



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
MAESTRÍA EN CIENCIAS (FÍSICA)
DEPARTAMENTO DE FÍSICA TEÓRICA, INSTITUTO DE FÍSICA

APRENDIZAJE AUTOMÁTICO EN LA PARAMETRIZACIÓN DEL ESPACIO DE
ORBIFOLIOS HETERÓTICOS

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS (FÍSICA)

PRESENTA:
JOSÉ ANTONIO MORALES ÁLVAREZ

TUTOR PRINCIPAL
DR. SAÚL NOÉ RAMOS SÁNCHEZ
(INSTITUTO DE FÍSICA)
MIEMBROS DEL COMITÉ TUTOR
DR. ERICK LEONARDO PATIÑO JAIDAR
(FACULTAD DE CIENCIAS)
DR. HERNANDO QUEVEDO CUBILLOS
(INSTITUTO DE CIENCIAS NUCLEARES)

CIUDAD DE MÉXICO, JULIO, 2021



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Dedicada a Ximena, mi hermana y mis padres

Quiero agradecer primeramente a mi asesor, el Dr. Saúl Ramos Sánchez, quien sin su guía académica y apoyo en todo momento, este trabajo no habría sido posible. Quiero agradecer también su paciencia y comprensión en todo momento durante momentos demasiado complicados y que a pesar de todo no se rindió conmigo. Al Dr. Hernando Quevedo y al Dr. Leonardo Patiño que, conformando mi comité tutor, siempre estuvieron atentos a mi progreso y avance semestre a semestre.

A todos mis profesores a lo largo de la Maestría, por sus cursos y por compartir sus conocimientos de una manera tan apasionada. Al Dr. Jorge Flores que siempre estuvo pendiente de mí y siempre trataba de motivarme a colaborar en proyectos interesantes.

A Ximena, que sin su amor incondicional y apoyo infinito, no habría salido adelante de situaciones complicadas y no podría haber salido adelante para continuar con este trabajo.

A mi hermana, que aunque ha estado un poco ausente, siempre está presente dándome ánimos y preocupándose por mi salud. A mis padres, por todo su apoyo y consejos empujándome siempre a ser mejor.

Al CONACyT por la beca recibida durante el periodo que estuve en el programa de maestría y con número de apoyo 717949.

Índice general

Resumen	II
Introducción	II
1. Fundamentos de teoría de cuerdas	1
1.1. La cuerda bosónica	3
1.1.1. La acción de cuerda bosónica	3
1.1.2. Ecuaciones de movimiento	7
1.1.3. Solución de las ecuaciones de campo	9
1.1.4. Cuantización en la norma de cono de luz	11
1.2. Supercuerdas	15
1.2.1. La acción de teoría de supercuerdas	15
1.2.2. Condiciones de frontera y expansión en modos	19
1.2.3. Cuantización canónica de la teoría de supercuerdas RNS	21
1.2.4. Estados base del sector Ramond vs. estados base del sector Neveu-Schwarz	22
1.2.5. Cuantización en la norma de cono de luz	23
1.2.6. Proyección GSO	23
1.2.7. Espectro para la supercuerda cerrada	26
1.3. La cuerda heterótica	28
1.3.1. La hoja de mundo	28
1.3.2. Descripción covariante	29
1.3.3. Norma de cono de luz	29

1.3.4.	Expansión en modos	30
1.3.5.	Compactificación en T^{16}	33
1.3.6.	Espectro no masivo de la cuerda heterótica	35
2.	Orbifolios heteróticos	37
2.1.	Compactificaciones toroidales	38
2.2.	Orbifolios toroidales	42
2.3.	Compactificación en orbifolios	47
2.3.1.	Espectro no masivo	53
2.3.2.	Vectores de traslación y espectro locales	57
2.3.3.	Clasificación de orbifolios simétricos toroidales	58
2.4.	Orbifolio \mathbb{Z}_3	61
2.5.	Orbifolios heteróticos \mathbb{Z}_6	69
3.	Machine Learning y el autoencoder	74
3.1.	Fundamentos de <i>machine learning</i>	74
3.1.1.	Capas comunes en redes neuronales	76
3.1.2.	Redes neuronales prealimentadas, capas completamente conectadas	77
3.1.3.	Funciones de activación	80
3.1.4.	Entrenamiento y evaluación de redes neuronales	83
3.1.5.	Gradiente descendente	88
3.1.6.	Inicialización de parámetros	89
3.1.7.	Funciones de pérdida	90
3.1.8.	Sobreentrenamiento y sobreajuste de Redes Neuronales	94
3.1.9.	Los <i>Autoencoders</i>	95
3.2.	El orbifold	97
3.3.	<i>Autoencoder</i> en orbifolios \mathbb{Z}_6 -II	99
3.3.1.	Preprocesamiento de datos	99
4.	Resultados	102
4.1.	Definición del autoencoder	102

4.2. Capa latente o cuello de botella	103
4.3. Modelos tipo MSSM, $SU(5)$ y Pati-Salam en la capa latente	107
4.4. Análisis del espacio latente	110
5. Conclusiones	116
A. El sistema de raíces $E_8 \times E_8$	123

Resumen

En esta tesis se presenta la implementación de un tipo especial de red neuronal llamada *autoencoder* para tratar de caracterizar el paisaje de modelos con geometría de compactificación en orbifolios heteróticos \mathbb{Z}_6 -II. El autoencoder parametriza a cada modelo con dos parámetros, los cuales viven en un espacio latente bidimensional. Generando un total de 4,000,000 de modelos con estas compactificaciones, empleamos 1,000,000 para el entrenamiento de la red neuronal y posteriormente graficamos los valores de los parámetros para 2,000,000 de modelos utilizando el autoencoder resultante. Identificamos un total de 30 regiones independientes en el espacio latente donde se acumula el total de modelos, acotando también el valor de todos los posibles parámetros. Utilizando el orbifold generamos modelos con este orbifolio heterótico en dos categorías fenomenológicamente interesantes, aquellos que tengan un espectro de partículas de 3 generaciones 4D de los grupos de norma Pati-Salam (16 modelos) y $SU(5)$ (457 modelos). Agregamos también un conjunto de modelos con esta geometría compatibles con el Modelos Estándar Mínimo Supersimétrico. Posteriormente, los modelos dentro de estas tres categorías son parametrizados con nuestro autoencoder. Contamos el número de modelos por categoría en cada una de las 30 regiones e identificamos aquellas que sean las más prometedoras para búsqueda de modelos con interés fenomenológico.

Introducción

La teoría de cuerdas logra incorporar una descripción cuántica de la gravedad. Al cuantizar las cuerdas, se encuentran modos de oscilación con masa cero, espín 1 y propiedades que permiten describir a las partículas mensajeras de las interacciones del Modelo Estándar y, de hecho, permite obtener de manera natural algunas propuestas de Gran Unificación (GUTs). Se pueden obtener también estados de vibración de la cuerda con masa cero y espín $\frac{1}{2}$, es decir, fermiones como los bloques básicos de la materia en el Modelo Estándar. También existen maneras de obtener ingredientes del tipo de el Higgs, teniendo que es posible que se puedan reproducir todos los ingredientes del Modelo Estándar (SM, por sus siglas en inglés) a través un planteamiento más fundamental. Una de las más grandes consecuencias de esta teoría es que predice (o requiere) la existencia de 6 dimensiones adicionales. Para explicar que estas dimensiones no han sido observadas, se postula que son muy pequeñas, y que están compactificadas en estructuras matemáticas muy específicas que determinan indirectamente la física de la teoría. Una de estas estructuras en las cuales se lleva a cabo la compactificación son los orbifolios toroidales que se definen como el cociente de un toro con una simetría discreta codificada en un grupo llamado grupo de punto. La elección de esta simetría discreta se puede restringir a sólo grupos abelianos \mathbb{Z}_N o $\mathbb{Z}_N \times \mathbb{Z}_M$ para ciertos valores de N y M . Además, una vez elegida la simetría discreta que rige a la compactificación, también se tiene constriñida la elección de la retícula que genera al toro inicial. Escogiendo cierta compactificación, las propiedades físicas de la teoría resultante están completamente definidas y en la mayoría de los casos estos modelos son diferentes de lo observado en el Modelo Estándar. La elección de la compactificación y la teoría 4D resultante no son únicas, dando lugar a un gran conjunto de modelos de cuerdas en 4D llamado generalmente como el paisaje de cuerdas. El número de modelos de cuerdas 4D se ha acotado a ser por lo menos 10^{1500} [1]. Lo deseable es que en este vasto paisaje de modelos se pueda realizar una búsqueda de aquellos que sean compatibles con el Modelo Estándar Mínimo Supersimétrico. Una búsqueda de dichos modelos fenomenológicamente interesantes implica una tarea demasiado compleja, pues implementar búsquedas en un conjunto de modelos tan enorme podría tomar demasiado tiempo incluso con el poder computacional con que se cuenta actualmente.

Los parámetros que definen por completo a un modelo de compactificación son del orden de $\mathcal{O}(100)$ los cuales especifican características físicas de la compactificación y cumplen con ciertas

condiciones de consistencia, tales como la invariancia modular. Una vez fijados estos parámetros, se puede calcular el grupo de norma 4D resultante y verificar si éste es de interés físico, como ser compatible con el Modelo Estándar Mínimo Supersimétrico (MSSM, por sus siglas en inglés). De gran interés fenomenológico también tenemos los casos en que los modelos tienen espectros con un número de 3 generaciones 4D de los grupos de norma para teorías de gran unificación (GUTs), tales como el grupo $SU(5)$ o el grupo $SU(2) \times SU(2) \times SU(4)$ (de Pati-Salam o PS). Una teoría de gran unificación constituye una suposición en la cual la naturaleza cuenta con "más simetría" que la del Modelo Estándar pero que se reduce a la del mismo a bajas energías motivada por la incompletez probada del SM. Las GUTs son de gran interés fenomenológico porque permiten explorar algunas ideas de cómo extender al Modelo Estándar. El estudio y búsqueda de una cantidad tan grande de modelos suele no ser ideal dado el enorme tiempo de cómputo que se requiere y por ello en los últimos años se han aplicado nuevas técnicas para llevar a cabo dichas tareas. Una de las técnicas actualmente empleadas es la del Aprendizaje Automático (*Machine Learning*), que consiste en utilizar un conjunto de datos de interés para que la computadora encuentre características ocultas, tales como regularidades, patrones o tendencias para que posteriormente logre identificar también características en conjuntos de datos ajenos al original. La ventaja de dicha técnica se fundamenta en el hecho de que es posible procesar una gran cantidad de información en poco tiempo, pues no se analiza dato por dato, si no al conjunto completo. Dentro de muchas de las técnicas que emplea Machine Learning se encuentra el diseño de Redes Neuronales (RNs), las cuales consisten en un conjunto capas contiguas, cada una de ellas con cierto número de nodos; la capa de entrada y salida corresponden a la información que podemos ver. Las RNs procesan la información recibida por la capa de entrada a través de las capas intermedias (la información de dichas intermedias está oculta y sólo se utiliza durante el procesamiento) hasta obtener un resultado en la capa de salida. La idea es entrenar a la RN utilizando dos conjuntos de datos: un conjunto que represente la información para la capa de entrada y un conjunto que represente el resultado que se desea obtener a partir de la información de entrada. El ajusta a los parámetros que definen el paso de la información a través de las capas hasta que los resultados de la capa de salida reproduzcan lo mejor posible el conjunto de datos que deseamos obtener a partir de los datos de entrada.

En los últimos años se ha utilizado esta técnica computacional para el estudio de diversos aspectos de la teoría de cuerdas [2, 3, 4, 5], además se puede encontrar un compendio de aplicaciones a dicha teoría en [6]. En [2] se estudia el paisaje de modelos de la teoría de cuerdas heteróticas con compactificación de grupo de punto \mathbb{Z}_6 -II (orbifolio heterótico \mathbb{Z}_6 -II). Esta geometría es de gran interés fenomenológico pues ha mostrado ser muy útil en la búsqueda de modelos compatibles con el MSSM [7, 8]. Allí, se diseña un tipo especial de red neuronal conocida como *autoencoder*, que consiste en una red simétrica cuyas capas de entrada y salida tienen la misma información, tratando de simular una función identidad. Generalmente, un autoencoder posee una capa intermedia (capa latente) con la menor cantidad de nodos respecto al resto y cuyos valores son de gran interés pues se les puede considerar como una parametrización de la información con menos parámetros que los de la información original. En [2] se diseña un autoencoder para modelos con orbifolios heteróticos \mathbb{Z}_6 -II, de tal manera que la capa latente sólo tenga dos nodos para así parametrizar a los modelos

con sólo dos parámetros. Al espacio bidimensional en el que viven los parámetros se le llama espacio latente. En la presente tesis tratamos de reproducir los resultados observados en dicho trabajo, y exploramos posibles conclusiones alternativas con el propósito de generalizar esta técnica para la clasificación de orbifolios toroidales con geometría arbitraria.

Estructura

La presente tesis está organizada como sigue. En el Capítulo 1 presentamos las generalidades de la Teoría de Cuerdas, comenzando por su construcción. Se presentan detalles de la Teoría de Cuerdas Bosónica y posteriormente se introduce su generalización al caso donde se incluye supersimetría, la Teoría de Supercuerdas. Por último detallamos aspectos de la Teoría de Cuerdas Heteróticas, las dimensiones espacio-temporales en las que se trabaja y un primer acercamiento a la compactificación toroidal de 16 grados de libertad bosónicos de norma. En el Capítulo 2 profundizamos la discusión sobre el proceso de compactificación en orbifolios toroidales simétricos, su clasificación y ventajas sobre las compactificaciones toroidales. Posteriormente, se detallan los aspectos geométricos específicos de los orbifolios heteróticos con grupo de punto \mathbb{Z}_3 y \mathbb{Z}_6 -II. En el Capítulo 3 hacemos un repaso de las generalidades de Machine Learning, detallando la construcción de diversas redes neuronales y sus parámetros intrínsecos, incluyendo sus funciones de activación. Se hace hincapié en un tipo de redes neuronales específico llamadas *autoencoder*, cuya funcionalidad es la de parametrizar un conjunto de datos con menos parámetros que los que originalmente posee. Proporcionamos la estructura de nuestro autoencoder con el cual se parametrizará mediante dos parámetros a una gran cantidad de modelos de orbifolios heteróticos \mathbb{Z}_6 -II. Mencionamos la herramienta *el orbifolder*, que es un programa que nos permite generar una gran cantidad de modelos con estructuras de compactificación específicas y que nos ayuda a generar aproximadamente $\mathcal{O}(4,000,000)$ de modelos con orbifolios \mathbb{Z}_6 -II inequivalentes para el entrenamiento del autoencoder. Finalmente, en dicho capítulo se detalla cómo es que el autoencoder recibe los datos de entrada y salida de cada uno de estos modelos. En el Capítulo 4 presentamos nuestros resultados. Se analiza la forma de la capa latente para un total de 2,000,000 de modelos y también se presenta la parametrización de modelos fenomenológicamente interesantes: aquellos que son compatibles con el MSSM y los que poseen espectros de partículas con 3 generaciones de los grupos de norma $SU(5)$ y de Pati-Salam. Se identifican regiones en las cuales hay mayor aglomeración de estos tres grupos de modelos y se caracteriza el espacio latente por la densidad de los mismos. Finalmente, en el Capítulo 5 presento mis conclusiones.

Capítulo 1

Fundamentos de teoría de cuerdas

A lo largo de la historia, el desarrollo de la física ha tenido eventos de suma importancia conocidos como unificaciones, evento en los cuales se descubre que diferentes fenómenos están relacionados entre sí y las teorías físicas son modificadas para tomar en cuenta esta relación. Uno de estos eventos reconoció la profunda relación que tienen la electricidad y el magnetismo. Desde hace milenios estos dos fenómenos fueron reconocidos como independientes, sin embargo, en 1819 Hans Christian Oersted descubrió que una corriente eléctrica en un alambre puede desviar la aguja de una brújula que se encuentra cerca de dicho alambre. Poco después, Jean-Baptiste Biot y Felix Savart (1820) y André-Marie Ampère (1820-1825) reportaron cuáles son las reglas para que corrientes eléctricas produzcan campos magnéticos. Michael Faraday (1831) demostró que un campo magnético variable genera campos eléctricos. Todos estos fenómenos y conexiones entre ellos fueron condensados por James Clerk Maxwell (1865) en un conjunto de ecuaciones e incluso lo condujeron a la predicción de las ondas electromagnéticas. Éstas fueron las ecuaciones de Maxwell y consideran a la electricidad y magnetismo como manifestaciones de un mismo fenómeno físico: el electromagnetismo. Esta unificación fue inminente, pues la descripción separada de estos fenómenos lleva a inconsistencias. [9]

Existen dos grandes teorías que explican la mayoría de los fenómenos físicos a grandes y pequeñas escalas. Por un lado tenemos la Teoría de la Relatividad General, formulada por Albert Einstein, que establece que la gravedad es consecuencia de la deformación del espacio-tiempo provocada por la mera existencia de masa o energía. Por otro lado, tenemos a la Mecánica Cuántica, desarrollada por Erwin Schrödinger, Werner Heisenberg, Paul Dirac y otros; dicho marco es el adecuado para describir fenómenos a escalas microscópicas. En adición a estos descubrimientos, 4 fuerzas fundamentales han sido identificadas en la naturaleza: la fuerza de gravedad, la fuerza electromagnética, la fuerza débil (responsable del decaimiento nuclear beta) y la fuerza fuerte (que mantiene unidos a los quarks de los nucleones). En 1960 ocurrió otro evento de unificación cuando el modelo de Weinberg-Salam de las interacciones electrodébiles describió en un mismo marco al electromagne-

tismo y la fuerza débil. El Modelo Estándar de la Física de Partículas describe de manera cuántica a las últimas tres interacciones en el marco de la Teoría Cuántica de Campos. En esta teoría se cuenta con la descripción de partículas con espín $\frac{1}{2}$ llamadas fermiones y que conforman los bloques básicos de la materia. También describe partículas de espín 1 llamados bosones y que son las mensajeras de las interacciones. También describe una partícula de espín cero llamada partícula de Higgs relacionada con el proceso de Higgs, el cual dota de masa a las partículas del Modelo Estándar. Sin embargo, se sabe que dicha teoría es incompleta pues existen aspectos de la Física de partículas que no logra incorporar: descripción de la materia oscura, masa de los neutrinos, problema de jerarquía, entre otros [9].

La gravedad debe ser incluida en el marco de la física de partículas, con o sin unificación, si uno desea tener una teoría completa. Los efectos de la fuerza gravitacional son considerablemente despreciables a nivel microscópico, pero son cruciales en estudios de cosmología del universo temprano. Como ya mencionamos, el Modelo Estándar es una teoría cuántica, mientras que la Relatividad General es una teoría clásica de la gravedad. Para incluir a la gravedad en el mismo marco de la física de partículas, la gravedad debe ser promovida a una teoría cuántica. Sin embargo, el proceso de cuantización de dicha interacción ha encontrado serias dificultades.

La teoría de cuerdas es un excelente candidato para una teoría unificadora de todas las fuerzas de la naturaleza. En la teoría de cuerdas, todas las fuerzas son unificadas de una manera profunda y significativa. La teoría de cuerdas es una teoría cuántica, y, porque incluye a la gravedad, es una teoría cuántica de la gravedad. En esta teoría, cada partícula se identifica como un modo de vibración de una cuerda microscópica elemental. Uno de los estados vibracionales de las cuerdas es el gravitón, el cuanto del campo gravitacional. Como sólo hay un tipo de cuerda, y todas las partículas provienen de las vibraciones de cuerdas, todas las partículas son naturalmente incorporadas en una sola teoría. Con el simple planteamiento de la existencia de un componente fundamental llamado cuerda se tienen varias consecuencias directas, algunas siendo virtudes que obligan a considerar a la teoría como la unificadora de las 4 interacciones fundamentales. La teoría de cuerdas predice la existencia de la gravedad al intentar describir consistentemente a una cuerda cuántica relativista. Se pueden obtener estados de vibración de la cuerda con masa cero y espín $\frac{1}{2}$, es decir, fermiones como los bloques de materia en el Modelo Estándar. Existe la manera de obtener ingredientes del tipo del Higgs, así que la teoría de cuerdas quizás sería capaz de reproducir todos los ingredientes del Modelo Estándar. La teoría de cuerdas no tiene ningún parámetro adimensional libre, a diferencia del problema que presenta el Modelo Estándar.

En este capítulo describiremos la fundamentación de la Teoría de Cuerdas, comenzando por la Teoría de Cuerdas Bosónica, que cuenta con un espectro de partículas bosónicas únicamente. Posteriormente, discutimos una generalización de la cuerda bosónica, conocida como Teoría de Supercuerdas, la cual, a partir de la introducción de supersimetría, permite la integración de fermiones en la teoría. Por último profundizamos en la llamada Teoría de Cuerdas Heteróticas, presentamos su planteamiento, características y consecuencias.

1.1. La cuerda bosónica

La teoría de cuerdas generaliza la idea de la partícula relativista que se mueve a través de un espacio tiempo D -dimensional, donde la acción en cuestión describe la dinámica de la línea de mundo (geodésica en el espacio-tiempo) trazada por la partícula. La acción respectiva está dada por la integral de la longitud invariante infinitesimal ds ,

$$S = -\alpha \int ds, \quad (1.1)$$

donde α es una constante, y hemos escogido unidades de tal manera que $c = \hbar = 1$. Para encontrar las ecuaciones de movimiento que describen a la geodésica (o línea de mundo) que va trazando la partícula en el espacio-tiempo intrínseco, se debe calcular la variación de la acción (1.1) e igualarla a cero. Entonces, el movimiento clásico de una partícula relativista es el camino que extremiza la distancia invariante. Esta idea será útil para generalizarlo al contexto de la teoría de cuerdas.

1.1.1. La acción de cuerda bosónica

La teoría de cuerdas se plantea generalizando la idea de la partícula puntual relativista. En lugar de describir sólo un punto a través del espacio-tiempo pensaremos en cuerdas minúsculas con cierta longitud (no nos preocuparemos de lo que estén hechas) que se mueven a través del espacio-tiempo, dichas cuerdas pueden ser de dos tipos: cerradas o abiertas. La trayectoria de estas cuerdas va trazando una superficie (en lugar de una línea para el caso de las partículas puntuales) (ver Figura 1.1). Se le llama *hoja de mundo* a la superficie 2-dimensional trazada por la trayectoria de la cuerda en el espacio-tiempo.

Para el caso de la partícula puntual relativista, vimos que la acción era la integral de la longitud invariante de la línea de mundo. Para las cuerdas definiremos el *área invariante* de Lorentz de la hoja de mundo de la cuerda. La acción de la cuerda relativista es proporcional a esta área y se conoce como acción de Nambu-Goto la cual tiene la forma [9]

$$S = -\beta \int dA, \quad (1.2)$$

donde dA es el área infinitesimal invariante de la hoja de mundo trazada y β es una constante. De ahora en adelante supondremos que el espacio-tiempo de fondo es plano.

Para el caso de una cuerda con cierta longitud inmersa en un espacio-tiempo de D dimensiones, se traza una hoja de mundo parametrizada por dos parámetros τ y σ , con τ siendo temporal y σ espacial. La inmersión de la cuerda en el espacio-tiempo D -dimensional estará dada por funciones o campos $X^\mu(\tau, \sigma)$. Los llamaremos grados de libertad de la cuerda (coordenadas que localizan a la hoja de mundo) son simbolizados por $X^\mu(\tau, \sigma)$ $\mu = 0, 1, \dots, D - 1$, que consiste en el barrido de curvas parametrizadas por σ a τ fijos (bajo cierta norma se puede interpretar como la fotografía

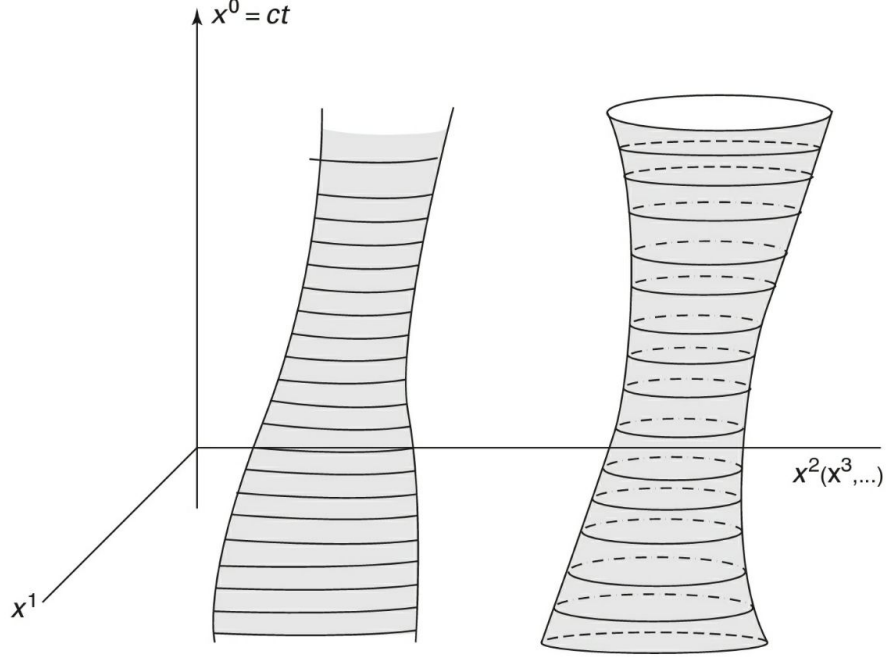


Figura 1.1: Hoja de mundo. A la izquierda apreciamos la hoja de mundo que va trazando una cuerda abierta, mientras que a la derecha apreciamos la hoja de mundo trazada por una cuerda cerrada y que se asemeja a un tubo. [9]

de una cuerda al tiempo τ) y puede ser una curva abierta o cerrada, por conveniencia se toma $0 \leq \sigma \leq \pi$. Así como en el caso de la partícula relativista en D dimensiones, queremos escribir una versión relativista de la cuerda inmersa en un espacio D -dimensional de forma covariante de Lorentz cuya física no dependa de alguna elección particular de los parámetros de la hoja de mundo σ y τ . Considerando que la métrica del espacio-tiempo plano es:

$$\eta_{\mu\nu} = \text{diag}(1, -1, \dots, -1) \quad \mu, \nu = 0, 1, \dots, D - 1. \quad (1.3)$$

Se puede demostrar entonces el área invariante se puede calcular con ayuda de estos parámetros y grados de libertad de la cuerda de tal manera que: [9]

$$dA = d\tau d\sigma \sqrt{(\dot{X} \cdot X')^2 - (\dot{X}^2)(X'^2)}, \quad (1.4)$$

donde $\dot{X}^\mu = \frac{\partial X^\mu}{\partial \tau}$ y $X'^\mu = \frac{\partial X^\mu}{\partial \sigma}$. Por lo que la acción de Nambu-Goto estaría dada por:

$$S_{NG} = -T \int d\tau d\sigma \sqrt{(\dot{X} \cdot X')^2 - (\dot{X}^2)(X'^2)}. \quad (1.5)$$

Donde T es una constante. Esta es la acción que describe a la cuerda relativista inmersa en un espacio-tiempo de D dimensiones. Sin embargo, calcular las ecuaciones de movimiento puede ser un poco complicado debido a la raíz cuadrada que tenemos en el integrando. Para deshacernos de ese problema podemos introducir un campo auxiliar $h_{\alpha\beta}(\tau, \sigma)$ (que se interpreta como la métrica inducida en la hoja de mundo) de signatura $(+,-)$ y donde $\alpha = 0$ y 1 se refiere a τ y σ respectivamente. La acción resultante es la de Polyakov y que caracteriza a la cuerda bosónica relativista: [10]

$$S = -\frac{T}{2} \int_{\tau_i}^{\tau_f} d\tau \int_0^\pi d\sigma \sqrt{-h} h^{\alpha\beta} \eta^{\mu\nu} \partial_\alpha X_\mu \partial_\beta X_\nu \quad (1.6)$$

donde h es el determinante de $h_{\alpha\beta}$. La acción cuenta con invariancia de Poincaré global del espacio-tiempo D -dimensional. Se puede demostrar que la acción de Nambu-Goto y de Polyakov son equivalentes [11]. Así que la acción que tomaremos como la que describe a la cuerda relativista será la acción de Polyakov-

A continuación enlistamos las simetrías locales de la hoja de mundo para la cuerda relativista: [11]

1. **Invariancia ante reparametrizaciones (también conocida con difeomorfismos):** La acción de Polyakov es invariante ante los cambios del parámetro σ a $\sigma' = f(\sigma)$ ya que $X^\mu(\tau, \sigma)$ transforma como escalar, mientras que el campo $h^{\alpha\beta}$ transforma como un 2-tensor, de tal forma que

$$X^\mu(\tau, \sigma) = X'^\mu(\tau, \sigma') \quad \text{y} \quad h_{\alpha\beta}(\tau, \sigma) = \partial_\alpha f^\gamma \partial_\beta f^\delta h'_{\gamma\delta}(\tau, \sigma'). \quad (1.7)$$

Con esta relación tenemos entonces que 1.6 es invariante ante reparametrizaciones. Como tenemos una redundancia en la teoría debido a esta simetría, entonces tenemos menos grados de libertad.

2. **Simetría de Weyl:** Estas transformaciones reescalan la métrica,

$$h_{\alpha\beta}(\tau, \sigma) \rightarrow h'_{\alpha\beta}(\tau, \sigma) = e^{2\phi(\sigma)} h_{\alpha\beta}(\tau, \sigma), \quad (1.8)$$

mientras que los escalares $X^\mu(\tau, \sigma)$ se mantienen invariantes bajo estos reescalamientos. Esta transformación es local pues el parámetro $\phi(\sigma)$ depende de las coordenadas de la hoja de mundo. Veamos cómo cambian las cosas en algunas partes de la acción. Por un lado tenemos que la transformación de $\sqrt{-h}$ está dada por

$$\begin{aligned} \sqrt{-h'} &= \sqrt{-\det(h'_{\alpha\beta})} \\ &= e^{2\phi(\sigma)/2} \sqrt{-\det(h_{\alpha\beta})} \\ &= e^{2\phi(\sigma)} \sqrt{-h}. \end{aligned} \quad (1.9)$$

Expandiendo 1.8 con ϕ como variable, obtendríamos $h'^{\alpha\beta} = e^{-2\phi}h^{\alpha\beta} = (1 - 2\phi + \dots)h^{\alpha\beta}$, por lo que la variación de la métrica respecto a esta simetría sería $\delta h^{\alpha\beta} = -2\phi h^{\alpha\beta}$ para $\phi \ll 1$, por lo que estamos hablando de una simetría local. Por lo que para $\sqrt{-h}h^{\alpha\beta}$ obtenemos que

$$\sqrt{-h'}h'^{\alpha\beta} = \sqrt{-h}e^{2\phi(\sigma)}e^{-2\phi(\sigma)}h^{\alpha\beta} = \sqrt{-h}h^{\alpha\beta}, \quad (1.10)$$

y por tanto, la variación de la acción bajo transformaciones de Weyl es nula. En consecuencia, el tensor de energía-momento asociado a esta teoría tiene traza nula, es decir, $h^{\alpha\beta}T_{\alpha\beta} = 0$. Para argumentar esta afirmación importante recordemos que, en general, el tensor de energía-momento se calcula como

$$T_{\alpha\beta} \equiv -\frac{2}{T} \frac{1}{\sqrt{h}} \frac{\delta S}{\delta h_{\alpha\beta}}. \quad (1.11)$$

Por lo que ante una transformación arbitraria de $h^{\alpha\beta}$, la variación de S puede ser escrita como

$$\delta S \equiv \int \frac{\delta S}{\delta h^{\alpha\beta}} \delta h^{\alpha\beta} = -\frac{T}{2} \int d\tau d\sigma \sqrt{-h} \delta h^{\alpha\beta} T_{\alpha\beta}, \quad (1.12)$$

por lo que con una transformación de Weyl vemos que la variación de la acción sería

$$\begin{aligned} \delta S &= -\frac{T}{2} \int d\tau d\sigma \sqrt{-h} \delta h^{\alpha\beta} T_{\alpha\beta} \\ &= -\frac{T}{2} \int d\tau d\sigma \sqrt{-h} (-2\phi) h^{\alpha\beta} T_{\alpha\beta} = 0, \end{aligned} \quad (1.13)$$

que como ya vimos antes, debe ser cero. Como $\sqrt{-h}$ y ϕ son arbitrarios, sólo podemos concluir que

$$h^{\alpha\beta} T_{\alpha\beta} = 0. \quad (1.14)$$

Tenemos entonces simetrías locales, o de norma, que indican la presencia de redundancias en la teoría respecto a sus grados de libertad y podemos utilizar dichas simetrías para remover estas redundancias. Se puede demostrar que si nuestra teoría es invariante bajo difeomorfismos y transformaciones de Weyl podemos fijar una norma de tal manera que la métrica intrínseca $h_{\alpha\beta}$ se vuelva plana. [11]

De tal manera que al fijar la norma, la métrica se convierte en

$$h_{\alpha\beta}(X) = \eta_{\alpha\beta} = \begin{pmatrix} -1 & 0 \\ 0 & 1. \end{pmatrix} \quad (1.15)$$

Debemos tener en cuenta que este proceso de fijar la norma fue gracias a la existencia de simetrías locales, y la posibilidad de considerar a la métrica como plana sólo es válido de manera local y, en general, no podemos hacer esta reducción de manera global a lo largo de toda la hoja de mundo.

En términos de la métrica plana proveniente de la norma fijada, tenemos que la acción de Polyakov 1.6 se reduce a

$$S = \frac{T}{2} \int d\tau d\sigma \left((\dot{X})^2 - (X')^2 \right) = \frac{T}{2} \int_{\tau_i}^{\tau_f} d\tau \int_0^\pi d\sigma \left((\dot{X})^2 - (X')^2 \right), \quad (1.16)$$

donde hemos denotado $\dot{X} \equiv \partial_\tau X^\mu = \frac{\partial X^\mu}{\partial \tau}$ y $X' \equiv \partial_\sigma X^\mu = \frac{\partial X^\mu}{\partial \sigma}$, notaciones que utilizaremos indistintamente de aquí en adelante. Es importante mencionar que estas dos simetrías locales no son las únicas presentes en la teoría, aunque hemos fijado las normas correspondientes a los difeomorfismos y a las simetrías de Weyl, aún es invariante ante otras simetrías locales conocidas como *transformaciones conformes*. [11]

1.1.2. Ecuaciones de movimiento

Supongamos que la topología de nuestra hoja de mundo nos permite extender de forma global la fijación de la norma en que la métrica $h^{\alpha\beta}$ es plana. Las ecuaciones de movimiento para los campos $X^\mu(\tau, \sigma)$ las obtenemos calculando la variación de la acción S de 1.6 respecto al campo $X^\mu \rightarrow X^\mu + \delta X^\mu$ y estableciendo que dicha variación es igual a cero. De tal manera que necesitamos:

$$\delta S = \frac{T}{2} \int d\tau d\sigma \left(2\dot{X}\delta\dot{X} - 2X'\delta X' \right), \quad (1.17)$$

integrando por partes ambos términos e igualando a cero la variación obtenemos las ecuaciones de movimiento para $X^\mu(\tau, \sigma)$:

$$(-\partial_\tau^2 + \partial_\sigma^2)X^\mu - T \int d\tau \left[X'\delta X^\mu \Big|_{\sigma=\pi} + X'\delta X^\mu \Big|_{\sigma=0} \right] = 0. \quad (1.18)$$

Es importante mencionar que esta ecuación se obtiene si anulamos $\delta X^\mu \Big|_{\partial\tau}$. Por lo que nos concentraremos en el caso para el cual $\delta X^\mu(\tau_i, \sigma) = \delta X^\mu(\tau_f, \sigma) = 0$ [9]. Esto quiere decir que la variación de X^μ es nula, y por tanto, dicha coordenada es fija en τ_i y fija en τ_f . Asumiremos siempre este hecho, por lo que siempre $\delta X^\mu \Big|_{\partial\tau} = 0$. Ahora bien, los términos de frontera asociados a σ dependen del tipo de cuerda en cuestión: abierta o cerrada. En cualquiera de los dos casos, veremos que los términos de frontera en 1.18 se anulan como sigue:

- **Cuerda Cerrada:** Para la cuerda cerrada consideramos condiciones de frontera periódica respecto al parámetro σ ,

$$X^\mu(\tau, \sigma + n) = X^\mu(\tau, \sigma) \quad (1.19)$$

donde n se interpreta como la longitud de la cuerda para un τ específico y casi siempre lo tomamos por convención como $n = \pi$. En los términos de frontera de 1.18 tenemos que

$\delta X^\mu(\tau, \sigma = 0) = \delta X^\mu(\tau, \sigma = n) = 0$ y por lo tanto las ecuaciones de movimiento quedan como:

$$(\partial_\tau^2 - \partial_\sigma^2)X^\mu(\tau, \sigma) = 0, \quad (1.20)$$

que es justamente la ecuación de onda unidimensional y cuyas soluciones conocemos en detalle.

- **Cuerda abierta (Condiciones de frontera de Neumann):** En estas condiciones de frontera exigimos lo siguiente:

$$\partial_\sigma X^\mu(\tau, \sigma = 0) = \partial_\sigma X^\mu(\tau, \sigma = n). \quad (1.21)$$

Bajo esta condición, también los términos de frontera de 1.18 se anulan, y por tanto las ecuaciones de movimiento de $X^\mu(\tau, \sigma)$ serían

$$(\partial_\sigma^2 - \partial_\tau^2)X^\mu(\tau, \sigma) = 0, \quad (1.22)$$

que es también la ecuación de onda plana bidimensional. Esta condición de frontera también se llama de puntos extremos libres y no impone ninguna condición sobre la variación $\delta X^\mu(\tau, \sigma^*)$, donde σ^* es el valor del parámetro σ en los extremos de la cuerda. Los puntos extremos son libres de moverse con una cierta restricción. Ver Figura 1.2. En [12] podemos encontrar el cálculo explícito que nos indica que siempre que tengamos condiciones de frontera de Neumann, necesariamente los puntos extremos $|dX^\mu/dt| = 1$, $\mu = 1, \dots, D - 1$ se mueven a la velocidad de la luz.

- **Cuerda abierta (Condiciones de frontera de Dirichlet):** En este caso exigimos que los puntos extremos tengan un valor constante, de tal manera que:

$$X^\mu(\tau, \sigma = 0) = X_0^\mu \quad X^\mu(\tau, \sigma = n) = X_n^\mu, \quad (1.23)$$

donde X_0^μ y X_n^μ son constantes. Igualmente, estas condiciones anulan los términos de frontera en la variación de la acción 1.18, de tal manera que las ecuaciones de movimiento quedan también como la ecuación de onda bidimensional:

$$(\partial_\sigma^2 - \partial_\tau^2)X^\mu(\tau, \sigma) = 0. \quad (1.24)$$

Una cosa importante a notar es que las condiciones de frontera de Neumann sí son invariantes de Poincaré, mientras que las condiciones de frontera de Dirichlet no lo son como se ve,

$$\begin{aligned} (X'^\mu) \Big|_{\sigma=0,n} &= (a^\mu{}_\nu X^\nu + b^\mu) \Big|_{\sigma=0,n} \\ &= a^\mu{}_\nu X_{0,n}^\nu + b^\mu \\ &\neq X_{0,n}^\mu \end{aligned} \quad (1.25)$$

Entonces bajo Poincaré, los puntos extremos realmente cambian y no se ven de la misma manera en cualquier sistema de referencia.

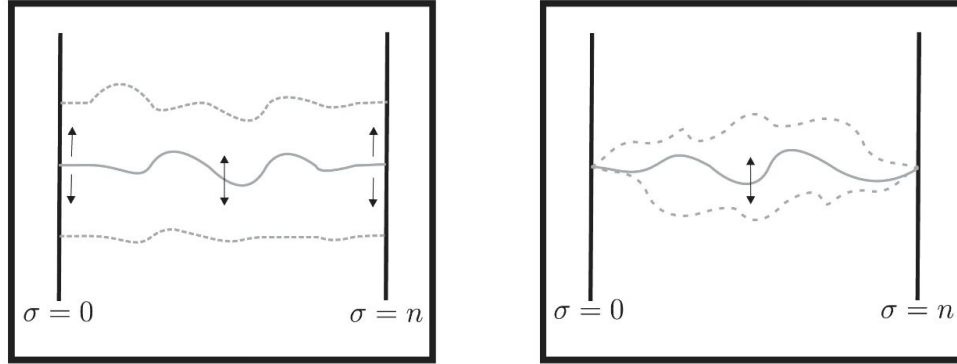


Figura 1.2: **Condiciones de frontera para la cuerda abierta.** Del lado izquierdo tenemos las condiciones de frontera de Neumann donde los extremos de la cuerda pueden moverse a lo largo de las fronteras siempre y cuando las derivadas respecto a σ se anulen. A la derecha tenemos la ilustración de las condiciones de frontera de Dirichlet y se observa que los extremos de la cuerda se mantienen fijos en la frontera. [11]

Vemos en los tres casos, las ecuaciones de movimiento corresponden a la ecuación de onda bidimensional para cada uno de los grados de libertad bosónico ($X^\mu(\tau, \sigma)$), lo único que cambian son las condiciones de frontera. En este trabajo nos concentraremos en la teoría de Cuerdas Heterótica (ver sección 1.3) la cual es una teoría de cuerdas cerradas, así que de aquí en adelante sólo nos concentraremos en este caso.

Adicionalmente a las ecuaciones de movimiento de los campos $X^\mu(\tau, \sigma)$, debemos imponer las ecuaciones de movimiento resultantes de calcular la variación de la acción de Polyakov con respecto a la métrica $h^{\alpha\beta}$ e igualarla a cero. Estas ecuaciones de campo están dadas por (ver [11]):

$$0 = T_{\alpha\beta} = \partial_\alpha X \cdot \partial_\beta X - \frac{1}{2} h_{\alpha\beta} h^{\gamma\delta} \partial_\gamma X \cdot \partial_\delta X, \quad (1.26)$$

que después de fijar la norma con métrica plana, estas condiciones se convierten en

$$T_{00} = T_{11} = -\frac{1}{2} \left(\frac{\partial X^\mu}{\partial \tau} \frac{\partial X_\mu}{\partial \tau} + \frac{\partial X^\mu}{\partial \sigma} \frac{\partial X_\mu}{\partial \sigma} \right) = 0 \quad (1.27)$$

y

$$T_{01} = T_{10} = -\frac{\partial X^\mu}{\partial \tau} \frac{\partial X_\mu}{\partial \sigma} = 0. \quad (1.28)$$

1.1.3. Solución de las ecuaciones de campo

Supondremos de nuevo que hemos fijado la norma de tal manera que la métrica plana es global sobre la hoja de mundo. Ahora resolveremos el sistema de ecuaciones introduciendo las coordenadas

de cono de luz para la hoja de mundo tales que

$$\sigma^\pm = (\tau \pm \sigma). \quad (1.29)$$

Con este cambio de variable, las ecuaciones de movimiento $(\partial_\tau^2 - \partial_\sigma^2)X^\mu = 0$ se reducen a

$$\partial_+ \partial_- X^\mu = 0, \quad (1.30)$$

donde $\partial_+ = \frac{\partial}{\partial \sigma^+}$ y $\partial_- = \frac{\partial}{\partial \sigma^-}$, mientras que las condiciones para el tensor de energía-momento se convierten en:

$$T_{++} = \partial_+ X^\mu \partial_+ X_\mu = 0, \quad (1.31)$$

$$T_{--} = \partial_- X^\mu \partial_- X_\mu = 0. \quad (1.32)$$

Estas últimas tres ecuaciones son las que debemos resolver para estudiar la dinámica de la cuerda relativista. La solución más general para (1.30) es la combinación lineal de dos funciones arbitrarias cuyos argumentos dependen solamente de una de las dos coordenadas de luz, es decir,

$$X^\mu(\sigma^+, \sigma^-) = X_R^\mu(\sigma^-) + X_L^\mu(\sigma^+). \quad (1.33)$$

Estas dos funciones independientes se pueden interpretar como la superposiciones ondas que se propagan hacia la izquierda (X_L^μ) y onda que se propagan hacia la derecha (X_R^μ). Por tanto, en general la solución está dada por

$$X^\mu(\tau, \sigma) = \underbrace{X_R^\mu(\tau - \sigma)}_{\text{oscilador derecho}} + \underbrace{X_L^\mu(\tau + \sigma)}_{\text{oscilador izquierdo}}. \quad (1.34)$$

Por simplicidad, a lo largo de este trabajo, a las ondas que se propagan a la izquierda les llamaremos osciladores izquierdos mientras que a la onda que se propagan hacia la derecha las llamaremos osciladores derechos.

Recordemos que en este trabajo nos estamos concentrando en sólo cuerdas cerradas, así que solo aplicaremos dichas condiciones de frontera. Recordemos que las condiciones de frontera para la cuerda cerrada exigen que $X^\mu(\tau, \sigma + n) = X^\mu(\tau, \sigma)$, es decir, condiciones periódicas y la expansión en modos de cada uno de los osciladores sería la siguiente:

$$X_R^\mu = \frac{1}{2}x^\mu + \frac{1}{2}l^2(\tau - \sigma)p^\mu + \frac{i}{2}l \sum_{n \neq 0} \frac{1}{n} \alpha_n^\mu e^{-2in(\tau - \sigma)}, \quad (1.35)$$

$$X_L^\mu = \frac{1}{2}x^\mu + \frac{1}{2}l^2(\tau + \sigma)p^\mu + \frac{i}{2}l \sum_{n \neq 0} \frac{1}{n} \tilde{\alpha}_n^\mu e^{-2in(\tau + \sigma)}, \quad (1.36)$$

donde x^μ es una constante y se le interpreta como el centro de masa de la cuerda, p^μ es también una constante y se relaciona con el momento de la cuerda, l_s es la longitud de la cuerda, que es constante, $T = \frac{1}{2n\alpha'}$ y $\alpha' = \frac{1}{2}l^2$. Entonces la solución global estaría dada por

$$X^\mu = X_R^\mu + X_L^\mu = \underbrace{x^\mu + \frac{1}{2}l^2 p^\mu \tau}_{\text{centro de masa de la cuerda}} + \frac{i}{2}l \underbrace{\sum_{n \neq 0} \left(\frac{1}{n} \alpha_n^\mu e^{-2in(\tau-\sigma)} + \frac{1}{n} \tilde{\alpha}_n^\mu e^{-2in(\tau+\sigma)} \right)}_{\text{oscilaciones de la cuerda}}. \quad (1.37)$$

Como X^μ debe ser real, es decir, $(X^\mu)^* = X^\mu$ y esto se cumple siempre que

$$(\alpha_n^\mu)^* = \alpha_{-n}^\mu \quad (\tilde{\alpha}_n^\mu)^* = \tilde{\alpha}_{-n}^\mu. \quad (1.38)$$

1.1.4. Cuantización en la norma de cono de luz

Anteriormente, se mencionó que la simetría de Weyl junto con la invariancia conforme, nos permitió reducir la métrica de la hoja de mundo en la acción de Polyakov 1.6 se podía reducir a la métrica plana $h_{\alpha\beta} = \eta_{\alpha\beta}$. La elección de esta norma no todos los grados de libertad de norma se han removido y es posible imponer condiciones de norma adicionales que reducen el número de componentes de X^μ dejando sólo grados de libertad físicos y dinámicos. La libertad de norma faltante proviene de los difeomorfismos generados por la invariancia ante reparametrizaciones de la acción de Polyakov. Empecemos definiendo las coordenadas de cono de luz para una cuerda en D dimensiones

$$X^\pm = \frac{1}{\sqrt{2}}(X^0 \pm X^{D-1}) \quad (1.39)$$

Ya que estamos tratando dos coordenadas espacio-temporales diferentes del resto, hemos perdido invariancia de Lorentz, entonces nuestra simetría $SO(1, D-1)$ se convierte en $SO(D-2)$. Ya hemos fijado la métrica $h_{\alpha\beta} = \eta_{\alpha\beta}$, sin embargo, esto no fija por completo la norma de los difeomorfismos y las simetrías de Weyl. Si consideramos las transformaciones dadas por

$$\delta_D \eta^{\alpha\beta} = -(\partial^\alpha \xi^\beta + \partial^\beta \xi^\alpha) \quad \delta_W \eta^{\alpha\beta} = \Lambda \eta^{\alpha\beta} \quad (1.40)$$

donde ξ^α es un parámetro infinitesimal de reparametrización, Λ es un parámetro infinitesimal del reescalamiento de Weyl, $\delta_D \eta^{\alpha\beta}$ proporciona la variación de la métrica bajo reparametrización y $\delta_W \eta^{\alpha\beta}$ la variación bajo el reescalamiento de Weyl. Si se combinan ambas transformaciones obtenemos que:

$$(\delta_D + \delta_W) \eta^{\alpha\beta} = (-\partial^\alpha \xi^\beta - \partial^\beta \xi^\alpha + \Lambda \eta^{\alpha\beta}) \quad (1.41)$$

Las soluciones de la ecuación anterior igualada a cero nos permitiría encontrar las simetrías adicionales de nuestro sistema. Esta invariancia de norma adicional nos permite tomar la elección trivial para X^+

$$X^+(\tau, \sigma) = x^+ + l^2 p^+ \tau \quad (1.42)$$

donde x^+ y p^+ son constantes.

La coordenada X^- ahora se puede calcular en términos de los grados de libertad transversos $X^i, i = 1, \dots, D - 2$ de la siguiente manera. Las ecuaciones de constricción dadas por 1.27 y 1.28 escritas en términos de las variables de cono de luz se convierten en

$$\frac{\partial X^+}{\partial \tau} \frac{\partial X^-}{\partial \tau} + \frac{\partial X^+}{\partial \sigma} \frac{\partial X^-}{\partial \sigma} = \frac{1}{2} \left(\left(\frac{\partial X^i}{\partial \tau} \right)^2 + \left(\frac{\partial X^i}{\partial \sigma} \right)^2 \right) \quad (1.43)$$

y

$$l^2 p^+ \frac{\partial X^-}{\partial \sigma} = \frac{\partial X^i}{\partial \tau} \frac{\partial X^i}{\partial \sigma} \quad (1.44)$$

con lo cual se puede resolver para obtener X^- en términos de los grados de libertad transversos y una constante de integración x^- . Por lo que en la norma de cono de luz, los únicos grados de libertad son los los grados de libertad transversos X^i .

En el caso de la cuerda cerrada, la ecuación 1.43 se puede escribir como

$$l^2 p^+ T^{-1} p^- = \frac{1}{2} \left(\left(\frac{\partial X^i}{\partial \tau} \right)^2 + \left(\frac{\partial X^i}{\partial \sigma} \right)^2 \right) \quad (1.45)$$

y usando la expansión en modos para los modos transversos a partir de la ecuación 1.37 e integran de desde 0 hasta π respecto a σ nos da

$$\frac{1}{8} M^2 = \frac{1}{8} p^2 = \frac{1}{8} (2p^+ p^- - (p^i)^2) = \frac{1}{2} \sum_{n=1}^{\infty} (\alpha_{-n}^i \alpha_n^i + \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i) - a \quad (1.46)$$

donde hemos impuesto que $l = 1$, o equivalentemente $T = \pi^{-1}$. La constante de orden normal a surge de las relaciones del reordenamiento de algunos términos de la suma usando los operadores de conmutación de la cuantización canónica La constante de ordenamiento sería:

$$a = -\frac{(D-2)}{2} \sum_{n=1}^{\infty} n. \quad (1.47)$$

Regularizando esta expresión divergente usando la función zeta $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ con una extensión analítica en $s = -1$, se tiene que:

$$a = -\frac{(D-2)}{2} \zeta(-1) = \frac{(D-2)}{24}. \quad (1.48)$$

Hagamos un breve paréntesis sobre el último artificio matemático. Como usualmente se hace en Física, cuando nos encontramos con cantidades divergentes en cálculos que se desean ser finitos, entonces se recurren a técnicas de regularización. Dichos procesos se refieren a reemplazar resultados

infinitos con resultados finitos con ayuda de resultados matemáticos. En el cálculo presentado en (1.48) decimos que en realidad estamos calculando la evaluación de la Función Zeta de Riemann en -1 y no la suma de todos los naturales. Las técnicas de regularización no son únicas, y los resultados físicos son independientes de la elección de dichas técnicas. Otra forma de regularizar el resultado anterior, y que es la técnica más utilizada en textos modernos, es imponiendo condiciones a los estados físicos al aplicar los operadores de Virasoro (que resultan ser los modos normales del tensor energía-momento) y dicha técnica también regulariza el valor de la constante de orden normal que acabamos de revisar [11].

Para el caso de la cuerda cerrada, la ecuación 1.44 se puede integrar de 0 a π respecto a σ , y la periodicidad de X^- en el intervalo de 0 a π se puede usar para deducir que

$$\int_0^\pi \frac{\partial X^i}{\partial \tau} \frac{\partial X^i}{\partial \sigma} d\sigma = 0 \quad (1.49)$$

y utilizando la expansión en modos de la cuerda cerrada 1.37 obtenemos la siguiente condición

$$\sum_{n=1}^{\infty} \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i = \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i. \quad (1.50)$$

Combinando las ecuaciones 1.46 y 1.50 podemos escribir entonces

$$M^2 = M_R^2 + M_L^2 \quad (1.51)$$

donde

$$\frac{1}{4}M_R^2 = \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i - a \quad \frac{1}{4}M_L^2 = \sum_{n=1}^{\infty} \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i - a \quad M_R^2 = M_L^2. \quad (1.52)$$

Un tratamiento similar se puede seguir para el caso de la cuerda abierta. Ahora bien, los estados de cuerda bosónica se pueden construir en la norma de cono de luz actuando sobre el producto de osciladores α_{-n}^i y $\tilde{\alpha}_{-n}^i$, $i = 1, \dots, D - 2$ sobre los estado base $|0\rangle_L$ y $|0\rangle_R$ para los osciladores izquierdos y derechos, respectivamente, y formando el producto de los estados derechos e izquierdos obtenidos. Para calcular las masas de estos estados, necesitamos evaluar la constante de orden normal que aparece en 1.52. Esta constante se puede determinar construyendo los generadores de Lorentz para la teoría de cuerdas D -dimensional, usando las corrientes de Noether para el grupo de Lorentz e imponiendo reglas de conmutación de los generadores de Lorentz estén libres de anomalías asegurando la invariancia de Lorentz.[10] La nulidad de dichas anomalías requiere que

$$D = 26 \quad a = 1 \quad (1.53)$$

Notemos que se puede definir el operador de número para los osciladores derechos e izquierdos

$$N = \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i - a \quad \tilde{N} = \sum_{n=1}^{\infty} \tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i - a \quad (1.54)$$

y cuentan el número de osciladores derechos e izquierdos hay. La última condición en 1.52 nos impone que $N = \tilde{N}$. Notemos que entonces la masa en términos de estos operadores serían

$$M^2 = 4(N - 1) = 4(\tilde{N} - 1) \quad (1.55)$$

Con esto en mente se puede empezar a generar el espectro de la cuerda bosónica:

- El estado base $|0\rangle_R |0\rangle_L$ que tiene masa $M^2 = -4$ que es un taquión
- Para el nivel $N = 1$ tenemos un conjunto de $24^2 = 576$ estados de la forma

$$|\Omega^{ij}\rangle = \alpha_{-1}^i |0\rangle_R \tilde{\alpha}_{-1}^j |0\rangle_L \quad (1.56)$$

que corresponde al producto tensorial de dos vectores no masivos. La parte de $|\Omega^{ij}\rangle$ que es simétrica y sin traza en i y j transforma bajo $SO(24)$ como una partícula de espín 2, el gravitón. El término de la traza $\delta_{ij} |\Omega^{ij}\rangle$ es un escalar no masivo que llamaremos dilatón, y la parte antisimétrica $|\Omega^{ij}\rangle = -|\Omega^{ji}\rangle$ transforma bajo $SO(24)$ como un tensor de rango 2 antisimétrico.

Los estados de cuerda cerrada siempre son multipletes de $SO(24)$ o $SO(25)$, dependiendo de si el estado es masivo o no. Esto se debe a que para un estado masivo podemos cambiar de sistema de referencia de tal manera que en dicho sistema el estado tome la forma

$$\left| E, E, \underbrace{0, 0, \dots, 0}_{25 \text{ veces}} \right\rangle, \quad (1.57)$$

Ahora, el grupo que deja invariantes todas las rotaciones de 25 dimensiones, el grupo pequeño, está dado por $SO(25)$, y entonces el estado masivo corresponde a alguna representación del grupo de rotación $SO(25)$. Mientras que para los estados no masivos podemos hacer una transformación de Lorentz a un sistema de referencia en el cual dicho estado tome la forma

$$\left| E, E, \underbrace{0, 0, \dots, 0}_{24 \text{ veces}} \right\rangle \quad (1.58)$$

entonces los estados no masivos tienen al grupo pequeño $SO(24)$, entonces dar un estado no masivo es dar una representación de $SO(24)$.

El problema de la aparición de un estado taquiónico se tratará de solventar en la siguiente sección con la teoría de supercuerdas. Físicamente, todos los estados representan a una cuerda cerrada con menor o mayor grado de excitación, dependiendo del número de operadores de creación que se apliquen sobre el vacío. Precisamente, en el espectro de cuerdas (cerradas o abiertas) obtenemos campos escalares como resultado de la excitación del vacío. En el caso de las cuerdas cerradas,

se ve que éstas son pequeñas fluctuaciones de una estructura (el campo de cuerdas cerradas) que incluye al propio espacio-tiempo 26D, junto con un número infinito de otros campos. El estado base taquiónico en nuestro espectro, que llamaremos $\phi(x)$, tiene masa cuadrada negativa. Recordemos que para un campo arbitrario $\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)^2 - V(\phi)$ con $V(\phi) = V(0) + V'(0)\phi + \frac{1}{2}V''(0)\phi^2 + \dots$ donde $V'(0) = 0$ y $V''(0) \equiv m^2$, por lo que tenemos un mínimo local en 0. Así que si el campo está asociado a un estado taquiónico, tenemos entonces que $m^2 < 0$, como en el caso del modo más bajo de la cuerda cerrada y por ende $V''(0) < 0$. Es decir, estamos perturbando alrededor de un extremo potencial que es un máximo, en lugar de un mínimo. La existencia del taquión nos informa entonces que el vacío $\langle\phi\rangle = 0$, alrededor del cual estamos perturbando, es inestable: el sistema tiende a cambiar de estado para adquirir $\langle\phi\rangle \neq 0$ (porque es energéticamente favorable), este proceso es conocido como *condensación del taquión* [13]. Las condiciones de Lorentz para las coordenadas X^μ se conservan porque en el estado base taquiónico aún no perturbamos al estado base con los modos asociados a dichas coordenadas.

1.2. Supercuerdas

La teoría de cuerdas bosónica presentada en la sección 1.1 no es completamente satisfactoria como una teoría completa, primero porque no contiene estados fermiónicos y segundo porque ya vimos que su estado base es un estado taquiónico. Estas dos dificultades se pueden solventar construyendo una teoría en la cual cada grado de libertad bosónico $X^\mu(\tau, \sigma)$, $\mu = 0, \dots, D - 1$, exista un grado de libertad fermiónico $\Psi^\mu(\tau, \sigma)$ de la hoja de mundo descrito por un espinor de Majorana con dos componentes.

1.2.1. La acción de teoría de supercuerdas

Requerimos de una acción que acople pares supersimétricos de los campos bosónicos y fermiónicos $X^\mu(\tau, \sigma)$ y $\Psi^\mu(\tau, \sigma)$ en una supergravedad 2-dimensional. Los espinores de $\Psi^\mu(\tau, \sigma)$ tiene dos componentes que se pueden escribir como

$$\Psi^\mu(\tau, \sigma) = \begin{pmatrix} \psi_-^\mu(\tau, \sigma) \\ \psi_+^\mu(\tau, \sigma) \end{pmatrix}, \quad (1.59)$$

y llamamos $\psi_A^\mu(\tau, \sigma)$ componentes quirales del espinor, donde, por ejemplo, $\psi_+^\mu(\tau, \sigma)$ es la componente de quiralidad positiva del espinor $\Psi^\mu(\tau, \sigma)$. Una teoría resultante del emparejamiento de los grados de libertad bosónicos de cuerdas y fermiones por medio de supersimetrías se llaman *teorías de supercuerdas*.

Para incorporar supersimetría en la teoría de cuerdas se pueden desarrollar dos caminos:

- **Formalismo de Ramond-Neveu-Schwarz (RNS):** en el cual la teoría es supersimétrica sobre la hoja de mundo.

- **Formalismo de Green-Schwarz (GS):** en el cual la teoría es supersimétrica en el espacio-tiempo de Mincowski de fondo 10-dimensional. Este formalismo se puede generalizar a otras geometrías para el espacio-tiempo de fondo.

Aquí nos concentraremos en el formalismo RNS, entonces a nuestra teoría bosónica en D dimensiones le agregaremos D campos fermiónicos libres $\Psi(\tau, \sigma)$. Estos campos son espinores de dos componentes que describen a fermiones viviendo en la hoja de mundo y que transforman como vectores bajo transformaciones de Lorentz en el espacio tiempo de fondo D -dimensional. Para incorporar estos campos fermiónicos en nuestra teoría se modifica la acción de la cuerda bosónica. La nueva acción ahora cuenta con la acción bosónica S_B , la acción de Polyakov, y la adición de la acción de Dirac para D fermiones no masivos libres, S_F , es decir,

$$\begin{aligned} S &= S_B + S_F \\ &= -\frac{1}{2\pi} \int d\tau d\sigma \partial_\alpha X^\mu \partial^\alpha X_\mu - \frac{1}{2\pi} \int d\tau d\sigma \bar{\Psi}^\mu \rho^\alpha \partial_\alpha \Psi_\mu, \end{aligned} \quad (1.60)$$

donde ρ^α , con $\alpha = 0, 1$, es la representación de las matrices de Dirac en dos dimensiones,

$$\rho^0 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \rho^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad (1.61)$$

y que satisfacen las relaciones de anticonmutación,

$$\{\rho^\alpha, \rho^\beta\} = 2\eta^{\alpha\beta} I. \quad (1.62)$$

Además, $\bar{\Psi}^\mu$ es el conjugado de Dirac para Ψ_μ , que está definido como

$$\bar{\Psi}^\mu = (\Psi^\mu)^\dagger i\rho^0, \quad (1.63)$$

con A^\dagger el hermitiano conjugado de A . De (1.61) podemos apreciar que las matrices de Dirac tienen componentes reales, y a esta representación de las mismas las llamaremos de Majorana. También, en dicha representación podemos imponer una condición en los espinores que los obliga a tener componentes reales,

$$(\Psi^\mu)^T C = (\Psi^\mu)^\dagger i\rho^0, \quad (1.64)$$

donde, en nuestro caso, C es una matrix de 2×2 llamada matrix de conjugación de carga. La condición (1.64) y sabiendo que la matrix de conjugación de carga está dada por $C = i\rho^0$ se puede demostrar que las componentes de los espinores en la representación de Majorana cumplen que

$$\psi_\pm^\mu = (\psi_\pm^\mu)^*, \quad (1.65)$$

donde $(A)^*$ es el conjugado complejo de A .

Ahora, usando explícitamente la forma de las matrices de Dirac (1.61) e introduciéndolas explícitamente en la acción (1.60) podemos escribir a la acción en términos de las coordenadas de los espinores de Majorana. En coordenadas de cono de luz obtenemos que la acción se escribe como

$$S = \frac{1}{\pi} \int d\sigma^+ d\sigma^- \partial_+ X^\mu(\sigma^-, \sigma^+) \partial_- X_\mu(\sigma^-, \sigma^+) + \frac{i}{2\pi} \int d\sigma^+ d\sigma^- (\psi_-^\mu(\sigma^-, \sigma^+) \partial_+ \psi_{-\mu}(\sigma^-, \sigma^+) + \psi_+^\mu(\sigma^-, \sigma^+) \partial_- \psi_{+\mu}(\sigma^-, \sigma^+)). \quad (1.66)$$

De la parte fermiónica de la acción vemos que las ecuaciones de movimiento para las dos componentes de los espinores son las ecuaciones de Dirac, que en términos de las coordenadas de cono de luz están dadas por

$$\partial_+ \psi_-^\mu = 0 \quad \text{y} \quad \partial_- \psi_+^\mu = 0. \quad (1.67)$$

Notemos que la primera ecuación describe al oscilador izquierdo, mientras que la segunda ecuación describe al oscilador derecho.

La acción S_0 posee invariancia bajo reparametrizaciones y supersimetría global en la hoja de mundo bajo la transformación supersimétrica

$$\delta X^\mu = \bar{\epsilon} \Psi^\mu \quad \delta \Psi^\mu = \rho^\alpha \partial_\alpha X^\mu \epsilon, \quad (1.68)$$

donde ϵ es un espinor constante de Majorana, es decir, de la forma

$$\epsilon = \begin{pmatrix} \epsilon_- \\ \epsilon_+ \end{pmatrix}. \quad (1.69)$$

Bajo esta transformación supersimétrica, la variación de la acción (1.60) es cero, $\delta S = 0$. Por lo tanto, tenemos una teoría de supercuerdas RNS ya que nuestra teoría tiene supersimetría en la hoja de mundo (ver [11]).

Notemos que la supersimetría que aquí introducimos es global, pues no depende de las coordenadas de la hoja de mundo. Lo que estamos describiendo aquí es una teoría con una norma fijada, lo que se llama *norma covariante*. En dicha norma explotamos la invariancia ante reparametrizaciones e invariancia conforme para escribir la métrica de la hoja de mundo en la forma

$$h_{\alpha\beta} = \eta_{\alpha\beta}, \quad (1.70)$$

donde la acción antes de fijar la norma es

$$S_0 = -\frac{1}{2\pi} \int d\tau d\sigma \sqrt{-\det h} (h^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X_\mu + \bar{\Psi}^\mu \rho^\alpha \partial_\alpha \Psi_\mu), \quad (1.71)$$

que también es invariante ante las transformaciones supersimétricas en (??). Sin embargo, la acción S_0 no es invariante si reemplazamos ϵ por una variable local $\epsilon(\tau, \sigma)$. La variación de la acción S_0 no se anularía en este caso y tendríamos entonces [10],

$$\delta S_0 \sim \int d\tau d\sigma (\partial_\alpha \bar{\epsilon}) \left(-\frac{1}{2} \rho^\beta \rho^\alpha \Psi_\mu \partial_\beta X^\mu \right). \quad (1.72)$$

Y podemos identificar entonces la supercorriente de mundo por (escrita con índices espinoriales)

$$J_A^\alpha = -\frac{1}{2} \left(\rho^\beta \rho^\alpha \Psi_\mu \right)_A \partial_\beta X^\mu. \quad (1.73)$$

Para cancelar esta variación de S_0 necesitamos introducir un supergravitino 2-dimensional ξ_α con la transformación

$$\delta \xi_\alpha = -\partial_\alpha \epsilon. \quad (1.74)$$

Entonces δS_0 es cancelada agregando a la acción el término

$$S_1 = -\frac{1}{\pi} \int d\tau d\sigma \sqrt{-\det h} \bar{\xi}_\alpha \rho^\beta \rho^\alpha \Psi^\mu \partial_\beta X_\mu. \quad (1.75)$$

Sin embargo, la transformación de X_μ proporciona un término adicional en la variación δS_1 ,

$$\begin{aligned} \delta S_1 &= -\frac{1}{\pi} \int d\tau d\sigma \sqrt{-\det h} \bar{\xi}_\alpha \rho^\beta \rho^\alpha \Psi^\mu \bar{\Psi}_\mu \partial_\beta \epsilon + \dots \\ &= \frac{1}{2\pi} \int d\tau d\sigma \sqrt{-\det h} \bar{\Psi}_\mu \Psi^\mu \bar{\xi}_\alpha \rho^\beta \rho^\alpha \partial_\beta \epsilon + \dots, \end{aligned} \quad (1.76)$$

y esta variación es cancelada agregando a la acción de cuerda supersimétrica el término

$$S_2 = -\frac{1}{4\pi} \int d\tau d\sigma \sqrt{-\det h} \bar{\Psi}_\mu \Psi^\mu \bar{x}^\alpha \rho^\beta \rho^\alpha \xi_\beta. \quad (1.77)$$

Por lo tanto, para tener una acción de supercuerdas invariante ante transformaciones supersimétricas locales, a S_0 debemos sumarle S_1 y S_2 , es decir,

$$S = S_0 + S_1 + S_2, \quad (1.78)$$

que sí es invariante bajo transformaciones locales

$$\delta X^\mu = \bar{\epsilon} \Psi^\mu \quad \delta \Psi^\mu = \rho^\alpha \partial_\alpha X^\mu \epsilon \quad \delta \xi_\alpha = -\partial_\alpha \epsilon \quad (1.79)$$

y

$$\delta e_\alpha^a = -2i \bar{\epsilon} \rho^a \xi_\alpha, \quad (1.80)$$

donde e_α^a es el *zweibein* que satisface $h_{\alpha\beta} = e_\alpha^a e_\beta^b \eta_{ab}$.

Ahora deseamos escribir la supercorriente de la hoja de mundo en términos de las coordenadas de cono de luz. Recordemos que para un espinor constante de Majorana ϵ sus dos componentes son ϵ_+ y ϵ_- como en (1.69) Haciendo un cambio de variable sobre la integral (1.72) se puede demostrar que la supercorriente asociada a ϵ_- se puede escribir como (ver [11])

$$j_+ \equiv \psi_+^\mu \partial_+ X_\mu. \quad (1.81)$$

Similarmente, la supercorriente asociada a ϵ_+ está dada por

$$j_- \equiv \psi_-^\mu \partial_- X_\mu. \quad (1.82)$$

Estas corrientes cumplirán que

$$\partial_- j_+ = \partial_+ j_- = 0, \quad (1.83)$$

que combinados nos indican que

$$\partial_\alpha J_A^\alpha = 0, \quad (1.84)$$

o que la supercorriente de la hoja de mundo, en efecto, se conserva.

La siguiente corriente en nuestra teoría es la correspondiente a la simetría traslacional de la acción RNS (1.60). Esta corriente es el tensor de energía-momento que está dado por (ver [11])

$$T_{\alpha\beta} = \partial_\alpha X^\mu \partial_\beta X_\mu + \frac{1}{4} \bar{\Psi}^\mu \rho_\alpha \partial_\beta \Psi_\mu + \frac{1}{4} \bar{\Psi}^\mu \rho_\beta \partial_\alpha \Psi_\mu - (\text{traza}). \quad (1.85)$$

Podemos escribir las componentes de este tensor en términos de las coordenadas de cono de luz y las componentes espinoriales como

$$T_{++} = \partial_+ X_\mu \partial_+ X^\mu + \frac{i}{2} \psi_+^\mu \partial_+ \psi_{+\mu} \quad (1.86)$$

$$T_{--} = \partial_- X_\mu \partial_- X^\mu + \frac{i}{2} \psi_-^\mu \partial_- \psi_{-\mu} \quad (1.87)$$

$$T_{-+} = T_{+-} = 0 \quad (1.88)$$

1.2.2. Condiciones de frontera y expansión en modos

Retomemos la acción de la teoría de supercuerdas ya con la norma fijada en la métrica de la hoja de mundo plana y en términos de las coordenadas de cono de luz (1.66), pero ahora escribamos el factor diferencial de la integral en términos de las coordenadas iniciales $d\tau d\sigma$ en lugar de $d\sigma^+ d\sigma^-$, pero escribiendo los integrandos en términos de las coordenadas de cono de luz. Tenemos entonces que la acción se escribe como

$$S = \frac{1}{\pi} \int d\tau d\sigma \partial_+ X^\mu(\sigma^-, \sigma^+) \partial_- X_\mu(\sigma^-, \sigma^+) + \frac{i}{2\pi} \int d\tau d\sigma (\psi_-^\mu(\sigma^-, \sigma^+) \partial_+ \psi_{-\mu}(\sigma^-, \sigma^+) + \psi_+^\mu(\sigma^-, \sigma^+) \partial_- \psi_{+\mu}(\sigma^-, \sigma^+)). \quad (1.89)$$

Las condiciones de frontera y las ecuaciones de campo para los grados de libertad bosónicos X^μ son las mismas que ya vimos en la sección 1.1, por lo que nos concentraremos sólo en la parte fermiónica de esta acción.

Para la acción fermiónica tenemos la siguiente acción después de suprimir los índices de Lorentz

$$S_F \sim \int d\sigma^2 (\psi_- \partial_+ \psi_- + \psi_+ \partial_- \psi_+), \quad (1.90)$$

donde $d\sigma^2 = d\tau d\sigma$. Calculando la variación de esta acción respecto a los campos fermiónicos tenemos

$$\delta S_F \sim \int d\tau (\psi_- \delta \psi_- - \psi_+ \delta \psi_+) \Big|_{\sigma=\pi} - \int d\tau (\psi_- \delta \psi_- - \psi_+ \delta \psi_+) \Big|_{\sigma=0}. \quad (1.91)$$

Ahora, como en el caso de la cuerda bosónica, deseamos que los términos de frontera se anulen para así tener una variación de la acción nula. Esta imposición a los términos de frontera nos llevan a las supercuerdas RNS abiertas y cerradas. Como ya dijimos anteriormente, en este trabajo nos concentraremos únicamente en las cuerdas cerradas, por lo que a continuación exponemos el caso que nos atañe.

Cuerdas cerradas RNS

Como en el caso de la cuerda bosónica, las condiciones de frontera para la cuerda cerrada da lugar a dos clases de modos fermiónicos desacoplados, llamados de igual manera, osciladores derechos e izquierdos. Existen dos condiciones periódicas posibles que anulan a los términos de frontera en (1.91),

$$\psi_{\pm}^{\mu}(\tau, \sigma) = \pm \psi_{\pm}^{\mu}(\tau, \sigma + \pi), \quad (1.92)$$

donde el signo positivo describe condiciones de frontera periódicas (condiciones de Ramond o R) mientras que el signo negativo describe condiciones de frontera antiperiódicas (condiciones de Neveu-Schwarz o NS). Esto anula los términos de frontera como deseamos pues si, por ejemplo, tomamos $\psi_{\pm}(\tau, \sigma) = -\psi_{\pm}(\tau, \sigma + \pi)$ se tiene que

$$\delta S_F \sim \int d\tau (\psi_- \delta \psi_- - \psi_+ \delta \psi_+) \Big|_{\sigma=\pi} - \int d\tau (\psi_- \delta \psi_- - \psi_+ \delta \psi_+) \Big|_{\sigma=\pi} = 0. \quad (1.93)$$

Ambos términos se hacen iguales y por ende se anulan con la diferencia. Debemos notar que podemos imponer condiciones R o NS a los osciladores derechos e izquierdos por separado. Esto nos conduce a dos posibles elecciones de la expansión en modos para los osciladores izquierdos

$$\psi_{+}^{\mu}(\tau, \sigma) = \sum_{n \in \mathbb{Z}} \tilde{d}_n^{\mu} e^{-2in(\tau+\sigma)} \quad \text{R} \quad \text{ó} \quad \psi_{+}^{\mu}(\tau, \sigma) = \sum_{r \in \mathbb{Z}+1/2} \tilde{b}_r^{\mu} e^{-2ir(\tau+\sigma)} \quad \text{NS}, \quad (1.94)$$

mientras que para los osciladores derechos tenemos las siguientes dos elecciones

$$\psi_{-}^{\mu}(\tau, \sigma) = \sum_{n \in \mathbb{Z}} d_n^{\mu} e^{-2in(\tau-\sigma)} \quad \text{R} \quad \text{ó} \quad \psi_{-}^{\mu}(\tau, \sigma) = \sum_{r \in \mathbb{Z}+1/2} b_r^{\mu} e^{-2ir(\tau-\sigma)} \quad \text{NS}, \quad (1.95)$$

El verdadero estado físico estará dado al hacer el producto tensorial de un oscilador derecho con un oscilador izquierdo, y como hay dos elecciones para cada uno de ellos, tenemos en total 4 diferentes sectores; el sector R-R, el sector R-NS, el sector NS-R y finalmente el sector NS-NS. Ya con las expresiones de los grados de libertad bosónicos y fermiónicos podemos cuantizar nuestra teoría de supercuerdas RNS.

1.2.3. Cuantización canónica de la teoría de supercuerdas RNS

Para cuantizar la teoría promoveremos los modos α y $\tilde{\alpha}$, que vienen de los campos bosónicos, y los modos b, \tilde{b}, d y \tilde{d} , que vienen de los campos fermiónicos, a operadores y además introducimos las siguientes relaciones de conmutación y anticonmutación para estos operadores

$$[\alpha_n^\mu, \alpha_n^\nu] = [\tilde{\alpha}_m^\mu, \tilde{\alpha}_n^\nu] = m\delta_{m,-n}\eta^{\mu\nu}, \quad (1.96)$$

$$\{b_r^\mu, b_s^\nu\} = \{\tilde{b}_r^\mu, \tilde{b}_s^\nu\} = \delta_{r,-s}\eta^{\mu\nu}, \quad (1.97)$$

$$\{d_m^\mu, d_n^\nu\} = \{\tilde{d}_m^\mu, \tilde{d}_n^\nu\} = \delta_{m,-n}\eta^{\mu\nu}, \quad (1.98)$$

con el resto de las relaciones nulas. El hecho de que aparezca la métrica espacio-temporal $\eta^{\mu\nu}$ en las relaciones de conmutación y anticonmutación nos indica que las componentes temporales de los osciladores bosónicos y fermiónicos dan origen a estados fantasma (estados con norma negativa). Estos estados fantasma serán removidos de la teoría con las constricciones de súper-Virasoro. Ahora veamos cuáles son los estados de los estados base para nuestra teoría RNS.

Tenemos 4 sectores para nuestras supercuerdas cerradas RNS y en cada uno de ellos los estados están conformados por el producto tensorial de dos estados, uno generado por operadores asociados a osciladores derechos y otro por operadores asociados a osciladores izquierdos. Así, tendremos entonces 4 estados base para cada oscilador derecho e izquierdo, dos para el sector Ramond (R) y dos para el sector Neveu-Schwarz (NS). Para el sector Ramond, los estados base están definidos como

$$\alpha_m^\mu |0\rangle_R = d_m^\mu |0\rangle_R = 0, \quad \text{para } m > 0, \quad \text{para los osciladores derechos,} \quad (1.99)$$

$$\tilde{\alpha}_m^\mu |0\rangle_R = \tilde{d}_m^\mu |0\rangle_R = 0, \quad \text{para } m > 0, \quad \text{para los osciladores izquierdos.} \quad (1.100)$$

Mientras que para el sector Neveu-Schwarz tenemos los estados base

$$\alpha_m^\mu |0\rangle_{NS} = b_r^\mu |0\rangle_{NS} = 0 \quad \text{para } m, r > 0, \quad \text{para los osciladores derechos,} \quad (1.101)$$

$$\tilde{\alpha}_m^\mu |0\rangle_{NS} = \tilde{b}_r^\mu |0\rangle_{NS} = 0 \quad \text{para } m, r > 0, \quad \text{para los osciladores izquierdos.} \quad (1.102)$$

Los estados base completos para nuestras supercuerdas cerradas serán los productos tensoriales de todos estos estados base izquierdos con derechos, salvo algunas restricciones importantes que mencionaremos más adelante. Por el momento, estudiemos la naturaleza de estos estados base.

1.2.4. Estados base del sector Ramond vs. estados base del sector Neveu-Schwarz

Los estados base del sector NS son únicos y corresponden a estados de espín 0, es decir, son bosones en el espacio-tiempo de fondo. Ahora bien, como todos los osciladores (α_n^μ , $\tilde{\alpha}_n^\mu$, b_r^μ y \tilde{b}_r^μ) transforman como vectores bajo transformaciones de Lorentz, los estados excitados del sector NS (que son generados por medio de la acción de operadores de creación sobre el estado base) también corresponden a bosones espacio-temporales. Además actuar sobre estos estados con operadores de creación va incrementando la masa como vimos en el caso de la cuerda bosónica.

El estado base en el sector R es degenerado. Para ver esto notemos que los operadores d_0^μ pueden actuar sobre el estado sin alterar su masa pues conmuta con el operador de número N que, veremos más adelante, está definido por (por ahora nos concentraremos sólo en los osciladores derechos)

$$N = \sum_{n=1}^{\infty} \alpha_n \cdot \alpha_n + \sum_{r=1/2}^{\infty} r b_r \cdot b_r, \quad (1.103)$$

cuyos eigenvalores nos ayudan a calcular la masa cuadrada de los estados. Ahora, por las relaciones de conmutación y anticonmutación que impusimos en la cuantización canónica, sabemos que d_0 obedece las mismas relaciones algebraicas del álgebra de Clifford, salvo un factor 2,

$$\{d_0^\mu, d_0^\nu\} = \delta_{0,0} \eta^{\mu\nu} = \eta^{\mu\nu}. \quad (1.104)$$

Recordemos rápidamente que el álgebra de Clifford está conformada por operadores Γ^μ que cumplen las reglas de anticonmutación

$$\{\Gamma^\mu, \Gamma^\nu\} = 2\eta^{\mu\nu}. \quad (1.105)$$

Por lo tanto, como el álgebra de Dirac es isomorfa al álgebra de Clifford, esto implica que el conjunto de estados base degenerados en el sector R debe satisfacer una representación del álgebra de Dirac. Esto implica que existe un conjunto de estados base degenerados que pueden ser escritos en la forma $|a\rangle$ con a un índice espinorial, tal que

$$d_0^\mu |a\rangle = \frac{1}{\sqrt{2}} \Gamma_{ba}^\mu |b\rangle, \quad (1.106)$$

donde Γ^μ es una matriz de dimensión a , en la representación de d_0^μ , es decir, una matriz de Dirac. Esta expresión define cómo el oscilador actúa sobre el espinor, esto es, proporciona una representación para d_0^μ en el espacio espinorial. Por ejemplo, si tuviéramos dos valores posibles para a , digamos $+$ y $-$, entonces Γ^μ será una matriz de 2×2 y la ecuación anterior podría expresarse como

$$d_0^\mu \begin{pmatrix} |+\rangle \\ |-\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Gamma_{++}^\mu & \Gamma_{+-}^\mu \\ \Gamma_{-+}^\mu & \Gamma_{--}^\mu \end{pmatrix} \begin{pmatrix} |+\rangle \\ |-\rangle \end{pmatrix}. \quad (1.107)$$

Como los operadores (α_n^μ , $\tilde{\alpha}_n^\mu$, d_n^μ y \tilde{d}_n^μ) transforman como vectores en el espacio-tiempo, y como cada estado en el sector R se puede obtener con operadores osciladores de creación (o negativos) sobre los estados base $|0\rangle_R$, vemos que todos los estados en el sector R son fermiones espacio-temporales.

1.2.5. Cuantización en la norma de cono de luz

Ya vimos en que para el caso de la cuerda bosónica (ver sección 1.1) que aunque habíamos fijado las normas en las cuales la métrica de la hoja de mundo era plana, todavía teníamos una simetría residual. Esta simetría hizo posible imponer la condición de la norma del cono de luz, la cual hace que

$$X^+(\tau, \sigma) = x^+ + p^+ \tau. \quad (1.108)$$

Esta condición también es cierta para la teoría de supercuerdas RNS. Sin embargo, ahora tenemos también una simetría fermiónica residual. Esta simetría fermiónica nos permite fijar más condiciones en nuestra teoría RNS. Usando esta simetría podemos fijar que en el sector NS

$$\Psi^+(\tau, \sigma) = 0, \quad (1.109)$$

mientras que en el sector R debemos mantener el modo nulo en la expansión de Ψ^+ , que es una matriz de Dirac. Por lo que en la teoría RNS podemos elegir las dos normas de cono de luz

$$X^+(\tau, \sigma) = x^+ + p^+ \tau \quad \text{y} \quad \Psi^+(\tau, \sigma) = 0, \quad (1.110)$$

donde

$$\Psi^\pm \equiv \frac{1}{\sqrt{2}}(\Psi^0 \pm \Psi^{D-1}). \quad (1.111)$$

En esta norma de cono de luz, las coordenadas X^- y Ψ^- no son grados de libertad independientes. Estas coordenadas pueden ser calculadas en términos de los grados de libertad transversos X^i y Ψ^- , $i = 1, \dots, D - 2$.

1.2.6. Proyección GSO

El operador de masa cuadrada para la cuerda cerrada se deduce de manera análoga a como se hizo con la cuerda bosónica en la sección 1.1 y se puede escribir de la manera

$$M^2 = M_R^2 + M_L^2, \quad (1.112)$$

pues los osciladores izquierdos y derechos contribuyen a la masa total de manera independiente para los estados completos. A diferencia de lo que vimos en la cuerda bosónica, ahora también debemos incluir los osciladores fermiónicos en la contribución del operador de masa. El operador de masa para los estados derechos está dado por

$$\frac{1}{4}M_R^2 = \begin{cases} \sum_{n=1}^{\infty} \alpha_{-n}^\mu \alpha_{\mu n} + \sum_{r=1/2}^{\infty} r b_{-r}^\mu d_{\mu r} - a_{NS} & \text{NS} \\ \sum_{n=1}^{\infty} \alpha_{-n}^\mu \alpha_{\mu n} + \sum_{n=1}^{\infty} n d_{-n}^\mu d_{\mu n} - a_R & \text{R} \end{cases} \quad (1.113)$$

mientras que para los estados izquierdos es

$$\frac{1}{4}M_L^2 = \begin{cases} \sum_{n=1}^{\infty} \tilde{\alpha}_{-n}^{\mu} \tilde{\alpha}_{\mu n} + \sum_{r=1/2}^{\infty} r \tilde{b}_{-r}^{\mu} \tilde{d}_{\mu r} - a_{NS} & \text{NS} \\ \sum_{n=1}^{\infty} \tilde{\alpha}_{-n}^{\mu} \tilde{\alpha}_{\mu n} + \sum_{n=1}^{\infty} n \tilde{d}_{-n}^{\mu} \tilde{d}_{\mu n} - a_R & \text{R} \end{cases} \quad (1.114)$$

donde a_R y a_{NS} son constantes derivadas del orden normal al hacer la cuantización de los modos. En el sector Ramond se cumple la condición de emparejamiento, en la cual

$$M_L^2 = M_R^2. \quad (1.115)$$

Además, en este sector, la constante de orden normal es nula pues se cancelan exactamente las contribuciones bosónicas y fermiónicas. Por tanto,

$$a_R = 0. \quad (1.116)$$

Como en el caso de la cuerda bosónica, el valor de las constantes de orden normal se determina imponiendo que los conmutadores de los generadores de Lorentz en la norma de cono de luz estén libres de anomalías cuánticas (ver [10]). Con esto en cuenta se puede demostrar que la constante de orden normal para el sector NS es

$$a_{NS} = \frac{1}{2}. \quad (1.117)$$

Analicemos ahora el estado base de los estados derechos para el sector NS. En este caso, el operador de masa correspondiente está dado por la primera ecuación en (1.113). El estado base de los estados derechos del sector NS es aniquilado por todos los modos osciladores positivos, es decir,

$$\alpha_n^i |0; k^{\mu}\rangle_{NS} = b_r^i |0; k^{\mu}\rangle_{NS} = 0 \quad (n, r > 0), \quad (1.118)$$

y tal que

$$\alpha_0^{\mu} |0; k^{\mu}\rangle_{NS} = \sqrt{2\alpha'} k^{\mu} |0; k^{\mu}\rangle_{NS}, \quad (1.119)$$

donde $\sqrt{2\alpha'}$ viene de la normalización del estado. Ahora calculemos la masa de dicho estado con el operador ya mencionado, tenemos

$$\begin{aligned} \frac{1}{4}M^2 |0; k^{\mu}\rangle_{NS} &= \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i |0; k^{\mu}\rangle_{NS} + \sum_{r=1/2}^{\infty} r b_{-r}^i b_r^i |0; k^{\mu}\rangle_{NS} - \frac{1}{2} |0; k^{\mu}\rangle_{NS} \\ &= -\frac{1}{2} |0; k^{\mu}\rangle_{NS}, \end{aligned} \quad (1.120)$$

que quiere decir que el estado base de los estados derechos en el sector NS tiene una masa dada por $\frac{1}{4}M^2 = -\frac{1}{2}$, lo que llamamos un estado taquiónico o un estado con masa imaginaria.

La última observación supone un problema para nuestra teoría RNS y queremos construir un mecanismo que saque de nuestra teoría a estos estados taquiónicos. Otro inconveniente que

encontramos en la teoría es que el espectro no es supersimétrico espacio-temporal. Por ejemplo, en el espectro no hay un fermión con la misma masa que el taquión. Se necesita una supersimetría no rota para una teoría de interacción consistente, ya que el espectro contiene un gravitino no masivo, que es el campo de norma para la supersimetría local. Para quitar estas inconsistencias de nuestra teoría RNS, proyectaremos el espectro en una forma muy específica que elimine el taquión y que lleve a una teoría supersimétrica en el espacio-tiempo 10-dimensional. Esta proyección se llama *proyección GSO*, pues fue introducida por Gliozzi, Scherk y Olive.

Para construir esta proyección primero definimos un operador que cuente el número de osciladores b de un estado en el sector NS, que está dado por

$$F_{NS} = \sum_{r=1/2}^{\infty} b_{-r}^i b_r^i, \quad (1.121)$$

junto con el operador

$$F_R = \sum_{r=1}^{\infty} d_{-n}^i d_n^i, \quad (1.122)$$

que cuenta el número de osciladores d de un estado en el sector R.

Ahora definimos el operador de G-paridad, que en el sector NS está dado por

$$G = (-1)^{F_{NS}+1} = (-1)^{\sum_{r=1/2}^{\infty} b_{-r}^i b_r^i + 1}, \quad (1.123)$$

mientras que en el sector R está dado por

$$G = \Gamma^{11} (-1)^{F_R} = \Gamma^{11} (-1)^{\sum_{r=1}^{\infty} d_{-n}^i d_n^i}. \quad (1.124)$$

El operador Γ^{11} está definido por

$$\Gamma^{11} = \Gamma^0 \Gamma^1 \dots \Gamma^{10} \quad (1.125)$$

es la versión análoga 10-dimensional de la matriz de Dirac γ^5 en 4 dimensiones y cuyos eigenvalores nos proporcionan la quiralidad de los espinores. Los espinores Ψ^μ satisfacen

$$\Gamma_{11} \Psi^\mu = \Psi^\mu, \quad (1.126)$$

se dicen de quiralidad positiva, mientras que los espinores que satisfacen

$$\Gamma_{11} \Psi^\mu = -\Psi^\mu, \quad (1.127)$$

son llamados de quiralidad negativa. Si un espinor dado tiene una quiralidad definida entonces decimos que es un espinor de Weyl.

Con estas herramientas ya podemos construir la proyección GSO. En el sector NS mantenemos sólo los estados que tengan una G-paridad positiva, o equivalentemente, la proyección GSO elimina del espectro a los estados con G-paridad negativa. Conservamos únicamente estados $|\Omega\rangle$ tales que

$$G|\Omega\rangle = (-1)^{F_{NS}+1}|\Omega\rangle = |\Omega\rangle. \quad (1.128)$$

Esto quiere decir que requerimos

$$1 = (-1)^{F_{NS}+1} = (-1)^{F_{NS}}(-1), \quad (1.129)$$

que se satisface únicamente si F_{NS} es un número impar. Así, en el sector NS estamos conservando sólo a los estados que tienen un número impar de excitaciones por osciladores b o estamos quitando del espectro a los estados que tienen un número par de excitaciones por osciladores b . En el sector R podemos proyectar fuera a estados con un número par o impar de excitaciones por osciladores d dependiendo de la quiralidad del espinor del estado base. La forma en que se proyecten fuera diferentes estados en ambos sectores da lugar a distintas teorías con distintas partículas y propiedades del espacio-tiempo. Se da a lugar a las teorías de supercuerdas tipo II-A y tipo II-B, las cuales tienen diferentes proyecciones GSO. El nombre de G-paridad fue introducido originalmente en el artículo original de NS, el cual deseó utilizar esta teoría para hadrones. Este operador fue identificado en ese artículo con el operador de G-paridad para hadrones. En lo que se presenta aquí, su rol es completamente diferente.

Acabamos de ver que el estado base, bajo la condición GSO, en la teoría de supercuerdas RNS tiene un número igual de grados de libertad bosónicos y fermiónicos, lo cual es suficiente pero no necesario para concluir que los dos estados base forman un supermultiplete. También se puede probar que la proyección GSO deja el mismo número de bosones y fermiones en cada nivel de masa del espectro para la teoría de supercuerdas RNS. Notemos que los dos estados base forman dos representaciones reales inequivalentes 8-dimensionales de $SO(8)$.

1.2.7. Espectro para la supercuerda cerrada

Como ya hemos visto, la cuerda cerrada tiene osciladores izquierdos y derechos. También está la posibilidad de que cada uno de esos osciladores tenga condiciones de frontera de Ramond (R) o de Neveu-Schwarz (NS). Por lo que para analizar el espectro la cuerda cerrada, debemos considerar los 4 posibles sectores: R-R, R-NS, NS-R y NS-NS. Proyectando hacia los estados con G-Paridad positiva en el sector NS, se remueve el estado base taquiónico que exhibimos en la sección pasada. Para el sector R podemos proyectar hacia los estados con G-paridad positiva o negativa dependiendo de la quiralidad del estado base del cual estén contruidos los estados. De esta elección surgen dos teorías de supercuerdas RNS, la tipo II-A y la tipo II-B, que dependen de si son iguales u opuestas la G-paridad en los sectores R derechos e izquierdos.

En la teoría tipo II-B, los estados base del sector R derechos o izquierdos tienen la misma quiralidad. Así que los dos sectores R tienen la misma G-paridad. Denotemos cada uno de ellos

como $|+\rangle_R$. Con estas consideraciones los estados no masivos de la teoría cerrada de cuerdas tipo II-B son

$$|+\rangle_R \otimes |+\rangle_R, \quad (1.130)$$

$$\tilde{b}_{-1/2}^i |0\rangle_{NS} \otimes b_{-1/2}^i |0\rangle_{NS}, \quad (1.131)$$

$$\tilde{b}_{-1/2}^i |0\rangle_{NS} \otimes |+\rangle_R, \quad (1.132)$$

$$|+\rangle_R \otimes b_{-1/2}^i |0\rangle_{NS}. \quad (1.133)$$

Como el estado $|+\rangle_R$ es un espinor con ocho componentes, cada uno de los 4 sectores en la teoría tipo II-B contiene $8 \times 8 = 64$ estados.

En la teoría tipo II-A los estados base derechos e izquierdos del sector R tienen quiralidad opuesta, y los denotaremos como $|+\rangle_R$ y $|-\rangle_R$. Los estados no masivos en la teoría cerrada de cuerdas tipo II-A son

$$|-\rangle_R \otimes |+\rangle_R, \quad (1.134)$$

$$\tilde{b}_{-1/2}^i |0\rangle_{NS} \otimes b_{-1/2}^i |0\rangle_{NS}, \quad (1.135)$$

$$\tilde{b}_{-1/2}^i |0\rangle_{NS} \otimes |+\rangle_R, \quad (1.136)$$

$$|-\rangle_R \otimes b_{-1/2}^i |0\rangle_{NS}. \quad (1.137)$$

Con argumentos similares a los vistos en la teoría tipo II-B, vemos que cada uno de los 4 sectores de la teoría tipo II-A tiene 64 estados. También se aprecia que en ambas teorías tipo II, uno tiene que el espectro no masivo contiene gravitinos de Majorana-Weyl. Entonces forman mutipletes supergravitacionales $\mathcal{N} = 2$.

Resumamos entonces los distintos tipos de estados no masivos en ambos tipos de teorías:

- Sector R-R: Se trata de bosones obtenidos al hacer el producto tensorial de un par de espinores Majorana-Weyl. En la teoría II-A los dos espinores tienen quiralidades opuestas y se obtiene una 1-forma (vector) correspondiente a 8 estados y una 3-forma (56 estados). En la teoría II-B los espinores tienen la misma quiralidad y se obtiene una 0-forma (un escalar) correspondiente a un estado, una 2-forma correspondiente a 28 estados y una 4-forma correspondiente a 35 estados.
- Sector NS-NS: Ambas teorías tipo II coinciden en este sector. El espectro contiene un escalar llamado dilatón (1 estado), una 2-forma de norma antisimétrica (28 estados) llamado campo de Kalb-Ramond y un tensor de rango 2 sin traza, el gravitón (35 estados).
- Sectores NS-R y R-NS: Cada uno de estos sectores contiene un gravitino de espín 3/2 (56 estados) y un fermión de espín 1/2 llamado dilatino (8 estados). En la teoría II-B los dos gravitinos tienen la misma quiralidad, mientras que en la teoría II-A tienen quiralidad opuesta.

1.3. La cuerda heterótica

Una teoría de supercuerdas realista debe dar origen a campos de norma para las interacciones débiles y fuertes y tal vez para teorías de gran unificación. La forma más sencilla de que esto pase es que estos campos de norma estén presentes en la teoría 10-dimensional. La teoría tipo II-A de cuerdas cerradas no parece ser el camino pues es una teoría no quirral. La teoría de cuerdas tipo II-B tampoco parece ser adecuada porque los únicos estados no masivos que contiene están en el multiplete de supergravedad con $\mathcal{N} = 2$, por lo que no se tienen campos de norma no abelianos. Para obtener una teoría de cuerdas cerradas con un grupo de norma no abeliano se puede usar el método que mezcla osciladores derechos de una teoría de supercuerdas tipo II y los osciladores izquierdos de una teoría de cuerdas bosónica. La proyección GSO se utiliza sobre los osciladores derechos para asegurar el mismo número de grados de libertad fermiónicos y bosónicos, como se requiere para tener supersimetría espacio-temporal. La libertad de escoger los osciladores derechos e izquierdos de distintos tipos de cuerdas viene del hecho de que los estados de una teoría de cuerdas bosónica o una teoría de supercuerdas tipo II son el producto directo de estados del espacio de Fock para osciladores derechos e izquierdos. Este tipo de teoría de cuerdas se le conoce como teoría de cuerdas heterótica. Tal vez puede parecer que no tiene sentido porque los osciladores derechos viven en un espacio-tiempo de 10 dimensiones, mientras que los osciladores izquierdos viven en uno de 26. Sin embargo, en esta aproximación lo que se hará es compactificar 16 de los osciladores izquierdos asociándolos con un toro 16-dimensional con un radio de escala extremadamente pequeño fijado por el valor de la longitud fundamental l_s .

1.3.1. La hoja de mundo

La teoría de cuerdas heterótica puede ser descrita como una teoría de campos 2-dimensional. Nuestro espacio base es parametrizada por una coordenada tipo temporal τ y una coordenada tipo espacial σ ; este espacio es la hoja de mundo trazada por la cuerda. Más precisamente, denotemos la hoja de mundo 2-dimensional como M_2 , un espacio-tiempo con métrica Lorentziana. Las simetrías de la acción de la cuerda clásica nos permite montarnos sobre el sistema de referencia Minkowskiano de la hoja de mundo, donde la acción toma la forma:

$$S = -\frac{1}{2\pi} \int d\tau d\sigma \left[\partial_\alpha X_\mu \partial^\alpha X^\mu + 4i\Psi^\mu (\partial_\tau - \partial_\sigma)\Psi_\mu + \sum_{I=1}^{16} \partial_\alpha X^I \partial^\alpha X^I \right] \quad (1.138)$$

En este sistema de referencia, la longitud o distancia invariante está dada por:

$$ds^2 = -d\tau^2 + d\sigma^2 \quad (1.139)$$

El parámetro τ simboliza al tiempo propio en el sistema de referencia de la cuerda. La coordenada de la hoja de mundo σ parametriza puntos a lo largo de la cuerda en su tiempo propio, con $\sigma \rightarrow \sigma + \pi$

cada que uno le da una vuelta completa a la cuerda. Se suele también trabajar con las coordenadas de hoja de mundo *oscilante izquierda* y *oscilante derecha*:

$$\sigma^- = \tau - \sigma \quad \sigma^+ = \tau + \sigma. \quad (1.140)$$

En la acción 1.138 podemos apreciar tres términos importantes. El primero es la ecuación de la cuerda bosónica para los grados de libertad $X^\mu, \mu = 0, 1, \dots, 9$, el segundo término es la ecuación de Dirac para la cuerda fermiónica y corresponde a los grados de libertad $\psi^\mu, \mu = 0, 1, \dots, 9$. Ya vimos que entre estos grados de libertad existe una supersimetría y ya describen a la cuerda supersimétrica. El último término es la ecuación de la cuerda bosónica para 16 grados de libertad bosónicos extra y que darán origen a simetrías de norma y por ende bosones de norma.

1.3.2. Descripción covariante

Los campos en esta teoría mapean de este espacio base a un espacio objetivo que resulta ser una *supervariedad Riemanniana*. Es decir, el espacio objetivo es el producto cartesiano de una variedad real y una variedad Grassmanniana. La cuerda heterótica es un mapeo descrito por

$$(\sigma^+, \sigma^-) \rightarrow (X^\mu(\sigma^+, \sigma^-), \psi(\sigma^-), X^I(\sigma^+)) \quad (1.141)$$

la parte derecha de la ecuación 1.141 pertenece al espacio $M_{10} \times G_{10}^- \times T^{16}$. Aquí, M_{10} es un espacio de Minkowski 10-dimensional. $X^\mu(\sigma^+, \sigma^-)$ transforma en la representación vectorial correspondiente al grupo de rotación $O(1, 9)$. G_{10}^- es un espacio vectorial impar, es decir, las funciones en G_{10}^- forman un álgebra de Grassmann. La simetría del espacio objetivo $O(1, 9)$ también se impone en estas coordenadas, y están en la representación vectorial. Sin embargo, cada componente $\psi^\mu(\sigma^-)$ está dada por un espinor de *Weyl-Majorana* $O(1, 1)$ con quiralidad negativa. Es decir, existe una base para las matrices de Dirac en dos dimensiones donde

$$(\psi^\mu(\tau, \sigma))^* = \psi^\mu(\tau, \sigma) \quad \gamma^3 \psi^\mu(\tau, \sigma) = -\psi^\mu(\tau, \sigma). \quad (1.142)$$

Aquí, $\gamma^3 = \gamma^0 \gamma^1$ es la matriz de quiralidad en dos dimensiones, que anticonmuta con las dos matrices 2-dimensionales de Dirac: $\{\gamma^3, \gamma^m\} = 0, (m = 0, 1)$. Las coordenadas $X^I(\sigma_+)$ ($I = 1, \dots, 16$) son la imagen de la cuerda en un toro $T^{16} = \mathbb{R}^6 / \Lambda$, con Λ una retícula 16-dimensional. En la teoría de cuerdas heterótica cuantizada, uno encuentra que para tener una consistencia interna se imponen fuertes restricciones para las posibles elecciones de Λ . Las únicas elecciones consistentes son las retículas raíz $E_8 \times E_8$ o $spin(32)/\mathbb{Z}_2$. En este trabajo nos concentraremos en el caso correspondiente a $E_8 \times E_8$.

1.3.3. Norma de cono de luz

Cuando uno aplica el formalismo canónico a lo que ya hemos descrito, uno se da cuenta de que no todos los momentos canónicos son independientes, por lo que tenemos libertad de norma. Las

ecuaciones que relacionan a dichos momentos forman un sistema de constricciones ($\partial_{\pm} = \partial/\partial\sigma_{\pm}$)

$$\psi^{\mu}\partial_{-}X^{\mu} = 0, \quad \partial_{+}X^{\mu}\partial_{+}X_{\mu} = 0, \quad \partial_{-}X^{\mu}\partial_{-}X_{\mu} + \frac{i}{2}\psi^{\mu}\partial_{-}\psi_{\mu} = 0, \quad (1.143)$$

que deben ser satisfechas por soluciones a las ecuaciones de movimiento de Euler-Lagrange

$$\partial_{+}\psi^{\mu} = 0, \quad \partial_{+}\partial_{-}X^{\mu} = 0. \quad (1.144)$$

Existe una arbitrariedad en cómo las ecuaciones 1.143 pueden ser satisfechas, correspondiendo a la parametrización de la acción. Explotando esta invariancia podemos fijar una norma del sistema y eliminar grados de libertad dependientes. Para fijar la norma, se ha probado que la norma de cono de luz es muy útil para dichos propósitos. En ésta, la dirección temporal $\mu = 0$ y una dirección espacial $\mu = 9$ se juntan para un tratamiento especial. Definimos

$$X^{\pm}(\sigma, \tau) = X^0(\sigma, \tau) \pm X^9(\sigma, \tau), \quad \psi^{\pm}(\sigma_{-}) = \psi^0(\sigma_{-}) \pm \psi^9(\sigma_{-}). \quad (1.145)$$

La norma del cono de luz utiliza la invariancia residual para fijar

$$X^{+}(\sigma, \tau) \equiv x^{+} + p^{+}\tau \quad \psi^{+}(\sigma, \tau) = 0, \quad (1.146)$$

para todo σ, τ . Aquí, x^{+}, p^{+} son constantes. Las ecuaciones de restricción entonces se satisfacen notando que podemos hacer $X^{-}(\sigma, \tau)$ y $\psi^{-}(\sigma, \tau)$ funciones de las coordenadas *transversas* $X^i(\sigma, \tau), \psi^i(\sigma_{-}), i = 1, 2, \dots, 8$:

$$X^{-}(\sigma, \tau) = F[X^i(\sigma, \tau), \psi^i(\sigma_{-})], \quad (1.147)$$

$$\psi^{-}(\sigma, \tau) = G[X^i(\sigma, \tau), \psi^i(\sigma_{-})]. \quad (1.148)$$

Por lo que las coordenadas transversas $X^i(\sigma, \tau), \psi^i(\sigma_{-}) (i = 1, \dots, 8)$ contienen la dinámica de la cuerda.

1.3.4. Expansión en modos

Para los osciladores derechos, la expansión en modos en la teoría heterótica de cuerdas son los mismos que obtuvimos en la teoría de cuerdas bosónica y las supercuerdas.

$$X_R^{\mu}(\tau - \sigma) = \frac{1}{2}x^{\mu} + \frac{1}{2}p^{\mu}(\tau - \sigma) + \frac{i}{2} \sum_{n \neq 0} \frac{\alpha_n^{\mu}}{n} e^{-2in(\tau - \sigma)} \quad (1.149)$$

y para los grados de libertad fermiónicos tenemos

$$\Psi_R^{\mu}(\tau - \sigma) = \sum_{n \in \mathbb{Z}} d_n^{\mu} e^{-2in(\tau - \sigma)} \quad \text{R} \quad (1.150)$$

y

$$\Psi_R^\mu(\tau - \sigma) = \sum_{r \in \mathbb{Z} + 1/2} b_r^\mu e^{-2ir(\tau - \sigma)} \quad \text{NS} \quad (1.151)$$

para cada $\mu = 0, 1, \dots, 9$. Ahora hagamos una distinción entre los primeros 10 grados de libertad bosónicos, que denotaremos por $X_L^\mu, \mu = 0, 1, \dots, 9$ y los últimos 16 grados de libertad, a los cuales llamaremos "internos" por medio de $X_L^I, I = 1, \dots, 16$. Para los primeros 10 grafos de libertad, la expansión de modos es la que ya hemos discutido para la parte izquierda de las cuerdas cerradas bosónicas

$$X_L^\mu(\tau + \sigma) = \frac{1}{2}x^\mu + \frac{1}{2}p^\mu(\tau + \sigma) + \frac{i}{2} \sum_{n \neq 0} \frac{\tilde{\alpha}_n^\mu}{n} e^{-2in(\tau + \sigma)}, \mu = 0, 1, \dots, 9. \quad (1.152)$$

Para el caso de los 16 grados de libertad internos tenemos la siguiente expansión en modos

$$X_L^I(\tau + \sigma) = x_L^I + p_L^I(\tau + \sigma) + \frac{i}{2} \sum_{n \neq 0} \frac{\tilde{\alpha}_n^I}{n} e^{-2in(\tau + \sigma)}, I = 1, \dots, 16 \quad (1.153)$$

La cuantización para los primeros 10 grados de libertad son los siguientes

$$[x^\mu, p^\nu] = -i\eta^{\mu\nu} \quad [\alpha_m^\mu, \alpha_n^\nu] = [\tilde{\alpha}_m^\mu, \tilde{\alpha}_n^\nu] = -m\delta_{m+n,0}\eta^{\mu\nu} \quad (1.154)$$

$$\{d_m^\mu, d_n^\nu\} = -\delta_{m+n,0}\eta^{\mu\nu} \quad \text{R} \quad \{b_r^\mu, b_s^\nu\} = -\delta_{m+n,0}\eta^{\mu\nu} \quad \text{NS}, \quad (1.155)$$

donde en cada caso $\mu, \nu = 0, 1, \dots, 9$. Para el caso de los grados de libertad $X_L^I, I = 1, \dots, 16$ se debe tener un poco más de cuidado, pues en la teoría no están los osciladores derechos con los cuales están emparejados. El tratamiento para este caso es como sigue. Tomamos la expresión completa para $X^I = X_L^I + X_R^I$ e imponemos la restricción

$$(\partial_\tau - \partial_\sigma)X^I = 0 \quad (1.156)$$

Aquí, utilizaremos los paréntesis de Poisson en la teoría clásica, donde incorporaremos la restricción de segunda clase 1.156 antes de pasar a la teoría cuántica reemplazando dichos paréntesis por un conmutador multiplicado por $-i$. El resultado de dicho cálculo sería como sigue

$$[P^I(\tau, \sigma), X^J(\tau, \sigma)] = \frac{i}{2}\delta(\sigma - \sigma')\eta^{IJ} = -\frac{i}{2}\delta(\sigma - \sigma')\delta^{IJ} \quad (1.157)$$

lo que nos impone la condición siguiente en el conmutador para el momento y posición de centro de masa

$$[X_L^I, P_L^I] = -\frac{i}{2}\eta^{IJ} = \frac{i}{2}\delta^{IJ} \quad (1.158)$$

mientras que los osciladores $\tilde{\alpha}_n^I$ tiene el conmutador

$$[\tilde{\alpha}_m^I, \tilde{\alpha}_n^J] = -m\delta_{m+n,0}\eta^{IJ} = m\delta_{m+n,0}\delta^{IJ} \quad (1.159)$$

sin $1/2$ porque no tenemos la contribución del oscilador derecho.

Ahora bien, tenemos entonces que los estados oscilantes derechos son perturbaciones al estado vacío $|0\rangle_R$, que se define como el estado tal que $b_r |0\rangle_R = d_n |0\rangle_R = 0$, para $r, n > 0$. La masa de los estados está dada por

$$\frac{M_R^2}{4} = N_R - a_R = \begin{cases} \sum_{n=1}^{\infty} (\alpha_{-n}^i \alpha_n^i + n d_n^i d_n^i) & \text{R,} \\ \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i + \sum_{r=1/2}^{\infty} r b_{-r}^i b_r^i - \frac{1}{2} & \text{NS} \end{cases} \quad (1.160)$$

donde a_R es la energía de punto cero que se origina por el orden normal de los osciladores durante la cuantización. Como ya vimos anteriormente, en el sector Ramond $a_R = 0$ mientras que en el Neveu-Shwarz $a_R = -\frac{1}{2}$. De la ecuación 1.162, observamos que $m_R = 0$ para estados con un oscilador $b_{-1/2}$ actuando sobre el vacío, $b_{-1/2} |0\rangle_R$, para el sector Neveu-Shwarz, y $d_0^i |0\rangle_R$ en el sector Ramond. Las ocho excitaciones transversas $b_{-1/2}^i (i = 3, \dots, 10)$ se comportan como bosones desde el punto de vista del espacio-tiempo y forman una representación vectorial 8_v de $SO(8)$, que es el grupo de Lorentz transversal del espacio no compactificado.

Por otro lado, después de la cuantización, los osciladores d_0^i cumplen con el álgebra de Clifford

$$\{i\sqrt{2}d_0^i, i\sqrt{2}d_0^j\} = 2\eta^{ij}, \quad (1.161)$$

el estado base $d_0^i |0\rangle_R$ forma una representación espinorial con 16 componentes. La proyección GSO (de la que hablamos en la sección anterior) ayuda a equiparar el número de grados de libertad bosónico y fermiónico. Después de la proyección GSO, nos quedamos con la 8_v representación de $SO(8)$ en el sector Ramond.

Por conveniencia, $|q\rangle_R$ denotará el estado base oscilante derecho para ambos sectores, donde q simboliza los pesos correspondientes a la representación de $SO(8)$ en los símbolos de Cartan-Weyl

$$|q\rangle_R = \begin{cases} |\pm 1, 0, 0, 0\rangle_R \sim 8_v & \text{NS,} \\ |\pm \frac{1}{2}, \pm \frac{1}{2}, \pm \frac{1}{2}, \pm \frac{1}{2}\rangle_R \sim 8_s & \text{R} \end{cases} \quad (1.162)$$

donde la representación espinorial tiene un número par de signos positivos. La línea por abajo denota permutación de las entradas. Por tanto, la ecuación 1.162 se convierte en

$$\frac{M_R^2}{4} = \frac{1}{2}q^2 - \frac{1}{2} \quad (1.163)$$

Como vimos en el caso de los osciladores derechos, los estados oscilantes izquierdos son perturbaciones del vacío $|0\rangle_L$. Las masas de estos estados después de compactificar los grados de libertad internos X^I (véase siguiente sección) están dadas por

$$\frac{M_L^2}{4} = \frac{1}{2}p^2 + \tilde{N} - 1 \quad (1.164)$$

donde \tilde{N} cuenta las excitaciones de los osciladores izquierdos y -1 es la energía de punto cero de la cuerda bosónica.

De acuerdo a esta ecuación, a nivel no masivo se tienen los siguientes estados oscilantes izquierdos:

$$\tilde{\alpha}_{-1}^i |0\rangle_L \quad i = 3, \dots, 10 \quad (1.165)$$

$$\tilde{\alpha}_{-1}^I |0\rangle_L \quad I = 1, \dots, 16 \quad (1.166)$$

$$|p\rangle_L \quad p^2 = 2 \quad (1.167)$$

Se tienen 480 momentos internos que cumplen la condición de que $p^2 = 2$. Los estados $|p\rangle_L$ corresponden a las partes izquierdas de los bosones de norma (los norminos) de esta teoría. Por esta razón, representaremos p por los vectores correspondientes a los bosones cargados

$$\mathcal{G} = E_8 \times E_8 : \quad \begin{cases} ((\pm 1)^2, 0^6), (0^8)((\pm 1)^2, 0^6) \\ ((\pm \frac{1}{2})^8)(0^8), (0^8)((\pm \frac{1}{2})^8) \end{cases} \quad \text{número par de } + \quad (1.168)$$

1.3.5. Compactificación en T^{16}

Como ya se mencionó antes, los 16 osciladores izquierdos adicionales de la teoría heterótica de cuerdas se compactifican en un toro para así obtener una teoría 10-dimensional con un grupo de norma emergente de las dimensiones compactificadas. Una teoría 10-dimensional con grados de libertad internos $X_L^I, I = 1, \dots, 16$ que proveerán un grupo de norma se podría construir compactificando estos grados de libertad en un toro 16-dimensional. Es necesario imponer condiciones de frontera en X^I para ambos, los modos derechos e izquierdos, y luego eliminar los derechos. Un toro 16 dimensional se define tomando en cuenta una retícula Γ con vectores base $e_a^I, a = 1, \dots, 16$ de tal manera que tengan una longitud $\sqrt{2}$ e imponiendo la identificación

$$x^I \equiv x^I + \sqrt{2}\pi \sum_{a=1}^{16} n_a R_a e_a^I \quad (1.169)$$

donde R_a son los radios y n_a son enteros arbitrarios. Por lo que habrá maneras adicionales de cumplir las condiciones de frontera para la cuerda cerrada enrollando a la cuerda en el toro, de tal manera que

$$X^I(\tau, \sigma + \pi) = X^I(\tau, \sigma) + \sqrt{2}\pi \sum_{a=1}^{16} n_a R_a e_a^I = X^I(\tau, \sigma) + 2\pi L^I \quad (1.170)$$

donde

$$L^I = \frac{1}{\sqrt{2}} \sum_{a=1}^{16} n_a R_a e_a^I \quad (1.171)$$

a estas últimas cantidades se les denomina números de enrollamiento. Por lo que al final, la expansión en modos queda como

$$X^I(\tau, \sigma) = x^I + p^I \tau + 2L^I \sigma + \frac{i}{2} \sum_{n \neq 0} \frac{1}{n} \left(\alpha_n^I e^{-2in(\tau-\sigma)} + \tilde{\alpha}_n^I e^{-2in(\tau+\sigma)} \right) = X_R^I + X_L^I \quad (1.172)$$

donde tenemos que los modos izquierdos son

$$X_L^I(\tau + \sigma) = x_L^I + p_L^I(\tau + \sigma) + \frac{i}{2} \sum_{n \neq 0} \frac{1}{n} \tilde{\alpha}_n^I e^{-2in(\tau+\sigma)} \quad (1.173)$$

donde el momento izquierdo son ahora

$$p_R^I = \frac{1}{2}(p^I - 2L^I) \quad P_L^I = \frac{1}{2}(p^I + 2L^I) \quad (1.174)$$

y

$$x^I = x_R^I + x_L^I. \quad (1.175)$$

Ahora bien, para eliminar los osciladores derechos, que significa hacer $p_R^I = 0$, con la consecuencia de que

$$p_L^I = 2L^I, \quad (1.176)$$

también debemos tomar $x_R^I = 0$, lo que implicará

$$x_L^I \equiv x^I + \sqrt{2\pi} \sum_{a=1}^{16} n_a R_a e_a^I \quad (1.177)$$

Las relaciones de conmutación implica que $2p_L^I$ es el que genera las traslaciones de x_L^I , por lo que el momento está definido por

$$p_L^I = \frac{1}{\sqrt{2}} \sum_{a=1}^{16} \frac{m_a}{R_a} e_a^{*I} \quad (1.178)$$

donde m_a son enteros arbitrarios y e_a^{*I} es la base de la retícula dual $\tilde{\Gamma}$. En la norma de cono de luz, el operador 10-dimensional de la masa cuadrada para los estados físicos está dado por

$$M^2 = M_R^2 + M_L^2. \quad (1.179)$$

Para los osciladores derechos de supercuerdas con $D = 10$ tenemos que

$$\frac{1}{4} M_R^2 = N, \quad (1.180)$$

con N el operador de número que tiene una expresión distinta dependiendo de si se está en el sector Ramond o Neveu-Schwartz. Para los osciladores izquierdos de la cuerda bosónica con 16 dimensiones compactificadas en un toro, el análogo sería

$$\frac{1}{4}M_L^2 = \frac{1}{2} \sum_{I=1}^{16} (p_L^I)^2 + \tilde{N} - 1 \quad (1.181)$$

con

$$\tilde{N} = \sum_{n=1}^{\infty} (\tilde{\alpha}_n^i \tilde{\alpha}_{in} + \tilde{\alpha}_{-n}^I \tilde{\alpha}_{In}) \quad (1.182)$$

También, para estados físicos, tenemos que

$$M_R^2 = M_L^2 \quad (1.183)$$

por lo que para los estado no masivos, requeriríamos que

$$M_R^2 = M_L^2 = 0 \quad (1.184)$$

Resulta que hay pocas opciones para la elección de la retícula Γ que sea consistente con una teoría de cuerdas aceptable. Existe una restricción fundamental en teoría de cuerdas llamada invariancia modular, la cual se requiere para asegurar la ausencia de anomalías gravitacionales y de norma y la finitud de las contribuciones de los lazos de cuerdos en las amplitudes de dispersión. Sólo dos retículas en 16 dimensiones que aseguran que esto es cumplido, denotadas por Γ_{16} y $\Gamma_8 \times \Gamma_8$. La primera contiene la retícula raíz de $SO(32)$ como subretícula. En lo que sigue nos concentraremos en la segunda posibilidad en la cual tenemos el producto directo de dos retículas raíz E_8 . Los momentos están entonces en la retícula raíz $E_8 \times E_8$.

Regresando a los grados de libertad X^μ, Ψ^μ , debemos notar que no todos son independientes. En la norma de cono de luz, las coordenadas

$$X^\pm = \frac{1}{\sqrt{2}}(X^0 \pm X^1), \quad \Psi^\pm = \frac{1}{\sqrt{2}}(\Psi^0 \pm \Psi^1) \quad (1.185)$$

están escritas en términos de otros grados de libertad, por lo tanto, sólo X^μ, Ψ^μ para $\mu = 2, \dots, 9$ son físicos. En esta norma, sólo el grupo transversal $SO(8)$ del grupo de Lorentz 10-dimensional $SO(9, 1)$ es relevante.

1.3.6. Espectro no masivo de la cuerda heterótica

Estados físicos deben cumplir con la condición de emparejamiento de niveles

$$M_R^2 = M_L^2 \quad (1.186)$$

que se sigue de la variación de la métrica de hoja de mundo. Esta restricción implica que, en contraste con otras teorías, la proyección GSO no se implementa para evitar la presencia de taquiones como hicimos en la teoría de supercuerdas. Como la masa del estado de energía más bajo para un estado oscilante izquierdo ($\tilde{N} = 0, p = 0$) es $m_L^2 = -1$ mientras que la masa del taquión oscilante derecho ($N_R = 0$) es $m_R^2 = -\frac{1}{2}$, entonces la ecuación 1.51 prohíbe la existencia de estados con masa cuadrada negativa en el espectro de la cuerda heterótica.

Combinando los osciladores izquierdos y derechos obtenemos los siguientes estados:

- un multiplete 10-dimensional de supergravedad \mathcal{N}

$$|q\rangle_R \tilde{\alpha}_{-1}^j |0\rangle_L, \quad i = 3, \dots, 10, \quad (1.187)$$

incluyendo el gravitón g^{ij} , el dilatón ϕ , el tensor antisimétrico B^{ij} y su compañero supersimétrico.

- 16 bosones de norma no cargados (y los norminos)

$$|0\rangle_R \tilde{\alpha}_{-1}^I |0\rangle_L, \quad I = 1, \dots, 16 \quad (1.188)$$

el cual contiene los generadores de Cartan H_I del grupo de norma \mathcal{G} ;

- 480 bosones de norma cargados (y norminos)

$$|q\rangle_R |p\rangle_L, \quad p^2 = 2 \quad (1.189)$$

con p dada como en 1.168.

Capítulo 2

Orbifolios heteróticos

En el capítulo pasado ya estudiamos la construcción de la teoría de cuerdas heterótica. Los grados de libertad de norma en dicha teoría están codificados en los osciladores bosónicos $X^I(\sigma, \tau)$, $I = 1, 2, \dots, 16$, los cuales compactificamos en un toro 16-dimensional con la retícula raíz $E_8 \times E_8$. Pero como veremos a continuación, eso no es suficiente si se desea probar que la teoría de cuerdas es una teoría que unifica a todas las interacciones[14]. Para poder argumentar lo anterior, debemos ser capaces de establecer contacto entre $SU(3) \times SU(2) \times U(1)$ del modelo estándar y la teoría de cuerdas. La compactificación por orbifolios proveen de un posible camino para esto [14]. Lo que se hace es empezar con una teoría de cuerdas consistente con *invariancia modular* y modificarla agregando ciertas simetrías para obtener otra teoría. Si lo que se desea es obtener una teoría de cuerdas 4-dimensional, entonces las modificaciones incluyen una *compactificación* de las dimensiones extra en un toro. Más aún, si se desea obtener una teoría quirral, se deben incluir condiciones de frontera de *torcedura* en los campos de la hoja de mundo.

Las construcciones de teorías de cuerdas con orbifolios son teorías 4-dimensionales. Se consideran una extensión a la construcción de la teoría de cuerdas heterótica, pues no sólo se compactifican los 16 osciladores izquierdos bosónicos en un toro, si no que además se compactifican los otros 6 osciladores izquierdos y derechos en un toro. Aunque las ventajas ya mencionadas derivadas de la construcción con orbifolios son útiles, los aspectos fenomenológicos de dichas construcciones no parecen muy prometedoras, pues tienden a dar un número muy grande de generaciones de quarks-leptones [15]. Esta situación cambia cuando hay campos de norma de fondo (líneas de Wilson) en el respectivo toro. En este caso, el grupo de norma inicial es más pequeño que $E_8 \times E_8$. Además, el sector no masivo del espectro de la cuerda cambia completamente y es más fácil obtener modelos de orbifolios con tres generaciones con grupos de norma realistas (por ejemplo, $SU(3) \times SU(2) \times U(1)^n$) [16].

2.1. Compactificaciones toroidales

Para tener una verdadera teoría realista es esencial compactificar 6 de las 9 dimensiones espaciales a una escala lo suficientemente pequeña, de tal manera que dichas dimensiones no sean observables para los aceleradores actuales. La manera más sencilla de hacer esto es haciendo la compactificación en un toro. Esto nos asegura que las simples ecuaciones lineales de movimiento de las cuerdas no se vean afectadas, pues los toros son planos. Trabajaremos en la norma del cono de luz, introducida en el capítulo pasado. Tenemos entonces 8 grados de libertad bosónicos transversos, denotados por $X^i(\tau, \sigma)$, $i = 1, 2$, y $X^k(\tau, \sigma)$, $k = 3, \dots, 8$. Estos grados de libertad representan las coordenadas de la hoja de mundo para cuerdas cerradas, y están separados en modos izquierdos y derechos como

$$X^{i(k)}(\tau, \sigma) = X_R^{i(k)}(\tau - \sigma) + X_L^{i(k)}(\tau + \sigma). \quad (2.1)$$

Adicionalmente, tenemos 8 modos derechos fermiónicos transversos, $\Psi_R^i(\tau - \sigma)$, $\Psi_R^k(\tau - \sigma)$, y 16 grados de libertad internos como modos izquierdos bosónicos $X^I(\tau + \sigma)$, ($I = 1, \dots, 16$) que generan el grupo de norma $E_8 \times E_8$ de la cuerda heterótica de cuerdas 10-dimensional. La compactificación de las 6 coordenadas espaciales $X^k(\tau, \sigma)$, ($k = 3, \dots, 8$) no afecta la expansión en modos de $X^i(\tau, \sigma)$, $\Psi_R^i(\tau - \sigma)$, $\Psi_R^k(\tau - \sigma)$ o $X_L^I(\tau + \sigma)$, así que las expansiones en modos son tales que

$$X^i(\tau, \sigma) = x^i + p^i \tau + \frac{i}{2} \sum_{n \neq 0} \left[\frac{1}{n} \alpha_n^i e^{-2in(\tau - \sigma)} + \frac{1}{n} \tilde{\alpha}_n^i e^{-2in(\tau + \sigma)} \right], \quad (2.2)$$

$$\begin{aligned} \Psi_R^{i(k)}(\tau - \sigma) &= \sum_{n \in \mathbb{Z}} d_n^{i(k)} e^{-2in(\tau - \sigma)}, & \text{R} \\ &= \sum_{r \in \mathbb{Z} + 1/2} b_r^{i(k)} e^{-2ir(\tau - \sigma)} & \text{NS.} \end{aligned} \quad (2.3)$$

Mientras que la expansión en modos de los grados de libertad de norma es la que vimos desde la cuerda bosónica

$$X_L^I(\tau + \sigma) = x_L^I + p_L^I(\tau + \sigma) + \frac{i}{2} \sum_{n \neq 0} \frac{\tilde{\alpha}_n^I}{n} e^{-2in(\tau + \sigma)}, \quad (2.4)$$

con p_L^I en la retícula raíz $E_8 \times E_8$.

En una base ortonormal, los vectores en E_8 son de la forma

$$l = (n_1, n_2, \dots, n_8) \quad \text{ó} \quad l = \left(n_1 + \frac{1}{2}, n_2 + \frac{1}{2}, \dots, n_8 + \frac{1}{2} \right), \quad (2.5)$$

con n_i enteros tales que (ver Apéndice A)

$$\sum_{n=1}^8 n_i = 0 \pmod{2}, \quad \sum_{i=1}^8 l_i^2 = 2. \quad (2.6)$$

Existe una formulación alternativa de estos grados de libertad internos que reemplaza a los 16 modos bosónicos izquierdos (X^I) compactificados en un toro con retícula $E_8 \times E_8$ por 32 modos izquierdo fermiónicos reales ($\lambda^A, \bar{\lambda}^A$, ($A = 1, \dots, 16$)) donde $\lambda^A, \bar{\lambda}^A$ pueden tener por separado del sector R o del sector NS. Entonces

$$\begin{aligned}\lambda^A &= \sum_{n \in \mathbb{Z}} \lambda_n^A e^{-2in(\tau+\sigma)} \quad \text{R} \\ &= \sum_{r \in \mathbb{Z}+1/2} \lambda_r^A e^{-2ir(\tau+\sigma)},\end{aligned}\tag{2.7}$$

y de manera similar para $\bar{\lambda}^A$. Tenemos entonces que $(\lambda^A, \bar{\lambda}^A)$ transforma como la representación $(\mathbf{16}, \mathbf{1}) + (\mathbf{1}, \mathbf{16})$ del grupo maximal $O(16) \times O(16) \subset E_8 \times E_8$.

Antes de describir el procedimiento para compactificar las coordenadas de los modos espaciales extra, presentemos de manera breve la definición de un toro D -dimensional. Sea Λ una retícula generada por una base de vectores D -dimensionales e_α de \mathbb{R}^n

$$\Lambda = \{n_\alpha e_\alpha \mid n_\alpha \in \mathbb{Z}, \alpha = 1, \dots, D\}.\tag{2.8}$$

De este modo, podemos definir nuestro toro T^D como el cociente

$$T^D = \mathbb{R}^D / \Lambda,\tag{2.9}$$

donde ahora los puntos son identificados en el espacio difiriendo por un vector dentro de la retícula, es decir,

$$X \sim X + n_\alpha e_\alpha.\tag{2.10}$$

Una característica importante que podemos notar aquí es que el espacio que describe a todo el toro se puede restringir a una región llamada *dominio fundamental*. El dominio fundamental es el conjunto más pequeño que está relacionado con el resto del espacio \mathbb{R}^D tal y como se aprecia en la figura 2.1. La métrica del toro está dada por

$$g_{\alpha\beta} = e_\alpha \cdot e_\beta\tag{2.11}$$

Ahora, la compactificación de los modos X^k incluyen la identificación de las coordenadas del centro de masa x^k con puntos que están separados de él por un vector en la retícula del toro. Entonces

$$x^k \equiv x^k + 2\pi L^k,\tag{2.12}$$

donde el factor 2π es por convención y el vector \mathbf{L} con coordenadas L^k pertenece a la retícula 6-dimensional Λ

$$\Lambda \equiv \left\{ \sum_{t=3}^8 r_t e_t \mid r_t \in \mathbb{Z} \right\},\tag{2.13}$$

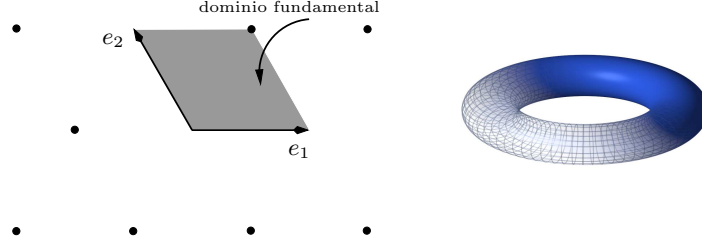


Figura 2.1: a) Retícula 2-dimensional y el dominio fundamental de un toro con esta retícula; b) una representación geométrica alternativa de dicho toro.

donde e_t ($t = 3, \dots, 8$) son vectores base de la retícula. Así pues, las condiciones de frontera de las coordenadas para la cuerda cerrada X^k deben satisfacer también

$$X^k(\tau, \pi) = X^k(\tau, 0) + 2\pi L^k, \quad (2.14)$$

correspondiendo al enrollamiento de la cuerda alrededor del toro. La compactificación también requiere la cuantización de los eigenvalores correspondientes a los operadores de momento p^k . Las eigenfunciones $\exp\{i \sum_k p^k x^k\}$ son univaluadas sólo si

$$\sum_{k=3}^8 p^k L^k \in \mathbb{Z}. \quad (2.15)$$

Entonces los momentos están cuantizados en la retícula Λ^* que es dual a Λ

$$\Lambda^* = \left\{ \sum_{t=3}^8 m_t e_t^* \mid m_t \in \mathbb{Z} \right\}, \quad (2.16)$$

donde los vectores base e_t^* de Λ^* satisfacen $\sum_{k=3}^8 e_t^{*k} e_u^k = \delta_{tu}$.

Así, la generalización de la expansión en modos se ve modificada de la manera

$$X_R^k(\tau, \sigma) = x_R^k + p_R^k(\tau - \sigma) + \frac{i}{2} \sum_{n \neq 0} \frac{1}{n} \alpha_n^k e^{-2in(\tau - \sigma)}, \quad (2.17)$$

$$X_L^k(\tau, \sigma) = x_L^k + p_L^k(\tau + \sigma) + \frac{i}{2} \sum_{n \neq 0} \frac{1}{n} \tilde{\alpha}_n^k e^{-2in(\tau + \sigma)}, \quad (2.18)$$

con

$$p_R^k \equiv \frac{1}{2}(p^k - 2L^k), \quad (2.19)$$

$$p_L^k \equiv \frac{1}{2}(p^k + 2L^k), \quad (2.20)$$

$$x^k = x_R^k + x_L^k, \quad (2.21)$$

donde $p \in \Lambda^*$ y $L \in \Lambda$.

La ecuación de masa para los modos derechos en la teoría de cuerdas heterótica 10-dimensional es modificada ahora a una fórmula de masa 4-dimensional con la compactificación toroidal que acabamos de comentar, de tal manera que, siguiendo la notación de Bailin [17]

$$\frac{1}{4}M_R^2 = N(\beta) + \frac{1}{2}p_R^k p_R^k - \alpha(\beta), \quad (2.22)$$

donde $\beta = R, NS$ rotula las condiciones de frontera correspondientes a los modos derechos fermiónicos, y el operador de número $N(\beta)$ está dado por

$$N(\beta) = N_B + N_F(\beta), \quad (2.23)$$

con

$$N_B = \sum_{n=1}^{\infty} (\alpha_{-n}^i \alpha_n^i + \alpha_{-n}^k \alpha_n^k), \quad (2.24)$$

$$N_F(R) = \sum_{r=1/2}^{\infty} (r b_{-r}^i b_r^i + r b_{-r}^k b_r^k). \quad (2.25)$$

Además, la constante $\alpha(\beta)$ es la constante originada del orden normal del operador de Virasoro L_0 y sus valores son

$$a(R) = 0, \quad a(NS) = \frac{1}{2}. \quad (2.26)$$

Análogamente, la ecuación de masa para los modos izquierdos es

$$\frac{1}{4}M_L^2 = \tilde{N} + \frac{1}{2}p_L^k p_L^k + \frac{1}{2}p_L^I p_L^I - 1, \quad (2.27)$$

donde la suma sobre el índice $I = 1, \dots, 16$ está implícita, también tenemos que el operador de número \tilde{N} es tal que

$$\tilde{N} = \sum_{n=1}^{\infty} (\tilde{\alpha}_{-n}^i \tilde{\alpha}_n^i + \tilde{\alpha}_{-n}^k \tilde{\alpha}_n^k + \tilde{\alpha}_{-n}^I \tilde{\alpha}_n^I). \quad (2.28)$$

Las ecuaciones de masa (2.22) y (2.27) incluyen la contribución de los momentos p_L^k, p_R^k en la variedad compacta del toro y están cuantizados. La retícula Λ y su dual Λ^* poseen factores de escala arbitrarios R_t , que son las longitudes de los vectores base e_t , y ángulos entre éstos. Así que, excepto para ciertos valores específicos de estos parámetros, estados no masivos sólo se dan cuando el momento y el número de enrollamiento en la variedad compacta son cero

$$p_R^k = 0 = p_L^k. \quad (2.29)$$

Pero ahora veremos que este tipo de compactificaciones (la toroidal) no son suficientes o aceptables debido a sus características fenomenológicas. Consideremos un estado no masivo, de tal manera que

$$M_L^2 = 0 = M_R^2. \quad (2.30)$$

Supongamos que fijamos el estado izquierdo no masivo con $\tilde{\alpha}_{-1} |0\rangle_L$ o usando el momento p_L^I en la retícula $E_8 \times E_0$ con $p_L^I p_L^I = 2$. Para cada uno de estos estados izquierdos podemos relacionarle un estado derecho no masivo tal que $b_{-1/2}^i |0\rangle_R$ ($i = 1, 2$), es decir, osciladores NS fermiónicos. Como los índices $i = 1, 2$ corresponde a las dos dimensiones transversas espacio-temporales, el estado completo transforma como un vector o un tensor espacio-temporal si el estado izquierdo es $\tilde{\alpha}_{-1}^j |0\rangle_L$ ($j = 1, 2$). Alternativamente, podemos hacer emparejar el estado base de Ramond no masivo $|0\rangle_R$ al estado izquierdo fijado. Esto transformaría como un espinor quiral 8-dimensional $SO(8)$, donde el espinor con quiralidad contraria ha sido eliminado por la proyección GSO que comentamos en el capítulo pasado. Este espinor de 8 componentes puede ser descompuesto en representaciones de $SO(2) \times SO(6) \subset SO(8)$, donde $SO(2)$ corresponde a las dos coordenadas espacio-temporales y $SO(6)$ a las seis coordenadas compactificadas. Entonces tendríamos

$$\mathbf{8}_L = \left(+\frac{1}{2}\right) \times \mathbf{4} + \left(-\frac{1}{2}\right) \bar{\mathbf{4}}, \quad (2.31)$$

haciendo evidente que hay 4 espinores espacio-temporales de cada quiralidad. Así que si tenemos un estado construido desde la teoría de cuerdas bosónica donde se tenía una partícula vectorial, el estado fermiónico que acabamos de construir corresponde a cuatro norminos; mientras que si de la teoría bosónica teníamos un tensor espacio-temporal, el gravitón, el estado fermiónico correspondiente corresponde a 4 gravitinos. Lo último es consecuencia de la compactificación toroidal que llevamos a cabo, llevándonos inevitablemente a una supersimetría espacio-temporal $\mathcal{N} = 4$, y por tanto una simetría de norma no quiral.

2.2. Orbifolios toroidales

En el capítulo 1 ya introdujimos brevemente la compactificación de los modos de norma para la teoría heterótica de cuerdas en un toro 16-dimensional y obtuvimos cierta intuición sobre cómo construir un toro. También analizamos el espectro no masivo proveniente de esta compactificación. Ahora, la teoría heterótica consta de 10 modos derechos bosónicos y 26 modos izquierdos bosónicos; de estos últimos 16 correspondientes a los grados de libertad de norma. Para obtener una teoría 4-dimensional a partir de la cuerda heterótica, podemos empezar modificándola agregando a la compactificación toroidal isometrías discretas M_R y M_L [14]

$$T_R^6 \otimes T_L^{22} \Rightarrow T_R^6/G_R \otimes T_L^{22}/G_L, \quad (2.32)$$

siendo G_R y G_L isometrías de los toros T_R^6 y T_L^{22} respectivamente. En la teoría heterótica de cuerdas, la supersimetría espacio-temporal viene solamente de los modos derechos. Si deseamos mantener

una supersimetría no rota, G_R tiene que ser un subgrupo discreto de $SU(3) \subset SO(6)$ [18], es decir,

$$G_R \subset SU(3). \quad (2.33)$$

Cuando la simetría discreta es la misma para las 6 dimensiones extra en los grados de libertad izquierdo y derechos, decimos que hemos construido un orbifolio simétrico, es decir,

$$O_s = T_R^6/P \otimes T_L^6/P \otimes T_L^{16}/G, \quad (2.34)$$

y llamamos a $P = G_R = G_L^{(6)}$ el grupo de punto del orbifolio (más adelante aclararemos que la isometría G no siempre coincide con la definición de grupo de punto); el grupo $G_L^{(6)}$ es la simetría discreta asociada a los 6 grados de libertad espaciales izquierdos. A las isometrías G las llamamos grupo de torsión de norma. Una definición más restrictiva del orbifolio simétrico requiere que T_R^6 y T_L^6 sean el mismo toro. Este último caso es el que consideraremos a partir de aquí. La forma en que se parte T_L^{22} en T_L^6 y T_L^{16} es por convención, así que en general, la separación entre las componentes de norma y las espacio de T_L^{22} no es muy clara.

Un orbifolio asimétrico es aquel en el cual las simetrías discretas que modifican a las 6 dimensiones extra de manera diferente para los modos izquierdos y derechos,

$$O_A = T_R^6/P_R \otimes T_L^6/P_L \otimes T_L^{16}/G. \quad (2.35)$$

Consideremos entonces orbifolios simétricos con las coordenadas izquierdas y derechas compactificadas en el mismo toro y modulado por el mismo grupo puntual. También, para ser consistentes con lo que introdujimos en el capítulo anterior respecto a la teoría de cuerdas heterótica, tomaremos $T_L^{16} = T_{E_8 \times E_8}$. Nuestro orbifolio simétrico para los modos izquierdos y derechos puede ser definido como

$$O = \mathbb{R}_{L+R}^6/S \otimes T_{E_8 \times E_8}^L/G. \quad (2.36)$$

Aquí, S es el llamado *grupo de espacio o espacial* del orbifolio e incluye el grupo puntual P más traslaciones en \mathbb{R}^6 , tal que $S = \{(\theta, l^k) | \theta \in P, k = 3, \dots, 8\}$. Un elemento (θ, l^k) del grupo de espacio actúa en las 6 coordenadas espacio-temporales a compactificar como

$$(sX)^k = (\theta X)^k + l^k, \quad i = 3, \dots, 8. \quad (2.37)$$

Es importante mencionar que los vectores de traslación l^i o también llamados *shifts* incluyen vectores e_a^k , $a = 1, \dots, 6$ base que definen al toro 6-dimensional. También estamos considerando que θ son automorfismos de la retícula 6-dimensional que construye a dicho toro. El *grupo de torsión de norma* $G = \{(\Theta, V^I) | I = 1, \dots, 16\}$ actúa sobre las 16 coordenadas de los grados de libertad extra como

$$(GX)^I = (\Theta X)^I + V^I, \quad I = 1, \dots, 16, \quad (2.38)$$

donde Θ representa a un automorfismo de la retícula $E_8 \times E_8$.

Con estos conceptos en mente, podemos entonces hacer una definición de los llamados *orbifolios toroidales* D -dimensionales. Se definen como el cociente de un toro T^D sobre un conjunto de isometrías discretas P

$$\mathcal{O} = T^D/P = \mathbb{R}^D/S, \quad (2.39)$$

donde P es llamado el *grupo de punto* y S es el *grupo de espacio*. Para que el grupo de punto sea aceptable, los elementos de P deben actuar de manera cristalográfica (como una isometría) sobre la retícula Λ , es decir, Λ debe ser invariante bajo P . Recordemos que el toro D -dimensional se define como $T^D = \mathbb{R}^D/\Lambda$. Sea S un grupo de espacio. El subgrupo Λ de S conformado por todas las traslaciones en S es la retícula del grupo de espacio. Notemos que para un elemento en general $g = (\theta, l) \in S$ el vector l no necesita ser un vector retícula. Elementos de la forma $g = (\theta, l) \in S$ con $l \notin \Lambda$ son llamados *roto-traslaciones* [19].

Como el grupo de espacio requiere tener D traslaciones linealmente independientes, toda retícula contiene una base $\mathfrak{e} = \{e_i\}_{i \in \{1, \dots, D\}}$ y la retícula completa es generada por los e_i (con coeficientes enteros), es decir, un elemento $l \in \Lambda$ puede ser escrito como $l = n_i e_i$, con suma sobre $i = 1, \dots, D$ y $n_i \in \mathbb{Z}$. Claramente, la elección de la base no es única. Por ejemplo, para una retícula Λ dada tomemos dos bases $\mathfrak{e} = \{e_1, \dots, e_D\}$ y $\mathfrak{f} = \{f_1, \dots, f_D\}$ y definimos $B_{\mathfrak{e}}$ y $B_{\mathfrak{f}}$ matrices cuyas columnas son los vectores base \mathfrak{e} y \mathfrak{f} , respectivamente. Entonces el cambio de base está dado por la matriz unimodular M (es decir, $M \in GL(D, \mathbb{Z})$) como

$$B_{\mathfrak{e}}M = B_{\mathfrak{f}}. \quad (2.40)$$

Por otro lado, uno puede decidir qué base utilizar para la retícula y generarla indistintamente con cualquiera de ellas.

Ahora, detallemos aspectos del grupo de punto. Para un grupo de espacio S con elementos de la forma (θ, l) , el conjunto P de todos los θ forman un grupo finito, el llamado grupo de punto de S . Los elementos de un grupo de punto también son llamados torsiones o rotaciones. Sin embargo, en general un grupo de punto también incluye inversiones y reflexiones, es decir, $\theta \in O(D)$. El grupo de punto P de S mapea la retícula de S a sí misma. Así, de manera similar al cambio de base de la retícula, los elementos del grupo de punto pueden ser representados por matrices de $GL(D, \mathbb{Z})$. Dadas estas definiciones, y dado que la retícula siempre es un subgrupo normal del grupo de espacio, el grupo S tiene estructura de producto semidirecto si y sólo si el grupo de punto P es un subgrupo del S , es decir, $P \subset S$. En ese caso,

$$S = P \ltimes \Lambda, \quad (2.41)$$

y uno puede definir al orbifolio como

$$\mathcal{O} = \mathbb{R}^D/(P \ltimes \Lambda) = T^D/P. \quad (2.42)$$

En general, sin embargo, el grupo de punto no es un subgrupo del grupo de espacio y por ende el grupo de espacio no es necesariamente un producto semidirecto de su grupo de punto con su retícula. Más precisamente, en general, el grupo de punto P no es igual al grupo de isometría G de (2.32) por la posible presencia de roto-traslaciones. Existen dos tipos de grupos de punto:

1. Abeliano: por ejemplo, $\mathbb{Z}_N, \mathbb{Z}_N \times \mathbb{Z}_M$
2. No abeliano: tales como $S_3, Q_8, D_4, A_4, \dots$

Los elementos del grupo \mathbb{Z}_N son la potencia de algún generador único θ tal que $\theta^N = 1$. En dos dimensiones, θ representa una rotación por $1\pi/N$. No cualquier \mathbb{Z}_N funciona como grupo de punto para un orbifolio, pues se requiere que θ actúe de manera cristalográfica sobre la retícula Λ . En la figura 2.2, podemos apreciar dos ejemplos de grupos de punto, uno que actúa cristalográficamente sobre Λ y otro que no lo hace. Dependiendo del grupo puntual que da origen a nuestro orbifolio,

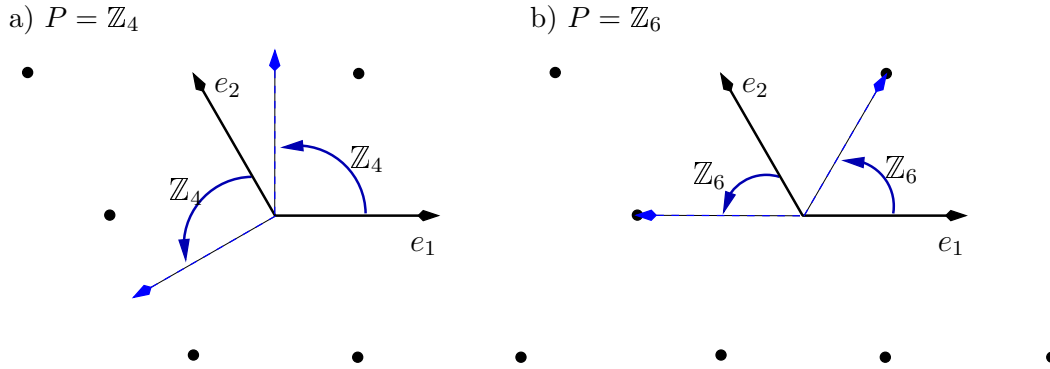


Figura 2.2: a) $P = \mathbb{Z}_4$ actúa cristalográficamente sobre Λ por lo que sí podríamos considerarlo como grupo de punto para la construcción de un orbifolio. b) en este caso el grupo $P = \mathbb{Z}_6$ no actúa cristalográficamente sobre Λ , por lo que la combinación de esa retícula y grupo de punto no pueden ser considerados para la construcción de un orbifolio.

éste se llamará abeliano o no abeliano. Aquí, nos restringiremos a los casos en que el grupo de punto P es abeliano, implicando que debe pertenecer al subálgebra de Cartan de $SO(6)$ asociada a X^k ($k = 3, \dots, 8$). Denotaremos a los generadores de esta subálgebra por M_{34}, M_{56}, M_{78} . Por consiguiente, definimos a la base compleja como

$$Z^1 \equiv \frac{1}{\sqrt{2}}(X^3 + iX^4), \quad (2.43)$$

$$Z^2 \equiv \frac{1}{\sqrt{2}}(X^5 + iX^6), \quad (2.44)$$

$$Z^3 \equiv \frac{1}{\sqrt{2}}(X^7 + iX^8). \quad (2.45)$$

El elemento de grupo θ actúa de manera diagonal y puede ser escrito como

$$\theta = \exp[2\pi i(v_1 M^{34} + v_2 M^{56} + v_3 M^{78})] \quad (2.46)$$

con $0 \leq |v_i| < 1$ ($i = 1, 2, 3$). La condición de que la ecuación (2.33) se satisfaga implica que [17]

$$\pm v_1 \pm v_2 \pm v_3 = 0, \quad (2.47)$$

esto se ve de que los eigenvalores de θ actuando sobre un espinor son $e^{i\pi(\pm v_1 \pm v_2 \pm v_3)}$.

En [20] se demuestra que el hecho de que θ actúe de manera cristalográfica sobre la retícula Λ que define al toro sumado a la condición (2.33) para el grupo de punto, se concluye que P debe ser \mathbb{Z}_N con $N = 3, 4, 6, 7, 8, 12$ o bien $\mathbb{Z}_N \times \mathbb{Z}_M$ con N múltiplo de M y $N = 2, 3, 4, 6$. Algunos de estos grupos tienen dos maneras inequivalentes en la elección de v_1, v_2, v_3 , es decir, dos formas de incrustar al grupo P contenido en $SU(3)$.

Grupo de punto	(v_1, v_2, v_3)
\mathbb{Z}_3	$\frac{1}{3}(1, 1, -2)$
\mathbb{Z}_4	$\frac{1}{4}(1, 1, -2)$
$\mathbb{Z}_6 - I$	$\frac{1}{6}(1, 1, -2)$
$\mathbb{Z}_6 - II$	$\frac{1}{6}(1, 2, -3)$
\mathbb{Z}_7	$\frac{1}{7}(1, 2, -3)$
$\mathbb{Z}_8 - I$	$\frac{1}{8}(1, 2, -3)$
$\mathbb{Z}_8 - II$	$\frac{1}{8}(1, 3, -4)$
$\mathbb{Z}_{12} - I$	$\frac{1}{12}(1, 4, -5)$
$\mathbb{Z}_{12} - II$	$\frac{1}{12}(1, 5, -6)$

Tabla 2.1: Generadores del grupo de punto para orbifolios $\mathbb{Z}_N \subset SU(3)$ con $\theta = \exp[2\pi i(v_1 M^{34} + v_2 M^{56} + v_3 M^{78})]$ [17]

Grupo de punto	(v_1, v_2, v_3)	(w_1, w_2, w_3)
$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\frac{1}{2}(1, 0, -1)$	$\frac{1}{2}(0, 1, -1)$
$\mathbb{Z}_3 \times \mathbb{Z}_3$	$\frac{1}{3}(1, 0, -1)$	$\frac{1}{3}(0, 1, -1)$
$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\frac{1}{2}(1, 0, -1)$	$\frac{1}{4}(0, 1, -1)$
$\mathbb{Z}_4 \times \mathbb{Z}_4$	$\frac{1}{4}(1, 0, -1)$	$\frac{1}{4}(0, 1, -1)$
$\mathbb{Z}_2 \times \mathbb{Z}_6 - I$	$\frac{1}{2}(1, 0, -1)$	$\frac{1}{6}(0, 1, -1)$
$\mathbb{Z}_2 \times \mathbb{Z}_6 - II$	$\frac{1}{2}(1, 0, -1)$	$\frac{1}{6}(1, 1, -2)$
$\mathbb{Z}_3 \times \mathbb{Z}_6$	$\frac{1}{3}(1, 0, -1)$	$\frac{1}{6}(0, 1, -1)$
$\mathbb{Z}_6 \times \mathbb{Z}_6$	$\frac{1}{6}(1, 0, -1)$	$\frac{1}{6}(0, 1, -1)$

Tabla 2.2: Generadores del grupo de punto para orbifolios $\mathbb{Z}_N \times \mathbb{Z}_M \subset SU(3)$ con $\theta = \exp[2\pi i(v_1 M^{34} + v_2 M^{56} + v_3 M^{78})]$ y $\phi = \exp[2\pi i(w_1 M^{34} + w_2 M^{56} + w_3 M^{78})]$ [17]

Los grupos \mathbb{Z}_N también son llamados grupos cíclicos de orden N , pues el grupo es generado por

un elemento llamado generador θ tal que $\mathbb{Z}_N = \{\theta^k | k \in \{0, 1, \dots, N-1\}\}$. Se puede demostrar que el producto directo de dos grupos cíclicos \mathbb{Z}_N y \mathbb{Z}_M es cíclico si y sólo si M y N son primos entre sí; en tal caso, el grupo obtenido será isomorfo a \mathbb{Z}_{NM} [21]. Por ejemplo, \mathbb{Z}_{12} es isomorfo a $\mathbb{Z}_3 \times \mathbb{Z}_4$, pero no a $\mathbb{Z}_6 \times \mathbb{Z}_2$. Otro ejemplo está dado por el grupo \mathbb{Z}_6 , el cual es isomorfo a $\mathbb{Z}_2 \times \mathbb{Z}_3$. El grupo \mathbb{Z}_3 no es isomorfo al producto directo que grupos cíclicos, pues 3 es un número primo.

Por el momento no hemos puesto alguna restricción a la elección de la retícula Λ que define a los toros espaciales, pues para los modos izquierdos ya hemos construido nuestra elección para $E_8 \times E_8$. Sin embargo, más adelante veremos que si al grupo espacial S de los toros espaciales se le exige estar incrustado dentro del grupo $E_8 \times E_8$, entonces las propiedades del orbifold dependen de la elección de la retícula, y además, sólo existen pocas opciones dependiendo del grupo de punto con el que se haya elegido trabajar.

2.3. Compactificación en orbifolios

La existencia del grupo de punto P para los orbifolios agregará formas en las que podemos imponer condiciones de frontera para la cuerda cerrada con respecto al método de compactificación en toros. Nos concentraremos a partir de aquí en grupos puntuales de la forma \mathbb{Z}_N . Un grupo discreto de este estilo es generado por un elemento $\theta \in \mathbb{Z}_N$, de tal manera que cualquier otro elemento del grupo es de la forma θ^n ($0 \leq n < N$). La identificación (2.37) en términos de las coordenadas espaciales X^k ($k = 3, \dots, 8$) implica que las condiciones de frontera para la cuerda cerrada están dadas por

$$\mathbf{X}(\tau, \sigma) = (\theta^n, \mathbf{v})\mathbf{X}(\theta, 0) = \theta^n \mathbf{X}(\tau, 0) + \mathbf{1}. \quad (2.48)$$

El sector *no torcido* ($n = 0$) corresponde justamente a la compactificación toroidal discutida anteriormente. Vemos ahora que para los casos en que $n \neq 0$ tenemos *sectores torcidos* adicionales que dará origen a estados físicos que no estaban presentes en el caso de la compactificación en toros. Para el sector torcido $n = 1$ debemos notar que las condiciones de frontera (2.48) modifican la expansión en modos de las coordenadas de la hoja de mundo, que en la base compleja definida por (2.43)-(2.45) se expande como

$$Z^\alpha = z_f^\alpha + \frac{i}{2} \sum_{n \neq 0} \left[\frac{1}{n + v_\alpha} \beta_{n+v_\alpha}^\alpha e^{-2i(n+v_\alpha)(\tau-\sigma)} + \frac{1}{n - v_\alpha} \tilde{\beta}_{n-v_\alpha}^\alpha e^{-2i(n-v_\alpha)(\tau+\sigma)} \right], \quad \alpha = 1, 2, 3. \quad (2.49)$$

Aquí, z_f^α es un punto fijo complejo, construido a partir de los puntos fijos en sus coordenadas reales. De la ecuación (2.49) podemos notar que la especificación completa de un sector torcido requiere no sólo del elemento del grupo de punto θ , si no también del punto fijo particular que aparece en la parte cero de la expansión en modos para las coordenadas de la hoja de mundo. Esto es completamente análogo para sectores torcidos con $n > 1$.

La expansión en modos conjugada compleja es

$$\bar{Z}_\alpha = \bar{z}_f^\alpha + \frac{i}{2} \sum_{n \neq 0} \left[\frac{1}{n - v_\alpha} \bar{\beta}_{n-v_\alpha}^\alpha e^{-2i(n-v_\alpha)(\tau-\sigma)} + \frac{1}{n + v_\alpha} \bar{\beta}_{n+v_\alpha}^\alpha e^{-2i(n+v_\alpha)(\tau+\sigma)} \right], \quad \alpha = 1, 2, 3, \quad (2.50)$$

y los operadores $\beta_{n+v_\alpha}^\alpha$, $\tilde{\beta}_{n-v_\alpha}^\alpha$, $\bar{\beta}_{n-v_\alpha}^\alpha$, $\tilde{\beta}_{n+v_\alpha}^\alpha$ que aparecen en las expansiones (2.49) y (2.50) cumplen con las relaciones de conmutación

$$[\beta_{n+v_\alpha}^\alpha, \bar{\beta}_{n-v_\alpha}^\alpha] = \delta^{\alpha\gamma} (n + v_\alpha) \delta_{m+n,0}, \quad (2.51)$$

$$[\tilde{\beta}_{n-v_\alpha}^\alpha, \tilde{\beta}_{n+v_\alpha}^\alpha] = \delta^{\alpha\gamma} (n - v_\alpha) \delta_{m+n,0}. \quad (2.52)$$

Por tanto, β_{n+v} con $n + v > 0$ son proporcionales a los operadores de aniquilación y $\bar{\beta}_{-n-v}$ se asocian a los operadores de creación. Análogamente, β_{n+v} con $n + v < 0$ son operadores de creación y $\bar{\beta}_{-n-v}$ son operadores de aniquilación. De manera similar, se puede concluir para $\tilde{\beta}_{n-v}$ y $\tilde{\beta}_{n+v}$.

El grupo puntual actúa de manera similar sobre los modos derechos fermiónicos, de tal manera que las condiciones de frontera en el primer sector torcido $n = 1$ se ven modificadas por

$$\Psi_R^\alpha(\tau - \sigma - \pi) = e^{2\pi i v_\alpha} \Psi_R^\alpha(\tau - \sigma) \quad \text{R}, \quad (2.53)$$

$$\Psi_R^\alpha(\tau - \sigma - \pi) = -e^{2\pi i v_\alpha} \Psi_R^\alpha(\tau - \sigma) \quad \text{NS}, \quad (2.54)$$

donde los modos fermiónicos complejos Ψ_R^α ($\alpha = 1, 2, 3$) se construyen a partir de los modos reales Ψ_R^k . Así, la expansión en modos para los grados de libertad fermiónicos está dada por

$$\begin{aligned} \Psi_R^\alpha(\tau - \sigma) &= \sum_n e_{n+v_\alpha}^\alpha e^{-2i(n+v_\alpha)(\tau-\sigma)} & \text{R} \\ &= \sum_r c_{r+v_\alpha}^\alpha e^{-2i(r+v_\alpha)(\tau-\sigma)} & \text{NS}, \end{aligned} \quad (2.55)$$

y

$$\begin{aligned} \bar{\Psi}_R^\alpha(\tau - \sigma) &= \sum_r \bar{e}_{n-v_\alpha}^\alpha e^{-2i(n-v_\alpha)(\tau-\sigma)} & \text{R} \\ &= \sum_r \bar{c}_{n-v_\alpha}^\alpha e^{-2i(r-v_\alpha)(\tau-\sigma)} & \text{NS}, \end{aligned} \quad (2.56)$$

donde

$$\{e_{n+v_\alpha}^\alpha, \bar{c}_{m-v_\alpha}^\beta\} = \delta^{\alpha\beta} \delta_{m+n,0}, \quad (2.57)$$

$$\{c_{r+v_\alpha}^\alpha, \bar{c}_{s-v_\beta}^\beta\} = \delta^{\alpha\beta} \delta_{r+s,0}. \quad (2.58)$$

El grupo de espacio S debe ser incrustado en los grados de libertad de norma, y veremos más adelante que esto, en general, es obligatorio. Lo anterior quiere decir que queremos mapear un

elemento arbitrario del grupo de espacio (θ, \mathbf{v}) a un (Θ, \mathbf{V}) , donde Θ es un automorfismo de la retícula $E_8 \times E_8$ y \mathbf{V} es un vector de traslación en la misma.

Incrustar un elemento del grupo de espacio de la forma $(\theta, \mathbf{0})$ dentro de la retícula $E_8 \times E_8$ es mapeándolo a un elemento $(1, \pi V^I)$ [17], de tal forma que las condiciones de frontera para X_L^I en el primer sector torcido se transforman en

$$X_L^I(\tau + \sigma + \pi) = X_L^I(\tau + \sigma) + \pi V^I, \quad (2.59)$$

de modo que la expansión en modos para los modos izquierdos bosónicos de norma son de la forma

$$X_L^I = x_l^I + (p_L + V^I)(\tau + \sigma) + \frac{i}{2} \sum_n \frac{1}{n} \tilde{\alpha}_n^I e^{-2in(\tau + \sigma)}. \quad (2.60)$$

El procedimiento anterior de incrustación para el grupo de espacio en la retícula $E_8 \times E_8$ se llama *incrustación estándar* [15]. De la expansión anterior podemos ver que el efecto del torcimiento por θ sobre el momento p_L^I es que se ve alterado por una traslación V^I , además

$$V^I = (v_1, v_2, v_3, 0^5)(0^8), \quad (2.61)$$

donde v_α , ($\alpha = 1, 2, 3$) son los torcimientos de las 3 coordenadas complejas compactificadas. Ahora, en el orbifolio \mathbb{Z}_N , el vector V^I está constraído a que NV^I viva en la retícula raíz de $E_8 \times E_8$, es decir,

$$NV^I \in \Lambda_{E_8 \times E_8}. \quad (2.62)$$

Las condiciones de (2.47) sobre las fases v_α , ($\alpha = 1, 2, 3$) asegura que la anterior constricción sobre el vector de traslación V^I siempre se cumpla en la incrustación estándar. Ahora bien, hasta el momento sólo mencionamos la forma en que se incrusta a los elementos de grupo de la forma $(\theta, 0)$ y ya vimos que estos son incrustados por medio de traslaciones en la retícula de raíces del grupo de norma como $(0, \pi V^I)$. Aún necesitamos analizar la incrustación de elementos de la forma $(0, \mathbf{v})$, que son traslaciones en la retícula del toro espacial y que llamaremos *líneas de Wilson*. Este último caso será discutido más adelante.

Las expansiones en modos que ya describimos hace unos momentos tiene consecuencias en los cálculos de los generadores del álgebra de Virasoro $L_m \tilde{L}_n$, y en particular para las expresiones de $L_0 \tilde{L}_0$ que son los que proporcionan las fórmulas de masa. Estas modificaciones ahora dan lugar a modos fraccionarios y éstos modifican el cálculo de las constantes de orden normal. La fórmula de masa modificada para los modos derechos modificada y para el primer sector torcido tiene la estructura general [17]

$$\frac{1}{4} M_R^2 = N_B + N_F(\beta) - a_B - a_F(\beta), \quad (2.63)$$

donde como ya habíamos establecido $\beta = R, NS$. Aquí también, N_B es el operador de número bosónico y N_F el correspondiente fermiónico. Las constantes de modo normal son

$$a_B = \frac{1}{3} - \frac{1}{2} \sum_{\alpha=1}^3 |v_\alpha|(1 - |v_\alpha|), \quad (2.64)$$

$$a_F(R) = -\frac{1}{3} + \frac{1}{2} \sum_{\alpha=1}^3 |v_\alpha|(1 - |v_\alpha|), \quad a_F(NS) = -\frac{5}{24} + \frac{1}{2} \sum_{\alpha=1}^3 \left| v_\alpha + \frac{1}{2} \right| \left(1 - \left| v_\alpha + \frac{1}{2} \right| \right). \quad (2.65)$$

Es importante recalcar que no hay contribución de momento para m_R , pues como ya vimos, p_R es nulo para los sectores torcidos, es decir, cuando $v_\alpha \neq 0$.

La ecuación de masa para los modos izquierdos en el primer sector torcido es

$$\frac{1}{4} M_L^2 = \tilde{N} + \frac{1}{2} (p_L^I + V^I)^2 - \tilde{a}, \quad (2.66)$$

donde \tilde{N} tiene la misma forma que N_B pero con los operadores reemplazados por las versiones izquierdas. La constante de orden normal correspondiente sería

$$\tilde{a} = 1 - \frac{1}{2} \sum_{\alpha} |v_\alpha|(1 - |v_\alpha|). \quad (2.67)$$

Entre las ventajas de la compactificación en orbifolios sobre la misma en toros está el hecho de que aseguramos la remoción de los gravitinos no deseados (en el sector no torcido), asegurando una sola supersimetría $\mathcal{N} = 1$. También tenemos que gracias a la invariancia del grupo de punto que restringe aún más al toro, el grupo de norma se ve reducido cuando el grupo de punto es incrustado en los grados de libertad de norma. Cuando tomamos la incrustación estándar podemos asegurar algunas cuestiones en la forma que el grupo de norma se ve reducido. La condición (2.33) asegura que el grupo de punto es un subgrupo discreto de $SU(3)$ de uno de los E_8 , pues

$$E_6 \times SU(3) \subset E_8, \quad (2.68)$$

por lo que la simetría de norma que sobrevive incluirá siempre a $E_6 \times E_8$ [17]. El rango del grupo de norma no se ve afectado por la incrustación, pues los bosones de norma asociados con el subálgebra de Cartan son invariantes bajo la acción del grupo de punto. Estos están dados por

$$b_{-1/2}^i |0\rangle_R \tilde{\alpha}_{-1}^I |0\rangle_L. \quad (2.69)$$

Hemos observado también que el estado derecho es invariante bajo la acción de P , y su incrustación como un vector de traslación V en la retícula $E_8 \times E_8$, implica que los modos $\tilde{\alpha}_n^I$ también son invariantes. Esto implica que la incrustación estándar nos reduce a un grupo de norma al menos $E_6 \times U(1)^2 \times E_8$. Los bosones de norma cargados de $E_8 \times E_8$ están dados por

$$b_{-1/2}^i |0\rangle_R |p_L\rangle^I, \quad (2.70)$$

con $(p_L^I)^2 = 2$, y los bosones de norma sobrevivientes una vez hecha la incrustación estándar satisfarán la condición [17]

$$p_L^I V^I = 0 \text{ mod } 1. \quad (2.71)$$

El hecho de haber incrustado elementos del grupo de espacio de la forma $(\theta, 0)$ quiere decir que sólo hemos incrustado al grupo de punto en el grupo de norma. Se puede romper aún más al grupo de norma incrustando al grupo de espacio completo. Es decir, no sólo incrustando elementos del grupo de punto como vectores de traslación en los grados de libertad bosónico $E_8 \times E_8$, sino también los vectores base de la retícula del toro que conforma al orbifolio. Este proceso no sólo provocará una ruptura del grupo de norma, si no que también modificará el contenido de materia, obteniendo así modelos de 3 generaciones.

Consideremos un sector torcido con condiciones de frontera torcidas por el elemento de grupo de espacio (θ, v) , donde θ es un elemento del grupo de punto con v vector en la retícula del toro en cuestión,

$$v = \sum r_\rho e_\rho, \quad (2.72)$$

donde r_ρ son coeficientes enteros y e_ρ son los vectores base de la retícula para el toro 6-dimensional. Para incrustar el grupo de espacio en el grupo de norma, el elemento del grupo de punto θ debe ser incrustado como el vector de traslación πV^I , como ya vimos, y el vector base de la retícula r_ρ como el vector de traslación πa_ρ^I . Para asegurarnos de que, en efecto, tenemos una incrustación, debemos corroborar que sí tenemos un homomorfismo. Así que debemos tener una imagen correcta para el producto de dos elementos de espacio (θ_1, l_1) y (θ_2, l_2) , de tal manera que

$$(\theta_1, l_1)(\theta_2, l_2) = (\theta_1\theta_2, l_1 + \theta_1 l_2). \quad (2.73)$$

Para un grupo de punto \mathbb{Z}_N generado por θ ,

$$(\theta, l)^N = (I, \mathbf{0}). \quad (2.74)$$

En consecuencia, debemos requerir que

$$N(V^I + r_\rho a_\rho^I) \in \Lambda_{E_8 \times E'_8}, \quad (2.75)$$

lo que implica que

$$NV^I \in \Lambda_{E_8 \times E_8} \quad \text{y} \quad Na_\rho^I \in \Lambda_{E_8 \times E'_8}, \quad (2.76)$$

así que por las propiedad de la retícula de raíces, automáticamente obtenemos que

$$N \sum_{i \in E_8} V^I = 0 \pmod{2} \quad N \sum_{I \in E'_8} V^I = 0 \pmod{2} \quad (2.77)$$

y

$$N \sum_{i \in E_8} a_\rho^I = 0 \pmod{2} \quad N \sum_{I \in E'_8} a_\rho^I = 0 \pmod{2}. \quad (2.78)$$

Adicionalmente a las condiciones anteriormente mencionadas, la incrustación del grupo de espacio debe ser elegido de tal manera que la propiedad fundamental de invariancia modular de la teoría

se preserve. La invariancia modular es una constricción que asegura que la teoría de campos compactificada está libre de anomalías. Términos de la función de partición a un lazo de orbifolios no abelianos adquieren, en general, una fase no trivial bajo transformaciones modulares. Si demandamos que la función de partición sea invariante ante estas transformaciones (invariante modular) aseguramos que la teoría resultante esté libre de anomalías. Así que debemos asegurarnos que la fase no trivial que aparece por las transformaciones modulares se anule.

Para asegurar que el término a un lazo de la función de partición se transforme hacia sí mismo, sin ningún factor de fase, requerimos que:

$$N \left(\sum_{i=1}^3 v^i (1 - v^i) - \sum_{I=1}^{16} \tilde{v}^I (1 - \tilde{v}^I) \right) = 0 \pmod{2}, \quad (2.79)$$

donde v^i son las torceduras en los modos derechos complejos fermiónicos, y $\tilde{v}^I = V^I + r_\rho a_\rho^I$ son las torceduras sobre los modos izquierdos fermiónicos de norma $E_8 \times E_8'$ para el sector torcido θ del orbifolio con líneas de Wilson. La condición (2.74) junto con la imposición de que

$$N \sum_{i=1}^3 v^i = 0 \pmod{2}, \quad (2.80)$$

para que la acción del grupo de punto sea de orden N al actuar sobre la representación espinorial de $SO(8)$ permite que la ecuación (2.79) se simplifique a

$$N \left(\sum_{i=1}^3 (v^i)^2 - \sum_{I=1}^{16} (\tilde{v}^I)^2 \right) = 0 \pmod{2}, \quad (2.81)$$

con $n = 0, \dots, N - 1$ y $r_\rho = 0, \dots, N - 1$. En particular, las incrustaciones del grupo de punto en el grupo de norma consistentes con invariancia modular requieren satisfacer que

$$N \left(\sum_{I=1}^{16} (V^I)^2 - \sum_{i=1}^3 (v^i)^2 \right) \equiv N(V^2 - v^2) = 0 \pmod{2}, \quad (2.82)$$

y las líneas de Wilson consistentes con invariancia modular requieren satisfacer

$$N \sum I = 1^{16} (a_\rho^I)^2 \equiv N a_\rho^2 = 0 \pmod{2}, \quad (2.83)$$

$$N \sum_{I=1}^{16} a_\rho^I a_\sigma^I \equiv a_\rho \cdot a_\sigma = 0 \pmod{1}, \quad \rho \neq \sigma, \quad (2.84)$$

y

$$N \sum_{I=1}^{16} V^I a_\rho^I \equiv N V \cdot a_\rho = 0 \pmod{1}. \quad (2.85)$$

2.3.1. Espectro no masivo

Recordemos que la materia en los orbifolios heteróticos son descritos por cuerdas cerradas y admiten dos tipos de cuerdas cerradas: torcidas y no torcidas (ver Figura 2.3). Las cuerdas no torcidas cumplen con las condiciones de frontera

$$Z(\tau, \sigma + \pi) = Z(\tau, \sigma) + n_\alpha e_\alpha, \quad n_\alpha \in \mathbb{Z}. \quad (2.86)$$

Esto nos indica que dichas cuerdas ya son cerradas en el toro y son libres de propagarse en el espacio compacto. Por otro lado, las cuerdas torcidas son cerradas sólo después de identificar puntos por medio de la torsión θ , de acuerdo a sus condiciones de frontera

$$Z(\tau, \sigma + \pi) = \theta^m Z(\tau, \sigma) + n_\alpha e_\alpha, \quad m, n_\alpha \in \mathbb{Z}, \quad (2.87)$$

de donde apreciamos que dichas cuerdas están entonces atadas a los puntos fijos. Las cuerdas no torcidas serán agrupadas en el sector no torcido U . Cuerdas torcidas de orbifolios \mathbb{Z}_N tendrán $N - 1$ sectores torcidos T_m , $m = 1, \dots, N - 1$.

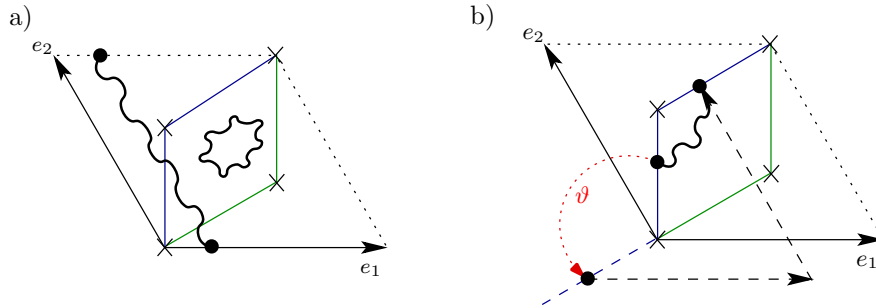


Figura 2.3: Cuerdas cerradas en una compactificación en orbifolios. a) Las cuerdas no torcidas son cerradas en el toro, mientras que b) cuerdas torcidas son cerradas sólo bajo la acción de la torcedura θ .

Se asocian diferentes estados de materia a diferentes elementos $g \in S$ correspondientes al sector no torcido y a puntos fijos. La cantidad de estos elementos es finita y se les suele llamar *elementos constructores*. Por otro lado, estados no torcidos son asociados al elemento constructor $g = (1, 0)$ y se pueden expresar como

$$|q\rangle_R \otimes \tilde{\alpha}_{-1}^x |p\rangle_L, \quad (2.88)$$

donde q es un pedo de $SO(8)$ y $p \in \Lambda$, $\tilde{\alpha}_{-1}^x$ representa la excitación de un oscilador en una dirección, ya sea del espacio de Minkowski ($\tilde{\alpha}_{-1}^\mu$) o del espacio compacto o de los grados de libertad de norma. Por otro lado, estados no masivos asociados a elementos constructores torcidos $g = (\theta^m, n_\alpha e_\alpha)$ (con $m = 1, \dots, N - 1$) son expresados en general como

$$|q_{sh}\rangle \otimes \tilde{\alpha} |p_{sh}\rangle \equiv |q + v_g\rangle_R \otimes \tilde{\alpha} |p + V_g\rangle_L, \quad (2.89)$$

donde $\tilde{\alpha}$ denota en este caso el producto de osciladores de la forma $\tilde{\alpha}_{-\eta^a}^a$ (o $\tilde{\alpha}_{-1+\eta^a}^{\bar{a}}$) en las direcciones complejas $a = 1, 2, 3$ (o de sus conjugados $\bar{a} = \bar{1}, \bar{2}, \bar{3}$), con $\eta^a = mv^a \bmod 1$, de tal manera que $0 < \eta^a \leq 1$. Aquí también hemos definido la torcedura local y los vectores de traslación locales asociados al elemento constructor g como

$$v_g = mv, \quad (2.90)$$

y,

$$V_g = mV + n_\alpha A_\alpha, \quad (2.91)$$

respectivamente. Consideremos ahora el conjunto de estados con condiciones de frontera

$$Z(\tau, \sigma + \pi) = gZ(\tau, \sigma), \quad (2.92)$$

para un elemento constructor g . Estos estados definen un espacio de Hilbert \mathcal{H}_g . Para asegurar la compatibilidad de los estados de dicho espacio con el orbifolio, dejamos que un elemento del grupo de espacio $h \in S$ tal que $[g, h] = 0$ actúe en las coordenadas que describen a las cuerdas. Las condiciones de frontera se convierten en $(hZ)(\tau, \sigma + \pi) = g(hZ)(\tau, \sigma)$, que indica que los estados descritos por las coordenadas hZ pertenecen al mismo espacio de Hilbert \mathcal{H}_g . En consecuencia, h debe actuar trivialmente en estados de \mathcal{H}_g ,

$$|q_{sh}\rangle_R \otimes \tilde{\alpha} |p_{sh}\rangle_L \rightarrow \Phi |q_{sh}\rangle_R \otimes \tilde{\alpha} |p_{sh}\rangle_L = |q_{sh}\rangle_R \otimes \tilde{\alpha} |p_{sh}\rangle_L, \quad (2.93)$$

donde la fase Φ depende de g y h . Estados que no cumplan la condición (2.93) deben ser excluidos.

Es conveniente entonces definir el *centralizador* \mathcal{Z}_g de un elemento constructor g como el conjunto de todos los elementos de espacio h que conmutan con g :

$$\mathcal{Z}_g = \{h \in S | [g, h] = 0\}. \quad (2.94)$$

El espectro de materia no masivo de las compactificaciones en orbifolios está formado entonces por estados no masivos que son invariantes bajo todos los elementos del centralizador.

Sector no torcido

En el sector no torcido U , la condición de emparejamiento para estados no masivos de los orbifolios coinciden con los de la cuerda heterótica no compactificada, es decir,

$$\frac{M_L^2}{2} = p^2 + \tilde{N} - 1 = 0 = \frac{1}{2}q^2 - \frac{1}{2} = \frac{M_R^2}{2}, \quad (2.95)$$

donde p es una raíz de $E_8 \times E_8$, q denota un vector de pesos de $SO(8)$ del oscilador derecho y \tilde{N} cuenta el número de excitaciones del oscilador. Como en la cuerda heterótica no compactificada, la

ecuación anterior tiene solución sólo si $q^2 = 1$ para los modos derechos y qu tanto $p^2 = 0$, $\tilde{N} = 1$ o $p^2 = 2$, $\tilde{N} = 0$ para los modos izquierdos. La diferencia nace de las proyecciones debidas al orbifolio.

Se puede demostrar que la fase adquirida por estados no masivos no torcidos bajo la acción de h está dada por

$$|q\rangle_R \otimes |p\rangle_L : \quad \Phi = e^{2\pi i[p \cdot V_h - q \cdot v_h]} \quad (2.96)$$

$$|q\rangle_R \otimes \tilde{\alpha}_{-1}^x |0\rangle_L : \quad \Phi = e^{2\pi i[-q \cdot v_h + (\delta_{x,a} - \delta_{x,\bar{a}})v_h^a]}, \quad (2.97)$$

donde x denota la dirección en la cual actúa la excitación del oscilador. El elemento constructor del sector no torcido es $g = (1, 0)$ y, por tanto, el centralizador asociado contiene todos los elementos del grupo de espacio. El espectro del sector no torcido está compuesto por aquellos elementos de la cuerda heterótica no compactificada y que es invariante bajo todos los elementos del grupo de espacio. Ahora evaluamos el efecto de la proyección del orbifolio sobre los estados no masivos del espectro de la cuerda heterótica.

El **multiplete de supergravedad 10-dimensional** de la cuerda heterótica se divide en:

- un **gravitón 4-dimensional** $g^{\mu\nu}$, **dilatón** ϕ , **tensor antisimétrico** $B^{\mu\nu}$ y sus supercompañeros. Están dados por las componentes del grupo de espacio invariante de

$$|q\rangle_R \otimes \alpha_{-1}^\nu |0\rangle_L, \quad (2.98)$$

con

$$p = \begin{cases} \pm(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}) \\ \pm(1, 0, 0, 0) \end{cases}. \quad (2.99)$$

- algunos **módulos geométricos**

$$|0\rangle_R \otimes \tilde{\alpha}_{-1}^{a/\bar{a}} |0\rangle_L \quad (2.100)$$

satisfaciendo la condición de invariancia

$$q \cdot v_h \pm v_h^a = 0 \text{ mod } 1, \quad (2.101)$$

donde el signo relativo $-(+)$ está asociado a un oscilador que acarrea un índice holomórfico a (índice antiholomórfico \bar{a}). Estos estados son singuletes de norma, provenientes de las componentes compactas del gravitón 10D y el tensor antisimétrico (y sus superparejas). La acción de h en los 16 **bosones de norma no cargados 10D** deja invariantes sólo sus componentes 4D especificadas por

$$|q\rangle_R \otimes \tilde{\alpha}_{-1}^I |0\rangle_L \quad (2.102)$$

con el momento del modo derecho q también como en (2.99). Estos estados son los 16 generadores de Cartan del grupo de norma 4D \mathcal{G}_{4D} .

- los **bosones de norma cargados** (y los norminos) de \mathcal{G}_{4D} son aquellos estados que satisfacen $q \cdot v_h = 0 \pmod{1}$ y $p \cdot V_h = 0 \pmod{1}$. Los únicos momentos de los modos derechos que cumplen con dicha condición son los que tenemos en (2.99). La condición para el momento del modo izquierdo debe ser cumplida por cualquier h , entonces replanteado como

$$p \cdot V = 0 \pmod{1}, \quad (2.103)$$

$$p \cdot A_\alpha = 0 \pmod{1}, \quad \alpha = 1, \dots, 6, \quad (2.104)$$

donde A_α son líneas de Wilson. Estos estados forman la representación adjunta de \mathcal{G}_{4D} . Dado que no todos los 480 momentos izquierdos de los bosones 10D satisfacen las condiciones anteriores, la simetría de norma es rota;

- aquellos estados cuyas componentes izquierda o derecha se transforman no trivialmente y satisfacen la condición de invariancia

$$p \cdot V_h - q \cdot v_h = 0 \pmod{1} \quad (2.105)$$

constituyen la materia cargada no torcida de los modelos con orbifolios. Sus propiedades de la transformación de norma respecto a \mathcal{G}_{4D} dependen de su momento.

Sector torcido

Los modos zero de los sectores torcidos T_k están asociados con los elementos constructores g de los puntos fijos. Con el requerimiento de que los estados sean no masivos se tienen las siguientes condiciones para los momentos de los modos izquierdo y derecho.

$$\frac{1}{4}M_L^2 = \frac{1}{2}p_{sh}^2 + \tilde{N} - 1 + \delta_C = 0, \quad (2.106)$$

$$\frac{1}{4}M_R^2 = \frac{1}{2}q_{sh}^2 - \frac{1}{2} + \delta_C = 0, \quad (2.107)$$

donde δ_C corresponde a un cambio en la energía de punto cero relacionado con la aparición de los osciladores torcidos $\alpha_{-\eta^a}^a, \alpha_{-1+\eta^a}^a$. Esta cantidad está expresada por

$$\delta_C = \frac{1}{2} \sum_a \eta^a (1 - \eta^a), \quad (2.108)$$

con $\eta^a = v_g^a \pmod{1}$, tal que $0 < \eta^a \leq 1$. Para estados no masivos, se puede escribir el operador de número para modos torcidos como

$$\tilde{N} = \sum_{a=1}^3 \eta^a (\tilde{N}_g^a + \tilde{N}_g^{*a}). \quad (2.109)$$

Aquí, \tilde{N}_g^a y \tilde{N}_g^{*a} son números de modos enteros, contando respectivamente el número de excitaciones de las direcciones holomórficas a y \bar{a} . La transformación de los estados oscilantes izquierdos y derechos $|p_{sh}\rangle_L, |q_{sh}\rangle_R$ bajo la acción de un elemento arbitrario h del centralizados sería

$$\Phi = e^{2\pi i[p_{sh}\cdot V_h - q_{sh}\cdot v_h + (\tilde{N}_g - \tilde{N}_g^*)\cdot v_h]} \Phi_{vac}, \quad (2.110)$$

donde la fase del vacío

$$\Phi_{vac} = e^{2\pi i[-\frac{1}{2}(V_g\cdot V_h - v_g\cdot v_h)]} \quad (2.111)$$

aparece como consecuencia de las propiedades geométricas de las cuerdas torcidas.

2.3.2. Vectores de traslación y espectro locales

Consideremos un modelo de orbifolio arbitrario con grupo de punto \mathbb{Z}_N o $\mathbb{Z}_N \times \mathbb{Z}_M$. Localmente, en un punto fijo con elemento constructor g , los 480 bosones de norma del grupo de norma 10-dimensional \mathcal{G} se ven afectados por la acción del vector de desplazamiento V_g . Entonces el grupo de norma local \mathcal{G}_g y el espectro de materia local coinciden con el grupo de norma 4-dimensional y el espectro en un punto fijo arbitrario de un modelo de orbifolio con vector de desplazamiento global $V = V_g$ (y sin líneas de Wilson). Las raíces del grupo de norma local serán aquellas raíces p del grupo de norma 10-dimensional \mathcal{G} que satisfagan

$$p \cdot V_g = 0 \text{ mod } 1 \quad p^2 = 2. \quad (2.112)$$

Claramente, los estados de materia transforman localmente bajo \mathcal{G}_g en lugar de transformarse bajo el grupo de norma 4-dimensional. Esto restringe el número de grupos de norma locales (y el espectro local) a ser el número de (inequivalentes) vectores de desplazamiento admisibles del orbifolio en cuestión.

La intersección de los grupos de norma locales de los diferentes puntos fijos en un modelo específico generan el grupo de norma 4D global \mathcal{G}_{4D} :

$$\mathcal{G}_{4D} = \mathcal{G}_{g_1} \cap \mathcal{G}_{g_2} \cap \mathcal{G}_{g_3} \cap \dots, \quad (2.113)$$

de donde se sigue que la simetría local de grupo es, en general, más grande que \mathcal{G}_{4D} , es decir,

$$\mathcal{G}_{4D} \subset \mathcal{G}_{g_i}, \quad i = 1, 2, 3, \dots \quad (2.114)$$

Desde una perspectiva global en 4 dimensiones, los estados de materia locales forman representaciones de \mathcal{G}_{4D} [22].

2.3.3. Clasificación de orbifolios simétricos toroidales

Lo que se desarrolla a continuación está basado principalmente en [19]. Las propiedades físicas de la teoría de cuerdas heterótica con compactificaciones en orbifolios dependen directamente en la elección de grupo de espacio. Dichas propiedades físicas son compartidas y comunes a conjuntos de grupos de espacio que pueden ser relacionados por algunas propiedades matemáticas. Por lo que se pueden definir clases de equivalencia de grupos de espacio. Existen tres tipos de clases de equivalencia que dependen de propiedades físicas y matemáticas de los grupos de espacio S . Las tres clases son:

1. Clases- \mathbb{Q} : determinan el grupo de punto P contenido en S y por ende el número de supersimetrías 4D y el número de módulos geométricos;
2. Clases- \mathbb{Z} : determinan la retícula Λ de S y por ende la naturaleza de los módulos geométricos;
3. Clases afines: determinan el grupo de sabor y la naturaleza de la ruptura de la simetría de norma (es decir, simetría de norma local y no local).

Cada clase- \mathbb{Q} contiene muchas clases- \mathbb{Z} y cada clase- \mathbb{Z} contiene bastantes clases afines.

$$\text{Clases Afines} \subset \text{Clases} - \mathbb{Z} \subset \text{Clases} - \mathbb{Q} \quad (2.115)$$

En otras palabras, para cada grupo de punto puede haber una gran cantidad de retículas inequivalentes y para cada retícula puede haber una gran cantidad de grupos orbifoldales (con o sin roto-traslaciones). La clasificación de orbifolios que se presentará aquí difiere de la hecha por Bailin y Love en [17].

Clases afines de grupos de espacio

Dos grupos de espacio S_1 y S_2 de grado D pertenecen a la misma clase afín (es decir, $S_1 \sim S_2$) si existe un mapeo afín $f : \mathbb{R}^D \rightarrow \mathbb{R}^D$ tal que

$$f^{-1}S_1f = S_2 \quad (2.116)$$

Un mapeo afín $f = (A, t)$ en \mathbb{R}^D consiste de una traslación t y un mapeo lineal A , es decir, permite reescalamientos y rotaciones. Esta definición permite distinguir entre espacios de grupo que realmente describan diferentes geometrías y aquellos que son los que ya conocemos visto desde un ángulo o distancia diferentes. Así, para un espacio de grupo representativo de una clase afín, una transformación afín no trivial que deja al grupo de punto invariante (es decir, $A^{-1}PA = P$) corresponde a un cambio de los datos geométricos. En el contexto de las compactificaciones de supercuerdas, esto corresponde a un cambio de los valores de los módulos geométricos. Así, estaremos interesados únicamente de un grupo de espacio representativo en cada una de las clases afines. Resulta que dada una dimensión D , existen sólo un número finito de clases afines de grupos de espacio. En la clasificación que se lleva a cabo en [19] sólo se concentran en el caso 6-dimensional.

Clases- \mathbb{Z} de grupos de espacio

Podemos refinar esta agrupación de grupos de espacio, agrupando las clases afines de acuerdo a propiedades comunes de sus grupos de punto correspondientes. Como ya vimos, los elementos del grupo de punto pueden escribirse en la base de la retícula como elementos de $GL(D, \mathbb{Z})$. Tomemos dos grupos de espacio S_1 y S_2 . Para $i = 1, 2$, el espacio de grupo S_i contiene una retícula Λ_i y su grupo de punto en la base de la retícula se denota por P_i , es decir, $P_i \subset GL(D, \mathbb{Z})$. Entonces, los dos grupos de espacio pertenecen a la misma clase- \mathbb{Z} si existe una matriz unimodular U (es decir, $U \in GL(D, \mathbb{Z})$) tal que

$$U^{-1}P_1U = P_2. \quad (2.117)$$

Es decir, si dos grupos de punto están relacionados por un cambio en la base de la retícula (utilizando U), el espacio de grupo pertenece a la misma clase- \mathbb{Z} , tienen la misma forma del espacio y, físicamente, tienen el mismo número de módulos y éstos también tienen la misma naturaleza geométrica. Sin embargo, los grupos de espacio de la misma clase \mathbb{Z} no son necesariamente equivalentes por la posible presencia de roto-traslaciones, en otras palabras, grupos de espacio en la misma clase- \mathbb{Z} pueden pertenecer a diferentes clases afines y por ende ser inequivalentes.

Clases- \mathbb{Q} de grupos de espacio

Tomemos dos grupos de espacio S_1 y S_2 . Para $i = 1, 2$, el grupo de punto en la base de la retícula asociada al grupo de espacio S_i denotado por P_i , es decir, $P_i \subset GL(D, \mathbb{Z})$. Los dos grupos de espacio pertenecen a la misma clase- \mathbb{Q} si existe una matriz $V \in GL(D, \mathbb{Z})$ tal que

$$V^{-1}P_1V = P_2. \quad (2.118)$$

Claramente, si dos grupos de espacio pertenecen a la misma clase- \mathbb{Z} también pertenecen a la misma clase- \mathbb{Q} . A diferencia de las clases- \mathbb{Z} , las clases- \mathbb{Q} no distinguen entre retículas inequivalentes. Sin embargo, si dos grupos de espacio pertenecen a la misma clase- \mathbb{Q} , las relaciones de conmutación y los órdenes de los correspondientes grupos de punto son iguales. Así, son grupos de punto isomorfos y cristalográficos. Estos grupos también poseen la misma forma de espacio, es decir, los mismos números para los módulos. Lo importante físicamente es que todos los grupos de espacio en la misma clase- \mathbb{Q} comparten un grupo de holonomía común [19]. Esto permite identificar las características que llevan a una $\mathcal{N} = 1$ supersimetría en 4D. En particular, para determinar el número de generadores de supersimetría, basta con considerar sólo un elemento representante de cada clase- \mathbb{Q} .

Clasificación de grupos de espacio

A continuación se describe brevemente la estrategia de clasificación de los grupos de espacio para orbifolios presentada en [19]. La cantidad de supersimetría residual que posee la teoría efectiva 4D

está relacionada con el grupo de holonomía del espacio compacto. En el contexto de los orbifolios, se relaciona el grupo de holonomía con el grupo de punto. Las compactificaciones en orbifolios perservan supersimetría 4D si el grupo de punto es un subgrupo discreto de $SU(3)$. Entonces la clasificación de la compactificación en orbifolios empieza con la identificación de todas las clases- \mathbb{Q} (es decir, los grupos de punto) que son subgrupos de $SU(3)$. Posteriormente, por cada clase- \mathbb{Q} identificamos todas las clases- \mathbb{Z} (es decir, las retículas) y finalmente para cada clase- \mathbb{Z} construimos todas las clases afines (es decir, roto-traslaciones). El proceso de clasificación es como sigue:

1. Escoger una clase- \mathbb{Q} y encontrar un P representativo de la misma.
2. Corroborar que P es un subgrupo de $SO(6)$ en lugar de $O(6)$.
3. Verificar que P es un subgrupo de $SU(3)$.
4. Encontrar todas las posibles clases- \mathbb{Z} dentro de esa clase- \mathbb{Q} .
5. Encontrar todas las clases afines posibles dentro de cada una de esas clases- \mathbb{Z} .

Existe un catálogo para todas las clases afines posibles hasta 6 dimensiones clasificadas dentro de clases- \mathbb{Z} y \mathbb{Q} . En [19] utilizan un software especializado llamado CARAT que permite tener acceso de dicho catálogo y en dicho trabajo se puede observar la clasificación completa de orbifolios heteróticos toroidales abelianos y no abelianos.

En este trabajo sólo mencionaremos dos orbifolios específicos, los correspondientes a las clases- \mathbb{Q} \mathbb{Z}_3 y \mathbb{Z}_6 -II cuyas características específicas profundizaremos más adelante. Debido a que en este trabajo no consideraremos roto-traslaciones, tenemos entonces que sólo hay una clase afín para cada uno de los orbifolios \mathbb{Z}_N .

clase- \mathbb{Q}	clase- \mathbb{Z}	correspondiente retícula
\mathbb{Z}_3	1	$SU(3)^3$
\mathbb{Z}_6 -II	1	$G_2 \times SU(3) \times SU(2)^2$
	2	-
	3	$SO(7) \times SU(3) \times SU(2)$
	4	$SU(6) \times SU(2)$

Tabla 2.3: Conexión en la clasificación de las clases- \mathbb{Q} de los grupos de espacio \mathbb{Z}_3 y \mathbb{Z}_6 -II. Los casos que no se conocían previamente están indicados con una raya. [19].

2.4. Orbifolio \mathbb{Z}_3

Desarrollemos como ejemplo el más sencillo de los orbifolios, que es el que tiene como grupo de punto a \mathbb{Z}_3 . En el orbifolio en cuestión \mathcal{O} , los puntos en el espacio se identifican entre sí por la relación

$$X \sim \Theta X + n_\alpha e_\alpha \quad (2.119)$$

donde $\Theta \in P$, es decir, tiene la forma $\Theta = \theta^k$ (con $0 \leq k < 3$) para $P = \mathbb{Z}_3$ y θ es el elemento generador de dicho grupo; además $e_\alpha \in \mathbb{R}^6$ ($\alpha = 1, \dots, 6$) es una base de \mathbb{R}^6 . La combinación lineal $n_\alpha e_\alpha = l \in \Lambda$ es un vector arbitrario en la retícula que construye al toro $T^6 = \frac{\mathbb{R}^6}{\Lambda}$.

Al combinar la acción del grupo de punto \mathbb{Z}_3 con las traslaciones dentro por vectores en la retícula Λ obtenemos el llamado *grupo espacial* S , y sus elementos se pueden definir de la siguiente manera

$$g = (\Theta, n_\alpha e_\alpha) \quad g \in S. \quad (2.120)$$

El grupo de espacio se define formalmente como el producto semidirecto $S = P \rtimes \Gamma$, lo que implica

$$g_1 g_2 = (\Theta_1, n_\alpha e_\alpha)(\Theta_2, m_\alpha e_\alpha) = (\Theta_1 \Theta_2, n_\alpha e_\alpha + m_\alpha \Theta_1 e_\alpha) \in S \quad (2.121)$$

$$g^{-1} = (\Theta, n_\alpha e_\alpha)^{-1} = (\Theta^{-1}, -n_\alpha \Theta^{-1} e_\alpha) \quad (2.122)$$

para $g, g_1, g_2 \in S$. De esta forma, el orbifolio puede ser redefinido como

$$\mathbb{O} = \mathbb{R}^6 / S \quad (2.123)$$

No todos los elementos del grupo de espacio S describen distintas acciones en el orbifolio. Los elementos equivalentes del grupo espacial se pueden agrupar en *clases de equivalencia* de la manera siguiente

$$[g] = \{hgh^{-1} | h \in S\} \quad (2.124)$$

donde todos los elementos hgh^{-1} para h arbitraria son llamados conjugados a g .

Puntos fijos

El grupo de espacio no actúa libremente en \mathbb{R}^D , es decir, existen algunos *puntos fijos*

$$X_f = g_f X_f = \Theta_f X_f + n_\alpha^f e_\alpha. \quad (2.125)$$

Estos puntos son invariantes por rotaciones $\Theta \in P$ salvo una traslación en la retícula Γ . Ya vimos anteriormente que si el grupo de punto es \mathbb{Z}_3 la retícula está constriñida a ser $\Gamma = SU(3)^3 = \Gamma_{SU(3)^3}$.

Un orbifolio es plano en todas partes, a excepción de los puntos fijos, que son singularidades de curvatura. Esto se puede verificar formalmente observando que el grupo local de holonomías en

las singularidades es no trivial. La ecuación 2.125 nos indica que para cada punto fijo X_f existe un único elemento del grupo espacial $g_f \in S$ que lo caracteriza. Refiriéndose a los puntos fijos, es preferible usar g_f en lugar de X_f , la clase de equivalencia de g_f , $[g_f]$, denota un conjunto infinito de puntos fijos equivalentes en \mathbb{R}^6

La construcción explícita del orbifolio con grupo de punto \mathbb{Z}_3 que se presenta a continuación está basada en [23]. Éste se construye a partir de un espacio Euclideo 6-dimensional \mathbb{R}^6 . Primero se define una base de vectores de dicho espacio e_1, \dots, e_6 tales que satisfagan

$$e_i^2 = e_{i+1}^2 = 2R_i^2, \quad e_i \cdot e_{i+1} = -1R_i^2, \quad i = 1, 3, 5 \quad (2.126)$$

con un vector $x \in \mathbb{R}^6$ de componentes reales:

$$x = \sum_{i=1}^6 x^i e_i, \quad x^i \in \mathbb{R} \quad \forall i = 1, \dots, 6 \quad (2.127)$$

Cada uno de los tres pares e_i, e_{i+1} ($i = 1, 3, 5$) define un subespacio bidimensional, que se puede visualizar como un plano complejo. El par i -ésimo también define una retícula raíz bidimensional $SU(3)$, obtenida del conjunto de todas las combinaciones lineales de la forma $n_i e_i + n_{i+1} e_{i+1}$ con n_i, n_{i+1} enteros. La retícula raíz $\Lambda_{SU(3)^3}$ es generada por todas las combinaciones lineales de los vectores base e_1, \dots, e_6 con coeficientes enteros, es decir,

$$\Lambda_{SU(3)^3} = \left\{ \sum_{i=1}^3 l^i e_i \mid l^i \in \mathbb{Z} \right\} \quad (2.128)$$

Notemos que el real R_i , llamado radio, no está fijo, tampoco los ángulos entre los vectores base como se definieron en la ecuación 2.126. Estos parámetros definen completamente las celdas unitarias de la retícula $\Lambda_{SU(3)^3}$ y están codificadas en los T-módulos T^{ij} . Salvo convenciones de normalización en los T^i y B_{ij} , la diagonal de los T-módulos están definidos como

$$T^i = \sqrt{\det G^{(i)}} + iB_{i,i+1}, \quad i = 1, 3, 5 \quad (2.129)$$

donde $G^{(i)}$ es la métrica del i -ésimo plano complejo

$$G^{(i)} = \begin{pmatrix} e_i \cdot e_i & e_i \cdot e_{i+1} \\ e_{i+1} \cdot e_i & e_{i+1} \cdot e_{i+1} \end{pmatrix} = R_i^2 \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad (2.130)$$

Las traslaciones en \mathbb{R}^6 por elementos en $\Lambda_{SU(3)^3}$,

$$x \rightarrow x + l, \quad l \in \Lambda_{SU(3)^3}, \forall x \in \mathbb{R}^6, \quad (2.131)$$

forman el grupo de la retícula. Una rotación θ en \mathbb{R}^6 es define completamente por la acción sobre los vectores base:

$$\theta \cdot e_i = e_{i+1}, \quad \theta \cdot e_{i+1} = -e_i - e_{i+1}, \quad i = 1, 3, 5 \quad (2.132)$$

Generalmente, θ es conocido como el operador de *torsión* del orbifolio. Es fácil corroborar que $\theta^3 = 1$. El operador de torsión θ genera el *grupo puntual* que ya habíamos introducido en la sección anterior.

$$\mathbb{Z}_3 = \{1, \theta, \theta^2\} \quad (2.133)$$

Con estas herramientas ya podemos definir el correspondiente grupo espacial S , cuyo elemento genérico puede escribirse como (ω, l) con $\omega \in \mathbb{Z}_3$ y $l \in \Lambda_{SU(3)^3}$. De tal manera que al actuar sobre un elemento $x \in \mathbb{R}^6$,

$$(\omega, l)x = \omega x + l = \sum_{i=1}^6 [x^i(\omega e_i) + l^i e_i] \quad (2.134)$$

donde ωe_i se puede obtener explícitamente de la ecuación 2.132. Por lo visto en la sección pasada, sabemos que la regla de multiplicación estaría dada por

$$(\omega, l)(\omega', l') = (\omega\omega', \omega l' + l) \quad (2.135)$$

El grupo espacial tiene 4 generadores: $(\theta, 0), (1, e_1), (1, e_3), (1, e_5)$. Y los puntos fijos cumplen con la condición de que $x_f \in \mathbb{R}^6$ sea invariante antes la acción de elementos del grupo espacial con $\omega = \theta$

$$(\theta, l)x_f = \theta x_f + l = x_f \quad (2.136)$$

Resolviendo esta ecuación, obtenemos que los puntos fijos están en perfecta correspondencia con elementos de $\Lambda_{SU(3)^3}$

$$x_f(l) = (1 - \theta)^{-1}l, \quad (2.137)$$

o dicho de otra manera, los puntos fijos sólo dependen de los elementos en la retícula. Ahora, para definir de manera completa el orbifolio, que denotaremos por $\Omega = (6, \mathbb{Z}_3) = \mathbb{R}^6/S$, debemos demandar que los puntos $x, x' \in \mathbb{R}^6$ son equivalentes si están relacionados bajo la acción del grupo espacial S . A continuación presentamos la definición formal.

Definición 2.4.1. El orbifolio $\Omega(6, \mathbb{Z}_3) = \mathbb{R}^6/S$ es el espacio cociente construido a partir del establecimiento de la equivalencia entre puntos $x, x' \in \mathbb{R}^6$ si estos están relacionados entre sí por la acción del grupo espacial S : $x \sim x'$ si y sólo si existe un elemento $(\omega, l) \in S$ tales que $x' = (\omega, l)x$.

Un dominio fundamental para el orbifolio, proyectado hacia cada uno de los planos complejos, se puede visualizar en la Figura 2.4. Las flechas indican identificación entre puntos en las fronteras, donde la zona abierta de la frontera es cerrada por dicha identificación. Como el orbifolio es 6-dimensional, tenemos en realidad 3 de estas almohadas de este estilo, asociadas a la proyección del orbifolio en cada uno de los planos complejos.

Ahora recordemos el concepto de *clase de equivalencia* pero aplicado en el contexto del grupo espacial S .

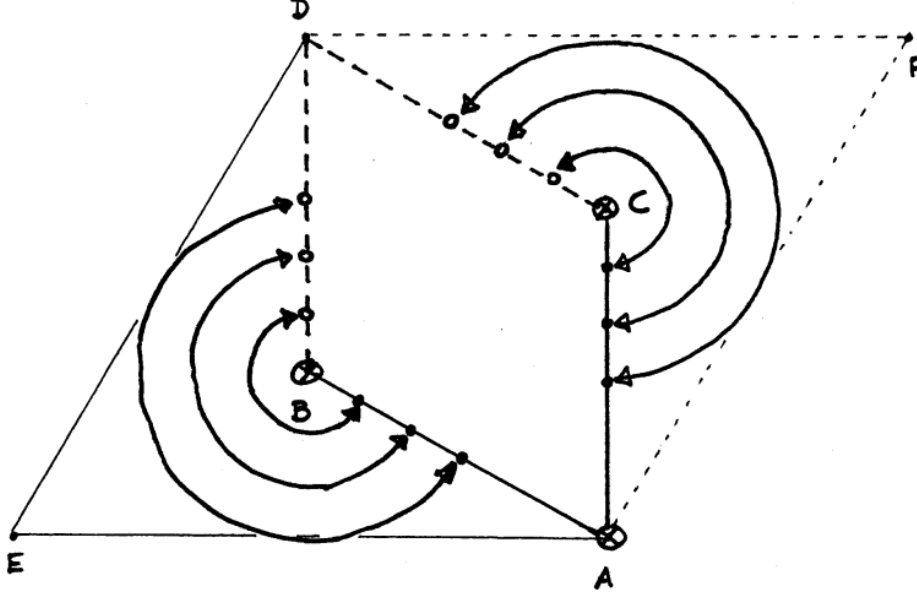


Figura 2.4: **Dominio fundamental del orbifolio $\Omega(6, \mathbb{Z}_3)$ proyectado en cada plano complejo.** El paralelogramo \overline{BACDB} ilustra el orbifolio \mathbb{Z}_3 2-dimensional proyectado a uno de los planos complejos. Aquí tenemos que $E = 0$, $A = 1$, $D = e^{i\pi/3}$, $F = A + D$, $B = F/3$ y $C = 2F/3$. Los puntos que se encuentran en la línea punteada están identificados con puntos en la línea sólida, cerrando el dominio completamente.

Definición 2.4.2 (Clase de equivalencia). La *clase de equivalencia* asociada al elemento del grupo espacial (ω, l) es el conjunto

$$\{(\omega', l') \cdot (\omega, l) \cdot (\omega', l')^{-1} \mid (\omega', l') \in S\} \quad (2.138)$$

Este concepto es de fundamental importancia pues permite clasificar a los puntos fijos del orbifolio, en el cual, la mayoría de los puntos fijos son equivalentes. Sólo hay 27 puntos fijos inequivalentes en el orbifolio $\Omega(6, \mathbb{Z}_3)$. Para poder argumentar esta afirmación, primero demostremos el siguiente resultado.

Teorema 1. *El orbifolio 6-dimensional \mathbb{Z}_3 , $\Omega(6, \mathbb{Z}_3) = \mathbb{R}^6/S$ tiene 27 puntos fijos inequivalentes que se pueden obtener con la ecuación 2.136 y usando los elementos de la red $\Lambda_{SU(3)^3}$ de la forma:*

$$l(n_1, n_3, n_5) = n_1 e_1 + n_3 e_3 + n_5 e_5, \quad n_i = 0, \pm 1. \quad (2.139)$$

de tal manera que las clases de equivalencia podremos denotarlas como

$$\{(\theta, l(n_1, n_3, n_5) + \sum_{i=1,3,5} m^i v_i) | m^i \in \mathbb{Z}\}_{n_i=0,\pm 1} \quad (2.140)$$

Demostración. Sean $\theta \in \mathbb{Z}_3$ y $l_0 \in \Lambda$, describamos la clase de equivalencia correspondiente al elemento $(\theta, l_0) \in S$ de nuestro grupo de espacio. De acuerdo a la regla del producto de los elementos del grupo de espacio 2.134, podemos determinar fácilmente que:

$$(\omega, l)^{-1} = (\omega^{-1}, -\omega^{-1}l), \quad \forall (\omega, l) \in S \quad (2.141)$$

Por lo que $\forall (\omega, l) \in S$ podemos calcular que

$$\begin{aligned} (\omega, l)(\theta, l_0)(\omega, l)^{-1} &= (\omega, l)(\theta, l_0)(\omega^{-1}, -\omega^{-1}l) \\ &= (\omega, l)(\theta\omega^{-1}, -\theta\omega^{-1}l + l_0) \\ &= (\omega\theta\omega^{-1}, -\omega\theta\omega^{-1}l + \omega l_0 + l) \\ &= (\theta, -\theta l + \omega l_0 + l) \\ &= (\theta, (1 - \theta)l + \omega l_0), \end{aligned} \quad (2.142)$$

donde el cuarto paso se debe a que el grupo \mathbb{Z}_3 es abeliano. Ahora analicemos el término $(1 - \theta)l$. En general, podemos expandir l como combinación lineal de los vectores base definidos en 2.132, de tal forma que

$$l \equiv \sum_{i=1}^6 l^i e_i \quad \Rightarrow \quad (1 - \theta)l = \sum_{i=1}^6 l^i (1 - \theta)e_i = \sum_{i=1}^6 l^i v_i, \quad (2.143)$$

$$\text{donde } v_i \equiv (1 - \theta)e_i = e_i - e_{i+1}, \quad v_{i+1} \equiv (1 - \theta)e_i = e_i + 2e_{i+1}, \quad i = 1, 3, 5.$$

Con esta expresión tenemos que en general:

$$(\omega, l)(\theta, l_0)(\omega, l)^{-1} = (\theta, l_0 + \sum_{i=1}^6 l^i v_i), \quad l^i \in \mathbb{Z} \quad (2.144)$$

Esto nos permite concluir, por tanto, que la clase de equivalencia de (θ, l_0) está dada por

$$\{(\theta, l_0 + \sum_{i=1}^6 m^i v_i)\}_{m^i \in \mathbb{Z}} \quad (2.145)$$

Ahora demostremos que existen 27 conjuntos inequivalentes de este estilo, de tal manera que

$$\{(\theta, l(n_1, n_3, n_5) + \sum_{i=1,3,5} m^i v_i) | m^i \in \mathbb{Z}\}_{n_i=0,\pm 1} \quad (2.146)$$

con

$$l(n_1, n_3, n_5) = n_1 e_1 + n_3 e_3 + n_5 e_5, \quad n_i = 0, \pm 1. \quad (2.147)$$

Sea $k \in \Lambda$ arbitrario, de tal forma que $k = \sum_{i=1}^6 m^i v_i$. Deseamos que el vector arbitrario k pueda ser expresado en términos de $l(n_1, n_3, n_5)$, se propone el ansatz

$$k = l(n_1, n_3, n_5) + \sum_{i=1}^6 m^i v_i = \sum_{i=1,3,5} [n_i e_i + m^i (e_i - e_{i+1}) + m^{i+1} (e_i + 2e_{i+1})]. \quad (2.148)$$

Con este ansatz tendremos las siguientes constricciones para ($i=1,3,5$):

$$\begin{aligned} 3m^{i+1} &= k^i + k^{i+1} - n_i \\ m^i &= k^i - n_i - m^{i+1}. \end{aligned} \quad (2.149)$$

Estas ecuaciones se cumplen sólo si $n_i = 0, \pm 1$ en 2.148. Entonces, siempre y cuando se cumplan las condiciones de la ecuación 2.149 podemos asegurar que k se puede escribir en términos de $l(n_1, n_3, n_5)$.

Sea $(\theta, k) \in S$, ya sabemos que entonces su clase de equivalencia correspondiente está dada por

$$\{(\theta, k + \sum_{i=1}^6 m^i v_i)\}_{m^i \in \mathbb{Z}}. \quad (2.150)$$

Por lo visto en la ecuación 2.148, tenemos por tanto que su clase de equivalencia es equivalente entonces a

$$\{(\theta, l(n_1, n_3, n_5) - \sum_{i=1}^6 m^i v_i)\}_{m^i \in \mathbb{Z}} = \{(\theta, l(n_1, n_3, n_5) + \sum_{i=1}^6 m^i v_i)\}_{m^i \in \mathbb{Z}}, \quad (2.151)$$

y como (θ, k) era arbitrario, entonces todas las clases de equivalencia en S se pueden escribir como en la ecuación anterior. Esto quiere decir que cada clase de equivalencia está caracterizada por un $l(n_1, n_3, n_5)$ que llamaremos líder de la clase en cuestión. Demostremos que cada líder pertenece a clases distintas.

Supongamos que $l(n_1, n_3, n_5)$ y $l(n'_1, n'_3, n'_5)$ corresponden a la misma clase. Entonces $l(n'_1, n'_3, n'_5)$ puede expresarse como en 2.148 y utilizando las constricciones 2.149 tenemos que ($i = 1, 3, 5$):

$$3m^{i+1} = n_i 0 - n_i, \quad (2.152)$$

como n_i y n'_i sólo pueden tomar valores $0, \pm 1$, la única solución es que $m^{i+1} = 0$ y por tanto $n'_i = n_i$. Por lo que podemos concluir que $l(n'_1, n'_3, n'_5) = l(n_1, n_3, n_5)$. Podemos concluir entonces que cada clase está caracterizada por un elemento de la forma $l(n_1, n_3, n_5)$ y como para cada n_i tenemos tres posibilidades entonces en total tenemos 27 líderes $l(n_1, n_3, n_5)$ distintos y por tanto 27 clases de equivalencia diferentes para S . \square

Con el teorema que acabamos de demostrar ya podemos contar el número de puntos fijos inequivalentes que tiene nuestro orbifolio. Sabemos que los puntos fijos son tales como se mostró en la ecuación 2.137 y como los $l \in \Lambda$ están en correspondencia uno a uno con los líderes de las 27 clases de equivalencia, tenemos entonces que los puntos fijos tienen la forma

$$f_{n_1, n_2, n_3} \equiv x_f(l(n_1, n_2, n_3)) = (1 - \theta)^{-1} l(n_1, n_2, n_3) \quad (2.153)$$

que nos permite concluir lo que ya habíamos afirmado sobre el número de puntos fijos inequivalentes.

Con esto terminamos de estudiar las características puramente geométricas del orbifolio con grupo de punto \mathbb{Z}_3 . Ahora estudiemos sus implicaciones físicas, principalmente las que tiene el hecho de considerar invariancia modular.

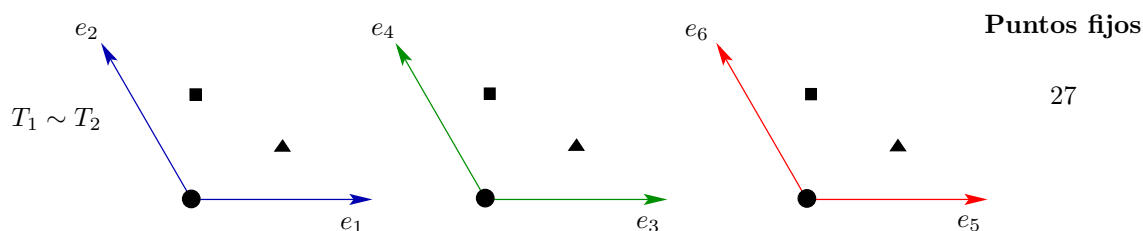


Figura 2.5: Geometría del orbifolio \mathbb{Z}_3 compactificado en una retícula $SU(3)^3$. Los 27 puntos fijos surgen del producto de los 3 planos que se aprecian.

Compactificaciones en el orbifolio \mathbb{Z}_3 invariantes modulares

Determinemos entonces las formas inequivalentes en las que podemos encajar al grupo de punto V^I . Primero escribimos a los vectores de traslación como

$$V = (V_1, V_2), \quad (2.154)$$

donde V_1 y V_2 son las componentes de V trasladando en las retículas de E_8 y E'_8 , respectivamente. Como ya vimos, las condiciones de homomorfismos construyen a estos vectores como

$$3 \sum_J V_1^J = 0 \pmod{2} \quad (2.155)$$

y

$$3 \sum_K V_2^K = 0 \pmod{2}. \quad (2.156)$$

Para el caso del orbifolio \mathbb{Z}_3 , tenemos que el vector de torcedura está dado por

$$v = \frac{1}{3}(1, 1, 2), \quad (2.157)$$

así que la condición para invariancia modular (2.82) se escribe como

$$3(V_1^2 + V_2^2) = 0 \pmod{2}. \quad (2.158)$$

Combinando las ecuaciones (2.155)-(2.157) requiere que

$$V_1^2 = \frac{2}{9}q_1, \quad V_2^2 = \frac{2}{9}q_2, \quad (2.159)$$

donde q_1 y q_2 son enteros.

Cualquier vector de traslación V_i , $i = 1, 2$ tiene una distancia a lo más 1 de cualquier punto en la retícula E_8 , entonces sustrayendo de manera adecuada un vector en la retícula, siempre podemos asegurar que

$$V_1^2 \leq 1, \quad V_2^2 \leq 1. \quad (2.160)$$

Así que salvo intercambios entre V_1 y V_2 las únicas posibilidades inequivalentes son

$$V_1^2 = V_2^2 = 0, \quad (2.161)$$

$$V_1^2 = \frac{2}{3}, \quad V_2^2 = 0, \quad (2.162)$$

$$V_1^2 = \frac{2}{9}, \quad V_2^2 = \frac{4}{9}, \quad (2.163)$$

$$V_1^2 = V_2^2 = \frac{2}{3} \quad (2.164)$$

y

$$V_1^2 = \frac{8}{9}, \quad V_2^2 = \frac{4}{9}. \quad (2.165)$$

Hay muchas opciones para las líneas de Wilson a_ρ . Como en el caso de V , tenemos las condiciones de homeomorfismos dadas por

$$2 \sum_{J \in E_8} a_\rho^J = 0 \pmod{2} \quad (2.166)$$

y

$$3 \sum_{K \in E'_8} a_\rho^K = 0 \pmod{2}. \quad (2.167)$$

De la misma forma, sustrayendo un vector de la retícula, podemos asegurar que

$$\sum_{J \in E_8} (a_\rho^J)^2 \leq 1, \quad \sum_{K \in E'_8} (a_\rho^K)^2 \leq 1. \quad (2.168)$$

No todas las líneas de Wilson que satisfacen (2.166)-(2.168) y las condiciones de invariancia modular

$$3a_\rho^2 = 0 \pmod{2} \quad (2.169)$$

y

$$3V \cdot a_\rho = 0 \pmod{2} \quad (2.170)$$

son independientes. Si la acción del elemento generador del grupo de punto en los vectores base para la retícula

$$\theta e_\rho = M_{\rho\sigma} e_\sigma, \quad (2.171)$$

entonces

$$a_\rho = M_{\rho\sigma} a_\sigma + \Lambda, \quad (2.172)$$

donde Λ es un vector en la retícula $E_8 \times E'_8$, indicando el hecho de que estos son caminos inequivalentes en el toro son equivalentes en el orbifolio.

2.5. Orbifolio \mathbb{Z}_6

Ya hemos visto la construcción explícita del orbifolio con grupo de punto \mathbb{Z}_3 , que es el que tienes las propiedades más sencillas de entre todos los orbifolios abelianos. En este trabajo nos concentraremos en el orbifolio con grupo de punto \mathbb{Z}_6 y vale la pena describirlo a fondo como sigue. Este orbifolio ya ha sido explorado anteriormente en [24, 22] y además se ha presentado un mecanismo para encajar el grupo del modelo estándar en el grupo de norma de la teoría con este orbifolio en [25]. Entonces la construcción de este orbifolio es para obtener una teoría 4-dimensional, por lo que 6 dimensiones de la teoría 10D heterótica de cuerdas son compactificadas en un orbifolio. En nuestro caso, esto es un orbifolio \mathbb{Z}_6 que se obtiene modificando un toro 6-dimensional junto con un toro de norma 16-dimensional por medio de una torsión dada por \mathbb{Z}_6 ,

$$\mathcal{O} = T^6 \otimes T_{E_8 \times E'_8} / \mathbb{Z}_6 \quad (2.173)$$

En las coordenadas complejas del toro 6-dimensional, $z^\alpha, \alpha = 1, 2, 3$, el grupo \mathbb{Z}_6 actúa como

$$z^\alpha \rightarrow e^{2\pi i v_k^\alpha} z^\alpha. \quad (2.174)$$

Aquí $6 \cdot v_k$ tiene componentes enteras. Las coordenadas compactas de cuerdas son descritas por las variables complejas que habíamos definido la sección pasada. El grupo \mathbb{Z}_6 actúa sobre estas coordenadas como

$$Z^\alpha(\tau + \sigma + \pi) = e^{2\pi i k v_k^\alpha} Z^\alpha(\tau + \sigma), \quad k = 0, 1, \dots, 5, \quad (2.175)$$

$$\Psi_R^\alpha(\tau - \sigma - \pi) = \Psi_R^\alpha(\tau - \sigma) - \pi k v_k^\alpha, \quad X^I(\tau + \sigma + \pi) = X^I(\tau + \sigma) + \pi k V_k^I, \quad (2.176)$$

donde $6 \cdot V^I$ es un vector que vive en la retícula $E_8 \times E_8$. Los vectores base e_k son tomados como raíces simples de un álgebra de Lie, cuya elección está constrañida por la simetría de la retícula. En nuestro caso, la retícula debe tener una simetría \mathbb{Z}_6 y permitir la existencia de 3 vectores de traslación independientes V_k (o dos líneas de Wilson de orden 2). Esto deja dos posibilidades para la retícula:

$$G_2 \times SU(3) \times SO(4) \quad \text{ó} \quad SU(3) \times SU(3) \times SO(4). \quad (2.177)$$

En la tabla 2.1 vimos hay dos generadores del grupo de punto \mathbb{Z}_6 ,

$$\theta = \exp[2\pi i(v_1 M^{34} + v_2 M^{56} + v_3 M^{78})], \quad (2.178)$$

con (v_1, v_2, v_3) parámetros de rotación o vector de torsión. Tenemos dos elecciones para dichos parámetros

$$v_k = \frac{1}{6}(1, 1, -2), \quad v_k = \frac{1}{6}(1, 2, -3), \quad (2.179)$$

donde la segunda elección corresponde al llamado grupo de punto \mathbb{Z}_6 -II, que es el grupo en el cual concentraremos nuestro trabajo. Discutamos con detalle la estructura de sus puntos fijos para el primer sector torcido, la cual está detallada en la figura 2.6. Consideremos que el orbifolio \mathbb{Z}_6 -II tiene como retícula a

$$\Lambda_{\mathbb{Z}_6-II} = G_2 \times SU(3) \times SO(4), \quad (2.180)$$

que es una retícula factorizable. Una ventaja de que dicha retícula sea factorizable es el hecho de que podemos encontrar los puntos fijos de manera independiente para cada uno de los factores sin ninguna pérdida de información, y después ponerlos juntos para encontrar la estructura de puntos fijos completa [22]. El grupo de punto \mathbb{Z}_6 -II es generado por θ como en (2.178), el cual actúa simultáneamente como una rotación por $2\pi/6$ en el plano G_2 , una rotación por $2\pi/3$ en el plano $SU(3)$ y una reflexión por el origen en la subretícula $SO(4)$ y con el segundo conjunto de parámetros que presentamos en la ecuación (2.179).

Para este grupo, tenemos un total de cinco sectores torcidos, correspondientes a diferentes potencias de θ . Por el teorema de puntos fijos de Lefschetz [22], el cual proporciona el número de puntos fijos para una retícula factorizable correspondiente al primer sector torcido (que sí es nuestro caso), concluimos que tenemos un total de 12 puntos fijos. Para localizarlos por completo en el espacio compacto (o dominio fundamental) debemos utilizar la relación para puntos fijos

$$Z_f = g_f Z_f = \theta Z_f + n_\alpha e_\alpha \Leftrightarrow (1 - \theta)Z_f \in \Lambda. \quad (2.181)$$

Los elementos constructores correspondientes a los puntos fijos están dados por

$$g_f^\theta \in \{(\theta, n_5 e_5 + n_6 e_6), (\theta, e_3 + n_5 e_5 + n_6 e_6), (\theta, e_3 + e_4 + n_5 e_5 + n_6 e_6)\}, \quad (2.182)$$

con $n_5, n_6 = 0, 1$. Para el siguiente sector torcido, es decir, θ^2 , tenemos que el vector de torcimiento corresponde a $2v = (1/3, 2/3, -1)$. Esto implica que todos los puntos de la subretícula $SO(4)$ se mantienen invariantes. Así, la acción de θ^2 en el espacio compacto introduce un toro fijo. El teorema de Lefshetz no puede ser aplicado en el segundo sector torcido. Sin embargo, sabemos de la construcción del orbifolio \mathbb{Z}_3 , que una rotación de $2 \times 2\pi/3$ en la subretícula $SU(3)$, se tienen 3 puntos fijos. Más aún, hay sólo dos puntos fijos inequivalentes en la subretícula G_2 . Conjuntando todo, encontramos entonces que hay 2×3 toros fijos en el segundo sector torcido. Los elementos constructores de este sector son

$$g_f^{\theta^2} \in \{(\theta^2, 0), (\theta^2, e_1), (\theta^2, e_4), (\theta^2, e_1 + e_4), (\theta^2, e_3 + e_4), (\theta^2, e_1 + e_3 + e_4)\}. \quad (2.183)$$

En el tercer sector torcido, la acción del grupo de punto, descrita por $3v = (1/2, 1, -3/2)$, deja invariante a la subretícula $SU(3)$, así que también obtenemos un toro fijo en este caso. Utilizando (2.181) podemos corroborar que existen 5 puntos invariantes bajo θ^3 en cada una de las dos subretículas restantes. Sin embargo, los puntos fuera del origen en la subretícula G_2 son equivalentes. Denotemos a los elementos de grupo de espacio asociados como $g_1 = (\theta^3, e_1)$, $g_2 = (\theta^3, e_1 + e_2)$ y $g_3 = (\theta^3, e_2)$. Se puede verificar que

$$g_1 \simeq (\theta, e_1)g_1(\theta, e_1)^{-1} = g_2 \quad \text{y} \quad g_2 \simeq (\theta, 2e_1 - e_2)g_2(\theta, 2e_1 + e_2)^{-1} = g_3, \quad (2.184)$$

concluyendo que los tres elementos pertenecen a la misma clase de equivalencia. Así, tenemos entonces un total de 2×4 toros fijos, descritos por los siguientes elementos constructores

$$g_f^{\theta^3} \in \{(\theta^3, n_5 e_5 + n_6 e_6), (\theta^3, e_2 + n_5 e_5 + n_6 e_6)\} \quad (2.185)$$

con $n_5, n_6 = 0, 1$. El cuarto y quinto sector torcido poseen la misma estructura que el segundo y primer sector, respectivamente. Así que es suficiente con estudiar los tres sectores torcidos ilustrados en la Figura 2.6.

Como ya mencionamos anteriormente, la compactificación en orbifolios permite romper al grupo de norma reduciendo el rango del mismo a través de la inclusión de Líneas de Wilson para dejar invariante un subgrupo del grupo de norma original. En [26] se construye una compactificación de la cuerda heterótica en un orbifolio T^6/\mathbb{Z}_6 conduciendo al espectro del Modelo Estándar con estados de materia vectoriales adicionales. En ese trabajo se proporciona la manera en que se puede obtener el grupo del modelo estándar a partir de la intersección de tres subgrupos $SO(10)$ de E_8 . También cuentan procesos de ruptura a partir de los cuales se obtienen Teorías de Gran Unificación (GUTs, por sus siglas en inglés) intermedias. Una GUT constituye una suposición en la cual la naturaleza cuenta con más simetría que la del Modelo Estándar pero que se reduce a la del mismo a bajas energías. Las simetrías y el contenido de partículas del SM apunta a las GUTs como el siguiente paso en la unificación de todas las fuerzas. Los quarks derechos e izquierdos y los leptones pueden ser agrupados en tres multipletes de $SU(5)$, $\mathbf{10} = (q_L, u_R^c, e_R^c)$, $\bar{\mathbf{5}} = (d_R^c, l_L)$ y $\mathbf{1} = \nu_R^c$. Todos los quarks y leptones de una generación pueden ser unificados en un solo multiplete del grupo GUT $SO(10)$,

$$\mathbf{16} = \mathbf{10} + \bar{\mathbf{5}} + \mathbf{1}. \quad (2.186)$$

El grupo $SO(10)$ contiene como subgrupos al grupo de Pati-Salam, $G_{PS} = SU(4) \times SU(2) \times SU(2)$ y al grupo Georgi-Glashow $SU(5)$. Es una propiedad remarcable que los campos de materia del Modelo Estándar forman multipletes completos de $SO(10)$ mientras que los campos de norma y el Higgs son multipletes divididos. Necesitan ser combinados con otros multipletes divididos, que no están presentes en el modelo estándar, para obtener una teoría unificada completa. También es sabido que grupos excepcionales juegan un rol excepcional en gran unificación, y el embebimiento

$$SO(10) \subset E_6 \subset E_8 \quad (2.187)$$

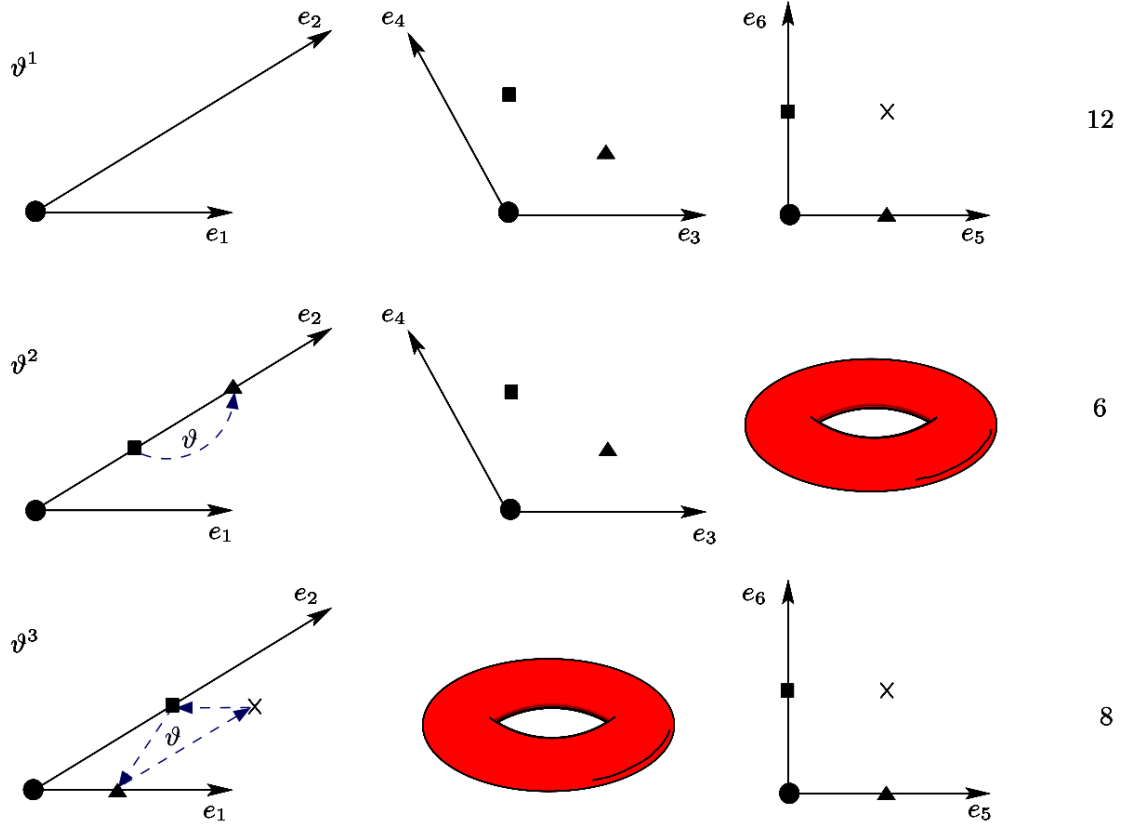


Figura 2.6: La geometría del orbifolio \mathbb{Z}_6 -II compactificado en una retícula $G_2 \times SU(3) \times SO(4)$. Puntos fijos en el primer sector torcido para el orbifolio \mathbb{Z}_6 . En total, identificamos 12 puntos fijos para el primer sector torcido. [22]

aparece, en particular, en las compactificaciones de la cuerda heterótica. En [26] se estudia el esquema a través del cual con orbifolios con torcimiento \mathbb{Z}_N permite la ruptura de E_8 al Modelo Estándar con las GUTs intermedias E_6 y $SO(10)$.

Volviendo al contexto del orbifolio \mathbb{Z}_6 -II, [26] también hace un análisis de rompimiento de simetría de norma hasta llegar al modelo estándar con materia vectorial adicional, utilizando la retícula $G_2 \times SU(3) \times SO(4)$. También en [27] hacen un análisis de la compactificación \mathbb{Z}_6 -II y con la misma retícula para obtener modelos con el grupo de norma Pati-Salam 4D. El orbifolio \mathbb{Z}_6 es equivalente a un orbifolio $\mathbb{Z}_2 \times \mathbb{Z}_3$, donde los dos vectores de torcedura son $v_2 = 3v_k = \frac{1}{2}(1, 2, -3)$ y $v_3 = 2v_k = \frac{1}{3}(1, 2, -3)$. En esta geometría hay, a lo más, 3 líneas de Wilson consistentes [28], una de grado 3 (W_3), a lo largo de la subretícula $SU(3)$, y dos de orden 2 (W_2, W'_2) a lo largo de la subretícula $SO(4)$. Recordemos que las líneas de Wilson son encajes de los 6 generadores de la

retícula en cuestión e_α , entonces uno podría pensar que como consecuencia habría seis líneas de Wilson a_α . Sin embargo, no todas las direcciones e_α son independientes en el orbifolio. Como en el caso de la subretícula $SU(3)$ de la Figura 2.6. En dicha retícula tenemos que el grupo \mathbb{Z}_6 -II actúa como una rotación $2\pi/3$, por lo que

$$e_3 \rightarrow e_4 \Rightarrow a_3 \rightarrow a_4, \tag{2.188}$$

por tanto, cuando hablamos de la línea de Wilson W_3 nos referimos a que $W_3 = a_3 = a_4$. Esta convención para las líneas de Wilson es la que utilizaremos más adelante. Dicha configuración la podemos apreciar en la Figura 2.7.

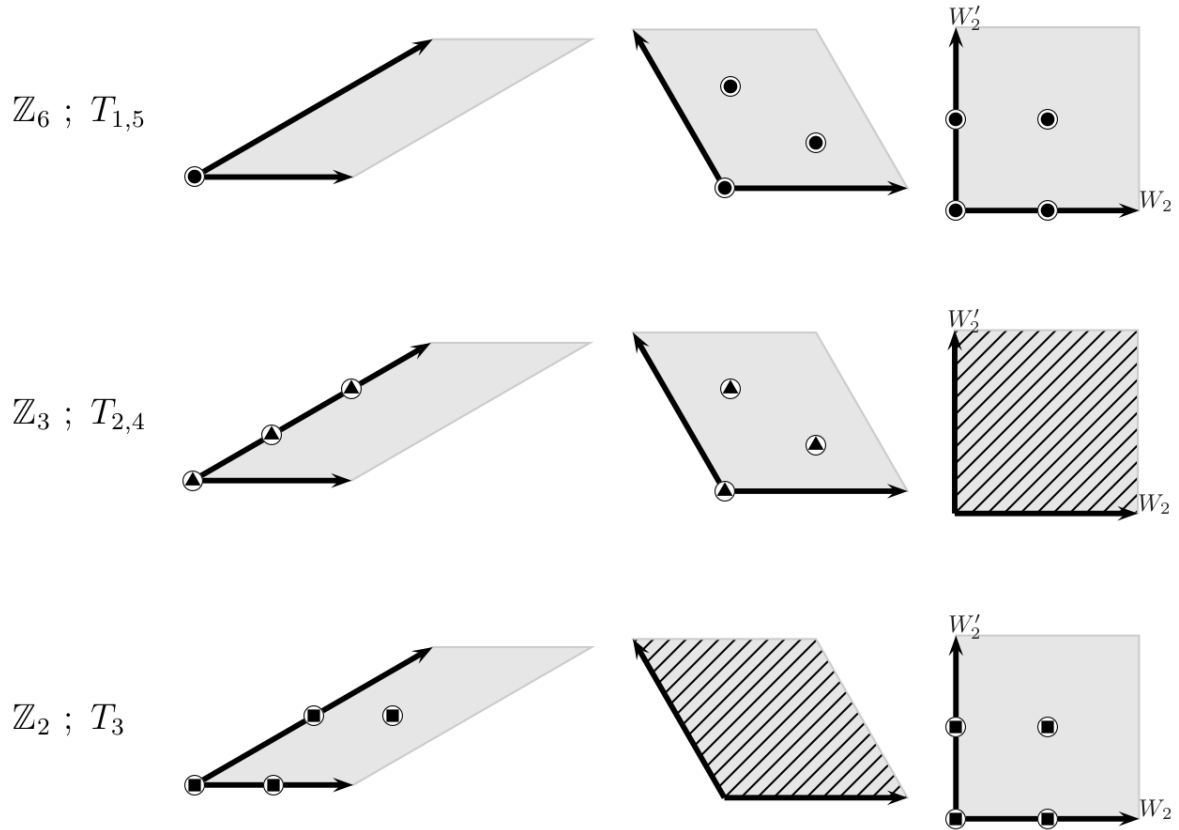


Figura 2.7: Puntos fijos y planos invariantes (con rayas) bajo el torcimiento \mathbb{Z}_6 y los subtorcimientos \mathbb{Z}_3 y \mathbb{Z}_2 describiendo la localización de los diferentes sectores torcidos. [26]

Capítulo 3

Machine Learning y el autoencoder

En la actualidad, muchas disciplinas científicas utilizan grandes cantidades de datos para poder obtener la caracterización de un fenómeno u obtener regularidades en sistemas complicados. Tal es el caso de sistemas complejos, que estudia sistemas conformados por una gran cantidad de unidades que interactúan entre sí. Ahora bien, aunque ya se cuenta con equipos de cómputo muy poderosos en la actualidad, muchas veces no son suficientes para procesar la enorme cantidad de datos con los que se posee para estudiar un problema específico. Esos conjuntos enormes de datos son tales que implicarían una enorme carga computacional y tiempos que podrían no ser prácticos. La ciencia de datos tiene como objetivo extraer características o información colectiva de dichos conjuntos, reduciendo así las exigencias de procesamiento. Uno de los métodos de la ciencia de datos para extraer esa información o características es por medio del aprendizaje automático (*machine learning*, en inglés) y cuyos fundamentos serán discutidos en este capítulo para así, posteriormente, aterrizarlo en una aplicación a nuestros intereses en teoría de cuerdas.

3.1. Fundamentos de *machine learning*

Esta sección se basa fundamentalmente en [29]. Como ya mencionamos, machine learning es una rama de la ciencia de datos y es un método de análisis de datos que automatiza la construcción de modelos analíticos. Es decir, el método tiene como objetivo desarrollar técnicas que permitan que un sistema aprenda por sí mismo a identificar de forma automática y sin intervención humana completa o parcial, patrones, tendencias, características o regularidades colectivas de un conjunto de datos. Machine learning utiliza heurística y aproximaciones para así llevar a cabo tareas cuyas soluciones exactas serían muy tardadas. Se distinguen al menos tres ramas de machine learning, las cuales están esquematizadas en la figura 3.1. Las tres ramas son:

- **Machine Learning no supervisado:** Esta rama, como indica su nombre, consiste en la

extracción de información de un conjunto de datos determinado sin intervención humana. Se pide a la máquina que por sí misma identifique características o información colectiva sin saber qué busca exactamente. Este tipo de machine learning es usado frecuentemente para identificar clústers para un conjunto de datos.

- Machine Learning supervisado:** Esta rama se refiere a entrenamientos de máquina de tal manera que ésta recibe un conjunto de datos, cuyos resultados son ya conocidos, así que la máquina aprende de esa entrada y salida conocidas para que al recibir un nuevo conjunto del cual no se conoce su resultado, la máquina lo prediga con cierta precisión. Este tipo de aprendizaje se aplica, por ejemplo, para regresiones o clasificaciones. Para el caso de clasificaciones, al algoritmo se le enseña por medio de ejemplos con datos de entrada de la misma clase. En regresiones, al algoritmo se le enseña por medio de datos que consisten en entradas y sus correspondientes salidas producidas por medio de alguna función o correspondencia. El algoritmo aprende ajustando parámetros internos para así saber qué resultados se le asignan a una entrada determinada.
- Aprendizaje reforzado:** En algunos casos, uno ya sabe lo que está buscando pero no sabe cómo encontrarlo. Dicha tarea de búsqueda es hecha de maneras satisfactorias utilizando aprendizaje reforzado. La idea es que al algoritmo no se le dice qué hacer exactamente, pero sí recibe una retroalimentación basada en su rendimiento, para así motivar a que las acciones que lleve a cabo sean las necesarias para llegar al resultado deseado.

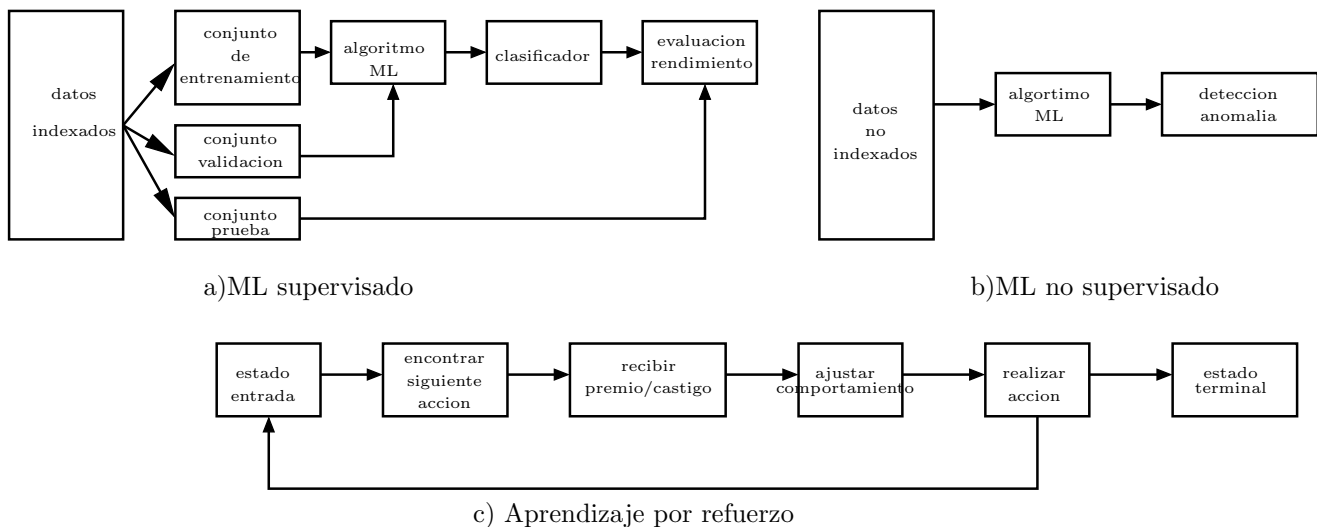


Figura 3.1: Diagramas que ilustran las generalidades de los algoritmos que implementan ML supervisado y no supervisado, también del aprendizaje reforzado.

La ciencia de datos no es muy nueva, de hecho, ha sido aplicada extensivamente en muchas

ramas científicas, tales como la biología, química, física de partículas experimental y cosmología. También se ha aplicado en ciencias sociales, como redes sociales, juegos o coches que se manejan solos. Sin embargo, no había sido aplicado en la teoría de cuerdas hasta junio de 2017, cuando 4 grupos publicaron de manera independiente trabajos aplicando machine learning a la teoría de cuerdas [2, 3, 4, 5]. Desde entonces, muchas técnicas de machine learning y ciencia de datos han sido aplicadas para:

- acelerar el cómputo de datos de teoría de cuerdas relacionados con geometrías algebraicas (o tóricas),
- identificar estructuras matemáticas en teoría de cuerdas,
- estudiar la geometría del vacío de cuerdas,
- encontrar nuevos vacíos de cuerdas.

3.1.1. Capas comunes en redes neuronales

La estructura de las redes neuronales está inspirada en el cerebro humano que tiene múltiples conexiones neuronales y que sabemos que interactúan entre sí con neuronas colindantes. Inspirados en esos procesos, la información en las redes neuronales (RNs) es procesada a través de capas conectadas entre sí, produciendo una capa saliente (resultado) a partir de una capa entrante.

Cada capa consiste de un conjunto de nodos y las conexiones entre los nodos de todas las capas son modeladas por medio de transformaciones lineales, es decir, multiplicación de matrices. Cuando la entrada o señal llega a un nodo, el nodo es activado o no (dependiendo del objetivo que tenga la RN y del conjunto de datos), es decir, una función de activación es aplicada a la señal o no. Después de eso, la señal activada es transformada por valores numéricos llamados *pesos* asociados a una transformación lineal y que configuran los valores en los nodos que configuran la siguiente capa. Esta es la arquitectura básica de las RNs.

A continuación se presenta un ejemplo de RN cuyo objetivo es el de aproximar una función escalar en dos variables. Si deseamos aproximar la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, podemos proceder generando una colección de datos que representen puntos en dos dimensiones, de la forma (x_1, x_2) . La función escalar tiene como entrada puntos de este estilo y regresa un número real. El procedimiento para construir una red neuronal que entrene a la máquina para aproximar a la función f sería el siguiente (la configuración de la RN no tiene que ser la que se presenta, puede modificarse de tal manera que la aproximación de esta función sea lo mejor posible):

- Una primera capa que esté conformada por dos nodos y donde cada uno de ellos represente a uno de los números reales que conforman al vector (x_1, x_2) ,

- Una segunda capa (la primera capa oculta), que consistiría de 5 nodos y que están conectados con la primera capa por medio de la multiplicación de una matriz de 5×2 , $w^{(1)}$, y la adición de un vector constante $b^{(1)}$. Posteriormente aplicamos la función de activación a cada nodo de la segunda capa. La función de activación $a(\cdot)$ tiene varias opciones, pero es una función del estilo $a : \mathbb{R} \rightarrow \mathbb{R}$; más adelante veremos opciones para esta función de activación.
- Una tercera capa, que de nuevo será una capa oculta, pero ahora contará con 3 nodos y la misma función de activación. La conexión de estos nodos con la segunda capa es por medio de una multiplicación de matriz de 3×5 , $w^{(2)}$ y un vector constante $b^{(2)}$.
- La cuarta capa sería la capa de salida, el resultado que nos interesa obtener a partir de la capa de entrada (la primera). Consiste de un solo nodo cuyo valor representa la salida \hat{y} de la red neuronal. Está conectado con la anterior capa por medio de una matriz de 1×3 , $w^{(3)}$ y la adición de un vector $b^{(3)}$. También tiene la misma función de activación que en las capas anteriores.

La arquitectura de la red neuronal anterior está ilustrada en la parte izquierda de la Figura 3.2. Las entradas de las matrices $w^{(i)}$ y de los vectores $b^{(i)}$ corresponden a parámetros de la red, que colectivamente se denotan por θ . Los valores de los parámetros varían durante el entrenamiento pero una vez terminado dicho proceso, éstos son fijos. Para entrenar a la red, se procede a que inicialmente se asignan valores aleatorios a θ y se hacen varias iteraciones de la red, de tal manera que se van obteniendo varios pares $((x_1, x_2), y)$ que son justamente los valores de los nodos en las capas de entrada y salida. Durante las iteraciones se van ajustando los parámetros θ de tal manera que las salidas $\hat{y} = h_\theta(x_1, x_2)$, donde h_θ es la función aplicada por toda la RN, aproxima el número de salida lo mejor posible. Durante la evolución del entrenamiento (a lo largo de las iteraciones), el error se va reduciendo como apreciamos en la parte derecha de la Figura 3.2. El error se refiere a la diferencia entre el valor obtenido y el valor deseado, éste se puede calcular por medio de las llamadas *funciones de pérdida* (MSE o MAE, por ejemplo, ver subsección de *Funciones de pérdida*). Más adelante describiremos el proceso de entrenamiento y las opciones disponibles para las funciones de pérdida.

3.1.2. Redes neuronales prealimentadas, capas completamente conectadas

Las redes neuronales prealimentadas son la implementación más sencilla de RNs y probablemente las más comunes. Este tipo de redes consiste en un conjunto de capas, una tras otra y en orden. Cada capa consiste de un conjunto de nodos, cuyo número puede variar conforme se avanza por capas. En el caso de este tipo de redes, como su nombre lo indica, tenemos una dirección y la información pasa desde la **capa de entrada** pasando por varias capas intermedias, **capas ocultas** hasta las **capas de salida**. En la Figura 3.3 podemos ver un diagrama que ilustra las características de este tipo de redes. El nombre de capas ocultas se deriva del hecho de que estas capas sólo son útiles

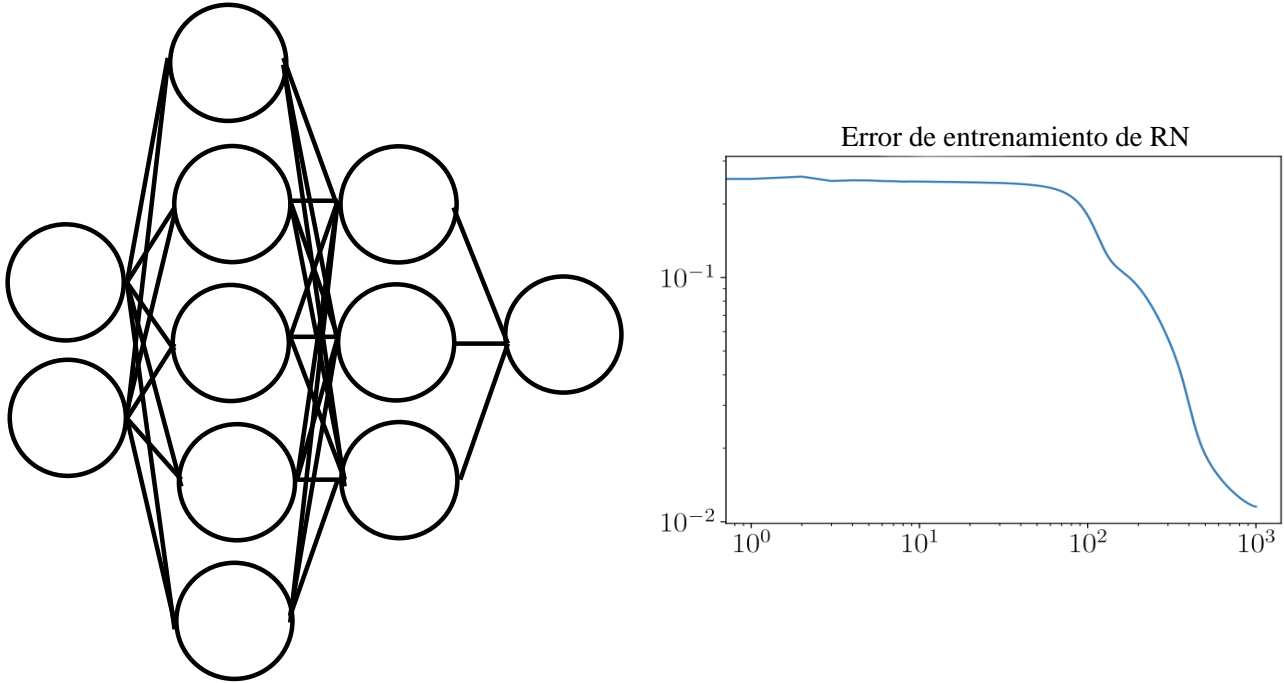


Figura 3.2: Del lado izquierdo se muestra la arquitectura de un ejemplo de RN. A la izquierda podemos apreciar cómo va decreciendo el error de entrenamiento de la RN a lo largo de las iteraciones.

durante el entrenamiento y no se aprecian en el resultado final. Cuando tenemos una gran cantidad de capas ocultas, decimos que tenemos una red neuronal **profunda**, mientras que si tenemos una gran cantidad de nodos, decimos que se trata de una red neuronal **amplia**. Como ya se mencionó, la conexión entre capas se lleva a cabo por medio de transformaciones lineales, las cuales son modeladas por medio de la multiplicación de matrices y la adición de un vector. Sea $l^{(i)}$ la salida de la capa i -ésima y denotemos por n_i el número de nodos de dicha capa. Así que el valor $z_\mu^{(i)}$ del μ -ésimo nodo de la i -ésima capa se obtiene después de la multiplicación matricial

$$z_\mu^{(i)} = \sum_{\nu=1}^{n_{i-1}} w_{\mu\nu}^{(i)} l_\nu^{(i-1)} + b_\mu^{(i)}, \quad \text{para } \mu = 1, 2, \dots, n_i. \quad (3.1)$$

Los índices latinos se aplicarán para indicar elementos de un conjunto y los índices griegos indexan componentes de vectores o matrices. Las entradas $w_{\mu\nu}^{(i)}$, de las matrices $w^{(i)}$ de $n_i \times n_{i-1}$ entradas, se les llama **pesos** y el vector $b^{(i)}$ se le llama **bias**. Aún falta aplicar a cada nodo de cada capa la acción de una función de activación de la forma $a(z)$. Con esto en mente, el resultado $l_\mu^{(i)}$ de μ -ésimo nodo de la i -ésima capa es

$$l_\mu^{(i)} = a^{(i)}(z_\mu^{(i)}), \quad (3.2)$$

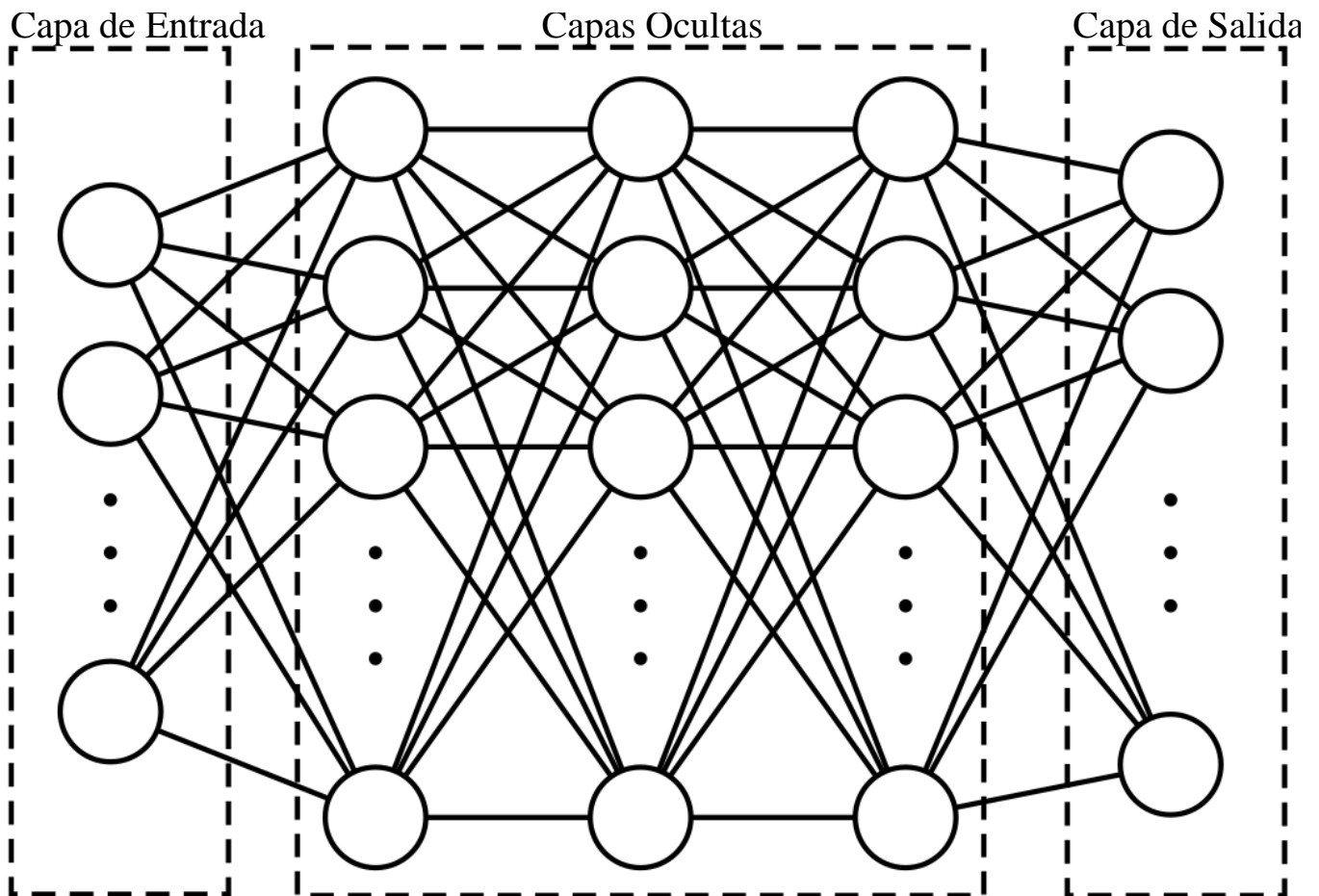


Figura 3.3: Arquitectura genérica de una red neuronal prealimentada. El cuadro izquierdo encierra la capa de entrada, el de en medio a las capas ocultas y el de la derecha encierra a la capa de salida que proporciona el resultado deseado y que se espera la máquina aprenda a predecir dada la capa de entrada.

dejando en claro que la función de activación es la misma para todos los nodos en la capa. Ahora denotaremos a la entrada de la red neuronal como x y la salida de la misma como \hat{y} . En general, nuestra RN tendrá $s + 1$ capas y comenzaremos contando desde cero, de tal manera que la capa de entrada está indexada por $i = 0$ y la capa de salida por $i = s$. Por la notación que tomamos en la (3.2) para la capa i -ésima, tenemos entonces que $x = l^{(0)}$ y $\hat{y} = l^{(s)} = h_{\theta}(x)$, donde ya habíamos visto que h_{θ} es la acción completa de la RN sobre x y que aproxima lo mejor posible a la salida \hat{y} con ayuda de los parámetros θ . De la misma forma, denotaremos la función de activación de la capa i -ésima por $a^{(i)}$ como en (3.2). Entonces, el mapeo de la RN h_{θ} tendría una acción neta sobre

vectores de entrada \mathbb{R}^{n_0} y los mapea a vectores de salida que viven en \mathbb{R}^{n_s} ,

$$h_\theta : \mathbb{R}^{n_0} \rightarrow \mathbb{R}^{n_s}, \quad (3.3)$$

$$\begin{aligned} l_{\mu_s}^{(s)} &= h_\theta(x) \\ &= a^{(s)} \left(\sum_{\mu_{s-1}=1}^{n_{s-1}} w_{\mu_s \mu_{s-1}}^{(s)} a^{(s-1)} \left(\dots a^{(1)} \left(\sum_{\mu_0=1}^{n_0} w_{\mu_1 \mu_0}^{(1)} l_{\mu_0}^{(0)} + b_{\mu_1}^{(1)} \right) \dots \right) + b_{\mu_s}^{(s)} \right). \end{aligned} \quad (3.4)$$

Generalmente, en ML, las entradas del vector de entrada son llamadas **características** y el espacio vectorial correspondiente \mathbb{R}^{n_0} es llamado **espacio de características**.

3.1.3. Funciones de activación

Las funciones de activación suelen ser funciones no lineales de los nodos, y se procura que sean **funciones unarias**. Las funciones unarias son aquellas que son aplicadas a la señal o salida de cada nodo μ de $l_\mu^{(i)}$ de manera separada. La elección de las funciones de activación depende de los objetivos de la RN, ya sea para hacer regresiones o clasificaciones, así como de la naturaleza de las características de entrada. Algunas de las funciones de activación unarias más populares son

- **Identidad:** $a : \mathbb{R} \rightarrow \mathbb{R}$, tal que

$$a(x) = x. \quad (3.5)$$

Estrictamente hablando, esta no es una función de activación. Además, carece de la introducción de no linealidad y provoca que la capa colapse con la siguiente; esto quiere decir que si tenemos una función de activación lineal $a^{(i)} = id$ en la i -ésima capa, entonces los dos productos de matrices que mapean desde la capa $i-1$ a i y desde la capa i a la $i+1$, puede ser escrita como un solo mapeo removiendo la capa i de la RN.

$$\begin{aligned} z_\mu^{(i+1)} &= \sum_{\nu=1}^{n_i} w_{\mu\nu}^{(i+1)} l_\nu^{(i)} + b_\mu^{(i+1)} = \sum_{\nu=1}^{n_i} w_{\mu\nu}^{(i+1)} a^{(i)}(z_\nu^{(i)}) + b_\mu^{(i+1)} \\ &= \sum_{\nu=1}^{n_i} \sum_{\rho=1}^{n_{i-1}} w_{\mu\nu}^{(i+1)} (w_{\nu\rho}^{(i)} z_\rho^{(i-1)} + b_\nu^{(i)}) + b_\mu^{(i+1)} \\ &= \sum_{\nu=1}^{n_i} \sum_{\rho=1}^{n_{i-1}} (w_{\mu\nu}^{(i+1)} w_{\nu\rho}^{(i)}) z_\rho^{(i-1)} + (w_{\mu\nu}^{(i+1)} b_\nu^{(i)} + b_\mu^{(i+1)}). \end{aligned} \quad (3.6)$$

- **Sigmoide logística:** $a : \mathbb{R} \rightarrow \mathbb{R}$, tal que

$$a(x) = \frac{1}{1 + e^{-x}}. \quad (3.7)$$

Esta función tiene amplias aplicaciones en la dinámica de poblaciones y está relacionada con los impulsos eléctricos de las neuronas ya sea para transmitir o recibir una señal. La sigmoide logística también se puede utilizar en aplicaciones de ajustes con el método llamado *máxima verosimilitud*. Una gran ventaja de esta función es que su derivada puede expresarse en términos de sí misma, es decir,

$$a'(x) = a(x)(1 - a(x)). \quad (3.8)$$

Más adelante hablaremos del proceso de entrenamiento con redes neuronales, y se verá que durante el procedimiento se necesita la razón de cambio de las funciones de activación utilizadas en cada una de las capas. El hecho de que la derivada de la sigmoide logística esté expresada en términos de sí misma permite que el proceso de entrenamiento sea más rápido. Una desventaja de la sigmoide logística es que podría llegar a atascarse durante el entrenamiento. Ya que esta función mapea valores negativos a valores cercanos a cero, tanto la función de activación como su gradiente se aproximan a cero. Como consecuencia, los nodos siguientes podrían no ser activados. Más aún, la actualización de los parámetros durante el entrenamiento, que son proporcionales a las derivadas, podría ser muy pequeña también en ese escenario, lo que hace muy complicado que la RN abandone una región inviable en el espacio paramétrico. Para evitar este comportamiento sería conveniente tener inicialmente una idea de los valores para los pesos $w^{(i)}$ de tal manera que la función de activación no reciba nodos con valores negativos.

- **Función hiperbólica Tanh:** $a : \mathbb{R} \rightarrow (-1, 1)$, tal que

$$a(x) = \tanh(x). \quad (3.9)$$

La forma de la gráfica de esta función es similar a la sigmoide logística, pero mapea a todos los reales al intervalo $(-1, 1)$. Como en el caso de la sigmoide logística, la derivada de esta función puede ser expresada en términos de sí misma, tal que

$$a'(x) = (1 - a(x)^2), \quad (3.10)$$

que como en el caso anterior, hace que el entrenamiento sea un proceso más rápido. A diferencia del caso anterior, esta función tiene menos tendencia a atascar el proceso de entrenamiento de las RN, pues sólo una pequeña región alrededor del cero es mapeada a valores nulos y así es poco probable que los nodos siguientes no sean activados.

- **Función rectificadora ReLU:** $a : \mathbb{R} \rightarrow \mathbb{R}$, tal que

$$a(x) = \begin{cases} x & \text{si } x \geq 0 \\ -cx & \text{si } x < 0. \end{cases} \quad (3.11)$$

Si, en esta ecuación tenemos el caso que $c = 0$ tenemos la función de activación más sencilla, su derivada es simplemente una constante, aunque no es derivable en $x = 0$. Tenemos entonces que si $c = 0$

$$a'(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x < 0. \end{cases} \quad (3.12)$$

La derivada es constante para valores positivos y durante el entrenamiento el valor de la misma nunca baja mientras nos acercamos a los valores de ajuste para los parámetros. Esta función, al igual que el caso de la sigmoide logística, su nulidad en los valores $x < 0$ puede provocar un atascamiento en el proceso de entrenamiento. Para combatir este problema, muchas veces se opta por fijar a la constante c en valores muy pequeños, de tal manera que no se anule su valor para x negativas.

- **Función softmax:** $a : \mathbb{R}^{n_i} \rightarrow (0, 1]$, tal que

$$a(x_\mu) = \frac{e^{x_\mu}}{\sum_{\nu=1}^{n_i} e^{x_\nu}}. \quad (3.13)$$

Estrictamente hablando, esta función no es unaria, pues no se aplica de manera independiente a cada uno de los nodos, si no que depende también de los valores del resto de los nodos en la capa para llevar a cabo una normalización, así asegurándose que la suma de todos los valores obtenidos de la acción sobre cada nodo sume un total de uno. Este tipo de función permite una interpretación probabilística en los valores de salida. Así que generalmente, esta función de activación se aplica a la capa de salida para propósitos de clasificación. Si en el problema de clasificación tenemos dos clases, decimos que tratamos con una **clasificación binaria**, pero si tenemos más de dos clases el problema se llama **clasificación de clases múltiples**. Para el caso de clasificación binaria suele recurrir a funciones como la sigmoide logística (donde tenemos dos clases indexadas por 0 y 1) o la función hiperbólica Tanh (donde las clases están indexadas por ± 1). Como en el caso de las funciones anteriores, la derivada es muy simple, pues puede expresarse en términos de sí misma,

$$\partial_{x_\mu} a(x_\nu) = a(x_\mu)(\delta_{\mu,\nu} - a(x_\nu)), \quad (3.14)$$

lo cual también contribuye a que el entrenamiento sea más rápido. En la Figura 3.4 apreciamos las 4 funciones que acabamos de explicar con sus respectivas derivadas.

- **Función Selu:** $a : \mathbb{R} \rightarrow \mathbb{R}$, tal que

$$a = \lambda \begin{cases} x & \text{si } x > 0 \\ \alpha e^x - \alpha & \text{si } x \leq 0. \end{cases} \quad (3.15)$$

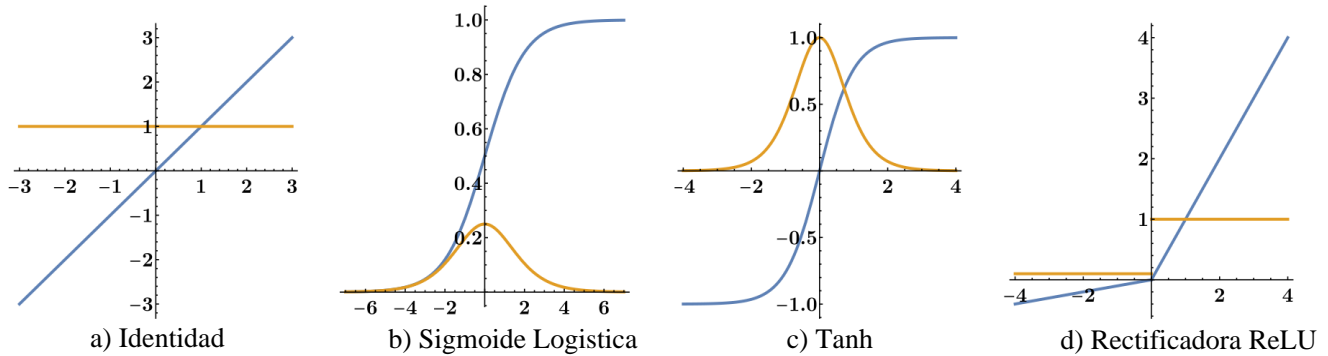


Figura 3.4: Gráficas de 4 funciones de activación y sus respectivas derivadas. En azul tenemos las gráficas y en dorado las gráficas de las correspondientes derivadas.

Por sus siglas en inglés *scaled exponential linear units*, es una función ampliamente utilizada en el contexto de la redes neuronales auto normalizadas, que detallaremos un poco más adelante. Sus ventajas y amplia utilización actual las explicaremos junto con el concepto de las redes neuronales auto normalizadas, pues van de la mano para una mejor implementación y la vamos mencionando aquí. Esta será la función de activación que más utilizaremos en nuestras redes neuronales para efectos de este trabajo. La gráfica de esta función y su respectiva derivada se pueden apreciar en la Figura 3.7.

3.1.4. Entrenamiento y evaluación de redes neuronales

Recordemos que el conjunto de parámetros de las redes neuronales los denotamos por θ y es el conjunto de todos los valores de los pesos y los biases, también pueden llegar a incluirse otros parámetros. Para poder entrenar una red neuronal, primero debemos especificar lo que se quiere obtener, es decir, qué salidas son las deseadas para determinadas entradas a la RN. Una vez determinadas entradas y salidas se modifican los parámetros θ hasta que el resultado obtenido sea razonablemente cercano al deseado. Se puede especificar un ejemplo en el contexto de ML supervisado, donde a una RN se le da de entrada un conjunto de N vectores x , cada uno de ellos con dimensión n_0 y sus correspondientes salidas u objetivos vectores y n_s -dimensionales. Durante el entrenamiento, la red va variando los parámetros de tal forma que la salida $\hat{y} = l^s = h_\theta(x)$ se acerque a los verdaderos valores deseados y . Este procedimiento se hace frecuentemente utilizando **gradiente descendente**, que describiremos en unos momentos. La función de **pérdida** o de **costo** L de la RN está dada como la desviación del valor de salida real y y el resultado obtenido por la RN \hat{y} .

El proceso de entrenamiento requiere que se vaya reduciendo la función de pérdida, o bien, optimizar una función no lineal en un espacio paramétrico con dimensión muy alta. Por ejemplo,

al actualizar el valor de un solo peso $w_{\mu\nu}^{(i+1)}$ cambiará el valor del nodo $l_{\mu}^{(i+1)}$ de la siguiente capa. Adicionalmente, si tenemos una RN completamente conectada, entonces el cambio en ese peso cambiará el valor de todos los nodos subsecuentes $l^{(k)}$ con $k > i + 1$ y por ende influenciará el valor de todos los pesos subsecuentes. Es importante hacer notar que este proceso de optimización es demasiado complicado debido al crecimiento exponencial del problema por la gran cantidad de parámetros a tomar en cuenta.

Durante el entrenamiento en ML supervisado, la RN tiene un par entrada-salida. El objetivo de entrenar a la RN es que eventualmente podamos proveerle información o datos para los cuales la salida o resultados no sean conocidos, y los prediga en función de lo que aprendió. Para evaluar qué tan bien se desempeña la RN en datos no conocidos por la misma, es importante reservar una cantidad determinada de pares entrada-salida para evaluar a la RN entrenada y así hacer una evaluación final de la eficacia de la RN. Para esto, los datos se dividen en dos conjuntos: el **conjunto de entrenamiento** y el **conjunto de prueba**. El cómo se dividen los datos en estos dos conjunto depende de la cantidad de información disponible, pero una división común es la siguiente en porcentajes

$$\text{entrenamiento} : \text{prueba} = 90 : 10. \quad (3.16)$$

A veces también se utiliza un **conjunto de validación**, con una división comúnmente tomada como

$$\text{entrenamiento} : \text{validación} : \text{prueba} = 80 : 10 : 10. \quad (3.17)$$

La idea de tener un conjunto de validación es para ir monitoreando el entrenamiento de la RN, es decir, poder ir evaluando al modelo de ajuste de parámetros de la red de manera no sesgada. La RN es entrenada con el conjunto de entrenamiento, que contiene la mayor parte de los datos, evaluada durante el entrenamiento con la ayuda del conjunto de validación y evaluada finalmente con el conjunto de prueba. La eficacia de la RN debe ser comparable tanto en el conjunto de entrenamiento y el conjunto de validación, de lo contrario los parámetros o hiperparámetros de la RN han sido tuneados sobreajustando dicha red. Esto quiere decir que la red aprendió bastante de los datos de entrenamiento y lo hizo tan bien, que sólo logra reproducir esos datos y no conjuntos externos ajenos al proceso.

Una vez que el entrenamiento y la evaluación son completados para la RN, ésta puede ser utilizada para procesar datos nuevos. La RN puede ser entrenada con todos los datos que se tengan a disposición. Como regla general, entre más datos sean utilizados durante el entrenamiento, mejor será su eficacia. Sin embargo, en caso de hacer eso, como uno no puede evaluar su rendimiento, entonces se debe tener cuidado de no subajustar o sobreajustar a la red.

El problema de sub o sobreajuste de un conjunto específico para futuras predicciones es conocido como el **problema de sesgo-varianza** en ML supervisado. Vale la pena utilizar un poco de estadística, pues se está intentando hacer futuras predicciones basadas en datos previamente disponibles. En estadística, el **sesgo** es un estimador de la diferencia entre el valor esperado $\mathbb{E}[\cdot]$, mientras que la **varianza** es el valor esperado del cuadrado de las desviaciones desde el valor promedio. De-

notando la función que estamos tratando de aprender a aproximar por $f(x)$ y su aproximación $\hat{f}(x)$, tenemos que

$$\text{Sesgo}[\hat{f}(x)] = \mathbb{E}[\hat{f}(x)] - \mathbb{E}[f(x)], \quad \text{Var}[\hat{f}(x)] = \mathbb{E}[\hat{f}(x)^2] - \mathbb{E}[\hat{f}(x)]^2, \quad (3.18)$$

o en notación de Bra-Kets

$$\text{Sesgo}[\hat{f}(x)] = \langle \hat{f}(x) \rangle - \langle f(x) \rangle, \quad \text{Var}[\hat{f}(x)] = \langle \hat{f}(x)^2 \rangle - \langle \hat{f}(x) \rangle^2. \quad (3.19)$$

Con esto, podemos escribir el valor esperado del error para datos no vistos x como

$$\langle (f(x) - \hat{f}(x))^2 \rangle = \left(\text{Sesgo}[\hat{f}(x)] \right)^2 + \text{Var}[\hat{f}(x)]. \quad (3.20)$$

Así que para reducir el error necesitamos reducir el sesgo o incrementar la varianza. Modelos sobreajustados tendrán un sesgo pequeño (predecirán los datos muy bien), pero una varianza muy alta (oscilan mucho para poder interpolar entre los datos de entrenamiento). Modelos subajustados no serán lo suficientemente complejos para modelar la función implícita dentro de la red, así que el sesgo será grande, pero tampoco fluctúan mucho de los datos, así que la varianza será pequeña. La complejidad del modelo está relacionada con la arquitectura de la RN, es decir, el número de nodos y capas que configuran a la red en cuestión. La Figura 3.5 ilustra estos conceptos en un diagrama.

Ahora hablemos del proceso de entrenamiento. Como ya se mencionó anteriormente, el objetivo de ML es poder conocer características de grandes cantidades de datos. Así que algunos métodos de entrenamiento actualizan los valores de los pesos después de que todos los datos de entrenamiento han sido vistos por la RN, pero otros métodos actualizan los valores de los pesos después de pasos intermedios en los cuales sólo se analizan partes específicas del conjunto de entrenamiento total. Los conceptos básicos del proceso de entrenamiento son los siguientes:

- Una **época** se refiere a completar el procesamiento de todo el conjunto de entrenamiento una vez, es decir, es una iteración sobre dicho conjunto. El número de épocas depende del problema en cuestión, una sola época no es suficiente, en general. Después de cada época, el conjunto de entrenamiento es revuelto aleatoriamente. En algunos casos es conveniente ordenar el conjunto de "fácil" a "difícil". Para identificar esa cualidad de "dificultad" se debe tener cierto conocimiento previo sobre la naturaleza del conjunto.
- Un **mini-batch** o llamado a veces simplemente **batch** (en español podríamos llamarlo lote) contiene un subconjunto del conjunto de entrenamiento. Al dividir el conjunto de entrenamiento en lotes, el entrenamiento se hace más óptimo. Un lote es un ensamble de datos, el cual se utiliza para calcular la función de pérdida en una iteración y a partir de dicho cálculo, modificar los valores de los pesos. Entre más grande es el lote, más podremos capturar el comportamiento de grupo. El tamaño de lote N_B se fija dependiendo de los datos en cuestión y de la cantidad de información con la que contamos y es la cardinalidad de los subconjuntos

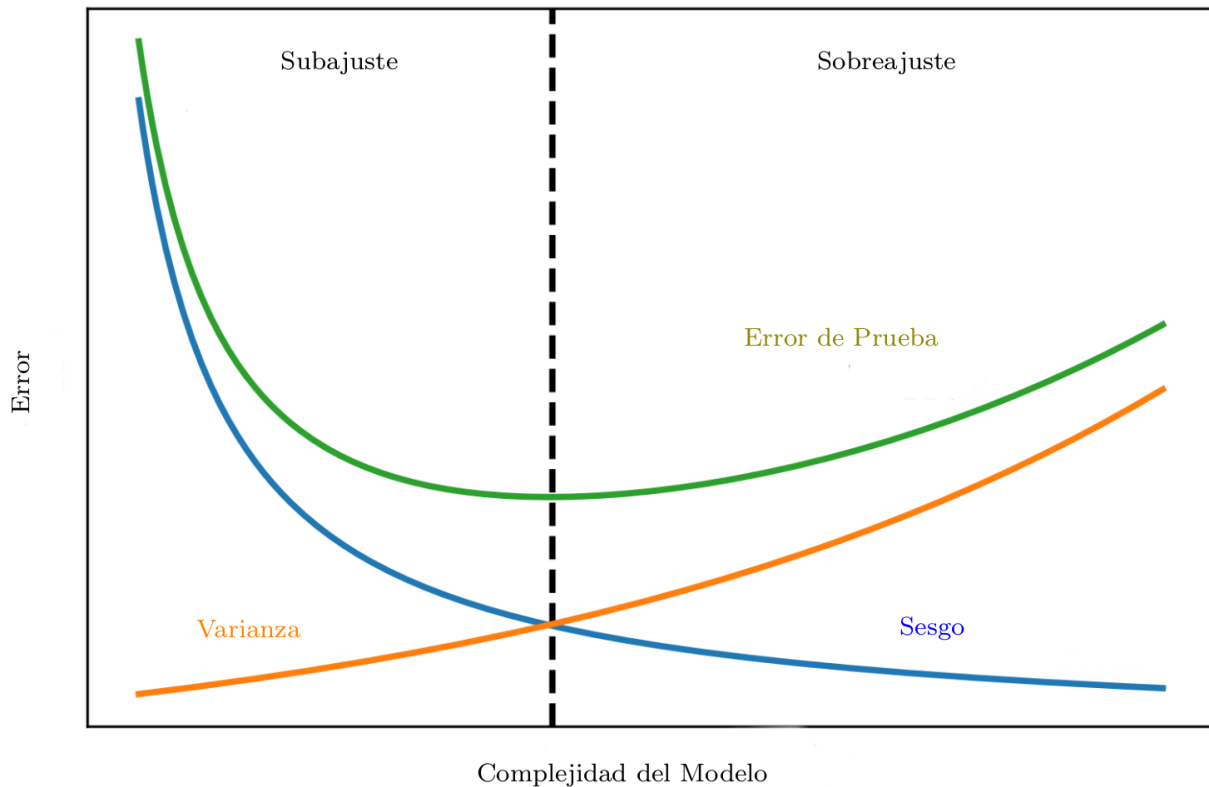


Figura 3.5: El error total al cuadrado en el conjunto de prueba no visto por la RN en relación con la varianza y el sesgo del entrenamiento. Para evitar tener sobreajuste debemos evitar que el error total en el conjunto del entrenamiento comience a crecer.

que dividen al de entrenamiento. También tenemos el número de lotes, que es el número de subconjuntos por cual queremos dividir al conjunto de entrenamiento total. En concordancia con el tamaño de lote, si queremos que éste último sea pequeño, entonces el número de batches debe ser grande.

- Una **iteración** es el acto de procesar un mini-batch. Entonces, denotando la cardinalidad de todo el conjunto de entrenamiento por N_T , el número de iteraciones en una época es N_T/N_B .

En resumen, para diferentes bases de datos, el número de épocas requerido es diferente. Sin embargo, el tamaño de la época está relacionado con la diversificación del conjunto de datos. Entre más diversificación se tenga, más grande debe ser la época. El llamado tamaño de lote es una parte de los datos que es enviada a la RN para entrenarla a cada tiempo, y es el número de muestras de entrenamiento en cada lote. Es crucial escoger un tamaño de lote adecuado para el entrenamiento pues puede incrementar la eficacia del proceso de entrenamiento de la RN. La Figura 3.6 muestra

un ejemplo de la convergencia de los hiperparámetros (asumiendo que son sólo dos) dependiendo del tamaño de batch. [30]

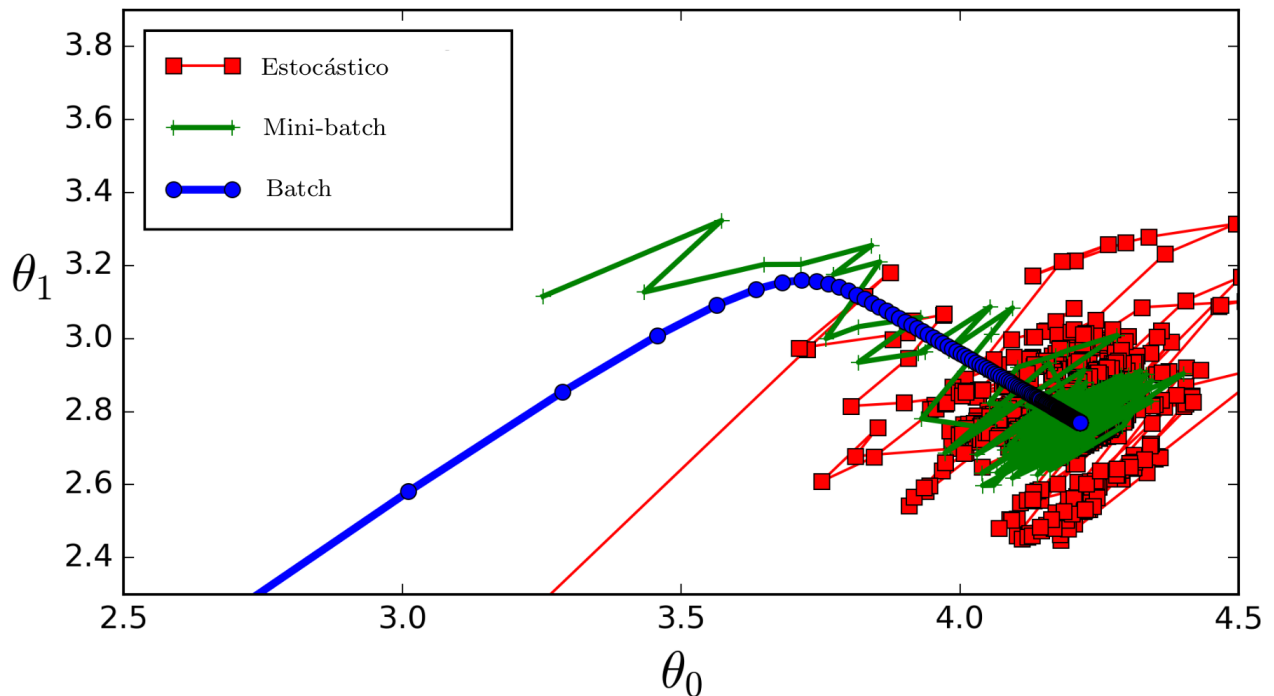


Figura 3.6: Variación de los hiperparámetros para distintos valores del tamaño de batch. Para el caso azul, todos los datos son enviados a un solo batch, y por tanto contiene a todas las muestras para el entrenamiento. La verde es el método de entrenamiento por mini-batch, es decir, todos los datos son divididos en muchos batches y cada uno contiene una pequeña parte del entrenamiento. La roja corresponde a entrenamiento aleatorio, es decir, hay una sola muestra en cada batch. [31]

Podemos apreciar de la Figura 3.6 vemos que en azul se aprecia la variación de los parámetros de la RN cuando todos los datos son utilizados para el entrenamiento, es decir, tenemos un solo lote y el lote contiene a todas las muestras de entrenamiento. En verde apreciamos el caso de mini-batch (o mini lote), es decir, todos los datos son divididos en varios lotes y cada lote contiene una pequeña parte de las muestras de entrenamiento. En rojo se aprecia el entrenamiento aleatorio, es decir, cuando tenemos una sola muestra de entrenamiento en cada lote. En los tres casos se vislumbra la variación de los parámetros a lo largo de las épocas, y conforme avanza el entrenamiento, se busca minimizar la función de error (o de pérdida) y que cuantifica la diferencia entre los resultados deseados y obtenidos a partir de la RN (ver más adelante). Como ya se mencionó, el número de lotes se interpreta como el número de iteraciones. Para poner un ejemplo sencillo, si tenemos 2000 datos, divididos en 4 lotes, entonces el tamaño de lote es 500. Corremos todos los datos para el entrenamiento. Para completar una época, se requieren 4 iteraciones.

3.1.5. Gradiente descendente

El método más empleado para el entrenamiento de RNs es el de **gradiente descendente**. Como su nombre lo indica, el mínimo de la función de pérdida se busca y encuentra siguiendo la dirección del decremento más pronunciado desde un punto específico, es decir, a lo largo de $\nabla_{\theta} J$, donde J es la **función de pérdida total**. El gradiente descendente se detendrá una vez que un punto crítico sea alcanzado, ya que en ese punto el gradiente se anula. Un posible inconveniente es que la complicada forma de la función de pérdida puede llegar a tener muchos puntos silla, y se puede presentar el escenario en que el gradiente descendente alcance un punto crítico de esta naturaleza. Verificar la naturaleza del punto crítico alcanzado implica el cálculo del Hessiano de la función de pérdida en dicho punto, y como se trata de un determinante, la complejidad computacional puede elevarse bastante.

El entrenamiento de RNs se compone de dos pasos conocidos como **paso hacia adelante** y el **paso hacia atrás** también conocido **propagación hacia atrás**. Para la explicación de este proceso, notemos que antes de empezar el entrenamiento, necesitamos iniciar la RN. Asumiremos que la RN ha sido inicializada, acción que discutiremos un poco más adelante.

En el paso hacia adelante, la entrada x es representada en la capa de entrada $l^{(0)}$ de la RN. Entonces los datos son propagados hacia adelante por medio de multiplicaciones de matrices, agregando el bias y aplicando la función de activación α_i , hasta que todos los valores $l^{(i)}$, $i = 0, 1, \dots, s$ hayan sido calculados y guardados. En ML supervisado, tenemos un par entrada-salida (x, y) , así que podemos comparar qué tanto difiere la salida $\hat{y} = l^{(s)}$ para una determinada entrada $x = l^{(0)}$. Esta diferencia J se calcula con la acción conjunta de una función de pérdida L más un término consistente de un **regularizador** o **decaimiento de los pesos** r ,

$$J(\theta, x, y) = L(\theta, x, y) + \lambda r(\theta). \quad (3.21)$$

Más adelante discutiremos algunas de las opciones que tenemos para la función de pérdida. El parámetro λ es llamado **parámetro de decaimiento de los pesos**. La elección para el valor de λ depende del problema en cuestión, pero usualmente tiene un orden de valores $\lambda \sim O(10^{-3})$. También existen múltiples elecciones para la función de decaimiento $r(\theta)$, siempre y cuando tengan la propiedad de conducir a que los valores de los pesos sean pequeños, además estas funciones sólo se encargan de regular la variación de los valores de los pesos y no del resto de los parámetros (tales como los biases) de la RN. La inclusión de estas funciones ayudan a prevenir el sobreajuste de las RNs.

En una iteración del gradiente descendente, los parámetros de la RN son actualizados. Si denotamos por $\theta^{(i)}$ a los parámetros de toda la i -ésima capa, la actualización de los mismos estará dada por

$$\theta_a^{(i)} \rightarrow \theta_a^{(i)} - \alpha \frac{\partial J(\theta, x, y)}{\partial \theta_a^{(i)}}. \quad (3.22)$$

El muy importante parámetro α se llama **tasa de aprendizaje**. Si α tiene un valor muy pequeño,

a la RN le toma mucho tiempo converger (a los parámetros). Si α tiene un valor muy grande, podría ser que se pase de la localización del mínimo y ascender o descender hacia la dirección equivocada. Algunos métodos también adaptan el valor de la tasa de aprendizaje durante el entrenamiento.

En el proceso de actualización de los parámetros, la desviación $\delta^{(s)}$ entre la salida deseada y y el valor real de la salida por la RN $l^{(s)} = h_{\theta}(x)$ se calcula y luego se propaga hacia atrás a través de la red para obtener recursivamente las otras desviaciones $\delta^{(i)}$ y las correspondientes actualizaciones hacia $\theta^{(i)}$ para las otras capas. En detalle, usamos los mapeos sucesivos capa a capa y con la regla de la cadena podemos calcular que

$$\frac{\partial J}{\partial \theta_a^{(i)}} = \frac{\partial L(\theta, x, y)}{\partial z^{(i+1)}} \frac{\partial z^{(i+1)}}{\partial \theta_a^{(i)}} + \lambda \frac{\partial r(\theta)}{\partial \theta_a^{(i)}}. \quad (3.23)$$

Recordemos que z es el vector obtenido después de aplicar el mapeo lineal entre capas pero antes de aplicar la función de activación. El segundo término y el segundo factor del primer término se calculan fácilmente. El primer factor del primer término, que esencialmente son las derivadas de la función de activación, proporcionan las desviaciones $\delta^{(i)}$ de las otras capas ocultas. Estas son utilizadas para propagar hacia atrás $\delta^{(s)}$ a través de las capas intermedias hasta $\delta^{(0)}$. Como ya se mencionó anteriormente, existen múltiples elecciones para las funciones de activación, algunas de las cuales tienen derivadas simples o proporcionales a sí mismas. Como la función de activación ha sido evaluada y los valores han sido guardados durante el paso hacia adelante, el resultado puede reutilizarse en el paso hacia atrás. Habiendo obtenido todas las desviaciones $\delta^{(i)}$, podemos finalmente calcular el gradiente $\nabla_{\theta} J$ y actualizar los valores de los parámetros.

3.1.6. Inicialización de parámetros

Para poder llevar a cabo los pasos hacia adelante/atrás, necesitamos inicializar los pesos. Una opción muy simple es inicializarla con puros ceros, sin embargo, establecer a todos los pesos con el mismo valor significa que todos los gradientes serán los mismos, lo que llevará a actualizaciones de los valores para los parámetros a ser las mismas. Así que lo más conveniente es inicializar los valores de los pesos con todos sus valores diferentes.

Dependiendo de la función de activación, también hay un problema de **gradientes nulos** o **gradientes divergentes**. Por otro lado, pesos con valores grandes también pueden conducir a grandes fluctuaciones de errores y por ende gradientes grandes, que son un problema para la convergencia de los parámetros. La regularización de los pesos que mencionamos en 3.1.5 contrarrestan estos efectos.

Una elección sencilla para la inicialización de los pesos es generando números aleatorios a partir de una distribución normal (o Gaussiana) con centro en cero y una varianza unitaria, es decir,

$$\theta \sim \mathcal{N}(0, 1), \quad (3.24)$$

debido a una falta de conocimiento previo de la verdadera forma en que se distribuyen los valores de los parámetros. Alternativamente, uno puede generar los parámetros a partir de una distribución uniforme,

$$\theta^{(i+1)} \sim \mathcal{U}\left(-\frac{1}{\sqrt{n_i}}, \frac{1}{\sqrt{n_i}}\right). \quad (3.25)$$

Nótese que el intervalo se escoge respecto al número de nodos que tiene la capa anterior. Existen otras distribuciones convenientes para la generación de los pesos.

Una técnica que hace más robusta a la RN respecto a la inicialización de parámetros y las tasa de aprendizaje y utilizar normalización de batches. Esto significa que después de cada mini-batch, la entrada es reescalada para que tenga promedio nulo y varianza unitaria. Los cambios se realizan a lo largo de la RN completa por propagación hacia atrás. También actúa como un regularizador y puede salvarnos de sobreentrenamiento de manera similar a como ocurre con la regularización de pesos. Estas condiciones se desarrollan en el contexto de las llamadas **redes neuronales autonormalizadas** [32].

Otra forma de lograr la normalización, es decir, tener desviación cero y varianza unitaria, es utilizando la función SELU que ya referenciamos anteriormente. Una gran ventaja es que la función es continua en $x = 0$. Recordemos que la función SELU está definida como

$$\text{SELU}(x) = \lambda \begin{cases} x & \text{si } x > 0 \\ \alpha e^x - \alpha & \text{si } x \leq 0. \end{cases} \quad (3.26)$$

con, $\alpha, \lambda > 0$ usualmente. A diferencia de la sigmoide logística, la función SELU puede producir tanto valores positivos como negativos, que es necesario para tener un promedio cero. Para cambiar la varianza a la unidad, la función de activación necesita tener regiones que tengan pendiente casi nula así como regiones con pendiente mayor que uno, entonces es necesario que $\lambda > 1$. Los hiperparámetros α y λ necesitan ser ajustados para asegurar la propiedad de autonormalización. Si los pesos son inicializados con una distribución con desviación cero y varianza unitaria, entonces los parámetros más convenientes son

$$\alpha \sim 1.6733, \quad \lambda \sim 1.0507. \quad (3.27)$$

En la Figura 3.7 podemos apreciar la gráfica de esta función con estos parámetros y su respectiva derivada.

3.1.7. Funciones de pérdida

La elección de la función de pérdida, como era de esperarse, también depende de la aplicación de la red neuronal. Presentamos aquí algunas funciones de pérdida que son utilizadas frecuentemente en regresiones.

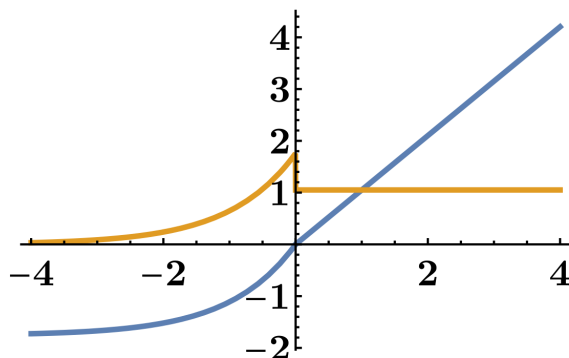


Figura 3.7: Gráfica de la función SELU (azul) y su respectiva derivada (dorado), para los valores $\alpha = 1.6733$, $\lambda = 1.0507$.

Funciones de error medio

Las funciones de error cuadrado medio, o por sus siglas en inglés, MSE son unas de las más utilizadas. Las funciones de error cuadrado medio se utilizan usualmente cuando las salidas de las RNs son valor numéricos (en lugar de interpretar la salida como probabilidades), siendo utilizadas entonces, por ejemplo en regresiones. La pérdida MSE se calcula como

$$L^{\text{MSE}}(y, \hat{y}) = \frac{1}{n_s} \sum_{i=1}^{n_s} (y_i - \hat{y}_i)^2. \quad (3.28)$$

A veces, la función de error raíz cuadrada medio es utilizado en lugar la MSE. Por supuesto que otras potencias también pueden utilizarse. Existe un problema al utilizar MSE junto con la función de activación sigmoide logística. Supongamos el caso más sencillo en que en el que tenemos una capa con un solo nodo y no vectores bias. En este caso $\theta = w$ y tenemos que el gradiente para la función de pérdida L^{MSE} respecto a θ es

$$\nabla_{\theta} L^{\text{MSE}}(x, y) = \frac{\partial L^{\text{MSE}}}{\partial w} = \frac{\partial [(y - \sigma(wx))^2]}{\partial w} = -2x(y - \sigma(wx))\sigma'(wx). \quad (3.29)$$

Recordemos que $\sigma'(z) = \sigma(z)(1 - \sigma(z))$, el término de la derivada es muy pequeño si el valor absoluto de wx es muy grande. Entonces, incluso si la diferencia $(y - \sigma(wx))$ es muy grande, la pérdida aún puede ser muy pequeña si el valor absoluto de wx es muy grande. Otras funciones de activación que no sufren este problema junto con la función de pérdida MSE.

Una alternativa a MSE es la función de error raíz cuadrada medio (MAE), que se calcula sumando el valor absoluto de la diferencia entre y_i y \hat{y}_i ,

$$L^{\text{MAE}}(y, \hat{y}) = \frac{1}{n_s} \sum_{i=1}^{n_s} |y_i - \hat{y}_i|. \quad (3.30)$$

La diferencia entre MSE y MAE es que MAE es más eficiente con los cálculos, pues la diferencia es lineal en lugar de cuadrática. Sin embargo, el gradiente para MSE se puede calcular más fácilmente que para MAE. Además, la función de pérdida MSE es más estable, en el sentido de que para determinados cambios de $y_i - \hat{y}_i$ la función MSE cambia menos que MAE.

Otra posibilidad es tomar la función de error logarítmica cuadrada, que por sus siglas en inglés se le conoce como MSLE,

$$L^{\text{MSLE}}(y, \hat{y}) = \frac{1}{n_s} \sum_{i=1}^{n_s} (\log(1 + y_i) - \log(1 + \hat{y}_i))^2. \quad (3.31)$$

Para pequeños valores, MSE y MSLE son comparables. La desventaja es que MSLE no está definida para valores de \hat{y} o y más pequeños que -1 . Finalmente, se tiene la **función de error de porcentaje absoluto**, o por sus siglas en inglés MAPE, y que está definida como

$$L^{\text{MAPE}}(y, \hat{y}) = \frac{1}{n_s} \sum_{i=1}^{n_s} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100. \quad (3.32)$$

Una desventaja clara es que esta función no está definida si algún valor de y_i es nulo.

Funciones de error con norma L_p

En lugar de tomar el promedio, podemos también utilizar la distancia entre los dos vectores \hat{y} y y en cualquier norma vectorial L_p , es decir,

$$L^{p\text{-NE}}(y, \hat{y}) = \|y - \hat{y}\|_p = \left(\sum_{i=1}^{n_s} (y_i - \hat{y}_i)^p \right)^{\frac{1}{p}}. \quad (3.33)$$

Para $p = 1$, esta es la norma valor absoluto, para $p = 2$ tenemos la distancia euclidiana, y para $p = \infty$, esta norma sólo da la componente máxima del vector.

Funciones de pérdida de entropía cruzada

Las funciones de pérdida de entropía cruzada, CEL por sus siglas en inglés, son utilizadas cuando la salida de las RNs son interpretadas como probabilidades como, por ejemplo, en tareas de clasificación. Para entender este tipo de funciones, debemos recordar de manera breve el concepto de entropía en el contexto de la teoría de la información. Dada una distribución de probabilidad P , la entropía de Shannon se define como

$$S(P) = \mathbb{E}_{x \sim P}[-\log_2 P(x)], \quad (3.34)$$

donde $x \sim P$ denota que x se obtiene de la distribución de probabilidad P . Como ejemplo simple, consideremos el caso discreto donde se tiene un conjunto P de cardinalidad N . Sean $K < P$ elementos diferentes que particionan P en K clases de elementos c_α , $a = 1, \dots, K$, cada una con n_α elementos. Entonces definimos $p_\alpha = n_\alpha/N$ y la entropía se calcula como

$$S(P) = - \sum_{a=1}^K p_\alpha \log_2(p_\alpha). \quad (3.35)$$

La entropía es una medida de cuántos bits son necesarios en promedio para describir la distribución de los valores x en X . La **entropía cruzada** es la comparación de dos distribuciones P y Q que están definidas con el mismo soporte promediando respecto a P pero contando los bits respecto a Q

$$S(P, Q) = \mathbb{E}_{x \sim P}[-\log_2 Q(x)]. \quad (3.36)$$

Notemos que $S(P, P) = S(P)$. en el caso discreto, la ecuación anterior se escribe simplemente como

$$S(P, Q) = - \sum_{a=1}^K p_\alpha \log_2(q_\alpha), \quad (3.37)$$

donde p_α y q_α son las frecuencias relativas de las clases en P y Q , respectivamente. La razón del por qué la teoría de la probabilidad entra en las RNs es debido a que podemos tener a la RN adivinando qué distribución o qué función está detrás de los datos discretos y finitos presentes durante el entrenamiento.

En el caso de la clasificación binaria, donde $y \in \{0, 1\}$ sólo puede estar en dos clases, la **función de pérdida de entropía cruzada binaria**, o BCEL por sus siglas en inglés, es simplemente

$$L^{\text{BCEL}}(y, \hat{y}) = -[y \log_2(\hat{y}) + (1 - y) \log_2(1 - \hat{y})]. \quad (3.38)$$

La salida $\hat{y} \in [0, 1]$ de la RN se interpreta como la probabilidad de que alguna entrada pertenezca a la clase cero o a la clase uno, donde $\hat{y} = 1/2$ indica completa incertidumbre al respecto. Notemos que como y es o cero o uno, sólo uno de los dos términos en (3.38) contribuye. Debido al logaritmo, la pérdida es muy grande si la red predice la clase equivocada con gran certidumbre. Si, por ejemplo, $y = 1$ pero la red predice un valor de salida \hat{y} cercano a cero, entonces el error se convierte en el logaritmo de un número pequeño (el segundo término en la suma es idénticamente cero) y por tanto el valor del error explota. También se puede definir una función de error para los casos de clasificación si se tienen múltiples clases, más de dos.

Existen otro tipo de funciones de pérdida, tales como la **pérdida de Hinge**, **pérdida de Kullback-Leiber** o la **pérdida de Huber** que no detallaremos aquí. La función de pérdida utilizada depende del uso que se le esté dando a la RN, ya sea para regresiones o clasificaciones. Las pérdidas para regresiones se basan en la distancia entre el valor real de la salida y y el valor predicho \hat{y} por la red. Para clasificaciones, las pérdidas tratan de minimizar la diferencia de entropías entre

la distribución de las clases del conjunto de entrenamiento y el utilizado por la RN para predecir las clases. Como las derivadas de las pérdidas aparecen en el gradiente descendente, todas las funciones de pérdida deberían tener derivadas que sean fáciles de calcular y que sean bien comportadas (no divergencias).

Las funciones de pérdida aquí mencionadas son usualmente utilizadas para problemas específicos a los que está destinada la RN en cuestión. Por ejemplo, las funciones MSE, MAE y MSLE son usualmente empleadas para análisis de regresión. Como ya mencionamos anteriormente, para problemas de clasificación binaria, la función de pérdida de entropía cruzada binaria es utilizada comúnmente, también suele utilizarse la función de pérdida de Hinge. Para problemas de clasificación de clases múltiples se emplean las funciones de entropía cruzada para clases múltiples o la de Kullback-Leiber. En este trabajo, destinaremos la utilidad de las RNs a la tarea de regresión, pues diseñaremos una RN para modelar un problema que implica predecir valores reales. De entre las tres opciones que tenemos (MSE, MAE y MSLE) para la función de pérdida, MAE y MSE son más rápidas de calcular, pues sólo implica el cálculo de diferencias de valores reales. La función MSE tiene la ventaja de que, al tener una ecuación de forma cuadrática, el proceso de gradiente descendente sólo buscará converger a un mínimo global pues la función cuadrática en cuestión no posee mínimos locales; además penaliza al modelo por tener errores grandes pues dichos errores los eleva al cuadrado. Por otro lado, la función MAE maneja mejor valores atípicos (errores grandes) por tratarse sólo del promedio de valores absolutos, sin embargo, entre sus desventajas se encuentra el hecho de que es computacionalmente costoso pues implica utilizar de manera implícita la función operador de módulo. Además, en el caso de la función MAE, podemos encontrar mínimos locales, lo cual puede entorpecer el proceso de entrenamiento de la RN.

3.1.8. Sobreentrenamiento y sobreajuste de Redes Neuronales

El especial interés por las Redes Neuronales yace en la capacidad que tienen de realizar aproximaciones no lineales. Redes con una sola capa oculta (y también con mayor número de capas) pueden interpolar cualquier función multidimensional con una precisión dada y pueden implementar de manera satisfactoria un conjunto finito de entrenamiento arbitrario. Sin embargo, enorme capacidad de las RNs en la aproximación no lineal puede resultar fácilmente en una mala generalización de la RN entrenada a nuevos datos; dando lugar a efectos azarosos. Ya que una RN con el suficiente número de nodos en las capas ocultas pueden implementar perfectamente un conjunto de entrenamiento arbitrario, puede aprender tanto las dependencias buscadas como un ruido inherente a los datos que disminuirán la capacidad predictiva de la red. Dos condiciones contribuyen en el problema [33]:

- arquitectura de la RN (complejidad del modelo),
- tiempo de entrenamiento de la RN.

El problema de sobreajuste se refiere al exceso del tamaño de la RN (número de capas y nodos),

mientras que el sobreentrenamiento se refiere al tiempo de entrenamiento de la RN que finalmente puede conducir a una peor capacidad de predictibilidad de la misma.

En 3.1.4 describimos los detalles y conceptos del proceso de entrenamiento de una RN y diseño de arquitectura de la misma. Allí se presentó el problema de sesgo-varianza en ML supervisado, el cual consiste en cómo identificar cuando tenemos sub o sobreajuste de una RN. El objetivo es que la complejidad del modelo de RN sea tal que el error en la predicción de valores sea mínima, o equivalentemente reduzcamos el sesgo o se incremente la varianza. Modelos sobreajustados tendrán un sesgo pequeño (predecirán los datos muy bien), pero una varianza muy alta (oscilan mucho para poder interpolar entre los datos de entrenamiento). La complejidad del modelo está relacionada con la arquitectura de la RN, es decir, el número de nodos y capas que configuran a la red en cuestión. En otras palabras, el sobreajuste es el uso de modelos que incluyen más parámetros ajustables de los necesarios para la optimización, o bien, utilizando un enfoque más complicado que el óptimo en última instancia.

Por otro lado, hablando de sobreentrenamiento, tenemos que si el aprendizaje de una red neuronal se desarrolla por mucho tiempo (muchos ciclos de entrenamiento, por ejemplo, épocas), la RN se ajusta a comportamientos aleatorios muy específicos de los datos de entrenamiento que no tienen una relación causal con la función objetivo. En este proceso de sobreentrenamiento, la precisión de la RN sobre las muestras del conjunto de entrenamiento se incrementa, mientras que la precisión sobre muestras de los datos no vistos por la red se vuelve cada vez peor (ver Figura 3.8). En otras palabras, en un escenario de sobreentrenamiento, hemos ajustado tantas veces los parámetros de la RN con el mismo conjunto de entrenamiento, de tal manera que la red tiene resultados muy óptimos para dichos conjuntos (ha aprendido de memoria dichos datos) y por ende no será capaz de tener una predictibilidad aceptable para un conjunto de datos ajeno al de entrenamiento. Un resultado interesante en [33] es que probando con varias Redes Neuronales con una capa oculta, pudieron concluir que el sobreajuste no tiene influencia sobre la predictibilidad de las redes cuando se evita el sobreentrenamiento.

3.1.9. Los *Autoencoders*

Son un tipo particular de redes neuronales y como su nombre lo dice, los autoencoders aprenden a codificar las entradas. De entrada tienen un vector característica y la RN neuronal es entrenada para reproducir exactamente el mismo vector pero como salida, es decir, la RN aprende y pretende reproducir la función identidad. Sin embargo, la RN en cuestión es diseñada de tal forma que la información se tiene que propagar de manera profunda a lo largo de la RN, de tal manera que las capas intermedias tengan menos nodos que las capas de entrada y salida. Generalmente, la arquitectura de la RN de los autoencoder es simétrica, con la capa intermedia siendo la más pequeña, es decir, **la capa cuello de botella** a través de la cual debe pasar la información. La arquitectura genérica de un autoencoder está ilustrada en la Figura 3.9. Como la presencia del cuello de botella significa que la RN debe aprender a reproducir la entrada con pocas características disponibles, es

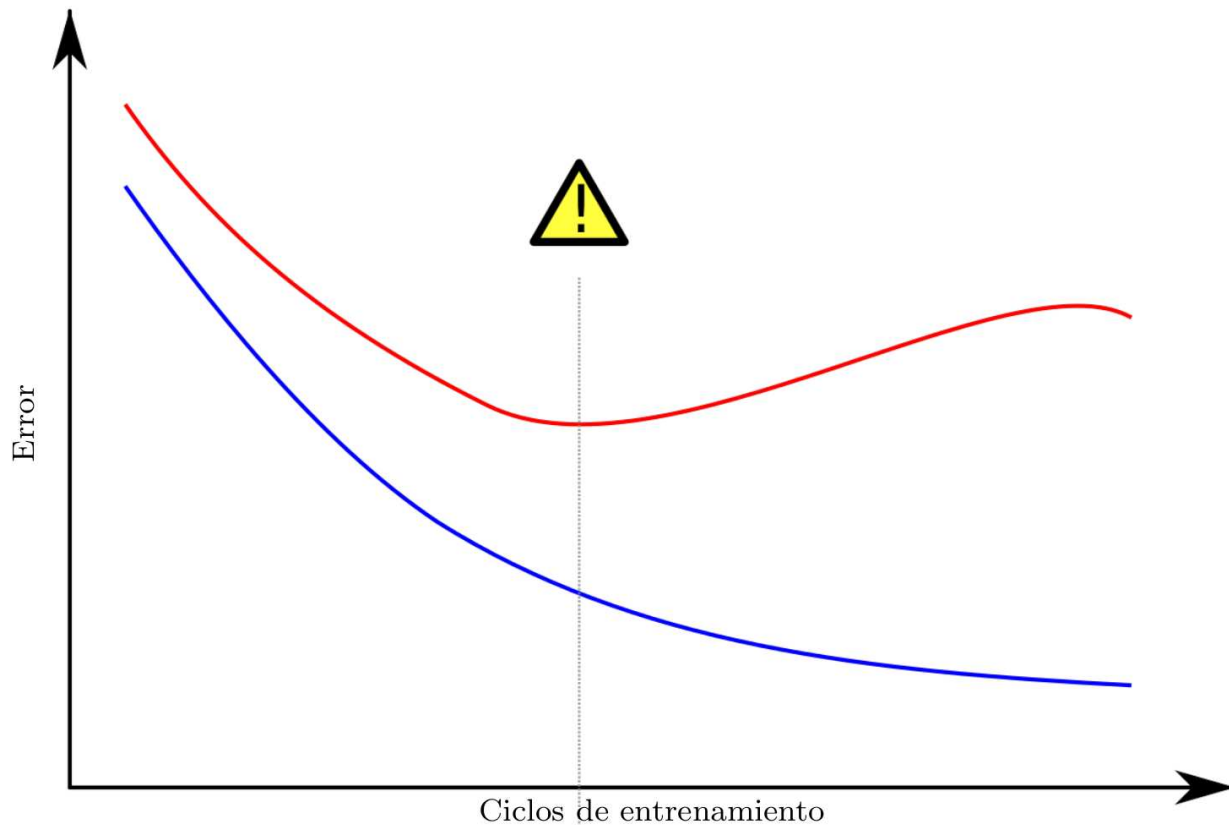


Figura 3.8: El sobreentrenamiento de ML supervisado. El error de entrenamiento se muestra en azul, el error de validación en rojo, ambos como función del número de ciclos de entrenamiento. Si el error de validación se incrementa (tasa de cambio positiva) mientras que el error de entrenamiento se mantiene en decremento (tasa de cambio negativa) entonces es posible que se haya presentado una situación de sobreentrenamiento. La mejor predictibilidad y ajuste se tendrá cuando el error de validación tiene su mínimo global. [34]

entrenada para codificar eficientemente la información contenida en el espacio de características. Uno no está interesado en la salida, pues es lo mismo que la entrada, sino la información comprimida en la capa cuello de botella. El estado oculto de la RN en el cuello de botella es utilizado para múltiples aplicaciones. Es una manera de **parametrizar** la información con menos grados de libertad respecto a la dimensión del espacio de características para la información de entrada.

Los tipos de funciones de activación que son utilizadas en autoencoders dependen de las características a codificar. El siguiente capítulo diremos qué funciones de activación utilizaremos en nuestro autoencoder. La arquitectura de la RN, especialmente el tamaño del cuello de botella, generalmente se determinan por prueba y error. La habilidad de la RN para reconstruir la entrada decrementará con un menor tamaño del cuello de botella. El punto crítico en el que uno se da

cuenta que se pierde mucha información en el cuello de botella y uno no puede reducir más su tamaño, depende de la aplicación en cuestión y de qué tan buena se necesita la codificación en el cuello de botella. La información de la capa cuello de botella puede ser utilizada para otras RNs para identificar la acumulación o clústers de modelos, o pueden ser analizados por medio de árboles de decisión, persistencia de homología u otras herramientas de análisis de datos. A continuación se presenta un ejemplo de la salida que tiene el orbifolder cuando se le pide generar de manera aleatoria modelos de orbifolios heteróticos con grupo de punto \mathbb{Z}_6 -II.

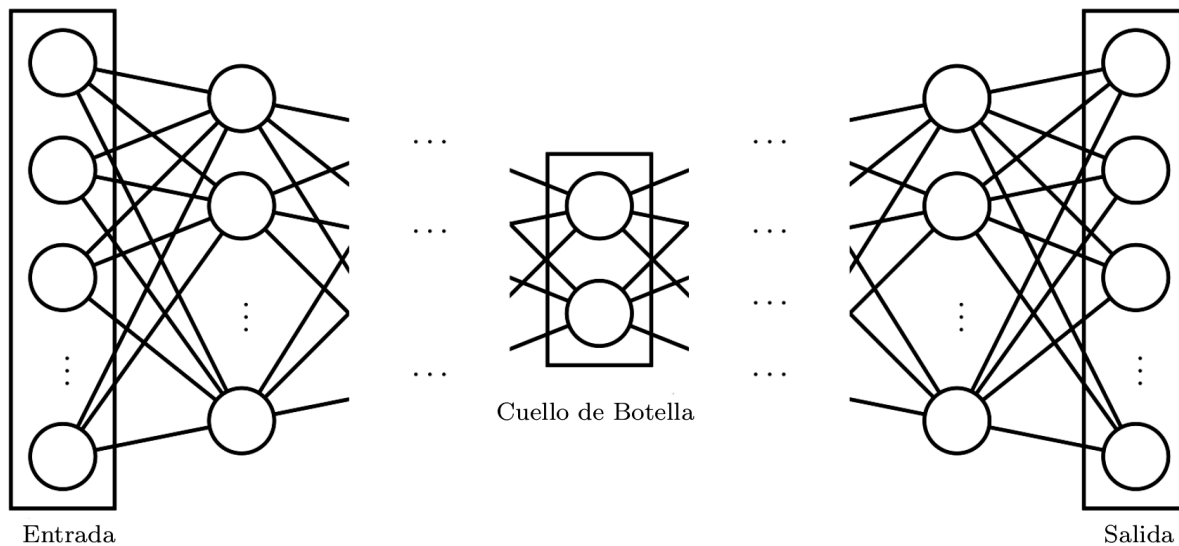


Figura 3.9: Estructura genérica de un autoencoder. Podemos apreciar las capas de salida y entrada, así como la capa cuello de botella que tiene el código de la información que se desea codificar.

3.2. El orbifolder

Como se dice en [35], el orbifolder es un programa desarrollado en C++ que calcula y analiza la teoría efectiva a bajas energías compactificadas en orbifolios heteróticos. El programa incluye rutinas que calculan el espectro no masivo, para identificar los acoplamientos en el superpotencial permitidos y para generar una gran cantidad de conjuntos de modelos en diferentes orbifolios para identificar modelos fenomenológicamente interesantes (como modelos que contengan un grupo de norma del Modelo Estándar Mínimo Supersimétrico), y para analizar sus configuraciones de vacío.

```
begin model
Label:Random1
SpaceGroup:Geometry/Geometry_Z6-II_1_1.txt
```

```

Lattice:E8xE8
Shifts and Wilsonlines:
-1/2, 0, 0, 0, 0, 0, 0, 0, 1/2, -1/12, -1/12, -1/12, -1/12, 1/12, 1/12, 1/12, 5/12
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
-1/6, -5/6, -1/6, 1/6, 1/6, 1/6, 1/6, -5/6, -1/3, 0, 0, 1/3, -2/3, -1/3, 0, 1
-1/6, -5/6, -1/6, 1/6, 1/6, 1/6, 1/6, -5/6, -1/3, 0, 0, 1/3, -2/3, -1/3, 0, 1
-5/4, -7/4, -7/4, -3/4, -1/4, 1/4, 5/4, -3/4, -3/2, 1, 2, 2, 1, 2, 2, 3/2
0, 1/2, 0, -3/2, -2, 3/2, -1, -1/2, 2, 2, -1, 3/2, 2, 3/2, -1/2, -3/2
end model

```

En el ejemplo de modelos generados por el orbifold que se presenta, vemos que se trata de un modelo aleatorio. En este caso, vemos que es un modelo generado aleatoriamente con un orbifold cuyo grupo de punto es \mathbb{Z}_6 -II, indicando que la retícula del toro para el grupo de norma es $E_8 \times E_8$. Después tenemos un total de 8 líneas que indican los parámetros correspondientes para el orbifold en cuestión. Las primeras dos líneas corresponden a los vectores de traslación, pero en nuestro caso sólo la primera línea tiene valores no nulos debido a la naturaleza del grupo de punto. Las siguientes 6 líneas representan los valores para las líneas de Wilson.

Esta valiosa herramienta permite investigar el paisaje de compactificaciones en orbifolds de la teoría heterótica de cuerdas. En este trabajo utilizamos el orbifold para generar $\mathcal{O}(4,000,000)$ modelos inequivalentes de orbifolds heteróticos con grupo de punto \mathbb{Z}_6 -II, es decir, $\mathcal{O}(4,000,000)$ combinaciones consistentes de vectores de traslación y líneas de Wilson que llevan a física inequivalente. Dos modelos son inequivalentes si se pueden distinguir los espectros no masivos a nivel de una teoría 4D, es decir, si tienen diferente grupo de norma y/o espectro no masivo diferentes. El orbifold permite elegir una generación de modelos inequivalentes.

Sin embargo, parametrizar a los modelos de orbifolds heteróticos por medio de sus vectores de traslación y líneas de Wilson tiene grandes redundancias, pues si éstos aparentan ser inequivalentes puede que difieran en realidad por:

- Combinaciones de enteros por la retícula de raíces en $E_8 \times E_8$,
- la acción del grupo de Weyl sobre $E_8 \times E_8$,

lo que en realidad los haría equivalentes. El grupo de simetría en este problema es muy grande, pues el grupo de Weyl de E_8 es enorme, con más de 10^{19} elementos. Cuando procesemos nuestros modelos con ayuda de un autoencoder, no utilizaremos directamente estos parámetros, si no que los transformaremos a una representación de vectores 16-dimensionales. Los detalles de esta transformación la describiremos en 3.3. Incluso en esta representación de nuestros modelos tendremos redundancias de las cuales no nos desharemos por completo y, de hecho, puede afectar entrenamiento del Autoencoder que se diseñará para nuestros propósitos (ver más detalles en 4). El conjunto total de modelos generados con el orbifold se dividirá para el conjunto de entrenamiento y prueba para nuestro autoencoder.

3.3. Autoencoder en orbifolios \mathbb{Z}_6 -II

El objetivo principal del presente trabajo es reproducir los resultados presentados en [36]. En dicha investigación se proporciona una representación bidimensional de múltiples modelos de orbifolios heteróticos con grupo de punto \mathbb{Z}_6 -II, dicha representación se obtiene de parametrizar a los modelos por medio del codificador de un autoencoder, cuya definición ya discutimos en 3.1.9. Dichos parámetros se obtienen de la capa latente o capa cuello de botella. Lo que haremos en la sección presente es discutir la construcción del autoencoder para modelos de orbifolios heteróticos con grupo de punto \mathbb{Z}_6 -II. Los modelos obtenidos por medios del autoencoder, cuyas características se explican en 3.2 deben ser preprocesados antes de darlos como entrada para la RN.

3.3.1. Preprocesamiento de datos

En 3.2 vimos que la generación de modelos nos proporciona los valores de los vectores de traslación y las líneas de Wilson correspondientes a un modelo de orbifolio heterótico con algún grupo de punto determinado. Para nuestro caso, los modelos con grupo de punto \mathbb{Z}_6 -II están parametrizados por 4 vectores 16-dimensionales (ver 2.5). Uno de ellos es el llamado vector de traslación V y el resto son las tres líneas de Wilson que denotamos por W_3 , W_2 y W'_2 . Así, tenemos un total de $4 \times 16 = 64$ parámetros de compactificación que determinan por completo a los modelos en orbifolios \mathbb{Z}_6 -II.

Como mencionamos en la sección anterior, estos 64 parámetros no serán los que reciba directamente el autoencoder como entrada, pues se pueden presentar grandes redundancias, y podemos tener modelos equivalentes físicamente pero con distintos parámetros. Para ello mapearemos estos 64 parámetros a vectores características 26-dimensionales. Recordemos que la geometría de los orbifolios \mathbb{Z}_6 -II tienen 12 puntos fijos, y cada punto fijo viene con una combinación única de vectores de traslación y líneas de Wilson. En consecuencia, en cada punto fijo, un conjunto diferente de raíces simples de $E_8 \times E_8$ transformarán covariantemente bajo las traslaciones y líneas de Wilson. Para nuestro grupo de punto, utilizaremos el vector de traslación V y las líneas de Wilson W_3 , W_2 y W'_2 para calcular los vectores de traslación locales V_g y así calcular el número de raíces simples que no se anulan de E_8 . Este número será invariante.

Consideraremos los 12 puntos fijos para el primer sector torcido θ . Cada punto fijo corresponde a un elemento constructor de la forma

$$g_a = (\theta, n_i^{(a)} e_i), \quad a = 1, 2, \dots, 12, \quad (3.39)$$

donde se tiene una suma implícita sobre $i = 1, 2, \dots, 6$ y para ciertos valores de $n_i^{(a)}$ cuyos valores se proporcionaron en 2.5. Para cada elemento constructor g_a definimos el correspondiente vector de traslación local como

$$V_{g_a} = V + (n_3^{(a)} + n_4^{(a)})W_3 + n_5^{(a)}W_2 + n_6^{(a)}W'_2. \quad (3.40)$$

Este vector 16-dimensional está dividido en dos vectores 8-dimensionales $V_{g_a} = (V_{g_a}^{(1)}, V_{g_a}^{(2)})$ correspondiendo cada uno a un factor E_8 del grupo de norma. Así, en cada punto fijo asociado a g_a calculamos el grupo de norma $G_a^{(\alpha)}$, ($\alpha = 1, 2$), llamado GUT local como sigue. Un vector raíz p de E_8 contribuye a la GUT local $G_a^{(\alpha)}$ si

$$V_{g_a}^{(\alpha)} \cdot p = 0 \pmod{1}. \quad (3.41)$$

Por cada una de estas 24 GUTs locales, contamos el número de raíces simples p no nulas (por ejemplo, 6 para $SU(3)$ y 240 para E_8) y almacenamos estos 24 números en un vector X de enteros, un entero por cada factor E_8 de los 12 puntos fijos. Es decir,

$$V_{g_a}^{(\alpha)} \cdot p = 0 \pmod{1} \Rightarrow \#p = N_{g_a}^\alpha \quad \alpha = 1, 2 \quad (3.42)$$

Agrupamos los 12 pares de $N_{g_a}^\alpha$

Adicionalmente a esto, los grupos de norma 4-dimensionales para los orbifolios \mathbb{Z}_6 -II $G_{4D} = G_{4D}^{(1)} \times G_{4D}^{(2)}$ está dado por la intersección de los 12 GUTs locales. Entonces, aún debemos agregar el número de raíces simples que sobreviven del grupo de norma 4D, es decir, dos enteros, uno para cada factor E_8 . Es decir, agregamos el número de raíces simples $N^{(\alpha)}$ tales que

$$p \cdot V^{(\alpha)} = 0 \pmod{1}, \quad p \cdot W_3^{(\alpha)} = 0 \pmod{1}, \quad p \cdot W_2^{(\alpha)} = 0 \pmod{1}, \quad p \cdot W_2'^{(\alpha)} = 0 \pmod{1}. \quad (3.43)$$

Así, tendremos entonces nuestro vector característica 26-dimensional de enteros invariante bajo la adición de vectores raíz de $E_8 \times E_8$ y reflexiones de Weyl como

$$X := (N^{(1)}, N^{(2)}, N_1^{(1)}, N_1^{(2)}, \dots, N_{12}^{(1)}, N_{12}^{(2)}). \quad (3.44)$$

Este es un mapeo del espacio de parámetros (vectores de traslación y líneas de Wilson) a un espacio llamado *espacio de características* del cual los vectores X , que acabamos de construir, forman parte. Las componentes de estos vectores son enteros llamados *características*. Dos vectores de características $X_{(1)}$ y $X_{(2)}$ del conjunto de datos no pueden representar el mismo modelo físico de compactificación \mathbb{Z}_6 -II al menos que $X_{(1)} = X_{(2)}$, un hecho que no sería cierto si solamente comparamos líneas de Wilson y vectores de traslación pues, en ese caso, hay varias combinaciones de estos parámetros diferentes que arrojan modelos de compactificación equivalentes, como explicamos en la sección anterior.

Ahora bien, en [35] se detecta que el entrenamiento de un autoencoder con estos parámetros mejora significativamente si, en lugar de dar de entrada este vector 26-dimensional de características, se realiza una codificación one-hot de dichos vectores. La codificación one-hot en nuestro caso es como sigue. Las características (entradas) de cada vector puede tener 37 valores diferentes, es decir, 37 maneras en las que el grupo E_8 puede ser roto en subgrupos en este orbifolio. Esta conclusión se obtiene después de contar los diferentes valores de los vectores de característica para una gran cantidad de modelos disponibles con compactificación en orbifolios \mathbb{Z}_6 -II. En la codificación

one-hot, cada característica diferente es reemplazada por un vector 37-dimensional que contiene sólo ceros y un sólo 1 en la posición correspondiente a un valor particular que tiene la característica correspondiente. Haciendo esto para las 26 características uno obtiene un vector mucho más grande de la forma $\{0, 1\}^{26 \times 37} = [0, 1]^{962}$. Así que este vector, que es la codificación one-hot del vector de características 26-dimensional es la entrada de nuestro autoencoder. La forma de nuestro autoencoder está resumida en la Figura 3.10.

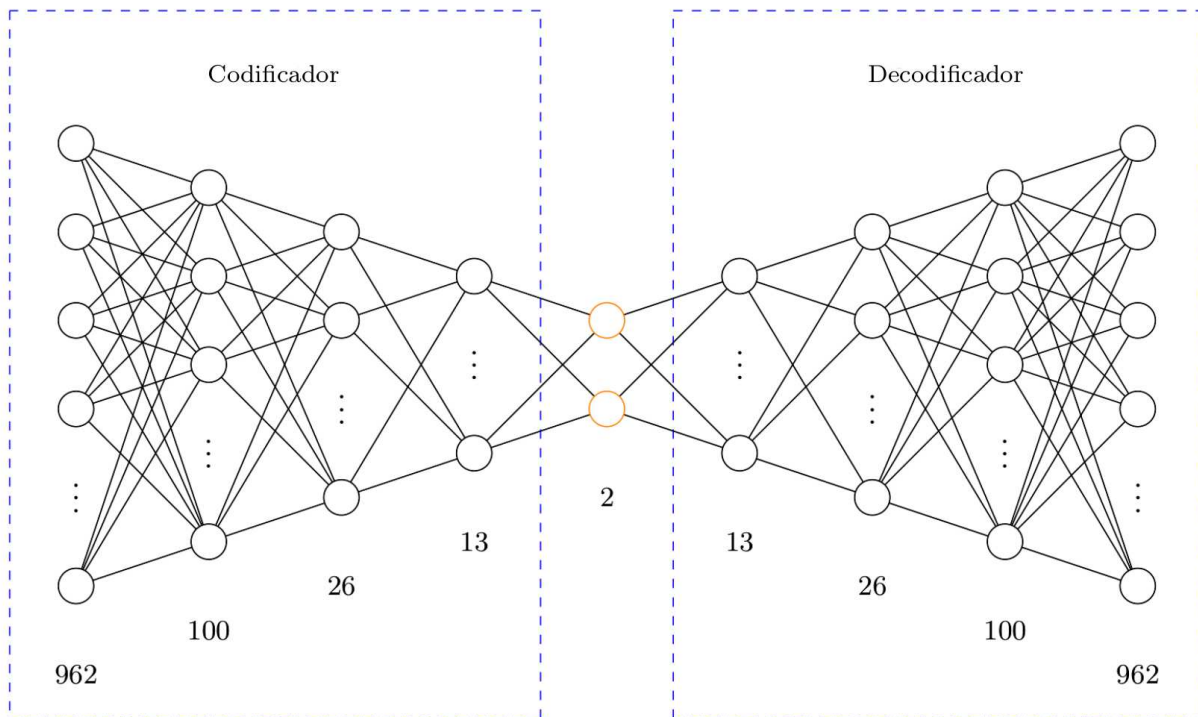


Figura 3.10: Arquitectura del autoencoder que utilizaremos para nuestros modelos de orbifolios heteróticos.

La arquitectura de nuestro autoencoder será la misma, en esencia, que la utilizada en [35] para reproducir los resultados obtenidos allí. La región del codificador nos permitirá reparametrizar a nuestros modelos a un espacio bidimensional, que son justamente los valores calculados en la capa cuello de botella (o capa latente). En el siguiente capítulo proporcionaremos en detalle los hiperparámetros de nuestro autoencoder, así como los resultados que obtuvimos en la capa latente.

Capítulo 4

Resultados

Ahora presentamos los resultados obtenidos con el autoencoder propio para codificar y decodificar nuestros modelos de orbifolios heteróticos \mathbb{Z}_6 -II. Primero daremos las características específicas de que tiene nuestro autoencoder, luego presentaremos la gráfica de la función de pérdida para verificar que la RN no está sobreentrenada. Posteriormente, presentaremos cómo se observa la capa latente en una gráfica bidimensional y codificaremos también modelos de orbifolios heteróticos \mathbb{Z}_6 -II que sean compatibles con el Modelo Estándar Mínimo Supersimétrico (MSSM), también otros dos conjuntos de modelos que tengan un espectro de partículas con tres generaciones de los grupos de norma $SU(5)$ y el grupo de norma Pati-Salam ($SU(4) \times SU(2) \times SU(2)$) más exóticos vectoriales. Por último, tratamos de identificar regiones prometedoras en la representación bidimensional de la capa latente para los tres grupos de modelos con un ejercicio de clústering.

4.1. Definición del autoencoder

La arquitectura de nuestro autoencoder está representada gráficamente por la Figura 3.10, y en 3.3 explicamos a detalle cómo es que los modelos generados por el Orbifolder serán procesados por la RN. Generamos una base de datos de $\mathcal{O}(4,000,000)$ modelos con geometría de compactificación \mathbb{Z}_6 -II con el orbifolder. Para entrenar a nuestro autoencoder, tomamos un total de $\mathcal{O}(1,000,000)$ modelos, donde la proporción de entrenamiento y prueba fue de

$$\text{entrenamiento : validación} = 85 : 15. \tag{4.1}$$

El autoencoder fue implementado con la paquetería TensorFlow [37]. La función de pérdida que utilizamos fue la función MSE, cuyos detalles explicamos en 3.1.7, también se le conoce como función de pérdida L_2 en las paqueterías que utilizamos en la implementación del autoencoder. Como mencionamos en 3.1.7, la función MSE pertenece al conjunto de funciones de pérdida utilizadas

para problemas de regresión (como en nuestro caso) y tiene la ventaja de sólo poseer un mínimo global que favorece al gradiente descendente durante el entrenamiento. El entrenamiento de la RN se hizo hasta que la función de pérdida convergiera a un mínimo global para el conjunto de validación, así evitando sobreentrenar a nuestra red. La función de activación utilizada para todas las capas ocultas fue la función SELU. Esta función de activación es importante, pues como se mencionó en 3.1.7 permite que la RN sea autonormalizada y el proceso de entrenamiento sea más rápido. El entrenamiento de la RN se realizó a través de 5,000 épocas y un tamaño de lote de 2,000. Después de este número de épocas obtuvimos una pérdida de 0.15 para el conjunto de entrenamiento, que equivale a una predicción correcta en promedio de 14 características. Este promedio se calcula en la última época del entrenamiento, promediando el número de características acertadas en cada modelo a lo largo del total de modelos analizados, que en nuestro caso corresponde a 850,000 modelos del conjunto de entrenamiento. Recordemos de 3.3.1 que cada característica de nuestros modelos se convierte en un vector de puros ceros y un uno con 37 entradas una vez que se ejecuta la codificación one-hot. Para verificar si acertamos la correspondiente característica comparamos la posición del 1 en el vector 37 dimensional correspondiente a dicha característica en el modelo de entrada con la posición del valor más grande para los vectores de salida predichos por nuestro autoencoder. En la Figura 4.1 se presenta la gráfica de la función de pérdida para el conjunto de entrenamiento y de validación durante el entrenamiento. Podemos apreciar que ambas pérdidas, tanto para el conjunto de entrenamiento y de prueba son decrecientes a lo largo de todas las épocas y, en particular, la pérdida para el conjunto de validación no tiene un punto de inflexión donde de repente empiece a crecer, indicando que no tenemos sobreentrenamiento tal y como explicamos en 3.1.8. Como no probamos con varias arquitecturas de RN, no podemos asegurar que la arquitectura propuesta en la Figura 3.10 es la óptima y no tenemos sobreajuste. Sin embargo, también mencionamos en 3.1.8 que asegurándonos de no tener sobreentrenamiento, es plausible también evitar un sobreajuste de nuestra red, pero insistimos en el hecho de que no se puede asegurar que la arquitectura sea la óptima para nuestro problema.

4.2. Capa latente o cuello de botella

Una vez completamente definido nuestro autoencoder, podemos ver la codificación de los modelos en el conjunto de entrenamiento y de validación en la capa latente. En la Figura 4.2 podemos apreciar la representación bidimensional de la codificación de los conjuntos de entrenamiento y de validación. Se aprecia en ambas gráficas que se tiene la misma configuración de regiones. La única diferencia que se puede observar es que las regiones son más densas en modelos para el caso del conjunto de entrenamiento con respecto al grupo de prueba, ésto debido a que evidentemente el conjunto de entrenamiento tiene más cardinalidad. Para observar que se tiene consistencia en ese aspecto, también graficamos la codificación de un conjunto ajeno al de entrenamiento y de validación de 2,000,000 de modelos en la Figura 4.3. Observamos la misma distribución que con el caso de los conjuntos de entrenamiento y validación, teniendo sólo como diferencia una mayor

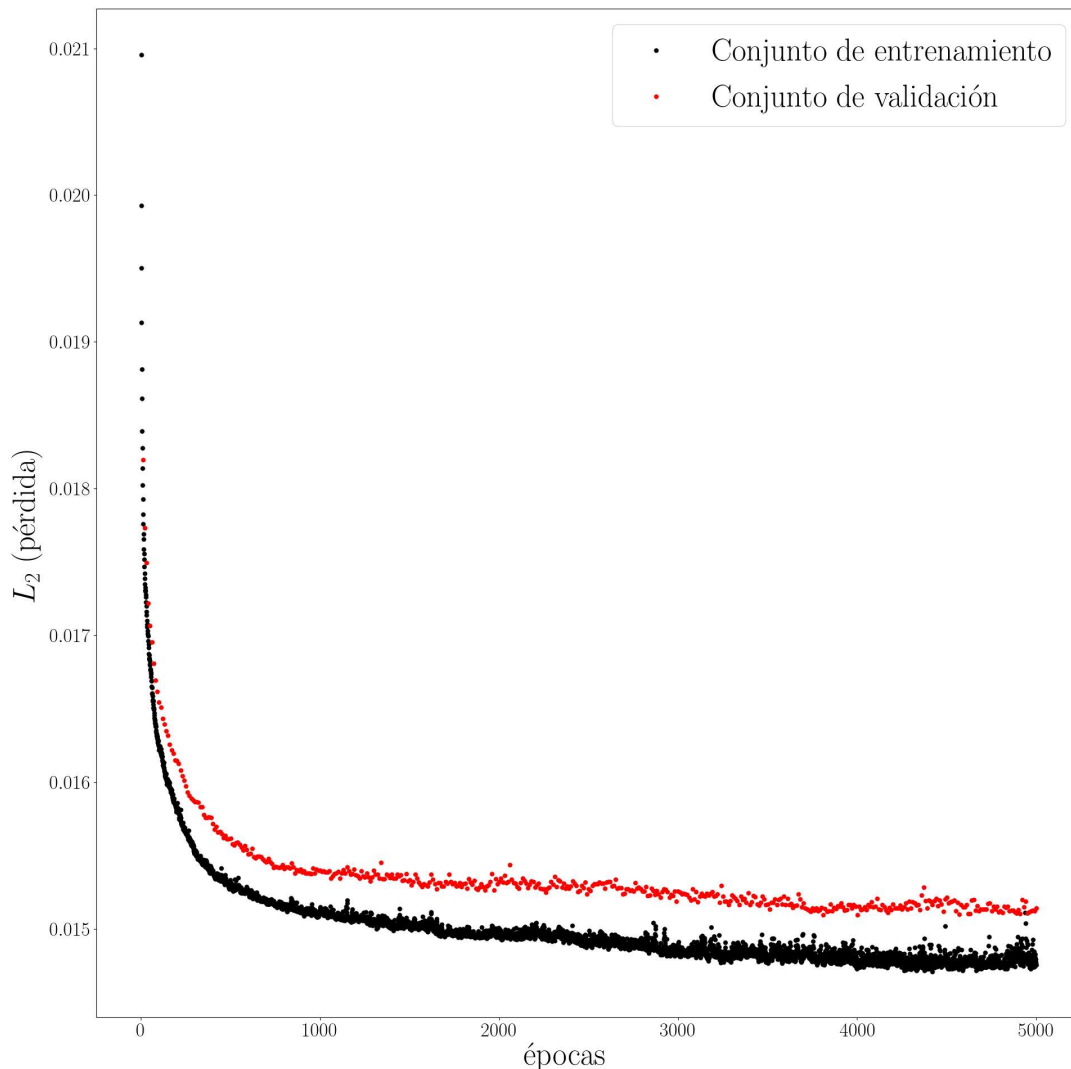


Figura 4.1: Gráfica del valor de la función de pérdida MSE para el conjunto de entrenamiento (puntos negros) y el conjunto de validación (puntos rojos).

densidad de elementos en todas las regiones. Con estas figuras podemos afirmar entonces que sólo en estas regiones se encuentran modelos consistentes con geometría de compactificación \mathbb{Z}_6 -II (1,1). Esta conclusión nos permite empezar a hacernos preguntas respecto a la localización de modelos con ciertas características, tales como su grupo de norma. Acabamos de concluir que el entrenamiento descrito nos permitió localizar la región bidimensional en la cual se encuentran todos los modelos

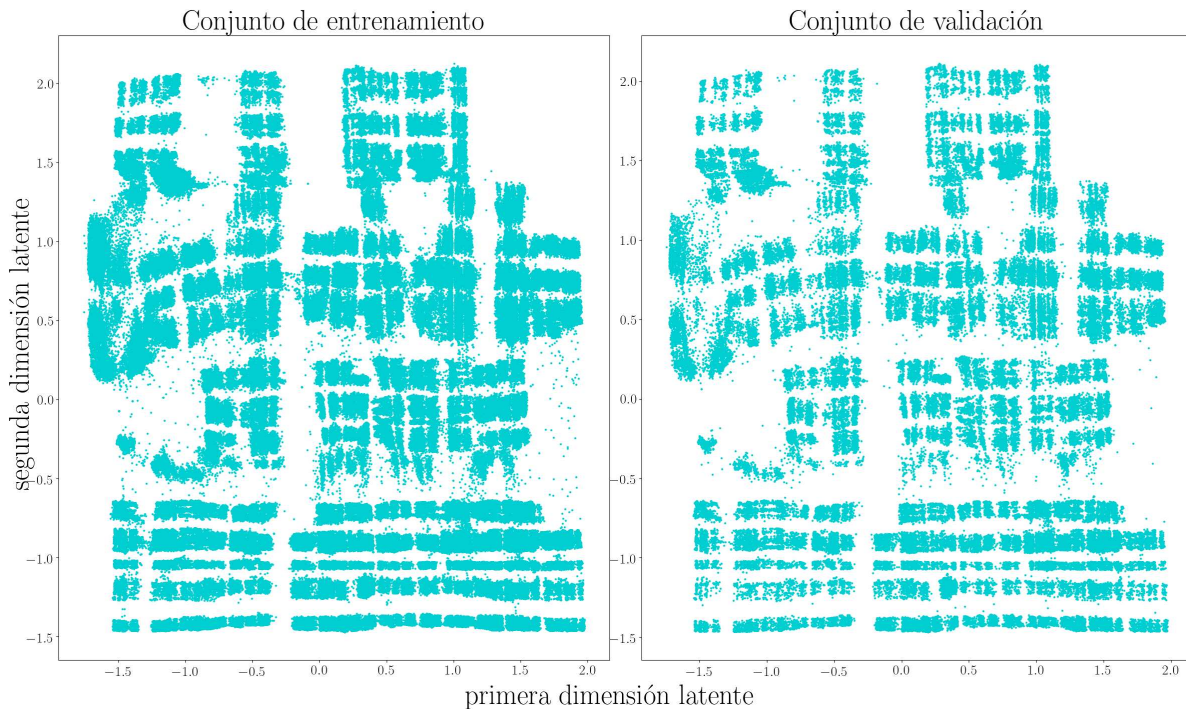


Figura 4.2: Representación bidimensional de la codificación del conjunto de entrenamiento y de validación bajo nuestro autoencoder.

consistentes de compactificación en orbifolios \mathbb{Z}_6 -II. Sin embargo, deseamos asegurar estabilidad de nuestro entrenamiento, es decir, estar seguros de que con un entrenamiento distinto que obtenga una pérdida mínima similar, las regiones de modelos consistentes sean similares. Para analizar consistencia en el entrenamiento, repetimos dicho entrenamiento con número de épocas y tamaño de batch diferentes para diversos conjuntos de entrenamiento y prueba. En total se hicieron otras 3 repeticiones donde no se tuviera sobreentrenamiento con las siguientes condiciones:

- Épocas: 7,000 Tamaño de batch: 1,000
- Épocas: 5,000 Tamaño de batch: 1,000
- Épocas: 2,000 Tamaño de batch: 3,000

En todos los casos se utilizó la misma función de activación y arquitectura de capas para la RN y se obtuvo la misma configuración de regiones de modelos consistentes en la capa latente, comprobando así que tenemos estabilidad ante repeticiones. Ya estando seguros de que nuestro entrenamiento

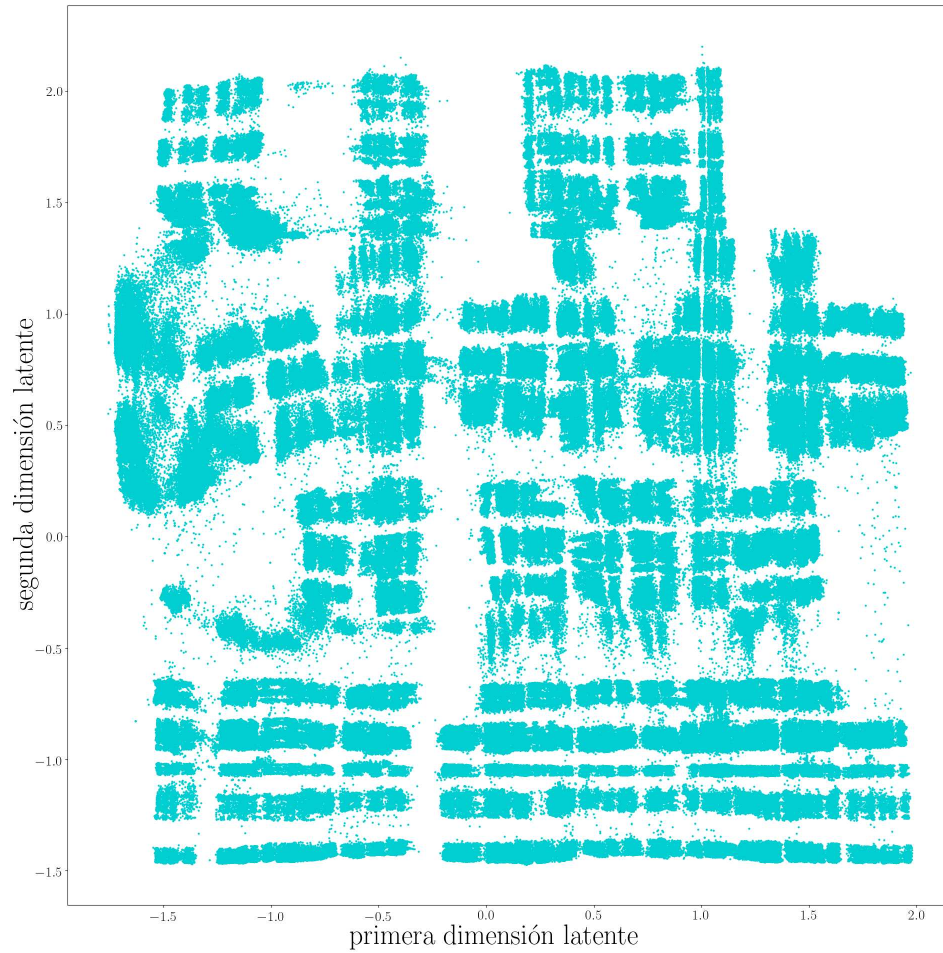


Figura 4.3: Representación bidimensional de la codificación de un conjunto de 2,000,000 de modelos ajenos al de prueba y entrenamiento. Podemos apreciar que se tiene consistencia en la forma de las regiones obtenidas.

cuenta con estabilidad ante repeticiones y habiendo identificado las regiones de modelos consistentes con geometría de compactificación \mathbb{Z}_6 -II, podemos comenzar a analizar las regiones tratando de determinar cuáles de ellas podrían arrojar información fenomenológica relevante.

4.3. Modelos tipo MSSM, $SU(5)$ y Pati-Salam en la capa latente

Ya que tenemos una parametrización en dos parámetros de modelos consistentes de orbifolios abelianos \mathbb{Z}_6 -II (1,1), deseamos poder explorar toda la región de los valores de los parámetros (espacio latente) y poder identificar algunas zonas específicas que contengan modelos fenomenológicamente interesantes. Para ello, lo que haremos es dividir todo el espacio de parámetros identificado en regiones claramente definidas. En la Figura 4.3 podemos apreciar que sí se pueden distinguir regiones ajenas entre sí y bien delimitadas, las cuales podemos identificar y numerar.

Un primer intento para poder identificar regiones de interés es contar con modelos que de facto sus características sean conocidas, tales como ser compatible con el espectro del modelo estándar mínimo supersimétrico (MSSM, por sus siglas en inglés), y codificarlos, así determinando la localización de los mismos dentro de la región de parámetros donde identificamos a todos los modelos de orbifolios heteróticos \mathbb{Z}_6 -II. La idea es tener varios de estos modelos y así identificar regiones que, en general, contengan la mayor cantidad de modelos de interés. Utilizaremos 3 conjuntos de modelos fenomenológicamente interesantes.

En [8] se hizo una clasificación de todas las simetrías de sabor 4D posibles del espectro no masivo proveniente de compactificaciones en orbifolios abelianos, incluyendo roto-traslaciones y toros no factorizables. En dicho trabajo, se concentraron en la teoría heterótica de cuerdas $E_8 \times E_8$ e identificaron un total de 121,000 modelos compatibles con el Modelo Estándar Mínimo Supersimétrico. Accediendo al material suplementario de dicho trabajo, tuvimos acceso los 121,000 modelos tipo MSSM, y de ese conjunto se contaron con 363 de ellos con geometría de compactificación abeliana \mathbb{Z}_6 -II (1,1). De la sección 2.3.3 recordemos que los otros dos números (1,1) corresponden a la clasificación de orbifolios simétricos toroidales. Quiere decir que estamos analizando los orbifolios con clase- \mathbb{Q} \mathbb{Z}_6 -II, a la clase- \mathbb{Z} correspondiente a la retícula del grupo de espacio $G_2 \times SU(3) \times SO(4)$ y a la primera clase afín (la única en nuestro caso) de acuerdo a la clasificación en [19]. Dichos modelos están en el mismo formato que el orbifolder provee al generarlos.

Vale la pena enfatizar a qué nos referimos con que un modelo de compactificación sea compatible con el MSSM. Un modelo específico es fenomenológicamente viable si su espectro no masivo 4D satisface las siguientes condiciones [8]:

- El grupo de norma no roto es $\mathcal{G}_{SM} \times \mathcal{G}_{oculto} = SU(3)_c \times SU(2)_L \times U(1)_Y \times \mathcal{G}_{oculto}$, donde el grupo \mathcal{G}_{oculto} contiene factores de norma continuos adicionales (abelianos y no abelianos), y el factor $U(1)_Y$ no tiene anomalías y es compatible con gran unificación;
- los estados efectivos (torcidos y no torcidos) incluyen campos que reproducen el espectro de materia del MSSM, y
- estados efectivos adicionales son vectoriales con respecto al factor \mathcal{G}_{SM} e incluye singuletes del Modelo Estándar que puedan jugar el rol de neutrinos derechos.

Además de los modelos tipo MSSM también contamos con otros dos conjuntos de modelos de interés. Utilizando el orbifolder, generamos modelos de orbifolios heteróticos \mathbb{Z}_6 -II (1,1) con espectros inequivalentes de 3 generaciones 4D de los grupos de norma Pati-Salam (PS) o $SU(5)$ más exóticos vectoriales. En el caso de modelos PS la búsqueda aleatoria con el orbifolder arrojó un total de 16 modelos inequivalentes y un total de 457 modelos inequivalentes para el caso de $SU(5)$. Los modelos pertenecientes a estas tres categorías son preprocesados como describimos en 3.3 y, por ende, mapeados al espacio de características, convirtiéndolos en un vector de enteros con 26 entradas. Posteriormente, a estos vectores se les hace pasar por la codificación one-hot, encontrando también que los valores de las correspondientes entradas (características) están dentro de las 37 posibilidades descritas también en 3.3 convirtiéndolos en vectores de la forma $\{0, 1\}^{26 \times 37} = [0, 1]^{962}$. De hecho, es un resultado inherente a esta geometría de compactificación que se encuentren 37 posibilidades para los valores de las características, resultado encontrado en [36], y que equivale al número de formas que se puede romper al grupo E_8 con orbifolios \mathbb{Z}_6 -II. Por último, los vectores 962 dimensionales son codificados con el autoencoder. La codificación de los modelos dentro de las tres categorías descritas se sobrepuso a la codificación del conjunto de 2,000,000 de modelos \mathbb{Z}_6 -II anteriormente utilizados. Dichas codificaciones se pueden apreciar en la Figura 4.4.

Para identificar regiones independientes en el espacio de parámetros, tratamos de identificar aglomeraciones de puntos utilizando la codificación de 2,000,000 de modelos. Identificar aglomeraciones o *clústers* de puntos en un espacio bidimensional es un problema de ML, incluso el número de aglomeraciones es encontrado por medio de aprendizaje automático si se quiere una clasificación más robusta. En este trabajo, para identificar aglomeraciones de modelos, utilizamos la paquetería K-means de Python [38]. El algoritmo K-Means agrupa datos tratando de separarlos en K grupos con la misma varianza, minimizando una función de distancia. Este algoritmo requiere que el número de *clústers* sea especificado; funciona bien para una gran cantidad de muestras. El agrupamiento se realiza minimizando la suma de distancias entre cada objeto y el centroide de su grupo o clúster. Se suele usar la distancia cuadrática. El algoritmo consta de los pasos:

1. **Inicialización:** Establecemos el número de grupos, K , se establecen K centroides en el espacio de datos, se pueden escoger aleatoria.
2. **Asignación de objetos a los centroides:** cada objeto de los datos es asignado a su centroide más cercano.
3. **Actualización de los centroides:** se actualiza la posición del centroide de cada grupo tomando como nuevo centroide la posición del promedio de objetos pertenecientes a dicho grupo.

Los pasos 2 y 3 se repiten hasta que los centroides no se mueven, o se mueven por debajo de una distancia umbral en cada paso. El algoritmo K-Means resuelve un problema de optimización, siendo la función a optimizar (minimizar) la suma de las distancias cuadráticas de cada objeto al centroide del clúster. Los objetos se representan como vectores d -dimensionales $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ y

4.3. MODELOS TIPO MSSM, $SU(5)$ Y PATI-SALAM EN LA CAPA LATENTE¹⁰⁹

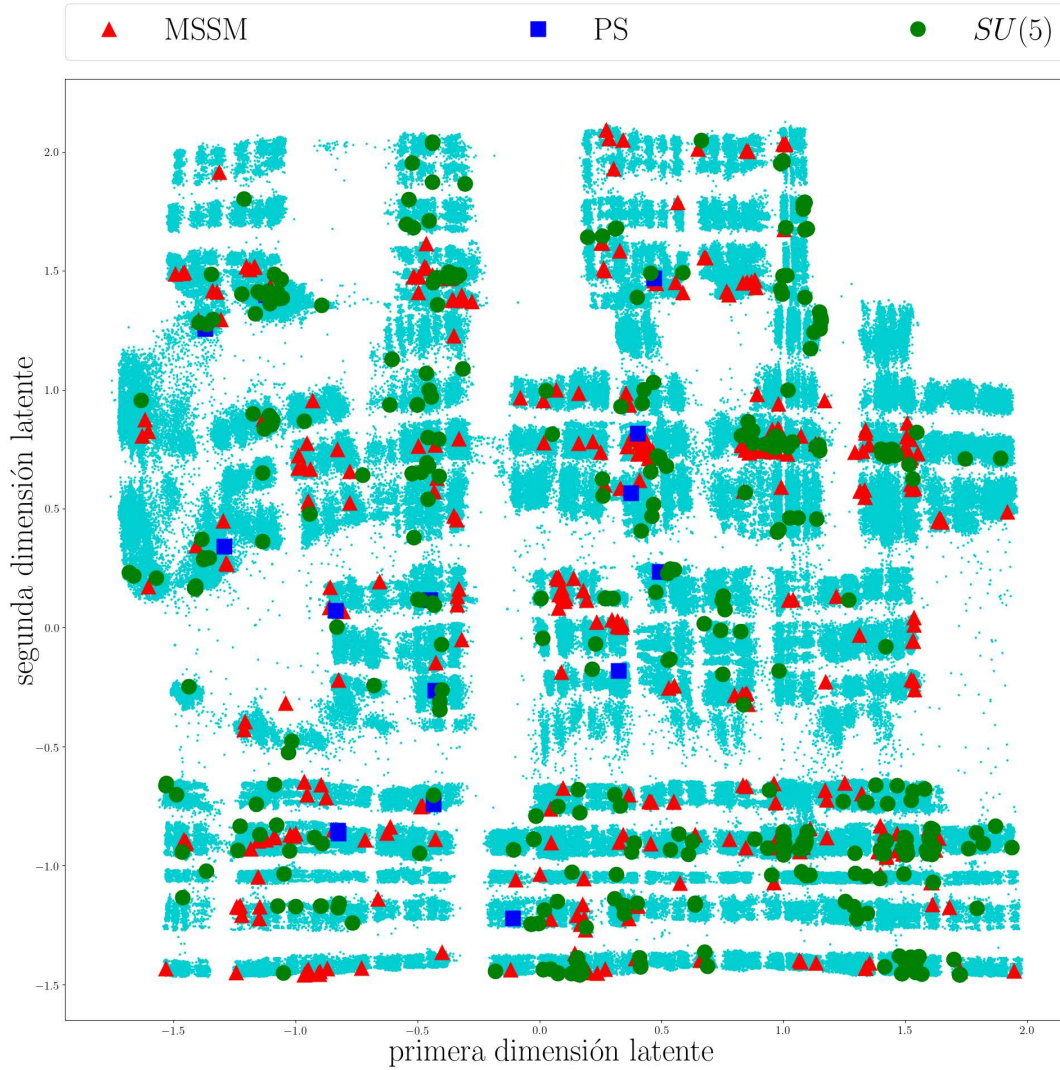


Figura 4.4: Configuración paramétrica de 2,000,000 de modelos ajenos al conjunto de entrenamiento y de prueba de nuestro autoencoder (puntos en color turquesa). También tenemos la parametrización de los modelos tipo MSSM, PS y $SU(5)$.

el algoritmo construye K grupos donde se minimiza la suma de distancias de los objetos de cada

grupo $\mathbf{S} = \{S_1, S_2, \dots, S_K\}$, a su centroide. Matemáticamente tenemos [39]

$$\min_{\mathbf{S}} E(\mu_i) = \min_{\mathbf{S}} \sum_{i=1}^K \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \mu_i\|^2. \quad (4.2)$$

Tendremos K grupos o clústers con su correspondiente centroide μ_i . En cada actualización de los centroides, desde el punto de vista matemático, imponemos la condición necesaria de extremo a la función $E(\mu_i)$ que, para la función cuadrática (4.2) es

$$\frac{\partial E}{\partial \mu_i} = 0 \Rightarrow \mu_i^{t+1} = \frac{1}{|S_i^{(t)}|} \sum_{\mathbf{x}_j \in S_i^{(t)}} \mathbf{x}_j, \quad (4.3)$$

y se toma el promedio de los elementos de cada grupo como nuevo centroide. Algunas desventajas de utilizar este algoritmo son las siguientes. La medida de distancia asume que los clústers son convexos e isotrópicos, lo cual no siempre es así. Responde de manera pobre ante clústers alargados o variedades con formas irregulares. La distancia cuadrática no es una métrica normalizada, sólo sabemos que los valores bajos son mejores y que los valores nulos son los óptimos. Pero en espacios multidimensionales, las distancias Euclidianas tienden a inflarse. El número de clústers tiene que darse inicialmente, pues este algoritmo no determina la cantidad de ellos, sino que dada esta cantidad divide al conjunto en ese número buscando los mejores centros posibles.

Así, dada la última desventaja, realizamos algunas pruebas sobre diferentes números de clústers que nuestra capa latente puede llegar a tener. Se corrió este algoritmo teniendo como datos de entrada a 2,000,000 de modelos codificados. Se probó con valores en número de clústers desde 25 a 35, concluyendo con meros criterios cualitativos que la división más adecuada fue con 30 clústers. En la Figura 4.5 apreciamos la división del espacio de parámetros en los 30 clústers, además también apreciamos la localización de modelos tipo MSSM, PS y $SU(5)$.

Una vez identificados los 30 clústers, contamos la cantidad de modelos de cada una de las tres categorías (MSSM, PS y $SU(5)$) en cada una de las regiones o clústers. Dicho conteo se aprecia en la Tabla 5.1. Todos estos datos nos permiten hacer un análisis de la distribución de modelos en el espacio latente.

4.4. Análisis del espacio latente

La Tabla 5.1 y la Figura 4.5 evidencian que en ciertas regiones hay una acumulación importante de modelos por categoría. Es importante notar que debemos estudiar la aglomeración de modelos por separado de cada categoría, pues cada conjunto de modelos es independiente entre sí. Por ejemplo, no necesariamente podemos reducir los modelos $SU(5)$ en su versión supersimétrica al MSSM si no contamos con los Higgs necesarios, por ende el MSSM no se puede considerar un subconjunto

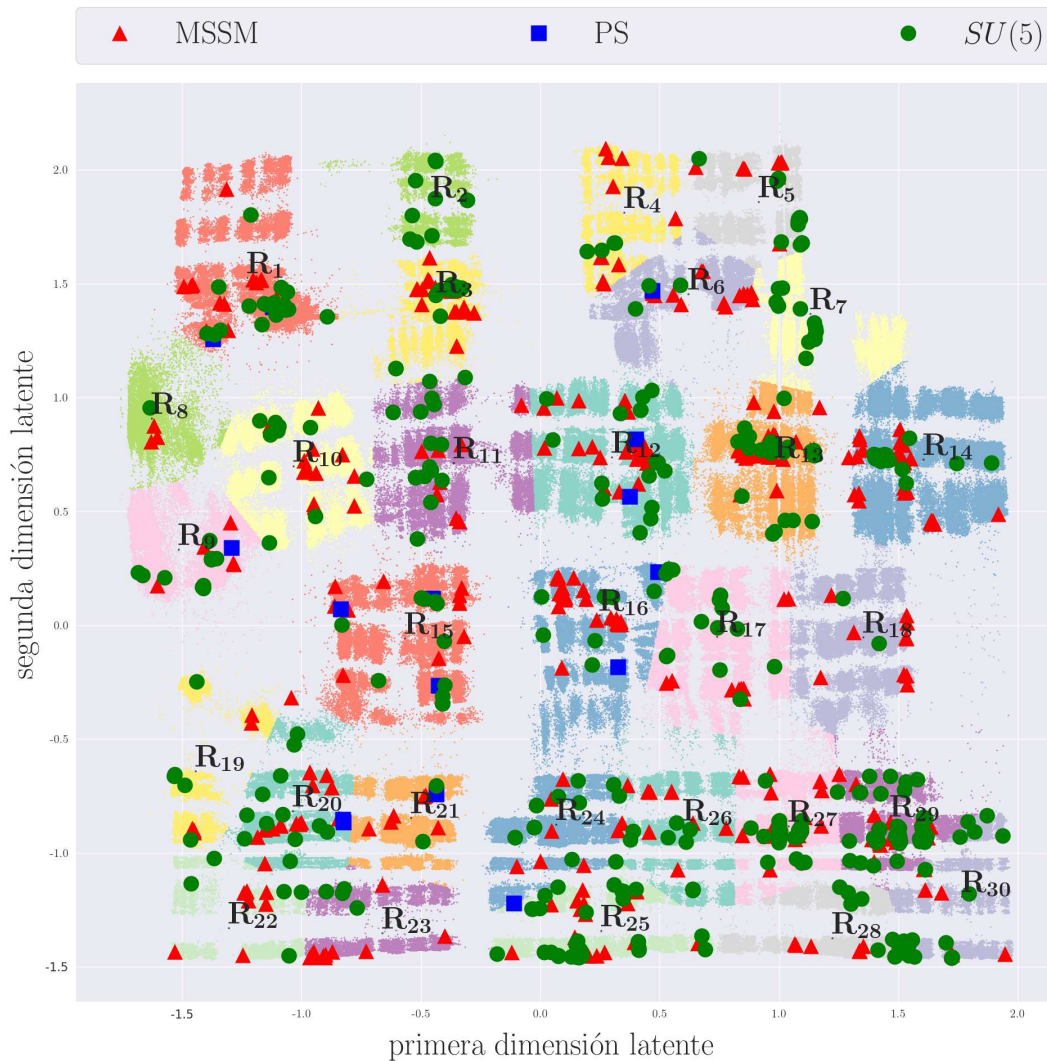


Figura 4.5: Espacio latente de parámetros dividido en 30 clústers. Notemos que la numeración de los mismos es en orden y en regiones colindantes. También apreciamos la localización de los modelos tipo MSSM, PS y $SU(5)$ en sus respectivos clústers.

de los modelos tipo $SU(5)$. Así que por cada conjunto (MSSM, $SU(5)$ y PS) buscamos los clústers con mayor acumulación de los mismos de manera independiente. A primera vista, pareciera que las regiones inferiores y las regiones 11, 12, 14 y 14 tienen mayor acumulación de modelos tipo MSSM y $SU(5)$. Ya con números duros directo de la tabla, apreciamos que para el caso de los modelos

	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}	R_{12}	R_{13}	R_{14}	R_{15}
MSSM	14	0	19	11	9	12	1	3	5	12	8	28	30	26	10
$SU(5)$	18	10	22	4	17	3	12	1	10	10	17	17	32	35	11
PS	2	0	0	0	0	1	0	0	1	0	0	2	0	0	3
	R_{16}	R_{17}	R_{18}	R_{19}	R_{20}	R_{21}	R_{22}	R_{23}	R_{24}	R_{25}	R_{26}	R_{27}	R_{28}	R_{29}	R_{30}
MSSM	22	9	10	6	17	6	8	11	7	20	11	15	6	19	7
$SU(5)$	14	12	2	5	12	2	5	5	16	44	10	31	12	51	16
PS	2	0	0	0	2	1	0	0	0	1	0	0	0	0	0

Tabla 4.1: Conteo de modelos por categoría en cada una de las 30 regiones en el espacio latente.

MSSM, se tiene una mayor concentración de ellos en las regiones 12, 13 y 14. Para los modelos tipo $SU(5)$ también se tiene gran concentración en las regiones 12, 13 y 14 pero aún mayor en las regiones 25, 27 y 29. Por último, los modelos tipo PS, que son escasos, se agrupan más en las regiones 1, 12, 15, 16 y 20.

Tomemos en cuenta que la separación por regiones de nuestro espacio espacio latente es artificial, sobre todo porque el número de clústers independientes lo establecimos nosotros a mano y de manera completamente cualitativa. Por otro lado, también apreciamos en la Figura 4.5 que no tenemos una clara jerarquía en cuanto a densidad de modelos por categoría se refiere. Aprovechando que la numeración de las regiones hace corresponder a regiones colindantes con indexación de números consecutivos, consideramos regiones más amplias constituidas por cúmulos de regiones dentro de las 30 ya identificadas. Tomamos entonces cúmulos de 5 en 5 regiones, teniendo en total 6 cúmulos, y en cada uno de ellos ahora contamos el número de modelos por categoría localizados allí. Se presenta una gráfica de barras en la Figura 4.6 con el conteo de modelos por categoría en las regiones agrupadas de 5 en 5.

Podemos observar de la figura que en las tres categorías el cúmulo de regiones 11-15 contiene el mayor número de modelos, permitiendo así concluir que si en una extensión de este trabajo se decide hacer una búsqueda aleatoria de modelos en cada categoría de manera independiente, entonces dicho cúmulo tiene la mayor probabilidad de dar una búsqueda exitosa. Para el caso de $SU(5)$ se observa que el cúmulo de mayor densidad de modelos de este tipo corresponde al cúmulo 26-30. Así, para el caso de modelos de tipo $SU(5)$, conviene hacer búsquedas en los cúmulos 11-15 y 26-30.

En la Figura 4.5 apreciamos que dentro de cada una de las regiones que localizamos, también hay vacíos de puntos en color gris, indicando que allí no podemos localizar nuestros modelos. Los valores de los parámetros en esos vacíos grises deberían ser desechados en el caso de que se realice una búsqueda aleatoria en el espacio latente delimitado a una región determinada, pues estamos seguros de que allí no podemos encontrar por el motivo siguiente. En [7] se hace una búsqueda de modelos compatibles con el MSSM y cuya geometría de compactificación es justamente la que

aquí estamos trabajando, encontrando un total de 300 modelos con estas características. En dicho trabajo, también determinan que el estimado de total de modelos inequivalentes con este orbifold es del orden de 10^7 . Considerando que nosotros estamos graficando la parametrización de 2,000,000 de modelos con este orbifold, es decir, el 20% del total de estos modelos, podemos concluir que la distribución de puntos que apreciamos en la Figura 4.5 ya proporciona la forma de la distribución puntos para el total de modelos con este orbifold. Es preciso enfatizar que el conjunto de modelos considerado aquí para visualizar el espacio latente es un muestreo aleatorio del conjunto total de modelos con la geometría de compactificación \mathbb{Z}_6 -II, pues fueron generados para este trabajo con el orbifold, el cual hace una búsqueda aleatoria de modelos con las características del orbifold que uno le indique.

Es importante recalcar que este trabajo de investigación es sólo un primer paso, pues a este nivel sólo se obtuvo la parametrización de modelos disponibles con geometría de compactificación \mathbb{Z}_6 -II (1,1). Posteriormente, agrupamos tres categorías de modelos con fenomenología específica (compatibles con el MSSM y con espectro de 3 generaciones de los grupos de norma $SU(5)$ y Pati-Salam) y los localizamos en el espacio latente parametrizándolos con el autoencoder que entrenamos. Lo deseable es llevar más allá este proceso encontrando un mapeo del espacio latente a modelos de la cuerda heterótica 4D, pudiendo conocer sus características físicas tales como su grupo de norma. Esto quiere decir que se busca un método que permitiera tomar un punto arbitrario de la capa latente y poder contar el número de raíces simples invariantes del grupo de norma 4D por cada factor E_8 del modelo correspondiente. Esto permitiría clasificar orbifolds mediante la exploración del espacio de parámetros disponible de la capa latente, facilitando enormemente la búsqueda de construcciones prometedoras.

Tras un par de días de ejecución ininterrumpida, el orbifold sólo pudo generar 16 modelos con espectros inequivalentes con 3 generaciones del grupo de norma Pati-Salam más exóticos vectoriales. Dichos modelos son posibles con estas construcciones de compactificación de la cuerda heterótica, de hecho, en [24] se obtienen modelos con grupo de norma Pati-Salam 4D, el cual posteriormente debe ser roto al modelo estándar por medio del mecanismo de Higgs. La poca cantidad de modelos generados con esta configuración por la búsqueda aleatoria del orbifold nos indica que modelos tipo PS no se ven favorecidos con esta geometría de compactificación respecto a las otras dos categorías: $SU(5)$ y MSSM.

Este trabajo está motivado en los resultados obtenidos en [36], donde también se trabaja con modelos de compactificación con nuestra misma geometría y en afán de reproducir los resultados, utilizamos la misma estructura del autoencoder. Sin embargo, es claro que no logramos reproducir lo observado en la capa latente por ese trabajo. La distribución de regiones ocupadas por modelos \mathbb{Z}_6 -II no se asemejan. Existen múltiples razones por las que dicha configuración no se reproduce. En dicho trabajo se logró un valor de pérdida mínimo de la función MSE de 0.013 con una predicción correcta en promedio de 16 características, mientras que en este trabajo el autoencoder entrenado alcanzó una pérdida de 0.015 y una predicción correcta en promedio de 14 características. Esto equivale a una predicción aproximada del 54% del total de características en cada modelo y es

la mejor precisión alcanzada con la técnica actual aquí presentada. Sin embargo, es importante enfatizar que este no es todavía el resultado final deseado; se quisiera alcanzar una precisión de al menos el 80 % para tener mayor confianza en la eficacia del modelo de autoencoder.

Expliquemos las razones por las cuales la técnica actual no tuvo un mejor desempeño en la reconstrucción de modelos con nuestra RN. La primera razón está relacionada con la complejidad del modelo para nuestro autoencoder; sólo probamos una arquitectura fija sin variar el número de nodos y capas. En particular, es posible que la eficacia aumente si incrementamos el número de nodos de la capa latente y, en consecuencia, tendríamos que explorar un espacio latente de mayor dimensionalidad. Al no haber variado la arquitectura de la RN, no podemos asegurar si la que utilizamos aquí es la óptima. Las siguientes razones están relacionadas con el preprocesamiento de los modelos, donde éstos son mapeados al espacio de características 26 dimensional. Estos vectores de características X tiene algunas ambigüedades: un modelo de cuerdas 4D es invariante bajo 1) intercambio entre los dos grupos de norma E_8 , 2) bajo ciertas permutaciones de los puntos fijos (ver Apéndice de [36]). En este trabajo sí realizamos un filtrado de modelos asegurándonos que no hubiera redundancias respecto al intercambio de los E_8 , pero no alcanzamos a hacer un filtrado respecto a la segunda redundancia, por lo que nuestro autoencoder pudo haber recibido modelos equivalentes y la eficacia del modelo de RN obtenido se pudo ver afectado. Otra redundancia de los vectores de características 26 dimensional se encuentra en que el grupo de Weyl para cada factor E_8 es del orden de $\approx 7 \times 10^8$. A pesar de que los vectores de características X son invariantes ante las reflexiones de Weyl, se puede dar el caso en que dentro del conjunto de modelos que utilizamos para el entrenamiento, existan modelos equivalentes ante estas reflexiones, agregando una redundancia más al proceso de entrenamiento (pues la RN interpreta a dos posibles vectores equivalentes físicamente como diferentes). Deshacernos de esta redundancia comparando vector a vector ante estas transformaciones no parece ser muy deseable debido al orden de dicho grupo. Como trabajo a futuro también se plantea buscar una solución óptima a este problema implementando también técnicas de machine learning, en el ámbito de clasificación, enseñándole a la máquina a identificar vectores equivalentes bajo reflexiones de Weyl.

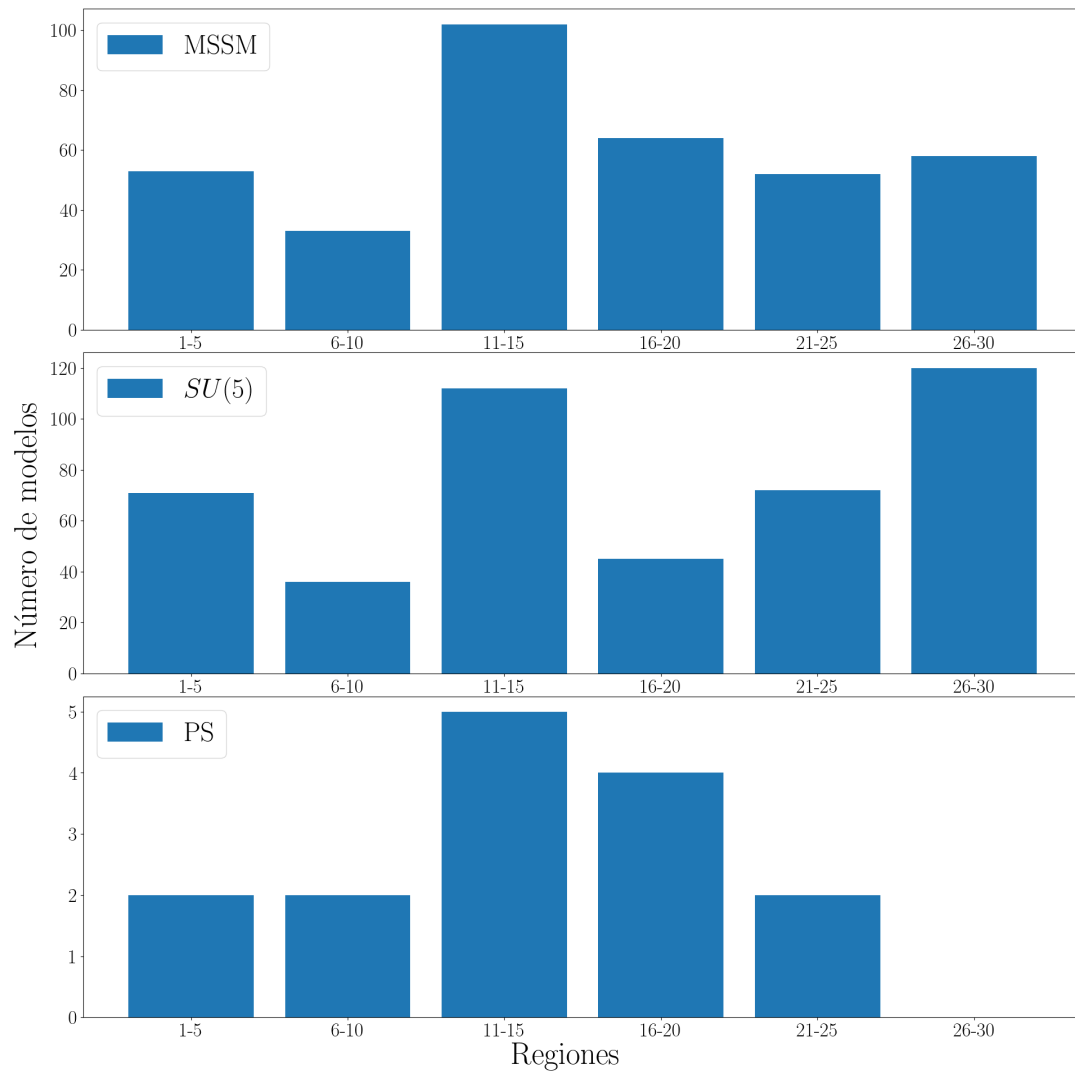


Figura 4.6: Gráfica de barras con el conteo de modelos por categoría en cúmulo de regiones de 5 en 5.

Capítulo 5

Conclusiones

El trabajo presentado en esta tesis surge motivado por la enorme cantidad de modelos de cuerdas con diversas propiedades comúnmente apodada "paisaje de cuerdas". Recientemente se han aplicado técnicas para el estudio de diversos aspectos de la teoría de cuerdas [2, 3, 4, 5]. Adicionalmente, podemos encontrar un compendio con las aplicaciones de dichas técnicas computacionales a la teoría de cuerdas en [6]. De nuestro particular interés fue [36], donde utilizando una herramienta de ML llamada *autoencoder*, se obtuvo una parametrización bidimensional de una gran cantidad de modelos de teoría de cuerda heterótica con geometría de compactificación \mathbb{Z}_6 -II. El objetivo de este trabajo fue reproducir esos resultados y tratar de identificar regularidades adicionales, localizando no sólo modelos compatibles con el Modelo Estándar Mínimo Supersimétrico (MSSM), si no también modelos con espectros de materia inequivalentes de 3 generaciones 4D de los grupos de norma $SU(5)$ o Pati-Salam.

Para entender el marco teórico de este trabajo, en el Capítulo 1 desarrollamos generalidades de la Teoría de Cuerdas empezando por la teoría bosónica, en la cual sólo se tiene un espectro de partículas bosónicas. La teoría de cuerdas bosónica se plantea como una generalización a la partícula puntual relativista donde, en lugar de considerar la línea de mundo de un punto en el espacio-tiempo, ahora analizamos la hoja de mundo que trazan objetos unidimensionales (llamados cuerdas) con cierta longitud. Estas cuerdas unidimensionales pueden ser cerradas (con condiciones de frontera periódicas) o abiertas (condiciones de frontera de Dirichlet o de Neumann) y en cada caso, la hoja de mundo tiene soluciones a las ecuaciones de movimiento en términos de la superposición de modos normales de oscilación. Dichos modos normales se promueven a operadores durante la cuantización de la teoría, los cuales generan estados bosónicos perturbando al vacío. Para que la teoría sea consistente, se requiere que la hoja de mundo esté inmersa en un espacio-tiempo $D = 26$ dimensional. En el caso de la cuerda bosónica cerrada tenemos dos tipos de modos desacoplados: izquierdos y derechos. La naturaleza requiere también de la existencia de fermiones, por lo que se debe modificar la teoría de cuerdas bosónica introduciendo supersimetría

a la respectiva acción. Con esta modificación, tenemos ahora grados de libertad fermiónicos cuyos modos normales cuantizados dan origen a estados de partículas fermiónicas. Esta teoría modificada es conocida como Teoría de Supercuerdas. Para que la teoría esté libre de anomalías se requiere que el espacio-tiempo de fondo sea $D = 10$ dimensional. Si dentro de esta teoría nos enfocamos en el caso de la cuerda cerrada, tenemos 10 modos bosónicos derechos y 10 modos bosónicos izquierdos desacoplados entre sí, de manera análoga al caso de la cuerda bosónica. También contamos con los pares supersimétricos correspondientes, es decir, 10 modos fermiónicos izquierdos y 10 modos fermiónicos derechos desacoplados. Para los grados de libertad fermiónicos de la cuerda cerrada tenemos condiciones de frontera periódicas (condiciones de Ramond, R) y antiperiódicas (condiciones de Neveu-Schwartz, NS), modificando así la expansión en modos normales. De la teoría de supercuerdas se derivan otros tipos de teorías como la tipo II-A y II-B. La libertad de escoger los modos derechos e izquierdos de distintos tipos de cuerdas se debe a que los estados de una teoría de cuerdas bosónica o de supercuerdas tipo II son el producto directo de estados del espacio de Fock para modos derechos e izquierdos. La teoría de cuerdas heterótica aprovecha esa independencia y combina los modos derechos fermiónicos (que viven en 10 dimensiones) con los modos bosónicos izquierdos de la teoría bosónica (que viven en 26 dimensiones). En dicha teoría se compactifican 16 modos bosónicos izquierdos (en un toro T^{16} , por ejemplo) para que haya una correspondencia entre los 10 modos bosónicos restantes y los 10 modos fermiónicos derechos. La idea de tener 16 grados de libertad internos es para dar lugar a simetrías de norma y por ende bosones de norma.

En el Capítulo 2 se describe el proceso de compactificación en estructuras toroidales. Aquí vemos cómo se define un toro de D dimensiones en general, siendo el cociente $T^D = \mathbb{R}^D/\Lambda$, donde Λ es una retícula generada por una base de vectores D -dimensionales e_α de \mathbb{R}^D y donde ahora los puntos en el espacio son identificados dentro de la retícula $X \sim X + n_\alpha e_\alpha$, $n_\alpha \in \mathbb{Z}$. En nuestro caso, compactificamos el espacio de los 16 modos bosónicos izquierdos extra en un toro $T^{16} = \mathbb{R}^{16}/\Lambda$. Para que la teoría sea consistente, sólo hay dos elecciones para dicha retícula: las retículas de raíces $E_8 \times E_8$ y $spin(32)/\mathbb{Z}_2$. En este trabajo nos concentramos en el caso que $\Lambda = E_8 \times E_8$. Hacer dicha compactificación para los modos bosónicos tiene implicaciones directas sobre la fórmula de masa para estados y por ende en el espectro de partículas no masivas. Sin embargo, dicha compactificación toroidal conduce a una teoría dotada de $\mathcal{N} = 4$ supersimetrías y, por tanto, una teoría de norma no quiral. El tratamiento para dicho comportamiento es compactificando en *orbifolios*. Estas estructuras se obtienen modificando a toros agregando simetrías discretas. En el contexto de la teoría de cuerdas heteróticas tenemos inicialmente una compactificación toroidal $T_R^6 \otimes T_L^6 \otimes T_L^{16}$, donde el primer toro compactifica los modos derechos fermiónicos, el segundo a 6 modos izquierdos bosónicos y el último es el toro que compactifica al resto de los modos bosónicos izquierdos que dan lugar a las simetrías de norma. Torciendo a dichos toros con simetrías discretas, tenemos que el orbifoldo de compactificación de nuestra teoría queda como

$$O = T_R^6/P \otimes T_L^6/P \otimes T_L^{16}/G, \quad (5.1)$$

donde a P se le conoce como grupo de punto y a las isometrías G las llamamos grupo de torcimiento de norma. Para que en teorías con estas compactificaciones se mantenga una supersimetría no rota,

el grupo de punto P debe ser un subgrupo discreto de $SU(3)$, es decir, $P \subset SU(3)$. También se puede definir dicho orbifoldo en términos del grupo de espacio $S = \{(\theta, l^k) | \theta \in P, l^k \in \Lambda, k = 2, \dots, 8\}$ como

$$O = \mathbb{R}_{L+R}^6 / S \otimes T_{E_8 \times E_8}^L / G, \quad (5.2)$$

de tal manera que un elemento del grupo de espacio $s = (\theta, l^k)$ actúa en las 6 coordenadas espaciales a compactificar como

$$(sX)^k = (\theta X)^k + l^k, k = 3, \dots, 8. \quad (5.3)$$

En analogía con el caso de las 16 coordenadas de los grados de libertad extra. También en dicho capítulo mencionamos que el grupo de punto P sólo tiene la opción de ser \mathbb{Z}_N con $N = 3, 4, 6, 7, 8, 12$ o bien $\mathbb{Z}_N \times \mathbb{Z}_M$ con N múltiplo de M y $N = 2, 3, 4, 6$. También en el Capítulo 2 se da una primera clasificación de dichos grupos con las elecciones de los vectores de torcedura (v_1, v_2, v_3) provenientes de la forma de un elemento θ arbitrario del grupo de punto. Evidentemente, dicha compactificación modificada lleva a cambios importantes en los espectros de partículas provenientes de las fórmulas de masa correspondientes. El grupo de espacio es, en general, encajado dentro del grupo $E_8 \times E_8$ por medio del encaje estándar. En este encaje las coordenadas de los modos izquierdos bosónicos de norma se ven alterados por vectores de traslación y líneas de Wilson. Estas líneas de Wilson y vectores de traslación cumplen ciertas características dependiendo del grupo de punto con el que se hace la compactificación y están constriñidas por condiciones que garantizan invariancia modular en la teoría. También discutimos en este capítulo las peculiaridades inherentes a compactificaciones con orbifoldos con grupo de punto \mathbb{Z}_3 y \mathbb{Z}_6 -II, el primero siendo el que tiene la estructura más sencilla y el segundo siendo de gran interés fenomenológico por sus aplicaciones fenomenológicas [36, 22, 7].

Después exploramos las generalidades de Machine Learning y vemos brevemente cuáles eran sus tres ramas: ML no supervisado, ML supervisado y ML reforzado. Cada una de sus ramas con diferentes paradigmas, todas con el objetivo de extraer patrones, tendencias, características o regularidades de un conjunto enorme de datos, facilitando así su interpretación. La idea básica de ML es construir una Red Neuronal (RN), formada de varias capas de nodos, incluyendo una capa de entrada y una de salida. Los nodos entre capas están conectados por medio de transformaciones lineales cuyas matrices tienen entradas $w^{(i)}$ para la i -ésima capa más la suma de un vector constante $b^{(i)}$. Para el proceso de aprendizaje, necesitamos contar con datos de entrada y datos de salida o resultantes, es decir, contar con el resultado a partir de cierto conjunto de datos. Los datos de entrada se encuentran cuantificados en los nodos de la capa de entrada y el resultado en los nodos de la capa de salida. Entrenar a la RN significa ajustar los hiperparámetros $w^{(i)}$ y $b^{(i)}$ para que las capas de entrada y salida coincidan con los datos de entrada y los de resultado. Este proceso de ajuste se itera a través de varias épocas hasta que la función de pérdida (diferencia entre valores de la capa obtenidos y los deseados) se minimice. En particular, existe un tipo especial de red neuronal llamada *autoencoder* en la cual la capa de salida y entrada recibe los mismos datos, es decir, dicha red se entrena para reproducir a la función identidad. Generalmente en los autoencoder, las capas de salida y entrada son las que más nodos tienen y las capas intermedias son simétricas, dejando

en medio la capa con menor cantidad de nodos llamada capa latente. La idea de un autoencoder es codificar la información que recibe de entrada en menos parámetros con los escasos nodos de la capa latente.

Utilizando el orbifold [35], se generaron $\mathcal{O}(4,000,000)$ de modelos inequivalentes de orbifolds heteróticos con geometría \mathbb{Z}_6 -II (1,1), de acuerdo a la clasificación [19], con combinaciones consistentes de vectores de traslación y líneas de Wilson que llevan a física inequivalente. Estos modelos se preprocesaron mapeando esta información (64 parámetros) a vectores de características 26-dimensionales (ver detalles en Sección 3.3) y posteriormente dichos vectores se procesaron con la codificación one-hot que reemplaza al vector 26-dimensional en un vector 962 dimensional de puros ceros y unos. Estos vectores codificados son justamente la capa de entrada de nuestro autoencoder cuya estructura se detalla en la Figura 3.10. Con la estructura del autoencoder definida, entrenamos a dicha RN con un conjunto de datos de $\mathcal{O}(1,000,000)$ de modelos procesados teniendo una proporción de conjunto de entrenamiento y prueba de

$$\text{entrenamiento} : \text{validación} = 85 : 15, \quad (5.4)$$

a lo largo de 5,000 épocas y con un tamaño de lote de 2,000. Obtuvimos un valor de pérdida mínimo de 0.15, que corresponde a una predicción correcta de en promedio 14 características de 26, y nos aseguramos de no tener sobreentrenamiento. Recordemos los vectores 26 dimensionales, sin tomar en cuenta la codificación one-hot, son la entrada y salida de información de nuestro autoencoder. Cada entrada de dichos vectores es una característica cuyo valor numérico cuenta el número de raíces simples no invariantes que auxilian en el cálculo de las GUTs locales y el grupo de norma 4D \mathcal{G}_{4D} . De tal manera que hemos logrado predecir en promedio 14 de 26 raíces simples en cada uno de los modelos de compactificación, aproximadamente el 54% de ellas. Esta fue la mejor precisión alcanzada con la técnica actual presentada en este trabajo. Este no es el resultado final deseado pues desearíamos alcanzar una precisión de al menos el 80% para tener una mayor confianza en la eficacia del modelo del autoencoder. Ahora mencionemos las razones por las cuales no obtuvimos una mayor eficacia en la predicción de las características de nuestros modelos. La primera razón está relacionada con la complejidad del modelo para nuestro autoencoder; sólo probamos una arquitectura fija sin variar el número de nodos y capas. En particular, es posible que la eficacia aumente si incrementamos el número de nodos de la capa latente y, en consecuencia, tendríamos que explorar un espacio latente de mayor dimensionalidad. Al no haber variado la arquitectura de la RN, no podemos asegurar si la que utilizamos aquí es la óptima. Las siguientes razones están relacionadas con el preprocesamiento de los modelos, donde éstos son mapeados al espacio de características 26 dimensional. Estos vectores de características X tiene algunas ambigüedades: un modelo de cuerdas 4D es invariante bajo 1) intercambio entre los dos grupos de norma E_8 , 2) bajo ciertas permutaciones de los puntos fijos (ver Apéndice de [36]). En este trabajo sí realizamos un filtrado de modelos asegurándonos que no hubiera redundancias respecto al intercambio de los E_8 , pero no alcanzamos a hacer un filtrado respecto a la segunda redundancia, por lo que nuestro autoencoder pudo haber recibido modelos equivalentes y la eficacia del modelo de la RN obtenido se pudo ver afectada. Otra redundancia de los vectores de características 26 dimensional se encuentra

en que éstos son equivalentes ante reflexiones del grupo de Weyl para cada factor E_8 y dicho grupo es del orden de $\approx 7 \times 10^8$ de la cual no nos hicimos cargo.

Posteriormente, codificamos un conjunto de 2,000,000 de modelos (ajenos al conjunto de entrenamiento y de prueba) con el autoencoder, pudiendo así observar la forma del espacio latente para todos nuestros modelos. Utilizando el orbifolder generamos dos grupos de modelos con la geometría de compactificación \mathbb{Z}_6 -II (1,1) con espectros inequivalentes para modelos con un número de 3 generaciones 4D de los grupos de norma Pati-Salam (PS) (16 modelos) y $SU(5)$ (457 modelos). En [8] se hizo una clasificación de todas las simetrías de sabor 4D posibles del espectro no masivo de compactificaciones en orbifolios abelianos, incluyendo roto-traslaciones y toros no factorizables y se encontraron un total de 121,000 compatibles con el Modelo Estándar Mínimo Supersimétrico (MSSM, por sus siglas en inglés). Accesando al material suplementario de dicho trabajo vemos que del conjunto total de modelos tipo MSSM se encontraron 363 con geometría de compactificación \mathbb{Z}_6 -II (1,1) y los utilizamos para codificarlos con nuestro autoencoder entrenado. Estos modelos son compatibles con el MSSM si se cumplen ciertas condiciones, tales como que su grupo de norma no roto tenga como factor al del Modelo Estándar, que los estados efectivos incluyan campos que reproducen el espectro de materia del MSSM y que los estados efectivos adicionales sean vectoriales respecto al factor \mathcal{G}_{SM} del grupo de norma para dichos modelos (ver más detalles en la Sección 4.4). Estos tres conjuntos de modelos fenomenológicamente interesantes fueron codificados junto con el conjunto de 2,000,000 de modelos previamente mencionados. La forma en que se ve el espacio latente la podemos apreciar en la Figura 5.1.

Para identificar cúmulos o *clusters* de modelos en los cuales se agrupen modelos de las tres categorías previamente mencionadas, utilizamos el algoritmo K-means de TensorFlow [37] y en total se estableció un número de 30 clústers. Contamos cuántos modelos de cada categoría se encuentra en cada clúster según la numeración que apreciamos en la Figura 5.1 y dicho conteo se resume en la Tabla 5.1. Concluimos que las regiones 11, 12, 13 y 14 tienen la mayor acumulación de modelos tipo MSSM y $SU(5)$. Para el caso de modelos tipo PS, tenemos que se agrupan mayoritariamente en las regiones 1, 12, 15, 16 y 20. Tomemos en cuenta que la separación por regiones de nuestro espacio espacio latente es artificial, sobre todo porque el número de clústers independientes lo establecimos nosotros a mano y de manera completamente cualitativa. También apreciamos en la Figura 5.1 que no tenemos una clara jerarquía en cuanto a densidad de modelos por categoría se refiere, sí podemos saber cuáles regiones tienen la mayor cantidad de modelos, pero podemos ver que en casi todo el espacio latente podemos encontrar modelos fenomenológicamente interesantes. Aprovechando que la numeración de las regiones hace corresponder a regiones colindantes con indexación de números consecutivos, consideramos regiones más amplias constituidas por cúmulos de regiones dentro de las 30 ya identificadas. Tomamos entonces cúmulos de 5 en 5 regiones, teniendo en total 6 cúmulos, y en cada uno de ellos ahora contamos el número de modelos por categoría localizados allí. Se presenta una gráfica de barras en la Figura 4.6 con el conteo de modelos por categoría en las regiones agrupadas de 5 en 5.

Notamos que el grupo de regiones 11-15 tiene el mayor número de modelos en las tres categorías,

concluyendo así que en una búsqueda aleatoria de modelos fenomenológicamente interesantes, este grupo de regiones son las más prometedoras. Para el caso de $SU(5)$ observamos una mayor densidad de los mismos en las regiones 26-30. También aseguramos que los valores de todos los parámetros se encuentran dentro del dominio que se aprecia en la Figura 5.1 pues hay más de 2,000,000 de modelos generados aleatoriamente con esta geometría de compactificación mientras que en [7] se identificó que en total hay un número del orden de 10^7 modelos con esta geometría, por lo que estamos apreciando un 20% de ese total, es decir, un conjunto que sí es representativo.

Este trabajo es un primer paso, pues llegamos sólo a localizar las regiones en las cuales, 3 categorías de modelos fenomenológicamente interesantes se localizan en el espacio latente y dar pauta para posibles búsquedas en el futuro de modelos físicos. Como trabajo futuro deseamos encontrar una manera en la cual podamos mapear un punto del espacio latente a modelos 4D y analizar sus características físicas, tales como su grupo de norma. Es decir, dado un punto del espacio latente, nos gustaría poder contar el número de raíces simples invariantes del grupo de norma 4D por cada factor E_8 correspondiente a ese modelo en la capa latente. Esto permitiría clasificar orbifolios mediante la exploración del reducido espacio paramétrico de la capa latente, facilitando enormemente la búsqueda de construcciones prometedoras.

	R₁	R₂	R₃	R₄	R₅	R₆	R₇	R₈	R₉	R₁₀	R₁₁	R₁₂	R₁₃	R₁₄	R₁₅
MSSM	14	0	19	11	9	12	1	3	5	12	8	28	30	26	10
$SU(5)$	18	10	22	4	17	3	12	1	10	10	17	17	32	35	11
PS	2	0	0	0	0	1	0	0	1	0	0	2	0	0	3
	R₁₆	R₁₇	R₁₈	R₁₉	R₂₀	R₂₁	R₂₂	R₂₃	R₂₄	R₂₅	R₂₆	R₂₇	R₂₈	R₂₉	R₃₀
MSSM	22	9	10	6	17	6	8	11	7	20	11	15	6	19	7
$SU(5)$	14	12	2	5	12	2	5	5	16	44	10	31	12	51	16
PS	2	0	0	0	2	1	0	0	0	1	0	0	0	0	0

Tabla 5.1: Conteo de modelos por categoría en cada una de las 30 regiones en el espacio latente.

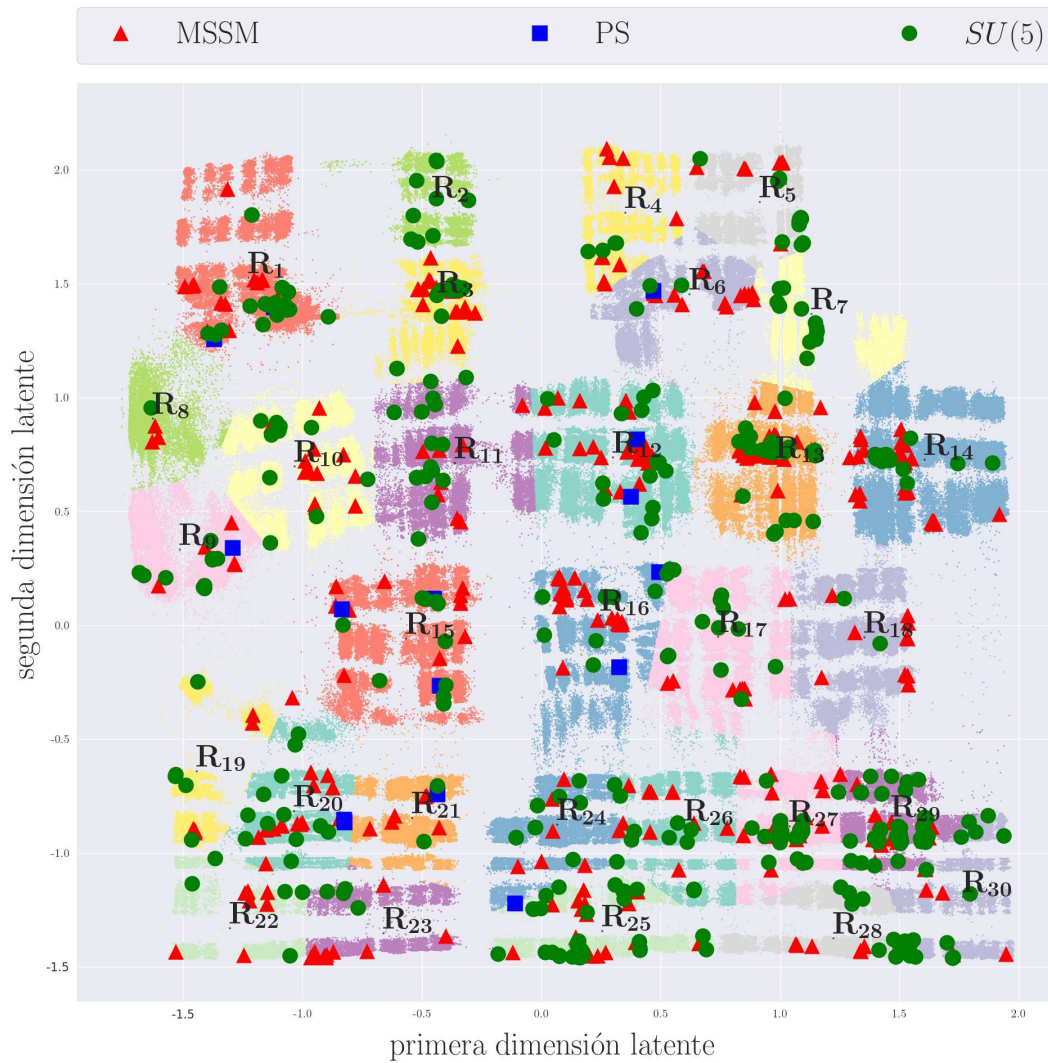


Figura 5.1: Espacio latente de parámetros dividido en 30 clústers. Notemos que la numeración de los mismos es en orden y en regiones colindantes. También apreciamos la localización de los modelos tipo MSSM, PS y $SU(5)$ en sus respectivos clústers.

Apéndice A

El sistema de raíces $E_8 \times E_8$

Aquí se presenta un breve resumen de los aspectos más fundamentales del sistema de raíces de $E_8 \times E_8$. Este apéndice está basado principalmente en [23]. El análisis necesario para determinar el espectro de estados en un modelo de orbifolio está íntimamente relacionado con los valores propio o pesos de estados $|W\rangle$ bajo elementos base H^I ($I = 1, \dots, 16$) del subálgebra de Cartan de $E_8 \times E_8$,

$$H^I |W\rangle = W^I |W\rangle. \quad (\text{A.1})$$

Los pesos de la representación adjunta son conocidos como raíces. Para $E_8 \times E_8$, la representación adjunta es la representación fundamental y representaciones de mayor dimensión son obtenidas del producto tensorial de la representación adjunta consigo misma. Los vectores de pesos suman cuando los productos tensoriales son utilizados para formar representaciones con dimensionalidad superior, entonces, los diagramas de pesos de representaciones de mayor dimensionalidad llenan una retícula de pesos generada por vectores base del diagrama de pesos de la representación adjunta. En el caso de $E_8 \times E_8$, esta es la retícula de raíces $\Lambda_{E_8 \times E_8}$. Una base en el espacio de raíces puede ser elegido de tal manera que la retícula de raíces E_8 puede ser escrita como el conjunto infinito de vectores 8-dimensionales

$$\Lambda_{E_8} = \left\{ (n_1, \dots, n_8), (n_1 + \frac{1}{2}, \dots, n_8 + \frac{1}{2}) \mid n_1, \dots, n_8 \in \mathbb{Z}, \sum_{i=1}^8 n_i = 0 \pmod{2} \right\}. \quad (\text{A.2})$$

Del conjunto anterior notamos que las componentes de un vector de raíces E_8 dado son todos enteros o todos medios enteros. Vectores l en la retícula Λ_{E_8} que satisfacen $l \cdot l = 2$ (donde el se supone implícitamente que el producto interior es el euclidiano 8-dimensional) producen las 240 raíces no nulas de E_8 , que denotamos por e_1, \dots, e_{240} . Por convención, tomamos como raíces positivas aquellas e_i cuya primer entrada no nula (de izquierda a derecha) es positiva. Una raíz simple es una raíz positiva que no puede ser obtenida de la suma de dos raíces positivas. Existen 8 raíces simples

para E_8 , que denotamos por $\alpha_1, \dots, \alpha_8$. Éstas forman una base de la retícula de raíces E_8 dada en (A.2), que alternativamente puede se escrita como

$$\Lambda_{E_8} = \left\{ \sum_{i=1}^8 m_i \alpha_i \mid m_i \in \mathbb{Z} \right\}. \quad (\text{A.3})$$

La retícula de raíces para $E_8 \times E_8$ se construye tomando la suma directa de dos copias de Λ_{E_8} , que distinguiremos por los superíndices (A) y (B),

$$\Lambda_{E_8 \times E_8} = \Lambda_{E_8}^{(A)} \oplus \Lambda_{E_8}^{(B)}. \quad (\text{A.4})$$

Así, un vector en la retícula de raíces de $E_8 \times E_8$ es un vector 16-dimensional que satisface

$$l = (l_A; l_B), \quad l_A \in \Lambda_{E_8}^{(A)}, \quad l_B \in \Lambda_{E_8}^{(B)}, \quad (\text{A.5})$$

donde heos denotado las primeras 8 entradas de l por l_A y las últimas 8 entradas. Las 480 raíces no nulas de $E_8 \times E_8$ están dadas por $(e_i; 0)$ y $(0; e_i)$, donde e_i es una de las 240 raíces no nulas de E_8 . De manera similar, las 16 raíces simples de $E_8 \times E_8$ están dadas por $\alpha_1, \dots, \alpha_{16}$. Tomando todas la combinaciones lineales de las 16 raíces simples con coeficiente enteros, se recupera la retícula de raíces $\Lambda_{E_8 \times E_8}$. Es decir,

$$\Lambda_{E_8 \times E_8} = \left\{ \sum_{i=1}^{16} m^i \alpha_i \mid m^i \in \mathbb{Z} \right\} \quad (\text{A.6})$$

Las 16 entradas de un vector en la retícula de raíces $(n_1, \dots, n_8; n_9, \dots, n_{16})$ corresponden a los valores propios respecto a la base del subálgebra de Cartan de $E_8 \times E_8$, que escribimos como H^I ($I = 1, \dots, 16$) y que cumplen con

$$\text{tr}_R(H^I H^J) = X(R) \delta^{IJ}, \quad (\text{A.7})$$

donde la traza se calcula sobre una irrep R de $E_8 \times E_8$. En particular, la representación adjunta (A) corresponde a las 480 raíces descritas arriba. No es difícil corroborar de (A.2) que $X(A) = 60$, que es dos veces el valor utilizado típicamente por los fenomenólogos. Es de particular importancia el mapeo de las raíces α_i dentro del subálgebra de Cartan por

$$H(\alpha_i) = \sum_{I=1}^{16} \alpha_i^I H^I. \quad (\text{A.8})$$

De esto, se define un producto interior en el espacio de raíces

$$\langle \alpha_i | \alpha_j \rangle \equiv \text{tr}_A[H(\alpha_i) \cdot H(\alpha_j)]. \quad (\text{A.9})$$

De (A.7), se puede calcular que

$$\langle \alpha_i | \alpha_j \rangle = X(A) \alpha_i \cdot \alpha_j. \quad (\text{A.10})$$

Se puede demostrar que el índice de Dynkin $X(A)'$ de la base (A.8) está relacionado con el índice de (A.7) por $X(A)' = 2X(A)$. Entonces los generadores (A.8) son más grandes por un factor de 2 respecto a la normalización fenomenológica. La matriz de Cartan de un álgebra de Lie está definida por

$$A_{ij} = \frac{2\langle\alpha_i|\alpha_j\rangle}{\langle\alpha_i|\alpha_j\rangle}, \quad (\text{A.11})$$

donde i, j son valores que corren sobre las raíces simples. Utilizando (A.10) y el hecho de que $\alpha_i^2 = 2$, es fácil corroborar que la matriz de Cartan (A.11) se puede calcular simplemente en términos de los vectores de raíces simple 16-dimensionales

$$A_{ij} = \alpha_i \cdot \alpha_j. \quad (\text{A.12})$$

En las construcciones con orbifolios, un subconjunto de raíces simples de $E_8 \times E_8$ sobreviven, y calculando las submatrices de acuerdo con (A.12), podemos identificar los factores no abelianos en el grupo de norma que sobrevive G .

Bibliografía

- [1] M. R. Douglas, *Journal of High Energy Physics* **2003** (2003), no. 05, 046–046.
- [2] Y.-H. He, *Deep-learning the landscape*, 2018.
- [3] D. Krefl and R.-K. Seong, *Physical Review D* **96** (2017), no. 6.
- [4] F. Ruehle, *Journal of High Energy Physics* **2017** (2017), no. 8.
- [5] J. Carifio, J. Halverson, D. Krioukov, and B. D. Nelson, *Journal of High Energy Physics* **2017** (2017), no. 9.
- [6] F. Ruehle, *Phys. Rept.* **839** (2020), 1–117.
- [7] O. Lebedev, H. P. Nilles, S. Ramos-Sánchez, M. Ratz, and P. K. Vaudrevange, *Physics Letters B* **668** (2008), no. 4, 331–335.
- [8] Y. Olguín-Trejo, R. Pérez-Martínez, and S. Ramos-Sánchez, *Physical Review D* **98** (2018), no. 10.
- [9] B. Zwiebach, *A first course in string theory*, Cambridge university press, 2004.
- [10] D. Bailin and A. Love, *Supersymmetric gauge field theory and string theory*, CRC Press, 1994.
- [11] K. Wray, Berkeley University (2011).
- [12] D. Tong, arXiv preprint arXiv:0908.0333 (2009).
- [13] A. Sen, *Tachyons in string theory*, From Fields to Strings: Circumnavigating Theoretical Physics: Ian Kogan Memorial Collection (In 3 Volumes), World Scientific, 2005, pp. 2035–2091.
- [14] L. E. Ibáñez, *The search for a standard model $su(3) \times su(2) \times u(1)$ superstring: an introduction to orbifold constructions*, Tech. report, 1987.
- [15] L. Dixon, J. Harvey, C. Vafa, and E. Witten, *Nuclear Physics B* **274** (1986), no. 2, 285–314.

- [16] F. Quevedo, *Heterotic strings in orbifold and wilson-line backgrounds*, Tech. report, CM-P00055691, 1987.
- [17] D. Bailin and A. Love, *Physics Reports* **315** (1999), no. 4-5, 285–408.
- [18] F. Gliozzi, J. Scherk, and D. Olive, *Nuclear Physics B* **122** (1977), no. 2, 253–290.
- [19] M. Fischer, M. Ratz, J. Torrado, and P. K. Vaudrevange, *Journal of High Energy Physics* **2013** (2013), no. 1.
- [20] L. Dixon, J. A. Harvey, C. Vafa, and E. Witten, *Nuclear Physics B* **261** (1985), 678–686.
- [21] *Group direct product of cyclic groups*, https://proofwiki.org/wiki/Group_Direct_Product_of_Cyclic_Groups, Acceso: 2021-06-04.
- [22] S. Ramos-Sánchez, *Fortschritte der Physik* **57** (2009), no. 10, 907–1036.
- [23] J. Giedt, arXiv preprint hep-ph/0204315 (2002).
- [24] T. Kobayashi, S. Raby, and R.-J. Zhang, *Physics Letters B* **593** (2004), no. 1-4, 262–270.
- [25] W. Buchmüller, K. Hamaguchi, O. Lebedev, and M. Ratz, *Nuclear Physics B* **712** (2005), no. 1-2, 139–156.
- [26] W. Buchmüller, K. Hamaguchi, O. Lebedev, and M. Ratz, *Nuclear Physics B* **712** (2005), no. 1-2, 139–156.
- [27] T. Kobayashi, S. Raby, and R.-J. Zhang, *Physics Letters B* **593** (2004), no. 1-4, 262–270.
- [28] T. Kobayashi and N. Ohtsubo, *Physics Letters B* **257** (1991), no. 1-2, 56–62.
- [29] F. Ruehle, *Physics Reports* **839** (2020), 1–117, Data science applications to string theory.
- [30] *The concepts of batch size, iteration and epoch in machine learning*, <https://www.programmersonsought.com/article/40565231807/>, Acceso: 2021-05-05.
- [31] *Epoch, batchsize, iterations*, <https://www.jianshu.com/p/e5076a56946c/>, Acceso: 2021-05-05.
- [32] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, *Self-normalizing neural networks*, 2017.
- [33] I. V. Tetko, D. J. Livingstone, and A. I. Luik, *Journal of chemical information and computer sciences* **35** (1995), no. 5, 826–833.
- [34] *Overfitting*, <https://en.wikipedia.org/wiki/Overfitting>, Acceso: 2021-06-04.

-
- [35] H. Nilles, S. Ramos-Sánchez, P. Vaudrevange, and A. Wingerter, *Computer Physics Communications* **183** (2012), no. 6, 1363–1380.
- [36] A. Mütter, E. Parr, and P. K. Vaudrevange, *Nucl. Phys. B* **940** (2019), 113–129, [1811.05993].
- [37] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *Tensorflow: Large-scale machine learning on heterogeneous distributed systems*, 2016.
- [38] *2.3. clustering*, <https://scikit-learn.org/stable/modules/clustering.html#k-means>, Acceso: 2021-05-05.
- [39] *El algoritmo k-means aplicado a clasificación y procesamiento de imágenes*, https://www.unioviado.es/compnum/laboratorios_py/kmeans/kmeans.html, Acceso: 2021-06-04.