_1a_2a_3}\epsilon,\\
J^{a_1\cdots a_5}_{(2,1)}\cdot\epsilon &= \frac{1}{2}\Gamma_{10}\Gamma^{a_1\cdots a_5}\epsilon, & J^{a_1\cdots a_6}_{(2,2)}\cdot\epsilon &= -\frac{1}{2}\Gamma^{a_1\cdots a_6}\epsilon,\\
J^{a_1\cdots a_7}_{(3,1)}\cdot\epsilon &= \frac{1}{2}\Gamma^{a_1\cdots a_7}\epsilon, & J^{a_0|a_1\cdots a_7}_{(3,2)}\cdot\epsilon &= \frac{7}{2}\Gamma_{10}\delta^{[a_1}_{a_0}\Gamma^{a_2\cdots a_7]}\epsilon,\\
J^{a_1\cdots a_8}_{(3,2)}\cdot\epsilon &= 0, & J^{a_1\cdots a_9}_{(4,1)}\cdot\epsilon &= \frac{1}{2}\Gamma_{10}\Gamma^{a_1\cdots a_9}\epsilon,
\end{aligned}
\tag{A.2.13}
$$

[1] We are grateful to Ulf Gran for generous help in modifying GAMMA to suit our needs.

where the generators above levels $(0,1)$ and $(1,0)$ are defined through the lower levels as follows

$$
\begin{aligned}
J^{a_1a_2a_3}_{(1,1)}\cdot\epsilon &:= \Big[J^{[a_1a_2}_{(1,0)}, J^{a_3]}_{(0,1)}\Big]\cdot\epsilon,\\
J^{a_1\cdots a_5}_{(2,1)}\cdot\epsilon &:= \Big[J^{[a_1a_2}_{(1,0)}, J^{a_3a_4a_5]}_{(1,1)}\Big]\cdot\epsilon,\\
J^{a_1\cdots a_6}_{(2,2)}\cdot\epsilon &:= \Big[J^{[a_1}_{(1,1)}, J^{a_2\ldots a_6]}_{(1,1)}\Big]\cdot\epsilon,\\
J^{a_1\cdots a_7}_{(3,1)}\cdot\epsilon &:= \Big[J^{[a_1a_2}_{(1,0)}, J^{a_3\cdots a_7]}_{(2,1)}\Big]\cdot\epsilon,\\
J^{a_0|a_1\cdots a_7}_{(3,2)}\cdot\epsilon &:= -\frac{7}{8}\Big(\Big[J^{a_0}_{(0,1)}, J^{a_1\cdots a_7}_{(3,1)}\Big]+\Big[J^{[a_1}_{(0,1)}, J^{a_2\cdots a_7]a_0}_{(3,1)}\Big]\Big)\cdot\epsilon,\\
J^{a_1\cdots a_8}_{(3,2)}\cdot\epsilon &:= \Big[J^{[a_1a_2}_{(1,0)}, J^{a_3\cdots a_8]}_{(2,2)}\Big]\cdot\epsilon,\\
J^{a_1\cdots a_9}_{(4,1)}\cdot\epsilon &:= \Big[J^{[a_1a_2}_{(1,0)}, J^{a_3\cdots a_9]}_{(3,1)}\Big]\cdot\epsilon.
\end{aligned}
\tag{A.2.14}
$$

We stress that this kind of construction is guaranteed to yield a consistent unfaithful representation of all of $\mathfrak{k}(\mathfrak{e}_{10})$ given that a few simple consistency conditions between the lowest level fundamental generators are satisfied [72]. That these conditions are satisfied here can be checked easily directly, but it also follows from the branching of the transformation rules given in [69, 70].

Vector-Spinor

By reduction of the transformation rules of [70, 71] one obtains that the fundamental $\mathfrak{k}(\mathfrak{e}_{10})$-generators act on the vector spinor representation as follows on the Ψ_{10} component

$$
\begin{aligned}
J^{a}_{(0,1)}\cdot\Psi_{10} &= \frac{1}{2}\Gamma_{10}\Gamma^{a}\Psi_{10}+\Psi^{a}\\
J^{a_1a_2}_{(1,0)}\cdot\Psi_{10} &= \frac{1}{6}\Gamma_{10}\Gamma^{a_1a_2}\Psi_{10}+\frac{4}{3}\Gamma^{[a_1}\Psi^{a_2]}.
\end{aligned}
$$

On the Ψ_a $(a=1,\ldots,9)$ component they act as

$$
\begin{aligned}
J^{a}_{(0,1)}\cdot\Psi_{b} &= \frac{1}{2}\Gamma_{10}\Gamma^{a}\Psi_{b}-\delta^{a}_{b}\Psi_{10}\\
J^{a_1a_2}_{(1,0)}\cdot\Psi_{b} &= \frac{1}{2}\Gamma_{10}\Gamma^{a_1a_2}\Psi_{b}-\frac{4}{3}\Gamma_{10}\delta^{[a_1}_{b}\Psi^{a_2]}+\frac{2}{3}\Gamma_{10}\Gamma_{b}{}^{[a_1}\Psi^{a_2]}\\
&\quad+\frac{4}{3}\delta^{[a_1}_{b}\Gamma^{a_2]}\Psi_{10}-\frac{1}{3}\Gamma_{b}{}^{a_1a_2}\Psi_{10}.
\end{aligned}
$$

The other levels action is then computed to be on Ψ_{10} as

$$\begin{aligned}
M^{a_1a_2}\cdot\Psi_{10} &= \frac{1}{2}\Gamma^{a_1a_2}\Psi_{10}\\
J_{(1,1)}^{a_1\cdots a_3}\cdot\Psi_{10} &= \frac{1}{2}\Gamma^{a_1\cdots a_3}\Psi_{10}-\Gamma_{10}\Gamma^{[a_1a_2}\Psi^{a_3]}\\
J_{(2,1)}^{a_1\cdots a_5}\cdot\Psi_{10} &= -\frac{1}{6}\Gamma_{10}\Gamma^{a_1\cdots a_5}\Psi_{10}-\frac{5}{3}\Gamma^{[a_1\cdots a_4}\Psi^{a_5]}\\
J_{(2,2)}^{a_1\cdots a_6}\cdot\Psi_{10} &= -\frac{1}{2}\Gamma^{a_1\cdots a_6}\Psi_{10}-4\Gamma_{10}\Gamma^{[a_1\cdots a_5}\Psi^{a_6]}\\
J_{(3,1)}^{a_1\cdots a_7}\cdot\Psi_{10} &= -\frac{3}{2}\Gamma^{a_1\cdots a_7}\Psi_{10}+7\Gamma_{10}\Gamma^{[a_1\cdots a_6}\Psi^{a_7]}\\
J_{(3,2)}^{a_1\cdots a_8}\cdot\Psi_{10} &= -\frac{4}{3}\Gamma_{10}\Gamma^{a_1\cdots a_8}\Psi_{10}-\frac{32}{3}\Gamma^{[a_1\cdots a_7}\Psi^{a_8]}\\
J_{(3,2)}^{a_0|a_1\cdots a_7}\cdot\Psi_{10} &= -\frac{7}{2}\Gamma_{10}\delta_{a_0}^{[a_1}\Gamma^{a_2\cdots a_7]}\Psi_{10}\\
J_{(4,1)}^{a_1\cdots a_9}\cdot\Psi_{10} &= \frac{9}{2}\Gamma_{10}\Gamma^{a_1\cdots a_9}\Psi_{10}-15\Gamma^{[a_1\cdots a_8}\Psi^{a_9]}.
\end{aligned}\tag{A.2.15}$$

On Ψ_a we obtain similarly

$$\begin{aligned}
M^{a_1a_2}\cdot\Psi_b &= \frac{1}{2}\Gamma^{a_1a_2}\Psi_b+2\delta_b^{[a_1}\Psi^{a_2]}\\
J_{(1,1)}^{a_1\cdots a_3}\cdot\Psi_b &= \frac{1}{2}\Gamma^{a_1\cdots a_3}\Psi_b+4\delta_b^{[a_1}\Gamma^{a_2}\Psi^{a_3]}-\Gamma_b{}^{[a_1a_2}\Psi^{a_3]}\\
J_{(2,1)}^{a_1\cdots a_5}\cdot\Psi_b &= \frac{1}{2}\Gamma_{10}\Gamma^{a_1\cdots a_5}\Psi_b+\frac{20}{3}\Gamma_{10}\delta_b^{[a_1}\Gamma^{a_2\cdots a_4}\Psi^{a_5]}-\frac{10}{3}\Gamma_{10}\Gamma_b{}^{[a_1\cdots a_4}\Psi^{a_5]}\\
&\quad+\frac{5}{3}\delta_b^{[a_1}\Gamma^{a_2\cdots a_5]}\Psi_{10}-\frac{2}{3}\Gamma_b{}^{a_1\cdots a_5}\Psi_{10}\\
J_{(2,2)}^{a_1\cdots a_6}\cdot\Psi_b &= -\frac{1}{2}\Gamma^{a_1\cdots a_6}\Psi_b+10\delta_b^{[a_1}\Gamma^{a_2\cdots a_5}\Psi^{a_6]}-4\Gamma_b{}^{[a_1\cdots a_5}\Psi^{a_6]}\\
J_{(3,1)}^{a_1\cdots a_7}\cdot\Psi_b &= \frac{1}{2}\Gamma^{a_1\cdots a_7}\Psi_b-7\Gamma_b{}^{[a_1\cdots a_6}\Psi^{a_7]}-2\Gamma_{10}\Gamma_b{}^{a_1\cdots a_7}\Psi_{10}\\
J_{(3,2)}^{a_1\cdots a_8}\cdot\Psi_b &= -\frac{4}{3}\Gamma_{10}\Gamma_b{}^{[a_1\cdots a_7}\Psi^{a_8]}-\frac{28}{3}\Gamma_{10}\delta_b^{[a_1}\Gamma^{a_2\cdots a_7}\Psi^{a_8]}\\
&\quad-\frac{4}{3}\Gamma_b{}^{a_1\cdots a_8}\Psi_{10}+\frac{4}{3}\delta_b^{[a_1}\Gamma^{a_2\cdots a_8]}\Psi_{10}\\
J_{(3,2)}^{a_0|a_1\cdots a_7}\cdot\Psi_b &= \frac{7}{2}\Gamma_{10}\delta_{a_0}^{[a_1}\Gamma^{a_2\cdots a_7]}\Psi_b-\frac{21}{2}\Gamma_{10}\delta_b^{[a_1}\Gamma_{a_0}{}^{a_2\cdots a_6}\Psi^{a_7]}-\frac{49}{4}\Gamma_{10}\delta_b^{a_0}\Gamma^{[a_1\cdots a_6}\Psi^{a_7]}\\
&\quad+42\Gamma_{10}\delta_{a_0}^{[a_1}\Gamma_b{}^{a_2\cdots a_6}\Psi^{a_7]}-7\delta_{a_0}^{[a_1}\Gamma_b{}^{a_2\cdots a_7]}\Psi_{10}\\
&\quad+\frac{7}{4}\Big(\Gamma_{10}\Gamma_b{}^{a_0[a_1\cdots a_6}\Psi^{a_7]}-\Gamma_{10}\delta_b^{[a_1}\Gamma^{a_2\cdots a_7]}\Psi^{a_0}\\
&\qquad+\Gamma_{10}\Gamma_b{}^{a_1\cdots a_7}\Psi^{a_0}+\delta_b^{[a_1}\Gamma^{a_2\cdots a_7]a_0}\Psi_{10}+\delta_b^{a_0}\Gamma^{a_1\cdots a_7}\Psi_{10}\Big)\\
J_{(4,1)}^{a_1\cdots a_9}\cdot\Psi_b &= \frac{1}{2}\Gamma_{10}\Gamma^{a_1\cdots a_9}\Psi_b-12\Gamma_{10}\Gamma_b{}^{[a_1\cdots a_8}\Psi^{a_9]}\\
&\quad-24\Gamma_{10}\delta_b^{[a_1}\Gamma^{a_2\cdots a_8}\Psi^{a_9]}-9\delta_b^{[a_1}\Gamma^{a_2\cdots a_9]}\Psi_{10}.
\end{aligned}\tag{A.2.16}$$

We note that the mixed symmetry generator at level $(3,2)$ indeed satisfies

$$J^{[a_0|a_1\cdots a_7]}_{(3,2)} \cdot \Psi_b = 0 \tag{A.2.17}$$

as desired.

A.3 Equations of Motion of the $E_{10}/K(E_{10})$ Coset Model

Using all the explicit commutators and representations of the Appendix A.2, we can now write out the bosonic and fermionic equations of motion in their full glory.

A.3.1 Bosonic Equations of Motion

The level zero equations of motion read

$$\begin{aligned}
\partial_t^2\phi &= e^{3\phi/2}P_aP_a - 2e^{-\phi}P_{a_1a_2}P_{a_1a_2} + \frac{1}{3}e^{\phi/2}P_{a_1a_2a_3}P_{a_1a_2a_3} \\
&\quad -\frac{2}{5!}e^{-\phi/2}P_{a_1\ldots a_5}P_{a_1\ldots a_5} + \frac{4}{6!}e^{\phi}P_{a_1\ldots a_6}P_{a_1\ldots a_6} \\
&\quad -\frac{6}{7!}e^{-3\phi/2}P_{a_1\ldots a_7}P_{a_1\ldots a_7} - \frac{10}{9!}e^{-5\phi/2}P_{a_1\ldots a_9}P_{a_1\ldots a_9} \qquad \text{(A.3.1a)}\\
D^{(0)}p_{ab} &= -\frac{1}{4}e^{3\phi/2}\delta_{ab}P_cP_c + 2e^{3\phi/2}P_aP_b \\
&\quad -\frac{1}{4}e^{-\phi}\delta_{ab}P_{cd}P_{cd} + 2e^{-\phi}P_{ca}P_{cb} \\
&\quad -\frac{1}{8}e^{\phi/2}\delta_{ab}P_{c_1c_2c_3}P_{c_1c_2c_3} + e^{\phi/2}P_{ac_1c_2}P_{bc_1c_2} \\
&\quad -\frac{1}{4\cdot 4!}e^{-\phi/2}\delta_{ab}P_{c_1\cdots c_5}P_{c_1\cdots c_5} + \frac{2}{4!}e^{-\phi/2}P_{ac_1\cdots c_4}P_{bc_1\cdots c_4} \\
&\quad -\frac{3}{2\cdot 6!}e^{\phi}\delta_{ab}P_{c_1\cdots c_6}P_{c_1\cdots c_6} + \frac{2}{5!}e^{\phi}P_{ac_1\cdots c_5}P_{bc_1\cdots c_5} \\
&\quad -\frac{1}{4\cdot 6!}e^{-3\phi/2}\delta_{ab}P_{c_1\cdots c_7}P_{c_1\cdots c_7} + \frac{2}{6!}e^{-3\phi/2}P_{ac_1\cdots c_6}P_{bc_1\cdots c_6} \\
&\quad -\frac{1}{8!}\delta_{ab}P_{c_1\cdots c_8}P_{c_1\cdots c_8} + \frac{1}{7!}P_{ac_1\cdots c_7}P_{bc_1\cdots c_7} \\
&\quad +\frac{1}{4\cdot 8!}\big(- \delta_{ab}P_{c_0|c_1\cdots c_7}P_{c_0|c_1\cdots c_7} + P_{a|c_1\cdots c_7}P_{b|c_1\cdots c_7} \\
&\quad +7P_{c_0|c_1\cdots c_6a}P_{c_0|c_1\cdots c_6b}\big) \\
&\quad -\frac{1}{4\cdot 8!}e^{-5\phi/2}\delta_{ab}P_{c_1\cdots c_9} + \frac{2}{8!}e^{-5\phi/2}P_{ac_1\cdots c_8}P_{bc_1\cdots c_8}\,. \qquad \text{(A.3.1b)}
\end{aligned}$$

For the higher level fields the equations of motion become

$$\begin{aligned}
D^{(0)}(e^{3\phi/2}P_a) &= -e^{\phi/2}P_{ac_1c_2}P_{c_1c_2} + \frac{2}{5!}e^{\phi}P_{ac_1\ldots c_5}P_{c_1\ldots c_5} + \frac{12}{8!}P_{ac_1\ldots c_7}P_{c_1\ldots c_7} \\
&\quad -\frac{7}{4\cdot 8!}(P_{c_1|ac_2\ldots c_7}P_{c_1\ldots c_7} + P_{a|c_1\ldots c_7}P_{c_1\ldots c_7})\,, \qquad \text{(A.3.2a)}
\end{aligned}$$

$$\begin{aligned}
D^{(0)}(e^{-\phi}P_{a_1a_2}) &= 2e^{\phi/2}P_{a_1a_2c}P_c + \frac{1}{3}e^{-\phi/2}P_{a_1a_2c_1c_2c_3}P_{c_1c_2c_3} \\
&\quad + \frac{2}{5!}e^{-3\phi/2}P_{a_1a_2c_1\dots c_5}P_{c_1\dots c_5} + \frac{2}{7!}e^{-5\phi/2}P_{a_1a_2c_1\dots c_7}P_{c_1\dots c_7} \\
&\quad + \frac{1}{6!}P_{a_1a_2c_1\dots c_6}P_{c_1\dots c_6} \\
&\quad + \frac{3}{8\cdot 6!}(P_{c_1|a_1a_2c_2\dots c_6}P_{c_1\dots c_6} + P_{a_1|a_2c_1\dots c_6}P_{c_1\dots c_6})\,, && \text{(A.3.2b)} \\
D^{(0)}(e^{\phi/2}P_{a_1a_2a_3}) &= -e^{-\phi/2}P_{a_1a_2a_3c_1c_2}P_{c_1c_2} - \frac{1}{3}e^{\phi}P_{a_1a_2a_3c_1c_2c_3}P_{c_1c_2c_3} \\
&\quad - \frac{1}{2\cdot 5!}P_{a_1a_2a_3c_1\dots c_5}P_{c_1\dots c_5} \\
&\quad + \frac{1}{256}(P_{c_1|a_1a_2a_3c_2\dots c_5}P_{c_1\dots c_5} + P_{a_1|a_2a_3c_1\dots c_5}P_{c_1\dots c_5})\,, && \text{(A.3.2c)} \\
D^{(0)}(e^{-\phi/2}P_{a_1\dots a_5}) &= 2e^{\phi}P_{a_1\dots a_5c}P_c - e^{-3\phi/2}P_{a_1\dots a_4c_1c_2}P_{c_1c_2} \\
&\quad - \frac{1}{12}P_{a_1\dots a_5c_1c_2c_3}P_{c_1c_2c_3} \\
&\quad + \frac{1}{28}(P_{c_1|a_1\dots a_5c_2c_3}P_{c_1\dots c_3} + P_{a_1|a_2\dots a_5c_1\dots c_3}P_{c_1\dots c_3})\,, && \text{(A.3.2d)} \\
D^{(0)}(e^{\phi}P_{a_1\dots a_6}) &= -\frac{1}{2}P_{a_1\dots a_6c_1c_2}P_{c_1c_2} \\
&\quad - \frac{3}{16}(P_{c_1|a_1\dots a_6c_2}P_{c_1c_2} + P_{a_1|a_2\dots a_6c_1c_2}P_{c_1c_2})\,, && \text{(A.3.2e)} \\
D^{(0)}(e^{-3\phi/2}P_{a_1\dots a_7}) &= -e^{-5\phi/2}P_{a_1\dots a_7c_1c_2}P_{c_1c_2} + \frac{3}{2}P_{a_1\dots a_7c}P_c \\
&\quad - \frac{7}{32}(P_{c|a_1\dots a_7}P_c + P_{a_1|a_2\dots a_7c}P_c)\,, && \text{(A.3.2f)} \\
D^{(0)}(e^{-5\phi/2}P_{a_1\dots a_9}) &= 0\,, && \text{(A.3.2g)} \\
D^{(0)}P_{a_1\dots a_8} &= 0\,, && \text{(A.3.2h)} \\
D^{(0)}P_{a_0|a_1\dots a_7} &= 0\,. && \text{(A.3.2i)}
\end{aligned}$$

A.3.2 Fermionic Equations of Motion

The fermionic sector of the E_{10}-invariant Lagrangian involves a Dirac-type kinetic term for the 320-dimensional vector-spinor representation Ψ of $\mathfrak{k}(\mathfrak{e}_{10})$ which was given in (7.2.36). The resulting Dirac equation can be evaluated for both the Ψ_a and the Ψ_{10} components as in (7.2.38) using the expressions for the $K(E_{10})$ action which were derived in Appendix A.2.2.

The result for the Ψ_{10} component up to level $(4,1)$ is

$$\begin{aligned}
0 &= \partial_t\Psi_{10} - \frac{1}{4}q_{a_1a_2}\Gamma^{a_1a_2}\Psi_{10} \\
&- \frac{1}{2}e^{3\phi/4}P_a\Gamma_{10}\Gamma^a\Psi_{10} - e^{3\phi/4}P_a\Psi^a \\
&- \frac{1}{12}e^{-\phi/2}P_{a_1a_2}\Gamma_{10}\Gamma^{a_1a_2}\Psi_{10} - \frac{2}{3}e^{-\phi/2}P_{a_1a_2}\Gamma^{a_1}\Psi^{a_2}
\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{12}e^{\phi/4}P_{a_1a_2a_3}\Gamma^{a_1\cdots a_3}\Psi_{10} + \frac{1}{6}e^{\phi/4}P_{a_1a_2a_3}\Gamma_{10}\Gamma^{a_1a_2}\Psi^{a_3}\\
&+ \frac{1}{6!}e^{-\phi/4}P_{a_1\cdots a_5}\Gamma_{10}\Gamma^{a_1\cdots a_5}\Psi_{10} + \frac{1}{3\cdot 4!}e^{-\phi/4}P_{a_1\cdots a_5}\Gamma^{a_1\cdots a_4}\Psi^{a_5}\\
&+ \frac{1}{2\cdot 6!}e^{\phi/2}P_{a_1\cdots a_6}\Gamma^{a_1\cdots a_6}\Psi_{10} + \frac{1}{180}e^{\phi/2}P_{a_1\cdots a_6}\Gamma_{10}\Gamma^{a_1\cdots a_5}\Psi^{a_6}\\
&+ \frac{3}{2\cdot 7!}e^{-3\phi/4}P_{a_1\cdots a_7}\Gamma^{a_1\cdots a_7}\Psi_{10} - \frac{1}{6!}e^{-3\phi/4}P_{a_1\cdots a_7}\Gamma_{10}\Gamma^{a_1\cdots a_6}\Psi^{a_7}\\
&+ \frac{4}{3\cdot 8!}P_{a_1\cdots a_8}\Gamma_{10}\Gamma^{a_1\cdots a_8}\Psi_{10} + \frac{4}{3\cdot 7!}P_{a_1\cdots a_8}\Gamma^{a_1\cdots a_7}\Psi^{a_8}\\
&+ \frac{7}{2\cdot 8!}P_{c|ca_1\cdots a_6}\Gamma_{10}\Gamma^{a_1\cdots a_6}\Psi_{10}\\
&- \frac{1}{2\cdot 8!}e^{-5\phi/4}P_{a_1\cdots a_9}\Gamma_{10}\Gamma^{a_1\cdots a_9}\Psi_{10} + \frac{15}{9!}e^{-5\phi/4}P_{a_1\cdots a_9}\Gamma^{a_1\cdots a_8}\Psi^{a_9} + \dots\,,
\end{aligned}
\tag{A.3.3}
$$

and is related to the dilatino equation of motion in the body of the article.

For the gravitino component Ψ_a one finds similarly

$$
\begin{aligned}
0 \;=\;& \partial_t\Psi_a - \frac{1}{4}q_{b_1b_2}\Gamma^{b_1b_2}\Psi_a - q_{ab}\Psi_b\\
&- \frac{1}{2}e^{3\phi/4}P_b\Gamma_{10}\Gamma^b\Psi_a + e^{3\phi/4}P_a\Psi_{10}\\
&- \frac{1}{4}e^{-\phi/2}P_{b_1b_2}\Gamma_{10}\Gamma^{b_1b_2}\Psi_a + \frac{2}{3}e^{-\phi/2}P_{ab}\Gamma_{10}\Psi^b - \frac{1}{3}e^{-\phi/2}P_{b_1b_2}\Gamma_{10}\Gamma_a{}^{b_1}\Psi^{b_2}\\
&-\frac{2}{3}e^{-\phi/2}P_{ab}\Gamma^b\Psi_{10} + \frac{1}{6}e^{-\phi/2}P_{b_1b_1}\Gamma_a{}^{b_1b_2}\Psi_{10}\\
&- \frac{1}{12}e^{\phi/4}P_{b_1b_2b_3}\Gamma^{b_1\cdots b_3}\Psi_a - \frac{2}{3}e^{\phi/4}P_{ab_1b_2}\Gamma^{b_1}\Psi^{b_2} + \frac{1}{6}e^{\phi/4}P_{b_1b_2b_3}\Gamma_a{}^{b_1b_2}\Psi^{b_3}\\
&- \frac{1}{2\cdot 5!}e^{-\phi/4}P_{b_1\cdots b_5}\Gamma_{10}\Gamma^{b_1\cdots b_5}\Psi_a - \frac{1}{18}e^{-\phi/4}P_{ab_1\cdots b_4}\Gamma_{10}\Gamma^{b_1\cdots b_3}\Psi^{b_4}\\
&+\frac{1}{36}e^{-\phi/4}P_{b_1\cdots b_5}\Gamma_{10}\Gamma_a{}^{b_1\cdots b_4}\Psi^{b_5} - \frac{1}{3\cdot 4!}e^{-\phi/4}P_{ab_1\cdots b_4}\Gamma^{b_1\cdots b_4}\Psi_{10}\\
&+\frac{2}{3\cdot 5!}e^{-\phi/4}P_{b_1\cdots b_5}\Gamma_a{}^{b_1\cdots b_5}\Psi_{10}\\
&+ \frac{1}{2\cdot 6!}e^{\phi/2}P_{b_1\cdots b_6}\Gamma^{b_1\cdots b_6}\Psi_a - \frac{1}{3\cdot 4!}e^{\phi/2}P_{ab_1\cdots b_5}\Gamma^{b_1\cdots b_4}\Psi^{b_5} + \frac{4}{6!}e^{\phi/2}P_{b_1\cdots b_6}\Gamma_a{}^{b_1\cdots b_5}\Psi^{b_6}\\
&- \frac{1}{2\cdot 7!}e^{-3\phi/4}P_{b_1\cdots b_7}\Gamma^{b_1\cdots b_7}\Psi_a + \frac{1}{6!}e^{-3\phi/4}P_{b_1\cdots b_7}\Gamma_a{}^{b_1\cdots b_6}\Psi^{b_7}\\
&+\frac{2}{7!}e^{-3\phi/4}P_{b_1\cdots b_7}\Gamma_{10}\Gamma_a{}^{b_1\cdots b_7}\Psi_{10}\\
&+ \frac{1}{6\cdot 7!}P_{b_1\cdots b_8}\Gamma_{10}\Gamma_a{}^{b_1\cdots b_7}\Psi^{b_8} + \frac{1}{6\cdot 6!}P_{ab_1\cdots b_7}\Gamma_{10}\Gamma^{b_1\cdots b_6}\Psi^{b_7}\\
&+\frac{1}{6\cdot 7!}P_{b_1\cdots b_8}\Gamma_a{}^{b_1\cdots b_8}\Psi_{10} - \frac{1}{6\cdot 7!}P_{ab_1\cdots b_7}\Gamma^{b_1\cdots b_7}\Psi_{10}\\
&- \frac{7}{2\cdot 8!}P_{c|cb_1\cdots b_6}\Gamma_{10}\Gamma^{b_1\cdots b_6}\Psi_a + \frac{21}{2\cdot 8!}P_{b_0|ab_1\cdots b_6}\Gamma_{10}\Gamma^{b_0b_1\cdots b_5}\Psi^{b_6}\\
&+\frac{49}{4\cdot 8!}P_{a|b_1\cdots b_7}\Gamma_{10}\Gamma^{b_1\cdots b_6}\Psi^{b_7} - \frac{42}{8!}P_{c|cb_1\cdots b_6}\Gamma_{10}\Gamma_a{}^{b_1\cdots b_5}\Psi^{b_6} + \frac{7}{8!}P_{c|cb_1\cdots b_6}\Gamma_a{}^{b_1\cdots b_6}\Psi_{10}
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{7}{4\cdot 8!}P_{b_0|b_1\cdots c_7}\Gamma_{10}\Gamma_a{}^{b_0b_1\cdots b_6}\Psi^{b_7}+\frac{7}{4\cdot 8!}P_{b_0|ab_1\cdots b_6}\Gamma_{10}\Gamma^{b_1\cdots b_6}\Psi^{b_0}\\
&-\frac{7}{4\cdot 8!}P_{b_0|b_1\cdots b_7}\Gamma_{10}\Gamma_a{}^{b_1\cdots b_7}\Psi^{b_0}-\frac{7}{4\cdot 8!}P_{b_0|ab_1\cdots b_6}\Gamma^{b_0b_1\cdots b_6}\Psi_{10}\\
&-\frac{7}{4\cdot 8!}P_{a|b_1\cdots b_7}\Gamma^{b_1\cdots b_7}\Psi_{10}\\
-\;&\frac{1}{2\cdot 9!}e^{-5\phi/4}P_{b_1\cdots b_9}\Gamma_{10}\Gamma^{b_1\cdots b_9}\Psi_a+\frac{12}{9!}e^{-5\phi/4}P_{b_1\cdots b_9}\Gamma_{10}\Gamma_a{}^{b_1\cdots b_8}\Psi^{b_9}\\
&+\frac{24}{9!}e^{-5\phi/4}P_{ab_1\cdots b_8}\Gamma_{10}\Gamma^{b_1\cdots b_7}\Psi^{b_8}+\frac{1}{8!}e^{-5\phi/4}P_{ab_1\cdots b_8}\Gamma^{b_1\cdots b_8}\Psi_{10}+\ldots\,. \qquad \text{(A.3.4)}
\end{aligned}
$$

A.3.3 Supersymmetry Variation

In the same fashion, the supersymmetry variation given in (7.2.39) can be written explicitly as

$$
\begin{aligned}
\delta\Psi_t \;=\;& \partial_t\epsilon-\frac{1}{4}q_{a_1a_2}\Gamma^{a_1a_2}\epsilon-\frac{1}{2}e^{3\phi/4}P_a\Gamma_{10}\Gamma^a\epsilon-\frac{1}{4}e^{-\phi/2}P_{a_1a_2}\Gamma_{10}\Gamma^{a_1a_2}\epsilon\\
&-\frac{1}{2\cdot 3!}e^{\phi/4}P_{a_1a_2a_3}\Gamma^{a_1a_2a_3}\epsilon-\frac{1}{2\cdot 5!}e^{-\phi/4}P_{a_1\cdots a_5}\Gamma_{10}\Gamma^{a_1\cdots a_5}\epsilon+\frac{1}{2\cdot 6!}e^{\phi/2}P_{a_1\cdots a_6}\Gamma^{a_1\cdots a_6}\epsilon\\
&-\frac{1}{2\cdot 7!}e^{-3\phi/4}P_{a_1\cdots a_7}\Gamma^{a_1\cdots a_7}\epsilon-\frac{7}{2\cdot 8!}P_{c|ca_1\cdots a_6}\Gamma_{10}\Gamma^{a_1\cdots a_6}\epsilon\\
&-\frac{1}{2\cdot 9!}e^{-5\phi/4}P_{a_1\cdots a_9}\Gamma_{10}\Gamma^{a_1\cdots a_9}\epsilon+\cdots \qquad \text{(A.3.5)}
\end{aligned}
$$

Bibliography

[1] E. Witten, *String theory dynamics in various dimensions*, Nucl. Phys. **B443**, 85–126 (1995) [hep-th/9503124].

[2] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*,. Cambridge, UK: Univ. Pr. (1998) 402 p.

[3] J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond*,. Cambridge, UK: Univ. Pr. (1998) 531 p.

[4] K. Becker, M. Becker and J. H. Schwarz, *String theory and M-theory: A modern introduction*,. Cambridge, UK: Cambridge Univ. Pr. (2007) 739 p.

[5] I. C. G. Campbell and P. C. West, $\mathcal{N} = 2, D = 10$ *Nonchiral Supergravity and Its Spontaneous Compactification*, Nucl. Phys. **B243**, 112 (1984).

[6] F. Giani and M. Pernici, $\mathcal{N} = 2$ *Supergravity in Ten Dimensions*, Phys. Rev. **D30**, 325–333 (1984).

[7] M. Huq and M. A. Namazie, *Kaluza-Klein Supergravity in Ten Dimensions*, Class. Quant. Grav. **2**, 293 (1985).

[8] C. M. Hull and P. K. Townsend, *Unity of superstring dualities*, Nucl. Phys. **B438**, 109–137 (1995) [hep-th/9410167].

[9] E. Cremmer and B. Julia, *The* $\mathcal{N} = 8$ *Supergravity Theory. 1: The Lagrangian*, Phys. Lett. **B80**, 48 (1978).

[10] E. Cremmer and B. Julia, *The* $SO(8)$ *Supergravity*, Nucl. Phys. **B159**, 141 (1979).

[11] N. A. Obers and B. Pioline, *U-duality and M-theory*, Phys. Rept. **318**, 113–225 (1999) [hep-th/9809039].

[12] E. Kiritsis and B. Pioline, *On* R^4 *threshold corrections in type IIB string theory and (p,q) string instantons*, Nucl. Phys. **B508**, 509–534 (1997) [hep-th/9707018].

[13] N. A. Obers and B. Pioline, *Eisenstein series and string thresholds*, Commun. Math. Phys. **209**, 275–324 (2000) [hep-th/9903113].

[14] M. B. Green and M. Gutperle, *Effects of D-instantons*, Nucl. Phys. **B498**, 195–227 (1997) [hep-th/9701093].

[15] M. B. Green and P. Vanhove, *D-instantons, strings and M-theory*, Phys. Lett. **B408**, 122–134 (1997) [hep-th/9704145].

[16] P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, *Vacuum Configurations for Superstrings*, Nucl. Phys. **B258**, 46–74 (1985).

[17] A. Strominger and E. Witten, *New Manifolds for Superstring Compactification*, Commun. Math. Phys. **101**, 341 (1985).

[18] J. Bagger and E. Witten, *Matter Couplings in N=2 Supergravity*, Nucl. Phys. **B222**, 1 (1983).

[19] B. de Wit, P. G. Lauwers and A. Van Proeyen, *Lagrangians of N=2 Supergravity - Matter Systems*, Nucl. Phys. **B255**, 569 (1985).

[20] P. Candelas, X. C. De La Ossa, P. S. Green and L. Parkes, *A pair of Calabi-Yau manifolds as an exactly soluble superconformal theory*, Nucl. Phys. **B359**, 21–74 (1991).

[21] C. Vafa and E. E. Zaslow, *Mirror Symmetry.* Clay Mathematics Monographs, Volume 1, 2003.

[22] S. Cecotti, S. Ferrara and L. Girardello, *Geometry of Type II Superstrings and the Moduli of Superconformal Field Theories*, Int. J. Mod. Phys. **A4**, 2475 (1989).

[23] B. Pioline, *Lectures on on black holes, topological strings and quantum attractors (2.0)*, Class. Quant. Grav. **23**, S981 (2006) [hep-th/0607227].

[24] N. Seiberg and E. Witten, *Gauge dynamics and compactification to three dimensions*, [hep-th/9607163].

[25] P. S. Aspinwall, *Compactification, geometry and duality: $\mathcal{N} = 2$*, [hep-th/0001001].

[26] K. Becker, M. Becker and A. Strominger, *Five-branes, membranes and nonperturbative string theory*, Nucl. Phys. **B456**, 130–152 (1995) [hep-th/9507158].

[27] D. Robles-Llana, M. Rocek, F. Saueressig, U. Theis and S. Vandoren, *Nonperturbative corrections to 4D string theory effective actions from $SL(2,\mathbb{Z})$ duality and supersymmetry*, Phys. Rev. Lett. **98**, 211602 (2007) [hep-th/0612027].

[28] D. Robles-Llana, F. Saueressig, U. Theis and S. Vandoren, *Membrane instantons from mirror symmetry*, Commun. Num. Theor. Phys. **1**, 681 (2007) [arXiv:0707.0838].

[29] S. Alexandrov, B. Pioline, F. Saueressig and S. Vandoren, *Linear perturbations of Hyperkähler metrics*, Lett. Math. Phys. **87**, 225–265 (2009) [arXiv:0806.4620].

[30] S. Alexandrov, B. Pioline, F. Saueressig and S. Vandoren, *Linear perturbations of quaternionic metrics*, [arXiv:0810.1675].

[31] S. Alexandrov, B. Pioline, F. Saueressig and S. Vandoren, *D-instantons and twistors*, JHEP **03**, 044 (2009) [arXiv:0812.4219].

[32] B. Julia, *Group disintegrations,*. Invited paper presented at Nuffield Gravity Workshop, Cambridge, Eng., Jun 22 - Jul 12, 1980.

[33] B. Julia, *Kac-Moody symmetry of gravitation and supergravity theories,*. Invited talk given at AMS-SIAM Summer Seminar on Applications of Group Theory in Physics and Mathematical Physics, Chicago, Ill., Jul 6-16, 1982.

[34] V. Kac, *Infinite dimensional Lie algebras.* Cambridge University Press, 1990.

[35] H. Nicolai, *The integrability of $\mathcal{N} = 16$ supergravity*, Phys. Lett. **B194**, 402 (1987).

[36] R. Geroch, *A method for generating solutions of Einstein's equations*, J. Math. Phys. **12**, 918–924 (1971).

[37] P. Breitenlohner and D. Maison, *On the Geroch Group*, Ann. Poincare **46**, 215 (1987).

[38] M. R. Gaberdiel, D. I. Olive and P. C. West, *A class of Lorentzian Kac-Moody algebras*, Nucl. Phys. **B645**, 403–437 (2002) [hep-th/0205068].

[39] T. Banks, W. Fischler and L. Motl, *Dualities versus singularities*, JHEP **01**, 019 (1999) [hep-th/9811194].

[40] J. Brown, O. J. Ganor and C. Helfgott, *M-theory and E_{10}: Billiards, branes, and imaginary roots*, JHEP **08**, 063 (2004) [hep-th/0401053].

[41] J. A. Harvey and G. W. Moore, *Algebras, BPS States, and Strings*, Nucl. Phys. **B463**, 315–368 (1996) [hep-th/9510182].

[42] J. A. Harvey and G. W. Moore, *On the algebras of BPS states*, Commun. Math. Phys. **197**, 489–519 (1998) [hep-th/9609017].

[43] O. Barwald, R. W. Gebert, M. Gunaydin and H. Nicolai, *Missing modules, the gnome Lie algebra, and E_{10}*, Commun. Math. Phys. **195**, 29–65 (1998) [hep-th/9703084].

[44] T. Damour and M. Henneaux, *E_{10}, BE_{10} and arithmetical chaos in superstring cosmology*, Phys. Rev. Lett. **86**, 4749–4752 (2001) [hep-th/0012172].

[45] V. A. Belinskii, I. M. Khalatnikov and E. M. Lifshitz, *Oscillatory approach to a singular point in the relativistic cosmology*, Adv. Phys. **19**, 525 (1970).

[46] V. Belinskii and I. Khalatnikov, *Effect of scalar and vector fields on the nature of the cosmological singularity*, Sov. Phys. JETP **36**, 591–597 (1973).

[47] V. A. Belinskii, I. M. Khalatnikov and E. M. Lifshitz, *A general solution of the einstein equations with a time singularity*, Adv. Phys. **31**, 639 (1982).

[48] C. W. Misner, *Mixmaster universe*, Phys. Rev. Lett. **22**, 1071–1074 (1969).

[49] J. Demaret, J. L. Hanquin, M. Henneaux and P. Spindel, *Cosmological models in eleven-dimensional supergravity*, Nucl. Phys. **B252**, 538–560 (1985).

[50] T. Damour and M. Henneaux, *Chaos in superstring cosmology*, Phys. Rev. Lett. **85**, 920–923 (2000) [hep-th/0003139].

[51] T. Damour, M. Henneaux and H. Nicolai, *Cosmological billiards*, Class. Quant. Grav. **20**, R145–R200 (2003) [hep-th/0212256].

[52] T. Damour, M. Henneaux, B. Julia and H. Nicolai, *Hyperbolic Kac-Moody algebras and chaos in Kaluza-Klein models*, Phys. Lett. **B509**, 323–330 (2001) [hep-th/0103094].

[53] T. Damour, M. Henneaux and H. Nicolai, *E_{10} and a 'small tension expansion' of M theory*, Phys. Rev. Lett. **89**, 221601 (2002) [hep-th/0207267].

[54] P. C. West, *Hidden superconformal symmetry in M-theory*, JHEP **08**, 007 (2000) [hep-th/0005270].

[55] P. C. West, *E_{11} and M theory*, Class. Quant. Grav. **18**, 4443–4460 (2001) [hep-th/0104081].

[56] P. C. West, *Very extended E_8 and A_8 at low levels, gravity and supergravity*, Class. Quant. Grav. **20**, 2393–2406 (2003) [hep-th/0212291].

[57] P. C. West, *E_{11}, $SL(32)$ and central charges*, Phys. Lett. **B575**, 333–342 (2003) [hep-th/0307098].

[58] P. West, *The IIA, IIB and eleven dimensional theories and their common E_{11} origin*, Nucl. Phys. **B693**, 76–102 (2004) [hep-th/0402140].

[59] F. Riccioni and P. C. West, *The E_{11} origin of all maximal supergravities*, JHEP **07**, 063 (2007) [arXiv:0705.0752].

[60] F. Riccioni and P. West, *Local E_{11}*, JHEP **04**, 051 (2009) [arXiv:0902.4678].

[61] A. Kleinschmidt and H. Nicolai, *E_{10} and $SO(9,9)$ invariant supergravity*, JHEP **07**, 041 (2004) [hep-th/0407101].

[62] A. Kleinschmidt and H. Nicolai, *IIB supergravity and E_{10}*, Phys. Lett. **B606**, 391–402 (2005) [hep-th/0411225].

[63] F. Englert and L. Houart, *$\mathcal{G}^{+++}$ invariant formulation of gravity and M-theories: Exact BPS solutions*, JHEP **01**, 002 (2004) [hep-th/0311255].

[64] F. Englert and L. Houart, *From brane dynamics to a Kac-Moody invariant formulation of M-theories*, [hep-th/0402076].

[65] F. Englert, M. Henneaux and L. Houart, *From very-extended to overextended gravity and m-theories*, JHEP **02**, 070 (2005) [hep-th/0412184].

[66] F. Englert and L. Houart, *$\mathcal{G}^{+++}$ invariant formulation of gravity and M-theories: Exact intersecting brane solutions*, JHEP **05**, 059 (2004) [hep-th/0405082].

[67] F. Englert and L. Houart, *$\mathcal{G}^{+++}$ invariant formulation of gravity and m-theories: Exact intersecting brane solutions*, JHEP **05**, 059 (2004) [hep-th/0405082].

[68] F. Englert, L. Houart, A. Kleinschmidt, H. Nicolai and N. Tabti, *An E_9 multiplet of BPS states*, JHEP **05**, 065 (2007) [hep-th/0703285].

[69] S. de Buyl, M. Henneaux and L. Paulot, *Hidden symmetries and Dirac fermions*, Class. Quant. Grav. **22**, 3595–3622 (2005) [hep-th/0506009].

[70] T. Damour, A. Kleinschmidt and H. Nicolai, *Hidden symmetries and the fermionic sector of eleven- dimensional supergravity*, Phys. Lett. **B634**, 319–324 (2006) [hep-th/0512163].

[71] S. de Buyl, M. Henneaux and L. Paulot, *Extended E_8 invariance of 11-dimensional supergravity*, JHEP **02**, 056 (2006) [hep-th/0512292].

[72] T. Damour, A. Kleinschmidt and H. Nicolai, *$K(E_{10})$, supergravity and fermions*, JHEP **08**, 046 (2006) [hep-th/0606105].

[73] J. E. Humphreys, *Reflection Groups and Coxeter Groups.* Cambridge University Press, 1990.

[74] J. G. Ratcliffe, *Foundations of Hyperbolic Manifolds.* Graduate Texts in Mathematics vol 149, Springer-Verlag, 1994.

[75] V. A. Gritsenko and V. V. Nikulin, *The Igusa modular forms and "the simplest" Lorentzian Kac–Moody algebras*, [alg-geom/9603010].

[76] R. Dijkgraaf, E. P. Verlinde and H. L. Verlinde, *Counting Dyons in $\mathcal{N}=4$ String Theory*, Nucl. Phys. **B484**, 543–561 (1997) [hep-th/9607026].

[77] M. C. N. Cheng and E. P. Verlinde, *Wall Crossing, Discrete Attractor Flow, and Borcherds Algebra*, [arXiv:0806.2337].

[78] M. C. N. Cheng and A. Dabholkar, *Borcherds-Kac-Moody Symmetry of $\mathcal{N}=4$ Dyons*, [arXiv:0809.4258].

[79] M. C. N. Cheng, *The Spectra of Supersymmetric States in String Theory*, [arXiv:0807.3099].

[80] A. Terras, *Harmonic Analysis on Symmetric Spaces and Applications I.* Springer Verlag, 1985.

[81] A. Terras, *Harmonic Analysis on Symmetric Spaces and Applications II.* Springer Verlag, 1985.

[82] A. Kleinschmidt, *Indefinite Kac-Moody Algebras in String Theory.* 2004. Ph. D. Thesis Cambridge.

[83] E. A. Bergshoeff, I. De Baetselier and T. A. Nutma, *E_{11} and the embedding tensor*, JHEP **09**, 047 (2007) [arXiv:0705.1304].

[84] A. J. Feingold and I. B. Frenkel, *A hyperbolic Kac-Moody algebra and the theory of Siegel modular forms of genus 2*, Math. Ann. **263**, 87 (1983).

[85] J. Fuchs, *Affine Lie Algebras and Quantum Groups.* Cambridge University Press, 1992.

[86] A. Kleinschmidt and H. Nicolai, *E_{10} and $SO(9,9)$ invariant supergravity*, JHEP **07**, 041 (2004) [hep-th/0407101].

[87] A. Kleinschmidt and H. Nicolai, *IIB supergravity and E_{10}*, Phys. Lett. **B606**, 391–402 (2005) [hep-th/0411225].

[88] T. Fischbacher, *The structure of E_{10} at higher a_9 levels: A first algorithmic approach*, JHEP **08**, 012 (2005) [hep-th/0504230].

[89] T. Damour and H. Nicolai, *Eleven dimensional supergravity and the $E_{10}/K(E_{10})$ sigma-model at low A_9 levels*, [hep-th/0410245].

[90] C. W. Misner, *The mixmaster cosmological metrics*, [gr-qc/9405068].

[91] E. Bergshoeff, M. de Roo, B. de Wit and P. van Nieuwenhuizen, *Ten-dimensional maxwell-einstein supergravity, its currents, and the issue of its auxiliary fields*, Nucl. Phys. **B195**, 97–136 (1982).

[92] G. F. Chapline and N. S. Manton, *Unification of Yang-Mills Theory and Supergravity in Ten- Dimensions*, Phys. Lett. **B120**, 105–109 (1983).

[93] T. Damour and S. de Buyl, *Describing general cosmological singularities in Iwasawa variables*, Phys. Rev. **D77**, 043520 (2008) [arXiv:0710.5692].

[94] M. Henneaux, D. Persson and P. Spindel, *Spacelike Singularities and Hidden Symmetries of Gravity*, Living Rev. Rel. **11**, 1 (2008) [arXiv:0710.1818].

[95] A. A. Kirillov Sov. Phys. JETP **76**, 355 (1993).

[96] V. D. Ivashchuk, A. A. Kirillov and V. N. Melnikov, *Stochastic behavior of multidimensional cosmological models near a singularity*, Russ. Phys. J. **37**, 1102–1106 (1994).

[97] V. D. Ivashchuk, V. N. Melnikov and A. A. Kirillov, *Stochastic properties of multidimensional cosmological models near a singular point*, JETP Lett. **60**, 235–239 (1994).

[98] A. A. Kirillov and V. N. Melnikov, *Dynamics of inhomogeneities of metric in the vicinity of a singularity in multidimensional cosmology*, Phys. Rev. **D52**, 723–729 (1995) [gr-qc/9408004].

[99] V. D. Ivashchuk and V. N. Melnikov, *Billiard representation for multidimensional cosmology with multicomponent perfect fluid near the singularity*, Class. Quant. Grav. **12**, 809–826 (1995) [gr-qc/9407028].

[100] V. D. Ivashchuk and V. N. Melnikov, *Billiard representation for multidimensional cosmology with intersecting p-branes near the singularity*, J. Math. Phys. **41**, 6341–6363 (2000) [hep-th/9904077].

[101] B. K. Berger, D. Garfinkle, J. Isenberg, V. Moncrief and M. Weaver, *The singularity in generic gravitational collapse is spacelike, local, and oscillatory*, Mod. Phys. Lett. **A13**, 1565–1574 (1998) [gr-qc/9805063].

[102] B. K. Berger, *Numerical approaches to space-time singularities*, Living Rev. Rel. **1**, 7 (1998).

[103] L. Andersson and A. D. Rendall, *Quiescent cosmological singularities*, Commun. Math. Phys. **218**, 479–511 (2001) [gr-qc/0001047].

[104] H. Ringstrom, *The Bianchi IX attractor*, Annales Henri Poincare **2**, 405–500 (2001) [gr-qc/0006035].

[105] T. Damour, M. Henneaux, A. D. Rendall and M. Weaver, *Kasner-like behaviour for subcritical Einstein-matter systems*, Annales Henri Poincare **3**, 1049–1111 (2002) [gr-qc/0202069].

[106] D. Garfinkle, *Numerical simulations of generic singuarities*, Phys. Rev. Lett. **93**, 161101 (2004) [gr-qc/0312117].

[107] J. M. Heinzle, C. Uggla and N. Rohr, *The cosmological billiard attractor*, Adv. Theor. Math. Phys. **13**, 293–407 (2009) [gr-qc/0702141].

[108] D. Garfinkle, *Numerical simulations of general gravitational singularities*, Class. Quant. Grav. **24**, S295–S306 (2007) [0808.0160].

[109] D. H. Wesley, P. J. Steinhardt and N. Turok, *Controlling chaos through compactification in cosmological models with a collapsing phase*, Phys. Rev. **D72**, 063513 (2005) [hep-th/0502108].

[110] D. H. Wesley, *Kac-Moody Algebras and Controlled Chaos*, Class. Quant. Grav. **24**, F7–F13 (2007) [hep-th/0610322].

[111] T. Damour and H. Nicolai, *Higher order M theory corrections and the Kac-Moody algebra E_{10}*, Class. Quant. Grav. **22**, 2849–2880 (2005) [hep-th/0504153].

[112] J. Khoury, B. A. Ovrut, P. J. Steinhardt and N. Turok, *The ekpyrotic universe: Colliding branes and the origin of the hot big bang*, Phys. Rev. **D64**, 123522 (2001) [hep-th/0103239].

[113] P. J. Steinhardt and N. Turok, *Cosmic evolution in a cyclic universe*, Phys. Rev. **D65**, 126003 (2002) [hep-th/0111098].

[114] A. Buonanno, T. Damour and G. Veneziano, *Pre-big bang bubbles from the gravitational instability of generic string vacua*, Nucl. Phys. **B543**, 275–320 (1999) [hep-th/9806230].

[115] M. Gasperini and G. Veneziano, *The pre-big bang scenario in string cosmology*, Phys. Rept. **373**, 1–212 (2003) [hep-th/0207130].

[116] T. Damour, S. de Buyl, M. Henneaux and C. Schomblond, *Einstein billiards and overextensions of finite-dimensional simple Lie algebras*, JHEP **08**, 030 (2002) [hep-th/0206125].

[117] Vinberg., *Geometry II - Spaces of Constant Curvature.* Springer Verlag, 1993.

[118] T. Damour and M. Henneaux, *Oscillatory behaviour in homogeneous string cosmology models*, Phys. Lett. **B488**, 108–116 (2000) [hep-th/0006171].

[119] P. G. Garrett, *Buildings and Classical Groups.* Chapman & Hall/CRC; 1 edition, 1997.

[120] K. S. Brown, *Buildings.* Springer-Verlag, 1998.

[121] A. J. Feingold and H. Nicolai, *Subalgebras of Hyperbolic Kac-Moody Algebras*, [math.qa/0303179].

[122] M. Henneaux, M. Leston, D. Persson and P. Spindel, *Geometric configurations, regular subalgebras of E_{10} and M-theory cosmology*, JHEP **10**, 021 (2006) [hep-th/0606123].

[123] N. Marcus and J. H. Schwarz, *Three-dimensional supergravity theories*, Nucl. Phys. **B228**, 145 (1983).

[124] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces.*

[125] A. Keurentjes, *Poincare duality and G^{+++} algebras*, [hep-th/0510212].

[126] B. S. DeWitt, *Quantum Theory of Gravity. 1. The Canonical Theory*, Phys. Rev. **160**, 1113–1148 (1967).

[127] E. Cremmer, B. Julia, H. Lu and C. N. Pope, *Dualisation of dualities. I*, Nucl. Phys. **B523**, 73–144 (1998) [hep-th/9710119].

[128] E. Cremmer, B. Julia, H. Lu and C. N. Pope, *Dualisation of dualities. II: Twisted self-duality of doubled fields and superdualities*, Nucl. Phys. **B535**, 242–292 (1998) [hep-th/9806106].

[129] T. Curtright, *Generalized gauge fields*, Phys. Lett. **B165**, 304 (1985).

[130] J. Brown, S. Ganguli, O. J. Ganor and C. Helfgott, *E_{10} orbifolds*, JHEP **06**, 057 (2005) [hep-th/0409037].

[131] M. Bagnoud and L. Carlevaro, *Hidden Borcherds symmetries in $\mathbb{Z}_n$ orbifolds of M-theory and magnetized D-branes in type 0' orientifolds*, JHEP **11**, 003 (2006) [hep-th/0607136].

[132] S. Kantor, *Die configurationen* $(3,3)_{10}$,. K. Academie der Wissenschaften, Vienna, Sitzungsbereichte der matematisch naturewissenshaftlichen classe, 84 II, 1291-1314 (1881).

[133] D. Hilbert and S. Cohn-Vossen, *Geometry and the imagination*,.

[134] W. Page and H. L. Dorwart, *Numerical patterns and geometric configurations'*, Mathematics Magazine **57 (No. 2)**, 82–92 (1984).

[135] A. Kleinschmidt and H. Nicolai, E_{10} *Cosmology*, JHEP **01**, 137 (2006) [hep-th/0511290].

[136] N. Ohta, *Accelerating cosmologies from S-branes*, Phys. Rev. Lett. **91**, 061303 (2003) [hep-th/0303238].

[137] J. C. Baez, *The Octonions*, [math.ra/0105155].

[138] R. Argurio, F. Englert and L. Houart, *Intersection rules for p-branes*, Phys. Lett. **B398**, 61–68 (1997) [hep-th/9701042].

[139] N. Ohta, *Intersection rules for S-branes*, Phys. Lett. **B558**, 213–220 (2003) [hep-th/0301095].

[140] A. Strominger, *Loop corrections to the universal hypermultiplet*, Phys. Lett. **B421**, 139–148 (1998) [hep-th/9706195].

[141] V. D. Ivashchuk and V. N. Melnikov, *On brane solutions related to non-singular Kac-Moody algebras*, [arXiv:0810.0196].

[142] L. J. Romans, *Massive $\mathcal{N} = 2a$ Supergravity in Ten-Dimensions*, Phys. Lett. **B169**, 374 (1986).

[143] E. Bergshoeff, M. de Roo, M. B. Green, G. Papadopoulos and P. K. Townsend, *Duality of Type II 7-branes and 8-branes*, Nucl. Phys. **B470**, 113–135 (1996) [hep-th/9601150].

[144] I. V. Lavrinenko, H. Lu, C. N. Pope and K. S. Stelle, *Superdualities, brane tensions and massive IIA/IIB duality*, Nucl. Phys. **B555**, 201–227 (1999) [hep-th/9903057].

[145] E. Bergshoeff, R. Kallosh, T. Ortin, D. Roest and A. Van Proeyen, *New Formulations of D=10 Supersymmetry and D8-O8 Domain Walls*, Class. Quant. Grav. **18**, 3359–3382 (2001) [hep-th/0103233].

[146] A. Kleinschmidt, I. Schnakenburg and P. C. West, *Very-extended Kac-Moody algebras and their interpretation at low levels*, Class. Quant. Grav. **21**, 2493–2525 (2004) [hep-th/0309198].

[147] E. Bergshoeff, Y. Lozano and T. Ortin, *Massive branes*, Nucl. Phys. **B518**, 363–423 (1998) [hep-th/9712115].

[148] E. Bergshoeff, M. de Roo, B. Janssen and T. Ortin, *The super D9-brane and its truncations*, Nucl. Phys. **B550**, 289–302 (1999) [hep-th/9901055].

[149] T. Damour, M. Henneaux and H. Nicolai, E_{10} *and a 'small tension expansion' of M theory*, Phys. Rev. Lett. **89**, 221601 (2002) [hep-th/0207267].

[150] H. Nicolai and H. Samtleben, *Maximal gauged supergravity in three dimensions*, Phys. Rev. Lett. **86**, 1686–1689 (2001) [`hep-th/0010076`].

[151] B. de Wit, H. Nicolai and H. Samtleben, *Gauged Supergravities, Tensor Hierarchies, and M-Theory*, JHEP **02**, 044 (2008) [`arXiv:0801.1294`].

[152] E. A. Bergshoeff, O. Hohm, A. Kleinschmidt, H. Nicolai, T. A. Nutma and J. Palmkvist, *E_{10} and Gauged Maximal Supergravity*, JHEP **01**, 020 (2009) [`arXiv:0810.5767`].

[153] I. Schnakenburg and P. C. West, *Massive IIA supergravity as a non-linear realisation*, Phys. Lett. **B540**, 137–145 (2002) [`hep-th/0204207`].

[154] A. Kleinschmidt and H. Nicolai, *IIA and IIB spinors from $K(E_{10})$*, Phys. Lett. **B637**, 107–112 (2006) [`hep-th/0603205`].

[155] M. B. Green and J. H. Schwarz, *Supersymmetrical Dual String Theory*, Nucl. Phys. **B181**, 502–530 (1981).

[156] M. B. Green and J. H. Schwarz, *Supersymmetrical Dual String Theory. 2. Vertices and Trees*, Nucl. Phys. **B198**, 252–268 (1982).

[157] M. B. Green and J. H. Schwarz, *Supersymmetrical Dual String Theory. 3. Loops and Renormalization*, Nucl. Phys. **B198**, 441–460 (1982).

[158] S. H. Shenker, *The Strength of nonperturbative effects in string theory*,. Presented at the Cargese Workshop on Random Surfaces, Quantum Gravity and Strings, Cargese, France, May 28 - Jun 1, 1990.

[159] M. Beneke, *Renormalons*, Phys. Rept. **317**, 1–142 (1999) [`hep-ph/9807443`].

[160] J. Harer and D. Zagier, *The Euler characteristic of the moduli space of curves*, Inv. Math. **85**, 457–485 (1986).

[161] R. Penner, *The decorated Teichmueller space of punctured surfaces*, Commun. Math. Phys. **113**, 299–339 (1987).

[162] J. Polchinski, *Dirichlet-Branes and Ramond-Ramond Charges*, Phys. Rev. Lett. **75**, 4724–4727 (1995) [`hep-th/9510017`].

[163] M. Dine, N. Seiberg, X. G. Wen and E. Witten, *Nonperturbative Effects on the String World Sheet*, Nucl. Phys. **B278**, 769 (1986).

[164] M. Dine, N. Seiberg, X. G. Wen and E. Witten, *Nonperturbative Effects on the String World Sheet. 2*, Nucl. Phys. **B289**, 319 (1987).

[165] B. Pioline and S. Vandoren, *Large D-instanton effects in string theory*, [`arXiv:0904.2303`].

[166] M. T. Grisaru, A. E. M. van de Ven and D. Zanon, *Two-Dimensional Supersymmetric Sigma Models on Ricci Flat Kahler Manifolds Are Not Finite*, Nucl. Phys. **B277**, 388 (1986).

[167] D. J. Gross and E. Witten, *Superstring Modifications of Einstein's Equations*, Nucl. Phys. **B277**, 1 (1986).

[168] M. T. Grisaru, A. E. M. van de Ven and D. Zanon, *Four Loop Divergences for the N=1 Supersymmetric Nonlinear Sigma Model in Two-Dimensions*, Nucl. Phys. **B277**, 409 (1986).

[169] N. Berkovits, *Construction of R^4 terms in $\mathcal{N} = 2, D = 8$ superspace*, Nucl. Phys. **B514**, 191–203 (1998) [`hep-th/9709116`].

[170] B. Pioline, *A note on non-perturbative R^4 couplings*, Phys. Lett. **B431**, 73–76 (1998) [`hep-th/9804023`].

[171] M. B. Green and S. Sethi, *Supersymmetry constraints on type IIB supergravity*, Phys. Rev. **D59**, 046006 (1999) [`hep-th/9808061`].

[172] M. B. Green, M. Gutperle and P. Vanhove, *One loop in eleven dimensions*, Phys. Lett. **B409**, 177–184 (1997) [`hep-th/9706175`].

[173] M. B. Green, H.-h. Kwon and P. Vanhove, *Two loops in eleven dimensions*, Phys. Rev. **D61**, 104010 (2000) [`hep-th/9910055`].

[174] M. B. Green and P. Vanhove, *The low energy expansion of the one-loop type II superstring amplitude*, Phys. Rev. **D61**, 104011 (2000) [`hep-th/9910056`].

[175] M. B. Green and P. Vanhove, *Duality and higher derivative terms in m theory*, JHEP **01**, 093 (2006) [`hep-th/0510027`].

[176] M. B. Green, J. G. Russo and P. Vanhove, *Non-renormalisation conditions in type II string theory and maximal supergravity*, JHEP **02**, 099 (2007) [`hep-th/0610299`].

[177] M. B. Green, J. G. Russo and P. Vanhove, *Ultraviolet properties of maximal supergravity*, Phys. Rev. Lett. **98**, 131602 (2007) [`hep-th/0611273`].

[178] M. B. Green, J. G. Russo and P. Vanhove, *Low energy expansion of the four-particle genus-one amplitude in type II superstring theory*, JHEP **02**, 020 (2008) [`arXiv:0801.0322`].

[179] M. B. Green, J. G. Russo and P. Vanhove, *Modular properties of two-loop maximal supergravity and connections with string theory*, JHEP **07**, 126 (2008) [`arXiv:0807.0389`].

[180] G. W. Gibbons, M. B. Green and M. J. Perry, *Instantons and Seven-Branes in Type IIB Superstring Theory*, Phys. Lett. **B370**, 37–44 (1996) [`hep-th/9511080`].

[181] P. Yi, *Witten Index and Threshold Bound States of D-Branes*, Nucl. Phys. **B505**, 307–318 (1997) [`hep-th/9704098`].

[182] S. Sethi and M. Stern, *D-brane bound states redux*, Commun. Math. Phys. **194**, 675–705 (1998) [`hep-th/9705046`].

[183] M. B. Green and M. Gutperle, *D-particle bound states and the D-instanton measure*, JHEP **01**, 005 (1998) [`hep-th/9711107`].

[184] I. K. Kostov and P. Vanhove, *Matrix string partition functions*, Phys. Lett. **B444**, 196–203 (1998) [`hep-th/9809130`].

[185] G. W. Moore, N. Nekrasov and S. Shatashvili, *D-particle bound states and generalized instantons*, Commun. Math. Phys. **209**, 77–95 (2000) [`hep-th/9803265`].

[186] E. Cremmer and B. Julia, *The $SO(8)$ Supergravity*, Nucl. Phys. **B159**, 141 (1979).

[187] R. P. Langlands, *On the Functional Relations Satisfied by Eisenstein Series*, Springer Lecture Notes in Mathematics **544**, 1–337 (1976).

[188] M. J. Duff and B. E. W. Nilsson, *Four-Dimensional String Theory from the $K3$ Lattice*, Phys. Lett. **B175**, 417 (1986).

[189] P. S. Aspinwall, *K3 surfaces and string duality*, [`hep-th/9611137`].

[190] A. Gregori *et al.*, *R^2 corrections and non-perturbative dualities of $\mathcal{N} = 4$ string ground states*, Nucl. Phys. **B510**, 423–476 (1998) [`hep-th/9708062`].

[191] E. Kiritsis, N. A. Obers and B. Pioline, *Heterotic/type II triality and instantons on K3*, JHEP **01**, 029 (2000) [`hep-th/0001083`].

[192] N. A. Obers and B. Pioline, *Exact thresholds and instanton effects in string theory*, Fortsch. Phys. **49**, 359–375 (2001) [`hep-th/0101122`].

[193] A. Borel, *Automorphic Forms on $SL(2,\mathbb{R})$*. Cambridge University Press, 2008.

[194] S. Miller, *Cusp Forms on $SL_3(\mathbb{Z})\backslash SL_3(\mathbb{R})/SO_3(\mathbb{R})$*. PhD Thesis, Princeton University, 1997.

[195] B. Pioline and A. Waldron, *Automorphic forms: A physicist's survey*, [`hep-th/0312068`].

[196] D. Kazhdan, B. Pioline and A. Waldron, *Minimal representations, spherical vectors, and exceptional theta series. I*, Commun. Math. Phys. **226**, 1–40 (2002) [`hep-th/0107222`].

[197] D. Kazhdan and A. Polishchik, *Minimal representations: spherical vectors and automorphic functionals*, [`math/0209315`].

[198] L. Brekke and P. G. O. Freund, *p-adic numbers in physics*, Phys. Rept. **233**, 1–66 (1993).

[199] S. S. Gelbart and S. D. Miller, *Riemann's Zeta Function and Beyond*, Bull. Amer. Math. Soc. **41**, 59–112 (2003).

[200] A. I. Vinogradov and L. A. Takhtadzhyan, *Theory of Eisenstein Series for the Group $SL(3,\mathbb{R})$ and its Application to a Binary Problem*, Translated from "Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta" im. V. A. Steklova AN USSR **116**, 119–141 (1982).

[201] B. Pioline and A. Waldron, *The automorphic membrane*, JHEP **06**, 009 (2004) [`hep-th/0404018`].

[202] J. P. Serre, *A Course in Arithmetic*. Graduate Texts in Mathematics, Springer-Verlag, 1973.

[203] A. Dhar and G. Mandal, *Probing 4-dimensional nonsupersymmetric black holes carrying D0- and D6-brane charges*, Nucl. Phys. **B531**, 256–274 (1998) [`hep-th/9803004`].

[204] D. Rasheed, *The Rotating dyonic black holes of Kaluza-Klein theory*, Nucl. Phys. **B454**, 379–401 (1995) [`hep-th/9505038`].

[205] F. Larsen, *Rotating Kaluza-Klein black holes*, Nucl. Phys. **B575**, 211–230 (2000) [`hep-th/9909102`].

[206] N. Lambert and P. C. West, *Enhanced coset symmetries and higher derivative corrections*, Phys. Rev. **D74**, 065002 (2006) [`hep-th/0603255`].

[207] N. Lambert and P. C. West, *Duality groups, automorphic forms and higher derivative corrections*, Phys. Rev. **D75**, 066002 (2007) [`hep-th/0611318`].

[208] L. Bao, M. Cederwall and B. E. W. Nilsson, *Aspects of higher curvature terms and U-duality*, Class. Quant. Grav. **25**, 095001 (2008) [`arXiv:0706.1183`].

[209] C. Colonnello and A. Kleinschmidt, *Ehlers symmetry at the next derivative order*, JHEP **08**, 078 (2007) [`arXiv:0706.2816`].

[210] Y. Michel and B. Pioline, *Higher Derivative Corrections, Dimensional Reduction and Ehlers Duality*, JHEP **09**, 103 (2007) [`arXiv:0706.1769`].

[211] M. B. Green, M. Gutperle and H.-h. Kwon, λ^{16} *and related terms in M-theory on* T^2, Phys. Lett. **B421**, 149–161 (1998) [`hep-th/9710151`].

[212] A. Keurentjes, *The group theory of oxidation*, Nucl. Phys. **B658**, 303–347 (2003) [`hep-th/0210178`].

[213] S. Alexandrov, F. Saueressig and S. Vandoren, *Membrane and fivebrane instantons from quaternionic geometry*, JHEP **09**, 040 (2006) [`hep-th/0606259`].

[214] G. Francsics and P. D. Lax, *A Fundamental Domain for the Picard Modular Group in* $\mathbb{C}^2$, Vienna, Preprint ESI 1273 (2003).

[215] C. Schoen, *On the geometry of a special determinantal hypersurface associated to the Mumford-Horrocks vector bundle*, J. Reine Angew. Math. **481**, 73 (1996).

[216] P. Candelas, E. Derrick and L. Parkes, *Generalized Calabi-Yau manifolds and the mirror of a rigid manifold*, Nucl. Phys. **B407**, 115–154 (1993) [hep-th/9304045].

[217] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, *Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes*, Commun. Math. Phys. **165**, 311–428 (1994) [hep-th/9309140].

[218] P. S. Aspinwall and J. Louis, *On the Ubiquity of K3 Fibrations in String Duality*, Phys. Lett. **B369**, 233–242 (1996) [hep-th/9510234].

[219] I. Antoniadis, R. Minasian, S. Theisen and P. Vanhove, *String loop corrections to the universal hypermultiplet*, Class. Quant. Grav. **20**, 5079–5102 (2003) [hep-th/0307268].

[220] K. Behrndt, I. Gaida, D. Lust, S. Mahapatra and T. Mohaupt, *From type IIA black holes to T-dual type IIB D-instantons in $\mathcal{N} = 2, D = 4$ supergravity*, Nucl. Phys. **B508**, 659–699 (1997) [hep-th/9706096].

[221] M. de Vroome and S. Vandoren, *Supergravity description of spacetime instantons*, Class. Quant. Grav. **24**, 509–534 (2007) [hep-th/0607055].

[222] S. Elitzur, A. Giveon, D. Kutasov and E. Rabinovici, *Algebraic aspects of matrix theory on T^d*, Nucl. Phys. **B509**, 122–144 (1998) [hep-th/9707217].

[223] K. Becker and M. Becker, *Instanton action for type II hypermultiplets*, Nucl. Phys. **B551**, 102–116 (1999) [hep-th/9901126].

[224] G. Francsics and P. D. Lax Private Communication.

[225] G. Zhang, *Shimura invariant differential operators and their eigenvalues*, Math. Ann. **319**, 235–265 (2001).

[226] G. Francsics and P. D. Lax, *Spectral analysis on complex hyperbolic spaces*, Proceedings of Hayama Symposium on Complex analysis in Several Variables 120–126 (2005).

[227] M. Gunaydin, A. Neitzke, O. Pavlyk and B. Pioline, *Quasi-conformal actions, quaternionic discrete series and twistors: $SU(2,1)$ and $G_{2(2)}$*, Commun. Math. Phys. **283**, 169–226 (2008) [arXiv:0707.1669].

[228] I. Antoniadis, S. Ferrara, R. Minasian and K. S. Narain, *R^4 couplings in M- and type II theories on Calabi-Yau spaces*, Nucl. Phys. **B507**, 571–588 (1997) [hep-th/9707013].

[229] H. Gunther, C. Herrmann and J. Louis, *Quantum corrections in the hypermultiplet moduli space*, Fortsch. Phys. **48**, 119–123 (2000) [hep-th/9901137].

[230] L. Anguelova, M. Rocek and S. Vandoren, *Quantum corrections to the universal hypermultiplet and superspace*, Phys. Rev. **D70**, 066001 (2004) [hep-th/0402132].

[231] D. Robles-Llana, F. Saueressig and S. Vandoren, *String loop corrected hypermultiplet moduli spaces*, JHEP **03**, 081 (2006) [hep-th/0602164].

[232] S. M. Salamon, *Quaternionic Kähler manifolds*, Invent. Math. **67**, 143–171 (1982).

[233] A. Swann, *Hyper-Kähler and quaternionic Kähler geometry*, Math. Ann. **289**, 421–450 (1991).

[234] C. LeBrun, *Fano manifolds, contact structures, and quaternionic geometry*, Internat. J. Math. **6**, 419–437 (1982) [`dg-ga/9409001`].

[235] B. de Wit, B. Kleijn and S. Vandoren, *Superconformal hypermultiplets*, Nucl. Phys. **B568**, 475–502 (2000) [`hep-th/9909228`].

[236] B. de Wit, M. Rocek and S. Vandoren, *Hypermultiplets, hyperkaehler cones and quaternion-Kähler geometry*, JHEP **02**, 039 (2001) [`hep-th/0101161`].

[237] B. de Wit and F. Saueressig, *Off-shell $N = 2$ tensor supermultiplets*, JHEP **09**, 062 (2006) [`hep-th/0606148`].

[238] J. B. Etnyre, *Introductory Lectures on Contact Geometry*, `[math/0111118]`.

[239] R. Bohm, H. Gunther, C. Herrmann and J. Louis, *Compactification of type IIB string theory on Calabi-Yau threefolds*, Nucl. Phys. **B569**, 229–246 (2000) [`hep-th/9908007`].

[240] S. Gelbart and P.-S. Shapiro, *Automorphic Forms and L-Functions for the Unitary Group*, Notes based on lectures given by S. G. at "The University of Maryland Special Year on Lie Group Representations" (1982-1983).

[241] Y. Ishikawa, *The Generalized Whittaker Functions for $SU(2,1)$ and the Fourier Expansion of Automorphic Forms*, J. Math. Sci. Univ. Tokyo **6**, 477–526 (1999).

[242] A. I. Vinogradov and L. A. Takhtadzhyan, *Theory of Eisenstein series for the group $SL(3,\mathbb{R})$ and its application to a binary problem*, Journal of Soviet Mathematics **18:3** (1983).

[243] N. V. Proskurin, *Expansions of Automorphic Functions*, Translated from "Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta" im. V. A. Steklova AN USSR **116**, 119–141 (1982).

[244] B. Pioline and D. Persson, *The automorphic NS5-brane (2.0)*, `[arXiv:0902.3274v2]`.

[245] B. H. Gross and N. R. Wallach, *On quaternionic discrete series representations, and their continuations*, J. Reine. Angew. Math. **481**, 73–123 (1996).

[246] M. Przanowski, *Locally Hermite-Einstein, selfdual gravitational instantons*, Acta Phys. Polon. **B14**, 625 (1983).

[247] M. Przanowski, *Killing vector fields in self-dual, Euclidean Einstein spaces with* $\Lambda \neq 0$, J. Math. Phys. **32**, 1004 (1991).

[248] U. Gran, *GAMMA: A Mathematica package for performing Gamma-matrix algebra and Fierz transformations in arbitrary dimensions*, `[hep-th/0105086]`.