Lecture Notes in Physics

The Editorial Policy for Edited Volumes

The series *Lecture Notes in Physics* (LNP), founded in 1969, reports new developments in physics research and teaching - quickly, informally but with a high degree of quality. Manuscripts to be considered for publication are topical volumes consisting of a limited number of contributions, carefully edited and closely related to each other. Each contribution should contain at least partly original and previously unpublished material, be written in a clear, pedagogical style and aimed at a broader readership, especially graduate students and nonspecialist researchers wishing to familiarize themselves with the topic concerned. For this reason, traditional proceedings cannot be considered for this series though volumes to appear in this series are often based on material presented at conferences, workshops and schools.

Acceptance

A project can only be accepted tentatively for publication, by both the editorial board and the publisher, following thorough examination of the material submitted. The book proposal sent to the publisher should consist at least of a preliminary table of contents outlining the structure of the book together with abstracts of all contributions to be included. Final acceptance is issued by the series editor in charge, in consultation with the publisher, only after receiving the complete manuscript. Final acceptance, possibly requiring minor corrections, usually follows the tentative acceptance unless the final manuscript differs significantly from expectations (project outline). In particular, the series editors are entitled to reject individual contributions if they do not meet the high quality standards of this series. The final manuscript must be ready to print, and should include both an informative introduction and a sufficiently detailed subject index.

Contractual Aspects

Publication in LNP is free of charge. There is no formal contract, no royalties are paid, and no bulk orders are required, although special discounts are offered in this case. The volume editors receive jointly 30 free copies for their personal use and are entitled, as are the contributing authors, to purchase Springer books at a reduced rate. The publisher secures the copyright for each volume. As a rule, no reprints of individual contributions can be supplied.

Manuscript Submission

The manuscript in its final and approved version must be submitted in ready to print form. The corresponding electronic source files are also required for the production process, in particular the online version. Technical assistance in compiling the final manuscript can be provided by the publisher's production editor(s), especially with regard to the publisher's own LaTeX macro package which has been specially designed for this series.

LNP Homepage (springerlink.com)

On the LNP homepage you will find:

—The LNP online archive. It contains the full texts (PDF) of all volumes published since 2000. Abstracts, table of contents and prefaces are accessible free of charge to everyone. Information about the availability of printed volumes can be obtained.

—The subscription information. The online archive is free of charge to all subscribers of the printed volumes.

—The editorial contacts, with respect to both scientific and technical matters.

—The author's / editor's instructions.

H. Ocampo S. Paycha A. Vargas (Eds.)

Geometric and Topological Methods for Quantum Field Theory

Springer

Editors

Professor Dr. Hernan Ocampo
Universidad del Valle
Apartado Aereo 25360
Cali, Columbia
hernan@univalle.edu.co

Sylvie Paycha
Université Blaise Pascal (Clermont II)
Laboratoire de Mathématiques
Complexe des Cézeaux
63177 Aubiere Cedex, France
Sylvie.Paycha@math.univ-bpclermont.fr

Andrés Vargas
University of Bonn
Institute of Applied Mathematics
Wegeler Str. 6
53115 Bonn, Germany
vargas@math.uni-bonn.de

H. Ocampo et al., *Geometric and Topological Methods for Quantum Field Theory*, Lect. Notes Phys. 668 (Springer, Berlin Heidelberg 2005),
DOI 10.1007/b104936

Library of Congress Control Number: 2005923771

ISSN 0075-8450

ISBN-10 3-540-24283-X Springer Berlin Heidelberg New York
ISBN-13 978-3-540-24283-3 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springeronline.com

Printed in The Netherlands

Typesetting: by the authors and TechBooks using a Springer LaTeX macro package
Cover production: *design &production* GmbH, Heidelberg

Printed on acid-free paper SPIN: 11374060 54/3141/jl 5 4 3 2 1 0

A la memoria de Chepe Escobar

Por su gran contribución al desarrollo de las matemáticas
en Colombia y a su reconocimiento a nivel internacional.

Preface

This volume offers an introduction to some recent developments in several active topics at the interface between geometry, topology and quantum field theory:

– String theory and Calabi-Yau manifolds
– Knot invariants and configuration spaces
– Euclidean quantum field theory and noncommutative geometry
– Index theorems and noncommutative topology

It is based on lectures delivered during a summer school "Geometric and Topological Methods for Quantum Field Theory" held at Villa de Leyva, Colombia, in July 2003. The short communications presented by participants of the school appeared separately in the Annales Mathématiques Blaise Pascal (Vol. 11 n.2, 2004) of the University Blaise Pascal in Clermont-Ferrand, France.

These lecture notes, which are aimed at graduate students in physics or mathematics, start with introductory material before presenting more advanced results. Each chapter is self-contained and can be read independently of the others.

The volume starts with an introductory course by Christine Lescop on knot invariants and configuration spaces, in which a universal finite type invariant for knots is constructed as a series of integrals over configuration spaces. It is followed by an introduction by Raimar Wulkenhaar to euclidean quantum field theory from a statistical physics point of view, after which the author discusses possible renormalization techniques on noncommutative spaces. The third lecture by Anamaría Font and Stefan Theisen presents an elementary introduction to (string) compactification with unbroken supersymmetry. The authors show how this requirement leads to internal spaces of special holonomy and describe Calabi-Yau manifolds in detail. String compactification on orbifolds are presented as particular examples. The last lecture by Thierry Fack is devoted to a K-theory proof of the Atiyah-Singer index theorem and discusses some applications of K-theory to noncommutative geometry.

We hope that these lectures will give – as much as the school itself seems to have given – young students the desire to pursue what might be their first acquaintance with some of the problems on the edge of mathematics and physics presented in this volume. On the other hand, we hope that the more advanced reader will find some pleasure reading about different outlooks on related topics and seeing how well-known geometric tools prove to be very useful in some areas of quantum field theory.

We are indebted to various organizations for their financial support for this school. Let us thank first of all the French organizations "C.I.M.P.A." and ECOS-Nord, without which this school would not have taken place. This school is part of a long-term scientific program supported by ECOS-Nord between the Université Blaise Pascal in Clermont-Ferrand and the Universidad de los Andes in Bogotá in the areas of mathematics and physics. We are also deeply grateful to the I.C.T.P. in Trieste, for its constant financial support over the years and especially for this school. We also address our thanks to the French Embassy in Bogotá, and particularly to the former cultural attaché, Jean-Yves Deler, for showing interest in this project and supporting us at difficult times. We are also very much indebted to the Universidad de los Andes, which was our main source of financial support in Colombia. Other organizations such as C.L.A.F. in Brasil, Colciencias, I.C.E.T.E.X., and I.C.F.E.S. in Colombia, also contributed in a substancial way to the financial support needed for this school.

Special thanks to Sergio Adarve (Universidad de los Andes), coorganizer of the school, who dedicated much time and energy to make this school possible in a country like Colombia where many difficulties are bound to arise along the way due to social, political and economic problems. This school would not have taken place without him. Our special thanks to José Rafael Toro, vice-rector of the Universidad de los Andes, for his constant encouragements and efficient support. From the Universidad de los Andes we also want to express our sincere gratitude to Carlos Montenegro, head of the Department of Mathematics, Rolando Roldán, dean of the Faculty of Science, and Bernardo Gómez, head of the Department of Physics.

We are extremely grateful to Juana Vall-Serra who did an extraordinary job for the practical organization of the school, the quality of which was very much appreciated by participants and lecturers. We are also very indebted to Marta Kovacsics, Alexandra Parra, Germán Barragán, María Fernanda Gómez, Andrés García, Juan Esteban Martín and Andersson García for their help with various essential tasks needed for the successful development of the school and to the academic coordinators, Alexander Cardona and Andrés Reyes, who jointly with Bernardo Uribe and Mario Fernando H. Arias participated actively with their permanent assistance and collaboration. Without the people named here, all of whom helped with the organization in some way or another, before, during and after the school, this scientific event would not

have left such vivid memories in the lecturers' and the participants' minds. Last but not least, thanks to all the participants who gave us all, lecturers and editors, the impulse to prepare this volume through the enthusiasm they showed us during the school.

February 2005

Hernán Ocampo
Sylvie Paycha
Andrés Vargas

Contents

List of Contributors

C. Lescop
CNRS, Institut Fourier (UMR 5582), B.P. 74,
38402 Saint-Martin-d'Hères Cedex, France
lescop@ujf-grenoble.fr

R. Wulkenhaar
Max-Planck-Institute for Mathematics in the Sciences,
Inselstraße 22-26, D-04103 Leipzig, Germany
raimar.wulkenhaar@mis.mpg.de

A. Font
Departamento de Física, Facultad de Ciencias, Universidad
Central de Venezuela, A.P. 20513,
Caracas 1020-A, Venezuela
afont@fisica.ciens.ucv.ve

S. Theisen
Albert-Einstein-Institut,
Am Mühlenberg 1,
D-14476 Golm, Germany
theisen@aei.mpg.de

T. Fack
Université Claude Bernard (Lyon 1),
UFR de Mathématiques,
Institut Girard Desargues (UMR 5028), bâtiment Doyen Jean Braconnier, 43 boulevard du 11 novembre 1918
69622 Villeurbanne Cedex, France
Thierry.Fack@univ-lyon1.fr

Knot Invariants and Configuration Space Integrals

C. Lescop

CNRS, Institut Fourier (UMR 5582), B.P. 74, 38402 Saint-Martin-d'Hères Cedex, France
lescop@ujf-grenoble.fr

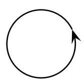
Summary. After a short presentation of the theory of Vassiliev knot invariants, we shall introduce a universal finite type invariant for knots in the ambient space. This invariant is often called the perturbative series expansion of the Chern-Simons theory of links in the euclidean space. It will be constructed as a series of integrals over configuration spaces.

1 First Steps in the Folklore of Knots, Links and Knot Invariants

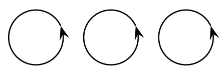
1.1 Knots and Links

Intuitively, a *knot* is a circle embedded in the ambient space up to elastic deformation; a *link* is a finite family of disjoint knots.

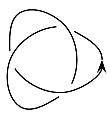
Examples 1.1. Here are some pictures of simple knots and links. More examples can be found in [32].



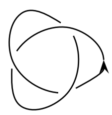
The *trivial knot*



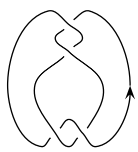
The *trivial 3-component link*



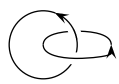
The *right-handed trefoil knot*



The *left-handed trefoil knot*

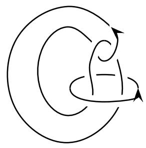


The *figure-eight knot*

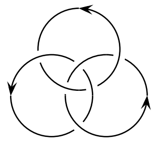


The *Hopf link*

C. Lescop: *Knot Invariants and Configuration Space Integrals*, Lect. Notes Phys. **668**, 1–57 (2005)
www.springerlink.com



The *Whitehead link*



The *Borromean rings*

Let $\coprod^k S^1$ denote the disjoint union of k **oriented** circles. We will represent a *knot* (resp. a k-component link) by a C^∞ embedding[1] of the circle S^1 (resp. of $\coprod^k S^1$) into the ambient space $\mathbf{R}^3$.

Definition 1.2. An *isotopy* of $\mathbf{R}^3$ is a C^∞ map

$$h : \mathbf{R}^3 \times I \longrightarrow \mathbf{R}^3$$

such that $h_t = h(.,t)$ is a diffeomorphism for all $t \in [0,1]$. Two embeddings f and g as above are said to be *isotopic* if there is an isotopy h of $\mathbf{R}^3$ such that $h_0 =$ Identity and $g = h_1 \circ f$. Link isotopy is an equivalence relation. (Checking transitivity requires smoothing... See [17].)

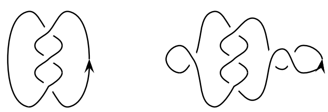
Definition 1.3. A *knot* is an isotopy class of embeddings of S^1 into $\mathbf{R}^3$. A *k-component link* is an isotopy class of embeddings of $\coprod^k S^1$ into $\mathbf{R}^3$.

Definition 1.4. Let $\pi : \mathbf{R}^3 \longrightarrow \mathbf{R}^2$ be the projection defined by $\pi(x,y,z) = (x,y)$. Let $f : \coprod^k S^1 \hookrightarrow \mathbf{R}^3$ be a representative of a link L. The *multiple* (resp. *double*) points of $\pi \circ f$ are the points of $\mathbf{R}^2$ that have several (resp. two) inverse images under $\pi \circ f$. A double point is said to be *transverse* if the two tangent vectors to $\pi \circ f$ at this point generate $\mathbf{R}^2$. $\pi \circ f : \coprod^k S^1 \hookrightarrow \mathbf{R}^2$ is a *regular projection* of L if and only if $\pi \circ f$ is an immersion whose only multiple points are transverse double points.

Proposition 1.5. *Any link L has a representative f whose projection $\pi \circ f$ is regular.*

A sketch of proof of this proposition is given in Subsect. 7.1.

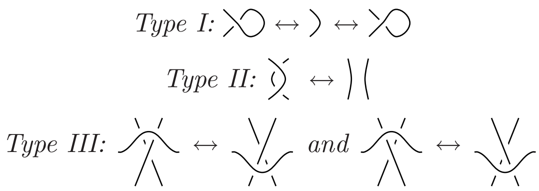
Definition 1.6. A *diagram* of a link is a regular projection equipped with the additional under/over information: at a double point, the strand that crosses under is broken near the crossing. Note that a link is well-determined by one of its diagrams. The converse is not true as the following diagrams of the right-handed trefoil knot show.



[1] The reader is referred to [17] for the basic concepts of differential topology as well as for more sophisticated ones.

Nevertheless, we have the following theorem whose proof is outlined in Sect. 7.1.

Theorem 1.7 (Reidemeister theorem (1926) [31]). *Up to orientation-preserving diffeomorphism of the plane, two diagrams of a link can be related by a finite sequence of* Reidemeister moves *that are local changes of the following type:*

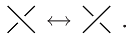


As it will often be the case during this course, these pictures represent local changes. They should be understood as parts of bigger diagrams that are unchanged outside their pictured parts.

Exercise 1.8. (**) Prove that there are at most 4 knots that can be represented with at most 4 crossings (namely, the trivial knot, the two trefoil knots and the figure-eight knot).

Exercise 1.9. (**) Prove that the set of knots is countable.

A *crossing change* is a local modification of the type



Proposition 1.10. *Any link can be unknotted by a finite number of crossing changes.*

PROOF: At a philosophical level, it comes from the fact that $\mathbf{R}^3$ is simply connected, and that a homotopy $h : S^1 \times [0,1] \longrightarrow \mathbf{R}^3$ that transforms a link into a trivial one can be replaced by a homotopy that is an isotopy except at a finite number of times where it is a crossing change. (Consider $h \times 1_{[0,1]} : (x,t) \mapsto (h(x,t),t) \in \mathbf{R}^3 \times [0,1]$. The homotopy h can be perturbed so that $h \times 1_{[0,1]}$ is an immersion with a finite number of multiple points that are transverse double points –[17, Exercise 1, p.82]–. . .)

However, an elementary constructive proof of this proposition from a link diagram is given in Subsect. 7.1. ◇

The general open problem in knot theory is to find a satisfactory classification of knots, that is an intelligent way of producing a complete and repetition-free list of knots. A less ambitious task is to be able to decide from two knot presentations whether these presentations represent the same knot. This will sometimes be possible with the help of knot invariants.

1.2 Link Invariants

Definition 1.11. A link *invariant* is a map from the set of links to another set.

Such a map can be defined as a function of diagrams that is invariant under the Reidemeister moves.

Definition 1.12. A *positive crossing* in a diagram is a crossing that looks like  (up to rotation of the plane). (The "shortest" arc that goes from the arrow of the top strand to the arrow of the bottom strand turns counterclockwise.) A *negative crossing* in a diagram is a crossing that looks like  .

Definition 1.13. The *linking number* for two-component links is half the number of the positive crossings that involve the two components minus half the number of negative crossings that involve the two components.

Exercise 1.14. Use Reidemeister's theorem to prove that the linking number is a link invariant. Use the linking number to distinguish the Hopf link from the Whitehead link.

1.3 An Easy-to-Compute Non-Trivial Link Invariant: The Jones Polynomial

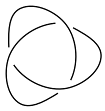
In this section, we will show an example of an easy-to-compute non-trivial link invariant: The Jones polynomial. We will follow the Kauffman approach that is different from the original approach of Jones who discovered this polynomial in 1984.

Let D be an unoriented link diagram. A crossing  of D can be removed in two different ways:
the left-handed one where  becomes ,
(someone walking on the upper strand towards the crossing turns left just before reaching the crossing)
and the right-handed one where  becomes .
Let $C(D)$ denote the set of crossings of D and let f be a map from $C(D)$ to $\{L, R\}$, then D_f will denote the diagram obtained by removing every crossing x, in the left-handed way if $f(x) = L$ and in the right-handed way if $f(x) = R$. D_f is nothing but a collection of $n(D_f)$ circles embedded in the plane.

We define the *Kauffman bracket* $< D >\in \mathbf{Z}[A, A^{-1}]$ of D as

$$< D >= \sum_{f:C(D)\longrightarrow\{L,R\}} A^{(\sharp f^{-1}(L)-\sharp f^{-1}(R))}\delta^{(n(D_f)-1)}$$

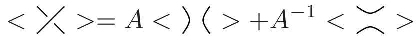
with $\delta = -A^2 - A^{-2}$.

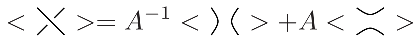
Exercise 1.15. Compute $<$  $>= -A^5 - A^{-3} + A^{-7}$.

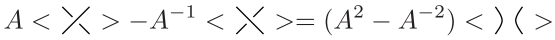
The Kauffman bracket satisfies the following properties:

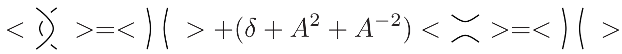
1. $<$ n disjoint circles $>= \delta^{n-1}$

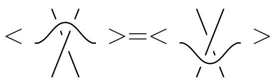
and we have the following equalities that relate brackets of diagrams that are identical anywhere except where they are drawn.

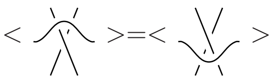
2. 

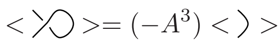
3. 

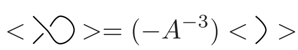
4. 

5. 

6. 

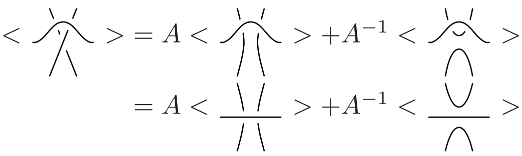
6′. 

7. 

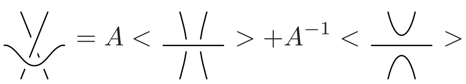
7′. 

8. The Kauffman bracket of the mirror image of a diagram D is obtained from $< D >$ by exchanging A and A^{-1}.

PROOF: The first five properties of the Kauffman bracket and the eighth one are straightforward. Property 5 shows that δ has been chosen to get invariance under the second Reidemeister move. Let us prove Equality 6.



The second equality comes from the above invariance under the second Reidemeister move. Rotating the picture by π yields:



and concludes the proof of Equality 6. Equality $6'$ is deduced from Equality 6 by mirror image (up to rotation). Equality 7 is easy and Equality $7'$ is its mirror image, too. ◇

Definition 1.16. The *writhe* $w(D)$ of an oriented link diagram D is the number of positive crossings minus the number of negative ones.

Theorem 1.17. *The Jones polynomial $V(L)$ of an oriented link L is the Laurent polynomial of $\mathbf{Z}[t^{1/2},\ t^{-1/2}]$ defined from an oriented diagram D of L by*

$$V(L) = (-A)^{-3w(D)} < D >_{A^{-2}=t^{1/2}}$$

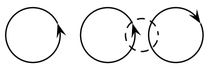
V is an invariant of oriented links. It is the unique invariant of oriented links that satisfies:

1. *V(trivial knot)=1,*
2. *and the* skein relation:

$t^{-1}V($$) - tV($$) = (t^{1/2} - t^{-1/2})V($$)$

Proof: It is easy to check that $(-A)^{-3w(D)} < D > \in \mathbf{Z}[A^2, A^{-2}]$. It is also easy to check that this expression is invariant under the Reidemeister moves since the only Reidemeister moves that change the writhe are the moves I, and the variation of $(-A)^{-3w(D)}$ under these moves makes up for the variation of $< D >$. The skein relation can be easily deduced from the fourth property of the Kauffman bracket, and it is immediate that $V(\text{trivial knot}) = 1$.

Let us prove that these two properties *uniquely* determine V. Applying the skein relation in the case when  is the surrounded part of a standard diagram of a trivial n-component link like:



and  and  are therefore (parts of) diagrams of trivial $(n-1)$-component links shows that

$$(t^{-1} - t)V(\text{trivial (n-1)-comp. link}) = (t^{1/2} - t^{-1/2})V(\text{trivial n-comp. link})$$

and thus determines

$$V(\text{trivial n-component link}) = (-t^{1/2} - t^{-1/2})^{n-1} .$$

By induction on the number of crossings, and, for a fixed number of crossings, by induction on the number of crossings to be changed in the projection to make the projected link trivial (see Proposition 1.10), the link invariant V is determined by its value at the trivial knot and the skein relation.

◇

Exercise 1.18. Compute $V(\text{right-handed trefoil}) = -t^4 + t^3 + t$ and show that the right-handed trefoil is not isotopic to the left-handed trefoil.

Knot invariants are especially interesting when they can be used to derive general properties of knots. The main known application of the Jones polynomial is given and proved in Subsect. 7.2. The Chern-Simons series defined in Sect. 4 will be used to prove the fundamental theorem of Vassiliev invariants 2.26.

2 Finite Type Knot Invariants

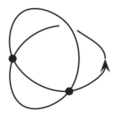
2.1 Definition and First Examples

Definition 2.1. A *singular knot* with n double points is represented by an immersion from S^1 to $\mathbf{R}^3$ with n transverse double points that is an embedding when restricted to the complement of the preimages of these double points.

Two such immersions f and g are said to be *isotopic* if there is an isotopy h of $\mathbf{R}^3$ such that $h_0 =$ Identity and $g = h_1 \circ f$.

A *singular knot* with n double points is an isotopy class of such immersions with n double points.

Example 2.2.



A *singular knot with two double points*

Definition 2.3. Let be a double point of a singular knot. This double point can disappear in a *positive way* by changing into , or in a *negative way* by changing into . Note that the sign of such a *desingularisation*, that can be seen in a diagram as above, is defined from the orientation of the ambient space. Choose one strand involved in the double point. Call this strand the first one. Consider the tangent plane to the double point, that is the vector plane equipped with the basis (tangent vector $\boldsymbol{v_1}$ to the first strand, tangent vector $\boldsymbol{v_2}$ to the second one). This basis orients the plane, and allows us to define a positive normal vector $\boldsymbol{n}$ to the plane (that is a vector $\boldsymbol{n}$ orthogonal to $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ such that the triple $(\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{n})$ is an oriented basis of $\mathbf{R}^3$). The n the *positive desingularisation* is obtained by pushing the first strand in the direction of $\boldsymbol{n}$. Note that this definition is independent of the choice of the *first* strand.

Notation 2.4. Let K be a singular knot with n double points numbered by 1,2 $\dots$, n. Let f be a map from $\{1, 2 \dots, n\}$ to $\{+, -\}$. Then K_f denotes the genuine knot obtained by removing every crossing i by the transformation: $_i$ becomes if $f(i) = +$, and $_i$ becomes if $f(i) = -$.

Let $\mathcal{K}$ denote the set of knots. Let $\mathbf{Z}[\mathcal{K}]$ denote the free $\mathbf{Z}$-module with basis $\mathcal{K}$. Then $[K]$ denotes the following element of $\mathbf{Z}[\mathcal{K}]$:

$$[K] = \sum_{f:\{1,2\ldots,n\}\longrightarrow\{+,-\}} (-1)^{\sharp f^{-1}(-)} K_f$$

where the symbol $\sharp$ is used to denote the cardinality of a set.

Proposition 2.5. *We have the following equality in V:*

[] = [] − []

that relates the brackets of three singular knots that are identical outside a ball where they look like in the above pictures.

PROOF: Exercise. ◇

Definition 2.6. Let G be an abelian group. Then any G-valued knot invariant I is extended to $\mathbf{Z}[\mathcal{K}]$, linearly. It is then extended to singular knots by the formula

$$I(K) \stackrel{\text{def}}{=} I([K])$$

Let n be an integer. A G-valued knot invariant I is said to be of *degree less or equal than* n, if it vanishes at singular knots with $(n+1)$ double points. Of course, such an invariant is of degree n if it is of degree less or equal than n without being of degree less or equal than $(n-1)$. A G-valued knot invariant I is said to be of *finite type* or of *finite degree* if it is of degree n for some n. Note that we could have defined the extension of a G-valued knot invariant to singular knots by induction on the number of double points using the induction formula:

$I($ $) = I($$) - I($$)$.

Performing the following change of variables:

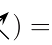
$$t^{1/2} = -\exp\left(-\frac{\lambda}{2}\right)$$

transforms the Jones polynomial into a series

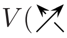
$$V(L) = \sum_{n=0}^{\infty} v_n(L)\lambda^n$$

Proposition 2.7. *The coefficient $v_n(L)$ of the renormalized Jones polynomial is a rational invariant of degree less or equal than n.*

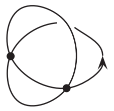
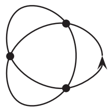
PROOF: Under the above change of variables, the skein relation satisfied by the Jones polynomial becomes

$\exp(\lambda)V($$) - \exp(-\lambda)V($$) = \left(\exp\left(\frac{\lambda}{2}\right) - \exp\left(-\frac{\lambda}{2}\right)\right) V($$)$

that is equivalent to

$V($$) = \left(1 - e^{\lambda}\right) V($$) + \left(e^{-\lambda} - 1\right) V($$) + \left(e^{\lambda/2} - e^{-\lambda/2}\right) V($$)$

This equality allows us to see that the series V of a singular knot with n double points has valuation (i.e. degree of the term of minimal degree) at least n by induction on n. ◇

Exercise 2.8. 1. Compute $v_2($$)$ and $v_3($$)$.

2. Show that v_2 and v_3 are exactly of degree 2 and 3, respectively.

Definition 2.9. Let V be a vector space. A *filtration* of V is a decreasing sequence of vector spaces

$$V = V_0 \supseteq V_1 \supseteq \cdots \supseteq V_n \supseteq V_{n+1} \supseteq \ldots$$

that begins with $V_0 = V$. The *graded space* associated to such a filtration is the vector space

$$\mathcal{G} = \oplus_{k=1}^{\infty} \mathcal{G}_k((V_n)_{n\in\mathbf{N}}) \text{ where } \mathcal{G}_k((V_n)_{n\in\mathbf{N}}) = \frac{V_k}{V_{k+1}}$$

Let $\mathcal{V}$ denote the $\mathbf{R}$-vector space freely generated by the knots. Let $\mathcal{V}_n$ denote the subspace of $\mathcal{V}$ generated by the brackets of the singular knots with n double points. The *Vassiliev filtration of* $\mathcal{V}$ is the sequence of the $\mathcal{V}_n$. Let $\mathcal{I}_n$ denote the set of real-valued invariants of degree less or equal than n. The set $\mathcal{I}_n$ is nothing but the dual vector space of $\frac{\mathcal{V}}{\mathcal{V}_{n+1}}$ that is the space of linear forms on $\frac{\mathcal{V}}{\mathcal{V}_{n+1}}$.

$$\mathcal{I}_n = \mathrm{Hom}\left(\frac{\mathcal{V}}{\mathcal{V}_{n+1}}; \mathbf{R}\right) = \left(\frac{\mathcal{V}}{\mathcal{V}_{n+1}}\right)^*$$

Proposition 2.10.

$$\frac{\mathcal{I}_n}{\mathcal{I}_{n-1}} = \left(\frac{\mathcal{V}_n}{\mathcal{V}_{n+1}}\right)^*$$

Proof: Exercise.

Exercise 2.11. (*) Let $\lambda \in \mathcal{I}_n$, let $\mu \in \mathcal{I}_m$. Define the invariant $\lambda\mu$ at genuine knots K by $\lambda\mu(K) = \lambda(K)\mu(K)$. Prove that $\lambda\mu \in \mathcal{I}_{n+m}$.

Exercise 2.12. Prove that

$$\dim\left(\frac{\mathcal{V}_2}{\mathcal{V}_3}\right) \geq 1 \quad \text{and} \quad \dim\left(\frac{\mathcal{V}_3}{\mathcal{V}_4}\right) \geq 1 \, .$$

We already know that there exist finite type invariants. Now, we ask the natural question.

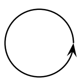
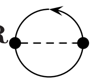
How many finite type invariants does there exist? In other words, what is the dimension of $\frac{\mathcal{V}_n}{\mathcal{V}_{n+1}}$?

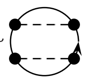
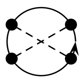
In the next subsection, we are going to bound this dimension from above. This dimension will then be theoretically given by the fundamental theorem of Vassiliev invariants 2.26 that will be roughly proved at the end of Sect. 5, as the dimension of a vector space presented by a finite number of generators and relators.

2.2 Chord Diagrams

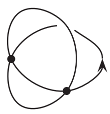
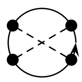
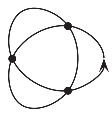
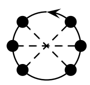
Definition 2.13. An n-chord diagram on a one-manifold M is an isotopy class of embeddings of a finite set U with $2n$ elements, equipped with a partition into n pairs, into the interior of M, up to a permutation of U that preserves the partition.

For us, so far, $M = S^1$. In particular, the datum of the above isotopy class is equivalent to the datum of a cyclic order on U where a *cyclic order* on a finite set U is a cyclic permutation σ of this set, that provides every element u of U with a unique successor (namely $\sigma(u)$). An n-chord diagram on M will be represented by an embedding of M equipped with $2n$ points (the image of the embedding of U) where the points of a pair are related by a dotted chord. The $\mathbf{R}$-vector space freely generated by the n-chord diagrams on M will be denoted by $\mathcal{D}_n(M)$.

Examples 2.14. $\mathcal{D}_0(S^1) = \mathbf{R}$ , $\mathcal{D}_1(S^1) = \mathbf{R}$ ,

$\mathcal{D}_2(S^1) = \mathbf{R}$  $\oplus\, \mathbf{R}$  .

Notation 2.15. Let K be a singular knot with n double points. The *n-chord diagram* $D(K)$ associated to K is the set of the $2n$ inverse images of the double points of K (cyclically ordered by the orientation of S^1), equipped with the partition into pairs where every pair contains the inverse images of one double point.

Examples 2.16. $D($  $) =$ , $D($  $) =$ 

Lemma 2.17. *Every chord diagram on S^1 is the diagram of a singular knot. Furthermore, two singular knots which have the same diagram are related by a finite number of crossing changes (where* a crossing change *is again a modification which transforms*  *into*  *).*

Sketch of proof: Consider a chord diagram D. First embed the neighborhoods in S^1 of the double points (that correspond to the chords) into disjoint balls of $\mathbf{R}^3$. Let C be the complement in $\mathbf{R}^3$ of these disjoint balls. We have enough room in C to join the neighborhoods of the double points by arcs embedded in C, in order to get a singular knot K_0 with associated diagram D.

Now, consider another singular knot K with the same diagram. There exists an isotopy that carries the neighborhoods of its double points on the corresponding neighborhoods for K_0. The arcs that join the double points in the simply connected C are homotopic to the former ones. Therefore, they can be carried to the former ones by a sequence of isotopies and crossing changes. A diagrammatic proof of this lemma is given in Subsect. 7.3. ◇

Notation 2.18. Let ϕ_n denote the linear map from $\mathcal{D}_n(S^1)$ to $\frac{\mathcal{V}_n}{\mathcal{V}_{n+1}}$ that maps an n-chord diagram D to the projection in $\frac{\mathcal{V}_n}{\mathcal{V}_{n+1}}$ of a singular knot K such that $D(K) = D$. The above lemma ensures that ϕ_n is well-defined. Indeed, if $D(K) = D(K')$, then, by the above lemma, K and K' are related by a finite number of crossing changes. Therefore, $[K] - [K']$ is an element of $\mathcal{V}_{n+1}$.

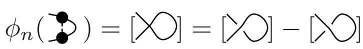
Furthermore, it is obvious that ϕ_n is onto. As a consequence, the dimension of $\frac{\mathcal{V}_n}{\mathcal{V}_{n+1}}$ is bounded from above by the number of n-chord diagrams.

Exercise 2.19. Prove that this number is bounded from above by $\frac{(2n)!}{2^n n!}$, and by $1 + (n-1)\frac{(2(n-1))!}{2^{(n-1)}(n-1)!}$ if $n \geq 1$. Improve these upper bounds.

The following lemmas will allow us to improve the upper bound on the dimension of $\frac{\mathcal{V}_n}{\mathcal{V}_{n+1}}$. In a chord diagram, an *isolated chord* is a chord that relates two consecutive points.

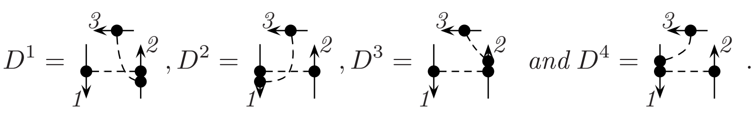
Lemma 2.20. *Let D be a diagram on S^1 that contains an isolated chord. Then $\phi_n(D) = 0$.*

Proof: In $\frac{\mathcal{V}_n}{\mathcal{V}_{n+1}}$, we can write:



◇

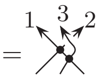
Lemma 2.21. *Let D^1, D^2, D^3, D^4 be four n-chord diagrams that are identical outside three portions of circles where they look like:*



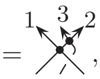
then

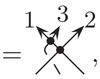
$$\phi_n(-D^1 + D^2 + D^3 - D^4) = 0 \; .$$

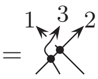
Proof: We may represent D^1 by a singular knot K^1 with n double points that intersects a ball like

$K^1 =$ 

Let K^2, K^3, K^4 be the singular knots with n double points that coincide with K^1 outside this ball, and that intersect this ball like in the picture:

$K^2 =$ 

, $K^3 =$ 

, $K^4 =$ 

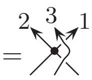
Then $D(K^2) = D^2$, $D(K^3) = D^3$ and $D(K^4) = D^4$. Therefore, $\phi_n(-D^1 + D^2 + D^3 - D^4) = -[K^1] + [K^2] + [K^3] - [K^4]$.

Thus, it is enough to prove that in $\mathcal{V}$ we have

$$-[K^1] + [K^2] + [K^3] - [K^4] = 0$$

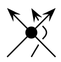
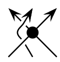
Let us prove this.

Let K_0 be the singular knot with $(n-1)$ double points that intersects our ball like

$K_0 =$ 

and that coincides with K^1 outside this ball.

The strands 1 and 2 involved in the pictured double point are in the horizontal plane and they orient it, the strand 3 is vertical and intersects the horizontal plane in a positive way between the tails of 1 and 2. Now, make 3 turn around the double point counterclockwise, so that it becomes successively the knots with $n-1$ double points:

$K_1 =$  , $K_2 =$  and $K_3 =$  .

On its way, it goes successively through our four knots K^1, K^2, K^3 and K^4 with n double points that appear inside matching parentheses in the following obvious identity in $\mathcal{V}_{n-1}$

$$([K_1] - [K_0]) + ([K_2] - [K_1]) + ([K_3] - [K_2]) + ([K_0] - [K_3]) = 0 \;.$$

Now, $[K^i] = \pm([K_i] - [K_{i-1}])$ where the sign $\pm$ is plus when the vertical strand goes through an arrow from K_{i-1} to K_i and minus when it goes through a tail. Therefore the above equality can be written as

$$-[K^1] + [K^2] + [K^3] - [K^4] = 0$$

and finishes the proof of the lemma. ◇

Notation 2.22. The relation

$$D^2 + D^3 - D^1 - D^4 = 0$$

that involves diagrams D^1, D^2, D^3, D^4 which satisfy the hypotheses of Lemma 2.21 is called the *four-term relation* and is denoted by $(4T)$. Let $\mathcal{A}_n(M)$ denote the quotient of $\mathcal{D}_n(M)$ by the relation $(4T)$, that is the quotient of $\mathcal{D}_n(M)$ by the subspace of $\mathcal{D}_n(M)$ generated by left-hand sides of $(4T)$, that are expressions of the form $(-D^1 + D^2 + D^3 - D^4)$, for a (D^1, D^2, D^3, D^4) satisfying the conditions of Lemma 2.21.

$$\mathcal{A}_n = \frac{\mathcal{D}_n}{4T}$$

Let $\overline{\mathcal{A}_n}(M)$ denote the quotient of $\mathcal{A}_n(M)$ by the subspace of $\mathcal{A}_n(M)$ generated by the projections of the diagrams with an isolated chord. The relation

$$D = 0 \text{ for a diagram } D \text{ with an isolated chord.}$$

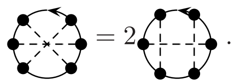
is sometimes called the one-term relation, and it is denoted by $(1T)$.

Lemmas 2.20 and 2.21 show the following proposition.

Proposition 2.23. *The map ϕ_n factors through $\overline{\mathcal{A}_n}(S^1)$ to define the surjective map:*

$$\overline{\phi_n} : \overline{\mathcal{A}_n}(S^1) \longrightarrow \frac{\mathcal{V}_n}{\mathcal{V}_{n+1}} .$$

Exercise 2.24. 1. Show that, in $\overline{\mathcal{A}_3}(S^1)$, we have



2. Show that

$$\dim\left(\frac{\mathcal{V}_0}{\mathcal{V}_1}\right) = 1 , \quad \dim\left(\frac{\mathcal{V}_1}{\mathcal{V}_2}\right) = 0 , \quad \dim\left(\frac{\mathcal{V}_2}{\mathcal{V}_3}\right) \leq 1 \quad \text{and} \quad \dim\left(\frac{\mathcal{V}_3}{\mathcal{V}_4}\right) \leq 1$$

3. Show that

$$\dim(\overline{\mathcal{A}_2})(S^1) = 1$$

Definition 2.25. The topological vector space $\prod_{k=0}^{\infty} \overline{\mathcal{A}_k}(M)$ will be denoted by $\overline{\mathcal{A}}(M)$. Similarly, the topological vector space $\prod_{k=0}^{\infty} \mathcal{A}_k(M)$ will be denoted by $\mathcal{A}(M)$. The *degree n part* of an element a of $\mathcal{A}(M)$ or $\overline{\mathcal{A}}(M)$ is the natural projection of a in $\mathcal{A}_n(M)$ or $\overline{\mathcal{A}}_n(M)$. It is denoted by a_n.

In the next sections, we shall prove that $\overline{\phi_n}$ is an isomorphism by constructing its inverse given by the Chern-Simons series. More precisely, we shall prove the following theorem.

Theorem 2.26 (Kontsevich, 92). *There exists a linear map*

$$\overline{Z} : \mathcal{V} \longrightarrow \overline{\mathcal{A}}(S^1)$$

such that, for any integer n *and for any singular knot* K *with* n *double points, if* $k < n$*, then*

$$\overline{Z}_k([K]) = 0 \ ,$$

and

$$\overline{Z}_n([K]) = D(K) \ .$$

In other words, the restriction to $\frac{\mathcal{V}_n}{\mathcal{V}_{n+1}}$ of $\overline{Z}_n$ will be the inverse of $\overline{\phi_n}$.

Corollary 2.27. *Any degree* n **R***-valued invariant* I *is of the form*

$$I = \psi \circ \overline{Z}$$

where ψ *is a linear form on* $\overline{\mathcal{A}}(S^1)$ *that vanishes on* $\prod_{k=n+1}^{\infty} \overline{\mathcal{A}_k}(S^1)$.

Proof: The Kontsevich theorem allows us to see, by induction on n, that $\overline{Z}$ induces the isomorphism

$$p_{\leq n} \circ \overline{Z} : \frac{\mathcal{V}}{\mathcal{V}_{n+1}} \longrightarrow \prod_{k=0}^{n} \overline{\mathcal{A}_k}(S^1)$$

where $p_{\leq n} : \overline{\mathcal{A}}(S^1) \longrightarrow \prod_{k=0}^{n} \overline{\mathcal{A}_k}(S^1)$ is the natural projection. Thus, any degree n **R**-valued invariant I is of the form

$$I = I \circ \left(p_{\leq n} \circ \overline{Z}\right)^{-1} \circ p_{\leq n} \circ \overline{Z} = \psi \circ \overline{Z}$$

where $\psi = I \circ \left(p_{\leq n} \circ \overline{Z}\right)^{-1} \circ p_{\leq n}$ is a linear form on $\overline{\mathcal{A}}(S^1)$ that vanishes on $\prod_{k=n+1}^{\infty} \overline{\mathcal{A}_k}(S^1)$. $\diamond$

In Sect. 4, we shall construct the Chern-Simons series

$$\overline{Z}^0_{CS} : \mathcal{V} \longrightarrow \overline{\mathcal{A}}(S^1)$$

as a series of integrals of configuration spaces. We shall show how to prove that the Chern-Simons series satisfies the properties of $\overline{Z}$ in the Kontsevich theorem above (2.26) in Sect. 5. Thus we will be able to construct all the real-valued finite type knot invariants as above, theoretically. Unfortunately, there are two problems. First, the Chern-Simons series is hard to compute explicitly. Second, we do not know a canonical basis for $\overline{\mathcal{A}_n}(S^1)$. The dimension of this vector space is unknown for a general n, although it can be computed by an algorithm that lists the finitely many diagrams and the finitely many relations and that computes the dimension of the quotient space. We first give another presentation of the vector spaces $\mathcal{A}_n(M)$ and study some of their properties in the next section.

3 Some Properties of Jacobi-Feynman Diagrams

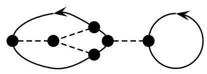
3.1 Jacobi Diagrams

Definition 3.1. Let M be an oriented one-manifold. A *diagram*, or a *Jacobi diagram,* Γ with support M is a finite uni-trivalent graph Γ such that every connected component of Γ has at least one univalent vertex, equipped with:

1. an isotopy class of injections i of the set U of univalent vertices of Γ into the interior of M,
2. an *orientation* of every trivalent vertex, that is a cyclic order on the set of the three half-edges which meet at this vertex,

Such a diagram Γ is again represented by a planar immersion of $\Gamma \cup M$ where the univalent vertices of U are located at their images under i, the one-manifold M is represented by solid lines, whereas the diagram Γ is dashed. The vertices are represented by big points. The local orientation of a vertex is represented by the counterclockwise order of the three half-edges that meet at it.

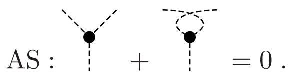
Here is an example of a diagram Γ on the disjoint union $M = S^1 \coprod S^1$ of two circles:

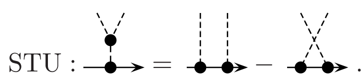


The *degree* of such a diagram is half the number of all the vertices of Γ.

Of course, a chord diagram is a diagram on a one-manifold M without trivalent vertices.

Let $\mathcal{D}^t_n(M)$ denote the real vector space generated by the degree n diagrams on M, and let $\mathcal{A}^t_n(M)$ denote the quotient of $\mathcal{D}^t_n(M)$ by the following relations AS and STU:

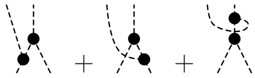




As usual, each of these relations relate diagrams which are identical outside the pictures where they are like in the pictures.

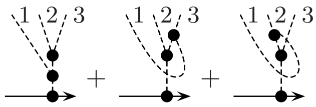
3.2 The Relation IHX in $\mathcal{A}^t_n$

Proposition 3.2. *Let M be a compact one-manifold, then the following relation IHX*

IHX :  $= 0$

(that relates three diagrams that are represented by three immersions which coincide outside a disk D where they are like in the pictures) is true in $\mathcal{A}_n^t(M)$.

PROOF: We want to prove that the relation IHX is true in the quotient of $\mathcal{D}_n^t(M)$ by AS and STU. Consider three diagrams that are represented by three immersions which coincide outside a disk D where they are like in the pictures involved in the relation IHX. Use STU as long as it is possible to remove all trivalent vertices that can be removed without changing the two vertices in D, simultaneously on the three diagrams. This transforms the relation IHX to be shown into a sum of similar relations, where one of the four entries of the disk is directly connected to M. Thus, since the four entries play the same role, we may assume that the relation IHX to be shown is:



$= 0$.

Using STU twice and AS transforms the summands of the left-hand side into diagrams that can be represented by three straight lines from the entries 1,2,3 to three fixed points of the horizontal line numbered from left to right. When the entry $i \in \{1,2,3\}$ is connected to the point $\sigma(i)$ of the horizontal plain line, where σ is a permutation of of $\{1,2,3\}$, the corresponding diagram will be denoted by $(\sigma(1)\sigma(2)\sigma(3))$. Thus, the expansion of the left-hand side of the above equation is

$$\begin{aligned} &((123)-(132)-(231)+(321)) \\ -&((213)-(231)-(132)+(312)) \\ -&((123)-(213)-(312)+(321)) \end{aligned}$$

that vanishes and the lemma is proved. ◇

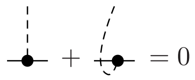
We shall prove the following proposition in Subsect. 3.4 below.

Proposition 3.3. *Let M be a compact one-manifold, then the natural map from $\mathcal{D}_n(M)$ to $\mathcal{A}_n^t(M)$ that maps a chord diagram to its class in $\mathcal{A}_n^t(M)$ induces an isomorphism from $\mathcal{A}_n(M)$ to $\mathcal{A}_n^t(M)$.*

Before, we need a technical lemma that will lead to other fundamental properties of the spaces of diagrams.

3.3 A Useful Trick in Diagram Spaces

We shall first adopt a convention. So far, in a diagram picture, or in a chord diagram picture, the dashed edge of a univalent vertex, has always been attached on the left-hand side of the oriented one-manifold. Now, if k dashed edges are attached on the other side on a diagram picture, then we agree that the corresponding represented element of $\mathcal{A}^t_n(M)$ or $\mathcal{A}_n(M)$ is $(-1)^k$ times the underlying diagram. With this convention, we have the new antisymmetry relation in $\mathcal{A}^t_n(M)$ or in $\mathcal{A}_n(M)$



and the STU relation can be drawn like the IHX relation:

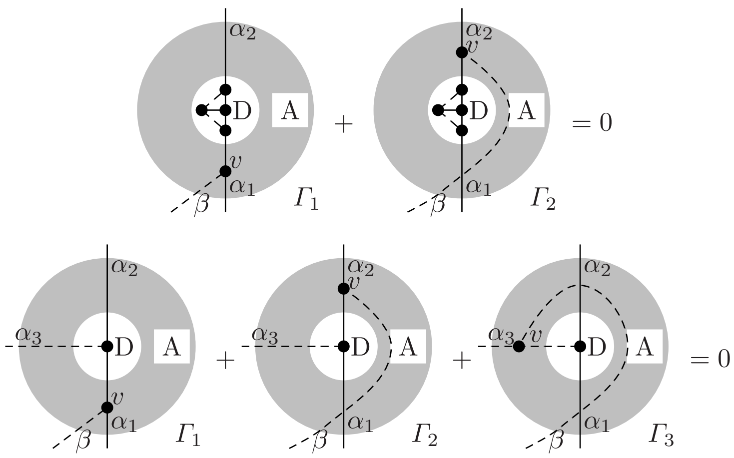
Lemma 3.4. *Let Γ_1 be a diagram (resp. a chord diagram) with support M. Assume that $\Gamma_1 \cup M$ is immersed in the plane so that $\Gamma_1 \cup M$ meets an open annulus A embedded in the plane exactly along $n+1$ embedded arcs α_1, α_2, …, α_n and β, and one vertex v so that:*

1. *The α_i may be dashed or solid, they run from a boundary component of A to the other one,*
2. *β is a dashed arc which runs from the boundary of A to $v \in \alpha_1$,*
3. *The bounded component D of the complement of A does not contain a boundary point of M.*

Let Γ_i be the diagram obtained from Γ_1 by attaching the endpoint v of β to α_i instead of α_1 on the same side, where the side of an arc is its side when going from the outside boundary component of A to the inside one ∂D. Then, we have in $\mathcal{A}^t(M)$, (resp. in $\mathcal{A}(M)$),

$$\sum_{i=1}^{n} \Gamma_i = 0 \; .$$

Examples 3.5.



Proof: The second example shows that STU is equivalent to this relation when the bounded component D of $\mathbf{R}^2 \setminus A$ intersects Γ_1 in the neighborhood of a univalent vertex on M. Similarly, IHX is easily seen as given by this relation when D intersects Γ_1 in the neighborhood of a trivalent vertex. Also note that AS corresponds to the case when D intersects Γ_1 along a dashed or solid arc. Now for the Bar-Natan [4, Lemma 3.1] proof. See also [35, Lemma 3.3]. Assume without loss that v is always attached on the right-hand-side of the α's. Add to the sum the trivial (by IHX and STU) contribution of the sum of the diagrams obtained from Γ_1 by attaching v to each of the three (dashed or solid) half-edges of each vertex w of $\Gamma_1 \cup M$ in D on the left-hand side when the half-edges are oriented towards w. Now, group the terms of the obtained sum by edges of $\Gamma_1 \cup M$ where v is attached, and observe that the sum is zero edge by edge by AS. $\diamond$

Assume that a one-manifold M is decomposed as a union of two one-manifolds $M = M_1 \cup M_2$ whose interiors in M do not intersect. Define the *product associated to this decomposition*:

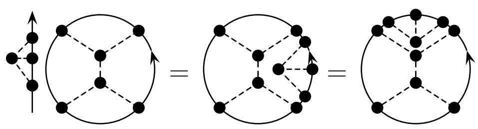
$$\mathcal{A}(M_1) \times \mathcal{A}(M_2) \longrightarrow \mathcal{A}(M)$$

as the continuous bilinear map which maps $([\Gamma_1], [\Gamma_2])$ to $[\Gamma_1 \coprod \Gamma_2]$, if Γ_1 is a diagram with support M_1 and if Γ_2 is a diagram with support M_2, where $\Gamma_1 \coprod \Gamma_2$ denotes their disjoint union. Let $I = [0, 1]$ be the compact oriented interval. If $I = M$, and if we identify I to $M_1 = [0, 1/2]$ and to $M_2 = [1/2, 1]$ with respect to the orientation, the above process turns $\mathcal{A}(I)$ into an algebra where the elements with non-zero degree zero part admit an inverse.

With each choice of a connected component C of M, associate an $\mathcal{A}(I)$-*module structure on* $\mathcal{A}(M)$, that is given by the continuous bilinear map:

$$\mathcal{A}(I) \times \mathcal{A}(M) \longrightarrow \mathcal{A}(M)$$

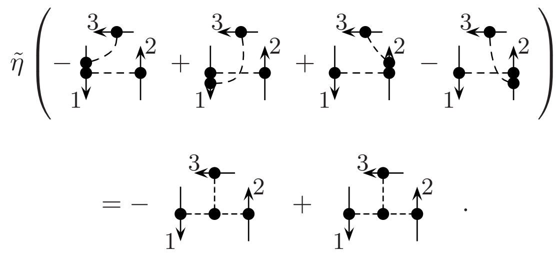
such that: If Γ' is a diagram with support M and if Γ is a diagram with support I, then $([\Gamma], [\Gamma'])$ is mapped to the class of the diagram obtained by inserting Γ along C outside the vertices of Γ, according to the given orientation. For example,



As shown in the first example that illustrates Lemma 3.4, the independence of the choice of the insertion locus is a consequence of Lemma 3.4 where Γ_1 is the disjoint union $\Gamma \coprod \Gamma'$ and intersects D along $\Gamma \cup I$. This also proves that $\mathcal{A}(I)$ is a commutative algebra. Since the morphism from $\mathcal{A}(I)$ to $\mathcal{A}(S^1)$ induced by the identification of the two endpoints of I amounts to quotient out $\mathcal{A}(I)$ by the relation that identifies two diagrams that are obtained from one another by moving the nearest univalent vertex to an endpoint of I near the other endpoint, a similar application of Lemma 3.4 also proves that this morphism is an isomorphism from $\mathcal{A}(I)$ to $\mathcal{A}(S^1)$. (In this application, β comes from the inside boundary of the annulus.) This identification between $\mathcal{A}(I)$ and $\mathcal{A}(S^1)$ will be used several times.

3.4 Proof of Proposition 3.3

Let $\tilde{\eta} : \mathcal{D}_n(M) \to \mathcal{A}_n^t(M)$ denote the above natural map. Let us show that it factors through $4T$. By STU, we have



Thus, $\tilde{\eta}(4T)$ vanishes and we are done. Note that the induced map η is surjective because STU allows one to express any diagram (whose components contain at least one univalent vertex!) as a combination of chord diagrams.

Let us try to construct an inverse $\bar{\iota}$ to the induced map $\eta : \mathcal{A}_n(M) \to \mathcal{A}_n^t(M)$. Let $\mathcal{D}_{n,k}(M)$ denote the subspace of $\mathcal{D}_n^t(M)$ generated by the diagrams on M that have at most k trivalent vertices.

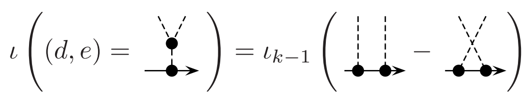
We shall define linear maps ι_k from $\mathcal{D}_{n,k}(M)$ to $\mathcal{A}_n(M)$ by induction on k so that

1. ι_0 is induced by the equality $\mathcal{D}_{n,0}(M) = \mathcal{D}_n(M)$,
2. the restriction of ι_k to $\mathcal{D}_{n,k-1}(M)$ is ι_{k-1}, and,
3. ι_k maps all the relations AS and STU that involve only elements of $\mathcal{D}_{n,k}(M)$ to zero.

It is clear that when we have succeeded in such a task, the linear map from $\mathcal{D}_n(M)$ that maps a diagram d with k trivalent vertices to $\iota_k(d)$ will factor through STU and AS, and that the induced map $\bar{\iota}$ will satisfy $\bar{\iota} \circ \eta = \mathrm{Id}$ and therefore provide the wanted inverse since η is surjective. Now, let us succeed!

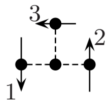
Let $k \geq 1$, assume that ι_{k-1} is defined on $\mathcal{D}_{n,k-1}(M)$ and that ι_{k-1} maps all the relations AS and STU that involve only elements of $\mathcal{D}_{n,k-1}(M)$ to zero. We want to extend ι_{k-1} on $\mathcal{D}_{n,k}(M)$ to a linear map ι_k that maps all the relations AS and STU that involve only elements of $\mathcal{D}_{n,k}(M)$ to zero.

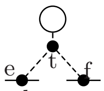
Let d be a diagram with k trivalent vertices, and let e be an edge of d that contains one univalent vertex and one trivalent vertex. Set



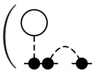
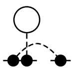
It suffices to prove that $\iota(d,e)$ is independent of our chosen edge e to conclude the proof by defining the linear map ι_k that will obviously satisfy the wanted properties by

$$\iota_k(d) = \iota(d,e)$$

Assume that there are two different edges e and f of d that connect a trivalent vertex to a univalent vertex. We prove that $\iota(d,e) = \iota(d,f)$. If e and f are disjoint, then the fact that ι_{k-1} satifies STU allows us to express both $\iota(d,e)$ and $\iota(d,f)$ as the same combination of 4 diagrams with $(k-2)$ vertices, and we are done. Thus, we assume that e and f are two different edges that share a trivalent vertex t. If there exists another trivalent vertex that is connected to M by an edge g, then $\iota(d,e) = \iota(d,g) = \iota(d,f)$ and we are done. Thus, we furthermore assume that t is the unique trivalent vertex that is connected to M by an edge. So, either t is the unique trivalent vertex, and its component is necessarily like 

and the fact that $\iota(d,e) = \iota(d,f)$ is a consequence of (4T), or the component of t is of the form 

where the circle represents a dashed diagram with only one pictured entry. Thus,

$$\iota(d,e) = \iota_{k-1}\Bigg($$  $$-$$  $$\Bigg)$$

Now, this is zero because the expansion of  as a sum of chord diagrams commutes with any vertex in $\mathcal{A}_n(M)$, according to Lemma 3.4. Similarly, $\iota(d,f) = 0$. Thus, $\iota(d,e) = \iota(d,f)$ in this last case and we are done. ⋄

Notation 3.6. Because of the canonical isomorphism of Proposition 3.3, $\mathcal{A}_n(M)$ will denote both $\mathcal{A}_n(M)$ and $\mathcal{A}^t_n(M)$ from now on.

4 The "Kontsevich, Bott, Taubes, Bar-Natan, Altschuler, Freidel, D. Thurston" Universal Link Invariant *Z*

4.1 Introduction to Configuration Space Integrals: The Gauss Integrals

In 1833, Carl Friedrich Gauss defined the first example of a *configuration space integral* for an oriented two-component link. Let us formulate his definition in a modern language. Consider a smooth (C^∞) embedding

$$L : S^1_1 \sqcup S^1_2 \hookrightarrow \mathbf{R}^3$$

of the disjoint union of two circles $S^1 = \{z \in \mathbf{C} \text{ s.t. } |z| = 1\}$ into $\mathbf{R}^3$. With an element (z_1, z_2) of $S^1_1 \times S^1_2$ that will be called a *configuration*, we associate the oriented direction

$$\Psi((z_1, z_2)) = \frac{1}{\parallel \overrightarrow{L(z_1)L(z_2)} \parallel} \overrightarrow{L(z_1)L(z_2)} \in S^2$$

of the vector $\overrightarrow{L(z_1)L(z_2)}$. Thus, we have associated a map

$$\Psi : S^1_1 \times S^1_2 \longrightarrow S^2$$

from a compact oriented 2-manifold to another one with our embedding. This map has an integral degree $\deg(\Psi)$ that can be defined in several equivalent ways. For example, it is the *differential degree* $\deg(\Psi, y)$ of any regular value y of Ψ, that is the sum of the ± 1 signs of the Jacobians of Ψ at the points of the preimage of y [28, §5]. Thus, $\deg(\Psi)$ can easily be computed from a regular diagram of our two-component link as the differential degree of a unit vector $\overrightarrow{v}$ pointing to the reader or as the differential degree of $(-\overrightarrow{v})$.

$$\deg(\Psi) = \deg(\Psi, \overrightarrow{v}) = \sharp \nearrow\!\!\!\!\nwarrow_{2\,1} - \sharp \nwarrow\!\!\!\!\nearrow_{1\,2} = \deg(\Psi, -\overrightarrow{v}) = \sharp \nearrow\!\!\!\!\nwarrow_{1\,2} - \sharp \nwarrow\!\!\!\!\nearrow_{2\,1} .$$

It can also be defined as the following *configuration space integral*

$$\deg(\Psi) = \int_{S^1 \times S^1} \Psi^*(\omega)$$

where ω is the homogeneous volume form on S^2 such that $\int_{S^2} \omega = 1$. Of course, this integral degree is an isotopy invariant of L, and the reader has recognized is nothing but the *linking number* of the two components of L.

We can again follow Gauss and associate the following similar *Gauss integral* $I(K;\theta)$ to a C^∞ embedding $K : S^1 \hookrightarrow \mathbf{R}^3$. (The meaning of θ will be specified later.) Here, we consider the *configuration space* $C(K;\theta) = S^1 \times]0, 2\pi[$, and the map

$$\Psi : C(K;\theta) \longrightarrow S^2$$

that maps (z_1, η) to the oriented direction of $\overrightarrow{K(z_1)K(z_1 e^{i\eta})}$, and we set

$$I(K;\theta) = \int_{C(K;\theta)} \Psi^*(\omega) \ .$$

Let us compute $I(K;\theta)$ in some cases. First notice that Ψ may be extended to the closed annulus

$$\overline{C}(K;\theta) = S^1 \times [0, 2\pi]$$

by the tangent map K' of K along $S^1 \times \{0\}$ and by $(-K')$ along $S^1 \times \{2\pi\}$. Then by definition, $I(K;\theta)$ is the algebraic area (the integral of the differential degree with respect to the measure associated with ω) of the image of the annulus in S^2. Now, assume that K is contained in a horizontal plane except in a neighborhood of crossings where it entirely lies in vertical planes. Such a knot embedding will be called *almost horizontal*. In that case, the image of the annulus boundary has the shape of the following bold line in S^2.

In particular, for each hemisphere, the differential degree of a regular value of Ψ does not depend on the choice of the regular value in the hemisphere. Assume that the orthogonal projection onto the horizontal plane is regular. Then $I(K;\theta)$ is the average of the differential degrees of the North Pole and the South Pole, and it can be computed from the horizontal projection as the writhe of the projection

$$I(K;\theta) = \sharp$$  $$- \sharp$$  .

This number can be changed without changing the isotopy class of the knot by local modifications where —— becomes  or . In particular, $I(K;\theta)$ can reach any integral value on a given isotopy class of knots, and since it varies continuously on such a class, it can reach any real value on any given isotopy class of knots. Thus, this Gauss integral is NOT an isotopy invariant.

However, we can follow Guadagnini, Martellini, Mintchev [14] and Bar-Natan [5] and associate configuration space integrals to any embedding L of an oriented one-manifold M and to any uni-trivalent diagram Γ without simple loop like  on M.

4.2 The Chern–Simons Series

Let M be an oriented one-manifold and let

$$L : M \longrightarrow \mathbf{R}^3$$

denote a C^∞ embedding from M to $\mathbf{R}^3$. Let Γ be a Jacobi diagram on M. Let $U = U(\Gamma)$ denote the set of univalent vertices of Γ, and let $T = T(\Gamma)$ denote the set of trivalent vertices of Γ. A *configuration* of Γ is an embedding

$$c : U \cup T \hookrightarrow \mathbf{R}^3$$

whose restriction $c_{|U}$ to U may be written as $L \circ j$ for some injection

$$j : U \hookrightarrow M$$

in the given isotopy class $[i]$ of embeddings of U into the interior of M. Denote the set of these configurations by $C(L;\Gamma)$,

$$C(L;\Gamma) = \left\{c : U \cup T \hookrightarrow \mathbf{R}^3 \, ; \exists j \in [i], c_{|U} = L \circ j\right\} \ .$$

In $C(L;\Gamma)$, the univalent vertices move along $L(M)$ while the trivalent vertices move in the ambient space, and $C(L;\Gamma)$ is naturally an open submanifold of $M^U \times (\mathbf{R}^3)^T$.

Denote the set of (dashed) edges of Γ by $E = E(\Gamma)$, and fix an orientation for these edges. Define the map $\Psi : C(L;\Gamma) \longrightarrow \left(S^2\right)^E$ whose projection to the S^2 factor indexed by an edge from a vertex v_1 to a vertex v_2 is the direction of $\overrightarrow{c(v_1)c(v_2)}$. This map Ψ is again a map between two orientable manifolds that have the same dimension, namely the number of dashed half-edges of Γ, and we can write the *configuration space integral:*

$$I(L;\Gamma) = \int_{C(L;\Gamma)} \Psi^* \left(\bigwedge^E \omega\right) \ .$$

Bott and Taubes have proved that this integral is convergent [8]. See also Subsects. 5.1, 5.2 below. Thus, this integral is well-defined up to sign. In fact, the orientation of the trivalent vertices of Γ provides $I(L;\Gamma)$ with a well-defined sign. Indeed, since S^2 is equipped with its standard orientation, it is enough to orient $C(L;\Gamma) \subset M^U \times (\mathbf{R}^3)^T$ in order to define this sign. This will be done by providing the set of the natural coordinates of $M^U \times (\mathbf{R}^3)^T$ with some order up to an even permutation. This set is in one-to-one correspondence with the set of (dashed) half-edges of Γ, and the vertex-orientation of the trivalent vertices provides a natural preferred such one-to-one correspondence up to some (even!) cyclic permutations of three half-edges meeting at a trivalent vertex. Fix an order on E, then the set of half-edges becomes ordered by (origin of the first edge, endpoint of the first edge, origin of the second edge, ..., endpoint of the last edge), and this order orients $C(L;\Gamma)$. The property of this sign is that the product $I(L;\Gamma)[\Gamma] \in \mathcal{A}(M)$ depends neither on our various choices nor on the vertex orientation of Γ. Check it as an exercise!

Now, the *perturbative series expansion of the Chern–Simons theory for one-manifold embeddings in* $\mathbf{R}^3$ is the following sum running over all the Jacobi diagrams Γ without vertex orientation[2]:

$$Z_{\mathrm{CS}}(L) = \sum_{\Gamma} \frac{I(L;\Gamma)}{\sharp \mathrm{Aut}\Gamma}[\Gamma] \quad \in \mathcal{A}(M)$$

[2] This sum runs over equivalence classes of Jacobi diagrams, where two diagrams are equivalent if and only if they coincide except possibly for their vertex orientation.

where $\sharp\mathrm{Aut}\Gamma$ is the number of automorphisms of Γ as a uni-trivalent graph with a given isotopy class of injections of U into M, but without vertex-orientation for the trivalent vertices.

More precisely, let Γ be a diagram on M, let $V(\Gamma)$ be the set of its vertices, let $U(\Gamma)$ be the set of its univalent vertices, and let $E(\Gamma)$ be the set of its edges. Define the set $H(\Gamma)$ of its half-edges as

$$H(\Gamma) = \{h = (v(h), e(h)) \in V(\Gamma) \times E(\Gamma); v(h) \in e(h)\} .$$

Let i denote an injection of $U(\Gamma)$ into M in the isotopy class that defines Γ. An *automorphism* of Γ is a permutation σ of $H(\Gamma)$ such that $v(h) = v(h') \Longrightarrow v(\sigma(h)) = v(\sigma(h'))$, $e(h) = e(h') \Longrightarrow e(\sigma(h)) = e(\sigma(h'))$, and i is isotopic to $i \circ \overline{\sigma}$ where $\overline{\sigma}$ denotes the permutation of $U(\Gamma)$ induced by σ.

Remark 4.1. Any configuration $c : U \cup T \hookrightarrow \mathbf{R}^3$ uniquely extends to a map $i(c) : \Gamma \longrightarrow \mathbf{R}^3$ that is linear along the edges of Γ. The configurations c such that two edges of Γ have colinear images under $i(c)$ do not contribute to the integral, because their images under Ψ lie in a codimension 2 subspace of $\left(S^2\right)^E$. (In particular, if two vertices of Γ are related by several edges, then $I(L;\Gamma) = 0$.) If $i(c)$ maps all the edges of Γ to pairwise non colinear segments of $\mathbf{R}^3$, there are exactly $\sharp Aut(\Gamma)$ configurations d of Γ with respect to $L(M)$ such that $i(c)(\Gamma) = i(d)(\Gamma)$. In other words, with the factor $\frac{1}{\sharp\mathrm{Aut}\Gamma}$, the image of an immersed unitrivalent graph contributes exactly once to the expression of $Z_{\mathrm{CS}}(L)$.

Let θ denote the Jacobi diagram

$\theta =$ 

on S^1. When L is a knot K, the degree one part of $Z_{CS}(K)$ is $\frac{I(K;\theta)}{2}[\theta]$ and therefore Z_{CS} is not invariant under isotopy. However, the evaluation Z^0_{CS} at representatives of knots with null Gauss integral is an isotopy invariant that is a universal Vassiliev invariant of knots. (All the real-valued finite type knot invariants factor through it.) This is the content of the following theorem, due independently to Altschuler and Freidel [1], and to D. Thurston [34], after the work of many people including Guadagnini, Martellini and Mintchev [14], Bar-Natan [5], Axelrod and Singer [2, 3], Kontsevich [18, 19], Bott and Taubes [8]...

Theorem 4.2 (Altschuler–Freidel [1], D. Thurston [34], 1995). *If $L = K_1 \cup \dots \cup K_k$ is a link, then $Z_{CS}(L)$ only depends on the isotopy class of L and on the Gauss integrals $I(K_i;\theta)$ of its components. In particular, the evaluation*

$$Z^0_{CS}(L) \in \prod_{n\in\mathbf{N}} \mathcal{A}_n(\sqcup_{i=1}^k S^1_i)$$

at representatives of L whose components have zero Gauss integrals is an isotopy invariant of L. Furthermore, Z^0_{CS} is a universal Vassiliev invariant *of links in the following sense. When L is a singular link with n double points, the degree k part $Z^0_{CS,k}(L)$ vanishes for $k < n$, while $Z^0_{CS,n}(L)$ is nothing but the chord diagram of L.*

This theorem implies the fundamental theorem 2.26 of Vasiliev invariants. The main ideas involved in its proof are sketched in the next section.

5 More on Configuration Spaces

In this section, we describe the main ideas involved in the proof of Theorem 4.2, and we sketch the proof of this theorem. First, we need to understand the compactifications of configuration spaces of [11]. We shall present them with the Poirier point of view [30].

5.1 Compactifications of Configuration Spaces

Definition 5.1. A homothety with ratio in $]0, +\infty[$ will be called a *dilation.*

Let $X = \{\xi_1, \xi_2, \dots, \xi_p\}$ be a finite set of cardinality $p \geq 2$, let k denote a positive integer. Let $C_0(X; \mathbf{R}^k)$ (resp. $C(X; \mathbf{R}^k)$) denote the set of injections (resp. of non-constant maps) f from X to $\mathbf{R}^k$, up to translations and dilations. $C_0(X; \mathbf{R}^k)$ is the quotient of

$$\{(x_1 = f(\xi_1), x_2 = f(\xi_2), \dots, x_p = f(\xi_p)) \in (\mathbf{R}^k)^p; x_i \neq x_j \text{ if } i \neq j\}$$

by the translations which identify $(x_1, x_2, \dots, x_p)$ to $(x_1 + T, x_2 + T, \dots, x_p + T)$ for all $T \in \mathbf{R}^k$ and by the dilations which identify $(x_1, x_2, \dots, x_p)$ to $(\lambda x_1, \lambda x_2, \dots, \lambda x_p)$ for all $\lambda > 0$.

Examples 5.2. 1. For example, $C_0(X; \mathbf{R})$ has $p!$ connected components corresponding to the possible orders of the set X. Each of its components can be identified with the interior $\{(x_2, x_3, \dots, x_{p-1}) \in \mathbf{R}^{p-2}; 0 < x_2 < x_3 < \dots < x_{p-1} < 1\}$ of a $(p-2)$ simplex.
2. As another example, $C(\{1,2\}; \mathbf{R}^k) = C_0(\{1,2\}; \mathbf{R}^k)$ is homeomorphic to the sphere S^{k-1}.

In general, the choice of a point $\xi \in X$ provides a homeomorphism

$$\begin{array}{rccl} \phi_\xi : & C(X, \mathbf{R}^k) & \longrightarrow & S^{kp-k-1} \\ & f & \mapsto & \left(x \mapsto \frac{f(x)-f(\xi)}{\|\sum_{i=1}^{p}(f(\xi_i)-f(\xi))\|}\right) \end{array}$$

where S^{kp-k-1} is the unit sphere of $\left(\mathbf{R}^k\right)^{p-1}$. These homeomorphisms equip $C(X, \mathbf{R}^k)$ with an analytic (C^ω) structure and make clear that $C(X, \mathbf{R}^k)$ is compact. There is a natural embedding

$$\begin{array}{rcl} i: C_0(X;\mathbf{R}^k) & \hookrightarrow & \prod_{A\subseteq X;\sharp A\geq 2} C(A;\mathbf{R}^k) \\ c_X & \mapsto & (c_{X|A})_{A\subseteq X;\sharp A\geq 2} \end{array}$$

where $c_{X|A}$ denotes the restriction of c_X to A. Define the compactification $C(X;k)$ of $C_0(X;\mathbf{R}^k)$ as

$$C(X;k)=\overline{i(C_0(X;\mathbf{R}^k))}\subseteq \prod_{A\subseteq X;\sharp A\geq 2} C(A;\mathbf{R}^k)$$

In words, in $C(X;k)$, some points of X are allowed to collide with each other, or to become infinitely closer to each other than they are to other points, but the compactification provides us with the magnifying glasses $C(A;\mathbf{R}^k)$ that allow us to see the infinitely small configurations at the scales of the collisions.

Observe that the elements $(c_A)_{A\subseteq X;\sharp A\geq 2}$ of $C(X;k)$ satisfy the following condition $(\star)$.

$(\star)$: If $B\subset A$, then the restriction $c_{A|B}$ of c_A to B is either constant or equal to c_B.

Indeed, the above condition holds for elements of $i(C_0(X;\mathbf{R}^k))$, and it can be rewritten as the following condition that is obviously closed. *For any two sets A and B such that $B\subset A$, if $x\in B$, the two vectors of $(\mathbf{R}^k)^{B\setminus\{x\}}$, $(c_{A|B}(y)-c_{A|B}(x))_{y\in B\setminus\{x\}}$ and $(c_B(y)-c_B(x))_{y\in B\setminus\{x\}}$, are colinear, and their scalar product is non negative.*

Proposition 5.3. *The set $C(X;k)$ has a natural structure of an analytic manifold with corners*[3] *and*

$$C(X;k)=\left\{(c_A)\in \prod_{A\subseteq X;\sharp A\geq 2} C\left(A;\mathbf{R}^k\right);\ (c_A)\ \textit{satisfies}\ (\star)\right\}.$$

PROOF: Set

$$\tilde{C}(X;k)=\left\{(c_A)\in \prod_{A\subseteq X;\sharp A\geq 2} C\left(A;\mathbf{R}^k\right);\ (c_A)\ \text{satisfies}\ (\star)\right\}.$$

We have already proved that $C(X;k)\subseteq \tilde{C}(X;k)$. In order to prove the reversed inclusion, we first study the structure of $\tilde{C}(X;k)$. Let

$$c^0=(c^0_A)_{A\subseteq X;\sharp A\geq 2}\in \tilde{C}(X;k)\ .$$

Given this point c^0, we construct the rooted tree $\tau(c^0)$, with oriented edges, whose vertices are some subsets of X with cardinality greater than 1,

[3] Every point c of $C(X;k)$ has a neighborhood diffeomorphic to $[0,\infty[^r\times\mathbf{R}^{n-r}$, and the transition maps are analytic.

in the following way. The root-vertex is X. The edges starting at a vertex $A \subseteq X$ are in one-to-one correspondence with the maximal subsets B of A with $\sharp B \geq 2$, such that $c^0_{A|B}$ is constant. The edge corresponding to a subset B goes from A to B. Note that the tree structure can be recovered from the set of vertices. Therefore, we identify the tree $\tau^0 = \tau(c^0)$ with its set of vertices and τ^0 is a set of subsets of X. For any strict subset A of X, $\hat{A}$ denotes the smallest element of τ^0 that strictly contains A.

Now, we construct a chart of $\tilde{C}(X;k)$ near c^0.

Let $B \in \tau^0$. Let $C_{\tau^0}(B;\mathbf{R}^k)$ be the subspace of $C(B;\mathbf{R}^k)$ made of maps from B to $\mathbf{R}^k$ such that two elements of B have the same image in $\mathbf{R}^k$ if and only if they belong to a common endpoint (subset of X) of an edge starting at B. (Note that $C_{\tau^0}(B;\mathbf{R}^k)$ is naturally homeomorphic to $C_0(B_{\tau^0};\mathbf{R}^k)$ where B_{τ^0} is the set obtained from B by identifying two elements of B that belong to a common strict subset of B in τ^0.) Let $V \subseteq \prod_{B\in\tau^0} C_{\tau^0}(B;\mathbf{R}^k)$ be an open neighborhood of $(c^0_B)_{B\in\tau^0}$ in $\prod_{B\in\tau^0} C_{\tau^0}(B;\mathbf{R}^k)$. Let $\varepsilon > 0$. When V and ε are small enough, define the map

$$\begin{aligned} F : [0,\varepsilon[^{\tau^0\setminus\{X\}}\times V &\longrightarrow \prod_{D\subseteq X;\sharp D\geq 2} C(D;\mathbf{R}^k) \\ P = \left((\lambda_B)_{B\in\tau^0\setminus\{X\}},(c_B)_{B\in\tau^0}\right) &\mapsto (F(P)_D)_{D\subseteq X;\sharp D\geq 2} \end{aligned}$$

where $F(P)_D$ will be equal to $F(P)_{\hat{D}|D}$ if $D \notin \tau^0$, and if $A \in \tau_0$, $F(P)_A$ is represented by the map

$$\tilde{F}(P)_A : A \longrightarrow \mathbf{R}^k$$

that maps an element $(x \in A)$ to the vector that admits the following recursive definition. Let

$$\{\hat{x}\} = B_1 \subset B_2 \subset \cdots \subset B_m \subset B_{m+1} = A$$

be the sequence of vertices of τ^0 such that $B_1 = \{\hat{x}\}$ and $B_{r+1} = \hat{B}_r$. Fix a point $\xi(B)$ in any subset $B \in \tau^0$ so that if $B' \in \tau^0$ and if $\xi(B) \in B' \subset B$, then $\xi(B') = \xi(B)$ (the $\xi(B)$ depend on τ^0 that is fixed). Then $\tilde{F}(P)_{B_1}(x) = \phi_{\xi(B_1)}(c_{B_1})(x)$ and

$$\tilde{F}(P)_{B_{k+1}}(x) = \phi_{\xi(B_{k+1})}(c_{B_{k+1}})(x) + \lambda_{B_k}\tilde{F}(P)_{B_k}(x)\ .$$

In particular, the small parameter $\lambda_B \in [0,\varepsilon[$ is the ratio of the scale of the representative $\tilde{F}(P)_{\hat{B}|B}$ in $\hat{B}$ by the scale of $\tilde{F}(P)_B$ in B.

Let $O(\tau^0)$ be the following open subset of $\prod_{A\subseteq X;\sharp A\geq 2} C(A;\mathbf{R}^k)$,

$$O(\tau^0) = \{(f_A)_{A\subseteq X;\sharp A\geq 2}; \forall A, \forall x,y \in A, c^0_A(x) \neq c^0_A(y) \Rightarrow f_A(x) \neq f_A(y)\}\ .$$

Since $F\left(P_0 = ((0)_{B\in\tau^0\setminus\{X\}},(c^0_B)_{B\in\tau^0})\right) = c^0 \in O(\tau^0)$, when V and ε are small enough, the image of F is in $O(\tau^0)$. In particular, $F(P)_D$ is never constant and F is well-defined. Furthermore, the image of F is in $\tilde{C}(X;k)$. Assume the following lemma that will be proved later.

Lemma 5.4. *There is an open neighborhood O of c^0 in $O(\tau^0)$, and an analytic map G from O to $\mathbf{R}^{\tau^0\setminus\{X\}} \times V$, such that the restriction of G to $O \cap \tilde{C}(X;k)$ is an inverse for $F : [0,\varepsilon[^{\tau^0\setminus\{X\}} \times V \longrightarrow O \cap \tilde{C}(X;k)$.*

Thus, F is a homeomorphism onto its image that is an open subset of $\tilde{C}(X;k)$. Also notice that the tree (or its vertices set) corresponding to a point in the image of $F((\lambda_B)_{B\in\tau^0\setminus\{X\}},(c_B)_{B\in\tau^0})$ is obtained from τ^0 by removing the subsets B such that $\lambda_B > 0$. The points of $i\left(C_0(X;\mathbf{R}^k)\right)$ are the points whose tree is reduced to its root X. In particular, the point c^0 we started with is in the closure of $F(]0,\varepsilon[^{B\in\tau^0\setminus\{X\}} \times V) \subseteq i(C_0(X;\mathbf{R}^k))$. This finishes proving that $C(X;k) \subseteq \tilde{C}(X;k)$. Furthermore, since F and G are analytic, the above local homeomorphisms equip $C(X;k)$ with the structure of a C^ω manifold with corners. This ends the proof of Proposition 5.3 assuming Lemma 5.4.$\diamond$

Proof of Lemma 5.4: Let $f = (f_A)_{A\subseteq X;\sharp A\geq 2} \in O(\tau^0)$. Define $G(f) = \left((\mu_B(f))_{B\in\tau^0\setminus\{X\}},(d_B(f))_{B\in\tau^0}\right) \in \mathbf{R}^{\tau^0\setminus\{X\}} \times \prod_{B\in\tau^0} C_{\tau^0}(B;\mathbf{R}^k)$ as follows. Assume that any f_B is represented by $f_B = \phi_{\xi(B)}(f_B)$. Define $\tilde{d}_B(f)$ by

$$\tilde{d}_B(f)(x) = \begin{cases} f_B(x) & \text{if } \widehat{\{x\}} = B \\ f_B(\xi(B')) & \text{if } x \in B' \text{ and } \hat{B}' = B \end{cases}$$

and $d_B(f) = \phi_{\xi(B)}(\tilde{d}_B(f))$. For $B \in \tau^0$, choose $\xi'(B) \neq \xi(B) \in B$ such that $c^0(\xi'(B)) \neq c^0(\xi(B))$, and either $\widehat{\{\xi'(B)\}} = B$ or there exists a B' such that $\hat{B}' = B$ and $\xi'(B) = \xi(B')$. Note that when $f_B = \phi_{\xi(B)}(\tilde{F}(P)_B)$, then

$$d_B = d_B(f) = \phi_{\xi(B)}(c_B)\ ,$$

$$f_B = \frac{\parallel f_B(\xi'(B)) \parallel}{\parallel \tilde{F}(P)_B(\xi'(B)) \parallel}\tilde{F}(P)_B = \frac{\parallel f_B(\xi'(B)) \parallel}{\parallel d_B(\xi'(B)) \parallel}\tilde{F}(P)_B\ ,$$

$$\tilde{F}(P)_{\hat{B}}(\xi'(B)) - \tilde{F}(P)_{\hat{B}}(\xi(B)) = \lambda_B \tilde{F}(P)_B(\xi'(B)) = \lambda_B d_B(\xi'(B))\ ,$$

$$\begin{aligned} \lambda_B &= \tfrac{<\tilde{F}(P)_{\hat{B}}(\xi'(B)) - \tilde{F}(P)_{\hat{B}}(\xi(B)), d_B(\xi'(B))>}{\|d_B(\xi'(B))\|^2} \\ &= \tfrac{\|d_{\hat{B}}(\xi'(\hat{B}))\|}{\|f_{\hat{B}}(\xi'(\hat{B}))\|}\tfrac{<f_{\hat{B}}(\xi'(B)) - f_{\hat{B}}(\xi(B)), d_B(\xi'(B))>}{\|d_B(\xi'(B))\|^2}\ . \end{aligned}$$

Therefore, we define

$$\mu_B(f) = \frac{\parallel d_{\hat{B}}(\xi'(\hat{B})) \parallel}{\parallel f_{\hat{B}}(\xi'(\hat{B})) \parallel}\frac{< f_{\hat{B}}(\xi'(B)) - f_{\hat{B}}(\xi(B)), d_B(\xi'(B)) >}{\parallel d_B(\xi'(B)) \parallel^2}\ .$$

Now, it is clear that G is analytic, and that $G \circ F = \text{Identity}$ on

$$O = G^{-1}(]-\varepsilon,\varepsilon[^{\tau^0\setminus\{X\}} \times V)\ .$$

Then it is enough to see that the restriction of $F \circ G$ to $O \cap \tilde{C}(X;k)$ is well-defined ($\forall B, \mu_B \geq 0$), and is the identity. This is left as an exercise for the reader. $\diamond$

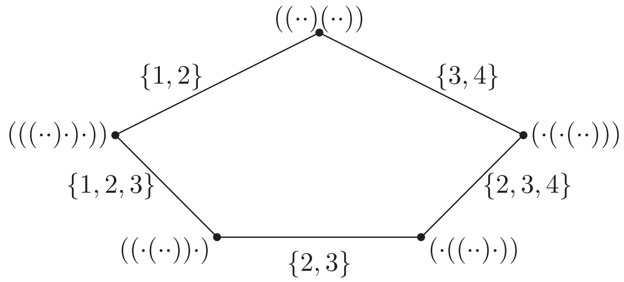
The space $C(X;k)$ is also equipped with a partition by the associated trees of the above proposition proof. Note that the part $F(\tau)$ corresponding to a given tree τ is a submanifold of dimension $\left(\dim(C_0(X;\mathbf{R}^k)) - (\sharp\tau - 1)\right)$ that is homeomorphic to $\prod_{B\in\tau} C_\tau(B;\mathbf{R}^k)$. In particular, the boundary of $C(X;k)$ has a partition into open *faces,* corresponding to trees τ with $\sharp\tau > 1$, of codimension $(\sharp\tau - 1)$.

Example 5.5. Let us again consider the case where $k = 1$. Let us fix one order on X and let us study the corresponding component $C_<(X;1)$ of $C(X;1)$. Here the trees that provide non-empty faces are those which are made of subsets containing consecutive elements, and all the faces are homeomorphic to open balls. When $X = \{1,2,3\}$, $C_<(X;1)$ is an interval whose endpoints correspond to the two trees $\{X, \{1,2\}\}$ and $\{X, \{2,3\}\}$. For $X = \{1,2,\dots,p\}$, we find a *Stasheff polyhedron* that is a polyhedron whose maximal codimension faces are points that can be described as *non-associative words* as in the following definition.

Definition 5.6. A *non-associative word* or *n.a. word* w in the letter $\cdot$ is an element of the free non-associative monoid generated by $\cdot$. The *length* of such a w is the number of letters of w. Equivalently, we can define a *non-associative word* by saying that each such word has an integral *length* $\ell(w) \in \mathbf{N}$, the only word of length 0 is the *empty word*, the only word of length 1 is $\cdot$, the product $w'w''$ of two n.a. words w' and w'' is a n.a. word of length $(\ell(w') + \ell(w''))$, and every word w of length $\ell(w) \geq 2$ can be decomposed in a unique way as the product $w'w''$ of two n.a. words w' and w'' of nonzero length.

Example 5.7. The unique n.a. word of length 2 is $(\cdot\cdot)$. The two n.a. words of length 3 are $((\cdot\cdot)\cdot)$ and $(\cdot(\cdot\cdot))$. There are five n.a. words of length 4 drawn in the following picture of $C_<(\{1,2,3,4\};1)$.
A n.a. word corresponds to the binary tree of subsets of points between matching parentheses.

In particular, $C_<(\{1,2,3,4\};1)$ is the following well-known pentagon, whose edges are labeled by the element of the corresponding $\tau \setminus \{\{1,2,3,4\}\}$.



As another example, the reader can recognize that $C(\{1,2,3\};2)$ is diffeomorphic to the product by S^1 of the complement of two disjoint open discs in the unit two-dimensional disk (or to the exterior of a 3-component Hopf link in S^3, made of three Hopf fibers).

5.2 Back to Configuration Space Integrals for Links

We can use these compactifications in order to study the configuration space integrals defined in Sect. 4.2 as in [30]. Indeed, there is a natural embedding

$$i : C(L;\Gamma) \hookrightarrow M^U \times (S^3 = \mathbf{R}^3 \cup \infty)^T \times C(U \cup T;3) \;.$$

Define the compactification $\overline{C}(L;\Gamma)$ of $C(L;\Gamma)$ as the closure of $i(C(L;\Gamma))$ in this compact space. As before, the compactification can be provided with a structure of a C^∞ manifold with corners, with a stratification that will again be given by trees recording the different relative collapses of points.[4] Furthermore, since Ψ is defined on $C(U \cup T;3)$ as the projection on $\prod_E C(E;3)$ where an edge E is seen as the pair of its endpoints ordered by the orientation, Ψ extends to $\overline{C}(L;\Gamma)$. This extension is smooth, and we have

$$I(L;\Gamma) = \int_{\overline{C}(L;\Gamma)} \Psi^*\left(\bigwedge^E \omega\right) \;.$$

In particular, this shows the convergence of the integrals of Sect. 4.2. The variation of $I(L;\Gamma)$ under a C^∞ isotopy

$$\begin{array}{rcl} L : M \times I & \longrightarrow & \mathbf{R}^3 \\ (m,t) & \mapsto & L^t(M) \end{array}$$

is computed with the help of the Stokes theorem. Since $\Psi^*(\bigwedge^E \omega)$ is a closed form defined on $\cup_{t\in I} C(L^t;\Gamma)$, the variation $(I(L_1;\Gamma) - I(L_0;\Gamma))$ is given by the sum over the codimension one faces $F(\tau)(L;\Gamma)$ of the

$$V(F(\tau)(L;\Gamma)) = \int_{\cup_{t\in I} F(\tau)(L^t;\Gamma)} \Psi^*\left(\bigwedge^E \omega\right) \;.$$

The Altschuler-Freidel proof and the Thurston proof that Z^0_{CS} provides a link invariant now rely on a careful analysis of the codimension one faces, and of the variations that they induce. See [1, 34, 30]. This analysis was successfully started by Bott and Taubes [8]. It shows that the faces that indeed contribute in the link case, where $M = \coprod_{i=1}^k S^1_i$, are of four possible forms.

[4] There are two main differences with the already studied case, due to the one-manifold embedding L. First, the univalent vertices vary along L, and when they approach each other, their direction that makes sense in the compactification approaches the direction of the tangent vector to L at the point where they meet. Second, there is a preferred observation scale namely the scale of the ambient space where the embedding lies.

1. Two trivalent vertices joined by an edge collide with each other.
2. Two univalent vertices consecutive on M collide with each other.
3. A univalent vertex and a trivalent vertex that are joined by an edge collide with each other.
4. The *anomalous faces* where some connected component of the dashed graph Γ collapses at one point.

The STU and IHX relation make the first three kinds of variations cancel. Let us see roughly how it works for the first kind of faces. Such a face is homeomorphic to the product of the sphere S^2 by the configuration space of the graph obtained from Γ by identifying the two colliding points (which become a four-valent vertex), where S^2 is the configuration space of the two endpoints of the infinitely small edge. Let Γ_1, Γ_2 and Γ_3 be three graphs related by an IHX relation so that $[\Gamma_1]+[\Gamma_2]+[\Gamma_3]=0$. Let τ_i be the tree made of $U \cup T$ and the visible edge of Γ_i. Then $V(F(\tau_i)(L;\Gamma_i))$ is independent of i. Therefore,

$$V(F(\tau_1)(L;\Gamma_1))[\Gamma_1]+V(F(\tau_2)(L;\Gamma_2))[\Gamma_2]+V(F(\tau_3)(L;\Gamma_3))[\Gamma_3]=0\ .$$

Thus, the sum[5] of these variations plugged into the Chern-Simons series is zero. The STU relation makes the variations of the second kind and the third kind of faces cancel each other in a similar way.

For the anomalous faces, we do not have such a cancellation. But, we are about to see that we have a formula like

$$\frac{\partial}{\partial t}\left(Z_{CS}(L^t)\right)=\left(\sum_{i=1}^{k}\frac{\partial}{\partial t}\frac{I(K_i^t;\theta)}{2}\alpha\sharp_i\right)Z_{CS}(L^t)\ . \qquad (1)$$

where α is the *anomaly* that is the constant of $\mathcal{A}(\mathbf{R})$ that is defined below, and $\sharp_i$ denotes the $\mathcal{A}(\mathbf{R})$-module structure on $\mathcal{A}(\coprod_{i=1}^k S_i^1)$ by insertion on the i^{th} component. See Subsect. 3.3.

5.3 The Anomaly

Let us define the anomaly. Let $v \in S^2$. Let D_v denote the linear map

$$\begin{array}{rcl} D_v : \mathbf{R} & \longrightarrow & \mathbf{R}^3 \\ 1 & \mapsto & v\ . \end{array}$$

Let Γ be a Jacobi-diagram on $\mathbf{R}$. Define $C(D_v;\Gamma)$ and Ψ as in Sect. 4.2. Let $\hat{C}(D_v;\Gamma)$ be the quotient of $C(D_v;\Gamma)$ by the translations parallel to D_v and by the dilations. Then Ψ factors through $\hat{C}(D_v;\Gamma)$ that has two dimensions less. Now, allow v to run through S^2 and define $\hat{C}(\Gamma)$ as the total space of

[5] To make this sketchy proof work and to avoid thinking of the $(1/\sharp \mathrm{Aut}\Gamma)$ factor, use Remark 4.1.

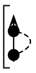
the fibration over S^2 where the fiber over v is $\hat{C}(D_v;\Gamma)$. The map Ψ becomes a map between two smooth oriented manifolds of the same dimension. Indeed, $\hat{C}(\Gamma)$ carries a natural smooth structure and can be oriented as follows. Orient $C(D_v;\Gamma)$ as before, orient $\hat{C}(D_v;\Gamma)$ so that $C(D_v;\Gamma)$ is locally homeomorphic to the oriented product (translation vector $(0,0,z)$ of the oriented line, ratio of homothety $\lambda \in]0,\infty[$) $\times \hat{C}(D_v;\Gamma)$ and orient $\hat{C}(\Gamma)$ with the (base$(=S^2)\oplus$fiber) convention[6]. Then we can again define

$$I(\Gamma) = \int_{\hat{C}(\Gamma)} \Psi^*\left(\bigwedge^E \omega\right) .$$

Now, the *anomaly* is the following sum running over all connected Jacobi diagrams Γ on the oriented line (again without vertex-orientation and without small loop):

$$\alpha = \sum \frac{I(\Gamma)}{\sharp \mathrm{Aut}\Gamma}[\Gamma] \quad \in \mathcal{A}(\mathbf{R})$$

Its degree one part is

$\alpha_1 = \Bigg[$  $\Bigg]$.

Then Formula 1 expresses the following facts. Let Γ be a connected dashed graph on the circle S^1_j. The set $U = \{u_1, u_2, \dots, u_k\}$ of its univalent vertices is cyclically ordered, and the anomalous faces for Γ correspond to the different total orders (which are visible at the scale of the collision) inducing the given cyclic order. Assume that $u_1 < u_2 < \dots < u_k = u_0$ is one of them. Denote the Jacobi diagram on $\mathbf{R}$ obtained by cutting the circle between u_{i-1} and u_i by Γ_i.

The anomalous face where Γ collapses with the total order induced by Γ_i fibers over $[0,1] \times S^1$. The fibration maps a limit configuration $c \in F(\tau)(L^t;\Gamma)$ to (t,z), where the collapse occurs at $K^t_j(z)$. The fiber over (t,z) is $\hat{C}(D_{(K^t_j)'(z)};\Gamma)$. In particular, the contribution of the collapse that orders U like Γ_i to the variation $(I(K^1_j;\Gamma) - I(K^0_j;\Gamma))$ during a knot isotopy $((z,t) \mapsto K^t_j(z))$ is proportional to the area covered by the unit derivative of K_j on S^2 during the isotopy, that is $\frac{I(K^1_j;\theta)-I(K^0_j;\theta)}{2}$. More precisely, it is

$$\frac{I(K^1_j;\theta) - I(K^0_j;\theta)}{2} I(\Gamma_i) .$$

Therefore, the contribution to the variation $\frac{(I(K^1_j;\Gamma)-I(K^0_j;\Gamma))}{\sharp \mathrm{Aut}(\Gamma)}$ of the anomalous faces is

$$\sum_{i=1}^{k} \frac{I(K^1_j;\theta) - I(K^0_j;\theta)}{2\sharp \mathrm{Aut}(\Gamma)} I(\Gamma_i) .$$

[6] This can be summarized by saying that the S^2-coordinates replace (z,λ).

The group of automorphisms of Γ_i is isomorphic to the subgroup $\mathrm{Aut}_0(\Gamma)$ of $\mathrm{Aut}(\Gamma)$ made of the automorphisms of Γ that fix U pointwise. The quotient $\frac{\mathrm{Aut}(\Gamma)}{\mathrm{Aut}_0(\Gamma)}$ is a subgroup of the cyclic group of the permutations of U that preserve the cyclic order of U, of order $\frac{k}{p}$, for some integer p that divides into k; and Γ_i is isomorphic to Γ_{i+p}, for any integer $i \leq (k-p)$. Thus, the contribution to the variation $\frac{(I(K_j^1;\Gamma)-I(K_j^0;\Gamma))}{\sharp \mathrm{Aut}(\Gamma)}$ of the anomalous faces is

$$\sum_{i=1}^{p} \frac{I(K_j^1;\theta) - I(K_j^0;\theta)}{2\sharp \mathrm{Aut}(\Gamma_i)} I(\Gamma_i) \,.$$

In general, one must multiply the infinitesimal variation due to the collapse of one connected component of the dashed graph by the contributions of the other connected components of the dashed graphs, and it is better to use Remark 4.1 to avoid thinking of the number of automorphisms.

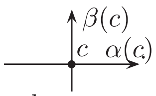
The integration of Formula 1 shows the Altschuler and Freidel formula:

$$Z_{CS}(L) = \exp\left(\frac{I(K_1;\theta)}{2}\alpha\right) \sharp_1 \exp\left(\frac{I(K_2;\theta)}{2}\alpha\right) \sharp_2 \dots$$
$$\exp\left(\frac{I(K_k;\theta)}{2}\alpha\right) \sharp_k Z^0_{CS}(L)$$

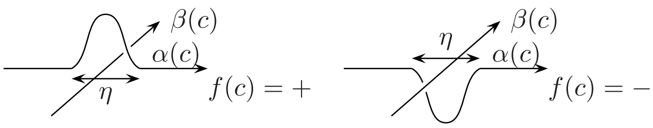
with respect to the structures defined in Subsect. 3.3.

5.4 Universality of Z^0_{CS}

In order to prove that Z^0_{CS} is a universal Vassiliev invariant, it is enough to compute its projection $\overline{Z}^0_{CS,k}(L)$ onto $\overline{\mathcal{A}}_k(S^1)$ when L is a singular link, that is a singular immersion of M with n double points, and when $k \leq n$.

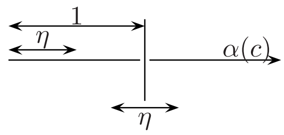
To do this, fix n disjoint balls $B(c)$ of radius 2 in $\mathbf{R}^3$ associated to the double points c of L, fix an almost planar representative of L that intersects each ball $B(c)$ as a pair of orthogonal linear horizontal arcs $\alpha(c)$ and $\beta(c)$ crossing at the center of $B(c)$ like 

The desingularisation of L associated to a map f will be obtained from this embedding by moving $\alpha(c)$ to one of the two following positions depending on the value $f(c)$ of the desingularisation f at c.

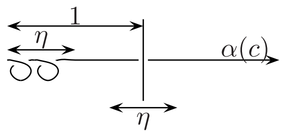


The obtained embeddings $L_0(f,\eta)$ depend on the small parameter η that is the diameter of the ball where $\alpha(c)$ has changed. But each of them will have integral Gauss integrals. Assume that the Gauss integrals of the positive desingularisation (where $f(c) = +$ for all c) of L vanish.

For every double point c that is a self-crossing of a component of L, cancel the Gauss integral modification of $L_0(f,\eta)$ if $f(c) = -$ by a modification of $\alpha(c)$ in $B(c)$ where



becomes



.

Let $L(f,\eta)$ denote the obtained almost planar representative of the desingularisation f of L. For any uni-trivalent diagram Γ on M, set

$$I(L(\eta);\Gamma) = \sum_{f:\{1,2\dots,n\}\longrightarrow\{+,-\}} (-1)^{\sharp f^{-1}(-)} I(L(f,\eta);\Gamma)\ .$$

It is enough to prove that for any given $\varepsilon > 0$, and for any uni-trivalent diagram Γ without isolated chord on M of degree $\leq n$, there exists η such that:
If $\Gamma \neq D(L)$, then $|I(L(\eta);\Gamma)| < \varepsilon$, and,
if $\Gamma = D(L)$, then $|I(L(\eta);\Gamma) - \sharp\mathrm{Aut}(\Gamma)| < \varepsilon$.

Fix a double point c of L and a uni-trivalent diagram Γ on M. The configurations of Γ that map no univalent vertices of Γ to $\alpha(c)$ will contribute in the same way to the integral $I(L(f,\eta);\Gamma)$ corresponding to a desingularisation f and to the integral $I(L(f^c,\eta);\Gamma)$ corresponding to a desingularisation f^c obtained from f by changing $f(c)$ into $(-f(c))$. Therefore, they will not contribute at all to $I(L(\eta);\Gamma)$. Thus, we shall only consider configurations with at least one univalent vertex $u(c)$ on each $\alpha(c)$.

Similarly, the contributions to $I(L(\eta);\Gamma)$ of the configurations where the given neighborhood of a given double point c does not contain any other vertex $u_2(c)$ in $B(c)$ related by an edge to $u(c)$ approach 0 when η approaches 0. Thus, we are left with the diagrams with at least 2 vertices $u(c)$ and $u_2(c)$ in every $B(c)$, and with at most $2n$ vertices. They are diagrams with exactly two vertices in each ball $B(c)$. If some vertex $u_2(c)$ is trivalent, then considering the value of $(u_2(c)-u(c))$ rather than the $\mathbf{R}^3$-position parameter of $u_2(c)$, we can conclude that $I(L(\eta);\Gamma)$ approaches 0 when η approaches 0. Thus, we are left with the case where all the $u_2(c)$ are univalent and therefore Γ must be a chord diagram with one chord between $u(c)$ and $u_2(c)$ for every double point c. Since we only consider diagrams without isolated chords, $u_2(c)$ must be on $\beta(c)$. Thus, we are left with the chord diagrams with one chord between $\alpha(c)$ and $\beta(c)$ for every double point c, that is by definition the chord diagram of the singular link L.

The number of embeddings (up to isotopy within the neighborhoods of the double points) of the vertices of this chord diagram Γ that respect the pairing of the chords is the number of automorphisms of Γ, and the contribution to $I(L(\eta);\Gamma)$ of each isotopy class of such embeddings is the product over the double points of the following algebraic areas. For each double point c, the algebraic area is the difference between the algebraic area covered by the directions of the segments from a point of $\alpha(c)$ to a point of $\beta(c)$ in the positive η-desingularisation of c minus the algebraic area covered by the directions of the segments from a point of $\alpha(c)$ to a point of $\beta(c)$ in the negative η-desingularisation of c. It is not hard to see that this algebraic area approaches one when η goes to zero... and to conclude this sketch of proof. $\diamond$

5.5 Rationality of Z^0_{CS}

As it has been first noticed by Dylan Thurston in [34], Z^0_{CS} is rational. This means that for any integer n, and for any link L, the degree n part $Z^0_{CS,n}(L)$ of $Z^0_{CS}(L)$ is in $\mathcal{A}^{\mathbf{Q}}_n(\coprod_{i=1}^k S^1_i)$ that is the quotient of the *rational* vector space generated by the diagrams on $\coprod_{i=1}^k S^1_i$ by the STU relation. Indeed, if L is almost horizontal, $Z^0_{CS,n}(L)$ may be interpreted as the following differential degree.

Let $e_n = 3n-2$ be a number of edges greater or equal than the number of edges of degree n diagrams that might contribute with a non zero integral to the Chern-Simons series. We wish to interpret $Z^0_{CS,n}(L)$ as the differential degree of a map to $\left(S^2\right)^{e_n}$. We first modify the configuration space $\overline{C}(L;\Gamma)$ of a degree n diagram Γ whose set of edges is $E(\Gamma)$ by

$$\hat{C}(L;\Gamma) = \overline{C}(L;\Gamma) \times \left(S^2\right)^{e_n - \sharp E(\Gamma)} .$$

Next, in order to be able to map it to $\left(S^2\right)^{e_n}$, we *label* Γ, that is we orient the edges of Γ and we define a bijection from $E(\Gamma) \cup \{1,2,\dots,e_n - \sharp E(\Gamma)\}$ to $\{1,2,\dots,e_n\}$. This bijection transforms the map

$$\Psi \times \mathrm{Identity}\left(\left(S^2\right)^{e_n-\sharp E(\Gamma)}\right) : \hat{C}(L;\Gamma) \longrightarrow \left(S^2\right)^{E(\Gamma)} \times \left(S^2\right)^{e_n-\sharp E(\Gamma)}$$

into a map

$$\hat{\Psi} : \hat{C}(L;\Gamma) \longrightarrow \left(S^2\right)^{e_n}$$

whose oriented image is independent of a possible labelling of the vertices. For a given degree n diagram Γ, there are $\frac{2^{\sharp E(\Gamma)} e_n!}{\sharp \mathrm{Aut}(\Gamma)}$ labeled diagrams. Now,

$$Z^0_{CS,n}(L) = \sum_{\Gamma \text{ labeled diagram of degree } n} \frac{1}{2^{\sharp E(\Gamma)} e_n!} \int_{\hat{C}(L;\Gamma)} \hat{\Psi}^*\left(\bigwedge^{e_n} \omega\right)[\Gamma]$$

Define the differential degree $\deg(\Psi, x)$ of Ψ over the formal union

$$\cup_{\Gamma \text{ labeled diagram of degree } n} \frac{1}{2^{\sharp E(\Gamma)} e_n!} [\Gamma] \hat{C}(L;\Gamma)$$

as follows for a regular[7] point $x \in (S^2)^{e_n}$:

$$\deg(\Psi, x) = \sum_{\Gamma \text{ labeled diagram of degree } n} \frac{1}{2^{\sharp E(\Gamma)} e_n!} \deg(\Psi_{|\hat{C}(L;\Gamma)}, x)[\Gamma]$$

where $\deg(\Psi_{|\hat{C}(L;\Gamma)}, x)$ is a usual differential degree. Then D. Thurston proved that $\deg(\Psi, x)$ does not vary across the images of the codimension one faces of the $\hat{C}(L;\Gamma)$. See also [30]. In other words, the above weighted union of configuration spaces behaves as a closed $2e_n$-dimensional manifold from the point of view of the differential degree theory. In particular, ω can be replaced by any volume form of S^2 with total volume 1. Computing $Z^0_{CS,n}$ as the degree of a generic point of $\left(S^2\right)^{e_n}$ shows that $Z^0_{CS,n}$ belongs to the lattice of $\mathcal{A}_n^{\mathbf{Q}}(\coprod_{i=1}^k S_i^1)$ generated by the $\frac{(e_n - \sharp E(\Gamma))!}{2^{\sharp E(\Gamma)} e_n!}[\Gamma]$, where the Γ's are the degree n graphs that may produce a nonzero integral. This interpretation is more convenient for computational purposes.

6 Diagrams and Lie Algebras. Questions and Problems

In this section, we are going to show how a Lie algebra equipped with a non-degenerate symmetric invariant bilinear form and some representation induces a linear form on $\mathcal{A}_n(S^1)$. In particular, such a datum allows one to deduce numerical knot invariants from the Chern-Simons series by composition.

First of all, we recall the needed background about Lie algebras.

6.1 Lie Algebras

Definition 6.1. A (finite-dimensional) *Lie algebra* over $\mathbf{R}$ is a vector space g over $\mathbf{R}$ of finite dimension equipped with a *Lie bracket* that is a bilinear map denoted by $[.,.] : g \times g \to g$ that satisfies:
the *antisymmetry relation:*

$$\forall x \in g, \quad [x,x] = 0 \quad (\Longrightarrow \quad \forall (x,y) \in g^2, \quad [x,y] = -[y,x])$$

and the *Jacobi relation:*

$$\forall (x,y,z) \in g^3, \quad [x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0\ .$$

[7] Here, regular means regular with respect to all the $\Psi_{|\hat{C}(L;\Gamma)}$.

Examples 6.2. 1. The vector space of the endomorphisms of $\mathbf{R}^N$, $gl_N = gl(\mathbf{R}^N)$, equipped with the bracket

$$\begin{array}{ll}[.,.] : gl_N \times gl_N & \to gl_N \\ \quad (x,y) & \mapsto xy - yx\end{array}$$

is a Lie algebra.

2. The vector space of the trace zero endomorphisms of $\mathbf{R}^N$, $sl_N = sl(\mathbf{R}^N)$, equipped with the restriction of the above Lie bracket is a Lie subalgebra of gl_N.

Definition 6.3. A (finite-dimensional) *representation* of such a Lie algebra g into a (finite-dimensional) $\mathbf{R}$-vector space E is a *Lie algebra morphism* ρ from g to $gl(E)$, where a Lie algebra morphism is a $\mathbf{R}$-linear map that preserves the Lie bracket ($\rho([x,y]) = [\rho(x),\rho(y)] \stackrel{\text{def}}{=} \rho(x)\rho(y) - \rho(y)\rho(x)$).

Examples 6.4. 1. The inclusion i_N of sl_N into gl_N is called the *standard representation* of sl_N into the vector space $\mathbf{R}^N$.
2. For any Lie algebra g, the morphism

$$\begin{array}{rl}\text{ad} : g & \longrightarrow gl(g) \\ x \mapsto & (\text{ad}(x) : y \mapsto ad(x)(y) = [x,y])\end{array}$$

is a representation of g (thanks to the Jacobi identity) that is called the *adjoint representation* of g.

Definition 6.5. A bilinear form $\beta : g \times g \longrightarrow \mathbf{R}$ is said to be *ad-invariant* or *invariant* if it satisfies

$$\forall (x,y,z) \in g^3, \quad \beta([x,z],y) + \beta(x,[y,z]) = 0 \;.$$

Example 6.6. When a Lie algebra is equipped with a representation (E,ρ), the *associated bilinear form*

$$\beta(\rho) : g \times g \longrightarrow \mathbf{R}$$

$$(x,y) \mapsto \text{trace}(\rho(x)\rho(y))$$

is a symmetric invariant bilinear form on g.

Definition 6.7. Let $\beta : g \times g \longrightarrow \mathbf{R}$ be a non-degenerate symmetric bilinear invariant form on a Lie algebra g. Then β induces the natural isomorphism $\hat{\beta} : g \longrightarrow g^*$ that maps x to $(\hat{\beta}(x) : y \mapsto \beta(x,y))$. The *Casimir element* Ω_β of β is the inverse of $\hat{\beta}$ viewed as an element of $g \otimes g$ with the help of the following canonical identifications.

$$g \otimes g \cong (g^*)^* \otimes g \cong \text{Hom}(g^*, g)$$

Exercise 6.8. Under the hypotheses of the above definition, let $(e_i)_{i=1,\dots,n}$ and $(e'_i)_{i=1,\dots,n}$ be two dual bases of g with respect to β that are two bases such that $\beta(e_i, e'_j)$ is the Kronecker symbol δ_{ij}. Show that

$$\Omega_\beta = \sum_{i=1}^{n} e_i \otimes e'_i$$

Note that this shows that the right-hand side of the above equality is independent of the choice of the two dual bases.

Proposition 6.9. *Let $(e_i)_{i=1,\dots,N}$ be a basis of $\mathbf{R}^N$, let $(e_i^*)_{i=1,\dots,N}$ denote its dual basis, and let e_{ij} be the element $e_j^* \otimes e_i$ of $gl_N \cong (\mathbf{R}^N)^* \otimes \mathbf{R}^N$. The form β_N of sl_N is non-degenerate,*

$$\Omega_{\beta_N} = \sum_{\substack{(i,j)\in\{1,\dots,N\}^2\\ i\neq j}} e_{ij} \otimes e_{ji} + \sum_{i=1}^{N} \left(e_{ii} - \frac{1}{N}\sum_{j=1}^{N} e_{jj} \right) \otimes \left(e_{ii} - \frac{1}{N}\sum_{j=1}^{N} e_{jj} \right)$$

and

$$\Omega_{\beta_N} = \sum_{(i,j)\in\{1,\dots,N\}^2} e_{ij} \otimes e_{ji} - \frac{1}{N}\left(\sum_{i=1}^{N} e_{ii}\right) \otimes \left(\sum_{i=1}^{N} e_{ii}\right)$$

PROOF: Let β be the form induced by the standard representation on gl_N that is $(x,y) \mapsto \mathrm{trace}(xy)$. The form β is non-degenerate on gl_N because the basis $(e_{ij})_{(i,j)\in\{1,\dots,n\}^2}$ of gl_N is dual to the basis (e_{ji}). Furthermore, $\beta\left(\sum_{i=1}^N e_{ii}, \sum_{i=1}^N e_{ii}\right) = N \neq 0$. Thus, since sl_N is the orthogonal of $\left(\sum_{i=1}^N e_{ii}\right)$ in gl_N, β_N that is the restriction of the symmetric form β to sl_N is non-degenerate. It is easy to check that the two proposed expressions of Ω_{β_N} are equal. The first one makes clear that our candidate belongs to $sl_N \otimes sl_N$. Now, it is enough to evaluate the second expression of our candidate viewed as an element of $\mathrm{Hom}(gl_N^*, gl_N)$ at the $\hat{\beta}(e_{ji})$, $j \neq i$ and at the $\hat{\beta}(e_{ii} - e_{jj})$ that are mapped to e_{ji} and $(e_{ii} - e_{jj})$, respectively, as they must be. ◇

6.2 More Spaces of Diagrams

We need to introduce more kinds of diagrams. Namely, we need to consider diagrams with free univalent vertices labeled by a finite set A.

Definition 6.10. Let M be an oriented one-manifold and let A be a finite set. A *diagram* Γ with support $M \cup A$ is a finite uni-trivalent graph Γ such that every connected component of Γ has at least one univalent vertex, equipped with:

1. a partition of the set U of univalent vertices of Γ also called *legs* of Γ into two (possibly empty) subsets U_M and U_A,
2. a bijection f from U_A to A,
3. an isotopy class of injections i of U_M into the interior of M,
4. an *orientation* of every trivalent vertex, that is a cyclic order on the set of the three half-edges which meet at this vertex.

Such a diagram Γ is again represented by a planar immersion of $\Gamma \cup M$ where the univalent vertices of U_M are located at their images under i, the one-manifold M is represented by solid lines, whereas the diagram Γ is dashed. The vertices are represented by big points. The local orientation of a trivalent vertex is again represented by the counterclockwise order of the three half-edges that meet at it.

Let $\mathcal{D}_n(M,A)$ denote the real vector space generated by the degree n diagrams on $M \cup A$, and let $\mathcal{A}_n(M,A)$ denote the quotient of $\mathcal{D}_n(M,A)$ by the relations AS, STU and IHX.

6.3 Linear Forms on Spaces of Diagrams

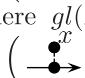
Notation 6.11. Let g be a finite-dimensional Lie algebra equipped with a finite-dimensional representation (E,ρ) and a non-degenerate bilinear symmetric invariant form β. For any oriented compact one-manifold M, the set of the boundary points of M where the corresponding components start as in $\bullet\!\!\to$ is denoted by ∂M^- whereas $\partial M \setminus \partial M^-$ is denoted by ∂M^+. When A is a finite set decomposed as the disjoint union of two subsets denoted by A^+ and A^-, and when M is an oriented compact one-manifold define

$$T(g,\rho,\beta)(M, A^- \coprod A^+) = \bigotimes^{\partial M^-} E^* \otimes \bigotimes^{\partial M^+} E \otimes \bigotimes^{A^-} g^* \otimes \bigotimes^{A^+} g$$

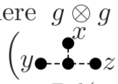
$T(g,\rho,\beta)(M, A^- \coprod A^+)$ is the tensor product of $\sharp \partial M^-$ copies of E^* indexed by the elements of ∂M^-, $\sharp\partial M^+$ copies of E indexed by the elements of ∂M^+, and copies of g^* and g indexed by the elements of A. In particular, when the boundary of M is empty $T(g,\rho,\beta)(M,\emptyset) = \mathbf{R}$.

Set

$T(g,\rho,\beta)\left(\longrightarrow \in \mathcal{A}([0,1])\right) = \text{Identity} \in gl(E)$
where $gl(E) \cong E^* \otimes E = T(g,\rho,\beta)([0,1])$,

$T(g,\rho,\beta)\Big($  $\in \mathcal{A}([0,1],\{x\}^-)\Big) = \rho \in \mathrm{Hom}(g, gl(E))$
where $\mathrm{Hom}(g,gl(E)) \cong g^* \otimes E^* \otimes E = T(g,\rho,\beta)([0,1],\{x\}^-)$,

$T(g,\rho,\beta)\left(x\bullet\text{---}\bullet y \in \mathcal{A}(\{x,y\}^+)\right) = \Omega_\beta \in g \otimes g$
where $g \otimes g = T(g,\rho,\beta)(\{x,y\}^+)$,

$T(g,\rho,\beta)\Big($  $\in \mathcal{A}(\{x,y,z\}^-)\Big) = \hat{\beta} \circ [.,.] \in \mathrm{Bil}(g_x \times g_y, g_z^*)$
where $\mathrm{Bil}(g_x \times g_y, g_z^*) \cong (g_x \otimes g_y)^* \otimes g_z^* \cong g_x^* \otimes g_y^* \otimes g_z^*$.

Note that $T(g,\rho,\beta)\left(x\bullet\!\text{---}\!\bullet y\right)$ is symmetric with respect to the permutation of the two factors. Also note that $T(g,\rho,\beta)\left(y\bullet\!\text{-}\!\overset{\overset{x}{\bullet}}{\bullet}\!\text{-}\!\bullet z\right)$ can be canonically identified with the trilinear form

$$g_x \times g_y \times g_z \longrightarrow \mathbf{R}$$

$$(a,b,c) \mapsto \beta([a,b],c)$$

In particular, it is antisymmetric with respect to any transposition of two factors g^*. Therefore, the above definitions make sense.

For any $\mathbf{R}$-vector space E, recall the natural linear *contraction map* from $E^* \otimes E$ to $\mathbf{R}$ that maps $f \otimes e$ to $f(e)$.

Theorem 6.12 (Bar-Natan [4]). *Let g be a finite-dimensional Lie algebra equipped with a finite-dimensional representation (E,ρ) and a non-degenerate bilinear symmetric invariant form β. Then there exists a unique family of linear maps*

$$T(g,\rho,\beta) : \mathcal{A}_n\left(M, A^- \coprod A^+\right) \longrightarrow T(g,\rho,\beta)\left(M, A^- \coprod A^+\right)$$

for any $n \in \mathbf{N}$, for any oriented compact one-manifold M, and for any finite set $A^- \coprod A^+$ such that:
1. *$T(g,\rho,\beta)$ takes the above values at the four diagrams above.*
2. *For any $(M, A^- \coprod A^+)$ as above, and for any $x^- \in \partial M^-$, and $x^+ \in \partial M^+$, the following diagram commutes:*

$$\begin{array}{ccc}
\mathcal{A}_n(M, A^- \coprod A^+) & \xrightarrow{T(g,\rho,\beta)} & T(g,\rho,\beta)(M, A^- \coprod A^+) \\
\Big\downarrow \pi_* & & \Big\downarrow \text{Contraction of the factors of } x^+ \text{ and } x^- \\
\mathcal{A}_n(\tilde{M}, A^- \coprod A^+) & \xrightarrow{T(g,\rho,\beta)} & T(g,\rho,\beta)(\tilde{M}, A^- \coprod A^+)
\end{array}$$

where $\tilde{M} = M/(x^+ \sim x^-)$ denotes the compact oriented one-manifold obtained from M by identifying x^+ and x^-, $\pi : M \longrightarrow \tilde{M}$ is the associated quotient map, and π_* is the induced map on diagram spaces.
3. *For any $(M, A^- \coprod A^+)$ as above, and for any $a^- \in A^-$, and $a^+ \in A^+$, the following diagram commutes:*

$$\begin{array}{ccc}
\mathcal{A}_n(M, A^- \coprod A^+) & \xrightarrow{T(g,\rho,\beta)} & T(g,\rho,\beta)(M, A^- \coprod A^+) \\
\Big\downarrow \iota & & \Big\downarrow \text{Contraction of the factors of } a^+ \text{ and } a^- \\
\mathcal{A}_{n-1}(M, \tilde{A}) & \xrightarrow{T(g,\rho,\beta)} & T(g,\rho,\beta)(M, \tilde{A})
\end{array}$$

where $\tilde{A} = (A^- \setminus \{a^-\}) \coprod (A^+ \setminus \{a^+\})$ and the map ι consists in identifying the two univalent vertices labeled by a^+ and a^-.

Proof: We first define $T = T(g,\rho,\beta)$ in a consistent way for diagrams, and next, we show that it factors through the relations AS, IHX, and STU. Except for the diagrams that have components like $x\bullet\text{---}\bullet y$ where x (by symmetry) belongs to the set labeled by $-$, any diagram can be decomposed into finitely many pieces like the four pieces where T has already been defined. The decomposition is unique when the number of $\longrightarrow$ is required to be minimal. Then the behaviour of T under gluings compels us to define T at any diagram as the tensor obtained by contracting the elementary tensors associated to the elementary pieces of such a decomposition, by the contractions corresponding to the gluings. To complete the definition, set

$$T(g,\rho,\beta)\left(x\bullet\text{---}\bullet y\ \in \mathcal{A}(\{x,y\}^-)\right) = \beta \in g^* \otimes g^*$$

and

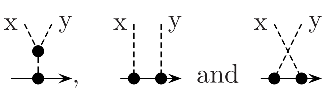
$$T(g,\rho,\beta)\left(x\bullet\text{---}\bullet y\ \in \mathcal{A}(\{x\}^- \coprod \{y\}^+)\right) = (\text{Identity} : g \to g) \in g^* \otimes g$$

Now, T is defined at any diagram.

When U, V and W are three **R**-vector spaces, if $f \in \mathrm{Hom}(U,V) \cong U^* \otimes V$ and $g \in \mathrm{Hom}(V,W) \cong V^* \otimes W$, then the contraction of $V \otimes V^*$ maps $(f \otimes g) \in U^* \otimes V \otimes V^* \otimes W$ to $g \circ f$. In particular, inserting trivial pieces like $\longrightarrow$ in the decomposition of a diagram does not change the resulting tensor, and the behaviour under the identification of boundary points is the desired one for diagrams that do not contain exceptional components like $x\bullet\text{---}\bullet y$. For the other ones, it is enough to notice that a contraction of $g^* \otimes g$ maps the tensor product of the two symmetric tensors $(\Omega_\beta = \hat{\beta}^{-1}) \otimes (\beta \cong \hat{\beta})$ to $(\text{Identity} : g \to g) \in g^* \otimes g$ that is mapped to the Identity map of g^* by the permutation of the factors. In fact, this just amounts to say that g and g^* are always identified via $\hat{\beta}$.

Now, T is consistently defined for diagrams. The antisymmetry relation for trivalent vertices comes from the total antisymmetry of $T\left(y\bullet\text{-}\overset{x}{\bullet}\text{-}\bullet z\right)$.

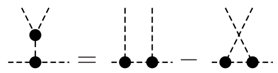
Let us show that T factors through STU. To do that, we view the values of T at the three local parts of STU



that are tensors in $g_x^* \otimes g_y^* \otimes E^* \otimes E$ as three bilinear maps from $g_x \times g_y$ to $gl(E)$, that map $(a,b) \in g_x \times g_y$ to $\rho([b,a])$, $\rho(b) \circ \rho(a)$ and $\rho(a) \circ \rho(b)$, respectively.

Thus, T factors through STU because the representation ρ represents.

By AS, the relation IHX can be redrawn as

IHX : 

and treated as a particular case of STU using the adjoint representation instead of ρ. Thus, T factors through IHX because the adjoint representation represents, and that is because of the Jacobi identity. $\diamond$

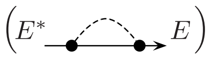
Remark 6.13. We could slightly generalize the maps T by *colouring* every component of the plain one-manifold with a different representation of g. Everything works similarly. Here gluings must respect colours.

Examples 6.14. Let us fix $N \in \mathbf{N}\setminus\{0,1\}$, and let us compute some examples when $g = sl_N$, $E = \mathbf{R}^N$, ρ is the standard representation that is the inclusion i_N of sl_N into gl_N, and $\beta = \beta_N$ is the associated invariant form. Set $T_N = T(sl_N, i_N, \beta_N)$.

1. The contraction from $gl(E) \cong E^* \otimes E$ to $\mathbf{R}$ is nothing but the trace of endomorphisms. In particular,

$$T_N(\bigcirc) = \text{trace}(\text{Identity of } \mathbf{R}^N) = N$$

2. Consider the following diagram $\Gamma_{11} \in \mathcal{A}(\uparrow)$ with only one chord.

$\Gamma_{11} =$ 

Let us compute $T_N(\Gamma_{11})$ viewed as an endomorphism of E. It is computed from the Casimir

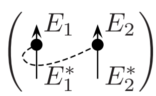
$$\Omega_{\beta_N} = \sum_{(i,j)\in\{1,\dots,N\}^2} e_{ij} \otimes e_{ji} - \frac{1}{N}\left(\sum_{i=1}^N e_{ii}\right) \otimes \left(\sum_{i=1}^N e_{ii}\right)$$

as

$$\begin{aligned} T_N(\Gamma_{11}) &= \textstyle\sum_{(i,j)\in\{1,\dots,N\}^2} e_{ji} \circ e_{ij} - \frac{1}{N}\left(\sum_{i=1}^N e_{ii}\right) \circ \left(\sum_{i=1}^N e_{ii}\right) \\ &= \left(N - \frac{1}{N}\right)\left(\text{Identity} = \textstyle\sum_{i=1}^N e_{ii}\right). \end{aligned}$$

Thus, $T_N(\Gamma_{11})$ is nothing but the multiplication by the number $(N^2-1)/N$.

3. Now, consider the following diagram $\Gamma_{12} \in \mathcal{A}(\uparrow\uparrow)$ with only one horizontal chord between the two strands.

$\Gamma_{12} =$ 

Let us compute $T_N(\Gamma_{12})$ viewed as an endomorphism of $(E_1 = E)\otimes(E_2 = E)$. It is computed from the Casimir Ω_{β_N} as

$$T_N(\Gamma_{12})(e_k \otimes e_l) = \sum_{(i,j)\in\{1,\ldots,N\}^2} e_{ij}(e_k) \otimes e_{ji}(e_l) - \frac{1}{N}(e_k \otimes e_l)$$

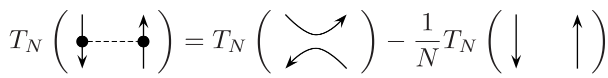
Thus,

$$T_N(\Gamma_{12})(e_k \otimes e_l) = e_l \otimes e_k - \frac{1}{N}(e_k \otimes e_l)$$

and, in general,

$$T_N(\Gamma_{12}) = \tau - \frac{1}{N}\text{Identity}$$

where $\tau : E \otimes E \longrightarrow E \otimes E$ is the transposition of the two factors that maps $(x \otimes y)$ to $(y \otimes x)$. This can be written as



and allows for a simple recursive computation of the evaluation of T_N at chord diagrams on disjoint union of circles by induction on the number of chords starting with

$$T_N\left(1 \in \mathcal{A}\left(\coprod^k S^1\right)\right) = N^k .$$

4. Let $\wedge^2(E)$ be the subspace of $E \otimes E$ generated by the antisymmetric elements of the form $(x \otimes y - y \otimes x)$, and let $S^2(E)$ be the subspace of $E \otimes E$ generated by the symmetric elements of the form $(x \otimes y + y \otimes x)$. Since $x \otimes y = \frac{1}{2}((x \otimes y - y \otimes x) + (x \otimes y + y \otimes x))$, $E \otimes E = \wedge^2(E) + S^2(E)$. Furthermore, the restriction of $T_N(\Gamma_{12})$ to $S^2(E)$ is the multiplication by $(1 - \frac{1}{N})$ while the restriction of $T_N(\Gamma_{12})$ to $\wedge^2(E)$ is the multiplication by $(-1-\frac{1}{N})$. In particular, $E \otimes E = \wedge^2(E) \oplus S^2(E)$, and this decomposes $E \otimes E$ into the two eigenspaces associated to the eigenvalues $(1-\frac{1}{N})$ and $(-1-\frac{1}{N})$. Let Γ_{12}^n denote the chord diagram with n horizontal chords between the two strands in $\mathcal{A}(\uparrow\uparrow)$. Then $T_N(\Gamma_{12}^n) = (1-\frac{1}{N})^n p_{S^2(E)} + (-1-\frac{1}{N})^n p_{\wedge^2(E)}$, and

$$\begin{aligned} T_N(\Gamma_{12}^n)(x \otimes y) = & \tfrac{(N-1)^n + (-N-1)^n}{2N^n}(x \otimes y) \\ & + \tfrac{(N-1)^n - (-N-1)^n}{2N^n}(y \otimes x) \end{aligned}$$

Remark 6.15. We can construct an injection Δ from $\overline{\mathcal{A}}_n(S^1)$ into $\mathcal{A}_n(S^1)$ as in [20] in order to be able to deduce linear forms on $\overline{\mathcal{A}}_n(S^1)$ from the above linear forms by composition.

We first contruct the linear map $\tilde{\delta} : \mathcal{D}_n(S^1) \longrightarrow \mathcal{A}_{n-1}(S^1)$ that maps a chord diagram D with n chords to the sum (over these n chords) of the n diagrams obtained by deleting one chord from D. This map factors through $\mathcal{A}_n(S^1)$. Let $\delta : \mathcal{A}_n(S^1) \longrightarrow \mathcal{A}_{n-1}(S^1)$ be the induced map. Then for any $(d, d') \in \mathcal{A}(S^1)^2$, with the algebra structure of Subsect. 3.3,

$$\delta(dd') = \delta(d)d' + d\delta(d') \; .$$

Now, define the linear map $\tilde{\Delta} : \mathcal{A}(S^1) \longrightarrow \mathcal{A}(S^1)$ that maps an element d of $\mathcal{A}_n(S^1)$ to

$$\tilde{\Delta}(d) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \theta^k \delta^k(d) \in \mathcal{A}_n(S^1) \; .$$

Note that $\tilde{\Delta}$ is a morphism of algebras that maps θ to 0. Therefore $\tilde{\Delta}$ factors through $\overline{\mathcal{A}}_n(S^1)$. Let Δ be the induced map and let $P : \mathcal{A}(S^1) \longrightarrow \overline{\mathcal{A}}(S^1)$ be the canonical projection. Since $P \circ \Delta$ is the identity, Δ is an injection.

6.4 Questions

Question 1: *Do real-valued finite type invariants distinguish knots?*

Note that this question is equivalent to the following one. Does the Chern-Simons series distinguish knots? The answer to this question is unknown. Nevertheless, very interesting results obtained by Goussarov [12], Habiro [15] and Stanford [33] present modifications on knots that do not change their invariants of degree less than any given integer n and such that any two knots with the same Vassiliev invariants of degree less than n can be obtained from one another by a sequence of these explicit modifications.

Question 2: *Relate the Chern-Simons series to other invariants.*

There is another famous universal Vassiliev link invariant –that is a possibly different map $\overline{Z}$ that satisfies the conclusions of Theorem 2.26– that is called the *Kontsevich integral* and will be denoted by Z_K (see [4, 10, 23]). The HOMFLY polynomial that is a generalization of the Jones polynomial for links discovered shortly after independently by Hoste, Ocneanu, Millet, Freyd, Lickorish, Yetter, Przytycki and Traczyk, can be expressed as a function of Z_K as $P = T^N \circ Z_K$ where

$$T^N \left(\sum_{n \in \mathbf{N}} d_n \in \mathcal{A}(M) \right) = \sum_{n \in \mathbf{N}} \lambda^n T_N(d_n)$$

with the notation of Subsect. 6.3. See [23]. The Homfly polynomial belongs to $\mathbf{R}[N][[\lambda]]$. The specialization of P at $N = 2$ is equal to the Jones polynomial. Setting $N = 0$ and performing the change of variables $t^{1/2} = \exp(\lambda/2)$ yields the famous *Alexander-Conway polynomial* discovered by Alexander in the beginning of the twentieth century. Le and Murakami [25] proved that all quantum invariants for knots can be obtained as the composition by Z_K of a linear map constructed as in Subsect. 6.3.

Is is conjectured but still unproved that the Kontsevich integral coincides with the Chern-Simons series. Sylvain Poirier proved that if the anomaly vanishes in degree greater than 6, then these two universal invariants coincide. In [24], using results of Poirier, I give the form of an isomorphism of $\overline{\mathcal{A}}$ that transforms one invariant into the other.

Let $\beta = (\beta_n)_{n\in\mathbf{N}}$ be an element of $\mathcal{A}(\emptyset, \{1,2\})$ that is symmetric with respect to the exchange of 1 and 2. (In fact, according to [35, Corollary 4.2], all two-leg elements are symmetric with respect to this symmetry modulo the standard AS and IHX relations. This is reproved as Lemma 7.18 below.) If Γ is a chord diagram, then $\Psi(\beta)(\Gamma)$ is defined by replacing each chord by β. By Lemma 7.16, $\Psi(\beta)$ is a well-defined morphism of topological vector spaces from $\mathcal{A}(M)$ to $\mathcal{A}(M)$ for any one-manifold M, and $\Psi(\beta)$ is an isomorphism as soon as $\beta_1 \neq 0$.

Theorem 6.16 ([24]). *There exists*

$$\beta = (\beta_n)_{n\in\mathbf{N}} \in \mathcal{A}(\emptyset, \{1,2\})$$

such that

– *the anomaly α reads $\alpha = \Psi(\beta)$* *,*
– *for any (zero-framed) link L, the Chern-Simons series $Z^0_{CS}(L)$ is equal to $\Psi(\beta)(Z_K(L))$.*

Of course, the following question is still open.
Question 3: *Compute the anomaly.*

After the articles of Axelrod, Singer [2, 3], Bott and Cattaneo [6, 7, 9], Greg Kuperberg and Dylan Thurston have constructed a universal finite type invariant for 3-dimensional homology spheres in the sense of [22, 29] as a series of configuration space integrals similar to Z^0_{CS}, in [21]. Their construction yields two natural questions:
Question 4: *Find a surgery formula for the Kuperberg-Thurston invariant in terms of the above Chern–Simons series.*
Question 5: *Compare the Kuperberg–Thurston invariant to the LMO invariant constructed in [26].*

7 Complements

7.1 Complements to Sect. 1

Remark 7.1. Definition 1.2 of isotopic embeddings is equivalent to the following one:
A *link isotopy* is a C^∞ map

$$h : \coprod^k S^1 \times I \longrightarrow \mathbf{R}^3$$

such that $h_t = h(.,t)$ is an embedding for all $t \in [0,1]$. Two embeddings f and g as above are said to be *isotopic* if there is an isotopy h such that $h_0 = f$ and $h_1 = g$.

The non-obvious implication of the equivalence comes from the isotopy extension theorem [17, Theorem 1.3, p.180].

Exercise 7.2. (**) Prove that for any C^∞ embedding $f : S^1 \longrightarrow \mathbf{R}^3$, there exists a continuous map $h : S^1 \times I \longrightarrow \mathbf{R}^3$ such that $h_t = h(.,t)$ is a C^∞ embedding for all $t \in [0,1]$, h_0 is a representative of the trivial knot, and $h_1 = f$. (Hint: Put the complicated part of f in a box, and shrink it.)

Sketch of proof of Proposition 1.5: In fact, it could be justified with the help of [17] that when the space of representatives of a given link is equipped with a suitable topology, the representatives whose projection is regular form a dense open subspace of this space. The reader can also complete the following sketch of proof. A *PL* or *piecewise linear* link representative is an embedding of a finite family of polygons whose restrictions to the polygon edges are linear. Such a PL representative can be *smoothed* by replacing a neighbborhood of a vertex like  by  in the same plane. It is a representative of our given link if the smooth representatives obtained by smoothing close enough to the vertices are representatives of our link. A planar linear projection of such a PL representative is *regular* if there are only finitely many multiple points which are only double points without vertices in their inverse image. Observe that an orthogonal projection of a generic PL representative is regular if the direction of the kernel of the projection avoids:

I. the vector planes parallel to the planes containing one edge and one vertex outside that edge.
II. the directions of the lines that meet the interiors of 3 distinct edges.

Fix a triple of pairwise non coplanar edges. Then for every point in the third edge there is at most one line intersecting this point and the two other edges. One can even see that the set of directions of lines intersecting these three edges is a dimension one compact submanifold of the projective plane $\mathbf{R}P^2$ parametrized by subintervals of this third edge. Thus, the set of allowed oriented directions for the kernel of the projection is the complement of a finite number of one-dimensional submanifolds of the sphere S^2. Therefore it is an open dense subset in S^2 according to a weak version of the Morse-Sard theorem [17, Proposition 1.2, p.69] or [28, p.16]. Note that changing the direction of the projection amounts to composing the embedding by a rotation of $SO(3)$. Now, it is easy to smooth the projection, and to get a smooth representative whose projection is regular. ◇

Lack of proof of the Reidemeister theorem 1.7: It could be proved by studying the topology of the space of representatives of a given link, that a generic path between two representatives whose projections are regular in this space (i.e. a generic isotopy) only meets a finite number of times three walls made of singular representatives. The first wall that leads to the first Reidemeister move is made of the representatives that have one vertical tangent vector (and nothing else prevents their projections from being regular).

The second wall that leads to the second move is made of embeddings whose projections have one non-transverse double point while the third wall (that leads to RIII) is made of embeddings whose projections have one triple point. ◇

Remark 7.3. The unproved Reidemeister theorem is not needed in Sects. 2 to 6 of the course. Nevertheless, the following exercise can help understanding a proof idea of the Reidemeister theorem.

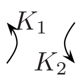
Exercise 7.4. (**) Say that a PL representative of a link is *generic* if its only pairs of coplanar edges are the pairs of edges that share one vertex. Consider a generic piecewise linear link representative. Prove that for any pair $(\rho,\sigma) \in SO(3)^2$ such that $\pi \circ \rho \circ f$ and $\pi \circ \sigma \circ f$ are regular, (smoothings of) $\pi \circ \rho \circ f$ and $\pi \circ \sigma \circ f$ are related by a finite sequence of Reidemeister moves.

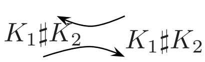
Elementary proof of Proposition 1.10: Number the components of the link from 1 to m. Choose an open arc of every component K_i of the link whose projection does not meet any crossing, and choose two distinct points b_i and a_i on this oriented arc so that a_i follows b_i on this arc. Then change the crossings in your link diagram if necessary so that:

1. If $i < j$, K_i crosses K_j under K_j.
2. When we follow the component K_i from a_i to b_i, we meet the lowest preimage of the crossing before meeting the corresponding highest preimage.

After these (possible) modifications, we get a diagram of a (usually different) link that is represented by an embedding whose first two coordinates can be read on the projection and whose third coordinate is given by a real-valued height function h that can be chosen so that $h(a_i) = 2i$, $h(b_i) = 2i+1$, and h is strictly increasing from a_i to b_i and strictly decreasing from b_i to a_i. (This is consistent with the above assumptions on the crossings.) Then the obtained link is a disjoint union of components (separated by horizontal planes) that have at most two points at each height and that can therefore not be knotted. ◇

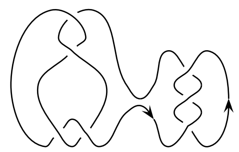
Definition 7.5. The *disjoint union* $K_1 \coprod K_2$ of two knots K_1 and K_2 is represented by two representatives of the knots sitting in two disjoint balls of the ambient space. The disjoint union of two regular projections of K_1 and K_2 – where the two projections lie in disjoint disks of the plane $\mathbf{R}^2$ – is a regular projection of $K_1 \coprod K_2$. The local change in such a projection of $K_1 \coprod K_2$:



becomes


transforms $K_1 \coprod K_2$ into the *connected sum* $K_1 \sharp K_2$ of K_1 and K_2. The connected sum of knots is a commutative well-defined operation. (Prove it as an exercise!) A knot is said to be *prime* if it cannot be written as a connected sum of two non-trivial knots. Modulo commutativity, every knot can be expressed in a unique way as the connected sum of a finite number of prime knots. (See [27, Theorem 2.12].) Let K be a knot represented by an embedding f from S^1 into $\mathbf{R}^3$. The *reverse* $-K$ of K is the knot represented by the embedding $f \circ \mathrm{conj}$ where conj is the complex conjugation acting on the unit circle S^1 of the complex plane. The *mirror image* $\overline{K}$ of K is the knot represented by the embedding $\sigma \circ f$ where σ is the reflection of $\mathbf{R}^3$ such that $\sigma(x, y, z) = (x, y, -z)$. If K is presented by a diagram D, $\overline{K}$ is presented by the *mirror image* $\overline{D}$ of D that is obtained from D by changing all its crossings.

Examples 7.6. The two trefoil knots are mirror images of each other. The figure-eight knot is its own mirror image. (You have surely proved it when solving Exercise 1.8!) There are knots which are not equivalent to their reverses, like the eight-crossing knot 8_{17} in [27, Table 1.1, p.5]. Here is a picture of the connected sum of the figure-eight knot and the right-handed trefoil knot.



The *connected sum of the figure-eight knot and the right-handed trefoil knot*

A table of prime knots with at most 9 crossings is given in [32]. In this table, knots are not distinguished from their reverses and their mirror images. According to Thistlethwaithe, with the same conventions, the table of prime knots with 15 crossings contains 253293 items [27, Table 1.2, p.6].

Remark 7.7. Every knot bounds an oriented[8] embedded surface in $\mathbf{R}^3$. See [32, p.120] or [27, Theorem 2.2]. Such a surface is called a Seifert surface of the knot. The linking number of two knots could be defined as the algebraic intersection number of a knot with a Seifert surface of the other one... and there are lots of other definitions like the original Gauss definition that is given in Sect. 4.1.

Examples 7.8. There are numerical knot invariants that are easy to define but difficult to compute like:

- the *minimal number* of crossings $m(L)$ in a projection,

[8] Boundaries are always oriented with the "outward normal first" convention.

- the *unknotting number* of a knot that is the minimal number of crossing changes to be performed in $\mathbf{R}^3$ to unknot the knot (i.e. to make it equivalent to the trivial knot),
- the *genus* of a knot that is the minimal genus of an oriented embedded surface bounded by the knot.

An invariant is said to be *complete* if it is injective. The knot itself is a complete invariant. There are invariants coming from algebraic topology like the fundamental group of the complement of the link. A *tubular neighborhood* of a knot is a solid torus $S^1 \times D^2$ embedded in $\mathbf{R}^3$ such that its core $S^1 \times \{0\}$ is (a representative of) the knot. (See [17, Theorem 5.2, p.110] for the existence of tubular neighborhoods.) A *meridian* of a knot is the boundary of a small disk that intersects the knot once transversally and positively. A *longitude* of the knot is a curve on the boundary of a tubular neighborhood of the knot that is parallel to the knot. Up to isotopy of the pair (knot, longitude), the longitudes of a knot are classified by their linking number with the knot. The *preferred longitude* of a knot is the one such that its linking number with the knot is zero. According to a theorem of Waldhausen [36], the fundamental group equipped with two elements that represent the oriented *meridian* of the knot and the preferred *longitude* is a complete invariant of the knot. (See also [16, Chap. 13]). According to a more recent difficult theorem of Gordon and Luecke [13], the *knot complement*, that is the compact 3-manifold that is the closure of the complement of a knot tubular neighborhood (viewed up to orientation-preserving homeomorphism), determines the knot up to orientation. Nevertheless, these meaningful invariants are hard to manipulate.

Exercise 7.9. The *segments* of a link diagram are the connected components of the link diagram that are segments between two undercrossings where the diagram is broken. An *admissible 3-colouring* of a link diagram is a function from the set of segments of a diagram to the three-element set {Blue, Red, Yellow} such that, for any crossing, the image of the set of (usually three) segments that meet at the crossing contains either one or three elements (exactly). Prove that the number of admissible 3-colourings is a link invariant. Use this invariant number to distinguish the trefoil knots from the figure-eight knot, and the Borromean link from the trivial 3-component link. (In fact the admissible 3-colourings of a link are in one-to-one correspondence with the representations of the fundamental group of the link complement to the group of permutations of 3 elements that map the link meridians to transpositions.)

The following additional properties of the Jones polynomial are not hard to check.

Proposition 7.10. *The Jones polynomial V satisfies the additional properties:*

3. For any link L,

$$V(\overline{L})(t) = V(L)(t^{-1}) \ .$$

4. For any two links L_1 *and* L_2,

$$V\left(L_1 \coprod L_2\right) = -(t^{1/2} + t^{-1/2})V(L_1)V(L_2)\ .$$

5. For any two knots K_1 *and* K_2,

$$V(K_1 \sharp K_2) = V(K_1)V(K_2)\ .$$

7.2 An Application of the Jones Polynomial to Alternating Knots

Definition 7.11. A link diagram is said to be *alternating* if the over-crossings and the under-crossings alternate as one travels along the link components. In this Subsect. 7.2, a *connected* link diagram is a diagram whose underlying knot projection is connected. In a connected link diagram, a crossing is said to be *separating* if one of the two transformations of the link diagram " becomes " or " becomes " makes the diagram disconnected, or, equivalently, if the transformation " becomes " makes the diagram disconnected.

The Jones polynomial allowed Kauffman and Murasugi to prove the following theorem in 1988, independently. This answered a Tait conjecture of 1898.

Theorem 7.12 (Kauffman-Murasugi, 1988). *When a knot* K *has a connected alternating diagram without separating crossing with* c *crossings, then* c *is the minimal number of crossings* $m(K)$ *of* K.

This theorem is a direct consequence of Proposition 7.14 about the properties of the breadth of the Jones polynomial.

Definition 7.13. The *breadth* $B(P)$ of a Laurent polynomial P is the difference between the maximal degree and the minimal degree occurring in the polynomial.

Proposition 7.14. *When a link* L *has a connected diagram with* c *crossings, then* $B(V(L)) \leq c$. *When a link* L *has a connected alternating diagram without separating crossing with* c *crossings, then* $B(V(L)) = c$.

The proof of this proposition will also yield the following obstruction for a link to have a connected alternating diagram without separating crossing.

Proposition 7.15. *When a link* L *has a connected alternating diagram without separating crossing, the coefficients of the terms of extremal degrees in its Jones polynomial* $V(L)$ *are* ± 1. *Furthermore, the product of these two coefficients is* $(-1)^{B(V(L))}$.

In particular, this obstruction shows that the eight-crossing knot 8_{21} has no connected alternating diagram without separating crossing. See [27, Table 1.1, p.5; Table 3.1, p.27].

PROOF OF PROPOSITIONS 7.14 AND 7.15: Let L be a link, and let D be one of its diagrams with c crossings. Observe that

$$B(V(L)) = \frac{B(< D >)}{4}$$

We will refer to the construction of the Kauffman bracket at the beginning of Subsect. 1.3. Let $C(D)$ denote the set of crossings of D. Let $f^L : C(D) \longrightarrow \{L, R\}$ be the constant map that maps every crossing to L, and let $f^R : C(D) \longrightarrow \{L, R\}$ be the constant map that maps every crossing to R. Set $n_L = n(D_{f^L})$ and $n_R = n(D_{f^R})$. We shall prove:

(i) $B(< D >) \leq 2c + 2n_L + 2n_R - 4$
(ii) The above inequality is an equality when D is a connected alternating diagram without separating crossing.
(iii) If D is a connected alternating diagram, then $n_L + n_R - 2 = c$.
(iv) If D is a connected diagram, then $n_L + n_R - 2 \leq c$.

It is clear that these four properties imply Proposition 7.14. Let us prove these properties.

A map f from $C(D)$ to $\{L, R\}$ gives rise to the term

$$A^{(\sharp f^{-1}(L) - \sharp f^{-1}(R))} \delta^{(n(D_f)-1)}$$

in $< D >$. This summand is a polynomial in A and A^{-1}
whose highest degree term is $(-1)^{(n(D_f)-1)} A^{h(f)}$ with

$$h(f) = \sharp f^{-1}(L) - \sharp f^{-1}(R) + 2(n(D_f) - 1)$$

and whose lowest degree term is $(-1)^{(n(D_f)-1)} A^{\ell(f)}$ with

$$\ell(f) = \sharp f^{-1}(L) - \sharp f^{-1}(R) - 2(n(D_f) - 1) \ .$$

Observe that $h(f^L) = c + 2n_L - 2$ and that $\ell(f^R) = -c - 2n_R + 2$.

Therefore, in order to prove (i), it is enough to show that for any map f from $C(D)$ to $\{L, R\}$, we have

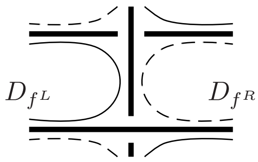
$$h(f) \leq h(f^L) \ \text{ and } \ \ell(f) \geq \ell(f^R)$$

Notice that if f and g are two maps from $C(D)$ to $\{L, R\}$ that coincide at every crossing but one, then $n(D_f) = n(D_g) \pm 1$. This property allows us to prove $h(f) \leq h(f^L)$ for all f by induction on $\sharp f^{-1}(R)$. Indeed, changing a value L of a map f into the value R removes 2 from $(\sharp f^{-1}(L) - \sharp f^{-1}(R))$ whereas $2(n(D_f) - 1)$ cannot increase by more than 2 by the above property. Thus, for all f, $h(f) \leq h(f^L)$. Similarly, $\ell(f) \geq \ell(f^R)$ and (i) is proved.

We prove (ii) and Proposition 7.15 for a connected alternating diagram D without separating crossing with c crossings. In view of the above arguments (and by definition of the Jones polynomial for Proposition 7.15), it is enough to prove that for any non constant map f from $C(D)$ to $\{L, R\}$, we have

$$h(f) < h(f^L) \text{ and } \ell(f) > \ell(f^R) .$$

We consider the underlying projection of D in the one-point compactification of the plane $\mathbf{R}^2$ that is the sphere S^2. The *faces* of the diagram D will be the connected components of $S^2 \setminus D$. Since D is connected, these components have only one boundary component. Therefore they are topological disks. The alternating nature of D allows us to push each connected component of D_{f^L} or D_{f^R} inside one face as the following picture shows:



This pushing defines a one-to-one correspondence between the faces of D and the connected components of $D_{f^L} \coprod D_{f^R}$. Choose a crossing x of D. Since x is not separating, the two parts of D_{f^L} near x bound distinct faces of D, and thus they belong to different components of D_{f^L}. Therefore, changing the value of f^L at x into R, changes D_{f^L} into a diagram D_f such that $n(D_f) = n_L - 1$, where f maps $(C(D) \setminus x)$ to L and x to R. Thus, $h(f) < h(f^L)$ for all the maps f such that $\sharp f^{-1}(R) = 1$, and for all the maps such that $\sharp f^{-1}(R) > 0$ by induction on $\sharp f^{-1}(R)$. Similarly, $\ell(f) > \ell(f^R)$ for all the maps f different from f^R. Thus (ii) and the first part of Proposition 7.15 are proved. The second part of 7.15 is a consequence of the above arguments and (iii) below.

(iii) is obtained by computing the Euler characteristic of the sphere as the number $(n_L + n_R)$ of faces of D plus the number (c) of crossings of D minus the number $(2c)$ of edges of the projection –that contain exactly two crossings which are at their extremities–. (It could also be proved by induction on c.)

We prove (iv). Let D be a connected diagram with c crossings. We want to prove:

$$n_L + n_R - 2 \leq c$$

by induction on c. This is true for the only connected diagram without crossing that is the diagram of the unknot. Let D be a connected diagram with $c \geq 1$ crossings. Let x be one of its crossings, and let D' be the diagram obtained from D by removing x in the left-handed way.

If D' is connected, then $n(D_{f^L}) = n(D'_{f'^L})$ while $D'_{f'^R}$ is equal to D_f where f maps $C(D) \setminus x$ to R and x to L, thus $n(D_{f^R}) = n(D'_{f'^R}) \pm 1$. Therefore, the inequality that is true for the diagram D' that has $(c-1)$ crossings, by induction hypothesis, implies that the inequality holds for D.

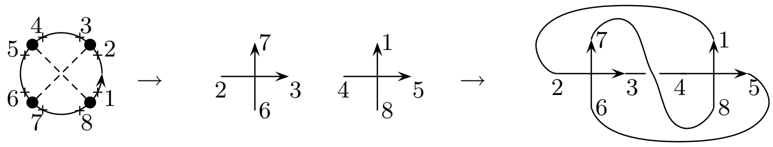
If D' is not connected, then the diagram $\overline{D}'$ obtained from the mirror image $\overline{D}$ of D by removing x in the left-handed way is connected, the inequality holds for $\overline{D}$ by the above argument, therefore it is true for D. ⋄

7.3 Complements to Sect. 2

Diagrammatic proof of Lemma 2.17: Though the first assertion is very easy, we prove it because its proof is the beginning of the proof of the second assertion. Let d be an n-chord diagram on S^1. Put 4n *cutting points,* numbered from 1 to 4n along S^1, on the support S^1 of d, one near each extremity of the 2n intervals separated by the vertices, so that $2i-1$ and $2i$ are on the same interval. Next embed the neighborhoods of the n double points, bounded by the 4n cutting points, into n fixed disjoint disks in $\mathbf{R}^2 = \mathbf{R}^2 \times \{0\}$ so that the cutting points lie on the boundaries of these disks. Choose n disjoint 3-dimensional balls that intersect $\mathbf{R}^2$ along these n disks. Let C be the closure of the complement of these n fixed balls in $\mathbf{R}^3$.

Then, in order to construct our first representative K^0 of d, it is enough to notice that we have enough room to embed the *remaining 2n intervals* of S^1 (the $[2i-1, 2i]$) into C. Next, the proof could be "concluded" as follows: Let K be another representative of d. After an isotopy, we may assume that K intersects our n balls like K^0 does. Then, since $\pi_1(C)$ is trivial, there is a boundary-fixing homotopy in C that maps the remaining 2n intervals for K to the remaining 2n intervals for K^0. Such a homotopy may be approximated by a finite sequence of (isotopies and) crossing changes, and we are done. However, we will again give a planar elementary proof.

We may demand that the orthogonal projection π of K^0 is regular and that the projections of the intervals $]2i-1, 2i[$ are embeddings that avoid the fixed neighborhoods of the double points, and that the interval $]2i-1, 2i[$ is under $]2j-1, 2j[$ if $i < j$. The projections of these three steps are represented in the following example:



Let K be another representative of d. After an isotopy, we may assume that K intersects our n balls like K^0 does and that the projection of K is regular and avoids the fixed neighborhoods of the double points. After some crossing changes, we may assume that $]1, 2[$ is above the other $]2i-1, 2i[$, we may unknot it as in the proof of Proposition 1.10 and we may assume that its projection coincides with the restriction of the projection of K^0. Do the same for $]3, 4[$: put it above everything else, unknot it, and make its projection coincide with the corresponding one for K^0, then for $]5, 6[, \ldots$, and finish with $]4n-1, 4n[$. $\diamond$

7.4 Complements to Subsect. 6.4

Let β be an element of $\mathcal{A}(\emptyset, \{1, 2\})$. Let Γ be a diagram with support M as in Definition 3.1. We define $\Psi(\beta)(\Gamma)$ to be the element of $\mathcal{A}(M)$ obtained by inserting β d times on each degree d component of Γ (where a *component* of Γ is a connected component of the dashed graph).

Lemma 7.16. *$\Psi(\beta)(\Gamma)$ does not depend on the choice of the insertion loci.*

Proof: It is enough to prove that moving β from an edge of Γ to another one does not change the resulting element of $\mathcal{A}(M)$, when the two edges share some vertex v. Since this move amounts to slide v through β, it suffices to prove that sliding a vertex from some leg of a two leg-diagram to the other one does not change the diagram modulo AS, IHX and STU, this is a direct consequence of Lemma 3.4 when the piece of diagram inside D is β. $\diamond$

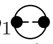
It is now easy to check that $\Psi(\beta)$ is compatible with the relations IHX, STU and AS. This allows us to define continuous vector space endomorphisms $\Psi(\beta)$ of the $\mathcal{A}(M)$ such that, for any diagram Γ:

$$\Psi(\beta)([\Gamma]) = \Psi(\beta)(\Gamma) \ .$$

Say that $(d_n)_{n \in \mathbf{N}} \in \mathcal{A}(M)$ is *of filtration at least d* if $d_k = 0$ for any $k < d$.

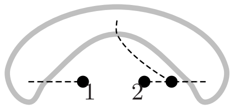
$\Psi(\beta)$ satisfies the following properties:

Lemma 7.17. *1. $\Psi(\beta)$ is compatible with the products of Subsect. 3.3. ($\Psi(\beta)(xy) = \Psi(\beta)(x)\Psi(\beta)(y)$.)*
2. If $\beta_1 \neq 0$, $\Psi(\beta)$ is an isomorphism of topological vector spaces such that $\Psi(\beta)$ and $\Psi(\beta)^{-1}$ map elements of filtration at least d to elements of filtration at least d.

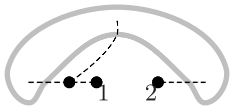
Proof: The first property is obvious. For the second one, first note that $\beta_1 = b_1$ for some non zero number b_1. Thus, for $x = \sum_{i=d}^{\infty} x_i$, $\Psi(\beta)(x) - (b_1^d x_d)$ is of filtration at least $d+1$. This shows that $\Psi(\beta)$ is injective and allows us to construct a preimage for any element by induction on the degree, proving that Ψ is onto. $\diamond$

Lemma 7.18 (Vogel). *Elements of $\mathcal{A}(\emptyset, \{1, 2\})$ are symmetric with respect to the exchange of 1 and 2.*

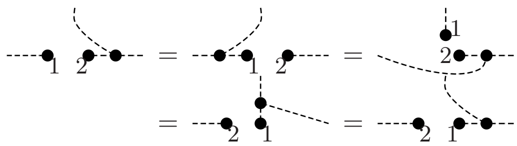
Proof: Since a chord is obviously symmetric, we can restrict ourselves to a two-leg diagram with at least one trivalent vertex and whose two univalent vertices are respectively numbered by 1 and 2. We draw it as



where the dashed trivalent part inside the thick topological circle is not represented. Applying Lemma 3.4 where the annulus is a neighborhood of the thick topological circle that contains the pictured trivalent vertex shows that this diagram is equivalent to



This yields the relations



◇

References

1. D. Altschuler, L. Freidel: *Vassiliev knot invariants and Chern-Simons perturbation theory to all orders*, Comm. Math. Phys. **187**, 2 (1997) pp 261–287 24, 30
2. S. Axelrod, I.M. Singer: Chern-Simons perturbation theory. In: *Proceedings of the XXth International Conference on Differential Geometric Methods in Theoretical Physics, vol. 1, 2 New York, 1991*, (World Sci. Publishing, River Edge, NJ 1992) pp 3–45 24, 45
3. S. Axelrod, I.M. Singer: *Chern-Simons perturbation theory. II*, J. Differential Geom. **39**, 1 (1994) pp 173–213 24, 45
4. D. Bar-Natan: *On the Vassiliev Knot Invariants*, Topology **34**, 2 (1995) pp 423–472 18, 40, 44
5. D. Bar-Natan: *Perturbative Chern-Simons Theory*, J. Knot Theory Ramifications **4**, 4 (1995) pp 503–548 22, 24
6. R. Bott, A.S. Cattaneo: *Integral invariants of* 3*-manifolds*, J. Differential Geom. **48**, 1 (1998) pp 91–133 45
7. R. Bott, A.S. Cattaneo: *Integral invariants of* 3*-manifolds. II*, J. Differential Geom. **53**, 1 (1999) pp 1–13 45
8. R. Bott, C. Taubes: *On the self-linking of knots*, Jour. Math. Phys. **35**, 10 (1994) pp 5247–5287 23, 24, 30
9. A.S. Cattaneo: Configuration space integrals and invariants for 3-manifolds and knots. In: *Low-dimensional topology (Funchal, 1998)*, (Contemp. Math. **233**, Amer. Math. Soc., Providence, RI, 1999) pp 153–165 45
10. S. Chmutov, S. Duzhin: *The Kontsevich integral*, Acta Appl. Math. **66**, 2 (2001) pp 155–190 44

11. W. Fulton, R. MacPherson: *A compactification of configuration spaces*, Ann. of Math. (2) **139**, 1 (1994) pp 183–225 25
12. S. Garoufalidis, M. Goussarov, M. Polyak: *Calculus of clovers and finite type invariants of 3-manifolds*, Geometry and Topology **5**, 3 (2001) pp 75–108 44
13. C. McA. Gordon, J. Luecke: *Knots are determined by their complements*, J. Amer. Math. Soc. **2**, 2 (1989) pp 371–415 49
14. E. Guadagnini, M. Martellini, M. Mintchev: *Chern-Simons field theory and link invariants*, Nucl.Phys **B330** (1990) pp 575–607 22, 24
15. K. Habiro: *Claspers and finite type invariants of links*, Geometry and Topology **4** (2000) pp 1–83 44
16. J. Hempel: 3-*Manifolds*, Ann. of Math. Studies **86**, (Princeton University Press 1976) 49
17. M. Hirsch: *Differential topology*, GTM, (Springer-Verlag 1976) 2, 3, 46, 49
18. M. Kontsevich: Jacobi diagrams and low-dimensional topology. In: *First European Congress of Mathematics **II***, (Birkhäuser Basel 1994) pp 97–121 24
19. M. Kontsevich: *Vassiliev's knot invariants*, Adv. in Sov. Math **16**, 2 (1993) pp 137–150 24
20. A. Kricker: *Alexander-Conway limits of many Vassiliev weight systems*, J. Knot Theory Ramifications **6**, 5 (1997) pp 687–714 43
21. G. Kuperberg, D.P. Thurston: *Perturbative 3-manifold invariants by cut-and-paste topology*, preprint, math.GT/9912167 45
22. T.T.Q. Le: *An invariant of integral homology 3-spheres which is universal for all finite type invariants*, Buchstaber, V. M. (ed.) et al., Solitons, geometry, and topology: on the crossroad. Providence, RI: American Mathematical Society, (ISBN 0-8218-0666-1/hbk). Transl., Ser. 2, Am. Math. Soc. **179**, 33 (1997) pp 75–100 45
23. C. Lescop: *Introduction to the Kontsevich Integral of Framed Tangles*, June 1999, Summer School in Grenoble, http://www-fourier.ujf-grenoble.fr/~lescop/ 44
24. C. Lescop: *About the uniqueness of the Kontsevich Integral*, J. Knot Theory Ramifications. **11**, 5 (2002) pp 759–780 44, 45
25. T.T.Q. Le, J. Murakami: *The universal Vassiliev-Kontsevich invariant for framed oriented links*, Compositio Mathematica **102**, (1996) pp 41–64 44
26. T.T.Q. Le, J. Murakami, T. Ohtsuki: *On a universal perturbative invariant of 3-manifolds*, Topology **37**, 3 (1998) pp 539–574 45
27. W.B.R. Lickorish: *An Introduction to Knot Theory*, GTM, (Springer-Verlag 1997) 48, 51
28. J. Milnor: *Topology from the Differentiable Viewpoint*, (The University Press of Virginia 1965, 1969) 21, 46
29. T. Ohtsuki: *Finite Type Invariants of Integral Homology 3-spheres*, J. Knot Theory Ramifications **5**, (1996) pp 101–115 45
30. S. Poirier: *The configuration space integral for links and tangles in* R^3, Algebraic and Geometric Topology **2**, 40 (2002) pp 1001–1050 25, 30, 36
31. K. Reidemeister: *Elementare Begründung der Knotentheorie*, Abh. math. Semin. Hamburg Univ. Bd. **5**, (1926) 3
32. D. Rolfsen: *Knots and links*, (Publish or Perish, Berkeley 1976) 1, 48
33. T. Stanford: *Vassiliev invariants and knots modulo pure braid subgroups*, math.GT/9805092 44
34. D. Thurston: *Integral Expressions for the Vassiliev Knot Invariants*, math.QA/9901110 24, 30, 35

35. P. Vogel: *Algebraic structures on modules of diagrams*, preprint 1997, http://www.math.jussieu.fr/~vogel/ 18, 45
36. F. Waldhausen: *On irreducible* 3*-manifolds which are sufficiently large*, Ann. of Math.**87**, 2 (1968) pp 56–88 49

More references on finite type invariants may be found in Dror Bar-Natan's bibliography on Vassiliev invariants at http://www.ma.huji.ac.il/~drorbn

Euclidean Quantum Field Theory on Commutative and Noncommutative Spaces

R. Wulkenhaar

Max-Planck-Institute for Mathematics in the Sciences, Inselstraße 22-26, 04103 Leipzig, Germany
raimar.wulkenhaar@mis.mpg.de

Summary. I give an introduction to Euclidean quantum field theory from the point of view of statistical physics, with emphasis both on Feynman graphs and on the Wilson-Polchinski approach to renormalisation. In the second part I discuss attempts to renormalise quantum field theories on noncommutative spaces.

1 From Classical Actions to Lattice Quantum Field Theory

1.1 Introduction

Ignoring gravity, space-time is described by Minkowski space given by the metric $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$. In particular, time plays a very different rôle than space. Looking at a classical field theory modelled on Minkowski space, the resulting field equations are hyperbolic ones. The formulation of the associated quantum field theory requires a sophisticated mathematical machinery. The classical reference is [1]. A comprehensive treatment can be found in [2].

From our point of view, it is much easier for a beginner to first study quantum field theory in Euclidean space E_4 given by the metric $g_{\mu\nu} = \delta_{\mu\nu} = \text{diag}(1, 1, 1, 1)$. Euclidean quantum field theory is more than just a bad trick. It has a physical interpretation as a spin system treated in the language of statistical mechanics [3, 4]. Applications to physical models are treated in [5]. There are rigorous theorems which under certain conditions allow to translate quantities computed within Euclidean quantum field theory to the Minkowskian version [6]. Eventually, from a practical point of view, computations of phenomenological relevance are almost exclusively performed in the Euclidean situation, making use of the possibility to translate them into the Minkowskian world. Our presentation of the subject is inspired by [7].

1.2 Classical Action Functionals

The starting point for both classical and quantum field theories are *action functionals*. We ignore topological questions and regard all fields as smooth (and integrable) functions on the Euclidean space E_4. The most important action functionals are the following ones.

R. Wulkenhaar: *Euclidean Quantum Field Theory on Commutative and Noncommutative Spaces*, Lect. Notes Phys. **668**, 59–100 (2005)
www.springerlink.com

– The real scalar field ϕ:

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

Here, m is the mass and λ the coupling constant. We use $\partial_\mu := \frac{\partial}{\partial x^\mu}$ and raise or lower indices with the metric tensor $g^{\mu\nu} = \delta^{\mu\nu}$ or $g_{\mu\nu} = \delta_{\mu\nu}$, such as in $\partial^\mu = g^{\mu\nu}\partial_\nu$. Summation over the same upper and lower greek index from 1 to 4 is self-understood (Einstein's sum convention).

– The Maxwell action for the electromagnetic field $A = \{A_\mu\}_{\mu=1,\dots,4}$ (the photon):

$$S[A] = \int d^4x\; \frac{1}{4g^2} F_{\mu\nu}F^{\mu\nu} , \qquad F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu , \tag{2}$$

where g is the electron charge. That action is invariant under a gauge transformation $A_\mu \mapsto A_\mu + \partial_\mu f$ for any smooth function f.

– The Dirac action for a spinor field (electron) ψ coupled to the electromagnetic field:

$$S[\psi, A] = \int d^4x\; \langle \psi, \mathrm{i}\gamma^\mu(\partial_\mu - \mathrm{i}A_\mu)\psi\rangle . \tag{3}$$

We regard the electron pointwise as $\psi(x) \in \mathbb{C}^4$ to be multiplied by the traceless (4×4)-matrices γ^μ which satisfy $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}1_{4\times 4}$. By $\langle\ ,\ \rangle$ we understand the scalar product in $\mathbb{C}^4$.

– There are matrix versions $\mathrm{i}A_\mu(x) \in \mathrm{su}(n)$ (gluon field) with $F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu - \mathrm{i}[A_\mu, A_\nu]$ of (2) where additionally the matrix trace must be taken. There is also a corresponding $\mathrm{su}(n)$-generalisation of (3).

The importance of action functionals in classical field theory is that they give rise to the equations of motion: A field configuration which satisfies the equation of motion minimases the action functional (Hamilton's principle). For example, to get the equation of motion for the electromagnetic field we vary (2) with respect to A and put the variation to zero:

$$0 = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\Big(S[A + \epsilon\tilde{A}] - S[A] \Big) = \int d^4x\; \frac{1}{g^2}\partial_\mu\tilde{A}_\nu\, F^{\mu\nu} . \tag{4}$$

If $\tilde{A}$ vanishes at infinity we can integrate (4) by parts and obtain, because the variation is zero for any $\tilde{A}$, the Euclidean version of Maxwell's equation in the vacuum $\partial_\mu F^{\mu\nu} = 0$.

1.3 A Reminder of Thermodynamics

The partition function is for the action what the free energy is for a thermodynamical system. Let us consider a system characterised by discrete energy levels E_i. Since the energy is bounded from below, there will be a ground state E_0 with $E_i \geq E_0$, with equality for $i = 0$ only (we assume the ground state to be non-degenerate). At zero temperature $T = 0$ and in thermodynamical equilibrium the system will be found with an probability $p_0 = 1$ in the ground state. At a temperature $T > 0$, however, there is due to thermal fluctuations some non-vanishing probability p_i to find the system in the energy state E_i. The probability distribution is governed by the entropy[1]

$$\Sigma(p) := -\sum_i p_i \ln p_i \tag{5}$$

and the requirement that the free energy

$$F_E(p) := \sum_i p_i E_i - k_B T \Sigma(p) \tag{6}$$

is minimal in the thermodynamical equilibrium. Here, k_B denotes Boltzmann's constant. The probability distribution $p^{\min}$ which minimises (6) is used to compute *expectation values* in thermodynamic equilibrium, such as the average energy $U = \sum_i p_i^{\min} E_i > E_0$.

1.4 The Partition Function for Discrete Actions

Whereas a thermodynamical system is described by its energy levels, a field theory is governed by its action. There is a striking similarity between the energy in thermodynamics and the action in field theory in the sense that the classical configuration is given by the minimum of the energy and the action, respectively. In the same way as the entropy term leads to thermal fluctuations away from the classical configuration if the reference energy $k_B T$ is different from zero, we should expect *quantum fluctuations* away from the classical field configuration if a reference action $\hbar$ is different from zero. Assuming for the moment that in the field theory only discrete actions $S_i \geq S_0$ are realised, we expect the quantum state to be given by the probability distribution $\{p_i\}$ which minimises the "free action"

$$F_S(p) := \sum_i p_i S_i - \hbar \Sigma(p) \;. \tag{7}$$

The entropy is given by (5). In the classical case $\hbar = 0$, the principle of minimising $F_S(p)$ reduces to Hamilton's principle of the minimal action, because $\min_p F_S(p) = S_0$ with $p_i = \delta_{i0}$.

We are going to compute the minimising probability distribution $\{p_i\}$ for $\hbar > 0$. For this purpose let us consider for two probability distributions $\{p_i\}$ and $\{\pi_i\}$ with $\sum_i p_i = \sum_i \pi_i = 1$ the relative entropy

[1] We avoid the standard symbol S for the entropy because S already denotes the action.

$$\Sigma(p|\pi) := -\sum_i p_i \ln \frac{p_i}{\pi_i} \; . \tag{8}$$

One has[2] $\Sigma(p|\pi) \leq 0$ with equality only for $p_i = \pi_i$. Let us consider

$$\pi_i = Z^{-1} \mathrm{e}^{-\frac{S_i}{\hbar}} \;, \qquad Z = \sum_i \mathrm{e}^{-\frac{S_i}{\hbar}} \; . \tag{9}$$

We get

$$F_S(p) = \sum_i p_i S_i + \hbar \sum_i p_i \ln p_i = -\hbar\Sigma(p|\pi) - \hbar \ln Z \geq -\hbar \ln Z \;, \tag{10}$$

with equality for $p_i = \pi_i$ only. Thus, the probability distribution $\{p_i^{\min}\}$ which minimises (7) is the distribution (9) and the minimum is given by $F_S(p^{\min}) = -\hbar \ln Z$.

Taking more and more states with decreasing difference we achieve in the limit a continuous probability density $p(i) \geq 0$ with $\int di\; p(i) = 1$ for the action $S(i)$. In this way we get for the free action

$$\begin{aligned} F_S(p) &= \int di\; p(i) S(i) - \hbar \Sigma(p) \geq F_S(p^{\min}) = -\hbar \ln Z \;, \\ \Sigma(p) &= -\int di\; p(i) \ln p(i) \;, \\ p^{\min}(i) &= Z^{-1} \mathrm{e}^{-\frac{S(i)}{\hbar}} \;, \qquad Z = \int di\; \mathrm{e}^{-\frac{S(i)}{\hbar}} \; . \end{aligned} \tag{11}$$

We first get $i \in \mathbb{R}^+$ but rearranging the indices we can also achieve $i \in \mathbb{R}^n$.

It is tempting now to identify the index i in (11) with the field ϕ in (1). Such an identification requires $\phi \in \mathbb{R}^n$, which we achieve by a lattice approximation to (1).

1.5 Field Theory on the Lattice

The following steps bring us from Euclidean space to a finite lattice. We first pass to the 4-torus by imposing periodic boundary conditions on the field, $\phi(x_1,\ x_2,\ x_3,\ x_4) = \phi(x_1 + L,\ x_2,\ x_3,\ x_4) = \cdots = \phi(x_1,\ x_2,\ x_3,\ x_4 + L)$. Next we restrict the 4-torus to the sublattice with equidistant spacing $a = L/N$. This lattice has N^4 points. If the field varies slowly, we can approximate

[2] In order to prove (8) we consider the convex function $f(u) = u \ln u$. As such, $f(\sum_i \pi_i u_i) \leq \sum_i \pi_i f(u_i)$. [Write $\pi_1 = \alpha_1$, $\pi_j = \prod_{i=1}^{j-1}(1-\alpha_i)\alpha_j$ for $2 \leq j \leq n$ and $\pi_n = \prod_{i=1}^{n-1}(1-\alpha_i)$, with $0 \leq \alpha_i \leq 1$, and use the definition of convexity $f(\alpha u_1 + (1-\alpha)u_2) \leq \alpha f(u_1) + (1-\alpha) f(u_2)$.] Taking $u_i = p_i/\pi_i$ we get with $\sum_i p_i = 1$ and $f(1) = 0$ the desired inequality $S(p|\pi) \leq 0$. To obtain $f(\sum_i \pi_i u_i) = \sum \pi_i f(u_i)$ we need $u_i = \text{const}$, i.e. $p_i = \pi_i$ due to the normalisation $\sum_i p_i = \sum_i \pi_i = 1$.

it by its values ϕ_q at these lattice points $q \in \mathbb{Z}_N^4$. The partial derivative is approximated by the difference quotient

$$(\partial_\mu \phi)(x) \mapsto \frac{1}{a}(\delta_\mu \phi)_q := \frac{1}{a}\big(\phi_{q+e_\mu} - \phi_q\big) , \qquad (12)$$

where $q + e_\mu$ denotes the neighboured lattice point of q in the μ^{th} coordinate direction. Now the action (1) can be approximated by

$$S[\phi] = a^4 \sum_{q \in \mathbb{Z}_N^4} \left(\frac{1}{2} a^{-2} (\delta_\mu \phi)_q (\delta^\mu \phi)_q + \frac{1}{2} m^2 \phi_q^2 + \frac{\lambda}{4!} \phi_q^4 \right) . \qquad (13)$$

Introducing dimensionless fields and masses $\tilde{\phi}_q := a\phi_q$ and $\tilde{m} := am$ (the coupling constant $\lambda = \tilde{\lambda}$ is already dimensionless), we obtain for the action

$$S[\tilde{\phi}] = \sum_{q \in \mathbb{Z}_N^4} \left(\frac{1}{2} (\delta_\mu \tilde{\phi})_q (\delta^\mu \tilde{\phi})_q + \frac{1}{2} \tilde{m}^2 \tilde{\phi}_q^2 + \frac{\tilde{\lambda}}{4!} \tilde{\phi}_q^4 \right) . \qquad (14)$$

We thus have $\tilde{\phi} \in \mathbb{R}^{N^4}$ with components $\tilde{\phi}_q \in \mathbb{R}$, for $q = 1, \dots, N^4$. We can now write the free action and its minimising probability distribution as follows:

$$F_S[p^{\min}] = \int_{\mathbb{R}^{N^4}} d\tilde{\phi}\; p^{\min}(\tilde{\phi}) S[\tilde{\phi}] - \hbar \Sigma(p^{\min}) = -\hbar \ln Z , \qquad (15)$$

where $d\tilde{\phi} = \prod_{q=1}^{N^4} d\tilde{\phi}_q$ and

$$\Sigma(p^{\min}) = - \int_{\mathbb{R}^{N^4}} d\tilde{\phi}\; p^{\min}(\tilde{\phi}) \ln p^{\min}(\tilde{\phi}) , \qquad (16)$$

$$p^{\min}(\tilde{\phi}) = Z^{-1} \mathrm{e}^{-S[\tilde{\phi}]/\hbar} , \qquad Z = \int_{\mathbb{R}^{N^4}} d\tilde{\phi}\; \mathrm{e}^{-S[\tilde{\phi}]/\hbar} . \qquad (17)$$

The interesting quantities in quantum field theory are the expectation values of products of fields at n points $q_1, \dots, q_n$ in quantum mechanical equilibrium:

$$\langle \tilde{\phi}_{q_1} \dots \tilde{\phi}_{q_n} \rangle := \int_{\mathbb{R}^{N^4}} d\tilde{\phi}\; p^{\min}(\tilde{\phi})\; \tilde{\phi}_{q_1} \cdots \tilde{\phi}_{q_n} = \frac{\displaystyle\int_{\mathbb{R}^{N^4}} d\tilde{\phi}\; \tilde{\phi}_{q_1} \cdots \tilde{\phi}_{q_n}\; \mathrm{e}^{-S[\tilde{\phi}]/\hbar}}{\displaystyle\int_{\mathbb{R}^{N^4}} d\tilde{\phi}\; \mathrm{e}^{-S[\tilde{\phi}]/\hbar}} . \qquad (18)$$

These expectation values can also be regarded as the correlation functions between the n fields at lattice sites $q_1, \dots, q_n$. The expectation values are most conveniently organised by the *generating functional*

$$Z[\tilde{j}] = \int_{\mathbb{R}^{N^4}} d\tilde{\phi}\; \mathrm{e}^{-\frac{1}{\hbar}(S[\tilde{\phi}] - \sum_{q \in \mathbb{Z}_N^4} \tilde{\phi}_q \tilde{j}_q)} . \qquad (19)$$

We thus get

$$\langle\tilde{\phi}_{q_1}\dots\tilde{\phi}_{q_n}\rangle := Z[0]^{-1}\hbar^n \frac{\partial^n Z[\tilde{j}]}{\partial\tilde{j}_{q_1}\dots\partial\tilde{j}_{q_n}}\Big|_{\tilde{j}_q=0} . \tag{20}$$

Exercise 1.1. Evaluate $\langle\tilde{\phi}_{q_1}\dots\tilde{\phi}_{q_n}\rangle$ for the classical case $\hbar = 0$. Hint: Insert the minimising probability distribution into the first equation of (18). ◁

At the end we are interested in a continuum field theory. This is, in principle, achieved by a limiting process for the expectation values (20). In the first step we pass to an infinite lattice $\mathbb{Z}^4$ by taking the limit $N \to \infty$. This process is referred to as the *thermodynamic limit*. In the second step we reintroduce the lattice spacing a by inverting the steps leading from (1) to (14). This provides the lattice $(a\mathbb{Z})^4$ embedded into the Euclidean space E_4. The difficulty is then to find an a-dependence of the dimensionless mass $\tilde{m}(a)$ and coupling constant $\tilde{\lambda}(a)$ so that the limit $a \to 0$ (the *continuum limit*) of the expectation values (20) exists. In this process the correlation length ξ of the two-point function, i.e. the inverse physical mass, is kept constant. We thus arrive at well-defined expectation values $\langle\tilde{\phi}(x_1)\dots\tilde{\phi}(x_n)\rangle$ for products of continuum fields at n positions $x_1,\dots,x_n$. Since the limit $a \to 0$ for constant λ means $\frac{\xi}{a} \to \infty$, we can equivalently regard the continuum limit as sending $\xi \to \infty$ on a lattice with constant spacing a. This means that the continuum limit corresponds to a *critical point* (where the correlation length diverges) of a lattice model.

This programme to produce continuum n-point function is called *construction of a quantum field theory*. So far this was successful in two and partly in three dimensions only [8]. Since we are particularly interested in four dimensions, a different (and less rigorous) treatment is required: *perturbative renormalisation*.

2 Field Theory in the Continuum

2.1 Generating Functionals

The idea is to perform the two limits $N \to \infty$ and $a \to 0$ *formally* in the partition function, giving

$$\langle\phi(x_1)\dots\phi(x_n)\rangle = \frac{\int\mathcal{D}\phi\;\phi(x_1)\cdots\phi(x_n)\;\mathrm{e}^{-S[\phi]/\hbar}}{\int\mathcal{D}\phi\;\mathrm{e}^{-S[\phi]/\hbar}} . \tag{21}$$

Here, the "measure" $\mathcal{D}\phi$ is the formal limit of the measure $a^{N^4}d\tilde{\phi}$ as $N \to \infty$ and $a \to 0$. Again it is useful to introduce a generating functional for the n-point functions (21),

$$Z[j] := \int \mathcal{D}\phi \; \mathrm{e}^{-\frac{1}{\hbar}(S[\phi]-\int d^4x\;\phi(x)j(x))} \; . \tag{22}$$

This generating functional has a formal meaning only and it is no surprise that it will produce problems. Using functional derivatives

$$\frac{\delta F[j(y)]}{\delta j(x)} := \lim_{\epsilon\to 0}\frac{1}{\epsilon}\big(F[j(y)+\epsilon\delta(x-y)]-F[j(y)]\big) \tag{23}$$

we can rewrite (21) as

$$\langle\phi(x_1)\dots\phi(x_n)\rangle = Z[0]^{-1}\hbar^n \frac{\delta^n Z[j]}{\delta j(x_1)\dots\delta j(x_n)}\Big|_{j(x)=0} \; . \tag{24}$$

There are two other important generating functionals derived from Z. The logarithm of Z generates (as we see later) connected n-point functions,

$$W[j] = \hbar \ln Z[j] \; . \tag{25}$$

By Legendre transformation we obtain the generating functional $\Gamma[\phi_{cl}]$ of one-particle-irreducible (1PI) n-point functions. This construction goes as follows: We first define the classical field ϕ_{cl} via

$$\phi_{cl}(x) := \frac{\delta W[j]}{\delta j(x)} \; . \tag{26}$$

Then $\Gamma[\phi_{cl}]$, which is also referred to as the effective action, is defined as

$$\Gamma[\phi_{cl}] := \int d^4x \; \phi_{cl}(x)j(x) - W[j] \; , \tag{27}$$

where $j(x)$ has to be replaced by the inverse solution of (26).

2.2 Perturbative Solution

The (perturbative) evaluation of (22) is most conveniently performed in momentum space obtained by Fourier transformation

$$\phi(x) = \int \frac{d^4p}{(2\pi)^4}\,\mathrm{e}^{-\mathrm{i}px}\,\hat{\phi}(p) \; , \qquad \hat{\phi}(p) = \int d^4x \; \mathrm{e}^{\mathrm{i}px}\,\phi(x) \; . \tag{28}$$

The action (1) reads in momentum space

$$S[\hat{\phi}] = \int \frac{d^4p}{(2\pi)^4}\,\frac{1}{2}(p^2+m^2)\hat{\phi}(p)\hat{\phi}(-p) + S_{\text{int}}[\hat{\phi}] \; , \tag{29}$$

$$S_{\text{int}}[\hat{\phi}] = \frac{\lambda}{4!}\int \left(\prod_{i=1}^{4}\frac{d^4p_i}{(2\pi)^4}\right)(2\pi)^4\delta\left(\sum_{j=1}^{4}p_j\right)\hat{\phi}(p_1)\hat{\phi}(p_2)\hat{\phi}(p_3)\hat{\phi}(p_4) \; . \tag{30}$$

For the free scalar field defined by $S_{\text{int}} = 0$ in (29) the generating functional $Z[j]$ is easy to compute:

$$\begin{aligned} Z_{\text{free}}[\hat{j}] &:= \int \mathcal{D}\hat{\phi}\, \mathrm{e}^{-\frac{1}{\hbar}\int \frac{d^4p}{(2\pi)^4}\,\left(\frac{1}{2}(p^2+m^2)\hat{\phi}(p)\hat{\phi}(-p)-\hat{\phi}(p)\hat{j}(-p)\right)} \\ &= \int \mathcal{D}\hat{\phi}\, \mathrm{e}^{-\frac{1}{\hbar}\int \frac{d^4p}{(2\pi)^4}\,\left(\frac{1}{2}(p^2+m^2)\hat{\phi}'(p)\hat{\phi}'(-p)-\frac{1}{2}(p^2+m^2)^{-1}\hat{j}(p)\hat{j}(-p)\right)} \\ &= Z[0]\, \mathrm{e}^{\frac{1}{2\hbar}\int \frac{d^4p}{(2\pi)^4}\,(p^2+m^2)^{-1}\hat{j}(p)\hat{j}(-p)}\;, \end{aligned} \tag{31}$$

where we have abbreviated $\hat{\phi}'(p) := \hat{\phi}(p) + (p^2+m^2)^{-1}\hat{j}(p)$ and used the invariance of the measure $\mathcal{D}\hat{\phi} = \mathcal{D}\hat{\phi}'$. The generating functional of free connected n-point functions becomes

$$W_{\text{free}}[\hat{j}] = W[0] + \frac{1}{2}\int \frac{d^4p}{(2\pi)^4}\hat{j}(-p)\frac{1}{(p^2+m^2)}\hat{j}(p)\;, \quad W[0] = \hbar \ln Z[0]\;. \tag{32}$$

The momentum space n-point functions are obtained from

$$(2\pi)^4\delta\left(\sum_{i=1}^{n} p_i\right)\langle\phi(p_1)\dots\phi(p_n)\rangle = \frac{1}{Z[0]}\frac{\hbar^n\delta^n Z[\hat{j}]}{\delta\hat{j}(-p_1)\dots\delta\hat{j}(-p_n)}\Big|_{\hat{j}(p)=0}\;, \tag{33}$$

where the functional derivation in momentum space is defined by

$$\frac{\delta F[\hat{j}(p)]}{\delta\hat{j}(q)} := \lim_{\epsilon\to 0}\frac{1}{\epsilon}\big(F[\hat{j}(p)+\epsilon(2\pi)^4\delta(p-q)] - F[\hat{j}(p)]\big)\;. \tag{34}$$

Exercise 2.1. Compute the effective action $\Gamma[\hat{\phi}_{cl}] = -\int d^4p\,\hat{\phi}_{cl}(p)\hat{j}(-p) + W[\hat{j}]$ for the free scalar field in momentum space. ◁

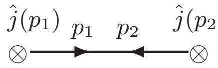
Let us now consider the full interacting ϕ^4-theory with $\lambda \neq 0$ in (30). We get formally

$$\begin{aligned} Z[\hat{j}] :&= \int \mathcal{D}\hat{\phi}\, \mathrm{e}^{-\frac{1}{\hbar}S_{\text{int}}[\hat{\phi}(q)]-\frac{1}{\hbar}\int \frac{d^4p}{(2\pi)^4}\,\left(\frac{1}{2}(p^2+m^2)\hat{\phi}(p)\hat{\phi}(-p)-\hat{\phi}(p)\hat{j}(-p)\right)} \\ &= \mathrm{e}^{-\frac{1}{\hbar}S_{\text{int}}\left[\hbar\frac{\delta}{\delta\hat{j}(-q)}\right]}\left(Z[0]\,\mathrm{e}^{\frac{1}{2\hbar}\int\frac{d^4p}{(2\pi)^4}(p^2+m^2)^{-1}\hat{j}(p)\hat{j}(-p)}\right)\;. \end{aligned} \tag{35}$$

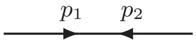
The generating functional for connected n-point function becomes

$$W[\hat{j}] = \hbar\ln\left(1 + Z_{\text{free}}[\hat{j}]^{-1}\left(\mathrm{e}^{-\frac{1}{\hbar}S_{\text{int}}\left[\hbar\frac{\delta}{\delta\hat{j}}\right]} - 1\right)Z_{\text{free}}[\hat{j}]\right) + W_{\text{free}}[\hat{j}]\;. \tag{36}$$

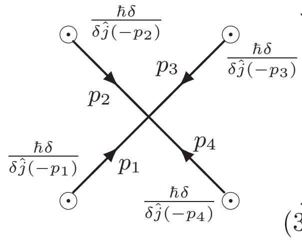
It is convenient now to introduce a graphical description for $W[\hat{j}]$. We symbolise the integrand in (32) by

$$W_{\text{free}}[\hat{j}] = \int \left(\prod_{i=1}^{2} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1+p_2) \left(\frac{1}{2} \right. $$ 

$$\left. \right) , \quad (37)$$

where the propagator 

stands for $(p_1^2+m^2)^{-1}$. The interaction part S_{int} of the action is represented by the vertex

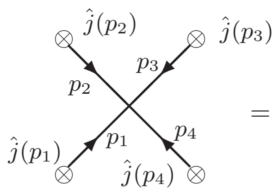
$$S_{\text{int}}\left[\hbar \frac{\delta}{\delta \hat{j}}\right] = \int \left(\prod_{i=1}^{4} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta\left(\sum_{j=1}^{4} p_j \right) \left(\right. $$ 

$$\left. \right) , \quad (38)$$

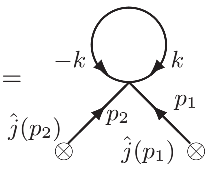
where the cross  stands for $\frac{\lambda}{4!}$. The idea is to expand in (36) both the exponential of (38) and the logarithm $\ln(1+\dots)$ into a Taylor series. In this way one obtains a formal power series in the coupling constant λ with coefficients given by *Feynman graphs*. To obtain a Feynman graph with V vertices one writes V vertices (38) next to each other and evaluates the functional derivations with respect to $\hat{j}(p)$ by their action to the $\hat{j}(q)$ in the exponent given by (37). Integrating out the resulting δ-distributions one arrives at order $V=1$ in λ at

$$\begin{aligned} W[\hat{j}]^{V=1} &= \int \left(\prod_{i=1}^{4} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1+p_2+p_3+p_4) W^{1,0}_{p_1,p_2,p_3,p_4}[\hat{j}] \\ &\quad + \int \left(\prod_{i=1}^{2} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1+p_2) W^{1,1}_{p_1,p_2}[\hat{j}] + W^{1,2} , \end{aligned} \quad (39)$$

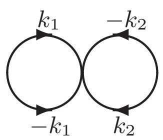
where

$$W^{1,0}_{p_1,p_2,p_3,p_4}[\hat{j}] = $$ 

$$= -\frac{\lambda}{4!} \left(\prod_{i=1}^{4} \frac{\hat{j}(p_i)}{p_i^2+m^2} \right) , \quad (40)$$

$$W^{1,1}_{p_1,p_2}[\hat{j}] = $$ 

$$= -\frac{\lambda\hbar}{4} \left(\prod_{i=1}^{2} \frac{\hat{j}(p_i)}{p_i^2+m^2} \right) \int \frac{d^4 k}{(2\pi)^4} \frac{1}{k^2+m^2} , \quad (41)$$

$$W^{1,2} = $$ 

$$= -\frac{\lambda\hbar^2}{8} \int \frac{d^4 k_1}{(2\pi)^4} \frac{1}{k_1^2+m^2} \int \frac{d^4 k_2}{(2\pi)^4} \frac{1}{k_2^2+m^2} . \quad (42)$$

Exercise 2.2. Verify (39)–(42). ◁

Exercise 2.3. Perform the Legendre transformation of the sum of (39) and (32) to $\Gamma[\phi_{cl}]^{V\leq 1}$. Compare the $\hbar$-independent part with (29) and (30). ◁

Exercise 2.4. Derive the graphical expression for $W[\hat{j}]^{V=2}$. Convince yourself that the resulting graphs are connected. ◁

Exercise 2.5. Prove that an L-loop graph contributing to $W[\hat{j}]$ leads to a factor $\hbar^L$. Hint: For a connected graph with L loops, I internal lines and V vertices the Euler characteristic reads $\chi = L - I + V = 1$. ◁

We thus deduce the following Feynman rules for the part $W[\hat{j}]^V$ of the generating functional with V vertices in a ϕ^4-model:

1. Draw the V vertices in a plane and connect in all possible ways the valences either with each other or with external sources $\hat{j}(p_i)$ such that the resulting graph is connected. The result is a sum of graphs with certain multiplicities.
2. If the graph has n sources, represent the sources and their attached lines by a factor $\prod_{i=1}^n (p_i^2 + m^2)^{-1}\hat{j}(p_i)$.
3. Represent each internal line connecting vertices by a factor $(q_j^2+m^2)^{-1}$ and determine the momenta q_j in terms of the external momenta p_i originating from the sources and L independent loop momenta k_l by the requirement that the total momentum flowing into each vertex is zero.
4. Add an integral operator $\int \prod_{l=1}^L \frac{d^4k_l}{(2\pi)^4}$ for the independent loop momenta and a factor $\int \left(\prod_{i=1}^n \frac{d^4p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1+\ldots+p_n)$ for the independent momenta of the sources. Multiply the result by a factor $\frac{(-\lambda)^V \hbar^L}{4!V!}$ for an L-loop graph with V vertices.

2.3 Calculation of Simple Feynman Graphs

When inserting $Z[\hat{j}] = \exp(W[\hat{j}]/\hbar)$ into (33), the sources $\hat{j}(p_i)$ and the integration operators $\int \frac{d^4p_i}{(2\pi)^4}$ are removed. It remains the integration over the internal momenta k_l. Due to momentum conservation the integration factorises into integrations over 1PI subgraphs. Let us thus consider an 1PI subgraph with L loops, I internal lines and E external lines. Scaling the independent loop momenta by a factor Λ, the integral will be scaled by a factor Λ^{4L-2I} for $\Lambda \gg 1$. Using Euler's formula $L = I - V + 1$ and the relation $4V = 2I + E$ (a valence of a vertex either attaches to one end of an internal line or to an external line) we get $\Lambda^{4L-2I} = \Lambda^{4-E}$. This means that the integral over the internal momenta of an 1PI graph with $E \leq 4$ external lines will be *divergent*. This divergence is due to the naïve way of performing the $N \to \infty$ and $a \to 0$ limits in the partition function (22).

In some cases (among them is the ϕ^4-model) it is possible to eliminate the divergences in a consistent way by expressing the perturbatively computed n-point functions in terms of a finite number of physically observable quantities. Such a model is called *perturbatively renormalisable*.

Let us describe the removal of divergences for the ϕ^4-model. The first step is to introduce a regulator ε which renders the integrals finite. There are many known possibilities. A common feature of these regularisations is that for dimensional reasons one also has to introduce a mass parameter μ. A very convenient regularisation is *dimensional regularisation* where the integration is performed in $4-2\varepsilon$ dimensions, $d^4k \mapsto \mu^{2\epsilon}d^{4-2\epsilon}k$. With this change of the integration one computes the generating functionals and one requires

$$\left(\frac{\delta^2\Gamma[\hat{\phi}_{cl}]}{\delta\hat{\phi}_{cl}(p_1)\delta\hat{\phi}_{cl}(p_2)}\Big|_{\hat{\phi}_{cl}=0}\right)'_{p_1=p_2=0} = m^2_{\text{phys}}\;, \tag{43}$$

$$\left(\frac{1}{2}\frac{\partial^2}{\partial p_1^2}\left(\frac{\delta^2\Gamma[\hat{\phi}_{cl}]}{\delta\hat{\phi}_{cl}(p_1)\delta\hat{\phi}_{cl}(p_2)}\Big|_{\hat{\phi}_{cl}=0}\right)'\right)_{p_1=-p_2=0} = 1\;, \tag{44}$$

$$\left(\frac{\delta^4\Gamma[\hat{\phi}_{cl}]}{\delta\hat{\phi}_{cl}(p_1)\delta\hat{\phi}_{cl}(p_2)\delta\hat{\phi}_{cl}(p_3)\delta\hat{\phi}_{cl}(p_4)}\Big|_{\hat{\phi}_{cl}=0}\right)'_{p_i=0} = \lambda_{\text{phys}}\;. \tag{45}$$

By $(\;)'$ we mean that the factor $(2\pi)^4\delta(p_1+\cdots+p_n)$ is omitted.

This means that the parts of the effective action which correspond to the mass, the amplitude of the kinetic term and the coupling constant are *normalised* to their physical values. The original parameters m, g and an additional wavefunction renormalisation factor $\mathcal{Z}$ are expressed in terms of m_{phys}, g_{phys}, ϵ and μ via the normalisation conditions (43)–(45). To prove renormalisability of the ϕ^4-model amounts to show that after that replacement the limit $\varepsilon \to 0$ of the n-point functions exists.

We will discuss in Sect. 3 another (more efficient) way to prove renormalisability. Here, we only demonstrate the method for a one-loop example. We determine $m[m_{\text{phys}}, g_{\text{phys}}, \epsilon, \mu]$ by computing the integral in (41) in dimensional regularisation.

Exercise 2.6. Prove that the surface of the sphere $x_1^2+\dots x_n^2 = 1$ equals $\frac{2\pi^{n/2}}{\Gamma(n/2)}$. Hint: compute $\int d^nx\,\mathrm{e}^{-x^2}$ both in cartesian and radial coordinates. ◁

Using the Schwinger trick $\frac{1}{A^n} = \frac{1}{\Gamma(n)}\int_0^\infty d\alpha\,\alpha^{n-1}\mathrm{e}^{-\alpha A}$ we have in $4-2\varepsilon$ dimensions

$$\begin{aligned}\int \frac{d^{4-2\varepsilon}k}{(2\pi)^4}\frac{\mu^{2\varepsilon}}{k^2+m^2} &= \frac{2\pi^{2-\varepsilon}\mu^{2\varepsilon}}{(2\pi)^4\Gamma(2-\varepsilon)}\int_0^\infty k^{3-2\varepsilon}dk\int_0^\infty d\alpha\;\mathrm{e}^{-\alpha(k^2+m^2)}\\ &= \frac{\pi^{2-\varepsilon}\mu^{2\varepsilon}}{(2\pi)^4\Gamma(2-\varepsilon)}\int_0^\infty u^{(1-\varepsilon)}du\int_0^\infty d\alpha\;\mathrm{e}^{-\alpha(u+m^2)}\end{aligned}$$

$$
\begin{aligned}
&= \frac{\pi^{2-\varepsilon}\mu^{2\varepsilon}}{(2\pi)^4 \Gamma(2-\varepsilon)} \Gamma(2-\varepsilon) \int_0^\infty \frac{d\alpha}{\alpha^{2-\epsilon}}\, \mathrm{e}^{-\alpha m^2} \\
&= \frac{m^2}{16\pi^2}\left(\frac{\mu^2}{\pi m^2}\right)^{\varepsilon} \Gamma(\varepsilon-1) = -\frac{m^2}{16\pi^2\varepsilon}\left(\frac{\mu^2}{\pi m^2}\right)^{\varepsilon} \frac{\Gamma(1+\varepsilon)}{(1-\varepsilon)} \\
&= -\frac{m^2}{16\pi^2\varepsilon} - \frac{m^2}{16\pi^2}\left(\ln\left(\frac{\mu^2}{\pi m^2}\right) + 1 + \psi(1)\right) + \mathcal{O}(\varepsilon)\,,
\end{aligned} \tag{46}
$$

where $\psi(x) = \frac{d}{dx}\ln\Gamma(x)$. We have exchanged the order of integrations. If we now reinsert the coupling constant and $\hbar$ and pass to the 1PI function (Exercises 2.1 and 2.3) we get from (43)

$$
\begin{aligned}
\left(\frac{\delta^2 \Gamma[\hat{\phi}_{cl}]}{\delta\hat{\phi}_{cl}(p_1)\delta\hat{\phi}_{cl}(p_2)}\Big|_{\hat{\phi}_{cl}=0}\right)'_{p_1=-p_2=0} &= m^2 - \frac{m^2\lambda\hbar}{32\pi^2\varepsilon}\left(\frac{\mu^2}{\pi m^2}\right)^{\varepsilon}\frac{\Gamma(1+\varepsilon)}{(1-\varepsilon)} + \mathcal{O}(\lambda^2) \\
&\equiv m^2_{\text{phys}}\,.
\end{aligned} \tag{47}
$$

Solving the formal power series for m^2 and using $\lambda = \lambda_{\text{phys}} + \mathcal{O}(\lambda^2_{\text{phys}})$ we get

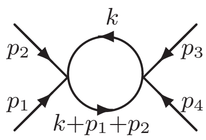
$$
m^2(\varepsilon) = m^2_{\text{phys}} + \frac{m^2_{\text{phys}}\lambda_{\text{phys}}\hbar}{32\pi^2\varepsilon}\left(\frac{\mu^2}{\pi m^2_{\text{phys}}}\right)^{\varepsilon}\frac{\Gamma(1+\varepsilon)}{(1-\varepsilon)} + \mathcal{O}(\lambda^2_{\text{phys}})\,. \tag{48}
$$

In other words, choosing the bare mass $m(\varepsilon)$ according to (48) removes the divergence of the two-point function at first order in λ_{phys}. For the treatment of subdivergences it is more convenient to perform the adjustment of m in two steps: In the first step we choose m such that the singular $\frac{1}{\varepsilon}$-term in (46) is compensated. In the second step we adjust the finite part of m to satisfy (43). Taking the limit $\varepsilon \to 0$ we get instead of (48)

$$
m^2 = m^2_{\text{phys}} + \frac{m^2_{\text{phys}}\lambda_{\text{phys}}\hbar}{32\pi^2}\left(\ln\left(\frac{\mu^2}{\pi m^2_{\text{phys}}}\right) + 1 + \psi(1)\right) + \mathcal{O}(\lambda^2_{\text{phys}})\,. \tag{49}
$$

More insight about this method is gained from Exercise 2.7. One sees that the adjustment of (45) removes the divergences from the four-point function at second order in λ_{phys} not only for zero momenta but for any momenta p_i.

Exercise 2.7. Compute the integral $\int \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+m^2)((k+p_1+p_2)^2+m^2)}$ arising in the Feynman graph

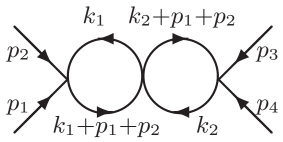


(50)

in dimensional regularisation. Hint: First bring the denominator using the Feynman trick $\frac{1}{AB} = \int_0^1 \frac{dy}{(Ay+B(1-y))^2}$ into the form $(k^2 + 2kq + r^2)^2$ and then use the Schwinger trick. Next perform the k-integration and finally the α-integration. The y-integral needs not to be computed. ◁

2.4 Treatment of Subdivergences

As part of the renormalisation process, the subtraction of divergences $\sim \frac{1}{\varepsilon^i}$ can only be carried out if these divergences appear in the n-point functions (43)–(45) for which we impose normalisation conditions. This means, in particular, that the coefficient of the $\frac{1}{\varepsilon^i}$-terms must not contain logarithms in the momenta. However, if one computes naïvely the integral corresponding to the graph



(51)

as in Exercise 2.7, one does get logarithms of momenta in front of $\frac{1}{\varepsilon}$. The solution of this problem is a different subtraction of divergences in presence of subdivergences.

The solution of this problem was found by Bogolyubov [1]. A review of the most important renormalisation schemes can be found in [9]. Instead of splitting the integral $I_{\mathcal{G}}$ associated with a Feynman graph $\mathcal{G}$ into convergent and divergent parts, there is a recursive construction of the integral to split. For a Laurent series in ε, let

$$T\left(\sum_{i=-r}^{\infty} a_i \varepsilon^i\right) := \sum_{i=-r}^{-1} a_i \varepsilon^i \tag{52}$$

be the projection to the divergent part. Then one defines for a graph $\mathcal{G}$ with disjoint subgraphs $\mathcal{G}_i$ the divergent part as

$$\mathcal{C}_{\mathcal{G}} := -T\left(\mathcal{I}_{\mathcal{G}} + \sum_{\substack{\{\mathcal{G}_1,\dots,\mathcal{G}_n\} \\ \mathcal{G}_i \in \mathcal{G},\ \mathcal{G}_i \cap \mathcal{G}_j = \emptyset}} \mathcal{C}_{\mathcal{G}_1} \cdots \mathcal{C}_{\mathcal{G}_n} \mathcal{I}_{\mathcal{G}/(\mathcal{G}_1 \cup \dots \cup \mathcal{G}_n)}\right) \tag{53}$$

and the convergent part as

$$\mathcal{R}_{\mathcal{G}} := (1-T)\left(\mathcal{I}_{\mathcal{G}} + \sum_{\substack{\{\mathcal{G}_1,\dots,\mathcal{G}_n\} \\ \mathcal{G}_i \in \mathcal{G},\ \mathcal{G}_i \cap \mathcal{G}_j = \emptyset}} \mathcal{C}_{\mathcal{G}_1} \cdots \mathcal{C}_{\mathcal{G}_n} \mathcal{I}_{\mathcal{G}/(\mathcal{G}_1 \cup \dots \cup \mathcal{G}_n)}\right) . \tag{54}$$

Here, a graph $\mathcal{G}$ is understood a the set of vertices and internal lines, and the sum runs over all sets of disjoint subgraphs. By $\mathcal{G}/(\mathcal{G}_1 \cup \cdots \cup \mathcal{G}_n)$ we mean the graph obtained by shrinking the subgraphs $\mathcal{G}_1, \dots, \mathcal{G}_n$ in $\mathcal{G}$ to a point.

Exercise 2.8. Using the result of Exercise 2.7, compute the integrals $\mathcal{I}_{\mathcal{G}}$ and $\mathcal{R}_{\mathcal{G}}$ for the graph $\mathcal{G}$ given by (51). Hint: The only subgraphs of $\mathcal{G}$ are either the left or the right one-loop subgraphs (50). ◁

There is an explicit solution of the recursion in terms of *forests* found by Zimmermann [10]. The process of passing from $\mathcal{I}_{\mathcal{G}}$ to $\mathcal{C}_{\mathcal{G}}$ and $\mathcal{R}_{\mathcal{G}}$ might seem quite unmotivated. However, as shown by Connes and Kreimer, there is the structure of a Hopf algebra behind the renormalisation process. The subtraction (53), (54) is actually a division of divergences via the antipode of the Hopf algebra, and the splitting (53), (54) is the Birkhoff decomposition solving a Riemann-Hilbert problem [11, 12].

3 Renormalisation by Flow Equations

3.1 Introduction

We have mentioned that the continuum limit of a lattice field theory corresponds to the critical point of the lattice model. One of the main tools to explore critical points in statistical physics are *renormalisation group* methods. This subject was mainly developed by Wilson [3]. A particular outcome was the understanding of renormalisation in terms of the scaling of effective actions. This idea was further developed by Polchinski to a very efficient proof that the ϕ^4-model is renormalisable to all orders in perturbation theory [13]. We refer to [14] for a textbook on this approach to renormalisation. Whereas renormalisability of the ϕ^4-model can also be proven in the previously presented Feynman graph approach, the superiority of Polchinski's method becomes manifest in the renormalisation problem of noncommutative field theories. We shall therefore present the main ideas of Polchinski's proof, following closely the original article.

3.2 Derivation of the Polchinski Equation

The starting point is a reformulation of the generating functional $Z[j]$ introduced in (22). First one brutally removes the modes[3] $\phi(p)$ with $p^2 > 2\Lambda^2$ in the measure $\mathcal{D}\phi$ of the partition function. The crucial idea is to take a *smooth cut-off* distributed over the interval $p^2 = \Lambda^2 \dots 2\Lambda^2$ which allows at a later step to differentiate with respect to the cut-off scale Λ. In this way

[3] In this section we work exclusively in momentum space so that we omit the hat over a Fourier-transformed field for simplicity. We also use natural units $\hbar = 1$.

one obtains a differential equation (the Polchinski equation) which governs the renormalisation flow of the effective action.

To be precise, we choose the cut-off function

$$K\left(\frac{p^2}{\Lambda^2}\right) = \begin{cases} 1 & \text{for } p^2 \leq \Lambda^2 \ , \\ 1 - \exp\left(-\dfrac{\exp\left(-\frac{1}{2\Lambda^2 - p^2}\right)}{p^2 - \Lambda^2}\right) & \text{for } \Lambda^2 < p^2 < 2\Lambda^2 \ , \\ 0 & \text{for } p^2 \geq 2\Lambda^2 \ . \end{cases} \tag{55}$$

We consider the generating functional

$$\begin{aligned} Z[j,\Lambda] &= \int \mathcal{D}\phi \;\exp(-S[\phi,j,\Lambda]) \ , \\ S[\phi,j,\Lambda] &:= \int \frac{d^4p}{(2\pi)^4}\left(\frac{1}{2}(p^2+m^2)K^{-1}\left(\frac{p^2}{\Lambda^2}\right)\phi(p)\phi(-p) - \phi(p)j(-p)\right) \\ &\qquad + L[\phi,\Lambda] + C[\Lambda] \ , \end{aligned} \tag{56}$$

with $L[0,\Lambda]=0$. Unless $\phi(\pm p)=0$ for $p^2 \geq 2\Lambda^2$, we have $S[\phi,j,\Lambda]=+\infty$, which means $Z[j,\Lambda]=0$. In other words, only modes $\phi(p)$ with momenta $p^2 < 2\Lambda^2$ contribute to $Z[j,\Lambda]$. Now we compute

$$\begin{aligned} &\left(2\phi(p)K^{-1}\left(\tfrac{p^2}{\Lambda^2}\right) - (p^2+m^2)^{-1}\frac{\delta S(\phi,j,\Lambda)}{\delta\phi(-p)}\right)\exp(-S[\phi,j,\Lambda]) \\ &\qquad = \left(\phi(p)K^{-1}(\tfrac{p^2}{\Lambda^2}) + (p^2+m^2)^{-1}j(p) - (p^2+m^2)^{-1}\frac{\delta L}{\delta\phi(-p)}\right) \\ &\qquad\qquad \times \exp(-S[\phi,j,\Lambda]) \ . \end{aligned} \tag{57}$$

Functional derivation with respect to $\phi(p)$ gives

$$\begin{aligned} &\frac{\delta}{\delta\phi(p)}\left\{\left(2\phi(p)K^{-1}(\tfrac{p^2}{\Lambda^2}) - (p^2+m^2)^{-1}\frac{\delta S(\phi,j,\Lambda)}{\delta\phi(-p)}\right)\exp\left(-S[\phi,j,\Lambda]\right)\right\} \\ &= \left\{\left(\phi(p)K^{-1}(\tfrac{p^2}{\Lambda^2}) + (p^2+m^2)^{-1}j(p) - (p^2+m^2)^{-1}\frac{\delta L}{\delta\phi(-p)}\right)\right. \\ &\times\left(-\phi(-p)(p^2+m^2)K^{-1}\left(\tfrac{p^2}{\Lambda^2}\right) + j(-p) - \frac{\delta L}{\delta\phi(p)}\right) \\ &+ \left.\left((2\pi)^4\delta(0)K^{-1}(\tfrac{p^2}{\Lambda^2}) - (p^2+m^2)^{-1}\frac{\delta^2 L}{\delta\phi(-p)\delta\phi(p)}\right)\right\}\exp(-S[\phi,j,\Lambda]) \ . \end{aligned} \tag{58}$$

For simplicity we choose $j(p) = 0$ for $p^2 > \Lambda^2$. This condition is not necessary, but it simplifies the following calculation considerably. We multiply (58) by $\Lambda \frac{\partial K(\frac{p^2}{\Lambda^2})}{\partial \Lambda}$, which is non-zero for $p^2 > \Lambda^2$ only, and therefore annihilates $j(\pm p)$. Next, we integrate over d^4p and finally apply the functional integration over $\mathcal{D}\phi$. This yields zero for the lhs of (58), because for each momentum p the derivative $\frac{\delta}{\delta\phi(p)}$ finds an integration over $d\phi(p)$ in the measure $\mathcal{D}\phi$ which by Stokes' theorem gives the value of the term in braces $\{\ \}$ on the lhs of (58) at the boundary $\phi(p) = \pm\infty$. But $\exp(-S[\infty, j, \Lambda]) = 0$. We thus have

$$0 = \int \mathcal{D}\phi \int \frac{d^4p}{(2\pi^4)} \Big\{ \phi(p)\phi(-p)(p^2+m^2)\Lambda \frac{\partial K^{-1}(\frac{p^2}{\Lambda^2})}{\partial \Lambda} + (2\pi)^4 \delta(0) K^{-1}(\tfrac{p^2}{\Lambda^2}) \Lambda \frac{\partial K(\frac{p^2}{\Lambda^2})}{\partial \Lambda} + \frac{1}{p^2+m^2} \Lambda \frac{\partial K(\frac{p^2}{\Lambda^2})}{\partial \Lambda} \Big(\frac{\delta L}{\delta\phi(p)} \frac{\delta L}{\delta\phi(-p)} - \frac{\delta^2 L}{\delta\phi(p)\delta\phi(-p)} \Big) \Big\} \exp(-S[\phi, j, \Lambda]) . \tag{59}$$

On the other hand, differentiating $Z[j, \Lambda]$ in (56) with respect to Λ we have

$$\Lambda \frac{\partial Z}{\partial \Lambda} = -\int \mathcal{D}[\phi] \Big\{ \Lambda \frac{\partial C}{\partial \Lambda} + \Lambda \frac{\partial L}{\partial \Lambda} + \int \frac{d^4p}{(2\pi)^4} \Big(\frac{1}{2} \phi(p)\phi(-p)(p^2+m^2) \Lambda \frac{\partial}{\partial \Lambda} K^{-1}(\tfrac{p^2}{\Lambda^2}) \Big) \Big\} \exp(-S[\phi, j, \Lambda]) . \tag{60}$$

Inserting (59) into (60) we arrive at

$$\Lambda \frac{\partial Z[j,\Lambda]}{\partial \Lambda} = 0 \qquad \text{if} \tag{61}$$

$$\Lambda \frac{\partial C[\Lambda]}{\partial \Lambda} = \frac{1}{2} \int \frac{d^4p}{(2\pi)^4} \Big((p^2+m^2)^{-1} \Lambda \frac{\partial K(\frac{p^2}{\Lambda^2})}{\partial \Lambda} \frac{\delta^2 L}{\delta\phi(p)\delta\phi(-p)} \Big|_{\phi=0} - (2\pi)^4 \delta(0)\, K^{-1}(\tfrac{p^2}{\Lambda^2}) \Lambda \frac{\partial K(\frac{p^2}{\Lambda^2})}{\partial \Lambda} \Big) , \tag{62}$$

$$\Lambda \frac{\partial L[\phi,\Lambda]}{\partial \Lambda} = -\frac{1}{2} \int \frac{d^4p}{(2\pi)^4} (p^2+m^2)^{-1} \Lambda \frac{\partial K(\frac{p^2}{\Lambda^2})}{\partial \Lambda} \Big\{ \frac{\delta L}{\delta\phi(p)} \frac{\delta L}{\delta\phi(-p)} - \Big[\frac{\delta^2 L}{\delta\phi(p)\delta\phi(-p)} \Big]_\phi \Big\} , \tag{63}$$

where $\big[F[\phi]\big]_\phi := F[\phi] - F[0]$.

We learn that the effect of restricting the integration modes in the partition function is undone if the effective action $L[\phi, \Lambda]$ and the vacuum energy $C[\Lambda]$ depend according to (63) and (62) on Λ. This means that instead of the original generating functional $Z[j] = Z[j, \infty]$ we can equally well work with $Z[j, \Lambda]$ for finite Λ. If one computes the Feynman rules from $Z[j, \Lambda]$, one finds that the propagator is given by $K(\frac{p^2}{\Lambda^2})(p^2 + m^2)^{-1}$ and the vertices by the expansion coefficients of $L[\phi, \Lambda]$. Since the loop integrations in these Feynman graphs have a finite range, we obtain finite n-point functions if the effective action $L[\phi, \Lambda]$ which evolves from $L[\phi, \infty]$ via the flow (63) is bounded. In other words, the problem to renormalise a quantum field theory is reduced to the proof that the renormalisation flow described by (63) does not produce singularities when starting from appropriate boundary conditions.

3.3 The Strategy of Renormalisation

If one naïvely integrates (63) from $L[\phi, \infty] = S_{int}[\phi]$ given by (30) down to Λ one will encounter the same divergences as found e.g. in (46) and Exercise 2.7, which must be removed by the normalisation similar to (43)–(45). The idea is thus to integrate in a first step the differential equation (63) between two scales $\Lambda_R \leq \Lambda \leq \Lambda_0$ where the initial values ρ^0 for L at Λ_0 are adjusted such that the distinguished functions (43)–(45) take given values at Λ_R.

Let us expand L into field monomials,

$$L[\phi, \Lambda, \Lambda_0, \rho^0] = \sum_{m=1}^{\infty} \frac{1}{(2m)!} \int \left(\prod_{i=1}^{2m} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1 + \cdots + p_{2m}) \times L_{2m}(p_1, \ldots, p_{2m}; \Lambda, \Lambda_0, \rho^0)\, \phi(p_1) \cdots \phi(p_{2m}) \;, \quad (64)$$

keeping the symmetry $\phi \mapsto -\phi$. The dependence on the initial conditions ρ_i^0 at $\Lambda = \Lambda_0$ is written explicitly. Thus, (63) becomes an infinite system of coupled differential equations for the amplitudes $L_{2m}(p_1, \ldots, p_{2m}; \Lambda, \Lambda_0, \rho^0)$ which at $\Lambda = \Lambda_0$ are parametrised by initial conditions ρ_i^0. Anticipating renormalisability we choose the initial condition

$$\begin{aligned} L[\phi, \Lambda_0, \Lambda_0, \rho^0] &= \int \frac{d^4 p}{(2\pi)^4} \left(\frac{1}{2} \rho_1^0 \phi(p)\phi(-p) + \frac{1}{2} p^2 \rho_2^0 \phi(p)\phi(-p) \right) \\ &+ \frac{1}{4!} \int \left(\prod_{i=1}^{4} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1 + \cdots + p_4)\, \rho_3^0\, \phi(p_1)\phi(p_2)\phi(p_3)\phi(p_4) \;. \end{aligned} \quad (65)$$

The evolution of $L[\phi, \Lambda]$ according to (63) will produce more complicated interactions than (65). Among these interactions we distinguish the same Taylor coefficients as in (65):

$$\begin{aligned}\rho_1[\Lambda,\Lambda_0,\rho^0] &:= L_2(0,0;\Lambda,\Lambda_0,\rho^0)\;,\\ \rho_2[\Lambda,\Lambda_0,\rho^0] &:= \frac{1}{2}\frac{\partial^2}{\partial p^2}L_2(p,-p;\Lambda,\Lambda_0,\rho^0)\Big|_{p=0}\;,\\ \rho_3[\Lambda,\Lambda_0,\rho^0] &:= L_4(0,0,0,0;\Lambda,\Lambda_0,\rho^0)\;,\end{aligned}\tag{66}$$

with

$$\rho_i[\Lambda_0,\Lambda_0,\rho^0]=\rho_i^0\;,\qquad i=1,2,3\;.\tag{67}$$

At the end we are interested in the limit $\Lambda_0\to\infty$. This limit requires a carefully chosen Λ_0-dependence of the initial data $\rho_i^0[\Lambda_0]$ such that $\rho_i[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]$ take the given normalised values. We consider the identity

$$\begin{aligned}&L[\phi,\Lambda_R,\Lambda_0',\rho^0[\Lambda_0']]-L[\phi,\Lambda_R,\Lambda_0'',\rho^0[\Lambda_0'']]\\ &\equiv\int_{\Lambda_0''}^{\Lambda_0'}\frac{d\Lambda_0}{\Lambda_0}\;\Lambda_0\frac{d}{d\Lambda_0}\Big(L[\phi,\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]\Big)\\ &=\int_{\Lambda_0''}^{\Lambda_0'}\frac{d\Lambda_0}{\Lambda_0}\;\Big(\Lambda_0\frac{\partial}{\partial\Lambda_0}L[\phi,\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]\\ &\qquad+\sum_{a=1}^{3}\frac{\partial}{\partial\rho_a^0}L[\phi,\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]\Lambda_0\frac{d\rho_a[\Lambda_0]}{d\Lambda_0}\Big)\;.\end{aligned}\tag{68}$$

On the other hand, we express the fact that $\rho_a[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]$ is kept fixed:

$$\begin{aligned}0=d\rho_a[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]&=\frac{\partial\rho_a[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]}{\partial\Lambda_0}d\Lambda_0\\ &\quad+\sum_{b=1}^{3}\frac{\partial\rho_a[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]}{\partial\rho_b^0}d\rho_b^0[\Lambda_0]\;.\end{aligned}\tag{69}$$

To be precise, we choose

$$\rho_1[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]=0\;,\quad\rho_2[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]=0\;,\quad\rho_3[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]=\lambda\;.\tag{70}$$

Assuming that we can invert the (3×3)-matrix $\frac{\partial\rho_a}{\partial\rho_b^0}$, which is always possible in perturbation theory, we can rewrite (69) as

$$\frac{d\rho_j^0[\Lambda_0]}{d\Lambda_0}=-\sum_{i=1}^{3}\frac{\partial\rho_j^0}{\partial\rho_i[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]}\frac{\partial\rho_i[\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]}{\partial\Lambda_0}\;.\tag{71}$$

Inserting this result into (68) we get

$$L[\phi,\Lambda_R,\Lambda_0',\rho^0[\Lambda_0']]-L[\phi,\Lambda_R,\Lambda_0'',\rho^0[\Lambda_0'']]=\int_{\Lambda_0''}^{\Lambda_0'}\frac{d\Lambda_0}{\Lambda_0}\;V[\phi,\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]]\;,\tag{72}$$

where

$$V[\phi,\Lambda,\Lambda_0,\rho^0[\Lambda_0]] := \Lambda_0 \frac{\partial L[\phi,\Lambda,\Lambda_0,\rho^0]}{\partial \Lambda_0} - \sum_{b=1}^{3} \frac{\partial L[\Lambda,\Lambda_0,\rho^0[\Lambda_0]]}{\partial \rho_b[\Lambda,\Lambda_0,\rho^0[\Lambda_0]]}\, \Lambda_0 \frac{\partial \rho_b[\Lambda,\Lambda_0,\rho^0[\Lambda_0]]}{\partial \Lambda_0}\;, \tag{73}$$

$$\frac{\partial L[\Lambda,\Lambda_0,\rho^0[\Lambda_0]]}{\partial \rho_b[\Lambda,\Lambda_0,\rho^0[\Lambda_0]]} := \sum_{a=1}^{3} \frac{\partial L[\Lambda,\Lambda_0,\rho^0[\Lambda_0]]}{\partial \rho_a^0} \frac{\partial \rho_a^0}{\partial \rho_b[\Lambda,\Lambda_0,\rho^0[\Lambda_0]]}\;. \tag{74}$$

The function V is linear in L and therefore in L_{2m} and its Taylor coefficients. Therefore, the projection of $V[\phi,\Lambda,\Lambda_0,\rho^0]$ to the initial field monomials as in (66) vanishes identically for all Λ, which means that V filters out power-counting divergent part of the effective action. It remains to show that the other coefficients of V have a Λ_0-dependence which lead to a convergence of (72) in the limit $\Lambda_0' \to \infty$.

For that purpose we compute the Λ-scaling of V. We define

$$M[V] := -\frac{1}{2}\int \frac{d^4p}{(2\pi)^4}(p^2+m^2)^{-1}\Lambda \frac{\partial K(\frac{p^2}{\Lambda^2})}{\partial \Lambda}\Big\{2\frac{\delta L}{\delta\phi(p)}\frac{\delta V}{\delta\phi(-p)} - \Big[\frac{\delta^2 V}{\delta\phi(p)\delta\phi(-p)}\Big]_\phi\Big\} \tag{75}$$

and expand M into field monomials,

$$M[V] = \sum_{m=1}^{\infty} \frac{1}{(2m)!}\int \Big(\prod_{i=1}^{2m}\frac{d^4p_i}{(2\pi)^4}\Big)(2\pi)^4\delta(p_1+\cdots+p_{2m}) \times \mathcal{M}_{2m}[V](p_1,\dots,p_{2m};\Lambda,\Lambda_0,\rho^0)\,\phi(p_1)\cdots\phi(p_{2m})\;. \tag{76}$$

As before we distinguish the coefficients

$$\begin{aligned} M_1[V] &:= \mathcal{M}_2[V](0,0;\Lambda,\Lambda_0,\rho^0)\;,\\ M_2[V] &:= \frac{1}{2}\frac{\partial^2}{\partial p^2}\mathcal{M}_2[V](p,-p;\Lambda,\Lambda_0,\rho^0)\Big|_{p=0}\;,\\ M_3[V] &:= \mathcal{M}_4[V](0,0,0,0;\Lambda,\Lambda_0,\rho^0)\;. \end{aligned} \tag{77}$$

Then one finds

$$\Lambda\frac{\partial V}{\partial\Lambda} = M[V] - \sum_{b=1}^{3}\frac{\partial L}{\partial\rho_b}M_b[V]\;, \tag{78}$$

$$\Lambda\frac{\partial}{\partial\Lambda}\Big(\frac{\partial L}{\partial\rho_b}\Big) = M\Big[\frac{\partial L}{\partial\rho_b}\Big] - \sum_{a=1}^{3}\frac{\partial L}{\partial\rho_a}M_a\Big[\frac{\partial L}{\partial\rho_b}\Big]\;, \tag{79}$$

where $M[\frac{\partial L}{\partial\rho_b}]$ arises from $M[V]$ by replacing V in (75) and (77) by $\frac{\partial L}{\partial\rho_b}$.

Exercise 3.1. Prove (78) and (79). Hint: First differentiate (73) with respect to Λ, taking into account the identity $(a^{-1})' = -(a^{-1})a'(a^{-1})$ in the $\rho[\Lambda]$-part. The Λ-derivatives act on derivatives of L or ρ with respect to Λ_0 or ρ^0. Since the derivatives commute, represent the second derivatives of L by the result of the differentiation of (63) with respect to Λ_0 or ρ^0 and the second derivatives of ρ by the projection similar to (77). Using the linearity of $M[?]$ and the ϕ-independence of the ρ-coefficients everything reassembles to (78). The proof of (79) is similar. ◁

By estimating V using the differential equation (78) and knowledge of the estimations of L and $\frac{\partial L}{\partial \rho_b}$ obtained by solving (63) and (79) before we can via (72) control the limit $\lim_{\Lambda_0\to\infty} L[\phi, \Lambda_R, \Lambda_0, \rho^0[\Lambda_0]]$.

3.4 Perturbative Solution of the Flow Equations

The evolution of the functions L, $\frac{\partial L}{\partial \rho_b}$ and V is essentially determined by the mass dimensions. We therefore absorb the dimensionality in an appropriate power of Λ. Expanding the functions also into a power series in the coupling constant we define

$$\begin{aligned} &L[\phi, \Lambda, \Lambda_0, \rho^0[\Lambda_0]] \\ &= \sum_{m=1}^{\infty} \frac{1}{(2m)!} \int \left(\prod_{i=1}^{2m} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1 + \cdots + p_{2m}) \\ &\quad \times \Lambda^{4-2m} \left(\sum_{r=1}^{\infty} \lambda^r A_{2m}^{(r)}(p_1, \dots, p_{2m}; \Lambda, \Lambda_0, \rho^0) \right) \phi(p_1) \cdots \phi(p_{2m}) , \end{aligned} \tag{80}$$

$$\begin{aligned} &\frac{\partial L}{\partial \rho_a}[\phi, \Lambda, \Lambda_0, \rho^0[\Lambda_0]] \\ &= \sum_{m=1}^{\infty} \frac{1}{(2m)!} \int \left(\prod_{i=1}^{2m} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1 + \cdots + p_{2m}) \\ &\quad \times \Lambda^{4-2m-2\delta_{a1}} \left(\sum_{r=1}^{\infty} \lambda^r B_{2m}^{a(r)}(p_1, \dots, p_{2m}; \Lambda, \Lambda_0, \rho^0) \right) \phi(p_1) \cdots \phi(p_{2m}) , \end{aligned} \tag{81}$$

$$\begin{aligned} &V[\phi, \Lambda, \Lambda_0, \rho^0[\Lambda_0]] \\ &= \sum_{m=1}^{\infty} \frac{1}{(2m)!} \int \left(\prod_{i=1}^{2m} \frac{d^4 p_i}{(2\pi)^4} \right) (2\pi)^4 \delta(p_1 + \cdots + p_{2m}) \\ &\quad \times \Lambda^{4-2m} \left(\sum_{r=1}^{\infty} \lambda^r V_{2m}^{(r)}(p_1, \dots, p_{2m}; \Lambda, \Lambda_0, \rho^0) \right) \phi(p_1) \cdots \phi(p_{2m}) . \end{aligned} \tag{82}$$

The functions $A_{2m}^{(r)}, B_{2m}^{a(r)}, V_{2m}^{(r)}$ are dimensionless. Inserting these definitions into (63), (79) and (78) one gets

$$\left(\Lambda\frac{\partial}{\partial\Lambda}+4-2m\right)A_{2m}^{(r)}(p_1,\ldots,p_{2m};\Lambda)$$
$$=\Bigg\{-\frac{1}{2}\sum_{l=1}^{m}\sum_{s=1}^{t-1}Q(P,\Lambda,m^2)A_{2l}^{(r-s)}(p_1,\ldots,p_{2l-1},-P;\Lambda)$$
$$\times A_{2m-2l+2}^{(s)}(p_{2l},\ldots,p_{2m},P;\Lambda)+\left(\binom{2m}{2l-1}-1\right)\text{ perm.}\Bigg\}$$
$$+\frac{1}{2}\int\frac{d^4p}{(2\pi\Lambda)^4}\;Q(p,\Lambda,m^2)A_{2m+2}^{(r)}(p_1,\ldots,p_{2m},p,-p;\Lambda)\;, \tag{83}$$

$$\left(\Lambda\frac{\partial}{\partial\Lambda}+4-2m-2\delta_{a1}\right)\left(B_{2m}^{a(r)}(p_1,\ldots,p_{2m};\Lambda)\right)$$
$$=\Bigg\{-\sum_{l=1}^{m}\sum_{s=0}^{r-1}Q(P,\Lambda,m^2)A_{2l}^{(r-s)}(p_1,\ldots,p_{2l-1},-P;\Lambda)$$
$$\times B_{2m-2l+2}^{a(s)}(p_{2l},\ldots,p_{2m},P;\Lambda)+\left(\binom{2m}{2l-1}-1\right)\text{ perm.}\Bigg\}$$
$$+\frac{1}{2}\int\frac{d^4p}{(2\pi\Lambda)^4}\;Q(p,\Lambda,m^2)\,B_{2m+2}^{a(r)}(p_1,\ldots,p_{2m},p,-p;\Lambda)$$
$$-\frac{1}{2}\sum_{s=0}^{r}B_{2m}^{1(r-s)}(p_1,\ldots,p_{2m};\Lambda)\;\int\frac{d^4q}{(2\pi\Lambda)^4}\;Q(q,\Lambda,m^2)\,B_4^{a(s)}(0,0,q,-q;\Lambda)$$
$$-\frac{\Lambda^2}{2}\sum_{s=0}^{r}B_{2m}^{2(r-s)}(p_1,\ldots,p_{2m};\Lambda)$$
$$\times\int\frac{d^4q}{(2\pi\Lambda)^4}\;Q(q,\Lambda,m^2)\,\frac{\partial^2B_4^{a(s)}(p,-p,q,-q;\Lambda)}{\partial p^2}\Big|_{p=0}$$
$$-\frac{1}{2}\sum_{s=0}^{r}B_{2m}^{3(r-s)}(p_1,\ldots,p_{2m};\Lambda)$$
$$\times\int\frac{d^4q}{(2\pi\Lambda)^4}\;Q(q,\Lambda,m^2)\,B_6^{a(s)}(0,0,0,0,q,-q;\Lambda)\;, \tag{84}$$

$$\left(\Lambda\frac{\partial}{\partial\Lambda}+4-2m\right)\left(V_{2m}^{(r)}(p_1,\ldots,p_{2m};\Lambda)\right)$$
$$=\Bigg\{-\sum_{l=1}^{m}\sum_{s=1}^{r-1}Q(P,\Lambda,m^2)A_{2l}^{(r-s)}(p_1,\ldots,p_{2l-1},-P;\Lambda)$$
$$\times V_{2m-2l+2}^{(s)}(p_{2l},\ldots,p_{2m},P;\Lambda)+\left(\binom{2m}{2l-1}-1\right)\text{ perm.}\Bigg\}$$
$$+\frac{1}{2}\int\frac{d^4p}{(2\pi\Lambda)^4}Q(p,\Lambda,m^2)\;V_{2m+2}^{(r)}(p_1,\ldots,p_{2m},p,-p;\Lambda)$$

$$
\begin{aligned}
&-\frac{1}{2}\sum_{s=1}^{r} B_{2m}^{1(r-s)}(p_1,\dots,p_{2m};\Lambda)\int\frac{d^4q}{(2\pi\Lambda)^4}\,Q(q,\Lambda,m^2)\,V_4^{(s)}(0,0,q,-q;\Lambda)\\
&-\frac{\Lambda^2}{2}\sum_{s=1}^{r} B_{2m}^{2(r-s)}(p_1,\dots,p_{2m};\Lambda)\\
&\qquad\qquad\times\int\frac{d^4q}{(2\pi\Lambda)^4}\,Q(q,\Lambda,m^2)\,\frac{\partial^2 V_4^{(s)}(p,-p,q,-q;\Lambda)}{\partial p^2}\Big|_{p=0}\\
&-\frac{1}{2}\sum_{s=1}^{r} B_{2m}^{3(r-s)}(p_1,\dots,p_{2m};\Lambda)\\
&\qquad\qquad\times\int\frac{d^4q}{(2\pi\Lambda)^4}\,Q(q,\Lambda,m^2)\,V_6^{(s)}(0,0,0,0,q,-q;\Lambda)\;, \qquad (85)
\end{aligned}
$$

where $Q(p,\Lambda,m^2)=\frac{1}{p^2+m^2}\Lambda^3\frac{\partial K(\frac{p^2}{\Lambda^2})}{\partial\Lambda}$ and $P:=p_1+\cdots+p_{2l-1}$. There are $\binom{2m}{2l-1}$ possibilities to assign $2l-1$ of the $2m$ momenta to the first function and the remaining ones to the second function.

Exercise 3.2. Prove (83), (84) and (85). Hint: In the last two equations a trilinear term in the functions disappears because at vanishing external momenta these functions are connected by $Q(0,\Lambda,m^2)=0$. (The support of $Q(p,\Lambda,m^2)$ is $\Lambda^2<p^2<2\Lambda^2$.) ◁

Due to the grading in the coupling constant, the differential equations (83), (84) and (85) allow us to recursively compute the functions $A_{2m}^{(r)},B_{2m}^{a(r)},V_{2m}^{(r)}$ starting from $A_4^{(1)}=1$. The concrete form is not necessary for the renormalisation proof. All we need are the norms

$$
\|f\|\equiv\|f(p_1,\dots,p_m;\Lambda)\|:=\max_{p_i^2\le 2\Lambda^2}|f(p_1,\dots,p_m;\Lambda)| \qquad (86)
$$

of these functions. The norms are computed in terms of $A_4^{(1)}=1$ and the bounds

$$
\int\frac{d^4p}{(2\pi)^4}|Q(p,\Lambda,m^2)|\le C\Lambda^4\;,\qquad \Big|\frac{\partial^n}{\partial p^n}Q(p,\Lambda,m^2)\Big|\le D_n\Lambda^{-n}\;, \qquad (87)
$$

for the propagator Q, for some constants C,D_n. Due to momentum conservation we also need a symbol $\partial^\mu_{i,j}:=\frac{\partial}{\partial p_{i\mu}}-\frac{\partial}{\partial p_{j\mu}}$ for the independent momentum derivatives.

Exercise 3.3. Verify (87). ◁

Now we can derive the estimations for the functions $A_{2m}^{(r)},B_{2m}^{a(r)},V_{2m}^{(r)}$.

Lemma 3.4.

$$
\|\partial^{\mu_1}_{i_1,j_1}\dots\partial^{\mu_d}_{i_d,j_d}A_{2m}^{(r)}(p_1,\dots,p_{2m};\Lambda)\|\begin{cases}\le\Lambda^{-d}P^{2r-m}\big[\ln\frac{\Lambda_0}{\Lambda_R}\big] & \textit{for } r{+}1\ge m\;,\\ =0 & \textit{for } r{+}1<m\;,\end{cases} \qquad (88)
$$

where $P^n[x]$ stands for a polynomial in x of degree n.

Remarks on the proof. The condition $A_{2m}^{(r)} \equiv 0$ for $r+1 < m$ is actually an additional requirement which guarantees a graphical interpretation of the A-functions: a connected graph with r vertices has at most $2r+2$ external legs. The Lemma is true for $m=2$ and $r=1$. Integrating the differential equation (83) either from Λ_0 down to Λ or from Λ_R up to Λ one obtains by induction upward in the number r of vertices and for given r downward in the number $2m$ of external legs one of the estimations

$$\begin{aligned}&\|\partial^{\mu_1}_{i_1,j_1} \dots \partial^{\mu_d}_{i_d,j_d} A_{2m}^{(r)}(p_1,\dots,p_{2m};\Lambda)\| \\ &\qquad \leq \|\partial^{\mu_1}_{i_1,j_1} \dots \partial^{\mu_d}_{i_d,j_d} A_{2m}^{(r)}(p_1,\dots,p_{2m};\Lambda_0)\| \\ &\qquad + \Lambda^{2m+d-4} \int_{\Lambda}^{\Lambda_0} d\Lambda' \Lambda'^{\,3-2m-d} P^{2r-m-1}\Big[\ln \frac{\Lambda_0}{\Lambda_R}\Big]\end{aligned} \tag{89}$$

or

$$\begin{aligned}&\|\partial^{\mu_1}_{i_1,j_1} \dots \partial^{\mu_d}_{i_d,j_d} A_{2m}^{(r)}(p_1,\dots,p_{2m};\Lambda)\| \\ &\qquad \leq \|\partial^{\mu_1}_{i_1,j_1} \dots \partial^{\mu_d}_{i_d,j_d} A_{2m}^{(r)}(p_1,\dots,p_{2m};\Lambda_R)\| \\ &\qquad + \Lambda^{2m+d-4} \int_{\Lambda_R}^{\Lambda} d\Lambda' \Lambda'^{\,3-2m-d} P^{2r-m-1}\Big[\ln \frac{\Lambda_0}{\Lambda_R}\Big]\ .\end{aligned} \tag{90}$$

For $m \geq 3$ we use (89) and the initial condition $A_{2m}^{(r)}(p_1,\dots,p_{2m};\Lambda_0) = 0$ to prove (88). In the same way we prove the (88) for $m=2, d \geq 1$. Then one uses (90) with the initial condition $A_4^{(r)}(0,0,0,0;\Lambda_R) = \delta^{r1}$ (which normalises the physical coupling constant at Λ_R to λ) to obtain $\|A_4^{(r)}(0,0,0,0;\Lambda)\| \leq P^{2r-2}\big[\ln \frac{\Lambda_0}{\Lambda_R}\big]$. The total four-point function is then reconstructed from Taylor's theorem

$$\begin{aligned}&A_4^{(2)}(p_1,p_2,p_3,p_4;\Lambda) \\ &= A_4^{(r)}(0,0,0,0;\Lambda) \\ &+ \sum_{i,j=1}^{3} p_{\mu,i}\, p_{\nu,j} \int_0^1 d\xi\,(1{-}\xi)\, \partial'^{\mu}_{i,4}\partial'^{\nu}_{j,4} A_4^{(r)}(p'_1,\dots,p'_4;\Lambda)\big|_{p'_k=\xi p_k}\ .\end{aligned} \tag{91}$$

The first derivative of $A_4^{(r)}$ at zero momentum vanishes. We thus get (88) for $m=2$. The extension to $m=1$ is similar, taking into account the initial conditions for ρ_1 and ρ_2 at Λ_0. The detailed proof is left as an exercise. □

Lemma 3.5.

$$\big\|\partial^{\mu_1}_{i_1,j_1} \dots \partial^{\mu_d}_{i_d,j_d} B_{2m}^{b(r)}(p_1,\dots,p_{2m};\Lambda)\big\| \begin{cases} \leq \Lambda^{-d} P^{2r-m+1+\delta^{b3}}\big[\ln \frac{\Lambda_0}{\Lambda_R}\big] & \textit{for } r{+}2 \geq m\ , \\ = 0 & \textit{for } r{+}2 < m\ . \end{cases} \tag{92}$$

Remarks on the proof. The first step is to derive the boundary condition

$$B^{a(r)}_{2m}(p_1,\dots,p_{2m};\Lambda_0) = \delta^{r0}\left(\delta_{m1}\delta^{b1} + \frac{p_1^2}{\Lambda_0^2}\delta_{m1}\delta^{b2} + \delta_{m2}\delta^{b3}\right) , \tag{93}$$

and the starting point $r=0$ in (92) of the induction by explicitly evaluating (74) for the initial data (65) at lowest order in the coupling constant. The further proof is similar to that of (88), with integration according to (89), apart from the problem of terms with $s=0$ and $s=r$ in the last three lines of (84). For $m \geq 3$ there is a problem with the third to last line only, which fortunately appears for $a=3$ only. One thus proves first (92) for $a=1,2$ and using this result one repeats to proof for $a=3$. Then one passes to $m=2, d\geq 1$ and again excludes $a=3$ to be processed later. The full function $B_4^{a(r)}$ is reconstructed from Taylor's theorem. The treatment of $m=1$ is similar. □

Lemma 3.6.

$$\|\partial^{\mu_1}_{i_1,j_1}\dots\partial^{\mu_d}_{i_d,j_d} V^{(r)}_{2m}(p_1,\dots,p_{2m};\Lambda)\| \begin{cases} \leq \Lambda^{-d}\left(\frac{\Lambda^2}{\Lambda_0^2}\right)P^{2r-m}\left[\ln\frac{\Lambda_0}{\Lambda_R}\right] \\ \qquad\qquad \text{for } r{+}1 \geq m\ , \\ = 0 \qquad\quad \text{for } r{+}1 < m\ . \end{cases} \tag{94}$$

Remarks on the proof. At $\Lambda=\Lambda_0$ one has

$$\begin{aligned} V[\phi,\Lambda_0,\Lambda_0,\rho^0] &:= \Lambda_0 \frac{\partial L[\phi,\Lambda,\Lambda_0,\rho^0]}{\partial\Lambda_0}\Big|_{\Lambda=\Lambda_0} \\ &\quad - \sum_{a=1}^{3}\frac{\partial L[\phi,\Lambda_0,\Lambda_0,\rho^0]}{\partial\rho^0_a}\Lambda_0\frac{\partial\rho_a[\Lambda,\Lambda_0,\rho^0]}{\partial\Lambda_0}\Big|_{\Lambda=\Lambda_0} . \end{aligned} \tag{95}$$

The result is zero for the distinguished coefficients (66). For all other interaction coefficients $\perp\rho$ we have $L[\phi,\Lambda_0,\Lambda_0,\rho^0]\big|_{\perp\rho}\equiv 0$ independent of Λ_0 and ρ^0. This means

$$\begin{aligned} 0 &= \Lambda_0\frac{\partial}{\partial\Lambda_0}L[\phi,\Lambda_0,\Lambda_0,\rho^0]\big|_{\perp\rho} \\ &= \Lambda_0\frac{\partial}{\partial\Lambda_0}L[\phi,\Lambda,\Lambda_0,\rho^0]\big|_{\perp\rho}\Big|_{\Lambda=\Lambda_0} + \Lambda\frac{\partial}{\partial\Lambda}L[\phi\Lambda,\Lambda_0,\rho^0]\big|_{\perp\rho}\Big|_{\Lambda=\Lambda_0} . \end{aligned} \tag{96}$$

Using (88) one gets for $m\leq r{+}1$

$$\|\partial^{\mu_1}_{i_1,j_1}\dots\partial^{\mu_d}_{i_d,j_d}V^{(r)}_{2m}(p_1,\dots,p_{2m};\Lambda_0)\big|_{\perp\rho}\| \leq \Lambda_0^{-d}P^{2r-m}\left[\ln\frac{\Lambda_0}{\Lambda_R}\right] . \tag{97}$$

Next one sees that for the first non-vanishing function $V_6^{(2)}(p_1,\dots,p_6;\Lambda)$ the rhs of (85) is zero so that $\Lambda^{-2}V_6^{(2)}(p_1,\dots,p_6;\Lambda) = \text{const}$. With the initial condition (97) one obtains

$$\|\partial^{\mu_1}_{i_1,j_1}\dots\partial^{\mu_d}_{i_d,j_d}V_6^{(2)}(p_1,\dots,p_{2m};\Lambda)\big|\| \leq \frac{\Lambda^2}{\Lambda_0^2}\Lambda^{-d}P^1\left[\ln\frac{\Lambda_0}{\Lambda_R}\right] . \tag{98}$$

Since (85) is a linear differential equation, the factor $\frac{\Lambda^2}{\Lambda_0^2}$ first appearing in (98) survives in all $V\big|_{\perp\rho}$ coefficients. □

Theorem 3.7. *The limit*

$$\lim_{\Lambda_0\to\infty} L^{(r)}_{2m}(p_1,\dots,p_{2m};\Lambda_R,\Lambda_0,\rho^0[\Lambda_0]) := L^{(r)}_{2m}(p_1,\dots,p_{2m};\Lambda_R,\infty)$$

exists (order by order in perturbation theory) and satisfies

$$\begin{aligned}&\big\|L^{(r)}_{2m}(p_1,\dots,p_{2m};\Lambda_R,\infty) - L^{(r)}_{2m}(p_1,\dots,p_{2m};\Lambda_R,\Lambda_0,\rho^0[\Lambda_0])\big\| \\ &\qquad\begin{cases} \leq \Lambda_R^{4-2m}\Big(\frac{\Lambda_R^2}{\Lambda_0^2}\Big)P^{2r-m}\big[\ln\frac{\Lambda_R}{\Lambda_0}\big] & \text{for } r{+}1 \geq m\,, \\ = 0 & \text{for } r{+}1 < m\,. \end{cases}\end{aligned} \tag{99}$$

Remarks on the proof. We reinsert the dimensional factors $L^{(r)}_{2m}[\Lambda_R] = \Lambda_R^{4-2m}A^{(r)}_{2m}[\Lambda_R]$. The existence of the limit and its property (99) follow from (94) inserted into (72) and Cauchy's criterion. □

Let us summarise what we have achieved. A quantum field theory is determined by an initial (classical) action S which gives rise to generating functionals $Z[j]$ for the n-point functions. Performing the integration of the generating functional gives meaningless results. Thus, one has to introduce a regularisation parameter ϵ and a mass scale μ and to fine-tune the initial action $S[\epsilon]$ in such a way that the limit $\epsilon \to 0$ for the n-point functions exists in perturbation theory. These renormalised n-point functions will now depend on μ. One possible regularisation is the momentum cut-off $p^2 \leq 2\Lambda_0^2$, where $\Lambda_0 = \mu/\epsilon$. Thus one would take a Λ_0-dependent initial action determined by the requirement that for n-point functions the limit $\Lambda_0 \to \infty$ exists in perturbation theory.

Here, a different approach is taken. We compare the cut-off theory at Λ_0 with another theory cut-off at Λ_R and require that the generating functionals of both theories coincide. This leads to a certain evolution of the initial interaction of the theory at Λ_0 to that of the other theory at Λ_R. The evolution is described by the Polchinski equation which is integrated from Λ_0 down to Λ_R. Integrating the differential equation requires the specification of initial conditions. It seems natural to take the given classical action at Λ_0 as initial condition. However, renormalisation requires a fine-tuning of the initial action at Λ_0, which gives certain parts of the initial condition at Λ_0 in terms of their normalised values at Λ_R.

In this way we obtain an effective action $L[\Lambda_R, \Lambda_0]$ for the theory at Λ_R which still depends on the initial cut-off scale Λ_0. At the end we want to send Λ_0 to ∞. It is then a rather long proof that the limit $\lim_{\Lambda_0 \to \infty} L[\Lambda_R, \Lambda_0]$ exists (in perturbation theory). However, the proof is technically very simple. All one needs are dimensional analysis and brutal majorisations, there is no need to evaluate Feynman graphs and to discuss overlapping divergences. One thus obtains a generating functional for *renormalised* n-point functions in which the propagator $(p^2 + m^2)^{-1} K(\frac{p^2}{\Lambda_R^2})$ cuts off momenta bigger than $\sqrt{2}\Lambda_R$. One still has to evaluate Feynman graphs in order to obtain the n-point functions. However, the loop momenta through the propagators are bounded so that there are no divergences any more in these n-point functions. The vertices in these Feynman graphs are given by the expansion coefficients of the effective action $L[\Lambda_R, \infty]$. In some sense, the effective action is obtained by integrating out in the partition function the fields with momenta bigger than $\sqrt{2}\Lambda_R$, avoiding the divergences by the mixed boundary conditions for the flow equation.

4 Quantum Field Theory on Noncommutative Geometries

4.1 Motivation

We have learned how the entropy term leads to quantum fluctuations about action functionals on Euclidean space E_4 and how to construct renormalised n-point functions. For suitably chosen action functionals one achieves a remarkable agreement of up to 10^{-11} between theoretical predictions derived from these n-point functions and experimental data. This shows that quantum field theories are very successful.

Unfortunately, this concept is inconsistent when taking gravity into account. The problem can not be cured by just developing the quantum field theory on a Riemannian manifold with general metric $g_{\mu\nu}$. The true problem is that combining the fundamental principles of both general relativity and quantum mechanics one concludes that space(-time) cannot be a differentiable manifold [15]. To the best of our knowledge, such a possibility was first discussed in [16].

To make this transparent, let us ask how we explore technically the geometry of space(-time). The building blocks of a manifold are the *points* labelled by coordinates $\{x^\mu\}$ in a given chart. Points enter quantum field theory via the *values* of the fields at the point labelled by $\{x^\mu\}$. This observation provides a way to “visualise” the points: we have to prepare a distribution of matter which is sharply localised around $\{x^\mu\}$. For a perfect visualisation we need a δ-distribution of the matter field. This is physically not possible, but one would think that a δ-distribution could be arbitrarily well approximated. However, that is not the case, there are limits of localisability long before the δ-distribution is reached.

Let us assume that there is a matter distribution which is believed to have two separated peaks within a space-time region R of diameter d. How do we test this conjecture? We perform a scattering experiment in the hope of finding interferences which tell us about the internal structure in the region R. We clearly need test particles of de Broglie wave length $\lambda = \frac{\hbar c}{E} \lesssim \mathrm{d}$, otherwise we can only resolve a single peak. For $\lambda \to 0$ the gravitational field of the test particles becomes important. The gravitational field created by an energy E can be measured in terms of the Schwarzschild radius

$$r_s = \frac{2G_N E}{c^4} = \frac{2G_N \hbar}{\lambda c^3} \gtrsim \frac{2G_N \hbar}{\mathrm{d} c^3} \;, \tag{100}$$

where G_N is Newton's constant. If the Schwarzschild radius r_s becomes larger than the radius $\frac{\mathrm{d}}{2}$, the inner structure of the region R can no longer be resolved (it is behind the horizon). Thus, $\frac{\mathrm{d}}{2} \geq r_s$ leads to the condition

$$\frac{\mathrm{d}}{2} \gtrsim \ell_P := \sqrt{\frac{G_N \hbar}{c^3}} \;, \tag{101}$$

which means that the Planck length ℓ_P is the fundamental length scale below of which length measurements become meaningless. Space-time cannot be a manifold.

Since geometric concepts are indispensable in physics, we need a replacement for the space-time manifold which still has a geometric interpretation. Quantum physics tells us that whenever there are measurement limits we have to describe the situation by non-commuting operators on a Hilbert space. Fortunately for physics, mathematicians have developed a generalisation of geometry, baptised noncommutative geometry [17], which is perfectly designed for our purpose. However, in physics we need more than just a better geometry: We need renormalisable quantum field theories modelled on such a noncommutative geometry.

Remarkably, it turned out to be very difficult to renormalise quantum field theories even on the simplest noncommutative spaces [18]. It would be a wrong conclusion, however, that this problem singles out the standard commutative geometry as the only one compatible with quantum field theory. The problem tells us that we are still at the very beginning of *understanding* quantum field theory. Thus, apart from curing the contradiction between gravity and quantum physics, in doing quantum field theory on noncommutative geometries we learn a lot about quantum field theory itself.

4.2 The Noncommutative $\mathbb{R}^D$

The simplest noncommutative generalisation of Euclidean space is the so-called noncommutative $\mathbb{R}^D$. Although this space arises naturally in a certain limit of string theory [19], we should not expect it to be a good model for nature. In particular, the noncommutative $\mathbb{R}^D$ does not allow for gravity. For us the main purpose of this space is to develop an understanding of quantum field theory which has a broader range of applicability.

The noncommutative $\mathbb{R}^D$, $D = 2, 4, 6, \dots$, is defined as the algebra $\mathbb{R}^D_\theta$ which as a vector space is given by the space $\mathcal{S}(\mathbb{R}^D)$ of (complex-valued) Schwartz class functions of rapid decay, equipped with the multiplication rule

$$(a \star b)(x) = \int \frac{d^D k}{(2\pi)^D} \int d^D y \; a(x + \tfrac{1}{2}\theta{\cdot}k)\, b(x+y)\, \mathrm{e}^{\mathrm{i}k\cdot y} \;, \tag{102}$$
$$(\theta{\cdot}k)^\mu = \theta^{\mu\nu} k_\nu \;, \qquad k{\cdot}y = k_\mu y^\mu \;, \qquad \theta^{\mu\nu} = -\theta^{\nu\mu} \;.$$

The entries $\theta^{\mu\nu}$ in (102) have the dimension of an area. The physical interpretation is $\|\theta\| \approx \ell_P^2$. Much information about the noncommutative $\mathbb{R}^D$ can be found in [20].

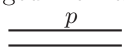
Exercise 4.1. Prove associativity $((a \star b) \star c)(x) = (a \star (b \star c))(x)$ of (102). Show that the product (102) is noncommutative, $a \star b \neq b \star a$ and that complex conjugation is an involution, $\overline{a \star b} = \overline{b} \star \overline{a}$. Show $\int d^D x \; (a \star b)(x) = \int d^D x \; a(x) b(x)$. Verify that partial derivatives are derivations, $\partial_\mu (a \star b) = (\partial_\mu a) \star b + a \star (\partial_\mu b)$. Hint: One often needs the identity $\int \frac{d^D k}{(2\pi)^D} \, \mathrm{e}^{\mathrm{i}k\cdot(x-y)} = \delta(x-y)$. ◁

Exercise 4.2. The multiplier algebra $\mathcal{M}(\mathbb{R}^D_\theta)$ consists of the distributions f which satisfy $f \star a \in \mathbb{R}^D_\theta$ and $a \star f \in \mathbb{R}^D_\theta$, for all $a \in \mathbb{R}^D_\theta$ and the same $\star$-product (102). Verify that for the coordinate functions $y^\mu \in \mathcal{M}(\mathbb{R}^D_\theta)$, $y^\mu(x) := x^\mu$, one has $\Big([y^\mu, y^\nu]_\star \star a\Big)(x) := \Big(y^\mu \star (y^\nu \star a) - y^\nu \star (y^\mu \star a)\Big)(x) = \mathrm{i}\theta^{\mu\nu} a(x)$. ◁

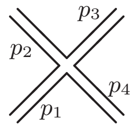
4.3 Field Theory on Noncommutative $\mathbb{R}^D$

A field theory is defined by an action functional. We obtain action functionals on $\mathbb{R}^D_\theta$ by replacing in action functionals on E_D the ordinary product of functions by the $\star$-product. For example, the noncommutative ϕ^4-action is given by

$$S[\phi] := \int d^D x \left(\frac{1}{2} \partial_\mu \phi \star \partial^\mu \phi + \frac{1}{2} m^2 \phi \star \phi + \frac{\lambda}{4!} \phi \star \phi \star \phi \star \phi \right) . \tag{103}$$

As described in Sect. 1, the entropy term leads to quantum fluctuations away from the minimum of (103). Expectation values are governed by the probability distribution which minimises the free action. A convenient way to organise the expectation values is the generating functional $Z[j]$ which is perturbatively solved by Feynman graphs. Due to $\int d^Dx\ (a\star b)(x) = \int d^Dx\ a(x)b(x)$, see Exercise 4.1, the propagator in momentum space is unchanged. For later purpose it is, however, convenient to write it as a double line, 

$= (p^2+m^2)^{-1}$. The novelty are phase factors in the vertices, which we also write in double line notation,



$$= \frac{\lambda}{4!}\,\mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_{i<j} p_i^\mu p_j^\nu \theta_{\mu\nu}}\ . \tag{104}$$

Exercise 4.3. Derive the Feynman rule (104) by repeating the steps leading to (38) for the interaction term in (103). ◁

The double line notation reflects the fact that the vertex (104) is invariant only under cyclic permutations of the legs (using momentum conservation). The resulting Feynman graphs are *ribbon graphs* which depend crucially on how the valences of the vertices are connected. For *planar graphs* the total phase factor of the integrand is independent of internal momenta, whereas *non-planar graphs* have a total phase factor which involves internal momenta. Planar graphs are integrated as usual. The resulting phase factor is precisely of the form of the original two-point function or vertex (104) so that the divergence can be removed via the normalisation conditions (43)–(45). Non-planar graphs require a separate treatment.

According to [21] there is a closed formula for the integral associated to a noncommutative Feynman graph in terms of the intersection matrices I, J, K which encode the phase factors and the incidence matrix $\mathcal{E}$. We give an orientation to each inner line l and let k_l be the momentum flowing through the line l. For each vertex v we define[4]

$$\mathcal{E}_{vl} = \begin{cases} 1 & \text{if } l \text{ leaves from } v\ , \\ -1 & \text{if } l \text{ arrives at } v\ , \\ 0 & \text{if } l \text{ is not attached to } v\ . \end{cases} \tag{105}$$

We let P_v be the total external momentum flowing into the vertex v. Restricting ourselves to 4 dimensions, an 1PI Feynman graph $\mathcal{G}$ with I internal lines and V vertices gives rise to the integral

$$\begin{aligned} \mathcal{I}_{\mathcal{G}} = \int \prod_{l=1}^{I} \frac{d^4k_l}{(k_l^2+m^2)} \prod_{v=1}^{V} (2\pi)^4 \delta\Big(P_v - \sum_{l=1}^{l} \mathcal{E}_{vl} k_l\Big) \\ \times \exp \mathrm{i}\theta_{\mu\nu}\Big(\sum_{m,n=1}^{I} I^{mn} k_m^\mu k_n^\nu + \sum_{m=1}^{I}\sum_{v=1}^{V} J^{mv} k_m^\mu P_v^\nu + \sum_{v,w=1}^{V} K^{vw} P_v^\mu P_w^\nu\Big)\ . \end{aligned} \tag{106}$$

[4] We assume that tadpoles (a line starting and ending at the same vertex) are absent.

One can show that $I^{mn}, J^{mv}, K^{vw} \in \{1, -1, 0\}$ after use of momentum conservation [22]. In terms of Schwinger parameters, this integral is evaluated to

$$\begin{aligned}\mathcal{I}_\mathcal{G} = (2\pi)^4\delta\Big(\sum_{v=1}^{V} P_v\Big)\frac{1}{16^I\pi^{2L}}\exp\Big(\mathrm{i}\theta_{\mu\nu}\sum_{v,w=1}^{V} K^{vw}P_v^\mu P_w^\nu\Big) \\ \times\int_0^\infty \prod_{l=1}^{I} d\alpha_l \,\frac{\mathrm{e}^{-\sum_{l=1}^{I}\alpha_l m^2}}{\sqrt{\det\mathcal{A}\det\mathcal{B}}}\exp\Big(-\frac{1}{4}(J\tilde{P})^T\mathcal{A}^{-1}(J\tilde{P}) \\ +\frac{1}{4}\Big(\bar{\mathcal{E}}\mathcal{A}^{-1}(J\tilde{P})+2\mathrm{i}P'\Big)^T\mathcal{B}^{-1}\Big(\bar{\mathcal{E}}\mathcal{A}^{-1}(J\tilde{P})+2\mathrm{i}P'\Big)\Big)\,, \qquad (107)\end{aligned}$$

where

$$\begin{aligned}\mathcal{A}^{mn}_{\mu\nu} &:= \alpha_m\delta^{mn}\delta_{\mu\nu} - \mathrm{i}I^{mn}\theta_{\mu\nu}\,, \qquad (J\tilde{P})^m_\mu := \sum_{v=1}^{V} J^{mv}\theta_{\mu\nu}P_v^\nu\,, \\ \bar{\mathcal{E}}^{\bar{v}l} &:= \mathcal{E}_{\bar{v}l}\,, \qquad P'^{\bar{v}}_\mu := P^\mu_{\bar{v}} \quad \text{for } \bar{v}=1,\dots,V{-}1\,, \\ \mathcal{B}^{\bar{v}\bar{w}}_{\mu\nu} &:= \sum_{m,n=1}^{I}\bar{\mathcal{E}}^{\bar{v}m}(\mathcal{A}^{-1})^{\mu\nu}_{mn}\bar{\mathcal{E}}^{\bar{w}n}\,. \qquad (108)\end{aligned}$$

The formula (107) is referred to as the parametric integral representation of a noncommutative Feynman graph.

Exercise 4.4. Verify (107). Hint: First introduce Schwinger parameters and the identity $\delta(q_v) = \int d^4y_v\ \mathrm{e}^{\mathrm{i}y_v q_v}$ for each vertex in (106). Complete the squares in k and perform the Gaußian k-integrations. Write $y_{\bar{v}} = y_V + z_{\bar{v}}$ for $\bar{v} = 1,\dots,V{-}1$ and notice that $\sum_{v=1}^{V} y_v\mathcal{E}_{vl} = \sum_{\bar{v}=1}^{V-1} z_{\bar{v}}\bar{\mathcal{E}}^{\bar{v}l}$. Then perform the y_V-integration, complete the squares for $z_{\bar{v}}$ and finally evaluate the Gaußian z_v-integrations. ◁

Possible divergences of (107) show up in the $\alpha_i \to 0$ behaviour. In order to analyse them one reparametrises the integration domain in (107), similar to the usual procedure described in [2]. For each sector

$$\alpha_{\pi_1} \le \alpha_{\pi_2} \le \dots \le \alpha_{\pi_I} \qquad \text{related to a permutation } \pi \text{ of } 1,\dots,I \qquad (109)$$

one defines $\alpha_{\pi_i} = \prod_{j=i}^{I}\beta_j^2$, with $0 \le \beta_I < \infty$ and $0 \le \beta_j \le 1$ for $j \ne I$. The leading contribution for small β_j has a topological interpretation.

A ribbon graph can be drawn on a genus-g Riemann surface with possibly several holes to which the external legs are attached [21, 23]. We say more on ribbon graphs on Riemann surfaces in Sect. 5. We explain, in particular, how a ribbon graph $\mathcal{G}$ defines a Riemann surface. On such a Riemann surface one considers *cycles*, i.e. equivalence classes of closed paths which cannot be contracted to a point. Actually one also factorises with respect to commutants, i.e. one considers the path $aba^{-1}b^{-1}$ involving two cycles a, b as trivial. We let $c_{\mathcal{G}}(\mathcal{G}_i)$ be the number of non-trivial cycles of the ribbon graph $\mathcal{G}$ wrapped by the subgraph $\mathcal{G}_i$. Next, there may exist external lines m, n such that the graph obtained by connecting m, n has to be drawn on a Riemann surface of genus $g_{mn} > g$. If this happens one declares an index $j(\mathcal{G}) = 1$, otherwise $j(\mathcal{G}) = 0$. The index extends to subgraphs by defining $j_{\mathcal{G}}(\mathcal{G}_i) = 1$ if there are external lines m, n of $\mathcal{G}$ which are already attached to $\mathcal{G}_i$ so that the line connecting m, n wraps a cycle of the additional genus $g \to g_{mn}$ of $\mathcal{G}$.

Now we can formulate the relation between the parametric integral representation and the topology of the ribbon graph. Each sector (109) of the α-parameters defines a sequence of (possibly disconnected) subgraphs $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \cdots \subset \mathcal{G}_I = \mathcal{G}$, where $\mathcal{G}_i$ is made of the i double-lines $\pi_1, \dots, \pi_i$ and the vertices to which these lines are attached. If $\mathcal{G}_i$ forms L_i loops it has a power-counting degree of divergence $\omega_i = 4L_i - 2i$. Then one has

$$\begin{aligned}\mathcal{I}_{\mathcal{G}} &= (2\pi)^4 \delta\Big(\sum_{v=1}^{V} P_v\Big) \frac{1}{8^I \pi^{2L} (\det\theta)^g} \exp\Big(\mathrm{i}\theta_{\mu\nu} \sum_{v,w=1}^{V} K^{vw} P_v^\mu P_w^\nu\Big) \\ &\times \int_0^\infty \frac{d\beta_I \, \mathrm{e}^{-\beta_I^2 m^2}}{\beta_I^{1+\omega_I - 4c_{\mathcal{G}}(\mathcal{G})}} \int_0^1 \Big(\prod_{i=1}^{I-1} \frac{d\beta_i}{\beta_i^{1+\omega_i - 4c_{\mathcal{G}}(\mathcal{G}_i)}}\Big) \\ &\qquad \times \exp\Big(-f(\pi, P) \prod_{i=1}^{I} \frac{1}{\beta^{2j_{\mathcal{G}}(\mathcal{G}_i)}}\Big) \Big(1 + \mathcal{O}(\beta^2)\Big) ,\end{aligned} \tag{110}$$

where $f(\pi, P) \geq 0$, with equality for exceptional momenta. The (very complicated) proof of (110) was given by Chepelev and Roiban [21, 23]. In order to obtain a finite integral $\mathcal{I}_{\mathcal{G}}$, we obviously need

1. $\omega_i - 4c_{\mathcal{G}}(\mathcal{G}_i) < 0$ for all i if $j(\mathcal{G}) = 0$ or $j(\mathcal{G}) = 1$ but the external momenta are exceptional,
2. $\omega_i - 4c_{\mathcal{G}}(\mathcal{G}_i) < 0$ or $j_{\mathcal{G}}(\mathcal{G}_i) = 1$ for all i if $j(\mathcal{G}) = 1$ and the external momenta are non-exceptional.

There are two types of divergences for which these conditions are violated.

First let the non-planarity be due to internal lines only, $j(\mathcal{G}) = 0$. Since the total graph $\mathcal{G}$ is non-planar, one has $c_{\mathcal{G}}(\mathcal{G}) > 0$ and therefore no superficial divergence. However, there might exist subgraphs $\mathcal{G}_i$ related to a sector of integration (109) where $\omega_i - 4c_{\mathcal{G}}(\mathcal{G}_i) \geq 0$. The standard example is a subgraph consisting of three or more *disconnected* loops wrapping the same handle of the Riemann surface. In this case the integral (107) does not exist unless one introduces a regulator. The problem is that such a subdivergence may appear in graphs with an arbitrary number of external lines. In the commutative theory this also happens, but there we renormalise already the subdivergence via the procedure described in Sect. 2.4. This procedure is based on normalisation conditions, which can only be imposed for *local* divergences. Since a non-planar graph wrapping a handle of a Riemann surface is clearly a non-local object (it cannot be reduced to a point, i.e. a counterterm vertex), it is not possible in the noncommutative case to remove that subdivergence. We are thus forced to use normalisation conditions for the total graph, but as the problem is independent of the number of external legs of the total graph, we finally need an infinite number of normalisation conditions. Hence, the model is not renormalisable in the standard way. The proposal to treat this problem is a reordering of the perturbation series [18].

The second class of problems is found in graphs where the non-planarity is at least partly due to the external legs, $j(\mathcal{G}) = 1$. This means that there is no way to remove possible divergences in these graphs by normalisation conditions. Fortunately, these graphs are superficially finite as long as the external momenta are non-exceptional. Subdivergences are supposed to be treated by a resummation. However, since the non-exceptional external momenta can become arbitrarily close to exceptional ones, these graphs are unbounded: For every $\delta > 0$ one finds non-exceptional momenta such that $\big|\langle\phi(x_1)\dots\phi(x_n)\rangle\big| > \delta^{-1}$. We present in Sect. 5 a different approach which solves all these problems.

5 Renormalisation Group Approach to Noncommutative Scalar Models

5.1 Introduction

We have seen that quantum field theories on noncommutative $\mathbb{R}^D$ are not renormalisable by standard Feynman graph evaluations. One may speculate that the origin of this problem is the too naïve way one performs the continuum limit. A way to treat the limit more carefully is the use of flow equations. We can therefore hope that applying Polchinski's method to the noncommutative ϕ^4-model we are able to prove renormalisability to all orders. There is, however, a serious problem of the momentum space proof. We have to guarantee that planar graphs only appear in the distinguished interaction coefficients for which we fix the boundary condition at Λ_R. Non-planar graphs have phase factors which involve inner momenta. Polchinski's method consists in taking norms of the interaction coefficients, and these norms ignore possible phase factors. Thus, we would find that boundary conditions for non-planar graphs at Λ_R are required. Since there is an infinite number of different non-planar structures, the model is not renormalisable in this way. A more careful examination of the phase factors is also not possible because the cut-off integrals prevent the Gaußian integration required for the parametric integral representation (107).

Fortunately, there is a matrix representation of the noncommutative $\mathbb{R}^D$ where the $\star$-product becomes a simple product of infinite matrices. The price for this simplification is that the propagator becomes complicated, but the difficulties can be overcome.

5.2 Matrix Representation

For simplicity we restrict ourselves to the noncommutative $\mathbb{R}^2$. There exists a matrix base $\{f_{mn}(x)\}_{m,n\in\mathbb{N}}$ of the noncommutative $\mathbb{R}^2$ which satisfies

$$(f_{mn}\star f_{kl})(x)=\delta_{nk}f_{ml}(x)\;,\qquad \int d^2x\; f_{mn}(x)=2\pi\theta_1\;, \tag{111}$$

where $\theta_1:=\theta_{12}=-\theta_{21}$. In terms of radial coordinates $x_1=\rho\cos\varphi$, $x_2=\rho\sin\varphi$ one has

$$f_{mn}(\rho,\varphi)=2(-1)^m\mathrm{e}^{\mathrm{i}(n-m)\varphi}\sqrt{\tfrac{m!}{n!}}\left(\sqrt{\tfrac{2\rho^2}{\theta_1}}\right)^{n-m}L_m^{n-m}\left(\tfrac{2\rho^2}{\theta_1}\right)\mathrm{e}^{-\frac{\rho^2}{\theta_1}}\;, \tag{112}$$

where $L_n^\alpha(z)$ are the Laguerre polynomials. The matrix representation was also used to obtain exactly solvable noncommutative quantum field theories [26, 27].

Exercise 5.1. Prove (111). [If you have a table of special functions you can also prove (112)]. Hint: First define $f_{00}(x_1,x_2):=2\,\mathrm{e}^{-\frac{(x_1^2+x_2^2)}{\theta_1}}$ and check $f_{00}\star f_{00}=f_{00}$. Define creation and annihilation operators $a=\frac{1}{\sqrt{2}}(x_1+\mathrm{i}x_2)$ and $\bar{a}=\frac{1}{\sqrt{2}}(x_1-\mathrm{i}x_2)$ and the corresponding derivatives $\frac{\partial}{\partial a}=\frac{1}{\sqrt{2}}(\frac{\partial}{\partial x_1}-\mathrm{i}\frac{\partial}{\partial x_2})$ and $\frac{\partial}{\partial \bar{a}}=\frac{1}{\sqrt{2}}(\frac{\partial}{\partial x_1}+\mathrm{i}\frac{\partial}{\partial x_2})$. Derive general rules for $a\star f$, $f\star a$, $\bar{a}\star f$, $f\star\bar{a}$ and prove that $f_{mn}=\frac{1}{\sqrt{m!n!\theta^{m+n}}}\bar{a}^{\star m}\star f_{00}\star a^{\star n}$ with $b^{\star n}:=b\star b^{\star(n-1)}$ satisfies (111). In order to obtain (112) one has to resolve the $\star$-product in favour of an ordinary product, pass to radial coordinates and compare the result with the definition of Laguerre polynomials. ◁

Now we can write down the noncommutative ϕ^4-action in the matrix base by expanding the field as $\phi(x)=\sum_{m,n\in\mathbb{N}}\phi_{mn}f_{mn}(x)$. It turns out, however, that in order to prove renormalisability we have to consider a more general action than (103) at the initial scale Λ_0. This action is obtained by adding a harmonic oscillator potential to the standard noncommutative ϕ^4-action:

$$\begin{aligned}S[\phi]&:=\int d^2x\bigg(\frac{1}{2}\partial_\mu\phi\star\partial^\mu\phi+\frac{1}{2}\Omega^2(\tilde{x}^\mu\phi)\star(\tilde{x}_\mu\phi)+\frac{1}{2}\mu_0^2\phi\star\phi\\&\qquad\qquad+\frac{\lambda}{4!}\phi\star\phi\star\phi\star\phi\bigg)(x)\\&=2\pi\theta_1\sum_{m,n,k,l}\left(\frac{1}{2}G_{mn;kl}\phi_{mn}\phi_{kl}+\frac{\lambda}{4!}\phi_{mn}\phi_{nk}\phi_{kl}\phi_{lm}\right)\;,\end{aligned} \tag{113}$$

where $\tilde{x}^\mu := 2(\theta^{-1})^{\mu\nu}x_\nu$ and

$$G_{mn;kl} := \int \frac{d^2x}{2\pi\theta_1}\Big(\partial_\mu f_{mn}\star\partial^\mu f_{kl} + \Omega^2(\tilde{x}^\mu f_{mn})\star(\tilde{x}_\mu f_{kl}) + \mu_0^2 f_{mn}\star f_{kl}\Big)\,. \tag{114}$$

We view Ω as a regulator and refer to the action (113) as describing a regularised ϕ^4-model. One finds

$$\begin{aligned} G_{mn;kl} &= \Big(\mu_0^2+(n+m+1)\mu^2\Big)\delta_{nk}\delta_{ml} \\ &\quad - \mu^2\sqrt{\omega\,(n+1)(m+1)}\,\delta_{n+1,k}\delta_{m+1,l} - \mu^2\sqrt{\omega\, nm}\,\delta_{n-1,k}\delta_{m-1,l}\,, \end{aligned} \tag{115}$$

where $\mu^2 = \frac{2(1+\Omega^2)}{\theta_1}$ and $\sqrt{\omega} = \frac{1-\Omega^2}{1+\Omega^2}$, with $-1 < \sqrt{\omega} \le 1$.

Exercise 5.2. Prove (115) using the formulae derived in Exercise 5.1. ◁

The kinetic matrix $G_{mn;kl}$ has the important property that $G_{mn;kl} = 0$ unless $m+k = n+l$. The same relation is induced for the propagator $\Delta_{nm;lk}$ defined by $\sum_{k,l=0}^\infty G_{mn;kl}\Delta_{lk;sr} = \sum_{k,l=0}^\infty \Delta_{nm;lk}G_{kl;rs} = \delta_{mr}\delta_{ns}$:

$$\begin{aligned} \Delta_{mn;kl} &= \frac{\delta_{m+k,n+l}}{(1+\sqrt{1-\omega})\mu^2}\sum_{v=\frac{|m-l|}{2}}^{\frac{\min(m+l,k+n)}{2}} B\Big(\tfrac{1}{2}+\tfrac{\mu_0^2}{2\sqrt{1-\omega}\,\mu^2}+\tfrac{1}{2}(m+k)-v,\,1+2v\Big) \\ &\quad\times\sqrt{\binom{n}{v+\frac{n-k}{2}}\binom{k}{v+\frac{k-n}{2}}\binom{m}{v+\frac{m-l}{2}}\binom{l}{v+\frac{l-m}{2}}}\left(\frac{(1-\sqrt{1-\omega})^2}{\omega}\right)^{v} \\ &\quad\times {}_2F_1\left(\begin{matrix}1+2v\,,\ \frac{1}{2}+\frac{\mu_0^2}{2\sqrt{1-\omega}\,\mu^2}-\frac{1}{2}(m+k)+v \\ \frac{3}{2}+\frac{\mu_0^2}{2\sqrt{1-\omega}\,\mu^2}+\frac{1}{2}(m+k)+v\end{matrix}\,\middle|\,\frac{(1-\sqrt{1-\omega})^2}{\omega}\right). \end{aligned} \tag{116}$$

Here, $B(a,b)$ is the Beta-function and $F\big(\genfrac{}{}{0pt}{}{a,\,b}{c};z\big)$ the hypergeometric function. The derivation of (116), which is performed in [28], involves Meixner polynomials [29] in a crucial way. We recall that in the momentum space version of the ϕ^4-model, the interactions contain oscillating phase factors which make a renormalisation by flow equations impossible. Here we use an adapted base which eliminates the phase factors from the interaction. We see from (116) that these oscillations do not reappear in the propagator. Note that all matrix elements $\Delta_{nm;lk}$ are non-zero for $m+k = n+l$.

5.3 The Polchinski Equation for Matrix Models

Introducing a cut-off for the matrix indices

$$\Delta^K_{nm;lk}(\Lambda) = K\Big(\tfrac{m\mu^2}{\Lambda^2}\Big)K\Big(\tfrac{n\mu^2}{\Lambda^2}\Big)K\Big(\tfrac{k\mu^2}{\Lambda^2}\Big)K\Big(\tfrac{l\mu^2}{\Lambda^2}\Big)\Delta_{nm;lk}\,, \tag{117}$$

for the same function K as in (55), one can derive in analogy to (63) the Polchinski equation in the matrix base of $\mathbb{R}^2_\theta$:

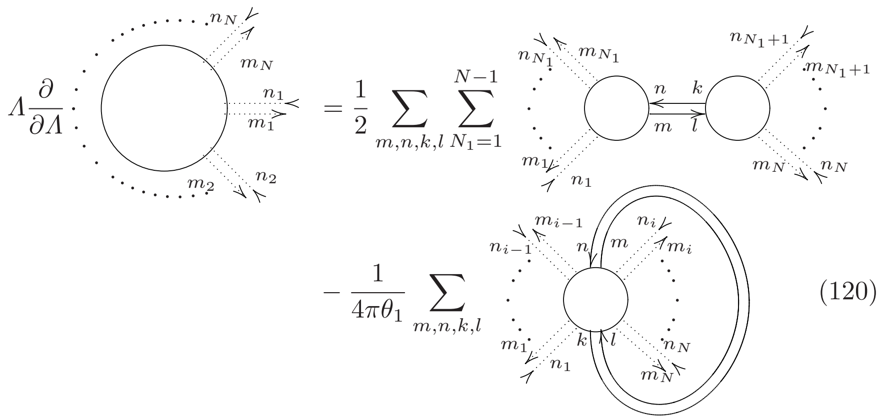
$$\Lambda\frac{\partial L[\phi,\Lambda]}{\partial\Lambda} = \sum_{m,n,k,l}\frac{1}{2}\Lambda\frac{\partial\Delta^K_{nm;lk}(\Lambda)}{\partial\Lambda}\bigg(\frac{\partial L[\phi,\Lambda]}{\partial\phi_{mn}}\frac{\partial L[\phi,\Lambda]}{\partial\phi_{kl}} - \frac{1}{2\pi\theta_1}\Big[\frac{\partial^2 L[\phi,\Lambda]}{\partial\phi_{mn}\,\partial\phi_{kl}}\Big]_\phi\bigg)\,. \tag{118}$$

Again, the differential equation (118) ensures (together with easier differential equations for functions such as C) that the partition function $Z[J,\Lambda]$ is actually independent of the cut-off Λ. This means that we can equally well evaluate the partition function for finite Λ, where it leads to Feynman graphs with vertices given by the Taylor expansion coefficients $A^{(V)}_{m_1n_1;\dots;m_Nn_N}$ in

$$L[\phi,\Lambda] = \lambda\sum_{V=1}^\infty\Big(2\pi\theta_1\lambda\Big)^{V-1}\sum_{N=2}^\infty\frac{1}{N!}\sum_{m_i,n_i}A^{(V)}_{m_1n_1;\dots;m_Nn_N}[\Lambda]\phi_{m_1n_1}\cdots\phi_{m_Nn_N}\,. \tag{119}$$

These vertices are connected with each other by internal lines $\Delta^K_{nm;lk}(\Lambda)$ and to sources j_{kl} by external lines $\Delta^K_{nm;lk}(\Lambda_0)$. Since the summation variables are cut-off in the propagator $\Delta^K_{nm;lk}(\Lambda)$, loop summations are finite, provided that the interaction coefficients $A^{(V)}_{m_1n_1;\dots;m_Nn_N}[\Lambda]$ are finite.

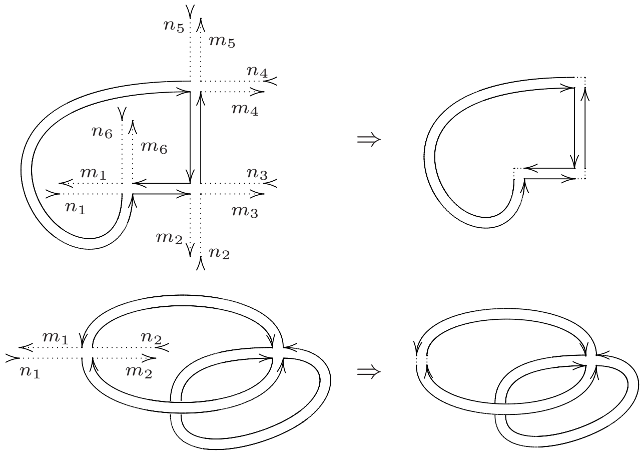
Inserting the expansion (119) into (118) and restricting to the part with N external legs we get the graphical expression



(120)

Combinatorial factors are not shown and symmetrisation in all indices m_in_i has to be performed. On the rhs of (120) the two valences mn and kl of subgraphs are connected to the ends of a *ribbon* which symbolises the differentiated propagator $\underset{m\quad l}{\overset{n\quad k}{\rightleftarrows}} = \Lambda\frac{\partial}{\partial\Lambda}\Delta^K_{nm;lk}$. We see that for the simple fact that the fields ϕ_{mn} carry two indices, the effective action is expanded into ribbon graphs.

In the expansion of L there will occur very complicated ribbon graphs with crossings of lines which cannot be drawn any more in a plane. A general ribbon graph can, however, be drawn on a *Riemann surface* of some *genus* g. In fact, a ribbon graph *defines* the Riemann surfaces topologically through the *Euler characteristic* χ. We have to regard here the external lines of the ribbon graph as amputated (or closed), which means to directly connect the single lines m_i with n_i for each external leg $m_i n_i$. A few examples may help to understand this procedure:



$$\tilde{L}=2,\quad I=3,\quad V=3,\quad g=0,\quad B=2,\quad N=6$$

$$\tilde{L}=1,\quad I=3,\quad V=2,\quad g=1,\quad B=1,\quad N=2 \tag{121}$$

The genus is computed from the number $\tilde{L}$ of single-line loops, the number I of internal (double) lines and the number V of vertices of the graph according to Euler's formula $\chi = 2-2g = \tilde{L} - I + V$. The number B of boundary components of a ribbon graph is the number of those loops which carry at least one external leg. There can be several possibilities to draw the graph and its Riemann surface, but $\tilde{L}, I, V, B$ and thus g remain unchanged. Indeed, the Polchinski equation (118) interpreted as in (120) tells us which external legs of the vertices are connected. It is completely irrelevant how the ribbons are drawn between these legs. In particular, there is no distinction between overcrossings and undercrossings.

We expect that non-planar ribbon graphs with $g > 0$ and/or $B > 1$ behave differently under the renormalisation flow than planar graphs having $B = 1$ and $g = 0$. This suggests to introduce a further grading in g, B in the interactions coefficients $A^{(V,B,g)}_{m_1n_1;\dots;m_Nn_N}$. Technically, our strategy is to apply the summations in (120) either to the propagator or the subgraph only and to maximise the other object over the summation indices. For that purpose one has to introduce further characterisations of a ribbon graph which disappear at the end, see [24].

5.4 ϕ^4-Theory on Noncommutative $\mathbb{R}^2$

First one estimates the A-functions by solving (118) perturbatively:

Lemma 5.3. *The homogeneous parts* $A^{(V,B,g)}_{m_1n_1;\dots;m_Nn_N}$ *of the coefficients of the effective action describing a regularised* ϕ^4*-theory on* $\mathbb{R}^2_\theta$ *in the matrix base are for* $2 \leq N \leq 2V{+}2$ *and* $\sum_{i=1}^N (m_i{-}n_i) = 0$ *bounded by*

$$\begin{aligned} &\big|A^{(V,B,g)}_{m_1n_1;\dots;m_Nn_N}[\Lambda,\Lambda_0,\omega,\rho_0]\big| \\ &\leq \Big(\frac{\Lambda^2}{\mu^2}\Big)^{2-V-B-2g}\Big(\frac{1}{\sqrt{1-\omega}}\Big)^{3V-\frac{N}{2}+B+2g-2} P^{2V-\frac{N}{2}}\Big[\ln\frac{\Lambda_0}{\Lambda_R}\Big]\,. \end{aligned} \tag{122}$$

We have $A^{(V,B,g)}_{m_1n_1;\dots;m_Nn_N} \equiv 0$ *for* $N > 2V{+}2$ *or* $\sum_{i=1}^N (m_i{-}n_i) \neq 0$.

The proof of (122) for general matrix models by induction goes over 20 pages! The formula specific for the ϕ^4-model on $\mathbb{R}^2_\theta$ follows from the asymptotic behaviour of the cut-off propagator (117,116) and a certain index summation, see [24, 25].

We see from (122) that the only divergent function is

$$\begin{aligned} A^{(1,1,0)}_{m_1n_1;m_2n_2} &= A^{(1,1,0)}_{00;00}\delta_{m_1n_2}\delta_{m_2n_1} \\ &\quad + \Big(A^{(1,1,0)}_{m_1n_1;m_2n_2}[\Lambda,\Lambda_0,\rho^0] - A^{(1,1,0)}_{00;00}\delta_{m_1n_2}\delta_{m_2n_1}\Big)\,, \end{aligned} \tag{123}$$

which is split into the distinguished divergent function

$$\rho[\Lambda,\Lambda_0,\rho^0] := A^{(1,1,0)}_{00;00}[\Lambda,\Lambda_0,\rho^0] \tag{124}$$

for which we impose the boundary condition $\rho[\Lambda_R,\Lambda_0,\rho^0] = 0$ and a convergent part with boundary condition at Λ_0.

One remarks that the limit $\omega \to 1$ in (122) is singular. In fact the estimation for $\omega = 1$ with an optimal choice of the ρ-coefficients (different than (124)!) would be

$$\begin{aligned} &\sum_{\mathcal{E}^s}\big|A^{(V,V^e,B,g,\iota)}_{m_1n_1;\dots;m_Nn_N}[\Lambda,\Lambda_0,1,\rho^0]\big| \\ &\leq \Big(\frac{\Lambda}{\mu}\Big)^{V-\frac{N}{2}-B-2g+2}\Big(\frac{\mu}{\mu_0}\Big)^{3V-\frac{N}{2}+B+2g-2} P^{2V-\frac{N}{2}}\Big[\ln\frac{\Lambda_0}{\Lambda_R}\Big]\,. \end{aligned} \tag{125}$$

Since the exponent of Λ can be arbitrarily large, there would be an infinite number of divergent interaction coefficients, which means that the ϕ^4-model is not renormalisable when keeping $\omega = 1$.

The limit $\Lambda_0 \to \infty$ is now governed by an identity like (72) and a ρ-subtracted function like (73) for which one has a differential equation like (78). It is then not difficult to see that the regularised ϕ^4-model with $\omega < 1$ is renormalisable. It turns out that one can even prove more [25]: On can endow the parameter ω for the oscillator frequency with an Λ_0-dependence so that in the limit $\Lambda_0 \to \infty$ one obtains a standard ϕ^4-model without the oscillator term:

Theorem 5.4. *The ϕ^4-model on $\mathbb{R}^2_\theta$ is (order by order in the coupling constant) renormalisable in the matrix base by adjusting the bare mass $\Lambda_0^2\rho[\Lambda_0]$ to give $A^{(1,1,0)}_{m_1n_1;m_2n_2}[\Lambda_R]=0$ and by performing the limit $\Lambda_0\to\infty$ along the path of regulated models characterised by $\omega[\Lambda_0]=1-(1+\ln\frac{\Lambda_0}{\Lambda_R})^{-2}$. The limit $A^{(V,B,g)}_{m_1n_1;\dots;m_Nn_N}[\Lambda_R,\infty] := \lim_{\Lambda_0\to\infty} A^{(V,B,g)}_{m_1n_1;\dots;m_Nn_N}[\Lambda_R,\Lambda_0,\omega[\Lambda_0],\rho^0[\Lambda_0]]$ of the expansion coefficients of the effective action $L[\phi,\Lambda_R,\Lambda_0,\omega[\Lambda_0],\rho^0[\Lambda_0]]$ exists and satisfies*

$$
\begin{aligned}
\Big|\lambda\Big(2\pi\theta_1\lambda\Big)^{V-1} & A^{(V,V^e,B,g,\iota)}_{m_1n_1;\dots;m_Nn_N}[\Lambda_R,\infty] \\
&-\Big(2\pi\theta_1\lambda\Big)^{V-1} A^{(V,V^e,B,g,\iota)}_{m_1n_1;\dots;m_Nn_N}\Big[\Lambda_R,\Lambda_0,1-\tfrac{1}{(1+\ln\frac{\Lambda_0}{\Lambda_R})^2},\rho^0\Big]\Big| \\
&\leq \frac{\Lambda_R^4}{\Lambda_0^2}\Big(\frac{\lambda}{\Lambda_R^2}\Big)^V\Big(\frac{\mu^2(1+\ln\frac{\Lambda_0}{\Lambda_R})}{\Lambda_R^2}\Big)^{B+2g-1} P^{5V-N-1}\Big[\ln\frac{\Lambda_0}{\Lambda_R}\Big] \,.
\end{aligned}
\tag{126}
$$

In this way we have proven that the real ϕ^4-model on $\mathbb{R}^2_\theta$ is perturbatively renormalisable when formulated in the matrix base. This proof was not simply a generalisation of Polchinski's original proof to the noncommutative case. The naïve procedure would be to take the standard ϕ^4-action at the initial scale Λ_0, with Λ_0-dependent bare mass to be adjusted such that at Λ_R it is scaled down to the renormalised mass. Unfortunately, this does not work. In the limit $\Lambda_0\to\infty$ one obtains an unbounded power-counting degree of divergence for the ribbon graphs. The solution is the observation that the cut-off action at Λ_0 is (due to the cut-off) not translation invariant. We are therefore free to break the translational symmetry of the action at Λ_0 even more by adding a harmonic oscillator potential for the fields ϕ. There exists a Λ_0-dependence of the oscillator frequency Ω with $\lim_{\Lambda_0\to\infty}\Omega=0$ such that the effective action at Λ_R is convergent (and thus bounded) order by order in the coupling constant in the limit $\Lambda_0\to\infty$. This means that the partition function of the original (translation-invariant) ϕ^4-model without cut-off and with suitable divergent bare mass can equally well be solved by Feynman graphs with propagators cut-off at Λ_R and vertices given by the bounded expansion coefficients of the effective action at Λ_R. Hence, this model is renormalisable, and in contrast to the naïve Feynman graph approach in momentum space [23] there is no problem with exceptional configurations. Whereas the treatment of the oscillator potential is easy in the matrix base, a similar procedure in momentum space will face enormous difficulties. This makes clear that the adaptation of Polchinski's renormalisation programme is the preferred method for noncommutative field theories.

5.5 ϕ^4-Theory on Noncommutative $\mathbb{R}^4$

The renormalisation of ϕ^4-theory on $\mathbb{R}^4_\theta$ in the matrix base is performed in an analogous way. We choose a coordinate system in which $\theta_1 = \theta_{12} = -\theta_{21}$ and $\theta_2 = \theta_{34} = -\theta_{43}$ are the only non-vanishing components of θ. Moreover, we assume $\theta_1 = \theta_2$ for simplicity. Then we expand the scalar field according to $\phi(x) = \sum_{m_1,n_1,m_2,n_2\in\mathbb{N}} \phi_{\substack{m_1 n_1\\ m_2 n_2}} f_{m_1n_1}(x_1,x_2) f_{m_2n_2}(x_3,x_4)$. The action (113) with integration over $\mathbb{R}^4$ leads then to a kinetic term generalising (115) and a propagator generalising (116). Using estimates on the asymptotic behaviour of that propagator one proves the four-dimensional generalisation of Lemma 5.3 on the power-counting degree of the N-point functions. For $\omega < 1$ one finds that all non-planar graphs ($B > 1$ and/or $g > 0$) and all graphs with $N \geq 6$ external legs are convergent.

The remaining infinitely many planar two- and four-point functions have to be split into a divergent ρ-part and a convergent complement. Using some sort of locality for the propagator (116) one proves that

$$
\begin{aligned}
A^{\text{planar}}_{\substack{m_1 n_1; k_1 l_1\\ m_2 n_2; k_2 l_2}} &- A^{\text{planar}}_{\substack{0\,0;0\,0\\ 0\,0;0\,0}} \delta_{m_1l_1}\delta_{n_1k_1}\delta_{m_2l_2}\delta_{n_2k_2} \\
&- m_1\Big(A^{\text{planar}}_{\substack{1\,0;0\,1\\ 0\,0;0\,0}} - A^{\text{planar}}_{\substack{0\,0;0\,0\\ 0\,0;0\,0}}\Big)\delta_{m_1l_1}\delta_{n_1k_1}\delta_{m_2l_2}\delta_{n_2k_2} \\
&- m_2\Big(A^{\text{planar}}_{\substack{0\,0;0\,0\\ 1\,0;0\,1}} - A^{\text{planar}}_{\substack{0\,0;0\,0\\ 0\,0;0\,0}}\Big)\delta_{m_1l_1}\delta_{n_1k_1}\delta_{m_2l_2}\delta_{n_2k_2} \\
&- A^{\text{planar}}_{\substack{1\,1;0\,0\\ 0\,0;0\,0}}\Big(\sqrt{(m_1{+}1)(n_1{+}1)}\delta_{m_1+1,l_1}\delta_{n_1+1,k_1}\delta_{m_2l_2}\delta_{n_2k_2} \\
&\qquad\qquad + \sqrt{m_1n_1}\delta_{m_1-1,l_1}\delta_{n_1-1,k_1}\delta_{m_2l_2}\delta_{n_2k_2}\Big) \\
&- A^{\text{planar}}_{\substack{0\,0;0\,0\\ 1\,1;0\,0}}\Big(\sqrt{(m_2{+}1)(n_2{+}1)}\delta_{m_2+1,l_2}\delta_{n_2+1,k_2}\delta_{m_1l_1}\delta_{n_1k_1} \\
&\qquad\qquad + \sqrt{m_2n_2}\delta_{m_2-1,2_1}\delta_{n_2-1,k_2}\delta_{m_1l_1}\delta_{n_1k_1}\Big)\,,
\end{aligned}
\tag{127}
$$

$$
A^{\text{planar}}_{\substack{m_1\ n_1\\ m_1'\ n_1'};\dots;\substack{m_4\ n_4\\ m_4'\ n_4'}} - A^{\text{planar}}_{\substack{0\,0\\ 0\,0};\dots;\substack{0\,0\\ 0\,0}}\Big(\frac{1}{6}\delta_{\substack{n_1\ m_2\\ n_1\ m_2}}\delta_{\substack{n_2\ m_3\\ n_2\ m_4}}\delta_{\substack{n_3\ m_4\\ n_3\ m_4}}\delta_{\substack{n_4\ m_1\\ n_4\ m_1}} + 5 \text{ perm's}\Big)\,,
\tag{128}
$$

are convergent functions, thus identifying

$$\begin{aligned}
\rho_1 &:= A^{\text{planar}}_{\substack{0\,0\,;\,0\,0\\0\,0\,;\,0\,0}}\,, \\
\rho_2 &:= A^{\text{planar}}_{\substack{1\,0\,;\,0\,1\\0\,0\,;\,0\,0}} - A^{\text{planar}}_{\substack{0\,0\,;\,0\,0\\0\,0\,;\,0\,0}} = A^{\text{planar}}_{\substack{0\,0\,;\,0\,0\\1\,0\,;\,0\,1}} - A^{\text{planar}}_{\substack{0\,0\,;\,0\,0\\0\,0\,;\,0\,0}}\,, \\
\rho_0 &:= A^{\text{planar}}_{\substack{1\,1\,;\,0\,0\\0\,0\,;\,0\,0}} = A^{\text{planar}}_{\substack{0\,0\,;\,0\,0\\1\,1\,;\,0\,0}} \\
\rho_3 &:= A^{\text{planar}}_{\substack{0\,0\,;\,0\,0\,;\,0\,0\,;\,0\,0\\0\,0\,;\,0\,0\,;\,0\,0\,;\,0\,0}}
\end{aligned} \tag{129}$$

as the distinguished divergent ρ-functions. Details are given in [28].

The function ρ_0 has no commutative analogue in (66). Due to (127) it corresponds to a normalisation condition for the frequency parameter ω in (115). This means that in contrast to the two-dimensional case we cannot remove the oscillator potential with the limit $\Lambda_0 \to \infty$. In other words, the oscillator potential in (113) is a necessary companionship to the $\star$-product interaction. This observation is in agreement with the UV/IR-entanglement first observed in [18]. Whereas the UV/IR-problem prevents the renormalisation of ϕ^4-theory on $\mathbb{R}^4_\theta$ in momentum space [23], we have found a self-consistent solution of the problem by providing the unique (due to properties of the Meixner polynomials) renormalisable extension of the action. We remark that the diagonalisation of the free action via the Meixner polynomials leads to discrete momenta as the only difference to the commutative case. The inverse of such a momentum quantum can be interpreted as the size of the (finite!) universe, as it is seriously discussed in cosmology [30].

Acknowledgement

I would like to thank the organisers of the Summer school "Geometric and Topological Methods for Quantum Field Theory" for invitation, the perfect organisation and the wonderful atmosphere in Villa de Leyva.

References

1. N. N. Bogolyubov, D. V. Shirkov, "Introduction to the theory of quantized fields," Interscience (1959). 59, 71
2. C. Itzykson, J.-B. Zuber, "Quantum field theory," McGraw-Hill (1980). 59, 88
3. K. G. Wilson, J. B. Kogut, "The Renormalization Group And The Epsilon Expansion," Phys. Rept. **12**, 75 (1974). 59, 72
4. J. Glimm, A. Jaffe, "Quantum Physics: a functional integral point of view," Springer-Verlag (1981). 59
5. H. Grosse, "Models in statistical physics and quantum field theory," Springer-Verlag (1988). 59
6. K. Osterwalder, R. Schrader, "Axioms For Euclidean Green's Functions. I, II," Commun. Math. Phys. **31**, 83 (1973); **42**, 281 (1975). 59
7. G. Roepstorff, "Path integral approach to quantum physics: an introduction," Springer-Verlag (1994). 59

8. B. Simon, "The $P(\Phi)_2$ Euclidean (Quantum) Field Theory," Princeton University Press (1974). 64
9. G. Velo, A. S. Wightman (eds), "Renormalization Theory," Reidel (1976). 71
10. W. Zimmermann, "Convergence Of Bogolyubov's Method Of Renormalization In Momentum Space," Commun. Math. Phys. **15**, 208 (1969) [Lect. Notes Phys. **558**, 217 (2000)]. 72
11. A. Connes, D. Kreimer, "Renormalization in quantum field theory and the Riemann-Hilbert problem. I: The Hopf algebra structure of graphs and the main theorem," Commun. Math. Phys. **210**, 249 (2000) [arXiv:hep-th/9912092]. 72
12. A. Connes, D. Kreimer, "Renormalization in quantum field theory and the Riemann-Hilbert problem. II: The beta-function, diffeomorphisms and the renormalization group," Commun. Math. Phys. **216**, 215 (2001) [arXiv:hep-th/0003188]. 72
13. J. Polchinski, "Renormalization And Effective Lagrangians," Nucl. Phys. B **231**, 269 (1984). 72
14. M. Salmhofer, "Renormalization: An Introduction," Springer-Verlag (1998). 72
15. S. Doplicher, K. Fredenhagen, J. E. Roberts, "The Quantum structure of spacetime at the Planck scale and quantum fields," Commun. Math. Phys. **172**, 187 (1995) [arXiv:hep-th/0303037]. 84
16. E. Schrödinger, "Über die Unanwendbarkeit der Geometrie im Kleinen," Naturwiss. **31**, 342 (1934). 84
17. A. Connes, "Noncommutative geometry," Academic Press (1994). 85
18. S. Minwalla, M. Van Raamsdonk, N. Seiberg, "Noncommutative perturbative dynamics," JHEP **0002**, 020 (2000) [arXiv:hep-th/9912072]. 85, 90, 98
19. N. Seiberg, E. Witten, "String theory and noncommutative geometry," JHEP **9909**, 032 (1999) [arXiv:hep-th/9908142]. 86
20. V. Gayral, J. M. Gracia-Bondía, B. Iochum, T. Schücker, J. C. Várilly, "Moyal planes are spectral triples," Commun. Math. Phys. **246**, 569 (2004) [arXiv:hep-th/0307241]. 86
21. I. Chepelev, R. Roiban, "Renormalization of quantum field theories on noncommutative $\mathbb{R}^d$. I: Scalars," JHEP **0005**, 037 (2000) [arXiv:hep-th/9911098]. 87, 88, 89
22. T. Filk, "Divergencies In A Field Theory On Quantum Space," Phys. Lett. B **376**, 53 (1996). 88
23. I. Chepelev, R. Roiban, "Convergence theorem for non-commutative Feynman graphs and renormalization," JHEP **0103**, 001 (2001) [arXiv:hep-th/0008090]. 88, 89, 96, 98
94, 95
24. Grosse, H., Wulkenhaar, R.: Renormalisation of 4 theory on noncommutative 4 to all orders. To appear in Lett. Math. Phys., http://arxiv.org/abs/hep-th/0403232, 2004
25. H. Grosse, R. Wulkenhaar, "Renormalisation of ϕ^4 theory on noncommutative $\mathbb{R}^2$ in the matrix base," JHEP **0312**, 019 (2003) [arXiv:hep-th/0307017]. 95
26. E. Langmann, R. J. Szabo, K. Zarembo, "Exact solution of noncommutative field theory in background magnetic fields," Phys. Lett. B **569**, 95 (2003) [arXiv:hep-th/0303082]. 91
27. E. Langmann, R. J. Szabo, K. Zarembo, "Exact solution of quantum field theory on noncommutative phase spaces," JHEP **0401**, 017 (2004) [arXiv:hep-th/0308043]. 91
28. Grosse, H., Wulkenhaar, R.: The -function in duality-covariant noncommutative 4-theory. Eur. Phys. J. C **35**, 277–282 (2004) 92, 98

29. R. Koekoek, R. F. Swarttouw, "The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue," arXiv:math.CA/9602214. 92
30. J. P. M. Luminet, J. Weeks, A. Riazuelo, R. Lehoucq, J. P. Uzan, "Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmic microwave background," Nature **425**, 593 (2003) [arXiv:astro-ph/0310253]. 98

Introduction to String Compactification

A. Font[1] and S. Theisen[2]

[1] Departamento de Física, Facultad de Ciencias, Universidad Central de Venezuela, A.P. 20513, Caracas 1020-A, Venezuela
afont@fisica.ciens.ucv.ve
[2] Albert-Einstein-Institut, Am Mühlenberg 1, 14476 Golm, Germany
theisen@aei.mpg.de

Summary. We present an elementary introduction to compactifications with unbroken supersymmetry. After explaining how this requirement leads to internal spaces of special holonomy we describe Calabi-Yau manifolds in detail. We also discuss orbifolds as examples of solvable string compactifications.

1 Introduction

The need to study string compactification is a consequence of the fact that a quantum relativistic (super)string cannot propagate in any space-time background. The dynamics of a string propagating in a background geometry defined by the metric G_{MN} is governed by the *Polyakov action*

$$S_{\rm P} = -\frac{1}{4\pi\alpha'}\int_\Sigma d^2\sigma\sqrt{-h}h^{\alpha\beta}\partial_\alpha X^M\partial_\beta X^N G_{MN}(X)\ . \qquad (1.1)$$

Here σ^α, $\alpha = 0,1$, are local coordinates on the string world-sheet Σ, $h_{\alpha\beta}$ is a metric on Σ with $h = \det h_{\alpha\beta}$, and X^M, $M = 0,\ldots,D-1$, are functions $\Sigma \hookrightarrow$ space-time $\mathcal{M}$ with metric $G_{MN}(X)$. α' is a constant of dimension $(\text{length})^2$. $S_{\rm P}$ is the action of a two-dimensional non-linear σ-model with target space $\mathcal{M}$, coupled to two-dimensional gravity ($h_{\alpha\beta}$) where the D-dimensional metric G_{MN} appears as a coupling function (which generalizes the notion of a coupling constant). For flat space-time with metric $G_{MN} = \eta_{MN}$ the two-dimensional field theory is a free theory. The action (1.1) is invariant under local scale (Weyl) transformations $h_{\alpha\beta} \to e^{2\omega}h_{\alpha\beta}$, $X^M \to X^M$. One of the central principles of string theory is that when we quantize the two-dimensional field theory we must not loose this local scale invariance. In the path-integral quantization this means that it is not sufficient if the action is invariant because the integration measure might receive a non-trivial Jacobian which destroys the classical symmetry. Indeed, for the Polyakov action anomalies occur and produce a non-vanishing beta function $\beta^{(G)}_{MN} \equiv \alpha' R_{MN} + \mathcal{O}(\alpha'^2)$. Requiring $\beta^{(G)}_{MN} = 0$ to maintain Weyl invariance gives the Einstein equations for the background metric: only their solutions are viable (perturbative) string backgrounds. But there are more restrictions.

Besides the metric, in the Polyakov action (1.1) other background fields can appear as coupling functions: an antisymmetric tensor-field $B_{MN}(X)$

A. Font and S. Theisen: *Introduction to String Compactification*,
Lect. Notes Phys. **668**, 101–181 (2005)
www.springerlink.com

and a scalar, the dilaton $\phi(X)$.[1] The background value of the dilaton determines the string coupling constant, i.e. the strength with which strings interact with each other. Taking into account the fermionic partners (under world-sheet supersymmetry) of X^M and $h_{\alpha\beta}$ gives beta functions for B_{MN} and ϕ that vanish for constant dilaton and zero antisymmetric field only if $D = 10$. This defines the *critical dimension* of the supersymmetric string theories. We thus have to require that the background space-time $\mathcal{M}_{10}$ is a ten-dimensional Ricci-flat manifold with Lorentzian signature. Here we have ignored the $\mathcal{O}(\alpha'^2)$ corrections, to which we will briefly return later. The bosonic string which has critical dimension 26 is less interesting as it has no fermions in its excitation spectrum.

The idea of compactification arises naturally to resolve the discrepancy between the critical dimension $D = 10$ and the number of observed dimensions $d = 4$. Since $\mathcal{M}_{10}$ is dynamical, there can be solutions, consistent with the requirements imposed by local scale invariance on the world-sheet, which make the world *appear* four-dimensional. The simplest possibility is to have a background metric such that space-time takes the product form $\mathcal{M}_{10} = \mathcal{M}_4 \times K_6$ where e.g. $\mathcal{M}_4$ is four-dimensional Minkowski space and K_6 is a compact space which admits a Ricci-flat metric. Moreover, to have escaped detection, K_6 must have dimensions of size smaller than the length scales already probed by particle accelerators. The type of theory observed in $\mathcal{M}_4$ will depend on properties of the compact space. For instance, in the classic analysis of superstring compactification of [2], it was found that when K_6 is a Calabi-Yau manifold, the resulting four-dimensional theory has a minimal number of supersymmetries [2]. One example of Calabi-Yau space discussed in [2] was the Z-manifold obtained by resolving the singularities of a $\mathrm{T}^6/\mathbb{Z}_3$ orbifold. It was soon noticed that string propagation on the singular orbifold was perfectly consistent and moreover exactly solvable [3]. These lectures provide an introduction to string compactifications on Calabi-Yau manifolds and orbifolds.

The outline is as follows. In Sect. 2 we give a short review of compactification à la Kaluza-Klein. Our aim is to explain how a particular choice of compact manifold imprints itself on the four-dimensional theory. We also discuss how the requirement of minimal supersymmetry singles out Calabi-Yau manifolds. In Sect. 3 we introduce some mathematical background: complex manifolds, Kähler manifolds, cohomology on complex manifolds. We then give a definition of Calabi-Yau manifolds and state Yau's theorem. Next we present the cohomology of Calabi-Yau manifolds and discuss their moduli

[1] There are other p-form fields, but their coupling to the world-sheet cannot be incorporated into the Polyakov action. The general statement is that the massless string states in the (NS,NS) sector, which are the metric, the anti-symmetric tensor and the dilaton, can be added to the Polyakov action. The massless p-forms in the (R,R) sector cannot. This can only be done within the so-called Green-Schwarz formalism and its extensions by Siegel and Berkovits; for review see [1].

spaces. As an application we work out the massless content of type II superstrings compactified on Calabi-Yau manifolds. In Sect. 4 we study orbifolds, first explaining some basic properties needed to describe string compactification on such spaces. We systematically compute the spectrum of string states starting from the partition function. The techniques are next applied to compactify type II strings on $\mathrm{T}^{2n}/\mathbb{Z}_N$ orbifolds that are shown to allow unbroken supersymmetries. These toroidal Abelian orbifolds are in fact simple examples of spaces of special holonomy and the resulting lower-dimensional supersymmetric theories belong to the class obtained upon compactification on Calabi-Yau n-folds. We end with a quick look at recent progress. In Appendix A we fix our conventions and recall a few basic notions about spinors and Riemannian geometry. Two technical results which will be needed in the text are derived in Appendices B and C.

In these notes we review well known principles that have been applied in string theory for many years. There are several important developments which build on the material presented here which will not be discussed: mirror symmetry, D-branes and open strings, string dualities, compactification on manifolds with G_2 holonomy, etc. The lectures were intended for an audience of beginners in the field and we hope that they will be of use as preparation for advanced applications. We assume that the reader is already familiar with basic concepts in string theory that are well covered in textbooks [4, 5, 6]. But most of Sects. 2 and 3 do not use string theory at all. We have included many exercises whose solutions will eventually appear on [7].

2 Kaluza-Klein Fundamentals

Kaluza and Klein unified gravity and electromagnetism in four dimensions by deriving both interactions from pure gravity in five dimensions. Generalizing this, one might attempt to explain all known elementary particles and their interactions from a simple higher dimensional theory. String theory naturally lives in ten dimensions and so lends itself to the Kaluza-Klein program.

The discussion in this section is relevant for the field theory limit of string theory, where its massive excitation modes can be neglected. The dynamics of the massless modes is then described in terms of a low-energy effective action whose form is fixed by the requirement that it reproduces the scattering amplitudes as computed from string theory. However, when we compactify a string theory rather than a field theory, there are interesting additional features to which we return in Sect. 4.

In the following we explain some basic results in Kaluza-Klein compactifications of field theories. For a comprehensive review see for instance [8] which cites the original literature. The basic material is well covered in [4] which also discusses the string theory aspects.

2.1 Dimensional Reduction

Given a theory in D dimensions we want to derive the theory that results upon compactifying $D-d$ coordinates on an internal manifold K_{D-d}. As a simple example consider a real massless scalar in $D=5$ with action

$$S_0 = -\frac{1}{2}\int d^5x\, \partial_M\varphi\, \partial^M\varphi \ , \tag{2.1}$$

where $\partial^M = \eta^{MN}\partial_N$ with $\eta_{MN} = \eta^{MN} = \mathrm{diag}(-,+,\cdots,+)$, $M,N = 0,\cdots,4$. The flat metric is consistent with the five-dimensional space $\mathcal{M}_5$ having product form $\mathcal{M}_5 = M_4 \times S^1$, where M_4 is four-dimensional Minkowski space and S^1 is a circle of radius R. We denote $x_M = (x_\mu, y)$, $\mu = 0,\cdots,3$, so that $y \in [0, 2\pi R]$. The field φ satisfies the equation of motion

$$\Box\varphi = 0 \quad \Rightarrow \quad \partial_\mu\partial^\mu\varphi + \partial_y^2\varphi = 0 \ . \tag{2.2}$$

Now, since $\varphi(x,y) = \varphi(x, y+2\pi R)$, we can write the Fourier expansion

$$\varphi(x,y) = \frac{1}{\sqrt{2\pi R}}\sum_{n=-\infty}^{\infty}\varphi_n(x)e^{iny/R} \ . \tag{2.3}$$

Notice that $Y_n(y) \equiv \frac{1}{\sqrt{2\pi R}}e^{iny/R}$ are the orthonormalized eigenfunctions of ∂_y^2 on S^1. Substituting (2.3) in (2.2) gives

$$\partial_\mu\partial^\mu\varphi_n - \frac{n^2}{R^2}\varphi_n = 0 \ . \tag{2.4}$$

This clearly means that $\varphi_n(x)$ are 4-dimensional scalar fields with masses n/R. This can also be seen at the level of the action. Substituting (2.3) in (2.1) and integrating over y (using orthonormality of the Y_n) gives

$$S_0 = -\sum_{n=-\infty}^{\infty}\frac{1}{2}\int d^4x\left[\partial_\mu\varphi_n\, \partial^\mu\varphi_n^* + \frac{n^2}{R^2}\varphi_n^*\varphi_n\right] \ . \tag{2.5}$$

This again shows that in four dimensions there is one massless scalar φ_0 plus an infinite tower of massive scalars φ_n with masses n/R. We are usually interested in the limit $R \to 0$ in which only φ_0 remains light while the φ_n, $n \neq 0$, become very heavy and are discarded. We refer to this limit in which only the *zero mode* φ_0 is kept as *dimensional reduction* because we could obtain the same results demanding that $\varphi(x_M)$ be independent of y. More generally, dimensional reduction in this restricted sense is compactification on a torus T^{D-d}, discarding massive modes, i.e. all states which carry momentum along the directions of the torus.

The important concept of zero modes generalizes to the case of curved internal compact spaces. However, it is only in the case of torus compactification that all zero modes are independent of the internal coordinates. This guarantees the consistency of the procedure of discarding the heavy modes in the sense that a solution of the lower-dimensional equations of motion is also a solution of the full higher-dimensional ones.

In D dimensions we can have other fields transforming in various representations of the Lorentz group $SO(1, D-1)$. We then need to consider how they decompose under the Lorentz group in the lower dimensions. Technically, we need to decompose the representations of $SO(1, D-1)$ under $SO(1, d-1) \times SO(D-d)$ associated to $M_d \times K_{D-d}$. For example, for a vector A_M transforming in the fundamental representation $\boldsymbol{D}$ we have the branching $\boldsymbol{D} = (\boldsymbol{d}, \mathbf{1}) + (\mathbf{1}, \boldsymbol{D-d})$. This just means that A_M splits into A_μ, $\mu = 0, \cdots, d-1$ and A_m, $m = d, \cdots, D-1$. A_μ is a vector under $SO(1, d-1)$ whereas A_m, for each m, is a singlet, i.e. the A_m appear as $(D-d)$ scalars in d dimensions. Similarly, a two-index antisymmetric tensor B_{MN} decomposes into $B_{\mu\nu}$, $B_{\mu m}$ and B_{mn}, i.e. into an antisymmetric tensor, vectors and scalars in d dimensions.

Exercise 2.1: Perform the dimensional reduction of:

– Maxwell electrodynamics.

$$S_1 = -\frac{1}{4}\int d^{4+n}x\, F_{MN}F^{MN} \ , \quad F_{MN} = \partial_M A_N - \partial_N A_M \ . \qquad (2.6)$$

– Action for a 2-form gauge field B_{MN}.

$$S_2 = -\frac{1}{12}\int d^{4+n}x\, H_{MNP}H^{MNP} \ , \quad H_{MNP} = \partial_M B_{NP} + \text{cyclic} \ . \qquad (2.7)$$

We also need to consider fields that transform as spinors under the Lorentz group. Here and below we will always assume that the manifolds considered are spin manifolds, so that spinor fields can be defined. As reviewed in Appendix A, in D dimensions, the Dirac matrices Γ^M are $2^{[D/2]} \times 2^{[D/2]}$-dimensional ($[D/2]$ denotes the integer part of $D/2$). The Γ^μ and Γ^m, used to build the generators of $SO(1, d-1)$ and $SO(D-d)$, respectively, then act on all $2^{[D/2]}$ spinor components. This means that an $SO(1, D-1)$ spinor transforms as a spinor under both $SO(1, d-1)$ and $SO(D-d)$. For example, a Majorana spinor ψ in $D=11$ decomposes under $SO(1,3) \times SO(7)$ as $\mathbf{32} = (\mathbf{4}, \mathbf{8})$, where $\mathbf{4}$ and $\mathbf{8}$ are respectively Majorana spinors of $SO(1,3)$ and $SO(7)$. Hence, dimensional reduction of ψ gives rise to eight Majorana spinors in $d = 4$.

We are mainly interested in compactification of supersymmetric theories that have a set of conserved spinorial charges Q^I, $I = 1, \cdots, \mathcal{N}$. Fields organize into supermultiplets containing both fermions and bosons that transform into each other by the action of the generators Q^I [9]. In each supermultiplet the numbers of on-shell bosonic and fermionic degrees of freedom do match and the masses of all fields are equal. Furthermore, the action that determines the dynamics of the fields is highly constrained by the requirement of invariance under supersymmetry transformations. For instance, for $D = 11$,

$\mathcal{N}=1$, there is a unique theory, namely eleven-dimensional supergravity. For $D=10$, $\mathcal{N}=2$ there are two different theories, non-chiral IIA (Q^1 and Q^2 are Majorana-Weyl spinors of opposite chirality) and chiral IIB supergravity (Q^1 and Q^2 of same chirality). For $D=10$, $\mathcal{N}=1$, a supergravity multiplet can be coupled to a non-Abelian super Yang-Mills multiplet provided that the gauge group is $E_8 \times E_8$ or $SO(32)$ to guarantee absence of quantum anomalies. The above theories describe the dynamics of M-theory and the various string theories at low energies.

One way to obtain four-dimensional supersymmetric theories is to start in $D=11$ or $D=10$ and perform dimensional reduction, i.e. compactify on a torus. For example, we have just explained that dimensional reduction of a $D=11$ Majorana spinor produces eight Majorana spinors in $d=4$. This means that starting with $D=11$, $\mathcal{N}=1$, in which Q is Majorana, gives a $d=4$, $\mathcal{N}=8$ theory upon dimensional reduction. As another interesting example, consider $D=10$, $\mathcal{N}=1$ in which Q is a Majorana-Weyl spinor. The **16** Weyl representation of $SO(1,9)$ decomposes under $SO(1,3) \times SO(6)$ as

$$\mathbf{16} = (\mathbf{2_L}, \bar{\mathbf{4}}) + (\mathbf{2_R}, \mathbf{4}) \ , \tag{2.8}$$

where $\mathbf{4}$, $\bar{\mathbf{4}}$ are Weyl spinors of $SO(6)$ and $\mathbf{2_{L,R}}$ are Weyl spinors of $SO(1,3)$ that are conjugate to each other. If we further impose the Majorana condition in $D=10$, then dimensional reduction of Q gives rise to four Majorana spinors in $d=4$. Thus, $\mathcal{N}=1,2$ supersymmetric theories in $D=10$ yield $\mathcal{N}=4,8$ supersymmetric theories in $d=4$ upon dimensional reduction.

Toroidal compactification of superstrings gives theories with too many supersymmetries that are unrealistic because they are non-chiral, they cannot have the chiral gauge interactions observed in nature. Supersymmetric extensions of the Standard Model require $d=4$, $\mathcal{N}=1$. Such models have been extensively studied over the last 25 years (for a recent review, see [10]). One reason is that supersymmetry, even if it is broken at low energies, can explain why the mass of the Higgs boson does not receive large radiative corrections. Moreover, the additional particles and particular couplings required by supersymmetry lead to distinct experimental signatures that could be detected in future high energy experiments.

To obtain more interesting theories we must go beyond toroidal compactification. As a guiding principle we demand that some supersymmetry is preserved. As we will see, this allows a more precise characterization of the internal manifold. Supersymmetric string compactifications are moreover stable, in contrast to non-supersymmetric vacua that can be destabilized by tachyons or tadpoles. Now, we know that in the real world supersymmetry must be broken since otherwise the superpartner of e.g. the electron would have been observed. Supersymmetry breaking in string theory is still an open problem.

2.2 Compactification, Supersymmetry and Calabi-Yau Manifolds

Up to now we have not included gravity. When a metric field G_{MN} is present, the fact that space-time $\mathcal{M}_D$ has a product form $\mathcal{M}_d \times K_{D\text{–}d}$, with $K_{D\text{–}d}$ compact, must follow from the dynamics. If the equations of motion have such a solution, we say that the system admits *spontaneous compactification*. The vacuum expectation value (vev) of G_{MN} then satisfies

$$\langle G_{MN}(x,y)\rangle = \begin{pmatrix} \bar{g}_{\mu\nu}(x) & 0 \\ 0 & \bar{g}_{mn}(y) \end{pmatrix} , \tag{2.9}$$

where x_μ and y_m are the coordinates of $\mathcal{M}_d$ and $K_{D\text{–}d}$ respectively. Note that with this Ansatz there are no non-zero components of the Christoffel symbols and the Riemann tensor which carry both Latin and Greek indices. An interesting generalization of (2.9) is to keep the product form but with the metric components on $\mathcal{M}_d$ replaced by $e^{2A(y)}\bar{g}_{\mu\nu}(x)$, where $A(y)$ is a so-called "warp factor" [11]. This still allows maximal space-time symmetry in $\mathcal{M}_d$. For instance $\langle G_{\mu\nu}(x,y)\rangle = e^{2A(y)}\eta_{\mu\nu}$ is compatible with d-dimensional Poincaré symmetry. In these notes we do not consider such warped product metrics.

We are mostly interested in D-dimensional supergravity theories and we will search for compactifications that preserve some degree of supersymmetry. Instead of analyzing whether the equations of motion, which are highly nonlinear, admit solutions of the form (2.9), it is then more convenient to demand (2.9) and require unbroken supersymmetries in $\mathcal{M}_d$. A posteriori it can be checked that the vevs obtained for all fields are compatible with the equations of motion.

We thus require that the vacuum satisfies $\bar{\epsilon}Q|0\rangle = 0$ where $\epsilon(x^M)$ parametrizes the supersymmetry transformation which is generated by Q, both Q and ϵ being spinors of $SO(1, D-1)$. This, together with $\delta_\epsilon \Phi = [\bar{\epsilon}Q, \Phi]$, means that $\langle \delta_\epsilon \Phi\rangle \equiv \langle 0|[\bar{\epsilon}Q, \Phi]|0\rangle = 0$ for every field generically denoted by Φ. Below we will be interested in the case where $\mathcal{M}_d$ is Minkowski space. Then, with the exception of a vev for the metric $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$ and a d-form $\bar{F}_{\mu_1\ldots\mu_d} = \epsilon_{\mu_1\ldots\mu_d}$, a non-zero background value of any field which is not a $SO(1, d-1)$ scalar, would reduce the symmetries of Minkowski space. In particular, since fermionic fields are spinors that transform non-trivially under $SO(1, d-1)$, $\langle \Phi_{\rm Fermi}\rangle = 0$. Hence, $\langle \delta_\epsilon \Phi_{\rm Bose}\rangle \sim \langle \Phi_{\rm Fermi}\rangle = 0$ and we only need to worry about $\langle \delta_\epsilon \Phi_{\rm Fermi}\rangle$. Now, among the $\Phi_{\rm Fermi}$ in supergravity there is always the gravitino ψ_M (or $\mathcal{N}$ gravitini if there are $\mathcal{N}$ supersymmetries in higher dimensions) that transforms as

$$\delta_\epsilon \psi_M = \nabla_M\, \epsilon + \cdots \ , \tag{2.10}$$

where ∇_M is the covariant derivative defined in Appendix A. The ... stand for terms which contain other bosonic fields (dilaton, B_{MN} and p-form fields) whose vevs are taken to be zero. Then, $\langle \delta_\epsilon \psi_M\rangle = 0$ gives

$$\langle \nabla_M \epsilon \rangle \equiv \bar{\nabla}_M \epsilon = 0 \qquad \Rightarrow \qquad \bar{\nabla}_m \epsilon = 0 \quad \text{and} \quad \bar{\nabla}_\mu \epsilon = 0 \; . \tag{2.11}$$

Notice that in $\bar{\nabla}_M$ there appears the vev of the spin connection $\bar{\omega}$. Spinor fields ϵ, which satisfy (2.11) are covariantly constant (in the vev metric); they are also called *Killing spinors*.

The existence of Killing spinors, which is a necessary requirement for a supersymmetric compactification, restricts the class of manifolds on which we may compactify. To see this explicitly, we iterate (2.11) to obtain the integrability condition (since the manipulations until (2.14) are completely general, we drop the bar)

$$[\nabla_m, \nabla_n]\epsilon = \frac{1}{4} R_{mn}{}^{ab} \Gamma_{ab} \epsilon = \frac{1}{4} R_{mnpq} \Gamma^{pq} \epsilon = 0 \; , \tag{2.12}$$

where $\Gamma_{ab} = \frac{1}{2}[\Gamma_a, \Gamma_b]$ and R_{mnpq} is the Riemann tensor on K_{D-d}.

Exercise 2.2: Verify (2.12) using (A.12).

Next we multiply by Γ^n and use the Γ property

$$\Gamma^n \Gamma^{pq} = \Gamma^{npq} + g^{np} \Gamma^q - g^{nq} \Gamma^p \; , \tag{2.13}$$

where Γ^{npq} is defined in (A.2). The Bianchi identity

$$R_{mnpq} + R_{mqnp} + R_{mpqn} = 0 \tag{2.14}$$

implies that $\Gamma^{npq} R_{mnpq} = 0$. In this way we arrive at

$$\bar{R}_{mq} \bar{\Gamma}^q \, \epsilon = 0 \; . \tag{2.15}$$

From the linear independence of the $\Gamma^q \, \epsilon$ it follows that a necessary condition for the existence of a Killing spinor on a Riemannian manifold is the vanishing of its Ricci tensor:

$$\bar{R}_{mq} = 0 \; . \tag{2.16}$$

Hence, the internal $K_{D\text{–}d}$ is a compact Ricci-flat manifold. This is the same condition as that obtained from the requirement of Weyl invariance at the level of the string world-sheet and it is also the equation of motion derived from the supergravity action if all fields except the metric are set to zero.

One allowed solution is $K_{D\text{–}d} = \mathrm{T}^{D\text{–}d}$, i.e. a $(D\text{–}d)$ torus that is compact and flat. This means that dimensional reduction is always possible and, since ϵ is constant because in this case $\bar{\nabla}_m \, \epsilon = \partial_m \epsilon = 0$, it gives the maximum number of supersymmetries in the lower dimensions. The fact that supersymmetry requires $K_{D\text{–}d}$ to be Ricci-flat is a very powerful result. For example, it is known that Ricci-flat compact manifolds do not admit Killing vectors other than those associated with tori. Equivalently, the Betti number b_1 only gets contributions from non-trivial cycles associated to tori factors in $K_{D\text{–}d}$. The fact that the internal manifold must have Killing spinors encodes much more information. To analyze this in more detail below we specialize to a six-dimensional internal K_6 which is the case of interest for string compactifications from $D=10$ to $d=4$.

Before doing this we need to introduce the concept of holonomy group $\mathcal{H}$ [12, 13]. Upon parallel transport along a closed curve on an m-dimensional manifold, a vector v is rotated into Uv. The set of matrices obtained in this way forms $\mathcal{H}$. The U's are necessarily matrices in $O(m)$ which is the tangent group of the Riemannian K_m. Hence $\mathcal{H} \subseteq O(m)$. For manifolds with an orientation the stronger condition $\mathcal{H} \subseteq SO(m)$ holds. Now, from (A.14) it follows that for a simply-connected manifold to have non-trivial holonomy it has to have curvature. Indeed, the Riemann tensor (and its covariant derivatives), when viewed as a Lie-algebra valued two-form, generate $\mathcal{H}$. If the manifold is not simply connected, the Riemann tensor and its covariant derivatives only generate the identity component of the holonomy group, called the *restricted holonomy group* $\mathcal{H}_0$ for which $\mathcal{H}_0 \subseteq SO(m)$. Non-simply connected manifolds can have non-trivial $\mathcal{H}$ without curvature, as exemplified in the following exercise.

Exercise 2.3: Consider the manifold $S^1 \otimes \mathbb{R}^n$ endowed with the metric

$$ds^2 = R^2 d\theta^2 + (dx^i + \Omega^i{}_j x^j d\theta)^2 \ , \tag{2.17}$$

where $\Omega^i{}_j$ is a constant anti-symmetric matrix, i.e. a generator of the rotation group $SO(n)$ and R is the radius of S^1. Show that this metric has vanishing curvature but that nevertheless a vector, when parallel transported around the circle, is rotated by an element of $SO(n)$.

Under parallel transport along a loop in K_6, spinors are also rotated by elements of $\mathcal{H}$. But a covariantly constant spinor such as ϵ remains unchanged. This means that ϵ is a singlet under $\mathcal{H}$. But ϵ is an $SO(6)$ spinor and hence it has right- and left-chirality pieces that transform respectively as **4** and $\bar{\mathbf{4}}$ of $SO(6) \simeq SU(4)$. How can ϵ be an $\mathcal{H}$-singlet? Suppose that $\mathcal{H} = SU(3)$. Under $SU(3)$ the **4** decomposes into a triplet and a singlet: $\mathbf{4}_{SU(4)} = (\mathbf{3} + \mathbf{1})_{SU(3)}$. Thus, if $\mathcal{H} = SU(3)$ there is one covariantly constant spinor of positive and one of negative chirality, which we denote $\epsilon_\pm$. If $\mathcal{H}$ were $SU(2)$ there would be two right-handed and two left-handed covariantly constant spinors since under $SU(2)$ the **4** decomposes into a doublet and two singlets. There could be as many as four covariantly constant spinors of each chirality as occurs when $K_6 = \mathrm{T}^6$ and $\mathcal{H}_0$ is trivial since the torus is flat.

Let us now pause to show that if K_6 has $SU(3)$ holonomy, the resulting theory in $d = 4$ has precisely $\mathcal{N} = 1$ supersymmetry if it had $\mathcal{N} = 1$ in $D=10$. Taking into account the decomposition (2.8) and the discussion in the previous paragraph, we see that the allowed supersymmetry parameter takes the form

$$\epsilon = \epsilon_R \otimes \epsilon_+ + \epsilon_L \otimes \epsilon_- \ . \tag{2.18}$$

Since ϵ is also Majorana it must be that $\epsilon_R = \epsilon_L^*$ and hence ϵ_R and ϵ_L form just a single Majorana spinor, associated to a single supersymmetry generator. Similarly, if K_6 has $SU(2)$ holonomy the resulting $d=4$ theory will have $\mathcal{N}=2$ supersymmetry. Obviously, the number of supersymmetries in $d=4$ is doubled if we start from $\mathcal{N}=2$ in $D=10$.

$2n$-dimensional compact Riemannian manifolds with $SU(n) \subset SO(2n)$ holonomy are *Calabi-Yau manifolds* CY_n. We have just seen that they admit covariantly constant spinors and that they are Ricci-flat. We will learn much more about Calabi-Yau manifolds in the course of these lectures and we will also make the definition more precise. For $n = 1$ there is only one CY_1, namely the torus T^2. The only CY_2 is the K3 manifold. For $n \geq 3$ there is a huge number. We will give simple examples of CY_3 in Sect. 3. Many more can be found in [14]. We want to remark that except for the trivial case $n = 1$ no metric with $SU(n)$ holonomy on any CY_n is known explicitly. Existence and uniqueness have, however, been shown (cf. Sect. 3).

Calabi-Yau manifolds are a class of *manifolds with special holonomy*. Generically on an oriented manifold one has $\mathcal{H} \simeq SO(m)$. Then the following question arises: which subgroups $G \subset SO(m)$ do occur as holonomy groups of Riemannian manifolds? For the case of simply connected manifolds which are neither symmetric nor locally a product of lower dimensional manifolds, this question was answered by Berger. His classification along with many of the properties of the manifolds is discussed at length in [12, 13]. All types of manifolds with special holonomy do occur in the context of string compactification, either as the manifold on which we compactify or as moduli spaces (cf. Sect. 3.6).

Exercise 2.4: Use simple group theory to work out the condition on the holonomy group of seven- and eight-dimensional manifolds which gives the minimal amount of supersymmetry if one compactifies eleven-dimensional supergravity to four or three dimensions or ten-dimensional supergravity to $d = 3$ and $d = 2$, respectively.

Going back to the important case, $\mathcal{N}=1$, $D=10$, $d=4$, and the requirement of unbroken supersymmetry we find the following possibilities. The internal K_6 can be a torus T^6 with trivial holonomy and hence ϵ leads to $d=4$, $\mathcal{N}=4$ supersymmetry. K_6 can also be a product $\mathrm{K3} \times \mathrm{T}^2$ with $SU(2)$ holonomy and ϵ leads to $\mathcal{N} = 2$ in $d = 4$. Finally, K_6 can be a CY_3 that has $SU(3)$ holonomy so that ϵ gives $d=4$, $\mathcal{N}=1$ supersymmetry. These are the results for heterotic and type I strings. For type II strings the number of supersymmetries in the lower dimensions is doubled since we start from $\mathcal{N}=2$ in $D=10$.

Let us also consider compactifications from $\mathcal{N} = 1$, $D = 10$ to $d = 6$. In this case unbroken supersymmetry requires K_4 to be the flat torus T^4 or the K3 manifold with $SU(2)$ holonomy. Toroidal compactification does not reduce the number of real supercharges (16 in $\mathcal{N} = 1$, $D = 10$), thus when the internal manifold is T^4 the theory in $d = 6$ has $\mathcal{N} = 2$, or rather (1,1), supersymmetry. Here the notation indicates that one supercharge is a left-handed and the other a right-handed Weyl spinor. The $SO(1, 5)$ Weyl spinors are complex since a Majorana-Weyl condition cannot be imposed in $d=6$. Compactification on K3 gives $d=6$, $\mathcal{N}=1$, or rather (1,0), supersymmetry. This can be understood from the decomposition of the **16** Weyl representation of $SO(1,9)$ under $SO(1,5)\times SO(4)$,

$$\mathbf{16} = (\mathbf{4_L}, \mathbf{2}) + (\mathbf{4_R}, \mathbf{2'}) \ , \tag{2.19}$$

where $\mathbf{4_{L,R}}$ and $\mathbf{2}$, $\mathbf{2'}$ are Weyl spinors of $SO(1,5)$ and $SO(4)$. In both groups each Weyl representation is its own conjugate. Since the supersymmetry parameter ϵ in $D=10$ is Majorana-Weyl, its $(\mathbf{4_L}, \mathbf{2})$ piece has only eight real components which form only one complex $\mathbf{4_L}$ and likewise for $(\mathbf{4_R}, \mathbf{2'})$. Then, if the holonomy is trivial, ϵ gives one $\mathbf{4_L}$ plus one $\mathbf{4_R}$ supersymmetry in $d=6$. Instead, if the holonomy is $SU(2)\subset SO(4)\simeq SU(2)\times SU(2)$, only one $SO(4)$ spinor, say $\mathbf{2}$, is covariantly constant and then ϵ gives only one $\mathbf{4_L}$ supersymmetry. Starting from $\mathcal{N}=2$ in $D=10$ there are the following possibilities. Compactification on T^4 gives (2,2) supersymmetry for both the non-chiral IIA and the chiral IIB superstrings. However, compactification on K3 gives (1,1) supersymmetry for IIA but (2,0) supersymmetry for IIB.

From the number of unbroken supersymmetries in the lower dimensions we can already observe hints of string dualities, i.e. equivalences of the compactifications of various string theories. For example, in $d=6$, the type IIA string on K3 is dual to the heterotic string on T^4 and in $d=4$, type IIA on CY_3 is dual to heterotic on K3 $\times$ T^2. On the heterotic side non-Abelian gauge groups are perturbative but on the type IIA side they arise from non-perturbative effects, namely D-branes wrapping homology cycles inside the K3 surface. We will not discuss string dualities in these lectures. For a pedagogical introduction, see [6].

2.3 Zero Modes

We now wish to discuss Kaluza-Klein reduction when compactifying on curved internal spaces. Our aim is to determine the resulting theory in d dimensions. To begin we expand all D-dimensional fields, generically denoted $\Phi^{mn\cdots}_{\mu\nu\cdots}(x,y)$, around their vacuum expectation values

$$\Phi^{mn\cdots}_{\mu\nu\cdots}(x,y) = \langle\Phi^{mn\cdots}_{\mu\nu\cdots}(x,y)\rangle + \varphi^{mn\cdots}_{\mu\nu\cdots}(x,y) \ . \tag{2.20}$$

We next substitute in the D-dimensional equations of motion and use the splitting (2.9) of the metric. Keeping only linear terms, and possibly fixing gauge, gives generic equations

$$\mathcal{O}_d\varphi^{mn\cdots}_{\mu\nu\cdots} + \mathcal{O}_{\mathrm{int}}\varphi^{mn\cdots}_{\mu\nu\cdots} = 0 \ , \tag{2.21}$$

where $\mathcal{O}_d$, $\mathcal{O}_{\mathrm{int}}$ are differential operators of order p ($p=2$ for bosons and $p=1$ for fermions) that depend on the specific field.

We next expand $\varphi^{mn\cdots}_{\mu\nu\cdots}$ in terms of eigenfunctions $Y^{mn\cdots}_a(y)$ of $\mathcal{O}_{\mathrm{int}}$ in $K_{D\text{–}d}$. This is

$$\varphi_{\mu\nu\cdots}^{mn\cdots}(x,y) = \sum_a \varphi_{a\mu\nu\cdots}(x)\; Y_a^{mn\cdots}(y)\ . \tag{2.22}$$

Since $\mathcal{O}_{\rm int} Y_a^{mn\cdots}(y) = \lambda_a Y_a^{mn\cdots}(y)$, from (2.21) we see that the eigenvalues λ_a determine the masses of the d-dimensional fields $\varphi_{a\mu\nu\cdots}(x)$. With R a typical dimension of $K_{D\text{–}d}$, $\lambda_a \sim 1/R^p$. We again find that in the limit $R \to 0$ only the zero modes of $\mathcal{O}_{\rm int}$ correspond to massless fields $\varphi_{0\mu\nu\cdots}(x)$.

To obtain the effective d-dimensional action for the massless fields φ_0 in general it is not consistent to simply set the massive fields, i.e. the coefficients of the higher harmonics, to zero [8]. The problem with such a truncation is that the heavy fields, denoted φ_h, might induce interactions of the φ_0 that are not suppressed by inverse powers of the heavy mass. This occurs for instance when there are cubic couplings $\varphi_0\,\varphi_0\,\varphi_h$. When the zero modes $Y_0(y)$ are constant or covariantly constant a product of them is also a zero mode and then by orthogonality of the $Y_a(y)$ terms linear in φ_h cannot appear after integrating over the extra dimensions, otherwise they might be present and generate corrections to quartic and higher order couplings of the φ_0. Even when the heavy fields cannot be discarded it might be possible to consistently determine the effective action for the massless fields [15].

We have already seen that for scalar fields $\mathcal{O}_{\rm int}$ is the Laplacian Δ. On a compact manifold Δ has only one scalar zero mode, namely a constant and hence a scalar in D dimensions produces just one massless scalar in d dimensions. An important and interesting case is that of Dirac fields in which both $\mathcal{O}_d$ and $\mathcal{O}_{\rm int}$ are Dirac operators $\Gamma \cdot \nabla$. The number of zero modes of $\not\nabla \equiv \Gamma^m \nabla_m$ happen to depend only on topological properties of the internal manifold $K_{D\text{–}d}$ and can be determined using index theorems [4]. When the internal manifold is Calabi-Yau we can also exploit the existence of covariantly constant spinors. For instance, from the formula $\not\nabla^2 = \nabla^m \nabla_m$, which is valid on a Ricci-flat manifold, it follows that when K_6 is a CY_3, the Dirac operator has only two zero modes, namely the covariantly constant ϵ_+ and ϵ_-.

Among the massless higher dimensional fields there are usually p-form gauge fields $A^{(p)}$ with field strength $F^{(p+1)} = dA^{(p)}$ and action

$$S_p = -\frac{1}{2(p+1)!}\int_{\mathcal{M}_D} F^{(p+1)} \wedge *F^{(p+1)}\ . \tag{2.23}$$

After fixing the gauge freedom $A^{(p)} \to A^{(p)} + d\Lambda^{(p-1)}$ by imposing $d^* A^{(p)} = 0$, the equations of motion are

$$\Delta_D A^{(p)} = 0\ , \qquad \Delta_D = dd^* + d^* d\ . \tag{2.24}$$

If the metric splits into a d-dimensional and a $(D-d)$-dimensional part, as in (2.9), the Laplacian Δ_D also splits $\Delta_D = \Delta_d + \Delta_{D-d}$. Then, $\mathcal{O}_{\rm int}$ is the Laplacian Δ_{D-d}. The number of massless d-dimensional fields is thus given by the number of zero modes of the internal Laplacian. This is a cohomology problem, as we will see in detail in Sect. 3. In particular, the numbers of zero modes are given by Betti numbers b_r. For example, there is a 2-form that decomposes $B_{MN} \to B_{\mu\nu} \oplus B_{\mu m} \oplus B_{mn}$. Each term is an n-form with respect to the internal manifold, where n is easily read from the decomposition. Thus, from $B_{\mu\nu}$ we obtain only one zero mode since $b_0 = 1$, from $B_{\mu m}$ we obtain b_1 modes that are vectors in d dimensions and from B_{mn} we obtain b_2 modes that are scalars in d dimensions. In general, from a p-form in D dimensions we obtain b_n massless fields, $n = 0, \cdots, p$, that correspond to $(p-n)$-forms in d dimensions.

Let us now consider zero modes of the metric that decomposes $g_{MN} \to g_{\mu\nu} \oplus g_{\mu m} \oplus g_{mn}$. From $g_{\mu\nu}$ there is only one zero mode, namely the lower dimensional graviton. Massless modes coming from $g_{\mu m}$, that would behave as gauge bosons in d dimensions, can appear only when $b_1 \neq 0$ and the internal manifold has continuous isometries. Massless modes arising from g_{mn} correspond to scalars in d dimensions. To analyze these modes we write $g_{mn} = \bar{g}_{mn} + h_{mn}$. We know that a necessary condition for the fluctuations h_{mn} not to break supersymmetry is $R_{mn}(\bar{g}+h) = 0$ just as $R_{mn}(\bar{g}) = 0$. Thus, the h_{mn} are degeneracies of the vacuum, they preserve the Ricci-flatness.

The h_{mn} are usually called moduli. They are free parameters in the compactification which change the size and shape of the manifold but not its topology. For instance, a circle S^1 has one modulus, namely its radius R. The fact that any value of R is allowed manifests itself in the space-time theory as a massless scalar field with vanishing potential. The 2-torus, that has one Kähler modulus and one complex structure modulus, is another instructive example. To explain its moduli we define T^2 by identifications in a lattice Λ. This means $\mathrm{T}^2 = \mathbb{R}^2/\Lambda$. We denote the lattice vectors e_1, e_2 and define a metric $G_{mn} = \mathrm{e}_m \cdot \mathrm{e}_n$. The Kähler modulus is just the area $\sqrt{\det G}$. If there is an antisymmetric field B_{mn} then it is natural to introduce the complex Kähler modulus T via

$$T = \sqrt{\det G} + iB_{12}\ . \tag{2.25}$$

The complex structure modulus, denoted U, is

$$U = -i\frac{|\mathrm{e}_2|}{|\mathrm{e}_1|}e^{i\varphi(\mathrm{e}_1,\mathrm{e}_2)} = \frac{1}{G_{11}}(\sqrt{\det G} - iG_{12})\ . \tag{2.26}$$

U is related to the usual modular parameter by $\tau = iU$. τ can be written as a ratio of periods of the holomorphic 1-form $\Omega = dz$. Specifically, $\tau = \int_{\gamma_2} dz / \int_{\gamma_1} dz$, where γ_1, γ_2 are the two non-trivial one-cycles (associated to $\mathrm{e}_1, \mathrm{e}_2$). While all tori are diffeomorphic as real manifolds, there is no holomorphic map between two tori with complex structures τ and τ' unless they are related by a $SL(2,\mathbb{Z})$ modular transformation, cf. (4.28). This is a consequence of the geometric freedom to make integral changes of lattice basis, as long as the volume of the unit cell does not change (see e.g. [16]).

Furthermore, in string theory compactification there is a T-duality symmetry, absent in field theory, that in circle compactification identifies R and $R' = \alpha'/R$, whereas in T^2 compactification identifies all values of T related by an $SL(2,\mathbb{Z})_T$ transformation (for review, see e.g. [17]). Compactification on a torus will thus lead to two massless fields, also denoted U and T, with completely arbitrary vevs but whose couplings to other fields are restricted by invariance of the low-energy effective action under $SL(2,\mathbb{Z})_U$ and $SL(2,\mathbb{Z})_T$.

The metric moduli of CY 3-folds are also divided into Kähler moduli and complex structure moduli. This will be explained in Sect. 3.6.

Our discussion of compactification so far has been almost entirely in terms of field theory, rather than string theory. Of course, what we have learned about compactification is also relevant for string theory, since at low energies, where the excitation of massive string modes can be ignored, the dynamics of the massless modes is described by a supergravity theory in ten dimensions (for type II strings) coupled to supersymmetric Yang-Mills theory (for type I and heterotic strings).

But there are striking differences between compactifications of field theories and string theories. When dealing with strings, it is not the classical geometry (or even topology) of the space-time manifold $\mathcal{M}$ which is relevant. One dimensional objects, such as strings, probe $\mathcal{M}$ differently from point particles. Much of the attraction of string theory relies on the hope that the modification of the concept of classical geometry to "string geometry" at distances smaller than the string scale $l_s = \sqrt{\alpha'}$ (which is of the order of the Planck length[2], i.e. $\sim 10^{-33}$cm) will lead to interesting effects and eventually to an understanding of physics in this distance range. At distances large compared to l_s a description in terms of point particles should be valid and one should recover classical geometry.

One particular property of string compactification as compared to point particles is that there might be more than one manifold K_m which leads to identical theories. This resembles the situation of point particles on so-called isospectral manifolds. However, in string theory the invariance is more fundamental, as no experiment can be performed to distinguish between the manifolds. This is an example of a duality, of which many are known. T-duality of the torus compactification is one simple example which was already mentioned. A particularly interesting example which arose from the study of Calabi-Yau compactifications is *mirror symmetry* . It states that for any Calabi-Yau manifold X there exists a mirror manifold $\hat{X}$, such that $\mathrm{IIA}(X) = \mathrm{IIB}(\hat{X})$. Here the notation $\mathrm{IIA}(X)$ means the full type IIA string theory, including all perturbative and non-perturbative effects, compactified on X. For the heterotic string with standard embedding of the spin connection

[2] This is fixed by the identification of one of the massless excitation modes of the closed string with the graviton and comparing its self-interactions, as computed from string theory, with general relativity. This leads to a relation between Newton's constant and α'.

in the gauge connection [2] mirror symmetry means $\text{het}(X) = \text{het}(\hat{X})$. The manifolds comprising a mirror pair are very different, e.g. in terms of their Euler numbers $\chi(X) = -\chi(\hat{X})$. The two-dimensional torus, which we discussed above, is its own mirror manifold, but mirror symmetry exchanges the two types of moduli: $U \leftrightarrow T$. In compactifications on Calabi-Yau 3-folds, mirror symmetry also exchanges complex structure and Kähler moduli between X and $\hat{X}$.

Mirror symmetry in string compactification is a rather trivial consequence of its formulation in the language of two-dimensional conformal field theory. However, when cast in the geometric language, it becomes highly non-trivial and has lead to surprising predictions in algebraic geometry. Except for a few additional comments at the end of Sect. 3.6 we will not discuss mirror symmetry in these lectures. An up-to-date extensive coverage of most mathematical and physical aspects of mirror symmetry has recently appeared [18].

3 Complex Manifolds, Kähler Manifolds, Calabi-Yau Manifolds

3.1 Complex Manifolds

In the previous chapter we have seen how string compactifications which preserve supersymmetry directly lead to manifolds with $SU(3)$ holonomy. These manifolds have very special properties which we will discuss in this chapter. In particular they can be shown to be complex manifolds. We begin this chapter with a review of complex manifolds and of some of the mathematics necessary for the discussion of CY manifolds. Throughout we assume some familiarity with real manifolds and Riemannian geometry. None of the results collected in this chapter are new, but some of the details we present are not readily available in the (physics) literature. Useful references are [4, 19, 20, 21, 22] (physics), [12, 23, 24, 25, 26, 27] (mathematics) and, in particular, [28]. In this section we use Greek indices for the (real) coordinates on the compactification manifold, which we will generically call M.

A *complex manifold* M is a differentiable manifold admitting an open cover $\{U_a\}_{a\in A}$ and coordinate maps $z_a : U_a \to \mathbb{C}^n$ such that $z_a \circ z_b^{-1}$ is holomorphic on $z_b(U_a \cap U_b) \subset \mathbb{C}^n$ for all a, b. $z_a = (z_a^1, \ldots, z_a^n)$ are *local holomorphic coordinates* and on overlaps $U_a \cap U_b$, $z_a^i = f_{ab}^i(z_b)$ are holomorphic functions, i.e. they do not depend on $\bar{z}_b^i$. (When considering local coordinates we will often drop the subscript which refers to a particular patch.) A complex manifold thus looks locally like $\mathbb{C}^n$. Transition functions from one coordinate patch to another are holomorphic functions. An *atlas* $\{U_a, z_a\}_{a\in A}$ with the above properties defines a *complex structure* on M. If the union of two such atlases has again the same properties, they are said to define the same complex structure; cf. differential structure in the real case, which is defined by (equivalence classes) of C^∞ atlases. n is called the *complex dimension* of M:

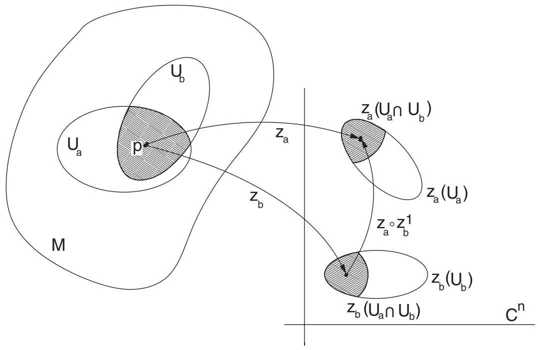


Fig. 1. Coordinate maps on complex manifolds

$n = \dim_{\mathbb{C}}(M)$. Clearly, a complex manifold can be viewed as a real manifold with even (real) dimension, i.e. $m = 2n$. Not all real manifolds of even dimension can be endowed with a complex structure. For instance, among the even-dimensional spheres S^{2n}, only S^2 admits a complex structure. However, direct products of odd-dimensional spheres always admit a complex structure ([24], p.4).

Example 3.1: $\mathbb{C}^n$ is a complex manifold which requires only one single coordinate patch. We can consider $\mathbb{C}^n$ as a real manifold if we identify it with $\mathbb{R}^{2n}$ in the usual way by decomposing the complex coordinates into their real and imaginary parts ($i = \sqrt{-1}$):

$$z^j = x^j + iy^j \ , \quad \bar{z}^j = x^j - iy^j \ , \quad j = 1 \ldots, n \ . \tag{3.1}$$

We will sometimes use the notation $x^{n+j} \equiv y^j$. For later use we give the decomposition of the partial derivatives

$$\partial_j \equiv \frac{\partial}{\partial z^j} = \frac{1}{2}\left(\frac{\partial}{\partial x^j} - i\frac{\partial}{\partial y^j}\right) \ , \quad \bar{\partial}_j \equiv \frac{\partial}{\partial \bar{z}^j} = \frac{1}{2}\left(\frac{\partial}{\partial x^j} + i\frac{\partial}{\partial y^j}\right) \ . \tag{3.2}$$

and the differentials

$$dz^j = dx^j + idy^j \ , \qquad d\bar{z}^j = dx^j - idy^j \ . \tag{3.3}$$

Locally, on any complex manifold, we can always choose real coordinates as the real and imaginary parts of the holomorphic coordinates. A complex manifold is thus also a real analytic manifold. Moreover, since $\det \frac{\partial(x_a^1,\ldots,x_a^{2n})}{\partial(x_b^1,\ldots,x_b^{2n})} = \left|\det \frac{\partial(z_a^i,\ldots,z_a^n)}{\partial(z_b^1,\ldots,\partial z_b^n)}\right|^2 > 0$ on $U_a \cap U_b$, any complex manifold is orientable.

Example 3.2: A very important example, for reasons we will learn momentarily, is n-dimensional *complex projective space* $\mathbb{CP}^n$, or, simply, $\mathbb{P}^n$. $\mathbb{P}^n$ is defined as the set of (complex) lines through the origin of $\mathbb{C}^{n+1}$. A line through the origin can be specified by a single point and two points z and w define the same line iff there exists $\lambda \in \mathbb{C}^* \equiv \mathbb{C} - \{0\}$ such that $z = (z^0, z^1 \ldots, z^n) = (\lambda w^0, \lambda w^1, \ldots, \lambda w^n) \equiv \lambda \cdot w$. We thus have

$$\mathbb{P}^n = \frac{\mathbb{C}^{n+1} - \{0\}}{\mathbb{C}^*} \tag{3.4}$$

The coordinates $z^0, \ldots, z^n$ are called *homogeneous coordinates* on $\mathbb{P}^n$. Often we write $[z] = [z^0 : z^1 : \cdots : z^n]$. $\mathbb{P}^n$ can be covered by $n+1$ coordinate patches $U_i = \{[z] : z^i \neq 0\}$, i.e. U_i consists of those lines through the origin which do not lie in the hyperplane $z^i = 0$. (Hyperplanes in $\mathbb{P}^n$ are $n-1$-dimensional submanifolds, or, more generally, codimension-one submanifolds.) In U_i we can choose local coordinates as $\xi_i^k = \frac{z^k}{z^i}$. They are well defined on U_i and satisfy

$$\xi_i^k = \frac{z^k}{z^i} = \frac{z^k}{z^j} \Big/ \frac{z^i}{z^j} = \frac{\xi_j^k}{\xi_j^i} \tag{3.5}$$

which is holomorphic on $U_i \cap U_j$ where $\xi_j^i \neq 0$. $\mathbb{P}^n$ is thus a complex manifold. The coordinates $\xi_i = (\xi_i^1, \ldots, \xi_i^n)$ are called *inhomogeneous coordinates*. Alternatively to (3.4) we can also define $\mathbb{P}^n$ as $\mathbb{P}^n = S^{2n+1}/U(1)$, where $U(1)$ acts as $z \to e^{i\phi} z$. This shows that $\mathbb{P}^n$ is compact.

Exercise 3.1: Show that $\mathbb{P}^1 \simeq S^2$ by examining transition functions between the two coordinate patches that one obtains after stereographically projecting the sphere onto $\mathbb{C} \cup \{\infty\}$.

A *complex submanifold* X of a complex manifold M^n is a set $X \subset M^n$ which is given locally as the zeroes of a collection $f_1, \ldots, f_k$ of holomorphic functions such that $\mathrm{rank}(J) \equiv \mathrm{rank}\left(\frac{\partial(f_1,\ldots,f_k)}{\partial(z^1,\ldots,z^n)}\right) = k$. X is a complex manifold of dimension $n-k$, or, equivalently, X has codimension k in M^n. The easiest way to show that X is indeed a complex manifold is to choose local coordinates on M such that X is given by $z^1 = z^2 = \cdots = z^k = 0$. It is then clear that if M is a complex manifold so is X. More generally, if we drop the condition on the rank, we get the definition of an *analytic subvariety*. A point $p \in X$ is a *smooth point* if $\mathrm{rank}(J(p)) = k$. Otherwise p is called a *singular point*. For instance, for $k = 1$, at a smooth point there is no simultaneous solution of $p = 0$ and $dp = 0$.

The importance of projective space, or more generally, of weighted projective space which we will encounter later, lies in the following result: there are no compact complex submanifolds of $\mathbb{C}^n$. This is an immediate consequence of the fact that any global holomorphic function on a compact complex manifold is constant, applied to the coordinate functions (for details, see [25], p.10). This is strikingly different from the real analytic case: any real analytic compact or non-compact manifold can be embedded, by a real analytic embedding, into $\mathbb{R}^N$ for sufficiently large N (Grauert-Morrey theorem).

An *algebraic variety* $X \subset \mathbb{P}^n$ is the zero locus in $\mathbb{P}^n$ of a collection of homogeneous polynomials $\{p_\alpha(z^0, \ldots, z^n)\}$. (A function $f(z)$ is homogeneous of degree d if it satisfies $f(\lambda z) = \lambda^d f(z)$. Taking the derivative w.r.t. λ and setting $\lambda = 1$ at the end, leads to the Euler relation $\sum z^i \partial_i f(z) = d \cdot f(z)$.)

More generally one would consider *analytic varieties*, which are defined in terms of holomorphic functions rather than polynomials. However by the *theorem of Chow* every analytic subvariety of $\mathbb{P}^n$ is in fact algebraic. In more sophisticated mathematical language this means that every analytic subvariety of $\mathbb{P}^n$ is the zero section of some positive power of the universal line bundle over $\mathbb{P}^n$, cf. e.g. [23].

An example of an algebraic submanifold of $\mathbb{P}^4$ is the *quintic hypersurface* which is defined as the zero of the polynomial $p(z) = \sum_{i=0}^{4}(z^i)^5$ in $\mathbb{P}^4$. We will see later that this is a three-dimensional Calabi-Yau manifold, and in fact (essentially) the only one that can be written as a hypersurface in $\mathbb{P}^4$, i.e. $X = \{[z^0 : \cdots : z^4] \in \mathbb{P}^4 | p(z) = 0\}$. We can get others by looking at hypersurfaces in products of projective spaces or as complete intersections of more than one hypersurface in higher-dimensional projective spaces and/or products of several projective spaces (here we need several polynomials, homogeneous w.r.t. each $\mathbb{P}^n$). The more interesting generalization is however to enlarge the class of ambient spaces and look at weighted projective spaces.

A weighted projective space is defined much in the same way as a projective space, but with the generalized $\mathbb{C}^*$ action on the homogeneous coordinates

$$\lambda \cdot z = \lambda \cdot (z^0, \ldots, z^n) = (\lambda^{w_0} z^0, \ldots, \lambda^{w_n} z^n) \tag{3.6}$$

where, as before, $\lambda \in \mathbb{C}^*$ and the non-zero integer w_i is called the weight of the homogeneous coordinate z^i. We will consider cases where all weights are positive. However, when one is interested in non-compact situations, one also allows for negative weights. We write $\mathbb{P}^n[w_0, \ldots, w_n] \equiv \mathbb{P}^n[\boldsymbol{w}]$.

Different sets of weights may give isomorphic spaces. A simple example is $\mathbb{P}^n[k\boldsymbol{w}] \simeq \mathbb{P}^n[\boldsymbol{w}]$. One may show that one covers all isomorphism classes if one restricts to so-called *well-formed* spaces [29]. Among the $n+1$ weights of a well-formed space no set of n weights has a common factor. E.g. $\mathbb{P}^2[1,2,2]$ is not well formed whereas $\mathbb{P}^2[1,1,2]$ is.

Weighted projective spaces are singular, which is most easily demonstrated by means of an example. Consider $\mathbb{P}^2[1,1,2]$, i.e. (z^0, z^1, z^2) and $(\lambda z^0, \lambda z^2, \lambda^2 z^2)$ denote the same point. For $\lambda = -1$ the point $[0 : 0 : z^2] \equiv [0 : 0 : 1]$ is fixed but λ acts non-trivially on its neighborhood: we have a $\mathbb{Z}_2$ orbifold singularity at this point. This singularity has locally the form $\mathbb{C}^2/\mathbb{Z}_2$, where $\mathbb{Z}_2$ acts on the coordinates (x^1, x^2) of $\mathbb{C}^2$ as $\mathbb{Z}_2 : (x^1, x^2) \mapsto -(x^1, x^2)$. In general there is a fixed point for every weight greater than one, a fixed curve for every pair of weights with a common factor greater than one and so on.

A hypersurface $X_d[\boldsymbol{w}]$ in weighted projective space is defined as the vanishing locus of a *quasi-homogeneous* polynomial, $p(\lambda \cdot z) = \lambda^d p(z)$, where d is the degree of $p(z)$, i.e.

$$X_d[\boldsymbol{w}] = \left\{ [z^0 : \cdots : z^b] \in \mathbb{P}^n[\boldsymbol{w}] \,\middle|\, p(z) = 0 \right\} \tag{3.7}$$

In this case the Euler relation generalizes to $\sum w_i z^i \partial_i p(z) = d \cdot p(z)$.

Exercise 3.2: Of how many points consist the following "hypersurfaces"? (1): $(z^0)^2 + (z^1)^2 = 0$ in $\mathbb{P}^1$; (2): $(z^0)^3 + (z^1)^2 = 0$ in $\mathbb{P}^1[2,3]$. The number of points is equal to the Euler number (the Euler number of a smooth point is one, as can be seen from the Euler formula $\chi = \#\text{vertices} - \#\text{edges} + \#\text{two dimensional faces} \mp \ldots$ of a triangulated space. This also follows from the familiar fact that after removing two points from a sphere with Euler number two one obtains a cylinder whose Euler number is zero).

It can happen that the hypersurface does not pass through the singularities of the ambient space. Take again the example $\mathbb{P}^2[1,1,2]$ and consider the quartic hypersurface. At the fixed point $[0 : 0 : z^2]$ only the monomial $(z^2)^2$ survives and the hypersurface constraint would require that $z^2 = 0$. But the point $z^0 = z^1 = z^2 = 0$ is not in $\mathbb{P}^2[1,1,2]$. As a second example consider $\mathbb{P}^3[1,1,2,2]$. We now find a singular curve rather than a singular point, namely $z^0 = z^1 = 0$ and a generic hypersurface will intersect this curve in isolated points. To obtain a smooth manifold one has to resolve the singularity, which in this example is a $\mathbb{Z}_2$ singularity. We will not discuss the process of resolution of the singularities but it is mathematically well defined and under control and most efficiently described within the language of toric geometry [30, 31].

Weighted projective spaces are still not the most general ambient spaces one considers in actual string compactifications, in particular when one considers mirror symmetry (see below). The more general concept is that of a toric variety. Toric varieties have some very simple features which allow one to reduce many calculations to combinatorics. Weighted projective spaces are a small subclass of toric varieties. For details we refer to Chap. 7 of [18] and to [31, 32].

We have seen that any complex manifold M can be viewed as a real (analytic) manifold. The tangent space at a point p is denoted by $T_p(M)$ and the tangent bundle by $T(M)$. The *complexified tangent bundle* $T_{\mathbb{C}}(M) = T(M) \otimes \mathbb{C}$ consists of all tangent vectors of M with complex coefficients, i.e. $v = \sum_{j=1}^{2n} v^j \frac{\partial}{\partial x^j}$ with $v^i \in \mathbb{C}$. With the help of (3.2) we can write this as

$$\begin{aligned} v &= \sum_{j=1}^{2n} v^j \frac{\partial}{\partial x^j} = \sum_{j=1}^{n} (v^j + i v^{n+j}) \partial_j + \sum_{j=1}^{n} (v^j - i v^{n+j}) \bar{\partial}_j \\ &\equiv v^{1,0} + v^{0,1} \end{aligned} \tag{3.8}$$

We have thus a decomposition

$$T_{\mathbb{C}}(M) = T^{1,0}(M) \oplus T^{0,1}(M) \tag{3.9}$$

into *vectors of type* $(1,0)$ and of type $(0,1)$: $T^{1,0}(M)$ is spanned by $\{\partial_i\}$ and $T^{0,1}(M)$ by $\{\bar{\partial}_i\}$. Note that $T_p^{0,1}(M) = \overline{T_p^{1,0}(M)}$ and that the splitting into the two subspaces is preserved under holomorphic coordinate changes. The transition functions of $T^{1,0}(M)$ are holomorphic, and we therefore call it the holomorphic tangent bundle. A holomorphic section of $T^{1,0}(M)$ is called a *holomorphic vector field*; its component functions are holomorphic.

$T^{1,0}(M)$ is just one particular example of a *holomorphic vector bundle* $E \xrightarrow{\pi} M$. Holomorphic vector bundles of rank k are characterized by their holomorphic transition functions which are elements of $Gl(k,\mathbb{C})$ (rather than $Gl(n,\mathbb{R})$ as in the real case) with holomorphic matrix elements.

In the same way as in (3.9) we decompose the dual space, the space of one-forms:

$$T^*_{\mathbb{C}}(M) = T^{*1,0}(M) \oplus T^{*0,1}(M)\ . \tag{3.10}$$

$T^{*1,0}(M)$ and $T^{*0,1}(M)$ are spanned by $\{dz^i\}$ and $\{d\bar{z}^i\}$, respectively. By taking tensor products we can define differential forms of type (p,q) as sections of $\overset{p}{\wedge} T^{*1,0}(M) \overset{q}{\wedge} T^{*0,1}(M)$. The space of (p,q)-forms will be denoted by $A^{p,q}$. Clearly $\overline{A^{p,q}} = A^{q,p}$. If we denote the space of sections of $\overset{r}{\wedge} T^*_{\mathbb{C}}(M)$ by A^r, we have the decomposition

$$A^r = \bigoplus_{p+q=r} A^{p,q}\ . \tag{3.11}$$

This decomposition is independent of the choice of local coordinate system.

Using the underlying real analytic structure we can define the exterior derivative d. If $\omega \in A^{p,q}$, then

$$d\omega \in A^{p+1,q} \oplus A^{p,q+1}\ . \tag{3.12}$$

We write $d\omega = \partial\omega + \bar{\partial}\omega$ with $\partial\omega \in A^{p+1,q}$ and $\bar{\partial}\omega \in A^{p,q+1}$. This defines the two operators

$$\partial :\ A^{p,q} \to A^{p+1,q}\ , \qquad \bar{\partial} :\ A^{p,q} \to A^{p,q+1}\ , \tag{3.13}$$

and

$$d = \partial + \bar{\partial}\ . \tag{3.14}$$

The following results are easy to verify:

$$d^2 = (\partial + \bar{\partial})^2 \equiv 0 \quad \Rightarrow \quad \partial^2 = 0\ , \quad \bar{\partial}^2 = 0\ , \quad \partial\bar{\partial} + \bar{\partial}\partial = 0\ . \tag{3.15}$$

Here we used that $\partial^2 : A^{p,q} \to A^{p+2,q}$, $\bar{\partial}^2 : A^{p,q} \to A^{p,q+2}$, $(\partial\bar{\partial} + \bar{\partial}\partial) : A^{p,q} \to A^{p+1,q+1}$, i.e. that the three operators map to three different spaces. They must thus vanish separately.

Equation (3.14) is not true on an almost complex manifold. Alternative to the way we have defined complex structures we could have started with an almost complex structure – a differentiable isomorphism $J : T(M) \to T(M)$ with $J^2 = -\mathbb{1}$ – such that the splitting (3.9) of $T(M)$ is into eigenspaces of J with eigenvalues $+i$ and $-i$, respectively. Then (3.12) would be replaced by $d\omega \in A^{p+2,q-1} \oplus A^{p+1,q} \oplus A^{p,q+1} \oplus A^{p-1,q+2}$ and (3.14) by $d = \partial + \bar{\partial} + \ldots$. Only if the almost complex structure satisfies an integrability condition – the vanishing of the Nijenhuis tensor – do (3.12) and (3.14) hold. A theorem of Newlander and Nierenberg then guarantees that we can construct on M an atlas of holomorphic charts and M is a complex manifold in the sense of the definition that we have given, see e.g. [13, 25].

ω is called a *holomorphic p-form* if it is of type $(p, 0)$ and $\bar{\partial}\omega = 0$, i.e. if it has holomorphic coefficient functions. Likewise $\bar{\omega}$ of type $(0, q)$ with $\bar{\partial}\bar{\omega} = 0$ is called *anti-holomorphic*. $\Omega^p(M)$ denotes the vector-space of holomorphic p-forms. We leave it as an exercise to write down the explicit expressions, in terms of coefficients, of $\partial\omega$, etc.

3.2 Kähler Manifolds

The next step is to introduce additional structures on a complex manifold: a hermitian metric and a hermitian connection.

A *hermitian metric* is a covariant tensor field of the form $\sum_{i,j=1}^{n} g_{i\bar{j}} dz^i \otimes d\bar{z}^j$, where $g_{i\bar{j}} = g_{i\bar{j}}(z)$ (here the notation is not to indicate that the components are holomorphic functions; they are not!) such that $g_{j\bar{i}}(z) = \overline{g_{i\bar{j}}(z)}$ and $g_{i\bar{j}}(z)$ is a positive definite matrix, that is, for any $\{v^i\} \in \mathbb{C}^n$, $v^i g_{i\bar{j}} \bar{v}^{\bar{j}} \geq 0$ with equality only if all $v^i = 0$. To any hermitian metric we associate a $(1, 1)$-form

$$\omega = i \sum_{i,j=1}^{n} g_{i\bar{j}} dz^i \wedge d\bar{z}^j \; . \tag{3.16}$$

ω is called the *fundamental form* associated with the hermitian metric g.

Exercise 3.3: Show that ω is a real $(1, 1)$-form, i.e. that $\omega = \bar{\omega}$.

One can introduce a hermitian metric on any complex manifold (see e.g. [26], p. 145).

Exercise 3.4: Show that

$$\begin{aligned}\frac{\omega^n}{n!} &= (i)^n g(z) dz^1 \wedge d\bar{z}^1 \wedge \cdots \wedge dz^n \wedge d\bar{z}^n \\ &= (i)^n (-1)^{n(n-1)/2} g(z) dz^1 \wedge \cdots \wedge dz^n \wedge d\bar{z}^1 \wedge \cdots \wedge d\bar{z}^n \\ &= 2^n g(z) dx^1 \wedge \cdots \wedge dx^{2n}\end{aligned} \tag{3.17}$$

where $\omega^n = \underbrace{\omega \wedge \cdots \wedge \omega}_{n \text{ factors}}$ and $g = \det(g_{i\bar{j}}) > 0$. ω^n is thus a good volume form on M. This shows once more that complex manifolds always possess an orientation.

The inverse of the hermitian metric is $g^{i\bar{j}}$ which satisfies $g^{j\bar{\imath}}g_{j\bar{k}} = \delta^{\bar{\imath}}_{\bar{k}}$ and $g_{i\bar{j}}g^{k\bar{j}} = \delta^k_i$ (summation convention used). We use the metric to raise and lower indices, whereby they change their type. Note that under holomorphic coordinate changes, the index structure of the metric is preserved, as is that of any other tensor field.

A hermitian metric g whose associated fundamental form ω is closed, i.e. $d\omega = 0$, is called a *Kähler metric*. A complex manifold endowed with a Kähler metric is called a *Kähler manifold*. ω is the *Kähler form*. An immediate consequence of $d\omega = 0 \Rightarrow \partial\omega = \bar{\partial}\omega = 0$ is

$$\partial_i g_{j\bar{k}} = \partial_j g_{i\bar{k}} \,, \qquad \bar{\partial}_{\bar{\imath}} g_{j\bar{k}} = \bar{\partial}_{\bar{k}} g_{j\bar{\imath}} \qquad \text{(Kähler condition)} \,. \tag{3.18}$$

From this one finds immediately that the only non-zero coefficients of the Riemannian connection are

$$\Gamma^k_{ij} = g^{k\bar{l}}\partial_i g_{j\bar{l}} \,, \qquad \Gamma^{\bar{k}}_{\bar{\imath}\bar{j}} = g^{l\bar{k}}\bar{\partial}_{\bar{\imath}} g_{l\bar{j}} \,. \tag{3.19}$$

The vanishing of the connection coefficients with mixed indices is a necessary and sufficient condition that under parallel transport the holomorphic and the anti-holomorphic tangent spaces do not mix (see below).

Note that while all complex manifolds admit a hermitian metric, this does not hold for Kähler metrics. Counterexamples are quaternionic manifolds which appear as moduli spaces of type II compactifications on Calabi-Yau manifolds. Another example is $S^{2p+1} \otimes S^{2q+1}$, $q > 1$. A complex submanifold X of a Kähler manifold M is again a Kähler manifold, with the induced Kähler metric. This follows easily if one goes to local coordinates on M where X is given by $z^1 = \cdots = z^k = 0$.

From (3.18) we also infer the *local* existence of a real *Kähler potential* K in terms of which the Kähler metric can be written as

$$g_{i\bar{j}} = \partial_i \bar{\partial}_{\bar{j}} K \tag{3.20}$$

or, equivalently, $\omega = i\partial\bar{\partial}K$. The Kähler potential is not uniquely defined: $K(z, \bar{z})$ and $K(z, \bar{z}) + f(z) + \bar{f}(\bar{z})$ lead to the same metric if f and $\bar{f}$ are holomorphic and anti-holomorphic functions (on the patch on which K is defined), respectively.

From now on, unless stated otherwise, we will restrict ourselves to Kähler manifolds; some of the results are, however, true for arbitrary complex manifolds. Also, if in doubt, assume that the manifold is compact.

Exercise 3.5: Determine a *hermitian connection* by the two requirements: (1) The only non-vanishing coefficients are Γ^i_{jk} and $\Gamma^{\bar{\imath}}_{\bar{j}\bar{k}}$ and (2) $\nabla_i g_{j\bar{k}} = 0$. Show that the connection is torsionfree, i.e. $T^k_{ij} \equiv \Gamma^k_{ij} - \Gamma^k_{ji} = 0$ if g is a Kähler metric. Check that the connection so obtained is precisely the Riemannian connection, i.e. the hermitian and the Riemannian structures on a Kähler manifold are compatible.

Exercise 3.6: Derive the components of the Riemann tensor on a Kähler manifold. Show that the only non-vanishing components of the Riemann tensor are those with the index structure $R_{i\bar{j}k\bar{l}}$ and those related by symmetries. In particular the components of the type R_{ij**} are zero. Show that the non-vanishing components are

$$R_{i\bar{j}k\bar{l}} = -\partial_i\bar{\partial}_{\bar{j}}g_{k\bar{l}} + g^{m\bar{n}}(\partial_i g_{k\bar{n}})(\bar{\partial}_{\bar{j}}g_{m\bar{l}}) \ . \tag{3.21}$$

Here the sign conventions are such that $[\nabla_i, \nabla_{\bar{j}}]V_k = -R_{i\bar{j}k}{}^l V_l$.

Exercise 3.7: The Ricci tensor is defined as $R_{i\bar{j}} \equiv g^{k\bar{l}}R_{i\bar{j}k\bar{l}}$. Prove that

$$R_{i\bar{j}} = -\partial_i\bar{\partial}_{\bar{j}}(\log\det g) \ . \tag{3.22}$$

Show that this is the same (up to a sign) as $R_{i\mu\bar{j}}{}^{\mu} = R_{i\mu\bar{j}\nu}g^{\mu\nu}$, $\mu = (k, \bar{k})$.

One also defines the *Ricci-form* (of type $(1,1)$) as

$$\mathcal{R} = iR_{j\bar{k}}dz^j \wedge d\bar{z}^k = -i\partial\bar{\partial}\log(\det\, g) \tag{3.23}$$

which satisfies $d\mathcal{R} = 0$. Note that $\log(\det\, g)$ is not a globally defined function since $\det g$ transforms as a density under change of coordinates. $\mathcal{R}$ is however globally defined (why?).

We learn from (3.23) that the Ricci form depends only on the volume form of the Kähler metric and on the complex structure (through ∂ and $\bar{\partial}$). Under a change of metric, $g \to g'$, the Ricci form changes as

$$\mathcal{R}(g') = \mathcal{R}(g) - i\partial\bar{\partial}\log\left(\frac{\det(g'_{k\bar{l}})}{\det(g_{k\bar{l}})}\right) \ , \tag{3.24}$$

where the ratio of the two determinants is a globally defined non-vanishing function on M.

Example 3.3: Complex projective space. To demonstrate that it is a Kähler manifold we give an explicit metric, the so called *Fubini-Study metric*. Recall that $\mathbb{P}^n = \{[z^0 : \cdots : z^n]; 0 \neq (z^0 : \cdots : z^n) \in \mathbb{C}^{n+1}\}$ and $U_0 = \{[1, z^1 : \cdots : z^n]\} \simeq \mathbb{C}^n$ is an open subset of $\mathbb{P}^n$. Set

$$g_{i\bar{j}} = \partial_i\partial_{\bar{j}}\log(1 + |z^1|^2 + \cdots + |z^n|^2) \equiv \partial_i\partial_{\bar{j}}\ln(1 + |z|^2) \tag{3.25}$$

or, equivalently,

$$\omega = i\partial\bar{\partial}\log(1 + |z|^2) = i\left(\frac{dz^i \wedge d\bar{z}^i}{1 + |z|^2} - \frac{\bar{z}^i dz^i \wedge z^j d\bar{z}^j}{(1 + |z|^2)^2}\right) \tag{3.26}$$

Closure of ω is obvious if one uses (3.15). From (3.25) we also immediately read off the Kähler potential of the Fubini-Study metric (cf. (3.20)) on U_0. Clearly this is only defined locally.

Exercise 3.8: Show that for any non-zero vector u, $u^i g_{i\bar{j}} \bar{u}^{\bar{j}} \geq 0$ to prove positive definiteness of the Fubini-Study metric.

On the other hand, ω is globally defined on $\mathbb{P}^n$. To see this, let $U_1 = \{(w^0, 1, w^2, \ldots, w^n)\} \subset \mathbb{P}^n$ and check what happens to ω on the overlap $U_0 \cap U_1 = \{[1 : z^1 : \cdots : z^n] = [w^0 : 1 : w^2 : \cdots : w^n]\}$, where $z^i = \frac{w^i}{w^0}$, for all $i \neq 1$ and $z^1 = \frac{1}{w^0}$. Then

$$\begin{aligned}\omega &= i\partial\bar{\partial}\log(1 + |z^1|^2 + \cdots + |z^n|^2) = i\partial\bar{\partial}\log\left(1 + \frac{1}{|w^0|^2} + \sum_{i=2}^{n} \frac{|w^i|^2}{|w^0|^2}\right) \\ &= i\Big(\partial\bar{\partial}\log(1 + |w|^2) - \partial\bar{\partial}\log(|w^0|^2)\Big) = i\partial\bar{\partial}\log(1 + |w|^2)\end{aligned} \tag{3.27}$$

since w^0 is holomorphic on $U_0 \cap U_1$. So ω and the corresponding Kähler metric are globally defined. Complex projective space is thus a Kähler manifold and so is every complex submanifold. With[3]

$$\det(g_{i\bar{j}}) = \frac{1}{(1 + |z|^2)^{n+1}} \tag{3.28}$$

one finds

$$R_{i\bar{j}} = -\partial_i \bar{\partial}_{\bar{j}} \log\left(\frac{1}{(1 + |z|^2)^{n+1}}\right) = (n+1) g_{i\bar{j}} \tag{3.29}$$

which shows that the Fubini-Study metric is a *Kähler-Einstein metric* and $\mathbb{P}^n$ a *Kähler-Einstein manifold*.

3.3 Holonomy Group of Kähler Manifolds

The next topic we want to discuss is the *holonomy group of Kähler manifolds*. Recall that the holonomy group of a Riemannian manifold of (real) dimension m is a subgroup of $O(m)$. It follows immediately from the index structure of the connection coefficients of a Kähler manifold that under parallel transport elements of $T^{1,0}(M)$ and $T^{0,1}(M)$ do not mix. Since the length of a vector does not change under parallel transport, the holonomy group of a Kähler manifold is a subgroup of $U(n)$ where n is the complex dimension of the manifold.[4] In particular, elements of $T^{1,0}(M)$ transform as $\boldsymbol{n}$ and elements of $T^{0,1}(M)$ as $\overline{\boldsymbol{n}}$ of $U(n)$. Consider now parallel transport around an infinitesimal loop in the (μ, ν)-plane with area $\delta a^{\mu\nu} = -\delta a^{\nu\mu}$. Under parallel transport around this loop a vector V changes by an amount δV given in (A.14). In complex coordinates this is $\delta V^i = -\delta a^{k\bar{l}} R_{k\bar{l}}{}^i{}_j V^j$. From what we

[3] To show this, use $\det(\delta_{ij} - v_i v_{\bar{j}}) = \exp\left(\mathrm{tr}\,\log(\delta_{ij} - v_i v_{\bar{j}})\right) = \exp\left(\log(1 - |v|^2)\right) = (1 - |v|^2)$.

[4] The unitary group $U(n)$ is the set of all complex $n \times n$ matrices which leave invariant a hermitian metric $\overline{g_{i\bar{j}}} = g_{j\bar{\imath}}$, i.e. $UgU^\dagger = g$. For the choice $g_{i\bar{j}} = \delta_{ij}$ one obtains the familiar condition $UU^\dagger = \mathbb{1}$.

said above it follows that on a Kähler manifold the matrix $-\delta a^{k\bar{l}} R_{k\bar{l}i}{}^{j}$ must be an element of the Lie algebra $u(n)$. The trace of this matrix, which is proportional to the Ricci tensor, generates the $u(1)$ part in the decomposition $u(n) \simeq su(n) \oplus u(1)$. We thus learn that the holonomy group of a Ricci-flat Kähler manifold is a subgroup of $SU(n)$. Conversely, one can show that any $2n$-dimensional manifold with $U(n)$ holonomy admits a Kähler metric and if it has $SU(n)$ holonomy it admits a Ricci-flat Kähler metric. This uses the fact that holomorphic and anti-holomorphic indices do not mix, which implies that all connection coefficients with mixed indices must vanish. One then proceeds with the explicitly construction of an almost complex structure with vanishing Nijenhuis tensor. Details can be found in [4, 28].

We should mention that strictly speaking the last argument is only valid for the restricted holonomy group $\mathcal{H}_0$ (which is generated by parallel transport around contractible loops). Also, in general only the holonomy around *infinitesimal* loops is generated by the Riemann tensor. For finite (but still contractible) loops, derivatives of the Riemann tensor of arbitrary order will appear [12]. For Kähler manifolds we do however have the $U(n)$ invariant split of the indices $\mu = (i, \bar{\imath})$ and $U(n)$ is a maximal compact subgroup of $SO(2n)$. Thus the restricted holonomy group is not bigger than $U(n)$ For simply connected manifolds the restricted holonomy group is already the full holonomy group. For non-simply connected manifolds the full holonomy group and the restricted holonomy group may differ. Their quotient is countable and the restricted holonomy group is the identity component of the full holonomy group, i.e. for a generic Riemannian manifold it is $SO(m)$ (cf. [12]).

3.4 Cohomology of Kähler Manifolds

Before turning to the next subject, *homology and cohomology* on complex manifolds, we will give a very brief and incomplete summary of these concepts in the real situation, which, of course, also applies to complex manifolds, if they are viewed as real analytic manifolds.

On a smooth, connected manifold M one defines p-chains a_p as formal sums $a_p = \sum_i c_i N_i$ of p-dimensional oriented submanifolds on M. If the coefficients c_i are real (complex, integer), one speaks of real (complex, integral) chains. Define ∂ as the operation of taking the boundary with the induced orientation. $\partial a \equiv \sum c_i \partial N_i$ is then a $p-1$-chain. Let $Z_p = \{a_p | \partial a_p = \emptyset\}$ be the set of *cycles*, i.e. the set of chains without boundary and let $B_p = \{\partial a_{p+1}\}$ be the set of *boundaries*. Since $\partial\partial a_p = \emptyset$, $B_p \subset Z_p$. The *p-th homology group of* M is defined as

$$H_p = Z_p / B_p \, . \tag{3.30}$$

Depending on the coefficient group one gets $H_p(M, \mathbb{R})$, $H_p(M, \mathbb{C})$, $H_p(M, \mathbb{Z})$, etc. Elements of H_p are equivalence classes of cycles $z_p \simeq z_p + \partial a_{p+1}$, called *homology classes* and denoted by $[z_p]$.

One version of *Poincaré duality* is the following isomorphism between homology groups, valid on orientable connected smooth manifolds of real dimension m:

$$H_r(M,\mathbb{R}) \simeq H_{m-r}(M,\mathbb{R}) \; . \tag{3.31}$$

One defines the *r-th Betti number* b_r as

$$b_r = \dim(H_r(M,\mathbb{R})) \; . \tag{3.32}$$

They are toplogical invariants of M. As a consequence of Poincaré duality,

$$b_r(M) = b_{m-r}(M) \; . \tag{3.33}$$

We now turn to de Rham cohomology, which is defined with the exterior derivative operator $d : A^r \to A^{r+1}$. Let Z^p be the set of *closed p-forms*, i.e. $Z^p = \{\omega_p | d\omega_p = 0\}$ and let B^p be the set of *exact p-forms* $B^p = \{d\omega_{p-1}\}$. The *de Rham cohomology groups* H^p are defined as the quotients

$$H^p_{\mathrm{D.R.}} = Z^p/B^p \; . \tag{3.34}$$

Elements of H^p are equivalence classes of closed forms $\omega_p \simeq \omega_p + d\alpha_{p-1}$, called cohomology classes and denoted by $[\omega_p]$. Each equivalence class possesses one harmonic representative, i.e. a zero mode of the Laplacian $\Delta = dd^* + d^*d$. The action of Δ on p-forms is

$$\Delta\omega_{\mu_1\cdots\mu_p} = -\nabla^\nu\nabla_\nu\omega_{\mu_1\cdots\mu_p} - pR_{\nu[\mu_1}\omega^\nu{}_{\mu_2\cdots\mu_p]} - \frac{1}{2}p(p-1)R_{\nu\rho[\mu_1\mu_2}\omega^{\nu\rho}{}_{\mu_3\cdots\mu_p]} \; . \tag{3.35}$$

Since the number of (normalizable) harmonic forms on a compact manifold is finite, the Betti numbers are all finite.

Exercise 3.9: Derive (3.35).

Given both the homology and the cohomology classes, we can define an inner product

$$\pi(z_p,\omega_p) = \int_{z_p} \omega_p \; , \tag{3.36}$$

where $\pi(z_p,\omega_p)$ is called a *period* (of ω_p). We speak of an *integral cohomology class* $[\omega_p] \in H_{\mathrm{D.R.}}(M,\mathbb{Z})$ if the period over any integral cycle is integer.

Exercise 3.10: Prove, using Stoke's theorem, that the integral does not depend on which representatives of the two classes are chosen.

A theorem of de Rham ensures that the above inner product between homology and cohomology classes is bilinear and non-degenerate, thus establishing an isomorphism between homology and cohomology. The following two facts are consequences of de Rham's theorem:

(1) Given a basis $\{z_i\}$ for H_p and any set of periods ν_i, $i = 1, \ldots, b_p$, there exists a closed p-from ω such that $\pi(z_i, \omega) = \nu_i$.
(2) If all periods of a p-form vanish, ω is exact.

Another consequence of de Rham's theorem is the following important result: Given any p-cycle z there exists a closed $(m-p)$-form α, called the *Poincaré dual of* z such that for any closed p-form ω

$$\int_z \omega = \int_M \alpha \wedge \omega \ . \tag{3.37}$$

Since ω is closed, α is only defined up to an exact form. In terms of their Poincaré duals α and β we can define the *intersection number* $A \cdot B$ between a p-cycle A and an $(m-p)$-cycle B as

$$A \cdot B = \int_M \alpha \wedge \beta \ . \tag{3.38}$$

This notion is familiar from Riemann surfaces.

So much for the collection of some facts about homology and cohomology on real manifolds. They are also valid on complex manifolds if one views them as real analytic manifolds. However one can use the complex structure to define (among several others) the so-called *Dolbeault cohomology* or $\bar\partial$-cohomology. As the (second) name already indicates, it is defined w.r.t. the operator $\bar\partial : A^{p,q}(M) \to A^{p,q+1}(M)$. A (p,q)-form α is $\bar\partial$-*closed* if $\bar\partial\alpha = 0$. The space of $\bar\partial$-closed (p,q)-forms is denoted by $Z^{p,q}_{\bar\partial}$. A (p,q)-form β is $\bar\partial$-*exact* if it is of the form $\beta = \bar\partial\gamma$ for $\gamma \in A^{p,q-1}$. Since $\bar\partial^2 = 0$, $\bar\partial(A^{p,q}(M)) \subset Z^{p,q+1}_{\bar\partial}(M)$. The Dolbeault cohomology groups are then defined as

$$H^{p,q}_{\bar\partial}(M) = \frac{Z^{p,q}_{\bar\partial}(M)}{\bar\partial(A^{p,q-1}(M))} \ . \tag{3.39}$$

There is a lemma (by Dolbeault) analogous to the Poincaré-lemma, which ensures that the Dolbeault cohomology groups (for $q \geq 1$) are locally[5] trivial. This is also referred to as the $\bar\partial$-*Poincaré lemma*.

The dimensions of the (p,q) cohomology groups are called *Hodge numbers*

$$h^{p,q}(M) = \dim_{\mathbb{C}}(H^{p,q}_{\bar\partial}(M)) \ . \tag{3.40}$$

They are finite for compact, complex manifolds [23]. The Hodge numbers of a Kähler manifold are often arranged in the *Hodge diamond*:

[5] More precisely, on polydiscs $P_r = \{z \in \mathbb{C}^n | z^i| < r, \text{ for all } i = 1, \ldots, n\}$.

$$\begin{array}{ccccccc} & & & h^{0,0} & & & \\ & & h^{1,0} & & h^{0,1} & & \\ & h^{2,0} & & h^{1,1} & & h^{0,2} & \\ h^{3,0} & & h^{2,1} & & h^{1,2} & & h^{0,3} \\ & h^{3,1} & & h^{2,2} & & h^{1,3} & \\ & & h^{3,2} & & h^{2,3} & & \\ & & & h^{3,3} & & & \end{array} \tag{3.41}$$

which we have displayed here for a three complex dimensional Kähler manifold. We will later show that for a Calabi-Yau manifold of the same dimension the only independent Hodge numbers are $h^{1,1}$ and $h^{2,1}$.

We can now define a scalar product between two forms, φ and ψ, of type (p,q):[6]

$$\psi = \frac{1}{p!q!}\psi_{i_1\ldots i_p\bar{\jmath}_1\ldots\bar{\jmath}_q}(z)dz^{i_1}\wedge\cdots\wedge dz^{i_p}\wedge d\bar{z}^{j_1}\wedge\cdots\wedge d\bar{z}^{j_q} \tag{3.42}$$

and likewise for φ. First define

$$(\varphi,\psi)(z) = \frac{1}{p!q!}\varphi_{i_1\ldots i_p\bar{\jmath}_1\ldots\bar{\jmath}_q}(z)\bar{\psi}^{i_1\ldots i_p\bar{\jmath}_1\ldots\bar{\jmath}_q}(z) \tag{3.43}$$

where

$$\bar{\psi}^{i_1\ldots i_p\bar{\jmath}_1\ldots\bar{\jmath}_q}(z) = g^{i_1\bar{k}_1}\cdots g^{i_p\bar{k}_p}g^{l_1\bar{\jmath}_1}\cdots g^{l_q\bar{\jmath}_q}\overline{\psi_{k_1\ldots k_p\bar{l}_1\ldots\bar{l}_q}(z)}\,. \tag{3.44}$$

Later we will also need the definition

$$\bar{\psi} = \frac{1}{p!q!}\overline{\psi_{i_1\ldots i_p\bar{\jmath}_1\ldots\bar{\jmath}_q}dz^{i_1}\wedge\cdots\wedge d\bar{z}^{j_q}} = \frac{1}{p!q!}\bar{\psi}_{j_i\ldots j_q\bar{\imath}_1\ldots\bar{\imath}_p}dz^{j_1}\wedge\cdots\wedge d\bar{z}^{i_p}\ , \tag{3.45}$$

where

$$\overline{\psi_{k_1\ldots k_p\bar{l}_1\ldots\bar{l}_q}} = (-1)^{pq}\bar{\psi}_{l_1\ldots l_q\bar{k}_1\ldots\bar{k}_p}\ . \tag{3.46}$$

The inner product $(\ ,\ ): A^{p,q}\times A^{p,q}\to\mathbb{C}$ is then

$$(\varphi,\psi) = \int_M(\varphi,\psi)(z)\frac{\omega^n}{n!}\ . \tag{3.47}$$

The following two properties are easy to verify:

$$\begin{aligned}&(\psi,\varphi) = \overline{(\varphi,\psi)}\ ,\\ &(\varphi,\varphi)\geq 0 \quad \text{with equality only for } \varphi=0\ .\end{aligned} \tag{3.48}$$

[6] A good and detailed reference for the following discussion is the third chapter of [26].

We define the Hodge-$*$ operator $* : A^{p,q} \to A^{n-q,n-p}$, $\psi \mapsto *\psi$ by requiring[7]

$$(\varphi, \psi)(z) \frac{\omega^n}{n!} = \varphi(z) \wedge *\bar{\psi}(z) \ . \tag{3.49}$$

Exercise 3.11: Show that for $\psi \in A^{p,q}$,

$$\begin{aligned} * \, \psi = & \tfrac{(i)^n (-1)^{n(n-1)/2+np}}{p!q!(n-p)!(n-q)!} g \epsilon^{m_1 \dots m_p}{}_{\bar{j}_1 \dots \bar{j}_{n-p}} \epsilon^{\bar{n}_1 \dots \bar{n}_q}{}_{l_1 \dots l_{n-q}} \\ & \cdot \psi_{m_1 \dots m_p \bar{n}_1 \dots \bar{n}_q} dz^{l_1} \wedge \dots \wedge dz^{l_{n-q}} \wedge d\bar{z}^{j_1} \wedge \dots \wedge d\bar{z}^{j_{n-p}} \in A^{n-q,n-p} \ . \end{aligned} \tag{3.50}$$

Here we defined $\epsilon_{i_1 \dots i_n} = \pm 1$ and its indices are raised with the metric, as usual; i.e. $\epsilon^{\bar{j}_1 \cdots \bar{j}_n} = \pm g^{-1}$.

Exercise 3.12: Prove the following properties of the $*$-operator:

$$\begin{aligned} * \, \bar{\psi} &= \overline{*\psi} \ , \\ **\psi &= (-1)^{p+q} \, \psi \ , \quad \psi \in A^{p,q} \ . \end{aligned} \tag{3.51}$$

Exercise 3.13: For ω the fundamental form and α an arbitrary real $(1,1)$-form, derive the following two identities, valid on a three-dimensional Kähler manifold:

$$\begin{aligned} * \, \alpha &= \frac{1}{2} (\omega, \alpha)(z) \, \omega \wedge \omega - \alpha \wedge \omega \ , \\ * \, \omega &= \frac{1}{2} \omega \wedge \omega \ . \end{aligned} \tag{3.52}$$

Exercise 3.14: Show that on a three-dimensional complex manifold for $\Omega \in A^{3,0}$ and $\alpha \in A^{2,1}$,

$$\begin{aligned} * \, \Omega &= -i \Omega \ , \\ * \, \alpha &= \quad i \alpha \ . \end{aligned} \tag{3.53}$$

Given the scalar product (3.47), we can define the adjoint of the $\bar{\partial}$ operator, $\bar{\partial}^* : A^{p,q}(M) \to A^{p,q-1}(M)$ via

$$(\bar{\partial}^* \psi, \varphi) = (\psi, \bar{\partial} \varphi) \ , \qquad \forall \varphi \in A^{p,q-1}(M) \ . \tag{3.54}$$

Exercise 3.15: Show that on M compact,

$$\bar{\partial}^* = - * \partial * \ . \tag{3.55}$$

Exercise 3.16: Show that, given a (p,q)-form ψ,

$$(\bar{\partial}^* \psi)_{i_1 \dots i_p \bar{j}_2 \dots \bar{j}_q} = (-1)^{p+1} \nabla^{\bar{j}_1} \psi_{i_1 \dots i_p \bar{j}_1 \dots \bar{j}_q} \ . \tag{3.56}$$

We now define the $\bar{\partial}$*-Laplacian* as

[7] Note that there are several differing notations in the literature; e.g. Griffiths and Harris define an operator $*_{\rm GH} : A^{p,q} \to A^{n-p,n-q}$. What they call $*_{\rm GH} \, \psi$ we have called $*\bar{\psi}$.

$$\Delta_{\bar\partial} = \bar\partial\bar\partial^* + \bar\partial^*\bar\partial\,, \qquad \Delta_{\bar\partial} : A^{p,q}(M) \to A^{p,q}(M) \tag{3.57}$$

and call ψ a ($\bar\partial-$)*harmonic form* if it satisfies

$$\Delta_{\bar\partial}\psi = 0\,. \tag{3.58}$$

The space of harmonic (p,q)-forms on M is denoted by $\mathcal{H}^{p,q}(M)$.

Exercise 3.17: Show that on a compact manifold, ψ is harmonic iff $\bar\partial\psi = \bar\partial^*\psi = 0$, i.e. a harmonic form has zero curl and zero divergence with respect to its anti-holomorphic indices. Show furthermore that a harmonic form is orthogonal to any exact form and is therefore never exact.

In analogy to de Rham cohomology, one has the (complex version of the) *Hodge Theorem*: $A^{p,q}$ has a unique orthogonal decomposition

$$A^{p,q} = \mathcal{H}^{p,q} \oplus \bar\partial A^{p,q-1} \oplus \bar\partial^* A^{p,q+1}\,. \tag{3.59}$$

In other words, every $\varphi \in A^{p,q}$ has a unique decomposition

$$\varphi = h + \bar\partial\psi + \bar\partial^*\eta \tag{3.60}$$

where $h \in \mathcal{H}^{p,q}$, $\psi \in A^{p,q-1}$ and $\eta \in A^{p,q+1}$. If $\bar\partial\varphi = 0$ then $\bar\partial^*\eta = 0$,[8] i.e. we have the unique decomposition of $\bar\partial$-closed forms

$$Z^{p,q}_{\bar\partial} = \mathcal{H}^{p,q} \oplus \bar\partial A^{p,q-1} \tag{3.61}$$

With reference to (3.39) we have thus shown that

$$H^{p,q}_{\bar\partial}(M) \simeq \mathcal{H}^{p,q}(M) \tag{3.62}$$

or, in words, every $\bar\partial$-cohomology class of (p,q)-forms has a unique harmonic representative $\in \mathcal{H}^{p,q}$. Conversely, every harmonic form defines a cohomology class.

The *Kähler class* of a Kähler form ω is the set of Kähler forms belonging to the cohomology class $[\omega]$ of ω.

Exercise 3.18: Prove that the Kähler form is harmonic.

In addition to the $\bar\partial$-Laplacian $\Delta_{\bar\partial}$, one defines two further Laplacians on a complex manifold: $\Delta_\partial = \partial\partial^* + \partial^*\partial$ and the familiar $\Delta_d = dd^* + d^*d$. The importance of the Kähler condition is manifest in the following result which is valid on Kähler manifolds but not generally on complex manifolds:

$$\Delta_{\bar\partial} = \Delta_\partial = \frac{1}{2}\Delta_d \tag{3.63}$$

i.e. the $\bar\partial-$, $\partial-$ and $d-$harmonic forms coincide. An elementary proof of (3.63) proceeds by working out the three Laplacians in terms of covariant

[8] From $\bar\partial\varphi = \bar\partial\bar\partial^*\eta$ it follows that $(\bar\partial\varphi, \eta) = (\bar\partial\bar\partial^*\eta, \eta) = (\bar\partial^*\eta, \bar\partial^*\eta)$.

derivatives and Riemann tensors on a Kähler manifold. For other proofs, see e.g. [23].

One immediate consequence of (3.63) is that Δ_d does not change the index type of a form. Another important consequence is that on Kähler manifolds every holomorphic p-form is harmonic and vice-versa, every harmonic $(p,0)$ form is holomorphic. Indeed, if $\alpha \in \Omega^p \subset A^{p,0}$, $\bar\partial\alpha = 0$ and $\bar\partial^*\alpha = 0$. The latter is true since $\bar\partial^* : A^{p,q} \to A^{p,q-1}$ and can also be seen directly from (3.56). Conversely, $\Delta\alpha = 0$ implies $\bar\partial\alpha = 0$ which, for $\alpha \in \mathcal{H}^{p,0}$, means $\alpha \in \Omega^p$.

It follows from (3.63) that on Kähler manifolds

$$
\begin{aligned}
&\sum_{p+q=r} h^{p,q} = b_r \, , \\
&\sum_{p,q} (-1)^{p+q} h^{p,q} = \sum_r (-1)^r b_r = \chi(M) \, ,
\end{aligned} \tag{3.64}
$$

where $\chi(M)$ is the Euler number of M. The decomposition of the Betti numbers into Hodge numbers corresponds to the $U(n)$ invariant decomposition $\mu = (i, \bar\imath)$. The second relation also holds in the non-Kähler case where the first relation is replaced by an inequality ($\geq$); i.e. the decomposition of forms (3.11) does not generally carry over to cohomology. Note that (3.64) relates real and complex dimensions.

In general, the Hodge numbers depend on the complex structure. On compact manifolds which admit a Kähler metric, these numbers do however not change under *continuous* deformations of the complex structure. They also do not depend on the metric. What does depend on the metric is the harmonic representative of each class, but the difference between such harmonic representatives is always an exact form.

The Hodge numbers of Kähler manifolds are not all independent. From $\overline{A^{p,q}} = A^{q,p}$ we learn

$$
h^{p,q} = h^{q,p} \, . \tag{3.65}
$$

This symmetry ensures that all odd Betti numbers of Kähler manifolds are even (possibly zero). Furthermore, since $[\Delta_d, *] = 0$ and since $* : A^{p,q} \to A^{n-q,n-p}$ we conclude

$$
h^{p,q} = h^{n-q,n-p} \overset{(3.65)}{=} h^{n-p,n-q} \, . \tag{3.66}
$$

The existence of a closed $(1,1)$-form, the Kähler form ω (which is in fact harmonic, cf. Exercise 3.18), ensures that

$$
h^{p,p} > 0 \quad \text{for} \quad p = 0, \ldots, n \, . \tag{3.67}
$$

Indeed, $\omega^p \in H^{p,p}(M)$ is obviously closed. If it were exact for some p, then ω^n were also exact. But this is impossible since ω^n is a volume form. $h^{0,0} = 1$ if the manifold is connected. The elements of $H^{0,0}(M, \mathbb{C})$ are the complex constants. One can show that on $\mathbb{P}^n$ the Kähler form generates the whole cohomology, i.e. $h^{p,p}(\mathbb{P}^n) = 1$ for $p = 0, \dots n$, with all other Hodge numbers vanishing.

For instance, on a connected three-dimensional Kähler manifold, these symmetries leave only five independent Hodge numbers, e.g. $h^{1,0}$, $h^{2,0}$, $h^{1,1}$, $h^{2,1}$ and $h^{3,0}$. For Ricci-flat Kähler manifolds, which we will consider in detail below, we will establish three additional restrictions on its Hodge numbers.

We have already encountered one important cohomology class on Kähler manifolds: from (3.23) we learn that $\mathcal{R} \in H^{1,1}(M, \mathbb{C})$ and from (3.24) that under change of metric $\mathcal{R}$ varies within a given cohomology class. In fact, one can show that, if properly normalized, the Ricci form defines an element on $H^{1,1}(M, \mathbb{Z})$. This leads us directly to a discussion of *Chern classes*.

Given a Kähler metric, we can define a matrix valued 2-form Θ of type $(1,1)$ by

$$\Theta_i^j = g^{j\bar{p}} R_{i\bar{p}k\bar{l}} dz_k \wedge d\bar{z}_l \; . \tag{3.68}$$

One defines the *Chern form*

$$c(M) = 1 + \sum_i c_i(M) = \det\left(\mathbb{1} + \frac{it}{2\pi}\Theta\right)|_{t=1} = (1 + t\phi_1(g) + t^2\phi_2(g) + \dots)|_{t=1} \tag{3.69}$$

which has the following properties (cf. e.g. [12, 27]):

- $d\phi_i(g) = 0$ and $[\phi_i] \in H^{i,i}(M, \mathbb{C}) \cap H^{2i}(M, \mathbb{R})$,
- $[\phi_i(g)]$ is independent of g,
- $c_i(M)$ is represented by $\phi_i(g)$.

$c_i(M)$ is the i^{th} Chern class of the manifold M. In these lectures we only need $c_1(M)$ which is expressed in terms of the Ricci form:

$$\phi_1(g) = \frac{i}{2\pi}\Theta_i^i = \frac{i}{2\pi} R_{k\bar{l}} dz^k \wedge d\bar{z}^l = \frac{1}{2\pi}\mathcal{R} = -\frac{i}{2\pi}\partial\bar{\partial}\log\det(g_{k\bar{l}}) \; .$$

For $c_1(M)$, the first two properties have been proven in (3.23) and (3.24). Moreover, if

$$dv = v dz^1 \wedge d\bar{z}^1 \wedge \dots \wedge dz^n \wedge d\bar{z}^n$$

is any volume form on M, we can represent $c_1(M)$ by

$$c_1(M) = -\left[\frac{i}{2\pi}\partial\bar{\partial}\log(v)\right] . \tag{3.70}$$

This is so since $v = f \det(g)$ for a non-vanishing positive function f on M.

Example 3.4: Let $M = \mathbb{P}^n$, endowed with the Fubini-Study metric. We then have (cf. (3.29)) $\mathcal{R} = (n+1)\omega$, i.e. $c_1(\mathbb{P}^n) = \frac{1}{2\pi}(n+1)[\omega]$.

We say that $c_1(M) > 0\,(< 0)$ if $c_1(M)$ can be represented by a positive (negative) form. In local coordinates this means

$$\phi_1 = i\phi_{k\bar{l}} dz^k \wedge d\bar{z}^l \ , \tag{3.71}$$

where $\phi_{k\bar{l}}$ is a positive (negative) definite matrix. We say that $c_1(M) = 0$ if the first Chern class is cohomologous to zero. Clearly $c_1(\mathbb{P}^n) > 0$. Note that, e.g. a $c_1(M) > 0$ means that $\int_{\mathcal{C}} c_1 > 0$ for any curve $\mathcal{C}$ in M.

3.5 Calabi-Yau Manifolds

We are now prepared to give a definition of a Calabi-Yau manifold:

A *Calabi-Yau manifold* is a compact Kähler manifold with vanishing first Chern class.

While it is obvious that any Ricci-flat Kähler manifold has vanishing first Chern class, the opposite is far from trivial. This problem was first considered by Calabi in a more general context. He asked the question whether any representative of $c_1(M)$ is the Ricci-form of some Kähler metric. (One can show that any two such representatives differ by a term of the form $\partial\bar{\partial}f$ where $f \in C^\infty(M, \mathbb{R})$. This is the content of the $\partial\bar{\partial}$-Lemma, cf. [12], 2.110.) Calabi also showed that if such a Kähler metric exists, then it must be unique. Yau provided the proof that such a metric always exists if M is compact.

The precise statement of Yau's theorem is: let M be a compact Kähler manifold, ω its Kähler form, $c_1(M)$ its first Chern class. Any closed real two-form of type (1,1) belonging to $2\pi c_1(M)$ is the Ricci form of one and only one Kähler metric in the class of ω.

For vanishing first Chern class, which is the case we are interested in, this means that given any Kähler metric g with associated Kähler form ω, one can always find a unique Ricci-flat Kähler metric g' with Kähler form ω' such that $[\omega] = [\omega']$, i.e. a Kähler manifold with $c_1(M) = 0$ admits a unique Ricci-flat Kähler form in each Kähler class.

Since the first Chern class is represented by the Ricci form and since the latter changes under change of metric by an exact form, i.e. $\mathcal{R}(g') = \mathcal{R}(g) + d\alpha$ (cf. (3.24)), vanishing of the first Chern class is necessary for having a Ricci-flat metric. This is the easy part of the theorem. To prove that this is also sufficient is the hard part. Yau's proof is an existence proof. In fact no Calabi-Yau metric has ever been constructed explicitly. In the non-compact case the situation in this respect is better; examples are the Eguchi-Hanson metrics, see e.g. [28], and the metric on the deformed and the resolved conifold [33]. They play a rôle in the resolution of singularities (orbifold and conifold singularities, respectively) which can occur in compact CY manifolds at special points in their moduli space.

The compact Kähler manifolds with zero first Chern class are thus precisely those which admit a Kähler metric with zero Ricci curvature, or equivalently, with restricted holonomy group contained in $SU(n)$. Following common practice we will talk about Calabi-Yau manifolds if the holonomy group is precisely $SU(n)$. This excludes tori and direct product spaces. We want to mention in passing that any compact Kähler manifold with

$c_1(M) = c_2(M) = 0$ is flat, i.e. $M = \mathbb{C}^n/\Gamma$. This shows that while Ricci-flatness is characterized by the first Chern class, flatness is characterized by the second Chern class.

We should mention that the analysis that we sketched in the introduction, which led to considering Ricci-flat manifolds, was based on a perturbative string theory analysis which was further restricted to lowest order in α'. If one includes α'-corrections, both the beta-function equations and the supersymmetry transformations will be corrected and the Ricci-flatness condition is also modified. One finds the requirement $R_{i\bar{j}} + \alpha'^3(R^4)_{i\bar{j}} = 0$, where (R^4) is a certain tensor composed of four powers of the curvature. It has been shown that the α'-corrections to the Ricci-flat metric, which one has at lowest order, do not change the cohomology class. They are always of the form $\partial\bar{\partial}(...)$ and are thus cohomologically trivial [34]. In other words, supersymmetry preserving string compactifications require manifolds which admit a Ricci-flat Kähler metric but the actual background configuration might have a metric with non-vanishing Ricci tensor.

One often defines Calabi-Yau manifolds as those compact complex Kähler manifolds with trivial canonical bundle. We now want to digress to explain the meaning of this statement and to demonstrate that it is equivalent to the definition given above. Chern classes can be defined for any complex vector bundle over M. By $c_i(M)$ as defined above we mean the Chern classes of the tangent bundle. Given a connection on the vector bundle, the Chern classes can be expressed by the curvature of the connection in the same way as for the tangent bundle with the hermitian connection.

Exercise 3.19: Show that $c_1(T^*M) = -c_1(TM)$.

A central property of Chern classes is that they do not depend on the choice of connection. They are topological cohomology classes in the base space of the vector bundle (see e.g. [25], p.90). An important class of vector bundles over a complex manifold are those with fibers of (complex) dimension one, the so called *line bundles* with fiber $\mathbb{C}$ (complex vector bundles of rank one). Holomorphic line bundles have holomorphic transition functions and a holomorphic section is given in terms of local holomorphic functions. Each holomorphic section defines a local holomorphic frame (which is, of course, one-dimensional for a line-bundle). One important and canonically defined line bundle is the *canonical line bundle* $K(M) = \overset{n}{\wedge} T^{*1,0}(M)$ whose sections are forms of type $(n, 0)$, where $n = \dim_{\mathbb{C}}(M)$. It is straightforward to verify that $[\nabla_i, \nabla_{\bar{j}}]\omega_{i_1\ldots,i_n} = -R_{i\bar{j}}\omega_{i_1\ldots i_n}$, i.e. its curvature form is the negative of the Ricci form of the Kähler metric. This shows that $c_1(M) = -c_1(K(M))$ and if $c_1(M) = 0$ the first Chern class of the canonical bundle also vanishes. For a line bundle this means that it is trivial. Consequently there must exist a globally defined nowhere vanishing section, i.e. globally defined nowhere vanishing holomorphic n-form on M. One finds from (3.35) that on a compact Ricci-flat Kähler manifold any holomorphic p-form is covariantly constant.

This means that the holonomy group $\mathcal{H}$ of a Calabi-Yau manifold is contained in $SU(n)$.

Example 3.5: In this example we consider complex hypersurfaces in $\mathbb{P}^n$ which are expressed as the zero set of a homogeneous polynomial. We already know that they are Kähler. We want to compute c_1 of the hypersurface as a function of the degree d of the polynomial and of n. From this we can read off the condition for the hypersurface to be a Calabi-Yau manifold. This can be done with the tools we have developed so far, even though more advanced and shorter derivations of the result can be found in the literature, see e.g. [23] or [12]. Later we will encounter another way to see that $d = n+1$ means $c_1(X) = 0$ by explicitly constructing the unique holomorphic n-form which, as we will see, must exist on a Calabi-Yau n-fold. The calculation is presented in Appendix C. The result we find there is

$$2\pi c_1(X) = (n+1-d)[\omega] \ . \tag{3.72}$$

It follows that the first Chern class $c_1(X)$ is positive, zero or negative according to $d < n+1$, $d = n+1$ and $d > n+1$, respectively.

We have thus found an easy way to construct Calabi-Yau manifolds. For one-folds, a cubic hypersurface in $\mathbb{P}^2$ is a 2-torus and for two-folds, a quartic hypersurface in $\mathbb{P}^3$ is a K3. If we are interested in three-folds, we have to choose the quintic hypersurface in $\mathbb{P}^4$. This is in fact the simplest example, which we will study further below.

The Calabi-Yau condition on the degree generalizes to the case of hypersurfaces in weighted projective spaces. Given a weighted projective space $\mathbb{P}^n[\boldsymbol{w}]$ and a hypersurface X specified by the vanishing locus of a quasi-homogeneous polynomial of degree d, we find

$$c_1(X) = 0 \quad \Leftrightarrow \quad d = \sum_{i=1}^{n} w_i \ . \tag{3.73}$$

The condition on the degrees and weights can also be easily written down for complete intersections in products of weighted projective spaces.

As we have discussed before, in the generic case the hypersurface will be singular. To get a smooth Calabi-Yau manifold one has to resolve the singularities in such a way that the canonical bundle remains trivial.

Example 3.6: An example of a CY_3 hypersurface in weighted projective space where no resolution is necessary is the sextic in $\mathbb{P}^4[1,1,1,1,2]$. The embedding space has only isolated singular points which are avoided by a generic hypersurface. On the other hand, the octic hypersurface in $\mathbb{P}^4[1,1,2,2,2]$ cannot avoid the singular $\mathbb{Z}_2$ surface of the embedding space and has thus itself a singular $\mathbb{Z}_2$ curve which must be “repaired” in order to obtain a smooth CY manifold.

We should mention that the construction of the first Chern class that we present in Appendix C does not provide the Ricci-flat metric. In fact, the Ricci-flat metric is never the induced metric. As we have mentioned once before, a Ricci-flat Kähler metric on a *compact* Kähler manifold has never been constructed explicitly. Interesting examples of non-compact Ricci-flat Kähler manifolds, which are of potential interest for M-theory and the AdS/CFT correspondence, are the cotangent bundles of spheres of any dimension and the complex cotangent bundle on $\mathbb{P}^n$ for any n. The latter are hyper-Kähler manifolds, which are always Ricci-flat. For these manifolds Ricci-flat metrics are known explicitly. For instance, T^*S^3 is the deformed conifold.

Let us come back to the fact that a compact Kähler manifold with $SU(n)$ holonomy always possesses a nowhere vanishing covariantly constant $(n,0)$-form Ω, called a *complex volume form* which is in fact unique (up to multiplication by a constant). Locally it can always be written as

$$\Omega_{i_1\dots i_n} = f(z)\epsilon_{i_1\dots i_n} \tag{3.74}$$

with f a non-vanishing holomorphic function in a given coordinate patch and $\epsilon_{i_1\dots i_n} = \pm 1$. Before proving this we want to derive two simple corollaries:
(1) Ω is holomorphic. Indeed, $\bar{\partial}_{\bar{\imath}}\Omega_{j_1\dots j_n} = \nabla_{\bar{\imath}}\Omega_{j_1\dots j_n} = 0$, because Ω is covariantly constant.
(2) Ω is harmonic. To show this we still have to demonstrate $\bar{\partial}^*\Omega = 0$. But this obvious since $\bar{\partial}^* = -*\partial*$ and $*: A^{n,0} \to A^{n,0}$ and $\partial A^{n,0} = 0$.

A simple argument that Ω always exists is the following [12, 35]. Start at any point p in M and define $\Omega_p = dz^1 \wedge \cdots \wedge dz^n$, where $\{z^i\}$ are local coordinates. Then parallel transport Ω to every other point on M. This is independent of the path taken, since when transported around a closed path (starting and ending at p), Ω is a singlet under $SU(n)$ and is thus unchanged. This defines Ω everywhere on M. Ω can also be constructed explicitly with the help of the covariantly constant spinor: $\Omega_{ijk} = \epsilon^{\mathrm{T}}\gamma_{ijk}\epsilon$. Here γ_{ijk} is the antisymmetrized product of three γ-matrices which satisfy $\{\gamma_i,\gamma_j\} = \{\gamma_{\bar{\imath}},\gamma_{\bar{\jmath}}\} = 0$, $\{\gamma_i,\gamma_{\bar{\jmath}}\} = 2g_{i\bar{\jmath}}$. The proof that Ω thus defined satisfies all the necessary properties is not difficult. It can be found in [4, 28].

We now show that Ω is essentially unique. Assume that given Ω there were a Ω' with the same properties. Then, since Ω is a form of the top degree, we must have $\Omega' = f\Omega$ where f is a non-singular function. Since we require $\bar{\partial}\Omega' = 0$, f must in fact be holomorphic. On a compact manifold this implies that f is constant.

Conversely, the existence of Ω implies $c_1 = 0$. Indeed, with (3.74), we can write the Ricci form as

$$\mathcal{R} = i\partial\bar{\partial}\log\det(g_{k\bar{l}}) = -i\partial\bar{\partial}\log\left(\Omega_{i_1\dots i_n}\bar{\Omega}_{\bar{\jmath}_1\dots\bar{\jmath}_n}g^{i_1\bar{\jmath}_1}\cdots g^{i_n\bar{\jmath}_n}\right) . \tag{3.75}$$

The argument of the logarithm is a globally defined function and the Ricci form is thus trivial in cohomology, implying $c_1 = 0$.

For hypersurfaces in weighted projective spaces one can explicitly construct Ω by extending the construction of holomorphic differentials on a Riemann surface (see e.g. [23]). Once constructed we know that Ω is essentially unique (up to a multiplicative constant on the hypersurface).

Consider first the torus defined as a hypersurface in $\mathbb{P}^2$ specified by the vanishing locus of a cubic polynomial, $f(x,y,z)=0$. This satisfies (3.72). The unique holomorphic differential (written in a patch with $z=1$) is $\omega = -dy/(\partial f/\partial x) = dx/(\partial f/\partial y) = dx/(2y)$. The first equality follows from $df=0$ along the hypersurface and the second equality if the hypersurface is defined by an equation of the form $f = zy^2 - p(x,z)$, e.g. the Weierstrass and Legendre normal forms. An interesting observation is that ω can be represented as a residue: $\omega = \frac{1}{2\pi i}\int_\gamma \frac{dx\wedge dy}{f(x,y)}$. The integrand is a two-form in the embedding space with a first order pole on the hypersurface $f=0$ and the contour γ surrounds the hypersurface. Changing coordinates $(x,y)\to(x,f)$ and using $\frac{1}{2\pi i}\int_\gamma \frac{df}{f} = 1$ we arrive at ω as given above.

The above construction of the holomorphic differential for a cubic hypersurface in $\mathbb{P}^2$ can be generalized to obtain the holomorphic three-form on a Calabi-Yau manifold realized as a hypersurface $p=0$ in weighted $\mathbb{P}^4[\boldsymbol{w}]$ [36, 37]. Concretely,

$$\Omega = \int_\gamma \frac{\mu}{p} \ , \tag{3.76}$$

where

$$\mu = \sum_{i=0}^{4} (-1)^i w_i z^i dz^0 \wedge \cdots \wedge \widehat{dz^i} \wedge \cdots \wedge dz^5 \ , \tag{3.77}$$

and the term under the $\widehat{\ }$ is omitted. The contour γ now surrounds the hypersurface $p=0$ inside the weighted projective space. Note that the numerator and the denominator in μ/p scale in the same way under (3.6). In the patch U_i where $z^i = \text{const}$, only one term in the sum survives. One can perform the integration by replacing one of the coordinates, say z^j, by p and using $\int_\gamma \frac{dp}{p} = 2\pi i$. In this way one gets an expression for Ω directly on the embedded hypersurface. For instance in the patch U_0 one finds (no sum on (i,j,k) implied)

$$\Omega = \frac{w_0 z^0 dz^i \wedge dz^j \wedge dz^k}{\Delta_0^{ijk}} \ , \tag{3.78}$$

where $\Delta_0^{ijk} = \frac{\partial(z^i,z^j,z^k,p)}{\partial(z^1,z^2,z^3,z^4)}$. From our derivation it is clear that this representation of Ω is independent of the choice of $\{i,j,k\}\subset\{1,2,3,4\}$ and of the choice of coordinate patch. Furthermore, it is everywhere non-vanishing and well defined at every non-singular point of the hypersurface. A direct verification of these properties can be found in [4, 38].

The existence of a holomorphic n-form then means that the holonomy group $\mathcal{H}$ (and not just $\mathcal{H}_0$) is contained in $SU(n)$.

Let us now complete the discussion of Hodge numbers of Calabi-Yau manifolds. We have just established the existence of a unique harmonic $(n,0)$-form, Ω, and thus

$$h^{n,0} = h^{0,n} = 1 \; . \tag{3.79}$$

With the help of Ω we can establish one further relation between the Hodge numbers. Given a holomorphic and hence harmonic $(p,0)$-form, we can, via contraction with Ω, construct a $(0,n-p)$-form, which can be shown to be again harmonic, as follows. Given

$$\alpha = \alpha_{i_1 \dots i_p} dz^{i_1} \wedge \dots \wedge dz^{i_p} \; , \qquad \bar{\partial}\alpha = 0 \; , \tag{3.80}$$

α being (Δ_∂)-harmonic means

$$\begin{aligned} \partial\, \alpha = 0 \quad &\Leftrightarrow \quad \nabla_{[j_i} \alpha_{j_2 \dots j_{p+1}]} = 0 \; , \\ \partial^* \alpha = 0 \quad &\Leftrightarrow \quad \nabla^{i_1} \alpha_{i_1 \dots i_p} = 0 \; . \end{aligned} \tag{3.81}$$

We then define the $(0,n-p)$-form

$$\beta_{\bar{\jmath}_{p+1} \dots \bar{\jmath}_n} = \frac{1}{p!} \bar{\Omega}_{\bar{\jmath}_1 \dots \bar{\jmath}_n} \alpha^{\bar{\jmath}_1 \dots \bar{\jmath}_p} \; . \tag{3.82}$$

This can be inverted to give (use (3.74))

$$\alpha^{\bar{\jmath}_1 \dots \bar{\jmath}_p} = \frac{1}{\|\Omega\|^2} \Omega^{\bar{\jmath}_1 \dots \bar{\jmath}_p \bar{\jmath}_{p+1} \dots \bar{\jmath}_n} \beta_{\bar{\jmath}_{p+1} \dots \bar{\jmath}_n} \; , \tag{3.83}$$

where we have defined

$$\|\Omega\|^2 = \frac{1}{n!} \Omega_{i_1 \dots i_n} \Omega^{i_1 \dots i_n} \; . \tag{3.84}$$

From this we derive

$$\nabla^{\bar{\jmath}_{p+1}} \beta_{\bar{\jmath}_{p+1} \dots \bar{\jmath}_n} = \frac{1}{p!} \bar{\Omega}_{\bar{\jmath}_1 \dots \bar{\jmath}_n} \nabla^{\bar{\jmath}_{p+1}} \alpha^{\bar{\jmath}_1 \dots \bar{\jmath}_p} = 0 \; , \tag{3.85}$$

using $(3.81)_1$. Similarly

$$\nabla_{\bar{\jmath}_1} \alpha^{\bar{\jmath}_1 \dots \bar{\jmath}_p} = \frac{1}{\|\Omega\|^2} \Omega^{\bar{\jmath}_1 \dots \bar{\jmath}_p \bar{\jmath}_{p+1} \dots \bar{\jmath}_n} \nabla_{\bar{\jmath}_1} \beta_{\bar{\jmath}_{p+1} \dots \bar{\jmath}_n} = 0 \tag{3.86}$$

by virtue of $(3.81)_2$. It follows that β is also harmonic.

We have thus shown the following relation between Hodge numbers

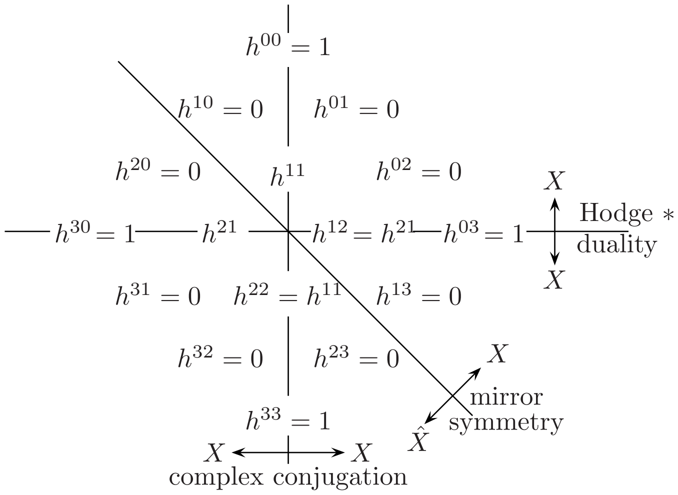
$$h^{p,0} = h^{0,n-p} = h^{n-p,0} \; . \tag{3.87}$$

Let us finally look at $h^{p,0}$. For this we need the Laplacian on p-forms. Specifying (3.35) for a harmonic $(p,0)$ form on a Ricci-flat Kähler manifold where $R_{i\bar{\jmath}} = R_{ij\bar{k}\bar{l}} \equiv 0$, we find $\nabla^\nu \nabla_\nu \omega_{i_1 \dots i_p} = 0$. On a compact manifold this means that ω is parallel, i.e. $\nabla_j \omega_{i_1 \cdots i_p} = 0$, $\bar{\partial}_{\bar{j}} \omega_{i_1 \cdots i_p} = 0$, the latter equality already being a consequence of harmonicity. But this means that ω transforms as a singlet under the holonomy group. We now assume that the holonomy group is exactly $SU(n)$, i.e. not a proper subgroup of it.[9] Since $\omega_{i_1 \cdots i_p}$ transforms in the $\wedge^p \boldsymbol{n}$ of $SU(n)$, the singlet only appears in the decomposition if $p = 0$ or $p = n$. We thus learn that on Calabi-Yau manifolds with holonomy group $SU(n)$

$$h^{p,0} = 0 \qquad \text{for} \qquad 0 < p < n \; . \tag{3.88}$$

Exercise 3.20: Show that $h^{1,0}(M) = 0$ implies that there are no continuous isometries on M.

If we collect the results on the Hodge numbers of Calabi-Yau manifolds for the case $n = 3$, we find that the only independent Hodge numbers are $h^{1,1} \geq 1$ and $h^{2,1} \geq 0$ and the Hodge diamond for Calabi-Yau three-folds is



(3.89)

The Euler number of a Calabi-Yau three-fold is then (cf. (3.64))

$$\chi(M_3) = 2(h^{1,1} - h^{1,2}) \; . \tag{3.90}$$

In higher dimensions there are more independent Hodge numbers, but this will not be covered here. For the case of CY four-folds, see [39]. The significance of $h^{1,1}$ and $h^{2,1}$ for Calabi-Yau three-folds will be explained in Sect. 3.6.

In (3.89) we have indicated operations which relate Hodge numbers to each other. In addition to complex conjugation (3.65) and the Hodge $*$-operation (3.66), which act on the Hodge numbers of a given CY manifold,

[9] In Chap. 4 we discuss orbifolds with discrete holonomy groups. There the condition will be that it is not contained in any continuous subgroup of $SU(n)$.

we have also shown the action of mirror symmetry: given a CY manifold X, there exists a mirror manifold $\hat{X}$ such that $h^{p,q}(X) = h^{3-p,q}(\hat{X})$. This in particular means that the two non-trivial Hodge numbers $h^{1,1}$ and $h^{2,1}$ are interchanged between X and $\hat{X}$ and that $\chi(X) = -\chi(\hat{X})$. Within the class of Calabi-Yau manifolds constructed as hypersurfaces in toric varieties (which we have not discussed) they manifestly come in mirror pairs [40]. So-called rigid manifolds, for which $h^{2,1}(X) = 0$, and consequently $h^{1,1}(\hat{X}) = 0$, are discussed in [41]. The Z-manifold described in example 4.4 is rigid.

As we have already mentioned at the end of Chap. 2, mirror symmetry is a much stronger statement than the mere existence of mirror pairs of CY manifolds. Its far-reaching consequences for both string theory and algebraic geometry are thoroughly covered in [18].

3.6 Calabi-Yau Moduli Space

In this section we will only treat three dimensional Calabi-Yau manifolds. References are [4, 21, 38, 42, 43]. The generalization to higher dimensions of most the issues discussed here is straightforward. The two-dimensional case (K3) is described in [44] in great detail.

In view of Yau's theorem, the parameter space of CY manifolds is that of Ricci-flat Kähler metrics. We thus ask the following question: given a Ricci-flat Riemannian metric $g_{\mu\nu}$ on a manifold M, what are the allowed infinitesimal variations $g_{\mu\nu} + \delta g_{\mu\nu}$ such that

$$R_{\mu\nu}(g) = 0 \quad \Rightarrow \quad R_{\mu\nu}(g + \delta g) = 0 \qquad ? \tag{3.91}$$

Clearly, if g is a Ricci-flat metric, then so is any metric which is related to g by a diffeomorphism (coordinate transformation). We are not interested in those δg which are generated by a change of coordinates. To eliminate them we have to fix the diffeomorphism invariance and impose a *coordinate condition*. This is analogous to fixing a gauge in electromagnetism. The appropriate choice is to demand that $\nabla^\mu \delta g_{\mu\nu} = 0$ (see e.g. [12], 4.62). Any $\delta g_{\mu\nu}$ which satisfies this condition also satisfies $\int_M \sqrt{g}\delta g^{\mu\nu}(\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu)d^d x = 0$, and is thus orthogonal to any change of the metric induced by a diffeomorphism generated by the vector field ξ_μ. Then, expanding (3.91) to first order in δg and using $R_{\mu\nu}(g) = 0$ and the coordinate condition, one finds

$$\nabla^\rho \nabla_\rho \delta g_{\mu\nu} - 2R_\mu{}^\rho{}_\nu{}^\sigma \delta g_{\rho\sigma} = 0 \, . \tag{3.92}$$

Exercise 3.21: Derive (3.92). Useful expansions of the curvature can be found in [45]. One needs to use that M is compact to eliminate a term $\nabla_\mu \nabla_\nu \mathrm{tr}(\delta g)$.

We now want to analyze (3.92) if (M, g) is a Kähler manifold. Given the index structure of the metric and the Riemann tensor on Kähler manifolds, one immediately finds that the conditions imposed on the components $\delta g_{i\bar{\jmath}}$

and δg_{ij} decouple and can thus be studied separately. This is what we now do in turn.

(1) $\delta g_{i\bar{\jmath}}$: With the help of (3.35) it is easy to see that the condition (3.92), which now reads $\nabla^\mu \nabla_\mu \delta g_{i\bar{\jmath}} - 2R_i{}^k{}_{\bar{\jmath}}{}^{\bar{l}} \delta g_{k\bar{l}} = 0$, $\mu{=}(k, \bar{k})$, is equivalent to $(\Delta \delta g)_{i\bar{\jmath}} = 0$. Here we view $\delta g_{i\bar{\jmath}}$ as the components of a $(1,1)$-form. We see that harmonic $(1,1)$-forms correspond to the metric variations of the form $\delta g_{i\bar{\jmath}}$ and to cohomologically non-trivial changes of the Kähler form. Of course, we already knew from Yau's theorem that for any $[\omega + \delta\omega]$ there is again a Ricci-flat Kähler metric. Expanding $\delta g_{i\bar{\jmath}}$ in a basis of real $(1,1)$-forms, which we will denote by b^α, $\alpha = 1, \ldots, h^{1,1}$, we obtain the following general form of the deformations of the Kähler structure of the Ricci flat metric:

$$\delta g_{i\bar{\jmath}} = \sum_{\alpha=1}^{h^{1,1}} \tilde{t}^\alpha b^\alpha_{i\bar{\jmath}} \,, \qquad \tilde{t}^\alpha \in \mathbb{R} \,. \tag{3.93}$$

Using (3.56) one may check that these δg satisfy the coordinate condition.

For $g + \delta g$ to be a Kähler metric, the *Kähler moduli* $\tilde{t}^\alpha$ have to be chosen such that the deformed metric is still positive definite. Positive definiteness of a metric g with associated Kähler form ω is equivalent to the condition

$$\int_C \omega > 0 \,, \qquad \int_S \omega^2 > 0 \,, \qquad \int_M \omega^3 > 0 \tag{3.94}$$

for all curves C and surfaces S on the Calabi-Yau manifold M. The subset in $\mathbb{R}^{h^{1,1}}$ spanned by the parameters $\tilde{t}^\alpha$ such that (3.94) is satisfied, is called the *Kähler cone*.

Exercise 3.22: Verify that this is indeed a cone.

(2) δg_{ij} : Now (3.92) reads $\nabla^\mu \nabla_\mu \delta g_{ij} - 2R_i{}^k{}_j{}^l \delta g_{kl} = 0$. With little work this can be shown to be equivalent to

$$\Delta_{\bar\partial} \delta g^i = (\bar\partial \bar\partial^* + \bar\partial^* \bar\partial) \delta g^i = 0 \tag{3.95}$$

where

$$\delta g^i = \delta g^i_{\bar{\jmath}} d\bar{z}^{\bar{\jmath}} \,, \qquad \delta g^i_{\bar{\jmath}} = g^{i\bar{k}} \delta g_{\bar{k}\bar{\jmath}} \tag{3.96}$$

is a $(0,1)$-form with values in $T^{1,0}(M)$. We conclude that (3.95) implies that $\delta g^i \in H^{(0,1)}_{\bar\partial}(M, T^{1,0})$. Again one may verify that these deformations of the metric satisfy the coordinate condition.

Exercise 3.23: Fill in the steps of the above argument.

What is the significance of these metric deformations? For the new metric to be again Kähler, there must be a coordinate system in which it has only mixed components. Since holomorphic coordinate transformations do not change the type of index, it is clear that δg_{ij} can only be removed by a non-holomorphic transformation. But this means that the new metric is

Kähler with respect to a different complex structure compared to the original metric. (Of course, this new metric cannot be obtained from the original undeformed metric by a diffeomorphism as they have been fixed by the coordinate condition. So while removing δg_{ij}, the non-holomorphic change of coordinates generates a $\delta g_{i\bar{\jmath}}$).

With the help of the unique holomorphic $(3,0)$ form we can now define an isomorphism between $H^{0,1}_{\bar\partial}(M,T^{1,0})$ and $H^{2,1}_{\bar\partial}(M)$ by defining the complex (2,1)-forms

$$\Omega_{ijk}\delta g^k_{\bar l} dz^i \wedge dz^j \wedge \bar z^{\bar l} \; . \tag{3.97}$$

which are harmonic if (3.92) is satisfied. These complex structure deformations can be expanded in a basis $b^a_{ij\bar k}$, $a=1,\ldots,h^{2,1}$, of harmonic $(2,1)$-forms:

$$\Omega_{ijk}\delta g^k_{\bar l} = \sum_{a=1}^{h^{2,1}} t^a b^a_{ij\bar l} \tag{3.98}$$

where the complex parameters t^a are called *complex structure moduli.*[10]

If we were geometers we would only be interested in the deformations of the metric and the number of real deformation parameters (moduli) would be $h^{1,1}+2h^{1,2}$. However, in string theory compactified on Calabi-Yau manifolds we have additional massless scalar degrees of freedom from the internal components of the antisymmetric tensor field in the (NS,NS) sector of the type II string. Its equations of motion in the gauge $d^*B=0$ are $\Delta B=0$, i.e. excitations of the B-field above the background where it vanishes are harmonic two-forms on the Calabi-Yau manifold. We can now combine these with the Kähler deformations of the metric and form

$$(i\delta g_{i\bar\jmath}+\delta B_{i\bar\jmath})dz^i\wedge d\bar z^{\bar\jmath} = \sum_{\alpha=1}^{h^{1,1}} \tilde t^\alpha b^\alpha \tag{3.99}$$

where the parameters $\tilde t^\alpha$ are now complex, their imaginary part still restricted by the condition discussed before. This is referred to as the *complexification of the Kähler cone.*

To summarize, there is a moduli space associated with the different Kähler and complex structures which are compatible with the Calabi-Yau condition. The former are parametrized by $H^{1,1}_{\bar\partial}(M)$ and the latter by of $H^{0,1}_{\bar\partial}(M,T^{1,0})\simeq H^{2,1}_{\bar\partial}(M)$. The moduli space of Ricci-flat Kähler metrics is parametrized by the harmonic representatives of these cohomology groups.

[10] Our discussion of complex structure moduli is not complete. We have only considered the linearized deformation equation. It still needs to be shown that they can be integrated to finite deformations. That this is indeed the case for Calabi-Yau manifolds has been proven by Tian [46] and by Todorov [47]. For a general complex manifold the number of complex structure deformations is less than $h^{2,1}$.

Let us now exemplify this discussion by the quintic in $\mathbb{P}^4$. Here we have $h^{1,1} = 1$: this is simply the Kähler form induced from the ambient space $\mathbb{P}^4$. (The metric induced from the Fubini-study metric is not the Ricci-flat one.) As shown in [26] and by more elementary means in [38], the complex structure parameters appear as coefficients in the most general quintic polynomial. One easily finds that there are 126 coefficients. However, polynomials which are related by a linear change of the homogeneous coordinates of $\mathbb{P}^4$ should not be counted as different. These are parametrized by $\dim_{\mathbb{C}}(GL(5,\mathbb{C})) =$ 25 coefficients. We therefore conclude that there are 101 complex structure moduli on the quintic hypersurface, i.e. $h^{2,1} = 101$. For special values of these coefficients the hypersurface is singular, i.e. there are solutions of $p = dp = 0$. With (3.90) we find that the Euler number of the quintic is -200.

The situation for hypersurfaces in weighted projective spaces is more complicated. If the hypersurface passes through the singular loci of the embedding space, they have to be "repaired". Care has to be taken that in doing this the Calabi-Yau condition $c_1 = 0$ is maintained. This introduces additional elements in the cohomology, so that in general $h^{1,1} > 1$. Also $h^{2,1}$ can no longer be counted as the number of coefficients in the defining polynomial: this counting falls short of the actual number of complex structure moduli. There are methods to compute the Hodge numbers of these manifolds. The most systematic and general one is by viewing them as hypersurfaces in toric varieties [40].

We will not address questions of global properties of the moduli space of string compactifications on Calabi-Yau manifolds, except for mentioning a few aspects. Mirror symmetry, which connects topologically distinct manifolds, is certainly relevant. Another issue is that of transitions among topologically different manifolds, the prime example being the conifold transition [33]. While one encounters singular geometries in the process, string theory is well behaved and the transition is smooth. Indeed, it has been speculated that the moduli space of all Calabi-Yau compactifications is smoothly connected [48].

3.7 Compactification of Type II Supergravities on a CY Three-Fold

Now that we know the meaning of the Hodge numbers $h^{1,1}$ and $h^{2,1}$, we can, following our general discussion in Sect. 2.3, examine the relevance of the existence of harmonic forms on Calabi-Yau manifolds for the massless spectrum of the compactified theory. We will consider the two ten-dimensional type II supergravities that are the field theory limits of type II strings. The discussion is thus also relevant for string compactification, as long as the restriction to the massless modes is justified, i.e. for energies $E^2\alpha' \ll 1$. However, there are string effects which are absent in field theory compactifications, such as topological non-trivial embeddings of the string world-sheet into the CY manifold. These stringy effects (world-sheet instantons) which are non-perturbative in

α', have an action which scales as R^2/α', where R is the typical size of the manifold. They are suppressed as $e^{-S_{\text{inst}}} \sim e^{-R^2/\alpha'}$ and are small for a large internal manifold but relevant for $R \sim \sqrt{\alpha'}$.

Type IIA supergravity is a non-chiral $\mathcal{N}=2$ theory with just a gravity multiplet whose field content is:

$$\mathcal{G}_{\text{IIA}}(10) = \left\{ G_{MN}, \psi_M^{(+)}, \psi_M^{(-)}, \psi^{(+)}, \psi^{(-)}, B_{MN}, A_{MNP}, V_M, \phi \right\} . \quad (3.100)$$

These fields correspond to the massless states of the type IIA string. The fermionic fields arise in the two Neveu-Schwarz-Ramond sectors, i.e. (NS,R) plus (R,NS), they are the two Majorana-Weyl gravitini of opposite chirality $\psi_M^{(\pm)}$, $M, N = 0, \cdots, 9$, and the two Majorana-Weyl dilatini $\psi^{(\pm)}$. The metric G_{MN}, the antisymmetric tensor B_{MN} and the dilaton ϕ come from the (NS,NS) sector. The remaining bosonic fields, the vector V_M and the 3-index antisymmetric tensor A_{MNP}, appear in the (R,R) sector.

Exercise 3.24: Show that (3.100) results upon circle compactification of the fields $\{G_{MN}, \psi_M, A_{MNP}\}$, with ψ_M Majorana. This is the field content of $D=11$ supergravity which is the low-energy limit of M-theory.

Type IIB supergravity has also $\mathcal{N}=2$ supersymmetry but it is chiral, i.e. the two gravitini have the same chirality. The gravity multiplet has content:

$$\mathcal{G}_{\text{IIB}}(10) = \left\{ G_{MN}, \psi_M^{(+)}, \tilde{\psi}_M^{(+)}, \psi^{(+)}, \tilde{\psi}^{(+)} B_{MN}, \tilde{B}_{MN}, A_{MNPQ}, \phi, a \right\} . \quad (3.101)$$

Now the bosonic fields from the (R,R) sector are the axion a, $\tilde{B}_{MN}$ and A_{MNPQ} which is completely antisymmetric and has self-dual field strength.

It is known that type IIA and type IIB strings compactified on a circle are related by T-duality [49]. Therefore, whenever the internal manifold contains a circle, type IIA and type IIB give T-dual theories that clearly must have the same supersymmetric structure. In particular, compactification on T^4 gives maximal (2,2) supersymmetry in $d=6$, compactification on T^6 gives maximal $\mathcal{N}=8$ supersymmetry in $d=4$ and compactification on $\text{K3} \times \text{T}^2$ gives $d=4$, $\mathcal{N}=4$ supersymmetry with 22 $U(1)$ vector multiplets. Below we examine compactification on CY_3 in some more detail. Our purpose is to determine the resulting massless fields by looking at the zero modes of the ten-dimensional multiplets given above. In the lower dimensions we will obtain a theory with a number of supersymmetries that depends on the internal manifold. Clearly, the zero modes must organize into appropriate multiplets whose structure is known beforehand.

Compactification of type IIA supergravity on a CY_3 was considered first in [50] and to greater extent in [51]. The resulting theory in $d=4$ has $\mathcal{N}=2$ supersymmetry. The massless fields belong to the gravity multiplet plus hypermultiplets and vector multiplets, which are the three possible irreducible representations with spins less or equal to two, cf. (4.65). To describe how the massless fields arise we split the ten-dimensional indices in a $SU(3)$ covariant way, $M = (\mu, i, \bar{\imath})$[11] and then use the known results for the number of harmonic (p,q) forms on the Calabi-Yau manifold. The zero modes of $G_{\mu\nu}$, $\psi_\mu^{(+)}$, $\psi_\mu^{(-)}$ and the graviphoton V_μ form the gravity multiplet. Both $\psi_\mu^{(\pm)}$ have an expansion of the form (2.18) so that we obtain two Majorana gravitini in four dimensions. For the remaining fields and components it is simpler to analyze the bosonic states. The fermions are most easily determined via $\mathcal{N}=2$ space-time supersymmetry and the known field content of the various multiplets. Of course they can also be obtained by a zero mode analysis. Altogether one finds for the bosons, in addition to those in the gravity multiplet,

$$A_\mu^\alpha,\ t^a,\ \tilde{t}^\alpha,\ C^a,\ S,\ C\ , \tag{3.102}$$

where A_μ^α arises from $A_{\mu i\bar{\jmath}}$ and the remaining fields are all complex scalars as follows. The $\tilde{t}^\alpha$ correspond to $G_{i\bar{\jmath}}$ and $B_{i\bar{\jmath}}$[12], the t^a to G_{ij}, C^a to the $A_{ij\bar{k}}$ modes, S to ϕ and $B_{\mu\nu}$ (which can be dualized to a pseudoscalar) and C to the A_{ijk} mode. We now group these fields into supermultiplets. A_μ^α and $\tilde{t}^\alpha$ combine to $h^{1,1}$ vector multiplets, whereas t^a and C^a to $h^{2,1}$ hypermultiplets. The two complex scalars S and C form an additional hypermultiplet, so there are $(h^{2,1}+1)$ hypermultiplets.

In the type IIB compactification the gravity multiplet is formed by the zero modes of $G_{\mu\nu}$, $\psi_\mu^{(+)}$, $\tilde{\psi}_\mu^{(+)}$ and $A_{\mu ijk}$. From the rest of the fields we obtain

$$A_\mu^a,\ t^a,\ \tilde{t}^\alpha,\ C^\alpha,\ S,\ C\ . \tag{3.103}$$

Here the fields A_μ^a arise from $A_{\mu ij\bar{k}}$ and t^a from G_{ij}; $(\tilde{t}^\alpha, C^\alpha)$ correspond to $G_{i\bar{\jmath}}$, $B_{i\bar{\jmath}}$, $\tilde{B}_{i\bar{\jmath}}$ and $A_{\mu\nu i\bar{\jmath}}$; (S,C) to ϕ, a, $B_{\mu\nu}$ and $\tilde{B}_{\mu\nu}$. The fields arising from the four-form are real, due to the self-duality constraint of its field-strength. Altogether the fields combine to $(h^{1,1}+1)$ hypermultiplets and $h^{2,1}$ vector multiplets. Notice that this is the same result as in the type IIA case upon exchanging $h^{1,1}$ and $h^{1,2}$. Indeed, it has been shown that compactification of type IIB strings on a CY three-fold X gives the same 4-dimensional theory that appears upon compactification of type IIA strings on the mirror $\hat{X}$ [50, 52].

The moduli of the Calabi-Yau manifold give rise to neutral massless scalars that will appear in the low-energy effective action of the string theory. Supersymmetry imposes stringent restrictions on the action and consequently on the geometry of the moduli spaces. In particular, the moduli fields have no potential and hence their vevs are free parameters. Moreover, in the kinetic terms scalars in vector multiplets do not mix with scalars in hypermultiplets. In fact, the interaction of vector multiplets and hypermultiplets consistent

[11] From now on we only use indices $(i, j, \ldots, \bar{\imath}, \bar{\jmath}, \ldots)$ for the internal space and μ for the four uncompactified space-time dimensions.

[12] Supersymmetry thus requires the complexification of the Kähler cone.

with $\mathcal{N}=2$ supergravity is a non-linear σ-model with a target-space geometry which is locally of the form [53, 54, 55]

$$\mathcal{M}_{\rm SK} \times \mathcal{Q} \tag{3.104}$$

where $\mathcal{M}_{\rm SK}$ is a (special) Kähler manifold (to be defined later) for the vector multiplets [53] and $\mathcal{Q}$ a quaternionic manifold for the hypermultiplets [53, 54, 55].[13] The manifolds $\mathcal{M}_{\rm SK}$ and $\mathcal{Q}$ are parametrized by the scalar fields inside the vector and hypermultiplets, respectively. The product structure is only respected for the gauge-neutral part of the theory. Nonabelian gauge symmetries and charged fields appear if we take non-perturbative effects into account, e.g. by wrapping branes around appropriate cycles. But this will not be considered in these lectures.

For the perturbative type IIA and IIB theories we thus have

$$\begin{aligned} \mathcal{M}^A &= \mathcal{M}^A_{h^{1,1}} \times \mathcal{Q}^A_{h^{2,1}+1} \ , \\ \mathcal{M}^B &= \mathcal{M}^B_{h^{2,1}} \times \mathcal{Q}^B_{h^{1,1}+1} \ . \end{aligned} \tag{3.105}$$

The indices give the complex and quaternionic dimensions, respectively. It is worth mentioning that while $\mathcal{M}_{\rm SK}$ contains only moduli fields, $\mathcal{Q}$ is obtained by combining moduli scalars with non-moduli scalars which, in string theory, come from the (R,R) sector of the left-right superconformal algebra.

The quaternionic dimension of the hypermultiplet moduli spaces is always ≥ 1. In both type II theories, there is at least the *universal hypermultiplet* with scalars (S, C). Its component fields are not related to the cohomology of a Calabi-Yau manifold. Most importantly, it contains the dilaton ϕ which organizes the string perturbation theory. This means that the hypermultiplet moduli space receives (perturbative and non-perturbative) stringy corrections in type IIA and IIB. In contrast to this, the vector multiplet moduli space is exact at string tree level. In types IIB and IIA this concerns the complex structure moduli and Kähler moduli, respectively. The metric of the Kähler moduli space of type IIA receives a perturbative correction at order $(\alpha'/R^2)^3$ [56] and non-perturbative corrections, powers of $e^{-R^2/\alpha'}$, from world-sheet instantons, i.e. topologically non-trivial embeddings of the world-sheet into the Calabi-Yau manifold. In contrast, the metric of the complex structure moduli space of type IIB is exact at both, string and world-sheet σ-model, tree level. It is thus determined by classical geometry. The vector multiplet moduli space of the type IIA theory, on the other hand, is not determined by classical geometry, but rather by "string geometry". The string effects are suppressed at large distances, i.e. when the Calabi-Yau manifold on which we compactify becomes large. At small distances, of the order of the string scale $l_s = 1/\sqrt{\alpha'}$, the intuition derived from classical geometry fails.

[13] A quaternionic manifold is a complex manifold of real dimension $4m$ and holonomy group $Sp(1) \times Sp(m)$.

It thus looks hopeless to compute the vector multiplet moduli space of the type IIA theory. Here mirror symmetry comes to rescue as was first shown, for the case of the quintic in $\mathbb{P}^4$, in [56]. It relates, via the mirror map, the vector multiplet moduli space of the type IIA theory on X to the vector multiplet moduli space of the type IIB theory on the mirror $\hat{X}$. Thus, with the help of mirror symmetry the structure of the vector multiplet moduli space of both type II theories is understood, as long as the conditions which lead to (3.104) are met. Due to lack of space we have to refer to the literature for any details [18, 21, 32].

One obtains the moduli space of the heterotic string by setting the (R,R) fields to zero. This gives

$$\mathcal{M}^{\text{het}} = \frac{SU(1,1)}{U(1)} \times \mathcal{M}_{h^{1,1}} \times \mathcal{M}_{h^{2,1}} \tag{3.106}$$

where the second and third factors are special-Kähler manifolds. (3.106), which was derived in [48, 50, 55, 57, 58], is only valid at string tree level. The loop corrections which destroy the product structure have been computed in [59].

We will now briefly explain the notion of a special Kähler manifold which arises in the construction of $\mathcal{N}=2$ supersymmetric couplings of vector multiplets to supergravity. It was found that the entire Lagrangian can be locally encoded in a holomorphic function $F(t)$, where t^a are (so-called special) coordinates on the space spanned by the scalar fields inside the vector multiplets. For instance, in type IIB compactification on a CY_3, this is the complex structure moduli space and $a = 1, \ldots, h^{2,1}$. Supersymmetry requires that this space is Kähler and furthermore, that its Kähler potential can be expressed through F via

$$\begin{aligned} K &= -\ln Y \\ Y &= 2(F - F) - (t^a - \bar{t}^a)(F_a + \bar{F}_a) \end{aligned} \tag{3.107}$$

where $F_a = \partial_a F$. For this reason F is called the (holomorphic) *prepotential*. If we introduce projective coordinates z via $t^a = z^a/z^0$ and define $\mathcal{F}(z) = (z^0)^2 F(t)$ we find that the Kähler potential (3.107) can be written, up to a Kähler transformation, as

$$K = \ln\left(\bar{z}^a \mathcal{F}_a - z^a \bar{\mathcal{F}}_a\right) \tag{3.108}$$

where now $a = 0, \ldots, h^{2,1}$, and $\mathcal{F}_a = \frac{\partial \mathcal{F}}{\partial z^a}$. Supersymmetry requires furthermore that $\mathcal{F}$ is a homogeneous function of degree two.

We will now show how these features are encoded in the CY geometry. We begin by introducing a basis of $H^3(X,\mathbb{Z})$ with generators α_a and β^b $(a, b = 0, \ldots, h^{2,1}(X))$ which are (Poincaré) dual to a canonical homology basis (B_a, A^b) of $H_3(X,\mathbb{Z})$ with intersection numbers $A^a \cdot A^b = B_a \cdot B_b = 0$, $A^a \cdot B_b = \delta^a_b$. Then

$$\int_{A^b} \alpha_a = \int_X \alpha_a \wedge \beta^b = -\int_{B_a} \beta^b = \delta_a^b \ . \tag{3.109}$$

All other pairings vanish. This basis is unique up to $Sp(2h^{2,1}+2,\mathbb{Z})$ transformations.

Following [60, 61], one can show that the A-periods of the holomorphic (3,0)-form Ω, i.e. $z^a = \int_{A^a} \Omega$ are local projective coordinates on the complex structure moduli space. We then have for the B-periods $\mathcal{F}_a = \int_{B_a} \Omega = \mathcal{F}_a(z)$. Note that $\Omega = z^a \alpha_a - \mathcal{F}_a \beta^a$. Furthermore, under a change of complex structure Ω, which was pure $(3,0)$ to start with, becomes a mixture of $(3,0)$ and $(2,1)$ (because dz in the old complex structure becomes a linear combination of dz and $d\bar{z}$ w.r.t. to the new complex structure): $\frac{\partial}{\partial z^a}\Omega \in H^{(3,0)} \oplus H^{(2,1)}$. In fact [27, 38, 42] $\frac{\partial \Omega}{\partial z^a} = k_a \Omega + b_a$ where $b_a \in H^{(2,1)}(X)$ and k_a is a function of the moduli but independent of the coordinates on X (since Ω is unique). One immediate consequence is that $\int \Omega \wedge \frac{\partial \Omega}{\partial z^a} = 0$. Inserting the expansion for Ω in the α_a, β^a basis into this equation, one finds $\mathcal{F}_a = \frac{1}{2}\frac{\partial}{\partial z^a}(z^b \mathcal{F}_b)$, or $\mathcal{F}_a = \frac{\partial \mathcal{F}}{\partial z^a}$ with $\mathcal{F} = \frac{1}{2} z^a \mathcal{F}_a$, $\mathcal{F}(\lambda z) = \lambda^2 \mathcal{F}(z)$. We thus identify z^a with the special coordinates of supergravity and $\mathcal{F}$ with the prepotential. It is easy to verify that the Kähler potential in the form (3.108) can be written as $K = -\ln \int \Omega \wedge \bar{\Omega}$. In fact, $\mathcal{F}$ can be explicitly computed for the complex structure moduli space of type IIB theory in terms of the periods of the holomorphic three-form. This is a calculation in classical geometry. The Kähler moduli space of type IIA theory is also characterized by a prepotential. However its direct calculation is very difficult since it receives contributions from world-sheet instantons. Mirror symmetry relates $\mathcal{F}^{\rm Kahler}(X)$ to the prepotential of the complex structure moduli space on the mirror manifold $\mathcal{F}^{\rm complex}(\hat{X})$ which can be computed and mapped, via the mirror map, to $\mathcal{F}^{\rm Kahler}(X)$ (see e.g. [21] for a review). In any case, it follows from this discussion that the metric on the Kähler part of the moduli space of type II Calabi-Yau compactifications can be computed explicitly.

In supergravity and superstring compactifications many other properties of special Kähler manifolds are relevant e.g. in the explicit construction of the mirror map, the computation of Yukawa couplings in heterotic compactifications, etc. All these details can be found in the cited references. Reference [62] discusses some subtle issues involving the existence of a prepotential (but see also [63] for their irrelevance in string compactification on CY manifolds once world-sheet instanton effects are included).

While, as we have seen, a great deal is known about the (local) geometry of the vector multiplet moduli space, the question about the structure of the hypermultiplet moduli space, except that it is a quaternionic manifold, is still largely unanswered and a subject of ongoing research. The difficulty comes, of course, from the fact that it receives perturbative and non-perturbative quantum corrections. Some partial results have been obtained e.g. in [64, 65].

4 Strings on Orbifolds

We now want to consider string compactifications in which the internal space belongs to a class of toroidal orbifolds that are analogous to Calabi-Yau spaces in that their holonomy group is contained in $SU(n)$ and therefore the theory in the lower dimensions has unbroken supersymmetry. Even though these orbifolds are singular, we will see that string propagation is perfectly consistent provided that twisted sectors are included. Moreover, since toroidal orbifolds are flat except at fixed points, the theory is exactly solvable. Indeed, the fields on the world-sheet satisfy free equations of motion with appropriate boundary conditions.

In this section we will first discuss some basic properties of orbifolds. We next describe in some detail the compactification of strings on orbifolds, introducing in the process the important concepts of partition function and modular invariance. Finally the general results are applied to type II theories. In appendix C we collect some useful results about the partition function of $\mathrm{T}^6/\mathbb{Z}_N$ orbifolds.

The standard references for strings on orbifolds are the original papers [3]. A concise review that also discusses conformal field theory aspects is [66].

4.1 Orbifold Geometry

In general, an orbifold $\mathcal{O}$ is obtained by taking the quotient of a manifold $\mathcal{M}$ by the action of a discrete group G that preserves the metric of $\mathcal{M}$. This means:

$$\mathcal{O} = \mathcal{M}/G \; . \tag{4.1}$$

For $g \in G$ and $x \in \mathcal{M}$, the points x and gx are equivalent in the quotient. Each point is identified with its orbit under G, hence the name orbifold. The fixed points of $\mathcal{M}$ under G are singular points of $\mathcal{O}$.

Perhaps the simplest example of an orbifold is the torus $\mathrm{T}^{\mathcal{D}}$ defined as

$$\mathrm{T}^{\mathcal{D}} = \mathbb{R}^{\mathcal{D}}/\Lambda \; , \tag{4.2}$$

where Λ is a $\mathcal{D}$-dimensional lattice. Hence, in $\mathrm{T}^{\mathcal{D}}$ the points x and $x+V$, $V \in \Lambda$, are identified. In the following we denote the basis of the torus lattice by e_a, $a = 1, \cdots, \mathcal{D}$. Figure 2 shows the case of T^2. Since the group of translations by lattice vectors acts freely, the torus has no singular points. However, when the discrete group leaves fixed points, the orbifold has singular points. A simple example is the cone obtained by taking the quotient of $\mathbb{C} \simeq \mathbb{R}^2$ by $\mathbb{Z}_N$ generated by multiplication by $e^{2i\pi/N}$. This is shown in Fig. 3. Notice that the origin, left fixed by $\mathbb{Z}_N$, is a singular point at which there is a deficit angle $2\pi(N-1)/N$.

Since we want compact spaces we are led to consider toroidal orbifolds $\mathrm{T}^{\mathcal{D}}/G_P$, where the so called point group $G_P \subset SO(\mathcal{D})$ is a discrete group that acts crystallographically on the torus lattice Λ. The elements of G_P are

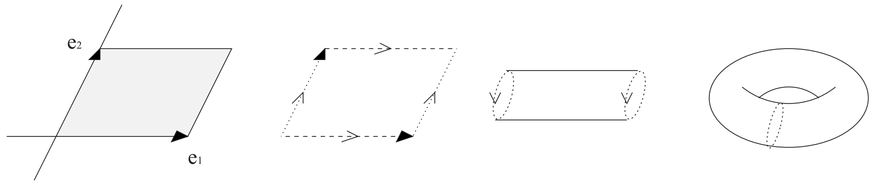


Fig. 2. $\mathrm{T}^2 = \mathbb{R}^2/\Lambda$

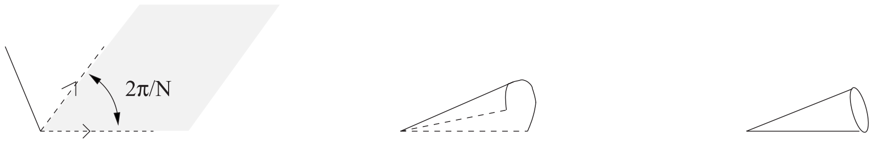


Fig. 3. $\mathbb{C}/\mathbb{Z}_N$

rotations denoted generically θ. Alternatively, toroidal orbifolds can be expressed as $\mathbb{R}^{\mathcal{D}}/S$, where S is the so-called space group that contains rotations and translations in Λ.

The point group is the holonomy group of the toroidal orbifold [3]. To show this, take two points x and y, distinct on the torus but such that $y = \theta x + V$. Then, x and y are identified on the orbifold and moreover the tangent vectors at x are identified with the tangent vectors at y rotated by θ. Next parallel-transport some vector along a path from x to y which is closed on the orbifold. The torus is flat and hence this vector remains constant but since the tangent basis is rotated by θ, the final vector is rotated by θ with respect to the initial vector. The loop from x to y necessarily encloses a singular point since otherwise there would be no curvature to cause the non-trivial holonomy.

In the following we will mostly consider point groups $G_P = \mathbb{Z}_N$. Then $\theta^N = \mathbb{1}$ and θ has eigenvalues $e^{\pm 2i\pi v_i}$, where $v_i = k_i/N$ for some integers k_i, $i = 1, \cdots, \mathcal{D}/2$ (we take $\mathcal{D}$ even). As we mentioned before, G_P must act crystallographically on the torus lattice. This means that for $V \in \Lambda$ and $\theta \in G_P$, $\theta V \in \Lambda$. Now, since $V = n_a \mathrm{e}_a$, with integer coefficients n_a, in the lattice basis θ must be a matrix of integers. Hence, the quantities

$$\mathrm{Tr}\,\theta = \sum_{i=1}^{\mathcal{D}/2} 2\cos 2\pi v_i$$

$$\chi(\theta) = \det(1-\theta) = \prod_{i=1}^{\mathcal{D}/2} 4\sin^2 \pi v_i \tag{4.3}$$

must be integers. Indeed, from Lefschetz fixed point theorem, $\chi(\theta)$ is the number of fixed points of θ. The upshot is that the requirement of crystallographic action is very restrictive. For instance, it is easy to find that for

$\mathcal{D} = 2$ only $N = 2, 3, 4, 6$ are allowed. In Table 1 we collect the irreducible possibilities for the v_i's when $\mathcal{D} = 2, 4, 6$ [67]. By irreducible we mean that the corresponding θ cannot be written in a block form. Notice that the case $\mathcal{D} = 2, v_1 = \frac{1}{2}$ is reducible since already in a one dimensional lattice a $\mathbb{Z}_2$ (only) is allowed.

Table 1. Irreducible crystallographic actions

$\mathcal{D} = 2$	$\mathcal{D} = 4$	$\mathcal{D} = 6$
(v_1)	(v_1, v_2)	(v_1, v_2, v_3)
$\frac{1}{3}(1)$	$\frac{1}{5}(1, 2)$	$\frac{1}{7}(1, 2, 3)$
$\frac{1}{4}(1)$	$\frac{1}{8}(1, 3)$	$\frac{1}{9}(1, 2, 4)$
$\frac{1}{6}(1)$	$\frac{1}{10}(1, 3)$	$\frac{1}{14}(1, 3, 5)$
	$\frac{1}{12}(1, 5)$	$\frac{1}{18}(1, 5, 7)$

Given the v_i's there remains the question of finding a concrete lattice Λ that has θ^n, $n = 1, \cdots, N$, as automorphisms. We refer the reader to [67, 68] for a discussion of these issues. Here we will mostly consider products of two-dimensional sub-lattices and for order two and order four rotations we take the $SO(4)$ root lattice whereas for order three and order six rotations we take the $SU(3)$ root lattice.

Let us now consider some examples.

Example 4.1: $\mathrm{T}^2(SO(4))/\mathbb{Z}_2$. Here $\mathbb{Z}_2$ has elements $\{\mathbb{1}, \theta\}$, where θ is a rotation by π. As Λ we take the root lattice of $SO(4)$ with basis $\mathrm{e}_1 = (1, 0)$ and $\mathrm{e}_2 = (0, 1)$. In T^2, $\mathbb{Z}_2$ has four fixed points:

$$f_0 = (0, 0) \; ; \quad f_1 = \left(\frac{1}{2}, 0\right) \; ; \quad f_2 = \left(0, \frac{1}{2}\right) \; ; \quad f_3 = \left(\frac{1}{2}, \frac{1}{2}\right) \; . \tag{4.4}$$

It is convenient to use a complex coordinate $z = x + iy$ so that $f_0 = 0$, $f_1 = \frac{1}{2}$, $f_2 = \frac{i}{2}$, $f_3 = \frac{1+i}{2}$.

The steps to construct the orbifold are shown in Fig. 4. To start, we take a fundamental cell defined by vertices $(0, 0), (1, 0), (0, 1), (1, 1)$. Given the identification $x \equiv \theta^n x + V$, we observe that it is actually enough to retain half of the fundamental cell, for instance the rectangle with vertices at f_0, f_1, i and $\frac{1}{2} + i$. Furthermore, since the edges are identified as indicated in Fig. 4 we must fold by the line joining f_2 and f_3. The resulting orbifold has singular points precisely at the f_i, each with a deficit angle of π.

Example 4.2: $\mathrm{T}^2(SU(3))/\mathbb{Z}_3$. Here $\mathbb{Z}_3$ has elements $\{\mathbb{1}, \theta, \theta^2\}$, where θ is a rotation by $2\pi/3$. As Λ we take the root lattice of $SU(3)$ with basis $\mathrm{e}_1 = (1, 0)$ and $\mathrm{e}_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$. The fixed points of θ are

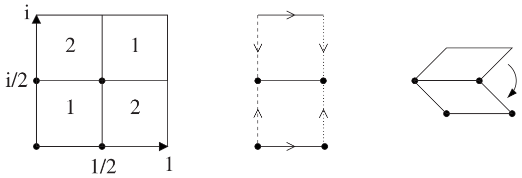


Fig. 4. $T^2/\mathbb{Z}_2$

$$f_0 = (0,0) \; ; \quad f_1 = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right) \; ; \quad f_2 = \left(0, \frac{1}{\sqrt{3}}\right) \; . \tag{4.5}$$

In terms of the complex coordinate z, θ acts by multiplication by $e^{2i\pi/3}$ and the fixed points are located at $0, \frac{1}{\sqrt{3}}e^{i\pi/6}, \frac{i}{\sqrt{3}}$. The element $\theta^2 = \theta^{-1}$ obviously has the same fixed points. In this case the resulting orbifold has singularities at the three fixed points, each with deficit angle $4\pi/3$.

In examples 4.1 and 4.2 the total deficit angle is 4π, i.e. the orbifold is topologically an S^2, as it is also clear from Fig. 4.

Example 4.3: $T^4(SO(4)^2)/\mathbb{Z}_2$. We take $T^4 = T^2 \times T^2$ and Λ the product of two 2-dimensional square $SO(4)$ root lattices. The $\mathbb{Z}_2$ action is just a rotation by π degrees in each square sub-lattice. In terms of $z^j = x^j + iy^j$ this means

$$\mathbb{Z}_2 : \quad (z^1, z^2) \to (-z^1, -z^2) \; . \tag{4.6}$$

In each sub-lattice there are four fixed points with complex coordinates $0, \frac{1}{2}, \frac{i}{2}, \frac{1+i}{2}$. Altogether the orbifold has then sixteen singular points.

Notice that there are no $\mathbb{Z}_2$ invariant $(1,0)$ harmonic forms and only one invariant $(2,0)$ harmonic form, namely $dz^1 \wedge dz^2$. This is an indication that the holonomy group of the orbifold is a subgroup of $SU(2)$. It turns out that the orbifold singularities at the fixed points can be "repaired" or "blown up" to produce a smooth manifold of $SU(2)$ holonomy, namely a smooth K3 [69]. Roughly, the idea is to excise the singular points and replace them by plugs that patch the holes smoothly. More precisely, the plugs are asymptotically Euclidean spaces (ALE) with metrics of $SU(2)$ holonomy that happen to be Eguchi-Hanson spaces. The claim that the resulting space is a smooth K3 manifold can be supported by a computation of the Hodge numbers of K3 in the orbifold picture. Firstly, the orbifold inherits the forms of T^4 that are invariant under $\mathbb{Z}_2$. Thus, the following are also harmonic forms on $T^4/\mathbb{Z}_2$:

$$1 \; , \quad dz^i \wedge d\bar{z}^j \; , \quad dz^1 \wedge dz^2 \; , \quad d\bar{z}^1 \wedge d\bar{z}^2 \; , \quad dz^1 \wedge dz^2 \wedge d\bar{z}^1 \wedge d\bar{z}^2 \; . \tag{4.7}$$

Secondly, the blowing up process gives a contribution of sixteen to $h^{1,1}$, one from the Eguchi-Hanson Kähler form at each fixed point. Then, altogether $h^{0,0} = h^{2,0} = h^{0,2} = h^{2,2} = 1$ and $h^{1,1} = 20$.

Example 4.4: $T^6(SU(3)^3)/\mathbb{Z}_3$. We take $T^6 = T^2 \times T^2 \times T^2$ and Λ the product of three $SU(3)$ root lattices. The $\mathbb{Z}_3$ group is generated by an order three rotation in each sub-lattice. In terms of complex coordinates the $\mathbb{Z}_3$ action is

$$(z^1, z^2, z^3) \to \left(e^{2i\pi/3} z^1, e^{2i\pi/3} z^2, e^{-4i\pi/3} z^3\right) . \tag{4.8}$$

In each sub-lattice there are three fixed points located at 0, $\frac{1}{\sqrt{3}} e^{i\pi/6}$, $\frac{i}{\sqrt{3}}$. The full orbifold has thus 27 singular points.

The singular points can be repaired to obtain a smooth manifold, the so-called Z-manifold that is a CY$_3$ [2]. The (3,0) harmonic form that must exist in every CY$_3$ is simply $dz^1 \wedge dz^2 \wedge dz^3$ that is $\mathbb{Z}_3$ invariant. The interesting Hodge numbers are computed as follows. Clearly, the nine $dz^i \wedge d\bar{z}^j$ forms are $\mathbb{Z}_3$ invariant. There are no (1,2) invariant forms on T^6. The blowing up process adds 27 (1,1) harmonic forms. Then, $h^{1,1} = 9 + 27$, $h^{1,2} = 0$ and $\chi = 72$.

To end this section we would like to address the question whether string compactification on a given orbifold can give a supersymmetric theory in the lower dimensions. We consider $\mathcal{D} = 6$, the results for $\mathcal{D} = 2, 4$ come as by-products. According to our discussion in Sect. 2.2, supersymmetry requires the existence of covariantly constant spinors. This means that there must exist spinors ϵ such that $\theta\epsilon = \epsilon$. In our case θ is an $SO(6)$ rotation with eigenvalues $e^{\pm 2i\pi v_i}$ acting on the vector representation that has weights $(\pm1, 0, 0)$, $(0, \pm1, 0)$ and $(0, 0, \pm1)$. In fact, we can write θ as

$$\theta = \exp\left(2\pi i(v_1 J_{12} + v_2 J_{34} + v_3 J_{56})\right) , \tag{4.9}$$

where the $J_{2i-1,2i}$ are the generators of the Cartan subalgebra. Now, since spinor weights of $SO(6)$ are $(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$, in this representation θ has eigenvalues $e^{i\pi(\pm v_1 \pm v_2 \pm v_3)}$. Hence, to have invariant spinors we need

$$\pm v_1 \pm v_2 \pm v_3 = 0 \bmod 2 \tag{4.10}$$

for some choice of signs. This condition guarantees that the holonomy group is contained in $SU(3)$. The additional condition $N(v_1 + v_2 + v_3) = 0$ mod 2, which follows from modular invariance, is derived in Appendix C. When $v_3 = 0$, from Table 1 we find that the only solutions are $v_1 = -v_2 = 1/N$, $N = 2, 3, 4, 6$. The case $N = 2$ is example 4.3 above, for other N's the corresponding orbifolds of T^4 are also singular limits of K3. For orbifolds of T^6, we can again use the data in Table 1 together with (4.10) to obtain all the allowed inequivalent solutions shown in Table 2 that were first found in [3]. The resulting T$^6/\mathbb{Z}_N$ orbifolds are generalizations of Calabi-Yau three-folds. In all cases it can be proved that the singular points can be resolved to obtain smooth manifolds of $SU(3)$ holonomy [68, 71].

4.2 Orbifold Hilbert Space

In this section we wish to discuss some general aspects of the propagation of closed strings on orbifolds [3]. We will explain how to determine the states

Table 2. Supersymmetric $\mathbb{Z}_N$ actions

$\mathbb{Z}_3$	$\frac{1}{3}(1,1,-2)$	$\mathbb{Z}_6'$	$\frac{1}{6}(1,-3,2)$	$\mathbb{Z}_8'$	$\frac{1}{8}(1,-3,2)$
$\mathbb{Z}_4$	$\frac{1}{4}(1,1,-2)$	$\mathbb{Z}_7$	$\frac{1}{7}(1,2,-3)$	$\mathbb{Z}_{12}$	$\frac{1}{12}(1,-5,4)$
$\mathbb{Z}_6$	$\frac{1}{6}(1,1,-2)$	$\mathbb{Z}_8$	$\frac{1}{8}(1,3,-4)$	$\mathbb{Z}_{12}'$	$\frac{1}{12}(1,5,-6)$

belonging to the physical Hilbert space, taking into account a projection on states invariant under the orbifold group, as well as including twisted sectors.

Let $X^m(\sigma^0,\sigma^1)$, $m=1,\cdots\mathcal{D}$, be bosonic coordinates depending on the world-sheet time and space coordinates σ^0 and σ^1. Since the string is closed, σ^1 is periodic, we take its length to be 2π. We assume that $\mathcal{M}$ is flat so that before taking the quotient to obtain the orbifold, X^m satisfies the free wave equation

$$(\partial_0^2-\partial_1^2)X^m=0\ . \tag{4.11}$$

Furthermore, there are boundary conditions

$$X^m(\sigma^0,\sigma^1+2\pi)=X^m(\sigma^0,\sigma^1)\ . \tag{4.12}$$

The equations of motion follow from the action

$$S=\int d^2\sigma\,\mathcal{L}=-\frac{1}{4\pi\alpha'}\int d^2\sigma\eta^{\alpha\beta}\partial_\alpha X^m\partial_\beta X_m\ . \tag{4.13}$$

This is the Polyakov action (1.1) in flat space-time and in conformal gauge $h_{\alpha\beta}=\eta_{\alpha\beta}=\mathrm{diag}(-1,1)$. The canonical conjugate momentum is $\Pi_m=\partial\mathcal{L}/\partial(\partial_0X^m)$. In the following we will drop the index m to simplify notation. The generator of translations in X is $P=\int_0^{2\pi}d\sigma^1\Pi$.

Now, in the orbifold we know that each point is identified with its orbit under $g\in G$. Hence, as physical states we should consider only the sub-space invariant under the action of g. The appropriate projection operator is

$$\mathcal{P}=\frac{1}{|G|}\sum_{g\in G}\bar{g}\ , \tag{4.14}$$

where $\bar{g}$ is the realization of g on the string states.

Exercise 4.1: Show that $\mathcal{P}^2=\mathcal{P}$.

For example, consider the quotient of $\mathbb{R}^{\mathcal{D}}$ by translations in a lattice Λ to obtain $\mathrm{T}^{\mathcal{D}}$. Since the generator of space-time translations is the momentum P, to each $W\in\Lambda$ the operator acting on states is $e^{2\pi iP\cdot W}$ (the factor of 2π is for convenience). Then, the sub-space of invariant states contains only strings whose center of mass momentum (the eigenvalue of P) belongs to the dual lattice Λ^*. Indeed, notice that $\sum_{W\in\Lambda}e^{2\pi iP\cdot W}$ vanishes unless $P\in\Lambda^*$. Recall that Λ^* is the set of all vectors that have integer scalar product with any vector in Λ. In this case $|G|$ is equal to the volume Vol(Λ) of the unit cell of Λ. It can be shown that $\mathrm{Vol}(\Lambda)\mathrm{Vol}(\Lambda^*) = 1$.

In the orbifoldized theory there appear naturally *twisted* sectors in which X closes up to a transformation $h \in G$. This is:

$$X(\sigma^0, \sigma^1 + 2\pi) = hX(\sigma^0, \sigma^1) \; . \tag{4.15}$$

The *untwisted* sector has $h = \mathbb{1}$. In the example of $\mathrm{T}^{\mathcal{D}}$, the twisted sectors have boundary conditions

$$X(\sigma^0, \sigma^1 + 2\pi) = X(\sigma^0, \sigma^1) + 2\pi W \; , \quad W \in \Lambda \; . \tag{4.16}$$

Thus, the twisted sectors are just the winding sectors in which the string wraps around the torus cycles.

The twisted states must be included in order to ensure modular invariance. It is instructive to see this in the $\mathrm{T}^{\mathcal{D}}$ compactification. To begin, consider the solution to (4.11) together with (4.16). Left and right moving modes are independent so that $X = X_L + X_R$, with expansions

$$\begin{aligned} X_L(\sigma^0, \sigma^1) &= x_L + P_L(\sigma^0 + \sigma^1) + i \sum_{n \neq 0} \frac{\alpha_n}{n} \, e^{-in(\sigma^0 + \sigma^1)} \\ X_R(\sigma^0, \sigma^1) &= x_R + P_R(\sigma^0 - \sigma^1) + i \sum_{n \neq 0} \frac{\tilde{\alpha}_n}{n} \, e^{-in(\sigma^0 - \sigma^1)} \; , \end{aligned} \tag{4.17}$$

where

$$(P_L, P_R) = \left(P + \frac{W}{2}, P - \frac{W}{2} \right) \; , \quad P \in \Lambda^* \; , \quad W \in \Lambda \; . \tag{4.18}$$

For simplicity we are setting $\alpha' = 2$ everywhere. The Fourier coefficients α_n and $\tilde{\alpha}_n$ are commonly called *oscillator modes*. Quantization proceeds in the standard way by promoting the expansion coefficients to operators and imposing equal time canonical commutation relations that imply $[\alpha_m, \alpha_n] = m\delta_{m,-n}$, $[\tilde{\alpha}_m, \tilde{\alpha}_n] = m\delta_{m,-n}$. Furthermore, $[x_L, P_L] = i$ and $[x_R, P_R] = i$. It is convenient to introduce the occupation number operators

$$\mathcal{N}_L = \sum_{n=1}^{\infty} \alpha_{-n} \, \alpha_n \quad , \quad \mathcal{N}_R = \sum_{n=1}^{\infty} \tilde{\alpha}_{-n} \, \tilde{\alpha}_n \; . \tag{4.19}$$

The vacuum state $|0, 0, k_L, k_R\rangle$ is defined to be annihilated by $\alpha_n, \tilde{\alpha}_n$, $n > 0$, and to be an eigenvector of the momenta (P_L, P_R) with eigenvalues (k_L, k_R) of the form (4.18). Acting on the vacuum with creation operators $\alpha_{-n}, \tilde{\alpha}_{-n}$, $n > 0$, gives states $|N_L, N_R, k_L, k_R\rangle$ that have generic eigenvalues N_L and N_R of the occupation number operators. For instance, $(\alpha_{-n_1})^{\ell_1} \, (\tilde{\alpha}_{-n_2})^{\ell_2} \, |0, 0, k_L, k_R\rangle$ has $N_L = n_1\ell_1$ and $N_R = n_2\ell_2$.

The Hamiltonian is

$$H = \int_0^{2\pi} d\sigma^1 (\Pi \cdot \partial_0 X - \mathcal{L}) = \frac{1}{8\pi} \int_0^{2\pi} d\sigma^1 \left[(\partial_0 X)^2 + (\partial_1 X)^2 \right] . \qquad (4.20)$$

Substituting the expansions (4.17) then gives

$$H = \frac{P_L^2}{2} + \frac{P_R^2}{2} + \mathcal{N}_L + \mathcal{N}_R - \frac{\mathcal{D}}{12} . \qquad (4.21)$$

The constant term comes from normal ordering all annihilation operators to the right and using the analytical continuation of the zeta function to regularize the sum $\sum_{n=1}^{\infty} n = \zeta(-1) = -1/12$. The Hamiltonian is the generator of translations in σ^0, meaning that $[H, X] = -i\partial_0 X$. The generator of translations in σ^1 is

$$P_\sigma = \int_0^{2\pi} d\sigma^1 \Pi \cdot \partial_1 X = \frac{P_L^2}{2} - \frac{P_R^2}{2} + \mathcal{N}_L - \mathcal{N}_R . \qquad (4.22)$$

Both H and P_σ can be written in terms of left and right moving Virasoro generators as

$$H = L_0 + \tilde{L}_0 \ , \quad P_\sigma = L_0 - \tilde{L}_0 \ . \qquad (4.23)$$

Then,

$$L_0 = \frac{P_L^2}{2} + \mathcal{N}_L - \frac{\mathcal{D}}{24} \ ; \quad \tilde{L}_0 = \frac{P_R^2}{2} + \mathcal{N}_R - \frac{\mathcal{D}}{24} \ . \qquad (4.24)$$

Since $\mathcal{D}$ free bosons have central charge $c = \mathcal{D}$, the constant term is the expected $-c/24$. The eigenvalue of L_0 ($\tilde{L}_0$) is the squared mass m_L^2 (m_R^2) of the given state. Invariance under translations along the closed string requires that P_σ vanishes acting on states. This implies the level-matching condition $m_R^2 = m_L^2$.

We next consider the partition function defined as

$$\mathcal{Z}(\tau, \bar{\tau}) = \mathrm{Tr}\, q^{L_0}\, \bar{q}^{\tilde{L}_0} \ ; \quad q \equiv e^{2i\pi\tau} \ ; \quad \tau \in \mathbb{C} \ , \qquad (4.25)$$

where the trace is taken over the states $|N_L, N_R, k_L, k_R\rangle$. Knowing the spectrum we can simply compute $\mathcal{Z}(\tau, \bar{\tau})$ by counting the number of states at each level of $L_0, \tilde{L}_0$. For the toroidal compactification one finds

$$\mathcal{Z}(\tau, \bar{\tau}) = \frac{1}{|\eta(\tau)|^{2\mathcal{D}}} \sum_{P \in \Lambda^*} \sum_{W \in \Lambda} q^{\frac{1}{2}(P + \frac{W}{2})^2}\, \bar{q}^{\frac{1}{2}(P - \frac{W}{2})^2} \ . \qquad (4.26)$$

The Dedekind eta function,

$$\eta(\tau) = q^{\frac{1}{24}} \prod_{k=1}^{\infty} (1 - q^k) \ , \qquad (4.27)$$

arises from the contribution of the oscillator modes.

Exercise 4.2: Show (4.26).

The partition function (4.26) has the remarkable property of being invariant under the modular transformations

$$\tau \to \frac{a\tau + b}{c\tau + d} \; ; \quad a, b, c, d \in \mathbb{Z} \; ; \quad ad - bc = 1 \; . \tag{4.28}$$

The $SL(2, \mathbb{Z})$ modular group is generated by the transformations $\mathcal{T} : \tau \to \tau + 1$ and $\mathcal{S} : \tau \to -1/\tau$. Invariance of (4.26) under $\mathcal{T}$ follows simply because $2P \cdot W =$ even. Invariance under $\mathcal{S}$ arises only because the partition function includes a sum over windings.

Exercise 4.3: Prove invariance of (4.26) under $\mathcal{S}$ using the property $\eta(-1/\tau) = \sqrt{-i\tau}\,\eta(\tau)$ and the Poisson resummation formula

$$\sum_{W \in \Lambda} e^{-\pi a (W+U)^2} \, e^{2i\pi Y \cdot (W+U)} = \frac{1}{\mathrm{Vol}(\Lambda)\, a^{\mathcal{D}/2}} \sum_{P \in \Lambda^*} e^{-\frac{\pi}{a}(P+Y)^2} \, e^{-2i\pi P \cdot U} \; , \tag{4.29}$$

where U and Y are arbitrary vectors and a is a positive constant.

Physically, the partition function $\mathcal{Z}(\tau, \bar{\tau})$ corresponds to the vacuum to vacuum string amplitude at one-loop. In this case the world-sheet surface is a torus T^2 that has precisely τ as modular parameter. From the brief discussion after (2.26) recall that T^2 with modular parameter $\tau = \tau_1 + i\tau_2$ can be defined by identifications in a lattice with basis $\mathrm{e}_1 = (1, 0)$, $\mathrm{e}_2 = (\tau_1, \tau_2)$. We can picture the T^2 as formed by a cylinder of length τ_2 in which we identify the string at the initial end with the string at the final end after translating by τ_1. Indeed, using (4.23) we find

$$\mathcal{Z}(\tau, \bar{\tau}) = \mathrm{Tr}\, e^{-2\pi\tau_2 H} \, e^{2i\pi\tau_1 P_\sigma} \; . \tag{4.30}$$

The first term in the trace is precisely what we expect of a partition function for a system propagating for Euclidean time $2\pi\tau_2$. The second term reflects a translation by $2\pi\tau_1$ in the coordinate σ^1 along the string. Now, the modular transformations (4.28) just correspond to an integral change of basis in the T^2 lattice. For example, Fig. 5 shows three equivalent lattices for T^2. All tori with τ's related by modular transformations are conformally equivalent and the partition function must therefore remain invariant.

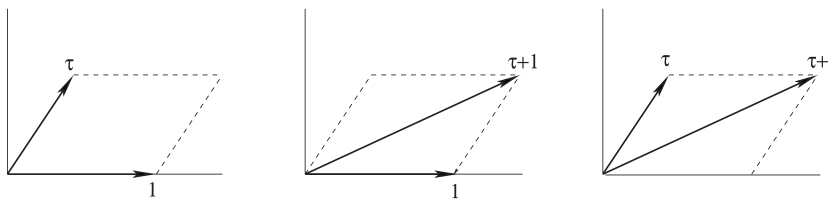


Fig. 5. Three equivalent T^2 lattices

In the example of toroidal compactification, the partition function is modular invariant only because the winding sectors are included. Indeed, $\mathcal{S}$ basically exchanges σ^0 and σ^1 so it transforms quantized momenta into windings. In general, the partition function for orbifold compactification is modular invariant only if twisted sectors are included. To see this, we start with the untwisted sector and implement the projection on invariant states according to (4.14). The partition function in the untwisted sector then becomes

$$\mathcal{Z}_{\mathbb{1}}(\tau,\bar{\tau}) = \mathrm{Tr}\left(\mathcal{P}\, q^{L_0(\mathbb{1})}\, \bar{q}^{\tilde{L}_0(\mathbb{1})}\right) = \frac{1}{|G|}\sum_{g\in G} \mathrm{Tr}\left(\bar{g}\, q^{L_0(\mathbb{1})}\, \bar{q}^{\tilde{L}_0(\mathbb{1})}\right) . \tag{4.31}$$

Due to the insertion of $\bar{g}$, the traces in the sum above are over states that satisfy not only the untwisted boundary condition (4.12) but also

$$X(\sigma^0 + 2\pi\tau_2, \sigma^1 + 2\pi\tau_1) = gX(\sigma^0, \sigma^1) . \tag{4.32}$$

We can then write schematically

$$\mathcal{Z}_{\mathbb{1}}(\tau,\bar{\tau}) = \frac{1}{|G|}\sum_{g\in G} \mathcal{Z}(\mathbb{1}, g) , \tag{4.33}$$

where $\mathcal{Z}(h,g)$ means partition function with boundary conditions (4.15) in σ^1 and (4.32) in σ^0. Now, under modular transformations the boundary conditions do change. For instance, under $\mathcal{T} : \tau \to \tau+1$, $(h,g) \to (h,gh)$, and under $\mathcal{TST} : \tau \to \tau/(\tau+1)$, $(h,g) \to (gh,g)$, as implied by the change of basis depicted in Fig. 5. Then, under $\mathcal{S} : \tau \to -1/\tau$, $(h,g) \to (g,h^{-1})$ and in particular $\mathcal{S}$ transforms the untwisted sector into a twisted sector. To obtain a modular invariant partition function we must include all sectors. More precisely, for Abelian G the full partition function has the form

$$\begin{aligned}\mathcal{Z}(\tau,\bar{\tau}) &= \frac{1}{|G|}\sum_{h\in G}\sum_{g\in G} \mathcal{Z}(h,g) \\ &= \sum_{h\in G}\left[\frac{1}{|G|}\sum_{g\in G} \mathrm{Tr}\left(\bar{g}\, q^{L_0(h)}\, \bar{q}^{\tilde{L}_0(h)}\right)\right] .\end{aligned} \tag{4.34}$$

The sum over h is a sum over twisted sectors while the sum over g implements the orbifold projection in each sector. For non-Abelian G we only sum over h and g such that $[h,g]=0$ since otherwise (4.15) and (4.32) are incompatible.

4.3 Bosons on $\mathbf{T}^{\mathcal{D}}/\mathbb{Z}_N$

We now wish to derive the partition function for bosonic coordinates compactified on $\mathrm{T}^{\mathcal{D}}/\mathbb{Z}_N$, with $\mathbb{Z}_N$ generated by θ as described in Sect. 4.1, and torus lattice Λ. We consider *symmetric* orbifolds in which θ acts equally on left and right movers. As we have explained, we need to include sectors twisted by $\theta^k, k = 0, \cdots, N-1$, in which the boundary conditions are

$$X(\sigma^0, \sigma^1 + 2\pi) = \theta^k X(\sigma^0, \sigma^1) + 2\pi V \; ; \quad V \in \Lambda \, . \tag{4.35}$$

The X's still satisfy the free equations of motion (4.11) so they have mode expansions of the form

$$X(\sigma^0, \sigma^1) = X_0 + 2P\sigma^0 + W\sigma^1 + \text{oscillators} \, . \tag{4.36}$$

To simplify the analysis we will assume that θ^k leaves no invariant directions so that the boundary conditions generically do not allow quantized momenta nor windings in the expansion. For the center of mass coordinate X_0 we find that it must satisfy $(1-\theta^k)X_0 = 0$ modulo $2\pi\Lambda$ which just means that X_0 is a fixed point of θ^k.

To find out the effect on the oscillator modes it is useful to define complex coordinates $z^j = \frac{1}{\sqrt{2}}(X^{2j-1} + iX^{2j})$, $j = 1, \cdots, \mathcal{D}/2$, such that $\theta z^j = e^{2i\pi v_j} z^j$ as we have seen in Sect. 4.1. Next write the z^j expansion as

$$z^j(\sigma^0, \sigma^1) = z_0^j + i\sum_t \frac{\alpha_t^j}{t}\, e^{-it(\sigma^0+\sigma^1)} + i\sum_s \frac{\tilde{\alpha}_s^j}{s}\, e^{-is(\sigma^0-\sigma^1)} \, , \tag{4.37}$$

where the frequencies t and s are to be determined by imposing the boundary condition (4.35). In this way we obtain $e^{-2i\pi t} = e^{2i\pi k v_j}$ and then $t = n - kv_j$, with n integer. Likewise, $s = n + kv_j$. For the complex conjugate $\bar{z}^j$ there is an analogous expansion with coefficients $\bar{\alpha}^j_{n+kv_j}$ and $\bar{\tilde{\alpha}}^j_{n-kv_j}$. Let us focus on the left-movers. After quantization, $[\bar{\alpha}^i_{m+kv_i}, \alpha^j_{n-kv_j}] = (m + kv_j)\delta^{i,j}\delta_{m,-n}$, with other commutators vanishing. There are now several Fock vacua $|f, 0\rangle_k$, where $f = 1, \cdots, \chi(\theta^k)$, is the fixed point label. Each vacuum is annihilated by all positive-frequency modes. The creation operators are thus $\alpha^j_{-kv_j}, \alpha^j_{-1-kv_j}, \cdots$ and $\bar{\alpha}^j_{-1+kv_j}, \bar{\alpha}^j_{-2+kv_j}, \cdots$ (assuming $0 < kv_j < 1$). The occupation number operator is

$$\mathcal{N}_L = \sum_{n=-\infty}^{\infty} :\, \alpha^j_{-n-kv_j}\, \bar{\alpha}^j_{n+kv_j} \,: \, , \tag{4.38}$$

where :: means normal ordering, i.e. all positive-frequency modes to the right. For right-movers the results are analogous.

We now construct the partition function that according to (4.34) has the form

$$\mathcal{Z} = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{\ell=0}^{N-1} \mathcal{Z}(\theta^k, \theta^\ell) \, ,$$
$$\mathcal{Z}(\theta^k, \theta^\ell) = \mathrm{Tr}\,(\theta^\ell\, q^{L_0(\theta^k)}\, \bar{q}^{\tilde{L}_0(\theta^k)}) \, . \tag{4.39}$$

The strategy is to start with the untwisted sector ($k = 0$) in which the Virasoro operators $L_0(\mathbb{1})$ and $\tilde{L}_0(\mathbb{1})$ are those given in (4.24). In particular, $\mathcal{Z}(\mathbb{1}, \mathbb{1})$ is just (4.26). For $\ell \neq 0$ we need to evaluate the trace with the θ^ℓ insertion. Since we are assuming that θ^ℓ leaves no unrotated directions, neither quantized momenta nor windings survive the trace. We only need to consider states obtained from the Fock vacuum by acting with creation operators which for the complex coordinates are eigenvectors of θ^ℓ. The Fock vacuum, denoted $|0\rangle_0$, is defined to be invariant under θ. Then, for instance, for the left movers in z^j we find the contribution

$$\mathrm{Tr}\left(\theta^\ell\, q^{L_0^j(\mathbb{1})}\right) = q^{-1/12}\left(1 + qe^{2i\pi\ell v_j} + qe^{-2i\pi\ell v_j} + \cdots\right) . \tag{4.40}$$

The first term comes from $|0\rangle_0$, the next two from states with α^j_{-1} and $\bar{\alpha}^j_{-1}$ acting on $|0\rangle_0$, and so on. In fact, the whole expansion can be cast as

$$\mathrm{Tr}\left(\theta^\ell\, q^{L_0^j(\mathbb{1})}\right) = q^{-1/12}\prod_{n=1}^{\infty}(1 - q^n e^{2i\pi\ell v_j})^{-1}\,(1 - q^n e^{-2i\pi\ell v_j})^{-1} . \tag{4.41}$$

This result can be conveniently written by using Jacobi ϑ functions that have the product representation

$$\frac{\vartheta\left[{\delta \atop \varphi}\right](\tau)}{\eta(\tau)} = e^{2i\pi\delta\varphi}\, q^{\frac{1}{2}\delta^2 - \frac{1}{24}}\prod_{n=1}^{\infty}(1+q^{n+\delta-\frac{1}{2}}e^{2i\pi\varphi})\,(1+q^{n-\delta-\frac{1}{2}}e^{-2i\pi\varphi}) . \tag{4.42}$$

Then,

$$\mathrm{Tr}\left(\theta^\ell\, q^{L_0^j(\mathbb{1})}\right) = -2\sin\ell\pi v_j \frac{\eta(\tau)}{\vartheta\left[{\frac{1}{2} \atop \frac{1}{2}+\ell v_j}\right](\tau)} . \tag{4.43}$$

Notice that for $\ell = 0$, (4.41) becomes $1/\eta^2$, as it should. Taking into account left and right movers for all coordinates we obtain

$$\mathcal{Z}(\mathbb{1}, \theta^\ell) = \chi(\theta^\ell)\left|\prod_{j=1}^{\mathcal{D}/2}\frac{\eta}{\vartheta\left[{\frac{1}{2} \atop \frac{1}{2}+\ell v_j}\right]}\right|^2 , \tag{4.44}$$

where $\chi(\theta^\ell) = \prod_{j=1}^{\mathcal{D}/2} 4\sin^2 \pi\ell v_j$ is the number of fixed points of θ^ℓ, cf. (4.3). We remark, as it is clear from (4.41), that the coefficient of the first term in the expansion in (4.44) is actually one. This means that in the full untwisted sector, i.e. fixing $k = 0$ and summing over ℓ, the untwisted vacuum appears with the correct multiplicity one.

To obtain other pieces $\mathcal{Z}(\theta^k, \theta^\ell)$ we take advantage of modular invariance. For example, $\mathcal{Z}(\theta^k, \mathbb{1})$ simply follows applying $\tau \to -1/\tau$ to (4.44). Using the modular properties of ϑ functions given in (C.4) gives

$$\begin{aligned}\mathcal{Z}(\theta^k, \mathbb{1}) &= \chi(\theta^k) \left| \prod_{j=1}^{\mathcal{D}/2} \frac{\eta}{\vartheta \left[\begin{smallmatrix} \frac{1}{2}+kv_j \\ \frac{1}{2} \end{smallmatrix} \right]} \right|^2 \\ &= \chi(\theta^k)\, (q\bar{q})^{-\frac{\mathcal{D}}{24}+E_k} \left| \prod_{j=1}^{\mathcal{D}/2} \prod_{n=1}^{\infty} (1-q^{n-1+kv_j})^{-1}\, (1-q^{n-kv_j})^{-1} \right|^2 ,\end{aligned} \tag{4.45}$$

where E_k is the twisted oscillator contribution to the zero point energy given by

$$E_k = \sum_{j=1}^{\mathcal{D}/2} \frac{1}{2} kv_j (1-kv_j)\ . \tag{4.46}$$

When $kv_j > 1$ we must substitute $kv_j \to (kv_j - 1)$ in (4.46).

Exercise 4.4: Derive (4.45).

The lowest order term in the expansion (4.45) does have coefficient $\chi(\theta^k)$ in agreement with the fact that in the θ^k sector the center of mass coordinate can be any fixed point. The q expansion also shows the contribution of the states created by operators $\alpha^j_{-kv_j}$, $\alpha^j_{-1-kv_j}, \cdots$ and $\bar{\alpha}^j_{-1+kv_j}$, $\bar{\alpha}^j_{-2+kv_j}, \cdots$. In fact, from the exponents of q we can read off the eigenvalues of $L_0(\theta^k)$, i.e. the squared masses $m_L^2(\theta^k)$. The general result can be written as

$$m_L^2(\theta^k) = N_L + E_k - \frac{\mathcal{D}}{24}\ . \tag{4.47}$$

Here N_L is the occupation number of the left-moving oscillators. For example, $\alpha^j_{-kv_j}|0\rangle_k$ and $\bar{\alpha}^j_{-1+kv_j}|0\rangle_k$ have $N_L = kv_j$ and $N_L = 1 - kv_j$, respectively. In the untwisted sector, or more generically in sectors in which quantized momenta or windings are allowed, m_L^2 also includes a term of the form $\frac{1}{2}P_L^2$. For particular shapes of the torus, $\frac{1}{2}P_L^2$ can precisely lead to extra massless states that signal enhanced symmetries as in the well known example of circle compactification at the self-dual radius. In these notes we will assume a generic point in the torus moduli space so that $\frac{1}{2}P_L^2$ does not produce new massless states. For right movers, $m_R^2(\theta^k)$ is completely analogous to (4.47). Notice that the level-matching condition becomes $N_L = N_R$.

We can continue generating pieces of the partition function by employing modular transformations. For example, applying $\tau \to \tau + 1$ to (4.45) gives $\mathcal{Z}(\theta^k, \theta^k)$. The general result can be written as

$$\mathcal{Z}(\theta^k, \theta^\ell) = \chi(\theta^k, \theta^\ell) \left| \prod_{j=1}^{\mathcal{D}/2} \frac{\eta}{\vartheta \left[\begin{smallmatrix} \frac{1}{2}+kv_j \\ \frac{1}{2}+\ell v_j \end{smallmatrix} \right]} \right|^2 , \tag{4.48}$$

where $\chi(\theta^k, \theta^\ell)$ is the number of simultaneous fixed points of θ^k and θ^ℓ. This formula is valid when θ^k leaves no fixed directions, otherwise a sum over momenta and windings could appear. This is important when determining the $\mathbb{Z}_N$-invariant states [67]. The correct result can be found by carefully determining the untwisted sector pieces and then performing modular transformations.

Exercise 4.5: Use (C.4) to show that (4.48) has the correct modular transformations, i.e. $\mathcal{Z}(\theta^k, \theta^\ell)$ transforms into $\mathcal{Z}(\theta^k, \theta^{k+\ell})$ under $\mathcal{T}$ and into $\mathcal{Z}(\theta^\ell, \theta^{-k})$ under $\mathcal{S}$.

Let us now describe the spectrum in a θ^k twisted sector. States are chains of left and right moving creation operators acting on the vacuum. Schematically this is

$$\alpha \cdots \bar{\alpha} \cdots \tilde{\alpha} \cdots \bar{\tilde{\alpha}} \cdots |f, 0\rangle_k \ . \tag{4.49}$$

Level-matching $N_L = N_R$ must be satisfied. States are further characterized by their transformation under a $\mathbb{Z}_N$ element, say θ^ℓ. The oscillator piece is just multiplied by an overall phase $e^{2i\pi\ell\rho}$, where $\rho = \rho_L + \rho_R$. In turn ρ_L (ρ_R) is found by adding the phases of all left (right) modes in (4.49). Concretely, each left-moving oscillator $\alpha^j_{-kv_j}, \alpha^j_{-1-kv_j}, \cdots$ (coming from z^j) adds v_j to ρ_L, whereas each $\bar{\alpha}^j_{-1+kv_j}, \bar{\alpha}^j_{-2+kv_j}, \cdots$ (coming from $\bar{z}^j$) contributes $-v_j$ to ρ_L. For right-movers, each mode $\bar{\tilde{\alpha}}^j_{-kv_j}, \bar{\tilde{\alpha}}^j_{-1-kv_j}, \cdots$, contributes $-v_j$ to ρ_R and each $\tilde{\alpha}^j_{-1+kv_j}, \tilde{\alpha}^j_{-2+kv_j}, \cdots$ adds v_j to ρ_R. Finally, the action on the fixed points must be $\theta^\ell |f, 0\rangle_k = |f', 0\rangle_k$, where f' is also a fixed point of θ^k.

Only states invariant under the full $\mathbb{Z}_N$ action survive in the spectrum. For example, in the untwisted sector ($k = 0$), both $\alpha^1_{-1}\tilde{\alpha}^1_{-1}|0\rangle_0$ and $\bar{\alpha}^1_{-1}\tilde{\alpha}^1_{-1}|0\rangle_0$ have $N_L = N_R = 1$ but the first is not invariant because it picks up a phase $e^{4i\pi v_1}$ under θ. For $k \neq 0$ there is a richer structure because states sit at fixed points. In the θ sector, $\chi(\theta, \theta^\ell) = \chi(\theta)$, i.e. all θ^ℓ leave the fixed points of θ invariant. Hence, $|f, 0\rangle_1$ and chain states (4.49) with $\rho_L + \rho_R = 0$ are invariant $\forall f$, meaning that there is one such state at each fixed point of θ. For N odd, $\chi(\theta^k, \theta^\ell) = \chi(\theta^k) = \chi(\theta)$, so that all twisted sectors are like the θ sector.

For N even, in general $\chi(\theta^k, \theta^\ell)$, $k \neq 1, N-1$, depends on ℓ. For example, take a $\mathrm{T}^2/\mathbb{Z}_4$ with square $SO(4)$ lattice (cf. Example 4.1) and θ a $\pi/2$ rotation ($v_1 = 1/4$). Then, θ^2 has the four fixed points in (4.4): f_0 and f_3 that are also fixed by θ, plus f_1 and f_2 that are exchanged by θ. Thus, in the θ^2 sector, there are three invariant vacua, namely $|f_0, 0\rangle_2$, $|f_3, 0\rangle_2$ and $[|f_1, 0\rangle_2 + |f_2, 0\rangle_2]$. Likewise, any level-matched chain, e.g. $\bar{\alpha}_{-\frac{1}{2}}\tilde{\alpha}_{-\frac{1}{2}}$, with $\rho_L + \rho_R = 0$, acting on the three vacua gives states that also survive in the spectrum. There are also invariant states of the form $\alpha_{-\frac{1}{2}}\tilde{\alpha}_{-\frac{1}{2}}[|f_1, 0\rangle_2 - |f_2, 0\rangle_2]$.

Conventionally, we drop the fixed point dependence and speak of states labeled by $(N_L, \rho_L; N_R, \rho_R)$, with $N_L = N_R$ determining the mass level, and having a degeneracy factor $\mathcal{F}_k(N_L, \rho_L; N_R, \rho_R)$ that might be zero when the state is not invariant. In the $\mathrm{T}^2/\mathbb{Z}_4$ example above, there are e.g. states $|0\rangle_2$ and $\bar{\alpha}_{-\frac{1}{2}}\tilde{\alpha}_{-\frac{1}{2}}|0\rangle_2$ with $\mathcal{F}_2 = 3$, $\alpha_{-\frac{1}{2}}\tilde{\alpha}_{-\frac{1}{2}}|0\rangle_2$ with $\mathcal{F}_2 = 1$, and so on.

A systematic way to determine the degeneracy factor is to implement the orbifold projection by performing the sum $\frac{1}{N}\sum_{\ell=1}^{N-1} \mathcal{Z}(\theta^k,\theta^\ell)$. Using (4.48) and (4.42) we obtain

$$\mathcal{F}_k(N_L,\rho_L;N_R,\rho_R) = \frac{1}{N}\sum_{\ell=0}^{N-1} \tilde{\chi}(\theta^k,\theta^\ell)\, e^{2i\pi\ell(\rho_L+\rho_R)} \,. \tag{4.50}$$

Here $\tilde{\chi}(\theta^k,\theta^\ell)$ is a numerical factor that counts the fixed point multiplicity. More concretely, $\tilde{\chi}(1,\theta^\ell)=1$, so that in the untwisted sector $\mathcal{F}_0$ projects out precisely the states non-invariant under θ that have $\rho_L+\rho_R$ not integer. In twisted sectors $\tilde{\chi}(\theta^k,\theta^\ell)$ is the number of simultaneous fixed points of θ^k and θ^ℓ in the sub-lattice effectively rotated by θ^k. $\tilde{\chi}(\theta^k,\theta^\ell)$ differs from $\chi(\theta^k,\theta^l)$ because when kv_j = integer, the expansion of $\vartheta[{}^{\frac{1}{2}+kv_j}_{\frac{1}{2}+\ell v_j}]/\eta$ has a prefactor $(-2\sin\pi\ell v_j)$, as follows using (4.42). Thus, the actual coefficient in the expansion of (4.48) is $\tilde{\chi}(\theta^k,\theta^l)=\chi(\theta^k,\theta^l)/\prod_{j,kv_j\in\mathbb{Z}} 4\sin^2\pi\ell v_j$.

4.4 Type II Strings on Toroidal $\mathbb{Z}_N$ Symmetric Orbifolds

The new ingredient is the presence of world-sheet fermions with boundary conditions

$$\begin{aligned} \Psi(\sigma^0,\sigma^1+2\pi) &= -e^{2\pi i\alpha}\theta^k\Psi(\sigma^0,\sigma^1)\,, \\ \Psi(\sigma^0+2\pi\tau_2,\sigma^1+2\pi\tau_1) &= -e^{2\pi i\beta}\theta^\ell\Psi(\sigma^0,\sigma^1)\,, \end{aligned} \tag{4.51}$$

where $\alpha,\beta=0,\frac{1}{2}$ are the spin structures. The full partition function has the form (4.39). Each contribution to the sum is explicitly evaluated as

$$\mathcal{Z}(\theta^k,\theta^\ell) = \mathrm{Tr}_{\,(\mathrm{NS}\oplus\mathrm{R})(\mathrm{NS}\oplus\mathrm{R})}\left\{\mathrm{P_{GSO}}\,\theta^\ell\, q^{L_0(\theta^k)}\,\bar{q}^{\tilde{L}_0(\theta^k)}\right\} \,. \tag{4.52}$$

The trace is over left and right Neveu-Schwarz (NS) and Ramond (R) sectors for the fermions. This is equivalent to summing over $\alpha=0,\frac{1}{2}$. Similarly, the GSO (Gliozzi-Scherk-Olive) projection is equivalent to summing over $\beta=0,\frac{1}{2}$ [4, 5, 6].

To find $\mathcal{Z}(\theta^k,\theta^\ell)$ we again start from the untwisted sector in which the Virasoro operators are known and then use modular invariance. The explicit form of $\mathcal{Z}(\theta^k,\theta^\ell)$ can be found in [70] and will be presented in Appendix C. It follows that the eigenvalues of $L_0(\theta^k)$ are

$$m_L^2(\theta^k) = N_L + \frac{1}{2}\,(r+k\,v)^2 + E_k - \frac{1}{2}\,. \tag{4.53}$$

Most terms in this formula arise as in the purely bosonic case of last section. In particular, E_k is given in (4.46). Notice that N_L and N_R also receive (integer) contributions from the fermionic degrees of freedom. The vector r is an $SO(8)$ weight as explained in Appendix C. The vector v is $(0,v_1,v_2,v_3)$, with the v_i specifying the $\mathbb{Z}_N$ action. When r belongs to the scalar or vector class, r takes the form (n_0, n_1, n_2, n_3), with n_a integer. This is the Neveu-Schwarz sector in which left-movers are space-time bosons. When r belongs to a spinorial class it takes the form $(n_0+\frac{1}{2}, n_1+\frac{1}{2}, n_2+\frac{1}{2}, n_3+\frac{1}{2})$. This is the Ramond sector in which left-movers are space-time fermions. For example, the weights of the fundamental vector and spinor representations are:

$$\begin{aligned} \mathbf{8}_v &= \underline{(\pm 1, 0, 0, 0)} \;\; ; \;\; \mathbf{8}_s = \pm \underline{\left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)} \ , \\ \mathbf{8}_c &= \left\{ \underline{\left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)}, \pm \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) \right\} \ , \end{aligned} \tag{4.54}$$

where underlining means permutations. As explained in Appendix C, the GSO projection turns out to be $\sum r_a =$ odd. Thus, in the untwisted sector, massless states must have $r^2 = 1$ and the possible solutions are $\mathbf{8}_v$ and $\mathbf{8}_s$.

For type II strings the mass formula for right-movers is completely analogous to (4.53):

$$m_R^2\left(\theta^k\right) = N_R + \frac{1}{2}\left(p + k\,v\right)^2 + E_k - \frac{1}{2} \ , \tag{4.55}$$

where p is an $SO(8)$ weight as well. In type IIB the GSO projection is also $\sum p_a =$ odd in both NS and R sectors. In type IIA one has instead $\sum p_a =$ even in the R sector. In the untwisted sector the spinor weights are then those of $\mathbf{8}_c$. Notice that upon combining left and right movers, states in (NS,NS) and (R,R) are space-time bosons, whereas states in (NS,R) and (R,NR) are space-time fermions.

States in a θ^k-twisted sector are characterized by $(N_L, \rho_L, r; N_R, \rho_R, p)$ such that the level-matching condition $m_L^2 = m_R^2$ is satisfied. Here ρ_L and ρ_R are due only to the internal bosonic oscillators as we explained in the previous section. The degeneracy factor of these states follows from the orbifold projection. Using the results in Sect. 4.3 and Appendix C we find

$$\mathcal{F}(N_L, \rho_L, r; N_R, \rho_R, p) = \frac{1}{N} \sum_{\ell=0}^{N-1} \tilde{\chi}\left(\theta^k, \theta^\ell\right)\, \Delta(k, \ell) \ , \tag{4.56}$$

where the phase Δ is

$$\Delta(k, \ell) = \exp\{2\pi\, i[(r + kv)\cdot \ell v - (p + kv)\cdot \ell v + \ell(\rho_L + \rho_R)]\} \ . \tag{4.57}$$

The factor $\tilde{\chi}(\theta^k, \theta^\ell)$ that takes into account the fixed point multiplicity was already introduced in (4.50).

Below we will consider examples of compactifications to six and four dimensions. We will find that, as expected, one obtains results similar to those found in K3 and CY_3 compactifications.

Six Dimensions

We first consider type IIA on $\mathrm{T}^4/\mathbb{Z}_3$. As torus lattice we take the product of two $SU(3)$ root lattices. The $\mathbb{Z}_3$ action has $v = (0, 0, \frac{1}{3}, -\frac{1}{3})$. The resulting theory in six dimensions has (1,1) supersymmetry that has gravity and vector multiplets with structure

$$\begin{aligned} \mathcal{G}_{11}(6) &= \left\{ g_{\mu\nu}, \psi_\mu^{(+)}, \psi_\mu^{(-)}, \psi^{(+)}, \psi^{(-)}, B_{\mu\nu}, V_\mu^a, \phi \right\} \ ; \quad a = 1, \cdots 4 \ , \\ \mathcal{V}_{11}(6) &= \{ A_\mu, \lambda^{(+)}, \lambda^{(-)}, \varphi^a \} \ , \end{aligned} \tag{4.58}$$

where $\lambda^{(\pm)}$ are Weyl spinors and the φ^a real scalars. Below we will see how the orbifold massless states fit into (1,1) supermultiplets.

In the untwisted sector, candidate massless states allowed by the orbifold projection (4.56) must have $r \cdot v = p \cdot v = 0, \pm 1/3$. With $r \cdot v = p \cdot v = 0$ there are

$$\begin{array}{cc} r & p \\ (\underline{\pm 1, 0}, 0, 0) & (\underline{\pm 1, 0}, 0, 0) \\ \pm \left(\underline{-\frac{1}{2}, \frac{1}{2}}, \frac{1}{2}, \frac{1}{2}\right) & \pm \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) \\ & \pm \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right) \end{array} \ . \tag{4.59}$$

The first two entries in r and p, corresponding to the non-compact coordinates, indicate the Lorentz representation under the little group $SO(4) \simeq SU(2) \times SU(2)$. The vector $(\underline{\pm 1, 0})$ of $SO(4)$ is the $(\frac{1}{2}, \frac{1}{2})$ representation of $SU(2) \times SU(2)$, whereas the spinors $(\underline{\frac{1}{2}, -\frac{1}{2}})$ and $\pm(\frac{1}{2}, \frac{1}{2})$ are the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations respectively. In (4.59) we thus have the product

$$\left[\left(\frac{1}{2}, \frac{1}{2}\right) \oplus 2 \left(\frac{1}{2}, 0\right) \right]_{\text{left}} \otimes \left[\left(\frac{1}{2}, \frac{1}{2}\right) \oplus 2 \left(0, \frac{1}{2}\right) \right]_{\text{right}} \ . \tag{4.60}$$

It is simple to check that the product gives rise to the representations that make up the gravity supermultiplet $\mathcal{G}_{11}(6)$ in (4.58).

In the untwisted sector with $r \cdot v = p \cdot v = 1/3$ we find

$$\begin{array}{cc} r & p \\ (0, 0, 1, 0) & (0, 0, 1, 0) \\ (0, 0, 0, -1) & (0, 0, 0, -1) \\ (-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}) & (-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}) \\ (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}) & (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}) \end{array} \ . \tag{4.61}$$

In terms of little group representations we have the product

$$\left[2\,(0, 0) \oplus \left(0, \frac{1}{2}\right) \right]_{\text{left}} \otimes \left[2\,(0, 0) \oplus \left(\frac{1}{2}, 0\right) \right]_{\text{right}} \ . \tag{4.62}$$

In this way we obtain the representations that fill a vector multiplet $\mathcal{V}_{11}(6)$. For $r \cdot v = p \cdot v = -1/3$ one also obtains a vector multiplet. In both cases the group is $U(1)$.

Let us now turn to the θ-twisted sector. Plugging v and $E_1 = 2/9$ we find that $m_R^2 = m_L^2 = 0$ implies the same r, p given in (4.61). Taking into account the fixed point multiplicity gives then 9 vector multiplets. In the θ^2 sector we find the same result.

In conclusion, type IIA compactification on $\mathrm{T}^4/\mathbb{Z}_3$ yields a (1,1) supersymmetric theory in six dimensions with one gravity multiplet and twenty vector multiplets. Other $\mathrm{T}^4/\mathbb{Z}_N$ orbifolds give exactly the same result which is also obtained in type IIA compactification on a smooth K3 manifold.

Compactification of type IIB on $\mathrm{T}^4/\mathbb{Z}_N$ follows in a similar way. We can obtain the results from the type IIA case noting that the different GSO projection for left-moving spinors simply amounts to changing the little group representation. For example, in the untwisted sector instead of (4.60) we have

$$\left[\left(\frac{1}{2},\frac{1}{2}\right) \oplus 2\left(\frac{1}{2},0\right)\right]_{\text{left}} \otimes \left[\left(\frac{1}{2},\frac{1}{2}\right) \oplus 2\left(\frac{1}{2},0\right)\right]_{\text{right}} . \tag{4.63}$$

In the product there are now two gravitini of the same chirality so that the resulting theory in six dimensions has $(2,0)$ supersymmetry with gravity and tensor multiplets having the field content

$$\begin{aligned}
\mathcal{G}_{20}(6) &= \{g_{\mu\nu}, \psi_\mu^{a(+)}, B_{\mu\nu}^{I(+)}\}\ ; \quad a = 1,2\ ; \quad I = 1,\cdots 5\ , \\
\mathcal{T}_{20}(6) &= \{B_{\mu\nu}^{(-)}, \psi^{a(-)}, \varphi^I\}\ ,
\end{aligned} \tag{4.64}$$

where the superscript $(+)$ or $(-)$ on the antisymmetric tensors indicates whether they have self-dual or anti-self-dual field strength. Altogether the product (4.63) gives a gravity multiplet $\mathcal{G}_{20}(6)$ together with a tensor multiplet $\mathcal{T}_{20}(6)$. Other states from the untwisted sector and the twisted sectors give rise to 20 tensor multiplets. In conclusion, compactification of type IIB on $\mathrm{T}^4/\mathbb{Z}_N$ gives $(2,0)$ supergravity with 21 tensor multiplets, exactly what is found in the compactification on K3 [72].

Four Dimensions

The resulting theory has $\mathcal{N} = 2$ supersymmetry. The massless fields must belong to the gravity multiplet or to hypermultiplets and vector multiplets. Schematically, the content of these multiplets is

$$\begin{aligned}
\mathcal{G}_2(4) &= \{g_{\mu\nu}, \psi_\mu^a, V_\mu\}\ ; \quad a,b = 1,2\ , \\
\mathcal{H}_2(4) &= \{\psi^a, \varphi^{ab}\}\ , \\
\mathcal{V}_2(4) &= \{A_\mu, \lambda^a, \varphi^a\}\ .
\end{aligned} \tag{4.65}$$

Note that $\mathcal{G}_2(4)$ contains the so-called graviphoton V_μ. Below we will group the orbifold massless states into these supermultiplets. We study type IIB on $\mathrm{T}^6/\mathbb{Z}_3$. The torus lattice is the product of three $SU(3)$ root lattices. The $\mathbb{Z}_3$ action has $v = (0, \frac{1}{3}, \frac{1}{3}, -\frac{2}{3})$.

Now the massless states are classified by the little group $SO(2)$, i.e. by helicity λ. For a given state, $\lambda = \lambda_r - \lambda_p$ where λ_r can be read from the first component of the $SO(8)$ weight r, and likewise for λ_p. In the untwisted sector, candidate massless states allowed by the orbifold projection must have $r \cdot v = p \cdot v = 0, \pm 1/3$. With $r \cdot v = p \cdot v = 0$ we find

$$\begin{array}{cc} r & p \\ (\pm 1, 0, 0, 0) & (\pm 1, 0, 0, 0) \\ \pm\left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) & \pm\left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) \end{array} . \tag{4.66}$$

Considering all possible combinations in (4.66) we find the helicities

$$\left\{ \pm 2, 2 \times \left(\pm \frac{3}{2} \right), \pm 1 \right\} \oplus \left\{ 2 \times \left(\pm \frac{1}{2} \right), 4 \times (0) \right\} . \tag{4.67}$$

Comparing with the structure of the $\mathcal{N} = 2$ supersymmetric multiplets in four dimensions, cf. (4.65), we observe that (4.67) includes a gravity multiplet $\mathcal{G}_2(4)$ plus a hypermultiplet $\mathcal{H}_2(4)$. The four real scalars in the hypermultiplet are the dilaton, the axion dual to $B_{\mu\nu}$, both arising from (NS,NS) (both r, p vectorial), plus a 0-form and another axion dual to $\tilde{B}_{\mu\nu}$, both arising from (R,R) (both r, p spinorial).

In the untwisted sector with $r \cdot v = p \cdot v = \pm 1/3$ we have

$$\begin{array}{ccc} & r & p \\ r \cdot v = \frac{1}{3} & (0, \underline{1, 0, 0}) & (0, \underline{1, 0, 0}) \\ & \left(-\frac{1}{2}, \underline{\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}}\right) & \left(-\frac{1}{2}, \underline{\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}}\right) \\ r \cdot v = -\frac{1}{3} & (0, \underline{-1, 0, 0}) & (0, \underline{-1, 0, 0}) \\ & \left(\frac{1}{2}, \underline{\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}}\right) & \left(\frac{1}{2}, \underline{\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}}\right) \end{array} . \tag{4.68}$$

Evaluating the helicities of all allowed combinations we find precisely nine hypermultiplets.

Consider now the θ-twisted sector. Plugging v and $E_1 = 1/3$ we find that $m_R^2 = m_L^2 = 0$ has solutions $r, p = (0, 0, 0, 1), (-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$. In the θ^{-1} sector the solutions are $r, p = (0, 0, 0, -1), (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$. According to the orbifold projection we can then combine the following

$$\begin{array}{ccc} & r & p \\ \theta & (0, 0, 0, 1) & (0, , 0, 0, 1) \\ & \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) & \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right) \\ \theta^{-1} & (0, 0, 0, -1) & (0, 0, 0, -1) \\ & \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right) & \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right) \end{array} . \tag{4.69}$$

Altogether we find the degrees of freedom of one hypermultiplet. Taking into account the fixed point multiplicity shows that 27 hypermultiplets originate in the twisted sectors.

In conclusion, compactification of type IIB on $T^6/\mathbb{Z}_3$ has massless content summarized by

$$\mathcal{G}_2(4) + \mathcal{H}_2(4) + 36\mathcal{H}_2(4) . \tag{4.70}$$

This result agrees with the general result for type IIB compactification on a CY_3 manifold. In fact, as we explained in Sect. 4.1, the $T^6/\mathbb{Z}_3$ orbifold has $h^{1,1} = 36$ and $h^{1,2} = 0$.

Compactification of type IIA on $T^6/\mathbb{Z}_3$ is completely analogous. The results are easily obtained changing the left-moving spinor helicities appropriately. In the untwisted sector with $r \cdot v = p \cdot v = 0$ there are no changes. In the untwisted sector with $r \cdot v = p \cdot v = \pm 1/3$, as well as in the twisted sectors, instead of hypermultiplets there appear vector multiplets. Hence, type IIA on $T^6/\mathbb{Z}_3$ has massless multiplets

$$\mathcal{G}_2(4) + \mathcal{H}_2(4) + 36\mathcal{V}_2(4) . \tag{4.71}$$

This agrees with the result for compactification on a CY_3.

5 Recent Developments

We have discussed basic aspects of supersymmetry preserving string compactifications. The main simplifying assumption was that the only background field allowed to have a non-trivial vacuum expectation value (vev) was the metric, for which the Ansatz (2.9) was made, and a constant dilaton ϕ_0 that fixes the string coupling constant as $g_s = e^{\phi_0}$. When we considered the complexification of the Kähler cone we also allowed a vev for the (NS,NS) antisymmetric tensor B_{MN}, but limited to vanishing field strength so that the equations of motion do not change. Restricting to these backgrounds means exploring only a small subspace of the moduli space of supersymmetric string compactifications. Type II string theory has several other massless bosonic excitations, the dilaton ϕ and the (R,R) p-form fields $A^{(p)}$ with p even for type IIB and p odd for type IIA, which could get non-vanishing vevs. The interesting situation is when the vevs for the field strengths $H = dB$ and $F^{(p+1)} = dA^{(p)}$ lead to non-vanishing fluxes through non-trivial homology cycles in the internal manifold. It is clearly important to examine the implications of these fluxes. One interesting result to date is that fluxes can generate a potential for moduli scalars [73]. This provides a mechanism for lifting flat directions in moduli space.

If the additional background fields are non-trivial they will have in general a non-zero energy-momentum tensor T_{MN} that will back-react on the geometry and distort it away from the Ricci-flat Calabi-Yau metric. At the level of the low-energy effective action this means that the lowest order (in α') equation of motion for the metric is no longer the vacuum Einstein equation $R_{MN} = 0$ but rather $R_{MN} = T_{MN}$. We also have to satisfy the equations of motion of the other background fields (setting them to zero is one solution, but we are interested in less trivial ones) and the Bianchi identities of their field strengths. Again, a practical way to proceed is to require unbroken supersymmetry, i.e. to impose that the fermionic fields have vanishing supersymmetric transformations which are now modified by the presence of additional background fields, cf. (2.10). It must then be checked that the Bianchi identities and the equations of motion are satisfied.

The effect of H flux was studied early on [74] and has lately attracted renewed attention. The upshot is that the supersymmetry preserving backgrounds are, in general, not Calabi-Yau manifolds. The analysis of these solutions is a current research subject. For recent papers that give references to the previous literature see e.g. [75].

The presence of (R,R) fluxes leads to an even richer zoo of possible type II string compactifications. One simple and well-studied example is the $AdS_5 \times S^5$ solution of type IIB supergravity which has, in addition to the metric, a non-trivial five-form field strength $F^{(5)}$ background. A general analysis of type IIB compactifications to four dimensions, including backgrounds for all bosonic fields as well as D-brane and orientifold plane sources, was given in [76]. Conditions for $\mathcal{N}=1$ supersymmetry of such configurations were found in [77]. These results have been applied in recent attempts to construct realistic models with moduli stabilization [78].

Compactification of M-theory, or its low-energy effective field theory, eleven-dimensional supergravity, on manifolds of G_2 holonomy, have also been much explored lately. These compactifications lead to $\mathcal{N}=1$ supersymmetry in four dimensions and are interesting in their own right and also in relation with various string dualities, such as compactification of M-theory on a manifold with G_2 holonomy and of the heterotic string on a Calabi-Yau manifold. See [79] for a recent review.

There are many other aspects which one could mention in the context of string compactifications. It is a vast and still growing subject with many applications in physics and mathematics. We hope that our lecture notes will be of use for those who are just entering this interesting and fascinating field.

Appendix A: Conventions and Definitions

A.1: Spinors

The Dirac matrices Γ^A, $A = 0, \cdots, D-1$, satisfy the Clifford algebra

$$\{\Gamma^A, \Gamma^B\} \equiv \Gamma^A \Gamma^B + \Gamma^B \Gamma^A = 2\eta^{AB} \ , \tag{A.1}$$

where $\eta^{AB} = \mathrm{diag}(-1, +1, \ldots, +1)$. The smallest realization of (A.1) is $2^{[D/2]} \times 2^{[D/2]}$-dimensional ($[D/2]$ denotes the integer part of $D/2$). One often uses antisymmetrized products

$$\Gamma^{A_1\cdots A_p} \equiv \Gamma^{[A_1}\cdots\Gamma^{A_p]} \equiv \frac{1}{p!}\left(\Gamma^{A_1}\cdots\Gamma^{A_p} \pm \text{permutations}\right) , \qquad (A.2)$$

with + (−) sign for even (odd) permutations.

The generators of $SO(1, D-1)$ in the spinor representation are

$$T^{AB} \equiv -\frac{i}{2}\Gamma^{AB} \equiv -\frac{i}{4}[\Gamma^A, \Gamma^B] . \qquad (A.3)$$

Spinor representations are necessary to describe space-time fermions. Strictly speaking, when discussing spinors we should go to the covering group, the spin group. We will not make this distinction here but it is always implied.

Exercise A.1: Verify that $T^{AB} \equiv -\frac{i}{2}\Gamma^{AB}$ are generators of $SO(1, D-1)$ in the spinor representation, i.e.

$$i[T^{AB}, T^{CD}] = \eta^{AC}T^{BD} - \eta^{AD}T^{BC} - \eta^{BC}T^{AD} + \eta^{BD}T^{AC} . \qquad (A.4)$$

Dirac spinors have then dimension $2^{[D/2]}$. For D even the Dirac representation is reducible since there exists a matrix that commutes with all generators. This is

$$\Gamma_{D+1} \equiv e^{-i\pi(D-2)/4}\Gamma^0 \ldots \Gamma^{D-1} . \qquad (A.5)$$

For D odd, $\Gamma_{D+1} \propto \mathbb{1}$.

Exercise A.2: Show that $\Gamma^2_{D+1} = \mathbb{1}$, $\{\Gamma_{D+1}, \Gamma^A\} = 0$, and $[\Gamma_{D+1}, \Gamma^{AB}] = 0$.

With the help of Γ_{D+1} we can define the irreducible inequivalent *Weyl representations*: if ψ is a Dirac spinor, the left and right Weyl spinors are

$$\psi_L = \frac{1}{2}(1 - \Gamma_{D+1})\psi , \qquad \psi_R = \frac{1}{2}(1 + \Gamma_{D+1})\psi . \qquad (A.6)$$

Note that $\Gamma_{D+1}\psi_R = \psi_R$ and $\Gamma_{D+1}\psi_L = -\psi_L$.

Dirac and Weyl spinors are complex but in some cases a *Majorana condition* of the form $\psi^* = B\psi$ with B a matrix such that $BB^* = \mathbb{1}$ is consistent with the Lorentz transformations $\delta\psi = i\omega_{MN}T^{MN}\psi$, i.e. B must satisfy $T^{*MN} = -BT^{MN}B^{-1}$. The Majorana condition is allowed for $D = 0, 1, 2, 3, 4$ mod 8. Majorana-Weyl spinors can be shown to exist only in $D = 2$ mod 8 [6].

$SO(D)$ spinors have analogous properties. For D even, there are two inequivalent irreducible Weyl representations of dimension $2^{D/2-1}$. A Majorana-Weyl condition can be imposed only for $D = 0$ mod 8.

A.2: Differential Geometry

We use $A, B, \ldots$ to denote flat tangent indices (raised and lowered with η^{AB} and η_{AB}) which are related to the curved indices $M, N, \ldots$ (raised and lowered with G^{MN} and G_{MN}) via the D-bein: e.g. $\Gamma^A = e^A_M \Gamma^M$ and the inverse D-bein, e.g. $\Gamma^M = e^M_A \Gamma^A$, where $G_{MN} = e^A_M e^B_N \eta_{AB}$ and $e^A_M e^M_B = \delta^A_B$, $e^A_M e^N_A = \delta^N_M$, $\eta^{AB}\eta_{BC} = \delta^A_C$. The Γ^M satisfy $\{\Gamma^M, \Gamma^N\} = 2G^{MN}$.

A Riemannian connection Γ^P_{MN} is defined by imposing

$$\begin{aligned} \nabla_P G_{MN} &\equiv \partial_p G_{MN} - \Gamma^Q_{PM} G_{QN} - \Gamma^Q_{PN} G_{MQ} = 0 \qquad \text{(metricity)} \\ \Gamma^P_{MN} &= \Gamma^P_{NM} \qquad \text{(no torsion)} \ . \end{aligned} \tag{A.7}$$

One finds for the *Christoffel symbols*

$$\Gamma^P_{MN} = \frac{1}{2} G^{PQ} \left(\partial_M G_{QN} + \partial_N G_{MQ} - \partial_Q G_{MN} \right) \ . \tag{A.8}$$

The *Riemann tensor* is

$$[\nabla_M, \nabla_N] V_P = -R_{MNP}{}^Q V_Q \ . \tag{A.9}$$

The *Ricci tensor* and the *Ricci scalar* are $R_{MN} = G^{PQ} R_{MPNQ}$ and $R = G^{MN} R_{MN}$. The *spin connection* is defined via the condition

$$\nabla_M e^A_N = \partial_M e^A_N - \Gamma^P_{MN} e^A_P + \omega_M{}^A{}_B e^B_N = 0 \tag{A.10}$$

which leads to the following explicit expression for its components

$$\omega^{AB}_M = \frac{1}{2} \left(\Omega_{MNR} - \Omega_{NRM} + \Omega_{RMN} \right) e^{NA} e^{RB} \tag{A.11}$$

where

$$\Omega_{MNR} = \left(\partial_M e^A_N - \partial_N e^A_M \right) e_{AR} \ .$$

In terms of ω^{AB}_M the components of the Lie-algebra valued curvature 2-form are

$$R_{MN}{}^{AB} = e^{AP} e^{BQ} R_{MNPQ} = \partial_M \omega^{AB}_N - \partial_N \omega^{AB}_M + \omega^{AC}_M \omega_{NC}{}^B - \omega^{AC}_N \omega_{MC}{}^B \ . \tag{A.12}$$

The covariant derivative, acting on an object with only tangent-space indices, is generically

$$\nabla_M = \partial_M + \frac{i}{2} \omega^{AB}_M T_{AB} \ , \tag{A.13}$$

where T_{AB} is a generator of the tangent space group $SO(1, D{-}1)$. For example, $i(T_{AB})_C{}^D = \eta_{AC}\delta^D_B - \eta_{BC}\delta^D_A$ for vectors and $iT_{AB} = \frac{1}{2}\Gamma_{AB}$ for spinors (spinor indices are suppressed).

Under infinitesimal parallel transport a vector V changes as $\delta V^M = -\Gamma^M_{NR} V^N dx^R$. When V is transported around an infinitesimal loop in the (M, N)-plane with area $\delta a^{MN} = -\delta a^{NM}$ it changes by the amount

$$\delta V^P = -\frac{1}{2} \delta a^{MN} R_{MN}{}^P{}_Q V^Q \, . \tag{A.14}$$

Notice that under parallel transport the length $|V|$ remains constant since $|V|^2 = V^M V^N G_{MN}$ and $\nabla_P G_{MN} = 0$. The generalization to the parallel transport of tensors and spinors is obvious.

Appendix B: First Chern Class of Hypersurfaces of $\mathbb{P}^n$

This Appendix is adopted from [27].

Let $X = \{z \in \mathbb{P}^n; f(z) = 0\}$, f a homogeneous polynomial of degree d, be a non-singular hypersurface in $\mathbb{P}^n$. From (3.70) we know that $c_1(X)$ can be expressed by any choice of volume element on X. As volume element we will use the pull-back of the $(n-1)$-st power of the Kähler form on $\mathbb{P}^n$. We will first compute this in general and will then use the Fubini-Study metric on $\mathbb{P}^n$. It suffices to do the calculation on the subset

$$U_0 \cap \{\partial_n f \neq 0\} \cap X \, . \tag{B.1}$$

Given that $\omega = i g_{i\bar{\jmath}} dz^i \wedge dz^{\bar{\jmath}}$ we compute (z^i are the inhomogeneous coordinates on U_0)

$$\omega^{n-1} = (i g_{i\bar{\jmath}} dz^i \wedge d\bar{z}^j)^{n-1} = i^{n-1} \sum_{i,j=1}^{n} (-1)^{i+j} \det(m_{ij})(\cdots \widehat{i} \cdots \widehat{\bar{\jmath}} \cdots) \, , \tag{B.2}$$

where m_{ij} is the (i,j)-minor of the metric $g_{i\bar{\jmath}}$. The notation $(\cdots \widehat{i} \cdots \widehat{\bar{\jmath}} \cdots)$ means that the hatted factors are missing in the product $dz^1 \wedge d\bar{z}^1 \wedge \cdots \wedge dz^n \wedge d\bar{z}^n$. In the next step we split the above sum according to how many powers of dz^n appear. We get

$$\begin{aligned} i^n \omega^{n-1} &= \sum_{i,j=1}^{n-1} (-1)^{i+j} \det(m_{ij})(\cdots \widehat{i} \cdots \widehat{\bar{\jmath}} \cdots) \\ &\quad + \sum_{i=1}^{n-1} (-1)^{n+i} \det(m_{in})(\cdots \widehat{i} \cdots \widehat{\bar{n}}) + \sum_{j=1}^{n-1} (-1)^{n+j} \det(m_{hj})(\cdots \widehat{\bar{\jmath}} \cdots \widehat{n} \cdot) \\ &\quad + \det(m_{nn})(\cdots \widehat{n}\, \widehat{\bar{n}}) \, . \end{aligned} \tag{B.3}$$

We now replace dz^n via the hypersurface constraint:

$$df = \sum_{i=1}^{n-1} \frac{\partial f}{\partial z^i} dz^i + \frac{\partial f}{\partial z^n} dz^n \quad \Rightarrow \quad dz^n = -\left(\frac{\partial f}{\partial z^n}\right)^{-1} \sum_{i=1}^{n-1} \frac{\partial f}{\partial z^i} dz^i \, . \tag{B.4}$$

Using this in (B.3), we find

$$(i)^n \omega^{n-1} = \left| \frac{\partial f}{\partial z^n} \right|^{-2} \sum_{i,j=1}^{n} \frac{\partial f}{\partial z^i} \overline{\frac{\partial f}{\partial z^j}} (-1)^{i+j} \det(m_{ij})(dz^1 \wedge \cdots \wedge d\bar{z}^{n-1}) \, . \tag{B.5}$$

Next we need the identity

$$g^{i\bar{\jmath}} \equiv (g^{-1})_{i\bar{\jmath}} = (-1)^{i+j} \det(m_{ij})(\det g)^{-1} \, . \tag{B.6}$$

Using this in (B.4), we obtain

$$(-i)^n \omega^{n-1} = (\det g) \left| \frac{\partial f}{\partial z^n} \right|^{-2} \sum_{i,j=1}^{n} g^{i\bar{j}} \frac{\partial f}{\partial z^i} \overline{\frac{\partial f}{\partial z^j}} (dz^1 \wedge \cdots \wedge d\bar{z}^n) \ . \tag{B.7}$$

We now specify to the Fubini-Study metric, for which

$$g^{i\bar{j}} = (1 + |z|^2)(\delta_{ij} + z^i \bar{z}^j) \ . \tag{B.8}$$

It follows that

$$\sum_{i,j=1}^{n} g^{i\bar{j}} \frac{\partial f}{\partial z^i} \overline{\frac{\partial f}{\partial z^j}} = (1 + |z|^2) \left(\sum_{i=1}^{n} \left| \frac{\partial f}{\partial z^i} \right|^2 + \sum_{i,j=1}^{n} z^i \frac{\partial f}{\partial z^i} \overline{z^j \frac{\partial f}{\partial z^j}} \right) \ . \tag{B.9}$$

Now, since f vanishes on X and since it is a homogeneous function of degree d, on X we get

$$0 = d \cdot f = \frac{\partial f}{\partial z^0} + \sum_{i=1}^{n} z^i \frac{\partial f}{\partial z^i} \ , \tag{B.10}$$

and therefore

$$\sum_{i,j=1}^{n} g^{i\bar{j}} \frac{\partial f}{\partial z^i} \overline{\frac{\partial f}{\partial z^j}} = (1 + |z|^2) \left(\sum_{i=0}^{n} \left| \frac{\partial f}{\partial z^i} \right|^2 \right) \ . \tag{B.11}$$

Because the determinant of the metric is (cf. (3.28))

$$\det(g_{i\bar{j}}) = \frac{1}{(1 + |z|^2)^{n+1}} \ , \tag{B.12}$$

we find

$$(i)^n \omega^{n-1} = \left| \frac{\partial f}{\partial z^n} \right|^{-2} \frac{\sum_{i=0}^{n} \left| \frac{\partial f}{\partial z^i} \right|^2}{(|z|^2)^n} \tag{B.13}$$

where now $|z|^2 = \sum_{i=0}^{n} |z^i|^2$. If we set

$$\psi = \log \left(\frac{\sum_{i=0}^{n} |\partial_i f|^2}{|z|^{2d-2}} \right) \ , \tag{B.14}$$

which is a globally defined function, i.e. it has a unique value on all overlaps, we can write

$$\begin{aligned} \partial\bar{\partial} \log \omega^{n-1} &= \partial\bar{\partial} \log \frac{\sum \left| \frac{\partial f}{\partial z_i} \right|^2}{(|z|^2)^n} - \partial\bar{\partial} \log \left| \frac{\partial f}{\partial z_n} \right|^2 \\ &= \partial\bar{\partial} \log e^{\psi} (|z|^2)^{d-n-1} \\ &= \partial\bar{\partial}\psi + i(n - d + 1)\omega \ . \end{aligned} \tag{B.15}$$

Recall that this is valid on the subset specified in (B.1), in particular that this expression is to be evaluated on the hypersurface $f(z) = 0$. Comparing this to (3.70) we realize that we have shown that

$$2\pi c_1(X) = (n + 1 - d)[\omega] \ . \tag{B.16}$$

Appendix C: Partition Function of Type II Strings on $\mathrm{T}^{10-d}/\mathbb{Z}_N$

The starting point is the partition function for the ten-dimensional type II strings that can be written as the product of a bosonic $\mathcal{Z}_B$ and a fermionic $\mathcal{Z}_F$ contribution [4, 5, 6]. Up to normalization:

$$\begin{aligned}
\mathcal{Z}(\tau,\bar\tau) &= \mathcal{Z}_B(\tau,\bar\tau)\,\mathcal{Z}_F(\tau,\bar\tau)\\
\mathcal{Z}_B(\tau,\bar\tau) &= \left(\frac{1}{\sqrt{\tau_2}\,\eta\,\bar\eta}\right)^8 \qquad\qquad (\text{C.1})\\
\mathcal{Z}_F(\tau,\bar\tau) &= \frac{1}{4}\left\{\frac{\vartheta^4\left[{0\atop 0}\right]}{\eta^4}-\frac{\vartheta^4\left[{0\atop \frac12}\right]}{\eta^4}-\frac{\vartheta^4\left[{\frac12\atop 0}\right]}{\eta^4}+\frac{\vartheta^4\left[{\frac12\atop \frac12}\right]}{\eta^4}\right\}\\
&\quad\times\left\{\frac{\bar\vartheta^4\left[{0\atop 0}\right]}{\bar\eta^4}-\frac{\bar\vartheta^4\left[{0\atop \frac12}\right]}{\bar\eta^4}-\frac{\bar\vartheta^4\left[{\frac12\atop 0}\right]}{\bar\eta^4}\pm\frac{\bar\vartheta^4\left[{\frac12\atop \frac12}\right]}{\bar\eta^4}\right\}\ ,
\end{aligned}$$

where $\eta(\tau)$ is the Dedekind function defined in (4.27) and the Jacobi theta functions are

$$\vartheta\left[{\delta\atop\varphi}\right](\tau)=\sum_n q^{\frac12(n+\delta)^2}\,e^{2i\pi(n+\delta)\varphi}\ ;\qquad q=e^{2i\pi\tau}\ . \qquad (\text{C.2})$$

The theta functions also have the product form (4.42) given in Sect. 4.3. In the following we will not write explicitly that ϑ and η are functions of τ.

Depending on the sign in the last term of the right-moving piece of $\mathcal{Z}_F$ we have type IIB (+ sign) or IIA (− sign) strings. In the following we consider type IIB so that $\mathcal{Z}_F(\tau,\bar\tau)=|\mathcal{Z}_F(\tau)|^2$. The left-moving $\mathcal{Z}_F(\tau)$ can be written as:

$$\mathcal{Z}_F(\tau)=\frac12\sum_{\alpha,\beta=0,\frac12} s_{\alpha\beta}\,\frac{\vartheta^4\left[{\alpha\atop\beta}\right]}{\eta^4}\ . \qquad (\text{C.3})$$

The $s_{\alpha\beta}$ are the spin structure coefficients. Modular invariance requires $s_{0\frac12}=s_{\frac12 0}=-s_{00}$. This can be checked using the transformation properties:

$$\begin{aligned}
\mathcal{T}:\tau\to\tau+1\ ;\quad &\eta\to e^{\frac{i\pi}{12}}\eta\ ;\quad \vartheta\left[{\alpha\atop\beta}\right]\to e^{-i\pi(\alpha^2-\alpha)}\vartheta\left[{\alpha\atop \alpha+\beta-\frac12}\right]\ ,\\
\mathcal{S}:\tau\to -1/\tau\ ;\ &\eta\to(-i\tau)^{\frac12}\eta\ ;\ \vartheta\left[{\alpha\atop\beta}\right]\to(-i\tau)^{\frac12}e^{2i\pi\alpha\beta}\vartheta\left[{\beta\atop -\alpha}\right]\ . \qquad (\text{C.4})
\end{aligned}$$

We take $s_{00}=1$ and choose $s_{\frac12\frac12}$ equal to s_{00} so that the GSO projections in the NS and R sectors turn out the same as we explain below.

The NS sector corresponds to $\alpha=0$. Using (C.2) we can write

$$\frac{1}{2}\left\{\frac{\vartheta^4\left[{0 \atop 0}\right]}{\eta^4} - \frac{\vartheta^4\left[{0 \atop \frac{1}{2}}\right]}{\eta^4}\right\} = \frac{1}{\eta^4}\sum_{r_a\in\mathbb{Z}} q^{\frac{1}{2}r^2}\left[\frac{1-e^{i\pi(r_0+r_1+r_2+r_3)}}{2}\right] . \tag{C.5}$$

This shows that left-moving fermionic degrees of freedom of a given NS state depend on a vector r with four integer entries. This is an $SO(8)$ weight in the scalar or vector class. Similarly, for the R sector with $\alpha = \frac{1}{2}$ we have

$$-\frac{1}{2}\left\{\frac{\vartheta^4\left[{\frac{1}{2} \atop 0}\right]}{\eta^4} - \frac{\vartheta^4\left[{\frac{1}{2} \atop \frac{1}{2}}\right]}{\eta^4}\right\} = -\frac{1}{\eta^4}\sum_{r_a\in\mathbb{Z}+\frac{1}{2}} q^{\frac{1}{2}r^2}\left[\frac{1-e^{i\pi(r_0+r_1+r_2+r_3)}}{2}\right] . \tag{C.6}$$

Now r has half-integer entries so that it corresponds to an $SO(8)$ spinor weight. Here we are actually exchanging the light-cone world-sheet fermions by four free bosons that have momentum r in the $SO(8)$ weight lattice. This equivalence between fermions and bosons is in fact seen in (C.5) and (C.6) when we write the left-hand-side using (4.42). Furthermore, because we have included both $\beta = 0$ and $\beta = \frac{1}{2}$, only states with $\sum r_a =$ odd do appear. This is the GSO projection. For instance, the tachyon $r = 0$ in the NS sector is eliminated from the spectrum. In the R sector one of the $SO(8)$ spinor representations with $r^2 = 1$ is also absent. For the right-moving piece we obtain completely analogous results in terms of an $SO(8)$ weight denoted p.

Let us now discuss the partition function for the orbifold that has the form (4.39). Each term $\mathcal{Z}(\theta^k, \theta^\ell)$ can be written as the product of bosonic and fermionic pieces. The bosonic piece is

$$\mathcal{Z}_B(\theta^k, \theta^\ell) = \left(\frac{1}{\sqrt{\tau_2}\,\eta\,\bar\eta}\right)^{d-2} \chi(\theta^k, \theta^\ell)\left|\prod_{j=1}^{5-\frac{d}{2}} \frac{\eta}{\vartheta\left[{\frac{1}{2}+kv_j \atop \frac{1}{2}+\ell v_j}\right]}\right|^2 , \tag{C.7}$$

where $\chi(\theta^k, \theta^\ell)$ is the number of simultaneous fixed points of θ^k and θ^ℓ. The first term is the contribution of the non-compact coordinates ($d-2$ in the light-cone gauge), whereas the second term comes from the $(10-d)$ compact coordinates as we have seen in Sect. (4.3). We are assuming d even.

For the fermionic piece we start with the untwisted sector. The insertion of θ^ℓ in the trace leads to

$$\mathcal{Z}_F(\mathbb{1}, \theta^\ell) = \frac{1}{4}\left|\sum_{\alpha,\beta=0,\frac{1}{2}} s_{\alpha\beta}(0,\ell)\,\frac{\vartheta\left[{\alpha \atop \beta}\right]}{\eta}\prod_{j=1}^{3}\frac{\vartheta\left[{\alpha \atop \beta+\ell v_j}\right]}{\eta}\right|^2 , \tag{C.8}$$

where $s_{\alpha\beta}(0,\ell) = s_{\alpha\beta}(0,0)$ are the spin structures in (C.3). We have specialized to $d = 4$, for other cases simply set $v_j = 0$ for $j > 5 - \frac{d}{2}$. To derive the remaining $\mathcal{Z}_F(\theta^k, \theta^\ell)$ we use modular transformations. In the end we obtain

$$\mathcal{Z}_F(\theta^k, \theta^\ell) = \frac{1}{4}\left|\sum_{\alpha,\beta=0,\frac{1}{2}} s_{\alpha\beta}(k,\ell)\,\frac{\vartheta\left[{\alpha \atop \beta}\right]}{\eta}\prod_{j=1}^{3}\frac{\vartheta\left[{\alpha+kv_j \atop \beta+\ell v_j}\right]}{\eta}\right|^2 . \tag{C.9}$$

Modular invariance imposes relations among the spin structure coefficients. We find:

$$s_{00}(k,\ell) = -s_{\frac{1}{2}0}(k,\ell) = 1 \; ; \quad s_{0\frac{1}{2}}(k,\ell) = -s_{\frac{1}{2}\frac{1}{2}}(k,\ell) = -e^{-i\pi k(v_1+v_2+v_3)} \; . \tag{C.10}$$

Setting $k = N$ then gives a further condition on the twist vector, namely:

$$N(v_1 + v_2 + v_3) = 0 \bmod 2 \; . \tag{C.11}$$

Notice that all twists in Table 2 do satisfy this condition.

Exercise C.1 : Use (C.4) to show that (C.9) has the correct modular transformations, i.e. $\mathcal{Z}_F(\theta^k, \theta^\ell)$ transforms into $\mathcal{Z}_F(\theta^k, \theta^{k+\ell})$ under $\mathcal{T}$ and into $\mathcal{Z}_F(\theta^\ell, \theta^{-k})$ under $\mathcal{S}$.

Acknowledgments

We are grateful to S. Adarve, H. Ocampo and S. Paycha for the invitation to participate at the Villa de Leyva School. A.F. thanks CLAF for travel aid, the Albert-Einstein-Institut for hospitality and support while finishing these notes, and A. Hernández for preparing figures.

References

1. N. Berkovits, *ICTP lectures on covariant quantization of the superstring*, hep-th/0209059. 102
2. P. Candelas, G. Horowitz, A. Strominger and E. Witten, *Vacuum Configurations for Superstrings*, Nucl. Phys. B256 (1985) 46–74. 102, 115, 153
3. L. Dixon, J.A. Harvey, C. Vafa and E. Witten, *Strings on Orbifolds*, Nucl. Phys. B261 (1985) 678–686; *Strings on Orbifolds II*, Nucl. Phys. B274 (1986) 285–314. 102, 149, 150, 153
4. M. B. Green, J. H. Schwarz and E. Witten, *Superstring Theory*, Vols. I and II, Cambridge University Press, 1987. 103, 112, 115, 125, 136, 137, 140, 163, 174
5. D. Lüst and S. Theisen, *Lectures On String Theory*, Lect. Notes Phys. 346, Springer-Verlag, 1989. 103, 163, 174
6. J. Polchinski, *String Theory*, Vols. I and II, Cambridge University Press, 1998. 103, 111, 163, 170, 174
7. `http://www.aei.mpg.de/~theisen`. 103
8. M. Duff, B.E.W. Nilsson and C. Pope, *Kaluza-Klein Supergravity* Phys. Rept. 130 (1986) 1–142. 103, 112
9. J. Wess and J. Bagger, *Supersymmetry and Supergravity*, Princeton University Press, 1992;
   J. D. Lykken, *Introduction to supersymmetry*, hep-th/9612114. 105
10. N. Polonsky, *Supersymmetry: Structure and phenomena. Extensions of the standard model*, Lect. Notes Phys. M68, Springer-Verlag, 2001 (hep-ph/0108236). 106

11. V. A. Rubakov and M. E. Shaposhnikov, *Extra Space-Time Dimensions: Towards a Solution to the Cosmological Constant Problem*, Phys. Lett. B125 (1983) 139–143. 107
12. A.L. Besse, *Einstein Manifolds*, Springer-Verlag, 1987. 109, 110, 115, 125, 132, 133, 135, 136, 140
13. D.D. Joyce, *Compact Manifolds with Special Holonomy*, Oxford University Press, 2000. 109, 110, 121
14. P. Candelas, A. M. Dale, C. A. Lütken and R. Schimmrigk, *Complete Intersection Calabi-Yau Manifolds*, Nucl. Phys. B298 (1988) 493–525;
    P. Candelas, M. Lynker and R. Schimmrigk, *Calabi-Yau Manifolds in Weighted* $\mathbb{P}^4$, Nucl. Phys. B341 (1990) 383–402;
    A. Klemm and R. Schimmrigk, *Landau-Ginzburg string vacua*, Nucl. Phys. B411 (1994) 559–583, hep-th/9204060;
    M. Kreuzer and H. Skarke, *No mirror symmetry in Landau-Ginzburg spectra!*, Nucl. Phys. B388 (1992) 113–130, hep-th/9205004;
    see also http://hep.itp.tuwien.ac.at/ kreuzer/CY/. 110
15. C. P. Burgess, A. Font and F. Quevedo, *Low-Energy Effective Action for the Superstring*, Nucl. Phys. B272 (1986) 661;
    A. Font and F. Quevedo, *N=1 Supersymmetric Truncations and the Superstring Low-Energy Effective Theory*, Phys. Lett. B184 (1987) 45–48. 112
16. R.C. Gunning, *Lectures on Riemann Surfaces*. Mathematical Notes, Princeton Univ. Press, Princeton 1967;
    M. Nakahara, *Geometry, Topology and Physics*, Adam Hilger, 1990. 113
17. A. Giveon, M. Porrati and E. Rabinovici, *Target space duality in string theory*, Phys. Rept. 244 (1994) 77–202, hep-th/9401139. 114
18. K. Hori, S. Katz, A. Klemm, R. Pandharipande, R. Thomas, C. Vafa, R. Vakil, E. Zaslow, *Mirror Symmetry*, Clay Mathematics Monographs, Vol. 1, AMS 2003. 115, 119, 140, 147
19. G. Horowitz, *What is a Calabi-Yau Space?*, in Proceedings of the workshop on Unified String Theories, Santa Barbara 1985, M. Green and D. Gross, eds. 115
20. T. Hübsch, *Calabi-Yau Manifolds*, World-Scientific, 1992. 115
21. S. Hosono, A. Klemm and S. Theisen, *Lectures on Mirror Symmetry*, hep-th/9403096. 115, 140, 147, 148
22. B. Greene, *String theory on Calabi-Yau manifolds*, hep-th/9702155. 115
23. P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, Wiley 1978. 115, 118, 127, 131, 135, 137
24. S. Chern, *Complex Manifolds without Potential Theory*, Springer-Verlag, 1979. 115, 116
25. R.O. Wells, *Differential Analysis on Complex Manifolds*, Springer-Verlag, 1979. 115, 117, 121, 134
26. K. Kodaira, *Complex Manifolds and Deformation of Complex Structures*, Springer-Verlag, 1986. 115, 121, 128, 143
27. G. Tian, *Canonical Metrics in Kähler Goemetry*, Birkhäuser, 2000. 115, 132, 148, 172
28. P. Candelas, *Lectures on Complex Manifolds*, published in *Superstrings' 87*, Proceedings of the 1987 ICTP Spring School, pp. 1–88. 115, 125, 133, 136
29. A. Dimca, *Singularities and Topology of Hypersurfaces*, Springer-Verlag, 1992;
    I. Dolgachev, *Weighted Projective Varieties*, in "Group Actions and Vector Fields, Proceedings 1981", LNM 959, Springer-Verlag, Berlin, 1982, pp 34–71. 118
30. M. Reid, *Young Persons Guide to Canonical Singularities*, Proc. Sym. Pure Math 46, AMS, 1987, 345–414; *Canonical 3-Folds*, in "Journées de géometrique algébrique d'Angers", (A. Beauville, ed.) 1990, 273–310. 119

31. W. Fulton, *Introduction to Toric Varieties*, Princeton University Press, 1993. 119
32. D.A. Cox and S. Katz, *Mirror Symmetry and Algebraic Geometry*, Mathematical Surveys and Monographs, Vol. 68, AMS, 1999. 119, 147
33. P. Candelas and X. C. de la Ossa, *Comments on Conifolds*, Nucl. Phys. B342 (1990) 246–268. 133, 143
34. P.S. Howe, G. Papadopoulos and K.S. Stelle, *Quantizing The N=2 Super Sigma Model In Two-Dimensions*, Phys. Lett. B176 (1986) 405–410. 134
35. A. Lichnerowicz, *Global theory of connections and holonomy groups*, Noordhoff International Publishing, 1976. 136
36. P. Griffiths, *On the periods of certain rational integrals I,II*, Ass. of Math. 90 (1969) 460–495, 498–541. 137
37. D. Morrison, *Picard-Fuchs equations and mirror maps for hypersurfaces* in Mirror Symmetry I (s.-T. Yau, ed.), AMS and International Press, 1998,p. 185–199, alg-geom/9202026. 137
38. P. Candelas, *Yukawa couplings between (2,1)-forms*, Nucl. Phys. B298 (1988) 458–492. 137, 140, 143, 148
39. P. Mayr, *Mirror Symmetry,* $N = 1$ *Superpotentials and Tensionless Strings on Calabi-Yau Fourfolds*, Nucl. Phys. B494 (1997) 489–545, hep-th/9610162;
A. Klemm, B. Lian, S.S. Roan and S.T. Yau, *Calabi-Yau Fourfolds for M-Theory and F-Theory Compactifications*, Nucl. Phys. B518 (1998) 515–574, hep-th/9701023. 139
40. V. Batyrev, *Variations of the Mixed Hodge Structure of Affine Hypersurfaces in Algebraic Tori*, Duke Math. J. 69 (1993) 349–409; *Dual Polyhedra and Mirror Symmetry for Calabi-Yau Hypersurfaces in Toric Varieties*, J. Algebraic Geometry 3 (1994) 493–535, alg-geom/9310003; *Quantum Cohomology Rings of Toric Manifolds*, Astérisque 218 (1993) 9–34, alg-geom/9310004. 140, 143
41. P. Candelas, E. Derrick and L. Parkes, *Generalized Calabi-Yau manifolds and the mirror of a rigid manifold*, Nucl. Phys. B407 (1993) 115–154, hep-th/9304045. 140
42. P. Candelas and X. de la Ossa, *Moduli Space of Calabi-Yau Manifolds*, Nucl. Phys. B355 (1991) 455–481. 140, 148
43. S. Ferrara and S. Theisen, *Moduli Spaces, Effective Actions and Duality Symmetry in String Compactification*, Proceedings of the 3rd Hellenic School on Elementary Particle Physics, Corfu 1989 (E.N. Argyres, N. Tracas and G. Zoupanos, eds.), p. 620–656. 140
44. P. S. Aspinwall, *K3 surfaces and string duality*, hep-th/9611137. 140
45. N. Barth and S. Christensen, *Quantizing Fourth Order Gravity Theories. 1. The Functional Integral*, Phys. Rev. D28 (1983) 1876–1893;
G. 't Hooft and M. Veltman, *One Loop Divergencies In The Theory Of Gravitation*, Annales Poincaré Phys. Theor. A20 (1974) 69–94. 140
46. G. Tian, *Smoothness of the Universal Deformation Space of Compact Calabi-Yau Manifolds and its Peterson-Weil Metric*, in S.T. Yau (ed), Mathematical Aspects of String Theory, World Scientific, 1987, p. 629–647. 142
47. A. Todorov. *The Weyl-Petersson Geometry of the Moduli-Space of* $SU(n \geq 3)$ *(Calabi-Yau) Manifolds I*, Commun. Math. Phys. 126 (1989) 325–346. 142
48. P. Candelas, P. S. Green and T. Hübsch, *Rolling Among Calabi-Yau Vacua*, Nucl. Phys. B330 (1990) 49–102. 143, 147

49. M. Dine, P. Huet and N. Seiberg, *Large and Small Radius in String Theory*, Nucl. Phys. B322 (1989) 301–316;
J. Dai, R.G. Leigh and J. Polchinski, *New Connections between String Theories*, Mod. Phys. Lett. A4 (1989). 144
50. N. Seiberg, *Observations on the Moduli Space of Superconformal Field Theories*, Nucl. Phys. B303 (1988) 286–304. 144, 145, 147
51. M. Bodner, A. C. Cadavid and S. Ferrara, *(2,2) Vacuum Configurations for Type IIA Superstrings: N=2 Supergravity Lagrangians and Algebraic Geometry*, Class. Quantum Grav. 8 (1991) 789–808. 144
52. D. Morrison, *Mirror Symmetry and the Type II String*, Nucl.Phys.Proc.Suppl. 46 (1996) 146–155, hep-th/9512016. 145
53. B. de Wit and A. Van Proeyen, *Potentials and Symmetries of General Gauged N=2 Supergravity - Yang-Mills Models*, Nucl. Phys. B245 (1984) 89–117;
B. de Wit, P. G. Lauwers and A. Van Proeyen, *Lagrangians Of N=2 Supergravity – Matter Systems*, Nucl. Phys. B255 (1985) 569–608;
E. Cremmer, C. Kounnas, A. Van Proeyen, J. P. Derendinger, S. Ferrara, B. de Wit and L. Girardello, *Vector Multiplets Coupled To N=2 Supergravity: Superhiggs Effect, Flat Potentials and Geometric Structure*, Nucl. Phys. B250 (1985) 385–426. 146
54. J. Bagger and E. Witten, *Matter Couplings In N=2 Supergravity*, Nucl. Phys. B222 (1983) 1–10. 146
55. S. Cecotti, S. Ferrara and L. Girardello, *Geometry of Type II Superstrings and the Moduli of Superconformal Field Theories*, Int. J. Mod. Phys. A4 (1989) 2475–2529. 146, 147
56. 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 (1991) 21–74. 146, 147
57. P. Candelas, T. Hübsch and R. Schimmrigk, *Relation Between the Weil-Petersson and Zamolodchikov Metrics*, Nucl. Phys. B329 (1990) 583–590. 147
58. L. J. Dixon, V. Kaplunovsky and J. Louis, *On Effective Field Theories Describing (2,2) Vacua Of The Heterotic String*, Nucl. Phys. B 329 (1990) 27–82. 147
59. I. Antoniadis, E. Gava, K. S. Narain and T. R. Taylor, *Superstring threshold corrections to Yukawa couplings*, Nucl. Phys. B407 (1993) 706–724, hep-th/9212045;
J. P. Derendinger, S. Ferrara, C. Kounnas and F. Zwirner, *On loop corrections to string effective field theories: Field dependent gauge couplings and sigma model anomalies*, Nucl. Phys. B372 (1992) 145–188;
G. Lopes Cardoso and B. A. Ovrut, *Coordinate and Kähler Sigma Model Anomalies and their Cancellation in String Effective Field Theories*, Nucl. Phys. B392 (1993) 315–344, hep-th/9205009. 147
60. R. Bryant and P. Griffiths, *Some Observations on the Infinitesimal Period Relations for Regular Threefolds with Trivial Canonical Bundle*, in Progress in Mathemtaics 36, (M. Artin and J. Tate, eds.), Birkhäuser 1983. 148
61. A. Strominger, *Special Geometry*, Commun. Math. Phys. 133 (1990) 163–180. 148
62. B. Craps, F. Roose, W. Troost and A. Van Proeyen, *The definitions of special geometry*, hep-th/9606073 and *What is special Kaehler geometry?*, Nucl. Phys. B 503 (1997) 565–613, hep-th/9703082. 148

63. P. Mayr, *On supersymmetry breaking in string theory and its realization in brane worlds*, Nucl. Phys. B593 (2001) 99–126, hep-th/0003198. 148
64. 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 (1997) 571–588, hep-th/9707013. 148
65. I. Antoniadis, R. Minasian, S. Theisen and P. Vanhove, *String loop corrections to the universal hypermultiplet*, Class. Quant. Grav. 20 (2003) 5079–5102, hep-th/0307268. 148
66. L. J. Dixon, *Some World Sheet Properties of Superstring Compactifications, on Orbifolds and Otherwise*, published in *Superstrings, Unified Theories and Cosmology 1987*, Proceedings of the 1987 ICTP Summer Workshop, pp. 67–126. 149
67. J. Erler and A. Klemm, *Comment on the Generation Number in Orbifold Compactifications*, Commun. Math. Phys. **153** (1993) 579–604, hep-th/9207111. 151, 162
68. D. G. Markushevich, M. A. Olshanetsky and A. M. Perelomov, *Description of a Class of Superstring Compactifications Related to Semisimple Lie Algebras*, Commun. Math. Phys. 111 (1987) 247–274. 151, 153
69. D.N. Page, *A Physical Picture of the K3 Gravitational Instanton*, Phys. Lett. B80 (1978) 55–57. 152
70. L.E. Ibáñez, J. Mas, H.P. Nilles and F. Quevedo, *Heterotic Strings in Symmetric and Asymmetric Orbifold Backgrounds*, Nucl. Phys. B301 (1988) 157–196; A. Font, L.E. Ibáñez, F. Quevedo and A. Sierra, *The Construction of "Realistic" Four-Dimensional Strings through Orbifolds* , Nucl. Phys. B331 (1990) 421–474. 163
71. S.-S. Roan, *On the Generalization of Kummer Surfaces*, J. Diff. Geom. 30 (1989) 523–537; G. Tian and S.-T. Yau, *Complete Kähler Manifolds with Zero Ricci Curvature I*, Invent. math. 106 (1991) 27–60. 153
72. P. Townsend, *A New Anomaly-Free Chiral Supergravity Theory from Compactification on K3*, Phys. Lett. B139 (1984) 283–287. 166
73. S. Gukov, C. Vafa and E. Witten, *CFT's from Calabi-Yau four-folds*, Nucl. Phys. B584 (2000) 69–108, hep-th/9906070. 168
74. A. Strominger, *Superstrings With Torsion*, Nucl. Phys. B274 (1986) 253–284. 169
75. S. Gurrieri, J. Louis, A. Micu and D. Waldram, *Mirror symmetry in generalized Calabi-Yau compactifications*, Nucl. Phys. B654 (2003) 61–113, hep-th/0211102.;
G. L. Cardoso, G. Curio, G. Dall'Agata, D. Lüst, P. Manousselis and G. Zoupanos, *Non-Kaehler string backgrounds and their five torsion classes*, Nucl. Phys. B652 (2003) 5–34, hep-th/0211118.;
K. Becker, M. Becker, K. Dasgupta and P. S. Green, *Compactifications of heterotic theory on non-Kaehler complex manifolds. I*, JHEP **0304** (2003) 007, hep-th/0301161;
J. P. Gauntlett, D. Martelli and D. Waldram, *Superstrings with intrinsic torsion*, hep-th/0302158;
K. Becker, M. Becker, P. S. Green, K. Dasgupta and E. Sharpe, *Compactifications of heterotic strings on non-Kaehler complex manifolds. II*, hep-th/0310058;
S. Fidanza, R. Minasian and A. Tomasiello, *Mirror symmetric SU(3)-structure manifolds with NS fluxes*, hep-th/0311122. 169
76. S. B. Giddings, S. Kachru and J. Polchinski, *Hierarchies from fluxes in string compactifications*, Phys. Rev. D66 (2002) 106006(16), hep-th/0105097. 169

77. M. Grana and J. Polchinski, *Supersymmetric three-form flux perturbations on* AdS^5, Phys. Rev. D63 (2001) 026001(8), hep-th/0009211. 169
78. S. Kachru, M. B. Schulz and S. Trivedi, *Moduli stabilization from fluxes in a simple IIB orientifold*, JHEP 0310 (2003) 007, hep-th/0201028;
R. Blumenhagen, D. Lüst and T. R. Taylor, *Moduli stabilization in chiral type IIB orientifold models with fluxes*, Nucl. Phys. B663 (2003) 319–342, hep-th/0303016;
J. F. G. Cascales and A. M. Uranga, *Chiral 4d* $N = 1$ *string vacua with D-branes and NSNS and RR fluxes*, JHEP 0305 (2003) 011, hep-th/0303024. 169
79. B. S. Acharya, *M Theory,* G_2*-Manifolds and Four-Dimensional Physics*, Class. Quant. Grav. 19 (2002) 5619–5653. 169

Index Theorems and Noncommutative Topology

T. Fack

Université Claude Bernard (Lyon 1), UFR de Mathématiques, Institut Girard Desargues (UMR 5028), bâtiment Doyen Jean Braconnier, 43 boulevard du 11 novembre 1918 69622 Villeurbanne Cedex, France
Thierry.Fack@univ-lyon1.fr

Summary. These lecture notes are mainly devoted to a K-theory proof of the Atiyah-Singer index theorem. Some applications of the K-theory to noncommutative topology are also given.

Introduction

Topological K-theory for locally compact spaces was introduced by Atiyah and Singer in their proof of the index theorem for elliptic operators. During the last two decades, topological K-theory and elliptic operators have become important tools in topology. For instance, index theory for C^*-algebras was used to compute the K-theory of many "noncommutative" spaces, leading to the so called *Baum-Connes conjecture*. Also, G. Kasparov used K-theory and K-homology for C^*-algebras to investigate the *Novikov's conjecture on higher signatures* for large classes of groups. With the emergence of A. Connes non-commutative geometry, one can say that K-theory and elliptic theory for complex algebras have become usual tools in topology.

The purpose of this course is to introduce the main ideas of the Atiyah-Singer index theorem for elliptic operators. We also show how K-theory for C^*-algebras can be used to study the leaf space of a foliation. We would like to thank the referee for many suggestions making these notes more transparent.

1 Index of a Fredholm Operator

1.1 Fredholm Operators

Recall that an endomorphism $T \in B(H)$ of a Hilbert space H is called *compact* if it is a norm limit of finite rank endomorphisms of H or, equivalently, if the image $T(B)$ of the unit ball B of H is compact. Let H_1, H_2 be two Hilbert spaces and $T : H_1 \longrightarrow H_2$ be a linear continuous map. We say that T is *Fredholm* if there exists a continuous linear map $S : H_2 \longrightarrow H_1$ such that the

T. Fack: *Index Theorems and Noncommutative Topology*, Lect. Notes Phys. **667**, 183–229 (2005)
www.springerlink.com

operators $ST - I \in B(H_1)$ and $TS - I \in B(H_2)$ are compact. For instance, if the operator $K \in B(H)$ is compact, $T = I + K \in B(H)$ is Fredholm.

Exercise 1.1.1. *Let* $K : [0,1] \times [0,1] \longrightarrow \mathbb{C}$ *be a continuous map. Show that the bounded operator* T *defined on* $H = L^2([0,1], dx)$ *by* $Tf(x) = \int_0^1 K(x,y)f(y)dy$ *is compact. Is* T *a Fredholm operator?*

Exercise 1.1.2. *Let* $(e_0, e_1, \ldots)$ *be the natural orthonormal basis of* $H = l^2(\mathbb{N})$*. Show that the unilateral shift* $S \in B(H)$ *defined by* $Se_n = e_{n+1}$ *is Fredholm.*

Let us give a characterization of Fredholm operators involving only the kernel and the image of the operator.

Theorem 1.1.3. *Let* $T : H_1 \longrightarrow H_2$ *be a bounded operator. The following conditions are equivalent:*

(i) *T is Fredholm;*
(ii) *$Ker(T)$ is a finite dimensional subspace of H_1 and $Im(T)$ is a closed finite codimensional subspace of H_2.*

Proof. $(i) \Longrightarrow (ii)$. If T is Fredholm, the restriction of the identity map of H_1 to $Ker(T)$ is compact since it is equal to the restriction of $I - ST$ to $Ker(T)$. It follows that the unit ball of $Ker(T)$ is compact and hence $Ker(T)$ is finite dimensional. On the other hand, since T^* is Fredholm, the subspace $Im(T)^\perp = Ker(T^*)$ is also finite dimensional so that we only have to prove that $Im(T)$ is closed. Let $y_n \in Im(T)$ be a sequence converging to $y \in H_2$, and write $y_n = Tx_n$ with $x_n \in Ker(T)^\perp$. The sequence (x_n) is bounded because if not we could choose a subsequence $(x_{n_k})_k$ with $\|x_{n_k}\| \xrightarrow[k\to+\infty]{} +\infty$. By compactnees of $ST - I$, we may assume in addition that

$$(ST - I)\left(\frac{x_{n_k}}{\|x_{n_k}\|}\right) \xrightarrow[k\to+\infty]{} z \in H_1 \,.$$

Since $ST\left(\frac{x_{n_k}}{\|x_{n_k}\|}\right) = \frac{S(y_{n_k})}{\|x_{n_k}\|} \xrightarrow[k\to+\infty]{} 0$, we would get $\frac{x_{n_k}}{\|x_{n_k}\|} \longrightarrow -z$, a fact which implies $z \in Ker(T)^\perp$ and $\|z\| = 1$. On the other hand, we have

$$\frac{y_{n_k}}{\|x_{n_k}\|} = T\left(\frac{x_{n_k}}{\|x_{n_k}\|}\right) \longrightarrow -T(z)$$

so that $T(z) = 0$. We thus would have $z \in Ker(T) \cap Ker(T)^\perp = \{0\}$, a fact which contradicts $\|z\| = 1$. The sequence (x_n) is thus bounded. By compactness of $ST - I$, we may choose a subsequence $(x_{n_k})_k$ such that

$$(ST - I)(x_{n_k}) \xrightarrow[k\to+\infty]{} z \in H_1 \ .$$

It follows that $x_{n_k} \xrightarrow[k\to+\infty]{} Sy - z \in H_1$ and hence $y = Tu$ with $u = Sy - z$. This shows that $Im(T)$ is closed.
$(ii) \Longrightarrow (i)$ By the Hahn-Banach theorem, $T_{|Ker(T)^\perp} : Ker(T)^\perp \longrightarrow Im(T)$ is an isomorphism. Let $S_1 : Im(T) \longrightarrow Ker(T)^\perp$ be the inverse of $T_{|Ker(T)^\perp}$, and consider the operator S which coincides with S_1 on $Im(T)$, and which is 0 on $Ker(T^*)$. We have $ST - I = -p_{Ker(T)}$, $\ TS - I = -p_{Ker(T^*)}$, where p_K denotes the orthogonal projection on the closed subspace K, so that $ST - I$ and $TS - I$ are finite rank operators. QED

In the sequel, we shall denote by $Fred(H_1, H_2)$ the set of Fredholm operators from H_1 to H_2. It is easy to see that $Fred(H_1, H_2)$ is an open subset of $B(H_1, H_2)$ equipped with the norm topology.

1.2 Toeplitz Operators

Toeplitz operators are good examples of "pseudodifferential" operators on $\mathbb{S}^1$. Let $H = H^2(\mathbb{S}^1)$ be the *Hardy* space, i.e. the subspace of $L^2(\mathbb{S}^1)$ generated by the exponentials $e_n(t) = e^{i2\pi nt}(n = 0, 1, \ldots)$, and denote by P the orthogonal projection onto $H^2(\mathbb{S}^1)$.

Definition 1.2.1. *Let $\varphi : \mathbb{S}^1 \longrightarrow \mathbb{C}$ be a continuous map. We call Toeplitz operator of symbol φ the bounded operator $T_\varphi \in B(H)$ defined by:*

$$T_\varphi(f) = P(\varphi f), \ f \in H^2(\mathbb{S}^1) \ .$$

For instance, T_{e_1} is the unilateral shift S and $T_{e_{-1}}$ its adjoint S^*.

Proposition 1.2.2. *Let $\varphi : \mathbb{S}^1 \longrightarrow \mathbb{C}$ be a non vanishing continuous map. Then T_φ is a Fredholm operator.*

Proof. It suffices to prove that $T_\varphi T_\psi - T_{\varphi\psi}$ is compact for any $\varphi, \psi \in C(\mathbb{S}^1)$. Indeed, for a non vanishing $\varphi : \mathbb{S}^1 \longrightarrow \mathbb{C}$, the inverse $\psi = \frac{1}{\varphi}$ is continuous and the above assertion will imply that $T_\varphi T_\psi - I$ and $T_\psi T_\varphi - I$ are compact. To show that $T_\varphi T_\psi - T_{\varphi\psi}$ is compact for any $\varphi, \psi \in C(\mathbb{S}^1)$ we may assume, by using the Stone-Weierstrass theorem and the continuity of the map

$$\varphi \in C(\mathbb{S}^1) \longrightarrow T_\varphi \in B(H) \ ,$$

that φ, ψ are trigonometric polynomials. By linearity, we are finally reduced to prove that $T_{e_n} T_{e_m} - T_{e_{n+m}}$ is compact for any $n, m \in \mathbb{Z}$. But we have:

$$(T_{e_n} T_{e_m} - T_{e_{n+m}})(e_k) = \begin{cases} -e_{n+m+k} & \text{if } -(n+m) \le k < -m \\ 0 & \text{if not} \end{cases}$$

so that $T_{e_n}T_{e_m} - T_{e_{n+m}}$ is a finite rank operator, and hence is compact. QED

Exercise 1.2.3. *Let $\varphi : \mathbb{S}^1 \longrightarrow \mathbb{C}$ be a continuous map such that T_φ is compact. Show that $\varphi = 0$.*

1.3 The Index of a Fredholm Operator

Definition 1.3.1. *Let $T : H_1 \longrightarrow H_2$ be a Fredholm operator. The integer:*

$$Ind(T) = dimKer(T) - codimKer(T) = dimKer(T) - dimKer(T^*) \in \mathbb{Z}$$

is called the index of T.

The main property of the index is its homotopy invariance:

Proposition 1.3.2. *For any norm continuous path $t \in [0,1] \longrightarrow T_t \in Fred(H_1, H_2)$ of Fredholm operators, we have $Ind(T_0) = Ind(T_1)$.*

The homotopy invariance of the index is a consequence of the continuity of the map $T \longrightarrow Ind(T)$ on $Fred(H_1, H_2)$. For a proof, see [15], theorem 2.3, page 224. Let us give a consequence of this homotopy invariance:

Corollary 1.3.3. *Let $T_1, T_2 \in B(H)$ be two Fredholm operators. Then, T_1T_2 is a Fredholm operator such that $Ind(T_1T_2) = Ind(T_1) + Ind(T_2)$.*

Proof. It is clear from the definition that T_1T_2 is Fredholm. For any $t \in [0, \pi/2]$, set $F_t = \begin{pmatrix} T_1 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} cost & sint \\ -sint & cost \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & T_2 \end{pmatrix} \begin{pmatrix} cost & -sint \\ sint & cost \end{pmatrix}$. We thus define a homotopy of Fredholm operators between $F_0 = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}$ and $F_{\pi/2} = \begin{pmatrix} T_1T_2 & 0 \\ 0 & I \end{pmatrix}$. By proposition 1.3.2., we get:

$$Ind(T_1) + Ind(T_2) = Ind(F_0) = Ind(F_{\pi/2}) = Ind(T_1T_2). \text{ QED}$$

Exercise 1.3.4. *Let $F, K \in B(H_1, H_2)$. Assume that F is Fredholm and K compact. Show that $F + K$ is Fredholm and $Ind(F + K) = Ind(F)$.*

1.3.5 Meaning of the Index. Let $T \in Fred(H_1, H_2)$. By the Hahn-Banach theorem, $T_{|Ker(T)^\perp} : Ker(T)^\perp \longrightarrow Im(T)$ is an isomorphism. If $Ind(T) = 0$, we can choose an isomorphism $R : Ker(T) \longrightarrow Im(T)^\perp$ and the operator $\widetilde{T}$ equal to T on $Ker(T)^\perp$ and to R on $Ker(T)$ is a finite rank perturbation of T which is an isomorphism. Conversely, if there exists a finite rank operator R such that $\widetilde{T} = T + R$ is an isomorphism, then $Ind(T) = Ind(\widetilde{T}) = 0$ (cf.

exercise 1.3.4.). This shows that $Ind(T)$ is the obstruction to make T an isomorphism by a finite rank perturbation. Let us give another interpretation of the index of a Fredholm operator. To avoid unnecessary technicalities, we shall assume that $H_1 = H_2 = H$. Denote by $Calk(H) = B(H)/K(H)$ the quotient of the algebra $B(H)$ by the closed ideal $K(H)$ of compact operators on H. The Calkin algebra $Calk(H)$ is a Banach*-algebra[1] with unit for the quotient norm. Denote by $\pi : B(H) \longrightarrow B(H)/K(H) = Calk(H)$ the canonical projection. By definition, an operator $T \in B(H)$ is Fredholm if and only if $\pi(T)$ is invertible in $Calk(H)$. The index of T is, by the preceding discussion, the obstruction to lift $\pi(T)$ to some invertible element in $B(H)$. The following proposition shows that we can always choose an invertible lift $X \in M_2(B(H))$ for the 2×2 matrix $\begin{pmatrix} \pi(T) & 0 \\ 0 & \pi(T)^{-1} \end{pmatrix} \in M_2(Calk(H))$ with coefficients in $Calk(H)$, and that $Ind(T)$ can be interpreted as the formal difference of the projections:

$$X \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X^{-1} \in M_2(I + K(H)) \text{ and } \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in M_2(I + K(H)) \, .$$

Proposition 1.3.6. *Let $T \in B(H)$ be a Fredholm operator. Denote by S the operator which is zero on $Ker(T^*)$ and which is equal on $Im(T)$ to the inverse of the isomorphism $T_{|Ker(T)^\perp} : Ker(T)^\perp \longrightarrow Im(T)$. Let e and f be the orthogonal projections on $Ker(T)$ and $Ker(T^*)$.*

(i) $X = \begin{pmatrix} T & f \\ e & S \end{pmatrix}$ *is an invertible element in* $M_2(B(H)) = B(H) \otimes M_2(\mathbb{C})$ *such that* $(\pi \otimes I_2)(X) = \begin{pmatrix} \pi(T) & 0 \\ 0 & \pi(T)^{-1} \end{pmatrix}$*;*

(ii) $X \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X^{-1} - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -f & 0 \\ 0 & e \end{pmatrix}$, *so that*

$$Ind(T) = Trace\left(X \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X^{-1} - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right).$$

Proof. (i) By direct calculation, we see that $\begin{pmatrix} S & e \\ -f & T \end{pmatrix} \in M_2(B(H))$ is an inverse for X. Since $ST = 1 - e$, we have $\pi(S)\pi(T) = 1$ and hence $\pi(S) = \pi(T)^{-1}$. It follows that $(\pi \otimes I_2)(X) = \begin{pmatrix} \pi(T) & 0 \\ 0 & \pi(T)^{-1} \end{pmatrix}$ and (i) is proved.
(ii) We get by direct computation:

[1] Recall that a Banach*-algebra B is a Banach algebra with an involution $x \longmapsto x^*$ such that $\|x^*\| = \|x\|$ for any $x \in B$. When B has a unit 1, we always ask that $\|1\| = 1$.

$$X\begin{pmatrix}1&0\\0&0\end{pmatrix}X^{-1}-\begin{pmatrix}1&0\\0&0\end{pmatrix}=\begin{pmatrix}1-f&0\\0&e\end{pmatrix}-\begin{pmatrix}1&0\\0&0\end{pmatrix}=\begin{pmatrix}-f&0\\0&e\end{pmatrix}.\ \text{QED}$$

Exercise 1.3.7. *Let $T \in B(H_1, H_2)$ and assume that there exists $S \in B(H_2, H_1)$ and a positive integer n such that $(ST - I)^n$ and $(TS - I)^n$ are trace class operators.*
Show that T is Fredholm and prove that:

$$Ind(T) = Trace((ST - I)^n) - Trace((TS - I)^n)\ .$$

Let us now compute the index of a Toeplitz operator. To this end, recall that the degree of a continuous map $\varphi : \mathbb{S}^1 \longrightarrow \mathbb{C}$ which does not vanish is by definition the degree $\frac{1}{2\pi i}\int_0^1 \frac{\psi'(t)}{\psi(t)}dt$ of any smooth function $\psi : \mathbb{S}^1 \longrightarrow \mathbb{C}$ sufficiently close to φ.

Theorem 1.3.8. *For any non vanishing continuous map $\varphi : \mathbb{S}^1 \longrightarrow \mathbb{C}$, we have:*

$$Ind(T_\varphi) = -deg(\varphi)\ .$$

Proof. Set $n = deg(\varphi)$. Since φ is homotopic to the map $t \in \mathbb{S}^1 \longrightarrow \frac{\varphi(t)}{|\varphi(t)|} \in \mathbb{S}^1$ whose degree is n, there exists by Hopf's theorem a continuous homotopy $(\varphi_t)_{0\le t\le 1}$ between $\varphi_0 = \varphi$ and $\varphi_1(s) = e^{i2\pi ns}$ such that $\varphi_t : \mathbb{S}^1 \longrightarrow \mathbb{C}$ is a continuous invertible map for each $t \in [0, 1]$. Since the index of a Fredholm operator and the degree of a continuous invertible map are homotopy invariants, we get:

$$Ind(T_\varphi) = Ind(T_{\varphi_1}) = Ind(T_{e_n}) = n.Ind(T_{e_1}) = -n = -deg(\varphi)\ ,$$

and the proof is complete. QED

Exercise 1.3.9. *Let T be the operator on $l^2(\mathbb{Z})$ defined by:*

$$T(e_n) = \begin{cases} \frac{n}{\sqrt{1+n^2}}e_{n-1} & if\ \ n \ge 0 \\ \frac{n}{\sqrt{1+n^2}}e_n & if\ \ n \le 0 \end{cases}$$

where $(e_n)_{n\in\mathbb{Z}}$ is the canonical orthonormal basis of $l^2(\mathbb{Z})$. Show that T is a Fredholm operator of index equal to 1.

2 Elliptic Operators on Manifolds

Elliptic operators on manifolds give rise to Fredholm operators, whose analytical index can be computed from the *principal symbol*, which is a purely topological data.

2.1 Pseudodifferential Operators on $\mathbb{R}^n$

Pseudodifferential operators of order m on $\mathbb{R}^n$ generalize differential operators. They are constructed from *symbols* of order m. In what follows, we shall write as usually: $D_x^\alpha = \frac{\partial^{|\alpha|}}{i^{|\alpha|}\partial x^\alpha}$.

Definition 2.1.1. *Let $m \in \mathbb{R}$. A smooth matrix-valued function $p = p(x, \xi)$ on $\mathbb{R}^n \times \mathbb{R}^n$ is said to be a symbol of order m if there is, for any pair (α, β) of multiindices, a constant $C_{\alpha,\beta} \geq 0$ such that:*

$$|D_x^\alpha D_\xi^\beta p(x,\xi)| \leq C_{\alpha,\beta}(1 + |\xi|)^{m-|\beta|} \; for \; all \; x, \xi \; .$$

We shall denote by S^m the space of symbols of order m. Note that $S^m \subset S^{m'}$ if $m \leq m'$. We shall say that a symbol p of order m has a formal development $p \sim \sum_{j=1}^{\infty} p_j$ with $p_j \in S^{m_j}$ if there exists, for each positive integer m, an integer N such that $p - \sum_{j=1}^{k} p_j \in S^{-m}$ for any $k \geq N$. The following result (see for instance [16], proposition 3.4, page 179) is very close to Borel's result on the existence of a smooth function having a given Taylor expansion at some point:

Proposition 2.1.2. *For any formal series $\sum_{j=1}^{\infty} p_j$ with $p_j \in S^{m_j}$ and $m_j \longrightarrow -\infty$, there exists a symbol p of order m such that we have $p \sim \sum_{j=1}^{\infty} p_j$.*

To any $p \in S^m$ with values in $M_k(\mathbb{C})$, we associate a linear operator:

$$P = Op(p) : S(\mathbb{R}^n) \otimes \mathbb{C}^k \longrightarrow S(\mathbb{R}^n) \otimes \mathbb{C}^k$$

by the formula:

$$Pu(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{i<x,\xi>} p(x,\xi)\widehat{u}(\xi) d\xi \; .$$

Here, $S(\mathbb{R}^n)$ is the Schwartz space of $\mathbb{R}^n$ and $\widehat{u}$ denotes the Fourier transform of $u \in S(\mathbb{R}^n) \otimes \mathbb{C}^k$. The fact that P defines a linear operator from the Schwartz space $S(\mathbb{R}^n) \otimes \mathbb{C}^k$ to itself is straightforward.

Definition 2.1.3. *The operators of the form $Op(p)$ with $p \in S^m$ are called pseudodifferential operators of order m on $\mathbb{R}^n$.*

The space of all pseudodifferential operators of order m will be denoted by Ψ^m. Differential operators on $\mathbb{R}^n$ are examples of pseudodifferential operators. Indeed, consider the symbol $p(x,\xi) = \sum_{|\alpha|\leq m} A^\alpha(x)\xi^\alpha$ of order m, where m is a positive integer and the $A^\alpha(x)$ are smooth matrix valued functions on $\mathbb{R}^n$. Since $D_x^\alpha u(\xi)\widehat{\ } = \xi^\alpha \widehat{u}(\xi)$, we have:

$$Op(p) = \sum_{|\alpha|\leq m} A^\alpha(x) D^\alpha \ .$$

For $s \in \mathbb{R}$, denote by $H^s(\mathbb{R}^n)$ the Sobolev space of exponent s in $\mathbb{R}^n$, i.e. the completion of the Schwartz space $S(\mathbb{R}^n)$ for the Sobolev s-norm:

$$\|u\|_s = \sqrt{\int (1+|\xi|)^{2s} |\widehat{u}(\xi)|^2 d\xi} \ .$$

Theorem 2.1.4. *For any $p \in S^m$ with values in $M_k(\mathbb{C})$ and compact x-support, the operator $P = Op(p)$ has a continuous extension:*

$$P : H^{s+m}(\mathbb{R}^n) \otimes \mathbb{C}^k \longrightarrow H^s(\mathbb{R}^n) \otimes \mathbb{C}^k \ .$$

For a proof of this theorem, see [16], proposition 3.2, page 178. Note that a pseudodifferential operator can have an order $-m < 0$. Such an operator is said to be *smoothing of order m.*

Definition 2.1.5. *A linear map $P : S(\mathbb{R}^n) \otimes \mathbb{C}^k \longrightarrow S(\mathbb{R}^n) \otimes \mathbb{C}^k$ that extends to a bounded linear map $P : H^{s+m}(\mathbb{R}^n) \otimes \mathbb{C}^k \longrightarrow H^s(\mathbb{R}^n) \otimes \mathbb{C}^k$ for all s and m is called infinitely smoothing.*

The space of all infinitely smoothing operators will be denoted by $\Psi^{-\infty}$. Since we have a continuous inclusion $H^s(\mathbb{R}^n) \subset C^q(\mathbb{R}^n)$ for any $s > (n/2) + q$ (Sobolev's embedding theorem), the image Pu of any $u \in H^s(\mathbb{R}^n) \otimes \mathbb{C}^k$ by an infinitely smoothing operator P is a smooth function. Two pseudodifferential operator P and P' will be called *equivalent* if $P - P' \in \Psi^{-\infty}$.

2.1.6 Kernel of a Pseudodifferential Operator. Any $P = Op(p) \in \Psi^m$ has a *Schwartz* (distribution) *kernel* $K_P(x,y)$ satisfying:

$$Pu(x) = < K_P(x,.), u(.) >$$

for any smooth compactly supported function u. Note that K_P is not a function on $\mathbb{R}^n \times \mathbb{R}^n$ in general. Its restriction to the complement of the diagonal of $\mathbb{R}^n \times \mathbb{R}^n$ is given by a smooth function, but it may have singularities on the diagonal (see for instance the case of differential operators). Formally, we have from the formula defining $P = Op(p)$:

$$K_P(x,y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{i<x-y,\xi>} p(x,\xi) d\xi \ ,$$

but we have to point out that this integral does not converge in general. For $m < -n$, this integral makes sense and defines a continuous function K_P on $\mathbb{R}^n \times \mathbb{R}^n$ which is smooth outside the diagonal. In this case, P is an ordinary integral operator. For $m \geq -n$, we get from the above remark, by writing $P = P(1+\Delta)^{-l}(1+\Delta)^l$ where l is a positive integer such that $m - 2l < -n$:

$$Pu(x) = \sum_{|\alpha| \leq 2l} \int K_{\alpha,P}(x,y) D^\alpha u(y) dy$$

where the $K_{\alpha,P}$ are continuous functions that are smooth outside the diagonal. We thus have, in the distributional sense:

$$K_P(x,y) = \sum_{|\alpha| \leq 2l} (-1)^{|\alpha|} D_y^\alpha K_{\alpha,P}(x,y) \ .$$

Definition 2.1.7. *Let $P = Op(p) \in \Psi^m$ be a pseudodifferential operator on $\mathbb{R}^n$.*

(i) We call P ε-local if we have:

$$supp(Pu) \subset \{x \in \mathbb{R}^n | dist(x, supp(u)) \leq \varepsilon\}$$

for any smooth compactly supported function u on $\mathbb{R}^n$:

(ii) We say that P has support in a compact set K if we have:

$$supp(Pu) \subset K \ \text{and}\ (supp(u) \cap K = \varnothing) \Longrightarrow Pu = 0$$

for any smooth compactly supported function u on $\mathbb{R}^n$.

The following proposition summarizes classical results used to construct pseudodifferential operators.

Theorem 2.1.8. *(i) For any formal series $\sum_{j=1}^{\infty} p_j$ with $p_j \in S^{m_j}$ and $m_j \longrightarrow -\infty$, there exists $P = Op(p) \in \Psi^{m_1}$, unique up to equivalence, such that $p \sim \sum_{j=1}^{\infty} p_j$;*

(ii) Let $a(x,y,\xi)$ be a smooth matrix-valued function on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$

with compact x and y-support. Assume that there exists $m \in \mathbb{R}$ and, for each α, β, γ, a constant $C_{\alpha\beta\gamma}$ such that:

$$|D_x^\alpha D_y^\beta D_\xi^\gamma a(x,y,\xi)| \leq C_{\alpha\beta\gamma}(1+|\xi|)^{m-|\gamma|} .$$

Then, the formula:

$$(Pu)(x) = \frac{1}{(2\pi)^n} \int\int e^{i<x-y,\xi>} a(x,y,\xi) u(y) dy d\xi$$

defines a pseudodifferential operator $P = Op(p)$ of order m whose symbol p has the asymptotic development $p \sim \sum_\alpha \frac{i^{|\alpha|}}{\alpha!}(D_\xi^\alpha D_y^\alpha a)(x,x,\xi)$;
(iii) For any $P = Op(p) \in \Psi^m$ whose symbol p has compact x-support and any $\varepsilon > 0$, there exists an ε-local pseudodifferential operator $P_\varepsilon = Op(p) \in \Psi^m$ such that $P - P_\varepsilon \in \Psi^{-\infty}$.

For a proof, we refer to [16], chap. III, § 3.

Exercise 2.1.9. *Let a be as in theorem 2.1.8 (ii), and assume in addition that $a(x,y,\xi)$ vanishes for all (x,y) in a neighbourhood of the diagonal. Show that $P = Op(p)$ is infinitely smoothing.*

Exercise 2.1.10. *Let $P = Op(p) \in \Psi^m$. Show that for any pair (φ, ψ) of smooth real valued functions with compact support, the operator $Q(u) = \psi P(\varphi u)$ is also pseudodifferential of order m. Deduce that if U is an open subset of $\mathbb{R}^n$, we have for any $u \in H^s(\mathbb{R}^n)$:*

$$u_{|U} \in C^\infty \Longrightarrow Pu_{|U} \in C^\infty .$$

The following theorem summarizes the main rules of symbolic calculus on pseudodifferential operators.

Theorem 2.1.11. *(i) Let $P = Op(p) \in \Psi^l$ and $Q = Op(q) \in \Psi^m$ be pseudodifferential operators. Then, the product $R = PQ$ is a pseudodifferential operator $R = Op(r) \in \Psi^{l+m}$ whose symbol r has the formal development:*

$$r \sim \sum_\alpha \frac{i^{|\alpha|}}{\alpha!} \left(D_\xi^\alpha p\right) \left(D_x^\alpha q\right) ;$$

(ii) Let $P = Op(p) \in \Psi^m$ be a pseudodifferential operator. Then, the formal adjoint P^* is a pseudodifferential operator $P^* = Op(p^*) \in \Psi^m$ whose symbol p^* has the formal development:

$$p^* \sim \sum_\alpha \frac{i^{|\alpha|}}{\alpha!} D_\xi^\alpha D_x^\alpha \overline{p}^t ;$$

(iii) Let $P = Op(p) \in \Psi^m$ be a pseudodifferential operator. Let $\varphi : U \longrightarrow V$ be a C^∞-diffeomorphism from an open subset U of $\mathbb{R}^n$ onto an open subset $V = \varphi(U) \subset \mathbb{R}^n$. For any pair (α, β) of smooth functions with compact support in U such that $\beta = 1$ in a neighbourhood of $supp(\alpha)$, the operator:

$$Q(u)(x) = (\alpha P \beta)(u \circ \varphi)(\varphi^{-1}(x))$$

is a pseudodifferential operator $Q = Op(q) \in \Psi^m$ whose symbol q has the formal development:

$$q(\varphi(x), \xi) \sim \alpha(x) \sum_\alpha \frac{i^{|\alpha|}}{\alpha!} D_y^\alpha e^{i(\varphi(y)-\varphi(x)-\frac{\partial\varphi}{\partial x}(y-x)).\xi}|_{y=x} D_\xi^\alpha p \left(x, \left(\frac{\partial \varphi}{\partial x}(x)\right)^t \xi \right),$$

where $\left(\frac{\partial\varphi}{\partial x}(x)\right)$ denotes the Jacobi matrix.

For a proof, see [8], chap. I, § 1.3. Assume in (iii) that P has compact support $K \subset U$, and denote by $\varphi_* P$ the compactly supported pseudodifferential operator defined by $(\varphi_* P)(u) = P(u \circ \varphi) \circ \varphi^{-1}$. By (iii), we get:

$$q\left(\varphi(x), \xi) \sim p(\varphi^{-1}(x), \left(\frac{\partial\varphi}{\partial x}(x)\right)^t \xi\right) \quad (mod\ S^{m-1}),$$

where q is the symbol of $\varphi_* P$. This shows that, modulo symbols of lower order, the symbol of a pseudodifferential operator transforms by change of variable like a function on the cotangent bundle. This observation leads to the following definition:

Definition 2.1.12. *Let $P = Op(p) \in \Psi^m$. The principal symbol $\sigma(P)$ of P is by definition the ressidue class of p in S^m/S^{m-1}.*

If $Op(p) = \sum_{|\alpha| \le m} A^\alpha(x) D^\alpha$, a representative of $\sigma(P)$ is given by the symbol:

$$\sigma_P(x, \xi) = \sum_{|\alpha|=m} A^\alpha(x) \xi^\alpha .$$

For instance, the principal symbol of the Cauchy-Riemann operator $\frac{\partial}{\partial \overline{Z}} = \frac{\partial}{\partial x_1} + i \frac{\partial}{\partial x_2}$ in $\mathbb{R}^2$ is: $\partial_{\frac{\partial}{\partial \overline{z}}}(x, \xi) = i\xi_1 - \xi_2$.

2.2 Pseudodifferential Operators on Manifolds

Let M be a n-dimensional smooth compact Riemannian manifold without boundary. Denote by $\pi : T^*M \longrightarrow M$ the canonical projection. Let E, F be smooth complex vector bundles over M.

Definition 2.2.1. *A linear operator $P : C^\infty(M,E) \longrightarrow C^\infty(M,F)$ is called pseudodifferential of order $m \in \mathbb{R}$ if, for every open chart U on M trivializing E and F and any $\varphi, \psi \in C_c^\infty(U)$, the localized operator $\varphi P \psi$ is pseudodifferential of order m (with compact x-support) on the chart U, viewed as an open subset of $\mathbb{R}^n$.*

As above, we identify here φ with the multiplication operator by φ. We shall denote by $\Psi^m(M; E, F)$ the space of all pseudodifferential operators of order m acting from the sections of E to the sections of F. Let $S^m(T^*M)$ be the set of all $p \in C^\infty(T^*M)$ such that the pullback to any local chart of M is in S^m, and define analogously $S^m(T^*M, Hom(\pi^*E, \pi^*F))$. By theorem 2.1.11 (iii), any $P \in \Psi^m(M; E, F)$ has a *principal symbol:*

$$\sigma(P) \in S^m(T^*M, Hom(\pi^*E, \pi^*F))/S^{m-1}(T^*M, Hom(\pi^*E, \pi^*F)) .$$

Let us denote by *dvol* the Riemannian measure on M and by $H^s(M,E)$ the Sobolev space of exponent $s \in \mathbb{R}$ for the sections of the vector bundle E over M. This space is equipped with the norm $\|u\|_s = \sum_{i=1}^{q} \|\varphi_i u\|_s$, where $(\varphi_1, \dots, \varphi_q)$ is a smooth partition of unity subordinate to a covering of M by charts trivializing E. From theorem 2.1.4, we get:

Theorem 2.2.2. *Any $P \in \Psi^m(M; E, F)$ extends, for any $s \in \mathbb{R}$, to a continuous map $P : H^{s+m}(M,E) \longrightarrow H^s(M,F)$.*

Exercise 2.2.3. *Show that any $P \in \Psi^m(M; E, E)$ with $m \le 0$ defines a bounded operator in $L^2(M,E)$.*

Definition 2.2.4. *A linear operator $P : C^\infty(M,E) \longrightarrow C^\infty(M,F)$ is called infinitely smoothing if it extends to a bounded map $P : H^{s+m}(M,E) \longrightarrow H^s(M,F)$ for any $s \in \mathbb{R}$.*

We shall denote by $\Psi^{-\infty}(M; E, F)$ the space of all infinitely smoothing operators $P : C^\infty(M,E) \longrightarrow C^\infty(M,F)$. It is straightforward to check that $\Psi^{-\infty}(M; E, F) = \bigcap_m \Psi^m(M; E, F)$.

Exercise 2.2.5. *Let $P : C^\infty(M,E) \longrightarrow C^\infty(M,F)$ be a linear operator. Show that $P \in \Psi^{-\infty}(M; E, F)$ if and only if P can be written as an integral operator:*

$$Pu(x) = \int_M K(x,y) u(y) dvol(y) ,$$

where $K(x,y) \in Hom_{\mathbb{C}}(E_y, F_x)$ varies smoothly with x and y on M.

Let us now collect some classical results about pseudodifferential operators on manifolds.

Theorem 2.2.6. *Let M, E, F be as above.*

(i) *Any $P \in \Psi^m(M; E, F)$ can be written as finite sum:*

$$P = \sum_{i=1}^{q} P_i + R \, ,$$

where R is an infinitely smoothing operator on M and each P_i is an order m pseudodifferential operator compactly supported by a local chart U_i trivializing E and F (more precisely, $P_i = \varphi_i P_i \psi_i$ where φ_i, ψ_i are smooth functions on M with compact support on U_i);

(ii) *For any $P \in \Psi^m(M; E, F)$, any open subset U of M and any $u \in H^s(M, E)$ such that $u_{|U}$ is C^∞, the section Pu of F is C^∞ over U;*

(iii) *Let $P \in \Psi^m(M; E, E)$ be a pseudodifferential operator of order $m \leq 0$, viewed as a bounded operator in $L^2(M, E)$ (cf. exercise 2.2.5).*
If $m < 0$, P is a compact operator;
If $m < -n/2$, P is Hilbert-Schmidt operator of the form:

$$Pu(x) = \int_M K_P(x, y) u(y) dvol(y) \, ,$$

where $K_P(x, y) \in Hom(E_y, F_x)$ varies continuously with x and y on M;
If $m < -n/p$, P belongs to the Schatten class C_p [2]*;*
In particular, if $m < -n$, P is a trace-class operator and $Tr(P) = \int_M tr(K_P(x, x)) dvol(x)$.

(iv) *If $P \in \Psi^m(M; E, F)$ and $Q \in \Psi^r(M; F, G)$, then:*

$$QP = Q \circ P \in \Psi^{m+r}(M; E, G) \, ;$$

(v) *Let $P \in \Psi^m(M; E, F)$ and assume that P has a formal adjoint i.e. there exists an operator $P^* : C^\infty(M, F) \longrightarrow C^\infty(M, E)$ such that:*

$$< Pu, v >_{L^2(M,F)} = < u, P^* v >_{L^2(M,E)}$$

for any smooth sections u and v. Then, $P^* \in \Psi^m(M; F^*, E^*)$.

Let us end up this section by giving a specific example of a differential operator on a manifold.

[2] Recall that C_p is the space of compact operators T acting on a separable Hilbert space H such that $Tr(| T |^p) < +\infty$. For more information on the Schatten class C_p, we refer to [19].

2.2.7 The Signature Operator. Let M be a compact oriented Riemannian manifold without boundary, of dimension $n = 4k$. Denote by d the exterior derivative $d : C^\infty(M, \Lambda^*(T^*M \otimes \mathbb{C})) \longrightarrow C^\infty(M, \Lambda^*(T^*M \otimes \mathbb{C}))$. The metric g on M induces a scalar product on $\Lambda^p(T_x^*M)$ by the formula:

$$< a_I dx^I | b_J dx^J >= p! g^{i_1 j_1} \dots g^{i_p j_p} a_I b_J \ ,$$

where $I = (i_1, \dots, i_p)$, $J = (j_1, \dots, j_p)$ and $g^{ij} =< dx^i | dx^j >$. Denote by δ the formal adjoint of d with respect to this scalar product. To describe δ, let us introduce the *Hodge-star* operation. Let $vol = \sqrt{g} dx^1 \wedge \dots \wedge dx^n$ be the volume form. The *Hodge star* operator $* : \Lambda^p(T_x^*M) \longrightarrow \Lambda^{n-p}(T_x^*M)$ is by definition the only linear map satisfying:

$$< \alpha | \beta > vol = \alpha \wedge *(\beta) \text{ for any } \alpha, \beta \in \Lambda^p(T_x^*M) \ .$$

It is easy to check that $\delta = - * d*$, so that δ is an order one differential operator on M. Set:

$$D = d + \delta \ ;$$

we thus define a self-adjoint order one diffferential operator. The signature operator is obtained by restriction of D to the positive part of some grading on $C^\infty(M, \Lambda^*(T^*M \otimes \mathbb{C}))$ that we shall now describe. To this end, note that we have $*(*\alpha) = (-1)^p id$, so that the operator:

$$\varepsilon = i^{2k+p(p-1)} * : \Lambda^p(T^*M \otimes \mathbb{C}) \longrightarrow \Lambda^{4k-p}(T^*M \otimes \mathbb{C})$$

defines a grading on $\Lambda^*(T^*M \otimes \mathbb{C})$, i.e. $\varepsilon^2 = 1$. The ± 1-eigenspaces $E_\pm = \Lambda^\pm(T^*M \otimes \mathbb{C})$ of ε give rise to a direct sum decomposition:

$$\Lambda^*(T^*M \otimes \mathbb{C}) = E_+ \oplus E_- \ ,$$

and since $D = d + \delta$ anticommutes with the grading ε, it decomposes to give rise to operators $D_\pm : C^\infty(M, \bigwedge^+(T^*M \otimes \mathbb{C})) \longrightarrow C^\infty(M, \Lambda^-(T^*M \otimes \mathbb{C}))$. The *signature operator* is by definition the operator

$$D_+ : C^\infty(M, E_+) \longrightarrow C^\infty(M, E_-) \ .$$

It is an order one differential operator on M with principal symbol:

$$\sigma(D_+)(x, \xi) = i(ext(\xi) - int(\xi)) \ ,$$

where $ext(\xi)$ is the exterior multiplication by ξ and $int(\xi)$ its adjoint. Note that $int(\xi)$ is just interior multiplication by ξ since we have for any local orthonormal basis $(e_1, \dots, e_n)$ for T^*M:

$$int(e_1)(e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_p}) = \begin{cases} e_{i_2} \wedge \dots \wedge e_{i_p} & if\ i_1 = 1 \\ 0 & if\ i_1 > 1 \ . \end{cases}$$

2.3 Analytical Index of an Elliptic Operator

Let M be a n-dimensional smooth compact manifold without boundary and denote by π the projection $T^*M \longrightarrow M$. Let $P : C^\infty(M, E) \longrightarrow C^\infty(M, F)$ be a pseudodifferential operator of order m, where E, F are smooth complex vector bundles over M. Recall that P has a principal symbol $\sigma_P \in S^m/S^{m-1}$, where $S^k = S^k(T^*M, Hom(\pi^*E, \pi^*F))$.

Definition 2.3.1. *We say that $P \in \Psi^m(M; E, F)$ is elliptic if its principal symbol σ_P has a representative $p \in S^m(T^*M, Hom(\pi^*E, \pi^*F))$ which is pointwise invertible outside a compact set in T^*M and satisfies the estimate*

$$|p(x,\xi)^{-1}| \leq C(1 + |\xi|)^{-m}$$

for some constant C and some riemannian metric on M.

Example 2.3.2. The signature operator D_+ on a $4k$-dimensional compact oriented manifold M without boundary is elliptic, since we have:

$$\sigma(D_+)(x,\xi)^2 = -(ext(\xi) - int(\xi))^2 = \|\xi\|^2 id \; .$$

The following result shows that an elliptic pseudodifferential operator on a compact manifold M is invertible modulo infinitely smoothing operators:

Theorem 2.3.3. *Let $P \in \Psi^m(M; E, F)$ be an order m elliptic pseudodifferential operator on M. Then, there exists $Q \in \Psi^{-m}(M; F, E)$ such that:*

$$QP - I \in \Psi^{-\infty}(M, E, E) \text{ and } PQ - I \in \Psi^{-\infty}(M, F, F) \; .$$

The operator Q is called a *parametrix* for P.

Sketch of Proof. We shall essentially prove that P admits a parametrix locally.

Case of an elliptic operator on $\mathbb{R}^n$. Let P be a pseudodifferential operator of order m on $\mathbb{R}^n$ whose principal symbol p is pointwise invertible outside a compact set in $T^*\mathbb{R}^n$ and satisfies the estimate $|p(x,\xi)^{-1}| \leq C(1 + |\xi|)^{-m}$ for some constant C. We are going to construct a parametrix Q for P from a formal development $q \sim \sum_{k=0}^{\infty} q_k$ of its symbol, where $q_k \in S^{-m-k}$. By adding some infinitely smoothing operator to P, we may assume that P is 1-local and ask that Q is 1-local too. Since the formal development of the symbol of QP is given by:

$$\sum_\alpha \frac{i^{|\alpha|}}{\alpha!}(D_\xi^\alpha q)(D_x^\alpha p) = \sum_k \sum_\alpha \frac{i^{|\alpha|}}{\alpha!}(D_\xi^\alpha q_k)(D_x^\alpha p) \; ,$$

where $(D_\xi^\alpha q_k)(D_x^\alpha p)$ is a symbol of order $-k-|\alpha|$, it is natural to determine $q_k \in S^{-m-k}$ in such a way that:

$$\begin{cases} q_0 p - I \in S^{-\infty} \\ q_k p + \sum_{j=0}^{k-1} \left[\sum_{|\alpha|+j=k} \frac{i^{|\alpha|}}{\alpha!} (D_\xi^\alpha q_j)(D_x^\alpha p) \right] \in S^{-\infty} \; for \; k = 1, 2, \dots \end{cases}$$

Let us now solve these equations. Note that we only have to choose q_0, because we can determine inductively $q_1, q_2, \dots$ from q_0 by setting:

$$(*) \qquad q_k = -\sum_{j=0}^{k-1} \left[\sum_{|\alpha|+j=k} \frac{i^{|\alpha|}}{\alpha!} \left(D_\xi^\alpha q_j\right) \left(D_x^\alpha p\right) \right] q_0 .$$

To get q_0, we set $q_0(x,\xi) = \theta(|\xi|)p(x,\xi)^{-1}$, where $\theta : \mathbb{R}^+ \longrightarrow [0,1]$ is a smooth function such that $\theta(t) = 0$ for $t \leq C$ and $\theta(t) = 1$ for $t \geq 2C$. It is easy to check that $q_0 \in S^{-m}$. Consider now the asymptotic series $\sum_{k=0}^{\infty} q_k$ where the q_k's are given by (*). There exists a 1-local pseudodifferential Q of order $-m$ with symbol q satisfying $q \sim \sum_{k=0}^{\infty} q_k$. From the formula for the symbol of a product, we get that $QP - I$ is infinitely smoothing. In the same way, we get a pseudodifferential operator Q' such that $PQ' - I$ is infinitely smoothing. But we have, modulo infinitely smoothing operators:

$$Q \equiv Q(PQ') = (QP)Q' \equiv Q' \ ,$$

so that Q is a parametrix for P.

General Case. To avoid technical difficulties, we shall only consider the case of a *differential* operator $P \in \Psi^m(M; E, F)$. Choose a covering of M by open charts U_α trivializing E and F, with local coordinates

$$x_\alpha : U_\alpha \longrightarrow x_\alpha(U_\alpha) = \mathbb{R}^n$$

such that the open subsets $\Omega_\alpha = \{m \in U_\alpha | \; |x_\alpha(m)| < 1\}$ cover M. Let $(\varphi_\alpha)_\alpha$ be a partition of unity subordinate to the Ω_α. Then the restriction P_α of P to U_α, viewed as differential operator on $\mathbb{R}^n$, has a parametrix Q_α that we may assume to be 1-local. Since Q_α is 1-local, the operators $\varphi_\alpha Q_\alpha$ and $Q_\alpha \varphi_\alpha$ have compact support in $\Omega'_\alpha = \{m \in U_\alpha | \; |x_\alpha(m)| < 2\} \subset U_\alpha$, and hence make sense as pseudodifferential operators in $\Psi^{-m}(M; E, F)$. Set:

$$Q = \sum_\alpha \varphi_\alpha Q_\alpha \in \Psi^{-m}(M; E, F) \; and \; Q' = \sum_\alpha Q_\alpha \varphi_\alpha \in \Psi^{-m}(M; E, F) \ .$$

We have:

$$PQ' - I = \sum_{\alpha}(PQ_\alpha\varphi_\alpha - \varphi_\alpha) = \sum_{\alpha}(P_\alpha Q_\alpha - I)\varphi_\alpha = \sum_{\alpha} R_\alpha\varphi_\alpha \ ,$$

where $R_\alpha = P_\alpha Q_\alpha - I$ is a 1-local infinitely smoothing operators in U_α, so that $R_\alpha\varphi_\alpha \in \Psi^{-\infty}(M; E, F)$. It follows that $PQ' - I$ is infinitely smoothing. In the same way, $QP - I$ is infinitely smoothing, and since we have

$$Q - Q' \in \Psi^{-\infty}(M; E, F)$$

as in the first step, the proof is complete. QED

Corollary 2.3.4. *Let $P \in \Psi^m(M; E, F)$ be an order m elliptic pseudodifferential operator on a compact manifold M. For any $s \in \mathbb{R}$, the operator P extends to a Fredholm operator $P_s : H^{s+m}(M, E) \longrightarrow H^s(M, F)$ whose index $Ind(P_s)$ is independent of s.*

Proof. By theorem 2.3.3, there exists $Q \in \Psi^{-m}(M; F, E)$ such that $PQ - I$ and $QP - I$ are infinitely smoothing. Denote by $Q_s : H^s(M, F) \longrightarrow H^{s+m}(M, E)$ the unique extension of Q to $H^s(M, F)$. Since an infinitely smoothing operator from $H^r(M, E)$ into itself is compact, we get that $Q_s P_s - I$ and $P_s Q_s - I$ are compact, and hence P_s is Fredholm. Since $Q_s P_s - I$ is infinitely smoothing, we have $u = -(Q_s P_s - I)u \in C^\infty(M, E)$ for any $u \in Ker(P_s)$, and hence:

$$Ker(P_s) = Ker(P) \subset C^\infty(M, E) \ .$$

In the same way, we get $Ker(P_s^*) = Ker(P^*) \subset C^\infty(M, F)$, where P^* is the formal adjoint of P, and hence $Ind(P_s)$ is independent of s. QED

Definition 2.3.5. *Let $P \in \Psi^m(M; E, F)$ be an elliptic operator on a compact manifold M. The index $Ind(P_s)$ of any extension $P_s : H^{s+m}(M, E) \longrightarrow H^s(M, F)$ is called the analytical index of P and is denoted by $Ind(P)$.* Let us now compute the analytical index of the signature operator on a compact riemannian manifold M without boundary. We assume here that M is $4k$-dimensional and oriented. Recall that the *signature* $\sigma(M)$ of M is by definition the signature of the symmetric bilinear form

$$\begin{aligned} H^{2k}(M, \mathbb{C}) \times H^{2k}(M, \mathbb{C}) &\longrightarrow \mathbb{C} \\ ([\omega_1], [\omega_2]) &\longrightarrow \int_M \omega_1 \wedge \omega_2 \end{aligned}$$

induced by the cup-product in cohomology.

Theorem 2.3.6. *The index of the signature operator on M is equal to the signature $\sigma(M)$ of M.*

Proof. Since $D = d + \delta$ anticommutes with the grading ε, we have:

$$D = \begin{pmatrix} 0 & D_+^* \\ D_+ & 0 \end{pmatrix}$$

where D_+ is the signature operator on M. But $D^2 = (d+\delta)^2 = d\delta + \delta d$ is nothing but the Hodge-Laplace operator Δ, so that we get:

$$\Delta = D^2 = \begin{pmatrix} D_+^* D_+ & 0 \\ 0 & D_+ D_+^* \end{pmatrix} ,$$

and hence $D_+^* D_+ = \Delta_+$ (resp.$D_+ D_+^* = \Delta_-$) is the restriction of Δ to the $+1$ (resp. -1) eigenspace of ε. It follows that:

$$\begin{aligned} Ind(D_+) &= dim(Ker D_+) - dim(Ker D_+^*) \\ &= dim(Ker D_+^* D_+) - dim(Ker D_+ D_+^*) \\ &= dim(Ker \Delta_+) - dim(Ker \Delta_-) \end{aligned}$$

and hence:

$$\begin{aligned} Ind(D_+) = &\Big(dim\left(Ker \Delta_+^{2k}\right) - dim\left(Ker \Delta_-^{2k}\right)\Big) \\ &+ \sum_{p=0}^{2k-1} \Big(dim\left(Ker \Delta_+^{p}\right) - dim\left(Ker \Delta_-^{p}\right)\Big), \end{aligned}$$

where we denote by $\Delta_\pm^p$ the restriction of $\Delta_\pm$ to the ε-invariant subspace

$$C^\infty(M, \Lambda_{\mathbb{C}}^p(T^*M) \oplus \Lambda_{\mathbb{C}}^{4k-p}(T^*M)) \text{ for } p = 0, 1, \ldots, 2k.$$

But we have:
$\omega \in Ker \Delta_\pm^p \Longleftrightarrow \omega = \alpha \pm \varepsilon(\alpha)$ for a harmonic p-form α,
and since the map $\alpha + \varepsilon(\alpha) \longrightarrow \alpha - \varepsilon(\alpha)$ induces an isomorphism between $Ker(\Delta_+^p)$ and $Ker\left(\Delta_-^p\right)$ for $p = 0, 1, \ldots, 2k-1$, we get:

$$Ind(D_+) = dim\left(Ker \Delta_+^{2k}\right) - dim\left(Ker \Delta_-^{2k}\right) .$$

By Hodge theory, $Ker(\Delta^{2k})$ identifies with $H^{2k}(M, \mathbb{C})$. Denote by $H_\pm$ the ± 1 eigenspace of $\varepsilon = *$ on $2k$-harmonic forms. Since we have:

$$\int_M \omega \wedge \omega = \pm \int_M \omega \wedge (*\omega) = \pm < \omega | \omega > \text{ for any } \omega \in H_\pm ,$$

the signature form is positive definite on H_+ and negative definite on H_-, so that finally $Ind(D_+) = \sigma(M)$. QED

Exercise 2.3.7. *Show that the analytical index of the De Rham operator D on a compact oriented Riemannian manifold M without boundary is given by:*

$$Ind(D) = \sum_{i=0}^{dim(M)} (-1)^i dim(H^i(M, \mathbb{C})) .$$

3 Topological K-Theory

The analytical index of an elliptic operator P of order m on a compact manifold M is computable from its principal symbol $\sigma(P)$ of order m. When $m = 0$, this principal symbol yields an element in the K-thery group (with compact support) of T^*M, and the analytical index can be viewed as a map $[\sigma(P)] \in K^0(T^*M) \longrightarrow Ind(P) \in \mathbb{Z}$. The aim of this section is to introduce the topological K-theory of locally compact spaces, in order to give a topological description of the above index map. Since it does not require more effort, we shall simultaneously introduce the topological K-functor $A \longrightarrow K_*(A) = K_0(A) \oplus K_1(A)$ for C^*-algebras.

3.1 The Group $\mathbf{K^0(X)}$

Definition 3.1.1. *Let X be a compact space. The K-theory group $K^0(X)$ is the abelian group generated by the isomorphism classes of complex vector bundles over X with the relations: $[E \oplus F] = [E] + [F]$ for any pair (E, F) of vector bundles over X.*

Every element of $K^0(X)$ is thus a difference $[E] - [F]$, where E and F are complex vector bundles over X. In this representation, we have (with obvious notations): $[E]-[F] = [E']-[F'] \Longleftrightarrow (\exists G)$ *such that* $E \oplus F' \oplus G \cong E' \oplus F \oplus G$. Denote by $[\tau]$ the K-theory class of the trivial bundle of rank 1 over X. Since there exists, for any complex vector bundle E over a compact space X, a vector bundle F over X such that $E \oplus F \cong X \times \mathbb{C}^n$ (trivial bundle of rank n), any element in $K^0(X)$ can be written in the form $[E] - n[\tau]$ for some bundle E over X and $n \in \mathbb{N}$.

Example 3.1.2. A vector bundle over a point is a finite dimensional complex vector space, and two such vector bundles are isomorphic if and only if they have same dimension. Henceforth, $K^0(p^t)$ is isomorphic to $\mathbb{Z}$.

Exercise 3.1.3. *Show that $K^0(S^1) \cong \mathbb{Z}$.*

Note that any continuous map $f : X \longrightarrow Y$ between compact spaces induces a group homomorphism $f^* : K^0(Y) \longrightarrow K^0(X)$ by $f^*([E]) = [f^*(E)]$, where $f^*(E) = \{(x, \zeta) \in X \times E | f(x) = \pi(\zeta)\}$ is the pull-back to X of the complex vector bundle $E \xrightarrow{\pi} Y$ over Y.

Definition 3.1.4. *If X is a locally compact space X, the K-theory group $K^0(X)$ (with compact support) is by definition the kernel of the map $K^0(\widehat{X}) \longrightarrow K^0(\{\infty\}) = \mathbb{Z}$ induced by the inclusion of $\{\infty\}$ into the one-point compactification $\widehat{X} = X \cup \{\infty\}$ of X.*

Racall that a continuous map $f : X \longrightarrow Y$ between locally compact spaces is called *proper* if $f^{-1}(K)$ is compact for any compact subset K of Y. From the definition of $K^0(X)$, it is clear that $X \longrightarrow K^0(X)$ is a contravariant functor from the category of locally compact spaces with proper continuous maps to abelian groups.

Exercise 3.1.5. *Show that we have* $K^0(\widehat{X}) = K^0(X) \oplus \mathbb{Z}$ *for any non-compact locally compact space* X. *Prove that* $K^0(\mathbb{R}) = \{0\}$.

Exercise 3.1.6. *Let* X *be a compact space that can be written as a disjoint union of two open subspaces* X_1 *and* X_2. *Prove that* $K^0(X) \cong K^0(X_1) \oplus K^0(X_2)$.

The main property of the functor $X \longrightarrow K^0(X)$ is its homotopy invariance:

Theorem 3.1.7. *Let* X, Y *be two locally compact spaces and* $f_t : X \longrightarrow Y$ $(0 \leq t \leq 1)$ *be a continuous path of proper maps from* X *to* Y. *Then, we have:*

$$f_0^* = f_1^* : K^0(Y) \longrightarrow K^0(X) .$$

For a proof of this result, see [10], theorem 1.25, p. 56.

Exercise 3.1.8. *By using theorem 3.1.7, show that* $K^0\left(]0,1]\right) = \{0\}$.

Exercise 3.1.9. *By using the correspondence between complex vector bundles over a compact space* X *and idempotents in matrix algebras over* $C(X)$, *try to give a proof of theorem 3.1.7.*

3.2 Fredholm Operators and Atiyah's Picture of $\mathbf{K^0(X)}$

Let X be a compact space and denote by H a separable infinite dimensional Hilbert space. Since the product of two Fredholm operators is Fredholm, the space $[X, Fred(H)]$ of homotopy classes of continuous functions from X to $Fred(H)$ has a natural semigroup structure. The following description of $K^0(X)$ by continuous fields of Fredholm operators can be found in [1]:

Theorem 3.2.1. *There is a group isomorphism* $Ind : [X, Fred(H)] \longrightarrow K^0(X)$ *such that:* $Ind \circ f_* = f^* \circ Ind$ *for any continuous map* $f : X \longrightarrow Y$ *of compact spaces.*

Sketch of Proof. Naively, we would like to define the index map Ind by setting:

(1) $$Ind([T]) = [(KerT_x)_{x\in X}] - [(KerT_x^*)_{x\in X}] \in K^0(X),$$

where $[T]$ denotes the the homotopy class of the continuous map

$$x \in X \longrightarrow T_x \in Fred(H) \; .$$

But since the dimension of $Ker(T_x)$ is not locally constant in general, $(KerT_x)_{x\in X}$ and $(KerT_x^*)_{x\in X}$ are not vector bundles over X so that the heuristic formula (1) does not make sense. To overcome this difficulty, fix a point $x_0 \in X$ and consider the map:

$$\widetilde{T}_x : (\zeta, \eta) \in Ker(T_{x_0}^*) \oplus H \longrightarrow \zeta + T_x\eta \in H \; ,$$

which is defined for x in a neighborhood of x_0. Since $\widetilde{T}_{x_0}$ is surjective, $\widetilde{T}_x$ is surjective for x in some neighborhood of x_0 and we get by the homotopy invariance of the index:

$$dim(Ker\widetilde{T}_x) = Ind(\widetilde{T}_x) = Ind(\widetilde{T}_{x_0}) = dim\; Ker(T_{x_0}) = Constant \; .$$

Now, by using a partition of unit, it is easy to patch together such local constructions (in the neighborhood of any point $x \in X$) to construct a finite number of continuous maps $\zeta_i : X \longrightarrow H (i = 1, 2, \ldots, N)$ satisfying the following two conditions:

(i) For any $x \in X$, the map

$$\widetilde{T}_x : (\lambda, \eta) \in \mathbb{C}^N \oplus H \longrightarrow \widetilde{T}_x(\lambda, \eta) = \Sigma_{i=1}^N \lambda_i \zeta_i(x) + T_x\eta \in H$$

is surjective;

(ii) The function $x \longrightarrow dim(Ker\widetilde{T}_x)$ is locally constant.

Now, $(Ker\widetilde{T}_x)_{x\in X}$ is a vector bundle over X by (ii) and we can define correctly the index map $Ind : [X, Fred(H)] \longrightarrow K^0(X)$ by setting (in view of (i)):

$$Ind([T]) = [(Ker\widetilde{T}_x)_{x\in X}] - [\mathbb{C}^N] \in K^0(X) \; .$$

It remains to check that Ind is a well defined map which is a group isomorphism. It is straghtforward to check that Ind is a well defined group homomorphism. To prove that Ind is an isomorphism, we can check that

$$[X, GL(H)] \longrightarrow [X, Fred(H)] \stackrel{Ind}{\longrightarrow} K^0(X) \longrightarrow 0$$

is an exact sequence (this is not hard) and use the contractibility of $GL(H)$ (Kuiper's theorem[3]) to get $[X, GL(H)] = 0$. QED

[3] For a proof of Kuiper's theorem, see [14].

3.3 Excision in K-Theory

Let Y be a closed subspace of a locally compact space X. The relative K-theory group $K^0(X,Y)$ is defined as a quotient of the set $Q(X,Y)$ of triples (E_0, E_1, σ) where E_0, E_1 are complex vector bundles over X that are direct factors of trivial bundles, and $\sigma \in Hom(E_0, E_1)$ is a morphism of vector bundles such that:

(i) There is a compact subset K of X such that $\sigma_{|X-K} : E_{0|X-K} \longrightarrow E_{1|X-K}$ is an isomorphism;
(ii) $\sigma_{|Y} : E_{0/Y} \longrightarrow E_{1|Y}$ is an isomorphism.

If σ is an isomorphism, the triple (E_0, E_1, σ) is called *degenerate*. There are obvious notions of sum, isomorphism and homotopy of pairs of triples in $Q(X,Y)$. Let us say that $(E_0, E_1, \sigma) \in Q(X,Y)$ and $(E_0', E_1', \sigma') \in Q(X,Y)$ are *equivalent* if there exist degenerate triples $(F_0, F_1, \rho), (F_0', F_1', \rho') \in Q(X,Y)$ and isomorphisms of bundles $\theta_0 : E_0 \oplus F_0 \longrightarrow E_0' \oplus F_0', \theta_1 : E_1 \oplus F_1 \longrightarrow E_1' \oplus F_1'$ such that $(E_0 \oplus F_0, E_1 \oplus F_1, \sigma \oplus \rho)$ is homotopic to $(E_0 \oplus F_0, E_1 \oplus F_1, \theta_1^{-1}(\sigma' \oplus \rho')\theta_0)$ in $Q(X,Y)$.

Definition 3.3.1. *Let Y be a closed subspace of a locally compact space X. The quotient of $Q(X,Y)$ by the above equivalence relation is denoted by $K^0(X,Y)$.*

$K^0(X,Y)$ is clearly an abelian group for the direct sum. The excision property can be expressed as follows:

Theorem 3.3.2. (Excision). *For any closed subspace Y of a locally compact space X, we have natural isomorphisms:*

$$K^0(X,Y) \cong K^0(X-Y) \approx K^0(X/Y, \{\infty\}) ,$$

where X/Y denotes the one-point compactification of $X-Y$ obtained by identifying all points in Y to a single point $\{\infty\}$.

Sketch of Proof. The second isomorphism is a tautology. To prove the first isomorphism, we may restrict our attention to the case where X is compact, since we have a natural isomorphism $K^0(X,Y) \cong K^0(\widehat{X}, \widehat{Y})$ where $\widehat{X}$ is the one point compactification of X. Let Z be the compact space obtained by gluing two copies $X_0 = X_1 = X$ of X along the common part $Y_0 = Y_1 = Y$ and denote by $i : X_1 \longrightarrow Z$ the natural inclusion. Since there is an obvious retraction $\rho : Z \longrightarrow X_1$, one can show that the natural exact sequence:

$$0 \longrightarrow K^0(Z - X_1) \cong K^0(X-Y) \xrightarrow{j^*} K^0(Z) \xrightarrow{i^*} K^0(X_1) \longrightarrow 0$$

is split exact, so that $K^0(Z) \cong K^0(X-Y) \oplus K^0(X_1)$. We can now construct the isomorphism $K^0(X,Y) \xrightarrow{\cong} K^0(X-Y)$. Let $[E_0, E_1, \sigma] \in K^0(X,Y)$ and consider the complex vector bundle F over Z obtained by identifying E_0 and E_1 over Y via the isomorphism $\sigma_{|Y}$. Since the element $[F]-[\rho^*(E_1)] \in K^0(Z)$ belongs to $Ker(i^*)$, there is a unique element $\chi(E_0, E_1, \sigma) \in K^0(X-Y)$ such that $j^*(\chi(E_0, E_1, \sigma)) = [F]-[\rho^*(E_1)]$. It is now straightforward to check that the map $(E_0, E_1, \sigma) \longrightarrow \chi(E_0, E_1, \sigma)$ defines an isomorphism from $K^0(X, Y)$ to $K^0(X - Y)$. QED

Definition 3.3.3. *Let X be a locally compact space. We call quasi isomorphism over X any triple (E_0, E_1, σ) where E_0, E_1 are complex vector bundles over X and $\sigma \in Hom(E_0, E_1)$ a morphism of vector bundles which is an isomorphism outside some compact subset of X.*

It follows from theorem 3.3.2 that any element in $K^0(X)$ can be represented by a quasi-isomorphism (E_0, E_1, σ) over X.

This slightly different point of view on K-theory with compact support allows to describe a multiplication in $K^0(X)$ by using the following heuristic formula:

$$[(E_0, E_1, \sigma)] \otimes [(F_0, F_1, \tau)]$$
$$= [((E_0 \otimes F_0) \oplus (E_1 \otimes F_1), (E_0 \otimes F_1) \oplus (E_1 \otimes F_0), \sigma \widehat{\otimes} 1 + 1 \widehat{\otimes} \tau)] ,$$

where $\sigma \widehat{\otimes} 1 + 1 \widehat{\otimes} \tau = \begin{pmatrix} \sigma \otimes 1 & -1 \otimes \tau^* \\ 1 \otimes \tau & \sigma^* \otimes 1 \end{pmatrix}$ (sharp product).

This product can be used to prove the Thom isomorphism for complex vector bundles. Let $\pi : E \longrightarrow X$ be a complex hermitian vector bundle over a compact space X, and consider the triple:

$$\Lambda_{-1}(E) = (E_0 = \pi^*(\Lambda^{ev}_{\mathbb{C}} E), E_1 = \pi^*(\Lambda^{odd}_{\mathbb{C}} E), \sigma) ,$$

where $\Lambda^{ev}_{\mathbb{C}} E = \bigoplus_p \Lambda^{2p}_{\mathbb{C}} E$ and $\sigma : E_0 \longrightarrow E_1$ is the morphism of bundles (over the total space of E) given by $\sigma(x, \zeta)(\omega) = \zeta \wedge \omega - \zeta^* \lfloor \omega$. Although $\Lambda_{-1}(E)$ is not a quasi-isomorphism over E, its sharp product $\pi_*(F_0, F_1, \varphi) \otimes \Lambda_{-1}(E)$ where (F_0, F_1, φ) is a quasi-isomorphism over X, yields an element in $K^0(E)$ which only depends on $[(F_0, F_1, \varphi)]$. Let us denote by $\pi_*([F_0, F_1, \varphi]) \otimes [\Lambda_{-1}(E)]$ this element. We have:

Theorem 3.3.4.(Thom isomorphism for complex hermitian bundles). *For any complex hermitian vector bundle $\pi : E \longrightarrow X$ over a compact space X, the map:*

$$[(F_0, F_1, \varphi)] \in K^0(X) \longrightarrow \pi_*([F_0, F_1, \varphi]) \otimes [\Lambda_{-1}(E)] \in K^0(E)$$

induces an isomorphism of K-theory groups.

For a proof of this result, see [10].

Exercise 3.3.5. *Show that* $K^0(\mathbb{R}^{2n}) \cong \mathbb{Z}$.

3.4 The Chern Character

For any locally compact space X, denote by $H^*(X, \mathbb{Q})$ the rational Čech chomology of X with compact support. We have by setting $H^{ev}(X, \mathbb{Q}) = \sum_k H^{2k}(X, \mathbb{Q})$:

Theorem 3.4.1. *There exists a natural homomorphism* $ch : K^0(X) \longrightarrow H^{ev}(X, \mathbb{Q})$*, called the Chern character, which satisfies the following properties:*

(i) $ch(f^*(x)) = f^*(ch(x))$ *for any proper map* $f : X \longrightarrow Y$ *and* $x \in K^0(Y)$*;*
(ii) $ch(x + y) = ch(x) + ch(y)$ *for* $x, y \in K^0(X)$*;*
(iii) $ch([L]) = e^{c_1(L)}$ *for any complex line bundle* L *over* X*, where* $c_1(L)$ *is the first Chern class of* L*, i.e. the image of the* 1*-cocycle associated with the* $\mathbb{S}^1$*-bundle* L *by the natural isomorphism* $c_1 : H^1(X, \mathbb{S}^1) \xrightarrow{\cong} H^2(X, \mathbb{Z})$*;*
(iv) $ch([E \otimes F]) = ch([E])ch([F])$ *for any pair* (E, F) *of complex vector bundles over* X*;*
(v) If X *is compact, the Chern character extends to an isomorphism:*

$$ch : K^0(X) \otimes \mathbb{Q} \xrightarrow{\cong} H^{ev}(X, \mathbb{Q}) .$$

Let us give a construction of the Chern character.

3.4.2 Construction of the Chern Character

Assume for simplicity that X is a compact manifold M. Let E be a k-dimensional complex vector bundle over M and choose a connexion ∇ on E. To any polynomial function $P : M_k(\mathbb{C}) \longrightarrow \mathbb{C}$ such that $P(XY) = P(YX)$ for any $X, Y \in M_k(\mathbb{C})$ (one say that P is an *invariant polynomial*), we can associate a closed differential form $P(E)$ on M by the formula $P(E) = P(\Omega)$, where Ω is the curvature of the connexion ∇, which is a 2-form on M with values in $End(E)$. Choosing a local framing for E, we may identify Ω with a matrix of ordinary 2-forms. Since P is an invariant polynomial, one can check that $P(\Omega)$ is a well defined differentiable form (independent of the choice of the local framing) whose cohomology class, again denoted by $P(E)$, does not depend on the connexion ∇ in E (see for instance [18], prop. 10.5, p. 112). The Chern character of E is defined by:

$$ch(E) := \sum_{k \geq 0} s_k(E) \in H^{ev}(M) ,$$

where s_k is the invariant polynomial $s_k(X) = \frac{1}{k!} Tr\left(\left(\frac{X}{4i\pi}\right)^k\right)$. We thus have formally:

$$ch(E) = \left[Tr\left(exp\left(\frac{\Omega}{4i\pi}\right)\right)\right] \in H^{ev}(M) ,$$

where Ω is the curvature of some connexion ∇ on E. One can prove that the Chern character only depends on the K-theory class of E, and extends to a homomorphism $ch : K^0(X) \longrightarrow H^{ev}(X, \mathbb{Q})$, which is the Chern character of theorem 3.4.1.

3.4.3 Computation of the Chern Character

Since the ring of invariant polynomials on $M_k(\mathbb{C})$ is generated by the polynomials $c_k(X) = \frac{Tr(\Lambda^k X)}{(4i\pi)^k}$, we can express $ch(E)$ from the corresponding *Chern classes* $c_k(E) \in H^{2k}(X, \mathbb{Q})$. We get:

$$ch(E) = dim(E) + c_1(E) + \frac{1}{2}(c_1^2(E) - 2c_2(E)) + \ldots \quad \text{(see below)} .$$

Exercise. *Show that $c_i(\overline{E}) = (-1)^i c_i(E)$ where $\overline{E}$ is the conjugate bundle of E.*

Let us denote by (x_i) the eigenvalues of $\frac{\Omega}{4i\pi}$. We have:

$$\begin{cases} x_1 + \ldots + x_m = c_1 \\ x_1 x_2 + \ldots + x_{m-1} x_{m-2} = c_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_1 \ldots x_m = c_m \end{cases}$$

where $m = dim(E)$, so that the Chern classes are the elementary symetric functions of the x_i. It follows that any symmetric formal power series in the x_i, which can therefore be expressed in terms of the elementary symmetric functions of the x_i, yields a cohomology class in $H^*(M, \mathbb{Q})$. For instance, any function $f(z)$ holomorphic near $z = 0$ gives rise to a cohomology class by the formula:

$$f(E) = \prod_{i=1}^{m} f(x_i) .$$

When $E = \bigoplus_{i=1}^{m} L_i$ is a sum of complex line bundles L_i, we can choose $x_i = c_1(L_i) \in H^2(M, \mathbb{Z})$ and we get from theorem 3.4.1:

$$ch(E) = \sum_{i+1}^{m} ch(L_i) = \sum_{i=1}^{m} e^{x_i} = m + \sum_{i=1}^{m} x_i + \frac{1}{2}\sum_{i=1}^{m} x_i^2 + \dots$$
$$= dim(E) + c_1(E) + \tfrac{1}{2}(c_1^2(E) - 2c_2(E)) + \dots$$

For instance, it follows from the relation $\Lambda^k(E \oplus F) = \sum_{i+j=k} \Lambda^i(E) \otimes \Lambda^j(F)$ that:

$$ch\left([\Lambda^{even}(E)] - [\Lambda^{odd}(E)]\right) = \prod_{i=1}^{m}(1 - e^{x_i}) \, .$$

Of course, a complex vector bundle E over M is not always isomorphic to a sum of complex line bundles. However, if we just want to identify $ch(E)$ with some naturally defined cohomology class (for instance, with $f(E)$ for some power series $f(z)$), we may use the following *splitting principle* (see [9], proposition 5.2, p. 237 for a proof):

Splitting Principle. *For any complex vector bundle over a manifold M, there exists a smooth fibration $f : N \longrightarrow M$ such that:*

(i) $f^*(E)$ *splits into a direct sum of complex line bundles;*
(ii) $f^* : H^*(M, \mathbb{Q}) \longrightarrow H^*(N, \mathbb{Q})$ *is injective.*

To define the higher K-theory groups $K^n(X)$ $(n \geqq 1)$ of a locally compact space X, we shall directly define the K-groups $K_n(A)$ of a C^*-algebra A and set $K^n(X) = K_n(C_0(X))$, where $C_0(X)$ is the C^*-algebra of continuous functions on X vanishing at the infinity.

3.5 Topological K-Theory for C*-Algebras

3.5.1 C*-Algebras. Recall that a C^*-algebra is a complex Banach algebra A with involution $x \in A \longrightarrow x^* \in A$ whose norm satisfies:

$$\|x^*x\| = \|x\|^2 \text{ for any } x \in A \, .$$

By Gelfand theory, any commutative C^*-algebra is isometrically isomorphic to the C^*-algebra $C_0(X)$ of complex continuous functions vanishing at infinity on some locally compact space X (the spectrum of A). If H is a Hilbert space, a closed $*$-subalgebra of $B(H)$ is a C^*-algebra, and any C^*-algebra can be realized as a closed $*$-subalgebra of $B(H)$ for some Hilbert space H. For instance, the algebra $K(H)$ of all compact operators on a separable Hilbert space H is a C^*-algebra. C^*-algebras naturally appear in non-commutative topology to describe "quantum spaces" like the quotient of a locally compact space X by a non proper action of a discrete group Γ For instance, the "dual" of a discrete groupe Γ. is described by the C^*-algebra $C^*(\Gamma)$ generated in $B(l^2(\Gamma))$ by the left regular representation λ defined by:

$$[\lambda(g)\xi](h) = \xi(g^{-1}h), \text{ where } g, h \in \Gamma \text{ and } \xi \in l^2(\Gamma).$$

Another example is the crossed product C^*-algebra $A \times_\alpha G$ of a C^*-algebra A by a continuous action $g \in G \longrightarrow \alpha_g \in Aut(A)$ of a locally compact group G acting on A by automorphisms. Here, we assume that $g \in G \longrightarrow \alpha_g(x) \in A$ is continuous for any $x \in A$ and we denote by Δ_G the modular function of G. The vector space $C_c(G, A)$ of continuous compactly supported functions on G with values in A has a natural structure of $*$-algebra. To describe this structure, it is convenient to write any element $a \in C_c(G, A)$ as a formal integral $a = \int a(g) U_g dg$, where U_g is a letter satisfying:

$$U_{gh} = U_g U_h,\ U_g^* = U_g^{-1} = U_{g^{-1}},\ \text{and } U_g x U_g^{-1} = \alpha_g(x) \text{ for any } x \in A.$$

Then, the product and the involution on $C_c(G, A)$ are given by:

$$\left(\int a(g)U_g dg\right)\left(\int b(g)U_g dg\right) = \int c(g)U_g dg\ ,$$

$$\text{where } c(g) = \int a(h)\alpha_h\left(b\left(h^{-1}g\right)\right) dh$$

$$\text{and } \left(\int a(g)\,U_g\,dg\right)^* = \int b(g)\,U_g\,dg, \text{ where } b(g) = \Delta_G(g)\alpha_g\left(a\left(g^{-1}\right)^*\right).$$

There are two natural ways of completing $C_c(G, A)$ to get a crossed product C^*-algebra $A \times_\alpha G$; they coincide when G is amenable. For more information on this subject, we refer to [17], p. 240.

Exercise 3.5.2. *Let θ be an irrational number and consider the action α of $\mathbb{Z}$ on $C(\mathbb{S}^1)$ defined by: $\alpha(f)(z) = f(e^{-2i\pi\theta} z)$. Show that $A_\theta = C(S^1) \times_\alpha \mathbb{Z}$ is the C^*-algebra generated by two unitaries U and V satisfying the commutation relation:*

$$UV = e^{2i\pi\theta} VU\ .$$

(non commutative 2-torus).

3.5.3 $\mathbf{K_0}$ of a $\mathbf{C^*}$-Algebra. Let A be a unital C^*-algebra. Recall that a finitely generated right module $\mathbf{E}$ over A is called *projective* if there exists a right module $\mathbf{F}$ over A such that $\mathbf{E} \oplus \mathbf{F} \cong A^n$. For instance, if $e \in M_n(A)$ is an idempotent, $\mathbf{E} = eA^n$ is a finitely generated projective module over A. Conversely, any finitely generated projective right A-module is of the above form.

Exercise 3.5.4. *Let X be a compact space. For any complex vector bundle E on X, denote by $\mathbf{E}$ the module of continuous sections of E. Show that $\mathbf{E}$ is a finitely generated projective module over $C(X)$.*

Let A be a unital C^*-algebra, and denote by $\mathbf{K}_0(A)$ the set of isomorphism classes of finitely generated projective A-modules. The direct sum of modules induces a commutative and associative sum on $\mathbf{K}_0(A)$.

Definition 3.5.5. *The group of formal differences* $[\mathbf{E}]$–$[\mathbf{F}]$ *of elements in* $\mathbf{K}_0(A)$ *is denoted by* $K_0(A)$.

Exercise 3.5.6. *Show that* $K_0(M_n(\mathbb{C})) \cong \mathbb{Z}$.

Note that any unital $*$-homomorphism $\pi : A \longrightarrow B$ between unital C^*-algebras induces a group homorphism $\pi_* : K_0(A) \longrightarrow K_0(B)$ by $\pi_*([\mathbf{E}]) = [\mathbf{E} \otimes_A B]$.

For a non unital C^*-algebra A, the unital morphism

$$\varepsilon : (a, \lambda) \in \widetilde{A} = A \oplus \mathbb{C} \longrightarrow \lambda \in \mathbb{C}$$

from the algebra $\widetilde{A}$ obtained by adjoining a unit to A to the scalars induces a K-theory map $\varepsilon_* : K_0(\widetilde{A}) \longrightarrow K_0(\mathbb{C})$. By definition, $K_0(A)$ is the kernel of ε_*.

3.5.7 $\mathbf{K_n}$ of a $\mathbf{C^*}$-Algebra ($\mathbf{n \leq 1}$). To define $K_n(A)$ ($n \in \mathbb{N} - \{0\}$) for a C^*-algebra A with unit, consider the group $GL_k(A)$ of invertible elements in $M_k(A)$. Let $GL_\infty(A)$ be the union of the $GL_k(A)$'s, where $GL_k(A)$ embeds in $GL_{k+1}(A)$ by the map $X \longrightarrow \begin{pmatrix} X & 0 \\ 0 & 1 \end{pmatrix}$. Note that $GL_\infty(A)$ is a topological group for the inductive limit topology.

Definition 3.5.8. For $n \geqq 1$, we set: $K_n(A) := \pi_{n-1}(GL_\infty(A))$.

We thus define a group which is abelian for $n \geqq 2$, since the homotopy group π_n of a topological group is abelian for $n \geqq 1$.

Exercise 3.5.9. *Show that* $\begin{pmatrix} X & 0 \\ 0 & 1 \end{pmatrix}$ *and* $\begin{pmatrix} 1 & 0 \\ 0 & X \end{pmatrix}$ *are in the same connected component of* $GL_{2n}(A)$ *for any* $X \in GL_n(A)$. *Deduce that* $K_1(A)$ *is abelian.*

Exercise 3.5.10. *Show that* $K_1(M_n(\mathbb{C})) = \{0\}$.

Any unital $*$-homomorphism $\pi : A \longrightarrow A$ between unital C^*-algebras yields a group homorphism $\pi_* : K_n(A) \longrightarrow K_n(B)$. For a non unital C^*-algebra A, the group $K_n(A)$ will be the kernel of the map $\varepsilon_* : K_n(\widetilde{A}) \longrightarrow K_n(\mathbb{C})$.

3.6 Main Properties of the Topological K-Theory for C*-Algebras

The following theorem summarizes the main properties of the K-theory for C^*-algebras:

Theorem 3.6.1. *The covariant functor* $A \longrightarrow K_n(A)$ *satisfies the following properties:*

(i) (Homotopy invariance). *We have* $(\pi_0)_* = (\pi_1)_* : K_n(A) \longrightarrow K_n(B)$ *for any path* $\pi_t : A \longrightarrow B$ $(t \in [0,1])$ *of unital* $*$*-homomorphisms from* A *to* B *such that* $t \longrightarrow \pi_t(x)$ *is norm continuous for any* $x \in A$*;*

(ii) (Stability). *There is a natural isomorphism* $K_n(A) \cong K_n(A \otimes K(H))$, *where* $K(H)$ *is the algebra of compact operators on a separable infinite dimensional Hilbert space* H*;*

(iii) For any $n \geqq 0$, there exists a natural isomorphism $\beta_n : K_n(A) \longrightarrow K_{n+2}(A)$;

(iv) (Six terms exact sequence). *Any short exact sequence of* C^**-algebras* $0 \longrightarrow J \xrightarrow{i} A \xrightarrow{p} B \longrightarrow 0$ *yields a cyclic exact sequence in* K*-theory:*

$$\begin{array}{ccccc} K_0(J) & \xrightarrow{i_*} & K_0(A) & \xrightarrow{p_*} & K_0(B) \\ \uparrow \delta & & & & \delta \downarrow \\ K_1(B) & \xleftarrow{p_*} & K_1(A) & \xleftarrow{i_*} & K_1(J) \end{array}$$

(v) (Bott periodicity). *For any* C^**-algebra* A, *there exists a natural isomorhism* $K_{i+n}(A) \xrightarrow{\approx} K_i(A \otimes C_0(\mathbb{R}^n))$*;*

(vi) (Thom isomorphism). *For any continuous action* α *of* $\mathbb{R}^n$ *by automorphisms of the* C^**-algebra* A, *there exists a natural isomorphism*

$$K_{i+n}(A) \xrightarrow{\approx} K_i(A \times_\alpha \mathbb{R}^n) ,$$

where $A \times_\alpha \mathbb{R}^n$ *is the crossed product* C^**-algebra of* A *by the action* α *of* $\mathbb{R}^n$.

Let us make some comments on the proofs.
Property (i) is obvious for $n \geq 1$. For $n = 0$, it follows from the fact that two nearby projections e, f in a C^*-algebra A are equivalent, i.e. there exists $u \in A$ such that $u^* = e$ and $u^*u = f$ (henceforth, $u : eA \longrightarrow fA$ is an A-module isomorphism).
Property (ii) is an immediate corollary of (i).
Property (iii) is a theorem, originally proved by Bott. It implies that the K-theory of a C^*-algebra A reduces to the groups $K_0(A)$ and $K_1(A)$. For $n = 0$, the isomorphism $\beta_0 : K_0(A) \longrightarrow K_2(A)$ is easy to describe: it sends the class of the module $eA^n (e = e^* = e^2 \in M_n(A))$ to the class of the loop:

$$z \in U(1) \longrightarrow ze + (1-e) \in GL_n(A) .$$

Property (iv) is a consequence of the long exact sequence for the homotopy groups of a fibration, which reduces here to a cyclic exact sequence in view of (iii).
Property (v) follows from (iv) for the exact sequence

$$0 \longrightarrow C_0(]0,1[, A) \xrightarrow{i} C_0(]0,1], A) \xrightarrow{p} A \longrightarrow 0 ,$$

where p is the evaluation at 1, since we have $K_n(C_0(]0,1], A)) = 0$.

Exercise 3.6.2. *Show that the path of $*$-morphisms $\pi_t : C_0(]0,1], A) \longrightarrow C_0(]0,1], A)$ $(t \in [0,1])$ defined by $\pi_t(f)(s) = \begin{cases} f(s-t) & if\ 0 \le t < s \le 1 \\ 0 & if\ 0 \le s \le t \le 1 \end{cases}$ yields a homotopy between 0 and Id. Deduce that $K_n(C_0(]0,1], A)) = 0$.*

Property (vi) was originally proved by A. Connes [5], and can be reduced to Bott periodicity (see for instance [7]).

3.7 Kasparov's Picture of $\mathbf{K_0(A)}$

In analogy with Atiyah's description of $K^0(X)$ for a compact space X, it is possible to describe $K_0(A)$ for any unital C^*-algebra A from *generalized A-Fredholm operators*. The main change consists in replacing the notion of Hilbert space by that of *Hilbert C^*-module.*

Definition 3.7.1. *Let A be a C^*-algebra. We call Hilbert C^*-module over A (or Hilbert A-module) any right A-module $\mathbf{E}$ equipped with an A-valued scalar product $< .,. >$ satisfying the following conditions:*

(i) $< \xi, \lambda\eta >= \lambda < \xi, \eta >$ *and* $< \xi, \eta a >=< \xi, \eta > a$ *for any* $\xi, \eta \in \mathbf{E}, a \in A$*;*
(ii) $< \xi, \eta >=< \eta, \xi >^*$ *for any* $\xi, \eta \in \mathbf{E}$*;*
(iii) $< \xi, \xi >\in A^+$ *for any* $\xi \in \mathbf{E}$ *and* $(< \xi, \xi >= 0 \Longrightarrow \xi = 0)$*;*
(iv) $\mathbf{E}$ *is complete for the norm* $\| \xi \| = \| < \xi, \xi > \|_A^{1/2}$*.*

A basic example of Hilbert A-module is the completion H_A of the algebraic direct sum $A \oplus A \oplus A \oplus \ldots$ for the norm associated with the A-valued scalar product
$< (x_n), (y_n) >= \sum_{n \geqq 1} x_n^* y_n \in A$. In fact, one can show in analogy with the Hilbert space theory [11] that the sum of any countably generated Hilbert A-module with H_A is isomorphic to H_A.

Exercise 3.7.2. *Let X be a compact space. Show that the Hilbert $C(X)$-modules are exactly the spaces of continuous sections of continuous fields of Hilbert spaces over X.*

Definition 3.7.3. *Let $\mathbf{E}$ be a Hilbert A-module. We shall call endomorphism of $\mathbf{E}$ any map $T : \mathbf{E} \longrightarrow \mathbf{E}$ such that there exists $T^* : \mathbf{E} \longrightarrow \mathbf{E}$ satisfying:*

$$< T\xi, \eta >=< \xi, T^*\eta > \ \ for\ any\ \xi, \eta \in \mathbf{E} \,.$$

An endomorphism of $\mathbf{E}$ is automatically A-linear and bounded. We denote by $B(\mathbf{E})$ the space of all endomorphisms of $\mathbf{E}$; it is a C^*-algebra for the operator norm. The closed ideal of $B(\mathbf{E})$ generated by the "rank one" operators $\xi \longrightarrow \xi_2 < \xi_1, \xi > (\xi_1, \xi_2 \in \mathbf{E})$ is denoted by $K(\mathbf{E})$; we call it the algebra of *compact operators of the Hilbert A-module* $\mathbf{E}$.

Definition 3.7.4. *Let A be a unital C^*-algebra and $\mathbf{E}$ a countably generated Hilbert A-module. A generalized A-Fredholm operator on $\mathbf{E}$ is by definition an endomophism $P \in B(\mathbf{E})$ such that there exists $Q \in B(\mathbf{E})$ with $R = I - PQ \in K(\mathbf{E})$ and $S = 1 - QP \in K(\mathbf{E})$.*

3.7.5 Generalized Fredholm A-index and Kasparov's Definition of $\mathbf{K_0(A)}$. Let A be a unital C^*-algebra and consider a *generalized A-Fredholm operator* $P \in B(\mathbf{E})$. In analogy with the description of the index of a Fredholm operator given in proposition 1.3.6, we shall define a *generalized A-index* $Ind_A(P) \in K_0(A)$. Since $\mathbf{E} \oplus H_A$ is isomorphic to H_A, we may assume, replacing P by $\begin{pmatrix} P & 0 \\ 0 & I \end{pmatrix}$ if necessary, that $\mathbf{E} = H_A$. In analogy with proposition 1.3.6, the generalized A-index $Ind_A(P) \in K_0(A) = K_0(K \otimes A)$ is defined by the formula:

$$Ind_A(P) = \left[X \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X^{-1} \right] - \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right] \in K_0(K(H_A)) = K_0(K \otimes A),$$

where $X \in B(H_A \oplus H_A)$ is some invertible lift of:

$$\begin{pmatrix} \dot{P} & 0 \\ 0 & \dot{P}^{-1} \end{pmatrix} \in B(H_A \oplus H_A)/K(H_A \oplus H_A) \,.$$

Here, $[e]$ is a shorthand for $[e\widetilde{B}^n]$ for any idempotent e of $M_n(\widetilde{B})$. Since the K-theory class of $\left[X \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X^{-1} \right] - \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right]$ does not depend on the choice of such an invertible lift X, the only point that we need to check is the existence of such a lift. With this aim in mind, consider the element:

$$X = \begin{pmatrix} P + (I - PQ)P & PQ - I \\ I - QP & Q \end{pmatrix} \in B(H_A \oplus H_A) \,.$$

It is equal to $\begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}$ modulo $K(H_A \oplus H_A)$ and the formula:

$$\begin{pmatrix} P + (I - PQ)P & PQ - I \\ I - QP & Q \end{pmatrix} = \begin{pmatrix} I & P \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ -Q & I \end{pmatrix} \begin{pmatrix} I & P \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$$

shows that it is invertible, with inverse:

$$X^{-1} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} I & -P \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ Q & I \end{pmatrix} \begin{pmatrix} I & -P \\ 0 & I \end{pmatrix} = \begin{pmatrix} Q & I - QP \\ PQ - I & P + (I - PQ)P \end{pmatrix} .$$

Since we have $X \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X^{-1} = \begin{pmatrix} I - R^2 & (I + R)PS \\ SQ & S^2 \end{pmatrix}$, the generalized A-index $Ind - A(P) \in K_0(A)$ will be finally defined by:

$$Ind_A(P) = \left[\begin{pmatrix} I - R^2 & (I + R)PS \\ SQ & S^2 \end{pmatrix} \right] - \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right] \in K_0(A) .$$

One can show that the index map $P \in Fred_A(H_A) \longrightarrow Ind_A(P) \in K_0(A)$ from the space $Fred_A(H_A)$ of generalized A-Fredholm operators to $K_0(A)$ induces a group isomorphism from $\pi_0(Fred_A(H_A))$ to $K_0(A)$. This leads Kasparov [12] to define $K_0(A)$ as the set of homotopy classes of triples $\left(\mathbf{E}^0, \mathbf{E}^1, F = \begin{pmatrix} 0 & Q \\ P & 0 \end{pmatrix} \right)$ where $\mathbf{E} = \mathbf{E}^0 \oplus \mathbf{E}^1$ is a graded Hilbert A-module and F an element in $B(\mathbf{E}^0 \oplus \mathbf{E}^1)$ such that:

$$F^2 - I = \begin{pmatrix} QP - I & 0 \\ 0 & PQ - I \end{pmatrix} \in K(\mathbf{E}^0 \oplus \mathbf{E}^1) .$$

This will be used in Sect. 5.

4 The Atiyah-Singer Index Theorem

4.1 Statement of the Theorem

We are now in position to state the Atiyah-Singer index theorem, which computes the analytical index of an elliptic operator P on M from its principal symbol $\sigma(P)$. Note that $\sigma(P)$ is a purely topological data which can be viewed as an element in $K^0(T^*M)$ (see remark after definition 3.3.3).

Theorem 4.1.1 (Atiyah-Singer index theorem). *Let P be an elliptic pseudodifferential operator on an n-dimensional compact oriented manifold M without boundary. Denote by $\sigma(P)$ the principal symbol of P, viewed as an element of the K-theory (with compact support) group $K^0(T^*M)$. Let $\pi! : H^*(T^*M) \longrightarrow H^*(M)$ be the integration's map (in cohomology with compact support) on the fibre of the canonical projection $\pi : T^*M \longrightarrow M$. Then, we have:*

$$Ind(P) = (-1)^{\frac{n(n+1)}{2}} \int_M ch_M(\sigma(P)) Td_{\mathbb{C}}(TM \otimes \mathbb{C}) ,$$

where $ch_M(\sigma(P)) = \pi! ch(\sigma(P))$ *is the image of* $ch(\sigma(P)) \in H^*(T^*M)$ *by* $\pi!$

The main steps of the proof are the following:

(i) *Construction of an analytical map* $Ind_a : K^0(T^*M) \longrightarrow \mathbb{Z}$, called the *analytical index,* such that $Ind_a(\sigma(P)) = Ind(P)$ for any elliptic pseudodifferential operator P on M with principal symbol $\sigma(P) \in K^0(T^*M)$;

(ii) *Construction of a topological map* $Ind_t : K^0(T^*M) \longrightarrow K^0(T^*\mathbb{R}^N) = \mathbb{Z}$, called the *topological index*, by using an embedding $M \longrightarrow \mathbb{R}^N$;

(iii) Proof of the equality $Ind_a = Ind_t$;

(iv) *Computation of the topological index* Ind_t by using the Chern character, to get the cohomological formula:

$$Ind_t(x) = (-1)^{\frac{n(n+1)}{2}} \int_M ch_M(x) Td_{\mathbb{C}}(TM \otimes \mathbb{C}) .$$

4.2 Construction of the Analytical Index Map

Let $P : C^\infty(M, E) \longrightarrow C^\infty(M, F)$ be an elliptic operator of order m on M, and consider its principal symbol $p = \sigma(P) \in C^\infty(T^*M, Hom(\pi^*E, \pi^*F))$, which is a bounded function in $S^m(T^*M, Hom(\pi^*E, \pi^*F))$. By ellipticity, there exists a bounded map $q \in S^{-m}(T^*M, Hom(\pi^*F, \pi^*E))$ such that $pq - I$ and $qp - I$ are bounded functions in $S^{-1}(T^*M, Hom(\pi^*F, \pi^*F))$ and $S^{-1}(T^*M, Hom(\pi^*E, \pi^*E))$ respectively. Note that the index of P only depends on the homotopy class of $\sigma(P)$:

Proposition 4.2.1. *Let* $P_0, P_1 : C^\infty(M, E) \longrightarrow C^\infty(M, F)$ *be two elliptic operators of order* m *on* M*. Assume that there exists a homotopy inside the symbols of elliptic operators of order* m *between the principal symbols*

$$p_0 = \sigma(P_0), p_1 = \sigma(P_1) \in C^\infty(T^*M, Hom(\pi^*E, \pi^*F))$$

of P_0 *and* P_1*. Then, we have:* $Ind(P_0) = Ind(P_1)$*.*

Proof. Any homotopy $t \in [0, 1] \longrightarrow p(t) \in C^\infty(T^*M, Hom(\pi^*E, \pi^*F))$ inside the symbols of elliptic operators of order m between p_0 and p_1 yields by pseudo-differential calculus a continuous field of Fredholm operators $P(t) : H^{s+m}(M, E) \longrightarrow H^s(M, F)$ with $P(0) = P_0$ and $P(1) = P_1$. By homotopy invariance of the index, we get $Ind(P_0) = Ind(P_1)$. QED

Let $P : C^\infty(M, E) \longrightarrow C^\infty(M, F)$ be an elliptic operator of order m on M, and choose a homogeneous function h of degree one on T^*M which is positive and C^∞ outside the zero section. By using proposition 4.2.1 we can show that, for any pseudodifferential operator $P_\tau : C^\infty(M,E) \longrightarrow C^\infty(M,F)$ with principal symbol $p_\tau(x,\xi) = p(x, \tau\frac{\xi}{h(\xi)})$, we have:

$$Ind(P_\tau) = Ind(P) \text{ for any } \tau > 0 \text{ sufficiently large.}$$

In other words, to determine the index of elliptic operators, it suffices to study operators with polyhomogeneous symbols of order 0. This leads to the following definition of the *analytical index map*. We shall use the following notation:

Notation. If $p \in C(T^*M, Hom(\pi^*E, \pi^*F))$ is a continuous section of the bundle $Hom(\pi^*E, \pi^*F)$ such that the set $\{(x,\xi) \in T^*M;\ p(x,\xi)$ is not invertible$\}$ is compact, we shall set: $p_\tau(x,\xi) = p(x, \tau\frac{\xi}{h(\xi)})$. We thus define a homogeneous continuous symbol p_τ which is invertible for τ sufficiently large.

Proposition and Definition 4.2.2. *Let $p \in C(T^*M, Hom(\pi^*E, \pi^*F))$ be such that the set of $(x,\xi) \in T^*M$ where $p(x,\xi)$ is not invertible is compact.*

(i) We have $Ind(P_1) = Ind(P_2)$ for any pair $P_1, P_2 : C^\infty(M,E) \longrightarrow C^\infty(M,F)$ of elliptic pseudodifferential with principal symbols $p_i(x,\xi)$ $(i = 1,2)$ satisfying: $\underset{x,\xi}{Sup}\, \|p_\tau(x,\xi)^{-1}p_i(x,\xi) - I\| < 1$ for $\tau > 0$ large enough;

*(ii) Set $Ind_a(p) := Ind(P)$, where $P : C^\infty(M,E) \longrightarrow C^\infty(M,F)$ is any elliptic pseudo-differential operator with principal symbol $p(x,\xi)$ satisfying $\underset{x,\xi}{Sup}\, \|p_\tau(x,\xi)^{-1}$
$p(x,\xi) - I\| < 1$ for $\tau > 0$ large enough. Then, $Ind_a(p)$ doesn't depend on the choice of h and τ. We call $Ind_a(p)$ the analytical index of p;*

(iii) We have $Ind_a(p) = Ind(P)$ if P is an elliptic pseudodifferential operator of order 0 with polyhomogeneous principal symbol p of order 0;

(iv) If $t \in [0,1] \longrightarrow p_t \in C(T^*M, Hom(\pi^*(E), \pi^*(F)))$ is a continuous path such that there exists a compact K in T^*M with $p_t(x,\xi)$ invertible for all t when $(x,\xi) \notin K$, then $Ind_a(p_t)$ is independent of t.

By (iv), the map $p \longrightarrow Ind_a(p)$ defined in (ii) yields a K-theory map $Ind_a : K^0(T^*M) \longrightarrow \mathbb{Z}$, called the *analytical index map.* Let us show that this K-theory map can be defined in a purely topological way.

4.3 Construction of the Topological Index Map

Let M be a compact oriented manifold without boundary of dimension n. There is a natural way to send $K^0(T^*M) \cong K^0(TM)$ to $K^0(pt) = \mathbb{Z}$ that we shall now describe. Choose an imbedding $i : M \longrightarrow \mathbb{R}^N$ of M into $\mathbb{R}^N$

(such an embedding always exist) and denote by $di : TM \longrightarrow T\,\mathbb{R}^N$ the corresponding proper imbedding of TM into $T\,\mathbb{R}^N$. The normal bundle to this embedding identifies wit the pull-back to TM of $N \oplus N$, where N is the normal bundle to the imbedding $i : M \longrightarrow \mathbb{R}^N$. Let us identify $N \oplus N$ with a tubular neighbourhood W of TM in $T\,\mathbb{R}^N$. Then, the Thom isomorphism for hermitian complex vector bundles (cf. [10]) yields a map $K^0(TM) \longrightarrow K^0(N \oplus N) \cong K^0(W)$.

Since W is an open subset of $T\,\mathbb{R}^N$, the natural inclusion $C_0(W) \longrightarrow C_0(T\,\mathbb{R}^N)$ yields a map $K^0(W) \longrightarrow K^0\left(T\,\mathbb{R}^N\right)$ and hence, by composition, a map:

$$i! : K^0(TM) \longrightarrow K^0(T\,\mathbb{R}^N) = K^0(\mathbb{R}^{2N}) \ .$$

Note that any smooth proper embedding $i : M \longrightarrow V$ of M into a smooth manifold V yields in the same way a natural map $i! : K^*(TM) \longrightarrow K^*(TV)$ which does not depend on the factorization of $di : TM \longrightarrow TV$ through the zero section associated with a tubular neighbourhood of TM into TV. Since $\mathbb{R}^{2N} = \mathbb{R}^N \oplus \mathbb{R}^N = \mathbb{C}^N \longrightarrow pt$ can be considered as a complex vector bundle over a point, we have a Thom isomorphism $K^0(pt) \longrightarrow K^0(\mathbb{R}^{2N})$ whose inverse is just the Bott periodicity isomorphism: $\beta : K^0(\mathbb{R}^{2N}) \longrightarrow K^0(pt) = \mathbb{Z}$. Taking $V = \mathbb{R}^{2N}$ for some large enough N, the composition map:

$$Ind_t = \beta \circ i! : K^0(T^*M) \cong K^0(TM) \longrightarrow \mathbb{Z}$$

is called the *topological index*. One can prove that it does not depend on the choice of the imbedding $i : M \longrightarrow \mathbb{R}^N$. The main content of the Atiyah-Singer index theorem is in fact the equality:

$$Ind_a = Ind_t : K^0(T^*M) \longrightarrow \mathbb{Z} \ ,$$

which allows computing the analytical index of an elliptic operator by a cohomological formula.

4.4 Coincidence of the Analytical and Topological Index Maps

The proof of the equality $Ind_a = Ind_t$ is based on the following two properties of the analytical index:

Property 1. For $M = pt$, the analytical index $Ind_a : K^0(pt) \longrightarrow \mathbb{Z}$ is the identity.

Property 2. For any smooth embedding $i : M \longrightarrow V$ between compact smooth manifolds, the following diagram is commutative:

$$\begin{array}{ccc} K^*(TM) & \xrightarrow{i!} & K^*(TV) \\ Ind_a \searrow & & \swarrow Ind_a \\ & \mathbb{Z} & \end{array}$$

To check that these properties imply the equality $Ind_a = Ind_t$, choose an embedding $i : M \longrightarrow \mathbb{R}^N \subset S^N = \mathbb{R}^N \cup \{\infty\}$ and denote by $j : \{\infty\} \longrightarrow S^N$ the inclusion of the point ∞. By property 2, we have for any $x \in K^0(TM)$:

$$Ind_a(x) = Ind_a(i!x) = Ind_a(j!^{-1}i!x)$$

and, since $Ind_a \circ j!^{-1} = j!^{-1}$ by property 1, and $j!^{-1}$ is just the Bott periodicity isomorphism on $K^0(\mathbb{R}^N)$, we get:

$$Ind_a(x) = j!^{-1} \circ i!(x) = Ind_t(x) \ .$$

Exercise 4.4.1. *Check directly property 1.*

To prove property 2, consider a tubular neighbourhood of M in V, which is diffeomorphic to the normal bundle N of M in V. To prove that

$$Ind_a(x) = Ind_a(i!x) \text{ for any } x \in K^0(TM),$$

one can show that it suffices to prove that $Ind_a(x) = Ind_a(j!x)$, where $j : M \longrightarrow N$ is the inclusion of the zero section. We may also replace (cf. [16]) the principal O_k-bundle $N = P \times_{O_k} \mathbb{R}^k$ by the associated sphere bundle $S_N = P \times_{O_k} S^k$, where O_k acts on S^k by trivially extending the natural representation on $\mathbb{R}^k$ to $\mathbb{R}^{k+1} = \mathbb{R}^k \times \mathbb{R}$ and then restricting to the unit sphere. In other word, we may compactify the fibre of N. If $j : M \longrightarrow S_N$ denotes the natural inclusion, we have for any $x \in K^0(TM) \cong K^0(T^*M)$:

$$j!(x) = x \otimes [D] \ ,$$

where $[D] \in K_{O_k}(T^*S^k)$ is the equivariant K-theory class of the de Rham-Hodge O_k-operator $d + d^* : \Lambda^{even} \longrightarrow \Lambda^{odd}$. Here, the product:

$$K(T^*M) \otimes K_{O_k}(T^*S^k) \longrightarrow K(T^*S_N)$$

is naturally defined by using a splitting $T^*S_N = \pi^*(T^*M) \oplus T(S_N/M)$, where $T(S_N/M) = T^*S_N/\pi^*(T^*M)$ denotes the tangent spaces along the fibres of the projection $\pi : S_N \longrightarrow M$, the obvious inclusion

$$K_{0_k}(T^*S^k) \longrightarrow K_{0_k}(P \times T^*S^k) \longrightarrow K(P \times_{O_k} T^*S^k) = K(T(S_N/M)) \ ,$$

and the external product: $K(T^*M) \otimes K(T(S_N/M)) \longrightarrow K(T^*S_N)$.
In this setting, property 2 follows from the multiplicativity property for sphere bundles:

Proposition 4.4.2. *Let S be an S^k-bundle over a compact manifold M. For any $x \in K(T^*M)$ and $[P] \in K_{O_k}(T^*S^k)$, we have:*

$$Ind_a(x.[P]) = Ind_a(x.Ind_{O_k}(P)) \ ,$$

where $Ind_{O_k}(P) \in R(O_k)$ is the equivariant index of the O_k-operator P. Here, the $R(O_k)$-module structure on $K(T^*M)$ (which is a $K(M)$-module in an obvious way) comes from the natural morphism $R(O_k) \longrightarrow K(M)$.

The equivariant index of an O_k-operator P is heuristically defined as the difference $Ind_{0_k}(P) = [Ker\ P] - [Ker\ P^*]$ of O_k-representations.

Since the equivariant index of the de Rham-Hodge O_k-operator $d + d^* : \Lambda^{even} \longrightarrow \Lambda^{odd}$ on S^k is equal to $1 \in R(O_k)$ (see for instance [16], p. 253), we get from proposition 4.4.2:

$$Ind_a(j!(x)) = Ind_a(x \otimes [D]) = Ind_a(x.Ind_{O_k}(D)) = Ind_a(x)$$

for any $x \in K^0(TM) \cong K^0(T^*M)$, and property 2 is proved. The technical proof of proposition 4.4.2 is modelled on the proof of the *multiplicativity of the analytical index:*

$$Ind_a([P] \otimes [Q]) = Ind_a(P)Ind_a(Q)\ , \tag{1}$$

for any pair of first order elliptic operators P and Q on compact manifolds M and N. Since $Ind_a([P] \otimes [Q]) = Ind_a(D)$ where D is the sharp product:

$$D = \begin{pmatrix} P \otimes 1 & -1 \otimes Q^* \\ 1 \otimes Q & P^* \otimes 1 \end{pmatrix}\ ,$$

this multiplicative property (1) is straightforward.

4.5 Cohomological Formula for the Topological Index

Let $\pi : E \longrightarrow M$ be a complex vector bundle of rank n over a smooth manifold M. Denote by i the zero section and consider the following diagram:

$$\begin{array}{ccc} K_0(M) & \xrightarrow{i!=Thom\ iso,\ in\ K-theory} & K_0(E) \\ ch\downarrow & & \downarrow ch \\ H^{ev}(M) & \xrightarrow{i!=Thom\ iso,\ in\ cohomology} & H^{ev}(E) \end{array}$$

where $i! : H^{ev}(M) \longrightarrow H^{ev}(E)$ is the inverse of the *"integration on the fibres"* $\pi! : H^{ev}(E) \longrightarrow H^{ev}(M)$, which is an isomorphism in cohomology. It turns out that this diagram is not commutative, since the cohomology class

$$\tau(E) = \pi!\, ch(i!(1)) \in H^{ev}(M)$$

is not trivial in general. This cohomology class really measures the defect of commutativity in the above diagram, since:

Proposition 4.5.1. *For any $x \in K^0(M)$, we have: $ch(i!(x)) = i!(ch(x)\ \tau(E))$.*

Exercise 4.5.2. *Check proposition 4.5.1.*

The computation of the obstruction class $\tau(E)$ follows from the formula:

$$\chi(E)\tau(E) = i_* i! \tau(E) = ch([\Lambda^{even}E] - [\Lambda^{odd}E]) ,$$

where $\chi(E)$ is the Euler class of E. If E is a complex bundle of dimension k over M, we get from a formal splitting $E \cong L_1 \oplus \ldots \oplus L_k$ of E into line bundles:

$$\chi(E)\tau(E) = \left(\prod_{i=1}^{k} c_1(L_i)\right) \tau(E) = \left(\prod_{i=1}^{k} x_i\right) \tau(E) ,$$

where $x_i = c_1(L_i)$. On the other hand, since $\Lambda^p(E \oplus F) = \bigoplus_{i+j=p} \Lambda^i(E) \otimes \Lambda^j(F)$, we get from the multiplicativity of the Chern character:

$$ch([\Lambda^{even}E] - [\Lambda^{odd}E]) = \prod_{i=1}^{k} (1 - e^{x_i}) .$$

We deduce that: $\tau(E) = \prod_{i=1}^{k} \frac{1-e^{x_i}}{x_i} = (-1)^k \prod_{i=1}^{k} \left(\frac{1-e^{-(-x_i)}}{-x_i}\right) = (-1)^{dim(E)}$ $Td_{\mathbb{C}}(\overline{E})^{-1}$, where $\overline{E}$ is the conjugate of E and the Todd class $Td_{\mathbb{C}}(E)$ is defined by the formal power series $\frac{x}{1-e^{-x}}$, i.e. $Td_{\mathbb{C}}(E) = \prod_{i=1}^{k} \left(\frac{x_i}{1-e^{-x_i}}\right)$.
We are now in position to give a cohomological formula for the topological index:

Theorem 4.5.3. *Let M be an n-dimensional compact oriented manifold without boundary. Denote by $\pi! : H^*(T^*M) \longrightarrow H^*(M)$ the integration's map (in cohomology with compact support) on the fiber of the canonical projection $\pi : T^*M \longrightarrow M$. Then, we have for any $x \in K^0(T^*M)$:*

$$Ind_t(x) = (-1)^{\frac{n(n+1)}{2}} \int_M ch_M(x) Td_{\mathbb{C}}(TM \otimes \mathbb{C}) ,$$

where $ch_M(x) = \pi! ch(x)$ is the image of $ch(x) \in H^*(T^*M)$ by $\pi!$.

Proof. In the following diagram, where N is the normal bundle to some inclusion $M \hookrightarrow \mathbb{R}^n$ as before:

$$\begin{array}{ccccccc}
K^0(TM) & \xrightarrow{i!} & K^0(N \oplus N) \cong K^0(W) & \longrightarrow & K^0(\mathbb{R}^{2N}) & \longrightarrow & K^0(p^t) = \mathbb{Z} \\
\downarrow ch & & \downarrow ch & & \downarrow ch & & id \downarrow \\
H^{ev}(TM) & \xrightarrow{i!} & H^{ev}(N \oplus N) \cong H^{ev}(W) & \longrightarrow & H^{ev}(\mathbb{R}^{2N}) & \xrightarrow{q!} & H^0(p^t) = \mathbb{Z}
\end{array}$$

the two squares on the right commute since $\tau(\mathbb{C}^N) = 1$ so that we have $Ind_t(x) = q!i!(ch(x)\tau(N \oplus N))$ for any $x \in K^0(T^*M)$. Since $TM \oplus N$ is trivial and $TM \otimes \mathbb{C}$ is self-conjugate, we get:

$$\tau(N \otimes \mathbb{C}) = \tau(TM \otimes \mathbb{C})^{-1} = (-1)^n Td_{\mathbb{C}}(TM \otimes \mathbb{C})$$

and hence:

$$Ind_t(x) = (-1)^n \int_{TM} ch(x) Td_{\mathbb{C}}(TM \otimes \mathbb{C}) \; .$$

Taking into account the difference between the orientation of TM induced by the one of M and the "almost complex" orientation of $T(TM)$, we get from the above formula by using the Thom isomorphism in cohomology:

$$Ind_t(x) = (-1)^{\frac{n(n+1)}{2}} \int_M ch_M(x) Td_{\mathbb{C}}(TM \otimes \mathbb{C}). \text{ QED}$$

One obtain various index theorems by applying the above formula to $x = \sigma$, the symbol of an elliptic pseudodifferential operator.

Exercise 4.5.4. *Show that the index of any elliptic differential operator P on an odd-dimensional compact manifold M is zero (Hint: use the Atiyah-Singer index theorem together with the formulas: $c_*[TM] = -[TM], c^*(\sigma(P)) = \sigma(P) \in K^0(T^*M)$ where c is the diffeomorphism $\xi \in TM \longrightarrow -\xi \in TM$).*

Exercise 4.5.5. *By using the Atiyah-Singer index formula for the de Rham operator on a compact oriented manifold M, prove the equality:*

$$\sum_{i=0}^{dim(M)} (-1)^i dim H^i(M, \mathbb{R}) = \int_M \chi(T_{\mathbb{C}} M) \; .$$

4.6 The Hirzebruch Signature Formula

From the Atiyah-Singer formula, we get the Hirzebruch formula:

Theorem 4.6.1. *The signature $\sigma(M)$ of any $4k$ oriented compact smooth manifold M is given by:*

$$\sigma(M) = \int_M L(M) \; ,$$

where $L(M)$ is the Hirzebruch-Pontrjagin class defined from a formal splitting $TM \otimes \mathbb{C} = \bigoplus_{i=1}^{2k} (L_i \oplus \overline{L}_i)$ into complex line bundles by: $L(M) = 2^{2k} \prod_{i=1}^{2k} \frac{c(L_i)/2}{th(c(L_i)/2)}$.

Proof. Let D_+ be the signature operator on M (cf. 2.3.2.). Since we have:

$$[\sigma(D_+)] = [(\Lambda^+(T^*M \otimes \mathbb{C}), \Lambda^-(T^*M \otimes \mathbb{C}), i(ext(\xi) - int(\xi)))] ,$$

we get from a formal splitting $TM \otimes \mathbb{C} = \bigoplus_{i=1}^{2k} (L_i \oplus \overline{L}_i)$ of $TM \otimes \mathbb{C}$ into complex line bundles, by setting $x_i = c(L_i)$:

$$\begin{aligned} ch_M(\sigma(D_+)) &= \frac{ch([\Lambda^+(T^*M \otimes \mathbb{C})] - [\Lambda^-(T^*M \otimes \mathbb{C})])}{\chi(TM)} \\ &= \prod_{i=1}^{2k} \frac{ch([\overline{L}_i] - [L_i])}{x_i} = \prod_{i=1}^{2k} \frac{e^{-x_i} - e^{x_i}}{x_i} = 2^{2k} \prod_{i=1}^{2k} \frac{e^{x_i} - e^{-x_i}}{2x_i} \\ &= 2^{2k} \prod_{i=1}^{2k} \frac{x_i/2}{th(x_i/2)} \left(\prod_{i=1}^{2k} \frac{x_i/2}{sh(x_i/2)} \right)^{-2} , \end{aligned}$$

and hence:

$$\begin{aligned} &ch_M(\sigma(D_+)) Td_{\mathbb{C}}(TM \otimes \mathbb{C}) \\ &\qquad = 2^{2k} \prod_{i=1}^{2k} \frac{x_i/2}{th(x_i/2)} \left(\prod_{i=1}^{2k} \frac{x_i/2}{sh(x_i/2)} \right)^{-2} \left(\prod_{i=1}^{2k} \frac{x_i/2}{sh(x_i/2)} \right)^{2} \\ &\qquad = 2^{2k} \prod_{i=1}^{2k} \frac{x_i/2}{th(x_i/2)} = L(M). \end{aligned}$$

Since the signature of M is equal to $Ind(D_+)$ by theorem 2.3.6, we get from the Atiyah-Singer index formula:

$$\sigma(M) = \int_M ch_M(x) Td_{\mathbb{C}}(TM \otimes \mathbb{C}) = \int_M L(M). \text{ QED}$$

5 The Index Theorem for Foliations

5.1 Index Theorem for Elliptic Families

5.1.1 Elliptic Families. Consider a smooth fibration $p : M \longrightarrow B$ with connected fiber F on a compact manifold M. For each $y \in B$, set $F_y = p^{-1}(y)$ and denote by $T_F^*(M)$ the bundle dual to the bundle $T_F(M)$ of vectors tangent to the fibres of the fibration. Let $q : T_F^*(M) \longrightarrow M$ be the projection map and consider a family $P = (P_y)_{y \in B}$ of zero order pseudodifferential operators:

$$P_y : C^\infty(F_y, E^0) \longrightarrow C^\infty(F_y, E^1)$$

on the fibres of the fibration $p : M \longrightarrow B$, where $E = E^0 \oplus E^1$ is a $\mathbb{Z}/2\mathbb{Z}$ graded hermitian vector bundle over M.

Definition 5.1.2. *The family* $(P_y)_{y\in B}$ *is said to be continuous if the map* P *defined on* $C^\infty(M,E^0)$ *by* $Pf(x) = (P_y f_y)(x)$ $(y = p(x), f_y = f|_{F_y})$ *sends* $C^\infty(M,E^0)$ *into* $C^\infty(M,E^1)$*.*

The *principal symbol* of such a continuous family $P = (P_y)_{y\in B}$ is by definition the family $\sigma(P) = (\sigma(P_y))_{y\in B}$ of the symbols of the P_y's. It can be viewed as a vector bundle morphism $\sigma(P) : q^*(E^0) \longrightarrow q^*(E^1)$. The family $P = (P_y)_{y\in B}$ is said to be *elliptic* if all the P_y's are elliptic. In this case, the principal symbol $\sigma(P)$ yields a K-theory class $[\sigma(P)] \in K^0(T_F^*M)$.

5.1.3 Analytical Index of a Family of Elliptic Operators. Let P be as above. By working locally as in the case of an elliptic operator, we can prove the existence of a continuous family $Q = (Q_y)_{y\in B}$ of zero order pseudodifferential operators such that $P_yQ_y - I = R_y$ and $Q_yP_y - I = S_y$ are continuous families of infinitely smoothing operators. In particular, the family $P = (P_y)_{y\in B}$ gives rise to a continuous field of Fredholm operators $P_y : L^2(F_y,E^0) \longrightarrow L^2(F_y,E^1)$, and the index $Ind(P) \in K_0(B)$ of this family of Fredholm operators makes sense by theorem 3.2.1.

5.1.4 Topological Index of a Family of Elliptic Operators. On the other hand, the principal symbol $\sigma(P)$ yields a K-theory class $[\sigma(P)] \in K^0(T_F^*M)$. To define a *topological index* $Ind_{top} : K^0(T_F^*M) \longrightarrow K^0(B)$, let us choose a smooth map $f : M \longrightarrow B \times \mathbb{R}^N$ which reduces for any $y \in B$ to a smooth embedding $f_y : F_y \longrightarrow \{y\} \times \mathbb{R}^N$. Such a map gives rise to a smooth embedding $f_* : T_FM \longrightarrow B \times T\,\mathbb{R}^N$ with normal bundle $N \oplus N$, where N_y is the normal bundle of $f_y(F_y)$ in $\{y\} \times \mathbb{R}^N$, pulled back to M. Since $N_y \oplus N_y \simeq N_y \otimes \mathbb{C}$, we have a well defined Gysin map:

$$\begin{aligned} p! : K^0(T_F^*M) &\longrightarrow K^0(N\otimes\mathbb{C}) \longrightarrow K^0(B\times T\,\mathbb{R}^N) \\ &= K^0(B\times\mathbb{R}^{2N}) \longrightarrow K^0(B)\ , \end{aligned}$$

where the last map on the right is the Bott periodicity isomorphism. We call *topological index* the map:

$$Ind_{top} = p! : K^0(T_F^*M) \longrightarrow K^0(B)\ ,$$

which is well defined and does not depend on the choice of $f : M \longrightarrow B\times\mathbb{R}^N$. At this point, it is almost clear that the proof of the Atiyah-Singer index theorem for elliptic operators on compact manifolds extends to the framework of fibrations to give the following index theorem:

Theorem 5.1.5. *Let* $p : M \longrightarrow B$ *be a smooth fibration with fiber* F *on a compact manifold* M*, and* $P = (P_y)_{y\in B}$ *be a continuous family of elliptic zero order pseudodifferential operators on the fibres* $F_y = p^{-1}(y)$*. Then, we have:*

(i) $Ind(P) = Ind_{top}([\sigma(P)]) \in K^0(B)$;
(ii) $ch(Ind(P)) = (-1)^{\frac{n(n+1)}{2}} \pi! \left(ch\left([\sigma(P)]\right) Td_{\mathbb{C}}(TM)\right) \in H^*(B)$,

where $n = dim(F)$ *is the dimension of the fibre and* $\pi : T_F^* M \longrightarrow B$ *is the natural projection.*

For a detailed proof of this result, see [2].

5.2 The Index Theorem for Foliations

5.2.1 Foliations. Let M be a smooth n-dimensional compact manifold. Recall that a smooth p-dimensional subbundle F of TM is called *integrable* if every $x \in M$ is contained in the domain U of a submersion $p : U \longrightarrow \mathbb{R}^{n-p}$ such that $F_y = Ker(p_*)_y$ for any $y \in U$. A p-dimensional *foliation* F on M is given by an integrable p-dimensional subbundle F of TM. We call *leaves* of the foliation (M, F) the maximal connected submanifolds L of M such that $T_x(L) = F_x$ for any $x \in L$. In any foliation (M, F), the equivalence relation on M corresponding to the partition into leaves is locally trivial, i.e. every point $x \in M$ has a neighborhood U with local coordinates $(x^1, \ldots, x^n) : U \longrightarrow \mathbb{R}^n$ such that the partition of U into connected components of leaves corresponds to the partition of $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$ into the plaques $\mathbb{R}^p \times \{y\}$. We shall called $p = dim(F)$ the *dimension* of the foliation, and $q = n - dim(F) = codim(F)$ the *codimension* of F. For instance, any smooth fibration $p : M \longrightarrow B$ with connected fiber on a compact manifold defines a foliation on M whose leaves are the fibres $p^{-1}(y), y \in B$. If θ is an irrational number, the flow of the differential equation $dy - \theta dx = 0$ on the two-dimensional torus $M = \mathbb{R}^2/\mathbb{Z}^2$ defines a codimension 1 foliation F_θ on M called the irrational *Kronecker foliation.*

Exercise 5.2.2. *Show that each leaf of the irrational Kronecker foliation* F_θ *on the* 2*-dimensional torus* $M = \mathbb{R}^2/\mathbb{Z}^2$ *is non compact and dense. Let* M/F *be the quotient of* M *by the equivalence relation corresponding to the partition into leaves. Show that the quotient topology on* M/F *is trivial (the only open subsets are* M/F *and* $\varnothing$*).*

5.2.3 Holonomy of a Foliation. Let $\gamma : [0,1] \longrightarrow M$ be a continuous path on a leaf L of F, and consider two q-dimensional submanifolds T, T' transverse to the foliation and whose interiors contain respectively the source $x = \gamma(0)$ and the range $y = \gamma(1)$ of γ. By "following the leaves" through some small enough tubular neighborhood of $\gamma([0,1])$, we get from γ a local diffeomorphism $\varphi_\gamma : Dom(\varphi_\gamma) \subset T \longrightarrow T'$ with $x = s(\gamma) \in Dom(\varphi_\gamma)$. The *holonomy germ* of γ is by definition the germ $h(\gamma) = [\varphi_\gamma]_x$ of φ_γ at $x = s(\gamma)$. Two paths $\gamma_1, \gamma_2 : [0,1] \longrightarrow L$ having the same source $x = s(\gamma_1) = s(\gamma_2)$ and the same range $y = r(\gamma_1) = r(\gamma_2)$ are said holonomy equivalent (and we write $\gamma_1 \sim \gamma_2$) if there exist transverse submanifolds T at x and T' at y such that

$h(\gamma_1) = h(\gamma_2)$. We thus define an equivalence relation on the set of all paths drawn on the leaves. The *holonomy groupoid* of the foliation is by definition the set G of all equivalence classes. Any $\gamma \in G$ is thus the holonomy class of a path on some leaf, with source x (denoted by $s(\gamma)$) and range y (denoted by $r(\gamma)$). For any $x \in M$, we shall set $G_x = \{\gamma \in G | r(\gamma) = x\}$. The composition of paths induces a natural structure of groupoid on G, and it can be shown that G has the structure of a smooth (possibly non Hausdorff) manifold. For more information on the holonomy groupoïd G of (M, F), see [3].

Exercise 5.2.4. *Show that the holonomy groupoid of the Kronecker foliation* F_θ *on the* 2*-dimensional torus* $M = \mathbb{R}^2/\mathbb{Z}^2$ *identifies with* $M \times \mathbb{R}$*. Describe its groupoid structure and its smooth structure.*

5.2.5 C*-Algebra of a Foliation. For a fibration, the space of leaves is a nice compact space which identifies with the base of the fibration. However, for a foliation with dense leaves such as the Kronecker foliation F_θ, the space of leaves can be very complicated although the local picture is that of a fibration. A. Connes [3] suggested describing the topology of the "leafspace" M/F of any foliation (M, F) by a C^*-algebra $C^*(M, F)$ whose noncommutativity tells us how far the foliation lies from a fibration. This C^*-algebra, which describes the non commutative space M/F, is obtained by quantization of the holonomy groupoid. More precisely, $C^*(M, F)$ is defined as the minimal C^*-completion of the algebra of continuous compactly supported sections $C_c(G, \Omega^{1/2})$ of the bundle $\Omega^{1/2}_\gamma = \Omega^{1/2}_{s(\gamma)} \otimes \Omega^{1/2}_{r(\gamma)}$ of half densities along the leaves of the foliation, endowed with the following laws:

$$\begin{cases} (f * g)(\gamma) = \int_{\gamma_1\gamma_2=\gamma} f(\gamma_1)g(\gamma_2) \\ f^*(\gamma) = \overline{f(\gamma^{-1})}\,. \end{cases}$$

When the foliation comes from a fibration $p : M \longrightarrow B$, the C^*-algebra $C^*(M, F)$ identifies with $C(B) \otimes K(L^2(F))$, where $K(H)$ denotes the algebra of compact operators on the Hilbert space H. In the case of the Kronecker foliation F_θ, we have $C^*(\mathbf{T}^2, F_\theta) \simeq A_\theta \otimes K(H)$ where H is a separable infinite dimensional Hilbert space and A_θ the irrational rotation algebra generated by two unitaries U and V in H satisfying the commutation relation $VU = exp(2i\pi\theta)UV$.

5.2.6 Elliptic Operators Along the Leaves of a Foliation. Let (M, F) be a smooth foliation on a compact manifold M and E^0, E^1 two smooth complex vector bundles over M. A *differential operator elliptic along the leaves* of (M, F) acting from the sections of E^0 to the sections of E^1 is a differential operator $D : C^\infty(M, E^0) \longrightarrow C^\infty(M, E^1)$ which restricts to the leaves and is elliptic when restricted to the leaves. Its principal symbol $\sigma(D)(x, \xi) \in Hom(E^0_x, E^1_x)$ is thus invertible for any non zero $\xi \in F^*_x$ and yields a K-theory class

$$[\sigma(D)] \in K^0(F^*) \ .$$

Since a foliation is locally a fibration, the notion of elliptic pseudodifferential operator along the leaves of (M, F) can be defined in a natural way. As in the case of fibrations, it generalizes the notion of differential elliptic operator along the leaves.

5.2.7 Analytical Index of an Operator Elliptic Along the Leaves. Assume for simplicity that the foliation (M, F) has no holonomy and consider an elliptic pseudodifferential operator P of order zero along the leaves of (M, F), acting from the sections of E^0 to the sections of E^1. The restriction P_L of P to the leaf L is a bounded operator in the Hilbert space $H_L = L^2(L, E^0 \oplus E^1)$. Moreover, the family $(P_L)_{L \in M/F}$ yields in a natural way an endomorphism of a Hilbert $C^*(M, F)$-module that we now describe. Let $\mathbf{E}$ be the Hilbert completion of the linear span of the 1/2 sections of the field $H_x = L^2(G_x, E^0 \oplus E^1)$ that have the from $x \longrightarrow \int_{G_x} (\xi \circ \gamma) f(\gamma)$ where ξ is a basic 1/2 section of H and $f \in C^*(M, F)$ an element with a square integrable restriction to G_x for any $x \in M$. There is an obvious structure of $C^*(M, F)$-module on $\mathbf{E}$. Moreover, since the coefficient $< \xi, \eta > (\gamma) =< \xi_{r(\gamma)} \circ \gamma, \eta_{s(\gamma)} >$ is required to be in $C^*(M, F)$ for any pair ξ, η of basic 1/2 sections of H, the $C^*(M, F)$-valued scalar product $< \xi, \eta >$ of two elements in $\mathbf{E}$ is well defined, and it is straightforward to check that $\mathbf{E}$ is a Hilbert $C^*(M, F)$-module (see [4] for more details).

Exercise 5.2.8. *Show that the family $(P_L)_{L \in M/F}$ yields an endomorphism P of the $C^*(M, F)$-module* $\mathbf{E}$.

By using the local construction of a parametrix for families of elliptic operators, one can show as in the case of families the existence of an endomorphism Q of $\mathbf{E}$ such that $PQ - I \in K(\mathbf{E})$ and $QP - I \in K(\mathbf{E})$. It follows that P is a generalized $C^*(M, F)$-Fredholm operator, and hence has an analytical index:

$$Ind_{C^*(M,F)}(P) \in K_0(C^*(M, F)) \ .$$

5.2.9 The Index Theorem for Foliations. To compute $Ind_{C^*(M,F)}(P)$ we shall define, as in the case of fibrations, a *topological index:*

$$Ind_t : K^0(F^*) \longrightarrow K_0(C^*(M, F))$$

by choosing an auxiliary embedding $i : M \longrightarrow \mathbb{R}^{2m}$. Let N be the total space of the normal bundle to the leaves (i.e. $N_x = i_*(F_x)^\perp$ for the Euclidean metric) and consider the product manifold $M \times \mathbb{R}^{2m}$ foliated by the $L \times \{t\}$'s (L = leaf of $F, t \in \mathbb{R}^{2m}$). The map $(x, \xi) \in N \longrightarrow (x, i(x) + \xi) \in M \times \mathbb{R}^{2m}$

sends a small neighborhood of the zero section of N into an open transversal T to the foliation $\widetilde{F}$ on $M \times \mathbb{R}^{2m}$. Putting T inside a small open tubular neighborhood Ω in $M \times \mathbb{R}^{2m}$ we get, from the inclusion of $C^*(\Omega, \widetilde{F}) \cong C_0(T) \otimes K(H)$ into $C^*(M \times \mathbb{R}^{2m}, \widetilde{F})$, a K-theory map:

$$K_0(C_0(T)) = K_0(C_0(T) \otimes K(H)) = K_0(C^*(\Omega, \widetilde{F})) \longrightarrow K_0(C^*(M \times \mathbb{R}^{2m}, \widetilde{F})) .$$

Since we have, by Bott periodicity:

$$K_0(C^*(M \times \mathbb{R}^{2m}, \widetilde{F})) = K_0(C^*(M, F) \otimes C_0(\mathbb{R}^{2m})) \cong K_0(C^*(M, F)) ,$$

we get by composition a K-theory map:

$$Ind_t : K^0(F^*) \cong K^0(N) \cong K^0(T) = K_0(C_0(T)) \longrightarrow K_0(C^*(M, F)) .$$

This map, which does not depend on the choices made, is called the *topological index.*

Theorem 5.2.9. (Index theorem for foliations). *For any zero order elliptic pseudodifferential operator P along the leaves of a foliation (M, F) with principal symbol $\sigma(P) \in K^0(F^*)$, we have:*

$$Ind_{C^*(M,F)}(P) = Ind_t(\sigma(P)) \in K_0(C^*(M, F)) .$$

For a proof of this result, see [6]. When the foliation (M, F) has an invariant transverse measure Λ, there exist a trace τ_Λ on $C^*(M, F)$ which yields a map from the finite part of $K_0(C^*(M, F))$ generated by trace-class projections in $M_n(C^*(M, F))$ to $\mathbb{R}$. This trace is given by:

$$\tau_\Lambda(k) = \int_{M/F} Trace(k_L) d\Lambda(L) \ (k \in C_c^\infty(G, \Omega^{1/2})) ,$$

where $Trace(k_L)$ is viewed as a measure on the leaf manifold. For a zero order elliptic pseudodifferential operator P along the leaves of (M, F), one can show that $Ind_{C^*(M,F)}(P)$ belongs to the finite part of $K_0(C^*(M, F))$. In this case, we get from theorem 5.2.9 (see [3] for the original proof):

Theorem 5.2.10. (Measured index theorem for foliations). *Let (M, F) be a p-dimensional smooth foliation on a compact manifold M. Assume that (M, F) has a holonomy invariant transverse measure Λ and denote by $[\Lambda]$ the associated Ruelle-Sullivan current. For any zero order elliptic pseudodifferential operator P along the leaves with principal symbol $\sigma(P) \in K^0(F^*)$, we have:*

$$\tau_\Lambda(Ind_{C^*(M,F)}(P)) = (-1)^{\frac{p(p+1)}{2}} < ch(\sigma(P)) Td_{\mathbb{C}}(F_{\mathbb{C}}), [\Lambda] > .$$

For two-dimensional leaves, this theorem gives in the case of the leafwise de Rham operator:

$$\beta_0 - \beta_1 + \beta_2 = \frac{1}{2\pi} \int K d\Lambda \,,$$

where the β_i are the Betti numbers of the foliation and K is the Gaussian curvature of the leaves. If the set of compact leaves is negligible, we have $\beta_0 = \beta_2 = 0$, and the above relation shows that $\int K d\Lambda \leq 0$. It follows that the condition $\int K d\Lambda > 0$ implies the existence of compact leaves.

References

1. M.F. Atiyah. *$K - Theory$*. Benjamin Press, New-York. 1967. 202
2. M.F. Atiyah and I.M. Singer. *The index of elliptic operators IV.* Ann. of Math. 93 (1971), 119–138. 224
3. A. Connes. *Sur la théorie non commutative de l'intégration.* In Lectures Notes in Mathematics. 725. Editor Pierre de la Harpe (1979), 19–143. 225, 227
4. A. Connes. *A survey of foliations and operator algebras.* Operator algebras and apllications, Part 1. Proc. Sympos. Pure Math., Amer. Math. Soc., Providence, R.I. 38 (1982), 521–628. 226
5. A. Connes. *An analogue of the Thom isomorphism for cross products of a C^*-algebra by an action of* $\mathbb{R}$. Adv. in Math. 39 (1981), 31–55. 212
6. A. Connes and G. Skandalis. *The longitudinal index theorem for foliations.* Publ. RIMS, 20 (No 6), 1984, 1139–1183. 227
7. T. Fack and G. Skandalis. *Connes' Analogue of the Thom Isomorphism of the Kasparov Groups.* Invent. Math. 64 (1981), 7–14. 212
8. P.B. Gilkey. *Invariance Theory, The Heat Equation, And the Atiyah-Singer Index Theorem.* Mathematics Lectures Series, 11, Publish or Perish, Inc. 1984. 193
9. D. Husemoller. *Fibre bundles.* 2^d Edition. Graduate Texts in Mathematics (20). Springer-Verlag. New York - Heidelberg - Berlin, 1975. 208
10. M. Karoubi. *K-Theory.* An introduction. Grundlehren der mathematischen Wissenschaften, 226 – Springer-Verlag. Berlin-Heidelberg – New York, 1978. 202, 206, 217
11. G.G. Kasparov. *Hilbert C^*-modules. Theorems of Stinespring and Voiculescu.* J. Op. Theory, 4 (1980), 133–150. 212
12. G.G. Kasparov. *The operator K-functor and extensions of C^*-algebras.* Math. U.S.S.R. – Izv. 16 (1981), 513–572. 214
13. G.G. Kasparov. *Equivariant KK-theory and the Novikov conjecture.* Invent. Math. 91 (1988), 147–201.
14. N. Kuiper. *The homotopy type of the unitary group of Hilbert space.* Topology 3 (1965), 19–30. 203
15. S. Lang, *Real Analysis.* Addison-Wesley Publishing Company. 1983. 186
16. H.B. Lawson, M-L. Michelson. *Spin Geometry.* Princeton Mathematical Series, 38. Princeton University Press. 1989. 189, 190, 192, 218, 219
17. G.K. Pedersen. *C^*-algebras and their Automorphism groups.* London Mathematical Society. Monographs no 14. Academic Press. London – New York – San Francisco. 1979. 209

18. J. Roe. *Elliptic operators, topology and asymptotic methods.* Pitman Research Notes in Mathematics Series (179). Longman Scientific and Technical. Longman Group UK Limited. 1988. 206
19. B. Simon. *Trace ideals and their applications.* London Math. Soc. Lecture Note Series, 35. Cambridge University Press. 1979. 195