Università degli Studi di Torino
Scuola di Dottorato

UNIVERSITÀ DEGLI STUDI DI TORINO

D-branes and Deep Learning:
Theoretical and Computational Aspects In String Theory

Riccardo Finotello

Università degli Studi di Torino
Scuola di Dottorato

Dottorato in Fisica ed Astrofisica

D-branes and Deep Learning: Theoretical and Computational Aspects In String Theory

Riccardo Finotello

Advisor: Igor Pesando

Abstract

We present topics of (semi-)phenomenological relevance in string theory ranging from particle physics amplitudes and Big Bang-like singularities to the study of state-of-the-art deep learning techniques for string compactifications based on recent advancements in artificial intelligence.

We show the computation of the leading contribution to amplitudes in the presence of non Abelian twist fields in intersecting D-branes scenarios in non factorised tori. This is a generalisation to the current literature which mainly covers factorised internal spaces. We also study a new method to compute amplitudes in the presence of an arbitrary number of spin fields introducing point-like defects on the string worldsheet. The procedure can then be treated as an alternative computation with respect to bosonization and approaches based on the Reggeon vertex. We then present an analysis of Big Bang-like cosmological divergences in string theory on time-dependent orbifolds. We show that divergences are not due to gravitational feedback but to the lack of an underlying effective field theory. We also introduce a new orbifold structure capable of fixing the issue and reinstate a distributional interpretation to field theory amplitudes.

We finally present a new artificial intelligence approach to algebraic geometry and string compactifications. We compute the Hodge numbers of Complete Intersection Calabi–Yau 3-folds using deep learning techniques based on computer vision and object recognition techniques. We also include a methodological study of machine learning applied to data in string theory: as in most applications machine learning almost never relies on the blind application of algorithms to the data but it requires a careful exploratory analysis and feature engineering. We thus show how such an approach can help in improving results by processing the data before utilising them. We then show that deep learning the configuration matrix of the manifolds reaches the highest accuracy in the task with smaller networks, less parameters and less data. This is a novel approach to the task: differently from previous attempts we focus on using convolutional neural networks capable of reaching higher accuracy on the predictions and ensuring phenomenological relevance to results. In fact parameter sharing and concurrent scans of the configuration matrix retain better generalisation properties and adapt better to the task than fully connected networks.

Acknowledgements

This manuscript signals the end of a fascinating journey I went through together with great people who supported me when I was at my worst and built in me the motivation to keep going. They were capable of giving me consciousness of what I was doing and guided me to this very moment. Words will certainly not render justice to their role, but I shall try nonetheless.

An important mention goes to my advisor Igor who taught me most of what I know now: with the trust he put in me, he made me feel what research is like, and introduced me to the dedication and determination needed to succeed. I most certainly would not be here without him, and this I shall always remember. I cannot help expressing the deepest gratitude to Harold: even though he was not my advisor, he trusted me, he guided me towards the best of all the worlds, and he supported me in every (im)possible way to the point words will never be enough. He is surely a brilliant scientist and a great collaborator, but if he is not first of all a friend then I do not know who can be. I also wish to thank Marco and Carlo for all the support they gave me whenever they could and the advice they provided at all times.

Life as a graduate student would have been incredibly tough were it not for the great people who were with me all the time. I was not yet officially inside my Ph.D. programme when Alberto welcomed me to the office and showed me how to survive as a student and get the most from the experience as a person. Even though he will not admit it, Riccardo helped me to discover dedication in the darkest moments, while contextually Giovanni taught me how to bring light and humanity in those, in the face of adversities. In all this, with his silence *à la Piemontese*, Francesco represented the needed balance. Life in the office would not have been the same were it not for the people who supported me in this adventure, from the overly calm Chiara and the elegantly mannered Michael to the extremely excitable Kostas, to colleagues such as Andrea with whom I was honoured to collaborate directly. Students and researchers met during these three years were absolutely great. Confrontation with them helped me discover a vast number of interconnected topics and fascinating subtleties, and it allowed me to meet and get to know fellow students such as Paolo, trusted and worthy accomplice in many adventures.

Even with all the help of an incredibly large number of great people, I do not think I could be here writing these without the support of Alessandra and my family. They had the most difficult job of all: they had to put up with me at my worst and to help me get through the difficulties. My parents and my brother showed me all the affection they possibly could, they provided me with everything I needed, and they unconditionally made sure I could be as happy as possible to succeed in what I really wanted to do. Not only did they succeed, they were absolutely amazing. Alessandra is my hero: though she went through the toughness of her Ph.D. programme as well, she was able to be there for me every time I needed confrontation or consolation, every time I wanted to share good or bad news, every single time I was happy for something or in distress. Alessandra is the strongest person I know. There are definitely no words to express how grateful I am to all of them, but I hope I can repay them in time as I could not be more joyfully blessed than I am right now.

To all of you: thank you.

Riccardo Finotello

Table of Contents

Outline

This thesis follows my research work as a Ph.D. student and candidate at the *Università degli Studi di Torino* in Italy. During my programme I mainly dealt with the topic of string theory and its relation with a viable formulation of phenomenology in this framework. I covered mathematical aspects related to amplitudes in intersecting D-branes scenarios and in the presence of defects on the worldsheet, but I also worked on computational issues such as the application of recent deep learning and machine learning techniques to the compactification of the extra-dimensions of superstrings. In this manuscript I present the original results obtained over the course of my Ph.D. programme. They are mainly based on published work [1], [2] and preprints [3]–[5]. However I also include some hints to future directions to cover which might expand the work shown here. The thesis is organised in three main parts plus a fourth with appendices and notes.

Part I of the manuscript is dedicated to set the stage for the entire discussion and to present mathematical tools used to compute amplitudes with (semi-)phenomenological relevance in string theory. Namely it starts with an introduction on conformal symmetry (clearly focusing only on aspects relevant to the discussion as many reviews on the subject have already been written) and the role of compactification and D-branes in replicating results obtained in particle physics. Then the analysis of a specific setup involving angled D6-branes intersecting in non factorised internal space is presented.[1] Here a general framework to deal with SO(4) rotated D-branes is introduced alongside the computation of the leading term of amplitudes involving an arbitrary number of non Abelian twist fields located at their intersection, that is the exponential contribution of the classical bosonic string in this geometry. Finally point-like defects in the time direction of the (super)string worldsheet are introduced and the propagation of fermions on such surface studied in detail. In this setup the stress-energy tensor has a time dependence but it still respects the usual operator product expansion. Thus the theory is still conformal though time dependence is due to the defects where spin fields are located. Through the study of the operator algebra the computation of amplitudes in the presence of spin fields and matter fields is eventually displayed by means of a method alternative to the usual bosonization, but which might be expanded also to twist fields and to more general configurations (e.g. non Abelian spin fields).

Part II deals with cosmological singularities in string and field theory. The main focus is on time-dependent orbifolds as simple models of Big Bang-like singularities in string theory: after a brief introduction on the concept of orbifold from the mathematical and the physical points of view, the Null Boost Orbifold is introduced as a first example. Differently from what usually referred in the literature, the divergences appearing in the amplitudes are not a consequence of gravitational feedback, but they are present also at the tree level of open string amplitudes. The source of the divergences are shown in string and field theory amplitudes due to the presence of the compact dimension and its conjugated momentum which prevents the interpretation of the amplitudes even as a distribution. In fact the introduction of a non compact direction of motion on the orbifold restores the physical significance of the amplitude, hence the origin of the divergences comes from geometrical aspects of the orbifold models. Namely it is hidden in contact terms and interaction with massive string states which are no longer spectators, thus invalidating the usual approach with the inheritance principle.

Part III is dedicated to state-of-the-art application of deep learning techniques to the field

[1] For instance this is a generalisation of the typical setup involving D6-branes filling entirely the 4-dimensional spacetime and embedded as lines in $T^2 \times T^2 \times T^2$, where the possible rotations performed by the D-branes are parametrised by Abelian SO(2) $\simeq$ U(1) rotations.

of string theory compactifications. The Hodge numbers of Complete Intersection Calabi–Yau 3-folds are computed through a rigorous data science and machine learning analysis. In fact the blind application of neural networks to the configuration matrix of the manifolds can be improved by exploratory data analysis and feature engineering, from which it is possible to infer behaviour and relations of topological quantities invisibly hidden in the raw data.[2] Deep learning techniques are then applied to the configuration matrix of the manifolds to obtain the Hodge numbers as a regression task.[3] A new neural network architecture based on recent computer vision advancements in the field of computer science is eventually introduced: it utilises parallel convolutional kernels to extract the Hodge numbers from the configuration matrix and it reaches near perfect accuracy on the prediction of $h^{1,1}$. Such model also leads to good preliminary results for $h^{2,1}$ which has been mostly ignored by previous attempts.

[2] Through accurate data analysis it is possible to reach the same performance of neural networks using simpler algorithms (in some cases even trivial linear models can reach the same accuracy in the predictions).

[3] Many previous approaches have proposed classification tasks to get the best performance out of machine learning models. This however implies specific knowledge of the definition interval of the Hodge numbers and does not generalise well to unknown samples.

Part I

Conformal Symmetry and Geometry of the Worldsheet

1 Introduction

In this first part we focus on aspects of string theory directly connected with its worldsheet description and symmetries. The underlying idea is to build technical tools to address the study of viable phenomenological models in this framework. The construction of realistic string models of particle physics is key to better understanding the nature of a theory of everything such as string theory.

As a first test of validity, the string theory should properly extend the known Standard Model (SM) of particle physics which is arguably one of the most experimentally backed theoretical frameworks in modern physics. In particular its description in terms of fundamental strings should be able to include a gauge algebra isomorphic to the algebra of the group

$$\mathrm{SU}(3)_{\mathrm{C}} \otimes \mathrm{SU}(2)_{\mathrm{L}} \otimes \mathrm{U}(1)_{\mathrm{Y}} \tag{1.1}$$

in order to reproduce known results. For instance a good string theory could provide a unified framework by predicting the existence of a larger gauge group containing the SM as a subset. In what follows we deal with the definition of mathematical tools to compute amplitudes to be used in phenomenological calculations related to the study of particles.

In this introduction we present instruments used throughout the manuscript as many aspects are strongly connected and their definitions are interdependent. In particular we recall some results on the symmetries of string theory and how to recover a realistic description of 4-dimensional physics.

1.1 Properties of String Theory and Conformal Symmetry

Strings are extended one-dimensional objects. They are curves in spacetime parametrized by a coordinate $\sigma \in [0, \ell]$. When propagating they span a two-dimensional surface, the *worldsheet*, described by the position of the string at given values of σ at a time τ, i.e. $X^\mu(\tau, \sigma)$ where $\mu = 0, 1, \ldots, D-1$ indexes the coordinates. Such surface can have different topologies according to the nature of the object propagating in spacetime: strings can be *closed* if $X^\mu(\tau, 0) = X^\mu(\tau, \ell)$ or *open* if the endpoints in $\sigma = 0$ and $\sigma = \ell$ do not coincide.

1.1.1 Action Principle

As the action of a point particle is proportional to the length of its trajectory (its *worldline*), the same object for a string is proportional to the area of the worldsheet in the original formulation by Nambu and Goto. The solutions of the classical equations of motion (E.O.M.) are therefore strings spanning a worldsheet of extremal area. While Nambu and Goto's formulation is fairly direct in its definition, it is usually best to work with Polyakov's action [6]:

$$S_P[\gamma, X] = -\frac{1}{4\pi\alpha'} \int_{-\infty}^{+\infty} \mathrm{d}\tau \int_0^{\ell} \mathrm{d}\sigma \, \sqrt{-\det \gamma(\tau, \sigma)} \, \gamma^{\alpha\beta}(\tau, \sigma) \, \partial_\alpha X^\mu(\tau, \sigma) \, \partial_\beta X^\nu(\tau, \sigma) \, \eta_{\mu\nu}. \tag{1.2}$$

The E.O.M. for the string $X^\mu(\tau,\sigma)$ are therefore:

$$\frac{1}{\sqrt{-\det\gamma}}\,\partial_\alpha\Big(\sqrt{-\det\gamma}\,\gamma^{\alpha\beta}\,\partial_\beta X^\mu\Big)=0,\qquad \mu=0,1,\dots,D-1,\qquad \alpha,\,\beta=0,1. \tag{1.3}$$

In this formulation $\gamma_{\alpha\beta}$ are the components of the worldsheet metric. As there are no derivatives of $\gamma_{\alpha\beta}$, the E.O.M. of the metric is a constraint ensuring the equivalence of Polyakov's and Nambu and Goto's formulations. In fact

$$\frac{\delta S_P[\gamma,\,X]}{\delta\gamma^{\alpha\beta}}=-\frac{1}{4\pi\alpha'}\sqrt{-\det\gamma}\left(\partial_\alpha X\cdot\partial_\beta X-\frac{1}{2}\,\gamma_{\alpha\beta}\,\gamma^{\lambda\rho}\,\partial_\lambda X\cdot\partial_\rho X\right)=0 \tag{1.4}$$

implies

$$S_P[\gamma,\,X]\Big|_{\frac{\delta S_P[\gamma,\,X]}{\delta\gamma^{\alpha\beta}}=0}=-\frac{1}{2\pi\alpha'}\int_{-\infty}^{+\infty}\mathrm{d}\tau\int_0^{\sigma}\mathrm{d}\sigma\,\sqrt{\dot X\cdot\dot X-X'\cdot X'}=S_{NG}[X], \tag{1.5}$$

where $S_{NG}[X]$ is the Nambu–Goto action of the classical string, and $\dot X=\partial_\tau X$ and $X'=\partial_\sigma X$.

The symmetries of $S_P[\gamma,\,X]$ are key to the success of the string theory framework [7]. Specifically (1.2) displays symmetries under:

- D-dimensional Poincaré transformations

$$\begin{aligned}X'^\mu(\tau,\sigma)&=\Lambda^\mu{}_\nu\,X^\nu(\tau,\sigma)+c^\mu,\\ \gamma'_{\alpha\beta}(\tau,\sigma)&=\gamma_{\alpha\beta}(\tau,\sigma)\end{aligned} \tag{1.6}$$

  where $\Lambda\in\mathrm{SO}(1,\,D-1)$ and $c\in\mathbb{R}^D$,

- 2-dimensional diffeomorphisms

$$\begin{aligned}X'^\mu(\tau',\sigma')&=X^\mu(\tau,\sigma)\\ \gamma'_{\alpha\beta}(\tau',\sigma')&=\frac{\partial\sigma'^\lambda}{\partial\sigma^\alpha}\,\frac{\partial\sigma'^\rho}{\partial\sigma^\beta}\,\gamma_{\lambda\rho}(\tau,\sigma)\end{aligned} \tag{1.7}$$

  where $\sigma^0=\tau$ and $\sigma^1=\sigma$,

- Weyl transformations

$$\begin{aligned}X'^\mu(\tau',\sigma')&=X^\mu(\tau,\sigma)\\ \gamma'_{\alpha\beta}(\tau,\sigma)&=e^{2\omega(\tau,\sigma)}\,\gamma_{\alpha\beta}(\tau,\sigma)\end{aligned} \tag{1.8}$$

  for an arbitrary function $\omega(\tau,\sigma)$.

Notice that the last is not a symmetry of the Nambu–Goto action and it only appears in Polyakov's formulation.

1.1.2 Conformal Invariance

The definition of the 2-dimensional stress-energy tensor is a direct consequence of (1.4) [8]. In fact the classical constraint on the tensor is simply

$$\mathcal{T}_{\alpha\beta} = \frac{4\pi}{\sqrt{-\det\gamma}}\frac{\delta S_P[\gamma,\, X]}{\delta\gamma^{\alpha\beta}} = -\frac{1}{\alpha'}\left(\partial_\alpha X\cdot\partial_\beta X - \frac{1}{2}\gamma_{\alpha\beta}\,\gamma^{\lambda\rho}\,\partial_\lambda X\cdot\partial_\rho X\right) = 0. \tag{1.9}$$

While its conservation $\nabla^\alpha T_{\alpha\beta} = 0$ is somewhat trivial, Weyl invariance also ensures the vanishing trace of the tensor

$$\operatorname{tr}\mathcal{T} = \mathcal{T}^\alpha{}_\alpha = 0. \tag{1.10}$$

In other words, the $(1+1)$-dimensional theory of massless scalars X^μ in (1.2) is *conformally invariant* (for review and details see [9]–[11]).

Using the invariances of the actions we set $\gamma_{\alpha\beta}(\tau,\sigma) = e^{\phi(\tau,\sigma)}\,\eta_{\alpha\beta}$, known as *conformal gauge* where $\eta_{\alpha\beta} = \operatorname{diag}(-1,1)$. This gauge choice is however preserved by the residual *pseudo-conformal* transformations

$$\tau \pm \sigma = \sigma_\pm \quad\mapsto\quad f_\pm(\sigma_\pm), \tag{1.11}$$

where $f_\pm$ is an arbitrary function of its argument (the subscript $\pm$ distinguishes the combination of the variables τ and σ in it).

It is natural to introduce a Wick rotation $\tau_E = i\tau$ and the complex coordinates $\xi = \tau_E + i\sigma$ and $\overline{\xi} = \xi^*$. The transformation maps the Lorentzian worldsheet to a new surface: an infinite Euclidean strip for open strings or a cylinder for closed strings. In these terms, the vanishing trace of the stress-energy tensor translates to

$$\mathcal{T}_{\xi\overline{\xi}} = 0, \tag{1.12}$$

while its conservation $\partial^\alpha\mathcal{T}_{\alpha\beta} = 0$ becomes:[4]

$$\partial_{\overline{\xi}}\mathcal{T}_{\xi\xi}(\xi,\,\overline{\xi}) = \partial_\xi\overline{\mathcal{T}}_{\overline{\xi}\,\overline{\xi}}(\xi,\,\overline{\xi}) = 0. \tag{1.13}$$

The last equation finally implies

$$\mathcal{T}_{\xi\xi}(\xi,\,\overline{\xi}) = \mathcal{T}_{\xi\xi}(\xi) = \mathcal{T}(\xi), \qquad \overline{\mathcal{T}}_{\overline{\xi}\,\overline{\xi}}(\xi,\,\overline{\xi}) = \overline{\mathcal{T}}_{\overline{\xi}\,\overline{\xi}}(\overline{\xi}) = \overline{\mathcal{T}}(\overline{\xi}), \tag{1.14}$$

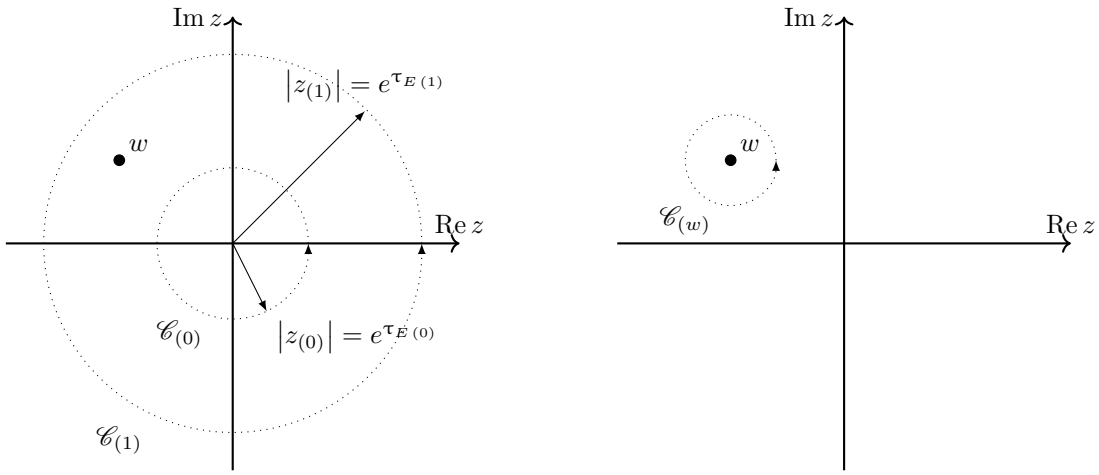
which are respectively the holomorphic and the anti-holomorphic components of the stress energy tensor.

The previous properties define what is known as a two-dimensional *conformal field theory* (CFT). Ordinary tensor fields

$$\phi_{\omega,\overline{\omega}}(\xi,\,\overline{\xi}) = \phi_{\underbrace{\xi\ldots\xi}_{\omega\ \text{times}}\,\underbrace{\overline{\xi}\ldots\overline{\xi}}_{\overline{\omega}\ \text{times}}}(\xi,\,\overline{\xi})\,(\mathrm{d}\xi)^\omega\,(\mathrm{d}\overline{\xi})^{\overline{\omega}} \tag{1.15}$$

are classified according to their weight $(\omega,\,\overline{\omega})$ referring to the holomorphic and anti-holomorphic parts respectively. In fact a transformation $\xi\mapsto\chi(\xi)$ and $\overline{\xi}\mapsto\overline{\chi}(\overline{\xi})$ maps the conformal fields to

$$\phi_{\omega,\overline{\omega}}(\chi,\overline{\chi}) = \left(\frac{\mathrm{d}\chi}{\mathrm{d}\xi}\right)^\omega\left(\frac{\mathrm{d}\overline{\chi}}{\mathrm{d}\overline{\xi}}\right)^{\overline{\omega}}\phi_{\omega,\overline{\omega}}(\xi,\,\overline{\xi}). \tag{1.16}$$



(a) Radial ordering.

(b) Difference of time ordered contours.

Figure 1.1: Map to the complex plane.

An additional conformal transformation

$$z = e^{\xi} = e^{\tau_e + i\sigma} \in \{z \in \mathbb{C} | \operatorname{Im} z \geq 0\}, \qquad \overline{z} = e^{\overline{\xi}} = e^{\tau_e - i\sigma} \in \{z \in \mathbb{C} | \operatorname{Im} z \leq 0\} \tag{1.17}$$

maps the worldsheet of the string to the complex plane. On this Riemann surface the usual time ordering becomes a *radial ordering* as constant time surfaces are circles around the origin (see the contours $\mathscr{C}_{(0)}$ and $\mathscr{C}_{(1)}$ in Figure 1.1). In these coordinates the conserved charge associated to the transformation $z \mapsto z + \epsilon(z)$ in radial quantization is:

$$Q_{\epsilon,\overline{\epsilon}} = \oint_{\mathscr{C}_0} \frac{\mathrm{d}z}{2\pi i} \epsilon(z)\, \mathcal{T}(z) + \oint_{\mathscr{C}_0} \frac{\mathrm{d}\overline{z}}{2\pi i} \overline{\epsilon}(\overline{z})\, \overline{\mathcal{T}}(\overline{z}), \tag{1.18}$$

where $\mathscr{C}_0$ is an anti-clockwise constant radial time path around the origin. The transformation on a field $\phi_{\omega,\overline{\omega}}$ of weight $(\omega, \overline{\omega})$ is thus given by the commutator with $Q_{\epsilon,\overline{\epsilon}}$:

$$\begin{aligned} \delta_{\epsilon,\overline{\epsilon}} \phi_{\omega,\overline{\omega}} &= \big[Q_{\epsilon,\overline{\epsilon}},\, \phi_{\omega,\overline{\omega}}(w,\overline{w})\big] \\ &= \oint_{\mathscr{C}_0} \frac{\mathrm{d}z}{2\pi i} \epsilon(z) \big[\mathcal{T}(z), \phi_{\omega,\overline{\omega}}(w,\overline{w})\big] + \oint_{\mathscr{C}_0} \frac{\mathrm{d}\overline{z}}{2\pi i} \overline{\epsilon}(\overline{z}) \big[\overline{\mathcal{T}}(\overline{z}), \phi_{\omega,\overline{\omega}}(w,\overline{w})\big] \\ &= \oint_{\mathscr{C}_w} \frac{\mathrm{d}z}{2\pi i} \epsilon(z)\, \mathrm{R}\big(\mathcal{T}(z)\, \phi_{\omega,\overline{\omega}}(w,\overline{w})\big) + \oint_{\mathscr{C}_{\overline{w}}} \frac{\mathrm{d}\overline{z}}{2\pi i} \overline{\epsilon}(\overline{z})\, \mathrm{R}\big(\overline{\mathcal{T}}(\overline{z})\, \phi_{\omega,\overline{\omega}}(w,\overline{w})\big), \end{aligned} \tag{1.19}$$

where in the last passage we used the fact that the difference of ordered integrals becomes the contour integral of the radially ordered product computed as a infinitesimally small anti-clockwise loop around w and $\overline{w}$. Equating the result with the expected variation

$$\begin{aligned} \delta_{\epsilon,\overline{\epsilon}} \phi_{\omega,\overline{\omega}} = {} & \omega\, \partial_w \epsilon(w)\, \phi_{\omega,\overline{\omega}}(w,\overline{w}) + \epsilon(w)\, \partial_w \phi_{\omega,\overline{\omega}}(w,\overline{w}) \\ & + \overline{\omega}\, \partial_{\overline{w}} \overline{\epsilon}(\overline{w})\, \phi_{\omega,\overline{\omega}}(w,\overline{w}) + \epsilon(\overline{w})\, \partial_{\overline{w}} \phi_{\omega,\overline{\omega}}(w,\overline{w}) \end{aligned} \tag{1.20}$$

[4]Since we fix $\gamma_{\alpha\beta}(\tau,\sigma) \propto \eta_{\alpha\beta}$ we do not need to account for the components of the connection and we can replace the covariant derivative ∇^{α} with a standard derivative ∂^{α}.

we find the short distance singularities of the components of the stress-energy tensor with the field $\phi_{\omega,\overline{\omega}}(w,\overline{w})$:

$$\begin{aligned}\mathcal{T}(z)\,\phi_{\omega,\overline{\omega}}(w,\overline{w}) &= \frac{\omega}{(z-w)^2}\,\phi_{\omega,\overline{\omega}}(w,\overline{w}) + \frac{1}{z-w}\,\partial_w\phi_{\omega,\overline{\omega}}(w,\overline{w}) + \mathcal{O}(1),\\ \overline{\mathcal{T}}(\overline{z})\,\phi_{\omega,\overline{\omega}}(w,\overline{w}) &= \frac{\overline{\omega}}{(\overline{z}-\overline{w})^2}\,\phi_{\omega,\overline{\omega}}(w,\overline{w}) + \frac{1}{\overline{z}-\overline{w}}\,\partial_{\overline{w}}\phi_{\omega,\overline{\omega}}(w,\overline{w}) + \mathcal{O}(1),\end{aligned} \tag{1.21}$$

where we drop the radial ordering symbol R for simplicity. Since the contour $\mathscr{C}_w$ is infinitely small around w, the conformal properties of $\phi_{\omega,\overline{\omega}}(w,\overline{w})$ are entirely defined by these relations. In fact $\phi_{\omega,\overline{\omega}}(w,\overline{w})$ is a *primary field* if its short distance behaviour with the stress-energy tensor is as in (1.21). This is an example of an *operator product expansion* (O.P.E.)

$$\phi^{(i)}_{\omega_i,\overline{\omega}_i}(z,\overline{z})\,\phi^{(j)}_{\omega_j,\overline{\omega}_j}(w,\overline{w}) = \sum_k \mathcal{C}_{ijk}(z-w)^{\omega_k-\omega_i-\omega_j}\,(\overline{z}-\overline{w})^{\overline{\omega}_k-\overline{\omega}_i-\overline{\omega}_j}\,\phi^{(k)}_{\omega_k,\overline{\omega}_k}(w,\overline{w}) \tag{1.22}$$

which is an asymptotic expansion containing full information on the singularities.[5] The constant coefficients $\mathcal{C}_{ijk}$ are subject to restrictive constraints given by the properties of the conformal theories to the point that a CFT is completely specified by the spectrum of the weights $(\omega_i,\overline{\omega}_i)$ and the coefficients $\mathcal{C}_{ijk}$ [9].

The O.P.E. can also be computed on the stress-energy tensor itself:

$$\begin{aligned}\mathcal{T}(z)\,\mathcal{T}(w) &= \frac{\frac{c}{2}}{(z-w)^4} + \frac{2}{(z-w)^2}\,\mathcal{T}(w) + \frac{1}{z-w}\,\partial_w\mathcal{T}(w),\\ \overline{\mathcal{T}}(\overline{z})\,\overline{\mathcal{T}}(\overline{w}) &= \frac{\frac{\overline{c}}{2}}{(\overline{z}-\overline{w})^4} + \frac{2}{(\overline{z}-\overline{w})^2}\,\overline{\mathcal{T}}(\overline{w}) + \frac{1}{\overline{z}-\overline{w}}\,\partial_{\overline{w}}\overline{\mathcal{T}}(\overline{w}).\end{aligned} \tag{1.23}$$

The components of the stress-energy tensor are therefore not primary fields and show an anomaly in the behaviour encoded by the constant $c\in\mathbb{R}$. This is a reflection of the anomalous algebra of the operator modes L_n and $\overline{L}_n$ computed from the Laurent expansion

$$\begin{aligned}\mathcal{T}(z) &= \sum_{n=-\infty}^{+\infty} L_n\,z^{-n-2} \Rightarrow L_n = \oint_{\mathscr{C}_0}\frac{\mathrm{d}z}{2\pi i}z^{n+1}\mathcal{T}(z),\\ \overline{\mathcal{T}}(\overline{z}) &= \sum_{n=-\infty}^{+\infty} \overline{L}_n\,\overline{z}^{-n-2} \Rightarrow \overline{L}_n = \oint_{\mathscr{C}_0}\frac{\mathrm{d}\overline{z}}{2\pi i}\overline{z}^{n+1}\overline{\mathcal{T}}(\overline{z}).\end{aligned} \tag{1.24}$$

This ultimately leads to the quantum algebra

$$\begin{aligned}[L_n,\,L_m] &= (n-m)\,L_{n+m} + \frac{c}{12}\,n\,(n^2-1)\,\delta_{n,-m},\\ [\overline{L}_n,\,\overline{L}_m] &= (n-m)\,\overline{L}_{n+m} + \frac{c}{12}\,n\,(n^2-1)\,\delta_{n,-m},\\ [L_n,\,\overline{L}_m] &= 0,\end{aligned} \tag{1.25}$$

[5] Expression (1.22) is valid assuming the normalisation of the 2-points function

$$\left\langle \phi^{(i)}_{\omega_i,\overline{\omega}_i}(z,\overline{z})\,\phi^{(j)}_{\omega_j,\overline{\omega}_j}(w,\overline{w})\right\rangle = \frac{\delta_{ij}}{(z-w)^{\omega_i+\omega_j}\,(\overline{z}-\overline{w})^{\overline{\omega}_i+\overline{\omega}_j}}.$$

known as Virasoro algebra, unique central extension of the classical de Witt algebra, with central charge c. Coefficients L_n and $\overline{L}_n$ are called Virasoro operators.[6] Notice that $L_0 + \overline{L}_0$ is the generator of the dilations on the complex plane. In terms of radial quantization this maps to time translations and $L_0 + \overline{L}_0$ can be considered to be the Hamiltonian of the theory.

In the same fashion as (1.24), fields can be expanded in modes:

$$\phi_{\omega,\overline{\omega}}(w,\overline{w}) = \sum_{n,\,m=-\infty}^{+\infty} \phi_{\omega,\overline{\omega}}^{(n,m)}\, z^{-n-\omega}\, \overline{z}^{-m-\overline{\omega}}. \tag{1.26}$$

From the previous relations we can finally define the "asymptotic" in-states as one-to-one correspondence with conformal operators:

$$\left|\phi_{\omega,\overline{\omega}}\right\rangle = \lim_{z,\,\overline{z}\to 0} \phi_{\omega,\overline{\omega}}\, |0\rangle_{\mathrm{SL}_2(\mathbb{R})}\,. \tag{1.27}$$

Regularity of (1.26) requires

$$\phi_{\omega,\overline{\omega}}^{(n,m)}\, |0\rangle_{\mathrm{SL}_2(\mathbb{R})} = 0, \qquad n > \omega, \qquad m > \overline{\omega}. \tag{1.28}$$

As a consequence also

$$L_n\, |0\rangle_{\mathrm{SL}_2(\mathbb{R})} = \overline{L}_n\, |0\rangle_{\mathrm{SL}_2(\mathbb{R})} = 0, \qquad n > -2. \tag{1.29}$$

Finally the definitions of the primary operators (1.21) define the *physical* states as

$$\begin{aligned} L_0 \left|\phi_{\omega,\overline{\omega}}\right\rangle &= \omega \left|\phi_{\omega,\overline{\omega}}\right\rangle, \\ \overline{L}_0 \left|\phi_{\omega,\overline{\omega}}\right\rangle &= \overline{\omega} \left|\phi_{\omega,\overline{\omega}}\right\rangle, \\ L_n \left|\phi_{\omega,\overline{\omega}}\right\rangle &= \overline{L}_n \left|\phi_{\omega,\overline{\omega}}\right\rangle = 0, \quad n \geq 1. \end{aligned} \tag{1.30}$$

From this definition we can build an entire representation of *descendant* states applying any operator L_{-n} (or $\overline{L}_{-n}$) with $n \geq 1$ to $\left|\phi_{\omega,\overline{\omega}}\right\rangle$. Let $\phi_\omega(w)$ be a holomorphic field in the CFT for simplicity, and let $|\phi_\omega\rangle$ be its corresponding state. The generic state at level m built from such state is

$$\left|\phi_\omega^{\{n_1,n_2,\ldots,n_m\}}\right\rangle = L_{-n_1}\, L_{-n_2}\, \ldots L_{-n_m}\, |\phi_\omega\rangle\,, \qquad \sum_{i=1}^{m} n_i = m \geq 0. \tag{1.31}$$

From the commutation relations (1.25) we finally compute its conformal weight as eigenvalue of the (holomorphic) Hamiltonian L_0:

$$L_0 \left|\phi_\omega^{\{n_1,n_2,\ldots,n_m\}}\right\rangle = (\omega + m) \left|\phi_\omega^{\{n_1,n_2,\ldots,n_m\}}\right\rangle. \tag{1.32}$$

States corresponding to primary operators have therefore the lowest energy (intended as eigenvalue of the Hamiltonian) in the entire representation. They are however called *highest weight* states from the mathematical literature which uses the opposite sign for the Hamiltonian operator.

[6] Notice that the subset of Virasoro operators $\{L_{-1},\, L_0,\, L_1\}$ forms a closed sub-algebra generating the group $\mathrm{SL}_2(\mathbb{R})$.

The particular case of the CFT in (1.2) can be cast in this language. In particular the solutions to the E.O.M. factorise into a holomorphic and an anti-holomorphic contributions:

$$\partial_z \partial_{\overline{z}} X(z,\overline{z}) = 0 \qquad \Rightarrow \qquad X(z,\overline{z}) = X(z) + \overline{X}(\overline{z}), \tag{1.33}$$

and the components of the stress-energy tensor (1.9) are

$$\begin{aligned} \mathcal{T}(z) &= \partial_z X(z) \cdot \partial_z X(z), \\ \overline{\mathcal{T}}(\overline{z}) &= \partial_{\overline{z}} \overline{X}(\overline{z}) \cdot \partial_{\overline{z}} \overline{X}(\overline{z}). \end{aligned} \tag{1.34}$$

Using the normalisation of the 2-points function $\langle X^\mu(z,\overline{z})\, X^\nu(w,\overline{w})\rangle = -\frac{1}{2}\eta^{\mu\nu} \ln|z-w|$ and the Wick theorem, we can prove that $c = D$ in (1.23), where D is the dimensions of spacetime (or equivalently the number of scalar fields X^μ in the action). It can be shown that in order to cancel the central charge in bosonic string theory we need to introduce a pair of conformal ghosts $b(z)$ and $c(z)$ with conformal weights $(2,0)$ and $(-1,0)$ respectively, together with their anti-holomorphic counterparts $\overline{b}(z)$ and $\overline{c}(z)$, as a first order Lagrangian theory. The non vanishing components of their stress-energy tensor can be computed as:[7]

$$\begin{aligned} \mathcal{T}_{\text{ghost}}(z) &= c(z)\, \partial_z b(z) - 2\, b(z)\, \partial_z c(z), \\ \overline{\mathcal{T}}_{\text{ghost}}(\overline{z}) &= \overline{c}(\overline{z})\, \partial_{\overline{z}} \overline{b}(\overline{z}) - 2\, \overline{b}(\overline{z})\, \partial_{\overline{z}} \overline{c}(\overline{z}). \end{aligned} \tag{1.35}$$

From their 2-points functions

$$\langle b(z)\, c(w)\rangle = \frac{1}{z-w}, \qquad \left\langle \overline{b}(\overline{z})\, \overline{c}(\overline{w})\right\rangle = \frac{1}{\overline{z}-\overline{w}}, \tag{1.36}$$

we get the O.P.E. of the components of their stress-energy tensor:

$$\begin{aligned} \mathcal{T}_{\text{ghost}}(z)\, \mathcal{T}_{\text{ghost}}(w) &= \frac{-13}{(z-w)^4} + \frac{2}{(z-w)^2}\, \mathcal{T}_{\text{ghost}}(z) + \frac{1}{z-w}\, \partial_z \mathcal{T}_{\text{ghost}}(z), \\ \overline{\mathcal{T}}_{\text{ghost}}(\overline{z})\, \overline{\mathcal{T}}_{\text{ghost}}(\overline{w}) &= \frac{-13}{(\overline{z}-\overline{w})^4} + \frac{2}{(\overline{z}-\overline{w})^2}\, \overline{\mathcal{T}}_{\text{ghost}}(\overline{z}) + \frac{1}{\overline{z}-\overline{w}}\, \partial_{\overline{z}} \overline{\mathcal{T}}_{\text{ghost}}(\overline{z}), \end{aligned} \tag{1.37}$$

[7] In general a system of ghosts $b(z)$ and $c(z)$ with weight $(\lambda,\, 0)$ and $(1-\lambda,\, 0)$ can be introduced as a standalone CFT with action [9]

$$S = \frac{1}{2\pi} \iint \mathrm{d}z\, \mathrm{d}\overline{z}\; b(z)\, \partial_{\overline{z}} c(z).$$

The E.O.M. are $\partial_{\overline{z}} c(z) = \partial_{\overline{z}} b(z) = 0$. The O.P.E. is

$$b(z)\, c(z) = \frac{\varepsilon}{z-w} + \mathcal{O}(1),$$

where $\varepsilon = +1$ for anti-commuting fields and $\varepsilon = -1$ for Bose statistic. Their stress-energy tensor is

$$\mathcal{T}_{\text{ghost}}(z) = -\lambda\, b(z)\, \partial_z c(z) - \varepsilon\, (1-\lambda)\, c(z)\, \partial_z b(z).$$

Their central charge is therefore $c_{\text{ghost}} = \varepsilon\, (1 - 3\mathcal{Q}^2)$, where $\mathcal{Q} = \varepsilon\, (1-2\lambda)$.

The ghost CFT has an additional *ghost number* U(1) symmetry generated by the current

$$j(z) = -b(z)\, c(z).$$

The current is a primary field (i.e. it is not anomalous) when $\mathcal{Q} = 0$ since

$$\mathcal{T}_{\text{ghost}}(z)\, j(w) = \frac{Q}{(z-w)^3} + \mathcal{O}\big((z-w)^{-2}\big).$$

This is the case of the worldsheet fermions in (1.39) for which $\lambda = \frac{1}{2}$. For instance the reparametrisation ghosts with $\lambda = 2$ have $Q = -3$, while the superghosts with $\lambda = \frac{3}{2}$ present $Q = 2$.

which show that $c_{\text{ghost}} = -26$. The central charge is therefore cancelled in the full theory (bosonic string and reparametrisation ghosts) when the spacetime dimensions are $D = 26$. In fact let $\mathcal{T}_{\text{full}} = \mathcal{T} + \mathcal{T}_{\text{ghost}}$ and $\overline{\mathcal{T}}_{\text{full}} = \overline{\mathcal{T}} + \overline{\mathcal{T}}_{\text{ghost}}$, then:

$$\begin{aligned} \mathcal{T}_{\text{full}}(z)\mathcal{T}_{\text{full}}(z)\Big|_{\mathcal{O}((z-w)^{-4})} = \overline{\mathcal{T}}_{\text{full}}(\overline{z})\overline{\mathcal{T}}_{\text{full}}(\overline{z})\Big|_{\mathcal{O}((\overline{z}-\overline{w})^{-4})} &= \frac{c + c_{\text{ghost}}}{2} \\ &= \frac{D}{2} - 13 = 0 \\ &\Leftrightarrow D = 26. \end{aligned} \tag{1.38}$$

$\mathcal{T}_{\text{full}}$ and $\overline{\mathcal{T}}_{\text{full}}$ are then primary fields with conformal weight -2.

1.2 Superstrings

As bosonic string theory deals with commuting fields X^μ, it is impossible to build spacetime fermions and consequently a consistent phenomenology. It is in fact necessary to introduce *worldsheet fermions* (i.e. anti-commuting variables on the string worldsheet) as an extension to the bosonic coordinates. We schematically and briefly recall some results due to the extension of bosonic string theory to the superstring as they will be used in what follows and mainly descend from the previous discussion.

The superstring action is built as an addition to the bosonic equivalent (1.2). In complex coordinates on the plane it is [12]:

$$S[X, \psi] = -\frac{1}{4\pi} \iint \mathrm{d}z\, \mathrm{d}\overline{z} \left(\frac{2}{\alpha'}\, \partial_{\overline{z}} X^\mu\, \partial_z X^\nu + \psi^\mu\, \partial_{\overline{z}} \psi^\nu + \overline{\psi}^\mu\, \partial_z \overline{\psi}^\nu \right) \eta_{\mu\nu}. \tag{1.39}$$

In the last expression ψ^μ are D two-dimensional holomorphic fermion fields with conformal weight $\left(\frac{1}{2}, 0\right)$ and $\overline{\psi}^\mu$ are their anti-holomorphic counterparts with weight $\left(0, \frac{1}{2}\right)$. Their short-distance behaviour is

$$\psi^\mu(z)\, \psi^\nu(w) = \frac{\eta^{\mu\nu}}{z - w}, \qquad \overline{\psi}^\mu(\overline{z})\, \overline{\psi}^\nu(\overline{w}) = \frac{\eta^{\mu\nu}}{\overline{z} - \overline{w}}. \tag{1.40}$$

In this case the components of the stress-energy tensor of the theory are:

$$\begin{aligned} \mathcal{T}(z) &= -\frac{1}{\alpha'}\, \partial_z X(z) \cdot \partial_z X(z) - \frac{1}{2}\, \psi(z) \cdot \partial_z \psi(z), \\ \overline{\mathcal{T}}(\overline{z}) &= -\frac{1}{\alpha'}\, \partial_{\overline{z}} \overline{X}(\overline{z}) \cdot \partial_{\overline{z}} \overline{X}(\overline{z}) - \frac{1}{2}\, \overline{\psi}(\overline{z}) \cdot \partial_{\overline{z}} \overline{\psi}(\overline{z}). \end{aligned} \tag{1.41}$$

The action (1.39) is also invariant under the *supersymmetric* transformations

$$\begin{aligned} \sqrt{\frac{2}{\alpha'}}\, \delta_{\epsilon,\overline{\epsilon}} X^\mu(z, \overline{z}) &= \epsilon(z)\, \psi^\mu(z) + \overline{\epsilon}(\overline{z})\, \overline{\psi}^\mu(\overline{z}), \\ \sqrt{\frac{2}{\alpha'}}\, \delta_\epsilon \psi^\mu(z) &= -\epsilon(z)\, \partial_z X^\mu(z), \\ \sqrt{\frac{2}{\alpha'}}\, \delta_{\overline{\epsilon}} \overline{\psi}^\mu(\overline{z}) &= -\overline{\epsilon}(\overline{z})\, \partial_{\overline{z}} \overline{X}^\mu(\overline{z}) \end{aligned} \tag{1.42}$$

generated by $J(z) = \epsilon(z)\,\mathcal{T}_F(z)$ and $\overline{J}(\overline{z}) = \overline{\epsilon}(\overline{z})\,\overline{\mathcal{T}}_F(\overline{z})$, where $\epsilon(z)$ and $\overline{\epsilon}(\overline{z}) = (\epsilon(z))^*$ are anti-commuting fermions and

$$\begin{aligned}\mathcal{T}_F(z) &= i\sqrt{\frac{2}{\alpha'}}\,\psi(z)\cdot\partial_z X(z),\\ \overline{\mathcal{T}}_F(\overline{z}) &= i\sqrt{\frac{2}{\alpha'}}\,\overline{\psi}(\overline{z})\cdot\partial_{\overline{z}}\overline{X}(\overline{z})\end{aligned} \tag{1.43}$$

are the *supercurrents*. The central charge associated to the Virasoro algebra is in this case given by both bosonic and fermionic contributions:

$$\begin{aligned}\mathcal{T}(z)\,\mathcal{T}(w) &= \frac{\frac{3D}{4}}{(z-w)^4} + \frac{2}{(z-w)^2}\mathcal{T}(w) + \frac{1}{z-w}\partial_w\mathcal{T}(w) + \mathcal{O}(1),\\ \overline{\mathcal{T}}(\overline{z})\,\overline{\mathcal{T}}(\overline{w}) &= \frac{\frac{3D}{4}}{(\overline{z}-\overline{w})^4} + \frac{2}{(\overline{z}-\overline{w})^2}\overline{\mathcal{T}}(\overline{w}) + \frac{1}{\overline{z}-\overline{w}}\partial_{\overline{w}}\overline{\mathcal{T}}(\overline{w}) + \mathcal{O}(1).\end{aligned} \tag{1.44}$$

The central charge is therefore $c = \frac{3}{2}D$ for the CFT defined in (1.39).

As in the case of the bosonic string, in order to cancel the central charge of superstring theory we introduce the reparametrisation anti-commuting ghosts $b(z)$ and $c(z)$ and their anti-holomorphic components as well as the commuting *superghosts* $\beta(z)$ and $\gamma(z)$ and their anti-holomorphic counterparts. These are conformal fields with conformal weights $\left(\frac{3}{2}, 0\right)$ and $\left(-\frac{1}{2}, 0\right)$. Their central charge becomes $c_{\text{ghost}} = c_{bc} + c_{\beta\gamma} = -26 + 11 = -15$ (see footnote 7 for the general computation). When considering the full theory $\mathcal{T}_{\text{full}} = \mathcal{T} + \mathcal{T}_{\text{ghost}}$ and $\overline{\mathcal{T}}_{\text{full}} = \overline{\mathcal{T}} + \overline{\mathcal{T}}_{\text{ghost}}$ the central charge vanishes only in 10-dimensional spacetime:

$$\begin{aligned}\mathcal{T}_{\text{full}}(z)\mathcal{T}_{\text{full}}(z)\Big|_{\mathcal{O}((z-w)^{-4})} = \overline{\mathcal{T}}_{\text{full}}(\overline{z})\overline{\mathcal{T}}_{\text{full}}(\overline{z})\Big|_{\mathcal{O}((\overline{z}-\overline{w})^{-4})} &= \frac{c + c_{\text{ghost}}}{2}\\ &= \frac{3}{4}D - \frac{15}{2} = 0\\ &\Leftrightarrow D = 10.\end{aligned} \tag{1.45}$$

1.3 Extra Dimensions and Compactification

We are ultimately interested in building a consistent phenomenology in the framework of string theory. Any theoretical infrastructure has to be able to support matter states made of fermions. In what follows we thus consider the superstring formulation in $D = 10$ dimensions even when we focus only on its bosonic components.

It is however clear that low energy phenomena need to be explained by a 4-dimensional theory in order to be comparable with other theoretical frameworks and experimental evidence. In this section we briefly review for completeness the necessary tools to be able to reproduce consistent models capable of describing particle physics and beyond. These results represent the background knowledge necessary to better understand more complicated scenarios involving strings. As we will never deal directly with 4-dimensional physics this is not a complete review and we refer to [13]–[16] for more in-depth explanations.

In general we consider Minkowski space in 10 dimensions $\mathscr{M}^{1,9}$. To recover 4-dimensional

spacetime we let it be defined as a product

$$\mathscr{M}^{1,9} = \mathscr{M}^{1,3} \otimes \mathscr{X}_6, \tag{1.46}$$

where $\mathscr{X}_6$ is a generic 6-dimensional manifold at this stage. This *internal* manifold $\mathscr{X}_6$ is however subject to very stringent restrictions due to mathematical consistency conditions and physical requests. In particular $\mathscr{X}_6$ should be a *compact* manifold to "hide" the 6 extra-dimensions computed in (1.45).[8] Moreover the geometry of $\mathscr{M}^{1,3}$ should be a maximally symmetric space and there should be a $N = 1$ unbroken supersymmetry in 4 dimensions. Finally the arising gauge group and the spectrum of fermions should be realistic (e.g. it should be possible to define chiral fermion states) [17]. These manifolds were first conjectured to exist by Eugenio Calabi [18] and their existence was later proved by Shing-Tung Yau [19], hence the name Calabi-Yau (CY) manifolds. They are defined as complex Ricci-flat Kähler manifolds M of dimensions $2m$ and with holonomy $\mathrm{SU}(m)$ (see for instance [20]–[22]). More on this topic is also presented in Part III of this thesis where we compute topological properties of a subset of CY manifolds.

1.3.1 Complex and Kähler Manifolds

In general an *almost complex structure* J is a tensor such that $J^a{}_b\, J^b{}_c = -\delta^a{}_c$. For any vector field $v_p \in \mathrm{T}_pM$ defined in $p \in M$ we then define $(Jv)^a = J^a{}_b\, v^b$, thus the tangent space T_pM has the structure of a *complex vector space*. The tensor J is called *complex structure* if

$$[v_p,\, w_p] + J([J\,v_p,\, w_p] + [v_p,\, J\,w_p]) - [J\,v_p,\, J\,w_p] = 0 \tag{1.47}$$

for any $v_p,\, w_p \in \mathrm{T}_pM$, where $[\cdot,\,\cdot] : \mathrm{T}_pM \times \mathrm{T}_pM \to \mathrm{T}_pM$ is the usual Lie braket of vector fields. A manifold M is a *complex* manifold if it is possible to define a complex structure J on it.[9]

Let then $(M,\, J,\, g)$ be a complex manifold with a Riemannian metric g. The metric is *Hermitian* if

$$g(v_p, w_p) = g(J\,v_p, J\,w_p) \quad \forall v_p,\, w_p \in \mathrm{T}_pM \tag{1.48}$$

In this case we can define a $(1,1)$-form ω as

$$\omega(v_p, w_p) = g(J\,v_p, w_p) \quad \forall v_p,\, w_p \in \mathrm{T}_pM. \tag{1.49}$$

(M, J, g) is a *Kähler* manifold if:

$$\mathrm{d}\omega = \big(\partial + \overline{\partial}\big)\omega(z, \overline{z}) = 0, \tag{1.50}$$

or equivalently $\nabla J = 0$ or $\nabla\omega = 0$, where ∇ is the connection of g. Notice that the operators ∂ and $\overline{\partial}$ are such that $\partial^2 = \overline{\partial}^2 = 0$: they replace the *de Rham cohomology* operator $\mathrm{d}^2 = 0$ in

[8] A compact manifold $\mathscr{X}$ is defined as a Hausdorff topological space whose open covers all have a finite subcover. In other words $\mathscr{X}$ is compact if for each covering atlas $\mathscr{A} = \{U_\alpha\}_{\alpha\in A}$ such that $\mathscr{X} = \bigcup_{\alpha\in A} U_\alpha$, then $\exists \mathscr{B} = \{V_\beta\}_{\beta\in B} \subset \mathscr{A}$ finite such that $\mathscr{X} = \bigcup_{\beta\in B} V_\beta$.

[9] Notice that a smooth function $f\colon M \to \mathbb{C}$ whose pushforward of $v_p \in \mathrm{T}_pM$ is $f_{*\,p}\colon \mathrm{T}_pM \to \mathrm{T}_{f(p)}\mathbb{C} \simeq \mathbb{C}$ is called holomorphic if $(J\, f_{*\,p}(v_p))^a = i(f_{*\,p}(v_p))^a$ as such expression encodes the Cauchy-Riemann equations. Let in fact $f(x,y) = f_1(x,y) + i\, f_2(x,y)$, then the expression implies

$$\begin{cases} \partial_x f_1(x,y) &= \partial_y f_2(x,y) \\ \partial_x f_2(x,y) &= -\partial_y f_1(x,y) \end{cases} \quad \Rightarrow \quad \partial_x f(x,y) = -i\partial_y f(x,y) \quad \Rightarrow \quad \partial_{\overline{z}} f(z,\overline{z}) = 0 \quad \Rightarrow \quad f(z,\overline{z}) \equiv f(z).$$

complex space with the holomorphic and anti-holomorphic *Dolbeault cohomology* operators. The covariant conservation of J and ω implies that the holonomy group must preserve these objects in $\mathbb{R}^{2m}$. Thus we have $\mathrm{Hol}(g) \subseteq \mathrm{U}(m) \subset \mathrm{O}(2m)$.

1.3.2 Calabi-Yau Manifolds

With the general definitions of the Kähler geometry we can now explicitly compute the conditions needed for a CY manifold. In local complex coordinates a Hermitian metric is such that

$$g = g_{a\bar{b}}\, \mathrm{d}z^a \otimes \mathrm{d}\overline{z}^{\bar{b}} + g_{\bar{a}b}\, \mathrm{d}\overline{z}^{\bar{a}} \otimes \mathrm{d}z^b\,, \tag{1.51}$$

thus the Kähler form becomes $\omega = i g_{a\bar{b}}\, \mathrm{d}z^a \wedge \mathrm{d}\overline{z}^{\bar{b}}$. Relation (1.50) translates into:

$$\mathrm{d}\omega = i\left(\partial + \overline{\partial}\right) g_{a\bar{b}}\, \mathrm{d}z^a \wedge \mathrm{d}\overline{z}^{\bar{b}} = 0 \quad \Leftrightarrow \quad \begin{cases} \partial_{z^c} g_{a\bar{b}} & = \partial_{z^a} g_{c\bar{b}} \\ \partial_{\overline{z}^c} g_{\bar{a}b} & = \partial_{\overline{z}^a} g_{\bar{c}b} \end{cases}. \tag{1.52}$$

The $(1,1)$-form ω can locally be written as $\omega = i\,\partial\overline{\partial}\,\phi(z,\overline{z})$ up to a constant. This ultimately leads to

$$g_{a\bar{b}} = \frac{\partial^2 \phi(z,\overline{z})}{\partial z^a \partial \overline{z}^{\bar{b}}} = \partial_{z^a}\partial_{\overline{z}^b}\, \phi(z,\overline{z}), \tag{1.53}$$

Since ω is the Kähler form then the Levi-Civita connection has only fully holomorphic and anti-holomorphic components:

$$\Gamma^a_{\ bc} = g^{a\bar{d}}\, \partial_{z^b}\, g_{\bar{d}c}, \qquad \Gamma^{\bar{a}}_{\ \bar{b}\bar{c}} = g^{\bar{a}d}\, \partial_{\overline{z}^b}\, g_{d\bar{c}}. \tag{1.54}$$

As a consequence the Ricci tensor becomes

$$R_{\bar{a}b} = -\frac{\partial \Gamma^{\bar{c}}_{\ \overline{ac}}}{\partial z^b}. \tag{1.55}$$

Since for CY manifolds $\mathrm{Hol}(g) \subseteq \mathrm{SU}(m)$, the trace part of the coefficients of the connection vanishes. CY manifolds thus have $R_{\bar{a}b} = 0$, that is they are complex Ricci-flat Kähler manifolds with $\mathrm{SU}(m)$ holonomy.

1.3.3 Cohomology and Hodge Numbers

CY manifolds M of complex dimension m present geometric characteristics of general interest both in pure mathematics and string theory. They can be characterised in different ways. For instance the study of the cohomology groups of the manifold has a direct connection with the analysis of topological invariants.

For real manifolds $\widetilde{M}$ of dimension $2m$, closed p-forms $\widetilde{\omega}$ are always defined up to an *exact* term. In fact:

$$\mathrm{d}\widetilde{\omega}'_{(p)} = \mathrm{d}\big(\widetilde{\omega}_{(p)} + \mathrm{d}\widetilde{\eta}_{(p-1)}\big) = 0 \tag{1.56}$$

implies an equivalence relation $\widetilde{\omega}'_{(p)} \sim \widetilde{\omega}_{(p)} + \mathrm{d}\widetilde{\eta}_{(p-1)}$. This translates to the fact that elements of the de Rham cohomology group $H^{(p)}_{\mathrm{d}}\left(\widetilde{M}, \mathbb{R}\right)$ are equivalence classes $[\widetilde{\omega}]$ computed through

the operator d. The term $b^p = \dim H_{\mathrm{d}}^{(p)}(\widetilde{M}, \mathbb{R})$ counts the total number of possible p-forms we can build on $\widetilde{M}$, up to *gauge transformations*. These are known as *Betti numbers*.

The extension to the Dolbeault cohomology in complex space is possible through the operators ∂ and $\overline{\partial}$ over (r, s)-forms on manifolds M of complex dimension m. The equivalence relation (1.56) has a similar expression in complex space as

$$\omega'_{(r,s)} \sim \omega_{(r,s)} + \overline{\partial}\eta_{(r,s-1)}, \tag{1.57}$$

or an equivalent formulation using ∂. The cohomology group in this case is $H_{\overline{\partial}}^{(r,s)}(M, \mathbb{C})$ and the relation with the real counterpart is

$$H_{\mathrm{d}}^{(p)}(M, \mathbb{R}) = \bigoplus_{p=r+s} H_{\overline{\partial}}^{(r,s)}(M, \mathbb{C}). \tag{1.58}$$

As in the case of Betti numbers, we can define the complex equivalents, the *Hodge numbers*, $h^{r,s} = \dim_{\mathbb{C}} H_{\overline{\partial}}^{(r,s)}(M, \mathbb{C})$ which count the number of harmonic (r, s)-forms on M. Notice that in this case $h^{r,s}$ is the complex dimension $\dim_{\mathbb{C}}$ of the cohomology group.

For CY manifolds it is possible to show that the $\mathrm{SU}(m)$ holonomy of g implies that the vector space of $(r, 0)$-forms is $\mathbb{C}$ if $r = 0$ or $r = m$. Therefore $h^{0,0} = h^{m,0} = 1$, while $h^{r,0} = 0$ if $r \neq 0,\ m$. Exploiting symmetries of the cohomology groups, Hodge numbers are usually collected in *Hodge diamonds*. In string theory we are ultimately interested in CY manifolds of real dimensions 6, thus we focus mainly on CY 3-folds (i.e. having $m = 3$). The diamond in this case is

$$\begin{array}{ccccccc} & & & h^{0,0} & & & \\ & & h^{1,0} & & h^{0,1} & & \\ & h^{2,0} & & h^{1,1} & & h^{0,2} & \\ h^{3,0} & & h^{2,1} & & h^{1,2} & & h^{0,3} \\ & h^{3,1} & & h^{2,2} & & h^{1,3} & \\ & & h^{3,2} & & h^{2,3} & & \\ & & & h^{3,3} & & & \end{array} \quad = \quad \begin{array}{ccccccc} & & & 1 & & & \\ & & 0 & & 0 & & \\ & 0 & & h^{1,1} & & 0 & \\ 1 & & h^{2,1} & & h^{2,1} & & 1 \\ & 0 & & h^{1,1} & & 0 & \\ & & 0 & & 0 & & \\ & & & 1 & & & \end{array}, \tag{1.59}$$

where we used $h^{r,s} = h^{m-r,m-s}$ to stress the fact that the only independent Hodge numbers are $h^{1,1}$ and $h^{2,1}$ for $m = 3$. These results will also be the starting point of Part III in which the ability to predict the values of the Hodge numbers using *artificial intelligence* is tested.

1.4 D-branes and Open Strings

Dirichlet branes, or *D-branes*, are another key mathematical object in string theory. They are naturally included as extended hypersurfaces supporting strings with open topology and as physical objects with charge and tension [23]–[27]. They are relevant in the definition of phenomenological models in string theory as they can be arranged to support chiral fermions and bosons in SM-like scenarios as well as beyond [28], [29]. We are ultimately interested in their study to construct Yukawa couplings in string theory.

1.4.1 Compactification of Closed Strings

As a first approach to the definition of D-branes, consider the action (1.2). The variation of such action with respect to δX leads to the equation of motion

$$\partial_\alpha \partial^\alpha X^\mu(\tau,\sigma) = 0 \qquad \mu = 0,1,\dots,D-1, \tag{1.60}$$

and naturally to the *Neumann* boundary conditions:[10]

$$\partial_\sigma X^\mu(\tau,\sigma)\Big|_{\sigma=0}^{\sigma=\ell} = 0, \qquad \mu = 0,1,\dots,D-1. \tag{1.61}$$

Closed strings are such that $X^\mu(\tau,\sigma+\ell) = X^\mu(\tau,\sigma)$. The usual mode expansion in conformal coordinates $X^\mu(z,\overline{z}) = X(z) + \overline{X}(\overline{z})$ leads to

$$\begin{aligned} X^\mu(z) &= x_0^\mu + i\sqrt{\frac{\alpha'}{2}}\left(-\alpha_0^\mu \ln z + \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{\alpha_n^\mu}{n} z^{-n}\right), \\ \overline{X}^\mu(\overline{z}) &= \overline{x}_0^\mu + i\sqrt{\frac{\alpha'}{2}}\left(-\overline{\alpha}_0^\mu \ln \overline{z} + \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{\overline{\alpha}_n^\mu}{n} \overline{z}^{-n}\right), \end{aligned} \tag{1.62}$$

where $\alpha_0^\mu = \overline{\alpha}_0^\mu$ and $\ell = 2\pi$. When the string is free to move in the entire D-dimensional space, then the momentum of the center of mass is $p^\mu = \frac{1}{\sqrt{2\alpha'}}(\alpha_0^\mu + \overline{\alpha}_0^\mu)$.

Now let

$$\mathscr{M}^{1,D-1} = \mathscr{M}^{1,D-2} \otimes S^1(R), \tag{1.63}$$

where $S^1(R)$ is a compact 1-dimensional circle of radius R such that the boundary conditions for the compact coordinate are

$$X^{D-1}(z\,e^{2\pi i}, \overline{z}\,e^{-2\pi i}) = X^{D-1}(z,\overline{z}) + 2\pi\, m\, R, \qquad m \in \mathbb{Z}. \tag{1.64}$$

This is cast into

$$\begin{aligned} \alpha_0^{D-1} + \overline{\alpha}_0^{D-1} &= \sqrt{\frac{2}{\alpha'}}\, n\, \frac{\alpha'}{R}, \qquad n \in \mathbb{Z}, \\ \alpha_0^{D-1} - \overline{\alpha}_0^{D-1} &= \sqrt{\frac{2}{\alpha'}}\, m\, R, \qquad m \in \mathbb{Z}, \end{aligned} \tag{1.65}$$

respectively encoding the quantisation of the momentum for a compact coordinate and the *winding* in the compact direction (1.64). We finally have

$$\begin{aligned} \alpha_0^{D-1} &= \frac{1}{\sqrt{2\alpha'}}\left(n\,\frac{\alpha'}{R} + m\,R\right), \\ \overline{\alpha}_0^{D-1} &= \frac{1}{\sqrt{2\alpha'}}\left(n\,\frac{\alpha'}{R} - m\,R\right), \end{aligned} \tag{1.66}$$

[10] As [24] shows, *Dirichlet* conditions can descend from T-duality which is introduced later.

An interesting phenomenon involving these quantities appears when computing the mass spectrum of the theory. From (1.24) and (1.34) we find

$$\begin{aligned} L_0 &= \frac{\alpha'}{2}\left(\left(\alpha_0^{D-1}\right)^2 + \sum_{i=0}^{D-2}\left(\alpha_0^i\right)^2 + \sum_{n=1}^{+\infty}\left(2\,\alpha_{-n}^{\mu}\alpha_n^{\nu}\,\eta_{\mu\nu} + a\right)\right), \\ \overline{L}_0 &= \frac{\alpha'}{2}\left(\left(\overline{\alpha}_0^{D-1}\right)^2 + \sum_{i=0}^{D-2}\left(\overline{\alpha}_0^i\right)^2 + \sum_{n=1}^{+\infty}\left(2\,\overline{\alpha}_{-n}^{\mu}\overline{\alpha}_n^{\nu}\,\eta_{\mu\nu} + a\right)\right), \end{aligned} \tag{1.67}$$

where a is constant given by normal ordering, representing the zero point energy of the theory. Imposing physical conditions (1.30) and the *level matching* $(L_0 - \overline{L}_0)\,|\phi\rangle = 0$ for closed strings, we find

$$\begin{aligned} M^2 &= \frac{1}{(\alpha')^2}\left(n\,\frac{\alpha'}{R} + m\,R\right)^2 + \frac{4}{\alpha'}\,(\mathrm{N} + a) \\ &= \frac{1}{(\alpha')^2}\left(n\,\frac{\alpha'}{R} - m\,R\right)^2 + \frac{4}{\alpha'}\,(\overline{\mathrm{N}} + a), \end{aligned} \tag{1.68}$$

where $\mathrm{N} = \sum_{n=1}^{+\infty}\alpha_{-n}\cdot\alpha_n$ and $\overline{\mathrm{N}} = \sum_{n=1}^{+\infty}\overline{\alpha}_{-n}\cdot\overline{\alpha}_n$. We then notice that as $R \to \infty$ all states with $m \neq 0$ become infinitely massive while the states for $m = 0$ and all values of n become a continuum. Conversely, as $R \to 0$ all states with $n \neq 0$ become infinitely heavy. In field theory this would translate into a reduction of the number of dimensions since the remaining fields would be independent of the compact coordinate. However in closed string theory as $R \to 0$ the compactified dimension is again present.

As seen in (1.68) the mass spectra of the theories compactified at radius R or $\alpha'\,R^{-1}$ are the same under the exchange of n and m. At the level of the modes this *T-duality* acts by swapping the sign of the right zero-modes in the compact direction

$$\alpha_0^{D-1} \overset{T}{\longmapsto} \alpha_0^{D-1}, \qquad \overline{\alpha}_0^{D-1} \overset{T}{\longmapsto} -\overline{\alpha}_0^{D-1}, \tag{1.69}$$

defining the dual coordinate

$$Y^{D-1}(z,\overline{z}) = Y^{D-1}(z) + \overline{Y}^{D-1}(\overline{z}) = X^{D-1}(z) - \overline{X}^{D-1}(\overline{z}). \tag{1.70}$$

1.4.2 D-branes from T-duality

Consider the case of open strings satisfying the E.O.M. (1.60) and the boundary conditions (1.61). The usual mode expansion (1.62) here leads to:

$$X^{\mu}(z,\overline{z}) = x_0^{\mu} - i\,\alpha'\,p^{\mu}\,\ln(z\overline{z}) + i\sqrt{\frac{\alpha'}{2}}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{\alpha_n^{\mu}}{n}\left(z^{-n} + \overline{z}^{-n}\right) \tag{1.71}$$

and $\ell = \pi$.

Under the compactification (1.63) open strings do not wind around the compact cycle. Thus they do not present a quantum number m as closed strings do. When $R \to 0$ modes with non

vanishing momentum (i.e. with $n \neq 0$) become infinitely massive:

$$p^{D-1} = \frac{n}{R} \xrightarrow{R\to 0} \infty. \tag{1.72}$$

The behaviour is similar to the traditional field theory: the compactified dimension disappears and open string endpoints live in a $(D-1)$-dimensional hypersurface. This is a consequence of the T-duality transformation applied on the compact direction. In fact the original Neumann boundary condition (1.61) becomes a *Dirichlet condition* for Y^{D-1} defined as in (1.70):

$$\begin{aligned}
\partial_\sigma X^{D-1}(\tau,\sigma)\Big|_{\sigma=0}^{\sigma=\pi} &= \partial_\sigma X^{D-1}(e^{\tau_E+i\sigma}) + \partial_\sigma \overline{X}^{D-1}(e^{\tau_E-i\sigma})\Big|_{\sigma=0}^{\sigma=\pi} \\
&= i\,\partial_\xi X^{D-1}(e^{\xi}) - i\,\partial_{\overline{\xi}}\overline{X}^{D-1}(e^{\overline{\xi}})\Big|_{\mathrm{Im}\,\xi=0}^{\mathrm{Im}\,\xi=\pi} \\
&= i\,\partial_{\tau_E} Y^{D-1}(e^{\tau_E+i\sigma}) + i\,\partial_{\tau_E}\overline{Y}^{D-1}(e^{\tau_E-i\sigma})\Big|_{\sigma=0}^{\sigma=\pi} \\
&= \partial_\tau Y^{D-1}(\tau,\sigma)\Big|_{\sigma=0}^{\sigma=\pi} \\
&= 0.
\end{aligned} \tag{1.73}$$

The coordinate of the endpoint in the compact direction is therefore fixed and constrained on a hypersurface called *Dp-brane*, where $p+1$ is the dimension of the surface (in this case $p = D-2$):

$$\begin{aligned}
Y^{D-1}(\tau,\pi) - Y^{D-1}(\tau,0) &= \int_0^\pi \mathrm{d}\sigma\,\partial_\sigma Y^{D-1}(\tau,\sigma) \\
&= i\int_0^\pi \mathrm{d}\sigma\,\partial_\tau X^{D-1}(\tau,\sigma) \\
&= 2\pi\alpha' p^{D-1} \\
&= 2\pi n\,\frac{\alpha'}{R} \\
&= 2\pi n\,R'.
\end{aligned} \tag{1.74}$$

The only difference in the position of the endpoints can only be a multiple of the radius of the compactified dimension. Otherwise they lie on the same hypersurface. The procedure can be generalised to p coordinates, constraining the string to live on a $(D-p-1)$-brane.

This geometric interpretation of the Dirichlet branes and boundary conditions is the basis for the definition of more complex scenarios in which multiple D-branes are inserted in spacetime. D-branes are however much more than mathematical entities. They also present physical properties such as tension and charge [23], [27], [30]. However these aspects will not be discussed here as the following analysis will mainly focus on geometrical aspects of D-branes in spacetime.

1.4.3 Gauge Groups from D-branes

As previously stated, in order to recover 4-dimensional physics we need to compactify the 6 extra-dimensions of the superstring. There are in general multiple ways to do such operation

consistently [31]–[33]. Reproducing the SM or beyond SM spectra are however strong constraints on the possible compactification procedures [29], [34]. Many of the physical requests usually involve the introduction of D-branes and the study of open strings in order to be able to define chiral fermions and realist gauge groups.

As seen in the previous section, D-branes introduce preferred directions of motion by restricting the hypersurface on which the open string endpoints live. Specifically a Dp-brane breaks the original ISO(1, $D-1$) symmetry to ISO(1, p) $\otimes$ SO($D-1-p$).[11] The massless spectrum of the theory on the D-brane is easily computed in lightcone gauge [35], [36]. Using the residual symmetries (1.11) of the two-dimensional diffeomorphism (i.e. harmonic functions of τ and σ) we can set

$$X^+(\tau,\sigma) = x_0^+ + 2\alpha' p^+ \tau, \tag{1.75}$$

where $X^\pm = \frac{1}{\sqrt{2}}(X^0 \pm X^{D-1})$. The vanishing of the stress-energy tensor fixes the oscillators in X^- in terms of the physical transverse modes. The mass shell condition for open strings then becomes:[12]

$$M^2 = \frac{1}{\alpha'}(N-1). \tag{1.76}$$

Consider for a moment bosonic string theory and define the usual vacuum as

$$\alpha_n^i \left|0\right\rangle_{\mathrm{SL}_2(\mathbb{R})} = 0, \qquad n \geq 0, \qquad i = 1, 2, \ldots, D-2, \tag{1.77}$$

we find that at the massless level we have a single U(1) gauge field in the representation of the Little Group SO($D-2$):

$$\mathcal{A}^i \qquad \rightarrow \qquad \alpha_{-1}^i \left|0\right\rangle_{\mathrm{SL}_2(\mathbb{R})}. \tag{1.78}$$

The introduction of a Dp-brane however breaks the Lorentz invariance. Thus the gauge field in the original theory is split into

$$\begin{aligned} \mathcal{A}^A \qquad &\rightarrow \qquad \alpha_{-1}^A \left|0\right\rangle_{\mathrm{SL}_2(\mathbb{R})}, \qquad A = 0, 1, \ldots, p, \\ \mathcal{A}^a \qquad &\rightarrow \qquad \alpha_{-1}^a \left|0\right\rangle_{\mathrm{SL}_2(\mathbb{R})}, \qquad a = 1, 2, \ldots, D-1-p. \end{aligned} \tag{1.79}$$

In the last expression $\mathcal{A}^A$ forms a representation of the Little Group SO($p-2$) and as such it is a vector gauge field in p dimensions. The field $\mathcal{A}^a$ forms a vector representation of the group SO($D-1-p$) and from the point of view of the Lorentz group they are $D-1-p$ scalars in the light spectrum.

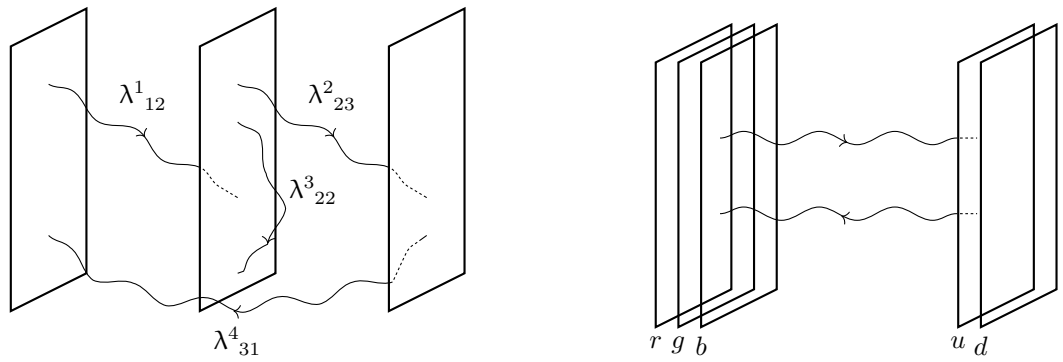
It is also possible to add non dynamical degrees of freedom (D.O.F.) to the open string endpoints. They are known as *Chan-Paton factors* [37]. They have no dynamics and do not spoil Poincaré or conformal invariance in the action of the string. Each state can then be labelled by i and j running from 1 to N. Matrices $\lambda^a{}_{ij}$ thus form a basis for expanding wave functions and states:

$$\left|n;\, a\right\rangle = \sum_{i,\,j=1}^{N} \left|n;\, i, j\right\rangle \lambda^a{}_{ij}. \tag{1.80}$$

In general Chan-Paton factors label the D-brane on which the endpoint of the string lives as in the left of Figure 1.2. Notice that strings stretching across different D-branes present an additional

[11] Notice that usually $D = 10$ in the superstring formulation ($D = 26$ for purely bosonic strings), but we keep a generic indication of the spacetime dimensions when possible.

[12] The constant a in (1.68) takes here the value -1 from the imposition of the canonical commutation relations and a ζ-regularisation.



(a) Chan-Paton factors labelling strings.

(b) Naive model of left handed massive quarks.

Figure 1.2: Strings attached to different D-branes.

term in the mass shell condition proportional to the distance between the hypersurfaces: fields built using strings with Chan-Paton factors $\lambda^a{}_{ij}$ for which $i \neq j$ will therefore be massive. However when N D-branes coincide in space and form a stack their mass vanishes again: it is then possible to organise the N^2 resulting massless fields in a representation of the gauge group $\mathrm{U}(N)$, thus promoting the symmetry $\bigotimes_{a=1}^{N} \mathrm{U}_a(1)$ of N separate D-branes to a larger gauge group. It is also possible to show that in the field theory limit the resulting gauge theory is a Super Yang-Mills gauge theory.

Eventually the massless spectrum of N coincident Dp-branes is formed by $\mathrm{U}(N)$ gauge bosons in the adjoint representation, $N^2 \times (D-1-p)$ scalars and N^2 sets of $(p+1)$-dimensional fermions [16]. These are the basic building blocks for a consistent string phenomenology involving both gauge bosons and matter.

1.4.4 Standard Model Scenarios

Being able to describe gauge bosons and fermions is not enough. Physics as we test it in experiments poses stringent constraints on what kind of string models we can use. For instance there is no way to describe chirality by simply using parallel D-branes and strings stretching among them, while requiring the existence of fermions transforming in different representations of the gauge group is necessary to reproduce SM results [38], [39].

For instance, in the low energy limit it is possible to build a gauge theory of the strong force using a stack of 3 coincident D-branes and an electroweak sector using 2 D-branes. These stacks would separately lead to a $\mathrm{U}(3) \times \mathrm{U}(2)$ gauge theory. It would however be a theory of pure force, without matter content. Moreover we should also worry about the extra U(1) groups appearing: these need careful consideration but go beyond the necessary analysis presented in what follows.

Matter fields are fermions transforming in the bi-fundamental representation $(\mathbf{N}, \mathbf{M})$ of the SM gauge group (1.1). For example left handed quarks in the SM transform under the $(\mathbf{3}, \mathbf{2})$ representation of the group $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L$. This is realised in string theory by a string stretched across two stacks of 3 and 2 D-branes as in the right of Figure 1.2. The fermion

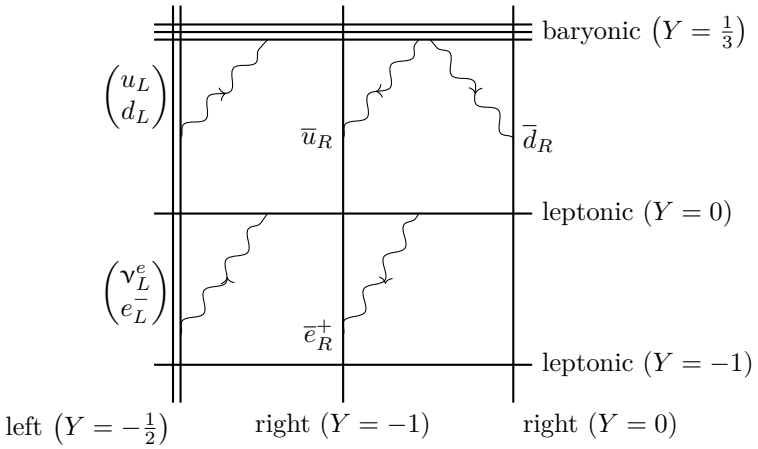


Figure 1.3: Example of SM-like construction using intersecting D-branes with the indications of the hypercharge Y. Perpendicular angles are only a matter of convenience: they are in principal arbitrary.

would then be characterised by the charge under the gauge bosons living on the D-branes. The corresponding anti-particle would then be modelled as a string oriented in the opposite direction. Things get complicated when introducing also left handed leptons transforming in the $(\mathbf{1},\mathbf{2})$ representation: they cannot have endpoints on the same stack of D-branes as quarks since they do not have colour charge. We therefore need to introduce more D-branes to account for all the possible combinations.

An additional issue comes from the requirement of chirality. Strings stretched across D-branes are naturally massive but, in the field theory limit, a mass term would mix different chiralities. We thus need to include a symmetry preserving mechanism for generating the mass of fermions. In string theory there are ways to deal with the requirement. These range from D-branes located at singular points of orbifolds to D-branes intersecting at angles [39]–[41]. In this manuscript we focus on intersecting D6-branes filling the 4-dimensional spacetime and whose additional 3 dimensions are embedded in a CY 3-fold (e.g. as lines in a factorised torus $T^6 = T^2 \times T^2 \times T^2$). This D-brane geometry supports chiral fermion states at their intersection: while some of the modes of the stretched string become indeed massive, the spectrum of the fields is proportional to combinations of the angles and some of the modes can remain massless. The light spectrum is thus composed of the desired matter content alongside with other particles arising from the string compactification.

It is therefore possible to recover a SM-like construction using multiple D-branes at angles as in Figure 1.3, where the angles have been drawn perpendicular but can in principle be arbitrary [42]–[44]. For instance quarks are localised at the intersection of the *baryonic* stack of D-branes, yielding the colour symmetry generators, with the *left* and *right* stacks, leading to the $(\mathbf{3},\mathbf{2})$ and $(\mathbf{3},\mathbf{1})$ representations. The same applies to leptons created by strings attached to the *leptonic* stack. Combinations of the additional U(1) factors in the resulting gauge group finally lead to the definition of the hypercharge Y.

Physics in four dimensions is eventually recovered by compactifying the extra-dimensions of

the superstring.[13] Fermions localised at the intersection of the D-branes are however naturally 4-dimensional as they only propagate in the non compact Minkowski space $\mathscr{M}^{1,3}$. The presence of compactified dimensions however leads to phenomena such as *family replications* of the fermions. With accurate calibration it is in fact possible to recover the quark and lepton families in the SM. Consider for example the simple CY factorised manifold $T^6 = T^2 \times T^2 \times T^2$ and introduce stacks of D6-branes as lines in each of the bi-torus. Even though the lines might never intersect on a plane, they can have points in common on a torus due to the identifications [41]. Since each intersections supports a different set of fermions with different spectrum, the angles of the intersecting branes can be calibrated to reproduce the mass separation of the families of quarks and leptons in the SM.

2 D-branes Intersecting at Angles

2.1 Motivation

As seen in the previous sections, the study of viable phenomenological models in string theory involves the analysis of the properties of systems of D-branes. The inclusion of the physical requirements deeply constrains the possible scenarios. In particular the chiral spectrum of the SM acts as a strong restriction on the possible setup. In this section we study *intersecting D-branes*, which represent a relevant class of models with interacting chiral matter. We focus on the development of technical tools for the computation of Yukawa interactions for D-branes at angles [45]–[51]. The fermion–boson couplings and the study of flavour changing neutral currents [52] are keys to the validity of the models. Furthermore these and many other computations require the ability to calculate correlation functions of (excited) twist and (excited) spin fields. The goal of the section is therefore to address such challenges in specific scenarios.

The computation of the correlation functions of Abelian twist fields can be found in literature and plays a role in many scenarios such as magnetic branes with commuting magnetic fluxes [53]–[56], strings in gravitational wave background [57], [58], bound states of D-branes [59], [60] and tachyon condensation in Superstring Field Theory [61]–[63]. A similar analysis can be extended to excited twist fields even though they are more subtle to treat and hide more delicate aspects [64]–[67]. Results were however found starting from dual models [68] up to modern interpretations of string theory. Correlation functions involving arbitrary numbers of plain and excited twist fields were more recently studied [69]–[71] blending the CFT techniques with the path integral approach and the canonical quantization [72]–[76].

We consider D6-branes intersecting at angles in the case of non Abelian relative rotations presenting non Abelian twist fields at the intersections. We try to understand subtleties and technical issues arising from a scenario which has been studied only in the formulation of non Abelian orbifolds [77]–[79] and for relative SU(2) rotations of the D-branes [80]. In this configuration we study three D6-branes in 10-dimensional Minkowski space $\mathscr{M}^{1,9}$ with an internal space of the form $\mathbb{R}^4 \times \mathbb{R}^2$ before the compactification. The D-branes are embedded as lines in $\mathbb{R}^2$ and as two-dimensional surfaces inside $\mathbb{R}^4$. We focus on the relative rotations which characterise each D-brane in $\mathbb{R}^4$ with respect to the others. In total generality, they are non commuting SO(4)

[13] We specifically reviewed particle physics interactions. Gravitational interactions in general remain untouched by D-branes constructions and still propagate in 10-dimensional spacetime unless the internal space is compact.

matrices. We study the classical solution of the bosonic string which dominates the behaviour of the correlator of twist fields. Using the path integral approach we can in fact separate the classical contribution from the quantum fluctuations and write the correlators of N_B twist fields $\sigma_{\mathrm{M}_{(t)}}\big(x_{(t)}\big)$ as:[14]

$$\left\langle \prod_{t=1}^{N_B} \sigma_{\mathrm{M}_{(t)}}(x_{(t)}) \right\rangle = \mathcal{N}\Big(\big\{x_{(t)},\, \mathrm{M}_{(t)}\big\}_{1\leq t\leq N_B}\Big)\, e^{-S_E\Big(\big\{x_{(t)},\mathrm{M}_{(t)}\big\}_{1\leq t\leq N_B}\Big)}, \tag{2.1}$$

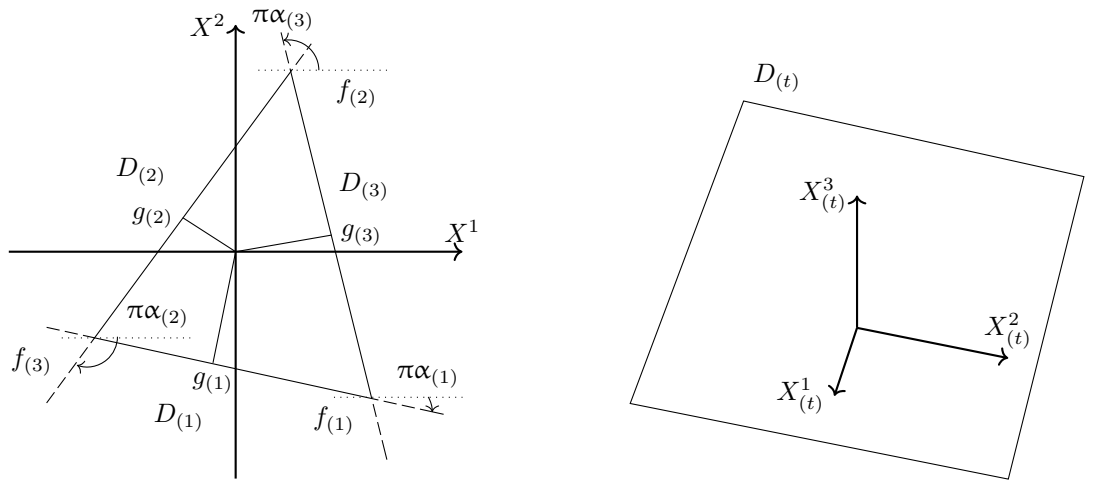
where $\mathrm{M}_{(t)}$ (for $1 \leq t \leq N_B$) are the monodromies induced by the twist fields, N_B is the number of D-branes and $x_{(t)}$ are the intersection points on the worldsheet. Even though quantum corrections are crucial to the complete determination of the normalisation of the correlator, the classical contribution of the Euclidean action represents the leading term of the Yukawa couplings. We focus on its contribution to better address the differences from the usual factorised case and generalise the results to non Abelian rotations of the D-branes. We do not consider the quantum corrections as they cannot be computed with the actual techniques. Their calculations requires the correlator of four twist fields which in turn requires knowledge of the connection formula for Heun functions which is not known.

We therefore study the boundary conditions for the open string describing the D-branes embedded in $\mathbb{R}^4$. In particular we first address the issue connected to the global description of the embedding of the D-branes. In conformal coordinates we rephrase such problem into the study of the monodromies acquired by the string coordinates. These additional phase factors can then be specialised to SO(4) and be studied in spinor representation as a tensor product of SU(2) elements. We thus recast the issue of finding the solution as 4-dimensional real vector to a tensor product of two solutions in the fundamental representation of SU(2). We then see that these solutions are well represented by hypergeometric functions, up to integer factors. Physical requirements finally restrict the number of possible solutions.

2.2 D-brane Configuration and Boundary Conditions

We focus on the bosonic string embedded in $\mathscr{M}^{1,d+4}$. The relevant configuration of the D-branes is seen as two-dimensional Euclidean planes in $\mathbb{R}^4$. We specifically concentrate on the Euclidean solution for the classical bosonic string in this scenario. The mathematical analysis is however more general and can be applied to any Dp-brane embedded in a generic Euclidean space $\mathbb{R}^q$. The classical solution can in principle be defined in this case provided the ability to write the explicit form of the basis of functions with the proper boundary and monodromy conditions. This is possible in the case of three intersecting D-branes but in general it is an open mathematical issue. In the case of three D-branes with generic embedding we can in fact connect a local basis around one intersection point to another, the third depending on the first two intersections, by means of Mellin-Barnes integrals. This way the solution can be explicitly and globally constructed. With more than three D-branes the situation is by far more difficult since the explicit form of the connection formulas is not known.

[14] Ultimately $N_B = 3$ in our case.



(a) D-branes as lines on $\mathbb{R}^2$.

(b) Well adapted system of coordinates.

Figure 2.1: Geometry of D-branes at angles.

2.2.1 Intersecting D-branes at Angles

Let N_B be the total number of D-branes and $t = 1, 2, \ldots, N_B$ be an index defined modulo N_B to label them. We describe one of the D-branes in a well adapted system of coordinates $X^I_{(t)}$, where $I = 1, 2, 3, 4$, as:

$$X^3_{(t)} = X^4_{(t)} = 0. \tag{2.2}$$

We thus choose $X^1_{(t)}$ and $X^2_{(t)}$ to be the coordinates parallel to the D-brane $D_{(t)}$ while $X^3_{(t)}$ and $X^4_{(t)}$ are the coordinates orthogonal to it.

The well adapted reference coordinates system is connected to the global $\mathbb{R}^4$ coordinates X^I by a transformation:

$$(X_{(t)})^I = (R_{(t)})^I{}_J \, X^J - (g_{(t)})^I, \qquad I, J = 1, 2, 3, 4, \tag{2.3}$$

where $R_{(t)}$ represents the rotation of the D-brane $D_{(t)}$ and $g_{(t)} \in \mathbb{R}^4$ its translation with respect to the origin of the global set of coordinates (see Figure 2.1 for a two-dimensional example).

While we could naively consider $R_{(t)} \in \mathrm{SO}(4)$, rotating separately the subset of coordinates parallel and orthogonal to the D-brane does not affect the embedding. In fact it just amounts to a trivial redefinition of the initial well adapted coordinates. The rotation $R_{(t)}$ is actually defined in the Grassmannian:

$$R_{(t)} \in \mathrm{Gr}(2,4) = \frac{\mathrm{SO}(4)}{\mathrm{S}(\mathrm{O}(2) \times \mathrm{O}(2))}, \tag{2.4}$$

that is we just need to consider the left coset where $R_{(t)}$ is a representative of an equivalence class

$$\big[R_{(t)}\big] = \big\{R_{(t)} \sim \mathcal{O}_{(t)} R_{(t)}\big\}, \tag{2.5}$$

where $\mathcal{O}_{(t)} \in \mathrm{S}(\mathrm{O}(2) \times \mathrm{O}(2))$ is defined as

$$\mathcal{O}_{(t)} = \begin{pmatrix} \mathcal{O}_{(t)}^{\parallel} & \\ & \mathcal{O}_{(t)}^{\perp} \end{pmatrix} \tag{2.6}$$

with $\mathcal{O}_t^{\parallel} \in \mathrm{O}(2)$, $\mathcal{O}_t^{\perp} \in \mathrm{O}(2)$ and $\det \mathcal{O}_{(t)} = 1$. The superscript $\parallel$ represents any of the coordinates parallel to the D-brane, while $\perp$ any of the orthogonal. Notice that the additional $\mathbb{Z}_2$ factor in $\mathrm{S}(\mathrm{O}(2) \times \mathrm{O}(2))$ with respect to $\mathrm{SO}(2) \times \mathrm{SO}(2)$ can be used to set $g_{(t)}^{\perp} \geq 0$.

2.2.2 Boundary Conditions for Branes at Angles

The peculiar embedding of the D-branes has natural consequences on the boundary conditions of the open strings. Let $\tau_E = i\tau$ be the Wick rotated time direction. We define the usual upper plane coordinates:

$$u = x + iy = e^{\tau_E + i\sigma} \quad \in \quad \mathscr{H} \cup \{z \in \mathbb{C} \mid \operatorname{Im} z = 0\}, \tag{2.7}$$

$$\overline{u} = x - iy = e^{\tau_E - i\sigma} \quad \in \quad \overline{\mathscr{H}} \cup \{z \in \mathbb{C} \mid \operatorname{Im} z = 0\}, \tag{2.8}$$

where $\mathscr{H} = \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}$ is the upper complex plane and $\overline{\mathscr{H}} = \{z \in \mathbb{C} \mid \operatorname{Im} z < 0\}$ is the lower complex plane. In conformal coordinates u and $\overline{u}$, D-branes at $\sigma = 0$ and $\sigma = \pi$ are mapped onto the real axis $\operatorname{Im} z = 0$. We use the symbol $D_{(t)}$ to label both the brane and the interval representing it on the real axis of the upper half plane:

$$D_{(t)} = \big[x_{(t)}, x_{(t-1)}\big], \qquad t = 2,\, 3,\, \ldots,\, N_B, \qquad x_{(t)} < x_{(t-1)}. \tag{2.9}$$

The points $x_{(t)}$ and $x_{(t-1)}$ represent the worldsheet intersection points of the brane $D_{(t)}$ with the branes $D_{(t+1)}$ and $D_{(t-1)}$ respectively. The choice of the intervals must be carefully considered: since the D-branes are defined modulo N_B, the shorthand for the interval $D_{(1)} = \big[x_{(1)}, x_{(N_B)}\big]$ should actually be:

$$D_{(1)} = \big[x_{(1)}, +\infty\big) \cup \big(-\infty, x_{(N_B)}\big]\,. \tag{2.10}$$

In the global coordinates system associated to the subspace $\mathbb{R}^4 \subset \mathscr{M}^{1,d+4}$ where D-branes intersect, the relevant part of the action in conformal gauge is:

$$\begin{aligned} S_{\mathbb{R}^4} &= \frac{1}{2\pi\alpha'} \iint\limits_{\mathscr{H}} \mathrm{d}u\, \mathrm{d}\overline{u}\; \partial_u X^I\, \partial_{\overline{u}} X^J\, \eta_{IJ} \\ &= \frac{1}{4\pi\alpha'} \iint\limits_{\mathbb{R}\times\mathbb{R}^+} \mathrm{d}x\, \mathrm{d}y\; \big(\partial_x X^I\, \partial_x X^J + \partial_y X^I\, \partial_y X^J\big)\, \eta_{IJ}, \end{aligned} \tag{2.11}$$

where $2\,\partial_u = \partial_x - i\,\partial_y$ and $2\,\partial_{\overline{u}} = \partial_x + i\,\partial_y$. The E.O.M. in these coordinates are:

$$\partial_u \partial_{\overline{u}} X^I(u, \overline{u}) = \frac{1}{4}\big(\partial_x^2 + \partial_y^2\big) X^I(x + iy, x - iy) = 0. \tag{2.12}$$

Their solution factorises as usual in holomorphic and anti-holomorphic components $X^I(u, \overline{u}) = X^I(u) + \overline{X}^I(\overline{u})$.

In the well adapted frame (2.2) we describe an open string with one of the endpoints on $D_{(t)}$ through the relations:

$$\partial_\sigma X^i_{(t)}(\tau,\sigma)\Big|_{\sigma=0} = \partial_y X^i_{(t)}(u,\overline{u})\Big|_{y=0} = 0, \qquad i=1,\,2, \tag{2.13}$$

$$X^m_{(t)}(\tau,0) = X^m_{(t)}(x,x) = 0, \qquad m=3,\,4, \tag{2.14}$$

where $x \in D_{(t)} = [x_t, x_{t-1}]$ and the index i labels the Neumann boundary conditions while m labels the Dirichlet coordinates associated to the direction orthogonal to the D-branes.

As the presence of $g^m_{(t)}$ in (2.3) and (2.14) may complicate the analysis, we consider the derivative along the boundary direction of (2.14) to remove the dependence on the translation vector. This procedure produces simpler boundary conditions which are nevertheless not equivalent to the original (2.13) and (2.14): they will be recovered later by adding further constraints. The simpler boundary conditions we consider in the global coordinates are:

$$\big(R_{(t)}\big)^i{}_J\,\partial_\sigma X^J(\tau,\sigma)\Big|_{\sigma=0} = i\,\big(R_{(t)}\big)^i{}_J\Big(\partial_u X^J(x+i\,0^+) - \partial_{\overline{u}}\overline{X}^J(x-i\,0^+)\Big) = 0, \tag{2.15}$$

$$\big(R_{(t)}\big)^m{}_J\,\partial_\tau X^J(\tau,\sigma)\Big|_{\sigma=0} = i\,\big(R_{(t)}\big)^m{}_J\Big(\partial_u X^J(x+i\,0^+) + \partial_{\overline{u}}\overline{X}^J(x-i\,0^+)\Big) = 0, \tag{2.16}$$

where $i=1,\,2$, $m=3,\,4$ and $x \in D_{(t)}$.

With the introduction of the target space embedding of the worldsheet interaction point between D-branes $D_{(t)}$ and $D_{(t+1)}$, $f_{(t)}$, we recover the full boundary conditions in terms of discontinuities on the real axis:

$$\begin{cases} \partial_u X^I(x+i\,0^+) &= \big(U_{(t)}\big)^I{}_J\,\partial_{\overline{u}}\overline{X}^J(x-i\,0^+), \qquad x \in D_{(t)} \\ X^I(x_{(t)},x_{(t)}) &= f_{(t)} \end{cases}. \tag{2.17}$$

In the last expression we introduced the matrix

$$U_{(t)} = \big(R_{(t)}\big)^{-1}\,\mathcal{S}\,R_{(t)} \in \frac{\mathrm{SO}(4)}{\mathrm{S}(\mathrm{O}(2)\times\mathrm{O}(2))}, \tag{2.18}$$

where

$$\mathcal{S} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{2.19}$$

embeds the difference between Neumann and Dirichlet conditions. Given its definition $U_{(t)}$ is such that $U_{(t)} = \big(U_{(t)}\big)^{-1} = \big(U_{(t)}\big)^T$.

The target space vector $f_{(t)}$ recovers the apparent loss of information suffered when losing $g_{(t)}$. Consider for instance the embedding equations (2.14) for any two intersecting D-branes $D_{(t)}$ and $D_{(t+1)}$. Introducing the auxiliary quantities

$$\mathcal{R}_{(t,\,t+1)} = \begin{pmatrix} R^m_{(t)} \\ R^n_{(t+1)} \end{pmatrix} \in \mathrm{GL}_4(\mathbb{R}), \qquad m,n=3,4, \tag{2.20}$$

$$\mathcal{G}_{(t,\,t+1)} = \begin{pmatrix} g^m_{(t)} \\ g^n_{(t+1)} \end{pmatrix} \in \mathbb{R}^4, \qquad m,n=3,4, \tag{2.21}$$

we can compute the intersection point as:

$$f_{(t)} = \left(\mathcal{R}_{(t,\,t+1)}\right)^{-1} \mathcal{G}_{(t,\,t+1)}. \tag{2.22}$$

Information on $g_{(t)}$ is thus recovered through the global boundary conditions in the second equation in (2.17).

2.2.3 Doubling Trick and Branch Cut Structure

In conformal coordinates we thus introduced the discontinuities (2.17) across each D-brane which define a non trivial cut structure on the plane. One way to deal with them is to introduce the *doubling trick* by gluing the relations along an arbitrary but fixed D-brane $D_{(\bar{t})}$:

$$\partial_z \mathcal{X}(z) = \begin{cases} \partial_u X(u) & \text{if} \quad z = u \quad \text{and} \quad \operatorname{Im} z > 0 \quad \text{or} \quad z \in D_{(\bar{t})} \\ U_{(\bar{t})}\, \partial_{\overline{u}} \overline{X}(\overline{u}) & \text{if} \quad z = \overline{u} \quad \text{and} \quad \operatorname{Im} z < 0 \quad \text{or} \quad z \in D_{(\bar{t})} \end{cases}. \tag{2.23}$$

Let then $\mathcal{U}_{(t,\,t+1)} = U_{(t+1)}\, U_{(t)}$ and $\widetilde{\mathcal{U}}_{(t,\,t+1)} = U_{(\bar{t})}\, U_{(t)}\, U_{(t+1)}\, U_{(\bar{t})}$. The boundary conditions in terms of the doubling field are:

$$\partial_z \mathcal{X}(x_{(t)} + e^{2\pi i}(\eta + i\,0^+)) \;=\; \mathcal{U}_{(t,\,t+1)}\, \partial_z \mathcal{X}(x_{(t)} + \eta + i\,0^+), \tag{2.24}$$

$$\partial_z \mathcal{X}(x_{(t)} + e^{2\pi i}(\eta - i\,0^+)) \;=\; \widetilde{\mathcal{U}}_{(t,\,t+1)}\, \partial_z \mathcal{X}(x_{(t)} + \eta - i\,0^+), \tag{2.25}$$

for $0 < \eta < \min\left(|x_{(t-1)} - x_{(t)}|, |x_{(t)} - x_{(t+1)}|\right)$ in order to consider only the two adjacent D-branes $D_{(t)}$ and $D_{(t+1)}$. Matrices $\mathcal{U}_{(t,\,t+1)}$ and $\widetilde{\mathcal{U}}_{(t,\,t+1)}$ are the non trivial monodromies arising from the rotation of the D-branes.

Since the relative rotations between consecutive D-branes are non Abelian, for each interaction point there are two monodromies $\mathcal{U}$ and $\widetilde{\mathcal{U}}$ depending on the location of the base point of the closed loop: one for paths starting in the upper plane $\mathscr{H}$ and one for paths starting in $\overline{\mathscr{H}}$. As a consequence of the geometry of the rotations of the D-branes, a path on the complex plane enclosing all of them does not present a monodromy:

$$\prod_{t=1}^{N_B} \mathcal{U}_{(\bar{t}-t,\,\bar{t}+1-t)} = \prod_{t=1}^{N_B} \widetilde{\mathcal{U}}_{(\bar{t}+t,\,\bar{t}+1+t)} = \mathbb{1}_4. \tag{2.26}$$

The complex plane has therefore branch cuts running between the D-branes at finite as shown in Figure 2.2. We thus translated the rotations of the D-branes encoded in the matrices $R_{(t)}$ in terms of $\mathcal{U}_{(t,\,t+1)}$ and $\widetilde{\mathcal{U}}_{(t,\,t+1)}$ which are matrix representations of the homotopy group of the complex plane with the described branch cut structure.

As a consistency check, the action (2.11) can be computed in terms of the doubling field $\mathcal{X}$. The map

$$x_{(t)} + \eta \pm i\,0^+ \quad \mapsto \quad x_{(t)} + e^{2\pi i}(\eta \pm i\,0^+) \tag{2.27}$$

must leave the action untouched since it does not depend on the branch cut structure. In fact we can show that

$$S_{\mathbb{R}^4} = \frac{1}{4\pi\alpha'} \iint_{\mathbb{C}} \mathrm{d}z\, \mathrm{d}\overline{z}\; \partial_z \mathcal{X}^T(z)\, U_{(\bar{t})}\, \partial_{\overline{z}} \mathcal{X}(\overline{z}). \tag{2.28}$$

As a matter of fact the action does not depend on the branch structure of the complex plane.

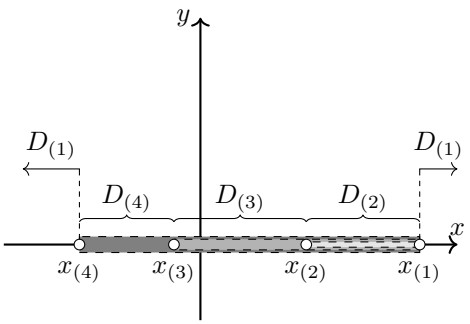


Figure 2.2: Branch cut structure of the complex plane with $N_B = 4$. Cuts are pictured as solid coloured blocks running from one intersection point to another at finite.

2.3 D-branes at Angles in Spinor Representation

In the previous section we showed that it is possible to map the information on the rotations of the D-branes to non trivial monodromies of the doubling field. We thus recast the issue of solving the E.O.M. of the string in the presence of rotated boundary conditions to the search for an explicit solution $\partial_z \mathcal{X}(z)$ reproducing the non trivial monodromies in (2.24) and (2.25).

The field $\partial_z \mathcal{X}(z)$ is technically a 4-dimensional *real* vector which has N_B non trivial monodromy factors represented by 4×4 real matrices, one for each interaction point $x_{(t)}$. A solution of the E.O.M. is encoded in four linearly independent functions with N_B branch points. In principle we could try to write it as a solution to fourth order differential equations with N_B finite Fuchsian points. This is however an open mathematical debate. In fact the basis of such functions around each branch point are usually complicated and defined up to several free parameters. Moreover the explicit connection formulae between any two of them is an unsolved mathematical problem. Using contour integrals and writing the functions as Mellin–Barnes integrals it might be possible to solve the issue in the case $N_B = 3$ but it is certainly not the best course of action.

On the other hand $N_B = 3$ is exactly the case we are investigating. In what follows we use the isomorphism

$$\mathrm{SO}(4) \cong \frac{\mathrm{SU}(2) \times \mathrm{SU}(2)}{\mathbb{Z}_2} \tag{2.29}$$

to map the problem of finding a 4-dimensional real solution to the E.O.M. to a quest for a 2×2 complex matrix. Such matrix is a linear superposition of tensor products of vectors in the fundamental representation of two different SU(2) groups. These vectors are solutions to second order differential equations with three Fuchsian points, possibly the hypergeometric equation. The task is then to find the parameters of the hypergeometric functions producing the spinor representation of the monodromies in (2.24) and (2.25).

2.3.1 Doubling Trick and Rotations in Spinor Representation

We recall some of the properties of the isomorphism (2.29) in Appendix A. We define the spinor representation of X as:

$$X_{(s)}(u, \overline{u}) = X^I(u, \overline{u})\, \tau_I, \tag{2.30}$$

where $\tau = (i\,\mathbb{1}_2,\, \vec{\sigma})$ and $\vec{\sigma}$ is the vector of the Pauli matrices. Consider then:

$$\partial_z \mathcal{X}_{(s)}(z) = \begin{cases} \partial_u X_{(s)}(u) & \text{if} \quad z \in \mathscr{H} \quad \text{or} \quad z \in D_{(\bar{t})} \\ U_L(\vec{n}_{(\bar{t})})\, \partial_{\overline{u}} X_{(s)}(\overline{u})\, U_R^\dagger(\vec{m}_{(\bar{t})}) & \text{if} \quad z \in \overline{\mathscr{H}} \quad \text{or} \quad z \in D_{(\bar{t})} \end{cases}. \tag{2.31}$$

As in the real representation the discontinuities on the D-branes can be cast into monodromy factors with respect to the D-brane $D_{(\bar{t})}$. Branch cut structure and considerations on the homotopy group are left unchanged as long as we consider both left and right sectors of $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$ at the same time. Let $0 < \eta < \min\left(\left|x_{(t)} - x_{(t-1)}\right|, \left|x_{(t+1)} - x_{(t)}\right|\right)$. We find:

$$\partial_z \mathcal{X}_{(s)}(x_t + e^{2\pi i}(\eta + i\,0^+)) = \mathcal{L}_{(t,\,t+1)} \partial_z\, \mathcal{X}_{(s)}(x_t + \eta + i\,0^+)\, \mathcal{R}^\dagger_{(t,\,t+1)}, \tag{2.32}$$

$$\partial_z \mathcal{X}_{(s)}(x_t + e^{2\pi i}(\eta - i\,0^+)) = \widetilde{\mathcal{L}}_{(t,\,t+1)}\, \partial_z \mathcal{X}_{(s)}(x_t + \eta - i\,0^+)\, \widetilde{\mathcal{R}}^\dagger_{(t,\,t+1)}, \tag{2.33}$$

where:

$$\mathcal{L}_{(t,\,t+1)} = U_L(\vec{n}_{(t+1)})\, U_L^\dagger(\vec{n}_{(t)}), \tag{2.34}$$

$$\widetilde{\mathcal{L}}_{(t,\,t+1)} = U_L(\vec{n}_{(\bar{t})})\, U_L^\dagger(\vec{n}_{(t)})\, U_L(\vec{n}_{(t+1)})\, U_L^\dagger(\vec{n}_{(\bar{t})}), \tag{2.35}$$

$$\mathcal{R}_{(t,\,t+1)} = U_R(\vec{m}_{(t+1)})\, U_R^\dagger(\vec{m}_{(t)}), \tag{2.36}$$

$$\widetilde{\mathcal{R}}_{(t,\,t+1)} = U_R(\vec{m}_{(\bar{t})})\, U_R^\dagger(\vec{m}_{(t)})\, U_R(\vec{m}_{(t+1)})\, U_R^\dagger(\vec{m}_{(\bar{t})}). \tag{2.37}$$

In spinor representation the action (2.11) becomes

$$\begin{aligned} S_{\mathbb{R}^4} &= \frac{1}{4\pi\alpha'} \iint\limits_{\mathscr{H}} \mathrm{d}u\, \mathrm{d}\overline{u}\ \mathrm{tr}\Big(\partial_u X_{(s)}(u,\overline{u}) \cdot \partial_{\overline{u}} X^\dagger_{(s)}(u,\overline{u})\Big) \\ &= \frac{1}{8\pi\alpha'} \iint\limits_{\mathbb{C}} \mathrm{d}z\, \mathrm{d}\overline{z}\ \mathrm{tr}\Big(U_L(\vec{n}_{(\bar{t})})\, \partial_z \mathcal{X}_{(s)}(z,\overline{z})\, U_R^\dagger(\vec{m}_{(\bar{t})})\, \partial_{\overline{z}} \mathcal{X}^\dagger_{(s)}(z,\overline{z})\Big). \end{aligned} \tag{2.38}$$

It is possible to show that the closed loop $x_t + \eta \pm i\,0^+ \mapsto x_t + e^{2\pi i}(\eta \pm i\,0^+)$ does not generate additional contributions in the action.

2.3.2 Special Form of Matrices for D-Branes at Angles

The SU(2) matrices involved in this scenario with D-branes intersecting at angles have a particular form. In the left sector (i.e. $\mathrm{SU}(2)_L$ matrices) we have:

$$\mathcal{L}_{(t,\,t+1)} = U_L(\vec{n}_{(t+1)})\, U_L^\dagger(\vec{n}_{(t)}) = -\vec{v}_{(t+1)} \cdot \vec{v}_{(t)} + i\,(\vec{v}_{(t+1)} \times \vec{v}_{(t)}) \cdot \vec{\sigma}, \tag{2.39}$$

with $\left\|\vec{v}_{(t)}\right\|^2 = 1$. This is a consequence of the peculiar properties of the SO(4) matrices $U_{(t)}$ defined in (2.18). Hence the corresponding $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$ element $(U_L(\vec{n}_{(t)}),\, U_R(\vec{m}_{(t)}))$ reflects such characteristics. In particular for the left part we have

$$U_L(\vec{n}_{(t)}) = i\,\vec{v}_{(t)} \cdot \vec{\sigma}, \qquad \left\|\vec{v}_{(t)}\right\|^2 = 1, \tag{2.40}$$

since $U^2_{(t)} = \mathbb{1}_4$ implies that $U_L^2 = \pm\mathbb{1}_2$. The right sector clearly follows the same discussion.

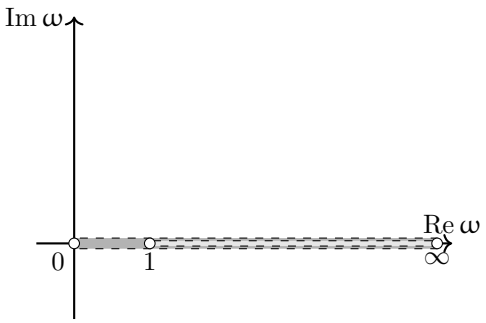


Figure 2.3: Fixing the $\mathrm{SL}_2(\mathbb{R})$ invariance for $N_B = 3$ and $\bar{t} = 1$ leads to a cut structure with all the cuts defined on the real axis towards $\omega_{\bar{t}} = \infty$.

In fact $\mathcal{S}$ in (2.19) can be represented as $U_L = U_R = i\,\sigma_1$. Then any matrix $U_L(\vec{n}_{(t)})$ is of the form $U_L(\vec{n}_{(t)}) = i\,U(\vec{r}_{(t)}) \cdot \sigma_1 \cdot U^\dagger(\vec{r}_{(t)})$, for some $\vec{r}_{(t)}$ as follows from (2.18). Such matrix has vanishing trace and squares to $-\mathbb{1}_2$ hence the term proportional to two-dimensional unit matrix in the expression of the generic SU(2) element given in Appendix A vanishes. As a consequence $n_{(t)} = \frac{1}{4}$ such that (2.40) follows.

2.4 The Classical Solution

In the previous sections we defined the principal tools to study the non Abelian embedding of the D-branes. In what follows we start the investigation of the relation between the hypergeometric solutions and the monodromies due to the geometry of the D-branes.

2.4.1 The Choice of Hypergeometric Functions

We build the spinorial representation with SU(2) matrices and solutions of Fuchsian equations with N_B regular singular points. We are specifically interested in a solution with $N_B = 3$. We fix the usual $\mathrm{SL}_2(\mathbb{R})$ invariance by mapping the three intersection points $x_{(\bar{t}-1)}$, $x_{(\bar{t}+1)}$ and $x_{(\bar{t})}$ to $\omega_{\bar{t}-1} = \omega_{x_{(\bar{t}-1)}} = 0$, $\omega_{\bar{t}+1} = \omega_{x_{(\bar{t}+1)}} = 1$ and $\omega_{\bar{t}} = \omega_{x_{(\bar{t})}} = \infty$ respectively through:

$$\omega_u = \frac{u - x_{(\bar{t}-1)}}{u - x_{(\bar{t})}} \cdot \frac{x_{(\bar{t}+1)} - x_{(\bar{t}-1)}}{x_{(\bar{t}+1)} - x_{(\bar{t})}} \tag{2.41}$$

The cut structure for this choice is presented in Figure 2.3. The map also defines $\arg(\omega_t - \omega_z) \in [0,\, 2\pi)$ for $t = \bar{t} - 1,\, \bar{t} + 1$. We choose $\bar{t} = 1$ in what follows.

The map (2.41) moves the generic Fuchsian singularities to known points on the complex plane. The functions reproducing the necessary monodromies are basis of hypergeometric functions. We define:

$$\mathrm{F}\,(a,\, b;\, c;\, z) = \sum_{k=0}^{+\infty} \frac{(a)_k\,(b)_k}{\Gamma\,(c+k)}\,\frac{z^k}{k!} = \frac{1}{\Gamma\,(c)}\,{}_2F_1\,(a,\, b;\, c;\, z), \tag{2.42}$$

where ${}_2F_1\,(a,\,b;\,c;\,z)$ is the Gauss hypergeometric function and $\Gamma\,(s)$ is the Euler Gamma function. The function $\mathrm{F}\,(a,\,b;\,c;\,z)$ is well defined for any value of its parameters.[15] We define a vector of independent hypergeometric functions:

$$B_0(z) = \begin{pmatrix} \mathrm{F}\,(a,\,b;\,c;\,z) \\ (-z)^{1-c}\ \mathrm{F}\,(a+1-c,\,b+1-c;\,2-c;\,z) \end{pmatrix} \tag{2.43}$$

as basis of functions around $z=0$ with a branch cut on the interval $[0,+\infty)$. The choice of the branch cuts follows from the cut on $[1,+\infty)$ coming from $\mathrm{F}\,(a,\,b;\,c;\,z)$ which has a singularity at $z=1$ and the cut on $[0,+\infty)$ from $(-z)^{1-c}$.

As argued in (2.26), the homotopy group of the complex plane with the branch cut structure of Figure 2.3 is such that a closed loop around all the singularities is homotopically trivial. The corresponding product of the monodromy matrices (2.26) is the unit matrix. Let for instance $\mathcal{M}^{\pm}_{\omega_z}$ be the monodromy matrix which represents a closed loop around ω_z (the $+$ sign represents a path starting in $\mathscr{H}$, while $-$ is a path with base point in $\overline{\mathscr{H}}$). The triviality property is realised through:

$$\mathcal{M}_0^+\,\mathcal{M}_1^+\,\mathcal{M}_\infty^+ = \mathcal{M}_\infty^-\,\mathcal{M}_1^-\,\mathcal{M}_0^- = \mathbb{1}_2 \tag{2.44}$$

The monodromy matrix $\omega_{\bar{t}+1}=1$ can thus be recovered as a product of monodromies around 0 and ∞ given the properties

$$\begin{aligned} \mathcal{M}_0^+ &= \mathcal{M}_0^- = \mathcal{M}_0, \\ \mathcal{M}_\infty^+ &= \mathcal{M}_\infty^- = \mathcal{M}_\infty, \end{aligned} \tag{2.45}$$

which encode the peculiar branch cut structure due to the doubling trick gluing the intervals on one arbitrary D-brane. These matrices are an abstract representation of the monodromy group since they are in an arbitrary basis.

Using the basis in $z=0$ (2.43) it is straightforward to find the explicit representation M_0 of the abstract monodromy $\mathcal{M}_0$:

$$\mathrm{M}_0(c) = \begin{pmatrix} 1 & \\ & e^{-2\pi i c} \end{pmatrix}. \tag{2.46}$$

The computation of the monodromy matrix M_∞ representing the monodromy in $\omega_z=\infty$ in the basis (2.43) requires to first compute the monodromy representation $\widetilde{\mathrm{M}}_\infty$ of the abstract monodromy $\mathcal{M}_\infty$ in the basis of hypergeometric functions around $z=\infty$:

$$B_\infty(z) = \begin{pmatrix} (-z)^{-a}\ \mathrm{F}\left(a,\,a+1-c;\,a+1-b;\,z^{-1}\right) \\ (-z)^{-b}\ \mathrm{F}\left(b,\,b+1-c;\,b+1-a;\,z^{-1}\right) \end{pmatrix}. \tag{2.47}$$

This basis is connected to (2.43) through the transition matrix

$$\mathcal{C}(a,\,b,\,c) = \frac{\pi}{\sin(\pi(a-b))} \begin{pmatrix} \frac{1}{\Gamma(b)\Gamma(c-a)} & -\frac{1}{\Gamma(a)\Gamma(c-b)} \\ \frac{1}{\Gamma(1-a)\Gamma(b+1-c)} & -\frac{1}{\Gamma(1-b)\Gamma(a+1-c)} \end{pmatrix}, \tag{2.48}$$

as $B_0(z) = \mathcal{C}(a,\,b,\,c)\ B_\infty(z)$. Through the loop $z \mapsto z e^{-2\pi i}$ we find:

$$\widetilde{\mathrm{M}}_\infty(a,\,b) = \begin{pmatrix} e^{2\pi i a} & \\ & e^{2\pi i b} \end{pmatrix}. \tag{2.49}$$

Finally we can build the desired monodromy:

$$\mathrm{M}_\infty = \mathcal{C}(a,\,b,\,c)\,\widetilde{\mathrm{M}}_\infty(a,\,b)\,\mathcal{C}^{-1}(a,\,b,\,c). \tag{2.50}$$

[15] It is not necessary to require $c \in \mathbb{Z}_+$ as in the definition of the Gauss hypergeometric function.

2.4.2 The Monodromy Factors

With the previous definitions we reproduce the monodromies of the doubling field in its spinor representation (2.32).[16] These monodromies are tensor products of two basis of hypergeometric functions: the first basis reproduces the monodromies defined as $\mathcal{L}$ and the second one those defined as $\mathcal{R}$ in (2.32). Since in principle there can be several combinations of parameters of the hypergeometric function yielding the same monodromy, we consider the full solution to be a linear superposition of all possible contributions:

$$\partial_z \mathcal{X}(z) = \frac{\partial \omega_z}{\partial z} \sum_{l,\,r} c_{lr}\, \partial_z \mathcal{X}_{l,r}(\omega_z), \tag{2.51}$$

where we drop the index representing the spinorial representation to lighten the notation. We write any possible solution in a factorised form as

$$\partial_z \mathcal{X}_{l,\,r}(\omega_z) = (-\omega_z)^{A_{lr}}\,(1-\omega_z)^{B_{lr}}\,\mathcal{B}^{(L)}_{0,\,l}(\omega_z)\Big(\mathcal{B}^{(R)}_{0,\,r}(\omega_z)\Big)^{T}, \tag{2.52}$$

where l and r label the parameters associates with the left and right sectors of the hypergeometric function. We introduce the left basis element

$$\begin{aligned}\mathcal{B}^{(L)}_{0,\,l}(\omega_z) &= D^{(L)}_l\; B^{(L)}_{0,\,l}(\omega_z) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & K^{(L)}_l \end{pmatrix} \begin{pmatrix} \mathrm{F}\left(a_l,\, b_l;\, c_l;\, \omega_z\right) \\ (-z)^{(1-c_l)}\,\mathrm{F}\left(a_l+1-c_l,\, b_l+1-c_l;\, 2-c_l;\, \omega_z\right) \end{pmatrix}\end{aligned} \tag{2.53}$$

where $D^{(L)}_l \in \mathrm{GL}_2(\mathbb{C})$ is a relative normalisation matrix weighting differently the components of the basis.[17] The right sector follows in a similar way. Notice that the matrices $D^{(L)}_l$ do not fix the absolute normalisation contained in c_{lr}.

2.4.3 Parameters of the Trivial Monodromy

Using the previous relations we can determine the possible $\partial_z \mathcal{X}_{l,r}(\omega_z)$ with the desired monodromies. In this section we study the case of the most general SU(2) matrices despite the fact that in Section 2.3.2 we argued that they have a specific form.

First of all consider the matrices in (2.46) and (2.50). We impose:

$$\begin{cases} D^{(L)}\,\mathrm{M}^{(L)}_0\left(D^{(L)}\right)^{-1} = e^{-2\pi i \delta^{(L)}_0}\,\mathcal{L}(\vec{n}_0) \\ D^{(R)}\,\mathrm{M}^{(R)}_0\left(D^{(R)}\right)^{-1} = e^{-2\pi i \delta^{(R)}_0}\,\mathcal{R}^*(\vec{m}_0) = e^{-2\pi i \delta^{(R)}_0}\,\mathcal{R}(\widetilde{\vec{m}}_0) \\ e^{2\pi i (A_{lr}-\delta^{(L)}_0-\delta^{(R)}_0)} = 1 \end{cases}, \tag{2.54}$$

$$\begin{cases} D^{(L)}, \mathrm{M}^{(L)}_\infty\left(D^{(L)}\right)^{-1} = e^{-2\pi i \delta^{(L)}_\infty}\,\mathcal{L}(\vec{n}_\infty) \\ D^{(R)}\,\mathrm{M}^{(R)}_\infty\left(D^{(R)}\right)^{-1} = e^{-2\pi i \delta^{(R)}_\infty}\,\mathcal{R}^*(\vec{m}_\infty) = e^{-2\pi i \delta^{(R)}_\infty}\,\mathcal{R}(\widetilde{\vec{m}}_\infty) \\ e^{2\pi i (A_{lr}+B_{lr}-\delta^{(L)}_\infty-\delta^{(R)}_\infty)} = 1 \end{cases}, \tag{2.55}$$

[16] In general we do not need to consider (2.33) since they are the same monodromies.
[17] In general they can be different for each solution.

where we defined

$$\mathcal{L}(\vec{n}_0) = \mathcal{L}_{(\bar{t}-1,\,\bar{t})} = U_L(\vec{n}_{(\bar{t})})\, U_L^\dagger(\vec{n}_{(\bar{t}-1)}), \tag{2.56}$$

$$\mathcal{L}(\vec{n}_\infty) = \mathcal{L}_{(\bar{t},\,\bar{t}+1)} = U_L(\vec{n}_{(\bar{t}+1)}) U_L^\dagger(\vec{n}_{(\bar{t})}), \tag{2.57}$$

$$\mathcal{R}(\vec{m}_0) = \mathcal{R}_{(\bar{t}-1,\,\bar{t})} = U_R(\vec{n}_{(\bar{t})}) U_R^\dagger(\vec{n}_{(\bar{t}-1)}), \tag{2.58}$$

$$\mathcal{R}(\vec{m}_\infty) = \mathcal{R}_{(\bar{t},\,\bar{t}+1)} = U_R(\vec{n}_{(\bar{t}+1)}) U_R^\dagger(\vec{n}_{(\bar{t})}). \tag{2.59}$$

The range of $\delta_0^{(L)}$ is

$$\alpha \leq \delta_0^{(L)} \leq \alpha + \frac{1}{2}, \tag{2.60}$$

that is the width of the range is only $\frac{1}{2}$ and not 1 as one would naively expect. This is a consequence of the fact that $e^{-4\pi i \delta_0^{(L)}}$ is the determinant of the right hand side of the first equation in (2.54). We then choose $\alpha = 0$ for simplicity. The same considerations hold true for all the other additional parameters $\delta_0^{(R)}$ and $\delta_\infty^{(L,\,R)}$.

Since we are interested in relative rotations of the D-branes, we choose the rotation in $\omega_{\bar{t}-1} = 0$ in the maximal torus of $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$ without loss of generality: as we have two independent groups, we can in fact fix the orientation of both vectors $\vec{n}_0$ and $\vec{m}_0$. In particular we set:

$$\vec{n}_0 = (0,\, 0,\, n_0^3) \in \mathbb{R}^3, \qquad 0 < n_0^3 < \frac{1}{2}, \tag{2.61}$$

$$\widetilde{\vec{m}}_0 = (0,\, 0,\, -m_0^3) \in \mathbb{R}^3, \qquad 0 < m_0^3 < \frac{1}{2}, \tag{2.62}$$

where $n_0^3 = 0$ is excluded to avoid considering a trivial rotation. We then define the parameters of the rotation in $\omega_{\bar{t}} = \infty$ to be the most general

$$\begin{aligned} \vec{n}_\infty &= (n_\infty^1,\, n_\infty^2,\, n_\infty^3), \\ \widetilde{\vec{m}}_\infty &= (-m_\infty^1,\, m_\infty^2,\, -m_\infty^3), \end{aligned} \tag{2.63}$$

We could actually set $n_\infty^2 = m_\infty^2 = 0$ since the choice of the "gauge" (2.61) and (2.62) is preserved by U(1) rotations mixing n_∞^1 and n_∞^2. We nevertheless keep the general expression in order to check the computations.

Solving (2.54) and (2.55) connects the parameters of the hypergeometric function to the parameter of the rotations (see Appendix B) thus reproducing the boundary conditions of the intersecting D-branes through the non trivial monodromies of the basis of hypergeometric functions. We find:

$$a_l^{(L)} = n_0 + (-1)^{f^{(L)}}\, n_1 + n_\infty + \mathfrak{a}_l^{(L)}, \qquad \mathfrak{a}_l^{(L)} \in \mathbb{Z}, \tag{2.64}$$

$$b_l^{(L)} = n_0 + (-1)^{f^{(L)}}\, n_1 - n_\infty + \mathfrak{b}_l^{(L)}, \qquad \mathfrak{b}_l^{(L)} \in \mathbb{Z}, \tag{2.65}$$

$$c_l^{(L)} = 2\, n_0 + \mathfrak{c}_l^{(L)}, \qquad \mathfrak{c}_l^{(L)} \in \mathbb{Z}, \tag{2.66}$$

$$\delta_0^{(L)} = n_0, \tag{2.67}$$

$$\delta_\infty^{(L)} = -n_0 - (-1)^{f^{(L)}}\, n_1, \tag{2.68}$$

$$K_l^{(L)} = -\frac{1}{2\pi^2}\, \mathcal{G}(a_l^{(L)},\, b_l^{(L)},\, c_l^{(L)})\, \mathcal{F}(a_l^{(L)},\, b_l^{(L)},\, c_l^{(L)})\, \frac{n_\infty^1 + i\, n_\infty^2}{n_\infty}, \tag{2.69}$$

where $f^{(L)} \in \{0, 1\}$. For the sake of brevity we defined two auxiliary functions, namely $\mathcal{G}(a, b, c) = \Gamma(1-a)\,\Gamma(1-b)\,\Gamma(a+1-c)\,\Gamma(b+1-c)$ and $\mathcal{F}(a, b, c) = \sin(\pi c)\sin(\pi(a-b))$. We also introduced the norm $n_1 = \|\vec{n}_1\|$ of the rotation vector around $\omega_{\bar{t}+1} = 1$. Its dependence on the other parameters is encoded in (2.44), where $\mathrm{M}_1^+ = \mathrm{M}_0^{-1}\,\mathrm{M}_\infty^{-1}$, and the composition rule (A.5):

$$\cos(2\pi n_1) = \cos(2\pi n_0)\,\cos(2\pi n_\infty) - \sin(2\pi n_0)\,\sin(2\pi n_\infty)\,\frac{n_\infty^3}{n_\infty}. \tag{2.70}$$

Relations for the right sector follow under the interchange of (L) with (R) and $\vec{n} \leftrightarrow \vec{m}$.

Parameters A_{lr} and B_{lr} follow the previous results and equations (2.54) and (2.55):

$$A_{lr} = n_0 + m_0 + \mathfrak{A}_{lr}, \qquad \mathfrak{A}_{lr} \in \mathbb{Z}, \tag{2.71}$$

$$B_{lr}(-1)^{f^{(L)}}\,n_1 + (-1)^{f^{(R)}}\,m_1 + \mathfrak{B}_{lr}, \qquad \mathfrak{B}_{lr} \in \mathbb{Z}. \tag{2.72}$$

2.4.4 Equivalent Solutions and Necessary Parameters

There are ambiguities in the equations presented in the previous section. In fact the choice of $f^{(L)}$ and $f^{(R)}$ looks arbitrary and leading to an undefined solution. We can use properties of the hypergeometric functions to show that any choice does not affect the final result. Specifically we can start with certain values but we can recover the others through:

$$\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ 0 & 0 & a & z \\ 1-c & c-a-b & b & \end{matrix}\right\} = (1-z)^{c-a-b}\,\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ 0 & 0 & c-b & z \\ 1-c & a+b-c & c-a & \end{matrix}\right\}, \tag{2.73}$$

where P is the Papperitz-Riemann symbol for the hypergeometric functions. We can then assign any admissible value to $f^{(L)}$ and $f^{(R)}$ and then recover the other through the identification:

$$f^{(L)\prime} = \left(1 + f^{(L)}\right) \bmod 2, \tag{2.74}$$

$$\mathfrak{a}_l' = \mathfrak{c}_l - \mathfrak{b}_l, \tag{2.75}$$

$$\mathfrak{b}_l' = \mathfrak{c}_l - \mathfrak{a}_l, \tag{2.76}$$

$$\mathfrak{c}_l' = \mathfrak{c}_l. \tag{2.77}$$

A similar procedure applies also for the right sector. Finally we also identified the "free" parameters:

$$\mathfrak{A}_{lr}' = \mathfrak{A}_{lr}, \tag{2.78}$$

$$\mathfrak{B}_{lr}' = \mathfrak{B}_{lr} - \mathfrak{a}_l^{(L)} - \mathfrak{a}_r^{(R)} - \mathfrak{b}_l^{(L)} - \mathfrak{b}_r^{(R)} + \mathfrak{c}_l^{(L)} + \mathfrak{c}_r^{(R)}. \tag{2.79}$$

The choice of $f^{(L,\,R)}$ is thus simply a convenient relabeling of parameters. We choose $f^{(L)} = f^{(R)} = 0$ for simplicity. Moreover in order to get a well defined solution we must impose constraints on the hypergeometric parameters. We require:

$$c_l^{(L)} \notin \mathbb{Z}, \tag{2.80}$$

$$a_l^{(L)} + b_l^{(L)} \notin \mathbb{Z} + \frac{1}{2}. \tag{2.81}$$

The relations between the parameters of the hypergeometric functions and the monodromies associated to the rotation of the intersecting D-brane are more general than needed. The number of parameters necessary to fix the configuration is 6, that is the amount of parameters to uniquely determine n_0^3, n_∞^1, n_∞^3 and m_0^3, m_∞^1, m_∞^3. As noticed before we can in fact fix $n_\infty^2 = m_\infty^2 = 0$. This is a consequence of the fact that all parameters depend on the norm of the rotation vectors exception made for $K^{(L)}$ and $K^{(R)}$. They depend on $n_\infty^1 + in_\infty^2$ and $m_\infty^1 + im_\infty^2$. Performing a $\mathrm{SU}(2)_L$ and $\mathrm{SU}(2)_R$ rotation around the third axis and a shift of the parameters δ_∞, the phases of the normalisation factors K can vanish.

2.4.5 The Importance of the Normalization Factors

Using the P symbol the solutions can be symbolically written as

$$\begin{aligned}&(-\omega)^{\mathfrak{A}}\,(1-\omega)^{\mathfrak{B}}\times\\ &\times \mathrm{P}\begin{Bmatrix} 0 & 1 & \infty & \\ n_0 & n_1 & n_\infty + \mathfrak{a}^{(L)} & \omega \\ -n_0 + 1 - \mathfrak{c}^{(L)} & -n_1 - \mathfrak{a}^{(L)} - \mathfrak{b}^{(L)} + \mathfrak{c}^{(L)} & -n_\infty + \mathfrak{b}^{(L)} & \end{Bmatrix}\\ &\times \mathrm{P}\begin{Bmatrix} 0 & 1 & \infty & \\ m_0 & m_1 & m_\infty + \mathfrak{a}^{(R)} & \omega \\ -m_0 + 1 - \mathfrak{c}^{(R)} & -m_1 - \mathfrak{a}^{(R)} - \mathfrak{b}^{(R)} + \mathfrak{c}^{(R)} & -m_\infty + \mathfrak{b}^{(R)} & \end{Bmatrix}.\end{aligned} \tag{2.82}$$

The normalisation parameters K cannot however be guessed from the P symbol.

As we are interested in finding the truly independent solutions to the original problem, we can use properties of the hypergeometric functions to reduce the number of possible choices of the integer factors in the definition of the parameters. It is possible to show that any hypergeometric function $\mathrm{F}\left(a+\mathfrak{a},\, b+\mathfrak{b};\, c+\mathfrak{c};\, z\right)$ can be written as a combination of $\mathrm{F}\left(a,\, b;\, c;\, z\right)$ and any of its contiguous functions [81]. For instance we can choose:

$$\mathrm{F}\left(a+\mathfrak{a},\, b+\mathfrak{b};\, c+\mathfrak{c};\, z\right) = h_1(a,\, b,\, c;\, z)\,\mathrm{F}\left(a+1,\, b;\, c;\, z\right) + h_2(a,\, b,\, c;\, z)\,\mathrm{F}\left(a,\, b;\, c;\, z\right), \tag{2.83}$$

where h_1 and h_2 are finite sums of integer (both positive and negative) powers of z and negative powers of $1-z$. For simplicity let:

$$\begin{aligned} F &= \mathrm{F}\left(a,\, b;\, c;\, z\right),\\ F(a+k) &= \mathrm{F}\left(a+k,\, b;\, c;\, z\right),\\ F(b+k) &= \mathrm{F}\left(a,\, b+k;\, c;\, z\right),\\ &\dots \end{aligned} \tag{2.84}$$

Similarly we use a shorthand notation for the basis of the hypergeometric functions:[18]

$$\mathcal{B}_0(a,\, b,\, c;\, z) = \begin{pmatrix} \mathrm{F}\left(a,\, b;\, c;\, z\right) \\ K_{a,b,c}\ (-z)^{(1-c)}\ \mathrm{F}\left(a+1-c,\, b+1-c;\, 2-c;\, z\right) \end{pmatrix}. \tag{2.85}$$

[18] In this expression we introduce a slight abuse of notation since $K_{a,b,c}$ depends on a phase which is not a function of a, b or c. See for instance (2.69) and (B.32).

We can then algorithmically apply the following relations

$$\begin{aligned}(c-a)\,F(a-1)+(2a-c+(b-a)z)\,F-a(1-z)\,F(a+1)=0.\\ (b-a)\,F+a\,F(a+1)-b\,F(b+1)=0,\\ (c-a-b)\,F+a(1-z)\,F(a+1)-(c-b)\,F(b-1)=0,\\ (a+(b-c)z)\,F-a(1-z)\,F(a+1)+(c-a)(c-b)z\,F(c+1)=0,\\ (c-a-1)\,F+a\,F(a+1)-F(c-1)=0,\end{aligned} \tag{2.86}$$

to eliminate unwanted integer factors and keep only $\mathrm{F}\,(a,\,b;\,c;\,z)$ and any of its contiguous functions.

Notice that $\mathcal{B}_0$ is a basis element of the possible solutions of the classical and quantum string E.O.M. Using any relation in (2.86) we can change a, b or c by one unit coherently in both hypergeometric functions contained in $\mathcal{B}_0$. For example from the first equation in (2.86) we expect:

$$(c-a)\,\mathcal{B}_0(a-1)+(2a-c+(b-a)z)\,\mathcal{B}_0-a(1-z)\,\mathcal{B}_0(a+1)=0, \tag{2.87}$$

which can be used to lower and rise the integer factors in a. The relation holds only because the normalisation factor K is present. In fact coefficients in this equation equal those in the relation for the first component of $\mathcal{B}_0$. It is not trivial for the second component where the factor K is key to the consistency. Similarly the relation needed to lower c reads:

$$(a-c)(b-c)\,\mathcal{B}_0(c+1)+(a+(b-c)z)\,\mathcal{B}_0-a(1-z)\mathcal{B}_0(a+1)=0. \tag{2.88}$$

2.4.6 Constraints from the Finite Euclidean Action

In previous sections we present a general procedure to write all possible independent solutions to the classical string E.O.M. However not all of them are physically acceptable. In fact we require the finiteness of the Euclidean action (2.38).

In principle it could appear obvious to use (2.86) to restrict the possible arbitrary integers to:

$$\mathfrak{a}^{(L)}\in\{-1,\,0\},\qquad \mathfrak{a}^{(R)}\in\{-1,\,0\}, \tag{2.89}$$

$$\mathfrak{b}^{(L)}=0,\qquad \mathfrak{b}^{(R)}=0, \tag{2.90}$$

$$\mathfrak{c}^{(L)}=0,\qquad \mathfrak{c}^{(R)}=0. \tag{2.91}$$

We could then use (2.83) to write the possible solution as

$$\begin{aligned}\partial_z\mathcal{X}(z)=\frac{\partial\omega_z}{\partial z}\,(-\omega_z)^{n_0+m_0}\,(1-\omega_z)^{n_1+m_1}\\ \times\sum_{\mathfrak{a}^{(L,\,R)}\in\{-1,0\}}h(\omega_z,\,\mathfrak{a}^{(L,R)})\times\\ \times\,\mathcal{B}_0^{(L)}(a^{(L)}+\mathfrak{a}^{(L)},\,b,\,c;\,\omega_z)\Big(\mathcal{B}_0^{(R)}(a^{(R)}+\mathfrak{a}^{(R)},\,b,\,c;\,\omega_z)\Big)^T.\end{aligned} \tag{2.92}$$

The issue is therefore to find an explicit form for $h(\omega_z,\,\mathfrak{a}^{(L,R)})$ yielding a finite action.

We could however use the symbolic solution (2.82) to find basis of solutions with finite action. As a matter of fact, finding the possible solutions with finite action can be recast to

finding conditions such that the field $\partial_z \mathcal{X}(z)$ is finite by itself. Linearity of this condition ensures a simpler approach with respect to the quadratic action of the string. From (2.38) it is clear that the action can be expressed as the sum of the product of any possible couple of elements of the expansion (2.51). We thus need to take into examination all possible pairs of contributions $\partial_z \mathcal{X}_{l_1 r_1}(z)\ \partial_{\bar{z}} \mathcal{X}_{l_2 r_2}(\bar{z})$. Near its singular points, the behavior of any element of solution (2.51) can be easily read from its symbolic representation (2.82):

$$\partial_z \mathcal{X}(z) \overset{\omega_z \to \omega_t}{\sim} \omega_t^{C_t} \begin{pmatrix} \omega_t^{k_{t_1}} \\ \omega_t^{k_{t_2}} \end{pmatrix} \begin{pmatrix} \omega_t^{h_{t_1}} & \omega_t^{h_{t_2}} \end{pmatrix}. \tag{2.93}$$

It can be verified that the convergence of the action both at finite and infinite intersection points is ensured by the same constraints found when imposing the convergence at any point of the classical solution

$$X_{(s)}(u,\, \overline{u}) = f_{(s)\,(\bar{t}-1)} + \int_{x_{(\bar{t}-1)}}^{u} \mathrm{d}u'\, \partial_{u'} \mathcal{X}_{(s)}(u') + U_L^{\dagger}(\vec{n}_{\bar{t}}) \left[\int_{x_{(\bar{t}-1)}}^{\overline{u}} \mathrm{d}\overline{u}'\, \partial_{\overline{u}'} \mathcal{X}_{(s)}(\overline{u}') \right] U_R(\vec{m}_{\bar{t}}), \tag{2.94}$$

which follows in spinor representation from (2.31) and where $f_{(s),\,(\bar{t}-1)} = f^I_{(\bar{t}-1)}\, \tau_I$. We specifically find:

$$\begin{aligned} &C_t + k_{t_i} + h_{t_j} > -1, \qquad i,\, j \in \{1,2\}, \qquad \omega_t \in \{0,1\}, \\ &C_t + k_{t_i} + h_{t_j} < -1, \qquad i,\, j \in \{1,2\}, \qquad \omega_t = \infty. \end{aligned} \tag{2.95}$$

For simplicity first consider the case of a trivial right rotation.[19] In this case (2.82) becomes

$$(-\omega)^{\mathfrak{A}}\, (1-\omega)^{\mathfrak{B}}\, \mathrm{P} \begin{Bmatrix} 0 & 1 & \infty & \\ n_0 & n_1 & n_\infty + \mathfrak{a}^{(L)} & \omega \\ -n_0 + 1 - \mathfrak{c}^{(L)} & -n_1 - \mathfrak{a}^{(L)} - \mathfrak{b}^{(L)} + \mathfrak{c}^{(L)} & -n_\infty + \mathfrak{b}^{(L)} & \end{Bmatrix}. \tag{2.96}$$

The only possible solution compatible with (2.95) is

$$\mathrm{P} \begin{Bmatrix} 0 & 1 & \infty & \\ n_0 - 1 & n_1 - 1 & n_\infty + 1 & \omega \\ -n_0 & -n_1 & -n_\infty + 2 & \end{Bmatrix}, \tag{2.97}$$

that is $\mathfrak{a}^{(L)} = -1$, $\mathfrak{b}^{(L)} = 0$, $\mathfrak{c}^{(L)} = 0$, $\mathfrak{A} = -1$ and $\mathfrak{B} = -1$.

In the general the solution is more complicated and it depends on the relation between the rotation vectors $\vec{n}_{0,\,1,\,\infty}$, $\vec{m}_{0,\,1,\,\infty}$. For each possible case the solution is however unique and it is given by

1. $n_0 > m_0$ and $n_1 > m_1$:

$$\mathrm{P} \begin{Bmatrix} 0 & 1 & \infty & \\ n_0 - 1 & n_1 - 1 & n_\infty + 1 & \omega \\ -n_0 & -n_1 & -n_\infty + 2 & \end{Bmatrix} \mathrm{P} \begin{Bmatrix} 0 & 1 & \infty & \\ m_0 & m_1 & m_\infty & \omega \\ -m_0 + 1 & -m_1 & -m_\infty + 1 & \end{Bmatrix}, \tag{2.98}$$

[19] That is require that U_R is proportional to the identity.

			$\mathfrak{A}$	$\mathfrak{B}$	$\mathfrak{a}^{(L)}$	$\mathfrak{b}^{(L)}$	$\mathfrak{c}^{(L)}$	$\mathfrak{a}^{(R)}$	$\mathfrak{b}^{(R)}$	$\mathfrak{c}^{(R)}$
$n_0 > m_0$	$n_1 > m_1$	$n_\infty \lesseqgtr m_\infty$	-1	-1	-1	0	0	0	+1	+1
$n_0 > m_0$	$n_1 < m_1$	$n_\infty > m_\infty$	-1	-1	-1	+1	0	0	0	+1
$n_0 > m_0$	$n_1 < m_1$	$n_\infty < m_\infty$	-1	-1	0	0	0	-1	+1	+1
$n_0 < m_0$	$n_1 > m_1$	$n_\infty > m_\infty$	-1	-1	-1	+1	+1	0	0	0
$n_0 < m_0$	$n_1 > m_1$	$n_\infty < m_\infty$	-1	-1	0	0	+1	-1	+1	0
$n_0 < m_0$	$n_1 < m_1$	$n_\infty \lesseqgtr m_\infty$	-1	-1	0	+1	+1	-1	0	0

Table 2.1: Integer shifts in the parameters of the hypergeometric function.

2. $n_0 > m_0$, $n_1 < m_1$ and $n_\infty > m_\infty$:

$$\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ n_0-1 & n_1 & n_\infty & \omega \\ -n_0 & -n_1 & -n_\infty+2 & \end{matrix}\right\}\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ m_0 & m_1-1 & m_\infty+1 & \omega \\ -m_0 & -m_1 & -m_\infty+1 & \end{matrix}\right\}, \tag{2.99}$$

3. $n_0 > m_0$, $n_1 < m_1$ and $n_\infty < m_\infty$:

$$\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ n_0-1 & n_1 & n_\infty+1 & \omega \\ -n_0 & -n_1 & -n_\infty+1 & \end{matrix}\right\}\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ m_0 & m_1-1 & m_\infty & \omega \\ -m_0 & -m_1 & -m_\infty+2 & \end{matrix}\right\}, \tag{2.100}$$

4. $n_0 < m_0$, $n_1 > m_1$ and $n_\infty > m_\infty$:

$$\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ n_0 & n_1-1 & n_\infty & \omega \\ -n_0 & -n_1 & -n_\infty+2 & \end{matrix}\right\}\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ m_0-1 & m_1 & m_\infty+1 & \omega \\ -m_0 & -m_1 & -m_\infty+1 & \end{matrix}\right\}, \tag{2.101}$$

5. $n_0 < m_0$, $n_1 > m_1$ and $n_\infty < m_\infty$:

$$\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ n_0 & n_1-1 & n_\infty+1 & \omega \\ -n_0 & -n_1 & -n_\infty+1 & \end{matrix}\right\}\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ m_0-1 & m_1 & m_\infty & \omega \\ -m_0 & -m_1 & -m_\infty+2 & \end{matrix}\right\}, \tag{2.102}$$

6. $n_0 < m_0$, $n_1 < m_1$:

$$\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ n_0 & n_1 & n_\infty & \omega \\ -n_0 & -n_1 & -n_\infty+1 & \end{matrix}\right\}\mathrm{P}\left\{\begin{matrix} 0 & 1 & \infty & \\ m_0-1 & m_1-1 & m_\infty+1 & \omega \\ -m_0 & -m_1 & -m_\infty+2 & \end{matrix}\right\}. \tag{2.103}$$

The parameters associated to this list of solutions are summarised in Table 2.1, where the symmetry under the exchange of n and m becomes evident.

2.4.7 The Basis of Solutions

In the previous section we produced one solution for each ordering of the n_{ω_z} with respect to m_{ω_z}. There are however other solutions connected to the $\mathbb{Z}_2$ equivalence class in the isomorphism

between SO(4) its double cover. Given a solution $(\vec{n}_0,\, \vec{n}_1,\, \vec{n}_\infty) \oplus (\vec{m}_0,\, \vec{m}_1,\, \vec{m}_\infty)$, we can in fact replace any couple of $\vec{n}$ and $\vec{m}$ by $\widehat{\vec{n}}$ and $\widehat{\vec{m}}$ and get an apparently new solution.[20] For instance we could consider $(\widehat{\vec{n}}_0,\, \widehat{\vec{n}}_1,\, \vec{n}_\infty) \oplus (\vec{m}_0,\, \widehat{\vec{m}}_1,\, \widehat{\vec{m}}_\infty)$. On the other hand the previous substitution would change the SO(4) in both $\omega = 0$ and $\omega = \infty$: it does not represent a new solution. We are left therefore with three possibilities besides the original one:

$$\begin{aligned} &(\widehat{\vec{n}}_0,\, \widehat{\vec{n}}_1,\, \vec{n}_\infty) \oplus (\widehat{\vec{m}}_0,\, \widehat{\vec{m}}_1,\, \vec{m}_\infty), \\ &(\widehat{\vec{n}}_0,\, \vec{n}_1,\, \widehat{\vec{n}}_\infty) \oplus (\vec{m}_0,\, \widehat{\vec{m}}_1,\, \widehat{\vec{m}}_\infty), \\ &(\widehat{\vec{n}}_0,\, \vec{n}_1,\, \widehat{\vec{n}}_\infty) \oplus (\vec{m}_0,\, \widehat{\vec{m}}_1,\, \widehat{\vec{m}}_\infty). \end{aligned} \tag{2.104}$$

We can gauge fix the $\mathbb{Z}_2$ choice by letting $\vec{n}_0^3,\, \vec{m}_0^3 > 0$ as required by (2.61) and (2.62). We are thus left with two possible solutions

$$(\vec{n}_0,\, \vec{n}_1,\, \vec{n}_\infty) \oplus (\vec{m}_0,\, \vec{m}_1,\, \vec{m}_\infty), \tag{2.105}$$

$$(\vec{n}_0,\, \widehat{\vec{n}}_1,\, \widehat{\vec{n}}_\infty) \oplus (\vec{m}_0,\, \widehat{\vec{m}}_1,\, \widehat{\vec{m}}_\infty). \tag{2.106}$$

These are the original and a modified solution obtained by acting with a parity operator P_2 on the rotation parameters at $\omega = 1, \infty$ on both left and right sector at the same time. We then need to ensure its independence in order to accept it as a possible solution.

As shown in Table 2.1, there are only two different cases up to left-right symmetry. The first is

$$\Big\{(n_0 > m_0,\, n_1 > m_1,\, n_\infty > m_\infty),\ (n_0 > m_0,\, \hat{n}_1 < \hat{m}_1,\, \hat{n}_\infty < \hat{m}_\infty)\Big\}, \tag{2.107}$$

which is mapped to

$$\Big\{(n_0 < m_0,\, n_1 < m_1,\, n_\infty < m_\infty),\ (n_0 < m_0,\, \hat{n}_1 > \hat{m}_1,\, \hat{n}_\infty > \hat{m}_\infty)\Big\} \tag{2.108}$$

by the left-right symmetry. The second is

$$\Big\{(n_0 > m_0,\, n_1 > m_1,\, n_\infty < m_\infty),\ (n_0 > m_0,\, \hat{n}_1 < \hat{m}_1,\, \hat{n}_\infty > \hat{m}_\infty)\Big\}, \tag{2.109}$$

which is mapped to

$$\Big\{(n_0 < m_0,\, n_1 < m_1,\, n_\infty > m_\infty),\ (n_0 < m_0,\, \hat{n}_1 > \hat{m}_1,\, \hat{n}_\infty < \hat{m}_\infty)\Big\} \tag{2.110}$$

by the same symmetry.

We can then study the two solutions in the two cases. We first perform the computations common to both cases and then we explicitly specialise the calculations. Computing the parameters of the hypergeometric functions of the first solution leads to:

$$\begin{cases} a^{(L)} &= n_0 + n_1 + n_\infty + \mathfrak{a}^{(L)} \\ b^{(L)} &= n_0 + n_1 - n_\infty + \mathfrak{b}^{(L)} \\ c^{(L)} &= 2\, n_0 + \mathfrak{c}^{(L)} \end{cases}, \qquad \begin{cases} a^{(R)} &= m_0 + m_1 + m_\infty + \mathfrak{a}^{(R)} \\ b^{(R)} &= m_0 + m_1 - m_\infty + \mathfrak{b}^{(R)} \\ c^{(R)} &= 2\, m_0 + 1 + \mathfrak{c}^{(R)} \end{cases}. \tag{2.111}$$

[20] We need to change two rotation vectors because the monodromies are constrained by (2.44).

The values of the constants are in Table 2.1. We then derive the factors $K^{(L)}$ and $K^{(R)}$ using (2.69). The first solution reads:

$$\begin{aligned}\partial_\omega \mathcal{X}_1 = {} & (-\omega)^{n_0+m_0-1}\,(1-\omega)^{n_1+m_1-1} \\ & \times \begin{pmatrix} \mathrm{F}\left(a^{(L)},\, b^{(L)};\, c^{(L)};\, \omega\right) \\ K^{(L)}\,(-\omega)^{1-c^{(L)}}\,\mathrm{F}\left(a^{(L)}+1-c^{(L)},\, b^{(L)}+1-c^{(L)};\, 2-c^{(L)};\, \omega\right) \end{pmatrix} \\ & \times \begin{pmatrix} \mathrm{F}\left(a^{(R)},\, b^{(R)};\, c^{(R)};\, \omega\right) \\ K^{(R)}\,(-\omega)^{1-c^{(R)}}\,\mathrm{F}\left(a^{(R)}+1-c^{(R)},\, b^{(R)}+1-c^{(R)};\, 2-c^{(R)};\, \omega\right) \end{pmatrix}\end{aligned} \tag{2.112}$$

The parameters of the second solution read

$$\begin{cases} \hat{a}^{(L)} &= n_0 + \hat{n}_1 + \hat{n}_\infty + \hat{\mathfrak{a}}^{(L)} = c^{(L)} - a^{(L)} + \mathfrak{a}^{(L)} - \mathfrak{c}^{(L)} + \hat{\mathfrak{a}}^{(L)} + 1 \\ \hat{b}^{(L)} &= n_0 + \hat{n}_1 - \hat{n}_\infty + \hat{\mathfrak{b}}^{(L)} = c^{(L)} - b^{(L)} + \mathfrak{b}^{(L)} - \mathfrak{c}^{(L)} + \hat{\mathfrak{b}}^{(L)} \\ \hat{c}^{(L)} &= 2\,n_0 + \hat{\mathfrak{c}}^{(L)} = c^{(L)} - \mathfrak{c}^{(L)} + \hat{\mathfrak{c}}^{(L)} \end{cases} \\ \begin{cases} \hat{a}^{(R)} &= m_0 + \hat{n}_1 + \hat{n}_\infty + \hat{\mathfrak{a}}^{(R)} = c^{(R)} - a^{(R)} + \mathfrak{a}^{(R)} - \mathfrak{c}^{(R)} + \hat{\mathfrak{a}}^{(R)} + 1 \\ \hat{b}^{(R)} &= m_0 + \hat{n}_1 - \hat{n}_\infty + \hat{\mathfrak{b}}^{(R)} = c^{(R)} - b^{(R)} + \mathfrak{b}^{(R)} - \mathfrak{c}^{(R)} + \hat{\mathfrak{b}}^{(R)} \\ \hat{c}^{(R)} &= 2\,m_0 + \hat{\mathfrak{c}}^{(R)} = c^{(R)} - \mathfrak{c}^{(R)} + \hat{\mathfrak{c}}^{(R)} \end{cases} \tag{2.113}$$

The two cases differ only for constant factors and not in structure.

Case 1 Consider $n_0 > m_0$, $n_1 > m_1$ and $n_\infty > m_\infty$. The associated second solution is $n_0 > m_0$, $\hat{n}_1 < \hat{m}_1$ and $\hat{n}_\infty < \hat{m}_\infty$. Its parameters are:

$$\begin{cases} \hat{a}^{(L)} &= c^{(L)} - a^{(L)} \\ \hat{b}^{(L)} &= c^{(L)} - b^{(L)} \\ \hat{c}^{(L)} &= c^{(L)} \end{cases}, \qquad \begin{cases} \hat{a}^{(R)} &= c^{(R)} - a^{(R)} \\ \hat{b}^{(R)} &= c^{(R)} - b^{(R)} + 1 \\ \hat{c}^{(R)} &= c^{(R)} + 1 \end{cases}, \tag{2.114}$$

The normalisation factors are

$$\hat{K}^{(L)} = K^{(L)}, \qquad \hat{K}^{(R)} = \frac{K^{(R)}}{a^{(R)}\left(c^{(R)} - b^{(R)}\right)}. \tag{2.115}$$

Using Euler relation

$$\mathrm{F}\left(a,\, b;\, c;\, \omega\right) = (1-\omega)^{c-a-b}\,\mathrm{F}\left(c-a,\, c-b;\, c;\, \omega\right), \tag{2.116}$$

we can write the second solution as

$$\begin{aligned}\partial_\omega \mathcal{X}_2 = {} & (-\omega)^{n_0+m_0-1}\,(1-\omega)^{n_1+m_1} \\ & \times \begin{pmatrix} \mathrm{F}\left(a^{(L)},\, b^{(L)};\, c^{(L)};\, \omega\right) \\ K^{(L)}\,(-\omega)^{1-c^{(L)}}\,\mathrm{F}\left(a^{(L)}+1-c^{(L)},\, b^{(L)}+1-c^{(L)};\, 2-c^{(L)};\, \omega\right) \end{pmatrix}. \\ & \times \begin{pmatrix} \mathrm{F}\left(a^{(R)}+1,\, b^{(R)};\, c^{(R)}+1;\, \omega\right) \\ \hat{K}^{(R)}\,(-\omega)^{-c^{(R)}}\,\mathrm{F}\left(a^{(R)}+1-c^{(R)},\, b^{(R)}-c^{(R)};\, 1-c^{(R)};\, \omega\right) \end{pmatrix}\end{aligned} \tag{2.117}$$

In this solution the left basis is exactly the same as in the first solution (2.112) while the right basis differs for $a^{(R)} \mapsto a^{(R)} + 1$ and $c^{(R)} \mapsto c^{(R)} + 1$.

Case 2 Consider now the second option $n_0 > m_0$, $n_1 > m_1$ and $n_\infty < m_\infty$. For the second solution we have $n_0 > m_0$, $\hat{n}_1 < \hat{m}_1$ and $\hat{n}_\infty > \hat{m}_\infty$ and the parameters are explicitly:

$$\begin{cases} \hat{a}^{(L)} &= c^{(L)} - a^{(L)} - 1 \\ \hat{b}^{(L)} &= c^{(L)} - b^{(L)} - 1 \\ \hat{c}^{(L)} &= c^{(L)} \end{cases}, \qquad \begin{cases} \hat{a}^{(R)} &= c^{(R)} - a^{(R)} \\ \hat{b}^{(R)} &= c^{(R)} - b^{(R)} \\ \hat{c}^{(R)} &= c^{(R)} \end{cases}, \tag{2.118}$$

The normalisation factors K are:

$$\hat{K}^{(L)} = K^{(L)} \frac{(b^{(L)} - 1)(c^{(L)} - a^{(L)} - 1)}{a^{(L)}(c^{(L)} - b^{(L)})}, \qquad \hat{K}^{(R)} = K^{(R)}. \tag{2.119}$$

Using Euler relation we write the second solution for the second case as

$$\begin{aligned} \partial_\omega \mathcal{X}_2 = (-\omega)^{n_0+m_0-1} (1-\omega)^{n_1+m_1} \\ \times \begin{pmatrix} \mathrm{F}\left(a^{(L)}+1,\, b^{(L)}-1;\, c^{(L)};\, \omega\right) \\ \hat{K}^{(L)} (-\omega)^{1-c^{(L)}} \mathrm{F}\left(a^{(L)}+2-c^{(L)},\, b^{(L)}-c^{(L)};\, 2-c^{(L)};\, \omega\right) \end{pmatrix} \\ \times \begin{pmatrix} \mathrm{F}\left(a^{(R)},\, b^{(R)};\, c^{(R)};\, \omega\right) \\ \hat{K}^{(R)} (-\omega)^{-c^{(R)}} \mathrm{F}\left(a^{(R)}+1-c^{(R)},\, b^{(R)}+1-c^{(R)};\, 2-c^{(R)};\, \omega\right) \end{pmatrix} \end{aligned} . \tag{2.120}$$

The right basis is the same as in the first solution while the left basis differs for $a^{(L)} \mapsto a^{(L)} + 1$ and $b^{(L)} \mapsto b^{(L)} - 1$.

2.4.8 The Solution

We showed that there are two independent solutions. The general solution for $\partial_\omega \mathcal{X}$ is therefore:

$$\partial_\omega \mathcal{X} = C_1\, \partial_\omega \mathcal{X}_1 + C_2\, \partial_\omega \mathcal{X}_2. \tag{2.121}$$

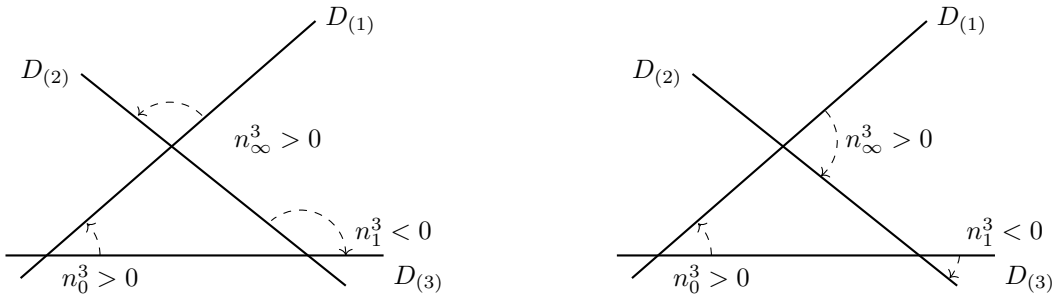
The final solution depends only on two complex constants, C_1 and C_2, which we can fix imposing the global conditions in (2.17), that is the second equation in the solution (2.94). As the three intersection points in target space always define a triangle on a 2-dimensional plane, we impose the boundary conditions knowing two angles formed by the sides of the triangle (i.e. the branes between two intersections) and the length of one of them. Since we already fixed the parameters associated to the rotations, we need to compute the length of one of the sides. Consider for instance the length of $X(x_{\bar{t}+1},\, x_{\bar{t}+1}) - X(x_{\bar{t}-1},\, x_{\bar{t}-1})$. Explicitly we impose the four real equations in spinorial formalism

$$\int_0^1 \mathrm{d}\omega\, \partial_\omega \mathcal{X}(\omega) + U_L^\dagger(\vec{n}_{\bar{t}}) \left[\int_0^1 \mathrm{d}\overline{\omega}\, \partial_{\overline{\omega}} \mathcal{X}(\overline{\omega}) \right] U_R(\vec{m}_{\bar{t}}) = f_{\bar{t}+1\,(s)} - f_{\bar{t}-1\,(s)}, \tag{2.122}$$

where we used the mapping (2.41) to write the integrals in the ω variables. This equation has enough D.O.F. to fix completely the two complex parameters C_1 and C_2. The final generic solution is thus uniquely determined.

$\vec{n}_0$	$\vec{n}_\infty$	relations		n_1	$\vec{n}_1$	$\sum_t \vec{n}_{\vec{t}}$
$n_0\,\vec{k}$	$n_\infty\,\vec{k}$	$n_0+n_\infty<\frac{1}{2}$	$n_0 \lesseqgtr n_\infty$	n_0+n_∞	$-n_1\,\vec{k}$	0
$n_0\,\vec{k}$	$n_\infty\,\vec{k}$	$n_0+n_\infty>\frac{1}{2}$	$n_0 \lesseqgtr n_\infty$	$1-(n_0+n_\infty)$	$n_1\,\vec{k}$	$\vec{k}$
$n_0\,\vec{k}$	$-n_\infty\,\vec{k}$	$n_0+n_\infty \lesseqgtr \frac{1}{2}$	$n_0>n_\infty$	n_0-n_∞	$-n_1\,\vec{k}$	0
$n_0\,\vec{k}$	$-n_\infty\,\vec{k}$	$n_0+n_\infty \lesseqgtr \frac{1}{2}$	$n_0<n_\infty$	$-n_0+n_\infty$	$n_1\,\vec{k}$	0

Table 2.2: Abelian limit of SU(2) monodromies



(a) Case 1.

(b) Case 2.

Figure 2.4: The Abelian limit when the triangle has all acute angles. This corresponds to the cases $n_0+n_\infty<\frac{1}{2}$ and $n_0<n_\infty$ which are exchanged under the parity P_2.

2.5 Recovering the SU(2) and the Abelian Solution

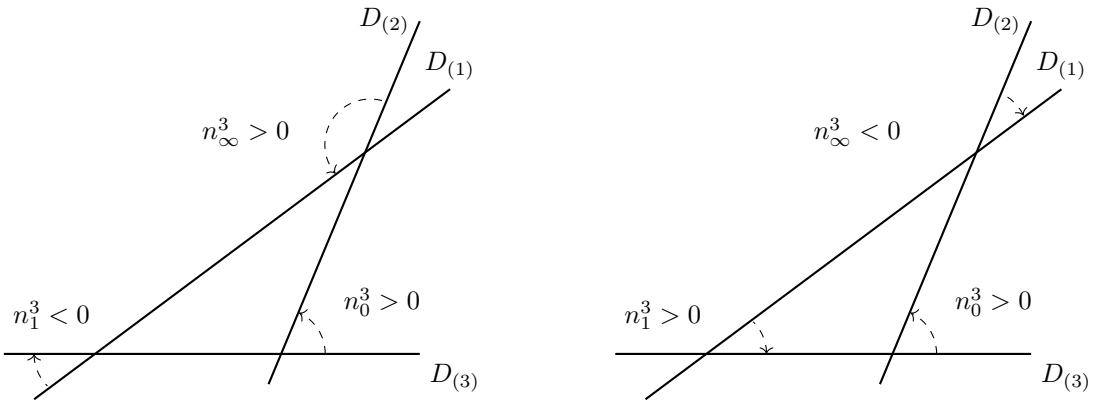
In this section we show how this general procedure includes both the solution with pure SU(2) rotation matrices and the solution with Abelian rotations of the D-branes. The Abelian solution emerges from this construction as a limit and produces the known result for Abelian SO(2) × SO(2) ⊂ SO(4) rotations in the case of a factorised space $\mathbb{R}^4=\mathbb{R}^2\times\mathbb{R}^2$.

2.5.1 Abelian Limit of the SU(2) Monodromies

Here we compute the parameter $\vec{n}_1$ given two Abelian rotation in $\omega=0$ and $\omega=\infty$ using the standard expression for two SU(2) element multiplication given in (A.5) in Appendix A. Results are shown in Table 2.2.

Under the parity transformation P_2 the previous four cases are grouped into two sets $\{n_1=n_0+n_\infty,\ \hat{n}_1=-n_0+\hat{n}_\infty\}$ and $\{n_1=1-(n_0+n_\infty),\ \hat{n}_1=n_0-\hat{n}_\infty\}$. Geometrically the first group corresponds to the same geometry which is depicted in Figure 2.4 while the second in Figure 2.5. We can in fact arbitrarily fix the orientation of $D_{(3)}$ to obtain these geometrical interpretations. Since $n_0^3>0$ we can then fix the orientation of D_1. D_2 is then fixed relatively to D_1 by the sign of n_∞^3. The sign of n_1^3 then follows.

Differently from the usual geometric Abelian case, this group analytical approach distinguishes between the possible orientations of the D-branes. In fact we can compare all possible



(a) Case 1. (b) Case 2.

Figure 2.5: The Abelian limit when the triangle has one obtuse angle. This corresponds to the cases $n_0 + n_\infty > \frac{1}{2}$ and $n_0 > n_\infty$ which are exchanged under the parity P_2.

D-brane orientation and the group parameter n^3 with the angles in the Abelian configuration in Figure 2.6. The relation between the usual Abelian paramter $\epsilon_{\vec{t}}$ and $n^3_{\vec{t}}$ is

$$\varepsilon_{\vec{t}} = n^3_{\vec{t}} + \theta(-n^3_{\vec{t}}), \tag{2.123}$$

when all $m = 0$.

2.5.2 The Abelian Limit of the Left Solutions

We can then compute the basis element for the entries of Table 2.1 for any value of n_1 given in Table 2.2. For simplicity we consider the left sector of the solution and drop the notation identifying it to avoid cluttering the equations. The right sector follows in a similar way.

In the Abelian limit either $K = 0$ or $K = \infty$. We can absorb the infinite divergence in a constant term globally multiplying the solution and use:

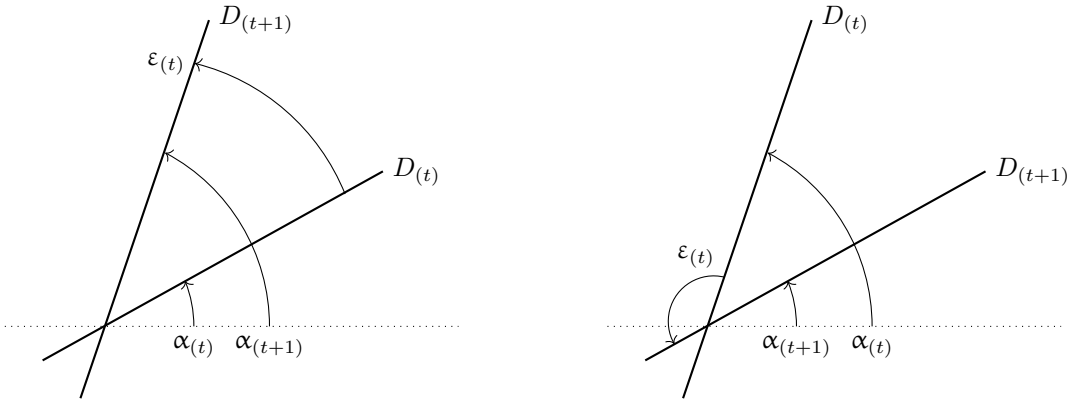
$$D\Big|_{K=0} = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \tag{2.124}$$

$$D\Big|_{K=\infty} = \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}. \tag{2.125}$$

Results are summarised in Table 2.3 where we left some hypergeometric functions in their symbolic form for compactness even though they are in fact elementary functions since either a or $c - b$ equal -1.

2.5.3 The $\mathrm{SU}(2)_L$ Limit

We recover the non Abelian SU(2) solution by considering $m_{\vec{t}} \sim 0$. This is the first specific case shown in Section 2.4.7. In this scenario the left solution $\mathcal{B}^{(L)}$ is always the same and matches the



(a) Case 1. (b) Case 2.

Figure 2.6: The geometrical angles used in the usual geometrical approach to the Abelian configuration do not distinguish among the possible branes orientations. In fact we have $0 \leq \alpha < 1$ and $0 < \varepsilon < 1$.

$\left(\mathfrak{a}^{(L)},\, \mathfrak{b}^{(L)},\, \mathfrak{c}^{(L)}\right)$	n_1	$\left(\mathcal{B}^{(L)}(z)\right)^T$
$(-1,\, 0,\, 0)$	$n_0 + n_\infty$	$\begin{pmatrix}(1-z)^{-2\,n_\infty - 2\,n_0 + 1} & 0\end{pmatrix}$
	$1 - (n_0 + n_\infty)$	$\begin{pmatrix}1 & 0\end{pmatrix}$
	$n_0 - n_\infty$	$\begin{pmatrix}1 & (-z)^{1-2\,n_0}\end{pmatrix}$
	$-n_0 + n_\infty$	$\begin{pmatrix}1 & 0\end{pmatrix}$
$(-1,\, 1,\, 0)$	$n_0 + n_\infty$	$\begin{pmatrix}\mathrm{F}\left(2\,n_\infty + 2\,n_0 - 1,\, 2\,n_0 + 1;\, 2\,n_0;\, z\right) & 0\end{pmatrix}$
	$1 - (n_0 + n_\infty)$	$\begin{pmatrix}1 & 0\end{pmatrix}$
	$n_0 - n_\infty$	$\begin{pmatrix}0 & (-z)^{1-2\,n_0}\end{pmatrix}$
	$-n_0 + n_\infty$	$\begin{pmatrix}0 & (1-z)^{2\,n_0 - 2\,n_\infty}\,(-z)^{1-2\,n_0}\end{pmatrix}$
$(0,\, 0,\, 0)$	$n_0 + n_\infty$	$\begin{pmatrix}(1-z)^{-2\,n_\infty - 2\,n_0} & 0\end{pmatrix}$
	$1 - (n_0 + n_\infty)$	$\begin{pmatrix}0 & (1-z)^{2\,n_\infty + 2\,n_0 - 2}\,(-z)^{1-2\,n_0}\end{pmatrix}$
	$n_0 - n_\infty$	$\begin{pmatrix}(1-z)^{2\,n_\infty - 2\,n_0} & 0\end{pmatrix}$
	$-n_0 + n_\infty$	$\begin{pmatrix}1 & 0\end{pmatrix}$
$(-1,\, 1,\, 1)$	$n_0 + n_\infty$	$\begin{pmatrix}(1-z)^{-2\,n_\infty - 2\,n_0} + 1 & 0\end{pmatrix}$
	$1 - (n_0 + n_\infty)$	$\begin{pmatrix}1 & 0\end{pmatrix}$
	$n_0 - n_\infty$	$\begin{pmatrix}0 & (-z)^{-2\,n_0}\,\mathrm{F}\left(-1,\, 1 - 2\,n_\infty;\, 1 - 2\,n_0;\, z\right)\end{pmatrix}$
	$-n_0 + n_\infty$	$\begin{pmatrix}0 & (1-z)^{-2\,n_\infty + 2\,n_0 + 1}\,(-z)^{-2\,n_0}\end{pmatrix}$
$(0,\, 0,\, 1)$	$n_0 + n_\infty$	$\begin{pmatrix}0 & (-z)^{-2\,n_0}\end{pmatrix}$
	$1 - (n_0 + n_\infty)$	$\begin{pmatrix}0 & (1-z)^{2\,n_\infty + 2\,n_0 - 1}\,(-z)^{-2\,n_0}\end{pmatrix}$
	$n_0 - n_\infty$	$\begin{pmatrix}0 & (-z)^{-2\,n_0}\end{pmatrix}\Big)$
	$-n_0 + n_\infty$	$\begin{pmatrix}1 & 0\end{pmatrix}$
$(0,\, 1,\, 1)$	$n_0 + n_\infty$	$\begin{pmatrix}(1-z)^{-2\,n_\infty - 2\,n_0} & 0\end{pmatrix}$
	$1 - (n_0 + n_\infty)$	$\begin{pmatrix}0 & (1-z)^{2\,n_\infty + 2\,n_0 - 2}\,(-z)^{-2\,n_0}\end{pmatrix}$
	$n_0 - n_\infty$	$\begin{pmatrix}0 & (-z)^{-2\,n_0}\end{pmatrix}\Big)$
	$-n_0 + n_\infty$	$\begin{pmatrix}0 & (1-z)^{2\,n_0 - 2\,n_\infty}\,(-z)^{-2\,n_0}\end{pmatrix}$

Table 2.3: Abelian limit of the solutions

previous computation, however the right sector seems to give different solutions when different Abelian limits are taken. Studying all possible solutions we find that all of them give the same answer in the limit $m_{\vec{t}} \to 0$, i.e. both $\mathcal{B}^{(R)} = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ and $\mathcal{B}^{(R)} = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$.[21] The difference is the solution obtained from $n_0 > m_0$, $n_1 > m_1$ and $n_\infty > m_\infty$ or $n_0 > m_0$, $\hat{n}_1 < \hat{m}_1$ and $\hat{n}_\infty < \hat{m}_\infty$. In any case the solution is factorised in the form $\mathcal{B}^{(L)}(z) \otimes \begin{pmatrix} C & C' \end{pmatrix}^T$ which is expected since the right sector plays no role.

2.5.4 Relating the Abelian Angles with the Group Parameters

Using the explicit form of the SO(4) and SU(2) × SU(2) matrices, we can verify that when the left and right SU(2) parameters are $\vec{n} = n^3\,\vec{k}$ and $\vec{m} = m^3\,\vec{k}$ the rotation of the D-branes in the plane 14 is a SO(2) element

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}, \qquad \theta = n^3 - m^3, \tag{2.126}$$

while in plane 23 the angle is $\theta = n^3 + m^3$. Comparing with the case $m = 0$ given in (2.123), we then guess that the general relation between the group parameters and the usual Abelian angles is:

$$\begin{aligned} \varepsilon_{\vec{t}} &= n^3_{\vec{t}} - m^3_{\vec{t}} + \theta(-(n^3_{\vec{t}} - m^3_{\vec{t}})), \\ \varphi_{\vec{t}} &= n^3_{\vec{t}} + m^3_{\vec{t}} + \theta(-(n^3_{\vec{t}} + m^3_{\vec{t}})). \end{aligned} \tag{2.127}$$

2.5.5 Recovering the Abelian Result: an Example

To show that the construction of the Abelian limit is indeed working, we consider the first case in Section 2.4.7 with $n_1 = 1 - (n_0 + n_\infty)$ and $m_1 = -m_0 + m_\infty$. This leads to two independent rational functions of ω_z:

$$\begin{aligned} \partial_{\omega_z}\mathcal{X}(\omega_z) &= \begin{pmatrix} i\partial_{\omega_z}\bar{\mathcal{Z}}^1(\omega_z) & \partial_{\omega_z}\mathcal{Z}^2(\omega_z) \\ \partial_{\omega_z}\bar{\mathcal{Z}}^2(\omega_z) & i\partial_{\omega_z}\mathcal{Z}^1(\omega_z) \end{pmatrix} \\ &= \begin{pmatrix} i\partial_{\omega_z}(\mathcal{X}^1(\omega_z) - i\mathcal{X}^4(\omega_z)) & \partial_{\omega_z}(\mathcal{X}^2(\omega_z) + i\mathcal{X}^3(\omega_z)) \\ \partial_{\omega_z}(\mathcal{X}^2(\omega_z) - i\mathcal{X}^3(\omega_z)) & i\partial_{\omega_z}(\mathcal{X}^1(\omega_z) + i\mathcal{X}^4(\omega_z)) \end{pmatrix} \\ &= \begin{pmatrix} 0 & C_1\,(1-\omega_z)^{\varepsilon_1 - 1}\,(-\omega_z)^{\varepsilon_0 - 1} \\ 0 & C_2\,(1-\omega_z)^{-\varphi_1}\,(-\omega_z)^{-\varphi_0} \end{pmatrix}, \end{aligned} \tag{2.128}$$

where C_1 and C_2 are constants as in (2.121). This is the known result in the presence of Abelian rotations of the D-branes: we have two different U(1) sectors undergoing two different rotations $\mathrm{U}(1)_1 \times \mathrm{U}(1)_2 \subset \mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$. In particular we used (2.127) to write the relation between the usual Abelian angles and the group parameters as

$$\varepsilon_0 = n_0 - m_0, \qquad \varepsilon_1 = n_1 - m_1, \qquad \varepsilon_\infty = n_\infty + m_\infty \tag{2.129}$$

such that $\sum_t \varepsilon_{\vec{t}} = 1$, and

$$\varphi_0 = n_0 + m_0, \qquad \varphi_1 = n_1 + m_1, \qquad \varphi_\infty = n_\infty - m_\infty, \tag{2.130}$$

[21] We write "possible solutions" because $m_1 = 1 - (m_0 + m_\infty)$ is not.

where $\sum_{\vec{t}} \varphi_{\vec{t}} = 2$, in order to approach the usual notation in the literature. As usual we have $\partial_{\omega_z} \mathcal{Z}^1(\omega_z) \neq \left[\partial_{\omega_z} \overline{\mathcal{Z}}^1(\omega_z)\right]^*$.

We can now build the Abelian solution to show the analytical structure of the limit. We have

$$\begin{pmatrix} i\overline{Z}^1(u,\,\overline{u}) & Z^2(u,\,\overline{u}) \\ \overline{Z}^2(u,\,\overline{u}) & iZ^1(u,\,\overline{u}) \end{pmatrix} = \begin{pmatrix} i\overline{f}^1_{(\bar{t}-1)} + i\int_0^{\overline{\omega}_{\overline{u}}} \mathrm{d}\omega\ \partial_\omega \mathcal{Z}^1 & f^2_{(\bar{t}-1)} + \int_0^{\omega_u} \mathrm{d}\omega\ \partial_\omega \mathcal{Z}^2 \\ \overline{f}^2_{(\bar{t}-1)} + \int_0^{\overline{\omega}_{\overline{u}}} \mathrm{d}\omega\ \partial_\omega \mathcal{Z}^2 & if^1_{(\bar{t}-1)} + i\int_0^{\omega_u} \mathrm{d}\omega\ \partial_{\omega_z} \mathcal{Z}^1 \end{pmatrix} \tag{2.131}$$

where we chose $R_{(\bar{t})} = \mathbb{1}_4$ so that $U_{(\bar{t})}$ in (2.18) is mapped to $(i\sigma_1, i\sigma_1) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$. Notice however that $\vec{n}_{\vec{t}} = n^3_{\vec{t}}\,\vec{k}$ implies that $v^3_{(t)} = 0$ in (2.40). Hence U_L and U_R are always off diagonal and their action on (2.128) is to fill the first column. From the previous relations we can also recover the usual holomorphicity $\overline{Z}^1(\overline{u}) = \left[Z^1(u)\right]^*$ of the sector with $\sum_{\vec{t}} \varepsilon_{\vec{t}} = 1$ and $\overline{Z}^2(\overline{u}) = \left[Z^2(u)\right]^*$ of the sector with $\sum_{\vec{t}} \varphi_{\vec{t}} = 2$.

2.5.6 Abelian Limits

From the example in the previous section it is possible to consider both cases given in Section 2.4.7 and all possible combinations of the expression of n_1 and m_1 for a total of $2 \times 4 \times 4 = 32$ possible combinations. In almost all cases (in fact all but six) the solution in spinorial formalism is a 2×2 matrix which has two non vanishing entries, hence two independent Abelian solutions. In the remaining cases the matrix has only one non vanishing entry but the constraints on n and m are not compatible, thus they should not be considered. In the first case encountered in Section 2.4.7 the inconsistent combinations are $\{n_1 = n_0 + n_\infty,\ m_1 = 1 - (m_0 + m_\infty)\}$ and $\{n_1 = 1 - (n_0 + n_\infty),\ m_1 = 1 - (m_0 + m_\infty)\}$. In the second case in Section 2.4.7 the incompatible constraints appear when $n_1 = -n_0 + n_\infty$.

2.6 The Physical Interpretation

In this section we show some consequences of the explicit classical solution for the phenomenology of models involving D-branes intersecting at angles. In particular we focus on the value of the action which plays a fundamental role in the hierarchy of the Yukawa couplings.

2.6.1 Rewriting the Action

Using the classical solution previously computed, it is possible to compute the classical action of the bosonic string and show its contribution to the correlation functions of twist fields and Yukawa couplings. We use the equations of motion (2.12) to simplify the action (2.11). We get:

$$4\pi\alpha'\, S_{\mathbb{R}^4}\Big|_{\text{on-shell}} = i\sum_{t=1}^{3}\sum_{m\in\{3,4\}} g_{(t)\,m}\int_{x_{(t)}}^{x_{(t-1)}} \mathrm{d}x\,\big(R_{(t)}\big)_{mI}\,\big(X'_L(x)-X'_R(x)\big)^I\Big|_{y=0^+}, \tag{2.132}$$

where indices $I = 1, 2, 3, 4$ are summed over and $m = 3, 4$ are the transverse directions in the well adapted frame with respect to the D-brane. As the number of D-branes is defined modulo $N_B = 3$, $D_{(1)}$ is split on two separate intervals:

$$\big[x_{(1)},\, x_{(3)}\big] = \big[x_{(1)},\, +\infty\big) \cup \big(-\infty,\, x_{(3)}\big]\,, \tag{2.133}$$

as it is visually shown in Figure 2.2. For $x_{(t)} < x < x_{(t-1)}$ we have:

$$X(x+iy,\, x-iy) = X^*(x+iy,\, x-iy) \quad \Rightarrow \quad X^*_L(x-iy) = X_R(x-iy) + Y, \tag{2.134}$$

where $Y \in \mathbb{R}$ is a constant factor which cannot depend on the particular D-brane $D_{(t)}$. In fact the continuity of $X_L(u)$ and $X_R(\overline{u})$ on the worldsheet intersection point ensures that

$$\lim_{x\to x^+_{(t)}} X(x,x) = \lim_{x\to x^-_{(t)}} X(x,x), \tag{2.135}$$

which does not allow Y to depend on the specific D-brane while the reality of $X(u,\overline{u})$ implies that $\operatorname{Im} Y = 0$. Now (2.132) becomes:

$$\begin{aligned} 4\pi\alpha'\, S_{\mathbb{R}^4}\Big|_{\text{on-shell}} &= -2\sum_{t=1}^{3}\sum_{m\in\{3,4\}} g_{(t)\,m}\,\operatorname{Im}\big(R_{(t)}\big)_{mI} X^I_L(x+i0^+)\Big|_{x=x_{(t)}}^{x=x_{(t-1)}} \\ &= -2\sum_{t=1}^{3} g^{(\perp)}_{(t)\,I}\,\operatorname{Im} X^I_L(x+i0^+)\Big|_{x=x_{(t)}}^{x=x_{(t-1)}} \in \mathbb{R}, \end{aligned} \tag{2.136}$$

where $g^{(\perp)}_{(t)\,I} = \sum_{m\in\{3,4\}}\left(R^{-1}_{(t)}\right)_{mI} g_{(t)\,m}$ is the transverse shift of $D_{(t)}$ in the global coordinates of the target space:

$$g^{(\perp)}_{(t)\,I}\,\big(f_{(t-1)} - f_{(t)}\big)^I = 0. \tag{2.137}$$

2.6.2 Holomorphic Case

In this case there are global complex coordinates for which the string solution is holomorphic:

$$Z^i(u,\overline{u}) = Z^i_L(u), \qquad \overline{Z}^i(u,\overline{u}) = \bar{Z}^i(\overline{u}) = \big(Z^i_L(u)\big)^*, \tag{2.138}$$

where $i = 1$ in the Abelian case and $i = 1, 2$ in the SU(2) case. We also have $f^i_{(t)} = Z^i_L(x_{(t)} + i\,0^+)$. Equations (2.137) and (2.136) then become

$$\operatorname{Re}\big(g^{(\perp)}_{(t)\,i}\,\big(f_{(t-1)} - f_{(t)}\big)^i\big) = 0, \tag{2.139}$$

$$4\pi\alpha'\, S_{\mathbb{R}^4}\Big|_{\text{on-shell}} = -2\sum_{t=1}^{3}\operatorname{Im}\big(g^{(\perp)}_{(t)\,i}\,\big(f_{(t-1)} - f_{(t)}\big)^i\big), \tag{2.140}$$

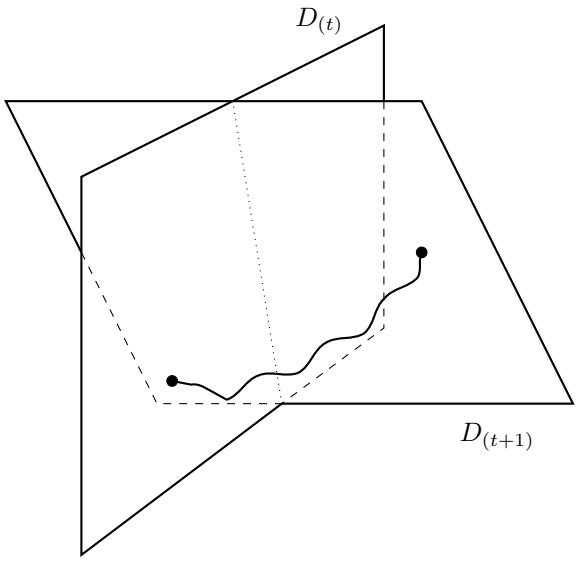


Figure 2.7: Pictorial 3-dimensional representation of two D2-branes intersecting in the Euclidean space $\mathbb{R}^3$ along a line (in $\mathbb{R}^4$ the intersection is a point since the co-dimension of each D-brane is 2): since it is no longer constrained on a bi-dimensional plane, the string must be deformed in order to stretch between two consecutive D-branes. Its action can be larger than the planar area.

where the last equation shows that the action can be expressed using just the global data.

In the Abelian scenario we can further simplify the action and give a clear geometrical meaning. Given to complex numbers $a, b \in \mathbb{C}$ such that $\mathrm{Re}(a^* b) = 0$ then $\mathrm{Im}(a^* b) = \pm |a||b|$. This can be seen either by direct computation or by using a U(1) rotation to set b equal to $|b|$. Since the action is positive then we can write

$$S_{\mathbb{R}^4}\Big|_{\text{on-shell}} = \frac{1}{2\pi\alpha'} \sum_{t=1}^{3} \left(\frac{1}{2} \left| g_{(t)}^{(\perp)} \right| \left| f_{(t-1)} - f_{(t)} \right| \right), \tag{2.141}$$

where a factor $\frac{1}{2}$ comes from raising the complex index in $g^{(\perp)}_{(t)\,1}$. The right hand side of the previous expression is the sum of the areas of the triangles having the interval between two intersection points on a given D-brane $D_{(t)}$ as base and the distance between the D-brane and the origin as height. A visual reference can be found in Figure 2.1.

For the SU(2) case we can use a rotation to map $(f_{(t-1)} - f_{(t)})^i$ to the form $\left\| f_{(t-1)} - f_{(t)} \right\| \delta^i_1$. Each term of the action can be interpreted again as an area of a triangle where the distance between the interaction points is the base.

2.6.3 General Case and Intuitive Explanation

In the general case there does not seem to be any possible way of computing the action (2.136) in term of the global data. Most probably the value of the action is larger than in the holomorphic

case since the string is no longer confined to a plane. Given the nature of the rotation its worldsheet has to bend in order to be attached to the D-brane as pictorially shown in Figure 2.7 in the case of a 3-dimensional space. The general case we considered then differs from the known factorised case by an additional contribution in the on-shell action which can be intuitively understood as a small "bump" of the string worldsheet in proximity of the boundary.

3 Fermions With Boundary Defects

3.1 Motivation

As previously pointed out, the computation of quantities such as Yukawa couplings involves correlators of excited spin and twist fields. After the analysis of the main contribution to amplitudes involving twist fields at the intersection of D-branes, we focus on the computation of correlators of (excited) spin fields. This has been a research subject for many years until the formulation found in the seminal paper by Friedan, Martinec and Shenker [9] based on bosonization. In general the available techniques allow to compute only correlators involving Abelian configurations, that is configurations which can be factorized in sub-configurations having U(1) symmetry. Non Abelian cases have also been considered [77]–[80], though their mathematical formulation is by far more complicated.

Despite the existence of an efficient method based on bosonization [9] and old methods based on the Reggeon vertex [68], [82]–[85], we take into examination the computation of spin field correlators and propose a new method to compute them. We hope to be able to extend this approach to correlators involving twist fields and non Abelian spin and twist fields. We would also like to investigate the reason of the non existence of an approach equivalent to bosonization for twist fields. At the same time we are interested to explore what happens to a CFT in presence of defects. It turns out that despite the defects it is still possible to define a radial time dependent stress-energy tensor which satisfies the canonical O.P.E. Moreover the boundary changing defects in the construction can be associated with excited spin fields enabling the computation of correlators involving excited spin fields without resorting to bosonization.

3.2 Point-like Defect CFT: the Minkowskian Formulation

Let $(\tau, \sigma) \in \Sigma = (-\infty, +\infty) \times [0, \pi]$ define a strip with Lorentzian metric and consider N_f massless complex fermions ψ^i such that $i = 1, 2, \ldots, N_f$. Their two-dimensional Minkowski action defined on the strip Σ is:

$$S = \frac{T}{2} \int_{-\infty}^{+\infty} d\tau \int_{0}^{\pi} d\sigma \left(\frac{1}{2} \overline{\psi}_i(\tau, \sigma) \left(-i\gamma^\alpha \overleftrightarrow{\partial}_\alpha \right) \psi^i(\tau, \sigma) \right), \tag{3.1}$$

where the gamma matrices are

$$\gamma^\tau = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} = -\gamma_\tau, \qquad \gamma^\sigma = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} = \gamma_\sigma, \tag{3.2}$$

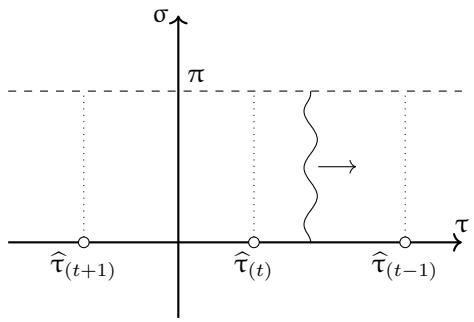


Figure 3.1: Propagation of the string in the presence of the worldsheet defects.

and the components of the massless fermions are

$$\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}, \qquad \overline{\psi} = \psi^\dagger \gamma^\tau = \begin{pmatrix} -\psi_-^* & \psi_+^* \end{pmatrix}. \tag{3.3}$$

We then define the lightcone coordinates $\xi_\pm = \tau \pm \sigma$ such that $\partial_\pm = \frac{1}{2}(\partial_\tau \pm \partial_\sigma)$. In components the action reads:

$$S = i\frac{T}{4} \int_{-\infty}^{+\infty} \mathrm{d}\xi_+ \int_{-\infty}^{+\infty} \mathrm{d}\xi_- \left(\psi^*_{-,i} \overset{\leftrightarrow}{\partial_+} \psi^i_- + \psi^*_{+,i} \overset{\leftrightarrow}{\partial_-} \psi^i_+ \right), \tag{3.4}$$

so the E.O.M. are:

$$\begin{aligned} \partial_- \psi^i_+(\xi_+, \xi_-) = \partial_+ \psi^i_-(\xi_+, \xi_-) = 0, \\ \partial_- \psi^*_{+,i}(\xi_+, \xi_-) = \partial_+ \psi^*_{-,i}(\xi_+, \xi_-) = 0. \end{aligned} \tag{3.5}$$

Their solutions are the "holomorphic" functions $\psi^i_+(\xi_+)$ and $\psi^i_-(\xi_-)$ and their complex conjugates.[22] The boundary conditions are instead:

$$\left(\delta\psi^*_{+,i}\, \psi^i_+ + \delta\psi^*_{-,i}\, \psi^i_- - \psi^*_{+,i}\, \delta\psi^i_+ - \psi^*_{-,i}\, \delta\psi^i_- \right) \Big|_{\sigma=0}^{\sigma=\pi} = 0. \tag{3.6}$$

We solve the constraint imposing the non trivial relations:

$$\begin{cases} \psi^i_-(\tau, 0) = \left(R_{(t)}\right)^i{}_j \psi^j_+(\tau, 0), & \tau \in \left(\widehat{\tau}_{(t)}, \widehat{\tau}_{(t-1)}\right), \\ \psi^i_-(\tau, \pi) = -\psi^i_+(\tau, \pi), & \tau \in \mathbb{R}, \end{cases} \tag{3.7}$$

where $t = 1, 2, \ldots, N$. This way we introduce N zero-dimensional defects on the boundary, pictorially shown in Figure 3.1. They are located on the strip at $(\widehat{\tau}_{(t)}, 0) \in \Sigma$ such that $\widehat{\tau}_{(t)} < \widehat{\tau}_{(t-1)}$ with $\widehat{\tau}_{N+1} = -\infty$ and $\widehat{\tau}_0 = +\infty$. Their characterisation is given by N matrices $R_{(t)} \in \mathrm{U}(N_f)$.

In most of this paper we want the in- and out-vacua to be the usual NS vacuum. We thus choose the boundary condition at $\sigma = \pi$ so that when there are no defects the system describes NS fermions. We require also the cancellation of the action of the defects at $\widehat{\tau} = \pm\infty$, that is:

$$R_{(N)} R_{(N-1)} \ldots R_{(1)} = \mathbb{1}. \tag{3.8}$$

[22] Notice that ψ^* is indeed the complex conjugate of the field ψ, while it will no longer be the case in the Euclidean formalism.

More general cases where the asymptotic vacua are twisted can be worked out in similar fashion.

In order to connect to the Euclidean formulation we introduce N_f "double fields" Ψ^i.[23] They can be obtained by gluing ψ^i_+ and ψ^i_- along the $\sigma = \pi$ boundary and labeled by an index $i = 1, 2, \ldots, N_f$:

$$\Psi^i(\tau, \phi) = \begin{cases} \psi^i_+(\tau, \phi), & 0 \le \phi \le \pi \\ -\psi^i_-(\tau, 2\pi - \phi), & \pi \le \phi \le 2\pi \end{cases} \tag{3.9}$$

where $0 \le \phi \le 2\pi$. The boundary conditions become:

$$\Psi^i(\tau, 2\pi) = -\big(R_{(t)}\big)^i{}_j \Psi^j(\tau, 0), \qquad \tau \in \big(\widehat{\tau}_{(t)}, \widehat{\tau}_{(t-1)}\big). \tag{3.10}$$

Using the equation of motion we get $\Psi^i(\tau, \phi) = \Psi^i(\tau + \phi)$ and the boundary conditions become the (pseudo-)periodicity conditions

$$\Psi^i(\tau + 2\pi) = -\big(R_{(t)}\big)^i{}_j \Psi^j(\tau), \qquad \tau \in \big(\widehat{\tau}_{(t)}, \widehat{\tau}_{(t-1)}\big). \tag{3.11}$$

The main issue is now to expand Ψ in a basis of modes and proceed to its quantization. Even in the simplest case $N_f = 1$ the task of finding the Minkowskian modes turns out to be fairly complicated. It is however possible to overcome the issue in the Euclidean formalism.

3.3 Conserved Product and Charges

In order to promote the theory to its quantum formulation, we define a procedure to build a Fock space of states in the Heisenberg formalism. Equal time anti-commutation relations must then be invariant in time. We thus need a time independent internal product to extract the creation and annihilation operators and expand the fields on the basis of modes.

3.3.1 Conserved Product and Current

Start from a generic conserved current

$$j(\tau, \sigma) = j_\tau(\tau, \sigma)\, \mathrm{d}\tau + j_\sigma(\tau, \sigma)\, \mathrm{d}\sigma\,, \tag{3.12}$$

and consider

$$\star j = j_\sigma\, \mathrm{d}\tau + j_\tau\, \mathrm{d}\sigma \quad \Rightarrow \quad \mathrm{d}(\star j) = (\partial_\tau j_\tau - \partial_\sigma j_\sigma)\, \mathrm{d}\tau\, \mathrm{d}\sigma\,, \tag{3.13}$$

where $\star$ is the Hodge dual operator. Integration over the surface $\Sigma' = [\tau_i, \tau_f] \times [0, \pi]$ yields:

$$\int_{\Sigma'} \mathrm{d}(\star j) = \int_{\partial\Sigma'} \star j = 0 \qquad \Leftrightarrow \qquad \int_0^\pi \mathrm{d}\sigma\, j_\tau \Big|^{\tau=\tau_i}_{\tau=\tau_f} = \int_{\tau_i}^{\tau_f} \mathrm{d}\tau\, j_\sigma \Big|^{\sigma=0}_{\sigma=\pi}. \tag{3.14}$$

The current $j_\tau(\tau, \sigma)$ is thus conserved in time if

$$\int_{\tau_i}^{\tau_f} \mathrm{d}\tau \left(j_\sigma \Big|_{\sigma=\pi} - j_\sigma \Big|_{\sigma=0} \right) = 0. \tag{3.15}$$

[23] In this case they correspond to the fields ψ^i_+.

In this case we can define

$$Q = \int_0^\pi \mathrm{d}\sigma\, j_\tau(\tau,\, \sigma) \tag{3.16}$$

as conserved quantity (that is $\partial_\tau Q = 0$). We now consider explicitly the symmetries of the action (3.4). We focus on diffeomorphism invariance and $\mathrm{U}(N_f)$ flavour symmetries of the bulk theory leading to the stress-energy tensor and a vector current.

3.3.2 Flavour Vector Current

Consider first the $\mathrm{U}(N_f)$ vector current of the flavour symmetry in (3.1). We write it as

$$j^a_\alpha(\tau,\, \sigma) = (\mathrm{T}^a)^i{}_j\, \overline{\psi}_i(\tau,\, \sigma)\, \gamma_\alpha\, \psi^j(\tau,\, \sigma), \tag{3.17}$$

where T^a is a generator of $\mathrm{U}(N_f)$ such that $a = 1,\, 2,\, \ldots,\, N_f^2$.[24] In components we have:

$$j^a_\tau(\tau,\, \sigma) = (\mathrm{T}^a)^i{}_j \left(\psi^*_{+,\,i}\psi^j_+ + \psi^*_{-,\,i}\psi^j_-\right) \tag{3.18}$$

$$j^a_\sigma(\tau,\, \sigma) = (\mathrm{T}^a)^i{}_j \left(\psi^*_{+,\,i}\psi^j_+ - \psi^*_{-,\,i}\psi^j_-\right). \tag{3.19}$$

In order to define a conserved charge, we require:

$$\int_{\tau_i}^{\tau_f} \mathrm{d}\tau \left(j^a_\sigma\Big|_{\sigma=\pi} - j^a_\sigma\Big|_{\sigma=0}\right) = 0, \tag{3.20}$$

where

$$j^a_\sigma(\tau,\, \sigma)\Big|_{\sigma=\pi} \equiv 0 \tag{3.21}$$

using the boundary conditions (3.6). Moreover we have:

$$j^a_\sigma(\tau,\, \sigma)\Big|_{\sigma=0} = \left[\psi^*_+ \left(\mathrm{T}^a - R^\dagger_{(t)}\mathrm{T}^a R_{(t)}\right)\psi_+\right]_{\sigma=0}, \qquad \tau \in \left(\widehat{\tau}_{(t)}, \widehat{\tau}_{(t-1)}\right). \tag{3.22}$$

In general

$$j^a_\sigma(\tau,\, \sigma)\Big|_{\sigma=0} = 0 \qquad \Leftrightarrow \qquad \mathrm{T}^a \propto \mathbb{1} \tag{3.23}$$

so that $R^\dagger_{(t)}\mathrm{T}^a = \mathrm{T}^a R^\dagger_{(t)}$. This shows that the presence of the point-like defects on the worldsheet breaks the $\mathrm{U}(N_f)$ symmetry down to a $\mathrm{U}(1)$ phase.[25] The $\mathrm{U}(1)$ vector current then defines a conserved charge for a restricted class of functions.

Let α and β be two arbitrary (bosonic) solutions to the E.O.M. (3.5), we define a product

$$\langle \alpha, \beta\rangle = \mathcal{N} \int_0^\pi \mathrm{d}\sigma \left(\alpha^*_{+,\,i}\beta^i_+ + \alpha^*_{-,\,i}\beta^i_-\right), \tag{3.24}$$

[24] The results however are more general and hold for a generic matrix M taking the place of any of the generators T^a. Spinors ψ and $\overline{\psi}$ can also be generalized to two different and arbitrary solutions of the E.O.M. (3.5) while keeping the current conserved.

[25] The symmetry is $\mathrm{SO}(N_f) \times \mathrm{SO}(N_f)$ if we consider Majorana-Weyl fermions.

where $\mathcal{N} \in \mathbb{R}$ is a normalisation constant and the integrand must not present non integrable singularities. The product is such that $\langle \alpha, \beta \rangle = \langle \alpha, \beta \rangle^*$. We can also rewrite the result to the double fields defined in (3.9). Let in fact A and B be the "double fields" corresponding to α and β respectively:

$$\langle \alpha, \beta \rangle = \mathcal{N} \int_0^{2\pi} \mathrm{d}\phi \; A_i^*(\tau + \phi)\, B^i(\tau + \phi). \tag{3.25}$$

3.3.3 Stress-Energy Tensor

Consider the stress-energy tensor of the bulk theory. Using the usual Nöther's procedure we get the on-shell non vanishing components:

$$\begin{aligned} \mathcal{T}_{++}(\xi_+) &= -i\frac{T}{4}\psi^*_{+,\,i}(\xi_+)\overset{\leftrightarrow}{\partial}_+\psi^i_+(\xi_+), \\ \mathcal{T}_{--}(\xi_-) &= -i\frac{T}{4}\psi^*_{-,\,i}(\xi_-)\overset{\leftrightarrow}{\partial}_-\psi^i_-(\xi_-). \end{aligned} \tag{3.26}$$

As always the boundary of Σ breaks the symmetry for translations in the σ direction, while the defects break the time translations: the Hamiltonian is therefore time-dependent but it is constant between two consecutive point-like defects. In fact, from the definition of the stress-energy tensor, we can in principle build the hypothetical charges:

$$\mathrm{H}(\tau) \;=\; \int_0^{\pi} \mathrm{d}\sigma\, \mathcal{T}_{\tau\tau}(\tau,\,\sigma) = \int_0^{\pi} \mathrm{d}\sigma\, (\mathcal{T}_{++}(\tau+\sigma) + \mathcal{T}_{--}(\tau-\sigma)), \tag{3.27}$$

$$\mathrm{P}(\tau) \;=\; \int_0^{\pi} \mathrm{d}\sigma\, \mathcal{T}_{\tau\sigma}(\tau,\,\sigma) = \int_0^{\pi} \mathrm{d}\sigma\, (\mathcal{T}_{++}(\tau+\sigma) - \mathcal{T}_{--}(\tau-\sigma)), \tag{3.28}$$

which are conserved if (3.15) holds.

We order the point-like defects as $\widehat{\tau}_{(t_0-1)} < \tau_i \leq \widehat{\tau}_{(t_0)} < \widehat{\tau}_{(t_N)} \leq \tau_f < \widehat{\tau}_{(t_N+1)}$. For the linear momentum P the condition of conservation reads:[26]

$$\begin{aligned} &\int_{\tau_i}^{\tau_f} \mathrm{d}\tau\, (\mathcal{T}_{++}(\tau+\sigma) + \mathcal{T}_{--}(\tau-\sigma))\Big|_{\sigma=0}^{\sigma=\pi} \\ &= -i\frac{T}{4}\int \Delta\tau \left(2\,\psi^*_{+,\,i}\overset{\leftrightarrow}{\partial}_\tau\psi^i_+\Big|_{\sigma=0}^{\sigma=\pi} - \psi^*_{+,\,i}\left(R^\dagger_{(t)}\overset{\leftrightarrow}{\partial}_\tau R_{(t)}\right)^i_j \psi^j_+\Big|_{\sigma=0} \right) \neq 0. \end{aligned} \tag{3.29}$$

The corresponding condition for the Hamiltonian H is:

$$\begin{aligned} &\int_{\tau_i}^{\tau_f} \mathrm{d}\tau\, (\mathcal{T}_{++}(\tau+\sigma) - \mathcal{T}_{--}(\tau-\sigma))\Big|_{\sigma=0}^{\sigma=\pi} \\ &= -i\frac{T}{4}\int \Delta\tau \left(\psi^*_{+,\,i}\left(R^\dagger_{(t)}\overset{\leftrightarrow}{\partial}_\tau R_{(t)}\right)^i_j \psi^j_+\Big|_{\sigma=0} \right) = 0 \quad \Leftrightarrow \quad (\tau_i, \tau_f) \in \left(\widehat{\tau}_{(t)}, \widehat{\tau}_{(t-1)}\right). \end{aligned} \tag{3.30}$$

[26] Notice that in the second term of the second line the differentiation with respect to τ is acting only on $R_{(t)}$ and $R^\dagger_{(t)}$.

In both cases we used the shorthand graphical notation

$$\int \Delta\tau = \left(\int_{\tau_i}^{\widehat{\tau}_{t_0}} + \sum_{t=t_0}^{t_N-1} \int_{\widehat{\tau}_{(t)}}^{\widehat{\tau}_{(t+1)}} + \int_{\widehat{\tau}_N}^{\tau_f} \right) \mathrm{d}\tau \tag{3.31}$$

for simplicity.

These relations therefore prove that the generator of the σ-translations (3.28) is not conserved in time because of the boundary conditions, while the time evolution operator H is only piecewise conserved and therefore globally time dependent.

3.4 Basis of Solutions and Dual Modes

Let $\{\psi^i_{n,\pm}\}_{n\in\mathbb{Z}}$ be a complete basis of modes such that:

$$\begin{cases} \psi^i_{n,+}(\tau,0) = \big(R_{(t)}\big)^i_{\ j}\psi^j_{n,-}(\tau,0) & \text{for} \quad \tau \in \big(\widehat{\tau}_{(t)}, \widehat{\tau}_{(t-1)}\big) \\ \psi^i_{n,+}(\tau,\pi) = -\psi^i_{n,-}(\tau,\pi) & \text{for} \quad \tau \in \mathbb{R} \end{cases} . \tag{3.32}$$

These fields are related to a complete basis of the modes of the "double field" Ψ^i_n as in (3.9). The modes ψ_n (and their counterparts Ψ_n) are a basis of solutions of the E.O.M. and the boundary conditions for $\tau \in \mathbb{R} \setminus \{\widehat{\tau}_{(t)}\}_{0\le t\le N}$. The fields ψ^i (and the fields Ψ^i) are then a superposition of such modes:

$$\psi^i_{\pm}(\xi_{\pm}) = \sum_{n=-\infty}^{+\infty} b_n\, \psi^i_{n,\pm}(\xi_{\pm}) \qquad \Rightarrow \qquad \Psi^i(\xi) = \sum_{n=-\infty}^{+\infty} b_n\, \Psi^i_n(\xi). \tag{3.33}$$

In order to extract the "coefficients" b_n we first introduce the dual basis ${}^*\psi_{n,\pm}$ (and ${}^*\Psi_n$) in an abstract sense such that:

- the dual fields ${}^*\psi_{n,\pm}$ (and ${}^*\Psi_n$) must be solutions to the E.O.M.,
- the dual fields ${}^*\psi_{n,\pm}$ (and ${}^*\Psi_n$) can differ from $\psi_{n,\pm}$ (and Ψ_n) in their behavior at the boundary,
- the functional form of ${}^*\psi_{n,\pm}$ (and ${}^*\Psi_n$) is fixed by the request of time invariance of the usual anti-commutation relations $\big[b_n,\, b^\dagger_m\big]_+$ (that is b_n and $b^\dagger_n$ can evolve in time, but their anti-commutation relations must remain constant).

We then define the conserved product for the "double fields" (3.25) in such a way that:

$$\langle\!\langle {}^*\Psi_n\ ,\ \Psi_m \rangle\!\rangle \Big|_{\tau=\tau_0} = \mathcal{N} \int_0^{2\pi} \mathrm{d}\sigma\ {}^*\Psi^*_{n,i}(\tau+\sigma)\, \Psi^i_m(\tau+\sigma) = \delta_{n,m}. \tag{3.34}$$

In the previous expression we changed the notation of the product. We are in fact dealing with the space of solutions whose basis is $\{\Psi_n\}$ and a dual space with basis $\{{}^*\Psi_n\}$ which is not

required to span entirely the original space but only to be a subset of it in order to be able to compute the anti-commutation relations among the annihilation and construction operators in a well defined way as in (3.38). Given the previous product we can extract the operators as

$$\langle\!\langle {}^*\Psi_n \, , \, \Psi \rangle = b_n, \tag{3.35}$$

$$\langle\!\langle {}^*\Psi^*_n \, , \, \Psi^* \rangle = b^\dagger_n. \tag{3.36}$$

As a consequence of the canonical anti-commutation relations

$$\big[\Psi^i(\tau,\, \sigma), \Psi^*_j(\tau,\, \sigma')\big]_+ = \frac{2}{T}\, \delta^i{}_j\, \delta(\sigma - \sigma'), \tag{3.37}$$

we have then:

$$\big[b_n,\, b^\dagger_m\big]_+ \Big|_{\tau=\tau_0} = \frac{2}{T} \mathcal{N}\, \langle\!\langle {}^*\Psi_n \, , \, {}^*\Psi_m \rangle \Big|_{\tau=\tau_0}. \tag{3.38}$$

By definition the product (3.34) is time independent as long as the integrand ${}^*\Psi^*_n \Psi_m$ is free of singularities at $\tau = \widehat{\tau}_{(t)}$ for $t = 1, 2, \ldots, N$. Such request on the dual basis automatically fixes its functional form. Clearly this does not exclude the possibility to have singularities in Ψ_m or ${}^*\Psi_n$ separately: they are instead deeply connected to the boundary changing primary operator hidden in the discontinuity of the boundary conditions, that is different singularities will be shown to be in correspondence to the excited spin fields. Using the definition of the conserved product we therefore moved the focus from finding a consistent definition of the Fock space to the construction of the dual basis of modes. This task is however easier to address in a Euclidean formulation.

3.5 Point-like Defect CFT: the Euclidean Formulation

In the Euclidean reformulation the solution to the E.O.M. might be easier to study than its Lorentzian worldsheet form. This is specifically the case when $R_{(t)} \in \mathrm{U}(1)^{N_f} \subset \mathrm{U}(N_f)$. The presence of a time dependent Hamiltonian is not standard and we can neither blindly apply the usual Wick rotation nor the usual CFT techniques.

3.5.1 Fields on the Strip

Performing the Wick rotation as $\tau_E = i\tau$ such that $e^{iS} = e^{-S_E}$, the Minkowskian action (3.4) becomes:

$$S_E = \frac{T}{2} \iint \mathrm{d}\xi\, \mathrm{d}\bar{\xi}\, \frac{1}{2} \left(\widehat{\psi}^*_{E,+,i} \overset{\leftrightarrow}{\partial}_{\bar{\xi}} \widehat{\psi}^i_{E,+} + \widehat{\psi}^*_{E,-,i} \overset{\leftrightarrow}{\partial}_{\xi} \widehat{\psi}^i_{E,-} \right), \tag{3.39}$$

where the Euclidean fermion on the strip is connected to the Minkowskian formulation through

$$\widehat{\psi}^i_{E,\pm}(\xi,\, \bar{\xi}) = \psi^i_\pm(-i\xi,\, -i\bar{\xi}). \tag{3.40}$$

In the previous expressions we defined the coordinates $\xi = \tau_E + i\sigma$, $\bar{\xi} = \tau_E - i\sigma$ such that $\bar{\xi} = \xi^*$. Moreover we get $\partial_\xi = \frac{\partial}{\partial \xi} = \frac{1}{2}(\partial_{\tau_E} - i\,\partial_\sigma)$, $\partial_{\bar{\xi}} = \frac{\partial}{\partial \bar{\xi}} = \frac{1}{2}(\partial_{\tau_E} + i\,\partial_\sigma)$. As a consequence the Euclidean "complex conjugation" $\star$ (defined off-shell) acts as

$$\Big[\widehat{\psi}^i_{E,\pm}(\xi,\, \bar{\xi})\Big]^\star = \widehat{\psi}^*_{E,\pm i}(-\bar{\xi},\, -\xi). \tag{3.41}$$

The E.O.M. are as usual

$$\partial_\xi \widehat{\psi}^i_{E,-}(\xi,\bar{\xi}) = \quad \partial_{\bar{\xi}} \widehat{\psi}^i_{E,+}(\xi,\bar{\xi}) = \quad 0, \tag{3.42}$$

$$\partial_\xi \widehat{\psi}^*_{E,-,i}(\xi,\bar{\xi}) = \quad \partial_{\bar{\xi}} \widehat{\psi}^*_{E,+,i}(\xi,\bar{\xi}) = \quad 0, \tag{3.43}$$

whose solutions are the holomorphic functions $\widehat{\psi}_{E,+}(\xi)$ and $\widehat{\psi}_{E,-}(\bar{\xi})$, together with $\widehat{\psi}^*_{E,+}(\xi)$ and $\widehat{\psi}^*_{E,-}(\bar{\xi})$. In these coordinates the boundary conditions (3.7) translate to:

$$\begin{cases} \widehat{\psi}^i_{E,-}(\tau_E - i\,0^+) & = \left(R_{(t)}\right)^i{}_j\, \widehat{\psi}^j_{E,+}(\tau_E + i\,0^+) \\ \widehat{\psi}^*_{E,-,i}(\tau_E - i\,0^+) & = \left(R^*_{(t)}\right)_i{}^j\, \widehat{\psi}^*_{E,+,j}(\tau_E + i\,0^+) \end{cases} \tag{3.44}$$

for $\tau_E \in \left(\widehat{\tau}_{E,(t)}, \widehat{\tau}_{E,(t-1)}\right)$ and

$$\begin{cases} \widehat{\psi}^i_{E,-}(\tau_E - i\,\pi) & = -\widehat{\psi}^i_{E,+}(\tau_E + i\,\pi) \\ \widehat{\psi}^*_{E,-,i}(\tau_E - i\,\pi) & = -\widehat{\psi}^*_{E,+,i}(\tau_E + i\,\pi) \end{cases}, \tag{3.45}$$

where $t = 1, 2, \ldots, N$ and $\widehat{\tau}_{E,(t)}$ are the Wick-rotated locations of the N zero-dimensional defects, analytically continued to a real value.

The conserved product on the strip becomes:

$$\left\langle \widehat{\alpha}^*_E, \widehat{\beta}_E \right\rangle = \mathcal{N} \int_0^\pi \mathrm{d}\sigma \left(\widehat{\alpha}^*_{E,+,i} \widehat{\beta}^i_{E,+} + \widehat{\alpha}^*_{E,-,i} \widehat{\beta}^i_{E,-} \right), \tag{3.46}$$

where $\widehat{\alpha}^*_E$ and $\widehat{\beta}_E$ are the Euclidean counterparts of the generic solutions in the original definition of the product in (3.24). In the Euclidean context we have to explicitly write $\widehat{\alpha}^*_E$ because it is no longer the "complex conjugate" of $\widehat{\alpha}_E$ in the traditional sense. The product is conserved only when it couples two solutions which have different boundary conditions as in (3.44).

The definition of the stress-energy tensor in (3.26) requires a change in the numerical prefactor to use the usual CFT normalisation. Introducing a spacetime variable central charge as well the components of the stress-energy tensor become:[27]

$$\begin{aligned} \mathcal{T}_{\xi\xi}(\xi) &= -\frac{\pi T}{2}\, \widehat{\psi}^*_{E,+,i}(\xi)\, \overleftrightarrow{\partial}_\xi \widehat{\psi}^i_{E,+}(\xi) + \widehat{\mathcal{C}}(\xi), \\ \mathcal{T}_{\bar{\xi}\bar{\xi}}(\bar{\xi}) &= -\frac{\pi T}{2}\, \widehat{\psi}^*_{E,-,i}(\bar{\xi}) \overleftrightarrow{\partial}_{\bar{\xi}} \widehat{\psi}^i_{E,-}(\bar{\xi}) + \widehat{\bar{\mathcal{C}}}(\bar{\xi}), \end{aligned} \tag{3.49}$$

[27] The canonical coefficient in front of the CFT stress-energy tensor is such that the Euclidean Hamiltonian L_0 is normalized such that

$$\mathcal{T}_{\zeta\zeta}(\zeta) = \sum_n \mathrm{L}_n e^{-n\zeta} \tag{3.47}$$

(we are anticipating the double strip notation defined in the next subsection for simplicity). We thus get:

$$\mathrm{H}_E = \mathrm{L}_0 = \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} \mathcal{T}_{\zeta\zeta}(\tau_E + i\,\phi) \tag{3.48}$$

therefore $\mathcal{T}_{\zeta\zeta}(\zeta) = 2\pi \mathcal{T}^{(can)}_{\zeta\zeta}(\zeta)$.

where $\widehat{\mathcal{C}}$ and $\widehat{\widehat{\mathcal{C}}}$ are the leftover terms after the regularization of the singularities due to the normal ordering. The canonical anti-commutation relations are then

$$\left[\widehat{\psi}^{i}_{E,\pm}(\xi_1,\overline{\xi}_1),\widehat{\psi}^{*}_{E,\pm,j}(\xi_2,\overline{\xi}_2)\right]_{+}\Big|_{\mathrm{Re}\,\xi_1=\mathrm{Re}\,\xi_2}=\frac{2}{T}\,\delta^{i}{}_{j}\,\delta(\mathrm{Im}\,\xi_1-\mathrm{Im}\,\xi_2). \tag{3.50}$$

Given the Euclidean modes $\widehat{\psi}^{i}_{E,\pm,n}$ and $\widehat{\psi}^{*}_{E,\pm,n,i}$ where $n\in\mathbb{Z}$, we can then define the dual modes ${}^{*}\widehat{\psi}^{i}_{E,n}$ and ${}^{*}\widehat{\psi}^{*}_{E,n,i}$ such that the conserved product (3.46) between them gives:

$$\left\langle\!\!\left\langle {}^{*}\widehat{\psi}^{*}_{E,n}\,,\,\widehat{\psi}_{E,m}\right\rangle\!\!\right\rangle=\left\langle\!\!\left\langle {}^{*}\widehat{\psi}_{E,n}\,,\,\widehat{\psi}^{*}_{E,m}\right\rangle\!\!\right\rangle=\delta_{n,m}. \tag{3.51}$$

We can then expand the fields as

$$\begin{cases}\widehat{\psi}^{i}_{E,+}(\xi) &= \sum\limits_{n=-\infty}^{+\infty} b_n\,\widehat{\psi}^{i}_{E,+,n}(\xi)\\ \widehat{\psi}^{i}_{E,-}(\overline{\xi}) &= \sum\limits_{n=-\infty}^{+\infty} b_n\,\widehat{\psi}^{i}_{E,-,n}(\overline{\xi})\end{cases} \tag{3.52}$$

and

$$\begin{cases}\widehat{\psi}^{*}_{E,+,i}(\xi) &= \sum\limits_{n=-\infty}^{+\infty} b^{*}_n\,\widehat{\psi}^{*}_{E,+,n,i}(\xi)\\ \widehat{\psi}^{*}_{E,-,i}(\overline{\xi}) &= \sum\limits_{n=-\infty}^{+\infty} b^{*}_n\,\widehat{\psi}^{*}_{E,-,n,i}(\overline{\xi})\end{cases} \tag{3.53}$$

in order to extract the operators through the conserved product

$$b_n=\left\langle\!\!\left\langle {}^{*}\widehat{\psi}^{*}_{E,n}\,,\,\widehat{\psi}_{E}\right\rangle\!\!\right\rangle,\qquad b^{*}_n=\left\langle\!\!\left\langle {}^{*}\widehat{\psi}_{E,n}\,,\,\widehat{\psi}^{*}_{E}\right\rangle\!\!\right\rangle, \tag{3.54}$$

and get the anti-commutation relations at fixed Euclidean time as

$$[b_n,\,b^{*}_m]_{+}\Big|_{\tau_E=\tau_{E,(0)}}=\frac{2\mathcal{N}}{T}\left\langle\!\!\left\langle {}^{*}\widehat{\psi}^{*}_{E,n}\,,\,{}^{*}\widehat{\psi}_{E,m}\right\rangle\!\!\right\rangle. \tag{3.55}$$

3.5.2 Double Strip Formalism and Doubling Trick

It is natural to use the doubling trick on the strip to simplify the previous expressions by gluing the holomorphic and anti-holomorphic fields along the $\sigma=\pi$ boundary. Define the coordinate $\zeta=\tau_E+i\,\phi$ with $0\leq\phi\leq 2\pi$. We then have

$$\widehat{\Psi}(\zeta)=\begin{cases}\widehat{\psi}_{E,+}(\zeta) & \text{for}\quad \phi=\sigma\in[0,\pi]\\ -\widehat{\psi}_{E,-}(\zeta-2\pi i) & \text{for}\quad \phi=2\pi-\sigma\in[\pi,2\pi]\end{cases} \tag{3.56}$$

on-shell (and similarly for $\widehat{\Psi}^{*}(\zeta)$ with the substitution $\widehat{\psi}_{E,\pm}\to\widehat{\psi}^{*}_{E,\pm}$). The "complex conjugation" $\star$ acts on the off-shell double fields as

$$\left[\widehat{\Psi}^{i}(\zeta,\overline{\zeta})\right]^{\star}=\widehat{\Psi}^{*}_{i}(-\overline{\zeta},-\zeta), \tag{3.57}$$

while the boundary conditions are translated into

$$\begin{cases} \widehat{\Psi}^{i}(\tau_E + 2\pi i^-) = -\big(R_{(t)}\big)^{i}{}_{j}\,\widehat{\Psi}^{j}(\tau_E + i\,0^+) \\ \widehat{\Psi}^{*i}(\tau_E + 2\pi i^-) = -\Big(R^{*}_{(t)}\Big)^{i}{}_{j}\,\widehat{\Psi}^{*j}(\tau_E + i\,0^+) \end{cases} \tag{3.58}$$

for $\tau_E \in \big(\widehat{\tau}_{E,\,(t)}, \widehat{\tau}_{E,\,(t-1)}\big)$. The conserved product can then be defined as

$$\Big\langle \widehat{A}^{*}, \widehat{B} \Big\rangle = \mathcal{N} \int_0^{2\pi} \mathrm{d}\phi\ \widehat{A}^{*}_{i}(\tau_E + i\,\phi)\,\widehat{B}^{i}(\tau_E + i\,\phi), \tag{3.59}$$

where $\widehat{A}^{*}$ and $\widehat{B}$ are the double fields connected to $\widehat{\alpha}^{*}_{E}$ and $\widehat{\beta}_{E}$ in the previous definition on the strip. The holomorphic stress-energy tensor is then

$$\mathcal{T}_{\zeta\zeta}(\zeta) = -\frac{\pi T}{2}\,\widehat{\Psi}^{*}_{i}(\zeta)\,\overset{\leftrightarrow}{\partial}_{\zeta}\widehat{\Psi}^{i}(\zeta) + \widehat{\mathcal{C}}(\zeta) \tag{3.60}$$

and the canonical anti-commutation relations are now

$$\Big[\widehat{\Psi}^{i}(\zeta_1), \widehat{\Psi}^{*}_{j}(\zeta_2)\Big]_{+}\Big|_{\mathrm{Re}\,\zeta_1 = \mathrm{Re}\,\zeta_2} = \frac{2}{T}\,\delta^{i}{}_{j}\delta(\mathrm{Im}\,\zeta_1 - \mathrm{Im}\,\zeta_2). \tag{3.61}$$

The double field formulation shows that we need only one coefficient b_n (or b^{*}_{n}) for both $\psi_{E,\,+}$ and $\psi_{E,\,-}$ (or for both $\psi^{*}_{E,\,+}$ and $\psi^{*}_{E,\,-}$). In fact, given the Euclidean modes $\widehat{\Psi}^{i}_{n}$ and $\widehat{\Psi}^{*}_{n,\,i}$ where $n \in \mathbb{Z}$, we define the dual modes ${}^{*}\widehat{\Psi}^{i}_{n}$ and ${}^{*}\widehat{\Psi}^{*}_{n,\,i}$ such that:

$$\Big\langle\!\Big\langle {}^{*}\widehat{\Psi}^{*}_{n}\ ,\ \widehat{\Psi}_{m} \Big\rangle = \Big\langle\!\Big\langle {}^{*}\widehat{\Psi}_{n}\ ,\ \widehat{\Psi}^{*}_{m} \Big\rangle = \delta_{n,m}. \tag{3.62}$$

We expand the double fields as

$$\widehat{\Psi}^{i}(\zeta) = \sum_{n=-\infty}^{+\infty} b_n \widehat{\Psi}^{i}_{n}(\zeta), \qquad \widehat{\Psi}^{*}_{i}(\zeta) = \sum_{n=-\infty}^{+\infty} b^{*}_{n}\widehat{\Psi}^{*}_{n}(\zeta) \tag{3.63}$$

Operators are then extracted as

$$b_n = \Big\langle\!\Big\langle {}^{*}\widehat{\Psi}^{*}_{n}\ ,\ \widehat{\Psi} \Big\rangle, \qquad b^{*}_{n} = \Big\langle\!\Big\langle {}^{*}\widehat{\Psi}_{n}\ ,\ \widehat{\Psi}^{*} \Big\rangle. \tag{3.64}$$

Finally we get the anti-commutation relations as

$$[b_n,\, b^{*}_{m}]_{+}\Big|_{\tau_E = \tau_{E,\,0}} = \frac{2\mathcal{N}}{T}\,\Big\langle\!\Big\langle {}^{*}\widehat{\Psi}^{*}_{n}\ ,\ {}^{*}\widehat{\Psi}_{m} \Big\rangle. \tag{3.65}$$

3.6 Fields on the Upper Half Plane

We consider a set of coordinates on the upper half $\mathscr{H}$ of the complex plane:

$$u = e^{\xi} \in \mathscr{H}, \tag{3.66}$$

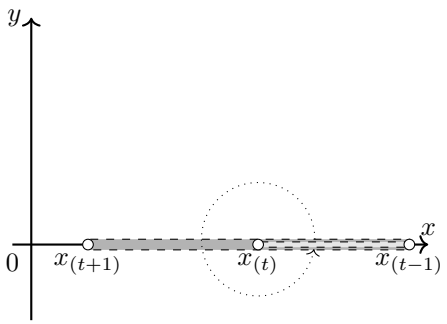


Figure 3.2: Fields are glued on the $x < 0$ semi-axis with non trivial discontinuities for $x_{(t)} < x < x_{(t-1)}$ for $t = 1, 2, \ldots, N$ and where $x_{(t)} = \exp\big(\widehat{\tau}_{E,(t)}\big)$.

where $\xi = \tau_E + i\sigma$ and $\sigma \in [0, \pi]$ define the usual strip, and $\mathscr{H} = \{w \in \mathbb{C} \mid \operatorname{Im} w \geq 0\}$. These coordinates can then be extended to the entire complex plane by considering

$$z = e^{\zeta} \in \mathbb{C}, \tag{3.67}$$

where $\zeta = \tau_E + i\phi$ and $\phi \in [0, 2\pi]$ define the double strip. Under this change of coordinates the Euclidean action (3.39) becomes

$$\begin{aligned} S_E &= \frac{T}{2} \iint \mathrm{d}u\, \mathrm{d}\overline{u}\, \frac{1}{2} \left(\frac{1}{u}\, \widehat{\psi}^*_{E,+,i} \overset{\leftrightarrow}{\partial_{\overline{u}}} \widehat{\psi}^i_{E,+} + \frac{1}{\overline{u}}\, \widehat{\psi}^*_{E,-,i} \overset{\leftrightarrow}{\partial_u} \widehat{\psi}^i_{E,-} \right) \\ &= \frac{T}{2} \iint \mathrm{d}u\, \mathrm{d}\overline{u}\, \frac{1}{2} \left(\psi^*_{E,+,i} \overset{\leftrightarrow}{\partial_{\overline{u}}} \psi^i_{E,+} + \psi^*_{E,-,i} \overset{\leftrightarrow}{\partial_u} \psi^i_{E,-} \right), \end{aligned} \tag{3.68}$$

where we introduce the off-shell field redefinitions:

$$\psi^i_{E,+}(u, \overline{u}) = \frac{1}{\sqrt{u}}\, \widehat{\psi}^i_{E,+}(\xi, \overline{\xi}), \qquad \psi^i_{E,-}(u, \overline{u}) = \frac{1}{\sqrt{\overline{u}}}\, \widehat{\psi}^i_{E,-}(\xi, \overline{\xi}). \tag{3.69}$$

Fields with the hat sign on top thus represent strip and double strip definitions, while fields without the hat sign are defined on $\mathscr{H}$ or $\mathbb{C}$.[28] Notice that this is the result one would expect from the engineering dimension: in this case it works since the theory is essentially free. Using the redefinitions (3.69), the off-shell "complex conjugation" $\star$ then becomes

$$[\psi_{E,+,i}(u, \overline{u})]^\star = \frac{1}{u}\, \psi^*_{E,+,i}\left(\frac{1}{\overline{u}}, \frac{1}{u}\right), \qquad [\psi_{E,-,i}(u, \overline{u})]^\star = \frac{1}{u}\, \psi^*_{E,-,i}\left(\frac{1}{\overline{u}}, \frac{1}{u}\right). \tag{3.71}$$

We choose to insert the cut of the square root on the real negative axis. The boundary conditions are then translated into

$$\begin{cases} \psi^i_{E,-}(x - i\,0^+) = \big(R_{(t)}\big)^i{}_j\, \psi^j_{E,+}(x + i\,0^+) \\ \psi^*_{E,-,i}(x - i\,0^+) = \Big(R^*_{(t)}\Big)^j{}_i\, \psi^*_{E,+,j}(x + i\,0^+) \end{cases} \tag{3.72}$$

[28] We could have anticipated these redefinitions from a CFT argument where

$$\psi(u) = \left(\frac{\mathrm{d}u}{\mathrm{d}\xi}\right)^{-\frac{1}{2}} \widehat{\psi}(\xi)\bigg|_{\xi = \ln(u)}, \tag{3.70}$$

but we cannot and do not rely on CFT properties since we have not shown that the theory is a CFT yet.

for $x \in \big(x_{(t)}, x_{(t-1)}\big)$, where $x_{(t)} = \exp\big(\widehat{\tau}_{E,\,(t)}\big) > 0$, and

$$\psi^{i}_{E,\,-}(x - i\,0^+) = \psi^{i}_{E,\,+}(x + i\,0^+), \qquad \psi^{*}_{E,\,-,\,i}(x - i\,0^+) = \psi^{*}_{E,\,+,\,i}(x + i\,0^+) \tag{3.73}$$

for $x < 0$.

The product (3.46) is then

$$\langle \alpha^*, \beta \rangle = -i\mathcal{N}\left[\int\limits_{\widehat{\Sigma}} \mathrm{d}u\, \alpha^{*}_{+,\,i}(u)\beta^{i}_{+}(u) - \int\limits_{\widehat{\overline{\Sigma}}} \mathrm{d}\overline{u}\, \alpha^{*}_{-,\,i}(\overline{u})\beta^{i}_{-}(\overline{u})\right], \tag{3.74}$$

where integrations are computed over two semi-circles $\widehat{\Sigma} = \{u \in \mathbb{C} \mid |u| = \exp(\widehat{\tau}_E),\ 0 \leq \operatorname{Im} u \leq \pi\}$ and $\widehat{\overline{\Sigma}} = \{u \in \mathbb{C} \mid |u| = \exp(\widehat{\tau}_E),\ -\pi \leq \operatorname{Im} u \leq 0\}$. The stress-energy tensor becomes:[29]

$$\begin{aligned} \mathcal{T}_{uu}(u) &= -\frac{\pi T}{2}\, \psi^{*}_{E,\,+,\,i}(u)\, \overset{\leftrightarrow}{\partial}_{u} \psi^{i}_{E,\,+}(u) + \widehat{\mathcal{C}}(u), \\ \mathcal{T}_{\overline{u}\overline{u}}(\overline{u}) &= -\frac{\pi T}{2}\, \psi^{*}_{E,\,-,\,i}(\overline{u})\, \overset{\leftrightarrow}{\partial}_{\overline{u}} \psi^{i}_{E,\,-}(\overline{u}) + \widehat{\overline{\mathcal{C}}}(\overline{u}). \end{aligned} \tag{3.77}$$

Finally the anti-commutation relations are

$$\begin{cases} \big[\psi^{i}_{E,\,+}(u_1,\, \overline{u}_1), \psi^{*}_{E,\,+,\,j}(u_2,\, \overline{u}_2)\big]_{+}\Big|_{|u_1|=|u_2|} &= \frac{2}{T u_1}\, \delta^{i}{}_{j}\, \delta(\arg(u_1) - \arg(u_2)) \\ \big[\psi^{i}_{E,\,-}(u_1,\, \overline{u}_1), \psi^{*}_{E,\,-,\,j}(u_2,\, \overline{u}_2)\big]_{+}\Big|_{|u_1|=|u_2|} &= \frac{2}{T \overline{u}_1}\, \delta^{i}{}_{j}\, \delta(\arg(u_1) - \arg(u_2)), \end{cases} \tag{3.78}$$

which despite the strange look of the expression are perfectly compatible with the definition (3.64) leading to:

$$[b_n,\, b^{*}_{m}]_{+} = \frac{2\mathcal{N}}{T}\, \Big\langle\!\Big\langle {}^{*}\widehat{\psi}^{*}_{E,\,n}\ ,\ {}^{*}\widehat{\psi}_{E,\,m} \Big\rangle = \frac{2\mathcal{N}}{T}\, \big\langle\!\big\langle {}^{*}\psi^{*}_{E,\,n}\ ,\ {}^{*}\psi_{E,\,m} \big\rangle \tag{3.79}$$

when the product $\langle\!\langle \cdot\ , \cdot \rangle$ is defined according to (3.74). We expand the fields in modes as:

$$\begin{cases} \psi^{i}_{E,\,+}(u) = \sum\limits_{n=-\infty}^{+\infty} b_n\, \psi^{i}_{E,\,+,\,n}(u) \\ \psi^{i}_{E,\,-}(\overline{u}) = \sum\limits_{n=-\infty}^{+\infty} b_n\, \psi^{i}_{E,\,-,\,n}(\overline{u}) \end{cases} \tag{3.80}$$

and

$$\begin{cases} \psi^{*}_{E,\,+,\,i}(u) = \sum\limits_{n=-\infty}^{+\infty} b^{*}_{n}\, \psi^{*}_{E,\,+,\,n,\,i}(u) \\ \psi^{*}_{E,\,-,\,i}(\overline{u}) = \sum\limits_{n=-\infty}^{+\infty} b^{*}_{n}\, \psi^{*}_{E,\,-,\,n,\,i}(\overline{u}) \end{cases} \tag{3.81}$$

and ${}^{*}\psi_{E,\,n}$ and ${}^{*}\psi^{*}_{E,\,n}$ are the corresponding dual modes on the upper half plane.

[29] While rewriting the operator part of the stress-energy tensor from the strip formulation into the coordinates in $\mathscr{H}$ we actually get

$$\mathcal{T}_{\xi\xi}(\xi(u)) = u^2\, \mathcal{T}_{uu}(u). \tag{3.75}$$

The reason of the presence of u^2 factor can be explained in two ways. Using GR we know that $\mathcal{T}_{\xi\xi}(\xi)\,\mathrm{d}\xi^2 = \mathcal{T}_{uu}(u)\,\mathrm{d}u^2$. Another way is to notice that a translation in ξ is a dilatation of u: the infinitesimal generator of ξ translation must be the infinitesimal generator of u dilatation, that is:

$$P_{\xi} \sim \int \mathrm{d}\sigma\ \mathcal{T}_{\xi\xi} \sim D_u \sim \int \mathrm{d}u\ u\, \mathcal{T}_{uu}. \tag{3.76}$$

3.7 Fields on the Complex Plane and the Doubling Trick

We use again the doubling trick to define the fields on the subset $\mathbb{C} \setminus \left[x_{(N)},\, x_{(1)}\right]$:

$$\Psi(z) = \begin{cases} \psi_{E,\,+}(u) & \text{for} \quad z = u \in \mathscr{H} \setminus \left[x_{(N)}, x_{(1)}\right] \\ \psi_{E,\,-}(\overline{u}) & \text{for} \quad z = \overline{u} \in \overline{\mathscr{H}} \setminus \left[x_{(N)}, x_{(1)}\right] \end{cases} \tag{3.82}$$

where $z = \exp(\tau_E + i\,\phi) = x + iy$ and $\overline{\mathscr{H}} = \{w \in \mathbb{C} \mid \operatorname{Im} w \leq 0\}$. The same procedure applies also to Ψ^* with the exchange $\psi_{E,\,\pm} \leftrightarrow \psi^*_{E,\,\pm}$.

In this case the "complex conjugation" $\star$ acts off-shell as

$$\left[\Psi^i(z, \overline{z})\right]^\star = \frac{1}{\overline{z}}\, \Psi^*_i\left(\frac{1}{\overline{z}}, \frac{1}{z}\right). \tag{3.83}$$

The boundary conditions then become:

$$\begin{cases} \Psi^i(x - i\,0^+) & = \left(R_{(t)}\right)^i{}_j\, \Psi^j(x + i\,0^+), \\ \Psi^{*\,i}(x - i0^+) & = \left(R^*_{(t)}\right)^i{}_j\, \Psi^{*\,j}(x + i\,0^+), \end{cases} \tag{3.84}$$

for $x \in \left(x_{(t)}, x_{(t-1)}\right)$, where $x_{(t)} = \exp\left(\widehat{\tau}_{E,\,(t)}\right) > 0$ for $t \in \{1,\, 2,\, \ldots,\, N\}$. When $x < 0$ we get

$$\begin{cases} \Psi(x - i\,0^+) & = \Psi(x + i\,0^+), \\ \Psi^*(x - i\,0^+) & = \Psi^*(x + i\,0^+) \end{cases}. \tag{3.85}$$

Given the relations $\mathrm{d}z = i\,z\;\mathrm{d}\phi$, we can write the conserved product (3.74) as:

$$\langle A^*, B\rangle = 2\pi\mathcal{N} \oint\limits_{|z| = \exp(\tau_E)} \frac{\mathrm{d}z}{2\pi i}\, A^*_i(z)\, B^i(z), \tag{3.86}$$

where we explicitly stressed that the integral has to be performed at a fixed Euclidean time τ_E: in the new coordinate on the plane the conserved product becomes a contour integral at a fixed radius from the origin.

In the same way we can recast the stress-energy tensor components (3.26) in the new coordinates:

$$\mathcal{T}(z) = -\frac{\pi T}{2}\, \Psi^*_i(z)\, \overset{\leftrightarrow}{\partial}_z \Psi^i(z) + \mathcal{C}(z), \tag{3.87}$$

where $\mathcal{T} = \mathcal{T}_{zz}$ for simplicity.

Finally the canonical anti-commutation relations between the fields are:

$$\left[\Psi^i(z_1),\, \Psi^*_j(z_2)\right]_+\Big|_{|z_1| = |z_2|} = \frac{2}{T z_1}\, \delta^i{}_j\, \delta(\arg(z_1) - \arg(z_2)). \tag{3.88}$$

The fields expansion in modes thus reads

$$\Psi^i(z) = \sum_{n=-\infty}^{+\infty} b_n\, \Psi^i_n(z), \qquad \Psi^*_i(z) = \sum_{n=-\infty}^{+\infty} b^*_n\, \Psi^*_{n.i}(z). \tag{3.89}$$

The anti-commutation relations among the operators are

$$[b_n,\, b^*_m]_+ = \frac{2\mathcal{N}}{T}\, \langle\!\langle {}^*\Psi^*_n\, ,\, {}^*\Psi_m \rangle\!\rangle\,, \tag{3.90}$$

when we introduce the dual modes ${}^*\Psi_n(z)$ and ${}^*\Psi^*_n(z)$ whose normalisation is

$$\langle\!\langle {}^*\Psi^*_n\, ,\, \Psi_m \rangle\!\rangle = \langle\!\langle {}^*\Psi_n\, ,\, \Psi^*_m \rangle\!\rangle = \delta_{m,n}. \tag{3.91}$$

3.8 Algebra of Creation and Annihilation Operators

In this section we find the explicit expression of the modes which satisfy the E.O.M. and the boundary conditions. We then compute the dual fields and finally the algebra of the creators and annihilators.

3.8.1 NS Complex Fermions

We start from NS complex fermions to show that the formalism can recover known results. Consider the usual definition:

$$\begin{cases} \psi^i_-(\tau, 0) & = \psi^i_+(\tau, 0), \\ \psi^i_-(\tau, \pi) & = -\psi^i_+(\tau, \pi) \end{cases} \tag{3.92}$$

for $\tau \in \mathbb{R}$, which can be recovered from (3.7) setting $R_{(t)} \equiv \mathbb{1}$. In the Euclidean formulation we use (3.84) and (3.85) to get:

$$\begin{cases} \Psi(x - i\,0^+) & = \Psi(x + i\,0^+) \\ \Psi^*(x - i\,0^+) & = \Psi^*(x + i\,0^+) \end{cases} \tag{3.93}$$

for $x \in \mathbb{R}$.

We define:

$$\Psi^i_{(n,\, i_0)}(z) \;=\; \mathcal{N}_\Psi\, \delta^i_{i_0}\, z^{-n}, \tag{3.94}$$

$${}^*\Psi_{(m,\, j_0),\, j}(z) \;=\; (2\pi\mathcal{N}\,\mathcal{N}_\Psi)^{-1}\, \delta_{j,j_0}\, z^{m-1} \tag{3.95}$$

to recover the definition of the dual modes (3.34) using the Euclidean conserved product (3.86). We then proceed similarly for Ψ^* in such a way that

$$\left\langle\!\!\left\langle {}^*\Psi^*_{(n,\, i_0)}\, ,\, \Psi_{(m,\, j_0)} \right\rangle\!\!\right\rangle = \left\langle\!\!\left\langle {}^*\Psi_{(m,\, j_0)}\, ,\, \Psi^*_{(n,\, i_0)} \right\rangle\!\!\right\rangle = \delta_{n,m}\, \delta_{i_0,j_0}. \tag{3.96}$$

As a consequence we find

$$\left\langle\!\!\left\langle {}^*\Psi^*_{(n,\, i_0)}\, ,\, {}^*\Psi_{(m,\, i_1)} \right\rangle\!\!\right\rangle = \frac{1}{2\pi\mathcal{N}\,\mathcal{N}_\Psi^2}\, \delta_{i_0,i_1}\, \delta_{n+m,1}. \tag{3.97}$$

Consider the NS expansion in modes of the double fields:

$$\Psi^i(z) = \sum_{n=-\infty}^{+\infty} \sum_{i_0} b_{(n,\,i_0)}\, \Psi^i_{(n,\,i_0)}(z), \tag{3.98}$$

$$\Psi^*_i(z) = \sum_{n=-\infty}^{+\infty} \sum_{i_0} b^*_{(n,\,i_0)}\, \Psi^*_{(n,\,i_0),\,i}(z), \tag{3.99}$$

then

$$b_{(n,\,i_0)} = \Big\langle\!\Big\langle {}^*\Psi^*_{(n,\,i_0)}\,,\,\Psi \Big\rangle\!\Big\rangle, \tag{3.100}$$

$$b^*_{(n,\,i_0)} = \Big\langle\!\Big\langle {}^*\Psi_{(n,\,i_0)}\,,\,\Psi^* \Big\rangle\!\Big\rangle, \tag{3.101}$$

and

$$\Big[b_{(n,i_0)},\, b^*_{(m,j_0)}\Big]_+ = \frac{1}{\pi T \mathcal{N}_\Psi^2}\, \delta_{i_0,j_0}\, \delta_{n+m,1}. \tag{3.102}$$

3.8.2 Twisted Complex Fermions: Preliminaries

We can then generalise the discussion about $N_f = 1$ complex fermions in the presence of N point-like defects which we will show to be primary boundary changing operators (i.e. plain and excited spin fields). Let

$$\begin{cases} R_{(t)} &= e^{i\pi\alpha_{(t)}} \in \mathrm{U}(1) \\ R^*_{(t)} &= e^{-i\pi\alpha_{(t)}} \in \mathrm{U}(1) \end{cases} \tag{3.103}$$

such that $0 < \alpha_{(t)} < 2$. The boundary conditions are:

$$\begin{cases} \Psi(x - i\,0^+) &= e^{i\pi\alpha_{(t)}} \Psi(x + i\,0^+) \\ \Psi^*(x - i\,0^+) &= e^{-i\pi\alpha_{(t)}} \Psi^*(x + i\,0^+) \end{cases}, \tag{3.104}$$

for $x \in (x_{(t)}, x_{(t-1)})$, and

$$\begin{cases} \Psi(x - i\,0^+) &= \Psi(x + i\,0^+) \\ \Psi^*(x - i\,0^+) &= \Psi^*(x + i\,0^+) \end{cases}, \tag{3.105}$$

for $x < 0$. These boundary conditions can be recast in the form of monodromy factors. With a loop around $x_{(t)}$ we find

$$\Psi\Big(x_{(t)} + \delta e^{i0^+}\Big) = e^{i\pi\left(\alpha_{(t)} - \alpha_{(t+1)}\right)}\, \Psi(x_{(t)} + \delta e^{2\pi i}), \tag{3.106}$$

where $\delta \in \mathbb{R}^+$ is small enough and the $\pm$ in the phase represents the position relative to the real axis ($+$ is in the upper half plane $\mathscr{H}$, while $-$ in the lower half plane $\overline{\mathscr{H}}$).[30] We then define the convenient combination:

$$\epsilon_{(t)} = \alpha_{(t+1)} - \alpha_{(t)} + \theta(\alpha_{(t)} - \alpha_{(t+1)} - 1) - \theta(\alpha_{(t+1)} - \alpha_{(t)} - 1) \tag{3.107}$$

[30] More precisely $0 < \delta < \min\Big(\big|x_{(t-1)} - x_{(t)}\big|, \big|x_{(t)} - x_{t+1}\big|\Big)$.

such that:[31]

$$-1 < \epsilon_{(t)} < 1, \qquad \forall t = 1, 2, \dots, N. \tag{3.108}$$

The previous loop around $x_{(t)}$ induces a monodromy

$$\begin{cases} \Psi(x_{(t)} + \delta e^{i0^+}) & = e^{-i\pi\epsilon_{(t)}} \Psi(x_{(t)} + \delta e^{2i\pi^+}) \\ \Psi^*(x_{(t)} + \delta e^{i0^+}) & = e^{-i\pi\overline{\epsilon}_{(t)}} \Psi^*(x_{(t)} + \delta e^{2i\pi^+}), \end{cases} \tag{3.109}$$

where $\overline{\epsilon}_{(t)} = -\epsilon_{(t)} \Rightarrow -1 < \overline{\epsilon}_{(t)} < 1$, thus showing a symmetry under the exchange of:

$$\Psi \longleftrightarrow \Psi^* \qquad \Rightarrow \qquad \epsilon_{(t)} \longleftrightarrow \overline{\epsilon}_{(t)}. \tag{3.110}$$

3.8.3 Usual Twisted Fermions

As a reference for future discussion, we consider the case of one complex fermion in the presence of one twisted boundary condition with the defects located at zero and infinity. We take $N = 2$ and $x_{(1)} = \infty$ and $x_{(2)} = 0$. For simplicity we denote with ϵ the argument of the monodromy factor arising from the presence of the cut on the interval $(0, +\infty)$.

In order to fulfill the requests (3.109) we write the modes as:

$$\begin{aligned} \Psi_n^{(\mathrm{E})} &= \mathcal{N}_\Psi \, z^{-n+\mathrm{E}}, \\ \Psi_n^{*\,(\overline{\mathrm{E}})} &= \mathcal{N}_\Psi \, z^{-n+\overline{\mathrm{E}}}, \end{aligned} \tag{3.111}$$

such that

$$\begin{aligned} \mathrm{E} &= n_{\mathrm{E}} + \frac{\epsilon}{2}, \quad n_{\mathrm{E}} \in \mathbb{Z}, \\ \overline{\mathrm{E}} &= n_{\overline{\mathrm{E}}} + \frac{\overline{\epsilon}}{2}, \quad n_{\overline{\mathrm{E}}} \in \mathbb{Z}. \end{aligned} \tag{3.112}$$

Together with the integer factor n_{E} and $n_{\overline{\mathrm{E}}}$ we also define a third integer for later convenience:[32]

$$\mathrm{L} = \mathrm{E} + \overline{\mathrm{E}} = n_{\mathrm{E}} + n_{\overline{\mathrm{E}}} \in \mathbb{Z}. \tag{3.113}$$

To extract creation and annihilation operators with the conserved product (3.86), we define the dual basis as:

$${}^*\Psi_n^{(\overline{\mathrm{E}})}(z) = \frac{1}{2\pi\mathcal{N}\mathcal{N}_\Psi} z^{n-1-\overline{\mathrm{E}}}, \tag{3.114}$$

$${}^*\Psi_n^{*\,(\mathrm{E})}(z) = \frac{1}{2\pi\mathcal{N}\mathcal{N}_\Psi} z^{n-1-\mathrm{E}}. \tag{3.115}$$

This way we compute the usual anti-commutation relations as

$$\left\langle\!\left\langle {}^*\Psi_n^{*\,(\mathrm{E})} \, , \, {}^*\Psi_m^{(\overline{\mathrm{E}})} \right\rangle\!\right\rangle = \frac{\delta_{n+m,1+\mathrm{L}}}{2\pi\mathcal{N}\mathcal{N}_\Psi^2} \quad \Rightarrow \quad [b_n, \, b_m^*]_+ = \frac{1}{\pi T \mathcal{N}_\Psi^2} \delta_{n+m,1+\mathrm{L}}, \tag{3.116}$$

[31] Notice that the choice of the range for $\epsilon_{(t)}$ is not unique. We can choose $0 < \alpha_{(t)} < 2$ leading to $\epsilon_{(t)} = \alpha_{(t+1)} - \alpha_{(t)} + 2\theta(\alpha_{(t)} - \alpha_{(t+1)})$. Then in this case $\epsilon_{(t)} = 2 - \overline{\epsilon}_{(t)}$ and $\epsilon_{(t)}, \, \overline{\epsilon}_{(t)} \in (0, 2)$. We will however stick to the first definition in the following sections since it allows us to consider the NS case as special.

[32] The choice discussed in Footnote 31 implies $\mathrm{L} = n_{\mathrm{E}} + n_{\overline{\mathrm{E}}} + 1$. We can swap the definitions by exchangin $\overline{\epsilon}_{(t)} \leftrightarrow \overline{\epsilon}_{(t)} + 2$ and $n_{\overline{\mathrm{E}}} \leftrightarrow n_{\overline{\mathrm{E}}} - 1$.

which are constant in time independently from E or $\overline{\mathrm{E}}$ since the only possible singularities are located at $z = 0$ and $z = \infty$. We can then expand the fields $\Psi(z)$ and $\Psi^*(z)$ using this or a more conventional basis:

$$\Psi(z) = \sum_{n=-\infty}^{+\infty} b_n^{(\mathrm{E})} \Psi_n^{(\mathrm{E})}(z) = \sum_{n=-\infty}^{+\infty} b_{n+n_{\mathrm{E}}} \Psi_n^{(\frac{\epsilon}{2})}(z), \tag{3.117}$$

$$\Psi^*(z) = \sum_{n=-\infty}^{+\infty} b_n^{*\,(\overline{\mathrm{E}})} \Psi_n^{*\,(\overline{\mathrm{E}})}(z) = \sum_{n=-\infty}^{+\infty} b^*_{n+n_{\overline{\mathrm{E}}}} \Psi_n^{*\,(-\frac{\epsilon}{2})}(z), \tag{3.118}$$

where we used the notation $b = b^{(\frac{\epsilon}{2})}$ and $b^* = b^{*\,(\frac{\epsilon}{2})}$.

3.8.4 Generic Case With Defects

We consider one complex fermion in the presence of N defects such that the modes satisfy:

$$\Psi_n(x_{(t)} + \delta e^{2\pi i^+}) = e^{i\pi\epsilon_{(t)}} \Psi_n(x_{(t)} + \delta e^{i0^+}) \tag{3.119}$$

for $t = 1, 2, \ldots, N$ and $\delta > 0$. We define the basis of solutions as:

$$\Psi_n(z; \{x_{(t)}, \mathrm{E}_{(t)}\}) = \mathcal{N}_\Psi\, z^{-n} \prod_{t=1}^{N} \left(1 - \frac{z}{x_{(t)}}\right)^{\mathrm{E}_{(t)}}, \tag{3.120}$$

$$\Psi_n^*(z; \{x_{(t)}, \overline{\mathrm{E}}_{(t)}\}) = \mathcal{N}_\Psi\, z^{-n} \prod_{t=1}^{N} \left(1 - \frac{z}{x_{(t)}}\right)^{\overline{\mathrm{E}}_{(t)}}, \tag{3.121}$$

where we generalise the definition of

$$\mathrm{E}_{(t)} = n_{\mathrm{E}_{(t)}} + \frac{\epsilon_{(t)}}{2}, \quad n_{\mathrm{E}_{(t)}} \in \mathbb{Z}, \tag{3.122}$$

$$\overline{\mathrm{E}}_{(t)} = n_{\overline{\mathrm{E}}_{(t)}} + \frac{\overline{\epsilon}_{(t)}}{2} \quad n_{\overline{\mathrm{E}}_{(t)}} \in \mathbb{Z} \tag{3.123}$$

and we define N integer factors analogous to (3.113):

$$\mathrm{L}_{(t)} = \mathrm{E}_{(t)} + \overline{\mathrm{E}}_{(t)} = n_{\mathrm{E}_{(t)}} + n_{\overline{\mathrm{E}}_{(t)}} \in \mathbb{Z}, \qquad \forall t \in \{1, 2, \ldots, N\}. \tag{3.124}$$

From the definition of the conserved product (3.86), we compute the dual basis:

$${}^*\Psi_n(z) = \frac{1}{2\pi\mathcal{N}\mathcal{N}_\Psi}\, z^{n-1} \prod_{t=1}^{N} \left(1 - \frac{z}{x_{(t)}}\right)^{-\overline{\mathrm{E}}_{(t)}}, \tag{3.125}$$

$${}^*\Psi_n^*(z) = \frac{1}{2\pi\mathcal{N}\mathcal{N}_\Psi}\, z^{n-1} \prod_{t=1}^{N} \left(1 - \frac{z}{x_{(t)}}\right)^{-\mathrm{E}_{(t)}}, \tag{3.126}$$

and the conserved products between dual modes:

$$\langle\!\langle {}^*\Psi_n^*\ ,\ {}^*\Psi_m \rangle = \frac{1}{2\pi\mathcal{N}\mathcal{N}_\Psi^2} \oint \frac{\mathrm{d}z}{2\pi i}\, z^{n+m-2} \prod_{t=1}^{N} \left(1 - \frac{z}{x_{(t)}}\right)^{-\mathrm{L}_{(t)}}. \tag{3.127}$$

Notice that the products are radially invariant only if

$$\mathrm{L}_{(t)} \leq 0, \qquad \forall t \in \{1,\, 2,\, \ldots,\, N\}, \tag{3.128}$$

since the integrand must not present time dependent singularities on the integration path, thus

$$\begin{aligned} \langle\!\langle {}^{*}\Psi_n^{*}\,,\, {}^{*}\Psi_m \rangle = & \frac{1}{2\pi\mathcal{N}\,\mathcal{N}_\Psi^2} \oint \frac{\mathrm{d}z}{2\pi i} \prod_{t=1}^{N} \sum_{k_t=0}^{|\mathrm{L}_{(t)}|} \binom{|\mathrm{L}_{(t)}|}{k_t} \left(-\frac{1}{x_{(t)}}\right)^{k_t} z^{k_t+n+m-2} \\ = & \frac{1}{2\pi\mathcal{N}\,\mathcal{N}_\Psi^2}\, p_{1-n-m}, \end{aligned} \tag{3.129}$$

where we defined

$$p_k = \prod_{t=1}^{N} \sum_{k_t=0}^{|\mathrm{L}_{(t)}|} \binom{|\mathrm{L}_{(t)}|}{k_t} \left(-\frac{1}{x_{(t)}}\right)^{k_t} \delta_{\sum_{t=1}^{N} k_t,\, k} \tag{3.130}$$

such that

$$p_{0 \leq k \leq \sum_{t=1}^{N} |\mathrm{L}_{(t)}|} \neq 0, \tag{3.131}$$

$$p_{k \leq -1} = p_{k \geq \sum_{t=1}^{N} |\mathrm{L}_{(t)}| + 1} = 0. \tag{3.132}$$

We can finally write

$$[b_n,\, b_m^{*}]_{+} = \frac{1}{\pi T \mathcal{N}_\Psi^2}\, p_{1-n-m}, \qquad 1 - \sum_{t=1}^{N} |\mathrm{L}_{(t)}| \leq n + m \leq 1. \tag{3.133}$$

3.9 Representation of the Algebra: Definition of the In-Vacuum

In the previous section we computed the algebra of the operators for different theories. We now define in-vacua where they are represented to be able to compute the relevant amplitudes. We show how to recover the NS vacuum and the usual twisted vacuum. Finally we discuss the vacuum in the presence of a generic number of defects.

3.9.1 NS Fermions

The in-vacuum can be correctly obtained either requiring $\Psi(z)$ and $\Psi^{*}(z)$ to be non singular as $z \to 0$ when applied on the vacuum. The same request can also be made on $\widehat{\Psi}(\xi)$ and $\widehat{\Psi}^{*}(\xi)$. In both cases we get the same vacuum which turns out to be $\mathrm{SL}_2(\mathbb{R})$ invariant:

$$b_{(n,\, i_0)}\, |0\rangle_{\mathrm{SL}_2(\mathbb{R})} = b^{*}_{(n,\, i_0)}\, |0\rangle_{\mathrm{SL}_2(\mathbb{R})} = 0, \qquad n \geq 1. \tag{3.134}$$

The spectrum of the theory is constructed acting with operators $b_{(n,\, i_0)}$ and $b^{*}_{(n,\, i_0)}$ with $n \leq 0$.

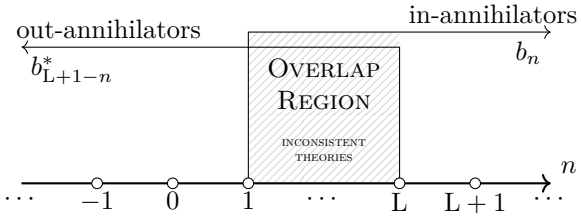


Figure 3.3: As a consistency condition, we have to exclude the values of L for which both $b_n^{(E)}$ and $b^{*\,(\overline{\mathrm{E}})}_{\mathrm{L}+1-n}$ are in-annihilators with a non vanishing anti-commutation relation.

3.9.2 Twisted Fermion

Consider the case of the usual twisted fermion in Section 3.8.3. Define the excited vacuum $\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle$ as:

$$b_n^{(\mathrm{E})}\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = b_n^{*\,(\overline{\mathrm{E}})}\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = 0, \qquad n \geq 1. \tag{3.135}$$

The introduction of E and $\overline{\mathrm{E}}$ is necessary to define the vacuum as in previous cases, that is with a range of n independent from them and without singularities as $z \to 0$. Explicitly we have:

$$\Psi(z)\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle \sim z^{\mathrm{E}}\,(\dots), \qquad \Psi^*(z)\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle \sim z^{\overline{\mathrm{E}}}\,(\dots). \tag{3.136}$$

Comparing with (3.120) and (3.121), the behavior suggests the existence of a hidden operator in $x_{(t)}$ creating $\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle$ with $\mathrm{E} = \mathrm{E}_{(t)}$ and $\overline{\mathrm{E}} = \overline{\mathrm{E}}_{(t)}$.

These relations are subject to consistency conditions since

$$\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = \pi T \mathcal{N}_\Psi^2 \left[b_n^{(E)},\, b^{*\,(\overline{\mathrm{E}})}_{\mathrm{L}+1-n}\right]_+ \left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle, \tag{3.137}$$

that is we cannot have two in-annihilators (namely both $b_n^{(\mathrm{E})}$ and $b^{*\,(\overline{\mathrm{E}})}_{\mathrm{L}+1-n}$) with non vanishing anti-commutation relations. Specifically we have:

$$1 \leq n \leq \mathrm{L} \qquad \Rightarrow \qquad b_n^{(\mathrm{E})}\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = 0, \quad b^{*\,(\overline{\mathrm{E}})}_{\mathrm{L}+1-n}\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = 0, \tag{3.138}$$

that is

$$\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = \pi T \mathcal{N}_\Psi^2 \left[b_n^{(\mathrm{E})},\, b^{*\,(\overline{\mathrm{E}})}_{\mathrm{L}+1-n}\right]_+ \left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = 0, \tag{3.139}$$

which is not consistent (see Figure 3.3 for a graphical reference): the theory does not exist. We shall therefore consider only cases such that

$$\mathrm{L} \leq 0, \tag{3.140}$$

analogously to (3.128).

Moreover notice that for $\mathrm{L} \leq -1$ both $b^{(\mathrm{E})}_{\mathrm{L}\leq n\leq 0}$ and $b^{*\,(\overline{\mathrm{E}})}_{\mathrm{L}\leq n\leq 0}$ are in- and out-creation operators.

The vacuum $\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle$ is not however associated to the lowest energy. In fact the usual way to build the vacuum would be to require $\Psi(z)$ and $\Psi^*(z)$ to be non singular as $z \to 0$ for the

in-vacuum so that $b_n^{(\mathrm{E})}\,|\mathrm{T}\rangle = 0$ for $n > \mathrm{E}$, and $b_n^{*\,(\overline{\mathrm{E}})}\,|\mathrm{T}\rangle = 0$ for $n > \overline{\mathrm{E}}$. However this procedure almost always fails to give a good definition of the vacuum. In fact it works only for NS fermions. For example when $\epsilon > 0$ we have:

$$0 = \pi T \mathcal{N}_\Psi^2 \left[b_{1+n_\mathrm{E}}^{(\mathrm{E})},\, b_{n_{\overline{\mathrm{E}}}}^{*\,(\overline{\mathrm{E}})} \right]_+ |\mathrm{T}\rangle = |\mathrm{T}\rangle\,, \tag{3.141}$$

which is not consistent since both $b_{1+n_\mathrm{E}}^{(\mathrm{E})}$ and $b_{n_{\overline{\mathrm{E}}}}^{*\,(\overline{\mathrm{E}})}$ are annihilators as $1 + n_\mathrm{E} > E$ and $n_{\overline{\mathrm{E}}} > \overline{\mathrm{E}}$.

The minimum energy vacuum is instead defined in a proper way on the strip. Requiring that the action of $\widehat{\Psi}(\xi)$ and $\widehat{\Psi}^*(\xi)$ for $\xi \to -\infty$ on the vacuum is well defined we get

$$b_n^{(\mathrm{E})}\,|\mathrm{T}\rangle = 0, \qquad n > \mathrm{E} + \frac{1}{2}, \tag{3.142}$$

$$b_m^{*\,(\overline{\mathrm{E}})}\,|\mathrm{T}\rangle = 0, \qquad m > \overline{\mathrm{E}} + \frac{1}{2}. \tag{3.143}$$

This is a good definition of the vacuum as $-\frac{1}{2} < \frac{\epsilon}{2} = -\frac{\overline{\epsilon}}{2} < \frac{1}{2}$ implies that $b_n^{(\mathrm{E})}$ and $b_m^{*\,(\overline{\mathrm{E}})}$ are annihilation operators for $n \geq n_\mathrm{E} + 1 > \mathrm{E} + \frac{1}{2}$ and $m \geq n_{\overline{\mathrm{E}}} + 1 > \overline{\mathrm{E}} + \frac{1}{2}$ so that

$$0 = \pi T \mathcal{N}_\Psi^2 \left[b_n^{(\mathrm{E})},\, b_m^{*\,(\overline{\mathrm{E}})} \right]_+ |\mathrm{T}\rangle = \delta_{n+m,\,\mathrm{E}+\overline{\mathrm{E}}+1}\,|\mathrm{T}\rangle = 0. \tag{3.144}$$

This way we get a consistent definition of the twisted vacuum. This is however not generally $\mathrm{SL}_2(\mathbb{R})$ invariant as we will show later when defining stress-energy tensor.[33]

The vacua $\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle$ and $|\mathrm{T}\rangle$ are related. Consider for example the case $n_\mathrm{E} \geq 1$ and the definition:

$$b_n^{(\mathrm{E})} \left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = 0, \qquad n \geq 1, \tag{3.145}$$

$$b_n^{(\mathrm{E})}\,|\mathrm{T}\rangle = 0, \qquad n \geq 1 + n_\mathrm{E}. \tag{3.146}$$

Then for $1 \leq n \leq n_\mathrm{E}$ the modes $b_n^{(\mathrm{E})}$ act as a annihilation operator on $\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle$ and as a creation operator on $|\mathrm{T}\rangle$:

$$\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle \propto b_{n_\mathrm{E}}^{(\mathrm{E})}\, b_{n_\mathrm{E}-1}^{(\mathrm{E})} \ldots b_1^{(\mathrm{E})}\,|\mathrm{T}\rangle\,. \tag{3.147}$$

Moreover, since $\mathrm{L} = n_\mathrm{E} + n_{\overline{\mathrm{E}}} \leq 0 \Rightarrow n_{\overline{\mathrm{E}}} \leq -1$, we have:

$$b_m^{*\,(\overline{\mathrm{E}})} \left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = 0, \qquad m \geq 1, \tag{3.148}$$

$$b_n^{*\,(\overline{\mathrm{E}})}\,|\mathrm{T}\rangle = 0, \qquad m \geq 1 - \left|n_{\overline{\mathrm{E}}}\right|, \tag{3.149}$$

which leads to:

$$|\mathrm{T}\rangle \propto b_0^{*\,(\overline{\mathrm{E}})}\, b_1^{*\,(\overline{\mathrm{E}})} \ldots b_{1-\left|n_{\overline{\mathrm{E}}}\right|}^{*\,(\overline{\mathrm{E}})} \left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle. \tag{3.150}$$

The consistency of the definition can be checked requiring

$$\left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle = \left(\pi T \mathcal{N}_\Psi^2\right)^{n_\mathrm{E}} b_{n_\mathrm{E}}^{(\mathrm{E})}\, b_{n_\mathrm{E}-1}^{(\mathrm{E})} \ldots b_1^{(\mathrm{E})}\, b_0^{*\,(\overline{\mathrm{E}})}\, b_1^{*\,(\overline{\mathrm{E}})} \ldots b_{1-\left|n_{\overline{\mathrm{E}}}\right|}^{*\,(\overline{\mathrm{E}})} \left|T_{\mathrm{E},\,\overline{\mathrm{E}}}\right\rangle, \tag{3.151}$$

[33] Notice that the second choice of ϵ interval discussed in Footnote 31 needs to distinguish between two cases: $0 < \frac{\epsilon}{2} < \frac{1}{2}$ (and $\frac{1}{2} < \frac{\overline{\epsilon}}{2} < 1$) and $\frac{1}{2} < \frac{\epsilon}{2} < 1$ (and $0 < \frac{\overline{\epsilon}}{2} < \frac{1}{2}$).

where the number of b operators has to match the number of b^* operators:

$$n_{\mathrm{E}} + n_{\overline{\mathrm{E}}} = \mathrm{E} + \overline{\mathrm{E}} = \mathrm{L} = 0. \tag{3.152}$$

The same procedure applies also in the case $n_{\mathrm{E}} \leq 0$, leading to the same result. As a consequence of (3.152), we express the twisted vacuum as:

$$b_n^{(\mathrm{E})} |\mathrm{T}\rangle = 0, \qquad n \geq 1 + n_{\mathrm{E}}, \tag{3.153}$$

$$b_m^{*(\overline{\mathrm{E}})} |\mathrm{T}\rangle = 0, \qquad m \geq 1 - n_{\mathrm{E}}. \tag{3.154}$$

3.9.3 Generic Case with Defects

Since the fields in presence of defects behave as NS fields in the limit $z \to 0$, we define the vacuum in the usual fashion by requiring a finite limit $\lim_{z\to 0} \Psi(z) \left|\Omega_{\{x_{(t)},\, \mathrm{E}_{(t)},\, \overline{\mathrm{E}}_{(t)}\}}\right\rangle$. As in the NS we get the representation:

$$b_n \left|\Omega_{\{x_{(t)},\, \mathrm{E}_{(t)},\, \overline{\mathrm{E}}_{(t)}\}}\right\rangle = b_n^* \left|\Omega_{\{x_{(t)},\, \mathrm{E}_{(t)},\, \overline{\mathrm{E}}_{(t)}\}}\right\rangle = 0, \qquad n \geq 1. \tag{3.155}$$

3.10 Asymptotic Fields and Vacua

In this section we define the asymptotic in-field and out-field and discuss how their vacua are related to the theory with defects. The relation is "radial time dependent" explicitly showing that an interaction is hidden in the defects. In particular the vacuum for the theory with defects can be identified with $\mathrm{SL}_2(\mathbb{R})$ in-field vacuum while it is connected by a Bogoliubov transformation to the $\mathrm{SL}_2(\mathbb{R})$ out-field vacuum.

In the following we use the expansion of

$$P\big(z;\, \{x_{(t)},\, \mathrm{E}_{(t)}\}\big) = \prod_{t=1}^{N} \left(1 - \frac{z}{x_{(t)}}\right)^{\mathrm{E}_{(t)}}, \tag{3.156}$$

around the origin and infinity with coefficients

$$\mathfrak{C}_k\big(0,\, \{x_{(t)},\, \mathrm{E}_{(t)}\}\big) = \sum_{\{k_t\}\in\mathbb{N}^N} \prod_{t=1}^{N} \left[\binom{\mathrm{E}_{(t)}}{k_t} \left(-\frac{1}{x_{(t)}}\right)^{k_t}\right] \delta_{\sum_{t=1}^{N} k_t,\, k} \tag{3.157}$$

$$\mathfrak{C}_k\big(\infty,\, \{x_{(t)},\, \mathrm{E}_{(t)}\}\big) = \sum_{\{k_t\}\in\mathbb{N}^N} \prod_{t=1}^{N} \left[\binom{\mathrm{E}_{(t)}}{k_t} \left(-x_{(t)}\right)^{k_t - \mathrm{E}_{(t)}}\right] \delta_{\sum_{t=1}^{N} k_t,\, k}, \tag{3.158}$$

so that we can write

$$P\big(z;\, \{x_{(t)}, \mathrm{E}_{(t)}\}\big) = \begin{cases} \sum_{k=0}^{+\infty} \mathfrak{C}_k\big(0,\, \{x_{(t)},\, \mathrm{E}_{(t)}\}\big)\, z^k, & \text{for} \quad |z| < x_{(N)} \\ \sum_{k=0}^{+\infty} \mathfrak{C}_k\big(\infty,\, \{x_{(t)},\, \mathrm{E}_{(t)}\}\big)\, z^{-k+\sum_{t=1}^{N} \mathrm{E}_{(t)}}, & \text{for} \quad |z| > x_{(1)} \end{cases} \tag{3.159}$$

We do not discuss intermediate fields, that is expansions for $x_{(t)} < |z| < x_{(t-1)}$, as it is not possible to disentangle the effects of defects before and after this range. The vacuum in the presence of defects is in fact related to the radial ordering of the operators associated with the defects as we argue later on.

3.10.1 Asymptotic In-field and Its Vacuum

Consider the definitions of the basis of solutions (3.120) and expand around $z = 0$.[34] We get for $0 \leq |z| < x_{(N)}$:

$$\Psi_n(z) = \sum_{k=0}^{+\infty} \mathfrak{C}_k\big(0,\, \{x_{(t)},\, \mathrm{E}_{(t)}\}\big)\, \Psi^{(0)}_{n-k}(z), \tag{3.160}$$

where $\Psi^{(0)}_n(z) = \mathcal{N}_\Psi z^{-n}$ as in (3.111) with $\mathrm{E} = 0$ which are the modes of a untwisted fermion, i.e. a plain NS fermion. The previous expansion connects the asymptotic behavior of the modes of the fermion with defects with the modes of a NS fermion close to the origin. We can now connect the operators of the system with defects with those of the asymptotic in-field. To this purpose we substitute the expansion (3.160) with the usual expression of the modes (3.89):

$$\Psi(z) = \sum_{n=-\infty}^{+\infty} b_n\, \Psi_n(z) \underset{|z|<x_{(N)}}{=} \Psi^{(\mathrm{in})}(z) = \sum_{n=-\infty}^{+\infty} b^{(0)}_n\, \Psi^{(0)}_n(z) \tag{3.161}$$

thus leading to

$$b^{(0)}_n = \sum_{k=0}^{+\infty} b_{n+k}\, \mathfrak{C}_k\big(0,\, \{x_{(t)},\, \mathrm{E}_{(t)}\}\big). \tag{3.162}$$

These expressions can be inverted writing $\Psi^{(0)}_n(z) = \Psi_n(z)\, P\big(z;\, \{x_{(t)},\, -\mathrm{E}_{(t)}\}\big)$. We then get:

$$b_n = \sum_{k=0}^{+\infty} \mathfrak{C}_k\big(0,\, \{x_{(t)},\, -\mathrm{E}_{(t)}\}\big)\, b^{(0)}_{n+k}, \tag{3.163}$$

Annihilation operators of the asymptotic theory, i.e. operators with positive index, are therefore expressed only using annihilation operators of the theory with defects. We can thus set

$$\Big|\Omega_{\{x_{(t)},\, \mathrm{E}_{(t)},\, \overline{\mathrm{E}}_{(t)}\}}\Big\rangle = \big|0_{(\mathrm{in})}\big\rangle_{\mathrm{SL}_2(\mathbb{R})}\,. \tag{3.164}$$

3.10.2 Asymptotic Vacua and Bogoliubov Transformations

Revisiting the previous section, we can also explicitly compute the expansion for $|z| > x_{(1)}$. Define for simplicity $\mathrm{M} = \sum_{t=1}^{N} \mathrm{E}_{(t)}$). We then get:

$$\Psi_n(z) = \sum_{k=0}^{+\infty} \mathfrak{C}_k\big(\infty,\, \{x_{(t)},\, \mathrm{E}_{(t)}\}\big)\, \Psi^{(0)}_{n+k-\mathrm{M}}(z). \tag{3.165}$$

[34] Similarly we could have considered (3.121) to begin with. Analogous relations can in fact be written for $b^{*\,(0)}_n$ with the substitutions of $\mathrm{E}_{(t)}$ with $\overline{\mathrm{E}}_{(t)}$.

The formula connects the asymptotic behavior of the modes of the fermion with defects to modes of a NS fermion which can be seen close to the infinity.

Out-operators can thus be connected to the theory with defect through:[35]

$$\Psi(z) = \sum_{n=-\infty}^{+\infty} b_n \, \Psi_n(z) \underset{|z|>x_1}{=} \Psi^{(\text{out})}(z) = \sum_{n=-\infty}^{+\infty} b_n^{(\infty)} \, \Psi_n^{(0)}(z). \tag{3.166}$$

We get:

$$b_n^{(\infty)} = \sum_{k=0}^{+\infty} b_{n+\mathrm{M}-k} \, \mathfrak{C}_k\big(\infty, \, \{x_{(t)}, \, \mathrm{E}_{(t)}\}\big). \tag{3.167}$$

The inverse of the expression is:

$$b_n = \sum_{k=0}^{+\infty} \mathfrak{C}_k\big(\infty, \, \{x_{(t)}, \, -\mathrm{E}_{(t)}\}\big) \, b_{n+\mathrm{M}-k}^{(\infty)}. \tag{3.168}$$

As we will show later however $\mathrm{M} = 0$. Annihilation operators of the asymptotic theory, i.e. operators with positive index, are thus expressed using both annihilation and creation operators of the theory with defects while creators, i.e. operators with non negative index, are expressed by means of creation operators only. It follows from the vacuum definition that:

$$\begin{aligned} \Big(\widetilde{\mathfrak{C}}_0 \, b_1^{(\infty)} + \text{creation op.}\Big) \Big|\Omega_{\{x_{(t)}, \, \mathrm{E}_{(t)}, \, \overline{\mathrm{E}}_{(t)}\}}\Big\rangle = 0, \\ \Big(\widetilde{\mathfrak{C}}_0 \, b_2^{(\infty)} + \widetilde{\mathfrak{C}}_1 \, b_1^{(\infty)} + \text{creation op.}\Big) \Big|\Omega_{\{x_{(t)}, \, \mathrm{E}_{(t)}, \, \overline{\mathrm{E}}_{(t)}\}}\Big\rangle = 0, \\ \vdots \end{aligned} \tag{3.169}$$

where $\widetilde{\mathfrak{C}}_n = \mathfrak{C}_n\big(\infty, \, \{x_{(t)}, \, -\mathrm{E}_{(t)}\}\big)$ for brevity. This means that the vacuum for the asymptotic out-field is non trivially connected to the vacuum of the theory with defects. More explicitly we have:

$$\begin{aligned} \Big|\Omega_{\{x_{(t)}, \, \mathrm{E}_{(t)}, \, \overline{\mathrm{E}}_{(t)}\}}\Big\rangle = \mathcal{N}_{(\text{out})}\big(\{x_{(t)}, \, \mathrm{E}_{(t)}, \, \overline{\mathrm{E}}_{(t)}\}\big) \\ \times e^{\sum_{m=-\infty}^{0} \sum_{n=-\infty}^{0} \mathcal{M}_{mn}^{(\text{out})}\left(\{x_{(t)}, \, \mathrm{E}_{(t)}, \, \overline{\mathrm{E}}_{(t)}\}\right) b_m^{*\,(\infty)} \, b_n^{(\infty)}} \Big|0_{(\text{out})}\Big\rangle_{\mathrm{SL}_2(\mathbb{R})}, \end{aligned} \tag{3.170}$$

so that the two $\mathrm{SL}_2(\mathbb{R})$ vacua are connected by a Bogoliubov transformation. More precisely we get (see appendix C for details)

$$\Big(\Psi^{(\text{out}, \, +)}(z) + \mathscr{L}_{(1)}\Big[\Psi^{(\text{out}, \, -)}\Big](z)\Big) \Big|\Omega_{\{x_{(t)}, \, \mathrm{E}_{(t)}, \, \overline{\mathrm{E}}_{(t)}\}}\Big\rangle = 0, \qquad |z| > x_{(1)}, \tag{3.171}$$

where

$$\mathscr{L}_{(t)}[\Psi](z) = \oint\limits_{|w|>x_{(t)}} \frac{\mathrm{d}w}{2\pi i} \frac{P\big(z; \, \{x_{(t)}, \, \mathrm{E}_{(t)}\}\big) \, P\big(w; \, \{x_{(t)}, \, -\mathrm{E}_{(t)}\}\big) - 1}{z - w} \, \Psi(w). \tag{3.172}$$

[35] To avoid a redundant notation we do not introduce an object $\Psi_n^{(\infty)}(z)$. Even though it would have been in principle correct, we would have also found that $\Psi_n^{(\infty)}(z) = \Psi_n^{(0)}(z)$.

The corresponding equation for $\Psi^{*\,(\mathrm{out})}(z)$ can be found with the substitution $\mathrm{E} \to \overline{\mathrm{E}}$. Notice that the kernel of the integral is nothing else (up to a multiplicative constant) but the regularised propagator, that is the propagator in the presence of the defects to which the NS propagator has been subtracted. The previous equation is solved explicitly by:

$$\begin{aligned}\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)},\,\overline{\mathrm{E}}_{(t)}\right\}}\right\rangle = \mathcal{N}\left(\left\{x_{(t)},\,\mathrm{E}_{(t)},\,\overline{\mathrm{E}}_{(t)}\right\}\right) \\ \times\, e^{\oint_{|z|>x_{(1)}} \frac{\mathrm{d}z}{2\pi i}\Psi^{*\,(\mathrm{out},\,-)}(z)\,\mathscr{L}\left[\Psi^{(\mathrm{out},\,-)}\right](z)} \left|0_{(\mathrm{out})}\right\rangle_{\mathrm{SL}_2(\mathbb{R})}.\end{aligned} \tag{3.173}$$

In the previous equation there is no need to specify the relation between $|z|$ and $|w|$ since $\Psi^{*\,(\mathrm{out},-)}(z)$ and $\Psi^{(\mathrm{out},-)}(w)$ anti-commute. From the same expression for $\Psi^{*\,(\mathrm{out},-)}(z)$ we deduce that $\overline{\mathrm{E}}_{(t)} = -\mathrm{E}_{(t)}$.

3.11 Contractions and Stress-energy Tensor

Given the definition of the algebra of the operators and its representation, we can finally define the normal ordering operation and proceed to compute the contractions and O.P.E. of the operators. The procedure ultimately leads to the definition of the stress-energy tensor. This is enough to show that the theory is a time dependent CFT since the stress-energy tensor satisfies the canonical O.P.E.

3.11.1 NS Complex Fermion

First of all we deal with the simple case of NS fermions. Using the algebra (3.102) we compute the O.P.E. of fermion fields as:

$$\Psi^i(z)\,\Psi^*_j(w) =: \Psi^i(z)\,\Psi^*_j(z):\, + \frac{1}{\pi T}\frac{\delta^i{}_j}{z-w}, \qquad |w|<|z|, \tag{3.174}$$

where the operation $:\cdot:$ is the normal ordering with respect to the $\mathrm{SL}_2(\mathbb{R})$ vacuum defined in (3.134). We then get the expression of the stress-energy tensor:

$$\begin{aligned}\mathcal{T}(z) &= \lim_{w\to z}\left[-\frac{\pi T}{2}\left(\Psi^*_i(z)\,\partial_w\Psi^i(w) - \partial_z\Psi^*_i(z)\,\Psi^i(w)\right) + \frac{N_f}{(z-w)^2}\right] \\ &= -\frac{\pi T}{2}:\Psi^*_i(z)\overset{\leftrightarrow}{\partial}_z\Psi^i(z):\,.\end{aligned} \tag{3.175}$$

We are then able to derive the necessary minimal subtraction:

$$\mathfrak{h}(z-w) = \frac{N_f}{(z-w)^2}. \tag{3.176}$$

3.11.2 Twisted Fermion

We now focus on $N_f = 1$ theories. First of all we consider the simplest case of the usual twisted fermion with the mode expansion (3.117) and (3.118). Using the anti-commutation

relations (3.116) we can compute the O.P.E.:

$$\Psi(z)\,\Psi^*(w) = \mathscr{N}_{\mathrm{E},\,\overline{\mathrm{E}}}[\Psi(z)\,\Psi^*(w)] + \frac{1}{\pi T}\Big(\frac{z}{w}\Big)^{\mathrm{E}}\frac{1}{z-w}, \qquad |w|<|z|, \tag{3.177}$$

and

$$\Psi^*(w)\,\Psi(z) = \mathscr{N}_{\mathrm{E},\,\overline{\mathrm{E}}}[\Psi^*(w)\,\Psi(z)] + \frac{1}{\pi T}\Big(\frac{w}{z}\Big)^{\overline{\mathrm{E}}}\frac{1}{w-z}, \qquad |w|>|z|. \tag{3.178}$$

If we require that the previous results can be assembled in a well defined continuous radial ordering $\mathrm{R}(\Psi(z)\Psi^*(w))$ we need to set $\mathrm{E} = -\overline{\mathrm{E}}$ so we can write

$$\mathrm{R}(\Psi(z)\Psi^*(w)) = \mathscr{N}_{\mathrm{E},\,\overline{\mathrm{E}}}[\Psi(z)\,\Psi^*(w)] + \frac{1}{\pi T}\Big(\frac{z}{w}\Big)^{\mathrm{E}}\frac{1}{z-w}. \tag{3.179}$$

The same result can be reached by computing the stress-energy tensor starting from the previous expressions. We have two ways to construct it depending on the ordering of the classical expression:

$$\begin{aligned}\mathcal{T}(z) &= \lim_{\substack{w\to z\\ |w|<|z|}}\left[-\frac{\pi T}{2}(\Psi^*(z)\partial_w\Psi(w) - \partial_z\Psi^*(z)\Psi(w)) + \frac{1}{(z-w)^2}\right]\\ &= -\frac{\pi T}{2}:\Psi^*(z)\overset{\leftrightarrow}{\partial}_z\Psi(z):\, + \frac{\mathrm{E}^2}{2z^2},\end{aligned} \tag{3.180}$$

or

$$\begin{aligned}\mathcal{T}(z) &= \lim_{\substack{w\to z\\ |w|<|z|}}\left[-\frac{\pi T}{2}(-\partial_z\Psi(z)\,\Psi^*(w) + \Psi(z)\,\partial_w\Psi^*(w)) + \frac{1}{(z-w)^2}\right]\\ &= -\frac{\pi T}{2}:\Psi^*(z)\overset{\leftrightarrow}{\partial}_z\Psi(z):\, + \frac{\overline{\mathrm{E}}^2}{2z^2},\end{aligned} \tag{3.181}$$

which however must coincide for consistency. Since

$$:\Psi(z)\overset{\leftrightarrow}{\partial}_z\Psi^*(z):\, =: \Psi^*(z)\overset{\leftrightarrow}{\partial}_z\Psi(z): \tag{3.182}$$

then we must then require $\mathrm{E}^2 = \overline{\mathrm{E}}^2$. We can get a stronger constraint by computing the O.P.E. $\mathcal{T}(z)\,\mathcal{T}(w)$. In fact the cancellation of the cubic divergence requires $\mathrm{E}+\overline{\mathrm{E}} = 0$. From now on we will therefore use the notation $|T_{\mathrm{E}}\rangle$ instead of $\Big|T_{\mathrm{E},\,\overline{\mathrm{E}}}\Big\rangle$.

From the usual definition of the stress-energy tensor in terms of the Virasoro generators $\mathcal{T}(z) = \sum_{k=-\infty}^{+\infty} L_k z^{-k-2}$, we extract the operators L_k from any of the previous definitions:

$$\begin{aligned}L_{(\mathrm{E})k} &= -\frac{\pi T}{2}\mathcal{N}_\Psi^2\sum_{k=-\infty}^{+\infty}\mathscr{N}_{\mathrm{E},\,\overline{\mathrm{E}}}\Big[b_n^{*\,(\overline{\mathrm{E}})}\,b_{k+1-n}^{(\mathrm{E})}\Big]\,(2n-k+2\mathrm{E}-1) + \frac{\mathrm{E}^2}{2}\delta_{k,0}\\ &= \frac{\pi T}{2}\mathcal{N}_\Psi^2\sum_{n=1}^{+\infty}\Big[(2n-k+2\mathrm{E}-1)\mathscr{N}_{\mathrm{E},\,\overline{\mathrm{E}}}\Big[b_{k+1-n}^{(\mathrm{E})}\,b_n^{*\,(\overline{\mathrm{E}})}\Big]\\ &+ (2n-k-2\mathrm{E}-1)\mathscr{N}_{\mathrm{E},\,\overline{\mathrm{E}}}\Big[b_{k+1-n}^{*\,(\overline{\mathrm{E}})}\,b_n^{(\mathrm{E})}\Big]\Big] + \frac{\mathrm{E}^2}{2}\delta_{k,0}.\end{aligned} \tag{3.183}$$

We already hinted to the fact that the vacua state involved are not in general $\mathrm{SL}_2(\mathbb{R})$ invariant. In particular we can see that that the excited vacua $|T_{\mathrm{E}}\rangle$ is a primary field

$$\begin{aligned} L_{(\mathrm{E})\,k}\,|T_{\mathrm{E}}\rangle &= 0, \qquad k>0, \\ L_{(\mathrm{E})\,0}\,|T_{\mathrm{E}}\rangle &= \frac{\mathrm{E}^2}{2}\,|T_{\mathrm{E}}\rangle\,, \end{aligned} \tag{3.184}$$

with non trivial conformal dimensions $\Delta(|T_{\mathrm{E}}\rangle) = \frac{\mathrm{E}^2}{2}$. This operator is an excited spin field $\mathrm{S}_{\mathrm{E}_{(t)}}(x)$ inserted at $x=0$. Its equivalent expression using bosonization is:

$$\mathrm{S}_{\mathrm{E}}(x) = e^{i\mathrm{E}\phi(x)}, \tag{3.185}$$

where ϕ is such that

$$\langle \phi(z)\phi(w)\rangle = -\frac{1}{(z-w)^2}. \tag{3.186}$$

The minimal conformal dimension is achieved for $n_{\mathrm{E}} = n_{\overline{\mathrm{E}}} = 0$, i.e. $\Delta(|\mathrm{T}\rangle) = \frac{\epsilon^2}{8}$, and identifies a plain spin field. We can further check this idea by showing that the conformal dimensions are consistent. Using (3.150) we get:

$$\begin{aligned} L_{(\mathrm{E})\,0}\,|\mathrm{T}\rangle &= L_0\left(b_0^{*\,(\overline{\mathrm{E}})}\,b_{-1}^{*\,(\overline{\mathrm{E}})}\,\dots\,b_{2-n_{\mathrm{E}}}^{*\,(\overline{\mathrm{E}})}\,|T_{\mathrm{E}}\rangle\right) \\ &= \left[\sum_{n=1}^{n_{\mathrm{E}}}\left(n-\frac{\mathrm{E}+1}{2}\right)+\frac{\mathrm{E}^2}{2}\right]|\mathrm{T}\rangle = \frac{\epsilon^2}{8}\,|\mathrm{T}\rangle\,. \end{aligned} \tag{3.187}$$

3.11.3 Generic Case With Defects

We then consider the generic case of one complex fermion in the presence of an arbitrary number of spin fields with respect to the vacuum we introduced in (3.155). We consider the mode expansion (3.120) and (3.121) as well as the anti-commutation relations (3.133).

As in the usual twisted case we first consider the contraction of the field Ψ and Ψ^* and then move to the stress-energy tensor. Using the anti-commutation relations and

$$\sum_{k=-\infty}^{+\infty} p_k z^k = \prod_{t=1}^{N}\left(1-\frac{z}{x_{(t)}}\right)^{-\mathrm{L}_{(t)}} \tag{3.188}$$

where p_k is defined in (3.130). We have:

$$\Psi(z)\,\Psi^*(w) =: \Psi(z)\,\Psi^*(w):\; + \frac{1}{\pi T}\,\frac{1}{z-w}\prod_{t=1}^{N}\left(1-\frac{z}{x_{(t)}}\right)^{\mathrm{E}_{(t)}}\left(1-\frac{w}{x_{(t)}}\right)^{-\mathrm{E}_{(t)}}, \tag{3.189}$$

as well as

$$\Psi^*(z)\,\Psi(w) =: \Psi^*(z)\,\Psi(w):\; + \frac{1}{\pi T}\,\frac{1}{z-w}\prod_{t=1}^{N}\left(1-\frac{z}{x_{(t)}}\right)^{\overline{\mathrm{E}}_{(t)}}\left(1-\frac{w}{x_{(t)}}\right)^{-\overline{\mathrm{E}}_{(t)}}, \tag{3.190}$$

both for $|w| < |z|$. To assemble the expressions in a well defined continuous radial ordering $\mathrm{R}[\Psi(z)\,\Psi^*(w)]$ we need to set $\mathrm{E}_{(t)} = -\overline{\mathrm{E}}_{(t)}$ such that we can write

$$\mathrm{R}[\Psi(z)\,\Psi^*(w)] =: \Psi(z)\,\Psi^*(w):\, + \frac{1}{\pi T}\,\frac{1}{z-w}\,\prod_{t=1}^{N}\left(1-\frac{z}{x_{(t)}}\right)^{\mathrm{E}_{(t)}}\left(1-\frac{w}{x_{(t)}}\right)^{-\mathrm{E}_{(t)}}. \tag{3.191}$$

We can then expand the results around z:

$$\begin{aligned}\mathrm{R}[\Psi(z)\,\Psi^*(w)] =:\, &(\Psi\Psi^*)(z):\, +:\, (\Psi\,\partial\Psi^*)(z):\, (w-z)\\ &+\frac{1}{\pi T}\left[\frac{-1}{w-z}+\sum_{t=1}^{N}\frac{\mathrm{E}_{(t)}}{z-x_{(t)}}\right.\\ &\left.-\frac{1}{2}\left(\sum_{t=1}^{N}\sum_{u\neq t}\frac{\mathrm{E}_{(t)}\,\mathrm{E}_{(u)}}{(z-x_{(t)})(z-x_{(u)})}+\sum_{t=1}^{N}\frac{\mathrm{E}_{(t)}\,(\mathrm{E}_{(t)}-1)}{(z-x_{(t)})^2}\right)(w-z)\right]\\ &+\mathcal{O}\big((w-z)^2\big),\end{aligned} \tag{3.192}$$

and around w

$$\begin{aligned}\mathrm{R}[\Psi(z)\,\Psi^*(w)] =:\, &(\Psi\Psi^*)(w):\, +:\, (\partial\Psi\,\Psi^*)(w):\, (z-w)\\ &+\frac{1}{\pi T}\left[\frac{1}{z-w}+\sum_{t=1}^{N}\frac{\mathrm{E}_{(t)}}{w-x_{(t)}}\right.\\ &\left.+\frac{1}{2}\left(\sum_{t=1}^{N}\sum_{u\neq t}\frac{\mathrm{E}_{(t)}\,\mathrm{E}_{(u)}}{(w-x_{(t)})(w-x_{(u)})}+\sum_{t=1}^{N}\frac{\mathrm{E}_{(t)}\,(\mathrm{E}_{(t)}-1)}{(w-x_{(t)})^2}\right)(z-w)\right]\\ &+\mathcal{O}\big((z-w)^2\big),\end{aligned} \tag{3.193}$$

so that the stress-energy tensor becomes:

$$\begin{aligned}\mathcal{T}(z) &= -\frac{\pi T}{2}:\Psi(z)\overset{\leftrightarrow}{\partial}_z\Psi^*(z):\, +\frac{1}{2}\left(\sum_{t=1}^{N}\frac{\mathrm{E}_{(t)}}{z-x_{(t)}}\right)^2\\ &= \frac{\pi T}{2}\,\mathcal{N}_\Psi^2\sum_{n,\,m=-\infty}^{+\infty}:b_n\,b_m^*:\,z^{-n-m}\left[\frac{m-n}{z}+2\sum_{t=1}^{N}\frac{\mathrm{E}_{(t)}}{z-x_{(t)}}\right]+\frac{1}{2}\left(\sum_{t=1}^{N}\frac{\mathrm{E}_{(t)}}{z-x_{(t)}}\right)^2.\end{aligned} \tag{3.194}$$

The last expression shows that the energy momentum tensor $\mathcal{T}(z)$ is radial time dependent but it satisfies the usual O.P.E.

First of all we notice that the vacuum $\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)},\,\overline{\mathrm{E}}_{(t)}\right\}}\right\rangle$ is actually $\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\}}\right\rangle$, i.e. it depends only on $x_{(t)}$ and $\mathrm{E}_{(t)}$. We can try to interpret the previous result in the light of the usual CFT approach. In particular we can refine the idea we discussed after (3.136) that the singularity in the modes (3.120) and (3.121) at the point $x_{(t)}$ is associated with a primary conformal operator which creates $|T_{\mathrm{E}}\rangle$ with $\mathrm{E} = \mathrm{E}_{(t)}$. By comparison with the stress energy tensor of an excited vacuum (3.180), from the second order singularity we learn that at the points $x_{(t)}$ there is an operator which creates the excited vacuum $\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\}}\right\rangle$from the $\mathrm{SL}_2(\mathbb{R})$ vacuum $|0\rangle_{\mathrm{SL}_2(\mathbb{R})}$. Given the discussion in the previous section this is an excited spin field $\mathrm{S}_{\mathrm{E}_{(t)}}\big(x_{(t)}\big) = e^{i\mathrm{E}_{(t)}\,\phi(x_{(t)})}$.

The first order singularities in $x_{(u)} - x_{(t)}$ are then the result of the interaction between two of the previous excited spin fields. Using the CFT operator approach we postulate that the following identification holds

$$\begin{aligned}\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\}}\right\rangle &= \mathcal{N}(\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\})\;\mathrm{S}_{\mathrm{E}_{(1)}}\left(x_{(1)}\right)\mathrm{S}_{\mathrm{E}_{(2)}}\left(x_{(2)}\right)\ldots\mathrm{S}_{\mathrm{E}_{(N)}}\left(x_{(N)}\right)|0\rangle_{\mathrm{SL}_2(\mathbb{R})}\\ &= \mathcal{N}(\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\})\;\mathrm{R}\left[\prod_{t=1}^{N}\mathrm{S}_{\mathrm{E}_{(t)}}\left(x_{(t)}\right)\right]|0\rangle_{\mathrm{SL}_2(\mathbb{R})}\,,\end{aligned} \tag{3.195}$$

then we get

$$\mathcal{T}(z)\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\}}\right\rangle = \mathcal{N}(\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\})\;\mathrm{R}\left[\mathcal{T}(z)\prod_{t=1}^{N}\mathrm{S}_{\mathrm{E}_{(t)}}\left(x_{(t)}\right)\right]|0\rangle_{\mathrm{SL}_2(\mathbb{R})}\,. \tag{3.196}$$

The fact that $\mathcal{T}(z)$ enters the radial ordering may seem strange but the left hand side is well defined for all z and the only well defined expression for the right hand side is with the radial ordering. In fact an operator expression like $\mathcal{T}(z)\mathrm{R}\left[\partial_{x_{(1)}}\phi(x_{(1)})\,\partial_{x_{(2)}}\phi(x_{(2)})\right]|0\rangle_{\mathrm{SL}_2(\mathbb{R})}$ is only defined for $|z| > x_{(1),\,(2)}$. It then follows that

$$\mathcal{T}(z)\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\}}\right\rangle = \sum_{t=1}^{N}\left(\frac{\mathrm{E}_{(t)}^2/2}{(z-x_{(t)})^2} + \frac{\partial_{x_{(t)}} - \partial_{x_{(t)}}\log\mathcal{N}}{z-x_{(t)}}\right)\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\}}\right\rangle + \mathcal{O}(1), \tag{3.197}$$

which allows us to write

$$\begin{aligned}&\mathcal{N}(\left\{x_{(t)},\mathrm{E}_{(t)}\right\})\;\mathrm{R}\left[\partial_{x_{(t)}}\mathrm{S}_{\mathrm{E}_{(t)}}\left(x_{(t)}\right)\prod_{u\neq t}\mathrm{S}_{\mathrm{E}_{(u)}}\left(x_{(u)}\right)\right]|0\rangle_{\mathrm{SL}_2(\mathbb{R})}\\ =&\mathrm{E}_{(t)}\left[\pi T\mathcal{N}_\Psi^2\sum_{n,\,m=0}^{+\infty}\frac{b_n\,b_m^*}{x_{(t)}^{n+m}} + \sum_{u\neq t}\frac{\mathrm{E}_{(u)}}{x_{(t)}-x_{(u)}}\right]\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)}\right\}}\right\rangle.\end{aligned} \tag{3.198}$$

This result shows the way non primary operators are represented in this formalism.

3.12 Hermitian Conjugation

Before we can define the amplitudes involving spin and matter fields, we still need to introduce some of the necessary tools. In this section we focus on the operation of "Hermitian conjugation" in a broad sense: the usual Hermitian conjugation requires the existence of an inner product which is not yet available since we have not defined the out-vacuum. The operation we define is similar to the $\star$ operator of $C^\star$ algebras even though the $\star$ operator sends an element of an algebra to another element of the same algebra. This is not what happens in the generic case since the $\star$ is essentially associated with the inversion $z \to \overline{z}^{-1}$, i.e. in evolving from $\tau = +\infty$ to $\tau = -\infty$ so that the order of boundary singularities is reversed.

3.12.1 Usual Twisted Fermions

In general for a chiral primary conformal operator of dimension Δ in z coordinates the Euclidean Hermitian conjugation is

$$[O(z)]^\dagger = \left(w^{2\Delta}O(w)\right)\Big|_{w=\overline{z}^{-1}}. \tag{3.199}$$

As a matter of fact we cannot use the words "Euclidean Hermitian conjugation" since we do not have an inner product. We define the operation $\star$ which mimics its behavior. Therefore we define

$$[\Psi(z;\mathrm{E})]^\star = \left[w\,\widetilde{\Psi}^*\left(w;-\widetilde{\mathrm{E}}\right)\right]\Big|_{w=\overline{z}^{-1}}, \qquad [\Psi^*(z;\mathrm{E})]^\star = \left(w\,\widetilde{\Psi}\left(w;\widetilde{\mathrm{E}}\right)\right)\Big|_{w=\overline{z}^{-1}}. \tag{3.200}$$

In the last expression we did not assume that the action of $\star$ is an automorphism and we wrote $\Psi(z;\mathrm{E})$ to explicitly show the dependence on the parameter E which enters in the modes. The previous action agrees with (3.83). In terms of the basis (3.111) we write:[36]

$$\left[\Psi_n^{(\mathrm{E})}(z)\right]^\star = \left[w\,\Psi^{*\,(-\mathrm{E})}_{1-n}(w)\right]\Big|_{w=\overline{z}^{-1}}, \qquad \left[\Psi_n^{(-\mathrm{E})}(z)\right]^\star = \left[w\,\Psi^{*\,(\mathrm{E})}_{1-n}(w)\right]\Big|_{w=\overline{z}^{-1}}, \tag{3.201}$$

which shows that in this case the image of the $\star$ operator is the same as the support. Using the mode expansion of (3.200) it follows that

$$\left[b_n^{(\mathrm{E})}\right]^\star = b^{*\,(\overline{\mathrm{E}})}_{1-n}, \qquad \left[b_n^{*\,(\overline{\mathrm{E}})}\right]^\star = b^{(\mathrm{E})}_{1-n}. \tag{3.202}$$

The $\star$ action is compatible with the anti-commutation relations as we can show by explicitly computing them:

$$\left(\left[b_n^{(\mathrm{E})},\, b_m^{*\,(\overline{\mathrm{E}})}\right]_+\right)^\star = \left[b^{*\,(\overline{\mathrm{E}})}_{1-n},\, b^{(\mathrm{E})}_{1-m}\right]_+ = \frac{1}{\pi T \mathcal{N}_\Psi^2}\delta_{n+m,1}. \tag{3.203}$$

Furthermore $\star$ is involutive since:

$$\left[\Psi_n^{(\mathrm{E})}(z)\right]^{\star\star} = \Psi_n^{(\mathrm{E})}(z) \quad \Rightarrow \quad \left[b_n^{(\mathrm{E})}\right]^{\star\star} = b_n^{(\mathrm{E})}. \tag{3.204}$$

3.12.2 Generic Case With Defects

Consider the modes (3.120). We define the action of the $\star$ operator on them as:

$$\begin{aligned}\left[\Psi_n\left(z;\,\{x_{(t)},\,\mathrm{E}_{(t)}\}\right)\right]^\star &= \mathcal{N}_\Psi\,\overline{z}^{-n}\prod_{t=1}^{N}\left(1-\frac{\overline{z}}{x_{(t)}}\right)^{\mathrm{E}_{(t)}} \\ &= \left(w\prod_{t=1}^{N}\left(-\frac{1}{x_{(t)}}\right)^{\mathrm{E}_{(t)}}\widetilde{\Psi}^*_{M+1-n}\left(w;\left\{\widetilde{x}_{(t)},\overline{\widetilde{\mathrm{E}}}_{(t)}\right\}\right)\right)\Big|_{w=\overline{z}^{-1}}\end{aligned} \tag{3.205}$$

where we used $\mathrm{M}=\sum_{t=1}^{N}\mathrm{E}_{(t)}$ and $\widetilde{\Psi}_l(w;\,\{y,\,\mathrm{F}\}) = \mathcal{N}_\Psi\,w^{-l}\prod_{t=1}^{N}\left(1-\frac{w}{y}\right)^{-\mathrm{F}}$. In this case the image of the $\star$ operator is a different space where the defects are located in $\widetilde{x}_{(t)}$ and the critical exponents are $\widetilde{\mathrm{E}}_{(t)}$ and $\overline{\widetilde{\mathrm{E}}}_{(t)}$ with

$$\widetilde{x}_{(t)} = \frac{1}{x_{(t)}}, \qquad \widetilde{\mathrm{E}}_{(t)} = -\mathrm{E}_{(t)} = \overline{\mathrm{E}}_{(t)}, \qquad \overline{\widetilde{\mathrm{E}}}_{(t)} = \mathrm{E}_{(t)} = -\overline{\mathrm{E}}_{(t)}, \tag{3.206}$$

[36] The second possibility $\left[\Psi_n^{(\mathrm{E})}(z)\right]^\star = \left[w\,\Psi^{*\,(-\mathrm{E}-1)}_{-n}(w)\right]\Big|_{w=\overline{z}^{-1}}$ is inconsistent with the anti-commutation relations.

where we used $\mathrm{E}_{(t)}+\overline{\mathrm{E}}_{(t)}=0$. We can then compute the action of the $\star$ operator on the creation and annihilation operators:

$$b_n^\star=\prod_{t=1}^{N}\left(-\frac{1}{x_{(t)}}\right)^{-\mathrm{E}_{(t)}}\widetilde{b}^*_{M+1-n},\qquad (b_n^*)^\star=\prod_{t=1}^{N}\left(-\frac{1}{x_{(t)}}\right)^{\mathrm{E}_{(t)}}\widetilde{b}_{-M+1-n}. \tag{3.207}$$

As in the previous situation the anti-commutation relations are preserved by the $\star$ operator. Explicitly we have:

$$\left([b_n,\,b^*_m]_+\right)^\star=\left[\widetilde{b}_{-M+1-m},\,\widetilde{b}^*_{M+1-n}\right]_+=\frac{1}{\pi T\mathcal{N}_\Psi^2}\,\delta_{n+m,1}. \tag{3.208}$$

Finally the $\star$ operator is involutive.

3.13 Definition of the Out-Vacuum

With the definition of the $\star$ operator we can now proceed to define the out-vacuum to as the Hermitian conjugation in the usual cases. It is a conceptually separated step from the definitions of the algebra of operators and their representation on the in-vacuum. We first consider the usual twisted theory from which we learn how to define the out-vacuum and then move to the generic case in the presence of multiple defects.

3.13.1 Usual Twisted Fermions

Consider the definition of the in-vacuum (3.135) for the fields image of the $\star$ operator, i.e. $\widetilde{\Psi}\left(w;\,\widetilde{\mathrm{E}}\right)$ and $\widetilde{\Psi}^*\left(w;\,\overline{\widetilde{\mathrm{E}}}\right)$. It is defined as

$$\widetilde{b}_n^{(\widetilde{\mathrm{E}})}\left|\widetilde{T}_{\widetilde{\mathrm{E}},\overline{\widetilde{\mathrm{E}}}}\right\rangle=\widetilde{b}^{*\,(\overline{\widetilde{\mathrm{E}}})}_n\left|\widetilde{T}_{\widetilde{\mathrm{E}},\overline{\widetilde{\mathrm{E}}}}\right\rangle=0,\qquad n\geq 1. \tag{3.209}$$

The usual Hermitian conjugation gives

$$\left\langle\widetilde{T}_{\widetilde{\mathrm{E}},\overline{\widetilde{\mathrm{E}}}}\right|\left(\widetilde{b}_n^{(\widetilde{\mathrm{E}})}\right)^\dagger=\left\langle\widetilde{T}_{\widetilde{\mathrm{E}},\overline{\widetilde{\mathrm{E}}}}\right|\left(\widetilde{b}^{*\,(\overline{\widetilde{\mathrm{E}}})}_n\right)^\dagger=0,\qquad n\geq 1. \tag{3.210}$$

Given the action of the $\star$ operator (3.202), the identification with the Hermitian conjugate is possible if

$$\langle T_{\mathrm{E}}|\,b_n^{(\mathrm{E})}=\langle T_{\mathrm{E}}|\,b^{*\,(\overline{\mathrm{E}})}_n=0,\qquad n\leq 0. \tag{3.211}$$

3.13.2 Generic Case With Defects

We can now analyse the case of an arbitrary number of defects. Following the steps of the previous section we define the in-vacuum for the tilded theory as

$$\widetilde{b}_n\left|\Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}}\right\rangle=\widetilde{b}_n\left|\Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}}\right\rangle=0,\qquad n\geq 1, \tag{3.212}$$

and interpret it as the out-vacuum of the initial theory. The definition of the out-vacuum is therefore:

$$\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right| b_n \;=\; 0, \qquad n \leq \mathrm{M}, \tag{3.213}$$

$$\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right| b^*_n \;=\; 0, \qquad n \leq -\mathrm{M}. \tag{3.214}$$

Since the action of the $\star$ operator is compatible with the anti-commutation relations, the definition of the out-states is consistent. If we assume that

$$\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle \neq 0, \tag{3.215}$$

using the anti-commutation relations we get

$$\begin{aligned} \frac{1}{\pi T \mathcal{N}_\Psi^2} \left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle &= \left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right| \left[b_{\mathrm{M}},\, b^*_{-\mathrm{M}+1} \right]_+ \left| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle \\ &= \left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right| b^*_{-\mathrm{M}+1}\, b_{\mathrm{M}} \left| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle \neq 0, \end{aligned} \tag{3.216}$$

which requires $b_{\mathrm{M}} \left| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle \neq 0$. A similar condition exists for $b^*_{-\mathrm{M}}$, thus we must require $\mathrm{M} \leq 0$ and $-\mathrm{M} \leq 0$:

$$\mathrm{M} = \sum_{t=1}^{N} \mathrm{E}_{(t)} = 0. \tag{3.217}$$

The situation is therefore analogous to the case depicted in Figure 3.3 where M and $\overline{\mathrm{M}}$ have the same role of L for the twisted fermion.

3.13.3 Asymptotic vacua

The discussion is essentially the same as in Section 3.10 with the role of asymptotic in- and out-fields swapped. In particular we get

$$\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right| = {}_{\mathrm{SL}_2(\mathbb{R})}\!\left\langle 0_{(\mathrm{out})} \right|, \tag{3.218}$$

and

$$\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right| = \mathcal{N}_{(in)}\big(\{x_{(t)},\,\mathrm{E}_{(t)}\}\big)\; {}_{\mathrm{SL}_2(\mathbb{R})}\!\left\langle 0_{(\mathrm{in})} \right| e^{\sum_{m,\,n=1}^{+\infty} \mathcal{M}_{mn}\left(\{x_{(t)},\,\mathrm{E}_{(t)}\}\right) b^{*\,(0)}_m b^{(0)}_n}. \tag{3.219}$$

3.14 Correlators in the Presence of Spin Fields

The definitions of the in- and out-vacua and the stress-energy tensor are critical to compute any correlation function of operators in the presence of the point-like defects. In fact we need to know both the algebra of the operators and their representation, usually defined on the in-vacuum (the ket vector), as well as their Hermitian conjugation in order to build the action of the operators on the out-vacuum (the bra vector).

Starting from (3.195) we can compute the spin field correlators

$$\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle = \mathcal{N}(\{x_{(t)},\,\mathrm{E}_{(t)}\}) \left\langle \mathrm{R}\left[\prod_{t=1}^{N} \mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)}) \right] \right\rangle . \tag{3.220}$$

At first sight both $\left|\Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}}\right\rangle$and $\left\langle\Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}}\right|$might seem to contain $\mathrm{R}\left[\prod_{t=1}^{N}\mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)})\right]$. However this it is not the case and it can be seen in different ways. The simplest is to realise that such a square would be divergent while the product seems to be perfectly finite. A more rigorous way is to consider what the previous product is from the point of view of asymptotic out field: we have $\left|\Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}}\right\rangle = \mathcal{N}_{(out)}\,\mathrm{R}\left[\prod_{t=1}^{N}\mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)})\right]|0_{(\mathrm{out})}\rangle_{\mathrm{SL}_2(\mathbb{R})}$ and $\left\langle\Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}}\right| = {}_{\mathrm{SL}_2(\mathbb{R})}\langle 0_{(\mathrm{out})}|$ so that $\mathcal{N}_{(out)} = \mathcal{N}$. Moreover $\mathcal{T}(z) \underset{|z|>x_{(1)}}{=} \mathcal{T}_{(out)}(z)$ when the two energy momentum tensors are normal ordered with respect to their different sets of operators which are related in (3.167). Hence all the expressions are surely valid for $|z| > x_{(1)}$ and can be analytically extended to the whole plane. The same result can be obtained from the point of view of asymptotic in-fields.

Unfortunately the normalisation factor cannot be uniquely fixed. The result depends on the normalisation chosen for the single spin field and effectively shows only when we relate the N points to $N-1$ points correlators, recursively down to two points correlators. Therefore we need to consider quantities where the normalisation is not present. In particular we consider

$$\begin{aligned}
&\frac{\partial}{\partial x_{(t)}} \ln \left\langle \mathrm{R}\left[\mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)}) \prod_{\substack{u=1\\u\neq t}}^{N} \mathrm{S}_{\mathrm{E}_{(u)}}(x_{(u)}) \right] \right\rangle \\
&= \oint_{|z|=x_{(t)}} \frac{\mathrm{d}z}{2\pi i} \frac{\left\langle \mathrm{R}\left[\mathcal{T}(z) \prod_{t=1}^{N} \mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)}) \right] \right\rangle}{\left\langle \mathrm{R}\left[\prod_{t=1}^{N} \mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)}) \right] \right\rangle} \\
&= \left(\oint_{|z|>x_{(t)}} \frac{\mathrm{d}z}{2\pi i} - \oint_{|z|<x_{(t)}} \frac{\mathrm{d}z}{2\pi i} \right) \frac{\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \mathcal{T}(z) \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle}{\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle} \\
&= \frac{\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \left(L_{-1}^{x_{(t)}^{+}} - L_{-1}^{x_{(t)}^{-}} \right) \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle}{\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle}
\end{aligned} , \tag{3.221}$$

since $[L_{-1},\,\mathcal{O}_h(z)] = \partial_z \mathcal{O}_h(z)$ for a quasi-primary operator $\mathcal{O}_h$. From the definition of $\mathcal{T}(z)$ it follows that:

$$L_{-1}^{x_{(t)}^{+}} - L_{-1}^{x_{(t)}^{-}} = \oint_{\mathcal{C}_{x_{(t)}}} \frac{\mathrm{d}z}{2\pi i} \mathcal{T}(z) = \pi T \mathcal{N}_{\Psi}^{2}\,\mathrm{E}_{(t)} \sum_{n,\,m=-\infty}^{+\infty} :\, b_n\, b_m^{*} :\, x_{(t)}^{-m-n} + \sum_{\substack{u=1\\u\neq t}}^{N} \frac{\mathrm{E}_{(u)}\mathrm{E}_{(t)}}{x_{(t)} - x_{(u)}}, \tag{3.222}$$

where $\mathcal{C}_{x_{(t)}}$ is a small loop around $x_{(t)}$. Therefore we have

$$\frac{\partial}{\partial x_{(t)}} \ln \left\langle \mathrm{R}\left[\mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)}) \prod_{\substack{u=1\\u\neq t}}^{N} \mathrm{S}_{\mathrm{E}_{(u)}}(x_{(u)}) \right] \right\rangle = \sum_{u\neq t} \frac{\mathrm{E}_{(u)}\mathrm{E}_{(t)}}{x_{(t)} - x_{(u)}}, \tag{3.223}$$

which can be solved by

$$\left\langle \mathrm{R}\left[\mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)}) \prod_{\substack{u=1 \\ u\neq t}}^{N} \mathrm{S}_{\mathrm{E}_{(u)}}(x_{(u)}) \right] \right\rangle = \mathcal{N}_0\big(\{\mathrm{E}_{(t)}\}\big) \prod_{\substack{t=1 \\ t>u}}^{N} \big(x_{(u)} - x_{(t)}\big)^{\mathrm{E}_{(u)}\mathrm{E}_{(t)}}. \tag{3.224}$$

The constant $\mathcal{N}_0\big(\{\mathrm{E}_{(t)}\}\big)$ which depends on the $\mathrm{E}_{(t)}$ only can be fixed by using the O.P.E. The last equation reproduces the usual bosonization procedure.

In similar way we can compute all correlators

$$\begin{aligned} &\frac{\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \mathrm{R}\left[\prod_i \Psi(x_i) \prod_j \Psi^*(x_j) \right] \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle}{\left\langle \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \middle| \Omega_{\{x_{(t)},\,\mathrm{E}_{(t)}\}} \right\rangle} \\ =&\frac{\left\langle \mathrm{R}\left[\prod_i \Psi(x_i) \prod_j \Psi^*(x_j) \prod_{t=1}^{N} \mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)}) \right] \right\rangle}{\left\langle \mathrm{R}\left[\prod_{t=1}^{N} \mathrm{S}_{\mathrm{E}_{(t)}}(x_{(t)}) \right] \right\rangle}, \end{aligned} \tag{3.225}$$

using Wick's theorem since the algebra and the action of creation and annihilation operators is the usual. In particular taking one $\Psi(z)$ and one $\Psi^*(w)$ we get the Green function which is nothing else but the contraction in equation (3.191).

4 Summary and Conclusion

We thus showed that the specific geometry of the intersecting D-branes leads to different results when computing the value of the classical action, that is the leading contribution to the Yukawa couplings in string theory. In particular in the Abelian case the value of the action is exactly the area formed by the intersecting D-branes in the $\mathbb{R}^2$ plane, that is the string worldsheet is completely contained in the polygon on the plane. The difference between the SO(4) case and SU(2) is more subtle as in the latter there are complex coordinates in $\mathbb{R}^4$ for which the classical string solution is holomorphic in the upper half plane. In the generic case presented so far this is in general no longer true. The reason can probably be traced back to supersymmetry, even though we only dealt with the bosonic string. In fact when considering SU(2) rotated D-branes part of the spacetime supersymmetry is preserved, while this is not the case for SO(4) rotations.

In the general case there does not seem to be any possible way of computing the action (2.136) in term of the global data. Most probably the value of the action is larger than in the holomorphic case since the string is no longer confined to a plane. Given the nature of the rotation its worldsheet has to bend in order to be attached to the D-brane as pictorially shown in Figure 2.7 in the case of a 3-dimensional space. The general case we considered then differs from the known factorised case by an additional contribution in the on-shell action which can be intuitively understood as a small "bump" of the string worldsheet in proximity of the boundary.

In a technical and direct way we also showed the computation of amplitudes involving an arbitrary number of Abelian spin and matter fields. The approach we introduced does not

generally rely on CFT techniques and can be seen as an alternative to bosonization and old methods based on the Reggeon vertex. Starting from this work the future direction may involve the generalisation to non Abelian spin fields and the application to twist fields. In this sense this approach might be the only way to compute the amplitudes involving these complicated scenarios. This analytical approach may also shed some light on the non existence of a technique similar to bosonisation for twist fields.

Part II

Cosmological Backgrounds and Divergences

5 Introduction

In the previous part we mainly focused on the mathematical tools needed to compute amplitudes in a (semi-)phenomenologically viable string theory framework of particle physics. This ultimately led to the introduction of intersecting D-branes and point-like defects to perform the calculation of correlation functions involving twist and spin fields, inevitably necessary when considering chiral matter fields. While this is indeed a good starting point to build an entire string phenomenology, the theory cannot be limited to the study of particle physics models. String theory is in fact considered to be one of the candidate theories for the description of quantum gravity alongside the nuclear interactions. As a *theory of everything* it is therefore fascinating to analyse cosmological implications as seen from its description. In this part of the thesis we focus on the implications of the string theory when considering for instance the Big Bang singularity, or, broadly speaking, singularities which exist in one point in time (i.e. space-like).[37] Among the different possible descriptions of such space-like singularities [86] we concentrate on string theory solutions on time-dependent orbifolds as they represent the simplest models describing such phenomena. Before delving into the subject we briefly present their definition and the reason behind their relevance in what follows [87]–[89].

5.1 Quotient Spaces and Orbifolds

First of all we recall the formal definition of orbifold to better introduce the idea of a manifold locally isomorphic to a quotient space. Let therefore M be a topological space and G a group with an action $\mathscr{G} : G \times M \to M$ defined by $\mathscr{G}(g,\, p) = gp$ for $g \in G$ and $p \in M$. Then the *isotropy subgroup* (or *stabiliser*) of $p \in M$ is $G_p = \{g \in G \mid gp = p\}$ such that $G_{gp} = g^{-1}\, G_p\, g$. Given an element $p \in M$ its *orbit* is $Gp = \{gp \in M \mid g \in G\}$. The action of the group is said *transitive* if $Gp = M$ and *effective* if its kernel is trivial, i.e. $\ker \mathscr{G} = \{\mathbb{1}\}_{\mathscr{M}}$. The *orbit space* M/G is the set of equivalence classes given by the orbital partitions and inherits the quotient topology from M.

Let now M be a manifold and G a Lie group acting continuously and transitively on M. For every point $p \in M$ we can define a continuous bijection $\lambda_p \colon G/G_p \to Gp = M$.[38] Such map is a diffeomorphism if M and G are locally compact spaces and M/G is in turn a manifold itself. If G is a discrete or finite group the action is called *properly discontinuous*, that is for every $U \subset M$ then $\{g \in G \mid U \cap gU \neq \emptyset\}$ is finite.

The definition of orbifold intuitively includes quotient manifolds such as M/G: analogously to manifold which are locally Euclidean, in the broad sense orbifolds are locally modelled by quotients with actions given by finite groups. An *orbifold chart* $\left(\widetilde{U},\, G,\, \phi\right)$ of dimension $n \in \mathbb{N}$ for an open subset $U \in M$ is made of:

- a connected open subset $\widetilde{U} \subset \mathbb{R}^n$,
- a finite group G acting acting on $\widetilde{U}$,

[37] They are intended as distinct from time-like singularities such as black holes which are present for extended periods of time in one spatial point. The space-like singularities we consider are the opposite: they exist in a given instant in time but could in principle cover an extended hypersurface in space.

[38] For any $U \subset M$ and a given $p \in M$ then $\lambda_p^{-1}(U) = \pi_p(\{g \in G \mid gp \in U\})$ where $\pi_p \colon G \to G/G_p$ is a projection map. Thus $\lambda_p^{-1}(U)$ is an open subset if U is open: the bijection is continuous.

- a map $\phi\colon \widetilde{U} \to M$ defined by the composition $\phi = \pi \circ \mathscr{P}$ where $\mathscr{P}\colon \widetilde{U} \to \widetilde{U}/G$ defines the orbits and $\pi\colon \widetilde{U}/G \to M$.

An embedding $\eta\colon \left(\widetilde{U}_2, G_1, \phi_1\right) \hookrightarrow \left(\widetilde{U}_2, G_2, \phi_2\right)$ between two charts is such that $\phi_2 \circ \eta = \phi_1$. Suppose now $U_i = \phi_i\left(\widetilde{U}_i\right)$ for $i = 1, 2$ and take $p \in U_1 \cap U_2$. The charts are *compatible* if there exist an open subset V such that $p \in V \subset U_1 \cap U_2$ and a chart $\left(\widetilde{V}, G, \phi\right)$ admitting two embeddings in the previous charts. A n-dimensional *orbifold atlas* is then a collection $\{(U_i, G_i, \phi_i)\}_{i\in I}$ of compatible n-dimensional orbifold charts covering M. The n-dimensional *orbifold* $\mathscr{O}$ is finally defined as a paracompact Hausdorff topological space together with a n-dimensional orbifold atlas.[39]

5.1.1 Orbifolds and Strings

In string theory the notion of orbifold has a more stringent characterisation with respect to pure mathematics. Differently from the general definition, orbifolds in physics usually appear as a global orbit space M/G where M is a manifold and G the group of its isometries, often leading to the presence of *fixed points* (i.e. points in the manifold which are left invariant by the action of G) where singularities emerge due to the presence of additional degrees of freedom given by *twisted states* of the string [90], [91]. They are commonly introduced as singular limits of CY manifolds [17], which in turn can be recovered using algebraic geometry to smoothen the singular points. However they can also be used to model peculiar time-dependent backgrounds [88], [89], [92]–[96]. They are in fact good toy models to study Big Bang scenarios in string theory. We focus specifically on the study of such cosmological singularities in the framework of string theory defined on time-dependent orbifolds.

6 Time Dependent Orbifolds

6.1 Motivation

The first attempts to consider space-like [97] or light-like singularities [98], [99] by means of orbifold techniques yielded divergent four points *closed string* amplitudes (see [95], [96] for reviews). These singularities are commonly assumed to be connected to a large backreaction of the incoming matter into the singularity due to the exchange of a single graviton [100], [101]. This claim was already questioned in the literature where the O-plane orbifold was constructed. This orbifold should in fact be stable against the gravitational collapse but it exhibits divergences in the amplitudes (see the discussion in [95]). In what follows we show a direct computation showing that the presence of the divergence is not related to a gravitational response.

Unnoticed in the Null Boost Orbifold (NBO) [98], even the four *open string* tachyons amplitude

[39] In this context paracompact refers to a topological space M which admits open covers with a *locally finite* refinement. In other words let $U = \{U_i\}_{i\in I}$ be a cover and $V = \{V_j\}_{j\in J}$ be its refinement (i.e. $\forall j \in J$, $\exists i \in I \mid V_j \subset U_i$). Then U is locally finite if $\forall p \in M$ there is a neighbourhood $B(p)$ of p such that $\{i \in I \mid U_i \cap B(p) \neq \emptyset\}$ is finite.

is divergent. Since we are working at tree level gravity is not an issue. In fact in [98] the four tachyons amplitude in the divergent region reads

$$A_4 \sim \int\limits_{q\sim\infty} \frac{\mathrm{d}q}{|q|} \mathscr{A}(q) \tag{6.1}$$

where $\mathscr{A}_{\text{closed}}(q) \sim q^{4-\alpha'\|\vec{p}_\perp\|^2}$ and $\mathscr{A}_{\text{open}}(q) \sim q^{1-\alpha'\|\vec{p}_\perp\|^2} \operatorname{tr}\left([T_1, T_2]_+ [T_3, T_4]_+\right)$ (T_i for $i = 1, 2, 3, 4$ are Chan-Paton matrices). Moreover divergences in string amplitudes are not limited to four points: interestingly we show that the open string three point amplitude with two tachyons and the first massive state may be divergent when some *physical* polarisations are chosen. The true problem is therefore not related to a gravitational issue but to the non existence of the effective field theory. In fact when we express the theory using the eigenmodes of the kinetic terms some coefficients do not exist, not even as a distribution. This holds true for both open and closed string sectors since it manifests also in the four scalar contact term. The issue can be roughly traced back to the vanishing volume of a subspace and the existence of a discrete zero mode of the Laplacian on this subspace.

As an introduction to the problem we first deal with singularities of the open string sector. We try to build a consistent scalar QED and show that the vertex with four scalar fields is ill defined. Divergences in scalar QED are due to the behaviour of the eigenfunctions of the scalar d'Alembertian near the singularity but in a somehow unexpected way. Near the singularity $u = 0$ in lightcone coordinates almost all eigenfunctions behave as $\frac{1}{\sqrt{|u|}} e^{i\frac{\mathcal{A}}{u}}$ with $\mathcal{A} \neq 0$. The product of N eigenfunctions gives a singularity $|u|^{-N/2}$ which is technically not integrable. However the exponential term $e^{i\frac{\mathcal{A}}{u}}$ allows for an interpretation as distribution when $\mathcal{A} = 0$ is not an isolated point. When $\mathcal{A} = 0$ is isolated the singularity is definitely not integrable and there is no obvious interpretation as a distribution. Specifically in the NBO we find $\mathcal{A} \sim \frac{l^2}{k_+}$ where l is the momentum along the compact direction. As a consequence we find the eigenfunction associated to the discrete momentum $l = 0$ along the orbifold compact direction with an isolated $\mathcal{A} = 0$. It is the eigenfunction which is constant along that direction and it is the root of all divergences.

We then check whether the most obvious ways of regularizing the theory by making $\mathcal{A}$ not vanishing may work. The first regularisation we try is to use a Wilson line along the compact direction even though the diverging three point string amplitude involves an anti-commutator of the Chan-Paton factor therefore it is divergent also for a neutral string, i.e. for a string with both ends attached to the same D-brane. This kind of string does not feel Wilson lines. Moreover anti-commutators are present in amplitudes with massive states in unoriented and supersymmetric strings and therefore neither worldsheet parity nor supersymmetry can help. The second obvious regularisation is the introduction of higher derivatives couplings to the Ricci tensor which is the only non vanishing tensor associated to the (regularised) metric. In any case it seems that a sensible regularisation must couple to all open string in the same way and this suggests a gravitational coupling. We then give a cursory look to whether closed string winding modes could help [102], as already suggested in [97], [99] in analogy to the resolution of static singularities. Twisted closed strings become massless near the singularity and they should in some way be included. They generate a background potential $B_{\mu\nu}$ which is equivalent to a electromagnetic background from the open string perspective. Under a plausible modification of the scalar action which is suggested by the two-tachyons—two-photons amplitude the problems seem to be solvable.

In any case the origin of the string divergence seems to originate from the lack of contact

terms in the effective field theory. Since these terms arise from string theory also through the exchange of massive string states we examine three point amplitudes with one massive state. A deeper understanding of the subject requires the study of the polarisations of the massive state on the orbifold as seen from the covering Minkowski space before the computation of the overlap of the wave functions. We then go back to string theory and we verify that in the NBO the open string three points amplitude with two tachyons and one first level massive string state does indeed diverge when some physical polarisation are chosen.

We then introduce the generalised Null Boost Orbifold (GNBO) as a generalisation of the NBO which still has a light-like singularity and is generated by one Killing vector. However in this model there are two directions associated with $\mathcal{A}$, one compact and one non compact. We can then construct the scalar QED and the effective field theory which extends it with the inclusion of higher order terms since all terms have a distributional interpretation. However if a second Killing vector is used to compactify the formerly non compact direction, the theory has again the same problems as in the NBO. In the literature there are however also other attempts at regularizing the NBO such as the Null Brane. This kind of orbifold was originally defined in [93], [95] and studied in perturbation theory in [99]. The Null Brane shares with the GNBO the existence of a non compact direction on the orbifold. In this case it is indeed possible to build single particle wave functions which leads to the convergence of the smeared amplitudes.

We finally present also a brief examination of the Boost Orbifold (BO) where the divergences are generally milder [92]. The scalar eigenfunctions behave in time t as $|t|^{\pm i\frac{l}{\Delta}}$ near the singularity but there is one eigenfunction which behaves as $\log(|t|)$ and again it is the constant eigenfunction along the compact direction which is the origin of all divergences. In particular the scalar QED on the BO can be defined and the first term which gives a divergent contribution is of the form $\left|\phi\ \dot{\phi}\right|^2$, i.e. divergences are hidden into the derivative expansion of the effective field theory. Again three points open string amplitudes with one massive state diverge.

6.2 Scalar QED on NBO and Divergences

As discussed the four open string tachyons amplitude diverges in the NBO. The literature on the subject (see for instance [95] and references therein) suggests that this can be cured by the eikonal resummation. We therefore consider the scalar QED on the NBO as a first approach. In this case all eigenmodes can be written using elementary functions thus making the issues even more evident. Its action is given by

$$\mathrm{S}_{\mathrm{sQED}} = \int_{\Omega} \mathrm{d}^D x\ \sqrt{-\det g}\Big(-(D^{\mu}\phi)^*\, D_{\mu}\phi - M^2(\phi^*)\,\phi - \frac{1}{4} f^{\mu\nu}\, f_{\mu\nu} - \frac{g_4}{4}|\phi|^4\Big), \tag{6.2}$$

with

$$D_{\mu}\phi = (\partial_{\mu} - i\, e\, a_{\mu})\phi, \qquad f_{\mu\nu} = \partial_{\mu} a_{\nu} - \partial_{\nu} a_{\mu}. \tag{6.3}$$

We reserve small letters for quantities defined on the orbifold and capital letters for those defined in flat space. Moreover Ω denotes the orbifold. We will construct directly both the scalar and the spin-1 eigenfunctions which we can use as a starting point for the perturbative computations.

6.2.1 Geometric Preliminaries

In Minkowski spacetime $\mathscr{M}^{1,D-1}$ with coordinates $(x^\mu) = \big(x^+, x^-, x^2, \vec{x}\big)$ and metric

$$\mathrm{d}s^2 = -2\,\mathrm{d}x^+\,\mathrm{d}x^- + \big(\mathrm{d}x^2\big)^2 + \eta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,, \tag{6.4}$$

we consider the following change of coordinates to $(x^\alpha) = (u,\, v,\, z,\, \vec{x})$

$$\begin{cases} x^- &= u \\ x^2 &= \Delta u z \\ x^+ &= v + \frac{1}{2}\Delta^2 u z^2 \end{cases} \quad \Leftrightarrow \quad \begin{cases} u &= x^- \\ z &= \frac{x^2}{\Delta\, x^-} \\ v &= x^+ - \frac{1}{2}\frac{(x^2)^2}{x^-} \end{cases}. \tag{6.5}$$

Then the metric becomes:

$$\mathrm{d}s^2 = -2\ \mathrm{d}u\ \mathrm{d}v + (\Delta u)^2(\mathrm{d}z)^2 + \eta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,, \tag{6.6}$$

along with the non vanishing geometrical quantities

$$-\det g = (\Delta u)^2, \tag{6.7}$$

and

$$\Gamma_z{}^v{}_z = \Delta^2 u, \qquad \Gamma_u{}^z{}_z = u^{-1}. \tag{6.8}$$

Riemann and Ricci tensor components however vanish since at this stage we only performed a change of coordinates from the original Minkowski spacetime. Locally it is the same as the NBO and they must have the same local differential geometry.

The NBO is introduced by identifying points along the orbits of the Killing vector:

$$\begin{aligned} \kappa &= -i(2\pi\Delta)J_{+2} \\ &= (2\pi\Delta)\,(x^2\partial_+ + x^-\partial_2) \\ &= 2\pi\partial_z, \end{aligned} \tag{6.9}$$

in such a way that

$$x^\mu \equiv \mathcal{K}^n\, x^\mu, \qquad n \in \mathbb{Z}, \tag{6.10}$$

where $\mathcal{K}^n = e^{n\kappa}$, leads to the identifications

$$x = \begin{pmatrix} x^- \\ x^2 \\ x^+ \\ \vec{x} \end{pmatrix} \equiv \mathcal{K}^n x = \begin{pmatrix} x^- \\ x^2 + n(2\pi\Delta)x^- \\ x^+ + n(2\pi\Delta)x^2 + \frac{1}{2}n^2(2\pi\Delta)^2 x^- \\ \vec{x} \end{pmatrix} \tag{6.11}$$

or to

$$(u,\, v,\, z,\, \vec{x}) \equiv (u,\, v,\, z + 2\pi n,\, \vec{x}) \tag{6.12}$$

in coordinates (x^α) where $\kappa = 2\pi\partial_z$ is a global Killing vector.

As a reference for the future, we notice that we could regularise the metric as

$$\mathrm{d}s^2 = -2\ \mathrm{d}u\ \mathrm{d}v + \Delta^2\big(u^2 + \epsilon^2\big)(\mathrm{d}z)^2 + \eta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,. \tag{6.13}$$

The non vanishing geometrical quantities are then:

$$-\det g = \Delta^2\big(u^2+\epsilon^2\big), \tag{6.14}$$

and

$$\Gamma_z{}^v{}_z = \Delta^2 u, \qquad \Gamma_u{}^z{}_z = \frac{u}{u^2+\epsilon^2}, \tag{6.15}$$

which lead to the following Riemann and Ricci tensor components:

$$R^z{}_{uzu} = -\frac{\epsilon^2}{\left(u^2+\epsilon^2\right)^2}, \quad R^v{}_{zzu} = -\frac{\Delta^2\epsilon^2}{u^2+\epsilon^2}, \quad Ric_{uu} = -\frac{\epsilon^2}{\left(u^2+\epsilon^2\right)^2}. \tag{6.16}$$

Since $\delta_{\text{reg}}(u) = \frac{1}{\pi}\frac{\epsilon}{u^2+\epsilon^2}$ then $R^z{}_{uzu} = -\pi^2[\delta_{\text{reg}}(u)]^2$.

6.2.2 Free Scalar Action

We study the eigenmodes of the Laplacian operator to diagonalize the scalar kinetic term given by:[40]

$$\begin{aligned}
\mathrm{S}^{(\text{kinetic})}_{\text{sQED}}[\phi] &= \int_\Omega \mathrm{d}^D x\ \sqrt{-\det g}\ \left(-g^{\alpha\beta}\partial_\alpha\phi^*\,\partial_\beta\phi - M^2\phi^*\,\phi\right) \\
&= \int_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{x} \int_{-\infty}^{+\infty}\mathrm{d}u \int_{-\infty}^{+\infty}\mathrm{d}v \int_0^{2\pi}\mathrm{d}z\,|\Delta u| \\
&\times\left(\partial_u\phi^*\,\partial_v\phi\ +\partial_v\phi^*\,\partial_u\phi\ -\frac{1}{(\Delta u)^2}\partial_z\phi^*\,\partial_z\phi\ -\partial_i\phi^*\,\partial_i\phi - M^2\phi^*\,\phi\right).
\end{aligned} \tag{6.17}$$

The solution to the equation of motion is enough when we want to perform the canonical quantization. Since we use Feynman diagrams we consider the path integral approach: we take off-shell modes and solve the eigenvalue problem $\Box\phi_r = r\phi_r$. Comparing with the flat case we see that r is $2\,k_-\,k_+ - \left\|\vec{k}\right\|^2$ when k is the impulse in flat coordinates. We therefore have

$$-2\partial_u\partial_v\phi_r - \frac{1}{u}\partial_v\phi_r + \frac{1}{(\Delta u)^2}\partial_z^2\phi_r + \partial_i^2\phi_r = r\phi_r. \tag{6.18}$$

Using Fourier transforms it follows that the eigenmodes are

$$\phi_{\{k_+,\,l,\,\vec{k},\,r\}}(u,\,v,\,z,\,\vec{x}) = e^{ik_+v+ilz+i\vec{k}\cdot\vec{x}}\,\widetilde{\phi}_{\{k_+,\,l,\,\vec{k},\,r\}}(u), \tag{6.19}$$

with

$$\widetilde{\phi}_{\{k_+,\,l,\,\vec{k},\,r\}}(u) = \frac{1}{\sqrt{(2\pi)^D\ |2\Delta k_+\,u|}}\,e^{-i\frac{l^2}{2\Delta^2 k_+}\frac{1}{u}+i\frac{\|\vec{k}\|^2+r}{2k_+}u}, \tag{6.20}$$

[40] The factor $-g^{\alpha\beta}$ is due to the choice of the East coast convention for the metric, namely:

$$-g^{\alpha\beta}\partial_\alpha\phi^*\,\partial_\beta\phi - M^2\phi^*\,\phi \sim \left|\dot{\phi}\right|^2 - M^2|\phi|^2 \sim \mathrm{E}^2 - M^2.$$

and

$$\phi^*_{\{k_+,\,l,\,\vec{k},\,r\}}(u,\,v,\,z,\,\vec{x}) = \phi_{\{-k_+,\,-l,\,-\vec{k},\,r\}}(u,\,v,\,z,\,\vec{x}). \tag{6.21}$$

We chose the numeric factor in order to get a canonical normalisation:

$$\begin{aligned}
&\left(\phi_{\{k_{(1)\,+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\}},\,\phi_{\{k_{(2)\,+},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\}}\right)\\
=&\int\limits_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{x} \int\limits_{-\infty}^{+\infty}\mathrm{d}u \int\limits_{-\infty}^{+\infty}\mathrm{d}v \int\limits_{0}^{2\pi}\mathrm{d}z\,|\Delta u|\,\phi_{\{k_{(1)\,+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\}}\,\phi_{\{k_{(2)\,+},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\}}\\
=&\delta^{D-3}(\vec{k}_{(1)}+\vec{k}_{(2)})\,\delta(r_{(1)}-r_{(2)})\,\delta(k_{(1)\,+}+k_{(2)\,+})\,\delta_{l_{(1)}+l_{(2)},\,0}.
\end{aligned} \tag{6.22}$$

We can then perform the off-shell expansion

$$\phi(u,\,v,\,z,\,\vec{x}) = \int\limits_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{k} \int\limits_{-\infty}^{+\infty}\mathrm{d}k_+ \int\limits_{-\infty}^{+\infty}\mathrm{d}r \sum_{l=-\infty}^{+\infty} \mathcal{A}_{\{k_+,\,l,\,\vec{k},\,r\}}\,\phi_{\{k_+,\,l,\,\vec{k},\,r\}}(u,\,v,\,z,\,\vec{x}), \tag{6.23}$$

such that the scalar kinetic term becomes

$$\mathrm{S}^{(\mathrm{kinetic})}_{\mathrm{sQED}}[\mathcal{A}] = \int\limits_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{k} \int\limits_{-\infty}^{+\infty}\mathrm{d}k_+ \int\limits_{-\infty}^{+\infty}\mathrm{d}r \sum_{l=-\infty}^{+\infty} (r-M^2)\,\mathcal{A}_{\{k_+,\,l,\,\vec{k},\,r\}}\,\mathcal{A}^*_{\{k_+,\,l,\,\vec{k},\,r\}}. \tag{6.24}$$

6.2.3 Free Photon Action

The action of the free photon can be written as

$$\mathrm{S}^{(\mathrm{kinetic})}_{\mathrm{sQED}}[a] = \int\limits_{\Omega} \mathrm{d}^D x\,\sqrt{-\det g}\left(-\frac{1}{2}g^{\alpha\beta}g^{\gamma\delta}D_\alpha a_\gamma (D_\beta a_\delta - D_\delta a_\beta)\right). \tag{6.25}$$

We choose to enforce the Lorenz gauge:[41]

$$D^\alpha a_\alpha = -\frac{1}{u}a_v - \partial_u a_v - \partial_v a_u + \frac{1}{\Delta^2 u^2}\partial_z a_z + \eta^{ij}\partial_i a_j = 0. \tag{6.26}$$

As covariant derivatives commute since we are locally flat, the E.O.M. read $(\Box a)_\alpha = 0$. Explicitly we have:

$$\begin{aligned}
(\Box a)_u &= \frac{1}{u^2}a_v - \frac{2}{\Delta^2 u^3}\partial_z a_z + \left[-2\partial_u\partial_v - \frac{1}{u}\partial_v + \frac{1}{\Delta^2 u^2}\partial_z^2 + \eta^{ij}\partial_i\partial_j\right]a_u,\\
(\Box a)_v &= \left[-2\partial_u\partial_v - \frac{1}{u}\partial_v + \frac{1}{\Delta^2 u^2}\partial_z^2 + \eta^{ij}\partial_i\partial_j\right]a_v,\\
(\Box a)_z &= -\frac{2}{u}\partial_z a_v + \left[-2\partial_u\partial_v + \frac{1}{u}\partial_v + \frac{1}{\Delta^2 u^2}\partial_z^2 + \eta^{ij}\partial_i\partial_j\right]a_z,\\
(\Box a)_i &= \left[-2\partial_u\partial_v - \frac{1}{u}\partial_v + \frac{1}{\Delta^2 u^2}\partial_z^2 + \eta^{ij}\partial_i\partial_j\right]a_i.
\end{aligned} \tag{6.27}$$

[41] Indeed it is exactly the usual Lorenz gauge since locally the space is Minkowski.

As in the previous scalar case we are actually interested in solving the eigenmodes problem $(\Box a_r)_\alpha = r\, a_{r\,\alpha}$. We proceed hierarchically: first we solve for a_v and a_i whose equations are the same as in the scalar field, then we insert the solutions as a source in the equation for a_z and eventually we solve for a_u.[42] We get the solutions:

$$\begin{aligned}
\left\|\widetilde{a}_{\{k_+,l,\vec{k},r\}\,\alpha}(u)\right\| = \begin{pmatrix}\widetilde{a}_u\\ \widetilde{a}_v\\ \widetilde{a}_z\\ \widetilde{a}_i\end{pmatrix} &= \sum_{\underline{\alpha}\in\{\underline{u},\underline{v},\underline{z},\underline{i}\}} \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{\alpha}} \left\|\widetilde{a}^{\underline{\alpha}}_{\{k_+,l,\vec{k},r\}\,\alpha}(u)\right\| \\
&= \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{u}} \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \widetilde{\Phi}_{\{k_+,l,\vec{k},r\}}(u) \\
&+ \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{v}} \begin{pmatrix}\frac{i}{2k_+u} + \frac{1}{2}\left(\frac{l}{\Delta k_+}\right)^2 \frac{1}{u^2}\\ 1\\ \frac{l}{k_+}\\ 0\end{pmatrix} \widetilde{\Phi}_{\{k_+,l,\vec{k},r\}}(u) \\
&+ \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{z}} \begin{pmatrix}\frac{l}{\Delta k_+|u|}\\ 0\\ \Delta|u|\\ 0\end{pmatrix} \widetilde{\Phi}_{\{k_+,l,\vec{k},r\}}(u) \\
&+ \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{j}} \begin{pmatrix}0\\0\\0\\ \delta_{\underline{ij}}\end{pmatrix} \widetilde{\Phi}_{\{k_+,l,\vec{k},r\}}(u),
\end{aligned} \tag{6.28}$$

then we can expand the off-shell fields as

$$a_\alpha(u,\, v,\, z,\, \vec{x}) = \int \mathscr{D}k \sum_{\underline{\alpha}\in\{\underline{u},\underline{v},\underline{z},\underline{i}\}} \sum_{l=-\infty}^{+\infty} \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{\alpha}}\, a^{\underline{\alpha}}_{\{k_+,l,\vec{k},r\}\,\alpha}(u,\, v,\, z,\, \vec{x}), \tag{6.29}$$

where

$$a^{\underline{\alpha}}_{\{k_+,l,\vec{k},r\}\,\alpha}(u,\, v,\, z,\, \vec{x}) = \widetilde{a}^{\underline{\alpha}}_{\{k_+,l,\vec{k},r\}\,\alpha}(u)\, e^{i\left(k_+v + lz + \vec{k}\cdot\vec{x}\right)} \tag{6.30}$$

and $\int \mathscr{D}k = \int\limits_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{k} \int\limits_{-\infty}^{+\infty} \mathrm{d}k_+ \int\limits_{-\infty}^{+\infty} \mathrm{d}r$.

We can also compute the normalisation as

$$\begin{aligned}
\left(a_{(1)},\, a_{(2)}\right) &= \int\limits_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{x} \int\limits_{-\infty}^{+\infty} \mathrm{d}u \int\limits_{-\infty}^{+\infty} \mathrm{d}v \int\limits_{0}^{2\pi} \mathrm{d}z\, |\Delta u| \\
&\times g^{\alpha\beta}\, a_{\{k_{(1)+},l_{(1)},\vec{k}_{(1)},r_{(1)}\}\,\alpha}\, a_{\{k_{(2)+},l_{(2)},\vec{k}_{(2)},r_{(2)}\}\,\beta} \\
&= \mathcal{E}_{\{k_{(1)+},l_{(1)},\vec{k}_{(1)},r_{(1)}\}} \circ \mathcal{E}_{\{k_{(2)+},l_{(2)},\vec{k}_{(2)},r_{(2)}\}} \\
&\times \delta^{D-3}(\vec{k}_{(1)} + \vec{k}_{(2)})\, \delta(r_{(1)} - r_{(2)})\, \delta(k_{(1)+} + k_{(2)+})\, \delta_{l_{(1)}+l_{(2)},0},
\end{aligned} \tag{6.31}$$

[42] Notice that inside the square brackets of the differential equation for a_z there is a different sign for the term $\frac{1}{u}\partial_v$ with respect to the equation for the scalar field.

where:[43]

$$\mathcal{E}_{(1)} \circ \mathcal{E}_{(2)} = -\mathcal{E}_{(1)\,\underline{u}}\,\mathcal{E}_{(2)\,\underline{v}} - \mathcal{E}_{(1)\,\underline{v}}\,\mathcal{E}_{(2)\,\underline{u}} + \mathcal{E}_{(1)\,\underline{z}}\,\mathcal{E}_{(2)\,\underline{z}} + \eta^{\underline{ij}}\,\mathcal{E}_{(1)\,\underline{i}}\,\mathcal{E}_{(2)\,\underline{j}}. \tag{6.32}$$

Finally the Lorenz gauge reads

$$\eta^{\underline{ij}}\,k_i\,\mathcal{E}_{\{k_+,\,l,\,\vec{k},\,r\}\,\underline{j}} - k_+\,\mathcal{E}_{\{k_+,\,l,\,\vec{k},\,r\}\,\underline{u}} - \frac{\left\|\vec{k}\right\|^2 + r}{2\,k_+}\mathcal{E}_{\{k_+,\,l,\,\vec{k},\,r\}\,\underline{v}} = 0, \tag{6.33}$$

which does not impose any constraint on the transverse polarisation $\mathcal{E}_{\{k_+,\,l,\,\vec{k},\,r\}\,\underline{z}}$. The photon kinetic term becomes

$$\mathrm{S}^{(\mathrm{kinetic})}_{\mathrm{sQED}}[\mathcal{E}] = \int\limits_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{k} \int\limits_{-\infty}^{+\infty} \mathrm{d}k_+ \int\limits_{-\infty}^{+\infty} \mathrm{d}r \sum_{l=-\infty}^{+\infty} \frac{r}{2}\,\mathcal{E}_{\{k_+,\,l,\,\vec{k},\,r\}} \circ \mathcal{E}^*_{\{k_+,\,l,\,\vec{k},\,r\}}. \tag{6.34}$$

6.2.4 Cubic Interaction

With the definition of the d'Alembertian eigenmodes we can now examine the cubic vertex which reads

$$\mathrm{S}^{(\mathrm{cubic})}_{\mathrm{sQED}}[\phi,\,a] = \int\limits_{\Omega} \mathrm{d}^D x\,\sqrt{-\det g}\,\big(-i\,e\,g^{\alpha\beta}a_\alpha(\phi^*\,\partial_\beta\phi - \partial_\beta\phi^*\,\phi)\big). \tag{6.35}$$

Its computation involves integrals such as

$$\int \mathrm{d}u\;|\Delta u|\left(\frac{l}{u}\right)^2 \prod_{i=1}^{3} \widetilde{\phi}_{\{k_{(i)}+,\,l_{(i)},\,\vec{k}_{(i)},\,r_{(i)}\}} \sim \int\limits_{u\sim 0} \mathrm{d}u\,\left(\frac{l^2}{|u|^{\frac{5}{2}}}\right) e^{-i\sum\limits_{i=1}^{3}\frac{l^2_{(i)}}{2\,\Delta^2 k_{(i)}+}\frac{1}{u}}, \tag{6.36}$$

and

$$\int \mathrm{d}u\;|\Delta u|\left(\frac{1}{u}\right) \prod_{i=1}^{3} \widetilde{\phi}_{\{k_{(i)}+,\,l_{(i)},\,\vec{k}_{(i)},\,r_{(i)}\}} \sim \int\limits_{u\sim 0} \mathrm{d}u\,\left(\frac{1}{u\,|u|^{\frac{1}{2}}}\right) e^{-i\sum\limits_{i=1}^{3}\frac{l^2_{(i)}}{2\,\Delta^2 k_{(i)}+}\frac{1}{u}}, \tag{6.37}$$

which can be interpreted as hints that the theory is troublesome. The first integral diverges if the exponential functions are all equal to unity. Fortunately it happens when all factors $l_{(i)}$ (where $i = 1, 2, 3$) vanish. In this case however the integral vanishes if we set $l_{(i)} = 0$ before its evaluation. This however suggests that when all $l_{(i)} = 0$, i.e. when the eigenfunctions are constant along the compact direction z, something suspicious is happening. On the other side when at least one l is different from zero we have an integral such as:

$$\int\limits_{u\sim 0} \mathrm{d}u\;|u|^{-\nu}\,e^{i\frac{\mathcal{A}}{u}} \sim \int\limits_{t\sim\infty} \mathrm{d}t\;t^{\nu-2}\,e^{i\mathcal{A}t}. \tag{6.38}$$

All $l_{(i)}$ are discrete but $k_{(i)\,+}$ are not thus $\mathcal{A}$ has an isolated zero. Otherwise it has continuous value and may be given a distributional meaning, similar to a derivative of the Dirac delta

[43] We use a shortened version of the polarisations $\mathcal{E}$ for the sake of readability. We write $\mathcal{E}_{(n)\,\underline{\alpha}} = \mathcal{E}_{\{k_{(n)}+,\,l_{(n)},\,\vec{k}_{(n)},\,r_{(n)}\}\,\underline{\alpha}}$ thus hiding the understood dependence of the components of $\mathcal{E}_{(n)}$ on the momenta.

function. The second integral has the same issues when all $l_{(*)} = 0$ but, since it is not proportional to any l as it stands, it is divergent unless we consider a principal part regularization.

We can give in any case meaning to the cubic terms and we get:[44]

$$\begin{aligned}
\mathrm{S}^{(\mathrm{cubic})}_{\mathrm{sQED}}[\mathcal{A},\,\mathcal{E}] = \prod_{i=1}^{3}\left[\int\limits_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{k}_{(i)} \int\limits_{-\infty}^{+\infty} \mathrm{d}r_{(i)} \int\limits_{-\infty}^{+\infty} \mathrm{d}k_{(i)+} \sum_{l_{(i)}}\right] (2\pi)^{D-1} \\
\times e\;\delta\left(\sum_{i=1}^{3}\vec{k}_{(i)}\right)\delta\left(\sum_{i=1}^{3}k_{(i)+}\right)\delta_{\sum\limits_{i=1}^{3} l_{(i)},\,0} \\
\times \left(\mathcal{A}_{\left\{-k_{(2)+},\,-l_{(2)},\,-\vec{k}_{(2)},\,r_{(2)}\right\}}\right)^{*}\mathcal{A}_{\left\{k_{(3)+},\,l_{(3)},\,\vec{k}_{(3)},\,r_{(3)}\right\}} \\
\times \Bigg\{\mathcal{E}_{\left\{k_{(1)+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}\underline{u}}\,k_{(2)+}\,\mathcal{I}^{[0]}_{\{3\}} \\
+\mathcal{E}_{\left\{k_{(1)+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}\underline{z}}\,\frac{k_{(2)+}l_{(1)}-l_{(2)}k_{(1)+}}{\Delta k_{(1)+}}\,\mathcal{J}^{[-1]}_{\{3\}} \\
+\mathcal{E}_{\left\{k_{(1)+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}\underline{v}}\,\mathscr{F}\left(k_{(1)+},\,l_{(1)},\,k_{(2)+},\,l_{(2)},\,r_{(2)},\,\vec{k}_{(2)}\right) \\
-\eta^{\underline{i}\,j}\,\mathcal{E}_{\left\{k_{(1)+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}\underline{i}}\,{k_{(2)}}_{j}\,\mathcal{I}^{[0]}_{\{3\}} - ((2)\to(3))\Bigg\},
\end{aligned} \tag{6.39}$$

where

$$\begin{aligned}
\mathscr{F}\left(k_{(1)+},\,l_{(1)},\,k_{(2)+},\,l_{(2)},\,r_{(2)},\,\vec{k}_{(2)}\right) = \frac{\left\|\vec{k}_{(2)}\right\|^2 + r_{(2)}}{2\,k_{(2)+}}\mathcal{I}^{[0]}_{\{3\}} + i\frac{k_{(2)+}}{2\,k_{(1)+}}\mathcal{I}^{[-1]}_{\{3\}} \\
+\frac{1}{2}\frac{k_{(2)+}}{\Delta^2}\left(\frac{l_{(1)}}{k_{(1)+}}-\frac{l_{(2)}}{k_{(2)+}}\right)^2\mathcal{I}^{[-2]}_{\{3\}}.
\end{aligned} \tag{6.40}$$

In the previous expressions we also defined for future use:

$$\mathcal{I}^{[\nu]}_{(1)\ldots(N)} = \mathcal{I}^{[\nu]}_{\{N\}} = \int\limits_{-\infty}^{+\infty}\mathrm{d}u\;|\Delta u|\,u^{\nu}\prod_{i=1}^{N}\widetilde{\phi}_{\left\{k_{(i)+},\,l_{(i)},\,\vec{k}_{(i)},\,r_{(i)}\right\}} \tag{6.41}$$

$$\mathcal{J}^{[\nu]}_{\{N\}} = \int\limits_{-\infty}^{+\infty}\mathrm{d}u\;|\Delta|\,|u|^{1+\nu}\prod_{i=1}^{N}\widetilde{\phi}_{\left\{k_{(i)+},\,l_{(i)},\,\vec{k}_{(i)},\,r_{(i)}\right\}}. \tag{6.42}$$

For the sake of brevity from now on we use

$$\widetilde{\phi}_{(i)} = \widetilde{\phi}_{\left\{k_{(i)+},\,l_{(i)},\,\vec{k}_{(i)},\,r_{(i)}\right\}}, \tag{6.43}$$

$$\widetilde{\phi}_{(i)} = \widetilde{\phi}_{\left\{k_{(i)+},\,l_{(i)},\,\vec{k}_{(i)},\,r_{(i)}\right\}} \tag{6.44}$$

when not causing confusion.

[44] The notation $(2)\to(3)$ meaning is that all previous terms inside the curly brackets appear again in exactly the same structure but with momenta of particle (3) in place of those of particle (2).

6.2.5 Quartic Interactions and Divergences

The issue with the divergent vertex is even more visible when considering the quartic terms:

$$\mathrm{S}^{(\mathrm{quartic})}_{\mathrm{sQED}}[\phi,\, a] = \int\limits_{\Omega} \mathrm{d}^D x\ \sqrt{-\det g}\left(e^2\, g^{\mu\nu}\, a_\mu a_\nu\, |\phi|^2 - \frac{g_4}{4}|\phi|^4\right), \tag{6.45}$$

which can be expressed using the modes as:

$$\begin{aligned}
\mathrm{S}^{(\mathrm{quartic})}_{\mathrm{sQED}}[\phi,\, a] = \prod_{i=1}^{4}\left[\int \mathrm{d}^{D-3}\vec{k}_{(i)}\, \mathrm{d}k_{(i)+}\, \mathrm{d}r_{(i)} \sum_{l_{(i)}}\right] (2\pi)^{D-1} \\
\times\, \delta\left(\sum_{i=1}^{4}\vec{k}_{(i)}\right)\delta\left(\sum_{i=1}^{4}k_{(i)+}\right)\delta_{\sum\limits_{i=1}^{4} l_{(i)},\,0} \\
\times \Bigg\{ e^2 \left(\mathcal{A}_{\left\{-k_{(3)+},\,-l_{(3)},\,-\vec{k}_{(3)},\,r_{(3)}\right\}}\right)^{*} \mathcal{A}_{\left\{k_{(4)+},\,l_{(4)},\,\vec{k}_{(4)},\,r_{(4)}\right\}} \\
\times \Bigg[\left(\mathcal{E}_{\left\{k_{(1)+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}} \circ \mathcal{E}_{\left\{k_{(2)+},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\right\}}\right) \mathcal{I}^{[0]}_{\{4\}} \\
- \frac{i}{2}\mathcal{E}_{\left\{k_{(1)+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}} \underline{v}\, \mathcal{E}_{\left\{k_{(2)+},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\right\}} \underline{v} \left(\frac{1}{k_{(2)+}} + \frac{1}{k_{(1)+}}\right) \mathcal{I}^{[-1]}_{\{4\}} \\
+ \frac{1}{2}\,\frac{\mathcal{E}_{\left\{k_{(1)+},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}} \underline{v}\, \mathcal{E}_{\left\{k_{(2)+},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\right\}} \underline{v}}{\Delta^2} \left(\frac{l_{(1)}}{k_{(1)+}} - \frac{l_{(2)}}{k_{(2)+}}\right)^2 \mathcal{I}^{[-2]}_{\{4\}} \Bigg] \\
- \frac{g_4}{4}\, \mathscr{A}\left(\left\{k_+,\, l,\, \vec{k},\, r\right\}\right) \mathcal{I}^{[0]}_{\{4\}} \Bigg\},
\end{aligned} \tag{6.46}$$

where

$$\begin{aligned}
\mathscr{A}\left(\left\{k_+,\, l,\, \vec{k},\, r\right\}\right) = \left(\mathcal{A}_{\left\{-k_{(1)+},\,-l_{(1)},\,-\vec{k}_{(1)},\,r_{(1)}\right\}}\right)^{*} \left(\mathcal{A}_{\left\{-k_{(2)+},\,-l_{(2)},\,-\vec{k}_{(2)},\,r_{(2)}\right\}}\right)^{*} \\
\times\, \mathcal{A}_{\left\{k_{(3)+},\,l_{(3)},\,\vec{k}_{(3)},\,r_{(3)}\right\}}\, \mathcal{A}_{\left\{k_{(4)+},\,l_{(4)},\,\vec{k}_{(4)},\,r_{(4)}\right\}}.
\end{aligned} \tag{6.47}$$

When setting $l_{(*)} = 0$ all the surviving terms are divergent. The explicit behaviour is $\mathcal{I}^{[0]}_{\{4\}} \sim \int \mathrm{d}u\ |u|^{1-4\times\frac{1}{2}}$ and $\mathcal{I}^{[-1]}_{\{4\}} \sim \int \mathrm{d}u\ u^{-1}\, |u|^{1-4\times\frac{1}{2}}$ since $\widetilde{\phi}\Big|_{l=0} \sim |u|^{-\frac{1}{2}}$. Higher order terms in the effective field theory have even worse behaviour. This makes the theory ill defined and the string theory which should give this effective theory ill defined too.

6.2.6 Failure of Obvious Divergence Regularizations

From the discussion in the previous section the origin of the divergences is the sector $l = 0$. When $l = 0$ the highest order singularity of the Fourier transformed d'Alembertian equation vanishes. Explicitly we have:

$$A\,\partial_u \widetilde{\phi}_{\left\{k_+,\,l,\,\vec{k},\,r\right\}} + B(u)\,\widetilde{\phi}_{\left\{k_+,\,l,\,\vec{k},\,r\right\}} = A\, e^{-\int^u \frac{B(u)}{A} du}\, \partial_u \left[e^{+\int^u \frac{B(u)}{A} du} \widetilde{\phi}_{\left\{k_+,\,l,\,\vec{k},\,r\right\}}\right] = 0, \tag{6.48}$$

with

$$A = (-2\,i\,k_+), \qquad B(u) = -\left(\left\|\vec{k}\right\|^2 + r\right) - i\,k_+\,\frac{1}{u} - \frac{l^2}{\Delta^2}\,\frac{1}{u^2}. \tag{6.49}$$

This implies the absence of the oscillating factor $e^{i\frac{A}{u}}$ when l vanishes. It follows that any deformation which prevents the coefficient of the highest order singularity from vanishing will do the trick.

The first and easiest possibility is to add a Wilson line along z, i.e. $a = \theta\,\mathrm{d}z$. This shifts $l \to l - e\,\theta$ and regularises the scalar QED. Unfortunately this does not work in the string theory where Wilson lines on D25-branes are not felt by the neutral strings starting and ending on the same D-brane. In fact not all interactions involve commutators of the Chan-Paton factors which vanish for neutral strings. For instance the interaction of two tachyons with the first massive state involves an anti-commutator as we discuss later. The anti-commutators are present also in amplitudes of supersymmetric strings with massive states and therefore the issue is not solved by supersymmetry.

A second possibility is to include higher derivative couplings to curvature as natural in the string theory. If we regularise the metric in a minimal way as shown at the end of Section 6.2.1, only Ric_{uu} does not vanish. We can introduce:

$$\begin{aligned} &S_{\mathrm{HE}}^{(\text{higher R})}[\phi,\, g] \\ &= \int_\Omega \mathrm{d}^D x\, \sqrt{-\det g}\left(\sum_{k=1}^{+\infty}(\alpha')^{2k-1}\prod_{j=1}^{k} g^{\mu_j\nu_j}\, g^{\rho_j\sigma_j}\, Ric_{\mu_j\rho_j}\left(\sum_{s=0}^{2k} c_{ks}\,\partial_{\nu_j}^{2k-s}\phi^*\,\partial_{\sigma_j}^{s}\phi\right)\right) \\ &= \int_\Omega \mathrm{d}^D x\, \sqrt{-\det g}\left(\alpha'\, g^{\mu\nu}\, g^{\rho\sigma}\, Ric_{\mu\rho}\left(c_{12}\phi^*\,\partial_{\nu\sigma}^2\phi + c_{11}\partial_\nu\phi^*\,\partial_\sigma\phi + c_{10}\partial_{\nu\sigma}^2\phi^*\,\phi\right)\right), \end{aligned} \tag{6.50}$$

where α' has been introduced after dimensional analysis and in order to have all adimensional c factors. Since only Ric_{uu} is non vanishing and it depends only on u, the regularised d'Alembertian eigenmode problem now reads:

$$\begin{aligned} -2\partial_u\partial_v\phi_r - \frac{u}{u^2+\epsilon^2}\partial_v\phi_r + \frac{1}{\Delta^2(u^2+\epsilon^2)}\partial_z^2\phi_r \\ + \sum_{k=1}^{+\infty}(\alpha')^{2k-1}\, C_k\, Ric_{uu}{}^k\,\partial_v^{2k}\phi + \partial_i^2\phi_r - r\,\phi_r = 0, \end{aligned} \tag{6.51}$$

with $C_k = \sum_{s=0}^{2k}(-1)^s\, c_{ks}$. We can perform the usual Fourier transform and the function $B(u)$ becomes

$$\begin{aligned} B(u) = -\left(\left\|\vec{k}\right\|^2 + r\right) - i\,k_+\,\frac{u}{u^2+\epsilon^2} - \frac{l^2}{\Delta^2}\,\frac{1}{u^2+\epsilon^2} \\ + \sum_{k=1}^{+\infty}(\alpha')^{2k-1}\, C_k\left(\frac{\epsilon^2}{(u^2+\epsilon^2)^2}\right)^k (-ik_+)^{2k}. \end{aligned} \tag{6.52}$$

When $u = 0$ we have:

$$B(0) \sim -\frac{l^2}{\Delta^2}\,\frac{1}{\epsilon^2} + \sum_{k=1}^{+\infty}(\alpha')^{2k-1}\, C_k\,\frac{(-ik_+)^{2k}}{\epsilon^{2k}}. \tag{6.53}$$

Though the correction seems to lead to a cure for the divergence, ff we consider α' and ϵ^2 uncorrelated we lose predictability. However if $\alpha' \sim \epsilon^2$ as natural in string theory we do not solve the problem since

$$B(0) \overset{\alpha' \sim \epsilon^2}{\sim} -\frac{l^2}{\Delta^2}\,\frac{1}{\epsilon^2} + \sum_{k=1}^{+\infty} C_k\,(-ik_+)^{2k}\epsilon^{2k-2} \tag{6.54}$$

and the curvature terms are no longer singular.

6.2.7 A Hope from Twisted State Background

The issue with the divergences is associated with the dipole string and its charge neutral states since the charged ones can be cured rather trivially by a Wilson line.

On the other hand we know that the usual time-like orbifolds are well defined because of a presence of a $B_{\mu\nu}$ background and this field is sourced by strings. We may switch on such a background in the open string. For open strings F is equivalent to such B field so we can consider what happens to an open string in an electromagnetic background.

The choice of such a background is limited first of all by the request that it must be an exact string solution, i.e. it needs to obey the E.O.M. derived from the Dirac–Born–Infeld action. If a closed string winds the compact direction z then it is coupled to B_{zu}, B_{zv} and B_{zi} but if we choose

$$\frac{1}{2\pi\alpha'}B(u) = f(u)\,\mathrm{d}u \wedge \mathrm{d}z\,. \tag{6.55}$$

then

$$\det(g + 2\pi\alpha' f(u)) = \det(g). \tag{6.56}$$

It is therefore a solution of the open string E.O.M. for any $f\big(u,\, v, z, x^i\big)$. As the two-tachyons—two-photons amplitude suggests, suppose that the action for a real neutral scalar ϕ is given by:

$$\begin{aligned} S^{(\text{kinetic})}_{\text{scalar}}[\phi] &= \int\limits_{\Omega} \mathrm{d}^D x\ \sqrt{-\det g}\,\frac{1}{2}\Big(-g^{\alpha\beta}\partial_\alpha\phi\,\partial_\beta\phi - M^2\phi^2 + c_1(\alpha')^2\,\partial_\mu\phi\,\partial_\nu\phi f^{\mu\kappa}f^{\nu}{}_{\kappa}\Big) \\ &= \int\limits_{\mathbb{R}^{D-3}} \mathrm{d}^{D-3}\vec{x} \int\limits_{-\infty}^{+\infty}\mathrm{d}u \int\limits_{-\infty}^{+\infty}\mathrm{d}v \int\limits_{0}^{2\pi}\mathrm{d}z\,|\Delta u|\,\frac{1}{2}\Bigg(2\,\partial_u\phi\,\partial_v\phi \\ &\quad - \frac{1}{(\Delta u)^2}(\partial_z\phi)^2 - \eta^{ij}\partial_i\phi\,\partial_j\phi - M^2\phi^2 + c_1(\alpha')^2\frac{1}{(\Delta u)^2}(\partial_v\phi)^2 f^2(u)\Bigg), \end{aligned} \tag{6.57}$$

Performing the same steps as before we get

$$B(u) = -\left(\left\|\vec{k}\right\|^2 + r\right) - i\,k_+\,\frac{1}{u} + \frac{\Big(c_1(\alpha')^2 f(u)^2 k_+^2 - l^2\Big)}{\Delta^2\,u^2}, \tag{6.58}$$

so even for a constant $f(u) = f_0$ we get a solution which solves the issues. Notice however that the "trivial" solution $f = f_0\,\mathrm{d}u \wedge \mathrm{d}z$ is not trivial in Minkowski coordinates where it reads $f = \frac{f_0}{x^-}\,\mathrm{d}x^- \wedge \mathrm{d}x^2$. Though appealing, the study of the string in the presence of this non trivial background needs a deeper analysis and it surely is a direction to cover in the future.

6.3 NBO Eigenfunction from the Covering Space

We recover the eigenfunctions from the covering Minkowski space in order to elucidate the connection between the polarisations in NBO and in Minkowski. Moreover we generalise the result to a symmetric two index tensor which is the polarisation of the first massive state to compute the two-tachyons–one-massive-state amplitude in the next section and to show that it diverges.

6.3.1 Spin-0 Wave Function from Minkowski space

We start with the usual plane wave in flat space and we express it in the new coordinates (we do not write the dependence on $\vec{x}$ since it is trivial):

$$\begin{aligned}\psi_{k_+\,k_-\,k_2}\big(x^+,\,x^-,\,x^2\big) &= e^{i\left(k_+x^+ + k_-x^- + k_2x^2\right)}\\ &= e^{\,i\left[k_+v + \frac{2\,k_+k_- - k_2^2}{2k_+}u + \frac{1}{2}\Delta^2 k_+ u\left(z + \frac{k_2}{\Delta k_+}\right)^2\right]}\\ &= \psi_{k_+\,k_-\,k_2}(u,\,v,\,z).\end{aligned} \tag{6.59}$$

The corresponding wave function on the NBO is obtained by the periodicity of z. This can be done in two ways either in (x^μ) coordinates or in $(x^\alpha) = (u\,v\,z)$. From the first we study how the map to the orbifold gives the function a dependence on the equivalence class of momenta. Implementing the projection on periodic z functions we get:

$$\begin{aligned}\Psi_{[k_+\,k_-\,k_2]}\big(\big[x^+,\,x^-,\,x^2\big]\big) &= \sum_{n=-\infty}^{+\infty}\psi_{k_+\,k_-\,k_2}\big(\mathcal{K}^n\big(x^+,\,x^-,\,x^2\big)\big)\\ &= \sum_{n=-\infty}^{+\infty}\psi_{\mathcal{K}^{-n}(k_+\,k_-\,k_2)}\big(x^+,\,x^-,\,x^2\big),\end{aligned} \tag{6.60}$$

where we write $[k_+\,k_-\,k_2]$ since the function depends on the equivalence class of $(k_+\,k_-\,k_2)$ only. The equivalence relation is given by

$$k = \begin{pmatrix}k_+\\ k_-\\ k_2\end{pmatrix} \equiv \mathcal{K}^{-n}k = \begin{pmatrix}k_+\\ k_- + n(2\pi\Delta)k_2 + \frac{1}{2}n^2(2\pi\Delta)^2k_+\\ k_2 + n(2\pi\Delta)k_+\end{pmatrix}. \tag{6.61}$$

It allows us to choose a representative with

$$\begin{cases}0 \le \frac{k_2}{\Delta|k_+|} < 2\pi, & k_+ \neq 0\\ 0 \le \frac{k_-}{\Delta|k_2|} < 2\pi, & k_+ = 0, \quad k_2 \neq 0\end{cases}. \tag{6.62}$$

If we perform the computation in $(u,\,v,\,z)$ coordinates we get:

$$\begin{aligned}\Psi_{[k_+\,k_-\,k_2]}(u,\,v,\,z) &= \sum_{n=-\infty}^{+\infty}\psi_{k_+\,k_-\,k_2}(u,\,v,\,z+2\pi n)\\ &= \sum_{n=-\infty}^{+\infty} e^{\,i\left[k_+v + \frac{r}{2k_+}u + \frac{1}{2}(2\pi\Delta)^2k_+u\left[n + \frac{1}{2\pi}\left(z + \frac{k_2}{\Delta k_+}\right)\right]^2\right]},\end{aligned} \tag{6.63}$$

with $r = 2\,k_+k_- - k_2^2$ and $\mathrm{Im}(k_+u) > 0$, i.e. $k_+u = |k_+u|e^{i\epsilon}$ and $\pi > \epsilon > 0$. There is no separate dependence on z and on $\frac{k_2}{\Delta k_+}$: we could fix the range $0 \leq z + \frac{k_2}{\Delta k_+} < 2\pi$. However this symmetry is broken when considering the photon eigenfunction.

We can now use the Poisson resummation

$$\sum_{n=-\infty}^{+\infty} e^{i\,a\,(n+b)^2} = \int \mathrm{d}s\ \delta_P(s) e^{i\,a\,(s+b)^2} = (2\pi)^2 \frac{e^{-i\left(\frac{\pi}{4}+\frac{1}{2}arg(a)\right)}}{2\sqrt{\pi|a|}} \sum_{m=-\infty}^{+\infty} e^{-i\,\frac{\pi^2 m^2}{a}+i\,2\pi bm}, \quad (6.64)$$

to finally get:[45]

$$\begin{aligned}\Psi_{[k_+\,k_-\,k_2\,\vec{k}]}(u,\,v,\,z,\,\vec{x}) &= \sqrt{2\pi}\ \frac{2e^{-i\,\frac{\pi}{4}}}{\Delta} \\ &\times \sum_{l=-\infty}^{+\infty} \left[\frac{1}{\sqrt{|k_+u|}} e^{\,i\left[k_+v+lz-\frac{l^2}{2\Delta^2 k_+}\,\frac{1}{u}+\frac{r+\|\vec{k}\|^2}{2k_+}u+\vec{k}\cdot\vec{x}\right]}\right] e^{i\,l\,\frac{k_2}{\Delta k_+}} \\ &= \mathcal{N} \sum_{l=-\infty}^{+\infty} \Phi_{\{k_+,\,l,\,\vec{k},\,r\}}(u,\,v,\,z,\,\vec{x}) e^{i\,l\,\frac{k_2}{\Delta k_+}},\end{aligned} \quad (6.65)$$

when $k_+ \neq 0$ and where

$$\mathcal{N} = \sqrt{\frac{(2\pi)^D}{\pi\Delta}}\,\frac{e^{-i\,\frac{\pi}{4}}}{\pi}. \quad (6.66)$$

The fact that Ψ depends only on the equivalence class $[k_+\,k_-\,k_2\,k]$ allows us to restrict $0 \leq \frac{k_2}{\Delta\,|k_+|} < 2\pi$ so that we can invert the previous expression and get:

$$\Phi_{\{k_+,\,l,\,\vec{k},\,r\}}(u,\,v,\,z,\,\vec{x}) = \frac{1}{\mathcal{N}}\,\frac{1}{2\pi\Delta|k_+|} \int_0^{2\pi\Delta|k_+|} \mathrm{d}k_2\, e^{-i\,l\,\frac{k_2}{\Delta k_+}}\,\Psi_{[k_+\,k_-\,k_2\,k]}(u,\,v,\,z,\,\vec{x}). \quad (6.67)$$

6.3.2 Spin-1 Wave Function from Minkowski space

We go through the steps in the previous case for an electromagnetic wave. We concentrate on x^+, x^- and x^2 coordinates and reinstate $\vec{x}$ at the end. We start with the usual plane wave in flat space $\psi^{[1]}_{k_+\,k_-\,k_2;\,\epsilon_+\,\epsilon_-\,\epsilon_2}$ and we express it in both Minkowski and orbifold coordinates. We use the notation $\psi^{[1]}_{k_+\,k_-\,k_2;\,\epsilon_+\,\epsilon_-\,\epsilon_2}$ to stress that it is the eigenfunction and not the field which is obtained as

$$A_\mu(x)\,\mathrm{d}x^\mu = \int_{\mathbb{R}^3} \mathrm{d}^3k \sum_{\{\epsilon_+,\,\epsilon_-,\,\epsilon_2\}} \psi^{[1]}_{k_+\,k_-\,k_2;\,\epsilon_+\,\epsilon_-\,\epsilon_2}, \quad (6.68)$$

[45] In the expression we insert the variables $\vec{k}$ and $\vec{x}$ for completeness. We also set $r = 2\,k_+k_- - k_2^2 - \left\|\vec{k}\right\|^2$.

where the sum is performed over ϵ_+, ϵ_-, ϵ_2 independent and compatible with k. The explicit expression for the eigenfunction with constant ϵ_+, ϵ_- and ϵ_2 is:[46]

$$\begin{aligned}\mathcal{N}\,\psi^{[1]}_{k_+\,k_-\,k_2;\,\epsilon_+\,\epsilon_-\,\epsilon_2}\big(x^+,\,x^-,\,x^2\big) &= \big(\epsilon_+\,\mathrm{d}x^+ + \epsilon_-\,\mathrm{d}x^- + \epsilon_2\,\mathrm{d}x^2\big)\,e^{i\left(k_+x^+ + k_-x^- + k_2x^2\right)}\\ &= (\epsilon_u\,\mathrm{d}u + \epsilon_z\,\mathrm{d}z + \epsilon_v\,\mathrm{d}v)\\ &\times e^{i\left[k_+v + \frac{2\,k_+k_- - k_2^2}{2k_+}u + \frac{1}{2}\Delta^2 k_+ u\left(z + \frac{k_2}{\Delta k_+}\right)^2\right]}\\ &= \mathcal{N}\,\psi^{[1]}_{k_+\,k_-\,k_2;\,\epsilon_+\,\epsilon_-\,\epsilon_2}(u,\,v,\,z),\end{aligned} \tag{6.69}$$

with

$$\begin{aligned}\epsilon_v &= \epsilon_+,\\ \epsilon_u(z) &= \epsilon_- + (\Delta z)\,\epsilon_2 + (\frac{1}{2}\Delta^2 z^2)\,\epsilon_+,\\ \epsilon_z(u,\,z) &= (\Delta u)\,(\epsilon_2 + \Delta z\,\epsilon_+).\end{aligned} \tag{6.70}$$

Notice that we are not imposing any gauge condition. Moreover if $(\epsilon_+,\,\epsilon_-,\,\epsilon_2)$ are constant then $(\epsilon_u,\,\epsilon_v,\,\epsilon_z)$ are generic functions. It is worth stressing that they are not the polarisations in the orbifold which are in any case constant: the fact that they depend on the coordinates is simply the statement that not all eigenfunctions of the vector d'Alembertian are equal.

Building the corresponding function on the orbifold amounts to summing the images created by the orbifold group:

$$\mathcal{N}\,\Psi^{[1]}_{[k,\,\epsilon]}([x]) = \sum_{n=-\infty}^{+\infty} \vec{\epsilon}\cdot(\mathcal{K}^n\,\mathrm{d}x)\;\psi_k(\mathcal{K}^n x) = \sum_{n=-\infty}^{+\infty} \mathcal{K}^{-n}\,\vec{\epsilon}\cdot\mathrm{d}x\;\;\psi_{\mathcal{K}^{-n}k}(x). \tag{6.71}$$

Under the action of the Killing vector ϵ transforms exactly as the k since it is induced by $\epsilon\cdot\mathcal{K}^n\,\mathrm{d}x = \mathcal{K}^{-n}\epsilon\cdot\mathrm{d}x$, that is:

$$\epsilon = \begin{pmatrix}\epsilon_+\\ \epsilon_2\\ \epsilon_-\end{pmatrix} \equiv \mathcal{K}^{-n}\epsilon = \begin{pmatrix}\epsilon_+\\ \epsilon_2 + n\,(2\pi\Delta)\,\epsilon_+\\ \epsilon_- + n\,(2\pi\Delta)\,\epsilon_2 + \frac{1}{2}n^2(2\pi\Delta)^2\epsilon_+\end{pmatrix}. \tag{6.72}$$

However the pair $\left(\vec{k},\,\vec{\epsilon}\right)$ transforms with the same n since both are "dual" to x, i.e. their transformation rules are dictated by x. There is therefore only one equivalence class $\left[\vec{k},\,\vec{\epsilon}\right]$ and not two separate classes $\left[\vec{k}\right]$, $[\vec{\epsilon}]$. In other words, a representative of the combined equivalence class is the one with $0 \le k_2 < 2\pi\Delta|k_+|$ when $k_+ \neq 0$.

In order to write the eigenfunctions on the orbifold in orbifold coordinates we notice that

[46] We introduce the normalisation factor $\mathcal{N}$ in order to have a less cluttered relation between ϵ and $\mathcal{E}$.

$\mathrm{d}u\,,\mathrm{d}v$ and $\mathrm{d}z$ are invariant. We write

$$\begin{aligned}
\mathcal{N}\Psi^{[1]}_{\left[\vec{k},\,\vec{\epsilon}\right]}([x]) &= \sum_{n=-\infty}^{+\infty} \epsilon\cdot(\mathcal{K}^n\,\mathrm{d}x)\,\psi_k(\mathcal{K}^n x)\\
&= \mathrm{d}v\,\left[\epsilon_+\sum_{n=-\infty}^{+\infty}\psi_k(\mathcal{K}^n x)\right]\\
&+ \mathrm{d}z\;(\Delta u)\left[\epsilon_2\sum_{n=-\infty}^{+\infty}\psi_k(\mathcal{K}^n x)+\epsilon_+\,\Delta\sum_{n=-\infty}^{+\infty}(z+2\pi n)\psi_k(\mathcal{K}^n x)\right]\\
&+ \mathrm{d}u\,\left[\epsilon_-\sum_{n=-\infty}^{+\infty}\psi_k(\mathcal{K}^n x)+\epsilon_2\,\Delta\sum_{n=-\infty}^{+\infty}(z+2\pi n)\psi_k(\mathcal{K}^n x)\right.\\
&\left.+\frac{1}{2}\epsilon_+\,\Delta^2\sum_{n=-\infty}^{+\infty}(z+2\pi n)^2\psi_k(\mathcal{K}^n x)\right].
\end{aligned}\tag{6.73}$$

From a direct computation we get:[47]

$$\begin{aligned}
\sum_{n=-\infty}^{+\infty}(z+2\pi n)\,\psi_k(\mathcal{K}^n x) &= \left(\frac{1}{i\,\Delta u}\frac{\partial}{\partial k_2}-\frac{k_2}{\Delta k_+}\right)\Psi_{[k]}([x])\\
\sum_{n=-\infty}^{+\infty}(z+2\pi n)^2\,\psi_k(\mathcal{K}^n x) &= \left(\frac{1}{i\,\Delta\,u}\frac{\partial}{\partial k_2}-\frac{k_2}{\Delta k_+}\right)^2\Psi_{[k]}([x]).
\end{aligned}\tag{6.74}$$

Then it follows that

$$\begin{aligned}
\mathcal{N}\Psi^{[1]}_{\left[\vec{k},\,\vec{\epsilon}\right]}([x]) &= \mathrm{d}v\;\left[\epsilon_+\,\Psi_{[k]}([x])\right]\\
&+ \mathrm{d}z\;(\Delta u)\left[\frac{\epsilon_2 k_+-\epsilon_+ k_2}{k_+}\,\Psi_{[k]}([x])-\epsilon_+\,\frac{i}{u}\frac{\partial}{\partial k_2}\Psi_{[k]}([x])\right]\\
&+ \mathrm{d}u\;\left[\left(\epsilon_- -\epsilon_2\frac{k_2}{k_+}+\frac{1}{2}\epsilon_+\left(\frac{k_2}{k_+}\right)^2\right)\Psi_{[k]}([x])+\frac{i}{2u}\frac{\epsilon_+}{k_+}\,\Psi_{[k]}([x])\right.\\
&\left.-\frac{\epsilon_2 k_+-\epsilon_+ k_2}{k_+}\,\frac{i}{u}\frac{\partial}{\partial k_2}\Psi_{[k]}([x])-\frac{1}{2}\epsilon_+\,\frac{1}{u^2}\,\frac{\partial^2}{\partial k_2^2}\Psi_{[k]}([x])\right].
\end{aligned}\tag{6.75}$$

Many coefficients of Ψ or its derivatives contain k_2. They cannot be expressed using the quantum numbers $\left\{k_+,\,l,\,\vec{k},\,r\right\}$ of the orbifold but are invariant on it. They are new orbifold quantities

[47] These expressions may be written using Hermite polynomials.

we interpret as orbifold polarisations. Using (6.65) we can finally write

$$\begin{aligned}\Psi^{[1]}_{[k,\,\epsilon]}([x]) = \sum_{l=-\infty}^{+\infty} \phi_{\{k_+,l,\vec{k},r\}}(u,\,v,\,z,\,\vec{x}) e^{i\,l\,\frac{k_2}{\Delta k_+}} \\ \times \Bigg\{ \mathrm{d}v\ \epsilon_+ \\ + \mathrm{d}z\ (\Delta u)\left[\frac{\epsilon_2 k_+ - \epsilon_+ k_2}{k_+} + \epsilon_+ \frac{1}{\Delta u}\,\frac{l}{k_+}\right] \\ + \mathrm{d}u\ \Bigg[\left(\epsilon_- - \epsilon_2\,\frac{k_2}{k_+} + \frac{1}{2}\epsilon_+\left(\frac{k_2}{k_+}\right)^2\right) + \frac{i}{2u}\,\frac{\epsilon_+}{k_+} \\ + \frac{\epsilon_2 k_+ - \epsilon_+ k_2}{k_+}\frac{1}{u}\,\frac{l}{\Delta k_+} + \epsilon_+\,\frac{1}{2u^2}\left(\frac{l}{\Delta k_+}\right)^2\Bigg]\Bigg\}.\end{aligned} \tag{6.76}$$

If we compare the last expression with (6.28) we find:

$$\begin{aligned}\mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{v}} &= \epsilon_+ \\ \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{z}} &= \mathrm{sign}(u)\frac{\epsilon_2 k_+ - \epsilon_+ k_2}{k_+} \\ \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{u}} &= \epsilon_- - \epsilon_2\,\frac{k_2}{k_+} + \frac{1}{2}\epsilon_+\left(\frac{k_2}{k_+}\right)^2,\end{aligned} \tag{6.77}$$

which implies that the true polarisations $(\epsilon_+,\,\epsilon_-,\,\epsilon_2)$ and $\mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{*}}$ are constant as it turns out from direct computation. A different way of reading the previous result is that the polarisations on the orbifold are the coefficients of the highest power of u.

We can also invert the previous relations to get:

$$\begin{aligned}\epsilon_+ &= \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{v}} \\ \epsilon_2 &= \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{z}}\,\mathrm{sign}(u) + \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{v}}\,\frac{k_2}{k_+} \\ \epsilon_- &= \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{u}} + \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{z}}\,\mathrm{sign}(u)\,\frac{k_2}{k_+} + \mathcal{E}_{\{k_+,l,\vec{k},r\}\,\underline{v}}\,\frac{1}{2}\left(\frac{k_2}{k_+}\right)^2,\end{aligned} \tag{6.78}$$

and use them in Lorenz gauge $\vec{k}\cdot\vec{\epsilon} = 0$ in order to get the gauge conditions expressed with the orbifold polarisations. If the definition of orbifold polarisations is right the result cannot depend on k_2 since it is not a quantum number of orbifold eigenfunctions. Taking into account $k_- = \frac{\|\vec{k}\|^2 + k_2^2 + r}{2k_+}$ in $\vec{k}\cdot\vec{\epsilon} = 0$ we get exactly the expression for the Lorenz gauge for orbifold polarisations (6.26).

6.3.3 Tensor Wave Function (Spin-2) from Minkowski space

We can use the analysis of the previous section in the case of a second order symmetric tensor wave function. Again we suppress the dependence on $\vec{x}$ and $\vec{k}$ with a caveat: the Minkowskian

polarisations $S_{+\,i}$, $S_{-\,i}$ and $S_{2\,i}$ transform non trivially, therefore we give the full expressions in Appendix D even if these components behave effectively as a vector of the orbifold.

We start with the usual wave in flat space and we express either in the Minkowskian coordinates

$$\begin{aligned}\mathcal{N}\,\psi^{[2]}_{k\,S}(x^+,\,x^-,\,x^2) &= S_{\mu\nu}\,\psi_k(x)\;\mathrm{d}x^\mu\;\mathrm{d}x^\nu\\ &= \Big(S_{++}\;\mathrm{d}x^+\;\mathrm{d}x^+ + 2\,S_{+\,x}\;\mathrm{d}x^+\;\mathrm{d}x^2 + 2\,S_{+\,-}\;\mathrm{d}x^+\;\mathrm{d}x^-\\ &\quad + 2\,S_{2\,2}\;\mathrm{d}x^2\;\mathrm{d}x^2 + 2\,S_{2\,-}\;\mathrm{d}x^2\;\mathrm{d}x^-\\ &\quad + 2\,S_{-\,-}\;\mathrm{d}x^-\;\mathrm{d}x^-\Big)e^{i\left(k_+x^+ + k_-x^- + k_2x^2\right)},\end{aligned} \tag{6.79}$$

or in orbifold coordinates

$$\begin{aligned}\mathcal{N}\,\psi^{[2]}_{k\,S}(x) &= S_{\alpha\,\beta}\,\psi_k(x)\;\mathrm{d}x^\alpha\;\mathrm{d}x^\beta\\ &= \Bigg\{\mathrm{d}v^2\;S_{+\,+}\\ &\quad + \mathrm{d}v\;\mathrm{d}z\;\Delta u[2\,S_{+\,2} + S_{+\,+}\Delta z]\\ &\quad + \mathrm{d}v\;\mathrm{d}u\;\left[2\,S_{+\,-} + 2\,S_{+\,2}\Delta z + S_{+\,+}\Delta^2 z^2\right]\\ &\quad + \mathrm{d}z^2\;\Delta^2u^2\left[S_{2\,2} + 2\,S_{+\,2}\Delta z + S_{+\,+}\Delta^2z^2\right]\\ &\quad + \mathrm{d}z\;\mathrm{d}v\;\Delta u\left[2\,S_{-\,2} + 2\,(S_{2\,2} + S_{+\,-})\,\Delta z + 3\,S_{+\,2}\Delta^2z^2 + S_{+\,+}\,\Delta^3z^3\right]\\ &\quad + \mathrm{d}u^2\;\left[S_{-\,-} + 2\,S_{-\,2}\Delta z + (S_{2\,2} + S_{+\,-})\,\Delta^2z^2 + S_{+\,2}\Delta^3z^3 + \frac{1}{4}S_{+\,+}\Delta^4z^4\right]\Bigg\}\\ &\quad \times e^{i\left[k_+v + \frac{2\,k_+k_- - k_2^2}{2k_+}u + \frac{1}{2}\Delta^2k_+u\left(z + \frac{k_2}{\Delta k_+}\right)^2\right]}.\end{aligned} \tag{6.80}$$

Now we define the tensor on the orbifold as a sum over all images as

$$\begin{aligned}\mathcal{N}\,\Psi^{[2]}_{[k\,S]}([x]) &= \sum_{n=-\infty}^{+\infty}(\mathcal{K}^n\,\mathrm{d}x)\cdot S\cdot(\mathcal{K}^n\,\mathrm{d}x)\;\psi_k(\mathcal{K}^n x)\\ &= \sum_{n=-\infty}^{+\infty}\mathrm{d}x\cdot(\mathcal{K}^{-n}\,S)\cdot\mathrm{d}x\;\;\psi_{\mathcal{K}^{-n}k}(x).\end{aligned} \tag{6.81}$$

In the last line we have defined the induced action of the Killing vector on $\left(\vec{k}, S\right)$ which can be explicitely written as:

$$\mathcal{K}^{-n}\begin{pmatrix}S_{+\,+}\\ S_{+\,2}\\ S_{+\,-}\\ S_{2\,2}\\ S_{2\,-}\\ S_{-\,-}\end{pmatrix} = \begin{pmatrix}S_{+\,+}\\ S_{+\,2} + n\Delta S_{+\,+}\\ S_{+\,-} + n\Delta S_{+\,2} + \frac{1}{2}n^2\Delta^2S_{+\,+}\\ S_{2\,2} + 2n\Delta S_{+\,2} + n^2\Delta^2S_{+\,+}\\ S_{2\,-} + n\Delta(S_{2\,2} + S_{+\,-}) + \frac{3}{2}n^2\Delta^2S_{+\,2} + \frac{1}{2}n^3\Delta^3S_{+\,+}\\ S_{-\,-} + 2n\Delta S_{-\,2} + n^2\Delta^2(S_{2\,2} + S_{+\,-}) + n^3\Delta^3S_{+\,2} + \frac{1}{4}n^4\Delta^4S_{+\,+}\end{pmatrix}. \tag{6.82}$$

Computing the tensor on the orbifold in its own coordinates is equivalent to summing over all the shifts $z \to (z + 2\pi n)$ and the use of a generalisation of (6.74), i.e. to substitute $(\Delta\,z)^j\psi_k \to$

$\left(\frac{1}{i\,u}\frac{\partial}{\partial k_2} - \frac{k_2}{\Delta\,k_+}\right)^j \Psi_{[k]}([x])$. When expressing all in the ϕ basis, the last step is equivalent to $(\Delta\,z)^j \psi_k \to \left(\frac{l}{\Delta\,u\,k_+}\right)^j + \ldots$. We identify the basic polarisations on the orbifold by considering the highest power in u:

$$\begin{aligned}
\mathcal{S}_{u\,u} &= \frac{1}{4}K^4\,S_{+\,+} + K^2\,S_{+\,-} - K^3\,S_{+\,2} + S_{-\,-} - 2\,K\,S_{-\,2} + S_{2\,2}\,K^2 \\
\mathcal{S}_{u\,v} &= \frac{1}{2}K^2\,S_{+\,+} + S_{+\,-} - K\,S_{+\,2} \\
\mathcal{S}_{u\,z} &= -\frac{1}{2}K^3\,S_{+\,+} - K\,S_{+\,-} + \frac{3}{2}K^2\,S_{+\,2} + S_{-\,2} - K\,S_{2\,2} \\
\mathcal{S}_{v\,v} &= S_{+\,+} \\
\mathcal{S}_{v\,z} &= S_{+\,2} - K\,S_{+\,+} \\
\mathcal{S}_{z\,z} &= K^2\,S_{+\,+} - 2\,K\,S_{+\,2} + S_{2\,2}.
\end{aligned} \tag{6.83}$$

where $K = \frac{k_2}{k_+}$. The previous equations can be inverted to get:

$$\begin{aligned}
S_{-\,-} &= K^2\,(\mathcal{S}_{z\,z} + \mathcal{S}_{u\,v}) + K^3\,\mathcal{S}_{v\,z} + \frac{1}{4}K^4\,\mathcal{S}_{v\,v} + 2\,K\,\mathcal{S}_{u\,z} + \mathcal{S}_{u\,u} \\
S_{+\,-} &= K\,\mathcal{S}_{v\,z} + \frac{1}{2}K^2\,\mathcal{S}_{v\,v} + \mathcal{S}_{u\,v} \\
S_{-\,2} &= K\,(\mathcal{S}_{z\,z} + \mathcal{S}_{u\,v}) + \frac{3}{2}\,K^2\,\mathcal{S}_{v\,z} + \frac{1}{2}K^3\,\mathcal{S}_{v\,v} + \mathcal{S}_{u\,z} \\
S_{+\,+} &= \mathcal{S}_{v\,v} \\
S_{+\,2} &= \mathcal{S}_{v\,z} + K\,\mathcal{S}_{v\,v} \\
S_{2\,2} &= \mathcal{S}_{z\,z} + 2\,K\,\mathcal{S}_{v\,z} + K^2\,\mathcal{S}_{v\,v}.
\end{aligned} \tag{6.84}$$

Since we plan to use the previous quantities in the case of the first massive string state we compute the relevant quantities. In particular we have the trace:

$$\mathrm{tr}(S) = \mathcal{S}_{z\,z} - 2\,\mathcal{S}_{u\,v} \tag{6.85}$$

and the transversality conditions

$$\begin{aligned}
\text{trans}\ \mathcal{S}_v &= \left(\vec{k}\cdot S\right)_+ = -\frac{\left(r + \left\|\vec{k}\right\|^2\right)}{2\,k_+}\,\mathcal{S}_{v\,v} - k_+\,\mathcal{S}_{u\,v}, \\
\text{trans}\ \mathcal{S}_z &= \left(\vec{k}\cdot S\right)_2 - K\,\left(\vec{k}\cdot S\right)_+ = -\frac{\left(r + \left\|\vec{k}\right\|^2\right)}{2\,k_+}\,\mathcal{S}_{v\,z} - k_+\,\mathcal{S}_{u\,z}, \\
\text{trans}\ \mathcal{S}_u &= \left(\vec{k}\cdot S\right)_- - K\,\left(\vec{k}\cdot S\right)_2 + \frac{1}{2}K^2\left(\vec{k}\cdot S\right)_+ = -\frac{\left(r + \left\|\vec{k}\right\|^2\right)}{2\,k_+}\,\mathcal{S}_{u\,v} - k_+\,\mathcal{S}_{u\,u}
\end{aligned} \tag{6.86}$$

where we used $k_- = \frac{\left(r+\left\|\vec{k}\right\|^2+k_2^2\right)}{(2k_+)}$. These conditions do not depend on K since k_2 is not an orbifold quantum number.

The final expression for the orbifold symmetric tensor is

$$
\begin{aligned}
\Psi^{[2]}_{\left[\vec{k},\,S\right]}([x]) = \sum_{l=-\infty}^{+\infty} & \Phi_{\{k_+,\,l,\,\vec{k},\,r\}}(u,\,v,\,z,\,\vec{x}) e^{i\,l\,\frac{k_2}{\Delta k_+}} \\
& \times \Bigg\{ \mathrm{d}v^2\ \mathcal{S}_{v\,v} \\
& + 2\,\Delta\,u\ \mathrm{d}v\ \mathrm{d}z\ \left[\mathcal{S}_{v\,z} + \left(\frac{L\mathcal{S}_{v\,v}}{\Delta}\right)\frac{1}{u}\right] \\
& + 2\ \mathrm{d}v\ \mathrm{d}u\ \left[\mathcal{S}_{u\,v} + \left(\frac{L\,\mathcal{S}_{v\,z}}{\Delta} + \frac{i\,\mathcal{S}_{v\,v}}{2\,k_+}\right)\frac{1}{u} + \left(\frac{L^2\,\mathcal{S}_{v\,v}}{2\,\Delta^2}\right)\frac{1}{u^2}\right] \\
& + (\Delta\,u)^2\,\mathrm{d}z^2\ \left[\mathcal{S}_{z\,z} + \left(\frac{2\,L\,\mathcal{S}_{v\,z}}{\Delta} + \frac{i\,\mathcal{S}_{v\,v}}{k_+}\right)\frac{1}{u} + \left(\frac{L^2\,\mathcal{S}_{v\,v}}{\Delta^2}\right)\frac{1}{u^2}\right] \\
& + 2\,\Delta\,u\ \mathrm{d}z\ \mathrm{d}u\ \Big[\mathcal{S}_{u\,z} + \left(\frac{L\,\mathcal{S}_{z\,z}}{\Delta} + \frac{3\,i\,\mathcal{S}_{v\,z}}{2\,k_+} + \frac{L\,\mathcal{S}_{u\,v}}{\Delta}\right)\frac{1}{u} + \left(\frac{3\,L^2\,\mathcal{S}_{v\,z}}{2\,\Delta^2} + \frac{3\,i\,L\,\mathcal{S}_{v\,v}}{2\,\Delta\,k_+}\right)\frac{1}{u^2} \\
& + \left(\frac{L^3\,\mathcal{S}_{v\,v}}{2\,\Delta^3}\right)\frac{1}{u^3}\Big] \\
& + \mathrm{d}u^2\ \Big[\mathcal{S}_{u\,u} + \left(\frac{i\,\mathcal{S}_{z\,z}}{k_+} + \frac{2\,L\,\mathcal{S}_{u\,z}}{\Delta} + \frac{i\,\mathcal{S}_{u\,v}}{k_+}\right)\frac{1}{u} \\
& + \left(\frac{L^2\,\mathcal{S}_{z\,z}}{\Delta^2} + \frac{3\,i\,L\,\mathcal{S}_{v\,z}}{\Delta\,k_+} - \frac{3\,\mathcal{S}_{v\,v}}{4\,k_+^2} + \frac{L^2\,\mathcal{S}_{u\,v}}{\Delta^2}\right)\frac{1}{u^2} \\
& + \left(\frac{L^3\,\mathcal{S}_{v\,z}}{\Delta^3} + \frac{3\,i\,L^2\,\mathcal{S}_{v\,v}}{2\,\Delta^2\,k_+}\right)\frac{1}{u^3} + \left(\frac{L^4\mathcal{S}_{v\,v}}{4\Delta^4}\right)\frac{1}{u^4}\Big]\Bigg\},
\end{aligned}
\tag{6.87}
$$

where $L = \frac{l}{k_+}$.

6.4 Overlaps of Wave Functions and Their Derivatives

In this section we compute overlaps of wave functions. We give their expressions using both integrals over the eigenfunctions and sums of products of delta functions. The latter is the expression which is naturally obtained by computing tree level string amplitudes on the orbifold when one starts with Minkowski amplitudes and adds the images. This is equivalent to computing emission vertices on the orbifold and then their correlation functions since this amounts to transfer the sum over the spacetime images to the sum of the polarisations images.

6.4.1 Overlaps Without Derivatives

We start from the simplest case of the overlap of N scalar wave functions. We compute the overlap of orbifold wave functions and then we express it as sum of images of the corresponding Minkowski overlap thus establishing a dictionary between Minkowski and orbifold spaces. Explicitly we

consider the following overlap where all the polarisations $\mathcal{A}_{(i)}$ have been set to one

$$\begin{aligned}
I^{(N)} &= \int_{\Omega} \mathrm{d}^3x\ \sqrt{-\det g}\ \prod_{i=1}^{N} \Psi_{\left[k_{(i)\,+}\ k_{(i)\,-}\ k_{(i)\,2}\right]}\left(\left[x^+,\, x^-,\, x^2\right]\right)\\
&= \int_{\mathscr{M}^{1,2}} \mathrm{d}^3x\ \sqrt{-\det g}\ \psi_{k_{(1)\,+}\ k_{(1)\,-}\ k_{(1)\,2}}\left(x^+,\, x^-,\, x^2\right)\\
&\times \prod_{i=2}^{N} \sum_{m_{(i)}=-\infty}^{+\infty} \psi_{k_{(i)\,+}\ k_{(i)\,-}\ k_{(i)\,2}}(\mathcal{K}^{m_{(i)}}\left(x^+,\, x^-,\, x^2\right))\\
&= \int_{\mathscr{M}^{1,2}} \mathrm{d}^3x\ \sqrt{-\det g}\ \psi_{k_{(1)\,+}\ k_{(1)\,-}\ k_{(1)\,2}}\left(x^+,\, x^-,\, x^2\right))\\
&\times \prod_{i=2}^{N} \sum_{m_{(i)}=-\infty}^{+\infty} \psi_{\mathcal{K}^{m_{(i)}}\left(k_{(i)\,+}\ k_{(i)\,-}\ k_{(i)\,2}\right)}\left(x^+,\, x^-,\, x^2\right)\\
&= (2\pi)^3 \delta\left(\sum_{i=-\infty}^{+\infty} k_{(i)\,+}\right)\\
&\times \prod_{i=2}^{N} \sum_{m_{(i)}=-\infty}^{+\infty} \delta\left(\sum_{i=-\infty}^{+\infty} \mathcal{K}^{m_{(i)}}\ k_{(i)\,2}\right) \delta\left(\sum_{i=-\infty}^{+\infty} \mathcal{K}^{m_{(i)}}\ k_{(i)\,-}\right)\Bigg|_{m_{(1)}=0},
\end{aligned} \tag{6.88}$$

where $\Omega = \mathscr{M}^{1,2}/\Gamma$ is the fundamental region identifying the orbifold. We used the unfolding trick to rewrite the integral as an integral over $\mathscr{M}^{1,2}$ thus dropping the sum over the images of particle (1). We then moved the action of the Killing vector from x to k and finally we used the usual δ definition. The previous integral can be expressed as:

$$\begin{aligned}
I^{(N)} &= \mathcal{N}^N \sum_{\{l_{(i)}\}\in\mathbb{Z}^N} e^{\,i\sum_{i=1}^{N} l_{(i)} \frac{k_{(i)\,2}}{\Delta k_{(i)\,+}}} \int_{\Omega} \mathrm{d}^3x\ \sqrt{-\det g} \prod_{i=1}^{N} \phi_{k_- r N i}([x])\\
&= \mathcal{N}^N \sum_{\{l_{(i)}\}\in\mathbb{Z}^N} e^{\,i\sum_{i=1}^{N} l_{(i)} \frac{k_{(i)\,2}}{\Delta k_{(i)\,+}}} (2\pi)^2 \delta\left(\sum_{i=1}^{N} k_{(i)\,+}\right) \delta_{\sum_{i=1}^{N} l_{(i)},\,0}\ \mathcal{I}^{[0]}_{\{N\}}.
\end{aligned} \tag{6.89}$$

From the last expression we recover the overlap of the wave functions as:

$$\begin{aligned}
&\int_{\Omega} \mathrm{d}^3x \prod_{i=1}^{N} \Phi_{k_- r N i}([x]) \\
=&\frac{1}{\mathcal{N}^N} \prod_{i=1}^{N} \int_0^{2\pi\Delta|k_{(i)+}|} \frac{\mathrm{d}k_{(i)\,2}}{2\pi\Delta|k_{(i)+}|}\, e^{-i\, l_{(i)} \frac{k_{(i)\,2}}{\Delta k_{(i)\,i}}}\, I^{(N)} \\
=&(2\pi)^3\, \delta\left(\sum_{i=-\infty}^{+\infty} k_{(i)+}\right) \frac{1}{\mathcal{N}^N} \prod_{i=1}^{N} \int_0^{2\pi\Delta|k_{(i)+}|} \frac{\mathrm{d}k_{(i)\,2}}{2\pi\Delta|k_{(i)+}|}\, e^{-i\, l_{(i)} \frac{k_{(i)\,2}}{\Delta k_{(i)+}}} \\
&\times \prod_{j=2}^{N} \sum_{m_{(j)}=-\infty}^{+\infty} \delta\left(\sum_{j=2}^{N} \mathcal{K}^{m_{(j)}} k_{(j)\,2}\right) \delta\left(\sum_{j=2}^{N} \mathcal{K}^{m_{(j)}} k_{(j)\,-}\right).
\end{aligned} \tag{6.90}$$

It follows from the explicit expression of $\mathcal{I}^{[0]}_{\{n\}}$ that all overlaps $I^{(N)}$ for $N \geq 4$ diverge.

Intuitively we are in fact summing over infinite distributions with accumulation points of their support. Nevertheless the existence of the accumulation point is not sufficient since the three scalars overlap, i.e. the three tachyons amplitude, converges.

6.4.2 An Overlap With One Derivative

Since we will also compute the amplitude involving two tachyons and one photon, as a preliminary step we consider the overlap in Minkowski space:

$$J_{Mink} = i\,(2\pi)^3 \left(\epsilon_{(1)} \cdot k_{(2)\,2}\right) \delta\left(\sum_{i=-\infty}^{+\infty} k_{(i)+}\right) \delta\left(\sum_{i=-\infty}^{+\infty} k_{(i)\,2}\right) \delta\left(\sum_{i=-\infty}^{+\infty} k_{(i)-}\right). \tag{6.91}$$

Summing over momenta and polarisations we then get to an expression which depends on equivalence classes as:

$$\begin{aligned}
J\big(\big[k_{(1)}, \epsilon_{(1)}\big],\, \big[k_{(2)}\big],\, \big[k_{(3)}\big]\big) = i\,(2\pi)^3 \delta&\left(\sum_{i=-\infty}^{+\infty} k_{(i)+}\right) \\
&\times \sum_{\{m_{(i)}\} \in \mathbb{Z}^3} \delta_{m_{(1)},\,1} \left(\mathcal{K}^{m_{(1)}} \epsilon_{(1)} \cdot \mathcal{K}^{m_{(2)}} k_{(2)\,2}\right) \\
&\times \delta\left(\sum_{i=-\infty}^{+\infty} \mathcal{K}^{m_{(i)}} k_{(i)\,2}\right) \delta\left(\sum_{i=-\infty}^{+\infty} \mathcal{K}^{m_{(i)}} k_{(i)\,-}\right).
\end{aligned} \tag{6.92}$$

The expression depends only on equivalence classes.[48]

The previous expression can be written as

$$J = \int_{\Omega} \mathrm{d}^3x\, \eta^{\mu\nu}\, \Psi^{[1]}_{\left[k_{(1)}, \epsilon_{(1)}\right]\,\mu}([x])\, \partial_\nu \Psi_{\left[k_{(2)}\right]}([x])\, \Psi_{\left[k_{(3)}\right]}([x]) \tag{6.93}$$

[48] In order to prove it, under $\left(k_{(1)},\, \epsilon_{(1)}\right) \to \mathcal{K}^s\left(k_{(1)},\, \epsilon_{(1)}\right)$ we can use $\mathcal{K}^s a \cdot b = a \cdot \mathcal{K}^{-s} b$ and the invariance of deltas $\delta^3(\mathcal{K}^s a) = \delta^3(a)$.

where we performed the unfolding using $a_{[k_{(1)},\,\epsilon_{(1)}]\,\mu}([x])$.[49] Notice that the previous expression is invariant despite the fact that the derivatives ∂_μ are not well defined on the orbifold. The fact that $\Psi^{[1]}_\mu$ is not invariant in turns helps in recovering the required invariance.

We can then evaluate the previous expression with Minkowskian polarisations using (6.76) which is nothing else but a rearrangement of terms of (6.92). We have:

$$\begin{aligned} J = i\,\mathcal{N}^2 \sum_{\{l_{(i)}\}\in\mathbb{Z}^3} & e^{\,i\sum_{i=1}^{3} l_{(i)}\frac{k_{(i)\,2}}{\Delta k_{(i)\,+}}}\,(2\pi)^2\,\delta\left(\sum_{i=1}^{3} k_{(i)\,+}\right)\delta_{\sum_{i=1}^{3} l_{(i)}} \\ & \times \int_\Omega \mathrm{d}^3x \prod_{i=1}^{3}\phi_{k_- r N i}([x]))\Bigg\{\epsilon_{(1)\,+}\left[\frac{i}{2u}+\frac{l_{(2)}^2}{k_{(2)\,+}}\frac{1}{2\Delta^2\,u^2}+\frac{r_{(i)}}{2k_{(2)\,+}}\right] \\ & + \frac{1}{\Delta\,u}\left[\epsilon_{(1)\,2}\,+\frac{1}{\Delta u}\epsilon_{(1)\,+}\frac{l_{(1)}}{k_{(1)\,+}}\right] l_{(2)} \\ & + \left[\epsilon_{(1)\,-}+\epsilon_{(1)\,2}\frac{1}{\Delta u}\frac{l_{(1)}}{k_{(1)\,+}}+\epsilon_{(1)\,+}\frac{1}{2(\Delta u)^2}\frac{l_{(1)}^2}{k_{(1)\,+}^2}\right] k_{(2)\,+}\Bigg\}. \end{aligned} \tag{6.94}$$

Divergences occur when $l=0$ because of the absence of the factor $e^{i\frac{A}{u}}$. However all explicit factors $\frac{1}{u}$ come always with l: when $l=0$ they do not give any contribution. The divergence in this case comes actually only from the contribution of the first line $\partial_u\phi\big|_{l=0} = -\frac{1}{2u}\,\phi\big|_{l=0}$. Since we still have to subtract the contribution of the exchange $(2)\leftrightarrow(3)$ the contribution is cancelled in scalar QED or with Abelian tachyons. It does not cancel when considering the non Abelian case and the related colour factors unless one uses a kind of principal part regularisation since replacing $\int_{-|a|}^{|b|}\mathrm{d}u\,\frac{\mathrm{sign}(u)}{|u|^{\frac{3}{2}}}$ with $\lim_{\delta\to 0}\left[\int_{-|a|}^{-|\delta|}\mathrm{d}u+\int_{-|\delta|}^{|b|}\mathrm{d}u\right]\frac{\mathrm{sign}(u)}{|u|^{\frac{3}{2}}}$ gives a finite result.

6.4.3 An Overlap With Two Derivatives

We can generalise the previous expressions to more general cases. Since we use the results from Section 6.3 we miss some non trivial contributions from polarisations like $\mathcal{S}_{v\,i}$. These contributions do not alter the final result. However for completeness we give the lengthy full expression in Appendix E.

We consider:[50]

$$K = \int_\Omega \mathrm{d}^3x\,\sqrt{-\det g}\;\eta^{\mu\nu}\,\eta^{\rho\sigma}\,\Psi^{[2]}_{[k_{(3)},S_{(3)}]\,\mu\rho}([x])\,\partial_\nu\partial_\sigma\Psi_{[k_{(2)}]}([x])\,\Psi_{[k_{(1)}]}([x]), \tag{6.95}$$

in Minkowskian coordinates or

$$K = \int_\Omega \mathrm{d}^3x\,\sqrt{-\det g}\;g^{\alpha\beta}\,g^{\gamma\delta}\,\Psi^{[2]}_{[k_{(3)},S_{(3)}]\,\alpha\gamma}([x])\,D_\beta\partial_\delta\Psi_{[k_{(2)}]}([x])\,\Psi_{[k_{(1)}]}([x]) \tag{6.96}$$

[49] Clearly we can perform the unfolding using whichever other field and this amount to keep the corresponding $m_{(i)}$ fixed in place of $m_{(1)}$.

[50] The underlying idea is to compute the amplitudes involving two tachyons and one massive state.

in orbifold coordinates where we need to use covariant derivatives. Using the unfolding trick over (3) we get

$$\begin{aligned} K = (2\pi)^3\,\delta\left(\sum_{i=-\infty}^{+\infty} k_{(i)\,+}\right)\prod_{i=2}^{N}\sum_{m_{(i)}=-\infty}^{+\infty} S_{(3)\mu\rho}\left(\mathcal{K}^{m_{(2)}}k_{(2)\,2}\right)^{\mu}\left(\mathcal{K}^{m_{(2)}}k_{(2)\,2}\right)^{\rho} \\ \times\,\delta\left(\sum_{i=-\infty}^{+\infty}\mathcal{K}^{m_{(i)}}k_{(i)\,2}\right)\delta\left(\sum_{i=-\infty}^{+\infty}\mathcal{K}^{m_{(i)}}k_{(i)\,-}\right). \end{aligned} \tag{6.97}$$

Explicitly in orbifold coordinates we can write

$$\begin{aligned} K = \int_{\Omega} \mathrm{d}^3x\,\sqrt{-\det g}\,\Bigg[& \Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]uu}\,\partial_v^2\Psi_{\left[k_{(2)}\right]} - \frac{2}{(\Delta u)^2}\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]uz}\,\partial_v\partial_z\Psi_{\left[k_{(2)}\right]} \\ & + 2\,\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]uv}\,\partial_v\partial_u\Psi_{\left[k_{(2)}\right]} + \frac{1}{(\Delta u)^4}\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]zz}\left(\partial_z^2\Psi_{\left[k_{(2)}\right]} - \Delta^2 u\,\partial_v\Psi_{\left[k_{(2)}\right]}\right) \\ & - \frac{2}{(\Delta u)^2}\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]zv}\left(\partial_z\partial_u\Psi_{\left[k_{(2)}\right]} - \frac{1}{u}\,\partial_z\Psi_{\left[k_{(2)}\right]}\right) + \Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]vv}\,\partial_u^2\Psi_{\left[k_{(2)}\right]}\Bigg]\Psi_{\left[k_{(1)}\right]}. \end{aligned} \tag{6.98}$$

Keeping the terms which do not vanish when all $l = 0$ and considering only the leading order in $\frac{1}{u}$ we get

$$K \sim \int \mathrm{d}u\;|u|\,\frac{3}{4}\frac{\left(k_{(2)\,+} + k_{(3)\,+}\right)^2}{k_{(3)\,+}^2}\,\mathcal{S}_{(3)\,vv}\;\frac{1}{u^2}\prod_{i=1}^{3}\Phi_{(i)}\Big|_{l_{(*)}=0}, \tag{6.99}$$

which is divergent as $|u|^{-\frac{5}{2}}$.

6.5 Three Points Amplitudes with One Massive State in String Theory

We consider string amplitudes including massive states. They are obtained using the inheritance principle and therefore they are connected to the integrals and relations derived in Section 6.4.3. In particular we want to use the inheritance principle on the momenta and polarisations, i.e. we start form amplitudes in Minkowski expressed with momenta and polarisations and then we implement on them the projection on the orbifold. In particular it is worth stressing that, as there is one Killing vector acting on the spacetime coordinates, there is only one common Killing vector action on all the momenta and polarisations of each field as discussed for spin-1 and spin-2 cases. Moreover this approach gives the complete answer only for tree level amplitudes since inside the loops twisted states may be created in pairs. The final result is that the open string amplitude with two tachyons and the first massive (level 2) state diverges and there is no obvious way of curing it since the divergence is also present in the Abelian sector. The open string expansion we use is

$$X(u,\,\overline{u}) = x_0 - i\,2\alpha'\,p\,\ln(|u|) + i\sqrt{\frac{\alpha'}{2}}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{\alpha_n}{n}\left(u^{-n} + \overline{u}^{-n}\right). \tag{6.100}$$

6.5.1 First Massive State in String Theory

Before computing the amplitude we would like to review the possible polarisations of the first massive state in open string. The first massive vertex is:

$$V_M(x;\, k,\, S,\, \xi) =: \left(\frac{i}{\sqrt{2\alpha'}} \xi \cdot \partial_x^2 X(x,\, x) + \left(\frac{i}{\sqrt{2\alpha'}} \right)^2 S_{\mu\nu}\, \partial_x X^\mu(x,\, x)\, \partial_x X^\nu(x,\, x) \right) \times e^{i\, k \cdot X(x,\, x)} : . \tag{6.101}$$

The corresponding state is:

$$\lim_{x \to 0} V_M(x;\, k,\, S,\, \xi)\, |0\rangle = |k,\, S,\, \xi\rangle = (\xi \cdot \alpha_{-2} + \alpha_{-1} \cdot S \cdot \alpha_{-1})\, |k\rangle\,. \tag{6.102}$$

For the state to be physical we require:

$$\begin{aligned} (L_0 - 1)\, |k,\, S,\, \xi\rangle = 0 \qquad &\Rightarrow \qquad \alpha' k^2 = -1 \\ L_1\, |k,\, S,\, \xi\rangle = 0 \qquad &\Rightarrow \qquad S \cdot k + \xi = 0 \\ L_2\, |k,\, S,\, \xi\rangle = 0 \qquad &\Rightarrow \qquad k \cdot \xi + \operatorname{tr} S = 0. \end{aligned} \tag{6.103}$$

String gauge invariance allows us to add:

$$L_{-1}(\chi \cdot \alpha_{-1}\, |k\rangle) = (\chi \cdot \alpha_{-2} + \chi \cdot \alpha_{-1}\, k \cdot \alpha_{-1})\, |k\rangle\,, \tag{6.104}$$

subject to the physical constraints $\alpha' k^2 + 1 = 0$ and $\chi \cdot k = 0$. In critical string theory there is another gauge invariance generated by $L_{-2} + \frac{3}{2} L_{-1}^2$. We can add a multiple of

$$\left(L_{-2} + \frac{3}{2} L_{-1}^2 \right) |k\rangle = \left(\frac{5}{2} k \cdot \alpha_{-2} + \frac{3}{2} (k \cdot \alpha_{-1})^2 + \frac{1}{2} \alpha_{-1}^2 \right)\, |k\rangle\,, \tag{6.105}$$

to set $a = 0$. Therefore the only non trivial D.O.F. refer to S^{TT}, that is:

$$\operatorname{tr} S^{TT} = k \cdot S^{TT} = \xi = 0. \tag{6.106}$$

We check that, given $k = \left(k_+,\, k_-,\, k_2,\, \vec{k} \right)$ such that $-2\, k_+ k_- + k_2^2 + \left\| \vec{k} \right\|^2 = -1$, we can find a non trivial S^{TT} with non vanishing components in the directions $\pm$, 2 only. In fact we find a two parameters family of solutions. The parameters may be taken to be S_{++} and $S_{+\,2}$. Explicitly we have

$$\begin{pmatrix} S_{++} \\ S_{+-} \\ S_{+\,2} \\ S_{--} \\ S_{-\,2} \\ S_{2\,2} \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{k_-}{k_+} \\ 0 \\ \frac{k_- \left(k_- k_+ - 2 k_2^2 \right)}{k_+^3} \\ -2 \frac{k_- k_2}{k_+^2} \\ -2 \frac{k_-}{k_+} \end{pmatrix} S_{++} + \begin{pmatrix} 0 \\ \frac{k_2}{k_+} \\ 1 \\ \frac{2 k_2 \left(-k_- k_+ + k_2^2 \right)}{k_+^3} \\ \frac{k_- k_+ - 2 k_2^2}{k_+^2} \\ 2 \frac{k_2}{k_+} \end{pmatrix} S_{+\,2} \tag{6.107}$$

There is even a non trivial solution for the special case $k = \left(k_+,\, k_- = \frac{1}{k_+},\, k_2 = 0, \vec{k} = \vec{0} \right)$.

Using the expressions for S^{TT} in orbifold coordinates, we check that there are two possible indepdendent polarisations $\mathcal{S}_{v\,v}$ and $\mathcal{S}_{v\,z}$ which correspond to the those used above. The non trivial solution is:

$$\begin{pmatrix} \mathcal{S}_{v\,v} \\ \mathcal{S}_{u\,v} \\ \mathcal{S}_{v\,z} \\ \mathcal{S}_{u\,u} \\ \mathcal{S}_{u\,z} \\ \mathcal{S}_{z\,z} \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{r+\|\vec{k}\|^2}{2k_+^2} \\ 0 \\ \left(\frac{r+\|\vec{k}\|^2}{2k_+^2}\right)^2 \\ 0 \\ -2\frac{r+\|\vec{k}\|^2}{2k_+^2} \end{pmatrix} \mathcal{S}_{v\,v} + \begin{pmatrix} 0 \\ -\frac{r+\|\vec{k}\|^2}{2k_+^2} \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \mathcal{S}_{v\,z}. \tag{6.108}$$

6.5.2 Two Tachyons and the First Massive State

This Minkowskian amplitude is given by the sum of two colour ordered sub-parts as:

$$\mathcal{A}_{TTM} = A_{T_{(1)}T_{(2)}M_{(3)}}\ \mathrm{tr}\big(T_{(1)}T_{(2)}T_{(3)}\big) + A_{T_{(2)}T_{(1)}M_{(3)}}\ \mathrm{tr}\big(T_{(2)}T_{(1)}T_{(3)}\big). \tag{6.109}$$

We find:

$$\begin{aligned} A_{T_{(1)}T_{(2)}M_{3)}} &= \big\langle\big\langle k_{(1)}\big|\ V_T\big(1;\,k_{(2)}\big)\left(\alpha_{-1}\cdot S_{(3)}^{TT}\cdot\alpha_{-1}\,\big|k_{(3)}\big\rangle\right) \\ &= \big\langle\big\langle k_{(1)}\big|\ e^{i\,k_{(2)}\cdot x_0}\,e^{-\sqrt{2\alpha'}k_{(2)}\cdot\alpha_1}\left(\alpha_{-1}\cdot S_{(3)}^{TT}\cdot\alpha_{-1}\,\big|k_{(3)}\big\rangle\right) \\ &= (2\pi)^D\left(\sqrt{2\alpha'}\right)^2\delta^D\left(\sum_{i=1}^3 k_{(i)}\right)k_{(2)}\cdot S_{(3)}^{TT}\cdot k_{(2)}. \end{aligned} \tag{6.110}$$

The transversality of $S_{(3)}^{TT}$ finally leads to:

$$\mathcal{A}_{TTM} = 2\,(2\pi)^D\left(\sqrt{2\alpha'}\right)^2\delta^D\left(\sum_{i=1}^3 k_{(i)}\right)k_{(2)}\cdot S_{(3)}^{TT}\cdot k_{(2)}\ \mathrm{tr}\left(\big[T_{(1)},\,T_{(2)}\big]_+\,T_{(3)}\right). \tag{6.111}$$

Then we can compute the orbifold amplitude as:

$$\begin{aligned} \mathcal{A}_{TTM} &= (2\pi)^{D-2}\delta^{D-3}\left(\sum_{i=1}^3\vec{k}_{(i)}\right)\delta\left(\sum_{i=1}^3 k_{(i)\,+}\right) \\ &\times 2\left(\sqrt{2\alpha'}\right)^2\sum_{\{m_{(1)},\,m_{(2)},\,m_{(3)}\}\in\mathbb{Z}^3}\delta_{m_{(3)},\,1}\left(\mathcal{K}^{m_{(2)}}k_{(2)}\right)\cdot S_{(3)}^{TT}\cdot\left(\mathcal{K}^{m_{(2)}}k_{(2)}\right) \\ &\times\delta\left(\sum_{i=1}^3\left(\mathcal{K}^{m_{(i)}}k_{(i)\,2}\right)\right)\delta\left(\sum_{i=1}^3\left(\mathcal{K}^{m_{(i)}}k_{(i)\,-}\right)\right)\ \mathrm{tr}\left(\big[T_{(1)},\,T_{(2)}\big]_+\,T_{(3)}\right). \end{aligned} \tag{6.112}$$

Such amplitude can then be expressed using an overlap:

$$\begin{aligned}\mathcal{A}_{TTM} &= 2\left(-i\sqrt{2\alpha'}\right)^2 \int_\Omega \mathrm{d}^3x\, g^{\mu\nu}\, g^{\rho\sigma}\, \Psi^{[2]}_{[k_{(3)},\, S_{(3)}]\,\mu\rho}([x])\, \partial_\nu\partial_\sigma \Psi_{[k_{(2)}]}([x])\, \Psi_{[k_{(1)}]}([x]) \\ &\quad \times \mathrm{tr}\left(\left[T_{(1)},\, T_{(2)}\right]_+ T_{(3)}\right), \\ &= 2\left(-i\sqrt{2\alpha'}\right)^2 \int_\Omega \mathrm{d}^3x\, g^{\alpha\beta}\, g^{\gamma\delta}\, \Psi^{[2]}_{[k_{(3)},\, S_{(3)}]\,\alpha\gamma}([x])\, D_\beta\partial_\delta \Psi_{[k_{(2)}]}([x])\, \Psi_{[k_{(1)}]}([x]) \\ &\quad \times \mathrm{tr}\left(\left[T_{(1)},\, T_{(2)}\right]_+ T_{(3)}\right).\end{aligned} \tag{6.113}$$

As discussed in Section 6.4.3 the integral is divergent when $S_{++} = \mathcal{S}_{v\,v} \neq 0$ and the divergence cannot be avoided even introducing a Wilson line around z since the amplitude involves an anticommutator which does not vanish in the Abelian sector.

6.6 Scalar QED on the Generalised NBO and Divergences

The issues related to the vanishing volume of the compact directions lead to incurable divergences. We introduce the GNBO by inserting one additional non compact direction with respect to the NBO and show that divergences no longer occur. As for the NBO, we first present the geometry of the GNBO and study scalar and spin-1 eigenfunctions to build the scalar QED on the orbifold. We then show how the presence of a non compact direction can cure the theory when considering amplitudes and overlaps.

6.6.1 Geometric Preliminaries

Consider Minkowski spacetime $\mathscr{M}^{1,D-1}$ and the change of coordinates from the lightcone set $(x^\mu) = (x^+, x^-, x^2, x^3, \vec{x})$ to $(x^\alpha) = (u, v, w, z, \vec{x})$:

$$\begin{aligned}&\begin{cases} x^- &= u \\ x^+ &= v + \frac{\Delta_2^2}{2} u(z+w)^2 + \frac{\Delta_3^2}{2} u(z-w)^2 \\ x^2 &= \Delta_2 u(z+w) \\ x^3 &= \Delta_3 u(z-w) \end{cases} \\ \Leftrightarrow &\begin{cases} u &= x^- \\ v &= x^+ - \frac{1}{2x^-}\left((x^2)^2 + (x^3)^2\right) \\ w &= \frac{1}{2x^-}\left(\frac{x^2}{\Delta_2} - \frac{x^3}{\Delta_3}\right) \\ z &= \frac{1}{2x^-}\left(\frac{x^2}{\Delta_2} + \frac{x^3}{\Delta_3}\right) \end{cases}\end{aligned} \tag{6.114}$$

where we do not perform any change on the transverse coordinates $\vec{x}$. The metric in these coordinates is non diagonal:

$$\mathrm{d}s^2 = -2\,\mathrm{d}u\,\mathrm{d}v + (\Delta_2^2 + \Delta_3^2)u^2(\mathrm{d}w^2 + \mathrm{d}z^2) + 2(\Delta_2^2 - \Delta_3^2)u^2\,\mathrm{d}w\,\mathrm{d}z + \eta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,, \tag{6.115}$$

and its determinant is:

$$-\det g = 4\,\Delta_2^2\Delta_3^2\,u^4. \tag{6.116}$$

From the previous expressions we can also derive the non vanishing Christoffel symbols:

$$\begin{aligned} \Gamma_{w\ w}^{\ v} &= \Gamma_{z\ z}^{\ v} = (\Delta_2^2 + \Delta_3^2)u, \\ \Gamma_{w\ z}^{\ v} &= (\Delta_2^2 - \Delta_3^2)u, \\ \Gamma_{u\ w}^{\ w} &= \Gamma_{u\ z}^{\ z} = u^{-1}, \end{aligned} \tag{6.117}$$

which however produce a vanishing Ricci tensor and curvature scalar since we are considering Minkowski spacetime anyway and (6.114) is just a map from $\mathscr{M}^{1,D-1}$ to the GNBO.

We introduce the GNBO by identifying points in space along the orbits of the Killing vector:

$$\begin{aligned} \kappa &= -2\pi i \left(\Delta_2 J_{+2} + \Delta_3 J_{+3}\right) \\ &= 2\pi \left(\Delta_2 x^2 + \Delta_3 x^3\right)\partial_+ + 2\pi\Delta_2 x^- \partial_2 + 2\pi\Delta_3 x^- \partial_3 \\ &= 2\pi\, \partial_z \end{aligned} \tag{6.118}$$

in such a way that

$$x^\mu \sim e^{n\kappa} x^\mu, \qquad n \in \mathbb{Z} \tag{6.119}$$

leads to the identifications

$$x = \begin{pmatrix} x^- \\ x^2 \\ x^3 \\ x^+ \\ \vec{x} \end{pmatrix} \equiv \mathcal{K}^n x = \begin{pmatrix} x^- \\ x^2 + 2\pi n \Delta_3^- \\ x^3 + 2\pi n \Delta_3 x^- \\ x^+ + 2\pi n \Delta_3^2 + 2\pi n \Delta_3 x^3 + (2\pi n)^2 \frac{\Delta_2^2 + \Delta_3^2}{2} x^- \\ \vec{x} \end{pmatrix}, \tag{6.120}$$

or to the simpler

$$(u,\, v,\, w,\, z) \sim (u,\, v,\, w,\, z + 2\pi n) \tag{6.121}$$

using the map to the orbifold coordinates (6.114) where the Killing vector $\kappa = 2\pi\,\partial_z$ does not depend on the local spacetime configuration. As in the previous case, the difference between Minkowski spacetime and the GNBO is therefore global.

The geodesic distance between the n-th copy and the base point on the orbifold can be computed in any set of coordinates and is:

$$\Delta s^2_{(n)} = \left(\Delta_2^2 + \Delta_3^2\right)\left(2\pi n x^-\right)^2 \geq 0. \tag{6.122}$$

Closed time-like curves are therefore avoided on the GNBO, but there are closed null curves on the surface $x^- = u = 0$ where the Killing vector κ vanishes.

6.6.2 Free Scalar Field

In order to build a quantum theory on the GNBO using Feynman's approach to quantization, we first solve the eigenvalue equations for the fields and then build their off-shell expansion. We start from a complex scalar field and then consider the free photon before moving to the scalar QED interactions on the GNBO.

Consider the action for a complex scalar field:

$$\begin{aligned}
\mathrm{S}_{\mathrm{sQED}}^{(\mathrm{kinetic})}[\phi] &= \int_{\Omega} \mathrm{d}^D x \, \sqrt{-\det g}\big(-g^{\mu\nu}\partial_\mu\phi^*\partial_\nu\phi - M^2\phi^*\phi\big) \\
&= \int_{\mathbb{R}^{D-4}} \mathrm{d}^{D-4}\vec{x} \int_{-\infty}^{+\infty}\mathrm{d}u \int_{-\infty}^{+\infty}\mathrm{d}v \int_{-\infty}^{+\infty}\mathrm{d}w \int_0^{2\pi}\mathrm{d}z\, 2\,|\Delta_2\Delta_3|\,u^2 \\
&\times \Bigg[\partial_u\phi^*\,\partial_v\phi + \partial_v\phi^*\,\partial_u\phi - \frac{1}{4u^2}\Bigg(\Big(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\Big)(\partial_w\phi^*\,\partial_w\phi + \partial_z\phi^*\,\partial_z\phi) \\
&+ \Big(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\Big)(\partial_w\phi^*\,\partial_z\phi + \partial_z\phi^*\,\partial_w\phi)\Bigg) - \eta^{ij}\,\partial_i\phi^*\,\partial_j\phi - M^2\phi^*\phi\Bigg].
\end{aligned} \tag{6.123}$$

As in the case of the NBO, the solutions to the E.O.M. are necessary to provide the modes of the quantum fields. We study the eigenvalue equation $\Box\phi_r = r\phi_r$, where r is $2\,k_+k_- - \vec{k}$ by comparison with the flat case (k is the momentum associated to the flat coordinates). We therefore need solve:

$$\begin{aligned}
\Bigg\{ &-2\,\partial_u\partial_v - \frac{2}{u}\partial_v + \frac{1}{4u^2}\Bigg[\Big(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\Big)\big(\partial_w^2+\partial_z^2\big) \\
&+ 2\Big(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\Big)\partial_w\partial_z\Bigg] + \eta^{ij}\,\partial_i\partial_j - r\Bigg\}\phi_r = 0.
\end{aligned} \tag{6.124}$$

To this purpose, we introduce a Fourier transformation over v, w, z, $\vec{x}$:

$$\begin{aligned}
&\phi_r(u,\,v,\,w,\,z,\,\vec{x}) \\
&= \sum_{l=-\infty}^{+\infty} \int_{\mathbb{R}^{D-4}} \mathrm{d}^{D-4}\vec{k} \int_{-\infty}^{+\infty}\mathrm{d}k_+ \int_{-\infty}^{+\infty}\mathrm{d}p\; e^{i\left(k_+v+pw+lz+\vec{k}\cdot\vec{x}\right)}\widetilde{\phi}_{\{k_+,\,p,\,l,\,\vec{k},\,r\}}(u),
\end{aligned} \tag{6.125}$$

where we defined k_+, p, l, $\vec{k}$ as associated momenta to v, w, z, $\vec{x}$ respectively. We find:

$$\phi_{\{k_+,\,p,\,l,\,\vec{k},\,r\}}(u,\,v,\,w,\,z,\,\vec{x}) = e^{i\left(k_+v+pw+lz+\vec{k}\cdot\vec{x}\right)}\widetilde{\phi}_{\{k_+,\,p,\,l,\,\vec{k},\,r\}}(u). \tag{6.126}$$

where

$$\widetilde{\phi}_{\{k_+,\,p,\,l,\,\vec{k},\,r\}}(u) = \frac{1}{2\sqrt{(2\pi)^D|\Delta_2\Delta_3k_+|}}\,\frac{1}{|u|}e^{-i\left(\frac{1}{8k_+u}\left[\frac{(l+p)^2}{\Delta_2^2}+\frac{(l-p)^2}{\Delta_3^2}\right]-\frac{\vec{k}^2+r}{2k_+}u\right)}. \tag{6.127}$$

These solutions present the right normalisation, as we can verify through the product:

$$\begin{aligned}
&\Big(\phi_{\{k_{(1)+},\,p_{(1)},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\}},\;\phi_{\{k_{(2)+},\,p_{(2)},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\}}\Big) \\
&= \int_{\mathbb{R}^{D-4}} \mathrm{d}^{D-4}\vec{x} \int_{-\infty}^{+\infty}\mathrm{d}u \int_{-\infty}^{+\infty}\mathrm{d}v \int_{-\infty}^{+\infty}\mathrm{d}w \int_0^{2\pi}\mathrm{d}z\; 2|\Delta_2\Delta_3|u^2 \\
&\times \phi_{\{k_{(1)+},\,p_{(1)},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\}}\;\phi_{\{k_{(2)+},\,p_{(2)},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\}} \\
&= \delta^{D-4}\Big(\vec{k}_{(1)}+\vec{k}_{(2)}\Big)\,\delta\big(k_{(1)+}+k_{(2)+}\big)\,\delta\big(p_{(1)}+p_{(2)}\big)\delta\big(r_{(i)}+r_{(i)}\big)\,\delta_{l_{(1)},\,l_{(2)}}.
\end{aligned} \tag{6.128}$$

Then we have the off-shell expansion:

$$\begin{aligned}\phi_r(u,\,v,\,w,\,z,\,\vec{x}) = &\frac{1}{2\sqrt{(2\pi)^D|\Delta_2\Delta_3 k_+|}}\sum_{l=-\infty}^{+\infty}\int_{\mathbb{R}^{D-4}}\mathrm{d}^{D-4}\vec{k}\int_{-\infty}^{+\infty}\mathrm{d}k_+\int_{-\infty}^{+\infty}\mathrm{d}p\int_{-\infty}^{+\infty}\mathrm{d}r\\ &\times\frac{\mathcal{A}_{\{k_+,\,p,\,l,\,\vec{k},\,r\}}}{|u|}e^{\,i\left(k_+v+pw+lz+\vec{k}\cdot\vec{x}-\frac{1}{8k_+u}\left[\frac{(l+p)^2}{\Delta_2^2}+\frac{(l-p)^2}{\Delta_3^2}\right]+\frac{\vec{k}^2+r}{2k_+}u\right)}.\end{aligned} \tag{6.129}$$

6.6.3 Free Photon Action

We then study the action of the free photon field a using the Lorenz gauge which in the orbifold coordinates it reads:

$$\begin{aligned}D^\alpha a_\alpha = &-\frac{2}{u}a_v-\partial_v a_u-\partial_u a_v\\ &+\frac{1}{4u^2}\left(\left(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\right)(\partial_w a_w+\partial_z a_z)+\left(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\right)(\partial_w a_z+\partial_z a_w)\right)\\ &+\eta^{ij}\,\partial_i a_j = 0.\end{aligned} \tag{6.130}$$

We then solve the eigenvalue equations $(\Box a_r)_\nu = r a_{r\,\nu}$, which in components read:

$$
\begin{aligned}
(\Box a_r)_u &= \frac{2}{u^2} a_{r\,v} \\
&\quad - \frac{1}{2u^3}\left[\left(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\right)(\partial_w a_{r\,w} + \partial_z a_{r\,z}) + \left(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\right)(\partial_w a_{r\,z} + \partial_z a_{r\,w})\right] \\
&\quad + \Bigg\{ -2\partial_u\partial_v - \frac{2}{u}\partial_v \\
&\quad + \frac{1}{4u^2}\left[\left(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\right)(\partial_w^2+\partial_z^2) + \left(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\right) 2\partial_w\partial_z\right] + \nabla_T^2 \Bigg\} a_{r\,u}, \\
(\Box a_r)_v &= \Bigg\{ -2\partial_u\partial_v - \frac{2}{u}\partial_v \\
&\quad + \frac{1}{4u^2}\left[\left(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\right)(\partial_w^2+\partial_z^2) + \left(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\right) 2\partial_w\partial_z\right] + \nabla_T^2 \Bigg\} a_{r\,v}, \\
(\Box a_r)_w &= -\frac{2}{u}\partial_w a_{r\,v} \\
&\quad + \Bigg\{ -2\partial_u\partial_v \\
&\quad + \frac{1}{4u^2}\left[\left(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\right)(\partial_w^2+\partial_z^2) + \left(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\right) 2\,\partial_w\partial_z\right] + \nabla_T^2 \Bigg\} a_{r\,w}, \\
(\Box a)_z &= -\frac{2}{u}\partial_z a_{r\,v} \\
&\quad + \Bigg\{ -2\partial_u\partial_v \\
&\quad + \frac{1}{4u^2}\left[\left(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\right)(\partial_w^2+\partial_z^2) + \left(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\right) 2\partial_w\partial_z\right] + \nabla_T^2 \Bigg\} a_{r\,z}, \\
(\Box a)_i &= \Bigg\{ -2\partial_u\partial_v - \frac{2}{u}\partial_v \\
&\quad + \frac{1}{4u^2}\left[\left(\frac{1}{\Delta_2^2}+\frac{1}{\Delta_3^2}\right)(\partial_w^2+\partial_z^2) + \left(\frac{1}{\Delta_2^2}-\frac{1}{\Delta_3^2}\right) 2\partial_w\partial_z\right] + \nabla_T^2 \Bigg\} a_{r\,i},
\end{aligned}
\tag{6.131}
$$

where $\nabla_T^2 = \eta^{ij}\,\partial_i\partial_j$ is the Laplace operator in the transverse coordinates $\vec{x}$. These equations can be solved using standard techniques through a Fourier transform:

$$
\begin{aligned}
a_{r\,\alpha}(u,\,v,\,w,\,z,\,\vec{x}) = \sum_{l=-\infty}^{+\infty} \int_{\mathbb{R}^{D-4}} \mathrm{d}^{D-4}\vec{k} \int_{-\infty}^{+\infty} \mathrm{d}k_+ \int_{-\infty}^{+\infty} \mathrm{d}p \\
\times\, e^{i\left(k_+ v + p w + l z + \vec{k}\cdot\vec{x}\right)} \widetilde{a}_{\{k_+,\,p,\,l,\,\vec{k},\,r\}\,\alpha}(u).
\end{aligned}
\tag{6.132}
$$

We first solve the equations for $\widetilde{a}_{\{k_+,\,p,\,l,\,\vec{k},\,r\}\,v}$ and $\widetilde{a}_{\{k_+,\,p,\,l,\,\vec{k},\,r\}\,i}$ since they are identical to the scalar equation (6.124). We then insert their solutions as sources for the equations for

$\widetilde{a}_{\{k_+,p,l,\vec{k},r\}\,u}$, $\widetilde{a}_{\{k_+,p,l,\vec{k},r\}\,w}$ and $\widetilde{a}_{\{k_+,p,l,\vec{k},r\}\,z}$. The solutions can be written as the expansion:

$$\begin{aligned}
\left\|\widetilde{a}_{\{k_+,p,l,\vec{k},r\}\,\alpha}(u)\right\| &= \begin{pmatrix} \widetilde{a}_u \\ \widetilde{a}_v \\ \widetilde{a}_w \\ \widetilde{a}_z \\ \widetilde{a}_i \end{pmatrix} \\
&= \sum_{\underline{\alpha}\in\{\underline{u},\underline{v},\underline{w},\underline{z},\underline{i}\}} \mathcal{E}_{\{k_+,p,l,\vec{k},r\}\,\underline{\alpha}} \left\|\widetilde{a}^{\underline{\alpha}}_{\{k_+,p,l,\vec{k},r\}\,\alpha}(u)\right\| \\
&= \mathcal{E}_{\{k_+,p,l,\vec{k},r\}\,\underline{u}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \widetilde{\Phi}_{\{k_+,p,l,\vec{k},r\}} \\
&+ \mathcal{E}_{\{k_+,p,l,\vec{k},r\}\,\underline{v}} \begin{pmatrix} \frac{i}{2k_+u} + \frac{1}{8k_+^2u^2}\left(\frac{(l+p)^2}{\Delta_2^2} + \frac{(l-p)^2}{\Delta_3^2}\right) \\ 1 \\ \frac{p}{k_+} \\ \frac{l}{k_+} \\ 0 \end{pmatrix} \widetilde{\Phi}_{\{k_+,p,l,\vec{k},r\}} \\
&+ \mathcal{E}_{\{k_+,p,l,\vec{k},r\}\,\underline{w}} \begin{pmatrix} \frac{1}{4k_+|u|}\left(\frac{l+p}{\Delta_2^2} - \frac{l-p}{\Delta_3^2}\right) \\ 0 \\ |u| \\ 0 \\ 0 \end{pmatrix} \widetilde{\Phi}_{\{k_+,p,l,\vec{k},r\}} \\
&+ \mathcal{E}_{\{k_+,p,l,\vec{k},r\}\,\underline{z}} \begin{pmatrix} \frac{1}{4k_+|u|}\left(\frac{l+p}{\Delta_2^2} + \frac{l-p}{\Delta_3^2}\right) \\ 0 \\ 0 \\ |u| \\ 0 \end{pmatrix} \widetilde{\Phi}_{\{k_+,p,l,\vec{k},r\}} \\
&+ \mathcal{E}_{\{k_+,p,l,\vec{k},r\}\,\underline{j}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \delta_{\underline{ij}} \end{pmatrix} \widetilde{\Phi}_{\{k_+,p,l,\vec{k},r\}}
\end{aligned} \tag{6.133}$$

Consider the Fourier transformed functions:

$$a^{\underline{\alpha}}_{\{k_+,p,l,\vec{k},r\}\,\alpha}(u,\,v,\,w,\,z,\,\vec{x}) = e^{i\left(k_+v+pw+lz+\vec{k}\cdot\vec{x}\right)}\widetilde{a}^{\underline{\alpha}}_{\{k_+,p,l,\vec{k},r\}\,\alpha}(u), \tag{6.134}$$

then we can expand the off shell fields as

$$\begin{aligned}
a_\alpha(x) = \sum_{l=-\infty}^{+\infty} \int_{\mathbb{R}^{D-4}} \mathrm{d}^{D-4}\vec{k} \int_{-\infty}^{+\infty} \mathrm{d}k_+ \int_{-\infty}^{+\infty} \mathrm{d}p \int_{-\infty}^{+\infty} \mathrm{d}r \\
\times \sum_{\underline{\alpha}\in\{\underline{u},\underline{v},\underline{w},\underline{z},\underline{i}\}} \mathcal{E}_{\{k_+,p,l,\vec{k},r\}\,\alpha}\, a^{\underline{\alpha}}_{\{k_+,p,l,\vec{k},r\}\,\alpha}(x).
\end{aligned} \tag{6.135}$$

We can compute the normalisation as:

$$\begin{aligned}\big(a_{(1)},\, a_{(2)}\big) = &\int\limits_{\mathbb{R}^{D-4}} \mathrm{d}^{D-4}\vec{x} \int\limits_{-\infty}^{+\infty} \mathrm{d}u \int\limits_{-\infty}^{+\infty} \mathrm{d}v \int\limits_{-\infty}^{+\infty} \mathrm{d}w \int\limits_{0}^{2\pi} \mathrm{d}z\; 2|\Delta_2\Delta_3|u^2 \\ &\times \Big(g^{\alpha\beta}\, a_{\{k_{(1)\,+},\, p_{(1)},\, l_{(1)},\, \vec{k}_{(1)},\, r_{(1)}\}\,\alpha}\, a_{\{k_{(2)\,+},\, p_{(2)},\, l_{(2)},\, \vec{k}_{(2)},\, r_{(2)}\}\,\beta}\Big) \\ &= \delta^{D-4}\Big(\vec{k}_{(1)} + \vec{k}_{(2)}\Big)\, \delta\big(p_{(1)} + p_{(2)}\big)\, \delta\big(k_{(1)\,+} + k_{(2)\,+}\big)\, \delta_{l_{(1)}+l_{(2)},0}\delta(r_1 - r_2) \\ &\times \mathcal{E}_{\{k_{(1)\,+},\, p_{(1)},\, l_{(1)},\, \vec{k}_{(1)},\, r_{(1)}\}} \circ \mathcal{E}_{\{k_{(2)\,+},\, p_{(2)},\, l_{(2)},\, \vec{k}_{(2)},\, r_{(2)}\}},\end{aligned} \tag{6.136}$$

where

$$\begin{aligned}\mathcal{E}_{(1)} \circ \mathcal{E}_{(2)} = &-\mathcal{E}_{(1)\,\underline{u}}\, \mathcal{E}_{(2)\,\underline{v}} - \mathcal{E}_{(1)\,\underline{v}}\, \mathcal{E}_{(2)\,\underline{u}} \\ &+ \frac{1}{4}\Bigg[\left(\frac{1}{\Delta_2^2} + \frac{1}{\Delta_3^2}\right)\big(\mathcal{E}_{(1)\,\underline{w}}\, \mathcal{E}_{(2)\,\underline{w}} + \mathcal{E}_{(1)\,\underline{z}}\, \mathcal{E}_{(2)\,\underline{z}}\big) \\ &+ \left(\frac{1}{\Delta_2^2} - \frac{1}{\Delta_3^2}\right)\big(\mathcal{E}_{(1)\,\underline{w}}\, \mathcal{E}_{(2)\,\underline{z}} + \mathcal{E}_{(1)\,\underline{z}}\, \mathcal{E}_{(2)\,\underline{w}}\big)\Bigg]\end{aligned} \tag{6.137}$$

is independent of the coordinates. The Lorenz gauge now reads:

$$\eta^{\underline{ij}}\, k_i\, \mathcal{E}_{\{k_+,\, p,\, l,\, \vec{k},\, r\}\,\underline{j}} - k_+ \mathcal{E}_{\{k_+,\, p,\, l,\, \vec{k},\, r\}\,\underline{u}} - \frac{\vec{k}^2 + r}{2k_+} \mathcal{E}_{\{k_+,\, p,\, l,\, \vec{k},\, r\}\,\underline{v}} = 0. \tag{6.138}$$

As in the previous case, the constraint equation does not pose any condition on the transverse polarisations $\mathcal{E}_{\{k_+,\, p,\, l,\, \vec{k},\, r\}\,\underline{w}}$ and $\mathcal{E}_{\{k_+,\, p,\, l,\, \vec{k},\, r\}\,\underline{z}}$.

6.6.4 Cubic Interaction

As previously studied on the NBO, we show the scalar QED 3-points vertex computation using the previously computed eigenmodes. The presence of a continuous momentum in the non compact direction plays a major role in saving the convergence of the integrals. In the case of the GNBO

we find:

$$
\begin{aligned}
\mathrm{S}^{(\mathrm{cubic})}_{\mathrm{sQED}}[\phi,\, a] &= \int\limits_{\Omega} \mathrm{d}^{D}x\, \sqrt{-\det g}\, (-ieg^{\mu\nu}\, a_{\mu}(\phi^{*}\, \partial_{\nu}\phi - \partial_{\nu}\phi^{*}\, \phi)) \\
&= \prod_{i=1}^{3} \sum_{l_{(i)}=-\infty}^{+\infty} \int\limits_{\mathbb{R}^{D-4}} \mathrm{d}^{D-4}\vec{k}_{(i)} \int\limits_{-\infty}^{+\infty} \mathrm{d}k_{(i)+} \int\limits_{-\infty}^{+\infty} \mathrm{d}p_{(i)} \int\limits_{-\infty}^{+\infty} \mathrm{d}r_{(i)} \\
&\times (2\pi)^{D-1}\, \delta^{D-4}\left(\sum_{i=1}^{3} \vec{k}_{(i)}\right) \delta\left(\sum_{i=1}^{3} p_{(i)}\right) \delta\left(\sum_{i=1}^{3} k_{(i)+}\right) \delta_{\sum\limits_{i=1}^{3} l_{(i)},\, 0} \\
&\times e\, \mathcal{A}^{*}_{\{-k_{(2)+},\, -p_{(2)},\, -l_{(2)},\, -\vec{k}_{(2)},\, r_{(2)}\}} \mathcal{A}_{\{k_{(3)+},\, p_{(3)},\, l_{(3)},\, \vec{k}_{(3)},\, r_{(3)}\}} \\
&\times \Bigg\{ \mathcal{E}_{\{k_{(1)+},\, p_{(1)},\, l_{(1)},\, \vec{k}_{(1)},\, r_{(1)}\}\, \underline{u}}\; k_{(2)+}\, \mathcal{I}^{[0]}_{\{3\}} \\
&+ \mathcal{E}_{\{k_{(1)+},\, p_{(1)},\, l_{(1)},\, \vec{k}_{(1)},\, r_{(1)}\}\, \underline{v}} \Bigg[\left(\frac{\vec{k}^{2}_{(2)} + r_{(2)}}{2k_{(2)+}} \right) \mathcal{I}^{[0]}_{\{3\}} + i \frac{k_{(2)+}}{k_{(1)+}}\, \mathcal{I}^{[-1]}_{\{3\}} \\
&+ \frac{k_{(2)+}}{8} \Bigg[\frac{1}{\Delta_2^2} \left(\frac{l_{(1)} + p_{(1)}}{k_{(1)+}} + \frac{l_{(2)} + p_{(2)}}{k_{(2)+}} \right)^{2} \\
&+ \frac{1}{\Delta_3^2} \left(\frac{l_{(1)} - p_{(1)}}{k_{(1)+}} + \frac{l_{(2)} - p_{(2)}}{k_{(2)+}} \right)^{2} \Bigg] \mathcal{I}^{[-2]}_{\{3\}} \Bigg] \\
&+ \left(\mathcal{E}_{\{k_{(1)+},\, p_{(1)},\, l_{(1)},\, \vec{k}_{(1)},\, r_{(1)}\}\, \underline{w}} - \mathcal{E}_{\{k_{(1)+},\, p_{(1)},\, l_{(1)},\, \vec{k}_{(1)},\, r_{(1)}\}\, \underline{z}} \right) \\
&\times \Bigg[\frac{1}{\Delta_2^2} \left(\frac{k_{(1)+}\big(l_{(2)} + p_{(2)}\big) + k_{(2)+}\big(l_{(1)} + p_{(1)}\big)}{k_{(1)+}} \right) \\
&- \frac{1}{\Delta_3^2} \left(\frac{k_{(1)+}\big(l_{(2)} - p_{(2)}\big) + k_{(2)+}\big(l_{(1)} - p_{(1)}\big)}{k_{(1)+}} \right) \Bigg] \mathcal{J}^{[-1]}_{\{3\}} \\
&+ ((2) \leftrightarrow (3)) \Bigg\}
\end{aligned}
\tag{6.139}
$$

where we defined:

$$
\begin{aligned}
\mathcal{I}^{[\nu]}_{\{N\}} &= \int\limits_{-\infty}^{+\infty} \mathrm{d}u\, 2\, |\Delta_2 \Delta_3| u^{2}\, u^{\nu} \prod_{i=1}^{N} \widetilde{\phi}_{\{k_{(i)+},\, p_{(i)},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\}}, \\
\mathcal{J}^{[\nu]}_{\{N\}} &= \int\limits_{-\infty}^{+\infty} \mathrm{d}u\, 2\, |\Delta_2 \Delta_3| u^{2}\, |u|^{\nu} \prod_{i=1}^{N} \widetilde{\phi}_{\{k_{(i)+},\, p_{(i)},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\}}.
\end{aligned}
\tag{6.140}
$$

While in the NBO case we need to regularise the integrals at least taking their principal part when all $l_{(*)} = 0$ in (6.37), the GNBO does not need any specific manipulation. In fact the form of $\widetilde{\phi}_{\{k_{(i)+},\, p_{(i)},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\}}$ in (6.127) prevents the formation of isolated zeros in the phase factor proportional to u^{-1}: the presence of the continuous momentum p, contrary to the NBO where all

momenta are discrete, gives the integrals a distributional interpretation, similar to a derivative of a Dirac δ function.

6.6.5 Quartic Interactions

As for the NBO, we consider the quartic interaction for the scalar QED action:

$$
\begin{aligned}
\mathrm{S}^{(\mathrm{quartic})}_{\mathrm{sQED}}[\phi,\, a] &= \int_{\Omega} \mathrm{d}^{D}x\, \sqrt{-\det g}\left(e^{2}\, g^{\mu\nu}\, a_{\mu}a_{\nu}|\phi|^{2} - \frac{g_{4}}{4}|\phi|^{4}\right) \\
&= \prod_{i=1}^{3}\left(\frac{1}{4\pi\sqrt{(2\pi)^{D}\left|\Delta_{2}\Delta_{3}k_{(i)+}\right|}}\right) \\
&\times \sum_{l_{(i)}=-\infty}^{+\infty} \int_{\mathbb{R}^{D-4}} \mathrm{d}^{D-4}\vec{k}_{(i)} \int_{-\infty}^{+\infty} \mathrm{d}k_{(i)+} \int_{-\infty}^{+\infty} \mathrm{d}p_{(i)} \int_{-\infty}^{+\infty} \mathrm{d}r_{(i)} \\
&\times (2\pi)^{D-1}\, \delta^{D-4}\left(\sum_{i=1}^{3}\vec{k}_{(i)}\right) \delta\left(\sum_{i=1}^{3}p_{(i)}\right) \delta\left(\sum_{i=1}^{3}k_{(i)+}\right) \delta_{\sum_{i=1}^{3} l_{(i)},\,0} \\
&\times \Bigg\{ e^{2}\mathcal{A}^{*}_{\left\{-k_{(3)+},\,-p_{(3)},\,-l_{(3)},\,-\vec{k}_{(3)},\,r_{(3)}\right\}} \mathcal{A}_{\left\{k_{(4)+},\,p_{(4)},\,l_{(4)},\,\vec{k}_{(4)},\,r_{(4)}\right\}} \\
&\times \Bigg[\mathcal{E}_{\left\{k_{(1)+},\,p_{(1)},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}} \circ \mathcal{E}_{\left\{k_{(2)+},\,p_{(2)},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\right\}} \mathcal{I}^{[0]}_{\{4\}} \\
&- i\mathcal{E}_{\left\{k_{(1)+},\,p_{(1)},\,l_{(1)},\,\vec{k}_{(1)},\,r_{(1)}\right\}\,\underline{v}}\, \mathcal{E}_{\left\{k_{(2)+},\,p_{(2)},\,l_{(2)},\,\vec{k}_{(2)},\,r_{(2)}\right\}\,\underline{v}} \\
&\times \Bigg(\left(\frac{1}{k_{(1)+}} + \frac{1}{k_{(2)+}}\right) \mathcal{I}^{[-1]}_{\{4\}} \\
&- i\left(\frac{\mathcal{G}_{+\,(1,2)}}{\Delta_{2}^{2}} + \frac{\mathcal{G}_{-\,(1,2)}}{\Delta_{3}^{2}}\right) \mathcal{I}^{[-2]}_{\{4\}} \Bigg) \\
&+ \frac{1}{4}\left(\tilde{\mathcal{E}}_{+\,(1,2)}\, \frac{\mathcal{G}_{+\,(1,2)}}{\Delta_{2}^{2}} - \tilde{\mathcal{E}}_{-\,(1,2)}\, \frac{\mathcal{G}_{-\,(1,2)}}{\Delta_{2}^{2}}\right) \mathcal{J}^{[-1]}_{\{4\}} \Bigg] \\
&- \frac{g_{4}}{4}\mathcal{A}^{*}_{\left\{-k_{(1)+},\,-p_{(1)},\,-l_{(1)},\,-\vec{k}_{(1)},\,r_{(1)}\right\}} \mathcal{A}^{*}_{\left\{-k_{(2)+},\,-p_{(2)},\,-l_{(2)},\,-\vec{k}_{(2)},\,r_{(2)}\right\}} \\
&\times \mathcal{A}_{\left\{k_{(3)+},\,p_{(3)},\,l_{(3)},\,\vec{k}_{(3)},\,r_{(3)}\right\}} \mathcal{A}_{\left\{k_{(4)+},\,p_{(4)},\,l_{(4)},\,\vec{k}_{(4)},\,r_{(4)}\right\}} \mathcal{I}^{[0]}_{\{4\}} \Bigg\},
\end{aligned}
\tag{6.141}
$$

where we defined:

$$\begin{aligned}
\mathcal{G}_{\pm\,(a,b)} &= \frac{l_{(a)} \pm p_{(a)}}{k_{(a)\,+}} - \frac{l_{(b)} \pm p_{(b)}}{k_{(b)\,+}}, \\
\tilde{\mathcal{E}}_{\pm\,(a,b)} &= \mathcal{E}_{\left\{k_{(a)\,+},\, p_{(a)},\, l_{(a)},\, \vec{k}_{(a)},\, r_{(a)}\right\}\underline{v}} \\
&\quad \times \left(\mathcal{E}_{\left\{k_{(b)\,+},\, p_{(b)},\, l_{(b)},\, \vec{k}_{(b)},\, r_{(b)}\right\}\underline{w}} \pm \mathcal{E}_{\left\{k_{(b)\,+},\, p_{(b)},\, l_{(b)},\, \vec{k}_{(b)},\, r_{(b)}\right\}\underline{z}}\right) \\
&\quad - \mathcal{E}_{\left\{k_{(b)\,+},\, p_{(b)},\, l_{(b)},\, \vec{k}_{(b)},\, r_{(b)}\right\}\underline{v}} \\
&\quad \times \left(\mathcal{E}_{\left\{k_{(a)\,+},\, p_{(a)},\, l_{(a)},\, \vec{k}_{(a)},\, r_{(a)}\right\}\underline{w}} \pm \mathcal{E}_{\left\{k_{(a)\,+},\, p_{(a)},\, l_{(a)},\, \vec{k}_{(a)},\, r_{(a)}\right\}\underline{z}}\right)
\end{aligned} \tag{6.142}$$

for simplicity.

As the four points function in the NBO case shows with clear evidence the presence of divergences when all $l_{(*)} = 0$, the GNBO allows a distributional interpretation of the integrals $\mathcal{I}^{[\nu]}_{\{N\}}$ and $\mathcal{J}^{[\nu]}_{\{N\}}$ in the previous expression. In fact the regularization occurs in the same way as in the three points function in the GNBO: the phase factor proportional to u^{-1} has a continuous value due to the continous momentum p and it does not present isolated zeros which would prevent the interpretation as distribution.

6.6.6 Resurgence of Divergences and Null Brane Regularisation

Looking back at the metric (6.115) and at the identifications (6.121) it seems reasonable to wonder what would happen if we acted in the same way over w, since $2\pi\partial_w$ is a Killing vector as well and it commutes with $2\pi\partial_z$. However from the analysis of NBO and GNBO, in the absence of at least one continuous transverse direction it is not possible to avoid the divergences associated with discrete zero energy modes and this is exactly what happens.

As mentioned in the introductory section, there have been attempts to regularise the NBO using the Null Brane. Differently from the NBO, in this case the orbifold generator (6.9) includes an additional translation along an extra spatial dimension, namely:

$$\begin{aligned}
\kappa &= -2\pi i\Delta\, J_{+2} - 2\pi i R P_3 \\
&= 2\pi\,(\Delta\,\partial_z + R\,\partial_3).
\end{aligned} \tag{6.143}$$

with metric

$$\mathrm{d}s^2 = -2\,\mathrm{d}u\,\mathrm{d}v + \Delta^2 u^2 (\mathrm{d}z)^2 + \left(\mathrm{d}x^3\right)^2 + \eta_{ij}\ \mathrm{d}x^i\,\mathrm{d}x^j\,. \tag{6.144}$$

Even though similar in appearance to the GNBO Killing vector, this Killing vector is substantially different from (6.118).

The scalar field satisfies the same equation of motion as in the NBO:

$$\left(-2\partial_u\partial_v - \frac{1}{u}\partial_v + \frac{1}{(u\Delta)^2}\partial_z^2 + \partial_{x^3}^2 + \eta^{ij}\,\partial_i\partial_j\right)\phi_r = r\phi_r, \tag{6.145}$$

where $i,\, j = 4,\, 5,\, \ldots D-1$. The solution is:

$$\tilde{\phi}_{\left\{k_+\ k_z\ k_3\ \vec{k}\ r\right\}}(u) \propto \frac{1}{\sqrt{|u|}} e^{-i\frac{k_z^2}{2k_+}\frac{1}{u} + i\frac{k_3^2+\vec{k}^2+r}{2k_+}u}. \tag{6.146}$$

but with different periodicity conditions:

$$e^{i2\pi n(\Delta k_z + R k_3)} = 1. \tag{6.147}$$

This obscures the issue of the presence of a non compact direction. To show the non compact direction hidden in this system we define the coordinates $\widehat{z} = \frac{1}{2}\left(\frac{x^3}{R} + \frac{z}{\Delta}\right)$ and $\widehat{x}^3 = \frac{1}{2}\left(\frac{x^3}{R} - \frac{z}{\Delta}\right)$ such that $\kappa = 2\pi\partial_{\widehat{z}}$ and

$$\begin{pmatrix} \widehat{z} \\ \widehat{x}^3 \end{pmatrix} \equiv \begin{pmatrix} \widehat{z} + 2\pi n \\ \widehat{x}^3 \end{pmatrix} \tag{6.148}$$

upon the orbifold identification. Then the momenta are $\widehat{k}_{\widehat{z}} = \widehat{l} \in \mathbb{Z}$ and $\widehat{k}_3 \in \mathbb{R}$ and they are related to the momenta of the other coordinates as:

$$k_3 = \frac{\widehat{l} + \widehat{k}_3}{2R}, \qquad k_z = \frac{\widehat{l} - \widehat{k}_3}{2\Delta}, \tag{6.149}$$

so that the solution can be written as

$$\tilde{\Phi}_{\left\{k_+ \widehat{l}\widehat{k}_3 \vec{k} r\right\}}(u) \propto \frac{1}{\sqrt{|u|}} e^{-i\frac{(\widehat{l}-\widehat{k}_3)^2}{8\Delta^2 k_+}\frac{1}{u} + i\frac{(2R)^{-2}(\widehat{k}_3-\widehat{l})^2 + \vec{k}^2 + r}{2k_+}u}, \tag{6.150}$$

which shows in a clear way that there is a non compact direction which allows a distributional interpretation as discussed in [99]. However this direction cannot be easily decoupled from the compact one.

6.7 Comments on the BO

In this section we would like to quickly show the analysis performed in the previous sections for the NBO but in the case of the BO. The results are not very different apart from the fact that divergences are milder. It is in fact possible to construct the full scalar QED but nevertheless it is impossible to consider higher derivative terms in the effective theory. Moreover some three point amplitudes with a massive state diverge.

6.7.1 Geometric Preliminaries

In $\mathscr{M}^{1,1}$ we consider the change of coordinates:

$$\begin{cases} x^+ = t\, e^{+\Delta\varphi} \\ x^- = \sigma_-\, t\, e^{-\Delta\varphi} \end{cases} \qquad \Leftrightarrow \qquad \begin{cases} t = \text{sign}(x^+)\sqrt{|x^+x^-|} \\ \varphi = \frac{1}{2\Delta}\log\left|\frac{x^+}{x^-}\right| \\ \sigma_- = \text{sign}(x^+x^-) \end{cases} \tag{6.151}$$

where $\sigma_- = \pm 1$ and $t,\ \varphi \in \mathbb{R}$. The metric reads:

$$\begin{aligned} \mathrm{d}s^2 &= -2\ \mathrm{d}x^+\,\mathrm{d}x^+ \\ &= -2\sigma_-\left(\mathrm{d}t^2 - (\Delta t)^2\ \mathrm{d}\varphi^2\right). \end{aligned} \tag{6.152}$$

Its determinant is:

$$-\det g = 4\Delta^2 t^2. \tag{6.153}$$

In orbifold coordinates the non vanishing Christoffel symbols are:

$$\Gamma_{\varphi\ \varphi}^{\ t} = \Delta^2 t, \qquad \Gamma_{t\ \varphi}^{\ \varphi} = t^{-1}. \tag{6.154}$$

Using the orbifold coordinates $(t,\, \varphi)$, the BO is obtained by requiring the identification $\varphi \equiv \varphi + 2\pi$ along the orbit of the global Killing vector $\kappa_\varphi = 2\pi\partial_\varphi$. We will therefore use the recurrent parameter $\Lambda = e^{2\pi\Delta}$ as shorthand notation.

6.7.2 Free Scalar Action

The action for a complex scalar ϕ is given by

$$\begin{aligned} \mathrm{S}_{\mathrm{SQED}}^{(\mathrm{kinetic})}[\phi] &= \int_\Omega \mathrm{d}^D x\ \sqrt{-\det g}\big(-g^{\mu\nu}\,\partial_\mu\phi^*\,\partial_\nu\phi - M^2\phi^*\phi\big) \\ &= \sum_{\sigma_-\in\{\pm1\}} \int_{\mathbb{R}^{D-2}} \mathrm{d}^{D-2}\vec{x} \int_{-\infty}^{+\infty} \mathrm{d}t \int_0^{2\pi} \mathrm{d}\varphi\ \Delta|t| \\ &\times \left(\frac{1}{2}\sigma_-\,\partial_t\phi^*\,\partial_t\phi\ + \frac{1}{2}\,\frac{\sigma_-}{(\Delta t)^2}\partial_\varphi\phi^*\,\partial_\varphi\phi\ - \eta^{ij}\,\partial_i\phi^*\partial_j\phi - M^2\phi^*\phi\right). \end{aligned} \tag{6.155}$$

As before we solve the associated eigenfunction problem for the d'Alembertian operator

$$\left(-\frac{1}{2}\sigma_-\,\partial_t^2 - \frac{1}{2}\sigma_-\,\frac{1}{t}\partial_t + \frac{1}{2}\sigma_-\,\frac{1}{(\Delta\, t)^2}\partial_\varphi^2 + \partial_i^2\right)\phi_r = r\phi_r. \tag{6.156}$$

with

$$r = 2\,k_+k_- - \vec{k}^2 = 2\zeta_- m^2 - \vec{k}^2 \tag{6.157}$$

where for later convenience (see the transformation of k under the induced action of the Killing vector (6.167)) we parameterise the momenta as:

$$\begin{cases} k_+ &= m\, e^{+\Delta\beta} \\ k_- &= \zeta_- m\, e^{-\Delta\beta} \end{cases} \quad \Leftrightarrow \quad \begin{cases} m &= \mathrm{sign}(k_+)\,\sqrt{|k_+k_-|} \\ \beta &= \frac{1}{2\Delta}\log\left|\frac{k_+}{k_-}\right| \\ \zeta_- &= \mathrm{sign}(k_+k_-) \end{cases} \tag{6.158}$$

where $\zeta_- = \pm 1$ and $m,\, \beta \in \mathbb{R}$. To solve the problem we use standard techniques and perform the Fourier transform with respect to $\vec{x}$ and ϕ as :

$$\phi(t,\, \varphi,\, \vec{x}) = \sum_{l=-\infty}^{+\infty} \int_{\mathbb{R}^{D-2}} \mathrm{d}^{D-2}\vec{x}\ e^{i\,\vec{k}\cdot\vec{x}} e^{i\,l\,\varphi} H_{\{l,\,\vec{k},\,r,\,\sigma_-\}}(t), \tag{6.159}$$

so that the new function $H_{\{l,\,\vec{k},\,r,\,\sigma_-\}}$ satisfies

$$\partial_t^2 H_{\{l,\,\vec{k},\,r,\,\sigma_-\}} + \frac{1}{t}\partial_t H_{\{l,\,\vec{k},\,r,\,\sigma_-\}} + \left[\frac{l^2}{(\Delta\, t)^2} + 2\sigma_-\Big(r + \vec{k}^2\Big)\right] H_{\{l,\,\vec{k},\,r,\,\sigma_-\}} = 0, \tag{6.160}$$

which, upon the introduction of the natural quantities (see also (6.169) for an explanation of the naturalness of λ)

$$\tau = m\,t, \qquad \lambda = e^{\Delta(\varphi+\beta)}, \qquad \widehat{\sigma}_- = \sigma_-\,\zeta_-, \tag{6.161}$$

shows that the actual dependence on parameters is

$$H_{\{l,\,\vec{k},\,r,\,\sigma_-\}}(t) = \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau), \tag{6.162}$$

so that

$$\partial_\tau^2\widetilde{\phi}_{l\,\widehat{\sigma}_-} + \frac{1}{\tau}\partial_\tau\widetilde{\phi}_{l\,\widehat{\sigma}_-} + \left[\frac{l^2}{(\Delta\,\tau)^2} + 4\widehat{\sigma}_-\right]\widetilde{\phi}_{l\,\widehat{\sigma}_-} = 0. \tag{6.163}$$

The asymptotic behaviour of the solutions is:

$$\widetilde{\phi}_{l\,\widehat{\sigma}_-} \sim \begin{cases} A_+|\tau|^{i\,\frac{l}{\Delta}} + A_-|\tau|^{-i\,\frac{l}{\Delta}} & \text{for} \quad l \neq 0 \\ A_+\log|\tau| + A_- & \text{for} \quad l = 0 \end{cases}. \tag{6.164}$$

6.7.3 Eigenmodes on BO from Covering Space

We now repeat the essential part of the analysis performed in the NBO case. As on the NBO we say "wave function" and not eigenfunction since eigenfunctions for non scalar states require some constraints on polarisations which we do not impose.

Scalar Wave Function We start as usual from the Minkowskian wave function and we write only the dependence on x^+ and x^- since all the other coordinates are spectators

$$\begin{aligned} \psi_{k_+\,k_-}\left(x^+,\,x^-\right) &= e^{i\left(k_+x^+ + k_-x^-\right)} \\ &= e^{i\,m\,t\left[e^{\Delta(\varphi+\beta)} + \widehat{\sigma}_-\,t\,e^{\Delta(\varphi-\beta)}\right]} \\ &= \psi_{k_+\,k_-}(t,\,\varphi,\,\sigma_-). \end{aligned} \tag{6.165}$$

We can compute the wave function on the orbifold by summing over all images:

$$\begin{aligned} \Psi_{[k_+\,k_-]}\left(\left[x^+,\,x^-\right]\right) &= \sum_{n=-\infty}^{+\infty} \psi_{k_+\,k_-}\left(\mathcal{K}^n\left(x^+,\,x^-\right)\right) \\ &= \sum_{n=-\infty}^{+\infty} \psi_{k_+\,k_-}\left(x^+e^{2\pi\Delta\,n},\,x^-e^{-2\pi\Delta\,n}\right) \\ &= \sum_{n=-\infty}^{+\infty} e^{i\left\{\left[k_+e^{2\pi\Delta\,n}\right]x^+ + \left[k_-e^{-2\pi\Delta\,n}\right]x^-\right\}} \\ &= \sum_{n=-\infty}^{+\infty} \psi_{\mathcal{K}^{-n}(k_+\,k_-)}\left(x^+,\,x^-\right), \end{aligned} \tag{6.166}$$

where we write $[k_+\,k_-]$ because the function depends on the equivalence class of k_+k_- only. The equivalence relation is given by

$$k = \begin{pmatrix} k_+ \\ k_- \end{pmatrix} \equiv \mathcal{K}^{-n}k = \begin{pmatrix} k_+e^{2\pi\Delta\,n} \\ k_-e^{-2\pi\Delta\,n} \end{pmatrix}. \tag{6.167}$$

The previous equation explains the rationale for the parametrization (6.158) so that we can always choose a representative

$$0 \leq \beta < 2\pi, \qquad m \neq 0, \tag{6.168}$$

or differently said $\beta \equiv \beta + 2\pi$ and therefore we can use the dual quantum number l using a Fourier transform. Using the well adapted set of coordinates we can write the spin-0 wave function in a way to show the natural variables as

$$\Psi_{[k_+\, k_-]}\big([x^+,\, x^-]\big) = \sum_{n=-\infty}^{+\infty} e^{i\,\tau\left[\lambda e^{+2\pi\Delta\, n} + \widehat{\sigma}_-\lambda^{-1} e^{-2\pi\Delta\, n}\right]} = \widehat{\Psi}(\tau,\, \lambda,\, \widehat{\sigma}_-). \tag{6.169}$$

Again the scalar eigenfunction has a unique equivalence class which mixes coordinates and momenta.

Now we use the basic trick used in Poisson resummation

$$\begin{aligned}
\Psi_{[k_+\, k_-]}\big([x^+,\, x^-]\big) &= \int_{-\infty}^{+\infty} \mathrm{d}s\, \delta_P(s)\, e^{i\left\{k_+ x^+ \Lambda^s + k_- x^- \Lambda^{-s}\right\}} \\
&= \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} \left|\frac{k_+ x^+}{k_- x^-}\right|^{-i\frac{l}{2\Delta}} \int_{-\infty}^{+\infty} \mathrm{d}s\, e^{i\,2\pi\, l\, s} \\
&\times e^{i\,\mathrm{sign}(k_+\, x^+)\sqrt{|k_+ k_- x^+ x^-|}\left\{\Lambda^s + \sigma_-\zeta_-\Lambda^{-s}\right\}} \\
&= \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} \left(e^{\Delta(\varphi+\beta)}\right)^{-i\frac{l}{\Delta}} \int_{-\infty}^{+\infty} \mathrm{d}s\, e^{i\,2\pi\, l\, s} e^{i\, m t\left\{\Lambda^s + \sigma_-\zeta_-\Lambda^{-s}\right\}} \\
&= \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} e^{i\, l\beta} \left[e^{i\, l\varphi} \int_{-\infty}^{+\infty} \mathrm{d}s\, e^{-i\,2\pi l s} e^{i\, m t\left\{\Lambda^s + \sigma_-\zeta_-\Lambda^{-s}\right\}}\right],
\end{aligned} \tag{6.170}$$

where the last line represents the change of quantum number from $m\,\beta$ to $m\,l$ and allows us to identify

$$\mathcal{N}_{\mathrm{BO}}\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}s\, e^{-i\,2\pi\, l\, s} e^{i\,\tau\left\{\Lambda^s + \widehat{\sigma}_-\Lambda^{-s}\right\}}, \tag{6.171}$$

where $\mathcal{N}_{\mathrm{BO}}$ is a constant which depends on the normalization chosen for $\widetilde{\phi}_{l\,\widehat{\sigma}_-}$. This expression gives an integral representation of the O.D.E. solutions.

Tensor Wave Function (Spin-2) We consider the tensor wave function in Minkowski space. We focus on x^+, x^- and x^2 since all other directions behave as x^2. Differently from scalar function we need to keep the dependence on x^2 since it is needed for non trivial physical polarisations

and it enters in the transversality conditions. Explicitly we find

$$\begin{aligned}\mathcal{N}_{\rm BO}\psi^{[2]}_{k\,S}\big(x^+,\,x^-,\,x^2\big) &= S_{\mu\nu}\;\mathrm{d}x^\mu\,\mathrm{d}x^\nu\;\psi_k(x)\\ &= \Big[S_{++}\left(\mathrm{d}x^+\right)^2 + 2\,S_{+-}\;\mathrm{d}x^+\,\mathrm{d}x^- + 2\,S_{+2}\;\mathrm{d}x^+\,\mathrm{d}x^2\\ &+ S_{--}\,(\mathrm{d}x^-)^2 + 2\,S_{-2}\;\mathrm{d}x^2\,\mathrm{d}x^2\\ &+ S_{22}\,(\mathrm{d}x^2)^2\Big]\,e^{i\left(k_+x^+ + k_-x^- + k_2x^2\right)},\end{aligned} \tag{6.172}$$

which we rewrite in orbifold coordinates

$$\begin{aligned}\mathcal{N}_{\rm BO}\,\psi^{[2]}_{k\,S}\big(t,\,\varphi,\,x^2,\,\sigma_-\big) &= S_{\alpha\beta}\;\mathrm{d}x^\alpha\,\mathrm{d}x^\beta\;\psi_k(x)\\ &\times\Big[\mathrm{d}t^2\;\left(2\,S_{+-}\,\sigma_- + S_{++}\,e^{2\,\Delta\,\varphi} + S_{--}\,e^{-2\,\Delta\,\varphi}\right)\\ &+ 2\,\Delta\,t\;\mathrm{d}t\;\mathrm{d}\varphi\;\left(S_{++}\,e^{2\,\Delta\,\varphi} - S_{--}\,e^{-2\,\Delta\,\varphi}\right)\\ &+ \Delta^2\,t^2\,\mathrm{d}\varphi^2\;\left(-2\,S_{+-}\,\sigma_- + S_{++}\,e^{2\,\Delta\,\varphi} + S_{--}\,e^{-2\,\Delta\,\varphi}\right)\\ &+ 2\;\mathrm{d}t\;\mathrm{d}x^2\;\left(S_{-2}\,e^{-\Delta\,\varphi}\,\sigma_- + S_{+2}\,e^{\Delta\,\varphi}\right)\\ &+ 2\Delta\,t\;\mathrm{d}x^2\;\mathrm{d}\varphi\;\left(S_{+2}\,e^{\Delta\,\varphi} - S_{-2}\,e^{-\Delta\,\varphi}\,\sigma_-\right)\\ &+ \left(\mathrm{d}x^2\right)^2\,S_{2\,2}\Big]\,e^{i\,m\,t\left[e^{\Delta(\varphi+\beta)} + \widehat{\sigma}_-\,e^{\Delta(\varphi-\beta)}\right] + i\,k_2x^2}.\end{aligned} \tag{6.173}$$

Now we define the tensor wave on the orbifold as a sum over all images as

$$\begin{aligned}\mathcal{N}_{\rm BO}\,\Psi^{[2]}_{[k\,S]}([x]) &= \sum_{n=-\infty}^{+\infty}\;(\mathcal{K}^n\;\mathrm{d}x)\cdot S\cdot(\mathcal{K}^n\;\mathrm{d}x)\,\psi_k(\mathcal{K}^n x)\\ &= \sum_{n=-\infty}^{+\infty}\mathrm{d}x\cdot\left(\mathcal{K}^{-n}\,S\right)\cdot\mathrm{d}x\;\psi_{\mathcal{K}^{-n}\,k}(x).\end{aligned} \tag{6.174}$$

In the last line we have defined the induced action of the Killing vector on $(k,\,S)$ which can be explicitly written as:

$$\mathcal{K}^{-n}\begin{pmatrix}S_{++}\\ S_{+-}\\ S_{--}\\ S_{+2}\\ S_{-2}\\ S_{22}\end{pmatrix} = \begin{pmatrix}e^{2n\Delta\varphi}\,S_{++}\\ S_{+-}\\ e^{-2n\Delta\varphi}\,S_{--}\\ e^{n\Delta\varphi}\,S_{+2}\\ e^{-n\Delta\varphi}\,S_{-2}\\ S_{22}\end{pmatrix}, \tag{6.175}$$

and it amounts to a trivial scaling. In orbifold coordinates computing the tensor wave simply amounts to sum over all the shifts $\varphi\to\varphi+2\pi n$. Then we have to give a close expression for the sum involving powers $e^{2\pi\Delta n}$. Explicitly we find:

$$\begin{aligned}&\sum_{n=-\infty}^{+\infty}\left(e^{2\pi\Delta n}\right)^N e^{i\,\tau\left[\lambda e^{+2\pi\Delta\,n} + \widehat{\sigma}_-\,\frac{1}{\lambda}e^{-2\pi\Delta\,n}\right]}\\ &= \begin{cases}\left[\frac{1}{2}\left(\frac{1}{\lambda}\partial_\tau + \frac{1}{\tau}\partial_\lambda\right)\right]^N\widehat{\Psi}(\tau,\,\lambda,\,\widehat{\sigma}_-) & \text{for}\quad N>0\\ \left[\frac{1}{2}\left(\lambda\partial_\tau - \frac{\lambda^2}{\tau}\partial_\lambda\right)\right]^N\widehat{\Psi}(\tau,\,\lambda,\,\widehat{\sigma}_-) & \text{for}\quad N<0\end{cases},\end{aligned} \tag{6.176}$$

where τ derivatives of $\widetilde{\phi}_{l\,\widehat{\sigma}_-}$ of order higher than 2 can be reduced with the help of the differential equation (6.163).

We now have to identify the basic polaritazions on the orbifold. However the quantum number β is no longer a good quantum number on the orbifold and it is replaced by l. The relations among orbifold polarisations and Minkowski polarisations may depend on β as long as the traceless and transversality conditions on the orbifold are independent of it.[51] Finally it seems reasonable to use the natural variable $\lambda = e^{\Delta(\varphi+\beta)}$. Therefore we have:

$$\begin{aligned}
\mathcal{S}_{t\,t} &= e^{-2\,\Delta\,\beta}\, S_{++},\\
\mathcal{S}_{t\,\varphi} &= S_{+-},\\
\mathcal{S}_{t\,2} &= e^{-\Delta\,\beta}\, S_{+\,2},\\
\mathcal{S}_{\varphi\,\varphi} &= e^{2\,\Delta\,\beta}\, S_{--},\\
\mathcal{S}_{\varphi\,2} &= e^{\Delta\,\beta}\, S_{-\,2},\\
\mathcal{S}_{2\,2} &= S_{2\,2},
\end{aligned} \tag{6.177}$$

which can be trivially inverted as

$$\begin{aligned}
S_{++} &= e^{2\,\Delta\,\beta}\, \mathcal{S}_{t\,t},\\
S_{+-} &= \mathcal{S}_{t\,\varphi},\\
S_{+\,2} &= e^{\Delta\,\beta}\, \mathcal{S}_{t\,2},\\
S_{--} &= e^{-2\,\Delta\,\beta}\, \mathcal{S}_{\varphi\,\varphi},\\
S_{-\,2} &= e^{-\Delta\,\beta}\, \mathcal{S}_{\varphi\,2},\\
S_{2\,2} &= \mathcal{S}_{2\,2}.
\end{aligned} \tag{6.178}$$

We can then compute the trace:

$$\mathrm{tr}(S) = -2\,\mathcal{S}_{t\,\varphi} + \mathcal{S}_{2\,2}, \tag{6.179}$$

while the transversality conditions become

$$\begin{aligned}
(k\cdot S)_+ &= -e^{\Delta\,\beta}\,(m\,\widehat{\sigma}_-\,\sigma_-\,\mathcal{S}_{t\,t} + m\,\mathcal{S}_{t\,\varphi} - k_2\,\mathcal{S}_{t\,2})\\
(k\cdot S)_- &= -e^{-\Delta\,\beta}\,(m\,\widehat{\sigma}_-\,\sigma_-\,\mathcal{S}_{t\,\varphi} + m\,\mathcal{S}_{\varphi\,\varphi} - k_2\,\mathcal{S}_{\varphi\,2})\\
(k\cdot S)_2 &= -(m\,\widehat{\sigma}_-\,\sigma_-\,\mathcal{S}_{t\,2} + m\,\mathcal{S}_{\varphi\,2} - k_2\,\mathcal{S}_{2\,2}),
\end{aligned} \tag{6.180}$$

which are independent from β when it is set to zero.

The final expression of the wave function for the symmetric tensor on the orbifold is:

$$\begin{aligned}
\Psi^{[2]}_{[k\,S]}([x]) = \sum_{l=-\infty}^{+\infty} e^{i\,l\,\beta}\Big[& S_{m\,l,\,tt}\ \mathrm{d}t^2 + 2\,S_{m\,l,\,t\varphi}\ \mathrm{d}t\,\mathrm{d}\varphi + 2\,S_{m\,l,\,t2}\ \mathrm{d}t\,\mathrm{d}x^2\\
& + S_{m\,l,\,\varphi\varphi}\ \mathrm{d}\varphi^2 + 2\,S_{m\,l,\,\varphi 2}\ \mathrm{d}\varphi\,\mathrm{d}x^2\\
& + S_{m\,l,\,22}\ \mathrm{d}x^{2^2}\Big],
\end{aligned} \tag{6.181}$$

[51] These conditions may be a linear combinations of the ones in Minkowski.

where the explicit expressions for the components are

$$\begin{aligned} S_{m\,l,\,tt} = & \left[-\frac{\widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, l\, \lambda^{\frac{i\,l}{\Delta}}\, (l\,\mathcal{S}_{t\,t} + i\,\Delta\,\mathcal{S}_{t\,t} + l\,\mathcal{S}_{\varphi\,\varphi} - i\,\Delta\,\mathcal{S}_{\varphi\,\varphi})}{2\,\Delta^2} \right] \frac{1}{\tau^2} \\ & + \left[\frac{1}{2\,\Delta} \frac{\mathrm{d}}{\mathrm{d}\tau}\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, \lambda^{\frac{i\,l}{\Delta}}\, (i\,l\,\mathcal{S}_{t\,t} - i\,l\,\mathcal{S}_{\varphi\,\varphi} - \Delta\,\mathcal{S}_{t\,t} - \Delta\,\mathcal{S}_{\varphi\,\varphi}) \right] \frac{1}{\tau} \\ & + \left[\widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, \lambda^{\frac{i\,l\,\Delta}{}}\, (\widehat{\sigma}_-\,\mathcal{S}_{t\,t} + 2\,\sigma_-\,\mathcal{S}_{t\,\varphi} + \widehat{\sigma}_-\,\mathcal{S}_{\varphi\,\varphi}) \right], \end{aligned} \tag{6.182}$$

$$\begin{aligned} S_{m\,l,\,t\varphi} = & \left[-\frac{\widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, l\, \lambda^{\frac{i\,l\,\Delta}{}}\, (l\,\mathcal{S}_{t\,t} + i\,\Delta\,\mathcal{S}_{t\,t} - l\,\mathcal{S}_{\varphi\,\varphi} + i\,\Delta\,\mathcal{S}_{\varphi\,\varphi})}{2\,\Delta\,m} \right] \frac{1}{\tau} \\ & + \left[\frac{\frac{\mathrm{d}}{\mathrm{d}\tau}\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, \lambda^{\frac{i\,l\,\Delta}{}}\, (i\,l\,\mathcal{S}_{t\,t} - \Delta\,\mathcal{S}_{t\,t} + i\,l\,\mathcal{S}_{\varphi\,\varphi} + \Delta\,\mathcal{S}_{\varphi\,\varphi})}{2\,m} \right] \\ & + \left[\frac{\Delta\,\widehat{\sigma}_-\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, \lambda^{\frac{i\,l\,\Delta}{}}\, (\mathcal{S}_{t\,t} - \mathcal{S}_{\varphi\,\varphi})}{m} \right] \tau, \end{aligned} \tag{6.183}$$

$$\begin{aligned} S_{m\,l,\,\varphi\,\varphi} = & \left[-\frac{1}{2\,m^2} \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, l\, \lambda^{\frac{i\,l\,\Delta}{}}\, (l\,(\mathcal{S}_{t\,t} + \mathcal{S}_{\varphi\,\varphi}) + i\,\Delta\,(\mathcal{S}_{t\,t} - \mathcal{S}_{\varphi\,\varphi})) \right] \\ & + \left[\frac{1}{2\,m^2} \Delta \left(\frac{\mathrm{d}}{\mathrm{d}\tau}\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau) \right) \lambda^{\frac{i\,l\,\Delta}{}}\, (i\,l\,\mathcal{S}_{t\,t} - i\,l\,\mathcal{S}_{\varphi\,\varphi} - \Delta\,\mathcal{S}_{t\,t} - \Delta\,\mathcal{S}_{\varphi\,\varphi}) \right] \tau \\ & + \left[\frac{1}{m^2} \Delta^2\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, \lambda^{\frac{i\,l\,\Delta}{}}\, (\widehat{\sigma}_-\,\mathcal{S}_{t\,t} + \widehat{\sigma}_-\,\mathcal{S}_{\varphi\,\varphi} - 2\,\sigma_-\,\mathcal{S}_{t\,\varphi}) \right] \tau^2, \end{aligned} \tag{6.184}$$

together with the effectively vector components in the orbifold directions:

$$\begin{aligned} S_{m\,l,\,t2} = & \left[\frac{i}{2\,\Delta} \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, l\, \lambda^{\frac{i\,l\,\Delta}{}}\, (\mathcal{S}_{t\,2} - \mathcal{S}_{\varphi\,2}\,\sigma_-) \right] \frac{1}{\tau} \\ & + \left[\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}\tau}\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, \lambda^{\frac{i\,l\,\Delta}{}}\, (\mathcal{S}_{t\,2} + \mathcal{S}_{\varphi\,2}\,\sigma_-) \right], \end{aligned} \tag{6.185}$$

and

$$\begin{aligned} S_{m\,l,\,\varphi 2} = & \left[\frac{i}{2\,m} \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, l\, \lambda^{\frac{i\,l\,\Delta}{}}\, (\mathcal{S}_{t\,2} + \mathcal{S}_{\varphi\,2}\,\sigma_-) \right] \\ & + \left[\frac{1}{2\,m} \Delta \left(\frac{d}{d\,\tau}\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau) \right) \lambda^{\frac{i\,l\,\Delta}{}}\, (\mathcal{S}_{t\,2} - \mathcal{S}_{\varphi\,2}\,\sigma_-) \right] \tau, \end{aligned} \tag{6.186}$$

and the effectively scalar component:

$$S_{m\,l,\,22} = \mathcal{S}_{2\,2}\, \widetilde{\phi}_{l\,\widehat{\sigma}_-}(\tau)\, \lambda^{\frac{i\,l}{\Delta}}. \tag{6.187}$$

6.8 Overlaps and Divergent Three Points String Amplitudes

We consider some overlaps as done for the NBO. The connection between the overlaps on the orbifold and the sums of images remains unchanged when we change the Killing vector $\mathcal{K}$, hence we can limit ourselves to discuss the integrals on the orbifold space.

6.8.1 Overlaps Without Derivatives

Let us start with the simplest case of the overlap of N scalar wave functions:

$$\begin{aligned} I^{(N)} &= \int_{\Omega} \mathrm{d}^3x\ \sqrt{-\det g}\prod_{i=1}^{N}\Psi_{\left[k_{(i)\,+}\,k_{(i)\,-}\right]}(\left[x^+,\,x^-,\,x^2\right])) \\ &= \mathcal{N}_{\mathrm{BO}}^{N}\sum_{\left\{l_{(i)}\right\}\in\mathbb{Z}^N} e^{\,i\sum_{i=1}^{N} l_{(i)}\beta_{(i)}}\int_{\Omega}\mathrm{d}^3x\ \sqrt{-\det g}\prod_{i=1}^{N}\phi_{l_{(i)}\,\widehat{\sigma}_{-\,(i)}}. \end{aligned} \tag{6.188}$$

This is always a distribution since the problematic $l_{(*)}=0$ sector gives a divergence $(\log|t|)^N$ when $t\sim 0$. All other sectors have no issues because of the asymptotic behaviours (6.164).

6.8.2 An Overlap With Two Derivatives

We consider in orbifold coordinates the overlap needed for the amplitude involving two tachyons and one massive state, i.e.:

$$K=\int_{\Omega}\mathrm{d}^3x\ \sqrt{-\det g}\,g^{\alpha\beta}\,g^{\gamma\delta}\,\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]\,\alpha\gamma}([x])\,D_\beta\partial_\delta\Psi_{\left[k_{(2)}\right]}([x])\,\Psi_{\left[k_{(1)}\right]}([x]). \tag{6.189}$$

Since we use the traceless condition we need to keep all momenta and polarisations. We write:

$$\begin{aligned} K=\int_{\Omega}\mathrm{d}^3x\ \sqrt{-\det g}\Bigg[&\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]\,tt}\,\partial_t^2\Psi_{\left[k_{(2)}\right]} \\ &-2\left(\frac{1}{\Delta t}\right)^2\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]\,t\varphi}\left(\partial_t\partial_\varphi\Psi_{\left[k_{(2)}\right]}-\frac{1}{t}\partial_\varphi\Psi_{\left[k_{(2)}\right]}\right) \\ &+\left(\frac{1}{\Delta t}\right)^4\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]\,\varphi\varphi}\left(\partial_\varphi^2\Psi_{\left[k_{(2)}\right]}-\Delta^2 t\partial_t\Psi_{\left[k_{(2)}\right]}\right) \\ &-2\,\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]\,t2}\,\partial_t\partial_2\Psi_{\left[k_{(2)}\right]} \\ &+2\left(\frac{1}{\Delta t}\right)^2\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]\,\varphi 2}\,\partial_\varphi\partial_2\Psi_{\left[k_{(2)}\right]} \\ &+\Psi^{[2]}_{\left[k_{(3)},S_{(3)}\right]\,22}\,\partial_2^2\Psi_{\left[k_{(2)}\right]}\Bigg]\Psi_{\left[k_{(1)}\right]}. \end{aligned} \tag{6.190}$$

Now consider the behaviour for $l_{(*)}=0$ for small t. All the ∂_φ can be dropped since they lower a $l_{(2)}$. The leading contributions from spin-2 components are $S_{m\,l\,tt}\sim\frac{1}{t^2}$, $S_{m\,l\,\varphi\varphi}, S_{m\,l\,2\,2}\sim 1$ and $S_{m\,l\,t2}\sim\frac{1}{t}$. The leading $\frac{1}{t^4}$ reads:

$$\begin{aligned} K\sim\int_{t\sim 0}\mathrm{d}t\ |t|\Bigg[&-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}\widetilde{\phi}_{l\,\widehat{\sigma}_-}\,(\mathcal{S}_{t\,t}+\mathcal{S}_{\varphi\,\varphi})\frac{1}{\tau}\,\partial_t^2\Psi_{\left[k_{(2)}\right]} \\ &+\left(\frac{1}{\Delta t}\right)^4\frac{-\Delta^2}{2\,m^2}\frac{\mathrm{d}}{\mathrm{d}\tau}\widetilde{\phi}_{l\,\widehat{\sigma}_-}\,(\mathcal{S}_{t\,t}+\mathcal{S}_{\varphi\,\varphi})\tau\left(-\Delta^2 t\partial_t\Psi_{\left[k_{(2)}\right]}\right)\Bigg]\Psi_{\left[k_{(3)}\right]} \end{aligned} \tag{6.191}$$

In the limit of our interest $\Psi_{[k]}\big|_{l=0} \sim \widetilde{\phi}_{l\,\widehat{\sigma}_-}\Big|_{l=0} \sim \log|t|$. The two terms add together because of sign of the covariant derivative to give:

$$K \sim \int\limits_{t\sim 0} \mathrm{d}t\ |t| \left[\left(\frac{1}{2}+\frac{1}{2}\right)\frac{\mathcal{S}_{t\,t}+\mathcal{S}_{\varphi\,\varphi}}{m^4}\frac{\log|t|}{t^4}+\mathcal{O}\left(\frac{(\log|t|)}{t}\right)\right], \tag{6.192}$$

which is divergent for the physical polarisation $\mathcal{S}_{t\,t}=\mathcal{S}_{\varphi\,\varphi}=-\widehat{\sigma}_-\sigma_-\mathcal{S}_{t\,\varphi}=-\frac{1}{2}\widehat{\sigma}_-\sigma_-\mathcal{S}_{22}$.

7 Summary and Conclusion

From the previous analysis it seems that string theory cannot do better than field theory when the latter does not exist, at least at the perturbative level where one deals with particles. Moreover when spacetime becomes singular, the string massive modes are not spectators anymore. Everything seems to suggest that issues with spacetime singularities are hidden into contact terms and interactions with massive states. This would explain in an intuitive way why the eikonal approach to gravitational scattering works well: it is indeed concerned with three point massless interactions. In fact it appears that the classical and quantum scattering on an electromagnetic wave [103] or gravitational wave [104] in BO and NBO are well behaved. From this point of view the ACV approach [105], [106] may be more sensible, especially when considering massive external states [107]. Finally it seems that all issues are related with the Laplacian associated with the space-like subspace with vanishing volume at the singularity. As a matter of fact if there is a discrete zero eigenvalue the theory automatically develops divergences.

Part III

Deep Learning the Geometry of String Theory

8 Introduction

In the previous parts we presented mathematical tools for the theoretical interpretation of amplitudes in field theory and string theory. The ultimate goal of the analysis is to provide some insights on the predictive capabilities of the string theory framework applied to (semi-)phenomenological data. As already argued in Section 1.3 the procedure is however quite challenging as there are different ways to match string theory with the experimental reality. There are in fact several different vacuum configurations arising from the compactification of the extra-dimensions. The investigation of feasible phenomenological models in a string framework has therefore to deal also with computational aspects related to the exploration of the *landscape* [108] of possible vacua. Unfortunately the number of possibilities is huge (numbers as high as $10^{272\,000}$ have been suggested for some models) [108]–[112], the mathematical objects entering the compactifications are complex and typical problems are often NP-complete, NP-hard, or even undecidable [113], [114], making an exhaustive classification impossible. Additionally there is no single framework to describe all the possible (flux) compactifications. As a consequence each class of models must be studied with different methods. This has in general discouraged, or at least rendered challenging or unfeasible, precise connections to the existing and tested theories (for instance the SM of particle physics).

Until recently the string landscape has been studied using different methods such as analytic computations for simple examples, general statistics, random scans or algorithmic enumerations of possibilities. This has been a large endeavor of the string community (see for instance [14], [115] and references therein). The main objective of such studies is to understand what are the generic predictions of string theory. The immediate conclusions are that compactifications giving an effective theory close to the Standard Model are scarce [116]–[119]. The approach however has limitations mainly given by lack of a general understanding or high computational power required to run the algorithms.

In reaction to these difficulties, and starting with the seminal paper by Abel and Rizos in 2014 [120], new investigations based on Machine Learning (ML) appeared in the recent years. The have been focusing on different aspects of the string landscape and of the geometries used in compactifications [121]–[139] (see [140] for a comprehensive summary of the state of the art). In fact ML is definitely adequate when it comes to pattern search or statistical inference starting from large amount of data. This motivates two main applications to string theory: the systematic exploration of the space of possibilities (if they are not random then ML should be able to find a pattern) and the deduction of mathematical formulas from the ML approximation. The last few years have seen a major uprising of ML, and more particularly of neural networks (NN) [141]–[143]. This technology is efficient at discovering and predicting patterns and now pervades most fields of applied sciences and of the industry. One of the most critical places where progress can be expected is in understanding the geometries used to describe string compactifications and this will be the object of study in the following analysis. We mainly refer to [141]–[143] for reviews in ML and deep learning techniques, and to [140], [144], [145] for applications of data science techniques.

We address the question of computing the Hodge numbers $h^{1,1} \in \mathbb{N}$ and $h^{2,1} \in \mathbb{N}$ for *complete intersection Calabi–Yau* (CICY) 3-folds [146] using different ML algorithms. A CICY is completely specified by its *configuration matrix* (whose entries are positive integers) which is the basic input of the algorithms. The CICY 3-folds are the simplest manifolds of their kind and

they have been well studied. In particular they have been completely classified and their topological properties computed [147]–[149]. For these reasons, they provide an excellent sandbox to test ML algorithms in a controlled environment.

The goal is therefore to predict two positive integers from a matrix of positive integers. This task is complicated by various redundancies in the description (such as an independence in the permutations of lines and columns). While usual physics application of ML reduces to feeding a (big) sequential neural network with raw data, real-world applications are built following a more general pipeline [143], [144]. In fact the first task after understanding of the problem would be to perform and exploratory data analysis (EDA) to highlight possible data which may help in getting a result. After the definition of a validation strategy, feature engineering can be used to improve the baseline computations and improve the design of ML models. While the first step is straightforward it is still interesting to notice that computations involved in string geometries (using algebraic topology) are far from standard applications of ML algorithms, which makes the problem even more interesting. EDA aims at better understanding the dataset showing how the variables are distributed, correlated, determining if outliers are present, etc. This analysis naturally leads to designing new variables during *feature engineering* which can be used in addition (or even substitution) of the original data. Adding derived features by hand may make the data more easily understandable by the ML algorithms for instance by emphasizing important properties.[52] This phase is followed by *feature selection*, where different set of features are chosen according to the need of each algorithm. A good validation strategy is also needed to ensure that the predictions appropriately reflect the real values, together with a baseline model, which gives a lower bound on the accuracy together with a working pipeline.[53] For instance, we find that a simple linear regression using the configuration matrix as input gives 43.6 % to 48.8 % for $h^{1,1}$ and 9.6 % to 10.4 % for $h^{2,1}$ using from 20% to 80% of data for training. Hence any algorithm *must* do better than this to be worth considering.

The datasets we use for task contains 7890 CICY 3-folds. Due to the freedom in representing the configuration matrix, we need to consider two datasets which have been constructed: the *original dataset* [147], [148] and the *favourable dataset* [149]. Our analysis continues and generalises [123], [126] at different levels. For example we compute $h^{2,1}$ which has been ignored in [123], [126], where the authors argue that it can be computed from $h^{1,1}$ and from the Euler characteristics (a simple formula exists for the latter). In our case, we want to push the idea of using ML to learn about the physics (or the mathematics) of CY to its very end: we assume that we do not know anything about the mathematics of the CICY, except that the configuration matrix is sufficient to derive all quantities. Moreover we have already mentioned that ML algorithms have rarely been used to derive data in algebraic topology, which can be a difficult task. Thus getting also $h^{2,1}$ from ML techniques is an important first step towards using ML for more general problems in string geometries. Finally regression is also more useful for extrapolating results: a classification approach assumes that we already know all the possible values of the Hodge numbers and has difficulties to predict labels which do not appear in the training set. This is necessary when we move to a dataset for which not all topological quantities have been computed, for instance CY constructed from the Kreuzer–Skarke list of polytopes [150].

The data analysis and ML are programmed in Python using known open-source packages such

[52] While one could expect ML algorithms to generate these features by themselves, this may complicate the learning process. So in cases where it is straightforward to compute meaningful derived features it is often worth considering them.

[53] For example the original work on this topic [123] did not set up a validation strategy and reported the accuracy over both the training and test data. Correcting this problem leads to an accuracy of 37% [126].

as `pandas` [151], `matplotlib` [152], `seaborn` [153], `scikit-learn` [154], `scikit-optimize` [155], `tensorflow` [156] (and its high level API *Keras*). Code is available on Github.

8.1 Complete Intersection Calabi–Yau Manifolds

As presented in Section 1.3, a CY n-fold is a n-dimensional complex manifold X with SU(n) holonomy (dimension in $\mathbb{R}$ is $2n$). An equivalent definition is the vanishing of its first Chern class. A standard reference for the physicist is [157] (see also [13] for useful references). The compactification on a CY leads to the breaking of large part of the supersymmetry which is phenomenologically more realistic than the very high energy description with intact supersymmetry.

CY manifolds are characterised by a certain number of topological properties (see Section 1.3.3), the most salient being the Hodge numbers $h^{1,1}$ and $h^{2,1}$, counting respectively the Kähler and complex structure deformations, and the Euler characteristics:[54]

$$\chi = 2\big(h^{1,1} - h^{2,1}\big). \tag{8.1}$$

Interestingly topological properties of the manifold directly translate into features of the 4-dimensional effective action (in particular the number of fields, the representations and the gauge symmetry) [157].[55] In particular the Hodge numbers count the number of chiral multiplets (in heterotic compactifications) and the number of hyper- and vector multiplets (in type II compactifications): these are related to the number of fermion generations (three in the Standard Model) and is thus an important measure of the distance to the Standard Model.

The simplest CY manifolds are constructed by considering the complete intersection of hypersurfaces in a product $\mathcal{A}$ of projective spaces $\mathbb{P}^{n_i}$ (called the ambient space) [146]–[149], [158]:

$$\mathcal{A} = \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}. \tag{8.2}$$

Such hypersurfaces are defined by homogeneous polynomial equations: a CICY manifold X is described by the solution to the system of equations, i.e. by the intersection of all these surfaces. The intersection is "complete" in the sense that the hypersurface is non-degenerate.

To gain some intuition, consider the case of a single projective space $\mathbb{P}^n$ with (homogeneous) coordinates Z^I, where $I = 0, 1, \ldots, n$. A codimension 1 subspace is obtained by imposing a single homogeneous polynomial equation of degree a on the coordinates:

$$\begin{aligned} p_a\big(Z^0, \ldots, Z^n\big) &= P_{I_1 \ldots I_a}\, Z^{I_1} \ldots Z^{I_a} = 0, \\ p_a\big(\lambda Z^0, \ldots, \lambda Z^n\big) &= \lambda^a\, p_a\big(Z^0, \ldots, Z^n\big). \end{aligned} \tag{8.3}$$

Each choice of the polynomial coefficients $P_{I_1 \ldots I_a}$ leads to a different manifold. However it can be shown that the manifolds are in general topologically equivalent. Since we are interested only in classifying the CY as topological manifolds and not as complex manifolds, the information on $P_{I_1 \ldots I_a}$ can be discarded and it is sufficient to keep track only of the dimension n of the projective space and of the degree a of the equation. The resulting hypersurface is denoted

[54] In full generality, the Hodge numbers $h^{p,q}$ count the numbers of harmonic (p, q)-forms.

[55] Another reason for sticking to topological properties is that there is no CY manifold for which the metric is known. Hence it is not possible to perform explicitly the Kaluza–Klein reduction in order to derive the 4-dimensional theory.

as $[\mathbb{P}^n \mid a] = [n \mid a]$. Notice that $[\mathbb{P}^n \mid a]$ is 3-dimensional if $n = 4$ (the equation reduces the dimension by one), and it is a CY (the "quintic") if $a = n + 1 = 5$ (this is required for the vanishing of its first Chern class). The simplest representative of this class if Fermat's quintic defined by the equation:

$$\sum_{I=0}^{4} \left(Z^I\right)^5 = 0. \tag{8.4}$$

This construction can be generalized to include m projective spaces and k equations which can mix the coordinates of the different spaces. A CICY 3-fold X as a topological manifold is completely specified by a *configuration matrix* denoted by the same symbol as the manifold:

$$X = \left[\begin{array}{c|ccc} \mathbb{P}^{n_1} & a_1^1 & \cdots & a_k^1 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{P}^{n_m} & a_1^m & \cdots & a_k^m \end{array}\right] \tag{8.5}$$

where the coefficients a_α^r are positive integers and satisfy the following constraints

$$\dim_{\mathbb{C}} X = \sum_{r=1}^{m} n_r - k = 3, \qquad n_r + 1 = \sum_{\alpha=1}^{k} a_\alpha^r, \quad \forall r \in \{1, 2, \ldots, m\}. \tag{8.6}$$

The first relation states that the difference between the dimension of the ambient space and the number of equations is the dimension of the CY 3-fold. The second set of constraints arises from the vanishing of its first Chern class. It implies that the n_i can be recovered from the matrix elements. Two manifolds described by the same configuration matrix but different polynomials are diffeomorphic as real manifold, and thus as topological manifolds, but they are different as complex manifolds. Hence it makes sense to write only the configuration matrix.

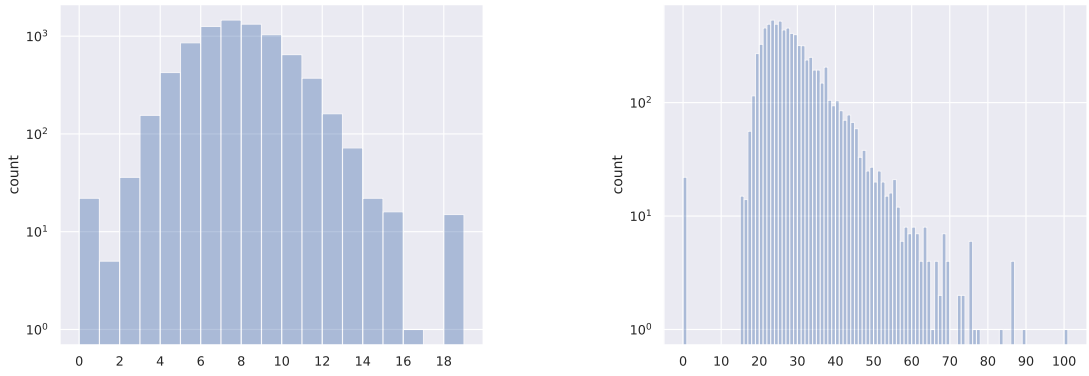
A given topological manifold is not described by a unique configuration matrix. First, any permutation of the lines and columns leave the intersection unchanged as it amounts to relabelling the projective spaces and equations. Secondly, two intersections can define the same manifold. The ambiguity in the line and column permutations is often fixed by imposing some ordering of the coefficients. Moreover there is an optimal representation of the manifold X, called *favourable* [149]: in such form topological properties of X can be more conveniently derived from the ambient space $\mathcal{A}$. Finally, simple arguments [146], [147], [159] show that the number of CICY is necessarily finite due to the constraints (8.6) together with identities between complete intersection manifolds.

8.2 Datasets

The classification of the CICY 3-folds has been tackled in [147]. The analysis established a dataset of 7890 CICY. The topological properties of each of these manifolds have been computed in [148]. More recently a new classification has been performed [149] in order to find the favourable representation of each manifold whenever it is possible.

Below we show a list of the CICY properties and of their configuration matrices:

- general properties:



(a) $h^{1,1}$ (b) $h^{2,1}$

Figure 8.1: Distribution of the Hodge numbers (log scale).

  - number of configurations: 7890
  - number of product spaces (block diagonal matrix): 22
  - $h^{11} \in [0, 19]$ with 18 distinct values (Figure 8.1a)
  - $h^{21} \in [0, 101]$ with 65 distinct values (Figure 8.1b)
  - unique Hodge number combinations: 266
- "original dataset" [147], [148]:
  - maximal size of the configuration matrices: 12×15
  - number of favourable matrices (excluding product spaces): 4874 (61.8%)
  - number of non-favourable matrices (excluding product spaces): 2994
  - number of different ambient spaces: 235
- "favourable dataset" [149]:
  - maximal size of the configuration matrices: 15×18
  - number of favourable matrices (excluding product spaces): 7820 (99.1%)
  - number of non-favourable matrices (excluding product spaces): 48
  - number of different ambient spaces: 126

We then move to the data science analysis of the data. To provide a good test case for the use of ML in context where the mathematical theory is not completely understood, we make no use of known formulas. In fact we try to push as far as possible the capabilities of ML algorithms to play a role in discovering patterns which can be used in phenomenology and algebraic geometry.

9 Machine Learning and Deep Learning for CICY Manifolds

In the following sections we present the preliminary analysis and the machine and deep learning study applied to the predictions of the Hodge numbers of CICY 3-folds. We use both a

"classical" approach to ML using *shallow learning* algorithm using geometrical methods to find new representations of the data and more moder approaches based on *computer vision* and recent developments in computer science techniques. We show how *deep learning* the geometry of string theory can help in providing good computational tools for phenomenology and algebraic geometry. We also stress future investigations which can be performed based on these results.

9.1 Exploratory Data Analysis

A typical ML project does not consist of feeding the raw data to the algorithm. It is instead preceded by a phase of exploration in order to better understand the data, which in turn can help to design the learning algorithms. We call *features* properties given as inputs, and *labels* the targets of the predictions. There are several phases in the exploratory data analysis (EDA) [144]:

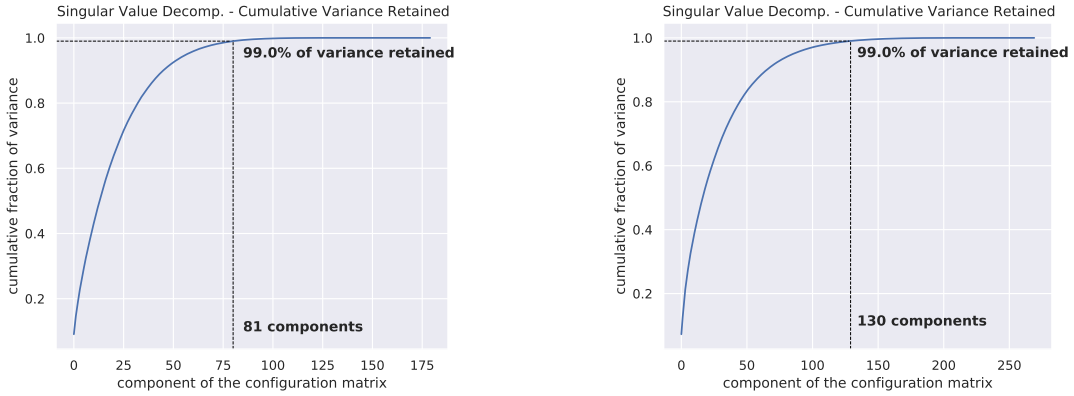
1. *feature engineering*: new features are derived from the inputs;
2. *feature selection*: the most relevant features are chosen to explain the targets;
3. *data augmentation*: new training data is generated from the existing ones;
4. *data diminution*: part of the training data is not used.

Engineered features are redundant by definition but they can help the algorithm learn more efficiently by providing an alternative formulation and by drawing attention on salient characteristics. A simple example is the following: given a series of numbers, one can compute different statistics, such as median, mean and variance, and add them to the inputs. It may happen that the initial series becomes then irrelevant once this new information is introduced. Another approach to improve the learning process is to augment or decrease the number of training samples artificially.[56] For example we could use invariances of the inputs to generate more training data. This however does not help in our case because the entries of the configuration matrices are partially ordered. Another possibility is to remove outliers which can damage the learning process by driving the algorithm far from the best solution. If there are few of them it is better to ignore them altogether during training since an algorithm which is not robust to outliers will in any case make bad predictions (a standard illustration is given by the Pearson and Spearman correlation coefficients, with the first not being robust to outliers [144]).

Before starting the EDA, the first step should be to split the data into training and validation sets to avoid biasing the choices of the algorithm and the strategy: the EDA should be performed only on the training set. However the dataset we consider is complete and quite uniform: a subset of it would display the same characteristics as the entire set.[57] To give a general overview of the properties we work with the full dataset.

[56] This is in general used in computer vision and object detection tasks where providing rotated, scaled and cropped versions of the input images can help the algorithms in learning more representations of the same object, thus creating more accurate predictions.

[57] A dataset is *tidy* if every column represents a separate variable and every row is a different observation. For instance every row could represent a date expressed in seconds from a reference instant and every column could be a separate sensor reading. However the "transposed" version of the dataset is not a tidy dataset because the observations are in the columns, thus representing a non standard form of the data. A tidy dataset is *complete* if there are no empty cells, that is there is no lack of data or information in the entire set. A *uniform* dataset can be understood as a complete dataset in which every variable is well distributed and does not present a lot of outliers or anomalies.



(a) original dataset

(b) favourable dataset

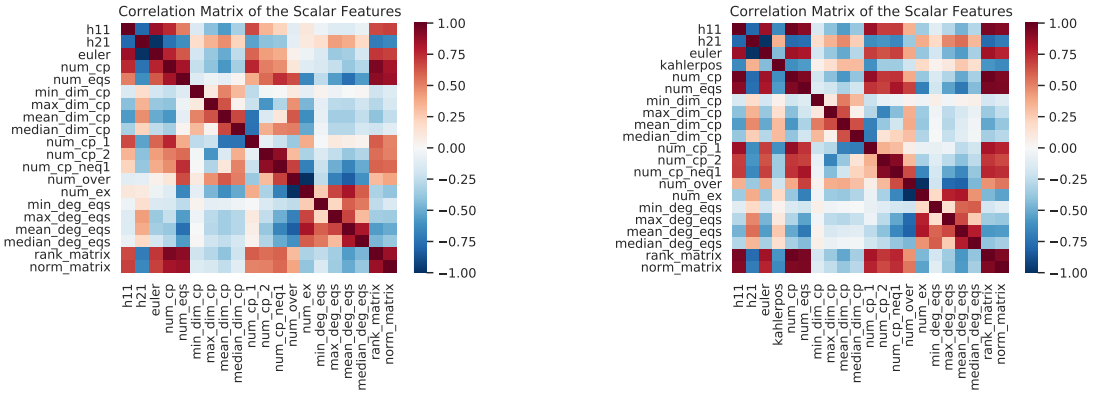
Figure 9.1: Cumulative retained variance of the principal components of the configuration matrix.

9.1.1 Engineering

Any transformation of the input data which has some mathematical meaning can be a useful feature. We establish the following list of useful quantities (most of them are already used to characterise CICY manifolds in the literature [157]):

- the number of projective spaces (rows), $m =$ `num_cp`;
- the number of equations (columns), $k =$ `num_eqs`;
- the number of $\mathbb{P}^1$, $f =$ `num_cp_1`;
- the number of $\mathbb{P}^2$, `num_cp_2`;
- the number of $\mathbb{P}^n$ with $n \neq 1$, $F =$ `num_cp_neq1`;
- the excess number $N_{ex} = \sum_{r=1}^{F} (n_r + f + m - 2k) =$ `num_ex`;
- the dimension of the cohomology group H^0 of the ambient space, `dim_h0_amb`;
- the Frobenius norm of the matrix, `norm_matrix`;
- the list of the projective space dimensions `dim_cp` and statistics thereof (min, max, median, mean);
- the list of the equation degrees `deg_eqs` and statistics thereof (min, max, median, mean);
- k-means clustering on the components of the configuration matrix (with a number of clusters going from 2 to 15);[58]
- principal components of the configuration matrix derived using a principal components analysis (PCA) with 99 % of the variance retained (see Figure 9.1).

[58] The algorithm determines the centroids of conglomerates of data called *clusters* using an iterative process which computes the distance of each sample from the center of the cluster. It then assigns the label of the cluster to the nearest samples. We use the class `cluster.KMeans` in `scikit-learn`.



(a) original dataset

(b) favourable dataset

Figure 9.2: Correlations between the engineered scalar features and the labels.

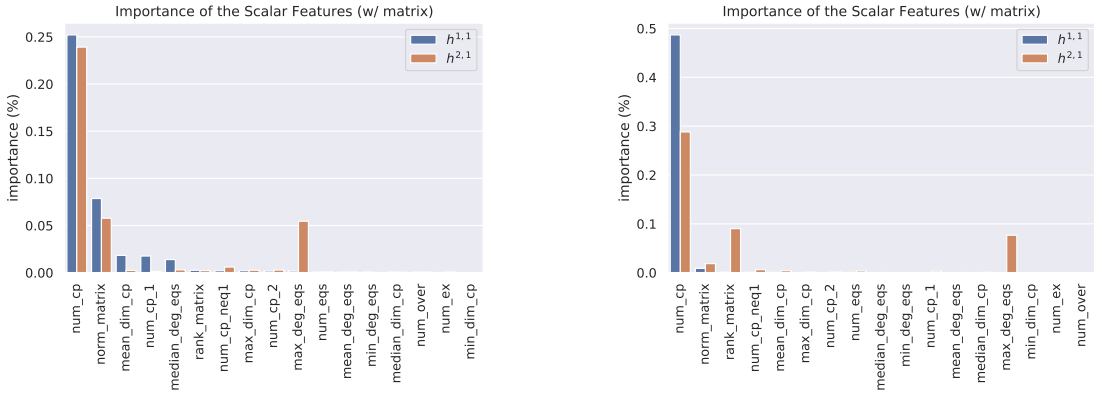
9.1.2 Selection

Correlations To get a first general idea it is useful to take a look at the correlation matrix of the features and the labels.[59] The correlation matrices for the scalar variables are displayed in Figure 9.2 for the original and favourable datasets (this excludes the configuration matrix). As we can see some engineered features are strongly correlated, especially in the favourable dataset. In particular $h^{1,1}$ (respectively $h^{2,1}$) is strongly correlated (respectively anti-correlated) with the number of projective spaces m and with the norm and rank of the matrix. This gives a first hint that these variables could help improve predictions by feeding them to the algorithm along with the matrix. On the other hand finer information on the number of projective spaces and equations do not correlate with the Hodge numbers.

From this analysis, and in particular from Figure 9.2, we find that the values of $h^{1,1}$ and $h^{2,1}$ are also correlated. This motivates the simultaneous learning of both Hodge numbers since it can increase chances for the neural network to learn more universal features. In fact this is something that often happens in practice: it has been found that multi-tasking enhances the ability to generalise [160]–[164].

Feature importance A second non-exclusive option is to sort the features by order of importance. This can be done using a decision tree which is capable to determine the weight of each variable towards making a prediction. One advantage over correlations is that the algorithm is non-linear and can thus determine subtler relations between the features and labels. To avoid biasing the results using only one decision tree, we trained a random forest of trees (using `ensemble.RandomForestRegressor` in `scikit-learn`). It consists in a large number of decision trees which are trained on different random subsets of the training dataset and averaged over the outputs (see Appendix F.3 for details on the implementation). The algorithm determines the importance of the different features to make predictions as a by-product of the learning process: the most relevant features tend to be found at the first branches (or to be used the

[59] The correlation is defined as the ratio between the covariance of two variables $\sigma(x,y) = \sum_i (x_i - \bar{x})(y_i - \bar{y})$ and the product of the standard deviations $\sigma(x)\sigma(y)$ (in this case $\bar{x}$ and $\bar{y}$ are the sample means).



(a) original dataset

(b) favourable dataset

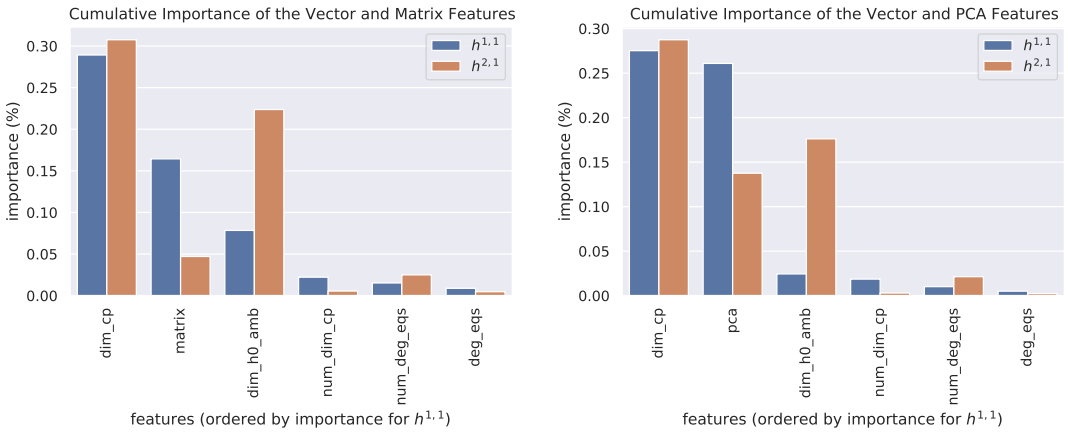
Figure 9.3: Importance of the scalar features in the datasets.

most) since they are the most important to make the prediction. The importance of a variable is a number between 0 and 1, and the sum over all of them must be 1. Since a random forest contains many trees the robustness of the variable ranking usually improves with respect to a single tree (Appendix F.3). Moreover, as the main objective is to obtain a qualitative preliminary understanding of the features, there is no need for fine tuning at this stage and we use the default parameters (specifically 100 decision trees). We computed feature importance for both datasets and for two different set of variables: one containing the engineered features and the configuration matrix, and one with the engineered features and the PCA components. In the following figures, we show several comparisons of the importance of the features, dividing the figures into scalars, vectors and configuration matrix (or its PCA), and clusters. The sum of importance of all features equals 1.

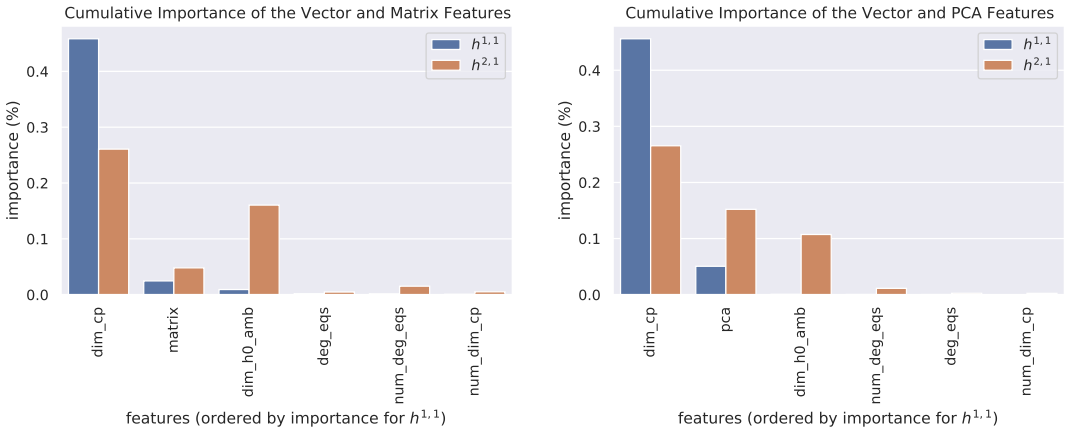
In Figure 9.3, we show the ranking of the scalar features in the two datasets (differences between the set using the configuration matrix and the other using the PCA are marginal and are not shown to avoid redundant plots). As already mentioned we find that the number of projective spaces is the most important feature by far. It is followed by the matrix norm in the original dataset, and by the matrix rank for $h^{2,1}$ in the favourable dataset. Finally the variable ranking points out that the other features have a negligible impact on the determination of the labels and may as well be ignored during training.

The same analysis can be repeated for the vector features and the configuration matrix component by component. In Figure 9.4 we show the cumulative importance of the features (i.e. the sum of the importance of each component). We can appreciate that the list of the projective space dimensions plays a major role in the determination of the labels in both datasets. In the case of $h^{2,1}$ we also have a large contribution from the dimensions of the cohomology group `dim_h0_amb`, as can be expected from algebraic topology [157].

In Figure 9.5 we show the importance associated to the number of clusters used during the EDA: no matter how many clusters we use, their relevance is definitely marginal compared to all other features used in the variable ranking (scalars, vectors, and the configuration matrix or its PCA) for both datasets.



(a) Original dataset



(b) Favourable dataset

Figure 9.4: Ranking of the vector features and the configuration matrix (or its PCA).

(a) Original dataset

(b) Favourable dataset

Figure 9.5: Incidence of the numbers of clusters on the variable ranking.

Discussion It seems therefore that the number of projective spaces plays a relevant role in the determination of $h^{1,1}$ and $h^{2,1}$ as well as the list of dimensions of the projective spaces. In order to validate this observation, in Figure 9.6 we present a scatter plot of the Hodge number distributions versus the number of projective spaces: it shows that there is indeed a linear dependence in m for $h^{1,1}$, especially in the favourable dataset. In fact the only exceptions to this pattern in the latter case are the manifolds which do not have a favourable embedding [149]. Hence, a simple data analysis hints naturally towards this mathematical result.

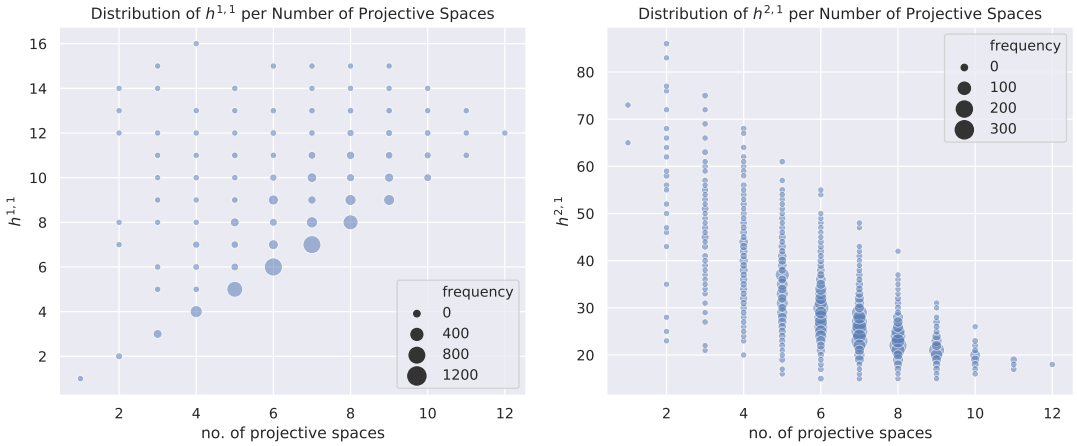
Finally we found other features which may be relevant and are worth to be included in the algorithm: the matrix rank and norm, the list of projective space dimensions and of the associated cohomology dimensions. However, we want to emphasize one caveat to this analysis: correlations look only for linear relations, and the random forest has not been optimized or could just be not powerful enough to make good predictions. This means that feature selection just gives a hint but it may be necessary to adapt it to different situations.

9.1.3 Removing Outliers

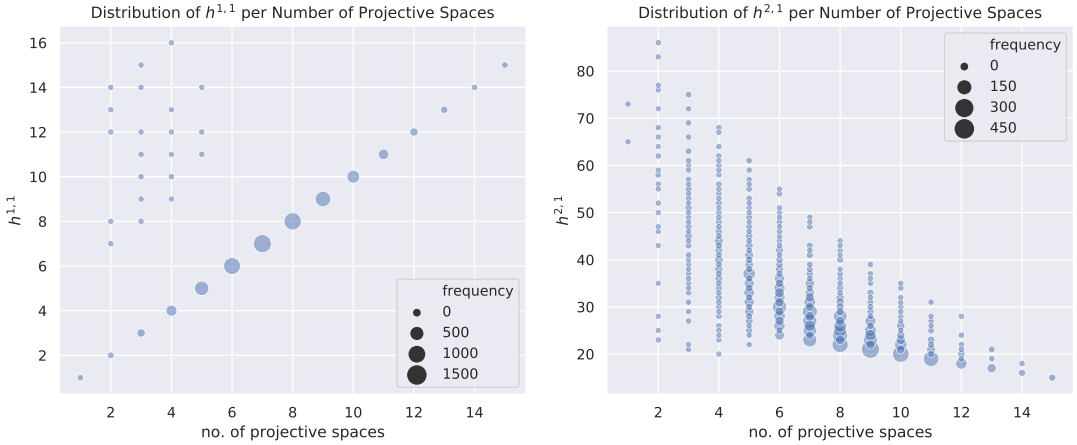
The Hodge number distributions (see Figures 8.1 and 9.7) display few outliers outside the tail of the main distributions. Such outliers may negatively impact the learning process and drive down the accuracy: it makes sense to remove them from the training set. It is easy to see that the 22 outlying manifolds with $h^{1,1} = h^{2,1} = 0$ are product spaces, recognisable from their block-diagonal matrix. We will also remove outliers with $h^{1,1} = 19$ and $h^{2,1} > 86$, which represent 15 and 2 samples. In total this represents 39 samples, or 0.49 % of the total data.

To simplify the overall presentation, since the dataset is complete we will mainly focus on the pruned subset of the data obtained by removing outliers, even from the test set.[60] This implies

[60] There is no obligation to use a ML algorithm to label outliers in the training set. It is perfectly fine to decide which data to include or not, even based on targets. However, for a real-world application, outliers in the test set should be labeled by some process based only on the input features. Flagging possible outliers may improve the predictions by helping the machine understand that such samples require more caution.

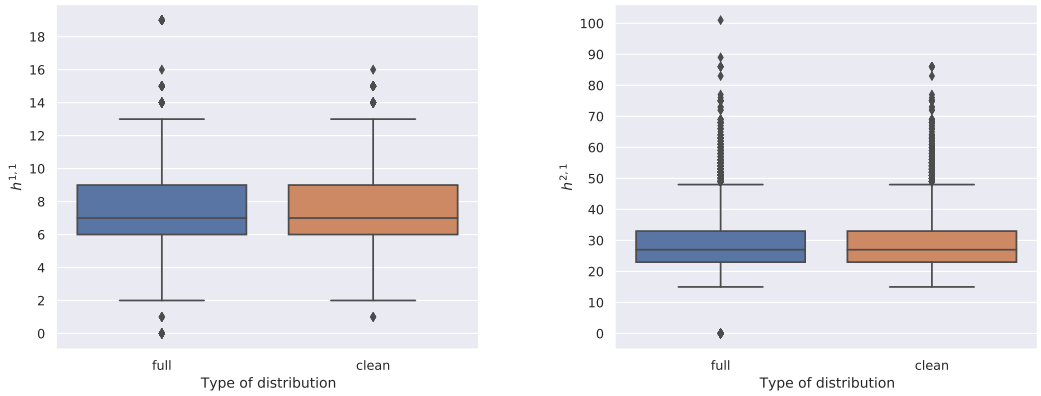


(a) Original dataset



(b) Favourable dataset

Figure 9.6: Distribution of the labels with respect to the number of projective spaces.



(a) $h^{1,1}$ (b) $h^{2,1}$

Figure 9.7: Summary of the statistics for the distributions of both Hodge numbers. The coloured box shows the three quartiles of the distributions, with the internal horizontal line corresponding to the median. The "whiskers" cover the interquartile range, i.e. 1.5 times the distance between the first and third quartiles from the lower and upper limits of the boxes. Isolated points show the remaining outliers which we however choose to keep to avoid excessively pruning the dataset.

that Hodge numbers lie in the ranges $1 \leq h^{1,1} \leq 16$ and $15 \leq h^{2,1} \leq 86$. Except when stated otherwise, accuracy is indicated for this pruned dataset. Obviously the very small percentage of outliers makes the effect of removing them from the test set negligible when stating accuracy.

9.2 Machine Learning Analysis

We compare the performances of different ML algorithms: linear regression, support vector machines (SVM), random forests, gradient boosted trees and (deep) neural networks. We obtain the best results using deep *convolutional* neural networks. In fact we present a new neural network architecture, inspired by the Inception model [165]–[167] which has been developed in the field of computer vision. We provide some details on the different algorithms in Appendix F and refer the reader to the literature [140]–[144] for more details.

9.2.1 Feature Extraction

In Section 9.1 the EDA showed that several engineered features are promising for predicting the Hodge numbers. In what follows we compare the performances of various algorithms using different subsets of features:

- only the configuration matrix (no feature engineering);
- only the number of projective spaces m;
- only a subset of engineered features and not the configuration matrix nor its PCA;

- a subset of engineered features and the PCA of the matrix.

Following the EDA and feature engineering, we finally select the features we use in the analysis by choosing the highest ranked features. We keep the number of projective spaces (`num_cp` in the dataset) and the list of the dimension of the projective spaces (`dim_cp`) for both $h^{1,1}$ and $h^{2,1}$). We also include the dimension of the cohomology group of the ambient space `dim_h0_amb` but only for $h^{2,1}$.[61]

9.2.2 Analysis Strategy

For the ML analysis, we split the dataset into training and test sets: we fit the algorithms on the first and then show the predictions on the test set, which will not be touched until the algorithms are ready.

Test split and validation The training set is made of 90 % of the samples for training, which leaves the remaining 10 % in the test set (i.e. 785 manifolds out of the 7851 in the set).[62] For most algorithms, we use *leave-one-out* cross-validation on the training set as evaluation of the algorithm: we subdivide the training set in 9 subsets, each of them containing 10 % of the *total* amount of samples, then we train the algorithm on 8 of them and evaluate it on the 9th. We then repeat the procedure changing the evaluation fold until the algorithm has been trained and evaluated on all of them. The performance measure in validation is given by the average over all the left out folds. When training neural networks, we will however use a single *holdout validation* set made of 10 % of the *total* samples.

Predictions and metrics Since we are interested in predicting exactly the Hodge numbers, the appropriate metric measuring the success of the predictions is the accuracy (for each Hodge number separately):

$$\text{accuracy} = \frac{1}{N}\sum_{i=1}^{N} \delta_{y^{\text{true}\,(i)},\, y^{\text{pred}\,(i)}}, \tag{9.1}$$

where N is the number of samples. In this analysis the accuracy of the predictions on the test set is rounded to the nearest integer.

Since the Hodge numbers are integers the problem of predicting them looks like a classification task. However, as argued in the introduction, we prefer to use a regression approach. Indeed regression does not require to specify the data boundaries and allows to extrapolate beyond them, contrary to a classification approach where the categories are fixed at the beginning.[63]

[61] Notice that providing different kinds of input features to the algorithm is fine as long as such variables come from the same training set. In other words, it is possible to provide different representations of the same set to different algorithms while retaining the same statistical relevance [143], [144].

[62] Remember that we have removed outliers, see Section 9.1.3. The interested reader can refer to [4] where outliers are kept in the test set.

[63] A natural way to transform the problem in a regression task is to *standardise* the Hodge numbers, for example by shifting by the mean value and diving by the standard deviation. Under this transformation, the Hodge numbers are mapped to real numbers. While standardisation often improves ML algorithms, we found that the impact was mild or even negative.

Most algorithms need a differentiable loss function since the optimisation of parameters (such as neural networks weights) uses some variant of the gradient descent method. For this reason the accuracy cannot be used and the models are trained by minimisation of the mean squared error (MSE), which is simply the squared ℓ_2-norm between of the difference between the predictions and the real values. There will however be also a restricted number of cases in which we will use either the mean absolute error (MAE), which is the ℓ_1-norm of the same difference, or a weighted linear combination of MSE and MAE (also known as *Huber* loss): we will point them out at the right time. When predicting both Hodge numbers together, the total loss is the sum of each individual loss with equal weight: $h^{1,1}$ is simpler to learn so it is useful to put emphasis on learning $h^{2,1}$, but the magnitudes of the latter are higher, such that the associated loss is naturally bigger (since we did not normalise the data).

In general predictions are real numbers: we need to turn them into integers. In general, rounding to the nearest integer gives the best result, but we found algorithms (such as linear regression) for which flooring to the integer below works better. The optimal choice of the integer function is found for each algorithm as part of the hyperparameter optimisation described below. The accuracy is computed after rounding.

Learning curves for salient models are displayed. They show how the performances of a model improves by using more training data, for fixed hyperparameters. To obtain it, we train models using from 10 % to 90 % of all the data ("training ratio") and evaluate the accuracy on the remaining data.[64]

To avoid redundant information and to avoid cluttering the paper with graphs, the results for models predicting separately the Hodge numbers for the test set are reported in tables, while the results for the models predicting both numbers together are reported in the learning curves. For the same reason, the latter are not displayed for the favourable dataset.

Visualisation of the performance Complementary to the predictions and the accuracy results, we also provide different visualisations of the performance of the models in the form of univariate plots (histograms) and multivariate distributions (scatter plots). In fact the usual assumption behind the statistical inference of a distribution is that the difference between the observed data and the predicted values can be modelled by a random variable called *residual* [144], [168].[65] As such we expect that its values can be sampled from a normal distribution with a constant variance (i.e. constant width), since it should not depend on specific observations, and centered around zero, since the regression algorithm tries to minimise the squared difference between observed and predicted values. Histograms of the residual errors should therefore exhibit such properties graphically. Another interesting kind of visual realisation of the residuals is to show their distribution against the variables used for the regression model: in the case of a simple regression model in one variable, it is customary to plot the residuals as a function of the independent variable, but in a multivariable regression analysis (such as the case at hand) the choice usually falls on the values predicted by the fit (not the observed data). We shall therefore

[64] Statistics are not provided due to the limitations of our available computational resources, namely a *Thinkpad t470p* laptop with *Intel i7-7700HQ* CPU, 16 GB RAM and *NVidia GeForce 940MX* GPU. However, we check manually on few examples that the reported results are typical.

[65] The difference between the non observable *true* value of the model and the observed data is known as *statistical error*. The difference between residuals and errors is subtle but the two definitions have different interpretations in the context of the regression analysis: in a sense, residuals are an estimate of the errors.

plot the residuals as functions of the predicted values.[66] Given the assumption of the random distribution of the residuals, they should not present strong correlations with the predictions and should not exhibit trends. In general the presence of correlated residuals is an indication of an incomplete or incorrect model which cannot explain the variance of the predicted data, meaning that the model is either not suitable for predictions or that we should add information (i.e. add features) to it.

Hyperparameter optimisation One of the key steps in a ML analysis is the optimisation of the *hyperparameters* of the algorithm. These are internal parameters of each estimator (such as the number of trees in a random forest or the amount of regularisation in a linear model): they are not modified during the training of the model, but they directly influence the process in terms of performance and outcome.

Hyperparameter optimisation is performed by training many models with different choices of their values. We then keep the values best performing according to some metric on the validation set(s).[67] As it does not need to be differentiable we use the accuracy as a scoring function to evaluate the models. There is however a subtle issue because it is not clear how to combine the accuracy of $h^{1,1}$ and $h^{2,1}$ to get a single metric. For this reason we perform the analysis on both Hodge numbers separately. Then we can design a single model computing both Hodge numbers simultaneously by making a compromise by hand between the hyperparameters found for the two models computing the Hodge numbers separately. The optimisation is implemented using the API in `scikit-learn`, using the function `metrics.make_scorer` and the accuracy as a custom scoring function.

There are several approaches to perform this search automatically, in particular: grid search, random search, genetic evolution, and Bayes optimisation. Grid and random search are natively implemented in `scikit-learn`. The first takes a list of possible discrete values of the hyperparameters and will evaluate the algorithm over all possible combinations. The second samples the values in both discrete sets and continuous intervals according to some probability distribution, repeating the process a fixed number of times. The grid search method is particularly useful for discrete hyperparameters, less refined searches or for a small number of combinations, while the second method can be used to explore the hyperparameter space on a larger scale [169]. Genetic algorithms are based on improving the choice of hyperparameters over *generations* that successively select only the most promising values: in general they require a lot of tuning and are easily influenced by the fact that the replication process can also lead to worse results totally at random [170]. They are however effective when dealing with very deep or complex neural networks. Bayes optimisation [171], [172] is a very well established mathematical procedure to find the stationary points of a function without knowing its analytical form [173]. It relies on assigning a *prior* probability to a given parameter and then multiply it by the probability distribution (or *likelihood*) of the scoring function to compute the probability of finding a better results given a set of hyperparameters. This has proven to be very effective in our case and we adopted this solution as it does not require fine tuning and leads to better results for models which are not deep neural networks. We choose to use `scikit-optimize` [155] whose method `BayesSearchCV` has a very well implemented Python interface compatible with `scikit-learn`. We will in general perform 50 iterations of the Bayes search algorithm, unless otherwise specified.

[66] We will use the same strategy also for the fit using just the number of projective spaces in order to provide a way to compare the plots across different models.

[67] Notice the importance of having a validation set separate from the test set: we must avoid adapting the algorithm to the same set we use for the predictions or the generalisation capabilities of the algorithm will suffer.

9.3 Linear Models

Linear models attempt to describe the labels as a linear combinations of the input features while keeping the coefficients at $\mathcal{O}(1)$ (see Appendix F.1). However non-linearity can still be introduced by engineering features which are non-linear in terms of the original data.[68]

From the results of Section 9.1, we made a hypothesis on the linear dependence of $h^{1,1}$ on the number of projective spaces m. As a first approach, we can try to fit a linear model to the data as a baseline computation and to test whether there is actual linear correlation between the two quantities. We will consider different linear models including their regularised versions.

Parameters The linear regression is performed with the class `linear_model.ElasticNet` in `scikit-learn`. The hyperparameters involved in this case are: the amount of regularisation α, the relative ratio (`l1_ratio`) between the ℓ_1 and ℓ_2 regularization losses, and the fit of the intercept. By performing the hyperparameter optimisation we found that ℓ_2 regularization has a minor impact and can be removed, which corresponds to setting the relative ratio to 1 (this is equivalent to using `linear_model.Lasso`). In Table 9.1 we show the choices of the hyperparameters for the different models we built using the ℓ_1 regularised linear regression.

For the original dataset, we floored the predictions to the integers below, while in the favourable we rounded to the next integer. This choice for the original dataset makes sense: the majority of the samples lie on the line $h^{1,1} = m$, but there are still many samples with $h^{1,1} > m$ (see Figure 9.6). As a consequence the ML prediction pulls the line up which can only damage the accuracy. Choosing the floor function is a way to counteract this effect. Note that accuracy for $h^{2,1}$ is only slightly affected by the choice of rounding, so we just choose the same one as $h^{1,1}$ for simplification.

Results In Table 9.2 we show the accuracy for the best hyperparameters. For $h^{1,1}$, the most precise predictions are given by the number of projective spaces which actually confirms the hypothesis of a strong linear dependence of $h^{1,1}$ on the number of projective spaces. In fact this gives close to 100 % accuracy for the favourable dataset which shows that there is no need for more advanced ML algorithms. Moreover adding more engineered features *decreases* the accuracy in most cases where regularization is not appropriate. The accuracy for $h^{2,1}$ remains low but including engineered features definitely improves it.

In Figure 9.8 we show the plots of the residual errors of the model on the original dataset. For the ℓ_1 regularised linear model, the univariate plots show that the errors seem to follow normal distributions peaked at 0 as they generally should: in the case of $h^{1,1}$, the width is also quite contained. The scatter plots instead show that in general there is no correlation between

[68] In general *linear model* is used to indicate that the coefficients β of the features appear linearly in the expression of the prediction of the i-th label:

$$y^{\text{pred}\,(i)} = \beta_0^{(i)} + \beta_1^{(i)} x_1^{(i)} + \cdots + \beta_F^{(i)} x_F^{(i)} = \sum_{j=0}^{F} \beta_j^{(i)} x_j^{(i)},$$

where m is the number of independent variables and $x_0^{(i)} = 1$ (i.e. $\beta_0^{(i)}$ is the intercept of the model and represents the value of the label without the contribution of any of the features). In other words, $\beta_j^{(i)}$ are used with unit exponent only once per model.

		matrix		**num_cp**		**eng. feat.**		**PCA**	
		old	*fav.*	*old*	*fav.*	*old*	*fav.*	*old*	*fav.*
α	$h^{1,1}$	2.0×10^{-6}	3.0×10^{-5}	0.10	2.0×10^{-6}	0.05	0.05	0.07	0.08
	$h^{2,1}$	1.0×10^{-6}	1.0×10^{-5}	0.1	1.0×10^{-6}	3.0×10^{-4}	1.2×10^{-3}	2.0×10^{-6}	1.2×10^{-3}
`fit_intercept`	$h^{1,1}$	False	False	True	False	True	True	False	True
	$h^{2,1}$	True	True	True	True	True	False	True	False
`normalise`	$h^{1,1}$	—	—	False	—	False	False	—	False
	$h^{2,1}$	False	True	False	False	False	—	True	—

Table 9.1: Hyperparameter choices of the ℓ_1 regression model used. In addition to the known hyperparameters α and `fit_intercept`, we also include the `normalise` parameter which indicates whether the samples have been centered and scaled by their ℓ_2 norm before the fit: it is ignored when the intercept is ignored.

		matrix	**num_cp**	**eng. feat.**	**PCA**
original	$h^{1,1}$	51 %	63 %	63 %	64 %
	$h^{2,1}$	11 %	8 %	21 %	21 %
favourable	$h^{1,1}$	95 %	100 %	100 %	100 %
	$h^{2,1}$	14 %	15 %	19 %	19 %

Table 9.2: Best accuracy of the linear model using ℓ_1 regularisation on the test split.

a particular sector of the predictions and the error made by the model, thus the variance of the residuals is in general randomly distributed over the predictions. Only the case of the fit of the number of projective spaces seems to show a slight correlation for $h^{2,1}$, signalling that the model using only one feature might be actually incomplete: in fact it is better to include also other engineered features.

The learning curves in Figure 9.9 show that the model underfits. We also notice that the models are only marginally affected by the number of samples used for training. In particular this provides a very strong baseline for $h^{1,1}$. For comparison, we also give the learning curve for the favourable dataset in Figure 9.10: this shows that a linear regression is completely sufficient to determine $h^{1,1}$ in that case.

9.4 Support Vector Machines

SVM are a family of algorithms which use a *kernel trick* to map the space of input data vectors into a higher dimensional space where samples can be accurately separated and fitted to an appropriate curve (see Appendix F.2). In this analysis we show two such kernels, namely a linear kernel (also known as *no kernel* since no transformations are involved) and a Gaussian kernel (known as `rbf` in ML literature as in *radial basis function*).

9.4.1 Linear Kernel

For this model we use the class `svm.LinearSVR` in `scikit-learn`.

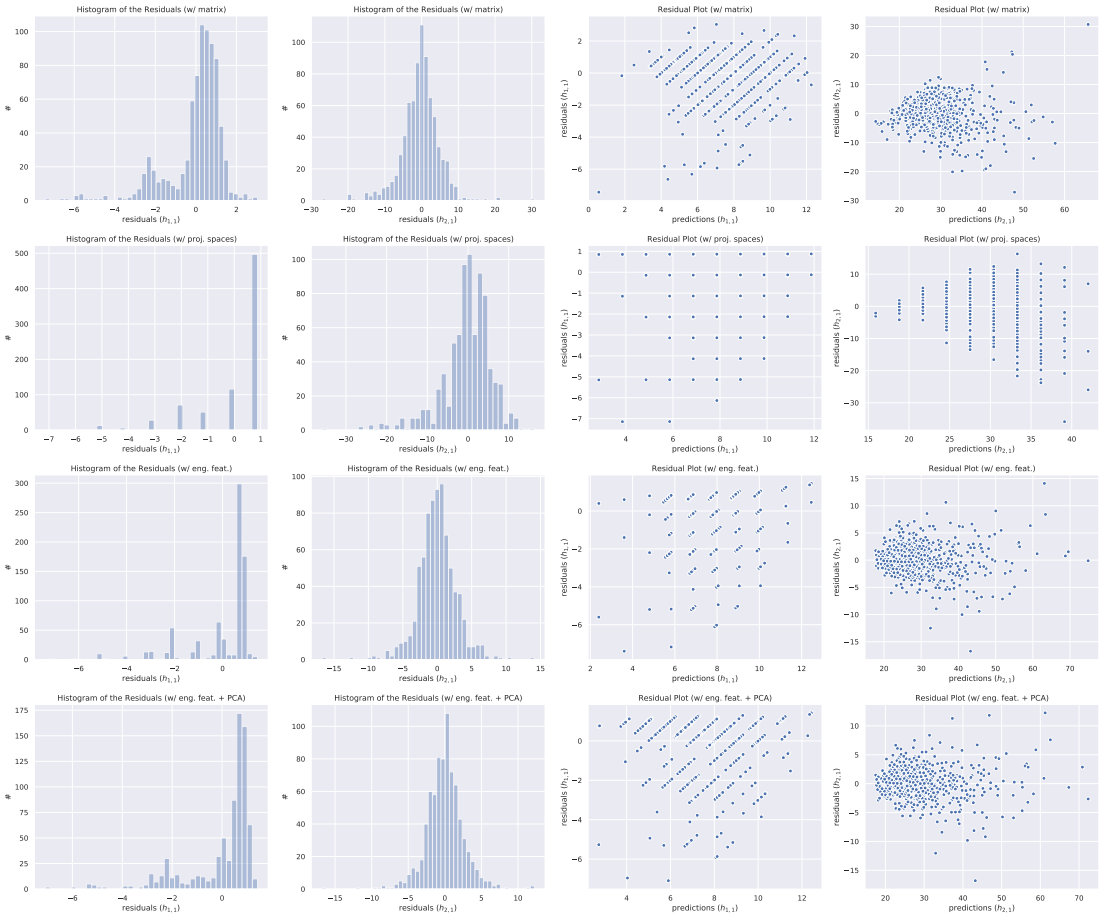
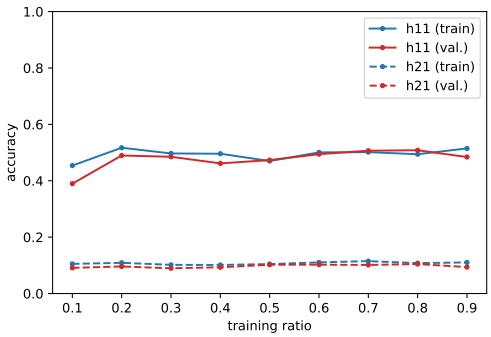
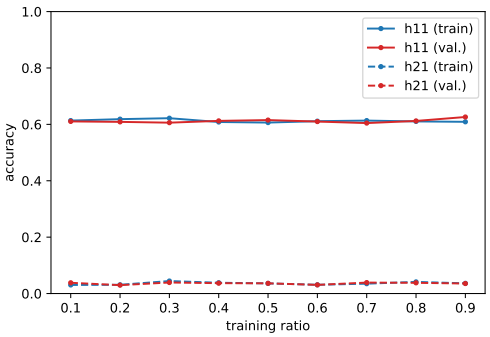


Figure 9.8: Plots of the residual error for the ℓ_1 regularised linear model: rows show the different scenarios (fit with only the matrix, with only the number of projective spaces, with the engineered features, with the engineered features and the PCA). Plots refer to the test split of the original dataset.



(a) input: `matrix`, $\alpha = 2 \times 10^{-4}$



(b) input: `num_cp`, $\alpha = 1$

Figure 9.9: Learning curves for the linear regression (original dataset), including outliers and using a single model for both Hodge numbers.

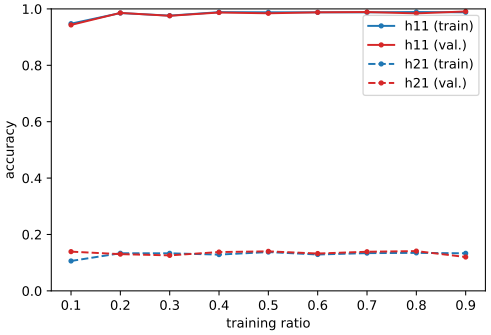


Figure 9.10: Learning curves for the linear regression (favourable dataset), including outliers and using a single model for both Hodge numbers. Input: `num_cp`, $\alpha = 1$.

		matrix		**num_cp**		**eng. feat.**		**PCA**	
		old	*fav.*	*old*	*fav.*	*old*	*fav.*	*old*	*fav.*
`C`	$h^{1,1}$	0.13	24	0.001	0.0010	0.13	0.001	0.007	0.4
	$h^{2,1}$	0.30	100	0.05	0.0016	0.5	0.4	1.5	0.4
ϵ	$h^{1,1}$	0.7	0.3	0.4	0.00	0.9	0.0	0.5	0.0
	$h^{2,1}$	0.0	0.0	10	0.03	0.0	0.0	0.0	0.6
`fit_intercept`	$h^{1,1}$	True	False	True	False	True	False	False	False
	$h^{2,1}$	True	False	True	True	True	True	True	False
`intercept_scaling`	$h^{1,1}$	0.13	—	100	—	0.01	—	—	—
	$h^{2,1}$	100	—	13	92	100	0.01	100	—
`loss`	$h^{1,1}$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$	$\lVert\epsilon\rVert^2$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$
	$h^{2,1}$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$	$\lVert\epsilon\rVert^2$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$	$\lvert\epsilon\rvert$

Table 9.3: Hyperparameter choices of the linear support vector regression. The parameter `intercept_scaling` is clearly only relevant when the intercept is used. The different losses used simply distinguish between the ℓ_1 norm of the ϵ-dependent boundary where no penalty is assigned and its ℓ_2 norm.

Parameters In Table 9.3 we show the choices of the hyperparameters used for the model. As we prove in Appendix F.2, parameters C and ϵ are related to the penalty assigned to the samples lying outside the no-penalty boundary (the loss in this case is computed according to the ℓ_1 or ℓ_2 norm of the distance from the boundary as specified by the `loss` hyperparameter). Other parameters are related to the use of the intercept to improve the prediction. We rounded the predictions to the floor for the original dataset and to the next integer for the favourable dataset.

Results In Table 9.4, we show the accuracy on the test set for the linear kernel. As we can see, the performance of the algorithm strongly resembles a linear model in terms of the accuracy reached. It is fascinating to notice that the contributions of the PCA do not improve the predictions using just the engineered features: it seems that the latter work better than using the configuration matrix or its principal components.

The residual plots in Figure 9.11 confirm what we already said about the linear models with regularisation: the model with only the number of projective spaces shows a tendency to heteroscedasticity which can be balanced by adding more engineered feature, also helping in having more precise predictions (translated into peaked univariate distributions).[69] In all cases, we notice that the model slightly overestimates the real values (residuals are computed as the difference between the prediction and the real value) as the second, small peaks in the histograms for $h^{1,1}$ suggest: this may also explain why flooring the predictions produces the highest accuracy. As in general for linear models, the influence of the number of samples used for training is marginal also in this case: we only noticed a decrease in accuracy when also including the PCA or directly the matrix.

[69] Heteroscedasticity refers to the tendency to have a correlation between the predictions and the residuals: theoretically speaking, there should not be any, since we suppose the residuals to be independent on the model

		matrix	**num_cp**	**eng. feat.**	**PCA**
original	$h^{1,1}$	61 %	63 %	65 %	62 %
	$h^{2,1}$	11 %	9 %	21 %	20 %
favourable	$h^{1,1}$	96 %	100 %	100 %	100 %
	$h^{2,1}$	14 %	14 %	19 %	20 %

Table 9.4: Accuracy of the linear SVM on the test split.

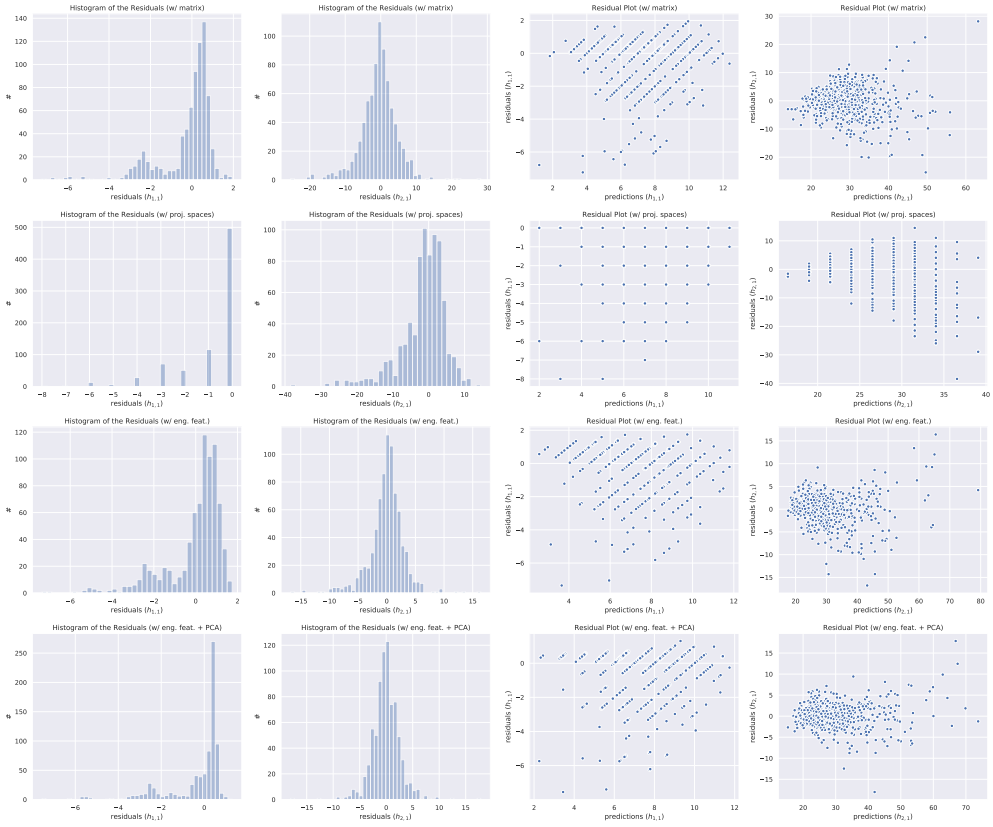


Figure 9.11: Plots of the residual errors for the SVM with linear kernel.

		matrix		**num_cp**		**eng. feat.**		**PCA**	
		old	*fav.*	*old*	*fav.*	*old*	*fav.*	*old*	*fav.*
C	$h^{1,1}$	14	1000	170	36	3	40	1.0	1000
	$h^{2,1}$	40	1000	1.0	1.0	84	62	45	40
ϵ	$h^{1,1}$	0.01	0.01	0.45	0.03	0.05	0.3	0.02	0.01
	$h^{2,1}$	0.01	0.01	0.01	0.09	0.29	0.10	0.20	0.09
γ	$h^{1,1}$	0.03	0.002	0.110	0.009	0.07	0.003	0.02	0.001
	$h^{2,1}$	0.06	0.100	0.013	1000	0.016	0.005	0.013	0.006

Table 9.5: Hyperparameter choices of the SVM regression with Gaussian kernel.

9.4.2 Gaussian Kernel

We then consider SVM using a Gaussian function as kernel. The choice of the function can heavily influence the outcome of the predictions since they map the samples into a much higher dimensional space and create highly non-linear combinations of the features before fitting the algorithm. In general this can help in the presence of "obscure" features which badly correlate one another. In our case we hope to leverage the already good correlations we found in the EDA with the kernel trick. The implementation is done with the class `svm.SVR` from `scikit-learn`.

Parameters As we show in Appendix F.2, this particular choice of kernel leads to profoundly different behaviour with respect to linear models: we will round the predictions to the next integer in both datasets since the loss function strongly penalises unaligned samples. In Table 9.5 we show the choices of the hyperparameters for the models using the Gaussian kernel. As usual the hyperparameter `C` is connected to the penalty assigned to the samples outside the soft margin boundary (see Appendix F.2) delimited by the ϵ. Given the presence of a non linear kernel we have to introduce an additional hyperparameter γ which controls the width of the Gaussian function used for the support vectors.

Results In Table 9.6 we show the accuracy of the predictions on the test sets. In the favourable dataset we can immediately appreciate the strong linear dependence of $h^{1,1}$ on the number of projective spaces: even though there are a few non favourable embeddings in the dataset, the kernel trick is able to map them in a better representation and improve the accuracy. The predictions for the original dataset have also improved: they are the best results we found using shallow learning. The predictions using only the configuration matrix matches [126] but we can slightly improve the accuracy by using a combination of engineered features and PCA. In Figure 9.12 we show the residual plots and their histograms for the original dataset: residuals follow peaked distributions which, in this case, do not present a second smaller peak (thus we need to round to the next integer the predictions) and good variate distribution over the predictions.

The Gaussian kernel is also more influenced by the size of the training set. Using 50 % of the samples as training set we witnessed a drop in accuracy of 3 % while using engineered features and

and normally sampled.

		matrix	num_cp	eng. feat.	PCA
original	$h^{1,1}$	70 %	63 %	66 %	72 %
	$h^{2,1}$	22 %	10 %	36 %	34 %
favourable	$h^{1,1}$	99 %	100 %	100 %	100 %
	$h^{2,1}$	22 %	17 %	32 %	33 %

Table 9.6: Accuracy of the Gaussian SVM on the test split.

the PCA, and around 1 % to 2 % in all other cases. The learning curves (presented in Figure 9.13) show that the accuracy improves by using more data. Interestingly, it shows that using all engineered features leads to an overfit on the training data since both Hodge numbers reach almost 100 %, while this is not the case for $h^{2,1}$. For comparison, we also display in Figure 9.14 the learning curve for the favourable dataset: this shows that predicting $h^{1,1}$ accurately works out-of-the-box.

9.5 Decision Trees

We now consider two algorithms based on decision trees: random forests and gradient boosted trees. Decision trees are powerful algorithms which implement a simple decision rule (in the style of an *if... then... else...* statement) to classify or assign a value to the predictions. However they have a tendency to adapt too well to the training set and to not be robust enough against small changes in the training data. We consider a generalisation of this algorithm used for *ensemble learning*: this is a technique in ML which uses multiple estimators (they can be the same or different) to improve the performances. We will present the results of *random forests* of trees which increase the bias compared to a single decision tree, and *gradient boosted* decision trees, which can use smaller trees to decrease the variance and learn better representations of the input data by iterating their decision functions and use information on the previous runs to improve (see Appendix F.3 for a more in-depth description).

9.5.1 Random Forests

The random forest algorithm is implemented with Scikit's `ensemble.RandomForestRegressor`.

Parameters Hyperparameter tuning for decision trees can in general be quite challenging. From the general theory on random forests (see Appendix F.3 for salient features) we can try and look for particular shapes of the trees: this ensemble learning technique usually prefers a small number of fully grown trees. We performed only 25 iterations of the optimisation process due to the very long time taken to train all the decision trees.

In Table 9.7 we show the hyperparameters used for the predictions. As we can see from `n_estimator`, random forests are usually built with a small number of fully grown (specified by `max_depth` and `max_leaf_nodes`) trees (not always the case, though). In order to avoid overfit we also tried to increase the number of samples necessary to split a branch or create a leaf node using `min_samples_leaf` and `min_samples_split` (introducing also a weight on the samples

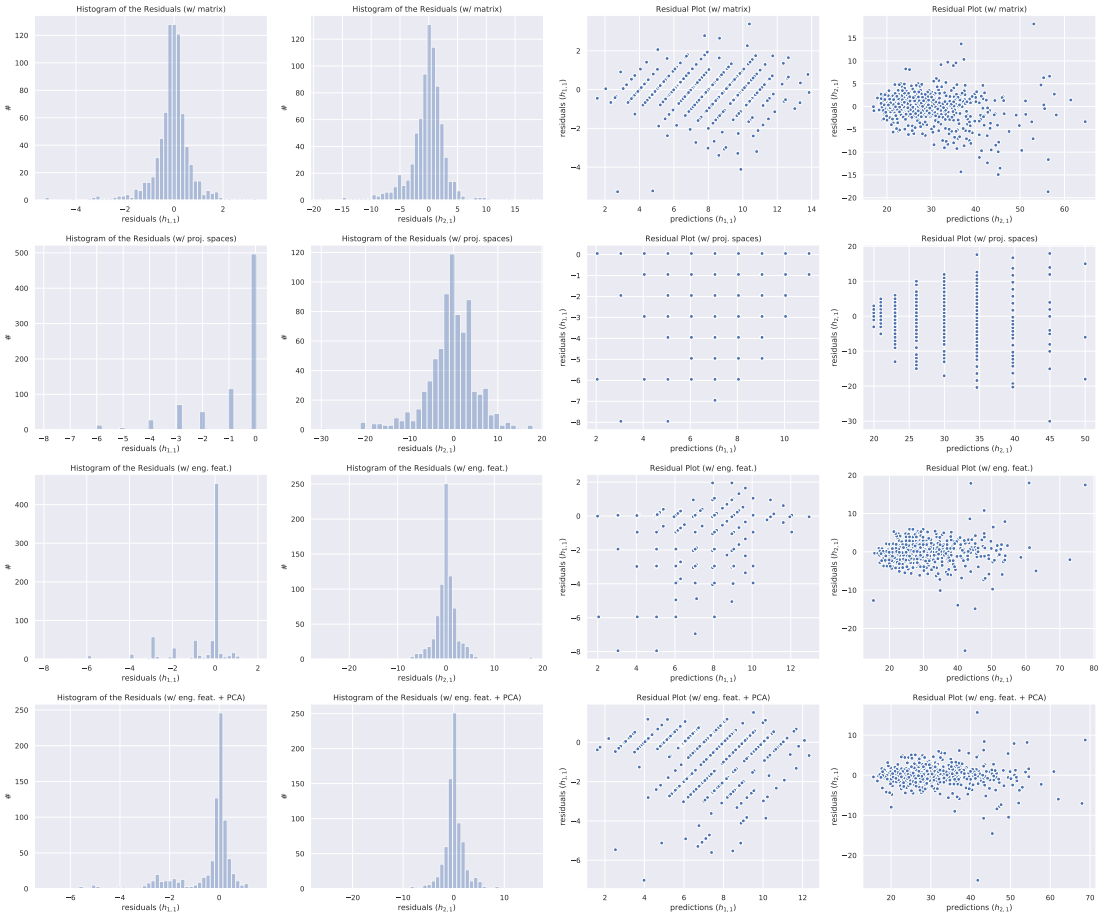
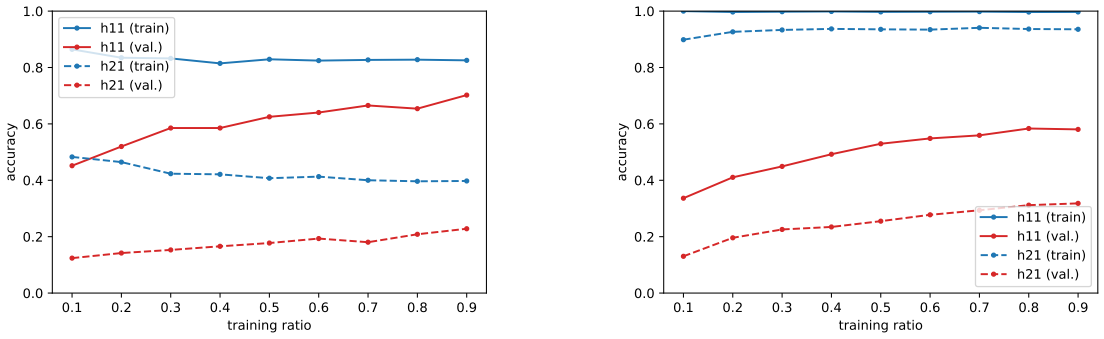


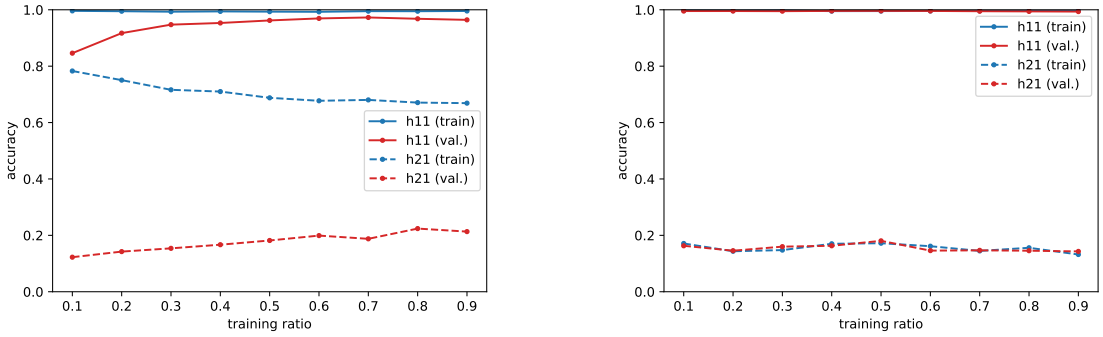
Figure 9.12: Plots of the residual errors for the SVM with Gaussian kernel.



(a) input: `matrix`, $C = 15, \gamma = 0.03, \epsilon = 0.1$

(b) input: all, $C = 10, \gamma = 0.03, \epsilon = 0.1$

Figure 9.13: Learning curves for the SVM with Gaussian kernel (original dataset), using a single model for both Hodge numbers.



(a) input: `matrix`, $C = 20, \gamma =$ `scale`, $\epsilon = 0.1$

(b) input: all, $C = 20, \gamma =$ `scale`, $\epsilon = 0.1$

Figure 9.14: Learning curves for the SVM with Gaussian kernel (favourable dataset), using a single model for both Hodge numbers.

in the leaf nodes specified by `min_weight_fraction_leaf` to balance the tree). Finally the `criterion` chosen by the optimisation reflects the choice of the trees to measure the impurity of the predictions using either the mean squared error or the mean absolute error of the predictions (see Appendix F.3).

Results In Table 9.8 we summarise the accuracy reached using random forests of decision trees as estimators. As we already expected, the contribution of the number of projective spaces helps the algorithm to generate better predictions. In general, it seems that the engineered features alone can already provide a good basis for predictions. In the case of $h^{2,1}$ the introduction of the principal components of the configuration matrix also increases the prediction capabilities. As in most other cases we used the floor function for the predictions on the original dataset and the rounding to next integer for the favourable one.

As usual in Figure 9.15 we show the histograms of the distribution of the residual errors and the scatter plots of the residuals. While the distributions of the errors are slightly wider than the SVM algorithms, the scatter plots of the residual show a strong heteroscedasticity in the case of the fit using the number of projective spaces: though quite accurate, the model is strongly incomplete. The inclusion of the other engineered features definitely helps and also leads to better predictions. Learning curves are displayed in Figure 9.16.

9.5.2 Gradient Boosted Trees

We used the class `ensemble.GradientBoostingRegressor` in `scikit-learn` to implement the gradient boosted trees.

Parameters Hyperparameter optimisation has been performed using 25 iterations of the Bayes search algorithm since by comparison the gradient boosting algorithms took the longest learning time. We show the chosen hyperparameters in Table 9.9.

		matrix		**num_cp**		**eng. feat.**		**PCA**	
		old	*fav.*	*old*	*fav.*	*old*	*fav.*	*old*	*fav.*
`criterion`	$h^{1,1}$	MSE	MSE	MAE	MAE	MAE	MSE	MAE	MAE
	$h^{2,1}$	MAE	MAE	MAE	MAE	MAE	MAE	MAE	MAE
`max_depth`	$h^{1,1}$	100	100	100	30	90	30	30	60
	$h^{2,1}$	90	100	90	75	100	100	100	60
`max_leaf_nodes`	$h^{1,1}$	100	80	90	20	20	35	90	90
	$h^{2,1}$	90	100	100	75	100	60	100	100
`min_samples_leaf`	$h^{1,1}$	1	1	1	15	1	15	1	1
	$h^{2,1}$	3	1	4	70	1	70	30	1
`min_samples_split`	$h^{1,1}$	2	30	20	35	10	10	100	100
	$h^{2,1}$	30	2	50	45	2	100	2	100
`min_weight_fraction_leaf`	$h^{1,1}$	0.0	0.0	0.0	1.7×10^{-3}	0.0	0.009	0.0	0.0
	$h^{2,1}$	3.0×10^{-4}	0.0	1.0×10^{-4}	0.13	0.0	0.0	0.0	0.0
`n_estimators`	$h^{1,1}$	10	100	45	120	155	300	10	300
	$h^{2,1}$	190	10	160	300	10	10	10	300

Table 9.7: Hyperparameter choices of the random forest regression.

		matrix	**num_cp**	**eng. feat.**	**PCA**
original	$h^{1,1}$	55 %	63 %	66 %	64 %
	$h^{2,1}$	12 %	9 %	17 %	18 %
favourable	$h^{1,1}$	89 %	99 %	98 %	98 %
	$h^{2,1}$	14 %	17 %	22 %	27 %

Table 9.8: Accuracy of the random forests on the test split.

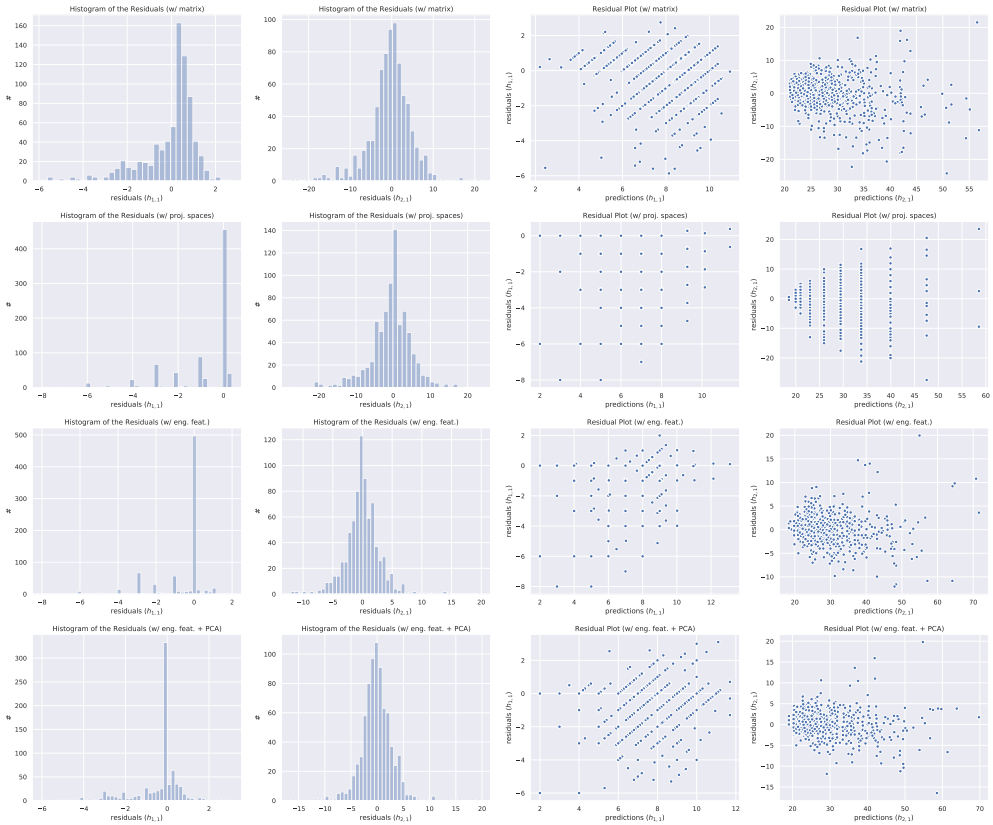
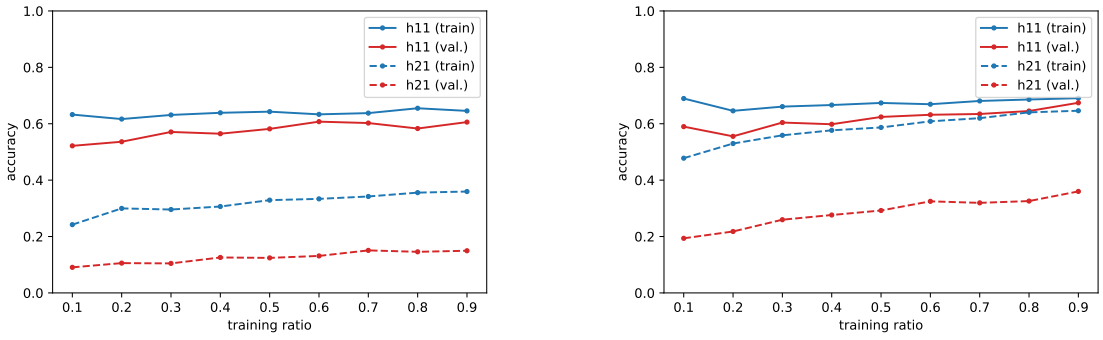


Figure 9.15: Plots of the residual errors for the random forests.



(a) input: `matrix`, default parameters

(b) input: all, default parameters

Figure 9.16: Learning curves for the random forest (original dataset) including outliers and using a single model for both Hodge numbers.

		matrix		num_cp		eng. feat.		PCA	
		old	*fav.*	*old*	*fav.*	*old*	*fav.*	*old*	*fav.*
α	$h^{1,1}$	0.4	—	—	—	—	—	—	—
	$h^{2,1}$	—	0.11	—	—	0.99	—	—	—
`criterion`	$h^{1,1}$	MAE	MAE	`friedman_mse`	MAE	`friedman_mse`	`friedman_mse`	MAE	MAE
	$h^{2,1}$	MAE	MAE	`friedman_mse`	MAE	MAE	MAE	MAE	MAE
`learning_rate`	$h^{1,1}$	0.3	0.04	0.6	0.03	0.15	0.5	0.04	0.03
	$h^{2,1}$	0.6	0.5	0.3	0.5	0.04	0.02	0.03	0.07
`loss`	$h^{1,1}$	huber	ls	lad	ls	ls	lad	ls	ls
	$h^{2,1}$	ls	huber	ls	ls	huber	ls	ls	lad
`max_depth`	$h^{1,1}$	100	100	15	60	2	100	55	2
	$h^{2,1}$	85	100	100	30	35	60	15	2
`min_samples_split`	$h^{1,1}$	2	30	20	35	10	10	100	100
	$h^{2,1}$	30	2	50	45	2	100	2	100
`min_weight_fraction_leaf`	$h^{1,1}$	0.03	0.0	0.0	0.2	0.2	0.0	0.06	0.0
	$h^{2,1}$	0.0	0.0	0.16	0.004	0.0	0.0	0.0	0.0
`n_estimators`	$h^{1,1}$	90	240	120	220	100	130	180	290
	$h^{2,1}$	100	300	10	20	200	300	300	300
`subsample`	$h^{1,1}$	0.8	0.8	0.9	0.6	0.1	0.1	1.0	0.9
	$h^{2,1}$	0.7	1.0	0.1	0.9	0.1	0.9	0.1	0.2

Table 9.9: Hyperparameter choices of the gradient boosted decision trees.

		matrix	num_cp	eng. feat.	PCA
original	$h^{1,1}$	50 %	63 %	61 %	58 %
	$h^{2,1}$	14 %	9 %	23 %	21 %
favourable	$h^{1,1}$	97 %	100 %	99 %	99 %
	$h^{2,1}$	17 %	16 %	35 %	22 %

Table 9.10: Accuracy of the gradient boosting on the test split.

With respect to the random forests, for the gradient boosting we also need to introduce the `learning_rate` (or *shrinking parameter*) which controls the gradient descent of the optimisation which is driven by the choice of the `loss` parameters (`ls` is the ordinary least squares loss, `lad` is the least absolute deviation and `huber` is a combination of the previous two losses weighted by the hyperparameter α). We also introduce the `subsample` hyperparameter which chooses a fraction of the samples to be fed into the algorithm at each iteration. This procedure has both a regularisation effect on the trees, which should not adapt too much to the training set, and speeds up the training (at least by a very small amount).

Results We show the results of gradient boosting in Table 9.10. As usual the linear dependence of $h^{1,1}$ on the number of projective spaces is evident and in this case also produces the best accuracy result (using the floor function for the original dataset and rounding to the next integer for the favourable dataset) for $h^{1,1}$. $h^{2,1}$ is once again strongly helped by the presence of the redundant features. In Figure 9.17 we finally show the histograms and the scatter plots of the residual errors for the original dataset showing that also in this case the choice of the floor function can be justified and that the addition of the engineered features certainly improves the overall variance of the residuals.

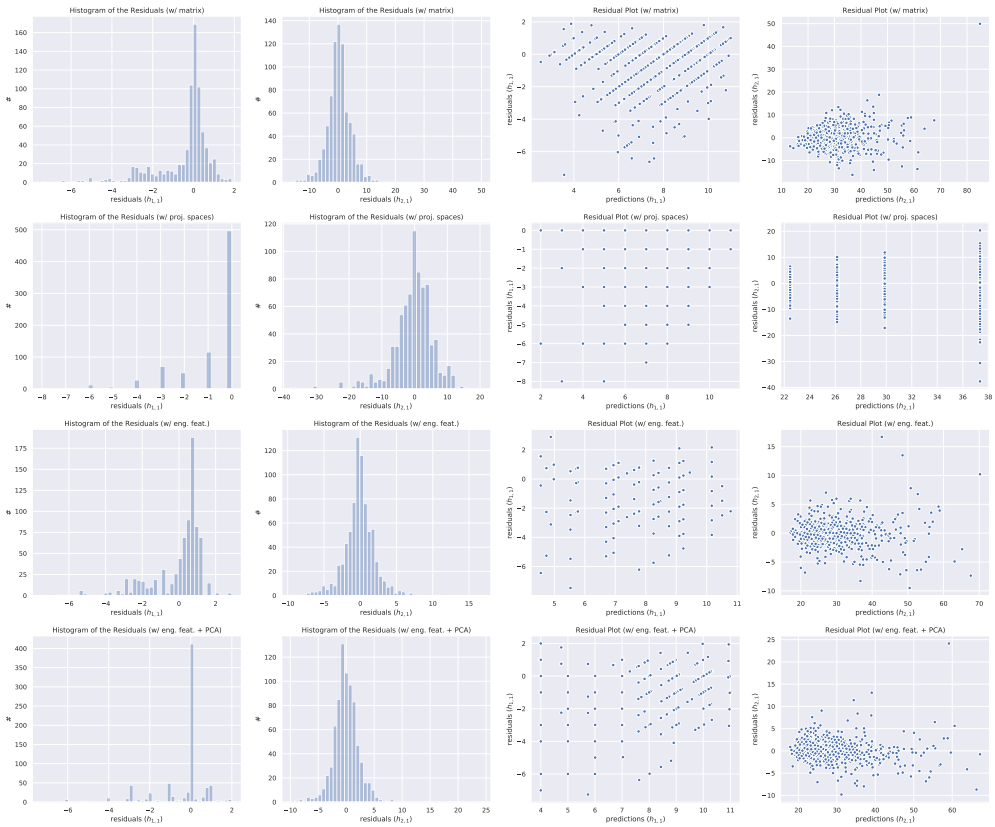


Figure 9.17: Plots of the residual errors for the gradient boosted trees.

9.6 Neural Networks

In this section we approach the problem of predicting the Hodge numbers using artificial neural networks (ANN), which we briefly review in Appendix F.4. We use Google's *Tensorflow* framework and *Keras*, its high-level API, to implement the architectures and train the networks [156]. We explore different architectures and discuss the results.

Differently from the previous algorithms, we do not perform a cross-validation scoring but we simply retain 10 % of the total set as a holdout validation set (also referred to as *development* set) due to the computation power available. Thus we use 80 % of the samples for training, 10 % for evaluation, and 10 % as a test set. For the same reason, the optimisation of the algorithm has been performed manually.

We always use the Adam optimiser with default learning rate 10^{-3} to perform the gradient descent and a fix batch size of 32. The network is trained for a large number of epochs to avoid missing possible local optima. In order to avoid overshooting the minimum of the loss function, we dynamically reduce the learning rate both using the *Adam* optimiser which implements learning rate decay, and through the callback `callbacks.ReduceLROnPlateau` in Keras, which scales the learning rate by a given factor when the monitored quantity (in our case the validation loss) does not decrease): we choose to reduce it by 0.3 when the validation loss does not improve for at least 75 epochs. Moreover we stop training when the validation loss does not improve during 200 epochs. We then keep only the weights of the neural networks which gave the best results. Batch normalisation layers are used with a momentum of 0.99. Training and evaluation were performed on a `NVidia GeForce 940MX` laptop GPU with 2 GB of RAM memory.

9.6.1 Fully Connected Network

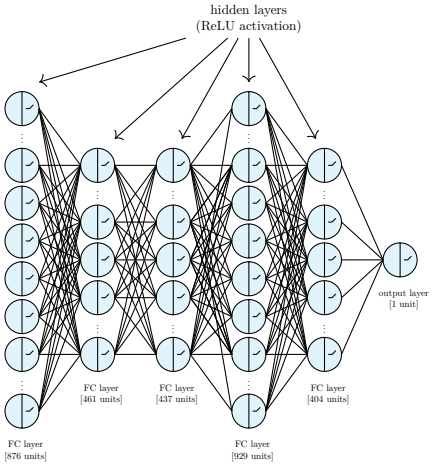
Before we move to the complete deep learning study using convolutional kernels, first we reproduce the analysis in [126] for the prediction of $h^{1,1}$. We will then use this as a baseline for our investigation.

Model The neural network presented in [126] for the regression task contains 5 hidden layers with 876, 461, 437, 929 and 404 units (Figure 9.18a). All layers (including the output layer) are followed by a ReLU activation and by a dropout layer with a rate of 0.2072. This network contains roughly 1.58×10^6 parameters.

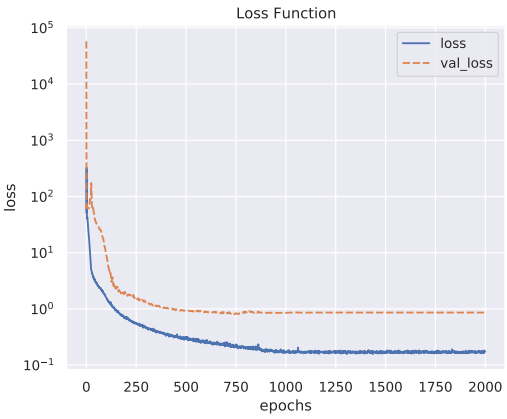
The other hyperparameters (like the optimiser, batch size, number of epochs, regularisation, etc.) are not mentioned. In order to reproduce the results, we fill the gap as follows:

- Adam optimiser with batch size of 32;
- maximal number epochs of 2000 without early stopping;[70]
- we implement learning rate reduction by 0.3 after 75 epochs without improvement of the validation loss;

[70] It took around 20 minutes to train the model.



(a) Architecture of the network.



(b) Loss function on the original dataset.

Figure 9.18: Fully connected network for the prediction of $h^{1,1}$. For simplicity we do not draw the dropout and batch normalisation layers present after every densely connected layer.

	training data				
	10 %	30 %	50 %	70 %	90 %
regression	58 %	68 %	72 %	75 %	75 %
classification	68 %	78 %	82 %	85 %	88 %

Table 9.11: Accuracy (approximate) for $h^{1,1}$ obtained in [126, Figure 1].

- no ℓ_1 or ℓ_2 regularisation;
- a batch normalisation layer [174] after each fully connected layer.

Results We reproduce the results from [126], which are summarised in Table 9.11. The training process was very quick and the loss function is reported in Figure 9.18b. We obtain an accuracy of 77 % both on the development and the test set of the original dataset with 80 % of training data (see Table 9.12). Using the same network we also achieve 97 % of accuracy in the favourable dataset.

9.6.2 Convolutional Network

We then present a new purely convolutional neural network (CNN) to predict $h^{1,1}$ and $h^{2,1}$, separately or together. The advantage of such networks is that it requires a smaller number of parameters and is insensitive to the size of the inputs. The latter point can help to work without padding the matrices (of the same or different representations), but the use of a flatten layer removes this benefit.

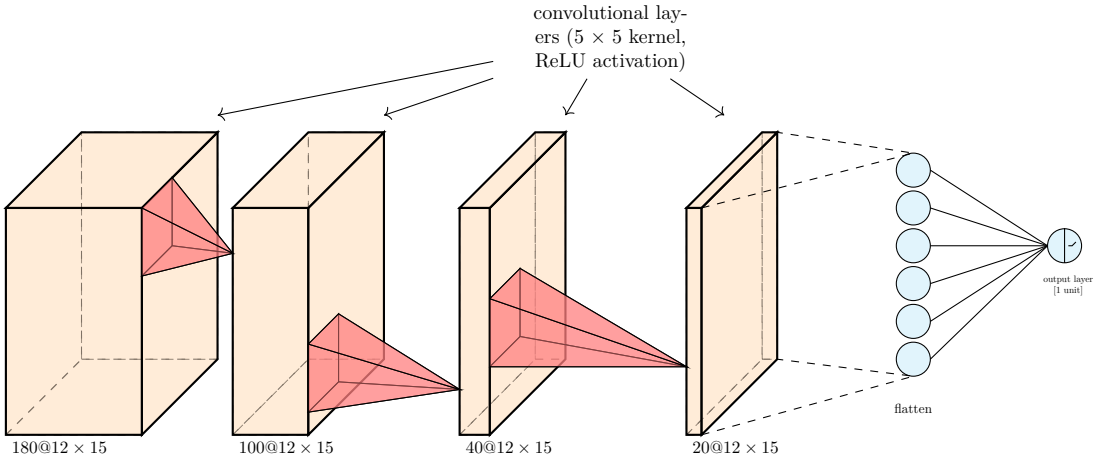
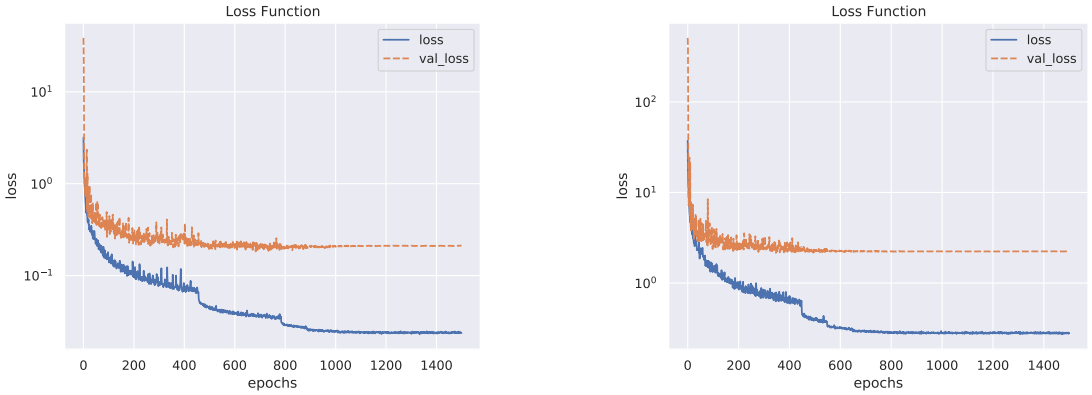


Figure 9.19: Pure convolutional neural network for redicting $h^{1,1}$. It is made of 4 modules composed by convolutional layer, ReLU activation, batch normalisation (in this order), followed by a dropout layer, a flatten layer and the output layer (in this order).

Model The neural network has 4 convolutional layers. They are connected to the output layer with a intermediate flatten layer. After each convolutional layer, we use the ReLU activation function and a batch normalisation layer (with momentum 0.99). Convolutional layers use the padding option `same` and a kernel of size $(5, 5)$ to be able to extract more meaningful representations of the input, treating the configuration matrix somewhat similarly to an object segmentation task [175]. The output layer is also followed by a ReLU activation in order to force the prediction to be a positive number. We use a dropout layer only after the convolutional network (before the flatten layer) but we introduced a combination of ℓ_2 and ℓ_1 regularisation to reduce the variance. The dropout rate is 0.2 in the original dataset and 0.4 for the favourable dataset, while ℓ_1 and ℓ_2 regularisation are set to 10^{-5}. We train the model using the *Adam* optimiser with a starting learning rate of 10^{-3} and a mini-batch size of 32.

The architecture is more similar in style to the old *LeNet* presented for the first time in 1998 by Y. LeCun during the ImageNet competition. In our implementation however we do not include the pooling operations and swap the usual order of batch normalisation and activation function by first putting the ReLU activation. In Figure 9.19 we show the model architecture in the case of the original dataset and of predicting $h^{1,1}$ alone. The convolution layers have 180, 100, 40 and 20 units each.

Results With this setup, we were able to achieve an accuracy of 94 % on both the development and the test sets for the "old" database and 99 % for the favourable dataset in both validation and test sets (results are briefly summarised in Table 9.12). We thus improved the results of the densely connected network and proved that convolutional networks can be valuable assets when dealing with the extraction of a good representation of the input data: not only are convolutional networks very good at recognising patterns and rotationally invariant objects inside pictures or general matrices of data, but deep architectures are also capable of transforming the input using non linear transformations [176] to create new patterns which can then be used for predictions.



(a) Loss function of $h^{1,1}$.

(b) Loss function of $h^{2,1}$.

Figure 9.20: Loss function of the networks for the prediction of $h^{1,1}$ and $h^{2,1}$. We can see that the validation loss flattens out while the training loss keeps decreasing: we took care of the overfit by using the weights of the network when the validation loss reached its minimum. The use of mini-batch gradient descent also completely spoils the monotonicity of the loss functions which can therefore increase moving from one epoch to the other, while keeping the descending trend for most of its evolution.

Even though the convolution operation is very time consuming another advantage of CNN is the extremely reduced number of parameters with respect to fully connected (FC) networks.[71] The architectures we used were in fact made of approximately 5.8×10^5 parameters: way less than half the number of parameters used in the FC network. Ultimately, this leads to a smaller number of training epochs necessary to achieve good predictions (see Figure 9.20).

Using this classic setup we tried different architectures. The network for the original dataset seems to work best in the presence of larger kernels, dropping by roughly 5 % in accuracy when a more "classical" 3×3 kernel is used. We also tried to use to set the padding to `valid`, reducing the input from a 12×15 matrix to a 1×1 feature map over the course of 5 layers with 180, 100, 75, 40 and 20 filters. The advantage is the reduction of the number of parameters (namely $\sim 4.9 \times 10^5$) mainly due to the small FC network at the end, but accuracy dropped to 87 %. The favourable dataset seems instead to be more independent of the specific architecture retaining accuracy also with smaller kernels.

The analysis for $h^{2,1}$ follows the same prescriptions. For both the original and favourable dataset, we opted for 4 convolutional layers with 250, 150, 100 and 50 filters and no FC network for a total amount of 2.1×10^6 parameters. In this scenario we were able to achieve 36 % of accuracy in the development set and 40 % on the test set for $h^{2,1}$ in the "old" dataset and 31 % in both development and test sets in the favourable set (see Table 9.12). The learning curves for both Hodge numbers are given in Figure 9.21. This model uses the same architecture as the one for predicting $h^{1,1}$ only, which explains why it is less accurate as it needs to also adapt to compute $h^{2,1}$ (see for example Figure 9.25).

[71] It took around 4 hours of training (and no optimisation) for each Hodge number in each dataset. The use use of modern generation GPUs with tensor cores can however speed up the training by order of magnitudes.

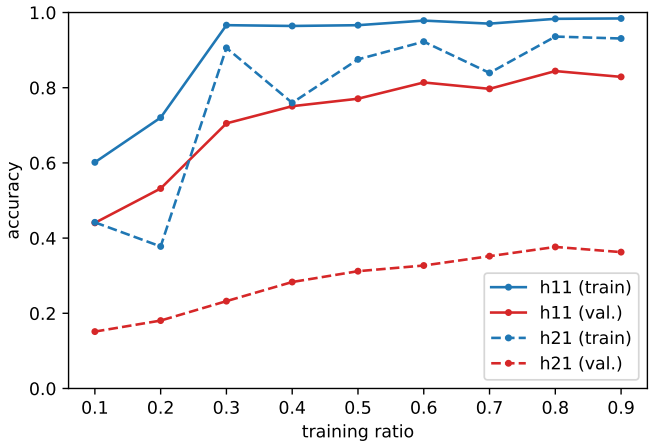


Figure 9.21: Learning curves for the classic convolutional neural network (original dataset), using a single model for both Hodge numbers.

9.6.3 Inception-like Neural Network

In the effort to find a better architecture, we took inspiration from Google's winning CNN in the annual *ImageNet challenge* in 2014 [165]–[167]. The architecture in its original presentation uses *inception* modules in which separate 1×1, 3×3 and 5×5 convolutions are performed side by side (together with *max pooling* operations) before recombining the outputs. The modules are then repeated until the output layer is reached. This has two evident advantages: users can avoid taking a completely arbitrary decision on the type of convolution to use since the network will take care of it tuning the weights, and the number of parameters is extremely restricted as the network can learn complicated functions using fewer layers. As a consequence the architecture of such models can be made very deep while keeping the number of parameters contained, thus being able to learn very difficult representations of the input and producing accurate predictions. Moreover while the training phase might become very long due to the complicated convolutional operations, the small number of parameters is such that predictions can be generated in a very small amount of time making inception-like models extremely appropriate whenever quick predictions are necessary. Another advantage of the architecture is the presence of different kernel sizes inside each module: the network automatically learns features at different scales and different positions thus leveraging the advantages of a deep architecture with the ability to learn different representations at the same time and compare them.

Model In Figure 9.22 we show a schematic of our implementation. Differently from the image classification task, we drop the pooling operation and implement two side-by-side convolution over rows (12×1 kernel for the original dataset, 15×1 for the favourable) and one over columns (1×15 and 1×18 respectively).[72] We use `same` as padding option. The output of the

[72] Pooling operations are used to shrink the size of the input. Similar to convolutions, they use a window of a given size to scan the input and select particular values inside. For instance, we could select the average value inside the small portion selected, performing an *average pooling* operation, or the maximum value, a *max pooling* operation. This usually improves image classification and object detection tasks as it can be used to sharpen edges and borders.

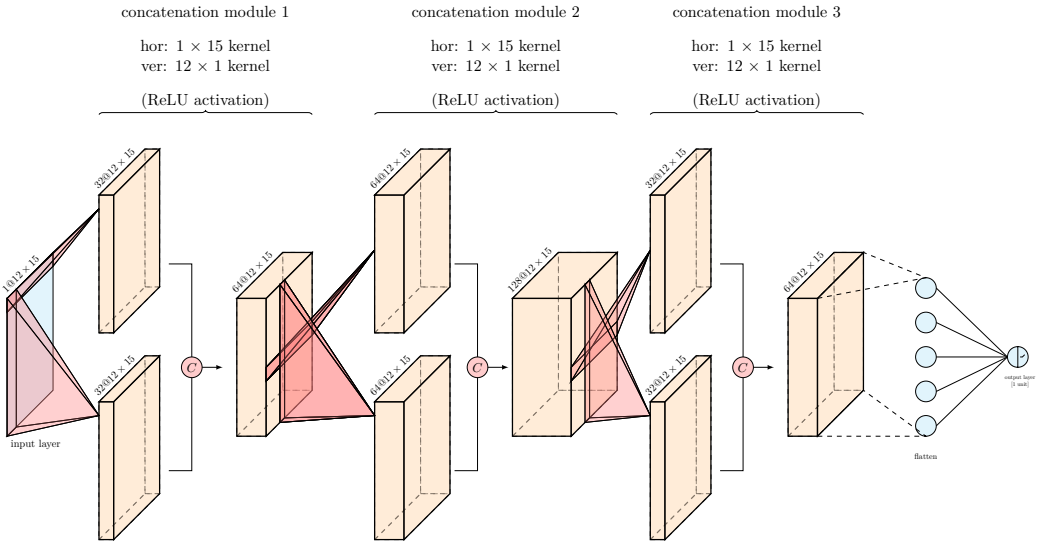


Figure 9.22: In each concatenation module (here shown for the "old" dataset) we operate with separate convolution operations over rows and columns, then concatenate the results. The overall architecture is composed of 3 "inception" modules made by two separate convolutions, a concatenation layer and a batch normalisation layer (strictly in this order), followed by a dropout layer, a flatten layer and the output layer with ReLU activation (in this order).

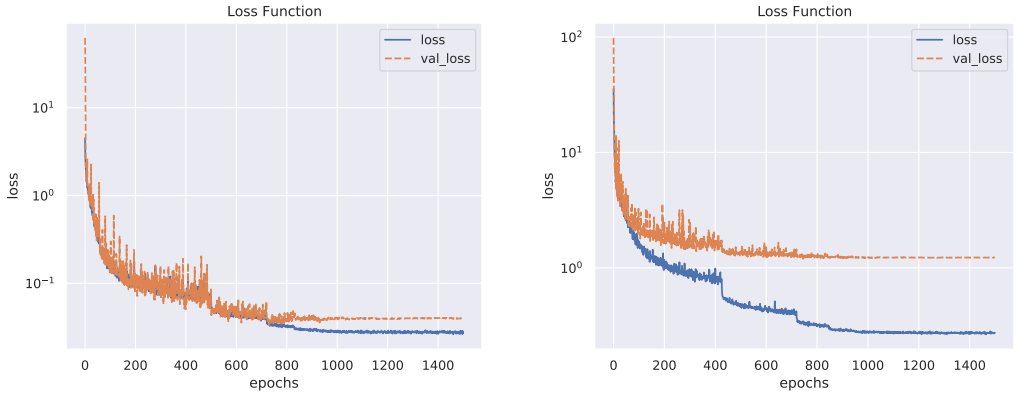
convolutions are then concatenated in the filter dimensions before repeating the "inception" module. The results from the last module are directly connected to the output layer through a flatten layer. In both datasets we use batch normalisation layers (with momentum 0.99) after each concatenation layer and a dropout layer (with rate 0.2) before the FC network.[73]

For both $h^{1,1}$ and $h^{2,1}$ (in both datasets) we used 3 modules made by 32, 64 and 32 filters for the first Hodge number, and 128, 128 and 64 filters for the second. We also included ℓ_1 and ℓ_2 regularisation of magnitude 10^{-4} in all cases. The number of parameters was thus restricted to 2.3×10^5 parameters for $h^{1,1}$ in the original dataset and 2.9×10^5 in the favourable set, and 1.1×10^6 parameters for $h^{2,1}$ in the original dataset and 1.4×10^6 in the favourable dataset. In all cases the number of parameters has decreased by a significant amount: in the case of $h^{1,1}$ they are roughly $\frac{1}{3}$ of the parameters used in the classical CNN and around $\frac{1}{6}$ of those used in the FC network. During training we used the *Adam* gradient descent with an initial learning rate of 10^{-3} and a batch size of 32. The callbacks helped to contain the training time (without optimisation) under 5 hours for each Hodge number in each dataset.

Results With these architectures we were able to achieve more than 99 % of accuracy for $h^{1,1}$ in the test set (same for the development set) and 50 % of accuracy for $h^{2,1}$ (a slightly smaller value for the development set). We report the results in Table 9.12.

We therefore increased the accuracy for both Hodge numbers (especially $h^{2,1}$) compared to

[73] The position of the batch normalisation is extremely important as the parameters computed by such layer directly influence the following batch. We however opted to wait for the scan over rows and columns to finish before normalising the outcome to avoid biasing the resulting activation function.



(a) Loss of $h^{1,1}$.

(b) Loss of $h^{2,1}$.

Figure 9.23: The loss functions of "inception" network for $h^{1,1}$ and $h^{2,1}$ in the original dataset show that the number of epochs required for training is definitely larger than for simpler architectures, despite the reduced number of parameters.

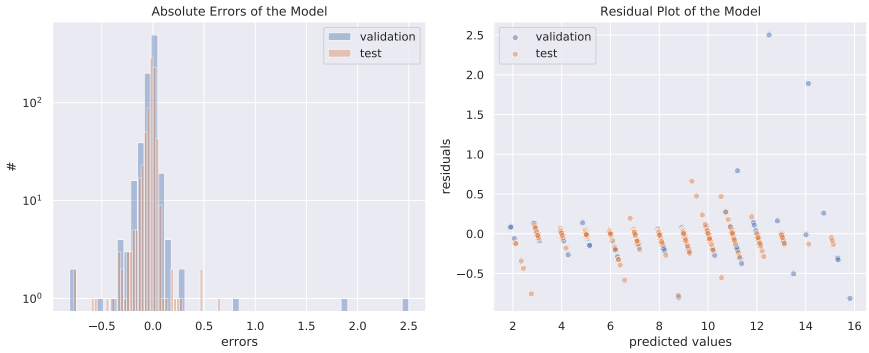
what can achieve a simple sequential network, while at the same time reducing significantly the number of parameters of the network.[74] This increases the robustness of the method and its generalisation properties.

In Figure 9.24 we show the distribution of the residuals and their scatter plot. The distribution of the errors does not present pathological behaviour and the variance of the residuals is well distributed over the predictions. In fact this neural network is much more powerful than the previous networks we considered, as can be seen by studying the learning curves in Figure 9.25. When predicting only $h^{1,1}$ it surpasses 97 % accuracy using only 30 % of the data for training. While it seems that the predictions suffer when using a single network for both Hodge numbers this remains much better than any other algorithm. It may seem counter-intuitive that convolutions work well on this data since they are not translation or rotation invariant but only permutation invariant. However convolution alone is not sufficient to ensure invariances under these transformations but it must be supplemented with pooling operations [141] which we do not use. Moreover convolution layers do more than just taking translation properties into account: they allow to make highly complicated combinations of the inputs and to share weights among components to find subtler patterns than standard fully connected layers. This network is more studied in more details in [4].

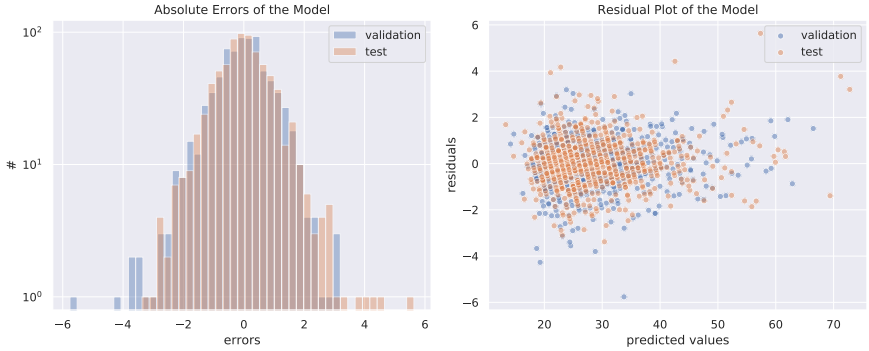
9.6.4 Boosting the Inception-like Model

To improve further the accuracy of $h^{2,1}$ we modify the network by adding engineered features as auxiliary inputs. This can be done by adding inputs to the inception neural network and merging

[74] In an attempt to improve the results for $h^{2,1}$ even further, we also considered to first predict $\ln\left(1+h^{2,1}\right)$ and then transform it back. However, the predictions dropped by almost 10 % in accuracy even using the "inception" network: the network seems to be able to approximate quite well the results (not better nor worse than simply $h^{2,1}$) but the subsequent exponentiation is taking apart predictions and true values. Choosing a correct rounding strategy then becomes almost impossible.

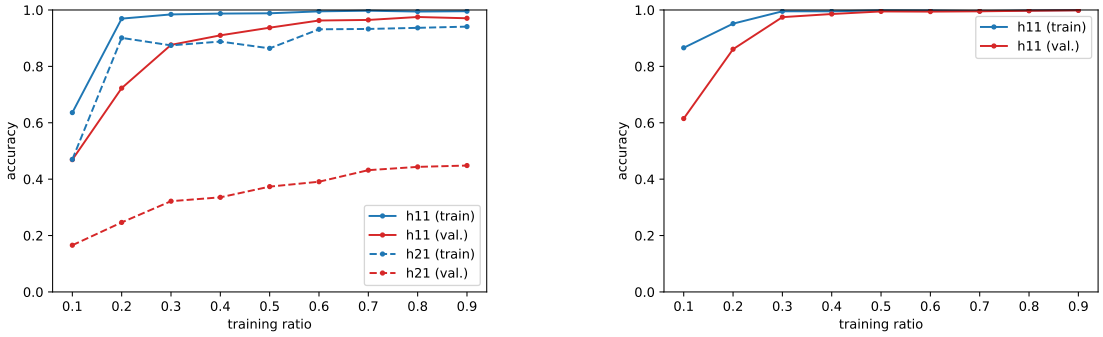


(a) Residuals of $h^{1,1}$.



(b) Residuals of $h^{2,1}$.

Figure 9.24: Histograms of the residual errors and residual plots of the Inception network.



(a) predicting both $h^{1,1}$ and $h^{2,1}$

(b) predicting $h^{1,1}$ only

Figure 9.25: Learning curves for the Inception neural network (original dataset).

	DenseNet		**classic ConvNet**		**inception ConvNet**	
	old	*fav.*	*old*	*fav.*	*old*	*fav.*
$h^{1,1}$	77 %	97 %	94 %	99 %	99 %	99 %
$h^{2,1}$	-	-	36 %	31 %	50 %	48 %

Table 9.12: Accuracy using *rint* rounding on the predictions of the ANNs on $h^{1,1}$ and $h^{2,1}$ on the test set.

the different branches at different stages. There are two possibilities to train such a network: train the whole network directly or train the inception network alone, then freeze its weights and connect it to the additional inputs, training only the new layer. We found that the architectures we tried did not improve the accuracy, but we briefly describe our attempts for completeness. We focused in particular on the number of projective spaces, the vector of dimensions of the projective spaces and the vector of dimensions of the principal cohomology group) and predicting $h^{1,1}$ and $h^{2,1}$ at the same time. The core of the neural network is the Inception network described earlier in Section 9.6.3. The engineered features are processed using fully connected layers and merged to the predictions from the Inception branch using a concatenation layer. Obviously output layers for $h^{1,1}$ and $h^{2,1}$ can be located on different branches which allow for different processing of the features.

As mentioned earlier, a possible approach is to first train the Inception branch alone, before freezing its weights and connecting it to the rest of the network. This can prevent spoiling the already good predictions and speed up the new learning process. This is a common technique called *transfer learning*: we can use a model previously trained on a slightly different task and use its weights as part of the new architecture. Our trials involved shallow fully connected layers (1 to 3 layers with 10 to 150 units) between the engineered features and after the concatenation layer. Since the EDA analysis in Section 9.1 shows a correlation between both Hodge numbers, we tried architectures where the result for $h^{1,1}$ is used to predict $h^{2,1}$. For the training phase we also tried an alternative to the canonical choice of optimising the sum of the losses. We first train the network and stop the process when the validation loss for $h^{1,1}$ does not longer improve, load back the best weights and save the results, keep training and stop when the loss for $h^{2,1}$ reaches a plateau.

With this setup we were able to slightly improve the predictions of $h^{1,1}$ in the original dataset, reaching almost 100 % of accuracy in the predictions, while the favourable dataset stayed at around 99 % of accuracy. The only few missed predictions (4 manifolds out of 786 in the test set) are in very peculiar regions of the distribution of the Hodge number. For $h^{2,1}$ no improvement has been noticed.

9.7 Ensemble Learning: Stacking

We conclude the ML analysis by describing a method very popular in ML competitions: ensembling. This consists in taking several ML algorithms and combining together the predictions of each individual model to obtain a more precise predictions. Using this technique it is possible to decrease the variance and improve generalization by compensating weaknesses of algorithms with strengths of others. Indeed the idea is to put together algorithms which perform best in

		$h^{1,1}$		$h^{2,1}$	
		old	*fav.*	*old*	*fav.*
1st level	EN	65 %	100 %	19 %	19 %
	SVM	70 %	100 %	30 %	34 %
	RF	61 %	98 %	18 %	24 %
	ANN	98 %	98 %	33 %	30 %
2nd level	Lasso	98 %	98 %	36 %	33 %

Table 9.13: Accuracy of the first and second level predictions of the stacking ensemble for elastic net regression (EN), support vector with `rbf` kernel (SVM), random forest (RF) and the artificial neural network (ANN) as first level learners and lasso regression as meta learner.

different zones of the label distribution in order to combine them to build an algorithm better than any individual component. The simplest such algorithm is *stacking* whose principle is summarised in Figure 9.26. First the original training set is split in two parts (not necessarily even). Second a certain number of *first-level learners* is trained over the first split and used to generate predictions over the second split. Third a "meta learner" is trained of the second split to combine the predictions from the first-level learners. Predictions for the test set are obtained by applying both level of models one after the other.

We have selected the following models for the first level: linear regression, SVM with the Gaussian kernel, the random forest and the "inception" neural network. The meta-learner is a simple linear regression with ℓ_1 regularisation (Lasso). The motivations for the first-level algorithms is that stacking works best with a group of algorithms which work in the most diverse way among them. Also in this case, we use a cross-validation strategy with 5 splits for each level of the training: from 90 % of total training set, we split into two halves containing each 45 % of the total samples and then use 5 splits to grade the algorithm, thus using 9 % of each split for cross correlation at each iteration) and the Bayes optimisation for all algorithms but the ANN (50 iterations for elastic net, SVM and lasso and 25 for the random forests). The ANN was trained using a holdout validation set containing the same number of samples as each cross-validation fold, namely 9 % of the total set. The accuracy is then computed as usual using `numpy.rint` for SVM, neural networks, the meta learner and $h^{1,1}$ in the original dataset in general, and `numpy.floor` in the other cases.

In Table 9.13, we show the accuracy of the ensemble learning. We notice that accuracy improves slightly only for $h^{2,1}$ (original dataset) compared to the first-level learners. However this is much lower than what has been achieved in Section 9.6.3. The reason is that the learning suffers from the reduced size of the training set. Another reason is that the different algorithms may perform similarly well in the same regions.

10 Summary and Conclusion

We have proved that a proper data analysis can lead to improvements in predictions of Hodge numbers $h^{1,1}$ and $h^{2,1}$ for CICY 3-folds. Moreover more complex neural networks inspired by computer vision applications [165]–[167] allowed us to reach close to 100 % accuracy for $h^{1,1}$ with

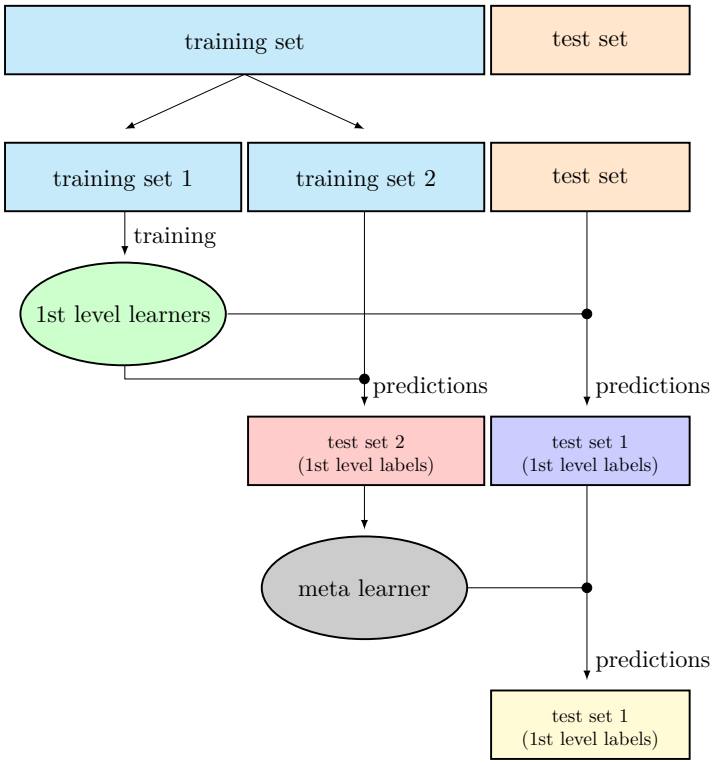


Figure 9.26: Stacking ensemble learning with two level learning.

much less data and less parameters than in previous works. While our analysis improved the accuracy for $h^{2,1}$ over what can be expected from a simple sequential neural network, we barely reached 50 %. Hence it would be interesting to push further our study to improve the accuracy. Possible solutions would be to use a deeper Inception network, find a better architecture including engineered features, and refine the ensembling.

Another interesting question to probe is related to representation learning, i.e. finding a better description of the CY. Indeed one of the main difficulty in making predictions is the redundancy of the possible descriptions of a single manifold. For instance we could try to set up a map from any matrix to its favourable representation (if it exists). This could be the basis for the use of adversarial networks [177] capable of generating the favourable embedding from the first. Or on the contrary one could generate more matrices for the same manifold in order to increase the size of the training set. Another possibility is to use the graph representation of the configuration matrix to which is automatically invariant under permutations [157] (another graph representation has been decisive in [136] to get a good accuracy). Techniques such as (variational) autoencoders [178], [179], cycle GAN [180], invertible neural networks [181], graph neural networks [182], [183] or techniques from geometric deep learning [184] could be helpful.

Finally our techniques apply directly to CICY 4-folds [185], [186]. However there are many more manifolds in this case (around 10^6) and more Hodge numbers, such that one can expect to reach a better accuracy for the different Hodge numbers (the different learning curves for the 3-folds indicate that the model training would benefit from more data). Another interesting class of manifolds to explore with our techniques are generalized CICY 3-folds [187]. These and others will indeed be ground for future investigations.

Part IV
Appendix

A The Isomorphism in Details

In this appendix we explain the conventions used for SU(2) and show the details of the isomorphism between SO(4) and a class of equivalence of SU(2) × SU(2).

A.1 Conventions

We parameterise SU(2) matrices U with a vector $\vec{n} \in \mathbb{R}^3$ such that:

$$U(\vec{n}) = \cos(2\pi n)\, \mathbb{1}_2 + i\, \frac{\vec{n}\cdot\vec{\sigma}}{n}\, \sin(2\pi n), \tag{A.1}$$

where $n = \|\vec{n}\|$ and $0 \le n \le \frac{1}{2}$. We also identify all $\vec{n}$ when $n = \frac{1}{2}$ since in this case $U(\vec{n}) = -\mathbb{1}_2$. The parametrisation is such that:

$$\begin{aligned} U^*(\vec{n}) &= \sigma^2\, U(\vec{n})\, \sigma^2 = U(\widetilde{\vec{n}}), & \text{(A.2)}\\ U^\dagger(\vec{n}) &= U^T(\widetilde{\vec{n}}) = U(-\vec{n}), & \text{(A.3)}\\ -U(\vec{n}) &= U(\widehat{\vec{n}}) & \text{(A.4)} \end{aligned}$$

where σ^2 is the second Pauli matrix, $\widetilde{\vec{n}} = \left(-n^1, n^2, -n^3\right)$ and $\widehat{\vec{n}} = -\left(\frac{1}{2} - n\right)\frac{\vec{n}}{n}$.

The group product of two elements $U(\vec{n}\circ\vec{m}) = U(\vec{n})\, U(\vec{m})$ has an explicit realisation as:

$$\begin{aligned} \cos(2\pi\|\vec{n}\circ\vec{m}\|) &= \cos(2\pi n)\, \cos(2\pi m) - \sin(2\pi n)\, \sin(2\pi m)\, \frac{\vec{n}\cdot\vec{m}}{n\, m},\\ \sin(2\pi\|\vec{n}\circ\vec{m}\|)\, \frac{\vec{n}\circ\vec{m}}{\|\vec{n}\circ\vec{m}\|} &= \cos(2\pi n)\, \sin(2\pi m)\, \frac{\vec{m}}{m} + \sin(2\pi n)\, \cos(2\pi m)\, \frac{\vec{n}}{n}. \end{aligned} \tag{A.5}$$

A.2 The Isomorphism

Let $I = 1, 2, 3, 4$ and define:

$$\tau_I = (i\, \mathbb{1}_2,\, \vec{\sigma}), \tag{A.6}$$

where $\vec{\sigma} = \left(\sigma^1,\, \sigma^2,\, \sigma^3\right)$ are the Pauli matrices. It is possible to show that:

$$\begin{aligned} (\tau_I)^\dagger &= \eta_{IJ}\, \tau^I,\\ \left(\tau^I\right)^* &= -\sigma_2\, \tau_I\, \sigma_2, \end{aligned} \tag{A.7}$$

where $\eta_{IJ} = \mathrm{diag}(-1, 1, 1, 1)$. The following relations are then a natural consequence:

$$\begin{aligned} \mathrm{tr}(\tau_I) &= 2\, i\, \delta_{I1}, & \text{(A.8)}\\ \mathrm{tr}(\tau_I \tau_J) &= 2\, \eta_{IJ}, & \text{(A.9)}\\ \mathrm{tr}\Big(\tau_I (\tau_J)^\dagger\Big) &= 2\, \delta_{IJ}. & \text{(A.10)} \end{aligned}$$

Now consider a vector in the spinor representation:

$$X_{(s)} = X^I\, \tau_I. \tag{A.11}$$

We can recover the components using the previous properties:

$$X^I = \frac{1}{2}\,\delta^{IJ}\ \mathrm{tr}\Big(X_{(s)}(\tau_J)^\dagger\Big) = \frac{1}{2}\,\eta^{IJ}\ \mathrm{tr}\big(X_{(s)}\tau_J\big), \tag{A.12}$$

where the trace acts on the space of the τ matrices. If the vector X^I is real, using (A.7) we have:

$$\begin{aligned} X^\dagger_{(s)} &= X^I\,\eta_{IJ}\,\tau^J = \frac{1}{2}\,\mathrm{tr}\big(X_{(s)}\tau_I\big)\,\tau^I, \\ X^*_{(s)} &= -\sigma_2\,X_{(s)}\,\sigma_2. \end{aligned} \tag{A.13}$$

A rotation in spinor representation is defined as:

$$X'_{(s)} = U_L(\vec{n})\,X_{(s)}\,U_R^\dagger(\vec{m}) \tag{A.14}$$

and it is equivalent to:

$$(X')^I = R^I{}_J\,X^J \tag{A.15}$$

through

$$R_{IJ} = \frac{1}{2}\,\mathrm{tr}\Big((\tau_I)^\dagger\,U_L(\vec{n})\,\tau_J\,U_R^\dagger(\vec{m})\Big). \tag{A.16}$$

The matrix R is the 4-dimensional rotation matrix we are looking for since:

$$\mathrm{tr}\Big(X'_{(s)}\,(X')^\dagger_{(s)}\Big) = \mathrm{tr}\Big(X_{(s)}\,X^\dagger_{(s)}\Big) \qquad \Rightarrow \qquad \sum_{K=1}^{4} R_{IK}R^*_{JK} = \delta_{I\,J}. \tag{A.17}$$

From the second equation in (A.7) and the first equation in (A.4) we then get the reality condition on R:

$$R_{NM} = \frac{1}{2}\,\eta_{NI}\,\eta_{MJ}\ \mathrm{tr}\Big(\tau_I^\dagger\,U_R\,\tau_J\,U_L^\dagger\Big) = \frac{1}{2}\,\mathrm{tr}\Big(\tau_N\,U_R\,\tau_M^\dagger\,U_L^\dagger\Big) = R^*_{NM}. \tag{A.18}$$

Furthermore the direct computation of the determinant of R using the parametrisation (A.1) shows that $\det R = 1$. Finally the explicit choice of the basis τ ensures R to be a real matrix which ensures $R \in \mathrm{SO}(4)$. Since $\{U_L, U_R\}$ and $\{-U_L, -U_R\}$ generate the same SO(4) matrix then the correct isomorphism takes the form:

$$\mathrm{SO}(4) \cong \frac{\mathrm{SU}(2)\times\mathrm{SU}(2)}{\mathbb{Z}_2}. \tag{A.19}$$

B The Parameters of the Hypergeometric Function

In this appendix we show the computation of the parameters of the hypergeometric functions and their relation with the rotation parameters.

B.1 Consistency Conditions of the Monodromy Matrices

In the main text we set

$$D\ \mathrm{M}_\infty\ D^{-1} = e^{-2\pi i\delta_\infty}\,\mathcal{L}(\vec{n}_\infty), \tag{B.1}$$

where $\mathcal{L}(\vec{n}_\infty) \in \mathrm{SU}(2)$. The previous equation implies

$$\left(D\,\mathrm{M}_\infty\,D^{-1}\right)^\dagger = \left(D\,\mathrm{M}_\infty\,D^{-1}\right)^{-1}, \tag{B.2}$$

which can be rewritten as

$$\widetilde{\mathrm{M}}_\infty^{-1}\;\mathcal{C}^\dagger\,D^\dagger\,D\,\mathcal{C} = \mathcal{C}^\dagger\,D^\dagger\,D\,\mathcal{C}\;\widetilde{\mathrm{M}}_\infty^{-1}. \tag{B.3}$$

As $\widetilde{\mathrm{M}}_\infty$ is a generic diagonal matrix, the previous equation implies that the off-diagonal elements of $\mathcal{C}^\dagger\,D^\dagger\,D\,\mathcal{C}$ must vanish. We therefore have

$$\begin{aligned} |K|^{-2} &= -\frac{\mathcal{C}_{21}\,\mathcal{C}_{22}^*}{\mathcal{C}_{11}\,\mathcal{C}_{12}^*} \\ &= -\frac{1}{\pi^4}\left|\Gamma\left(a\right)\Gamma\left(b\right)\Gamma\left(c-a\right)\Gamma\left(c-b\right)\right|^2\times \\ &\times \sin(\pi a)\,\sin^*(\pi(c-a))\,(\sin(\pi b)\,\sin^*(\pi(c-b)))^*. \end{aligned} \tag{B.4}$$

When a, b, $c \in \mathbb{R}$ this ultimately means that

$$\sin(\pi a)\,\sin(\pi(c-a))\,\sin(\pi b)\,\sin(\pi(c-b)) < 0. \tag{B.5}$$

Since the previous equation is invariant under integer shift of any of its parameters, we can consider just the fractional parts $0 \leq \{a\},\,\{b\},\,\{c\} < 1$. In order to have U(2) monodromies finally requires

$$0 \leq \{b\} < \{c\} < \{a\} < 1 \quad \text{or} \quad 0 \leq \{a\} < \{c\} < \{b\} < 1. \tag{B.6}$$

Should we request $\mathrm{U}(1,1)$ monodromies as in moving rotated branes then we get:

$$|K|^{-2} = \frac{\mathcal{C}_{21}\,\mathcal{C}_{22}^*}{\mathcal{C}_{11}\,\mathcal{C}_{12}^*}. \tag{B.7}$$

This would then imply

$$0 \leq \{c\} < \{a\},\,\{b\} < 1 \quad \text{or} \quad 0 \leq \{a\},\,\{b\} < \{c\} < 1. \tag{B.8}$$

B.2 Fixing the Parameters

We can finally show in details the computation of the parameters of the basis of hypergeometric functions used in the main text. The relation between these and the SU(2) matrices can be computed requiring that the monodromies induced by the choice of the parameters equal the monodromies produced by the rotations of the D-branes.

The monodromy in $\omega_{\bar{t}-1} = 0$ is simpler to compute given that we choose $\mathcal{L}(\vec{n}_0)$ and $\mathcal{R}(\widetilde{\vec{m}}_0)$ to be diagonal. We impose:

$$\begin{pmatrix} 1 & \\ & e^{-2\pi i c^{(L)}} \end{pmatrix} = e^{-2\pi i \delta_0^{(L)}} \begin{pmatrix} e^{2\pi i n_0} & \\ & e^{-2\pi i n_0} \end{pmatrix}, \tag{B.9}$$

$$\begin{pmatrix} 1 & \\ & e^{-2\pi i c^{(R)}} \end{pmatrix} = e^{-2\pi i \delta_0^{(R)}} \begin{pmatrix} e^{-2\pi i m_0} & \\ & e^{2\pi i m_0} \end{pmatrix}, \tag{B.10}$$

where $n_0^3 = \|\vec{n}_0\| = n_0$ and $m_0^3 = \|\vec{m}_0\| = m_0$ with $0 \leq n_0,\, m_0 < 1$ due to the conventions (2.61) and (2.62). We thus have:

$$\begin{aligned} \delta_0^{(L)} &= n_0 + k_{\delta_0^{(L)}}, \qquad k_{\delta_0^{(L)}} \in \mathbb{Z}, \\ c^{(L)} &= 2n_0 + k_c, \qquad k_c \in \mathbb{Z}. \end{aligned} \tag{B.11}$$

Since the determinant of the right hand side is $e^{-4\pi i \delta_0^{(L)}}$, the range of definition of $\delta_0^{(L)}$ is $\alpha \leq \delta_0^{(L)} \leq \alpha + \frac{1}{2}$. Given that $0 \leq n_0 < \frac{1}{2}$ we simply take $\alpha = 0$ and set $\delta_0^{(L)} = n_0$. Analogous results hold in the right sector. Furthermore from the third equation in (2.54) and from the first equation in (B.11) we can restrict:

$$n_0 + m_0 - A \in \mathbb{Z}. \tag{B.12}$$

We then need to find 3 equations to determine $a^{(L)}$, $b^{(L)}$ and $\delta_\infty^{(L)}$. After that we then fix the remaining factors in B and $\left|K^{(L)}\right|$. The equations follow from (2.55). The first two equations for $a^{(L)}$, $b^{(L)}$ and $\delta_\infty^{(L)}$ follow by considering the trace of (2.55):

$$e^{\pi i (a^{(L)} + b^{(L)})} \cos\Big(\pi(a^{(L)} - b^{(L)})\Big) = e^{-2\pi i \delta_\infty^{(L)}} \cos(2\pi n_\infty), \tag{B.13}$$

which is satisfied by:

$$\begin{aligned} \delta_\infty^{(L)} &= -\frac{1}{2}(a^{(L)} + b^{(L)}) + \frac{1}{2} k_{\delta_\infty^{(L)}}, \qquad k_{\delta_\infty} \in \mathbb{Z}, \\ a^{(L)} - b^{(L)} &= 2\,(-1)^{p^{(L)}}\, n_\infty + (-1)^{q^{(L)}}\, k_{\delta_\infty^{(L)}} + 2\,k'_{ab}, \qquad k'_{ab} \in \mathbb{Z}, \end{aligned} \tag{B.14}$$

where $p^{(L)},\, q^{(L)} \in \{0, 1\}$. Notice that changing the value of $p^{(L)}$ corresponds to swapping a and b: since the hypergeometric function is symmetric in those parameters we can fix $p^{(L)} = 0$. Redefining k' we can always set $q^{(L)} = 0$. We therefore have:

$$a^{(L)} - b^{(L)} = 2\, n_\infty + k_{\delta_\infty^{(L)}} + 2k_{ab}, \qquad k_{ab} \in \mathbb{Z}. \tag{B.15}$$

The allowed values for $k_{\delta_\infty^{(L)}}$ follow a construction similar to the monodromy around $\omega_{\bar{t}-1} = 0$. The main difference is given by the fact that $\frac{1}{2}(a^{(L)} + b^{(L)})$ may a priori take values in an interval of width 1. As in the previous case we have $\alpha \leq \delta_\infty^{(L)} \leq \alpha + \frac{1}{2}$ with α technically arbitrary. We cannot thus choose a vanishing $k_{\delta_\infty^{(L)}}$ but we have to consider $k_{\delta_\infty^{(L)}} = 0,\, 1$.

We find a third relation by considering the entry

$$\mathrm{Im}\left(e^{+2\pi i \delta_\infty^{(L)}}\, D^{(L)}\, \mathrm{M}_\infty^{(L)} \left(D^{(L)}\right)^{-1}\right)_{11} = \mathrm{Im}\left(\mathcal{L}(n_\infty)\right)_{11}. \tag{B.16}$$

Using

$$\det \mathcal{C} = \frac{\sin\big(\pi c^{(L)}\big)}{\sin\big(\pi(a^{(L)} - b^{(L)})\big)}, \tag{B.17}$$

and the second equation in (B.11) and (B.15) leads to:

$$\cos\Big(\pi(a^{(L)} + b^{(L)} - c^{(L)})\Big) = (-1)^{k_c + k_{\delta_\infty^{(L)}}} \cos\Big(2\pi \mathcal{A}^{(L)}\Big), \tag{B.18}$$

where

$$\cos\Big(2\pi\mathcal{A}^{(L)}\Big) = \cos(2\pi n_0)\,\cos(2\pi n_\infty) - \sin(2\pi n_0)\,\sin(2\pi n_\infty)\,\frac{n_\infty^3}{n_\infty}. \tag{B.19}$$

This expression is connected with rotation parameter in the third interaction point $\omega_{\bar{t}+1} = 1$. In fact $\cos\big(2\pi\mathcal{A}^{(L)}\big) = \cos(2\pi n_1)$. We then write

$$a^{(L)} + b^{(L)} - c^{(L)} = 2\,(-1)^{f^{(L)}}\,n_1 + k_c + k_{\delta_\infty^{(L)}} + 2\,k_{abc}, \qquad k_{abc} \in \mathbb{Z}, \tag{B.20}$$

with $f^{(L)} \in \{0, 1\}$. The request

$$A + B - n_0 - m_0 - (-1)^{f^{(L)}}\,n_1 - (-1)^{f^{(R)}}\,m_1 \in \mathbb{Z} \tag{B.21}$$

finally fixes the B parameter in the third equation of (2.55).

So far we can summarise the results in

$$a = n_0 + (-1)^{f^{(L)}}\,n_1 + n_\infty + m_a, \qquad m_a \in \mathbb{Z}, \tag{B.22}$$

$$b = n_0 + (-1)^{f^{(L)}}\,n_1 - n_\infty + m_b, \qquad m_b \in \mathbb{Z}, \tag{B.23}$$

$$c = 2\,n_0 + m_c, \qquad m_c \in \mathbb{Z}, \tag{B.24}$$

$$\delta_0^{(L)} = n_0, \tag{B.25}$$

$$\delta_\infty^{(L)} = -n_0 - (-1)^{f^{(L)}}\,n_1 + m_c + 2\,m_\delta, \qquad m_\delta \in \mathbb{Z}, \tag{B.26}$$

$$A = n_0 + m_0 + m_A, \qquad m_A \in \mathbb{Z}, \tag{B.27}$$

$$B = (-1)^{f^{(L)}}\,n_1 + (-1)^{f^{(R)}}\,m_1 + m_B, \qquad m_B \in \mathbb{Z}. \tag{B.28}$$

$K^{(L)}$ is finally determined from

$$\Big(D^{(L)}\,\mathrm{M}_\infty\,\Big(D^{(L)}\Big)^{-1}\Big)_{21} = e^{-2\pi i\delta_\infty^{(L)}}\,(\mathcal{L}(n_\infty))_{21}, \tag{B.29}$$

and get:

$$K^{(L)} = -\frac{(-1)^{m_a+m_b+m_c}}{2\pi^2}\,\mathcal{G}(a^{(L)},\,b^{(L)},\,c^{(L)})\,\sin(2\pi n_0)\sin(2\pi n_\infty)\frac{n_\infty^1 + i\,n_\infty^2}{n_\infty}, \tag{B.30}$$

where $\mathcal{G}(a,\,b,\,c) = \Gamma\,(1-a)\;\Gamma\,(1-b)\;\Gamma\,(a+1-c)\;\Gamma\,(b+1-c)$.

B.3 Checking the Consistency of the Solution

We check the consistency condition (B.6) using (A.5). The result is

$$\begin{aligned}\Big(K^{(L)}\Big)^{-1} = {} & \frac{(-1)^{m_a+m_b+m_c}}{2\pi^2}\,\mathcal{G}(1-a^{(L)},\,1-b^{(L)},\,2-c^{(L)}) \\ & \times \sin(2\pi n_0)\,\sin(2\pi n_\infty)\,\frac{n_\infty^1 - i n_\infty^2}{n_\infty},\end{aligned} \tag{B.31}$$

where the function $\mathcal{G}(a,\,b,\,c)$ was defined at the end of the previous section. Compatibility with (B.30) requires

$$\frac{(n_\infty^1)^2 + (n_\infty^2)^2}{n_\infty^2} = -4\frac{\sin(\pi a)\sin(\pi(c-a))\sin(\pi b)\sin(\pi(c-b))}{\sin^2(\pi c)\sin^2(\pi(a-b))}. \tag{B.32}$$

We can then rewrite (B.19) as

$$\frac{(n_\infty^3)^2}{n_\infty^2} = \frac{(\cos(\pi(a-b))\cos(\pi c) - \cos(\pi(a+b-c)))^2}{\sin^2(\pi c)\sin^2(\pi(a-b))}. \tag{B.33}$$

It is then possible to verify that the sum of the left and right hand sides of (B.32) and the last equation are equal to 1. The same consistency check can also be performed by computing $K^{(L)}$ from

$$\left(D^{(L)}\,\mathrm{M}_\infty\left(D^{(L)}\right)^{-1}\right)_{12} = e^{-2\pi i\delta_\infty^{(L)}}\,(\mathcal{L}(n_\infty))_{12}, \tag{B.34}$$

instead of (B.29).

C Reflection Conditions on the Vacuum

We provide details on how (3.171) can be computed. First we introduce the projector of positive frequency and negative frequency modes for the NS fermion as

$$P^{(+,0)}(z,w) = \frac{+1}{z-w}, \qquad |z|>|w| \tag{C.1}$$

$$P^{(-,0)}(z,w) = \frac{-1}{z-w}, \qquad |z|<|w|, \tag{C.2}$$

such that

$$\oint_{|z|>|w|}\frac{\mathrm{d}w}{2\pi i}P^{(+,0)}(z,w)\,\Psi^{(0)}(0) = \Psi^{(0,+)}(z), \tag{C.3}$$

and similarly for the negative frequency modes. Likewise we introduce the projectors for the field with defects as

$$P^{(+)}(z,\,w) = \frac{P\big(z;\,\{x_{(t)},\,\mathrm{E}_{(t)}\}\big)\,P\big(w;\,\{x_{(t)},\,-\mathrm{E}_{(t)}\}\big)}{z-w}, \qquad |z|>|w| \tag{C.4}$$

$$P^{(-)}(z,\,w) = -\frac{P\big(z;\,\{x_{(t)},\,\mathrm{E}_{(t)}\}\big)\,P\big(w;\,\{x_{(t)},\,-\mathrm{E}_{(t)}\}\big)}{z-w}, \qquad |z|<|w|, \tag{C.5}$$

with $P\big(z;\,\{x_{(t)},\,\mathrm{E}_{(t)}\}\big) = \prod_{t=1}^{N}\left(1-\frac{z}{x_{(t)}}\right)^{\mathrm{E}_{(t)}}$ as in the main text.

We then compute

$$\begin{aligned}\left(P^{(+)}\,P^{(+,0)}\right)(z,\,w) &= \oint_{|z|>|\zeta|>|w|}\frac{\mathrm{d}z}{2\pi i}P^{(+)}(z,\,\zeta)\,P^{(+,0)}(\zeta,\,w) = P^{(+,0)}(z,\,w)\\ \left(P^{(+)}\,P^{(-,0)}\right)(z,\,w) &= \frac{P\big(z;\,\{x_{(t)},\,\mathrm{E}_{(t)}\}\big)\,P\big(w;\,\{x_{(t)},\,-\mathrm{E}_{(t)}\}\big)-1}{z-w}.\end{aligned} \tag{C.6}$$

The last equation is valid when $\mathrm{M} = \sum_{t=1}^{N}\mathrm{E}_{(t)} \le 0$ and for $|z|$ and $|w|$ arbitrary. Specializing the previous expressions to $\Psi^{(out)}(z)$, we need to constrain $|z|>x_{(1)}$ and $|w|>x_{(1)}$.

Finally the vacuum in presence of defects can be described by

$$\begin{aligned}
\Psi^{(+)}(z)\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)},\,\overline{\mathrm{E}}_{(t)}\right\}}\right\rangle &= \Big(P^{(+)}\,\Psi\Big)(z)\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)},\,\overline{\mathrm{E}}_{(t)}\right\}}\right\rangle \\
&= \Big(P^{(+)}\,\Psi^{(out)}\Big)(z)\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)},\,\overline{\mathrm{E}}_{(t)}\right\}}\right\rangle \\
&= \Big\{\Big(P^{(+)}\,P^{(+,\,0)}\,\Psi^{(out)}\Big)(z) \\
&+ \Big(P^{(+)}\,P^{(-,\,0)}\,\Psi^{(out)}\Big)(z)\Big\}\left|\Omega_{\left\{x_{(t)},\,\mathrm{E}_{(t)},\,\overline{\mathrm{E}}_{(t)}\right\}}\right\rangle \\
&= 0,
\end{aligned} \tag{C.7}$$

where we assumed $|z| > x_{(1)}$. The expression finally becomes (3.171).

D Tensor Wave Functions on NBO

For the sake of completeness we report the expression of the full NBO tensor wave function. In what follows $L = \frac{l}{k_+}$. We have

$$
\begin{aligned}
\begin{pmatrix} S_{u\,u} \\ S_{u\,v} \\ S_{u\,z} \\ S_{u\,i} \\ S_{v\,v} \\ S_{v\,z} \\ S_{v\,i} \\ S_{z\,z} \\ S_{z\,i} \\ S_{i\,i} \end{pmatrix} = \Bigg\{ & \mathcal{S}_{u\,u} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mathcal{S}_{u\,v} \begin{pmatrix} \frac{i}{k_+\,u} + \frac{L^2}{\Delta^2\,u^2} \\ 1 \\ L \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mathcal{S}_{u\,z} \begin{pmatrix} \frac{2\,L}{\Delta\,u} \\ 0 \\ \Delta\,u \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \mathcal{S}_{u\,i} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
& + \mathcal{S}_{v\,v} \begin{pmatrix} -\frac{3}{4\,k_+^2\,u^2} + \frac{3\,i\,L^2}{2\,\Delta^2\,k_+\,u^3} + \frac{L^4}{4\,\Delta^4\,u^4} \\ \frac{i}{2\,k_+\,u} + \frac{L^2}{2\,\Delta^2\,u^2} \\ \frac{3\,i\,L}{2\,k_+\,u} + \frac{L^3}{2\,\Delta^2\,u^2} \\ 0 \\ 1 \\ L \\ 0 \\ \frac{i\,\Delta^2\,u}{k_+} + L^2 \\ 0 \\ 0 \end{pmatrix} + \mathcal{S}_{v\,z} \begin{pmatrix} \frac{3\,i\,L}{\Delta\,k_+\,u^2} + \frac{L^3}{\Delta^3\,u^3} \\ \frac{L}{\Delta\,u} \\ \frac{3\,L^2}{2\,\Delta\,u} + \frac{3\,i\,\Delta}{2\,k_+} \\ 0 \\ 0 \\ \Delta\,u \\ 0 \\ 2\,\Delta\,L\,u \\ 0 \\ 0 \end{pmatrix} \\
& + \mathcal{S}_{v\,i} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{i}{2\,k_+\,u} + \frac{L^2}{2\,\Delta^2\,u^2} \\ 0 \\ 0 \\ 1 \\ 0 \\ L \\ 0 \end{pmatrix} + \mathcal{S}_{z\,z} \begin{pmatrix} \frac{i}{k_+\,u} + \frac{L^2}{\Delta^2\,u^2} \\ 0 \\ L \\ 0 \\ 0 \\ 0 \\ 0 \\ \Delta^2\,u^2 \\ 0 \\ 0 \end{pmatrix} + \mathcal{S}_{z\,i} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{L}{\Delta\,u} \\ 0 \\ 0 \\ 0 \\ 0 \\ \Delta\,u \\ 0 \end{pmatrix} \\
& + \mathcal{S}_{i\,j} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \delta_{ij} \end{pmatrix} \Bigg\} \Phi_{\{k_+, l, \vec{k}, r\}}.
\end{aligned}
\tag{D.1}
$$

E Overlap of Second Level Massive States on NBO

We report the full expression of the overlap with two derivatives considered in the main text. It corresponds to the colour ordered amplitude of two tachyons and one level-2 massive state:

$$\begin{aligned} K = \mathcal{N}^2 & \int \mathrm{d}^D x \ \sqrt{-\det g} \\ & \times \Bigg[u^{-3}\, \mathfrak{s}^{(-3)}_{\{\mathcal{S}\};\left\{k_{(i)+},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\right\}} + u^{-2}\, \mathfrak{s}^{(-2)}_{\{\mathcal{S}\};\left\{k_{(i)+},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\right\}} \\ & + u^{-1}\, \mathfrak{s}^{(-1)}_{\{\mathcal{S}\};\left\{k_{(i)+},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\right\}} + \mathfrak{s}^{(0)}_{\{\mathcal{S}\};\left\{k_{(i)+},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\right\}} \\ & + u\, \mathfrak{s}^{(1)}_{\{\mathcal{S}\};\left\{k_{(i)+},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\right\}} \Bigg] \prod_{j=1}^{3} \Phi_{\left\{k_{(j)+},\, l_{(j)},\, \vec{k}_{(j)},\, r_{(j)}\right\}} \end{aligned} \tag{E.1}$$

where $i = 1,\, 2,\, 3$ and:

$$\begin{aligned} \mathfrak{s}^{(-3)}_{\{\mathcal{S}\},\left\{k_{(i)+},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\right\}} = \Bigg(& -\frac{k^4_{(2)+} + l^4_{(3)} - 4\, k^3_{(2)+}\, k_{(3)+}\, l_{(2)}\, l^3_{(3)}}{4\, k^2_{(2)+}\, k^4_{(3)+} + \Delta^3} \\ & - \frac{6\, k^2_{(2)+}\, k^2_{(3)+} + l^2_{(2)}\, l^2_{(3)} + k^4_{(3)+} + l^4_{(2)}}{4\, k^2_{(2)+}\, k^4_{(3)+} + \Delta^3} \Bigg) \mathcal{S}_{v\,v}, \end{aligned} \tag{E.2}$$

$$\begin{aligned} \mathfrak{s}^{(-2)}_{\{\mathcal{S}\},\left\{k_{(i)+},\, l_{(i)},\, \vec{k}_{(i)},\, r_{(i)}\right\}} = \Bigg(& -\frac{3i\left(k^2_{(2)+}\, k_{(3)} + l^2_{(3)} + k^3_{(2)+}\, l^2_{(3)}\right)}{2\, k_{(2)+}\, k^3_{(3)+} + \Delta} \\ & + \frac{i\left(2\, k_{(2)+}\, k^2_{(3)+}\, l_{(2)}\, l_{(3)} + 3\, k^2_{(2)+}\, k_{(3)+}\, l_{(2)}\, l_{(3)}\right)}{k_{(2)+}\, k^3_{(3)+} + \Delta} \\ & - \frac{3i\left(k^3_{(3)+}\, l^2_{(2)} + k_{(2)+}\, k^2_{(3)+}\, l^2_{(2)}\right)}{2\, k_{(2)+}\, k^3_{(3)+} + \Delta} \Bigg) \mathcal{S}_{v\,v} \\ & - \left(\frac{l_{(3)}\left(k^2_{(2)+}\, l^2_{(3)} - 3\, k_{(2)+}\, k_{(3)+}\, l_{(2)}\, l_{(3)} + 3\, k^2_{(3)+}\, l^2_{(2)}\right)}{k^3_{(3)+}\, \Delta^2} \right) \mathcal{S}_{v\,z}, \end{aligned} \tag{E.3}$$

$$
\begin{aligned}
\mathfrak{s}^{(-1)}_{\{\mathcal{S}\},\left\{k_{(i)+},l_{(i)},\vec{k}_{(i)},r_{(i)}\right\}} = & \left(-\frac{\left(k_{(2)+}l_{(3)}-k_{(3)+}l_{(2)}\right)^2}{k^2_{(3)+}\Delta}\right)\mathcal{S}_{u\,v} \\
& + \left(-\frac{k^2_{(2)+}l^2_{(3)}\left(r_{(2)}+\left\|\vec{k}_{(2)}\right\|^2\right)+k^2_{(3)+}l^2_{(2)}\left(r_{(2)}+\left\|\vec{k}_{(2)}\right\|^2\right)}{2\,k^2_{(2)+}k^2_{(3)+}\Delta}\right. \\
& + \frac{2\,k^3_{(2)+}k_{(3)+}l_{(2)}\,l_{(3)}}{k^2_{(2)+}k^2_{(3)+}\Delta} \\
& \left. + \frac{3\,k^2_{(2)+}k^2_{(3)+}\Delta 6\,k^3_{(2)+}k_{(3)+}\Delta 3\,k^4_{(2)+}\Delta}{4\,k^2_{(2)+}k^2_{(3)+}}\right)\mathcal{S}_{v\,v} \\
& - \left(\frac{i\left(3\,k_{(2)+}k_{(3)+}l_{(3)}+3\,k^2_{(2)+}l_{(3)}\right)}{k^2_{(3)+}}\right. \\
& \left. + \frac{i\left(2\,k^2_{(3)+}l_{(2)}+3\,k_{(2)+}k_{(3)+}l_{(2)}\right)}{k^2_{(3)+}}\right)\mathcal{S}_{v\,z} \\
& + \left(\frac{k_{(2)\,i}\,l_{(3)}\left(k_{(2)+}l_{(3)}-2\,k_{(3)+}l_{(2)}\right)}{k^2_{(3)+}\Delta}\right)\mathcal{S}_{v\,i} \\
& + \left(-\frac{\left(k_{(2)+}l_{(3)}-k_{(3)+}l_{(2)}\right)^2}{k^2_{(3)+}\Delta}\right)\mathcal{S}_{z\,z},
\end{aligned}
\tag{E.4}
$$

$$
\begin{aligned}
\mathfrak{s}^{(0)}_{\{\mathcal{S}\},\left\{k_{(i)+},l_{(i)},\vec{k}_{(i)},r_{(i)}\right\}} = & \left(-\frac{i\,k_{(2)+}\left(k_{(3)+}+k_{(2)+}\right)\Delta}{k_{(3)+}}\right)\mathcal{S}_{u\,v} \\
& + \left(-\frac{2\,k_{(2)+}\left(k_{(2)+}l_{(3)}-k_{(3)+}l_{(2)}\right)}{k_{(3)+}}\right)\mathcal{S}_{u\,z} \\
& + \left(-\frac{i\left(k_{(3)+}+k_{(2)+}\right)\Delta\left(r_{(2)}+\left\|\vec{k}_{(2)}\right\|^2\right)}{2\,k_{(2)+}k_{(3)+}}\right)\mathcal{S}_{v\,v} \\
& + \left(-\frac{l_{(3)}\left(r_{(2)}+\left\|\vec{k}_{(2)}\right\|^2\right)-2\,k_{(2)+}k_{(3)+}l_{(2)}}{k_{(3)+}}\right)\mathcal{S}_{v\,z} \\
& + \left(\frac{i\,k_{(2)\,i}\,k_{(2)+}\Delta}{k_{(3)+}}\right)\mathcal{S}_{v\,i} \\
& + \left(-\frac{i\,k_{(2)+}\left(k_{(3)+}+k_{(2)+}\right)\Delta}{k_{(3)+}}\right)\mathcal{S}_{z\,z} \\
& + \left(\frac{2\,k_{(2)\,i}\left(k_{(2)+}l_{(3)}-k_{(3)+}l_{(2)}\right)}{k_{(3)+}}\right)\mathcal{S}_{z\,i},
\end{aligned}
\tag{E.5}
$$

$$
\begin{aligned}
\mathfrak{s}^{(1)}_{\{\mathcal{S}\},\left\{k_{(i)+},l_{(i)},\vec{k}_{(i)},r_{(i)}\right\}} &= \left(-k^2_{(2)+}\Delta\right)\mathcal{S}_{u\,u} \\
&+ \left(-\Delta\left(r_{(2)}+\left\|\vec{k}_{(2)}\right\|^2\right)\right)\mathcal{S}_{u\,v} \\
&+ \left(2\,k_{(2)\,i}\,k_{(2)+}\Delta\right)\mathcal{S}_{u\,i} \\
&+ \left(-\frac{\Delta\left(r_{(2)}+\left\|\vec{k}_{(2)}\right\|^2\right)^2}{4\,k^2_{(2)+}}\right)\mathcal{S}_{v\,v} \\
&+ \left(2\,k_{(2)\,i}\,k_{(2)+}\Delta\right)\mathcal{S}_{v\,i} \\
&+ \left(-k_{(2)\,i}k_{(2)\,j}\,\Delta\right)\mathcal{S}_{i\,j}.
\end{aligned}
\tag{E.6}
$$

F Machine Learning Algorithms

In this appendix we give a brief review and definition of the main ML algorithms used in the text. We highlight the specific characteristics of interest in the analysis.

F.1 Linear regression

Consider a set of F features $\{x_n\}$ where $n = 1, \ldots, F$. A linear model learns a function

$$
f(x_n) = \sum_{n=1}^{F} w_n x_n + b, \tag{F.1}
$$

where w and b are the *weights* and *intercept* of the fit.

One of the key assumptions behind a linear fit is the independence of the residual error between the predicted point and the value of the model, which can therefore be assumed to be sampled from a normal distribution peaked at the average value [144], [168]. The parameters of the fit are then chosen to maximise their *likelihood* function, or conversely to minimise its logarithm with a reversed sign (the χ^2 function). A related task is to minimise the mean squared error without assuming a statistical distribution of the residual error: ML for regression usually implements this as loss function of the estimators. In this sense loss functions for regression are more general than a likelihood approach but they are nonetheless related. For plain linear regression the associated loss is:

$$
\mathcal{L}(w,\,b) = \frac{1}{2N}\sum_{i=1}^{N}\sum_{n=1}^{F}\left(y^{(i)} - (w_n x_n^{(i)} + b)\right)^2, \tag{F.2}
$$

where N is the number of samples and $x_n^{(i)}$ the nth feature of the i-th sample. The values of the parameters will therefore be:

$$
(w, b) = \underset{w,\,b}{\operatorname{argmin}}\ \mathcal{L}(w, b). \tag{F.3}
$$

This usually requires looping over all samples and all features, thus the *least squares* method has a time complexity of $\mathcal{O}(F \times N)$: while the increase of the number of samples might be an issue, the number of engineered features and matrix components usually does not change and does not represent a huge effort in terms of rescaling the algorithm.

There are however different versions of possible regularisation which we might add to constrain the parameters of the fit and avoid adapting too well to the training set. In particular we may be interested in adding a ℓ_1 regularisation:

$$\mathcal{L}_1(w) = \sqrt{\sum_{n=1}^{F} w_n^2}, \tag{F.4}$$

or the ℓ_2 version:

$$\mathcal{L}_2(w) = \sum_{n=1}^{F} w_n^2. \tag{F.5}$$

Notice that in general we do not regularise the intercept. These terms can be added to the plain loss function to try and avoid large parameters to influence the predictions and to keep better generalisation properties:

- add both ℓ_1 and ℓ_2 regularisation (this is called *elastic net*):

$$\mathcal{L}_{\text{EN}}(w, b;\ \alpha_{\text{EN}}, L) = \mathcal{L}(w, b) + \alpha_{\text{EN}} \cdot L \cdot \mathcal{L}_1(w) + \frac{\alpha_{\text{EN}}}{2} \cdot (1 - L) \cdot \mathcal{L}_2(w), \tag{F.6}$$

- keep only ℓ_1 regularisation (i.e. the *lasso* regression):

$$\mathcal{L}_{\text{LSS}}(w, b;\ \alpha_{\text{LSS}}) = \mathcal{L}(w, b) + \alpha_{\text{LSS}} \cdot \mathcal{L}_1(w), \tag{F.7}$$

- keep only ℓ_2 regularisation (*ridge* regression):

$$\mathcal{L}_{\text{RDG}}(w, b;\ \alpha_{\text{RDG}}) = \mathcal{L}(w, b) + \alpha_{\text{RDG}} \cdot \mathcal{L}_2(w). \tag{F.8}$$

The role of the hyperparameter L is to balance the contribution of the additional terms. For larger values of the hyperparameter α, w (and b) assume smaller values and adapt less to the particular training set.

F.2 Support Vector Machines for Regression

This family of supervised ML algorithms was created with classification tasks in mind [188] but have proven to be effective also for regression problems [189]. Differently from the linear regression, instead of minimising the squared distance of each sample, the algorithm assigns a penalty to predictions of samples $x^{(i)} \in \mathbb{R}^F$ (for $i = 1, 2, \ldots, N$) which are further away than a certain hyperparameter ε from their true value y, allowing however a *soft margin* of tolerance represented by the penalties ζ above and ξ below. This is achieved by minimising w, b, ζ and ξ

in the function:[75]

$$\begin{aligned}\mathcal{L}(w,b,\zeta,\xi) = \frac{1}{2}\sum_{n=1}^{F'} w_n^2 + C\sum_{i=1}^{N}\left(\zeta^{(i)}+\xi^{(i)}\right)\\ +\sum_{i=1}^{N}\sum_{n=1}^{F'}\alpha^{(i)}\left(y^{(i)} - w_n\phi_n\left(x^{(i)}\right) - b - \varepsilon - \zeta^{(i)}\right)\\ +\sum_{i=1}^{N}\sum_{n=1}^{F'}\beta^{(i)}\left(w_n\phi_n\left(x^{(i)}\right) + b - y^{(i)} - \varepsilon - \xi^{(i)}\right)\\ -\sum_{i=1}^{N}\left(\rho^{(i)}\zeta^{(i)} + \sigma^{(i)}\xi^{(i)}\right)\end{aligned} \tag{F.9}$$

where $\alpha^{(i)}$, $\beta^{(i)}$, $\rho^{(i)}$, $\sigma^{(i)} \geq 0$ such that the previous expression encodes the constraints

$$\begin{cases} y^{(i)} - \sum\limits_{n=1}^{F'} w_n\phi_n\left(x^{(i)}\right) - b & \leq \varepsilon + \zeta^{(i)}, \qquad \varepsilon \geq 0, \quad \zeta^{(i)} \geq 0, \quad i = 1,2,\ldots,N \\ \sum\limits_{n=1}^{F'} w_n\phi_n\left(x^{(i)}\right) + b - y^{(i)} & \leq \varepsilon + \xi^{(i)}, \qquad \varepsilon \geq 0, \quad \xi^{(i)} \geq 0, \quad i = 1,2,\ldots,N \end{cases} \tag{F.10}$$

and where $\phi\left(x^{(i)}\right) \in \mathbb{R}^{F'}$ is a function mapping the feature vector $x^{(i)} \in \mathbb{R}^F$ in a higher dimensional space ($F' > F$), whose interpretation will become clear in an instant. The minimisation problem leads to

$$\begin{cases} w_n - \sum\limits_{i=1}^{N}\left(\alpha^{(i)} - \beta^{(i)}\right)\phi_n\left(x^{(i)}\right) = 0 \\ \sum\limits_{i=1}^{N}\left(\alpha^{(i)} - \beta^{(i)}\right) = 0 \\ \sum\limits_{i=1}^{N}\left(\alpha^{(i)} + \rho^{(i)}\right) = \sum\limits_{i=1}^{N}\left(\beta^{(i)} + \sigma^{(i)}\right) = C \end{cases} \tag{F.11}$$

such that $0 \leq \alpha^{(i)}, \beta^{(i)} \leq C$, $\forall\, i = 1,2,\ldots,N$. This can be reformulated as a *dual* problem in finding the extrema of $\alpha^{(i)}$ and $\beta^{(i)}$ in

$$W(\alpha,\beta) = \frac{1}{2}\sum_{i,j=1}^{N}\theta^{(i)}\theta^{(j)}\mathrm{K}(x^{(i)},x^{(j)}) - \varepsilon\sum_{i=1}^{N}\left(\alpha^{(i)}+\beta^{(i)}\right) + \sum_{i=1}^{N} y^{(i)}\theta^{(i)}, \tag{F.12}$$

where $\theta = \alpha - \beta$ are called *dual coefficients* (accessible through the attribute `dual_coef_` of `svm.SVR` in `scikit-learn`) and $\mathrm{K}\left(x^{(i)},x^{(j)}\right) = \sum\limits_{n=1}^{F'}\phi_n\left(x^{(i)}\right)\phi_n\left(x^{(j)}\right)$ is the *kernel* function. Notice that the Lagrange multipliers $\alpha^{(i)}$ and $\beta^{(i)}$ are non vanishing only for particular sets of vectors $l^{(i)}$ which lie outside the ε dependent bounds of (F.10) and operate as landmarks for the others. They are called *support vectors* (accessible using the attribute `support_vectors_` in `svm.SVR`), hence the name of the algorithm. There can be at most N when $\varepsilon \to 0^+$. As a

[75] In a classification task the training objective would be the minimisation of the opposite of the log-likelihood function of predicting a positive class, that is $y^{(i)}\left(w_n\phi_n\left(x^{(i)}\right)+b\right)$, which should equal the unity for good predictions (we can consider $\varepsilon = 1$), instead of the regression objective $y^{(i)} - w_n\phi_n\left(x^{(i)}\right) - b$. The differences between SVM for classification purposes and regression follow as shown.

consequence any sum involving $\alpha^{(i)}$ or $\beta^{(i)}$ can be restricted to the subset of support vectors. Using the kernel notation, the predictions will therefore be

$$y^{(i)}_{\text{pred}} = y_{\text{pred}}\left(x^{(i)}\right) = \sum_{n=1}^{F'} w_n \phi_n\left(x^{(i)}\right) + b = \sum_{a\in A} \theta^{(a)} \mathrm{K}\left(x^{(i)}, l^{(a)}\right) + b, \tag{F.13}$$

where $A \subset \{1, 2, \dots, N\}$ is the subset of labels of the support vectors.

In Section 9.4 we consider two different implementations of the SVM algorithm:

- the *linear kernel*, namely the case when $K \equiv \mathrm{id}$ and the loss, in the `scikit-learn` implementation of `svm.LinearSVR`, can be simplified to

$$\mathcal{L}(w, b) = \mathrm{C} \sum_{i=1}^{N} \sum_{n=1}^{F'} \max\left(0, \left|y^{(i)} - w_n \phi_n\left(x^{(i)} - b\right)\right| - \varepsilon\right) + \frac{1}{2} \sum_{n=1}^{F'} w_j^2, \tag{F.14}$$

  without resolving to the dual formulation of the problem.

- the Gaussian kernel (called `rbf`, from *radial basis function*) in which

$$\mathrm{K}\left(x^{(i)}, l^{(a)}\right) = \exp\left(-\gamma \sum_{n=1}^{F} \left(x_n^{(i)} - l_n^{(a)}\right)^2\right). \tag{F.15}$$

From the definition of the loss function in (F.9) and the kernels, we can appreciate the role of the main hyperparameters of the algorithm. While the interpretation of ε is straightforward as the margin allowed without penalty for the prediction, γ represents the width of the normal distribution used to map the features in the higher dimensional space. Furthermore, C plays a similar role to the l_2 additional term in (F.8) by controlling the entity of the penalty for samples outside the ε-dependent bound, however its relation to the linear regularisation is $\alpha_{ridge} = C^{-1}$, thus $C > 0$ by definition.

Given the nature of the algorithm, support vectors are powerful tools which usually grant better results in both classification and regression tasks with respect to logistic and linear regression, but they scale poorly with the number of samples used during training. In particular the time complexity is at worst $\mathcal{O}\left(F \times N^3\right)$ due to the quadratic nature of (F.12) and the computation of the kernel function for all samples: for large datasets ($N \gtrsim 10^4$) they are usually outperformed by neural networks.[76]

F.3 Decision Trees, Random Forests and Gradient Boosting

Decision trees are supervised ML algorithms which model simple decision rules based on the input data [190], [191]. They are informally referred to with the acronym CART (as in *Classification And Regression Trees*) and their name descends from the binary tree structure coming from such decision functions separating the input data at each iteration (*node*), thus creating a bifurcating

[76] In general it is plausible that the time complexity is $\mathcal{O}\left(F \times N^2\right)$ based on good implementations of caching in the algorithm.

structure with *branches* (the different paths, or decisions made) and *leaves* (the samples in each branch): the basic idea behind them is an *if... then... else* structure. In `scikit-learn` this is implemented in the classes `tree.DecisionTreeClassifier` and `tree.DecisionTreeRegressor`.

The idea behind it is to take input samples $x^{(i)} \in \mathbb{R}^F$ (for $i = 1, 2, \dots, N$) and partition the space in such a way that data with the same label $y^{(i)} \in \mathbb{R}$ is on the same subset of samples (while for classification this may be natural to visualise, for regression this amounts to approximate the input data with a step function whose value is constant inside the partition). Let in fact $j = 1, 2, \dots, F$ be a feature and $x_j^{(i)}$ the corresponding value for the sample i, at each node n of the tree we partition the set of input data $\mathcal{M}_n$ into two subsets:

$$\begin{aligned} \mathcal{M}_n^{[1]}(t_{j,\,n}) &= \left\{ \left(x^{(i)},\, y^{(i)} \right) \in \mathbb{R}^F \times \mathbb{R} \quad | \quad x_j^{(i)} < t_{j,\,n} \quad \forall i \in A_n \right\}, \\ \mathcal{M}_n^{[2]}(t_{j,\,n}) &= \mathcal{M}_n \setminus \mathcal{M}_n^{[1]}(t_{j,\,n}), \end{aligned} \tag{F.16}$$

where A_n is the full set of labels of the data samples in the node n and $t_{j,\,n} \in \mathbb{R}$ is a threshold value for the feature j at node n.

The measure of the ability of the split to reach the objective (classifying or creating a regression model to predict the labels) is modelled through an *impurity* function (i.e. the measure of how often a random data point would be badly classified or how much it would be badly predicted). Common choices in classification tasks are the Gini impurity, a special quadratic case of the Tsallis entropy (which in turn is a generalisation of the Boltzmann-Gibbs entropy, recovered as the first power of the Tsallis entropy) and the information theoretic definition of the entropy. In regression tasks it is usually given by the l_1 and l_2 norms of the deviation from different estimators (mean and median) for each node n:

- *mean absolute error*

$$H_n^{[l]}(x;\, t_{j,\,n}) = \frac{1}{\left| \mathcal{M}_n^{[l]}(t_{j,\,n}) \right|} \sum_{i \in A_n^{[l]}} \left| y^{(i)} - \tilde{y}_{\text{pred},\,n}^{[l]}(x) \right|, \quad \left(x^{(i)},\, y^{(i)} \right) \in \mathcal{M}_n(t_{j,\,n}), \tag{F.17}$$

- *mean squared error*:

$$H_n^{[l]}(x;\, t_{j,\,n}) = \frac{1}{\left| \mathcal{M}_n^{[l]}(t_{j,\,n}) \right|} \sum_{i \in A_n^{[l]}} \left(y^{(i)} - \bar{y}_{\text{pred},\,n}^{[l]}(x) \right)^2, \quad \left(x^{(i)}, y^{(i)} \right) \in \mathcal{M}_n(t_{j,\,n}), \tag{F.18}$$

where $\left| \mathcal{M}_n^{[l]}(t_{j,\,n}) \right|$ is the cardinality of the set $\mathcal{M}_n^{[l]}(t_{j,\,n})$ for $l = 1, 2$ and

$$\tilde{y}_{\text{pred},\,n}^{[l]}(x) = \underset{i \in A_n^{[l]}}{\text{median}}\; y_{\text{pred}}\left(x^{(i)} \right), \qquad \bar{y}_{\text{pred},\,n}^{[l]}(x) = \frac{1}{\left| A_n^{[l]} \right|} \sum_{i \in A_n^{[l]}} y_{\text{pred}}\left(x^{(i)} \right), \tag{F.19}$$

where $A_n^{[l]} \subset A_n$ are the subset of labels in the left and right splits ($l = 1$ and $l = 2$, that is) of the node n.

The full measure of the impurity of the node n and for a feature j is then:

$$G_{j,\,n}(\mathcal{M};\, t_{j,\,n}) = \frac{\left| \mathcal{M}_n^{[1]}(t_{j,\,n}) \right|}{|\mathcal{M}_n|} H_n^{[1]}(x;\, t_{j,\,n}) + \frac{\left| \mathcal{M}_n^{[2]}(t_{j,\,n}) \right|}{|\mathcal{M}_n|} H_n^{[2]}(x;\, t_{j,\,n}), \tag{F.20}$$

from which we select the parameters

$$\widehat{t}_{j,\,n} = \underset{t_{j,\,n}}{\operatorname{argmin}}\; G_n(\mathcal{M}_n;\, t_{j,\,n}). \tag{F.21}$$

We then recurse over all $\mathcal{M}_n^{[l]}\big(\widehat{t}_{j,\,n}\big)$ (for $l = 1, 2$) until we reach the maximum allowed depth of the tree (at most $|\mathcal{M}_n| = 1$).

Other than just predicting a class or a numeric value, decision trees provide a criterion to assign the importance of each feature appearing in the nodes. The implementation of the procedure can however vary between different libraries: in `scikit-learn` the importance of a feature is computed by the total reduction in the objective function due to the presence of the feature, normalised over all nodes. Namely it is defined as the difference between the total impurity normalised by the total amount of samples in the node and the sum of the separate impurities of the left and right split normalised over the number of samples in the respective splits, summed over all the nodes. Thus features with a high *variable ranking* (or *variable importance*) are those with a higher impact in reducing the loss of the algorithm and can be expected to be seen in the initial branches of the tree. A measure of the variable importance is in general extremely useful for feature engineering and feature selection since it gives a natural way to pick features with a higher chance to provide a good prediction of the labels.

By nature decision trees have a query time complexity of $\mathcal{O}(\log(N))$ as most binary search algorithms. However their definition requires running over all F features to find the best split for each sample thus increasing the time complexity to $\mathcal{O}(F \times N \log(N))$. Summing over all samples in the whole node structure leads to the worst case scenario of a time complexity $\mathcal{O}\big(F \times N^2 \log(N)\big)$. Well balanced trees (that is, nodes are approximately symmetric with the same amount of data samples inside) can usually reduce that time by a factor N, but it may not always be the case.

Decision trees have the advantage to be very good at classifying or creating regression relations in the presence of "well separable" data samples and they usually provide very good predictions in a reasonable amount of time (especially when balanced). However if F is very large, a small variation of the data will almost always lead to a huge change in the decision thresholds and they are usually prone to overfit. There are however smart ways to compensate this behaviour based on *ensemble* learning such as *bagging* and *boosting* as well as *pruning* methods such as limiting the depth of the tree or the number of splits and introducing a dropout parameter to remove certain nodes of the tree.[77] Also random forests of trees provide a variable ranking system by averaging the importance of each feature across all base estimators in the bagging aggregator.

As a reference, *random forests* of decision trees (as in `ensemble.RandomForestRegressor` in `scikit-learn`) are ensemble learning algorithms based on fully grown (deep) decision trees. They were created to overcome the issues related to overfitting and variability of the input data and are based on random sampling of the training data [192]. The idea is to take K random partitions of the training data and train a different decision tree for each of them and combine the results: for a classification task this would resort to averaging the *a posteriori* (or conditional) probability of predicting the class c given an input x (i.e. the Bayesan probability $P(c \mid x)$) over

[77] The term *bagging* comes from the contraction of *bootstrap* and *aggregating*: predictions are in fact made over randomly sampled partitions of the training set with substitution (i.e. samples can appear in different partitions, known as *bootstrap* approach) and then averaged together (*aggregating*). Random forests are an improvement to this simple idea and work best for decision trees: while it is possible to bag simple trees and take their predictions, using the random subsampling as described usually leads to better performance and results.

the K trees, while for regression this amount to averaging the predictions of the trees $y^{(i)\,\{k\}}_{\text{pred},\,\widehat{n}}$ where $k = 1, 2, \ldots, K$ and $\widehat{n}$ is the final node (i.e. the node containing the final predictions). This defines what has been called a *random forest* of trees which can usually help in improving the predictions by reducing the variance due to trees adapting too much to training sets.

Boosting methods are another implementation of ensemble learning algorithms in which more *weak learners*, in this case shallow decision trees, are trained over the training dataset [193], [194]. In general parameters $\widehat{t}_{j,\,n}$ in (F.21) can be approximated by an expansion

$$t_{j,\,n}(x) = \sum_{m=0}^{M} t^{\{m\}}_{j,\,n}(x) = \sum_{m=0}^{M} \beta^{\{m\}}_{j,\,n}\, g(x;\, a^{\{m\}}_{j,\,n}), \tag{F.22}$$

where $g(x;\, a^{\{m\}}_{j,\,n})$ are called *base learners* and M is the number of iterations.[78] The values of $a^{\{m\}}_{j,\,n}$ and $\beta^{\{m\}}_{j,\,n}$ are enough to specify the value of $t_{j,\,n}(x)$ and can be compute by iterating (F.21):

$$\left(a^{\{m\}}_{j,\,n},\, \beta^{\{m\}}_{j,\,n}\right) = \underset{\{a_{j,\,n};\, \beta_{j,\,n}\}}{\operatorname{argmin}} G_{j,\,n}\Big(\mathcal{M}_n;\, t^{\{m-1\}}_{j,\,n}(x) + \beta_{j,\,n}\, g(x;\, a_{j,\,n})\Big). \tag{F.23}$$

The specific case of boosted trees is simpler since the base learner predicts a constant value $g\left(x;\, a^{\{m\}}_{j,\,n}\right)$, thus (F.23) simplifies to

$$\gamma^{\{m\}}_{j,\,n} = \underset{\gamma_{j,\,n}}{\operatorname{argmin}}\, G_{j,\,n}\Big(\mathcal{M}_n;\, t^{\{m-1\}}_{j,\,n}(x) + \gamma_{j,\,n}\Big). \tag{F.24}$$

Ultimatelythe value of the parameters in (F.22) are updated using gradient descent as

$$t^{\{m\}}_{j,\,n}(x) = t^{\{m-1\}}_{j,\,n}(x) + \nu\, \gamma^{\{m\}}_{j,\,n}, \tag{F.25}$$

where $0 \leq \nu \leq 1$ is the *learning rate* which controls the magnitude of the update. Through this procedure, boosted trees can usually vastly improve the predictions of very small decision trees by increasing variance over bias. Another way to prevent overfitting the training set is to randomly *subsample* the features vector by taking a subset of them (in `scikit-learn` it is represented as a percentage of the total number of features). Moreover `scikit-learn` introduces various ways to control the loss of gradient boosting: apart from the aforementioned *least squares* and *least absolute deviation*, we can have hybrid versions of these such as the *huber* loss which combines the two previous losses with an additional hyperparameter α [195]. While more implementations are present, also the boosted trees provide a way to measure the importance of the variables as any decision tree algorithm.

F.4 Artificial Neural Networks

ANN are a state of the art algorithm in ML. They usually outperform any other algorithm in very large datasets (the size of our dataset is roughly at the threshold) and can learn very complicated

[78] Different implementations of the algorithm refer to the number of iterations in different way. For instance `scikit-learn` calls them `n_estimators` in the class `ensemble.GradientBoostingRegressor` in analogy to the random forest where the same name is given to the number of trained decision trees, while `XGBoost` prefers `num_boost_rounds` and `num_parallel_tree` to name the number of boosting rounds (the iterations) and the number of trees trained in parallel in a forest.

decision boundaries and functions.[79] In the main text we used two types of neural networks: *fully connected* (FC) networks and *convolutional neural networks* (CNN). They both rely on being built in a layered structure, starting from the input layers (e.g. the configuration matrix of CY manifolds or an RGB image or several engineered features) going towards the output layers (e.g. the Hodge numbers or the classification class of the image).

In FC networks the input of layer l is a feature vector $a^{(i)\,\{l\}} \in \mathbb{R}^{n_l}$ (for $i = 1, 2, \ldots, N$) and, as shown in Figure 9.18a, each layer is densely connected to the following.[80] In other words, each entry of the vectors $a_j^{(i)\,\{l\}}$ (for $j = 1, 2, \ldots, n_l$) is mapped through a function ψ to all the components of the following layer $a^{\{l+1\}} \in \mathbb{R}^{n_{l+1}}$:

$$\begin{array}{rccc} \psi\colon & \mathbb{R}^{n_l} & \longrightarrow & \mathbb{R}^{n_{l+1}} \\ & a^{(i)\,\{l\}} & \longmapsto & a^{(i)\,\{l+1\}} = \psi_j\big(a^{(i)\,\{l\}}\big), \end{array} \tag{F.26}$$

such that

$$a_j^{(i)\,\{l+1\}} = \psi_j(a^{(i)\,\{l\}}) = \phi\left(\sum_{k=1}^{n_l} a_k^{(i)\,\{l\}} W_{kj}^{\{l\}} + b^{\{l\}}\,\mathbb{1}_j\right), \tag{F.27}$$

where $\mathbb{1} \in \mathbb{R}^{n_{l+1}}$ is an identity vector. The matrix $W^{\{l\}}$ is *weight matrix* and $b^{\{l\}}$ is the *bias* term. The function ϕ is a non linear function and plays a fundamental role: without it the successive application of the linear map $a^{\{l\}} \cdot W^{\{l\}} + b\,g$ would prevent the network from learning more complicated decision boundaries or functions as the ANN would only be capable of reproducing linear relations. ϕ is known as *activation function* and can assume different forms, as long as its non linearity is preserved (e.g. a *sigmoid* function in the output layer of a network squeezes the results in the interval $[0, 1]$ thus reproducing the probabilities of of a classification). A common choice is the *rectified linear unit* (ReLU) function

$$\phi(z) = \mathrm{ReLU}(z) = \max(0, z), \tag{F.28}$$

which has been proven to be better at training deep learning architectures [196], or its modified version $\mathrm{LeakyReLU}(z) = \max(\alpha z, z)$ which introduces a slope $\alpha > 0$ to improve the computational performance near the non differentiable point in the origin.

CNN architectures rose to fame in the context of computer vision and object localisation [197]. As one can suspect looking at Figure 9.19 for instance, the fundamental difference with FC networks is that they use a convolution operation $K^{\{l\}} * a^{(i)\,\{l\}}$ instead of a linear map to transform the output of the layers, before applying the activation function.[81] This way the network is no longer densely connected, as the results of the convolution (*feature map*) depends only on a restricted neighbourhood of the original feature, depending on the size of the *kernel* window $K^{\{l\}}$ used and the shape of the input $a^{(i)\{l\}}$, which is no longer limited to flattened vectors. In turn its size influences the convolution operator which we can compute: one way to see this is to visualise an image being scanned by a smaller window function over all pixels or by skipping some a certain number of them (the length of the *stride* of the kernel). In general the

[79] Despite their fame in the face of the general public, even small networks can prove to be extremely good at learning complicated functions in a small amount of time.

[80] The input vector $x \in \mathbb{R}^F$ is equivalent to the vector $a^{\{0\}}$ and $n_0 = F$. Inputs to each layer are here represented as a matrix $a^{\{l\}}$ whose columns are made by samples and whose rows are filled with the values of the features.

[81] In general the input of each layer can be a generic tensor with an arbitrary number of axis. For instance, an RGB image can be represented by a three dimensional tensor with indices representing the width of the image, its height and the number of filters (in this case 3, one for each colour channel).

output will therefore be different than the input, unless the latter is *padded* (with zeros usually) before the convolution. The size of the output is therefore:

$$O_n = \frac{I_n - k_n + 2p_n}{S_n} + 1, \qquad n = 1, 2, \ldots, \tag{F.29}$$

where O is the output size, I the input size, k the size of the kernel used, p the amount of padding (symmetric at the start and end of the axis considered) and S the stride. In the formula, n runs over the number of components of the input tensor. While any padding is possible, we are usually interested in two kinds of possible convolutions:

- "same" convolutions for which $O_n = I_n$, thus $p_n = \frac{I_n(S_n-1)-S_n+k_n}{2}$,
- "valid" convolutions for which $O_n < I_n$ and $p_n = 0$.

In both cases the learning process aims to minimise the loss function defined for the task: in our regression implementation of the architecture we used the mean squared error of the predictions. The objective is to find best possible values of weight and bias terms $W^{\{l\}}$ and $b^{\{l\}}$) or to build the best filter kernel $K^{\{l\}}$ through *backpropagation* [198], that is by reconstructing the gradient of the loss function climbing back the network from the output layer to the input and then using the usual gradient descent procedure to select the optimal parameters. For instance, in the case of FC networks we need to find

$$\left(\widehat{W}^{\{l\}}, \widehat{b}^{\{l\}}\right) = \underset{W^{\{l\}}, b^{\{l\}}}{\operatorname{argmin}} \frac{1}{2N} \sum_{i=1}^{N} \left(y^{(i)} - a^{(i)\,\{L\}}\right)^2 \quad \forall l = 1, 2, \ldots, L, \tag{F.30}$$

where L is the total number of layers in the network. A similar relation holds in the case of CNN architectures. In the main text we use the *Adam* [199] implementation of gradient descent and add batch normalisation layers to improve the convergence of the algorithm.

As we can see from their definition, neural networks are capable of learning very complex structures at the cost of having a large number of parameters to tune. The risk of overfitting the training set is therefore quite evident. There are in general several techniques to counteract the tendency to adapt the training set, one of them being the introduction of regularisation (l_2 and l_1) in the same fashion of a linear model (we show it in Appendix F.1). Another successful way is to introduce *dropout* layers [200] where connections are randomly switched off according to a certain retention probability (or its complementary, the dropout *rate*): this regularisation technique allows to keep good generalisation properties since the prediction can rely in a less incisive way on the particular architecture since which is randomly modified during training (dropout layers however act as the identity during predictions to avoid producing random results).

References

[1] R. Finotello and I. Pesando, 'The Classical Solution for the Bosonic String in the Presence of Three D-Branes Rotated by Arbitrary SO(4) Elements', *Nuclear Physics B*, vol. 941, pp. 158–194, 2019, ISSN: 05503213. DOI: 10/gf66b3. arXiv: 1812.04643.

[2] A. Arduino, R. Finotello and I. Pesando, 'On the Origin of Divergences in Time-Dependent Orbifolds', *The European Physical Journal C*, vol. 80, no. 5, p. 476, 2020, ISSN: 1434-6044, 1434-6052. DOI: 10/gg54bw. arXiv: 2002.11306.

[3] R. Finotello and I. Pesando. (2019). '2D Fermion on the Strip with Boundary Defects as a CFT with Excited Spin Fields'. arXiv: 1912.07617.

[4] H. Erbin and R. Finotello. (2020). 'Inception Neural Network for Complete Intersection Calabi-Yau 3-Folds'. arXiv: 2007.13379.

[5] ——, (2020). 'Machine Learning for Complete Intersection Calabi-Yau Manifolds: A Methodological Study'. arXiv: 2007.15706.

[6] A. M. Polyakov, 'Quantum geometry of bosonic strings', *Physics Letters B*, vol. 103, no. 3, pp. 207–210, 1981, ISSN: 03702693. DOI: 10/cq538n.

[7] J. Polchinski, *String Theory. An Introduction to the Bosonic String*. Cambridge University Press, 1998, vol. 1, ISBN: 978-0-521-67227-6. [Online]. Available: https://www.cambridge.org/academic/subjects/physics/theoretical-physics-and-mathematical-physics/string-theory-volume-1.

[8] M. B. Green, J. H. Schwarz and E. Witten, *Superstring Theory. Introduction*. Ser. Cambridge Monographs on Mathematical Physics. 1988, vol. 1, ISBN: 978-0-521-35752-4.

[9] D. Friedan, E. Martinec and S. Shenker, 'Conformal Invariance, Supersymmetry and String Theory', *Nuclear Physics B*, vol. 271, no. 3-4, pp. 93–165, 1986, ISSN: 05503213. DOI: 10/fq6qtz.

[10] P. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal Field Theory*, ser. Graduate Texts in Contemporary Physics. New York: Springer, 1997, ISBN: 978-1-4612-7475-9. [Online]. Available: http://link.springer.com/10.1007/978-1-4612-2256-9.

[11] R. Blumenhagen and E. Plauschinn, *Introduction to Conformal Field Theory*, ser. Lecture Notes in Physics. Berlin, Heidelberg: Springer, 2009, vol. 779, ISBN: 978-3-642-00449-0. [Online]. Available: http://link.springer.com/10.1007/978-3-642-00450-6.

[12] J. Polchinski, *String Theory. Superstring Theory and Beyond*. Cambridge University Press, 1998, vol. 2, ISBN: 978-0-521-63304-8. [Online]. Available: https://www.cambridge.org/academic/subjects/physics/theoretical-physics-and-mathematical-physics/string-theory-volume-2.

[13] L. B. Anderson and M. Karkheiran. (2018). 'TASI Lectures on Geometric Tools for String Compactifications'. arXiv: 1804.08792.

[14] M. Graña, 'Flux compactifications in string theory: A comprehensive review', *Physics Reports*, vol. 423, no. 3, pp. 91–158, 2006, ISSN: 03701573. DOI: 10/bdzzsc. arXiv: hep-th/0509003.

[15] M. Graña and H. Triendl, *String Theory Compactifications*, ser. SpringerBriefs in Physics. Cham: Springer, 2017, ISBN: 978-3-319-54315-4. [Online]. Available: http://link.springer.com/10.1007/978-3-319-54316-1.

[16] A. M. Uranga, 'TASI Lectures on String Compactification, Model Building, and Fluxes', 2005. [Online]. Available: http://cds.cern.ch/record/933469/files/cer-002601054.pdf.

[17] P. Candelas, G. T. Horowitz, A. Strominger *et al.*, 'Vacuum configurations for superstrings', *Nuclear Physics B*, vol. 258, pp. 46–74, 1985, ISSN: 05503213. DOI: 10/bxjjzx.

[18] E. Calabi, 'On Kähler Manifolds with Vanishing Canonical Class', in *Algebraic Geometry and Topology. A Symposium in Honor of S. Lefschetz.*, vol. 12, 1957, pp. 78–89. DOI: 10/ghf4nv.

[19] S.-T. Yau, 'Calabi's conjecture and some new results in algebraic geometry', *Proceedings of the National Academy of Sciences*, vol. 74, no. 5, pp. 1798–1799, 1977, ISSN: 0027-8424, 1091-6490. DOI: 10/cbzd2h.

[20] D. Joyce, *Compact Manifolds with Special Holonomy*. Oxford University Press, 2000, ISBN: 978-0-19-850601-0.

[21] ——, (2002). 'Lectures on Calabi-Yau and Special Lagrangian Geometry'. arXiv: math/0108088.

[22] B. Greene. (1997). 'String Theory on Calabi-Yau Manifolds'. arXiv: hep-th/9702155.

[23] J. Polchinski, 'Dirichlet Branes and Ramond-Ramond Charges', *Physical Review Letters*, vol. 75, no. 26, pp. 4724–4727, 1995, ISSN: 00319007. DOI: 10/bxcwrv. arXiv: hep-th/9510017.

[24] ——, 'TASI Lectures on D-Branes', *New Frontiers in Fields and Strings*, pp. 75–136, 1996. arXiv: hep-th/9611050.

[25] P. Di Vecchia and A. Liccardo, 'D-Branes in String Theory II', in *YITP Workshop on Developments in Superstring and M-Theory*, 1999, pp. 7–48. arXiv: hep-th/9912275.

[26] ——, 'D Branes in String Theory I', *NATO Sci. Ser. C*, vol. 556, no. NORDITA-1999-77-HE, L. Thorlacius and T. Jonsson, Eds., pp. 1–60, 2000. DOI: 10/gf66cb. arXiv: hep-th/9912161.

[27] P. Di Vecchia, M. Frau, I. Pesando *et al.*, 'Classical P-Branes from Boundary State', *Nuclear Physics B*, vol. 507, no. 1-2, pp. 259–276, 1997, ISSN: 05503213. DOI: 10/bjkrq3. arXiv: hep-th/9707068.

[28] G. Honecker and J. Vanhoof, 'Towards the Field Theory of the Standard Model on Fractional D6-Branes on T6/Z6: Yukawa Couplings and Masses', *Fortschritte der Physik*, vol. 60, no. 9-10, pp. 1050–1056, 2012, ISSN: 00158208. DOI: 10/fz4mj9. arXiv: 1201.5872.

[29] D. Lüst, S. Stieberger and T. R. Taylor, 'The LHC String Hunter's Companion', *Nuclear Physics B*, vol. 808, no. 1-2, pp. 1–52, 2009, ISSN: 05503213. DOI: 10/fdcxt4. arXiv: 0807.3333.

[30] P. Di Vecchia, A. Liccardo, R. Marotta *et al.*, 'Boundary State for Magnetized D9 Branes and One-Loop Calculation', in *Proceedings of Sense of Beauty in Physics: Miniconference in Honor of Adriano Di Giacomo on His 70th Birthday*, 2006. arXiv: hep-th/0601067.

[31] R. Bousso and J. Polchinski, 'Quantization of Four-Form Fluxes and Dynamical Neutralization of the Cosmological Constant', *Journal of High Energy Physics*, vol. 2000, no. 06, pp. 006–006, 2000, ISSN: 1029-8479. DOI: 10/fwrxdr. arXiv: hep-th/0004134.

[32] L. Susskind. (2003). 'The Anthropic Landscape of String Theory'. arXiv: hep-th/0302219.

[33] S. Kachru, R. Kallosh, A. Linde *et al.*, 'De Sitter Vacua in String Theory', *Physical Review D*, vol. 68, no. 4, p. 046 005, 2003, ISSN: 0556-2821, 1089-4918. DOI: 10/bfvjj7. arXiv: hep-th/0301240.

[34] G. B. Cleaver. (2007). 'In Search of the (Minimal Supersymmetric) Standard Model String'. arXiv: hep-ph/0703027.

[35] P. Goddard, J. Goldstone, C. Rebbi *et al.*, 'Quantum dynamics of a massless relativistic string', *Nuclear Physics B*, vol. 56, no. 1, pp. 109–135, 1973, ISSN: 05503213. DOI: 10/ccmwgf.

[36] C. Angelantonj and A. Sagnotti. (2002). 'Open Strings'. arXiv: hep-th/0204089.

[37] H.-M. Chan and J. E. Paton, 'Generalized Veneziano model with isospin', *Nuclear Physics B*, vol. 10, no. 3, pp. 516–520, 1969, ISSN: 05503213. DOI: 10/chp79v.

[38] L. E. Ibanez and A. M. Uranga, *String Theory and Particle Physics: An Introduction to String Phenomenology*. Cambridge University Press, 2012, ISBN: 978-0-521-51752-2.

[39] G. Aldazabal, L. E. Ibanez, F. Quevedo *et al.*, 'D-Branes at Singularities : A Bottom-Up Approach to the String Embedding of the Standard Model', *Journal of High Energy Physics*, vol. 2000, no. 08, pp. 002–002, 2000, ISSN: 1029-8479. DOI: 10/dxtkc6. arXiv: hep-th/0005067.

[40] A. M. Uranga, 'Chiral Four-Dimensional String Compactifications with Intersecting D-Branes', *Classical and Quantum Gravity*, vol. 20, no. 12, S373–S393, 2003, ISSN: 0264-9381, 1361-6382. DOI: 10/cr8vtd. arXiv: hep-th/0301032.

[41] B. Zwiebach, *A First Course in String Theory*. Cambridge University Press, 2009, ISBN: 978-0-521-88032-9. [Online]. Available: https://www.cambridge.org/academic/subjects/physics/theoretical-physics-and-mathematical-physics/first-course-string-theory-2nd-edition.

[42] L. E. Ibanez, F. Marchesano and R. Rabadán, 'Getting Just the Standard Model at Intersecting Branes', *Journal of High Energy Physics*, vol. 2001, no. 11, pp. 002–002, 2001, ISSN: 1029-8479. DOI: 10/drzgmv. arXiv: hep-th/0105155.

[43] M. M. Sheikh-Jabbari, 'Classification of Different Branes at Angles', *Physics Letters B*, vol. 420, no. 3-4, pp. 279–284, 1998, ISSN: 03702693. DOI: 10/dbhwc6. arXiv: hep-th/9710121.

[44] M. Berkooz, M. R. Douglas and R. G. Leigh, 'Branes Intersecting at Angles', *Nuclear Physics B*, vol. 480, no. 1-2, pp. 265–278, 1996, ISSN: 05503213. DOI: 10/fjff5p. arXiv: hep-th/9606139.

[45] N. Chamoun, S. Khalil and E. Lashin, 'Fermion masses and mixing in intersecting brane scenarios', *Physical Review D*, vol. 69, no. 9, p. 095 011, 2004, ISSN: 1550-7998, 1550-2368. DOI: 10/dbfm3j. arXiv: hep-ph/0309169.

[46] D. Cremades, L. E. Ibanez and F. Marchesano, 'Yukawa Couplings in Intersecting D-Brane Models', *Journal of High Energy Physics*, vol. 2003, no. 07, pp. 038–038, 2003, ISSN: 1029-8479. DOI: 10/bpp94m. arXiv: hep-th/0302105.

[47] M. Cvetič, I. García-Etxebarria and R. Richter, 'Branes and Instantons Intersecting at Angles', *Journal of High Energy Physics*, vol. 2010, no. 1, p. 5, 2010, ISSN: 1029-8479. DOI: 10/dg2zz6. arXiv: 0905.1694.

[48] S. A. Abel and M. D. Goodsell, 'Realistic Yukawa Couplings through Instantons in Intersecting Brane Worlds', *Journal of High Energy Physics*, vol. 2007, no. 10, pp. 034–034, 2007, ISSN: 1029-8479. DOI: 20071006031312. arXiv: hep-th/0612110.

[49] C.-M. Chen, T. Li, v. E. Mayes *et al.*, 'A Realistic World from Intersecting D6-Branes', *Physics Letters B*, vol. 665, no. 4, pp. 267–270, 2008, ISSN: 03702693. DOI: 10/b25m9j. arXiv: hep-th/0703280.

[50] ——, 'Realistic Yukawa Textures and SUSY Spectra from Intersecting Branes', *Physical Review D*, vol. 77, no. 12, p. 125 023, 2008, ISSN: 1550-7998, 1550-2368. DOI: 10/c3gb98. arXiv: 0711.0396.

[51] S. A. Abel and B. W. Schofield, 'One-Loop Yukawas on Intersecting Branes', *Journal of High Energy Physics*, vol. 2005, no. 06, pp. 072–072, 2005, ISSN: 1029-8479. DOI: 10/bt34kd. arXiv: hep-th/0412206.

[52] S. A. Abel, M. Masip and J. Santiago, 'Flavour Changing Neutral Currents in Intersecting Brane Models', *Journal of High Energy Physics*, vol. 2003, no. 04, pp. 057–057, 2003, ISSN: 1029-8479. DOI: 10/b39bd8. arXiv: hep-ph/0303087.

[53] C. Angelantonj, I. Antoniadis, E. Dudas *et al.*, 'Type-I Strings on Magnetised Orbifolds and Brane Transmutation', *Physics Letters B*, vol. 489, no. 1-2, pp. 223–232, 2000, ISSN: 03702693. DOI: 10/cc5skj. arXiv: hep-th/0007090.

[54] M. Bianchi and E. Trevigne, 'The Open Story of the Magnetic Fluxes', *Journal of High Energy Physics*, vol. 2005, no. 08, pp. 034–034, 2005, ISSN: 1029-8479. DOI: 10/bjnzck. arXiv: hep-th/0502147.

[55] I. Pesando, 'Open and Closed String Vertices for Branes with Magnetic Field and T-Duality', *Journal of High Energy Physics*, vol. 2010, no. 2, p. 64, 2010, ISSN: 1029-8479. DOI: 10/fr82w9. arXiv: 0910.2576.

[56] S. Forste and C. Liyanage, 'Yukawa Couplings from Magnetized D-Brane Models on Non-Factorisable Tori', *Journal of High Energy Physics*, vol. 2018, no. 8, p. 169, 2018, ISSN: 1029-8479. DOI: 10/gf66b4. arXiv: 1802.05136.

[57] E. Kiritsis and C. Kounnas, 'String Propagation in Gravitational Wave Backgrounds', *Physics Letters B*, vol. 320, no. 3-4, pp. 264–272, 1994, ISSN: 03702693. DOI: 10/bswvpn. arXiv: hep-th/9310202.

[58] G. D'Appollonio and E. Kiritsis, 'String Interactions in Gravitational Wave Backgrounds', *Nuclear Physics B*, vol. 674, no. 1-2, pp. 80–170, 2003, ISSN: 05503213. DOI: 10/cwvfg9. arXiv: hep-th/0305081.

[59] E. Gava, K. S. Narain and H. M. Sarmadi, 'On the Bound States of P- and (P+2)-Branes', *Nuclear Physics B*, vol. 504, no. 1-2, pp. 214–238, 1997, ISSN: 05503213. DOI: 10/djsprb. arXiv: hep-th/9704006.

[60] D. Duo, R. Russo and S. Sciuto, 'New Twist Field Couplings from the Partition Function for Multiply Wrapped D-Branes', *Journal of High Energy Physics*, vol. 2007, no. 12, pp. 042–042, 2007, ISSN: 1029-8479. DOI: 10/dptb7d. arXiv: 0709.1805.

[61] J. R. David, 'Tachyon Condensation in the D0/D4 System', *Journal of High Energy Physics*, vol. 2000, no. 10, pp. 004–004, 2000, ISSN: 1029-8479. DOI: 20050405175528. arXiv: hep-th/0007235.

[62] ——, 'Tachyon Condensation Using the Disc Partition Function', *Journal of High Energy Physics*, vol. 2001, no. 07, pp. 009–009, 2001, ISSN: 1029-8479. DOI: 10/fb5m3w. arXiv: hep-th/0012089.

[63] K. Hashimoto and S. Nagaoka, 'Recombination of Intersecting D-Branes by Local Tachyon Condensation', *Journal of High Energy Physics*, vol. 2003, no. 06, pp. 034–034, 2003, ISSN: 1029-8479. DOI: 10/fd2kjv. arXiv: hep-th/0303204.

[64] T. T. Burwick, R. K. Kaiser and H. F. Müller, 'General Yukawa couplings of strings on orbifolds', *Nuclear Physics B*, vol. 355, no. 3, pp. 689–711, 1991, ISSN: 05503213. DOI: 10/c56dxj.

[65] S. Stieberger, D.-U. Jungnickel, J. Lauer *et al.*, 'Yukawa Couplings for Bosonic Z_N Orbifolds: Their Moduli and Twisted Sector Dependence', *Modern Physics Letters A*, vol. 07, no. 33, pp. 3059–3070, 1992, ISSN: 0217-7323, 1793-6632. DOI: 10/d9jgv3. arXiv: hep-th/9204037.

[66] P. Anastasopoulos, M. Bianchi and R. Richter, 'Light Stringy States', *Journal of High Energy Physics*, vol. 2012, no. 3, p. 68, 2012, ISSN: 1029-8479. DOI: 10/f3sx4b. arXiv: 1110.5424.

[67] P. Anastasopoulos, M. D. Goodsell and R. Richter, 'Three- and Four-Point Correlators of Excited Bosonic Twist Fields', *Journal of High Energy Physics*, vol. 2013, no. 10, p. 182, 2013, ISSN: 1029-8479. DOI: 10/gf66b5. arXiv: 1305.7166.

[68] S. Sciuto, 'The general vertex function in dual resonance models', *Lettere al Nuovo Cimento*, vol. 2, no. 9, pp. 411–418, 1969, ISSN: 0375-930X, 1827-613X. DOI: 10/drsft8.

[69] I. Pesando, 'Correlators of Arbitrary Untwisted Operators and Excited Twist Operators for N Branes at Angles', *Nuclear Physics B*, vol. 886, pp. 243–287, 2014, ISSN: 05503213. DOI: 10/gf66cg. arXiv: 1401.6797.

[70] ——, 'Green Functions and Twist Correlators for N Branes at Angles', *Nuclear Physics B*, vol. 866, no. 2, pp. 87–123, 2012, ISSN: 05503213. DOI: 10/gf66ch. arXiv: 1206.1431.

[71] ——, (2011). 'The Generating Function of Amplitudes with N Twisted and M Untwisted States'. arXiv: 1107.5525.

[72] ——, 'Multi-Branes Boundary States with Open String Interactions', *Nuclear Physics B*, vol. 793, no. 1-2, pp. 211–245, 2008, ISSN: 05503213. DOI: 10/bh6q64. arXiv: hep-th/0310027.

[73] P. Di Vecchia, A. Liccardo, R. Marotta *et al.*, 'Wrapped Magnetized Branes: Two Alternative Descriptions?', *Journal of High Energy Physics*, vol. 2007, no. 11, pp. 100–100, 2007, ISSN: 1029-8479. DOI: 10/c7xmmn. arXiv: 0709.4149.

[74] I. Pesando, 'Strings in an Arbitrary Constant Magnetic Field with Arbitrary Constant Metric and Stringy Form Factors', *Journal of High Energy Physics*, vol. 2011, no. 6, p. 138, 2011, ISSN: 1029-8479. DOI: 10/ddhtfm. arXiv: 1101.5898.

[75] P. Di Vecchia, R. Marotta, I. Pesando *et al.*, 'Open Strings in the System D5/D9', *Journal of Physics A: Mathematical and Theoretical*, vol. 44, no. 24, p. 245 401, 2011, ISSN: 1751-8113, 1751-8121. DOI: 10/brf7sk. arXiv: 1101.0120.

[76] I. Pesando, 'Light Cone Quantization and Interactions of a New Closed Bosonic String Inspired to D1 String', *Nuclear Physics B*, vol. 876, no. 1, pp. 1–15, 2013, ISSN: 05503213. DOI: 10/f5cppv. arXiv: 1305.2710.

[77] K. Inoue, M. Sakamoto and H. Takano, 'Non-Abelian Orbifolds', *Progress of Theoretical Physics*, vol. 78, no. 4, pp. 908–922, 1987, ISSN: 0033-068X, 1347-4081. DOI: 10/bfp9q4.

[78] K. Inoue and S. Nima, 'String Interactions on Non-Abelian Orbifold', *Progress of Theoretical Physics*, vol. 84, no. 4, pp. 702–727, 1990, ISSN: 0033-068X, 1347-4081. DOI: 10/ghf4n4.

[79] B. Gato, 'Vertex operators, non-Abelian orbifolds and the Reimann–Hilbert problem', *Nuclear Physics B*, vol. 334, no. 2, pp. 414–430, 1990, ISSN: 05503213. DOI: 10/chx7rp.

[80] I. Pesando, 'Towards a Fully Stringy Computation of Yukawa Couplings on Non-Factorized Tori and Non-Abelian Twist Correlators (I): The Classical Solution and Action', *Nuclear Physics B*, vol. 910, pp. 618–664, 2016, ISSN: 05503213. DOI: 10/f82v5m. arXiv: 1512.07920.

[81] 'NIST Digital Library of Mathematical Functions', F. W. J. Olver, A. Olde Daalhuis, D. W. Lozier *et al.*, Eds., 2020. [Online]. Available: http://dlmf.nist.gov.

[82] P. Di Vecchia, R. Madsen, K. Hornfeck *et al.*, 'A vertex including emission of spin fields', *Physics Letters B*, vol. 235, no. 1-2, pp. 63–70, 1990, ISSN: 03702693. DOI: 10/fr3zhk.

[83] B. E. W. Nilsson and A. K. Tollstén, 'General NSR string reggeon vertices from a dual ramond vertex', *Physics Letters B*, vol. 240, no. 1-2, pp. 96–104, 1990, ISSN: 03702693. DOI: 10/fbgqqd.

[84] N. Di Bartolomeo, P. Di Vecchia and R. Guatieri, 'General properties of vertices with two Ramond or twisted states', *Nuclear Physics B*, vol. 347, no. 3, pp. 651–686, 1990, ISSN: 05503213. DOI: 10/ctk5bp.

[85] J. L. Petersen, J. R. Sidenius and A. K. Tollsten, 'Covariant super-reggeon calculus for superstrings', *Nuclear Physics B*, vol. 317, no. 1, pp. 109–146, 1989, ISSN: 05503213. DOI: 10/bpfrkw.

[86] M. Berkooz and D. Reichmann, 'A Short Review of Time Dependent Solutions and Space-like Singularities in String Theory', *Nuclear Physics B - Proceedings Supplements*, vol. 171, pp. 69–87, 2007, ISSN: 09205632. DOI: 10/fnjt32. arXiv: 0705.2146.

[87] F. C. Caramello Jr. (2019). 'Introduction to Orbifolds'. arXiv: 1909.08699.

[88] C. Bachas and C. Hull, 'Null Brane Intersections', *Journal of High Energy Physics*, vol. 2002, no. 12, pp. 035–035, 2002, ISSN: 1029-8479. DOI: 10/b4h4q3. arXiv: hep-th/0210269.

[89] C. Bachas, 'Relativistic String in a Pulse', *Annals of Physics*, vol. 305, no. 2, pp. 286–309, 2003, ISSN: 00034916. DOI: 10/dwshqh. arXiv: hep-th/0212217.

[90] L. J. Dixon, J. A. Harvey, C. Vafa *et al.*, 'Strings on orbifolds', *Nuclear Physics B*, vol. 261, pp. 678–686, 1985, ISSN: 05503213. DOI: 10/fnwxts.

[91] ——, 'Strings on orbifolds (II)', *Nuclear Physics B*, vol. 274, no. 2, pp. 285–314, 1986, ISSN: 05503213. DOI: 10/bv9nwg.

[92] G. T. Horowitz and A. R. Steif, 'Singular string solutions with nonsingular initial data', *Physics Letters B*, vol. 258, no. 1-2, pp. 91–96, 1991, ISSN: 03702693. DOI: 10/dzbvx3.

[93] J. Figueroa-O'Farrill and J. Simón, 'Generalised Supersymmetric Fluxbranes', *Journal of High Energy Physics*, vol. 2001, no. 12, pp. 011–011, 2001, ISSN: 1029-8479. DOI: 10/c2bshm. arXiv: hep-th/0110170.

[94] L. Cornalba and M. S. Costa, 'A New Cosmological Scenario in String Theory', *Physical Review D*, vol. 66, no. 6, p. 066 001, 2002, ISSN: 0556-2821, 1089-4918. DOI: 10/c4nfnn. arXiv: hep-th/0203031.

[95] ——, 'Time-Dependent Orbifolds and String Cosmology', *Fortschritte der Physik*, vol. 52, no. 2-3, pp. 145–199, 2004, ISSN: 00158208. DOI: 10/brv6jj. arXiv: hep-th/0310099.

[96] B. Craps, 'Big Bang Models in String Theory', *Classical and Quantum Gravity*, vol. 23, no. 21, S849–S881, 2006, ISSN: 0264-9381, 1361-6382. DOI: 10/cx8s3s. arXiv: hep-th/0605199.

[97] B. Craps, D. Kutasov and G. Rajesh, 'String Propagation in the Presence of Cosmological Singularities', *Journal of High Energy Physics*, vol. 2002, no. 06, pp. 053–053, 2002, ISSN: 1029-8479. DOI: 10/cjdtw7. arXiv: hep-th/0205101.

[98] H. Liu, G. Moore and N. Seiberg, 'Strings in a Time-Dependent Orbifold', *Journal of High Energy Physics*, vol. 2002, no. 06, pp. 045–045, 2002, ISSN: 1126-6708. DOI: 10/b2d2mj. arXiv: hep-th/0204168.

[99] ——, 'Strings in Time-Dependent Orbifolds', *Journal of High Energy Physics*, vol. 2002, no. 10, pp. 031–031, 2002, ISSN: 1126-6708. DOI: 20050405175528. arXiv: hep-th/0206182.

[100] M. Berkooz, B. Craps, D. Kutasov *et al.*, 'Comments on Cosmological Singularities in String Theory', *Journal of High Energy Physics*, vol. 2003, no. 03, pp. 031–031, 2003, ISSN: 1029-8479. DOI: 10/c9rfmc. arXiv: hep-th/0212215.

[101] G. T. Horowitz and J. Polchinski, 'Instability of spacelike and null orbifold singularities', *Physical Review D*, vol. 66, no. 10, p. 103 512, 2002, ISSN: 0556-2821, 1089-4918. DOI: 10/frkxrv. arXiv: hep-th/0206228.

[102] M. Berkooz and B. Pioline, 'Strings in an Electric Field, and the Milne Universe', *Journal of Cosmology and Astroparticle Physics*, vol. 2003, no. 11, pp. 007–007, 2003, ISSN: 1475-7516. DOI: 10/bh47tt. arXiv: hep-th/0307280.

[103] R. Jackiw, D. Kabat and M. Ortiz, 'Electromagnetic Fields of a Massless Particle and the Eikonal', *Physics Letters B*, vol. 277, no. 1, pp. 148–152, 1992, ISSN: 0370-2693. DOI: 10/fs877h. arXiv: hep-th/9112020.

[104] G. 't Hooft, 'Graviton Dominance in Ultra-High-Energy Scattering', *Physics Letters B*, vol. 198, no. 1, pp. 61–63, 1987. DOI: 10/fkzsbw.

[105] M. Soldate, 'Partial-wave unitarity and closed-string amplitudes', *Physics Letters B*, vol. 186, no. 3-4, pp. 321–327, 1987, ISSN: 03702693. DOI: 10/dnd5gz.

[106] D. Amati, M. Ciafaloni and G. Veneziano, 'Superstring collisions at planckian energies', *Physics Letters B*, vol. 197, no. 1-2, pp. 81–88, 1987, ISSN: 03702693. DOI: 10/bcmx6s.

[107] W. Black and C. Monni, 'High energy string–brane scattering for massive states', *Nuclear Physics B*, vol. 859, no. 3, pp. 299–320, 2012, ISSN: 05503213. DOI: 10/gg66ng. arXiv: 1107.4321.

[108] M. R. Douglas, 'The Statistics of String/M Theory Vacua', *Journal of High Energy Physics*, vol. 2003, no. 05, pp. 046–046, 2003, ISSN: 1029-8479. DOI: 10/fktbj2. arXiv: hep-th/0303194.

[109] S. K. Ashok and M. R. Douglas, 'Counting Flux Vacua', *Journal of High Energy Physics*, vol. 2004, no. 01, pp. 060–060, 2004, ISSN: 1029-8479. DOI: 10/fqsf7n. arXiv: hep-th/0307049.

[110] W. Taylor and Y.-N. Wang, 'The F-theory geometry with most flux vacua', *Journal of High Energy Physics*, vol. 2015, no. 12, pp. 1–21, 2015, ISSN: 1029-8479. DOI: 10/ghf4n7. arXiv: 1511.03209.

[111] ——, 'Scanning the skeleton of the 4D F-theory landscape', *Journal of High Energy Physics*, vol. 2018, no. 1, p. 111, 2018, ISSN: 1029-8479. DOI: 10/ghf4n8. arXiv: 1710.11235.

[112] A. Constantin, Y.-H. He and A. Lukas, 'Counting string theory standard models', *Physics Letters B*, vol. 792, pp. 258–262, 2019, ISSN: 03702693. DOI: 10/ghf4nx. arXiv: 1810.00444.

[113] F. Denef and M. R. Douglas, 'Computational complexity of the landscape: Part I', *Annals of Physics*, vol. 322, no. 5, pp. 1096–1142, 2007, ISSN: 00034916. DOI: 10/bp2wbs. arXiv: hep-th/0602072.

[114] J. Halverson and F. Ruehle, 'Computational complexity of vacua and near-vacua in field and string theory', *Physical Review D*, vol. 99, no. 4, p. 046 015, 2019, ISSN: 2470-0010, 2470-0029. DOI: 10/gg66j7. arXiv: 1809.08279.

[115] T. D. Brennan, F. Carta and C. Vafa, 'The String Landscape, the Swampland, and the Missing Corner', in *Proceedings of Theoretical Advanced Study Institute Summer School 2017 "Physics at the Fundamental Frontier"*, Boulder, Colorado: Sissa Medialab, 2018, p. 015. DOI: 10/ghf4nt. arXiv: 1711.00864.

[116] T. P. T. Dijkstra, L. R. Huiszoon and A. N. Schellekens, 'Chiral supersymmetric Standard Model spectra from orientifolds of Gepner models', *Physics Letters B*, vol. 609, no. 3-4, pp. 408–417, 2005, ISSN: 03702693. DOI: 10/frz278. arXiv: hep-th/0403196.

[117] R. Blumenhagen, F. Gmeiner, G. Honecker *et al.*, 'The statistics of supersymmetric D-brane models', *Nuclear Physics B*, vol. 713, no. 1-3, pp. 83–135, 2005, ISSN: 05503213. DOI: 10/dxmsqx. arXiv: hep-th/0411173.

[118] M. R. Douglas and W. Taylor, 'The Landscape of Intersecting Brane Models', *Journal of High Energy Physics*, vol. 2007, no. 01, pp. 031–031, 2007, ISSN: 1029-8479. DOI: 10/dm8rxf. arXiv: hep-th/0606109.

[119] L. B. Anderson, A. Constantin, J. Gray *et al.*, 'A Comprehensive Scan for Heterotic SU(5) GUT models', *Journal of High Energy Physics*, vol. 2014, no. 1, p. 47, 2014, ISSN: 1029-8479. DOI: 10/ghf4nq. arXiv: 1307.4787.

[120] S. Abel and J. Rizos, 'Genetic algorithms and the search for viable string vacua', *Journal of High Energy Physics*, vol. 2014, no. 8, p. 10, 2014, ISSN: 1029-8479. DOI: 10/f6v8g4. arXiv: 1404.7359.

[121] D. Krefl and R.-K. Seong, 'Machine learning of Calabi-Yau volumes', *Physical Review D*, vol. 96, no. 6, p. 066 014, 2017, ISSN: 2470-0010, 2470-0029. DOI: 10/gcpp5w.

[122] F. Ruehle, 'Evolving neural networks with genetic algorithms to study the string landscape', *Journal of High Energy Physics*, vol. 2017, no. 8, p. 38, 2017, ISSN: 1029-8479. DOI: 10/gbss9c. arXiv: 1706.07024.

[123] Y.-H. He, 'Machine-learning the string landscape', *Physics Letters B*, vol. 774, pp. 564–568, 2017, ISSN: 03702693. DOI: 10/gcqfzv.

[124] J. Carifio, J. Halverson, D. Krioukov *et al.*, 'Machine learning in the string landscape', *Journal of High Energy Physics*, vol. 2017, no. 9, p. 157, 2017, ISSN: 1029-8479. DOI: 10/gb4szm. arXiv: 1707.00655.

[125] R. Altman, J. Carifio, J. Halverson *et al.*, 'Estimating Calabi-Yau hypersurface and triangulation counts with equation learners', *Journal of High Energy Physics*, vol. 2019, no. 3, p. 186, 2019, ISSN: 1029-8479. DOI: 10/gg66h4. arXiv: 1811.06490.

[126] K. Bull, Y.-H. He, V. Jejjala *et al.*, 'Machine learning CICY threefolds', *Physics Letters B*, vol. 785, pp. 65–72, 2018, ISSN: 03702693. DOI: 10/gfm446. arXiv: 1806.03121.

[127] A. Mütter, E. Parr and P. K. S. Vaudrevange, 'Deep learning in the heterotic orbifold landscape', *Nuclear Physics B*, vol. 940, pp. 113–129, 2019, ISSN: 05503213. DOI: 10/gg66kz. arXiv: 1811.05993.

[128] A. Ashmore, Y.-H. He and B. Ovrut. (2020). 'Machine Learning Calabi-Yau Metrics'. arXiv: 1910.08605.

[129] C. R. Brodie, A. Constantin, R. Deen *et al.*, 'Machine Learning Line Bundle Cohomology', *Fortschritte der Physik*, vol. 68, no. 1, p. 1 900 087, 2020, ISSN: 0015-8208, 1521-3978. DOI: 10/gg66m6. arXiv: 1906.08730.

[130] K. Bull, Y.-H. He, V. Jejjala *et al.*, 'Getting CICY high', *Physics Letters B*, vol. 795, pp. 700–706, 2019, ISSN: 03702693. DOI: 10/gg66m5. arXiv: 1903.03113.

[131] A. Cole, A. Schachner and G. Shiu, 'Searching the landscape of flux vacua with genetic algorithms', *Journal of High Energy Physics*, vol. 2019, no. 11, p. 45, 2019, ISSN: 1029-8479. DOI: 10/ghf4nw. arXiv: 1907.10072.

[132] A. E. Faraggi, G. Harries, B. Percival *et al.*, 'Towards Machine Learning in the Classification of $Z_2 \times Z_2$ Orbifold Compactifications', *Journal of Physics: Conference Series*, vol. 1586, p. 012 032, 2020, ISSN: 1742-6588, 1742-6596. DOI: 10/ghf4nz. arXiv: 1901.04448.

[133] J. Halverson, B. Nelson and F. Ruehle, 'Branes with brains: Exploring string vacua with deep reinforcement learning', *Journal of High Energy Physics*, vol. 2019, no. 6, p. 3, 2019, ISSN: 1029-8479. DOI: 10/gg66j8. arXiv: 1903.11616.

[134] N. C. Bizet, C. Damian, O. Loaiza-Brito *et al.*, 'Testing swampland conjectures with machine learning', *The European Physical Journal C*, vol. 80, no. 8, p. 766, 2020, ISSN: 1434-6044, 1434-6052. DOI: 10/ghf4ns. arXiv: 2006.07290.

[135] J. Halverson and C. Long, 'Statistical Predictions in String Theory and Deep Generative Models', *Fortschritte der Physik*, vol. 68, no. 5, p. 2 000 005, 2020, ISSN: 0015-8208, 1521-3978. DOI: 10/gg66j9. arXiv: 2001.00555.

[136] S. Krippendorf and M. Syvaeri. (2020). 'Detecting Symmetries with Neural Networks'. arXiv: 2003.13679.

[137] H. Otsuka and K. Takemoto, 'Deep learning and k-means clustering in heterotic string vacua with line bundles', *Journal of High Energy Physics*, vol. 2020, no. 5, p. 47, 2020, ISSN: 1029-8479. DOI: 10/gg66k2. arXiv: 2003.11880.

[138] E. Parr and P. K. Vaudrevange, 'Contrast data mining for the MSSM from strings', *Nuclear Physics B*, vol. 952, p. 114 922, 2020, ISSN: 05503213. DOI: 10/ghf4n5. arXiv: 1910.13473.

[139] E. Parr, P. K. S. Vaudrevange and M. Wimmer, 'Predicting the Orbifold Origin of the MSSM', *Fortschritte der Physik*, vol. 68, no. 5, p. 2 000 032, 2020, ISSN: 0015-8208, 1521-3978. DOI: 10/ghf4n6. arXiv: 2003.01732.

[140] F. Ruehle, 'Data science applications to string theory', *Physics Reports*, vol. 839, pp. 1–117, 2020, ISSN: 03701573. DOI: 10/ggwkvm.

[141] I. Goodfellow, B. Yousha and A. Courville, *Deep Learning*. MIT press, 2017, vol. 1, ISBN: 978-0-262-33737-3. [Online]. Available: https://www.deeplearningbook.org/.

[142] F. Chollet, *Deep Learning with Python*. Shelter Island, New York: Manning Publications Co., 2018, 361 pp., ISBN: 978-1-61729-443-3. [Online]. Available: https://www.manning.com/books/deep-learning-with-python.

[143] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. 2019, ISBN: 978-1-4920-3264-9. [Online]. Available: http://shop.oreilly.com/product/0636920142874.do.

[144] S. S. Skiena, *The Data Science Design Manual*. Springer, 2017, ISBN: 978-3-319-55444-0. [Online]. Available: https://link.springer.com/book/10.1007/978-3-319-55444-0.

[145] A. Zheng and A. Casari, *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*. O'Reilly Media, Inc., 2018, ISBN: 978-1-4919-5324-2. [Online]. Available: https://www.oreilly.com/library/view/feature-engineering-for/9781491953235.

[146] P. Green and T. Hübsch, 'Calabi-Yau manifolds as complete intersections in products of complex projective spaces', *Communications in Mathematical Physics*, vol. 109, no. 1, pp. 99–108, 1987, ISSN: 0010-3616, 1432-0916. DOI: 10/bb29bx.

[147] P. Candelas, A. M. Dale, C. A. Lütken *et al.*, 'Complete intersection Calabi-Yau manifolds', *Nuclear Physics B*, vol. 298, no. 3, pp. 493–525, 1988, ISSN: 05503213. DOI: 10/cbx253.

[148] P. S. Green, T. Hübsch and C. A. Lütken, 'All the Hodge Numbers for All Calabi-Yau Complete Intersections', *Classical and Quantum Gravity*, vol. 6, no. 2, pp. 105–124, 1989, ISSN: 0264-9381, 1361-6382. DOI: 10/c6d47n.

[149] L. B. Anderson, X. Gao, J. Gray *et al.*, 'Fibrations in CICY threefolds', *Journal of High Energy Physics*, vol. 2017, no. 10, p. 77, 2017, ISSN: 1029-8479. DOI: 10/ggkmrn. arXiv: 1708.07907.

[150] M. Kreuzer and H. Skarke, 'Complete classification of reflexive polyhedra in four dimensions', *Advances in theoretical and mathematical physics*, vol. 4, no. 6, pp. 1209–1230, 2000, ISSN: 10950761, 10950753. DOI: 10.4310/ATMP.2000.v4.n6.a2. arXiv: hep-th/0002240.

[151] Wes McKinney, 'Data Structures for Statistical Computing in Python', in *Proceedings of the 9th Python in Science Conference*, S. van der Walt and Jarrod Millman, Eds., 2010, pp. 56–61. DOI: 10/ggr6q3.

[152] J. D. Hunter, 'Matplotlib: A 2D Graphics Environment', *Computing in Science Engineering*, vol. 9, no. 3, pp. 90–95, 2007. DOI: 10/drbjhg.

[153] M. Waskom, O. Botvinnik, M. Gelbart *et al.*, *Mwaskom/Seaborn: V0.11.0 (September 2020)*, version v0.11.0, Zenodo, 2020. [Online]. Available: https://zenodo.org/record/4019146.

[154] F. Pedregosa, G. Varoquaux, A. Gramfort *et al.*, 'Scikit-Learn: Machine Learning in Python', *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011. [Online]. Available: http://jmlr.org/papers/v12/pedregosa11a.html.

[155] T. Head, M. Kumar, H. Nahrstaedt *et al.*, *Scikit-Optimize/Scikit-Optimize*, version v0.8.1, Zenodo, 2020. [Online]. Available: https://zenodo.org/record/4014775.

[156] M. Abadi, A. Agarwal, P. Barham *et al.*, 'TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems', 2015. [Online]. Available: https://www.tensorflow.org/.

[157] T. Hübsch, *Calabi-Yau Manifolds: A Bestiary for Physicists*. World Scientific, 1992, ISBN: 978-981-02-1927-7. [Online]. Available: https://www.worldscientific.com/worldscibooks/10.1142/1410.

[158] P. Green and T. Hübsch, 'Polynomial deformations and cohomology of Calabi-Yau manifolds', *Communications in Mathematical Physics*, vol. 113, no. 3, pp. 505–528, 1987, ISSN: 0010-3616, 1432-0916. DOI: 10/fjxkft.

[159] C. A. Lütken, 'Recent Progress in Calabi-Yauology', 1988. DOI: 10/cw4cz2.

[160] S. Thrun, 'Is Learning the N-Th Thing Any Easier than Learning the First?', in *Advances in Neural Information Processing Systems*, 1996, pp. 640–646.

[161] R. Caruana, 'Multitask Learning', *Machine learning*, vol. 28, no. 1, pp. 41–75, 1997. DOI: 10/d3gsgj.

[162] J. Baxter, 'A Model of Inductive Bias Learning', *Journal of artificial intelligence research*, vol. 12, pp. 149–198, 2000. DOI: 10/gg66h8.

[163] A. Maurer, M. Pontil and B. Romera-Paredes, 'The Benefit of Multitask Representation Learning', *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2853–2884, 2016.

[164] A. Ndirango and T. Lee, 'Generalization in Multitask Deep Neural Classifiers: A Statistical Physics Approach', in *Advances in Neural Information Processing Systems*, 2019, pp. 15 862–15 871. arXiv: 1910.13593.

[165] C. Szegedy, W. Liu, Y. Jia *et al.*, 'Going Deeper with Convolutions', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9. arXiv: 1409.4842.

[166] C. Szegedy, V. Vanhoucke, S. Ioffe *et al.*, 'Rethinking the Inception Architecture for Computer Vision', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826. arXiv: 1512.00567.

[167] C. Szegedy, S. Ioffe, V. Vanhoucke *et al.* (2016). 'Inception-v4, Inception-Resnet and the Impact of Residual Connections on Learning'. arXiv: 1602.07261.

[168] B. Caffo, J. Leek and R. D. Peng. (). 'Data Science Specialization', Coursera, [Online]. Available: https://www.coursera.org/specializations/jhu-data-science.

[169] J. Bergstra and Y. Bengio, 'Random Search for Hyper-Parameter Optimization', *Journal of Machine Learning Research*, vol. 13, pp. 281–305, Feb 2012.

[170] G. Rudolph, 'Convergence Analysis of Canonical Genetic Algorithms', *IEEE transactions on neural networks*, vol. 5, no. 1, pp. 96–101, 1994. DOI: 10/fw8z8k.

[171] J. Snoek, H. Larochelle and R. P. Adams, 'Practical Bayesian Optimization of Machine Learning Algorithms', in *Advances in Neural Information Processing Systems*, 2012, pp. 2951–2959.

[172] B. Shahriari, K. Swersky, Z. Wang *et al.*, 'Taking the Human out of the Loop: A Review of Bayesian Optimization', *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2015. DOI: 10/f75n9c.

[173] J. Močkus, 'On bayesian methods for seeking the extremum', in *Optimization Techniques IFIP Technical Conference Novosibirsk, July 1–7, 1974*, G. I. Marchuk, Ed., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 1975, pp. 400–404, ISBN: 978-3-540-37497-8. DOI: 10/dh23rk.

[174] S. Ioffe and C. Szegedy. (2015). 'Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift'. arXiv: 1502.03167.

[175] C. Peng, X. Zhang, G. Yu *et al.*, 'Large Kernel Matters. Improve Semantic Segmentation by Global Convolutional Network', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4353–4361.

[176] S. Mallat, 'Understanding Deep Convolutional Networks', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20 150 203, 2016. DOI: 10/gcsgwj.

[177] I. Goodfellow, J. Pouget-Abadie, M. Mirza *et al.*, 'Generative Adversarial Nets', in *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes *et al.*, Eds., Curran Associates, Inc., 2014, pp. 2672–2680. [Online]. Available: http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf.

[178] D. P. Kingma and M. Welling. (2014). 'Auto-Encoding Variational Bayes'. arXiv: 1312.6114.

[179] D. J. Rezende, S. Mohamed and D. Wierstra. (2014). 'Stochastic Backpropagation and Approximate Inference in Deep Generative Models'. arXiv: 1401.4082.

[180] J.-Y. Zhu, T. Park, P. Isola *et al.*, 'Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks', in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2223–2232.

[181] L. Ardizzone, J. Kruse, S. Wirkert *et al.* (2019). 'Analyzing Inverse Problems with Invertible Neural Networks'. arXiv: 1808.04730, [Online]. Available: http://arxiv.org/abs/1808.04730.

[182] M. Gori, G. Monfardini and F. Scarselli, 'A New Model for Learning in Graph Domains', in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, vol. 2, 2005, pp. 729–734. DOI: 10/cr2f33.

[183] F. Scarselli, A. C. Tsoi, M. Gori *et al.*, 'Graphical-Based Learning Environments for Pattern Recognition', in *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, 2004, pp. 42–56, ISBN: 978-3-540-22570-6 978-3-540-27868-9. DOI: 10/dtgbgk.

[184] F. Monti, D. Boscaini, J. Masci *et al.*, 'Geometric Deep Learning on Graphs and Manifolds Using Mixture Model Cnns', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5115–5124.

[185] J. Gray, A. S. Haupt and A. Lukas, 'All Complete Intersection Calabi-Yau Four-Folds', *Journal of High Energy Physics*, vol. 2013, no. 7, p. 70, 2013, ISSN: 1029-8479. DOI: 10/ghf4n2. arXiv: 1303.1832.

[186] ——, 'Topological Invariants and Fibration Structure of Complete Intersection Calabi-Yau Four-Folds', *Journal of High Energy Physics*, vol. 2014, no. 9, p. 93, 2014, ISSN: 1029-8479. DOI: 10/ghf4n3. arXiv: 1405.2073.

[187] L. B. Anderson, F. Apruzzi, X. Gao *et al.*, 'A New Construction of Calabi-Yau Manifolds: Generalized CICYs', *Nuclear Physics B*, vol. 906, pp. 441–496, 2016, ISSN: 05503213. DOI: 10/f8kd3j. arXiv: 1507.03235.

[188] C. Cortes and V. Vapnik, 'Support-Vector Networks', *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[189] H. Drucker, C. J. Burges, L. Kaufman *et al.*, 'Support Vector Regression Machines', in *Advances in Neural Information Processing Systems*, 1997, pp. 155–161.

[190] J. R. Quinlan, 'Induction of decision trees', *Machine learning*, vol. 1, no. 1, pp. 81–106, 1986, ISSN: 1573-0565. DOI: 10/ctd6mv.

[191] K. Wittkowski, 'Classification and Regression Trees - l. Breiman, j. H. Friedman, r. A. Olshen and c. J. Stone.', *Metrika*, vol. 33, pp. 128–128, 1986, ISSN: 0026-1335. [Online]. Available: http://eudml.org/doc/176041.

[192] T. K. Ho, 'Random Decision Forests', in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, vol. 1, 1995, 278–282 vol.1. DOI: 10/c7x7s8.

[193] J. H. Friedman, 'Greedy Function Approximation: A Gradient Boosting Machine', *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001, ISSN: 0090-5364. DOI: 10/fbgj35. JSTOR: 2699986.

[194] ——, 'Stochastic Gradient Boosting', *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367–378, 2002, ISSN: 0167-9473. DOI: 10/fxb956.

[195] T. Fawcett, 'Using Rule Sets to Maximize ROC Performance', in *Proceedings 2001 IEEE International Conference on Data Mining*, 2001, pp. 131–138. DOI: 10/d7q2hk.

[196] X. Glorot, A. Bordes and Y. Bengio, 'Deep Sparse Rectifier Neural Networks', in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, 2011, pp. 315–323.

[197] J. Tompson, R. Goroshin, A. Jain *et al.*, 'Efficient Object Localization Using Convolutional Networks', 2015. DOI: 10/ggtmv2. arXiv: 1411.4280.

[198] D. E. Rumelhart, G. E. Hinton and R. J. Williams, 'Learning Representations by Back-Propagating Errors', vol. 323, no. 6088, pp. 533–536, 1986. DOI: 10/cvjdpk.

[199] D. P. Kingma and J. Ba. (2017). 'Adam: A Method for Stochastic Optimization'. arXiv: 1412.6980.

[200] N. Srivastava, G. Hinton, A. Krizhevsky *et al.*, 'Dropout: A Simple Way to Prevent Neural Networks from Overfitting', *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014. [Online]. Available: http://jmlr.org/papers/v15/srivastava14a.html.