Some Aspects of Field Theory

A thesis submitted to

Tata Institute of Fundamental Research, Mumbai, India

for the degree of

Doctor of Philosophy in Physics

By

Lavneet Janagal

Department of Theoretical Physics

Tata Institute of Fundamental Research

Mumbai-400005, India

June 2019

DECLARATION

I hereby declare that this thesis, "Some Aspects of Field Theory", is a presentation of my original research work, that I carried out with my collaborators. The work presented here does not feature as the work of someone else in obtaining a degree in this or any other institute. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

The work was done under the guidance of Professor Shiraz Minwalla, at the Tata Institute of Fundamental Research, Mumbai.

Date: 14th June 2019

Lavneet Janagal

Lavneet Janagal

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.

Prof. Shiraz Minwalla

Acknowledgements

I will always be grateful to my thesis advisor, Prof. Shiraz Minwalla for his constant support, guidance and encouragement. It has been an amazing experience to discover new and exciting things together with him. I am constantly amazed by his unique insights into Physics and other things. He was always patient and supportive during the times when my progress in research work was moving slow. The discussions with him and other members of string group at the lunch table were a lot of fun. His enthusiasm has been a constant source of motivation for me. I learned a lot about life from him.

I would like to thank Prof. Gautam Mandal and Prof. Sandip Trivedi for their support and stimulating discussions. I would also like to thank Prof. Abhijit Gadde for his support and guidance. It has been a wonderful experience to discuss Physics and collaborate with him.

I would like to thank all the faculty members and office staff in DTP for being supportive. It was nice to interact with my colleagues and friends in DTP, especially Subham, Indranil, Sounak, Geet, Bhawik, Yogesh, Arunabha, Nilakash, Mangesh, Ashish, Anshuman.

I would like to acknowledge the essential support and encouragement I received from my parents and my sister for pursuing research.

Collaborators

This thesis is based on work done in collaboration with several people.

- The work presented in chapter 1 was done in collaboration with Sayantan Choudhury, Anshuman Dey, Indranil Halder, Shiraz Minwalla, and Rohan Poojary.
- The work presented in chapter 2 was done in collaboration with Anshuman Dey, Indranil Halder, Sachin Jain, Shiraz Minwalla, and Naveen Prabhakar.
- The work presented in chapter 3 was done in collaboration with Subham Dutta Chowdhury, Abhijit Gadde, Indranil Haldar, and Shiraz Minwalla.

List of Publications

Publications relevant to the thesis

1. Sayantan Choudhury, Anshuman Dey, Indranil Halder, Lavneet Janagal, Shiraz Minwalla, and Rohan Poojary. "Notes on melonic $O(N)^{q-1}$ tensor models." JHEP, **06:094**, 2018.

2. Anshuman Dey, Indranil Halder, Sachin Jain, Lavneet Janagal, Shiraz Minwalla, and Naveen Prabhakar. "Duality and an exact Landau-Ginzburg potential for quasi-bosonic Chern- Simons-Matter theories." JHEP, **11:020**, 2018.

3. Subham Dutta Chowdhury, Abhijit Gadde, Indranil Haldar, Lavneet Janagal and Shiraz Minwalla. '–to-appear–'

Other publications

1. Sayantan Choudhury, Anshuman Dey, Indranil Halder, Sachin Jain, Lavneet Janagal, Shiraz Minwalla, and Naveen Prabhakar. "Bose-Fermi Chern-Simons Dualities in the Higgsed Phase." JHEP, **11:177**, 2018.

To

my family.

Contents

Chapter 1

Introduction

The work presented in this thesis is diverse. So, a separate primer is given for each of the chapters. The thesis first starts with a study of a class of quantum mechanical models, called tensor models. Various interesting aspects of these models are explored - emergence of new light modes, unusual thermodynamics. Then the next chapter dives into the rich sea of Chern-Simons matter theories. We strengthen the belief in the famous Chern-Simons matter dualities by providing more evidence for their validity, by explicitly verifying the conjectured free energy from these dualities. In the last chapter, we address a classification problem, namely 'what are all the classical $2-2$ s matrices coming from a generic theory of electrodynamics and gravity in arbitrary dimensions?'. Although, this problem is simple to state and fundamental in nature, there has not been much work in the literature regarding this. We solved it explicitly, along with special cases for low enough dimensions. Now, some detailed overview of our work in each of the chapters.

(Chapter 1) $O(N)^{q-1}$ Tensor Models

SYK model is a quantum mechanical model of N fermions with a 4 fermion interaction term. The model is a statistical average over couplings randomly chosen from a Gaussian ensemble. This model has attracted a lot of interest recently, because for one its exactly solvable in the large N limit. One can write down a Schwinger-Dyson equation for the two point correlator and solve it in the low energy limit as a perturbation expansion. It is quite remarkable that the exact solution displays extremely nontrivial dynamics. This model shows self equilibration over the time scales set by inverse temperature. It is already very interesting that a phenomena such as self thermalization is captured by an exactly solvable theory. Even more interestingly, the SYK model turns out to be maximally chaotic; out-of-time-ordered four point function grows with time; the Lyapunov exponent which measures this growth turns out to saturate the 'bound on chaos' proposed by Maldacena-Shenkar-Stanford [2]. This fact is particularly interesting because the only other known systems that also saturate this bound are field theories that have a gravitational dual description and whose thermodynamics is dominated by black holes. This coincidence has motivated the proposal that SYK theories have a (highly stringy and as yet unknown in detail) gravity

dual description whose thermal behavior is dominated by black holes (see e.g [9, 10]. If this conjecture is correct the fact that the SYK model is exactly solvable holds out the hope that the intense study of this model could allow us to address several key questions and puzzles about black hole physics, such as 'are black hole horizons smooth or do they have a firewall?' and (assuming that black hole event horizons are indeed smooth) 'what is the solution to Hawking's information paradox'. Over the last few years this hope has triggered a major research effort (spanning several hundreds of papers).

But this model is not strictly a unitary. After we average over the couplings it is not clear how to associate a Hilbert Space with the SYK system, and its energy spectrum (obtained by inverting the thermal partition function) is continuous rather than discrete. There is no useful sense in which time evolution in the SYK theory is generated by a unitary operator. As several of the deepest issues relating to black hole physics are intimately connected with unitarity, this observation seems worrying.

Witten [19] noticed that there exists another theory which shares several features of the SYK model, but has the virtue of being a unitary quantum theory in its own right. These are the tensor models. Its a model of fermions charged under the global/gauge symmetry group $O(N)^{q-1}$ which agrees with the large N limit of the SYK model. The leading order Feynman graphs for the two point function are exactly the same, i.e. the Schwinger Dyson equation of the tensor model is same as that of SYK model. Witten's observation sparked the following hope. Perhaps the tensor model - which is as solvable as the SYK model at large N and share several properties of this large N solution - have all the good features of the SYK model while also being genuine quantum mechanical systems. Perhaps, therefore, the tensor models are the useful toy models to study in order to draw lessons about black holes.

Motivated by this hope, in [3] we investigated several dynamical aspects of tensor models in detail and found, to our disappointment, that when the dynamics of these theories differ from that of the SYK theory in at least two qualitative ways, moreover the differences in question make the tensor models less like the dual of a gravitational theory. In the rest of this chapter we outline and explain our main results.

First, we argue that the spectrum of fluctuations about the finite temperature saddle point in these theories has $(q-1)\frac{N^2}{2}$ new light modes in addition to the light Schwarzian mode that exists even in the SYK model, suggesting that the bulk dual description of theories differ significantly if they both exist.

Second, we study the thermal partition function of a mass deformed version of the SYK model. At large mass we show that the effective entropy of this theory grows with energy like $E \log E$ (i.e. faster than Hagedorn) up to energies of order N^2. The canonical partition function of the model displays a deconfinement or Hawking Page type phase transition at temperatures of order $1/\log N$. We derive these results in the large mass limit but argue that they are qualitatively robust to small corrections in J/m.

(Chapter 2) Chern-Simons Matter Duality

The chapter in my thesis on Chern-Simons matter theories has its genesis in a paper

[54] I was a co author on but is not included in my thesis. First, I briefly summarize the contents of [54].

The paper [54] studied the conjectured duality between Fermions minimally coupled to Chern-Simons gauge fields (Regular Fermion or RF theory) and Wilson Fisher bosons minimally coupled to level rank dual gauge fields (Critical Boson or CB theory). At zero mass both theories described above define conformal field theories (CFTs). Each of these CFTs admits a a single relevant mass deformation parameterized by a real mass.

Let m denote the mass of the theories above. By scaling, the low energy physics (i.e. the 'phase' in condensed matter language) is invariant under the scaling $m \to \lambda m$ where λ is a positive real number. However the phase can and does depend on the sign of m. On the fermionic side this comes about as follows. When the fermion is massive it can be integrated out and the low energy physics in the IR is topological and is governed by pure Chern Simons theory. The rank of the low energy Topological Field Theory is same as the rank of the original UV Chern Simons theory. However, it is well known that the level of the low energy CFT differs from the UV level by the formula

$$k_{IR} = K_{UV} \pm \frac{\mathrm{sgn}(m)}{2}.$$

It follows that the low energy Chern Simons theories that govern the dynamics of theories with fermionic masses of different signs have levels that differ by unity.

If the conjectured level rank duality between the RF theories and the CB theories is indeed correct, then this must imply that, on the bosonic side of this duality, the low energy theories that govern the dynamics of the bosonic theory with positive and negative masses must have ranks that differ by unity. This is indeed the case and comes about as follows. When the bosons have positive (Wilson Fisher) masses, the low energy Chern Simons theory has both the same level and the same rank as the UV theory. When the bosons have negative Wilson Fisher masses, on the other hand, the bosons condense. The condensation breaks the $SU(N)$ gauge symmetry down to $SU(N-1)$. As a consequence the low energy Chern Simons theory has the same level as the UV theory, but a rank that is one less than that of the UV theory, in perfect agreement with the predictions of duality [58].

While the picture spelt out in the previous paragraph sounds compelling, it is qualitative in nature, and one would like to find quantitative evidence for its validity. When both masses are positive, such evidence was obtained in earlier work, when, for instance, the thermal partition functions of both theories was computed in detail at large N and found to match perfectly under the conjectured duality [83, 56].

When the masses were negative, the fermionic thermal partition function had, once again been computed in earlier work. However the bosonic partition function had not yet been obtained. This major deficiency was remedied in our paper [54] in which we performed this computation. The computation required many new ideas - symmetry breaking triggered the Higgs mechanism. In order to find the thermal free energy in this situation we needed to solve the Schwinger Dyson equation for the new physical particles in our problem, namely the W bosons. We were able to accomplish this in [54] and demonstrate

that the final result for the thermal partition function agrees perfectly under duality with that for the fermionic theory under duality.

In this chapter of my thesis we consider the follow-up [43] to the work mentioned above. The similar conjecture exists for the duality between the Chern-Simons (CS) gauged regular bosons in the fundamental representation and the level-rank dual theory- CS gauged critical fundamental fermions. As in the previous subsection, the conjecture of duality implies that the set of all RG flows that originate in one theory are dual to the RG flows from the dual theory. Generic RG flows from these quasi-bosonic fixed points lead to gapped phases, whose low energy dynamics are governed by topological field theory. There are two inequivalent topological phases, unHiggsed phase, and the Higgsed phase. In the unHiggsed phase the bosonic (resp. fermionic) theory is governed at long distances by pure $SU(N_B)_{k_B}$ (resp. $U(N_F)_{k_F}$) topological field theory with two theories level-rank dual to each other. In the Higgsed phase the bosonic (resp. fermionic) theory is governed in the IR by a pure $SU(N_B - 1)_{k_B}$ (resp. $U(N_F)_{\text{sgn}\, k_F(|k_F|-1)}$) Chern-Simons theory with the two topological field theories once again level-rank dual to each other.

As in the previous subsection the qualitative guesses of the last paragraph had been made quantitative in earlier work only in one half of the phase diagram. In particular the large N thermal free energy for the bosonic theory (RB) in the unHiggsed phase had earlier been demonstrated to match the corresponding fermionic results under duality [83, 56]. In this paper we complete this qualitative matching by evaluating the large N thermal free energy of the bosonic theory in the Higgsed phase and demonstrate that our results, again, perfectly match the predictions of duality. Our computation is performed in a unitary gauge (as mentioned before) by integrating out the physical excitations of the theory - i.e. W bosons - at all orders in the 't Hooft coupling.

The results and calculations are very similar to those of [54] except one difference. In the critical boson theory, in [54], the 'classical' potential for the scalar field is infinitely deep. This potential freezes the magnitude of the scalar field in the Higgsed phase to its classical minimum even in the quantum theory. It follows that the Higgs vev is independent of the temperature and has a simple dependence on the 't Hooft coupling. In the regular boson theory, however, the classical potential for the scalar field is finite and receives nontrivial quantum corrections. The value of the scalar condensate is determined extremizing the quantum effective action for the scalar field. The result of this minimization yields a scalar vev that is a nontrivial function of both the 't Hooft coupling and the temperature. We are able to compute the smooth 'quantum effective potential' for the RB theory as a function of the Higgs vev. More precisely we compute the quantum effective potential for the composite field $(\overline{\phi}\phi)$. In the Higgsed phase and in the unitary gauge this quantity reduces to a potential - an exact Landau-Ginzburg effective potential. The extremization of this potential determines the Higgs vev.

We emphasize that one of the key results of our work is an exact computation of the quantum effective potential for $\overline{\phi}\phi$ in the RB and CF theories, to all orders in the coupling constants. This quantum effective potential is rich and contains a lot of physics, some of which is explored in this chapter. We find that fact that we are able to compute such a nontrivial and dynamically rich quantity exactly quite remarkable.

In addition to giving us quantitative evidence for duality, the new exact result for the quantum effective potential for the Higgs field gives us new understanding of the RB and CF theories in a global manner. In particular it turns out that this quantum effective potential is bounded from below only for a certain range of x_6 (the parameter that governs sextic interactions of ϕ), demonstrating that the regular boson theory has a stable vacuum only when x_6 lies in this range.

Moreover, the form of the quantum effective potential allows us to compute the phase diagrams of the RB and CF theory in a simple and intuitive manner.

Finally, in the the zero temperature limit this exact Landau-Ginzburg potential is non-analytic at $\overline{\phi}\phi = 0$. The extrema of this effective potential at positive $\overline{\phi}\phi$ solve the gap equations in the Higgsed phase while the extrema at negative $\overline{\phi}\phi$ solve the gap equation in the unHiggsed phase.

(Chapter 3) Classifying and Constraining 4 Photon and 4 Graviton classical S-Matrices.

Consider the S matrix of a theory of four identical massless scalars. It is well known that the most general S matrix in such a theory is specified by a single function

$$A(s,t)$$

which enjoys the following symmetry properties

$$A(s,t) = A(t,s) = A(s,u), \quad u = -s - t$$

where s, t and u are the usual Mandlestam invariants.

Now consider a slightly more complicated kinematical setting - the scattering of four identical photons or four identical gravitons. What is the most general S matrix for such a theory allowed simply on kinematical grounds? While this question is a very simple and natural one, but has never been systematically addressed in the literature. In this chapter we supply the answer to this question both for photons and for gravitons. We find that there are 3 basic structures in terms of which any S-matrix of a classical theory of electrodynamics can be written, and for gravity we find that the number of basic structures for classical $2-2$ scattering is 10, in high enough dimensions. We solved the classification question described above as preparatory material in order to address a question that we believe is of great physical importance, namely the classification of consistent local classical theories of gravity and electromagnetism.

Chapter 2

Tensor Models

2.1 Introduction

It has recently been demonstrated that the dynamically rich Sachdev-Ye-Kitaev model - a quantum mechanical model of fermions interacting with random potentials - is solvable at large N [1, 4, 5]. This model is interesting partly because its thermal properties have several features in common with those of black holes. The SYK model self equilibrates over a time scale of order the inverse temperature and has a Lyapunov index that saturates the chaos bound [5, 4]. Moreover the long time behaviour of this model at finite temperature is governed by an effective action that has been reinterpreted as a particular theory of gravity expanded about AdS_2 background solution [1, 6, 7, 8, 10, 11, 12, 13].

These facts have motivated the suggestion that the SYK model is the boundary dual of a highly curved bulk gravitational theory whose finite temperature behaviour is dominated by a black hole saddle point. If this suggestion turns out to be correct, the solvability of the SYK model at large N - and its relative simplicity even at finite N- could allow one to probe old mysteries of black hole physics in a manner that is nonperturbative in $\frac{1}{N}$, the effective dual gravitational coupling (see e.g. [14, 15, 16] for recent progress).

There is, however, a potential fly in the ointment. While the SYK model - defined as a theory with random couplings - is an average over quantum systems, it is not a quantum system by itself. One cannot, for instance, associate the SYK model with a Hilbert space in any completely precise manner, or find a unitary operator that generates time evolution in this model. As several of the deepest puzzles of black hole physics concern conflicts with unitarity, this feature of the SYK model is a concern.

Of course any particular realization of the couplings drawn from the SYK ensemble is a genuine quantum theory. It is plausible that several observables - like the partition function - have the same large N limit when computed for any given typical member of the ensemble as they do for the SYK model defined by averaging over couplings [15, 17, 18]. It might thus seem that *every* typical realization of random couplings is an in-equivalent consistent quantization of classical large N SYK system. As the number of such quantizations is very large, this would be an embarrassment of riches. The potential issue here is that if we work

with any given realization of the SYK model, it appears inconsistent to restrict attention to averaged observables for any finite N no matter how large. On the other hand correlators of individual ψ_i operators (as opposed to their averaged counterparts) presumably do not have a universal large N limit (and so are not exactly solvable even at large N). [1]

In order to address these concerns some authors have recently [19, 20, 21] (based on earlier work [22, 23, 24, 25, 26, 27]) studied a related class of models. These models are ordinary quantum mechanical systems; in fact they describe the global or gauged quantum mechanics of a collection of fermions in 0+1 dimensions. In this chapter we will focus our attention on the model

$$\begin{aligned} S &= \int dt \sum_{a=1}^{N_F} [\overline{\psi}_a D_0 \psi_a - (g\ \psi_a^q + h.c.)], \\ D_0 &= \partial_0 + iA_0\ ,\ g = \frac{J}{N^{\frac{(q-1)(q-2)}{4}}}, \end{aligned} \tag{2.1}$$

that was first proposed - at least in the current context - in [20] . In (2.1) ψ_a are a collection of complex gauged fermionic fields in $0+1$ dimensions that transform in the fundamental of each of the $q-1$ copies of $O(N)$. The index a is a flavour index that runs from $1 \ldots N_F$. [2]. J is a coupling constant with dimensions of mass and ψ^q is a schematic for a q vertex generalization of a 'tetrahedronal' interaction term between q copies of the fermionic fields, whose gauge index contraction structure is explained in detail in [19, 20] and will be elaborated on below.

The tetrahedral structure of the interaction[19, 20] is such that for any even number of fermions q each fermion has $q-1$ indices each in a different $O(N)(or\, U(N))$. The indices among the q fermions are contracted such that every fermion is index contracted with an index of the same gauge group on one of the remaining fermions. Moreover, given any - and every - 2 fermions have a single index (of some gauge group) contracted between them. For $q=4$ it is easy to check that these words define a unique contraction structure which may be viewed as a tetrahedral contraction among the 4 fermions each with $q-1=3$ indices(legs) with every fermion(point or vertex of the tetrahedron) connected to 3 different coloured legs. For $q \geq 6$ it is not clear that the words above define a unique contraction structure. In case the contraction structure is not unique, we pick one choice - for example the Round-Robin scheduling process to define our interaction [28, 29]. [3]

The connection between the quantum mechanical theories (2.1) and the SYK model itself is the following; it has been demonstrated (subject to certain caveats) that sum over Feynman graphs of the theory (2.1) coincides with the sum over Feynman graphs of the SYK model at at leading order at large N (see [19] for the argument in a very similar model), even though these two sums differ at finite values of N (see e.g. the recent paper [30] and references therein). It follows that the quantum mechanical models (2.1) are

[1] We thank S. Sachdev for discussion on this point.

[2] For this simplest case $N_F = 1$ this model was presented in Eq 3.23 of [20].

[3] We would like to thank J. Yoon for explaining the Round Robin scheduling process to us and clearing up our misconceptions about uniqueness of the contraction structure for $q > 4$.

exactly as solvable as the SYK model at large N; moreover they also inherit much of the dynamical richness of the SYK model. In other words the models (2.1) are solvable at large N, are unitary and are potentially boundary duals of (highly curved) black hole physics.

Motivated by these considerations, in this chapter we study the effective theory that governs the long time dynamics of the model (2.1) at finite temperature. We focus attention on dynamical aspects of (2.1) that have no counterpart in the already well studied dynamics of the original SYK model. [4]

In the rest of this introduction we will explain and describe our principal observations and results.

2.1.1 New light modes

The thermal behaviour of both the theory (2.1) and the original SYK model is determined by the path integral of these theories on a circle of circumference β.

It was demonstrated in [4, 5] that, in the case of the original SYK model, this path integral is dominated by a saddle point of an effective action whose fields are the two point function and self energy of the fermions. An extremization of this effective action determines both the fermionic two point function at finite temperature as well as the free energy of the system at leading order at large N.

In a similar manner, the thermal behaviour of the quantum mechanical systems (2.1) is dominated by a saddle point at large N. Under appropriate assumptions it may be shown that resultant effective action has the same minimum as that of the original SYK theory [19]. [5]. Specializing to the case $N_F = 1$, the leading order fermionic two point function of the quantum mechanical system is given by

$$\langle \overline{\psi}^a(t)\psi_b(t')\rangle = \delta^a_b G^{SYK}(t-t'), \tag{2.2}$$

where a and b denote the (collection of) vector indices for the fermions and $G^{SYK}(t)$ is the thermal propagator of the original SYK model. [6]

While the thermal behaviour of the model (2.1) is thus indistinguishable from that of the SYK model at leading order in the large N limit, the dynamics of the quantum

[4] See [31, 32, 33, 34, 35, 36, 37, 28, 38, 39, 40] for other recent work on the model (2.1) and its close relatives.

[5] A potential subtlety is that path integral of the quantum mechanical system (2.1) has a degree of freedom that is absent in the original SYK model, namely the holonomy of the gauge group $O(N)^{q-1}$. As for the SYK model, integrating out the fermions leads to an effective action - proportional to N^{q-1} - whose fields are a two point function of the fermions, a self energy and the holonomy of the gauge group. As in the case of the original SYK model, at leading order in the large N limit the free energy of the system is captured by the saddle point of this effectively classical action. If we work at temperatures that are held fixed as $N \to \infty$ it is highly plausible that this effective action is minimized when the holonomy is the identity matrix (see 2.3 below). Under this assumption the saddle point of the quantum mechanical system coincides with that of the SYK model.

[6] (2.2) applies both the the case that the group $O(N)^{q-1}$ is global and local. In the latter case this equation applies in the gauge $\partial_0 A_0 = 0$. Assuming that the holonomy degree of freedom is frozen to identity at large N, the gauged and global model coincide.

mechanical model (2.1) differs from that of the SYK model at subleading orders in $1/N$. The first correction to leading large N thermal behaviour may be obtained by performing a one loop path integral over quadratic fluctuations around the saddle point. In the long time limit, correlators are dominated by the lightest fluctuations around the saddle point.

Recall that in the UV (i.e. as $\beta J \to 0$) the fermions of (2.1) have dimension zero. The term proportional to ψ^q in (2.1) represents a dimension zero relevant deformation of this UV fixed point. The resultant RG flow ends in the deep IR in a conformal field theory in which the fermions have dimension $\frac{1}{q}$. [4, 5]. In this IR limit (relevant to thermodynamics when $\beta J \to \infty$) ψ^q is marginal while the kinetic term in (2.1) is irrelevant [4, 5]. The fact that the kinetic term is irrelevant in the IR - and so can effectively be ignored in analyzing the symmetries of (2.1) at large βJ - has important implications for the structure of light fluctuations about the thermal saddle point.

The first implication of the irrelevance of the kinetic term occurs already in the SYK model and was explored in detail in [4, 5, 8]. The main point is that the action (2.1), with the kinetic term omitted, enjoys invariance under conformal diffeomorphisms (i.e. diffeomorphisms together with a Weyl transformation). However the saddle point solution for the Greens function $G^{\rm SYK}(t)$ is not invariant under conformal diffeomorphisms. It follows immediately that the action of infinitesimal conformal diffeomorphisms on this solution generates zero modes in the extreme low energy limit.

At any finite temperature, no matter how small, the kinetic term in (2.1) cannot completely be ignored and conformal invariance is broken; the action of conformal diffeomorphisms on the SYK saddle point consequently produces anomalously light (rather than exactly zero) modes. The action for these modes was computed in [4, 5, 8] and takes the form of the Schwarzian for the conformal diffeomorphisms.

A very similar line of reasoning leads to the conclusion that the model (2.1) has $(q-1)\frac{N^2}{2}$ additional light modes in the large βJ limit, as we now explain. Let us continue to work in the gauge $A_0 = 0$. In this gauge the action (2.1) is obviously invariant under the global rotations $\psi \to V\psi$, $\overline{\psi} \to \overline{\psi} V^\dagger$ where V is an arbitrary time independent $O(N)^{q-1}$ rotation. In the global model (2.1) the rotation by V is the action of a global symmetry. In gauged model on the other hand, these rotations are part of the gauge group and do not generate global symmetries of our model; the Gauss law in the theory ensures that all physical states are uncharged under this symmetry.

Let us now consider the transformation $\psi \to V(t)\psi$ together with $\overline{\psi} \to \overline{\psi} V(t)^\dagger$ where $V(t)$ is an arbitrary time *dependent* $O(N)^{q-1}$ rotation. In the case of the gauged models, this transformation is not accompanied by a change in A_0 ($A_0 = const$ throughout) so the rotation is not a gauge transformation.

At finite βJ the rotation by a time dependent $V(t)$ is not a symmetry of the action (2.1) in either the global or the gauged theory as the kinetic term in (2.1) is not left invariant by this transformation. As we have explained above, however, the kinetic term is irrelevant in the low temperature limit $\beta J \to \infty$. It follows that the time dependent transformation is an effective symmetry of dynamics this strict low temperature limit.

However the saddle point solution (2.2) is clearly not invariant under the time dependent rotations by $V(t)$. It follows that, as in the discussion for conformal diffeomorphisms above, the action of $V(t)$ on (2.2) generates exact zero modes in the strict limit $\beta J \to \infty$ and anomalously light modes at any finite βJ. We emphasize that this discussion applies both to the global model where $O(N)^{q-1}$ is a global symmetry, and the gauged model where it is not.

In section 2.2.2 below we argue that the dynamics of our new light modes is governed by the effective sigma model on the group manifold

$$S = -\mathcal{A}\ \frac{N^{q-2}}{|J|} \int dt\ \sum_{l=1}^{q-1}\ Tr\ \left[\left(V_l^{-1}(t)\ \frac{\partial}{\partial t} V_l(t)\right)^2\right], \tag{2.3}$$

where $V_l(t)$ is an arbitrary element of the group $O(N)$ and $\mathcal{A}$ is a number of order unity that we have not been able to determine.

The formula (2.3) has appeared before in a closely related context. The authors of [41] (see also [17]) studied the a complex version of the SYK model. Their model had an exact $U(1)$ symmetry at all energies, which - using the arguments presented in the previous paragraphs - was approximately enhanced to a local $U(1)$ symmetry at low energies. The authors of [41] argued the long distance dynamics of the new light modes is governed by a sigma model on the group manifold $U(1)$. [7] Given these results, the appearance of a low energy sigma model in the large βJ finite temperature dynamics of the theory (2.1) seems natural.

We would, however, like to emphasize two qualitative differences between the sigma model (2.3) and the model that appeared in [41]. First (2.3) is a sigma model for a group $O(N)^{q-1}$ whose dimensionality goes to infinity in the large N limit, $N \to \infty$. Second that we find the new light modes of the action even of the gauged model (2.1) even though $O(N)^{q-1}$ is not a global symmetry of this theory.

The new modes governed by (2.3) are approximately as light - and so potentially as important to long time dynamics - as the conformal diffeomorphisms described above. Note, however, that there are $(q-1)\frac{N^2}{2}$ light time dependent $O(N)^{q-1}$ modes but (as far as we can tell) only one conformal diffeomorphism.

We have already remarked above that the light diffeomorphism degree of freedom described above has been given an interpretation as a particular gravitational action in an AdS_2 background. It seems likely to us that the effective action (2.3) will, in a similar way, admit a bulk interpretation as a gauge field propagating in AdS_2. The Yang Mills coupling of this gauge field - like Newton's constant for the gravitational mode - will be of order $\frac{1}{N^{q-1}}$ (this is simply a reflection of the fact that our model has N^{q-1} degrees of freedom). This means that the t' Hooft coupling of all the gauge fields in the bulk will be of order $g_{YM}^2 N \sim \frac{1}{N^{q-2}}$. The fact that this coupling goes to zero in the large N limit implies that

[7] They also argued for some mixing between the diffeomorphism and $U(1)$ long distance modes.

the bulk gauge fields are classical even though there are so many of them. [8]

It has been established that the light diffeomorphism degree of freedom has a qualitatively important effect on out of time ordered thermal correlators; it leads to exponential growth in such correlators at a rate that saturates the chaos bound $G \sim e^{2\pi T t}$. When we include the contribution of the new light modes described in this subsection, we expect this growth formula to be modified to

$$G(t) \sim \left(e^{2\pi T t} + N^2 f(t)\right). \tag{2.4}$$

[9] The factor of N^2 is a reflection of the fact that our new modes are N^2 in number, whereas - as far as we can tell - there is only a single light mode corresponding to conformal diffeomorphisms.

Given that the solutions of the equations of motion to the Sigma model (2.3) grow no faster than linearly in time, we expect $f(t)$ to grow at most polynomial in time. This suggests it that the light modes (2.3) will dominate correlators up to a time of order $\frac{1}{\pi T}\log N$. At later times the exponentially growing diffeomorphism mode will dominate, leading to exponential growth and a Lyapunov index that saturates the chaos bound.

To end this subsection let us return to a slightly subtle point in our discussion. In order to derive the effective action for $V(t)$ we worked in the gauge $A_0 = 0$. As our theory is on a thermal circle, in the case of the gauged model (2.1) we have missed a degree of freedom - the gauge holonomy - by working in the gauge $A_0 = 0$. This, however, is easily corrected for. Even in the presence of a holonomy, we can set the gauge field A_0 to zero by a gauge transformation provided we allow ourselves to work with gauge transformations that are not single valued on the circle. The net effect of working with such a gauge transformation is that the matter fields are no longer periodic around the thermal circle but obey the boundary conditions

$$\psi(\beta) = -U\psi(0), \tag{2.5}$$

where U is the holonomy around the thermal circle. For the fields of the low energy effective action (2.3) this implies the boundary conditions

$$V(\beta) = UV(0)U^{-1}. \tag{2.6}$$

Recall we are instructed to integrate over all values of the holonomy U. Consequently we must integrate over the boundary conditions (2.6) with the Haar measure. See section A.3 for some discussion of this point.

In summary, the discussion of this subsection suggests that the bulk low energy effective action 'dual' to the gauged/global quantum mechanics of (2.1) differs from the low energy effective action 'dual' to the SYK model in an important way; in addition to the gravitational field it contains gauge fields of a gauge group whose rank is a positive fractional power of the inverse Newton (and Yang Mills) coupling constant of the theory. In

[8] We would like to thank J. Maldacena for a discussion of this point.

[9] See [29] for related work.

the classical limit in which Newton's constant is taken to zero, the rank of the low energy gauge fields also diverges. Nonetheless the limits are taken in such a way that the effective bulk theory remains classical.

2.1.2 Holonomy dynamics and the spectrum at large mass

Our discussion up to this point has applied equally to the 'global' and 'gauged' quantum mechanical models (2.1). In the rest of this introduction we focus attention on the gauged models, i.e. the models in which the $O(N)^{q-1}$ symmetry algebra is gauged. In this case the thermal path integral of our system includes an integral over gauge holonomies over the thermal circle. We wish to study the effect of this holonomy integral on the dynamics of our system.

In order to do this in the simplest and clearest possible way we deform the model (2.1) in a way that trivializes the dynamics of all non holonomy modes in the theory. This is accomplished by adding a mass to the fermions. For concreteness we work with the $O(N)^{q-1}$ model

$$\begin{aligned} S &= \int dt \sum_{a=1}^{N_F} \left[\left(\overline{\psi}_a D_0 \psi_a + m \overline{\psi}_a \psi_a \right) - \left(g \ \psi_a^q + h.c. \right) \right], \\ D_0 &= \partial_0 + i A_0 \ , \ g = \frac{J}{N^{\frac{(q-1)(q-2)}{4}}}, \end{aligned} \tag{2.7}$$

where m, the mass of the fermion is taken to be positive. [10] We work the large mass limit, i.e. the limit $\frac{m}{J} \gg 1$. The effective interaction between fermions in (2.7), $\frac{J}{m}$, is small in this limit and can be handled perturbatively. In the strict $m \to \infty$ limit the only interaction that survives in the system is that between the (otherwise free) matter fields and the holonomy U. [11]

Let us first work in the strict limit $\frac{m}{J} \to \infty$. In this limit the dynamics of the holonomy field U in this theory is governed by an effective action obtained by integrating out the matter fields at one loop. [12] The resultant effective action is easily obtained and is given

[10] In the case that the mass is negative, most of our formulae below go through once under the replacement $m \to |m|$.

[11] We emphasize that, in the limit under consideration, modes corresponding to diffeomorphisms or $V(t)$ are no longer light - and so are irrelevant. However the holonomy continues to be potentially important.

[12] For orientation, we remind the reader that the integral over the holonomy is the device the path integral uses to ensure that the partition function only counts those states that obey the A_0 equation of motion, i.e. the Gauss law constraint. Restated, the integral over holonomies ensures that the partition function only counts those states in the matter Hilbert space that are singlets under the gauge group.

by ([42])

$$
\begin{aligned}
Z = \mathrm{Tr} x^{\frac{H}{m}} &= \int \prod_{i=1}^{q-1} dU_i \exp(-S_{\text{eff}}(U_i)), \\
S_{\text{eff}}(U_i) &= -N_F \sum_{n=1}^{\infty} \frac{(-x)^n \left(\prod_{i=1}^{q-1} \mathrm{Tr} U_i^n\right)}{n}, \\
x &= e^{-\beta |m|},
\end{aligned}
\tag{2.8}
$$

where H is the Hamiltonian of our theory. [13]

Each U_i is an $O(N)$ matrix that represents the holonomy in the i^{th} factor in the gauge group $O(N)^{q-1}$. dU is the Haar measure over the group $O(N)^{q-1}$ normalized so that the total group volume is unity.

Notice that when x is of order unity, $S_{\text{eff}} \sim N^{q-1}$ in (2.8). On the other hand the contribution of the group measure to the 'effective' action is of order N^2. The integral in (2.8) is interesting when these two contributions are comparable. This is the case if we scale temperatures so that

$$
x = e^{-\beta |m|} = \frac{\alpha}{N_F N^{q-3}}, \tag{2.10}
$$

with α held fixed as N is taken to infinity. In this limit the terms in the second of (2.8) with $n > 1$ are subleading and can be ignored. Effectively

$$
\begin{aligned}
Z(x) &= \int \prod_{i=1}^{q-1} dU_i \exp(-S_{\text{eff}}(U_i)) \\
S_{\text{eff}} &= \frac{\alpha}{N^{q-3}} \left(\prod_{i=1}^{q-1} \mathrm{Tr} U_i \right).
\end{aligned}
\tag{2.11}
$$

In the large N limit the matrix integral (2.11) is equivalent - as we show below - to the well known Gross Witten Wadia model and is easily solved. The solution - presented in detail below - has the following features

- 1. In the canonical ensemble, the partition function undergoes a deconfinement type phase transition at $\alpha = \alpha_{1pt}$ where the value of α_{1pt} is given in (2.77). At smaller values of α the system is dominated by the 'confining' saddle point in which U is the clock matrix. At larger values of α_{1pt} the system is dominated by a more complicated

[13] The generalization of these results to a model with N_B bosons and N_F fermions yields the holonomy effective action

$$
S_{\text{eff}}(U_i) = \sum_{n=1}^{\infty} (N_B + (-1)^{n+1} N_F)\ x^n \frac{\left(\prod_{i=1}^{q-1} \mathrm{Tr} U_i^n\right)}{n}. \tag{2.9}
$$

As we will see below, in the scaling limit of interest to this chapter, only the term with $n = 1$ is important. In the strictly free limit it follows that most the results presented above apply also to a theory with N_F fermions and N_B bosons once we make the replacement $N_F \to N_F + N_B$.

'deconfined' or black hole saddle point. The phase transition is reminiscent of the transitions described in [44, 45]. [14]

- 2. In the microcanonical ensemble, the scaling limit described above captures the density of states of the system at energies less than or of order N^2. Over the range of energies $1 \ll E < \frac{N^2}{4}$, the entropy S is given by the simple formula

$$S(E) = (q-3)\left[\frac{E}{2}\log\left(\frac{E}{2}\right) - \frac{E}{2}\right] + E\ \log N_F + (q-3)\frac{E}{2}\log(2). \tag{2.12}$$

The saddle point that governs the density of states of the theory changes in a non analytic manner at $E = \frac{N^2}{4}$. For $E > \frac{N^2}{4}$ the formula for the entropy is more complicated. For energies $E \gg (q-2)\frac{N^2}{4}$, however, the entropy simplifies to the formula for $n_B N^{q-1}$ complex bosonic and $n_F N^{q-1}$ free complex fermionic harmonic oscillators

$$S(E) = E\left[1 - \log\left(\frac{E}{pN^{q-1}}\right)\right]. \tag{2.13}$$

The complicated formula that interpolates between these special results is presented in (2.99).

The formula (2.12) suggests that if a dual bulk interpretation of the theory (2.8) exists, it is given in terms of a collection of bulk fields whose number grows faster than exponentially with energy. It would be fascinating to find a bulk theory with this unusual behaviour.

Moreover the existence of a Hawking Page type phase transition in this model - and in particular the existence of a subdominant saddle point even at temperatures at which the dominant phase is a black hole phase - opens the possibility of the subdominant phase playing a role in effectively unitarizing correlators about the black hole saddle point by putting a floor on the decay of the amplitude of correlators as in [46].

The results presented above apply only in the limit $\frac{m}{J} \to \infty$. We have also investigated how these results are modified at very weak (rather than zero) coupling. We continue to work at low temperatures, in a manner we now describe in more detail. It turns that $S_{\text{eff}}(U)$ takes the schematic form

$$S_{\text{eff}}(U) = \sum_{a=1}^{\infty} x^a f_a(\beta, U). \tag{2.14}$$

Working to any given order in perturbation theory, the functions $f_a(\beta)$ are all polynomials of bounded degree in β. We work at temperatures low enough so that we can truncate

[14] We note that the first order phase transitions described in [45] were strongly first order (i.e. not on the edge between first and second order) only after turning on gauge interactions. In the current context, in contrast, the phase transition in our system is strongly first order even in the absence of interactions.

(2.14) to its first term. In other words the terms we keep are all proportional to x multiplied by a polynomial dressing in β.

We demonstrate below that within this approximation the partition function (2.14) takes the form

$$-S_{\text{eff}}(U) = N^{q-1}x \left(\prod_{m=1}^{q-1} \rho_m^1 \right) \left(\sum_{k=0}^{\infty} \left(\frac{J}{m} \right)^{2k} \widetilde{H}_k(\frac{J^2\beta}{m}) \right). \tag{2.15}$$

Note that (2.15) asserts that the interacting effective action has the same dependence on x and U as its free counterpart did. The only difference between the interacting and free effective action is a prefactor which is a function of the two effective couplings $\frac{J}{m}$ and $\frac{J^2\beta}{m}$. Below we have summed an infinite class of graphs and determined the function $\widetilde{H}_0$. Working at $N_F = 1$ we find

$$\widetilde{H}_0 = 2 \left[\frac{1}{2} + 2\gamma(q) \, \frac{(-\beta)}{m} |J|^2 e^{\gamma(q) \, \frac{(-\beta)}{m} |J|^2} - \frac{1}{2} e^{2\gamma(q) \, \frac{(-\beta)}{m} |J|^2} - \frac{(-1)^{q/2}}{2} q \, \beta \frac{|J|^2}{m} \right], \tag{2.16}$$

where $\gamma(q)$ is defined in (2.138).

(2.15) and (2.16) determine the effective action of our system whenever the terms proportional to $\widetilde{H}_m$ ($m = 1, 2 \ldots$) in the second line of (2.140) can be ignored compared to the term proportional to $\widetilde{H}_0$. This is always the case at weak enough coupling; the precise condition on the coupling when this is the case depends on the nature of the as yet unknown large argument behaviour of the functions $\widetilde{H}_m$.

The partition function that follows from the action (2.15) is identical to the free partition function described above under the replacement $\alpha \to \alpha \widetilde{H}_0$. It follows that the interacting partition function is essentially identical to the free one in the canonical ensemble. The β dependence of the effective value of α leads to some differences in the microcanonical ensemble that turn out not to impact the main qualitative conclusions of the analysis of the free theory. For instance the super Hagedorn growth of the entropy persists upon including the effects of interaction.

Note Added: 'We have recently become aware of the preprint [47] which overlaps with this chapter in multiple ways. We hope it will prove possible to combine the results of this chapter with the methods of [47] to better understand the new light modes discussed earlier in this introduction'.

2.2 Light thermal modes of the Tensor models

In this section we consider the Gurau-Witten-Klebanov-Tarnopolsky model (or Tensor model) at finite temperature. The Lagrangians for the specific theories we study was listed in (2.1). As we have explained in the introduction, this model has a new set of light modes

parameterized by $V(t)$, an arbitrary group element as a function of time, where V belongs to $O(N)^{q-1}$. In this section we will present an argument that suggests that the dynamics of these light modes is governed by a (quantum mechanical) sigma model on the group manifold. We will also present an estimate for the coupling constant of this sigma model.

That the dynamics of $V(t)$ should be governed by a sigma model is very plausible on general grounds. Recall that in the formal IR limit, $V(t)$ is an exact zero mode of dynamics. It follows that $V(t)$ picks up dynamics only because of corrections to extreme low energy dynamics. From the point of view of the low energy theory these corrections are UV effects, and so should lead to a local action for $V(t)$. The resultant action must be invariant under global shifts $V(t) \to V_0 V(t)$. We are interested in the term in the action that will dominate long time physics, i.e. the action with this property that has the smallest number of time derivatives. Baring a dynamical coincidence (that sets the coefficient of an apparently allowed term to zero) the action will be that of the sigma model.

In the rest of this section we will put some equations to these words. We would like to emphasize that the 'derivation' of the sigma model action presented in this section is intuitive rather than rigorous - and should be taken to be an argument that makes our result highly plausible rather than certain.

2.2.1 Classical effective action

In [5] the effective large N dynamics of the SYK model was recast as the classical dynamics of two effective fields; the Greens function $G(t)$ and the self energy $\Sigma(t)$. The action for Σ and G derived in [5] was given by

$$S = N^{q-1}\left(-\log Pf(\partial_t - \widetilde{\Sigma}) + \int dt_1\ dt_2\ \left[-\widetilde{\Sigma}(t_1,t_2)\widetilde{G}(t_2,t_1) - \frac{J^2}{q}\widetilde{G}^q(t_1,t_2)\right]\right). \quad (2.17)$$

The utility of the action (2.17) was twofold. First, the solutions to the equations of motion that follow from varying (2.17) are the saddle point that govern thermal physics of the SYK model. Second, an integral over the fluctuations in (2.17) also correctly captures the leading order (in $\frac{1}{N}$) correction to this saddle point result. In order to obtain these corrections, one simply integrates over the quadratic fluctuations about this saddle point. In particular the action (2.17) was used to determine the action for the lightest fluctuations about the saddle point (2.17), namely conformal diffeomorphism [5].

In this section we wish to imitate the analysis of [5] to determine the action for fluctuations of the new zero modes - associated with time dependent $O(N)^{q-1}$ rotations - described in the introduction. The action (2.17) is not sufficient for this purpose. As explained in the introduction, the low energy fluctuations we wish to study are obtained by acting on the saddle point Greens function with time dependent $O(N)^{q-1}$ rotations; however the fields G and Σ that appear in (2.17) have no indices and so cannot be rotated.

As the first step in our analysis we proceed to generalize the effective action (2.17) to an action whose variables are the matrices G_a^b and Σ_a^b. The indices a and and b are both

fundamental indices of the group $O(N)^{q-1}$. Our generalized action is given by

$$S = -\log Pf(D_0 - \widetilde{\Sigma}) + \int dt_1\ dt_2\ \left[-\widetilde{\Sigma}_a^{\ b}(t_1,t_2)\widetilde{G}_b^{\ a}(t_2,t_1) - \frac{|g|^2}{q}\widetilde{G}^q(t_1,t_2)\right]. \qquad (2.18)$$

In this action, the expression $\widetilde{G}^q$ is a product of q copies of $\widetilde{G}_b^a$ where all gauge indices are contracted in a manner we now describe. Recall that a and b are fundamental indices for the group $O(N)^{q-1}$. Each of these indices may be thought of as a collection of $q-1$ fundamental indices

$$a = (a_1 a_2 \dots a_{q-1}), \quad b = (b_1 b_2 \dots b_{q-1}),$$

where a_i and b_i are fundamental indices in the (i^{th} factor of) $O(N)$. In the contraction $\widetilde{G}^q$, a type indices are contracted with each other while b type indices are also contracted with each other - there is no cross contraction between a and b type indices. The structure of contractions is as follows; the a indices of precisely one of the $O(N)$ factors of the gauge group are contracted between any two (and every two) Gs and, simultaneously, the b indices of the same two $O(N)$ factors are also contracted between the same two $\widetilde{G}s$. [15]

As a quick check note that the total number of contraction of a (or b) indices, according to our rule, is the number of ways of choosing two objects from a group of q, or, $\frac{q(q-1)}{2}$. As each pair hit two indices, we see that the pairing rule described in this paragraph saturates the indices present q copies of $\widetilde{G}$ (there are a total of $q(q-1)$ a type indices).

The contraction structure described for a type indices in the previous paragraph is precisely the contraction structure for the interaction term ψ^q in the action (2.1).

We regard (2.18) as a phenomenological action with the following desirable properties. First it is manifestly invariant under global $O(N)^{q-1}$ transformations. Second if we make the substitutions $\widetilde{G}_b^a \to \widetilde{G}\delta_b^a$, $\widetilde{\Sigma}_b^a \to \widetilde{\Sigma}\delta_b^a$ into (2.18) we recover the action (2.17). It follows in particular that, if G and Σ denote the saddle point values of (2.17) then

$$G_b^a = \delta_b^a G, \quad \Sigma_b^a = \delta_b^a \Sigma, \qquad (2.19)$$

are saddle points of (2.18). This point can also be verified directly from the equations of motion that follow from varying (2.18), i.e.

$$\begin{aligned} G_a^{\ b}(t_1,t_2) =& ((D_0 - \Sigma)^{-1})_a^{\ b}(t_1,t_2), \\ \Sigma_a^{\ b}(t_1,t_2) =& |g|^2\ (G^{q-1})_a^{\ b}(t_1,t_2). \end{aligned} \qquad (2.20)$$

While (2.18) correctly reproduces finite temperature saddle point of the the model (2.1), it does not give us a weakly coupled description of arbitrary fluctuations about this saddle point. The fact that (2.18) has $N^{2(q-1)}$ fields makes the action very strongly coupled. The

[15] These rules have their origin in the generalized 'tetrahedronal' contraction structure described in the introduction. For values of q at which the basic interaction structure has an ambiguity, we make one choice; for instance we adopt the 'Round Robin' scheme to fix the ambiguities. As far as we can tell, none of our results depend on the details of the choice we make.

key assumption in this section - for which we will offer no detailed justification beyond its general plausibility - is that the action (2.18) can, however, be reliably used to obtain the effective action for the very special manifold of configurations described in the introduction, namely

$$\begin{aligned}\widetilde{G}_b{}^a(t_1,t_2) =& V_b{}^{b'}(t_1)G(t_1,t_2)V_{b'}{}^a(t_2),\\ \widetilde{\Sigma}_b{}^a(t_1,t_2) =& V_b{}^{b'}(t_1)\Sigma(t_1,t_2)V_{b'}{}^a(t_2),\end{aligned} \tag{2.21}$$

where the index free functions $G(t,t_2)$ and $V(t_1,t_2)$ are the solutions to the SYK gap equations and $V(t)$ is an arbitrary $O(N)^{q-1}$ group element. The RHS in (2.21) is the result of performing a time dependent $O(N)^{q-1}$ rotation on the saddle point solution (2.19).

The fact that we have only $(q-1)\frac{N^2}{2}$ fields ($V(t)$) on this manifold of solutions - at least formally makes the action restricted to this special manifold weakly coupled, as we will see below.

In the rest of this section we will use the action (2.18) to determine the effective action that controls the dynamics of the matrices $V(t)$ at leading order in the long wavelength limit.

2.2.2 Effective action

In order to study quadratic fluctuations about (2.19), we follow [5] to insert the expansion [16]

$$\begin{aligned}\widetilde{G}_a{}^b(t_1,t_2) =& G_a{}^b(t_1,t_2) + |G(t_1,t_2)|^{\frac{q-2}{2}} g_a{}^b(t_1,t_2),\\ \widetilde{\Sigma}_a{}^b(t_1,t_2) =& \Sigma_a{}^b(t_1,t_2) + |G(t_1,t_2)|^{\frac{2-q}{2}} \sigma_a{}^b(t_1,t_2),\end{aligned} \tag{2.22}$$

into (2.18) and work to quadratic order in $g_a{}^b(t_1,t_2)$ and $\sigma_a{}^b(t_1,t_2)$. Integrating out $\sigma_a{}^b(t_1,t_2)$ using the linear equations of motion, we find an effective action of the general structure

$$\begin{aligned}S(\widetilde{G},\widetilde{\Sigma}) = S(G,\Sigma) &+ \frac{1}{2}\int dt_1..dt_4\ g_a{}^b(t_1,t_2)\widetilde{K}^{-1}(t_1,t_2;t_3,t_4)g_b{}^a(t_3,t_4)\\ &- \frac{|g|^2}{q}\ \frac{q}{2}N^{\frac{1}{2}(q-1)(q-4)+1}\int dt_1\ dt_2\ g(t_1,t_2)g(t_1,t_2).\end{aligned} \tag{2.23}$$

The expression in the first line of (2.23) results from varying the first two terms in (2.18), while the second line is the variation of the $\widetilde{G}^q$ term in (2.18). This term denotes the a sum of different contraction of indices between the two gs

$$g(t_1,t_2)g(t_1,t_2) = \sum_{k=1}^{q-1} g^{c_1c_2...c_{k-1}a_kc_{k+1}....c_{q-1}}_{c_1c_2...c_{k-1}b_kc_{k+1}....c_{q-1}} g^{d_1d_2...d_{k-1}a_kd_{k+1}....d_{q-1}}_{d_1d_2...d_{k-1}b_kd_{k+1}....d_{q-1}}. \tag{2.24}$$

[16] Note that we have scaled G fluctuations and Σ fluctuations with factors that are inverses of each other ensures that our change of variables does not change the path integral measure. The scalings of fluctuations in (2.22) are chosen to ensure that the second line of (2.23) takes the schematic form gg rather than $gK'G$ where K' is an appropriate Kernel. We emphasize that the scaling factor $|G(t_1,t_2)|^{\pm\frac{q-2}{2}}$ in (2.22) represents the power of a function; no matrices are involved.

In the special case that the fluctuation fields g are taken to be of the form $g^a_b = \delta^a_b g$, the matrix contractions in (2.23) give appropriate powers of N, and (2.23) reduces to the effective action for g presented in [5].

It was demonstrated in [5] that

$$\widetilde{K}(t_1,t_2;t_3,t_4) = -|G(t_1,t_2)|^{\frac{q-2}{2}} G(t_1,t_3)G(t_2,t_4)|G(t_3,t_4)|^{\frac{q-2}{2}}. \tag{2.25}$$

In the long distance limit the Greens function can be expanded as

$$\begin{aligned} G =& G_c + \delta G + ..., \\ \delta G(t_1,t_2) \equiv& \, G_c(t_1,t_2)\ f_0(t_1,t_2), \end{aligned} \tag{2.26}$$

where G_c is the Greens function in the conformal limit and δG is the first correction to G_c in a derivative expansion. It follows that f_0 is an even function of the time difference, an approximate form of which is given in [5]. Plugging this expansion into (2.25) it follows that $\widehat{K}$ can be expanded as

$$\widetilde{K} = \widetilde{K}_c + \delta\widetilde{K} + ..., \tag{2.27}$$

where [5]

$$\delta\widetilde{K}(t_1,t_2;t_3,t_4) = \ \widetilde{K}_c(t_1,t_2;t_3,t_4)\left[\frac{q-2}{2}(f_0(t_1,t_2)+f_0(t_3,t_4)) + f_0(t_1,t_3) + f_0(t_2,t_4)\right]. \tag{2.28}$$

The first two contributions have their origin in the factors of $G^{\frac{q-2}{2}}$ in (2.25) and were called rung contributions in [5] (2.25). The remaining two contributions have their origin in the factors of G in (2.25) and were called rail contributions in [5]. We note that for rung contributions f_0 appears with either first two times or last two times of the kernel. On the other hand the two times in rail contributions are one from the first set and one from the second.

Our discussion so far has applied to general fluctuations about the saddle point, and has largely been a review of the general results of [5] with a few extra indices sprinkled in. In the rest of this subsection we now focus attention on the specific fluctuations of interest to us, namely those generated by the linearized form of (2.21) around conformal solution

$$(g_c)_a{}^b(t_1,t_2) = |G_c(t_1,t_2)|^{\frac{q-2}{2}} G_c(t_1,t_2)\left[H_a{}^b(t_1) - H_a{}^b(t_2)\right]. \tag{2.29}$$

Notice that the fluctuations (2.29) represent the change of the propagator under a time dependent $O(N)^{q-1}$ rotation. The form of (2.29) is similar in some respects to the variation of the propagator under diffeomorphisms, studied in [5], with one important difference; the factors of $H^b_a(t_1)$ and $H^b_a(t_2)$ appear with a relative negative sign in (2.29), whereas the infinitesimal diffeomorphism fields in the light fluctuations of [5] appeared with a relative positive sign in [5]. The fact that our fluctuations are 'anti-symmetric' rather than' symmetric' will play an important role below.

Specializing to this particular fluctuation, It can be shown (see Appendix A.1) that g_c *is an eigenfunction of* $\widetilde{K}_c^{-1}$ *with eigenvalue* $|J|^2$ more clearly

$$\int dt_3\ dt_4\ \widetilde{K_c}^{-1}(t_1,t_2;t_3,t_4)(g_c)_a{}^b(t_3,t_4) = |J|^2\ (g_c)_a{}^b(t_1,t_2). \tag{2.30}$$

It follows immediately from (2.30) that

$$\frac{1}{2}\ g_c\ \widetilde{K}_c^{-1}\ g_c = \frac{|g|^2}{q}\ g_c\ g_c. \tag{2.31}$$

Using this equation it may be verified that for the for the particular fluctuations under study- the second line of (2.23) simply cancels the part of the term in the first line obtained by replacing $\widetilde{K}$ with $\widetilde{K}_c$.

It follows that the action (2.23) evaluated on the modes (2.29) is nonzero only because K^{-1} differs from K_c^{-1}. Recall $K = K_c + \delta K$ (see (2.27)). Using $\delta K^{-1} = -K\delta K K^{-1}$ that the action for our special modes evaluates at quadratic order to

$$S_{\text{eff}} = -\frac{1}{2} g_c\ \widetilde{K}_c^{-1}\ \delta\widetilde{K}\ \widetilde{K}_c^{-1}\ g_c. \tag{2.32}$$

Using the fact that $\widetilde{K}^{-1}$ is Hermitian ([5]) and the eigenvalue equation (2.30), the action simplifies to

$$S_{\text{eff}} = -\frac{1}{2}|J|^4 \int dt_1..dt_4\ (g_c)_a{}^b(t_1,t_2)\ \delta\widetilde{K}(t_1,t_2;t_3,t_4)\ \ (g_c)_b{}^a(t_3,t_4). \tag{2.33}$$

Plugging the specific form of our fluctuations (2.29) into this expression we find [17]

$$S_{\text{eff}} = -\frac{1}{2}N^{q-2}\sum_{l=1}^{q-1}\sum_{(i,k)\ pair}(-1)^{i-k}\int dt_i\ dt_k\ (H_l)_a{}^b(t_i)(H_l)_b{}^a(t_k)L_{ik}(t_i,t_k), \tag{2.34}$$

where $i \in (1,2)$, $k \in (3,4)$ and

$$\begin{aligned} L_{ik}(t_i,t_k) &= \int A(t_1,..,t_4)\prod_{m\neq i, m\neq k} dt_m, \\ A(t_1,..t_4) &= |J|^4\ G_c(t_1,t_2)|G_c(t_1,t_2)|^{\frac{q-2}{2}}\delta\widetilde{K}(t_1,t_2;t_3,t_4)|G_c(t_3,t_4)|^{\frac{q-2}{2}}G_c(t_3,t_4). \end{aligned} \tag{2.35}$$

The expression (2.34) is not yet completely explicit, as L_{ik} in (2.35) is given in terms of δK which is given in terms of the first correction to the conformal propagator G_c which,

[17] Here factors of N comes from trace over other colour index δ-functions that multiply H_l of any colour.

in turn, is not explicitly known. Luckily δG can be eliminated from (2.34) as we now demonstrate. [18]

While we do not know the explicit form of the correction to the conformal two-point function $\delta G(t_1, t_2)$, we know that it satisfies the equation

$$\Sigma_c * \delta G + \delta\Sigma * G_c + s * G_c = 0. \tag{2.38}$$

This is simply the gap equation expanded around the conformal point. Here $s(t_1, t_2) = -\dfrac{\partial}{\partial t_1}\delta(t_1 - t_2)$ is a local differential operator.

In order to make the expression (2.34) explicit we first simplify the formulae (2.35) for L_{ij}. Plugging the expansion $G = G_c + \delta G$ into (2.25), and using properties of conformal solutions, it may be verified after some algebra that for odd $i - k$ [19]

$$L_{ik}(t_i, t_k) = 2\ \delta(t_i - t_k)\left[\frac{q-2}{2} G_c * \frac{\delta\Sigma}{q-1} + \Sigma_c * \delta G\right](t_i, t_k). \tag{2.39}$$

The fact that L_{ik} is proportional to a δ function establishes that the contribution of terms with odd $i - k$ to the action is local. (2.39) may be further simplified using the relation

$$\delta(t_i - t_k) G_c * \frac{\delta\Sigma}{q-1}(t_i, t_k) = \delta(t_i - t_k)\Sigma_c * \delta G(t_i, t_k), \tag{2.40}$$

and to give

$$L_{ik}(t_i, t_k) = q\delta(t_i - t_k)\Sigma_c * \delta G(t_i, t_k). \tag{2.41}$$

Multiplying δ-function on both sides of (2.38) and using (2.40), we find

$$L_{ik}(t_i, t_k) = -\delta(t_i - t_k) s * G_c(t_i, t_k) = \delta(t_i - t_k)\frac{\partial}{\partial t_i} G_c(t_i, t_k). \tag{2.42}$$

---

[18] Using the fact that g_c is an eigenfunction of $\widetilde{K}_c$ with eigenvalue $\dfrac{1}{|J|^2}$ rung contributions can easily be summed up to

$$S_{\text{eff}}^{\text{rung}} = -\frac{1}{2}(q-2)\frac{1}{|J^2|}\int (g_c)_a{}^b(t_1, t_2)\ f_0(t_1, t_2)\ (g_c)_b{}^a(t_1, t_2)\ dt_1 dt_2. \tag{2.36}$$

This expression is not by itself useful as the integral that appears in it has a log divergence once numerically determined form of $f_0(\tau_1, \tau_2) \xrightarrow[|\tau_1 - \tau_2| \to 0]{} \dfrac{1}{|\tau_1 - \tau_2|}$ (from [5]) is used; follows from

$$g_c(\tau_1, \tau_2) = |G_c(\tau_1, \tau_2)|^{\frac{q-2}{2}} G_c(\tau_1, \tau_2)[H(\tau_1) - H(\tau_2)] \xrightarrow[|\tau_1 - \tau_2| \to 0]{} \frac{\text{sgn}(\tau_1 - \tau_2)}{|\tau_1 - \tau_2|} H'(\tau_1)(\tau_1 - \tau_2) \sim O(|\tau_1 - \tau_2|^0). \tag{2.37}$$

[19] Here overall factor of 2 comes from symmetry of the integrations and $\dfrac{q-2}{2}$ comes from rung part.

On the other hand when $i-k$ is even, using properties of conformal solutions [20]

$$L_{ik}(t_i,t_k) = -\left[\frac{q-2}{2}\times 2+1\right]\Sigma_c(t_i,t_k)\delta G(t_i,t_k) + (\Sigma_c * \delta G * \Sigma_c)(t_i,t_k)G_c(t_i,t_k). \quad (2.43)$$

(2.43) can be further simplified by substituting

$$\Sigma_c * \delta G * \Sigma_c = \delta\Sigma + s, \quad (2.44)$$

and then using the linearized form of the gap equation

$$\delta\Sigma\ G_c = (q-1)\delta G\ \Sigma_c, \quad (2.45)$$

to give

$$L_{ik}(t_i,t_k) = -G_c(t_i,t_k)\frac{\partial}{\partial t_i}\delta(t_i-t_k). \quad (2.46)$$

Adding together the contributions of $i-k$ even and $i-k$ odd we have a manifestly local effective action, whose structure accounts for the fact that we have worked beyond the purely conformal limit (recall that in the purely conformal limit our fluctuation action simply vanished) even though the final expression makes no reference to the explicit form of the correction δG to the conformal propagator G_c.

$$S_{\text{eff}} = -\,N^{q-2}\sum_{l=1}^{q-1}\int dt_i\ dt_k G_c(t_i-t_k)\delta(t_i-t_k)\text{Tr}\ \left(\frac{\partial}{\partial t_i}H_l(t_i)H_l(t_k)\right) \quad (2.47)$$

Expanding $H_l(t_k)$ in a Taylor series expansion about t_i

$$H_l(t_k) = \sum_{n=0}^{\infty}\frac{\partial^n}{\partial t^n}H_l(t_i)\frac{(t_k-t_i)^n}{n!}$$

allows us to recast (2.47) into the form

$$S_{\text{eff}} = -N^{q-2}\int dt\ \sum_{l=1}^{q-1}\sum_{n=0}^{\infty}C_n\ Tr\ \left(\frac{\partial}{\partial t}H_l(t)\ \frac{\partial^n}{\partial t^n}H_l(t)\right). \quad (2.48)$$

where

$$C_n = \frac{1}{n!}\int dt\ G_c(t)\delta(t)t^n. \quad (2.49)$$

The term in the sum (2.48) with $n=0$ is a total derivative and so can be ignored. It follows that

$$S_{\text{eff}} = -\int dt\ \sum_{l=1}^{q-1}\sum_{n=1}^{\infty}C_n\text{Tr}\ \left(\frac{\partial}{\partial t}H_l(t)\ \frac{\partial^n}{\partial t^n}H_l(t)\right). \quad (2.50)$$

[20] As before $\frac{q-2}{2}\times 2$ comes from rung part.

Our final result (2.50) for the effective action, has now been arranged as an expansion over terms with increasing numbers of derivatives.

Recall that all the results of this section have been obtained after expanding the Greens function

$$G(t_1, t_2) = G_c(t_1, t_2) + \delta G(t_1, t_2), \tag{2.51}$$

and assumed that $\delta G \ll G_c$. This assumption is only valid when $t_1 - t_2 \gg \frac{1}{J}$, but are not valid for $t_1 - t_2 \sim \frac{1}{J}$. All potential non localities in the effective action for H presumably have their origin in regions where our approximations are valid. It thus seems plausible that the central result of this section - namely the absence of non-localities in the effective action on length scales large compared to $\frac{1}{J}$ - which therefore takes the form (2.50) - is a reliable result.

On the other hand the precise expressions for the coefficient functions C_n involve integrals over a function - namely the delta function - which varies over arbitrarily small distances - and so is not reliable (it uses our approximations in a regime where they are not valid). We would expect the correct versions of (2.49) to be given by smeared out versions of the integrals in (2.49). On general dimensional grounds it follows that

$$C_n \to \frac{A_n}{|J|^n}. \tag{2.52}$$

We will make the replacement (2.52) in what follows. The numbers A_n could presumably be computed by studying four point correlators of appropriate operators at finite temperature. We will not attempt this exercise in this chapter.

For the purposes of long time physics we are interested only in the term with the leading number of derivatives, i.e. with the term with $n = 1$ in (2.50). The coefficient of our action in this case is proportional to $A_1 \equiv \mathcal{A}$. [21] and the effective action of our theory at leading order in the derivative expansion takes the form

$$S = -\mathcal{A}\, \frac{N^{q-2}}{|J|} \int dt\, \sum_{l=1}^{q-1} \text{Tr}\, \left(\frac{\partial}{\partial t} H_l(t)\, \frac{\partial}{\partial t} H_l(t) \right). \tag{2.56}$$

[21] Note that

$$C_1 = \int dt\, \delta(t) G_c(t)\, t. \tag{2.53}$$

Plugging the formula

$$G_c = b \frac{\text{sgn}(t)}{|Jt|^{\frac{2}{q}}}, \tag{2.54}$$

into (2.53) we find, formally, that

$$C_1 \propto \int dt |t|^{1-\frac{2}{q}} \delta(t) = 0, \tag{2.55}$$

(where we have used the fact that $q > 2$). As explained above, we expect that the vanishing of C_1 is not a physical result but rather is a consequence of inappropriate use of approximations. We assume that $C_1 \to \frac{\mathcal{A}}{|J|}$ in what follows where $\mathcal{A}$ is an unknown dimensionless number.

In the analysis presented so far we have determined the form of the effective action for infinitesimal group rotations H. The group invariant extension of our result to finite group rotations is the sigma model action

$$S = -\mathcal{A}\,\frac{1}{|J|}\int dt\;\;\sum_{l=1}^{q-1}\mathrm{Tr}\;\left[\left(V_l^{-1}(t)\;\frac{\partial}{\partial t}V_l(t)\right)^2\right], \tag{2.57}$$

where $V_l \in SU(N)$ whose infinitesimal form is $V_l = 1 + H_l + \mathcal{O}(H_l^2)$. (2.45) is simply the action for a free particle moving on the group manifold $O(N)^{q-1}$ [22]. As explained in the introduction, the structure of this action could have been anticipated on general grounds. The fact that the action is proportional to $\frac{1}{J}$ follows largely on grounds of dimensional analysis.

As we have already seen in the introduction, once we have established that the action for $V(t)$ is local the form of the low energy effective action (2.3) for our system is almost inevitable using the general principles of effective field theory. The main accomplishment of the algebra presented in this section is the demonstration that the effective action for $V(t)$ is, indeed, local.

Note that the Sigma model action (2.45) has an $O(N)^{q-1} \times O(N)^{q-1}$ global symmetry under which

$$V_l \to A V_l B, \tag{2.58}$$

where A and B both belong to $O(N)^{q-1}$. The rotations by A are simply the global symmetry that the microscopic SYK model possesses. Rotations by B are an emergent symmetry of the low energy effective action. The corresponding conserved quantities are $L_l = \dot{V}_l V_l^{-1}$, and $R_l = V_l^{-1}\dot{V}_l$ [23]. Choosing a basis (T_a) [24] of Lie algebra $\mathcal{O}(N)$ it can be shown that Hamiltonian vector fields corresponding to group functions $L_{l,a} = Tr\;(T_a L_l)$, $R_{l,a} = Tr\;(T_a R_l)$ give two copies of $\mathcal{O}(N)$ (at both classical and quantum level), both of which commutes with the Hamiltonian which is the quadratic Casimir of the algebra.

2.3 Holonomy dynamics and density of states at large mass

We now switch gears; in this section and next we discuss a the mass deformed SYK theory (2.7) in the large mass limit. We work with the theory based on the $O(N)^{q-1}$ symmetry where this symmetry is gauged. The large mass limit is of interest because it allows us to focus on the dynamics of the holonomy at finite temperature, and also allows us to compute the growth of states in the theory as a function of energy in a very simple setting.

[22] Non-trivial holonomy can be turned on for these new light modes, details of contribution of these light modes to effective action for holonomy is presented in Appendix A.3.

[23] A dot over a quantity indicates derivative with respect to time.

[24] It is assumed in what follows that this basis puts the Killing form in a form proportional to identity.

2.3.1 Scaling limit

As explained in the introduction, in this section we will compute the finite temperature partition function

$$Z = \mathrm{Tr}\ x^{\frac{H}{m}},$$

for the mass deformed gauged $O(N)^{q-1}$ melonic theory (2.7).

In the large mass limit all fields in (2.7) except the holonomies of the gauge group can be integrated out at quadratic order. The result of this integration is easily obtained using the formulae of [42], and is given by (2.8).

Notice that the effective action $S_{\text{eff}}(U_i)$ presented in (2.8) is invariant under the global 'gauge transformations' $U_i \to V_i U_i V_i^{-1}$ for arbitrary orthogonal matrices V_i. This invariance may be used to diagonalize each U_i. The integral in (2.8) may then be recast as an integral over the eigenvalues of each of the holonomy matrices U_i with the appropriate measure. As U_m are each unitary, their eigenvalues take the form $e^{i\theta_m^n}$ where n runs from 1 to N. We define the eigen value density functions

$$\rho_m(\theta) = \frac{1}{N}\sum_{n=1}^{N}\delta(\theta - \theta_m^n). \tag{2.59}$$

As we are dealing with orthogonal matrices, the eigenvalues of our matrix occurs in equal and opposite pairs $(\theta_a, -\theta_a)$ and so the eigenvalue density function defined in (2.59) is an even function.

As usual the rather singular looking sum over delta functions in (2.59) morphs into an effectively smooth function at large N as the individual eigenvalues merge into a continuum. Note that

$$\frac{\mathrm{Tr}U_m^n}{N} = \frac{\sum_{j=1}^{N} e^{in\theta_m^j}}{N} = \int \rho_i(\theta)e^{in\theta} \equiv \rho_m^n, \tag{2.60}$$

where the last equality defines the symbol ρ_i^n. Note that the subscript m on ρ runs from $1 \ldots q-1$ and labels the $O(N)$ factor under study, while the superscript n runs from $1 \ldots \infty$ and labels the Fourier mode of the eigenvalue distribution. Using the fact that $\rho_i(\theta) = \rho_i(-\theta)$ it follows that

$$\rho_i^n = \int d\theta \rho(\theta)\cos n\theta. \tag{2.61}$$

It follows that ρ_i^n are all real numbers and that $\rho_i^n = \rho_i^{-n}$.

In the large N limit the integral over the eigenvalues θ_m^n may be recast, in the large N limit into a path integral over the eigenvalue functions $\rho_m(\theta)$ given by

$$Z(x) = \int \prod_{i=1}^{q-1} D\rho_i \exp\left[\frac{1}{2}\sum_{n=1}^{\infty}\left(-N^2\sum_{m=1}^{q-1}\frac{|\rho_m^n|^2}{n} - 2N_F N^{q-1}(-x)^n\frac{\left(\prod_{m=1}^{q-1}\rho_m^n\right)}{n}\right)\right], \tag{2.62}$$

[25] where the path integral is now taken over the eigenvalue density functions ρ_m with a measure which descends from the flat integration measure for individual eigenvalues θ_m^j. As we have only $(q-1)N$ eigenvalues, the Jacobian of this variable change is of order N in the exponent and so is subleading at large N and will not concern us.

Notice that the effective action in (2.62) is a sum of two kinds of terms; those proportional to N^2 (we call these terms the contribution of the measure) and those proportional to N^{q-1} (we call these terms the contribution of the energy). As $q \geq 4$ the energy overwhelms the measure at large N if x is taken to be of order unity. In order to work in a regime in which the measure and the energy compete with each other we define

$$x = \frac{\alpha}{pN^{q-3}}, \tag{2.63}$$

where [26]

$$p = N_F,$$

and take the limit $N \to \infty$ with α held fixed. In this limit the 'energy' term with $n = 1$ in (2.62) is of order N^2 and so competes with the measure. All energy terms with $n > 1$ are, however, subleading compared to the measure and can be dropped at large N. In the limit under consideration, in other words, the effective action in (2.62) simplifies to

$$\begin{aligned} Z(\alpha) &= \int \prod_{i=1}^{q-1} dU_i \exp(-S_{\text{eff}}(U_i)), \\ S_{\text{eff}} &= -\frac{\alpha}{N^{q-3}} \left(\prod_{i=1}^{q-1} \text{Tr} U_i \right). \end{aligned} \tag{2.64}$$

We will now evaluate the integral (2.62) at large N with the effective action replaced by the simplified effective action (2.64). In order to facilitate comparison with the matrix model literature, it is useful to note that the matrix integral (2.64) is closely related to the

[25] Let us focus on the special case $N_F = 1$. In this case the Hilbert space of our quantum mechanical problem is simply the sum of q forms of the group $O(N^3)$ with q running from 1 to N^3. The exponential in (2.62) is the character of this Hilbert space w.r.t the subgroup $O(N)^3$, with representations coming from q forms in $O(N^3)$ graded by x^q. (In order to view the exponential as a character one must use (2.60)). The integral in (2.62) projects onto the singlet subspace, and so counts the number of $O(N)^3$ singlets. Note that it was very important for this discussions that the fundamental fermions in this chapter are complex. The case of real fermions was studied from this point of view in [48]. In this case the Hilbert space of the $N_F = 1$ theory consists of spinors of $O(N^3)$, and the decomposition of this representation content into representations of $O(N)^3$ appears to be a very different problem; it was suggested in [48] that this decomposition contains no singlets. We thank C. Krishnan for discusssions on this point.

[26] As explained in the introduction, in the free limit we could as well study bosons coupled to the gauge field in which case we would have $p = N_B + N_F$ where N_B is the number of bosons.

following integral over unitary matrices

$$\begin{aligned} Z_{SU}(\alpha) &= \int \prod_{i=1}^{q-1} dU_i \exp(-S_{\text{eff}}(U_i)), \\ S_{\text{eff}} &= -\frac{\alpha}{N^{q-3}} \left(\prod_{i=1}^{q-1} \text{Tr} U_i + \prod_{i=1}^{q-1} \text{Tr} U_i^\dagger \right). \end{aligned} \tag{2.65}$$

Where the integral is now taken over unitary matrices. In the large N limit the two matrix models have the same gap equation (see below) and

$$\log Z_{SU}(\alpha) = 2 \log Z(\alpha). \tag{2.66}$$

2.3.2 Determination of saddle points

The matrix model (2.65) (and so (2.64)) is easily solved in the large N limit using the usual saddle point method. In order to see how this can be done note that as far as the integral over the eigenvalues of U_1 are concerned, $\text{Tr}U_2$, $\text{Tr}U_3$ $\ldots$ $\text{Tr}U_{q-1}$ are all constants. Focusing only on the integral over U_1, (2.64) reduces to

$$\begin{aligned} Z_{SU} &= \int dU_1 \exp\left(\frac{N}{g_1} \left(\text{Tr} U_1 + \text{Tr} U_1^\dagger \right) \right), \\ \frac{1}{g_1} &= \alpha \rho_2^1 \rho_3^1 \ldots \rho_{q-1}^1 = \alpha u_2 u_3 \ldots u_{q-1}, \end{aligned} \tag{2.67}$$

where in order to lighten the notation we have defined

$$\rho_m^1 = u_m \tag{2.68}$$

A similar statement applies to the integral over all U_i for $i = 1 \ldots q-1$. However (2.67) is precisely the celebrated Gross Witten Wadia matrix integral [49, 50, 51]. Recall that the saddle point that dominates the integral (2.67) (and its counterparts for U_2 etc) is given by [49, 50, 51]

$$\rho_m(\theta) = \begin{cases} \dfrac{1}{2\pi} \left[1 + \dfrac{2}{g_m} \cos\theta \right], & g_m \geq 2, \quad |\theta| \leq \pi \\ \dfrac{2}{\pi g_m} \cos\dfrac{\theta}{2} \sqrt{\dfrac{g_m}{2} - \sin^2 \dfrac{\theta}{2}}, & g_m < 2, \quad |\theta| < 2 \sin^{-1} \left(\dfrac{g_m}{2} \right)^{1/2}, \end{cases} \tag{2.69}$$

where[27]

$$\frac{1}{g_m} = \alpha \prod_{j \neq m} u_j. \tag{2.70}$$

[27] This eigenvalue densities produced above solve the GWW saddle point equations

$$\frac{2N}{g_m} \sin\theta_m^n = \sum_{j \neq n} \cot\left(\frac{\theta_m^j - \theta_m^n}{2} \right),$$

in the large N limit.

Taking the Fourier transform of (2.69) it follows that

$$u_m = \begin{cases} \frac{1}{g_m}, & g_m \geq 2 \\ 1 - \frac{g_m}{4}, & g_m < 2. \end{cases} \tag{2.71}$$

We refer to the solution $u_m = \frac{1}{g_m}$ as the wavy phase while the solution $u_m = 1 - \frac{g_m}{4}$ as the gapped phase.

(2.70) and (2.71) may be regarded as a set of $2(q-1)$ equations for the $2(q-1)$ variables u_m and g_m. In order to complete the evaluation of our matrix integrals we will now solve these equations.

Let us first demonstrate that the variables g_m are either all greater than 2 or all less than two simultaneously; (2.70) and (2.71) admit no solutions in which some of the g_m are greater than 2 while others are less than 2. [28]

Let us assume that $g_m \geq 2$. It follows from (2.70) and (2.71) that

$$\alpha u_1 u_2 \dots u_{q-1} = \frac{u_m}{g_m} = \frac{1}{g_m^2} \leq \frac{1}{4}. \tag{2.72}$$

On the other hand let us suppose that $g_k < 2$ Then it follows from (2.70) and (2.71) that

$$\alpha u_1 u_2 \dots u_{q-1} = \frac{u_k}{g_k} = \frac{1}{g_k} - \frac{1}{4} > \frac{1}{4}. \tag{2.73}$$

As (2.72) and (2.73) contradict each other it follows that either all $g_m \geq 2$ or all $g_m < 2$ as we wanted to show. Moreover it follows immediately from (2.73) that when all $g_m \leq 2$ they are in fact all equal. Similarly it follows from (2.72) that when all $g_m \geq 2$ then once again they are all equal. [29] It follows that in either case all u_m and all g_m are equal. Let us refer to the common saddle point value of u_m as u. The saddle point equations (2.71) now simplify to

$$u = \begin{cases} \alpha u^{q-2} & u \leq \frac{1}{2} \\ 1 - \frac{1}{4\alpha u^{q-2}}, & u > \frac{1}{2}. \end{cases} \tag{2.74}$$

Once we have determined the solution to (2.74) value of the partition function (2.64), in

[28] Equivalently u_ms are either all less than half or all greater than half. Equivalently the matrix models for U_m are all simultaneously in the wavy phase or simultaneously in the gapped phase.

[29] Actually all solutions are equal up to sign - however saddle points that differ by sign assignments are actually essentially identical - they can be mapped to each other by $U \to -U$, so we ignore this issue.

the large N limit under consideration, is given by

$$\begin{aligned} Z(\alpha) &= \exp\left(-\frac{N^2}{2}V(u)\right), \\ V(u) &= (q-1)f(u) - 2\alpha\ u^{q-1}, \\ f(u) &= \begin{cases} u^2, & u \leq \frac{1}{2} \\ \frac{1}{4} - \frac{1}{2}\log\left[2(1-u)\right], & u > \frac{1}{2}. \end{cases} \end{aligned} \tag{2.75}$$

[30] Indeed the saddle point equation (2.74) is simply the condition that the 'potential' $V(u)$ in (2.75) is extremist. In other words the saddle point solutions of our matrix integral are in one to one correspondence with the saddle points (or extrema) of $V(u)$; the contribution of each saddle point to the matrix integral is simply given by $e^{-N^2\frac{V(u)}{2}}$.

At every positive value of α, $V(u) = 0$ when $u = 0$ and $V(u)$ diverges as u approaches unity from below. [31] However the qualitative behaviour of the function $V(u)$ for values between zero and unity depends sensitively on α.

It is easily verified that for $\alpha \leq \alpha_c = \frac{(q-1)^{q-1}}{4(q-2)^{q-2}}$ the function $V(u)$ increases monotonically as u increases from 0 to unity (see Fig 2.1 (a)). It follows that when $\alpha \leq \alpha_c$ the only saddle point lies at $u = 0$. In this case the saddle point value of the partition function is $Z(x) = 1$ (see below for a discussion of fluctuations about this saddle point value).

At $\alpha = \alpha_c$ the potential $V(u)$ develops a point of inflection at $u = u_c = \frac{q-2}{q-1}$ (see Fig. 2.1 (b)). Note that $u_c > \frac{1}{2}$. At this value of α we have a new saddle point in the gapped phase.

As α is increased above α_c the point of inflection at $u = u_c$ splits up into two saddle points; a local maximum at $u = u_{max} < u_c$ and a local minimum at $u = u_{min} > u_c$ (see Fig. 2.1 (c)). To start with both saddle points are in the gapped phase. We refer to the saddle point at u_{max} as the upper saddle and the saddle point at u_{min} as the lower saddle.

As α is increased further the value of u_{min} continues to increase while the value of u_{max} continues to decrease. At $\alpha = \alpha_{pt} = 2^{q-3} > \alpha_c$, $u_{max} = \frac{1}{2}$. For $\alpha > \alpha_{pt}$, $u_{max} < \frac{1}{2}$ and the upper saddle makes a Gross Witten Wadia phase transition into the wavy phase (see Fig 2.1 (d)). [32]

Finally, when the new saddle point at $u = u_c$ is first nucleated, we have $V(u_c) > 0$. As α is increased $V(u_{min})$ decreases below this value. At $\alpha = \alpha_{1pt} =$ we have $V(u_{min}) = 0$

[30] The factor of $\frac{1}{2}$ in the exponent of the first equation in (2.64) is a consequence of the fact that we are working with the orthogonal model. The analogous formula for the partition function of the unitary model, (2.65), is the square of the partition function listed here and so does not have the factor of half in the exponential.

[31] Note that $u = \frac{\mathrm{Tr}U}{N} \leq 1$.

[32] The formula for u_{max}, u_{min} as a function of α is complicated in general. However the formula simplifies

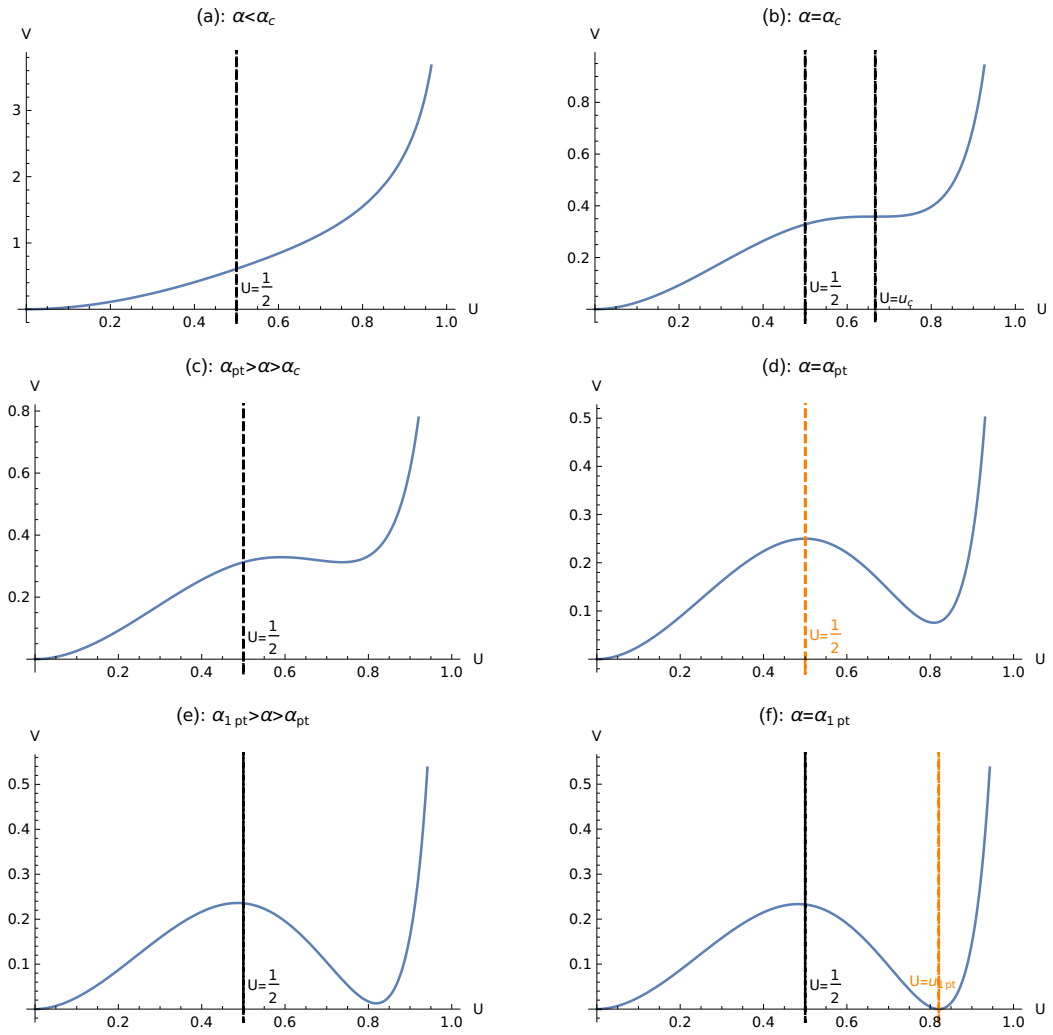


Figure 2.1: Effective potential for different values of temperature and associated phase transitions. The graphs are drawn for $q = 4$.

(see Fig 2.1(f)). For larger values of α, $V(u_{min}) < 0$ and our matrix model undergoes a first order phase transition from the saddle at $u = 0$ to the saddle at $u = u_{min}$. Note that at $\alpha = \alpha_{1pt}$ (i.e. at the 'Hawking Page transition temperature') the saddle at $u = u_{max}$ is already in the the wavy phase when $q = 4$ but is still in the gapped phase for $q > 4$.

at large α and we find

$$u_{max} = \left(\frac{1}{\alpha}\right)^{\frac{1}{q-3}} \ , \ u_{min} = 1 + \frac{1}{-4\alpha + q - 2} + \frac{q^2 - 3q + 2}{2(-4\alpha + q - 2)^3} + \dots \tag{2.76}$$

2.3.3 Thermodynamics in the canonical ensemble

The thermodynamics of our system in the canonical ensemble follows immediately from the nature of the function $V(u)$ as a function of α described at the end of the last section. For convenience we discuss the phase diagram of our system as a function of α rather than temperature (recall that α is defined by the relations $e^{-\beta m} = x = \frac{\alpha}{pN^{q-3}}$).

For $\alpha < \alpha_c$ the saddle at $u = 0$ is the only saddle point in the theory (see Fig 2.1 (a)). For $\alpha_c < \alpha < \alpha_{pt}$ [33] there are two additional saddle points at $u = u_{min}$ and $u = u_{max}$ with $\frac{1}{2} < u_{max} < u_{min} < 1$. The saddle point at $u = u_{max}$ is a local maximum and $V(u_{max}) > 0$ (see Fig 2.1 (c)). The saddle point at $u = u_{min}$ is a local minimum and however $V(u_{min}) > 0$. Both these saddles are subdominant compared to the flat saddle in this range of α.

For $\alpha_{pt} < \alpha < \alpha_{1pt}$ the two new phases continue to be subdominant compared to the phase at $u = 0$; in this range, however, the solution at $u = u_{max} < \frac{1}{2}$ is now in the wavy phase (see Fig 2.1 (e)).

At $\alpha = \alpha_{1pt}$ we have $V(u_{min}) = 0$. For $\alpha > \alpha_{1pt}$ $V(u_{min}) < 0$, so the solution at $u = u_{min}$ is the dominant saddle point. Our system undergoes a phase transition at $\alpha = \alpha_{1pt}$ (see Fig 2.1 (e)). The value of α_{1pt} is given as a function of q by

$$\alpha_{1pt} = \frac{1}{4}(q-1)w\left[1 - \frac{1}{(q-1)w}\right]^{-(q-2)}, \quad w = -W_{-1}\left[-\frac{2\exp\left[-\frac{(q+1)}{2(q-1)}\right]}{q-1}\right], \tag{2.77}$$

where W_n is the productlog function.

2.3.4 Thermodynamics in the microcanonical ensemble

In this subsection we compute the density of states as a function of energy corresponding to each of the saddle points described in the previous subsection. In order to do this we use the thermodynamical relations

$$E(\alpha) = \alpha\partial_\alpha \log Z(\alpha) \quad S(\alpha) = \left(\log Z(\alpha) - E(\alpha)\log\frac{\alpha}{N^{q-3}p}\right), \tag{2.78}$$

where E is the eigenvalue of $\frac{H}{m}$. We invert the first of these equations to solve for $\alpha(E)$, and then plug this solution into the second equation to obtain $S = S(E)$. For the trivial saddle, the saddle value of $S(E)$ is trivial, so we include the contribution of fluctuations around this saddle.

[33] In the text of this paragraph and the next we have assumed that $\alpha_{pt} < \alpha_{1pt}$ as is the case for $q = 4$. For $q \geq 6$ the order above is reversed, and the discussion has obvious modifications.

The saddle at $u = 0$

The saddle point at $u = 0$ exists at every value of α. In this case the saddle point values of the energy and entropy both vanish so the first nontrivial contribution to the thermodynamics comes from the study of fluctuations about the saddle point. In this subsection - which is a bit of a deviation from the main flow of the (otherwise purely saddle point) computations of this chapter we describe the relevant computations. For the purposes of this subsection - and this subsection only - we retreat away from the scaling limit (2.10) and work with the full matrix model (2.8) - or more precisely with its generalization (2.9) which allows for bosonic as well as fermionic harmonic oscillators. Working with this generalized model we compute the fluctuations around the trivial saddle point $\mathrm{Tr}U_m^n = 0$, i.e. $\rho_m^n = 0$.

For the purposes of studying small fluctuations around this saddle point we work with the integral (2.62). The integral (2.62) can be simplified by making the variable change

$$\rho_m^n = \frac{\beta_m^n}{N} \tag{2.79}$$

The point of the scaling (2.79) is that it eliminates all explicit factors of N from the integral (2.62). It follows that - at least for the purposes of the perturbative Wick contraction evaluations we perform in this subsection - at any finite order in perturbation theory the integral over β_n^m receives significant contributions only from values of β_n^m of order unity. Note however that if β_n^m are of order unity then ρ_n^m are of order $\frac{1}{N}$ and so are very small. We can thus safely integrate over all values of β_n^m without worrying about boundaries to the domain of integration. [34] In other words (2.62) may be rewritten in terms of these scaled variables

$$\begin{aligned} Z(x) &= \prod_{n=1}^{\infty} F_n(x), \\ F_n(x) &= \begin{cases} M_n \int \prod_{m=1}^{q-1} d\beta_m^n \exp\left(-\sum_{m=1}^{q-1} \frac{|\beta_m^n|^2}{2n} + N_F x^n \frac{\left(\prod_{m=1}^{q-1} \beta_m^n\right) + \text{c.c}}{n} \right) & n \text{ odd} \\ M_n \int \prod_{m=1}^{q-1} d\beta_m^n \exp\left(-\sum_{m=1}^{q-1} \frac{|\beta_m^n|^2}{2n} - N_F x^n \frac{\left(\prod_{m=1}^{q-1} \beta_m^n\right) + \text{c.c}}{n} \right) & n \text{ even.} \end{cases} \end{aligned} \tag{2.80}$$

The expressions for F_n above involve an integral with the usual measure $dzd\bar{z}$ for the complex variable β_m^n. The integral is taken over the whole complex plane[35]. The x independent normalization constant M_n above are chosen to ensure that normalization of Haar measure,i.e, $F_n(0) = 1$.

[34] More generally the variables ρ_n^m are constrained by the requirement that the function $\rho_m(\theta) = \frac{1}{2\pi} \sum_n \rho_m^n e^{-in\theta}$, is positive for every value of θ. This constraint is trivial when all ρ_m^n as is effectively the case for the perturbative evaluations discussed above.

[35] As mentioned above, the difference between this measure and $d\theta_m^j$ is sub-dominant in large N limit.

The expressions for F_n presented in (2.80) are formal as the integrals that define F_n do not converge. However this fact does not bother us, as we are not really interested in the the expression for $Z(x)$ but only in the coefficients in of x^k for each k in that expression. Each of these coefficients is easily obtained (by Taylor expanding the non Gaussian terms in the integrands in the formulas for F_n above and performing all integrals using Wicks theorem. We find

$$F_n(x) = \sum_{k=0}^{\infty} x^{2kn} \left(p^2 (2n)^{q-3}\right)^k (k!)^{q-3}, \quad p \equiv N_F, \tag{2.81}$$

Let E denote the eigenvalues of $\frac{H}{m}$; in other words E is the energy of our theory in units of the oscillator mass (or frequency). It follows from (2.81) that the functions $F_n(x)$ represent the partition function of a system whose entropy as a function of energy is given by $S_n(E)$ where

$$e^{S_n(E)} = \left(\frac{E}{2n}!\right)^{q-3} \left(p^2 (2n)^{q-3}\right)^{\frac{E}{2n}}. \tag{2.82}$$

At large E (i.e. when $E \gg 2n$) we may use Sterling's approximation to simplify (2.82) to obtain the asymptotic formula

$$S_n(E) = (q-3)\frac{E}{2n} \log\left(\frac{E}{2n}\right) + \frac{E}{2n}\left(-(q-3) + 2\log p + (q-3)\log(2n)\right). \tag{2.83}$$

Notice that the density of states grows faster than exponentially as a function of energy, explaining the divergence of the integrals that define F_n (or, equivalently, explaining why the sums in (2.81) are divergent at every x no matter how small.

As the partition function of our system is simply the product over the functions F_n, the entropy of our system at large energies is obtained by distributing the available energy E among the various systems S_n in such a way as to maximize the entropy. A glance at (2.83) is sufficient to convince oneself that the best one can do is to put all available energy into the 'system' S_1. It follows that for $E \gg 1$, the contribution of the saddle point at $u_1 = 0$ to the entropy of the system is

$$S(E) = S_1(E) = (q-3)\frac{E}{2} \log\left(\frac{E}{2}\right) + \frac{E}{2}\left((\log(2) - 1)(q-3) + 2\log p\right). \tag{2.84}$$

The saddle at $u = 0$ is exceptional in that it is trivial as a saddle point; in order to determine the thermodynamics of this 'phase' we had to perform the one loop expansion about this saddle point. The remaining saddle points we will study in this section are nontrivial even at leading order, and so will be analyzed only within the strict saddle point approximation. In the rest of this subsection we also return to the study of the strict scaling limit (2.10).

The wavy phase

In this subsection we study the thermodynamics of the wavy saddle, i.e. the saddle point at $u = u_{max}$ for $\alpha > \alpha_{pt} = 2^{q-3}$. The contribution of this saddle point to partition function is

$$\log Z(\alpha) = -\frac{N^2}{2}(q-3)\alpha^{-\frac{2}{q-3}}. \tag{2.85}$$

The energy of the corresponding phase is given by

$$E(\alpha) = \alpha\partial_\alpha \log Z(\alpha) = \frac{N^2}{\alpha^{\frac{2}{q-3}}}, \tag{2.86}$$

Note that the energy is a decreasing function of α so that this phase has a negative specific heat. As this phase exists only for $\alpha > \alpha_{pt}$ it follows that the energy in this phase is bounded from above by

$$E_{pt} \equiv E(\alpha_{pt}) = \frac{N^2}{4}. \tag{2.87}$$

The entropy of this phase is given by

$$S(\alpha) = \left(\log Z(\alpha) - E(\alpha)\log\frac{\alpha}{N^{q-3}p}\right). \tag{2.88}$$

Eliminating α between (2.86) and (2.88) we obtain

$$S(E) = (q-3)\left[\frac{E}{2}\log\left(\frac{E}{2}\right) - \frac{E}{2}\right] + E\ \log p + (q-3)\frac{E}{2}\log(2). \tag{2.89}$$

Note that (2.89) is in perfect agreement with (2.84). This match strongly suggests that the formula (2.89) is correct for all values of E in the range

$$1 \ll E < \frac{N^2}{4}. \tag{2.90}$$

The gapped phase

The analysis of this section applies to the saddle point at $u = u_{max}$ for $\alpha \leq \alpha_{pt}$ and to the saddle point at u_{min}. The partition function of this saddle is given by plugging the solution of the equation

$$u = 1 - \frac{1}{4\alpha u^{q-2}}\ , \quad u \geq \frac{1}{2} \tag{2.91}$$

into the formula

$$\log Z = -\frac{N^2}{2}\left[(q-1)\left(\frac{1}{4} - \frac{1}{2}\log\left[2(1-u)\right]\right) - 2\alpha u^{q-1}\right]. \tag{2.92}$$

As we have explained above, for $\alpha < \alpha_c = \dfrac{(q-1)^{q-1}}{4(q-2)^{q-2}}$ there are no legal solutions to (2.91). For $\alpha_c < \alpha < \alpha_{pt} = 2^{q-3}$ there are two legal solutions and for $\alpha > \alpha_{pt}$ there is a

single legal solution to this equation. After the partition function is obtained one obtains the energy and entropy of the solution using the thermodynamical formulae

$$E(\alpha) = \alpha \partial_\alpha \log Z(\alpha) \ , \quad S(\alpha) = \left(\log Z(\alpha) - E(\alpha) \log \frac{\alpha}{N^{q-3} p} \right). \tag{2.93}$$

Eliminating α from the expressions obtained in (2.93) we find the entropy S as a function of the energy. This function $S = S(E)$ is difficult to find explicitly simply because (2.91) is difficult to solve. The procedure described above, however, implicitly defines this function. It is not difficult to convince oneself that there is a single saddle point of this nature for every energy $E > \frac{N^2}{4}$ and that the function $S(E)$ is an analytic function of energy for every energy greater than $\frac{N^2}{4}$.

While explicit formulae are difficult to obtain in general, they are easy to obtain in three special limits which we now describe

A. The solutions with α near α_{pt} i.e. (E near E_{pt}):

At $\alpha = \alpha_{pt}$ (2.91) admits the solution $u = \frac{1}{2}$. (This is a solution at $u = u_{max}$, i.e. the solutions that is a local maximum). It follows that at $\alpha = \alpha_{pt} - \delta\alpha$, (2.91) admits a solution with $u = \frac{1}{2} + \delta u$. Here δu is solved order by order in $\delta\alpha$. A few lines of standard algebra gives:

$$\begin{aligned} S\left(E_{pt} + \delta E\right) = & -\frac{1}{4} N^2 \left[\log \left(\frac{2^{q-3} N^{3-q}}{p} \right) + \frac{q-3}{2} \right] - \log \left[\frac{2^{(q-3)/2} N^{3-q}}{p} \right] \delta E \\ & + \frac{2(q-3)}{2N^2} \left(\delta E\right)^2 + \frac{4(7-3q)}{6N^4} \left(\delta E\right)^3 + \dots \end{aligned} \tag{2.94}$$

Comparing (2.94) and (2.89), it is easily verified that while $S(E)$, $S'(E)$ and $S''(E)$ are continuous at $E = \frac{N^2}{4}$, $S'''(E)$ is discontinuous. In that sense the function $S(E)$ has a third order phase transition' at $E = \frac{N^2}{4}$. Further taking the limit:

$$\lim_{\epsilon \to 0^+} S''' \left(\frac{N^2}{4} - \epsilon \right) = \frac{4(6-2q)}{N^4}, \quad \lim_{\epsilon \to 0^+} S''' \left(\frac{N^2}{4} + \epsilon \right) = \frac{4(7-3q)}{N^4} \tag{2.95}$$

This discontinuity is a consequence of the fact that the saddle point undergoes a Gross Witten Wadia transition at this energy.

B. The solutions with α near α_c (i.e. E near E_c):

At $\alpha = \alpha_c$ (2.91) admits the solution $u = \frac{q-2}{q-1}$. For $\alpha = \alpha_c + \delta\alpha$ the (2.91) admits two solutions near this critical solution at $u = u_c + \delta u$; these are the solutions at $u = u_{max}$

and $u = u_{min}$ respectively. A careful calculation shows E, S as a function of α are different for this two branches but S as a function of E is same for both of them and given by:

$$\begin{aligned} S\left(E_c + \delta E\right) =& \frac{1}{4}N^2\left[-(q-2)\log\left(\frac{(q-2)^{2-q}(q-1)^{q-1}N^{3-q}}{4p}\right) + (q-1)\log\left(\frac{2}{q-1}\right) + \frac{(q-3)}{2}\right] \\ & - \log\left[\frac{(q-2)^{2-q}(q-1)^{q-1}N^{3-q}}{2^{(q+1)/2}p}\right](\delta E) + \left[-\frac{8}{3N^4(q-2)(q-1)}\right](\delta E)^3 + \dots \end{aligned} \tag{2.96}$$

Note that (2.96) is completely smooth around $E = E_c = \frac{1}{4}N^2(q-2)$.

C. The solutions with $\alpha \gg 1$ (i.e. $E \gg \frac{N^2}{2}$):

At $\alpha \gg 1$ (2.91) admits the solution near $u = 1$; this is the thermodynamically dominant saddle at $u = u_{max}$. Setting $u = 1 - \delta u$, δu is solved to give as series in $\frac{1}{\alpha}$:

$$\delta u = \left(\frac{1}{4}\right)\alpha^{-1} + \left(\frac{q-2}{16}\right)\alpha^{-2} + \dots \tag{2.97}$$

It follows that:

$$\begin{aligned} \log Z(\alpha) &= N^2\alpha + \left(-\frac{1}{4}N^2(q-1)\right)\log(\alpha) + \dots, \\ E(\alpha) &= N^2\alpha + \left(-\frac{N^2(q-2)(q-1)}{32}\right)\alpha + \dots, \\ S(\alpha) &= \left(-N^2\right)\alpha\log(\alpha) + \left(N^2\left(1 + \log\left(2pN^{q-3}\right)\right)\right)\alpha + \dots \end{aligned} \tag{2.98}$$

which gives

$$S(E) \;=\; E - E\log\left(\frac{E}{pN^{q-1}}\right) - \frac{N^2}{2}\left[\frac{q-1}{2}\log\left(\frac{2E}{N^2}\right) + \frac{3}{4}(q-1) + \frac{1}{8}(q-1)\left(\frac{2E}{N^2}\right)^{-1}\right] + \dots \tag{2.99}$$

Entropy as a function of energy for $E \gg \frac{N^2}{2}$

We have verified above that for $E \gg \frac{N^2}{2}$ the saddle point for the eigenvalue distribution function becomes very peaked and so is well approximated by a delta function. Whenever the eigenvalue distribution becomes so peaked effect of the holonomies on the partition function of the system can be ignored. It follows that for energies much greater than N^2 the partition function of our system is simply that of $N_F N^{q-1}$ complex fermionic oscillators. The partition function for our system thus reduces to

$$\log Z(x) = N_F N^{q-1}\log(1+x), \tag{2.100}$$

For $x \ll 1$ (2.100) reduces to

$$\log Z(x) = \ x \ p \ N^{q-1}. \tag{2.101}$$

Substituting $x = \frac{\alpha}{N^{q-3}p}$ we find that (2.101) agrees precisely with the leading term in the first line of (2.98):

$$\log Z(\alpha) = N^2\alpha. \tag{2.102}$$

The energy of the corresponding phase is given by

$$E(\alpha) = \alpha\partial_\alpha \log(Z(\alpha)) = N^2\alpha. \tag{2.103}$$

The entropy of this phase is given by

$$S(\alpha) = \log Z(\alpha) - E(\alpha)\log\frac{\alpha}{N^{q-3}p} = N^2\left(1 - \log\left(\frac{\alpha}{pN^{q-3}}\right)\right)\alpha. \tag{2.104}$$

Eliminating α between (2.103) and (2.104) we obtain

$$S(E) = E\left(1 - \log\left(\frac{E}{pN^{q-1}}\right)\right). \tag{2.105}$$

Note that (2.105) matches with the leading and 1st subleading term in (2.99).

2.4 The holonomy effective action with weak interactions

In the previous section we studied the free energy of the mass deformed SYK model in the zero coupling $\frac{J}{m} = 0$. In this section we will study corrections to the results of the previous section in a power series expansion in the coupling constant. For simplicity we also study the special case $N_F = 1$ in (2.7) .

In principle the leading large N contribution to S_{eff} is given as follows (we restrict attention to the massless case for simplicity in this paragraph). Consider the gap equation (2.20). We are instructed to solve this gap equation on a thermal circle, subject to the requirement that the solution respect the boundary conditions

$$\begin{aligned}
G\left(t_1 + \frac{\beta}{2}, t_2\right) &= -UG\left(t_1 - \frac{\beta}{2}, t_2\right) \\
G\left(t_1, t_2 + \frac{\beta}{2}\right) &= -G\left(t_1, t_2 - \frac{\beta}{2}\right)U^{-1} \\
\Sigma\left(t_1, t_2 + \frac{\beta}{2}\right) &= -U\Sigma\left(t_1, t_2 - \frac{\beta}{2}\right) \\
\Sigma\left(t_1 + \frac{\beta}{2}, t_2\right) &= -\Sigma\left(t_1 - \frac{\beta}{2}, t_2\right)U^{-1}
\end{aligned} \tag{2.106}$$

We must then plug this solution into (2.18) and the corresponding result is represented by $S_{\text{eff}}(U)$. While this prescription is clear it is rather difficult to implement in practice. In order to get some intuition for the effect of interactions on $S_{\text{eff}}(U)$ present some perturbative results for this object.

The thermal partition function of theory (2.7) is given, as usual, by the Euclidean path integral of the theory on a thermal circle of circumference β. The free result (2.8) is obtained by integrating out all fermions at at 'one loop' (i.e. by computing fermionic determinants -we explain how this works in more detail below). Corrections to (2.8) are obtained by including the contribution of more general diagrams.

It was demonstrated in [19] that, in the strict large N limit of interest to this chapter, the only graphs that contribute are melonic graphs. One way of organizing the graphs that contribute to our computation is by the number of melons a graph contains. We will refer to a graph with n melons as an n^{th} order graph. Such graphs are proportional to J^{2n}. As in the previous section we will be interested in the effective action as a function of holonomies, $S_{\text{eff}}(U)$. Let the contribution to $S_{\text{eff}}(U)$ from graphs of n^{th} order be denoted by $S_n(U)$. We have

$$S_{\text{eff}}(U) = \sum_{n=0}^{\infty} S_n(U). \tag{2.107}$$

As in the previous section we will principally be interested in the partition function in the scaling limit (2.10). In this limit the temperature is very small and so β is very large $\beta \sim \log N$. For this reason it is important to keep track of explicit multiplicative factors of β (as opposed, for instance, to factors of $x = e^{-\beta m}$) in our results. Below we will demonstrate that n^{th} order graphs have at least one and at most n explicit multiplicative factors of β. It follows that the contributions of n^{th} order graphs to the effective action can be organized in series

$$S_n(U) = \frac{J^{2n}\beta^n}{m^n} \sum_{a=0}^{n-1} \left(\frac{1}{m\beta}\right)^a f_a^n(x,U) \equiv -\left(\frac{J}{m}\right)^{2n} F_{2n}(m\beta, x, U). \tag{2.108}$$

Substituting (2.108) into (2.107), we can rearrange the sum over graphs as

$$\begin{aligned} S_{\text{eff}}(U) &= \sum_{k=0}^{\infty} \left(\frac{J}{m}\right)^{2k} H_k(\frac{J^2\beta}{m}, x, U), \\ H_k\left(\frac{J^2\beta}{m}, x, U\right) &= \sum_{n=k}^{\infty} \left(\frac{J^2\beta}{m}\right)^{n-k} f_k^n(x,U). \end{aligned} \tag{2.109}$$

As we are interested in the scaling limit (2.10) it follows that:

$$\begin{aligned} H_k\left(\frac{J^2\beta}{m}, x, U\right) &= \widetilde{H}_k\left(\frac{J^2\beta}{m}\right) x \prod_{i=1}^{q-1} \mathrm{Tr} U_i, \\ f_k^n(x,U) &= f_k^n x \prod_{i=1}^{q-1} \mathrm{Tr} U_i, \\ H_k\left(\frac{J^2\beta}{m}\right) &= \sum_{n=k}^{\infty} \left(\frac{J^2\beta}{m}\right)^{n-k} \widetilde{f}_k^n, \\ S_{\text{eff}}(U) &= x \prod_{i=1}^{q-1} \mathrm{Tr} U_i \times \sum_{k=0}^{\infty} \left(\frac{J}{m}\right)^{2k} \widetilde{H}_k\left(\frac{J^2\beta}{m}\right), \end{aligned} \tag{2.110}$$

where we will present an argument for the u dependences asserted here below.

(2.110) represents an interesting reorganization of usual perturbation theory. This reorganization is particularly useful at small $\frac{J^2}{m^2} \ll 1$ but finite values of $\frac{J^2\beta}{m}$. As $\beta \sim \frac{m}{\log N}$ in the scaling limit, it follows that $\frac{J^2\beta}{m}$ is fixed only for $\frac{J^2}{m^2} \sim \frac{1}{\log N}$. At these small values of the coupling, $S_{\text{eff}}(U)$ is well approximated by the first term in the expansion in (2.110), i.e. by the term proportional to $\widetilde{H}_0$. We will explicitly evaluate $\widetilde{H}_0$ in this section and so reliably determine the partition function when $\frac{J^2}{m^2}$ is in the parametric range described above. [36]

In the rest of this section we present the results of our explicit perturbative computations. Although we are principally interested in the function H_0 in the scaling limit, to set notations and for practice we first present the results of simpler computations. To start with we work out the partition function at level zero and recover the free partition function of the previous section. We then work out the partition function at level 1 (i.e. including graphs with a single melon). Next we present our results at level 2 (i.e. including all graphs with two melons). Finally we turn to the problem of principal interest to us,

[36] Although this is far from guaranteed, it is possible that the approximation $S_{\text{eff}} \sim H_0$ has a larger range of validity. Let us consider the parametric regime in which $\frac{J^2}{m^2}$ is small compared to unity but large compared to $\frac{1}{m\beta}$. In this regime $\frac{J^2\beta}{m}$ is effectively scaled to infinity. Let us define

$$r_k = \lim_{\frac{J^2\beta}{m} \to \infty} \frac{H_k}{H_0}. \tag{2.111}$$

If it turns out that r_k is bounded (finite) then it follows that H_0 is in fact also a good approximation to the partition function for all values of β assuming only that $\frac{J^2}{m^2} \ll 1$. It would be interesting to investigate whether r_k above are actually bounded for all k; however we leave that to future work.

namely the sum of the infinite set of graphs that generates H_0. As preparation for all these computations we first briefly discuss the structure of the free Greens function.

2.4.1 Free Greens Function

Consider the free fermionic Greens function

$$\langle \psi^a(t)\overline{\psi}_b(0)\rangle.$$

We work in a colour basis in which the holonomy U is diagonal. In this basis in which the action of holonomies on the fermions is given by

$$\begin{aligned} U\psi^a &= e^{i\theta_a}\psi^a, \\ U\overline{\psi}_a &= e^{-i\theta_a}\psi_a. \end{aligned} \tag{2.112}$$

The free fermionic Greens function at finite temperature is given by

$$\langle \psi^a(t)\overline{\psi}_b(0)\rangle = G_0(t)\delta^a_b,\ G_0(t) = f(t,m,\theta_a),\quad for\ \ -\beta \le t \le \beta \tag{2.113}$$

where

$$\begin{aligned} f(t,m,\theta_a) =& \frac{1}{2}e^{-(m+i\theta_a)t}\left[\mathrm{sgn}(t) + \tanh\left(\frac{1}{2}(m+i\theta_a)\beta\right)\right] \\ =& \frac{e^{-(m+i\theta_a)t}}{1+x\ e^{-i\theta_a\beta}}\left[\Theta(t) - \Theta(-t)\ x\ e^{-i\theta_a\beta}\right]. \end{aligned} \tag{2.114}$$

Note that the function f obeys the identity

$$f\left(\frac{\beta}{2}+t,m,\theta_a\right) = -f\left(-\frac{\beta}{2}+t,m,\theta_a\right)\quad for\ \ 0 \le t < \frac{\beta}{2}, \tag{2.115}$$

from which it follows that the Greens function is antiperiodic on the circle as required on physical grounds.

Note that we have presented the Greens function only in the 'fundamental domain' $-\beta < t < \beta$. Our fermionic Greens function is taken by definition to be a periodic function of t with period 2β; this property plus the explicit results (2.113) and (2.114) can be used to define the Greens function at every value of Euclidean time as required. The extended Greens function defined in this manner has singularities at $t = n\beta$ for every integral value of n, and is smooth everywhere else.

Note also that the 'reversed' Greens function $\langle\overline{\psi}_a(t)\psi^b b(0)\rangle$ is also given in terms of the function G_0 by the formula [37]

$$\langle\overline{\psi}_a(t)\psi^b(0)\rangle = -G_0(-t)\delta^b_a. \tag{2.116}$$

This formula is also manifestation of symmetry of mass deformed SYK Lagrangian under the simultaneous swaps $\overline{\psi} \leftrightarrow \psi$, $U \leftrightarrow U^{-1}$, $m \leftrightarrow -m$.

[37] Owing to time translation symmetry.

2.4.2 Level zero: Free theory

In this brief subsection we compute S_{eff} at one loop, i.e. in the free theory. The result for $S_{\text{eff}}(U)$ was already presented in the previous subsection; we obtain that result here from a one loop computation as a simple practice exercise. Let

$$\omega_n = \frac{2\pi}{\beta}\left(n + \frac{1}{2}\right). \tag{2.117}$$

The fields ψ^a and $\overline{\psi}_a$ can be independently expanded in Fourier space as

$$\psi(t) = \sum_n \psi^n e^{-i\omega_n t} \ , \ \overline{\psi}(t) = \sum_n \overline{\psi}_n e^{+i\omega_n t}. \tag{2.118}$$

When substituted the free part of action (2.7) becomes

$$S = \sum_{n,a} \overline{\psi}_{a,n}[\beta(-i\omega_n + m + i\theta_a)]\psi^{a,n}. \tag{2.119}$$

Fermionic integration gives:

$$\begin{aligned}
Z_F =& \prod_a \prod_{n=-\infty}^{n=+\infty} [\beta(-i\omega_n + m + i\theta_a)] \\
=& \prod_a \prod_{n=-\infty}^{n=+\infty} [-i(2\pi n + \pi) + m\beta + i\theta_a\beta] \\
=& \prod_a \prod_{n=-\infty}^{n=+\infty} [-i2\pi n + c(\theta_a)] \\
=& \prod_a c(\theta_a)^2 \prod_{n=1}^{n=+\infty} [(-i2\pi n + c(\theta_a))(+i2\pi n + c(\theta_a))] \\
=& \prod_a c(\theta_a)^2 \prod_{n=1}^{n=+\infty} \left[(2\pi n)^2 + c(\theta_a)^2\right] \\
=& \prod_a c(\theta_a)^2 \left(\prod_{n=1}^{n=+\infty} (2\pi n)^2\right) \prod_{n=1}^{n=+\infty} \left[1 + \left(\frac{c(\theta_a)/2}{\pi n}\right)^2\right] \\
=& \prod_a c(\theta_a)^2 \left(\prod_{n=1}^{n=+\infty} (2\pi n)^2\right) \left[\frac{\sinh\frac{c(\theta_a)}{2}}{\frac{c(\theta_a)}{2}}\right] \\
=& N \prod_a \left[\sinh\frac{c(\theta_a)}{2}\right] \sim \prod_a e^{\frac{c(\theta_a)}{2}}(1 - e^{-c(\theta_a)}),
\end{aligned} \tag{2.120}$$

where

$$c(\theta_a) = m\beta + i\theta_a\beta - i\pi$$

and

$$N = \prod_a \prod_{n=1}^{n=+\infty} (2\sqrt{2}\pi n)^2$$

is an infinite holonomy independent constant. As for every θ_a there is $-\theta_a$ to be taken into account $\prod_a e^{\frac{c(\theta_a)}{2}}$ becomes independent of holonomy. Keeping only holonomy dependent terms [38]

$$\log Z = \sum_a \log[1 + xe^{-i\theta_a\beta}] \ , \quad x = e^{-m\beta}, \tag{2.121}$$

In other words

$$\log Z = \mathrm{Tr} \log[1 + xU]. \tag{2.122}$$

Expanding (2.122) in a power series in x we recover (2.8) at $N_F = 1$. In the scaling limit we recover (2.64).

2.4.3 Level one: single melon graphs

The contributions of graphs with a single melon to the Free energy is given by

$$F_2 = \frac{1}{2!} \ {}^2C_{\frac{2}{2}} \ (-1)^{q/2} q \ m^2 \int \prod_{k=1}^{q} G_0(t_1 - t_2, \theta_{a_k}) \ dt_1 dt_2, \tag{2.123}$$

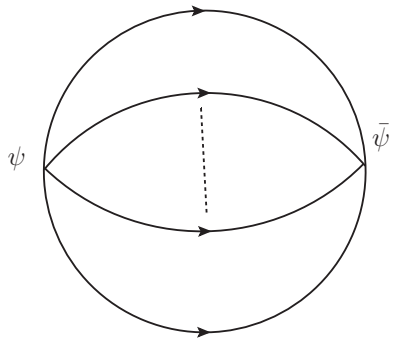


Figure 2.2: Single loop contribution to free energy.

In this graph we contract each of fields in the interaction vertex ψ^q with one of the fields in $\overline{\psi}^q$. Consider any particular ψ field. This ψ field has to contract with one of the q $\overline{\psi}$ fields in the second interaction vertex. It is thus clear that there are q choices for this contraction (the choices of which $\overline{\psi}$ our specified ψ pairs up with). Once this choice has been made, if we are interested - as we are- in graphs that contribute only at leading order

[38] Note that this also ensures for $\beta \to \infty$ partition function is 1 and for $\beta \to 0$ total number of states for a given a are 2.

in large N there are no further choices in our contraction. Recall that every one of the remaining ψ's (respectively $\overline{\psi}$'s) has exactly one colour common with the ψ (resp $\overline{\psi}$) that we have just contracted together. The leading large N behaviour is obtained only if the ψ that shares any given gauge index with our special contracted ψ is now contracted with the $\overline{\psi}$ that shares the same gauge index with the special contracted $\overline{\psi}$. This rule specifies a unique contraction structure for the remaining fields. It follows that, up to a sign, the symmetry factor is simply q. The sign in question is simply $(-1)^{(q-1)+(q-2)+..+1}$ Recalling that q is even, it is easy to see that this phase $= (-1)^{q/2}$.

The integral in (2.123) is very easy to perform. To see this note that the analytic structure of the integrand as a function of $t = t_1 - t_2$ takes the form

$$e^{-qmt}(A_q \, \mathrm{sgn}(t) + B_q),$$

for various different values of q. The integrand is integrated from $-\frac{\beta}{2}$ to $\frac{\beta}{2}$. The integral over $t_1 + t_2$ produces an overall factor of β. The integrals are all trivial to do; evaluating them we find the final answer

$$F_2 = \frac{(-1)^{q/2}}{2!} \, {}^2C_{\frac{2}{2}} \, q \, m\beta \, I_1^{(2)}(q, x), \tag{2.124}$$

where

$$I_1^{(2)}(q, x) = \frac{1 - x^q}{q} \mathrm{Tr}_F \prod_{k=1}^{q} \left(\frac{1}{1 + x\widetilde{U}_k} \right). \tag{2.125}$$

The expression $\mathrm{Tr}_F(...)$ in the equation above represents the trace over an operator built on a particular Auxiliary Hilbert space. The operator in question is a function of the elementary operators $\widetilde{U}_k$ that act on this Hilbert space. We will now carefully define the relevant Hilbert space and the operators $\widetilde{U}_k$ and so give meaning to (2.125).

The operators $\widetilde{U}_k$ in (2.125) have the following meaning. These operators are unitary operators that act on a vector space whose dimensionality is $N^{\frac{q}{2}(q-1)}$. The vector space in question is the tenor product of $q-1$ factors, each of which has dimension $N^{\frac{q}{2}}$. Each factor described above is associated with one of the $q-1$ gauge groups. Let us focus on any one gauge group, say the first. The factor associated with this gauge group consists of $\frac{q}{2}$ distinct factors of isomorphic N dimensional spaces on which the $N \times N$ holonomy matrices of the first gauge group naturally act.

Recall that each ψ field that appears in an interaction has exactly one gauge index contraction with every other ψ field. This means, in particular, that the indices of gauge group 1 are contracted between $\frac{q}{2}$ pairs of ψs. This fact is the origin of the $\frac{q}{2}$ distinct factors of the space on which the holonomy matrices of the first gauge group act.

With all this preparation we now explain the form of the operators $\widetilde{U}_k$. Each $\widetilde{U}_k$ acts as U_1 (the holonomy of the first $O(N)$ gauge group) on one of the $\frac{q}{2}$ copies of the N dimensional vector space associated with the first $O(N)$, and as identity on the remaining $\frac{q}{2} - 1$ copies of this space. In a similar fashion it acts as U_2 on one of the $\frac{q}{2}$ copies of

the N dimensional vector space associated with the second $O(N)$, and as identity on the remaining $\frac{q}{2}-1$ copies of this space. And so on. Exactly two $\widetilde{U}_k$s act as U_1 on the same Hilbert space. Exactly two $\widetilde{U}_k$s act as U_2 on the same Hilbert space, etc. Finally every two $\widetilde{U}_k$s act on the same Hilbert space for one and only one gauge group.[39] The symbol Tr_F in that equation denotes the trace over the full $N^{\frac{q}{2}(q-1)}$ dimensional Hilbert space.

From a practical point of view it is less complicated to use the definitions of the $\widetilde{U}_k$ operators than it might at first seem. We could, for instance, expand the result (2.125) in a power series in x. The formal looking expressions of traces of sums of products of $\widetilde{U}_k$ operators that appear as coefficients in this expansion can easily be evaluated in terms of traces of powers of the holonomy matrices $U_1 \dots U_{q-1}$ of the factors of $O(N)$.

A little thought will allow the reader to convince herself that the rules described above imply that, for instance

$$\mathrm{Tr}\left(\sum_{k=1}^{q}\widetilde{U}_k\right) = qN^{\frac{(q-2)(q-1)}{2}}\mathrm{Tr}U_1\mathrm{Tr}\ U_2\dots\mathrm{Tr}U_{q-1}, \tag{2.126}$$

$$\begin{aligned}\mathrm{Tr}\left(\sum_{k_1\neq k_2}^{q}\widetilde{U}_{k_1}\widetilde{U}_{k_2}\right) &= qN^{\frac{q^2-5q+6}{2}}\left[\prod_{k=1}^{q-1}\mathrm{Tr}U_1^2\ (\mathrm{Tr}U_2)^2\dots(\mathrm{Tr}U_{q-1})^2\right.\\ &\qquad \left.+(1\leftrightarrow 2)+(1\leftrightarrow 3)+\dots(1\leftrightarrow q-1)\right],\end{aligned} \tag{2.127}$$

$$\mathrm{Tr}\left(\sum_{k=1}^{q}\widetilde{U}_k^2\right) = qN^{\frac{(q-2)(q-1)}{2}}\left(\mathrm{Tr}U_1^2\ \mathrm{Tr}U_2^2\dots\mathrm{Tr}U_{q-1}^2\right). \tag{2.128}$$

As an illustration of these rules let us evaluate the partition function in the low energy scaling limit described in the previous section. Recall that in the limit of interest $x\sim\frac{1}{N^{q-3}}$ and we are instructed to retain only those contributions to $S_{\mathrm{eff}}(U)$ that are linear in x; terms of higher order in x can be discarded. It follows that the partition function in this limit may be evaluated by Taylor expanding (2.125) in x and discarding all terms that are quadratic or higher order in x. Using the first of (2.126) we conclude immediately that

$$F_2 = \frac{(-1)^{q/2}}{2!}\ {}^2C_{2/2}\ q\,m\beta\ N^{q-1}(-x)\prod_{m=1}^{q-1}\rho_m^1. \tag{2.129}$$

where $\rho_m^1=\frac{\mathrm{Tr}U_m}{N}$ as in the previous section, and we have dropped the terms of order x^0 which are independent of U_m.

[39] This means that if U_1 and U_3 act on the same copy of the Hilbert space for gauge group 1, then they necessarily act on different copies of the Hilbert space for all the other gauge groups.

2.4.4 Level 2: 2 melon graphs

At level 2 we once again have contributions from a single Feynman diagram Fig 2.3. In order to evaluate this graph we must evaluate in integral

$$F_4 = \frac{1}{4!}\,{}^4C_{4/2}\,(-1)2(q^2)^2 \int \prod_{i=1}^{4} dt_i \left(\prod_{i=1}^{q-1} G_0(t_{12},\theta_{a_i})\right)\left(\prod_{i=1}^{q-1} G_0(t_{34},\theta_{b_i})\right) \text{`} \times G_0(t_{32},\theta_{c_2})G_0(t_{14},\theta_{c_1}). \quad (2.130)$$

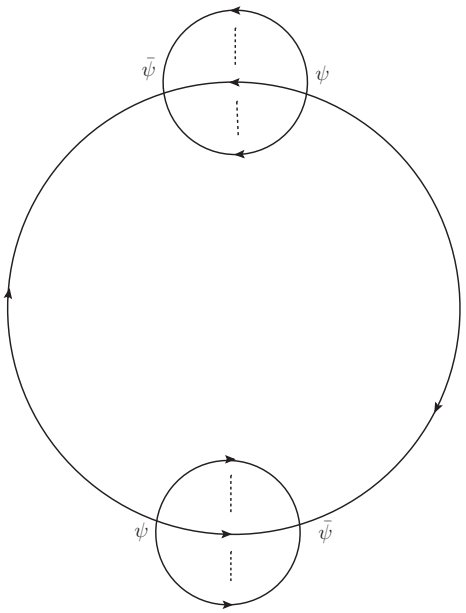


Figure 2.3: Two loop contribution to free energy.

We give some details of this expression and the evaluation of this integral in the Appendix A.2. We have completely evaluated this integral with the help of mathematica (see Appendix A.2.2 for arbitrary number of melons), but the final result for $S_{\text{eff}}(U)$ in the general case is too complicated to transfer to text. As before, however, the answer simplifies dramatically in the low energy scaling limit of the previous section (see Appendix A.2.1) and we find

$$F_4 = \frac{(-1)}{4!}\,{}^4C_{4/2}\,2(q^2)^2[m\beta\; I_1^{(4)}(q) + m^2\beta^2\; I_2^{(4)}(q)]\; N^{q-1}x \prod_{m=1}^{q-1} \rho_m^1, \quad (2.131)$$

where

$$\begin{aligned} I_1^{(4)}(q) &= -\frac{2}{q}(2q-3), \\ I_2^{(4)}(q) &= -1. \end{aligned} \quad (2.132)$$

Note that the final answer had two terms; one proportional to an overall factor of β and the second proportional to β^2. In the next subsection we will argue that a graph at level n, in the low temperature scaling limit, has terms proportional to β^q for $q = 1 \ldots n$.

2.4.5 The infinite sum H_0

We will now turns to a study of the free energy at level n. As in the previous subsection we will focus on the start at the low energy scaling limit of the previous section, and so retain only those terms in all graphs that are proportional to x. As we will see below, general graphs in the scaling limit and at level n break up into different pieces that are proportional to β^k for $k = 1 \ldots n$.[40] We will further focus our attention on the graph with the largest power of β, i.e. in this subsection we will contribute that piece of the level n answer that scales like $\beta^n x$. It turns out that this piece is rather easy to extract as we now explain.

Let us first recall that the propagator in our theory takes the following form:

$$\langle \psi_a(t) \overline{\psi}^b(0) \rangle = \frac{e^{-(m+i\theta_a)t}}{1 + x\ e^{-i\theta_a \beta}} \left[\Theta(t) - \Theta(-t)\ x\ e^{-i\theta_a \beta} \right]. \tag{2.133}$$

It will turn out (and we will see explicitly below) that the denominator in (2.133) only contributes at order β^{n-1} or lower in free energy linear in x. For the purposes of the current subsection, therefore (where we wish to ignore terms at order x^2 or higher and only keep highest power of β) this denominator can be dropped, and we can work with the simplified propagator[41]

$$\langle \psi_a(t) \overline{\psi}^b(0) \rangle = e^{-(m+i\theta_a)t} \left[\Theta(t) - \Theta(-t)\ x\ e^{-i\theta_a \beta} \right]. \tag{2.134}$$

In this subsection we assume $m > 0$; the case $m < 0$ can be argued in a completely analogous manner with the role of ψ and $\overline{\psi}$ reversed in the analysis below. In the computation of Feynman diagrams on the circle we will need to choose a 'fundamental domain' on the circle; our (arbitrary but convenient) choice of fundamental domain is

$$-\frac{\beta}{2} < t < \frac{\beta}{2} \tag{2.135}$$

[40] For instance the level one graph computed above was proportional to β while the level two graph was the sum of one term proportional to β and another term proportional to β^2.

[41] The role that of the overall holonomy dependent phase factors above is quite subtle. Naively these overall factors can be dropped in their contribution to free energy diagrams. The naive argument for this is that the net contribution to of these phase factors at any interaction vertex is proportional to $\prod_a e^{i(\theta_a)t_1}$ where the sum runs over the phases θ_a of all the q propagators that end at that interaction vertex. As the interaction vertex is a gauge singlet, $\sum \theta_a$ vanishes, so it might at first seem that the contribution of all these phase factors drops out. This is in general incorrect. The subtlety is that t_1 is not single valued on the circle. In diagrams in which propagators 'wind' as they go around the circle, one of the factors in the product may effectively be evaluated at, e.g. $t_1 + \beta$ and so the net contribution of this phase factor could turn out to be $e^{i\beta\theta_a}$. While this contribution is constant (independent of t_1), it is nontrivial in nonzero winding sectors. Such a contribution will play an important role in our computation below.

Finally some terminology. We will call the part of the propagator (2.134) that is proportional to $\theta(t)$ the 'forward' ('normal') part of the propagator, and the part of the propagator proportional to $\theta(-t)$ the 'reverse' part of the propagator. Note that the normal part of the propagator ranges is modulus from 1 to $\sqrt{x}$; it is maximum (i.e. unity) at $t = 0$ and minimum (i.e. $\sqrt{x}$) at $t = \frac{\beta}{2}$. The modulus of the reverse part of the propagator varies in magnitude from $\sqrt{x}$ to x. It is minimum (i.e. equal to x) at $t = 0$ and maximum (i.e. equal to $\sqrt{x}$) at $t = -\frac{\beta}{2}$.

With all this preparation we are now ready to isolate the parts of the level n diagrams whose contribution is proportional to $x\beta^n$.

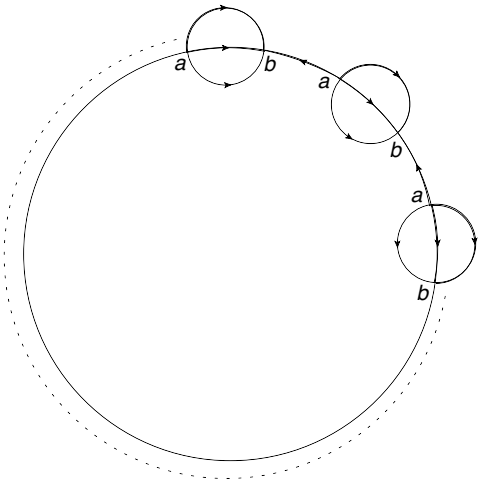


Figure 2.4: Circle diagram: a, b represents respectively insertions of ψ, $\overline{\psi}$. Direction of arrow is from ψ to $\overline{\psi}$. The diagram is drawn for $q = 4$.

To start with let us consider the simple n^{th} level ring diagram depicted in Fig. 2.4. In this diagram we have n a type vertices and n b type vertices. In this graphs we have $q-1$ propagators connecting adjacent a and b type vertices, but only a single propagator connecting b to a type vertices.

Consider any propagator between a and b type vortices - which has a type vertex A at time t_1 and its adjacent b type vertex B at time t_2. Depending on whether $t_1 > t_2$ or $t_1 < t_2$, all the $q-1$ propagators from A to B are either simultaneously all reverse or simultaneously all normal. If all propagators are reverse, the modulus of these propagators is less than $\sqrt{x}^{q-1} < x$ (recall $q \geq 4$). It follows that configurations in which the propagators from

A to B do not contribute in the scaling limit, and so all propagators from A to B must be normal. Given that these propagators are all normal their modulus is proportional to $e^{-m(q-1)|t_1-t_2|}$. It is intuitively clear that separating t_1 from t_2 over a finite fraction of the circle forces us to pay a high cost in factors of x; it can be shown (this will be clearer in a bit) that such configurations do not contribute to the result in the scaling limit. In the scaling limit we only receive contributions from configurations in which $|t_1 - t_2|$ is of order $\frac{1}{m}$. It follows that for parametric purposes, we can simply regard t_1 and t_2 as the same point, replacing the integral over $t_1 - t_2$ by $\frac{1}{m}$. For parametric purposes, in other words, each of the melons in Fig. 2.4 can be thought of as a single interaction vertex, inserted at a single 'self energy vertex', inserted at a single time, with effective an effective insertion factor of order $\frac{J^2}{m}$.

Now let us turn to the propagators between b and a type vertices. These are now n different propagators connecting the effective self energy blobs described in the previous paragraph. Let the effective times of insertions of these self energy blobs be T_1, $T_2 \ldots T_n$. Our graph is proportional to the product of n propagators, the first from T_1 to T_2, the second from T_2 to T_3 ... and the last from T_n to $T_1 + w\beta$ where w is an integer. As each reverse propagator contributes a factor of at least $\sqrt{x}$ to the integrand, no more than two of these propagators can be reverse.

Let us first consider diagrams in which all propagators are forward. As all propagators move forward in time, the final propagator in the sequence must end not at time T_1 but at time $T_1 + w\beta$ where w is a positive integer. The modulus of the product of these propagators is then easily seen to be proportional to $e^{-wm\beta} = x^w$. In the scaling limit of interest to us, the only option is $w = 1$. Once we set $w = 1$, the integrand of the diagram is now independent of the effective insertion times T_i. The integral over these n insertion times thus gives a factor β^n, and the contribution of the graph in question is proportional to $x\beta^n$ as desired.

Now let us consider diagrams in which one of the propagators between the effective self energy vertices is reverse, and the rest are forward. It is easy to verify that the modulus of the product of propagators in such a graphs is proportional to $xe^{-wm\beta}$ where $w = 0, 1, \ldots$. In the scaling limit under consideration we are interested only in $w = 0$. Once again the modulus of these graphs is independent of the insertion positions of the effective self energy vertices, and integration over their locations produces a result proportional to $x\beta^n$ as required.

Diagrams in which two of the propagators are reverse are kinematically very constrained. Similar argument as above shows these graphs are proportional to x only if $w = -1$,i.e, if the two reverse propagators each have length $\frac{\beta}{2}$ (up to corrections of order $\frac{1}{m}$) and so all the forward propagators have length zero, again up to corrections of order $\frac{1}{m}$. These constraints ensure that such graphs are proportional to β but no higher power

of β (certainly not β^n) and so are not of interest to the current section.

In summary, graphs of the form depicted in Fig. 2.4 only contribute at order $x\beta^n$ if all propagators from a to adjacent b type vertices are normal, if the separation between a and adjacent b type vertices is of order $\frac{1}{m}$, and if the propagators between adjacent melons are either all normal with net winding number one or one reverse and the rest normal with net winding number zero. Once we have identified the parts of these graphs that contribute at order $x\beta^n$, the computation of these contributions is very simple (see below).

Let us now turn to more general graphs than those drawn in Fig. 2.4. All graphs that contribute to the free energy at leading order in the large N limit are of the general structure depicted in 2.4, but with the melons in Fig 2.4 replaced by effective melons or 'cactus graphs'. The net effect of this is to replace the bare propagators between a and b type vertices in Fig. 2.4 by exact propagators. Recall that we are only interested in the propagator corrections at times $t = |t_1 - t_2| \sim \frac{1}{m} \ll \beta$. The k^{th} order correction to the forward propagator at short times takes the schematic form

$$G(t) \sim \frac{|J|^{2k} t^k}{m^k} \sum_{n=0}^{k} C_n \left(\frac{m}{t}\right)^n \tag{2.136}$$

As all values of t that contribute to our integrand in the low energy scaling limit of interest to this chapter are of order $\frac{1}{m}$, it follows that all terms on the RHS of (2.136) are of order $\frac{J^{2k}}{m^{2k}}$. As compared to the contribution of the graphs of Fig. 2.4, in other words, these graphs have extra powers of J^2 but no compensating factors of β. It follows that The contribution of such graphs at level n is always of the form $x\beta^h$ with h strictly less than n. Consequently all such graphs can be ignored.

In summary, the only graphs that contribute at terms proportional to $x\beta^n$ at level n are the very simple 'necklace' graphs depicted in Fig. 2.4. We have already explained above that the contribution of each of these graphs is easily evaluated in the scaling limit. It follows that the computation of the sum of these graphs is a relatively simple job.

Relegating all further details to the Appendix A.2.1 we simply list our results. The contribution of order $x\beta^n$ to $S_{\text{eff}}(U)$ from graphs of level n is given, for $n \geq 2$ by

$$\frac{J^{2n}}{m^{2n}} F_{2n} = 2x \; N^{q-1} \left(\prod_{m=1}^{q-1} \rho_m^1\right) \; \frac{1}{(n-1)!} \left[\gamma(q) \; \frac{(-\beta)}{m} |J|^2\right]^n \left(2 - \frac{2^{n-1}}{n}\right) + \mathcal{O}(\beta^{n-1}), \tag{2.137}$$

where

$$\gamma(q) = (-1)^{\frac{q}{2}(q-1)} \frac{q}{2}. \tag{2.138}$$

Summing these contributions over all $n = 2$ to infinity and adding the separate contribution

of $n = 1$ we find H_0.

$$\begin{aligned} H_0 =& 2x\ N^{q-1}\left(\prod_{m=1}^{q-1}\rho_m^1\right)\left[\sum_{n=2}^{\infty}\frac{1}{(n-1)!}\left[\gamma(q)\ \frac{(-\beta)}{m}|J|^2\right]^n\left(2-\frac{2^{n-1}}{n}\right)-\frac{(-1)^{q/2}}{2}q^2\ \beta\frac{|J|^2}{m}\right] \\ =& 2x\ N^{q-1}\left(\prod_{m=1}^{q-1}\rho_m^1\right)\left[\frac{1}{2}+2\gamma(q)\ \frac{(-\beta)}{m}|J|^2 e^{\gamma(q)\ \frac{(-\beta)}{m}|J|^2}-\frac{1}{2}e^{2\gamma(q)\ \frac{(-\beta)}{m}|J|^2}-\frac{(-1)^{q/2}}{2}q\ \beta\frac{|J|^2}{m}\right], \end{aligned} \tag{2.139}$$

so that the free holonomy effective action takes the form (2.108) with F_0 in that equation given by H_0 in (2.139).

Note that $\gamma(q)$ is positive for $q = 4, 8, 12\ldots$ but is negative for $q = 6, 10, 14\ldots$. It follows that the exponential terms in (2.139) decay at large $\frac{J^2\beta}{m}$ for the first set of values of q but blow up for the second set of values of q. It would be interesting to better understand the meaning and consequences of this observation.

2.4.6 Thermodynamics

At sufficiently weak coupling we have demonstrated in the previous subsection that the free result for $S_{\text{eff}}(U)$ in the scaling limit, (2.64), is replaced by the formula

$$-S_{\text{eff}}(U) = N^{q-1}\left(\prod_{m=1}^{q-1}\rho_m^1\right)x\widetilde{H}_0, \tag{2.140}$$

where H_0 was computed in the previous subsection.

Note that (2.140) has the same structure of U dependence as (2.64); it follows that the partition function obtained by integrating $e^{-S_{eff}(U)}$ over U is simply $Z(\widetilde{x})$ (where the function $Z(x)$ was defined in (2.64)). At small enough coupling $\widetilde{x}$ is close to x, and the structure of the canonical partition function generated by (2.140) is very similar to the results described in detail for the free theory in the previous section.

What does consequence does the replacement of x by $\widetilde{x}$ have for the micro canonical partition function? Let us first recall a simple formal result. Let

$$e^{-\beta m} \to e^{-m\beta}(1 + \epsilon\ h_0(\beta)).$$

By linearizing the usual thermodynamical formulae it is easy to show that this replacement results in the replacement

$$S(E) = S_0(E) + \epsilon\frac{E}{m}h_0\left[\frac{\partial S_0(E)}{\partial E}\right] + \mathcal{O}(\epsilon^2), \tag{2.141}$$

(this result holds provided we expand about an analytic point in the phase diagram, i.e. away from phase transitions). Clearly in our context this result applies if $\frac{J^2}{m^2} \sim \frac{\alpha}{\log N}$ and α is taken to be small. However our results for the partition function are valid over a larger

parametric regime; they are definitely valid whenever $\frac{J^2}{m^2} \sim \frac{\alpha}{\log N}$ even at finite values of α. In order to understand the effect $\widetilde{H}_0$ has on the entropy as a function of energy at such values of α we take a slightly different route.

Define a 'particle mass probability function' $p(m)$ by the following requirement

$$\int dm' e^{-\beta m'} p(m') = x H_0. \tag{2.142}$$

Intuitively $p(m)$ denotes a spread in the mass density function (which was a δ function for the free theory) that mimics the effects of interactions in thermodynamics.

A little thought demonstrates that the following ansatz for $p(m')$ reproduces the structure of our perturbative expansion for xH_0

$$p(m') = \sum_{k=0}^{\infty} \frac{1}{m} g_k \left(\frac{m'-m}{|J|^2/m} \right) \left(\frac{|J|}{m} \right)^{2k-2}, \tag{2.143}$$

where the functions $g_k(y)$ do not depend on J. Working with the probability distribution (2.143) is equivalent to replacing x by

$$x \to \int_0^{\infty} e^{-\beta m'} p(m') dm' = x \sum_{n,k=0}^{\infty} \frac{1}{n!} \left(-\frac{|J|^2 \beta}{m} \right)^n \left(\frac{|J|}{m} \right)^{2k} \int_{-m^2/J^2}^{\infty} u^n g_k(u) du. \tag{2.144}$$

The lower limit of the integration in (2.144) can safely be approximated by $-\infty$. If we want the RHS of (2.144) to equal $\widetilde{x}$ we must choose

$$\begin{aligned} \int_{-\infty}^{\infty} g_0(u) du = 1 \ , \ \int_{-\infty}^{\infty} u \ g_0(u) du = \frac{2}{q} \gamma(q), \\ \int_{-\infty}^{\infty} u^n \ g_0(u) du = 4n \left(1 - \frac{2^{n-2}}{n} \right) \gamma(q)^n \ , \ n \geq 2 \end{aligned} \tag{2.145}$$

These relationships determine the moments the as yet unknown g_0. Inverting these relations we find

$$\begin{aligned} p(m') =& 2\delta(m'-m) - \delta \left(m' - m - 2\gamma(q) \tfrac{|J|^2}{m} \right) \\ & - 4\gamma(q) \frac{|J|^2}{m} \delta' \left(m' - m - \gamma(q) \tfrac{|J|^2}{m} \right) \end{aligned} \tag{2.146}$$

Recall that the function $p(m')$ in the free theory was just a δ function localized at $m' = m$. The interaction effects considered in this section split this δ function into a set of 4 localized δ (or δ') spikes, distributed in a width of order $\frac{J^2}{m}$ around $m' = m$. As an aside we note the striking fact that interaction effects - at least at the order we have computed them - do not smoothen the free spectral function out.

It is not difficult to convince oneself that the function $S(E)$ that follows from (2.146) is qualitatively similar to the entropy as a function of energy derived in detail for the free theory in the previous section, and in particular displays faster than Hagedorn growth.

2.5 Discussion

In these notes we have argued that the quantum mechanical model (2.1) - which is known to agree with the SYK model in the strict large N limit - displays qualitatively new dynamics at subleading orders in $\frac{1}{N}$. We argued that the fluctuation spectrum about the finite temperature saddle point in this theory has new light modes - that originate in time dependent $O(N)^{q-1}$ transformations - in addition to the modes that arise from conformal diffeomorphisms and that were present also in the original SYK theory. The total number of new light modes is $(q-1)\frac{N^2}{2}$ and so is very large in the large N limit. We have also proposed that the dynamics of these new modes is governed by the sigma action (2.3), with a normalization constant $\mathcal{A}$ whose value we have not been able to calculate.

Assuming that our proposal for the new light modes is correct, it raises several interesting questions. It should be possible to check our proposal for the structure for the effective action (2.3) by performing an independent computation of the four point function of four operators in the theory (2.1) (by summing ladder diagrams) and comparing the long time behaviour of this computation with what one obtains directly from (2.3). Such a procedure should also permit the direct computation of the as yet unknown constant $\mathcal{A}$.

It is also natural to attempt to find a bulk interpretation of our new modes. One natural suggestion is that these modes are dual to gauge fields in AdS_2 [42] If this is the case it is interesting that the rank of the bulk gauge fields diverges in the effectively classical $N \to \infty$ limit. In other words the bulk classical dual of this theory is given in terms of a weakly coupled theory of an *infinite* number of classical fields. The situation is somewhat reminiscent of the proliferation of 'light states' in the duality of [52], and also the situation with ABJ 'triality' in the ABJM limit [53] (although in this context the number of bulk Vasiliev fields is never both parametrically large and parametrically weakly coupled). It would be very interesting to investigate this further.

We have also shown that the density of states in an extreme mass deformation of the model (2.1) displays a faster than Hagedorn growth at energies of order N^2. In our opinion this is also a very striking result; the phase that displays this rapid growth is the 'thermal graviton' or 'string gas' phase. The rapid growth in the density of states of this phase presumably means it cannot thermally equilibriate with another system. It would be interesting to understand what consequences this rapid growth has for potential bulk duals of mass deformed versions of the theory (2.1).

Finally we have performed detailed calculations for the holonomy effective action of the mass deformed theory (2.1) away from the strict large mass limit. In a particular scaling limit that zooms in on the dynamics of the theory at energies of order N^2 we demonstrated that the holonomy effective action of our theory, $S_{\text{eff}}(U)$ takes a simple universal form. We were able to capture the leading interaction effects by summing the appropriate infinite class of graphs and obtain a very simple effective action that captures the leading deviation away from free behaviour. It should certainly be possible to generalize our perturbative

[42] We thank J. Maldacena for this suggestion.

computation of $\widetilde{H}_0$ to a computation of $\widetilde{H}_1$. More ambitiously, it may eventually prove possible to completely sum this perturbative expansion. We leave investigation of this possibility to the future.

Chapter 3

Chern-Simons Matter Theory

3.1 Introduction

There is now considerable evidence that a single fermionic field in the fundamental of $U(N_F)$ minimally coupled to $U(N_F)$ Chern-Simons gauge theory at level[1] k_F is dual to vector $SU(N_B)$ Wilson-Fisher scalars minimally coupled to $SU(N_B)$ Chern-Simons gauge theory at level $k_B = -\text{sgn}(k_F)N_F$ with $N_B = |k_F|$ [54]-[55] [2]. This (almost standard by now) duality asserts that the two so-called *quasi-fermionic* CFTs i.e. Chern-Simons gauged 'regular fermions' (RF) and 'critical bosons' (CB) - are secretly the same theory.

It has also been conjectured (see [56] and references therein) that the 'quasi-fermionic' duality of the previous paragraph follows as the infrared limit of a duality between pairs of fermionic and bosonic RG flows. The fermionic RG flows are obtained by starting in the ultraviolet with the Chern-Simons gauged Gross-Neveu or 'critical fermion' (CF) theory and deforming this theory with relevant operators fine tuned to ensure that the IR end point of the RG flow is the RF theory. In a similar manner the conjecturally dual bosonic flows are obtained by starting in the ultraviolet with the gauged 'regular boson' (RB) theory deformed with the fine tuning that ensures that the RG flow ends in the CB theory.

The UV starting points of the flows described above define dual pairs of conformal field theories. These RB and CF theories - so-called *quasi-bosonic* theories - are conjectured to be dual to each other [3]. If valid, this conjecture implies that the set of all RG flows that originate in the RB theory are dual to the set of all RG flows that originate in the CF theory. The duality of the pair of specially tuned RG flows of the last paragraph- those

[1] In our conventions the level of a Chern-Simons theory coupled to fermions is defined to be the level of the low energy gauge group obtained after deforming the theory with a fermion mass of the same sign as the fermion level.

[2] See the introduction to the recent paper [54] for a more more detailed description of earlier work.

[3] At leading order in large N - the order to which we work in this chapter - the RB and CF theories appear as a line of fixed points parameterized by the single parameter x_6, the coefficient of the ϕ^6 coupling of the bosonic theory (see below for the dual statement in the fermionic theory). In other words the one parameter set of RB and CF theories (and flows originating therein) that we study in this chapter are actually only physical at three particular values of the parameter x_6. See the very recent paper [57] for a computation of the beta function for x_6 that establishes this point.

that end in the IR in the quasi-fermionic conformal field theories - is a special case of this general phenomenon.

Generic RG flows that originate at quasi-bosonic fixed points lead to gapped phases, or more accurately, phases whose low energy behaviour is governed by a topological field theory. There are two inequivalent topological phases. In the unHiggsed phase the bosonic (resp. fermionic) theory is governed at long distances by pure $SU(N_B)_{k_B}$ (resp. $U(N_F)_{k_F}$) topological field theory with the two theories being level-rank dual to each other. In the Higgsed phase the bosonic (resp. fermionic) theory is governed in the IR by a pure $SU(N_B - 1)_{k_B}$ (resp. $U(N_F)_{\text{sgn}\, k_F(|k_F|-1)}$) Chern-Simons theory with the two topological field theories once again being level-rank dual to each other. [4]

The most compelling evidence for the scenarios spelt out above comes from explicit results of direct all-orders calculations that have been performed separately in the fermionic and bosonic theories in the large N limit. In particular, the thermal partition function of deformed RF and CB theories have both been computed in the unHiggsed phases to all orders in the 't Hooft coupling, and have been shown to match exactly with each other for all relevant deformations that end up in this phase [56] [5]. While impressive, this matching is incomplete, as the restriction to the unHiggsed phase covers only half of the phase diagram of these theories.

The authors of [56] (and references therein) were also able to compute the thermal partition function of the CF theory in the 'Higgsed' phase. However, they were unable to perform the analogous computation in the bosonic theory in this phase and so were unable to verify the matching of thermal free energies in this phase. In this chapter we fill the gap described above. We present an explicit all-orders computation of the thermal free energy of the RB theory in the Higgsed phase. Under duality our final results exactly match the free energy of the fermionic theory in the Higgsed phase, completing the large N check of the conjectured duality in a satisfying manner.

At the technical level, the computation described in the previous paragraph (and presented in detail in section 3.3) is a relatively straightforward generalization of the computations presented in the recent paper [54]. In [54] the large N free energy of the Higgsed phase of the Chern-Simons gauged Wilson-Fisher boson theory was computed for the first time. As we describe in much more detail below, the computation of the free energy in the Higgsed phase of the RB theory can be divided into two steps. In the first step we compute the thermal free energy (or equivalently, the gap equation) of the CB theory as a function of its Higgs vev. We are able to import this computation directly from [54]. In the relatively simple second step carried through in this chapter, we derive a second gap

[4] As first explained in [58], the reduction in rank of the bosonic Chern-Simons theory compared to the unHiggsed phase is a consequence of the Higgs mechanism in the bosonic field theory. The reduction in level of the fermionic Chern-Simons theory is a consequence of the switch in sign of the mass of the fermion - level of the pure Chern-Simons theory obtained by integrating out a negative mass fermion is one unit smaller than the level obtained by integrating out a positive mass fermion.

[5] A similar matching has also been performed for the S-matrix in the unHiggsed phase [59, 60]. The generalization of this match to the Higgsed phase is also an interesting project, but one that we will not consider in this chapter.

equation that determines the effective value of the Higgs vev.

The second step described at the end of the last paragraph had no counterpart in [54]. In the critical bosonic theory the 'classical' potential for the scalar field is infinitely deep. This potential freezes the magnitude of the scalar field in the Higgsed phase to its classical minimum even in the quantum theory. It follows that the Higgs vev is independent of the temperature and has a simple dependence on the 't Hooft coupling in the critical boson theory. In the regular boson theory, on the other hand, the classical potential for the scalar field is finite and receives nontrivial quantum corrections. The value of the scalar condensate is determined extremizing the quantum effective action for the scalar field. The result of this minimization yields a scalar vev that is a nontrivial function of both the 't Hooft coupling and the temperature. It follows that the computations of this chapter give us a bonus: we are able to compute the smooth 'quantum effective potential' for the RB theory as a function of the Higgs vev. More precisely we compute the quantum effective potential for the composite field $(\overline{\phi}\phi)$. In the Higgsed phase and in the unitary gauge employed in the computations of this chapter, this quantity reduces to a potential - an exact Landau-Ginzburg effective potential - for the Higgs vev. The extremization of this potential determines the Higgs vev - and is an equivalent and intuitively satisfying way of obtaining the gap equations - in the Higgsed phase. The later sections of this paper - sections 3.4 and 3.5 are devoted to the study of the exact quantum effective potential of the theory and its physical consequences.

Let us denote the expectation value of $(\overline{\phi}\phi)$ by $(\overline{\phi}\phi)_{\rm cl}$. At zero temperature it turns out that the quantum effective action is non-analytic at $(\overline{\phi}\phi)_{\rm cl} = 0$. For this reason the domain of the variable in our effective potential - namely $\overline{\phi}\phi$ - naturally splits into two regions. We refer to the region $\overline{\phi}\phi > 0$ as the Higgsed branch of our effective potential. On the other hand the region $\overline{\phi}\phi < 0$ is the unHiggsed branch of our effective potential. On the Higgsed branch the quantum effective potential for $\overline{\phi}\phi$ is simply a quantum corrected version of the classical potential of the theory. Classically $\overline{\phi}\phi$ always positive, and so the potential for the theory on the unHiggsed branch (i.e. at negative $\overline{\phi}\phi$) has no simple classical limit and is purely quantum in nature. The extremization of the effective potential on the Higgsed/unHiggsed branches exactly reproduces the gap equations in the Higgsed/unHiggsed phases.

In both phases the extrema of this effective potential are of two sorts; local maxima and local minima. Local maxima clearly describe unstable 'phases'. The instability of these phases has an obvious semiclassical explanation in the Higgsed phase; it is a consequence of the fact that we have chosen to expand about a maximum of the potential for the Higgs vev. In this chapter we find an analogous physical explanation for the instability of the 'maxima' in the unHiggsed phase. In Section 3.5.3 we use the results for exact S-matrices in these theories [59, 60] to demonstrate that the 'phases' constructed about maxima in the unHiggsed branch always have bound states of one fundamental particle (created by ϕ) and one antifundamental particle (created by $\overline{\phi}$) in the so called 'singlet' channel. Moreover we demonstrate in Section 3.5.3 that these bound states are always tachyonic (i.e. have negative squared mass). As a consequence, such expansion points are maxima in the potential of the field that creates these bound states (in this case $\overline{\phi}\phi$),

explaining the instability of the corresponding solutions of the gap equation.

It turns out that our effective potential is unbounded from below in the limit $(\overline{\phi}\phi)_{\rm cl} \to +\infty$ when $x_6 < \phi_1$. Here x_6 is the parameter that governs the ϕ^6 interaction of the theory defined precisely in (3.2), and ϕ_1 is a particular function of the 't Hooft coupling λ_B of this theory listed in (3.70). When $(\overline{\phi}\phi)_{\rm cl} \to -\infty$, on the other hand, the potential turns out to be unbounded from below when $x_6 > \phi_2$; ϕ_2 is given in (3.70). It follows that the RB theory is unstable - i.e. does not have a stable vacuum state - if either of the conditions above are met. Happily it turns out that $\phi_1 < \phi_2$ so that there is a range of values for x_6, namely

$$\phi_1 \leq x_6 \leq \phi_2 \ , \tag{3.1}$$

over which the regular boson theory is stable.

The zero temperature phase diagram of the RB theory was worked out in great detail in the recent paper [57]. In order to accomplish this, the authors of [57] evaluated every solution of the gap equation of the RB theory and then compared their free energies. The dominant phase at any given values of microscopic parameters is simply the solution with the lowest free energy; this dominant solution was determined in [57] by performing detailed computations. In Section 3.5.4 of this chapter we demonstrate that the structure of the phase diagrams presented in [57] has a simple intuitive explanation in terms of the exact Landau-Ginzburg effective potential for $(\overline{\phi}\phi)_{\rm cl}$ described above. As we explain in Section 3.5.4 below, the general structure of the phase diagram follows from qualitative curve plotting considerations and can be deduced without performing any detailed computations. Moreover the analysis of the current paper has an added advantage; it allows us to distinguish regions of the phase diagram where the dominant phase is merely metastable (this happens when $x_6 > \phi_2$ or $x_6 < \phi_1$) from regions in the phase diagram in which the dominant solution of the gap equation is truly stable (this happens in the range (3.1)).

3.2 Review of known results

3.2.1 Theories and the conjectured duality map

The RB theory is defined by the action

$$S_B = \int d^3x \Bigg[i\varepsilon^{\mu\nu\rho} \frac{\kappa_B}{4\pi} \mathrm{Tr}(X_\mu \partial_\nu X_\rho - \frac{2i}{3} X_\mu X_\nu X_\rho) + D_\mu \overline{\phi} D^\mu \phi \\ + m_B^2 \overline{\phi}\phi + \frac{4\pi b_4}{\kappa_B} (\overline{\phi}\phi)^2 + \frac{(2\pi)^2}{\kappa_B^2} \left(x_6^B + 1\right) (\overline{\phi}\phi)^3 \Bigg], \tag{3.2}$$

while the ζ_F and ζ_F^2 deformed critical fermion (CF) theory is defined by the Lagrangian

$$S_F = \int d^3x \Bigg[i\varepsilon^{\mu\nu\rho} \frac{\kappa_F}{4\pi} \mathrm{Tr}(X_\mu \partial_\nu X_\rho - \frac{2i}{3} X_\mu X_\nu X_\rho) + \overline{\psi} \gamma_\mu D^\mu \psi \\ - \frac{4\pi}{\kappa_F} \zeta_F (\overline{\psi}\psi - \frac{\kappa_F y_2^2}{4\pi}) - \frac{4\pi y_4}{\kappa_F} \zeta_F^2 + \frac{(2\pi)^2}{\kappa_F^2} x_6^F \zeta_F^3 \Bigg] \ . \tag{3.3}$$

In these formulae

$$\kappa_B = \mathrm{sgn}(k_B)\,(|k_B| + N_B)\,, \quad \kappa_F = \mathrm{sgn}(k_F)\,(|k_F| + N_F)\ . \tag{3.4}$$

The levels k_F and k_B are defined to be the levels of the WZW theory dual to the pure Chern-Simons theory (throughout this chapter we work with the dimensional regularization scheme). For concreteness, in this chapter we will assume that the bosonic theory gauge group is $SU(N_B)$ while the fermionic gauge group is $U(N_F)$ with 'equal' levels k_F for the $SU(N_F)$ and $U(1)$ parts of the gauge group. The generalization to $U(N_B) \leftrightarrow SU(N_F)$ and $U(N_B) \leftrightarrow U(N_F)$ dualities is straightforward at large N and will not be explicitly considered in this chapter.

In the rest of this chapter we will present our formulae in terms of the 't Hooft couplings defined by

$$\lambda_B = \frac{N_B}{\kappa_B}, \quad \lambda_F = \frac{N_F}{\kappa_F}\ . \tag{3.5}$$

We have already mentioned in the introduction that the two theories above have been conjectured to be dual to each other under the level-rank duality map

$$N_B = |\kappa_F| - N_F, \quad \kappa_B = -\kappa_F\ . \tag{3.6}$$

This implies that the bosonic 't Hooft coupling is given in terms of its fermionic counterpart by

$$\lambda_B = \lambda_F - \mathrm{sgn}(\lambda_F), \tag{3.7}$$

The relations (3.6) and (3.7) are expected to hold even at finite N. On the other hand the map between deformations of these two theories is conjectured to be[6]

$$x_6^F = x_6^B\ , \quad y_4 = b_4\ , \quad y_2^2 = m_B^2\ . \tag{3.8}$$

The above equation (3.8) is known to hold only in the large N limit; this relationship may well receive corrections in a power series expansion in $\dfrac{1}{N}$.

To end this subsection, let us note that under the field redefinition

$$\phi = \sqrt{\kappa_B}\,\varphi\ , \tag{3.9}$$

the action (3.2) turns into

$$S_B = \frac{N_B}{\lambda_B}\int d^3x \Big[i\varepsilon^{\mu\nu\rho}\frac{1}{4\pi}\mathrm{Tr}(X_\mu\partial_\nu X_\rho - \frac{2i}{3}X_\mu X_\nu X_\rho) + D_\mu\overline{\varphi}D^\mu\varphi \\ + m_B^2\overline{\varphi}\varphi + 4\pi b_4(\overline{\varphi}\varphi)^2 + (2\pi)^2\left(x_6^B+1\right)(\overline{\varphi}\varphi)^3\Big]\ . \tag{3.10}$$

It follows immediately that in the limit

$$\lambda_B \to 0\ , \quad m_B^2,\ b_4,\ x_6 = \text{fixed}\ , \tag{3.11}$$

the theory (3.2) should reduce to a nonlinear but classical theory of the fields φ and X_μ. We will return to this point below.

[6]Since $x_6^B = x_6^F$, we drop the superscript ^{B} or ^{F} on x_6 often in the paper when referring to this coupling.

3.2.2 Structure of the thermal partition function

As explained in e.g. [54], the large N thermal free energy of either of these theories on $S^2 \times S^1$ can be obtained following a two step process. In the first step we compute the free energy of the theory in question on $\mathbb{R}^2 \times S^1$, at a fixed value of the gauge holonomies around S^1. The result is a functional of the holonomy eigenvalue distribution function $\rho(\alpha)$ and is given by the schematic equation

$$e^{-\mathcal{V}_2 T^2 v[\rho]} = \int_{\mathbb{R}^2 \times S^1} [d\phi]\ e^{-S[\phi,\rho]}\ . \tag{3.12}$$

where $\mathcal{V}_2$ is the volume of two dimensional space and T is the temperature.

In order to complete the evaluation of the $S^2 \times S^1$ partition function of interest, in the next step we are instructed to evaluate the unitary matrix integral

$$\mathcal{Z}_{S^2 \times S^1} = \int [dU]_{\rm CS}\ e^{-\mathcal{V}_2 T^2 v[\rho]}. \tag{3.13}$$

where $[dU]_{\rm CS}$ is the Chern-Simons modified Haar measure over $U(N)$ described in [61].

It was demonstrated in [61] that the thermal partition functions (3.13) of the bosonic and fermionic theories agree with each other in the large N limit provided that under duality

$$v_F[\rho_F] = v_B[\rho_B]\ , \tag{3.14}$$

where the bosonic and fermionic eigenvalue distribution functions, ρ_B and ρ_F, are related via

$$|\lambda_B|\rho_B(\alpha) + |\lambda_F|\rho_F(\pi - \alpha) = \frac{1}{2\pi}. \tag{3.15}$$

In this chapter we will evaluate the free energy $v_B[\rho_B]$ of the bosonic theory in the Higgsed phase and verify (3.14), thus establishing the equality of thermal free energies of the RB and CF theories in the Higgsed phase. We summarize the map between the parameters (3.6), (3.7), (3.8) and the holonomy distributions (3.15):

$$\begin{aligned} &N_F = |\kappa_B| - N_B\ , \quad \kappa_F = -\kappa_B\ , \quad \lambda_F = \lambda_B - \mathrm{sgn}(\lambda_B)\ , \\ &x_6^F = x_6^B\ , \quad y_4 = b_4\ , \quad y_2^2 = m_B^2\ , \quad |\lambda_B|\rho_B(\alpha) + |\lambda_F|\rho_F(\pi - \alpha) = \frac{1}{2\pi}. \end{aligned} \tag{3.16}$$

In Appendix B.1 we provide a comprehensive review of everything that is known about the large N thermal free energies of the CF and RB theories (the appendix also contains a formula for a 'three variable off-shell' free energy functional of the CF theory that is valid in both phases (B.3)). In the rest of this section we only present those results that will be of relevance for the computations in the paper.

The free energy $v_F[\rho_F]$ in the critical fermion theory has been computed in both fermionic phases in [56]. The result is given in terms of an auxiliary off-shell free energy[7]

[7] We put a hat over a particular quantity (e.g. $\widehat{c}_F$) to denote the dimensionless version of that quantity (e.g. c_F) obtained by multiplying by appropriate powers of the temperature T.

(equation (B.12) in Appendix B.1)

$$F_F(c_F,\zeta_F) = \frac{N_F}{6\pi}\Bigg[\frac{|\lambda_F| - \text{sgn}(\lambda_F)\text{sgn}(X_F)}{|\lambda_F|}\widehat{c}_F^3 - \frac{3}{2\lambda_F}\left(\frac{4\pi}{\kappa_F}\widehat{\zeta}_F\right)\widehat{c}_F^2 + \frac{1}{2\lambda_F}\left(\frac{4\pi}{\kappa_F}\widehat{\zeta}_F\right)^3 + \frac{6\pi\widehat{y}_2^2}{\kappa_F\lambda_F}\widehat{\zeta}_F - \frac{24\pi^2\widehat{y}_4}{\kappa_F^2\lambda_F}\widehat{\zeta}_F^2 + \frac{24\pi^3 x_6^F}{\kappa_F^3\lambda_F}\widehat{\zeta}_F^3 - 3\int_{-\pi}^{\pi} d\alpha\rho_F(\alpha)\int_{\widehat{c}_F}^{\infty} dy\ y\ \left(\log\left(1+e^{-y-i\alpha}\right)+\log\left(1+e^{-y+i\alpha}\right)\right)\Bigg]. \tag{3.17}$$

The above free energy is a function of two variables c_F and ζ_F. Extremizing F_F with respect to these variables and plugging back in the extremum values gives us the free energy $v_F[\rho_F]$. The physical interpretation of the variable c_F is that its value at the extremum of F_F coincides with the pole mass of the fermion.

The free energy (3.17) assumes two different analytic expressions depending on the sign $\text{sgn}(\lambda_F)\text{sgn}(X_F)$ and governs the dynamics of the two different phases. The phase in which $\text{sgn}(X_F)\text{sgn}(\lambda_F) = \pm 1$ is referred to as the unHiggsed phase and the Higgsed phase respectively. In equation (B.3) in Appendix B.1, we give an off-shell free energy in terms of three variables (which include c_F and ζ_F) which is analytic in all three variables and encompasses the behaviour of both phases.

The free energy (3.17) in the unHiggsed phase of the CF theory matches the free energy of the regular boson theory in the unHiggsed phase (equation (B.13) in Appendix B.1) computed in [56] under the duality map (3.16). The free energy (3.17) with $\text{sgn}(\lambda_F)\text{sgn}(X_F) = -1$ gives a prediction for the regular boson theory in the Higgsed phase. Applying the duality transformation (3.16) and making the following 'field' redefinitions:

$$c_F = c_B\ , \quad \frac{4\pi\zeta_F}{\kappa_F} = -2\lambda_B\sigma_B\ , \tag{3.18}$$

we get the following prediction for the free energy in the Higgsed phase (equation (B.28)):

$$F_B(c_B,\sigma_B) = \frac{N_B}{6\pi}\Bigg[-\frac{\lambda_B - 2\text{sgn}(\lambda_B)}{\lambda_B}\widehat{c}_B^3 - 3\widehat{\sigma}_B(\widehat{c}_B^2 - \widehat{m}_B^2) + 6\widehat{b}_4\lambda_B\widehat{\sigma}_B^2 + (3x_6^B+4)\lambda_B^2\widehat{\sigma}_B^3 + 3\int_{-\pi}^{\pi} d\alpha\rho_B(\alpha)\int_{\widehat{c}_B}^{\infty} dy\ y\left(\log\left(1-e^{-y-i\alpha}\right)+\log\left(1-e^{-y+i\alpha}\right)\right)\Bigg]. \tag{3.19}$$

The extremum value of the variable c_B corresponds to the pole mass of the W boson excitation in the Higgsed phase. In the next section we will independently compute the off-shell free energy of the RB theory, and demonstrate that our answer agrees with (3.19) once we identify the field σ_B with

$$\sigma_B = 2\pi\frac{\overline{\phi}\phi}{N_B}\ . \tag{3.20}$$

where $\overline{\phi}$ and ϕ respectively stand for the saddle point values of the corresponding fields denoted by the same letters (recall these fields have nonzero saddle point values in the Higgsed phase).

Note: Here and in the rest of the paper, we define the quantities c_F and c_B to be always positive. In other words, $c_{F,B}$ is shorthand for $|c_{F,B}|$. This is the same convention used in [56].

3.3 The Higgsed Phase of the regular boson theory

3.3.1 Lagrangian in Unitary gauge

Consider the following action for the $SU(N_B)$ regular boson theory:

$$\begin{aligned} S_{\rm E} = \int d^3x \Big[\, i\epsilon^{\mu\nu\rho}\frac{\kappa_B}{4\pi}\,{\rm Tr}(X_\mu\partial_\nu X_\rho - \frac{2i}{3}X_\mu X_\nu X_\rho) + D_\mu\overline{\phi}D^\mu\phi \\ + m_B^2\overline{\phi}\phi + \frac{4\pi b_4}{\kappa_B}(\overline{\phi}\phi)^2 + \frac{(2\pi)^2}{\kappa_B^2}(x_6^B+1)(\overline{\phi}\phi)^3\Big] \, , \end{aligned} \tag{3.21}$$

with $D_\mu = \partial_\mu - iX_\mu$. The above action can be reorganized as follows in the Higgsed phase where we anticipate $\langle\overline{\phi}\phi\rangle \neq 0$. Following [54] we work in the unitary gauge

$$\phi^i(x) = \delta^{iN_B}\sqrt{|\kappa_B|}\,V(x) \, . \tag{3.22}$$

For future reference we note also that (3.22) implies the following for the 'classical' field φ defined in (3.9):

$$\varphi^i(x) = \delta^{iN_B}\,\sqrt{{\rm sgn}(\kappa_B)}\,V(x) \, . \tag{3.23}$$

The field $V(x)$ shall be termed the *Higgs field*. The above gauge choice lets us decompose the gauge field X_μ as

$$X_\mu = \begin{pmatrix} (A_\mu)^a{}_b - \frac{\delta^a{}_b}{N_B-1}Z_\mu & \frac{1}{\sqrt{\kappa_B}}(W_\mu)^a \\ \frac{1}{\sqrt{\kappa_B}}(\overline{W}_\mu)_b & Z_\mu \end{pmatrix} \, , \tag{3.24}$$

where the indices a, b run over $1, \ldots, N_B - 1$. In terms of these variables, the action can be rewritten as follows[8]:

$$\begin{aligned} S_{\rm E}[A,W,Z,V] = \frac{i\kappa_B}{4\pi}\int {\rm Tr}(AdA - \tfrac{2i}{3}AAA) \\ + \frac{i}{4\pi}\int \left(2\overline{W}_a dW^a + \kappa_B Z dZ - 2iZ\overline{W}_a W^a - 2i\overline{W}_a A^a{}_b W^b\right) \\ + \int d^3x\,(|\kappa_B|V^2 Z_\mu Z^\mu + {\rm sgn}(\kappa_B)V^2\overline{W}_{a\mu}W^{a\mu}) \\ + |\kappa_B|\int d^3x\left(\partial_\mu V\partial^\mu V + m_B^2V^2 + 4\pi b_4{\rm sgn}(\kappa_B)V^4 + 4\pi^2(x_6^B+1)V^6\right) \, . \end{aligned} \tag{3.25}$$

[8]The notation ABC stands for $d^3x\epsilon^{\mu\nu\rho}A_\mu B_\nu C_\rho$.

3.3.2 An effective action for the Higgs field V

We will now compute the thermal partition function of the regular boson theory in the Higgsed phase, i.e. we will compute $v_B[\rho_B]$ defined by

$$e^{-\mathcal{V}_2 T^2 v_B[\rho_B]} = \int_{\mathbb{R}^2\times S^1} [dV dW dZ dA]\; e^{-S_{\rm E}[A,W,Z,V]} \;, \tag{3.26}$$

where $S_{\rm E}[A,W,Z,V]$ was defined in (3.25). For this purpose it is convenient to break up the effective action $S_{\rm E}[A,W,Z,V]$ into two parts

$$S_{\rm E}[A,W,Z,V] = S_1[A,W,Z,V] + S_2[V] \tag{3.27}$$

where

$$\begin{aligned} S_1[A,W,Z,V] = \frac{i\kappa_B}{4\pi}\int d^3x\; \mathrm{Tr}(AdA - \tfrac{2i}{3}AAA) \\ + \frac{i}{4\pi}\int \left(2\overline{W}_a dW^a + \kappa_B Z dZ - 2iZ\overline{W}_a W^a - 2i\overline{W}_a A^a{}_b W^b\right) \\ + \int d^3x\, \left(|\kappa_B| V^2 Z_\mu Z^\mu + \mathrm{sgn}(\kappa_B) V^2 \overline{W}^\mu_a W^a_\mu\right) \;, \end{aligned} \tag{3.28}$$

and

$$\begin{aligned} S_2[V] = \int d^3x \Big(|\kappa_B|\partial_\mu V \partial^\mu V + U_{\rm cl}(V)\Big) \;, \\ U_{\rm cl}(V) = |\kappa_B| m_B^2 V^2 + 4\pi b_4 \kappa_B V^4 + 4\pi^2 |\kappa_B| (x_6^B + 1) V^6 \;. \end{aligned} \tag{3.29}$$

The path integral (3.12) can be rewritten as

$$e^{-\mathcal{V}_2 T^2 v_B[\rho_B]} = \int [dV]\, e^{-S_2[V]} \int [dW dZ dA]\, e^{-S_1[A,W,Z,V]} \tag{3.30}$$

Let us first study 'inner' path integral i.e.

$$e^{-\mathcal{V}_2 T^2 v_{\rm CB}[\rho_B,V]} \equiv \int [dW dZ dA]\, e^{-S_1[A,W,Z,V]} \;, \tag{3.31}$$

where the right hand side defines the quantity $v_{\rm CB}[\rho_B, V]$. As far as the path integral in (3.31) is concerned, $V(x)$ is a background field. The path integral (3.31) is difficult to evaluate for arbitrary $V(x)$ even in the large N_B limit[9]. This problem simplifies, however, in the special case that $V(x)$ is a constant. In fact, precisely in this limit, the path integral (3.31) has been evaluated in the recent paper [54]. Luckily, it will turn out that, in the large N limit, the integral over $V(x)$ in (3.30) localises to a saddle point at which $V(x)$ is constant (see below). As a consequence we only need the result of the path integral (3.31)

[9] (3.31) is effectively the generating function of all correlation functions of the dimension two scalar J_0 in the large N critical boson theory, and so contains a great deal of information.

for constant $V(x)$; we are able to read off this result directly from [54] which we now pause to recall.

The authors of [54] studied the critical boson theory in its Higgsed phase. Working in unitary gauge and following manipulations essentially identical to those outlined in the previous subsection, they found that the CB theory in the Higgsed phase can be rewritten as effective theory of interacting massive W bosons, Z bosons and $SU(N_B-1)$ gauge fields, whose action is given by

$$\begin{aligned} S_{\rm E}[A,W,Z] = \frac{i\kappa_B}{4\pi}\int {\rm Tr}(AdA - \tfrac{2i}{3}AAA) \\ + \frac{i}{4\pi}\int \left(2\overline{W}_a dW^a + \kappa_B Z dZ - 2iZ\overline{W}_a W^a - 2i\overline{W}_a A^a{}_b W^b\right) \\ - \int d^3x \left(\frac{N_B}{4\pi} m_B^{\rm cri} Z_\mu Z^\mu + \frac{\lambda_B}{4\pi} m_B^{\rm cri}\overline{W}_{a\mu} W^{a\mu}\right) . \end{aligned} \tag{3.32}$$

The authors of [54] were then able to evaluate the finite temperature partition function for the theory defined by (3.32). Their final result for $v_{\rm CB}[\rho_B]$ is given as follows. One obtains $v_{\rm CB}[\rho_B]$ by extremizing an off-shell free energy $F_{\rm CB}(c_B)$ given by

$$\begin{aligned} F_{\rm CB}(c_B) = \frac{N_B}{6\pi}\Bigg[- \frac{\lambda_B - 2{\rm sgn}(\lambda_B)}{\lambda_B}\widehat{c}_B^3 + \frac{3}{2}\widehat{m}_B^{\rm cri}\widehat{c}_B^2 + \Lambda\left(\widehat{m}_B^{\rm cri}\right)^3 \\ + 3\int_{-\pi}^{\pi} d\alpha \rho_B(\alpha)\int_{\widehat{c}_B}^{\infty} dy\, y\left(\log\left(1-e^{-y-i\alpha}\right) + \log\left(1-e^{-y+i\alpha}\right)\right)\Bigg], \end{aligned} \tag{3.33}$$

Here, Λ is an undetermined constant; shifts in Λ correspond to shifts in the cosmological constant counterterm in the starting action for the CB theory (see [54] for a discussion).

Note that the action S_1 in (3.25) agrees precisely with the action (3.32) reported in [54] if we replace $m_B^{\rm cri}$ by the quantity

$$m_B^{\rm cri} = -\frac{4\pi}{|\lambda_B|}V^2 \quad \text{with } V \text{ constant.} \tag{3.34}$$

It follows that for the special case that $V(x)$ is constant, the path integral (3.31) is given by the extremum value of (3.33) with the replacement (3.34) and the path integral over V in (3.30) takes the form

$$\int [dV] e^{-S_{\rm eff}[V]} \quad \text{with} \quad S_{\rm eff}[V] = S_2[V] + \mathcal{V}_2 T^2 v_{\rm CB}[\rho_B, V] . \tag{3.35}$$

From the expressions for $S_2[V]$,[10] and $F_B(c_B)$ in (3.29) and (3.33), it is clear that there is an overall factor of N_B in front of the effective action $S_{\rm eff}[V]$. In the large N_B limit the

[10] In [54], the free energy $v_{\rm CB}[\rho_B]$ depended on $m_B^{\rm cri}$ which was a parameter in the theory. After the replacement (3.34), the dependence on the parameter $m_B^{\rm cri}$ is replaced by a dependence on the field $V(x)$. We have included an explicit V in the notation for $v_{\rm CB}[\rho_B, V]$ to highlight this dependence on the Higgs field V.

path integral over V may be evaluated in the saddle-point approximation. We expect the dominant minima of the effective action to occur at constant values of V since the kinetic term $\partial_\mu V \partial^\mu V$ adds a positive definite piece to the action. For this reason it is sufficient to have the expression $v_{\rm CB}[\rho_B, V]$ only at constant V. As we have already explained above, this result is given by extremizing (3.33) w.r.t. c_B after making the replacement (3.34). It follows that the final result for $v_B[\rho_B]$ in (3.26) is obtained by extremizing the regular boson off-shell free energy

$$F_B(c_B, V) = F_{\rm CB}(c_B) + \frac{1}{\mathcal{V}_2 T^2} S_2[V] \ ,$$

with respect to both c_B *and* V.[11] Using the explicit expressions (3.29), (3.33) and (3.34) we find the following explicit result for the off-shell free energy of the RB theory:

$$\begin{aligned} F_B(c_B, V) = \frac{N_B}{6\pi} \Bigg[& - \frac{(\lambda_B - 2{\rm sgn}(\lambda_B))}{\lambda_B} \widehat{c}_B^3 - \frac{8\Lambda}{|\lambda_B|^3} (2\pi \widehat{V}^2)^3 \\ & - \frac{3}{|\lambda_B|} (\widehat{c}_B^2 - \widehat{m}_B^2) 2\pi \widehat{V}^2 + \frac{6\widehat{b}_4}{\lambda_B} (2\pi \widehat{V}^2)^2 + \frac{(3x_6^B + 3)}{|\lambda_B|} (2\pi \widehat{V}^2)^3 \\ & + 3 \int_{-\pi}^{\pi} \rho_B(\alpha) d\alpha \int_{\widehat{c}_B}^{\infty} dy y \left(\log\left(1 - e^{-y - i\alpha}\right) + \log\left(1 - e^{-y + i\alpha}\right) \right) \Bigg] \ . \end{aligned} \tag{3.36}$$

We compare the result (3.36) with the prediction (3.19) of duality for the Higgsed phase free energy given by

$$\begin{aligned} F_B(c_B, V) = \frac{N_B}{6\pi} \Bigg[& - \frac{\lambda_B - 2{\rm sgn}(\lambda_B)}{\lambda_B} \widehat{c}_B^3 \\ & - \frac{3}{|\lambda_B|} (\widehat{c}_B^2 - \widehat{m}_B^2) 2\pi \widehat{V}^2 + \frac{6\widehat{b}_4}{\lambda_B} (2\pi \widehat{V}^2)^2 + \frac{(3x_6^B + 4)}{|\lambda_B|} (2\pi \widehat{V}^2)^3 \\ & + 3 \int_{-\pi}^{\pi} d\alpha \rho_B(\alpha) \int_{\widehat{c}_B}^{\infty} dy \ y \ \left(\log\left(1 - e^{-y - i\alpha}\right) + \log\left(1 - e^{-y + i\alpha}\right) \right) \Bigg] \ , \end{aligned} \tag{3.37}$$

where, to get the above expression, we have used (3.22) and (3.20) to write the field σ_B in (3.19) as

$$\sigma_B = \frac{2\pi V^2}{|\lambda_B|} \ . \tag{3.38}$$

We see that (3.36) agrees precisely with (3.37) provided we choose the as yet undetermined parameter Λ as

$$\Lambda = -\frac{1}{8} \lambda_B^2 \ . \tag{3.39}$$

[11] It is understood that $F_{\rm CB}[\rho_B, c_B]$ is evaluated after making the replacement (3.34).

In the next subsection we will verify that the result (3.39) - which is so far just a prediction of duality - can also be obtained by direct computation within the Higgsed CB theory. The strategy we employ is the following. We first note that the gap equation corresponding to stationarity of (3.36) with respect to $\widehat{V}^2$ is

$$\widehat{c}_B^2 - \widehat{m}_B^2 + 8\Lambda\widehat{\sigma}_B^2 - 4\widehat{b}_4\lambda_B\widehat{\sigma}_B - (3x_6^B + 3)\lambda_B^2\widehat{\sigma}_B^2 = 0 \ . \tag{3.40}$$

where σ_B is given in terms of V^2 by (3.38). The equation (3.40) merely simply expresses the condition that the tadpole of the fluctuation of the scalar field V vanishes when the field V is expanded around its true solution v. In the next subsection we directly evaluate this 'tadpole vanishing condition' in the RB theory in the Higgsed phase and thereby determine Λ by comparison with (3.40).

3.3.3 Tadpole cancellation for V

As we have explained above, in the Higgsed phase our scalar field V gets the expectation value v. It is useful to define

$$V(x) = v + H(x) \ . \tag{3.41}$$

The condition that v is the correct vacuum expectation value of $V(x)$ is equivalent to the condition that the expectation value (i.e. one point function, i.e. tadpole) of the fluctuation $H(x)$ vanishes. In other words we require that

$$\int_{\mathbb{R}^2\times S^1} [dVdWdZdA]\ H(x)\ e^{-S_{\rm E}[A,W,Z,V]} = 0 \ . \tag{3.42}$$

Using the explicit form of $S_{\rm E}[A,W,Z,V]$ in (3.25), equation (3.42) can be rewritten as

$$\mathrm{sgn}(\kappa_B)\langle\overline{W}_{a\mu}(x)W^{a\mu}(x)\rangle + |\kappa_B|\langle Z_\mu(x)Z^\mu(x)\rangle \ + \frac{\partial}{\partial(v^2)}U_{\rm cl}(v^2) = 0 \ . \tag{3.43}$$

where $U_{\rm cl}(V)$ is the potential for the Higgs field V given in (3.29) and all expectation values are evaluated about the 'vacuum' where $V(x) = v$. While the first and third terms in (3.43) above are both of order N_B, it is easily verified that the second term in this equation - the term proportional to $\langle Z_\mu(x)^2\rangle$ is of order unity[12] and so can be dropped in the large N_B limit. At leading order in the large N_B limit, it follows that the tadpole cancellation condition (3.42) can be rewritten as

$$\frac{\lambda_B}{2\pi\mathcal{V}_3}\int d^3x\langle\overline{W}_{a\mu}(x)W^{a\mu}(x)\rangle + \frac{\partial U_{\rm cl}(\sigma_B)}{\partial\sigma_B} = 0 \ , \tag{3.44}$$

where we have integrated the equation (3.43) over spacetime and have divided the resulting expression by the volume of spacetime $\mathcal{V}_3$. We have also changed variables to $\sigma_B = 2\pi v^2/|\lambda_B|$ defined in (3.38). Moving to momentum space, (3.44) turns into

$$\frac{\lambda_B}{2\pi}\int\frac{\mathcal{D}^3p}{(2\pi)^3}\eta^{\mu\nu}\,G^a_{a\mu\nu}(p) + \frac{\partial U_{\rm cl}(\sigma_B)}{\partial\sigma_B} = 0 \ , \tag{3.45}$$

[12]This follows from the observation that the Z propagator scales like $1/N_B$.

where

$$\langle \overline{W}_{a\mu}(-p) W^b_\nu(p') \rangle = G^b_{a\mu\nu}(p)\,(2\pi)^3\delta^{(3)}(p-p') = \delta_a{}^b\; G_{\mu\nu}(p)\,(2\pi)^3\delta^{(3)}(p-p')\ , \qquad (3.46)$$

and the measure $\mathcal{D}$ is the natural measure in momentum space at finite temperature.[13] Happily, the exact all-orders formula for the propagator $G_{\mu\nu}$ was computed in [54]. In Appendix B.2 we proceed to plug the explicit expression for $G_{\mu\nu}$ and evaluate the first term in (3.45). We are able to evaluate all the relevant summations and integrals analytically, and demonstrate that in our choice of regularization scheme (3.44) takes the explicit form

$$-\frac{N_B}{2\pi}\left(c_B^2 - \lambda_B^2\sigma_B^2\right) + \frac{\partial U_{\rm cl}(\sigma_B)}{\partial\sigma_B} = 0\ . \qquad (3.48)$$

Recall the expression for $U(\sigma_B)$ from (3.29):

$$U_{\rm cl}(\sigma_B) = \frac{N_B}{2\pi}\left(m_B^2\sigma_B + 2b_4\lambda_B\sigma_B^2 + (x_6^B+1)\lambda_B^2\sigma_B^3\right)\ . \qquad (3.49)$$

Plugging this back into (3.48) we find

$$c_B^2 - m_B^2 - 4b_4\lambda_B\sigma_B - (3x_6^B+4)\lambda_B^2\sigma_B^2 = 0\ , \qquad (3.50)$$

Comparing this with the gap equation obtained earlier in (3.40), we see that (3.40) matches (3.50) for the predicted value of $\Lambda = -\lambda_B^2/8$ in (3.39) as expected.

3.4 A three variable off-shell free energy

The finite temperature unHiggsed phase is governed by the two-variable off-shell free energy (equation (B.14) in Appendix B.1)

$$F_B(c_B,\widetilde{\mathcal{S}}) = \frac{N_B}{6\pi}\Bigg[-\widehat{c}_B^3 + 3\widetilde{\mathcal{S}}\left(\widehat{c}_B^2 - \widehat{m}_B^2\right) + 6\widehat{b}_4\lambda_B\widetilde{\mathcal{S}}^2 - (4+3x_6^B)\lambda_B^2\widetilde{\mathcal{S}}^3$$
$$+ 3\int_{-\pi}^{\pi} d\alpha\rho_B(\alpha)\int_{\widehat{c}_B}^{\infty} dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right) + \log\left(1-e^{-y+i\alpha}\right)\right)\Bigg]\ , \qquad (3.51)$$

On the other hand, we have demonstrated in this chapter that the finite temperature Higgsed phase is governed by the two-variable off-shell free energy (3.19). As these are

[13] Explicitly, the notation $\mathcal{D}^3p$ signifies that we work at finite temperature i.e. on the spacetime $\mathbb{R}^2\times S^1_\beta$ where the third direction x^3 is a circle of circumference β. The measure $\mathcal{D}^3p$ is then given by

$$\int\frac{\mathcal{D}^3p}{(2\pi)^3}f(p) = \int\frac{dp_1dp_2}{(2\pi)^2}\int_{-\pi}^{\pi}d\alpha\,\rho_B(\alpha)\frac{1}{\beta}\sum_{n=-\infty}^{\infty} f\left(\frac{2\pi n+\alpha}{\beta}\right)\ , \qquad (3.47)$$

where $\rho_B(\alpha)$ is the distribution of the eigenvalues α of the gauge field holonomy around S^1_β.

two separate 'phases' of the same theory it is somewhat unsatisfying that the off-shell 'Landau-Ginzburg' free energies used to describe them are different. The reader may wonder whether there exists a single master off-shell free energy functional - analytic in all 'fields' - which encompasses the physics of both (3.51) and (3.19). At least at the algebraic level there is a simple affirmative answer to this question as we now describe.

Consider the off-shell free energy

$$F(c_B,\sigma_B,\widetilde{\mathcal{S}}) = \frac{N_B}{6\pi}\Bigg[-\widehat{c}_B^3 - 4\widetilde{\mathcal{S}}^3\lambda_B^2 - 3\widehat{c}_B^2\widehat{\sigma}_B - 12\widetilde{\mathcal{S}}^2\lambda_B^2\widehat{\sigma}_B - 12\widetilde{\mathcal{S}}\lambda_B^2\widehat{\sigma}_B^2 + 6\widehat{c}_B|\lambda_B|(\widetilde{\mathcal{S}}+\widehat{\sigma}_B)^2 + 3\left(\widehat{m}_B^2\widehat{\sigma}_B + 2\lambda_B\widehat{b}_4\widehat{\sigma}_B^2 + \lambda_B^2 x_6^B\widehat{\sigma}_B^3\right) + 3\int_{-\pi}^{\pi} d\alpha\rho_B(\alpha)\int_{\widehat{c}_B}^{\infty} dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right)+\log\left(1-e^{-y+i\alpha}\right)\right)\Bigg]. \tag{3.52}$$

Note that (3.52) is a function of three 'field' variables, namely c_B, $\widetilde{\mathcal{S}}$ and σ_B. Extremizing (3.52) w.r.t. $\widetilde{\mathcal{S}}$, c_B and σ_B respectively yields the equations

$$\begin{aligned}
&(\widetilde{\mathcal{S}}+\widehat{\sigma}_B)(-\widehat{c}_B+|\lambda_B|(\widetilde{\mathcal{S}}+\widehat{\sigma}_B)) = 0\ ,\\
&\widehat{c}_B(\mathcal{S}(c_B)+\widehat{\sigma}_B) - |\lambda_B|(\widetilde{\mathcal{S}}+\widehat{\sigma}_B)^2 = 0\ ,\\
&\widehat{c}_B^2 - \widehat{m}_B^2 - 4\widehat{c}_B|\lambda_B|(\widetilde{\mathcal{S}}+\widehat{\sigma}_B) + \lambda_B\left(4\widetilde{\mathcal{S}}^2\lambda_B - 4\widehat{b}_4\widehat{\sigma}_B + 8\lambda_B\widehat{\sigma}_B\widetilde{\mathcal{S}} - 3\lambda_B\widehat{\sigma}_B^2 x_6^B\right) = 0\ .
\end{aligned} \tag{3.53}$$

The quantity $\mathcal{S}(c_B)$ that appears in the second of (3.53) is defined in (B.1). Off-shell, the objects $\mathcal{S}(c_B)$ and $\widetilde{\mathcal{S}}$ are completely distinct. $\mathcal{S}(c_B)$ is a function of c_B while $\widetilde{\mathcal{S}}$ is an independent variable. However it is easy to see (by subtracting the first two equations in (3.53)) that these two quantities are, in fact, equal on-shell.

Note in particular that the first of (3.53) - the equation that follows upon extremizing (3.52) w.r.t. $\widetilde{\mathcal{S}}$ - is the product of two factors. This equation is satisfied either if

$$\widetilde{\mathcal{S}}+\widehat{\sigma}_B = 0\ , \tag{3.54}$$

or if

$$-\widehat{c}_B+|\lambda_B|(\widetilde{\mathcal{S}}+\widehat{\sigma}_B) = 0\ . \tag{3.55}$$

(clearly (3.54) and (3.55) cannot simultaneously be obeyed unless $c_B = 0$). Let us first suppose that (3.54) is obeyed. Using (3.54) to eliminate σ_B from (3.52) yields an off-shell free energy that now depends only on $\widetilde{\mathcal{S}}$ and c_B. It is easily verified that the resultant free energy agrees exactly with the two-variable free energy (3.51) in the unHiggsed phase. It follows that solutions of (3.54) parameterize - and govern the physics of - the unHiggsed phase of the RB theory.

In a similar manner let us now suppose that (3.55) is obeyed in which case we use it to eliminate $\widetilde{\mathcal{S}}$. It is easily verified that the resultant two-variable free energy - which depends

on c_B and σ_B - agrees exactly with (3.37) with the identification (3.38):

$$\sigma_B = \frac{2\pi v^2}{|\lambda_B|} \ . \tag{3.56}$$

It follows that solutions of (3.55) parameterize - and govern the physics of - the Higgsed phase of the RB theory.

The identification (4.84) has a simple explanation. Recall that the bare mass m_B^2 appeared in the action (3.2) as the coefficient of $\overline{\phi}\phi$. It follows that the Legendre transform of the free energy of our theory w.r.t. m_B^2 yields the exact quantum corrected effective potential of our theory as a function of the composite field $(\overline{\phi}\phi)_{\rm cl}$. This Legendre transform may be computed by adding the term

$$-m_B^2(\overline{\phi}\phi)_{\rm cl}$$

to (3.52) and then treating m_B^2 as a new dynamical field w.r.t. which (3.52) has to be extremized (of course we also continue to extremize (3.52) w.r.t. c_B, σ_B and $\widetilde{\mathcal{S}}$ as before). Note that the dependence of (3.52) on m_B^2 is extremely simple; it occurs entirely through the term $\frac{N_B}{2\pi}\sigma_B m_B^2$. As a consequence, extremizing w.r.t. m_B^2 sets

$$\sigma_B = \frac{2\pi(\overline{\phi}\phi)_{\rm cl}}{N_B} \ . \tag{3.57}$$

In the Higgsed phase it follows from $\phi^i = \delta^{iN_B}\sqrt{|\kappa_B|}\,v$ (equation (3.22)) that

$$(\overline{\phi}\phi)_{\rm cl} = |\kappa_B| v^2 \ . \tag{3.58}$$

(in obtaining (3.58) we use the fact that the Higgs field V is effectively classical in the large N_B limit). Inserting (3.58) into (3.57) yields (4.84). We note, however, that (3.57) is more general than (4.84) because it applies even in the unHiggsed phase. We will make use of this fact in the next section.

We have thus found a simple single off-shell free energy - namely (3.52) - that captures the physics of both the Higgsed and the unHiggsed phases. We have also explained that one of the three variables that appears in this free energy - namely σ_B - has a simple direct physical interpretation given by (3.57). It follows, in particular, that if we integrate c_B and $\widetilde{\mathcal{S}}$ out from (3.52), the resultant free energy (which is a function of σ_B) can be reinterpreted as the quantum effective potential of the theory as a function of $(\overline{\phi}\phi)_{\rm cl}$. In the next section we will explicitly undertake this exercise in the zero temperature limit.

It is easily verified that the duality map (3.16) between parameters together with the field redefinitions

$$\lambda_B\widetilde{\mathcal{S}} = \lambda_F\widetilde{\mathcal{C}} - \frac{\mathrm{sgn}(\lambda_F)}{2}\widehat{c}_F \ , \quad \lambda_B\sigma_B = -\frac{2\pi\zeta_F}{\kappa_F} \ , \quad c_B = c_F \ . \tag{3.59}$$

turns the bosonic off-shell free energy (3.52) into the fermionic off-shell free energy (B.3). This match captures the Bose Fermi duality between RB and CF theories at the level of the

complete thermal off-shell free energies of the two theories; note that each of these off-shell free energies is analytic in all 'fields'. In Appendix B.3 we investigate the behaviour of our three-variable off-shell free energy in the so called critical boson scaling limit of the RB theory.

3.5 The exact Landau-Ginzburg effective potential

In this section we integrate out the variables $\widetilde{\mathcal{S}}$ and c_B out from the effective action (3.51) and obtain an off-shell free energy for the field σ_B. We work at zero temperature throughout this section. In this simple - and physically especially important - limit we obtain a simple analytic expression for the resultant free energy as a function of σ_B. As we have explained in the previous section, this free energy is simply related to the quantum effective potential of the RB theory as a function of the field $(\overline{\phi}\phi)_{\rm cl}$.

After having obtained this exact Landau-Ginzburg potential we study and use it in various ways. First we note that this effective potential has extrema of two sorts - local maxima and local minima. Local maxima represent unstable saddle point solutions of the theory. In the case of the unHiggsed branch (see below) we present an interpretation of the resultant instability in terms of the tachyonic bound states of the system. We also use the exact Landau-Ginzburg effective action that we obtain to understand the zero temperature phase diagram of the RB theory (as a function of its microscopic parameters) in a simple and intuitive way. Finally we also make a prediction for the range of the parameter x_6 over which the RB theory is stable, i.e. has a stable vacuum.

3.5.1 An effective potential for σ_B

In the zero temperature limit the three-variable off-shell free energy (3.52) simplifies to

$$F(c_B,\sigma_B,\widetilde{\mathcal{S}}) = \frac{N_B}{6\pi}\Big[-\widehat{c}_B^3 - 4\widetilde{\mathcal{S}}^3\lambda_B^2 - 3\widehat{c}_B^2\widehat{\sigma}_B - 12\widetilde{\mathcal{S}}^2\lambda_B^2\widehat{\sigma}_B - 12\widetilde{\mathcal{S}}\lambda_B^2\widehat{\sigma}_B^2 + 6\widehat{c}_B|\lambda_B|(\widetilde{\mathcal{S}}+\widehat{\sigma}_B)^2 + 3\left(\widehat{m}_B^2\widehat{\sigma}_B + 2\lambda_B\widehat{b}_4\widehat{\sigma}_B^2 + x_6^B\lambda_B^2\widehat{\sigma}_B^3\right)\Big] . \quad (3.60)$$

Varying this free energy w.r.t. $\widetilde{\mathcal{S}}$ produces the first of the gap equations in (3.53) which we repeat here for convenience

$$(\widetilde{\mathcal{S}}+\widehat{\sigma}_B)(-\widehat{c}_B + |\lambda_B|(\widetilde{\mathcal{S}}+\widehat{\sigma}_B)) = 0 . \quad (3.61)$$

As we have discussed above, this equation has two solutions corresponding to the unHiggsed and Higgsed branches:

$$\text{unHiggsed}: \quad \widetilde{\mathcal{S}} = -\widehat{\sigma}_B , \qquad \text{Higgsed}: \quad \widetilde{\mathcal{S}} = -\widehat{\sigma}_B + \frac{\widehat{c}_B}{|\lambda_B|} . \quad (3.62)$$

Plugging these solutions back into the expression for the free energy, we have, in the unHiggsed phase,

$$F^{(\mathrm{uH})}(c_B,\sigma_B)=\frac{N_B}{6\pi T^3}\left(-c_B^3+4\lambda_B^2\sigma_B^3-3(c_B^2-m_B^2)\sigma_B+6b_4\lambda_B\sigma_B^2+3x_6^B\lambda_B^2\sigma_B^3\right)\ , \quad (3.63)$$

and in the Higgsed phase,

$$F^{(\mathrm{H})}(c_B,\sigma_B)=\frac{N_B}{6\pi T^3}\left(\tfrac{2-|\lambda_B|}{|\lambda_B|}c_B^3+4\lambda_B^2\sigma_B^3-3(c_B^2-m_B^2)\sigma_B+6b_4\lambda_B\sigma_B^2+3x_6^B\lambda_B^2\sigma_B^3\right)\ . \quad (3.64)$$

We then extremize the above free energies with respect to c_B to get

$$\text{unHiggsed}: \quad c_B=-2\sigma_B\ , \qquad \text{Higgsed}: \quad c_B=\frac{|\lambda_B|}{2-|\lambda_B|}2\sigma_B\ . \quad (3.65)$$

Recall that c_B is positive by definition. It follows that the solutions (3.65) exist only when σ_B is positive (negative) in the Higgsed (unHiggsed) phase respectively. Plugging back the above expressions into the free energies in (3.63) and (3.64), we get

$$F^{(\mathrm{uH})}(\sigma_B)=\frac{N_B}{2\pi T^3}\left[\left((1+x_6^B)-\frac{4-\lambda_B^2}{3\lambda_B^2}\right)\lambda_B^2\sigma_B^3+2b_4\lambda_B\sigma_B^2+m_B^2\sigma_B\right]\ , \quad (3.66)$$

and in the Higgsed phase,

$$F^{(\mathrm{H})}(\sigma_B)=\frac{N_B}{2\pi T^3}\left[\left((1+x_6^B)-\frac{|\lambda_B|(4-|\lambda_B|)}{3(2-|\lambda_B|)^2}\right)\lambda_B^2\sigma_B^3+2b_4\lambda_B\sigma_B^2+m_B^2\sigma_B\right]\ . \quad (3.67)$$

The quantum effective potential for the field $(\overline{\phi}\phi)_{\mathrm{cl}}$ is related to the above free energies as

$$U_{\mathrm{eff}}((\overline{\phi}\phi)_{\mathrm{cl}})=T^3F(\sigma_B) \quad \text{with the replacement} \quad \sigma_B\to\frac{2\pi(\overline{\phi}\phi)_{\mathrm{cl}}}{N_B}\ . \quad (3.68)$$

We continue to use the variable σ_B as the argument of the effective potential U_{eff} to avoid clutter, with the understanding that all instances of σ_B in U_{eff} are to be replaced with $2\pi(\overline{\phi}\phi)_{\mathrm{cl}}/N_B$. Explicitly, we have

$$\boxed{\begin{aligned}U_{\mathrm{eff}}(\sigma_B)=&\begin{cases}\frac{N_B}{2\pi}\left[(x_6-\phi_2)\lambda_B^2\sigma_B^3+2\lambda_Bb_4\sigma_B^2+m_B^2\sigma_B\right] & \text{for } \sigma_B<0\ ,\\ \frac{N_B}{2\pi}\left[(x_6-\phi_1)\lambda_B^2\sigma_B^3+2\lambda_Bb_4\sigma_B^2+m_B^2\sigma_B\right] & \text{for } \sigma_B>0\ ,\end{cases}\\ &\text{with the replacement} \quad \sigma_B\to\frac{2\pi(\overline{\phi}\phi)_{\mathrm{cl}}}{N_B}\ .\end{aligned}} \quad (3.69)$$

The constants ϕ_1 and ϕ_2 are given by

$$\phi_1=\frac{4}{3}\left(\frac{1}{(2-|\lambda_B|)^2}-1\right)\ , \quad \phi_2=\frac{4}{3}\left(\frac{1}{\lambda_B^2}-1\right)\ . \quad (3.70)$$

Observe that the effective potential (3.69) is bounded from below for positive values of σ_B if the coefficient of σ_B^3 is positive in the second of (3.69), i.e. when

$$x_6 > \phi_1 \ . \tag{3.71}$$

Similarly, the effective potential is bounded from below for negative values of σ_B if the coefficient of the σ_B^3 term is negative in the first of (3.69), i.e. when

$$x_6 < \phi_2 \ . \tag{3.72}$$

Note that $\phi_1 < \phi_2$.

Note that the terms proportional to σ_B^2 and σ_B are identical for the two ranges of σ_B but the coefficients of the σ_B^3 terms are different: this non-analyticity in the cubic term is what gives a sharp distinction between the Higgsed and unHiggsed branches of the effective potential at zero temperature. When we turn on temperature we expect this non-analyticity to be smoothed out. [14]

We also give a slightly different expression for the Landau-Ginzburg potential in terms of the variable c_B which is useful for the analysis of the gap equations as performed in Section 4 of [57]. For this purpose, we substitute back the expressions for σ_B in terms of c_B from (3.65):

$$U_{\rm eff}(c_B) = \begin{cases} \dfrac{N_B}{2\pi}\left[A_u\dfrac{c_B^3}{6} + B_{4,u}\dfrac{c_B^2}{2} - m_B^2\dfrac{c_B}{2}\right] & \text{unHiggsed} \ , \\ \dfrac{N_B(2-|\lambda_B|)}{2\pi|\lambda_B|}\left[-A_h\dfrac{c_B^3}{6} - B_{4,h}\dfrac{c_B^2}{2} + m_B^2\dfrac{c_B}{2}\right] & \text{Higgsed} \ . \end{cases} \tag{3.73}$$

Here, A_u, $B_{4,u}$ and A_h, $B_{4,h}$ are constants defined by

$$\begin{aligned} A_u &= 1 - \left(1 + \frac{3x_6}{4}\right)\lambda_B^2 \ , & B_{4,u} &= \lambda_B b_4 \ , \\ A_h &= 1 - \left(1 + \frac{3x_6}{4}\right)(2-|\lambda_B|)^2 \ , & B_{4,h} &= -\mathrm{sgn}(\lambda_B)(2-|\lambda_B|)b_4 \ . \end{aligned} \tag{3.74}$$

The off-shell variable c_B has the following advantage; its on-shell value coincides with the pole mass (the *gap*) of the fundamental excitation in the corresponding phase. We record the gap equations that follow from extremizing (3.73) w.r.t the variable c_B:

$$\begin{aligned} \text{unHiggsed}: \quad & A_u c_B^2 + 2B_{4,u}c_B - m_B^2 = 0 \ , \\ \text{Higgsed}: \quad & A_h c_B^2 + 2B_{4,h}c_B - m_B^2 = 0 \ . \end{aligned} \tag{3.75}$$

Solutions to the quadratic equations in (3.75) above correspond to candidates for the Higgsed or unHiggsed phases of the theory. The very recent paper [57] analysed the solutions of (3.75) in detail and used information about the free energy at these solutions to obtain the phase structure as a function of the parameters x_6, $\lambda_B b_4$ and m_B^2. In the next subsection we will extract the same physical information from the exact Landau-Ginzburg effective potential (3.69).

[14] This should be easy to verify - and seems to follow from the fact that the finite temperature free energy is an analytic function of its variables - but we have not verified it in detail.

3.5.2 Higgsed branch

In this subsection we study the effective potential (3.69) in more detail on the Higgsed branch.

Potential for the Higgs vev

We have already noted around (4.84) that the variable σ_B has a simple interpretation in terms of the Higgs vev on the Higgsed branch. Making the replacement (4.84), i.e. $\sigma_B = 2\pi v^2/|\lambda_B|$ in the second line of (3.69) we find

$$U_{\rm eff}(v) = \frac{N_B}{|\lambda_B|}\left(m_B^2 v^2 + 4\pi \mathrm{sgn}(\lambda_B) b_4 v^4 + \left((1+x_6^B) - \tfrac{|\lambda_B|(4-|\lambda_B|)}{3(2-|\lambda_B|)^2}\right) 4\pi^2 v^6\right) \ , \tag{3.76}$$

It is easily verified that (3.76) may also be obtained by taking the zero temperature limit of (3.37) (i.e. dropping the last line in that formula) and integrating c_B out of that equation. It follows that the true value of the Higgs vev in the vacuum is obtained by extremizing (3.76).

Note that (3.76) manifestly reduces to the classical potential

$$U_{\rm cl}(v) = \frac{N_B}{|\lambda_B|}\left(m_B^2 v^2 + 4\pi \mathrm{sgn}(\lambda_B) b_4 v^4 + (2\pi)^2 (x_6+1) v^6\right) \ , \tag{3.77}$$

in the classical limit (3.11) (i.e. $\lambda_B \to 0$ with m_B^2, b_4 and x_6 fixed) as expected on general grounds.

Graphs of $U_{\rm eff}(\sigma_B)$ in various cases

In this subsubsection we will study the graphs of $U_{\rm eff}(\sigma_B)$ on the Higgsed branch for various ranges of values of microscopic parameters. The results of this subsubsection will prove useful in sketching the phase diagram of the RB theory in later subsections.

Recall that, on the Higgsed branch, $U_{\rm eff}(\sigma_B)$ is given by the expression

$$U_{\rm eff}(\sigma_B) = \frac{N_B}{2\pi}\left(-\tfrac{|\lambda_B|^2}{(2-|\lambda_B|)^2} A_h \frac{4\sigma_B^3}{3} - \tfrac{|\lambda_B|}{(2-|\lambda_B|)} 2 B_{4,h} \sigma_B^2 + m_B^2 \sigma_B\right) \ , \quad \sigma_B > 0 \ , \tag{3.78}$$

(A_h, $B_{4,h}$ were defined in the second line of (3.74)). The extremization of $U_{\rm eff}(\sigma_B)$ produces the gap equation in the second of (3.75) which we reproduce here:

$$A_h c_B^2 + 2 B_{4,h} c_B - m_B^2 = 0 \ . \tag{3.79}$$

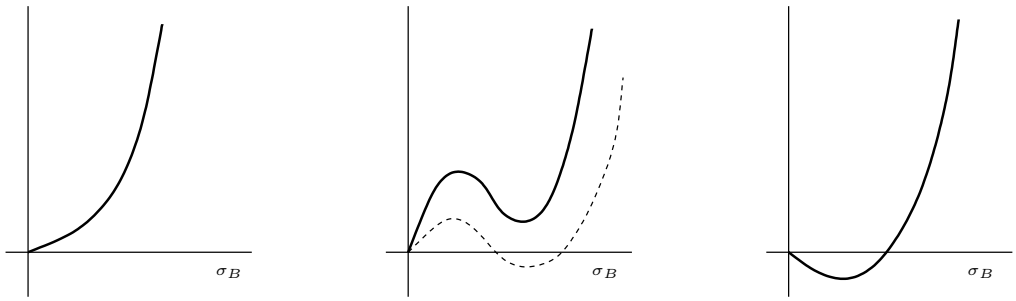
The structure of the curves for $U_{\rm eff}$ above turn out to depend sensitively on discriminant D_h of this gap equation (3.79):

$$D_h = 4(B_{4,h}^2 + m_B^2 A_h) \ . \tag{3.80}$$

As discussed below (3.69), the above effective potential is bounded below when $x_6 > \phi_1$ where ϕ_1 was defined in (3.70). In other words $A_h < 0$ when $x_6 > \phi_1$ and $A_h > 0$ when $x_6 < \phi_1$. [15]

We have the following cases as depicted in Figures 3.1 and 3.2:

1. A_h *is negative*: the potential U_{eff} increases at large σ_B.



1(a).i. 1(a).ii. 1(b)

Figure 3.1: Effective potential in the Higgsed phase for A_h negative.

   (a) m_B^2 *is positive*:

      i. $B_{4,h}$ *is negative, or* $B_{4,h}$ *is positive such* D_h *is negative* : the potential rises monotonically as σ_B increases from zero to infinity, and there are no nontrivial positive solutions of the gap equation (3.79).

      ii. $B_{4,h}$ *is positive such that* D_h *is positive*: As σ_B is increased from zero, U_{eff} initially increases, reaches a local maximum and then decreases, reaches a minimum and then increases without bound. In this case the gap equation has two solutions; the larger of which is the candidate for a stable phase (the smaller solution presumably describes unstable dynamics since it occurs at a local maximum of the effective potential).

   (b) m_B^2 *is negative*: For either sign of $B_{4,h}$, U_{eff} initially decreases, reaches a minimum and then turns and increases indefinitely. The gap equation has exactly one legal solution (i.e. a solution for c_B which is positive) which is the candidate for a stable phase.

2. A_h *is positive*: the potential U_{eff} decreases at large σ_B.

   (a) m_B^2 *is negative*

[15] Note that ϕ_1 is an increasing function of λ_B. In particular $\phi_1 = -1$ for the free theory ($\lambda_B = 0$), whereas for the strongly coupled case ($|\lambda_B| = 1$) we have $\phi_1 = 0$. The fact that ϕ_1 increases as we increase $|\lambda_B|$ indicates that coupling effects increase the propensity of our theory to develop a runaway instability along the v direction.

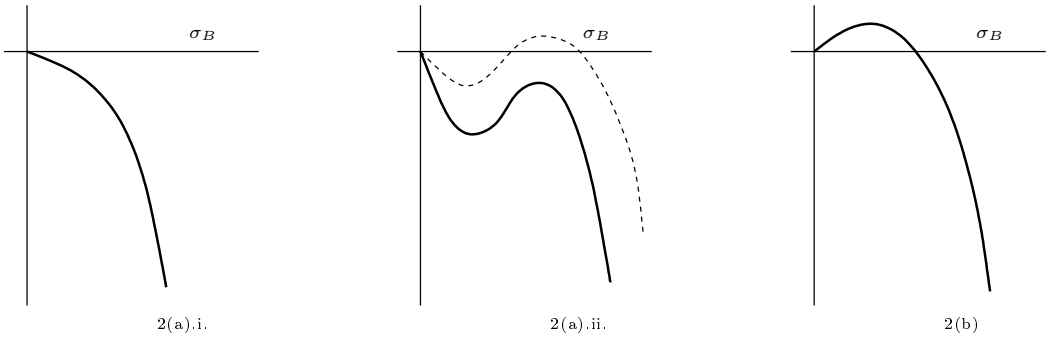


Figure 3.2: Effective potential in the Higgsed phase for A_h negative.

i. *$B_{4,h}$ is positive, or $B_{4,h}$ is negative such that D_h is negative*: The potential decreases monotonically as σ_B increases from zero to infinity, and there are no nontrivial positive solutions of the gap equation (3.79).

ii. *$B_{4,h}$ is negative such that D_h is positive*: As σ_B is increased from zero U_{eff} initially decreases, reaches a local minimum and then increases till it reaches a local maximum after which it decreases without bound. In this case the gap equation has two solutions; the smaller of which is the candidate metastable phase (the larger solution presumably describes unstable dynamics again since it occurs at a local maximum of the potential).

(b) *m_B^2 is positive*: For either sign of $B_{4,h}$, U_{eff} initially increases, reaches a maximum and then turns and decreases indefinitely. The gap equation has exactly one legal solution; this is a local maximum and so presumably describes an unstable 'phase'.

In the last two paragraphs above we have encountered three examples of solutions (1.a.ii, 2.a.ii, 2.b) to the gap equations that describe unstable 'phases'. For future use we note that these three unstable solutions are all given by the following root of (3.79):

$$c_B = \frac{-B_{4,h} + \sqrt{B_{4,h}^2 + A_h m_B^2}}{A_h} \ . \tag{3.81}$$

It is also easy to check that the three cases described above are the only three legal roots of the form (3.81). In other words every local maximum of the potential (3.76) is a root of the form (3.81), and a legal root (i.e. a root for which the RHS is positive) of the form (3.81) is one of the three 'local maxima' situations described above [16].

[16] To repeat, these three cases are as follows. First when A_h is negative, m_B^2 positive and $B_{4,h}$ positive such that D_h is positive. Second when A_h is positive, m_B^2 is negative and $B_{4,h}$ negative such that D_h is positive. Lastly when A_h is positive and m_B^2 positive for either sign of $B_{4,h}$.

3.5.3 unHiggsed branch

Graphs of $U_{\rm eff}(\sigma_B)$ in various cases

In this subsubsection we plot $U_{\rm eff}(\sigma_B)$ on the unHiggsed branch for various ranges of microscopic parameters.

We start with the following expression for the Landau-Ginzburg potential in the unHiggsed branch in terms of the constants A_u and $B_{4,u}$:

$$U_{\rm eff}(\sigma_B) = \frac{N_B}{2\pi}\left(-A_u\frac{4\sigma_B^3}{3} + 2B_{4,u}\sigma_B^2 + m_B^2\sigma_B\right) \ , \tag{3.82}$$

where A_u and $B_{4,u}$ are as in the first line of (3.74). Now recall that σ_B is necessarily negative in the unHiggsed phase. As a consequence (3.82) may be rewritten as

$$U_{\rm eff}(\sigma_B) = \frac{N_B}{2\pi}\left(A_u\frac{4|\sigma_B|^3}{3} + 2B_{4,u}|\sigma_B|^2 - m_B^2|\sigma_B|\right) \ . \tag{3.83}$$

The gap equation that follows by varying (3.83) w.r.t. σ_B is given in (3.75) and is reproduced below:

$$A_u c_B^2 + 2B_{4,u}c_B - m_B^2 = 0 \ , \tag{3.84}$$

where $c_B = 2|\sigma_B|$. Note the formal and notational similarity with the analogous equation (3.79) in the Higgsed phase. As in the previous subsection we briefly analyze the behaviour of (3.83) as a function of σ_B in all the various cases. We define the discriminant D_u of (3.84):

$$D_u = 4(B_{4,u}^2 + m_B^2 A_u) \ . \tag{3.85}$$

We then have the following cases as depicted in Figures 3.3 and 3.4:

1. A_u *is positive*: the potential $U_{\rm eff}$ increases at large $|\sigma_B|$.

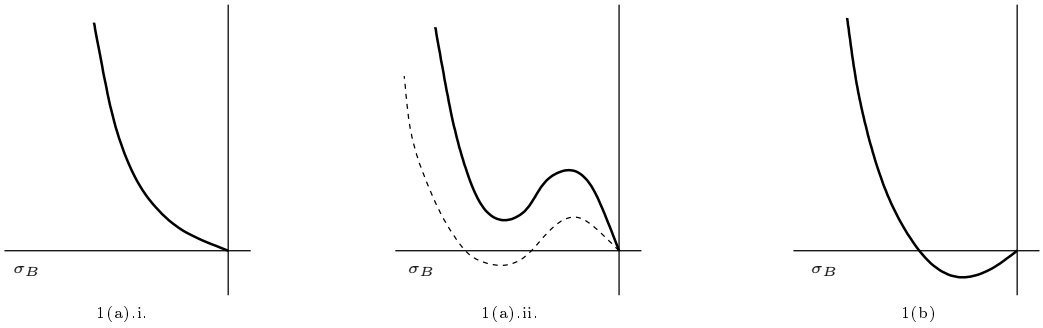


Figure 3.3: Effective potential in the unHiggsed phase for A_u positive.

(a) m_B^2 *is negative*:

i. $B_{4,u}$ *is positive, or* $B_{4,u}$ *is negative such that* D_u *is negative*: the potential rises monotonically as $|\sigma_B|$ increases from zero to infinity, and there are no nontrivial positive solutions of the gap equation (3.84).

ii. $B_{4,u}$ *is negative such that* D_u *is positive*: As $|\sigma_B|$ is increased from zero, U_{eff} initially increases up to a local maximum and then decreases down to a local minimum and then increases without bound. The gap equation has two solutions the larger of which is the dominant stable phase (the smaller solution presumably describes unstable dynamics since it occurs at a local maximum of the effective potential).

(b) m_B^2 *is positive*: For either sign of $B_{4,u}$, U_{eff} initially decreases, reaches a minimum and then turns and increases indefinitely. The gap equation has exactly one legal solution which is the dominant stable phase.

2. A_u *is negative*: the potential U_{eff} decreases at large $|\sigma_B|$.

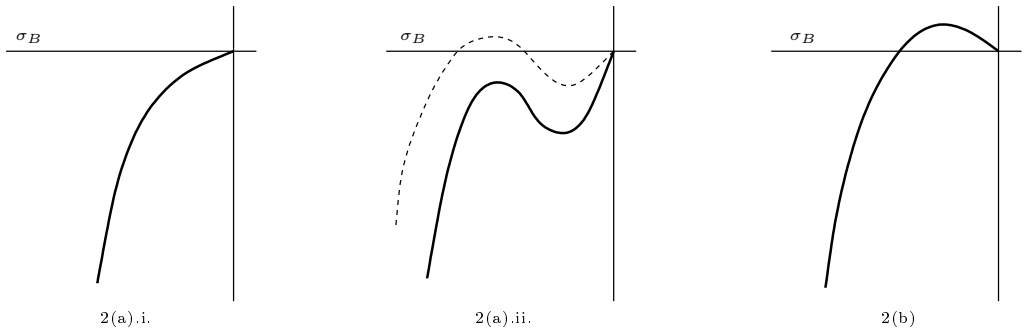


Figure 3.4: Effective potential in the unHiggsed phase for A_u negative.

(a) m_B^2 *is positive*

i. $B_{4,u}$ *is negative, or* $B_{4,u}$ *is positive such that* D_u *is negative*: The potential decreases monotonically as $|\sigma_B|$ increases from zero to infinity, and there are no nontrivial positive solutions of the gap equation (3.84).

ii. $B_{4,u}$ *is positive such that* D_u *is positive*: As $|\sigma_B|$ is increased from zero U_{eff} initially decreases down to a local minimum and then increases up to a local maximum after which it decreases without bound. The smaller of the two solutions to the gap equation (3.84) is the 'dominant' metastable phase (the larger solution presumably describes unstable dynamics again since it occurs at a local maximum of the potential).

(b) m_B^2 *is negative*: For either sign of $B_{4,u}$, U_{eff} initially increases, reaches a maximum and then turns and decreases indefinitely. The gap equation has exactly one legal solution; this is a local maximum and so presumably describes an unstable 'phase'.

In the paragraphs above we have encountered three examples of 'unstable phases' (1.a.ii, 2.a.ii, 2.b). The solution of the gap equation (3.84) associated with each of these phases is easily verified to be

$$c_B = \frac{-B_{4,u} - \sqrt{B_{4,u}^2 + A_u m_B^2}}{A_u} \ . \tag{3.86}$$

Moreover it is also easy to check that every legal (i.e. positive) solution of the form (3.86) is one of the three local maxima of the paragraphs described above.

Explanation for the instability of local maxima

We have already explained above that

$$|\sigma_B| = -\frac{2\pi(\overline{\phi}\phi)_{\rm cl}}{N_B} \ , \tag{3.87}$$

where the RHS of this equation should be interpreted in quantum rather than semiclassical terms since semiclassically the variable $(\overline{\phi}\phi)_{\rm cl}$ is given by $\langle\overline{\phi}\rangle\langle\phi\rangle$ and is positive. The operator $\overline{\phi}\phi$ however is not necessarily positive (this follows because the subtraction that is used to give this operator meaning is not positive). Equation (3.87) effectively asserts that the unHiggsed branch explores only negative values of the operator $\overline{\phi}\phi$. We have noted above the effective potential as a function of $(\overline{\phi}\phi)_{\rm cl}$ has unstable 'phases' that sit at local maxima of the effective potential. In the rest of this subsection we will present an explanation of these instabilities.

Our proposal for the mechanism of the instability of the 'local maxima' phases is that it is the tachyonic instability of a bound state of a single fundamental and antifundamental field in the singlet channel. We claim that the solutions (3.86) are all unstable in this sense, while none of the stable phases - i.e. the phases that occur at legal values of

$$c_B = \frac{-B_{4,u} + \sqrt{B_{4,u}^2 + A_u m_B^2}}{A_u} \ , \tag{3.88}$$

suffer from such an instability.

In order to see that this is indeed the case let us recall that bound states do occur as poles in the S-matrix of a fundamental ϕ field scattering off an antifundamental $\overline{\phi}$ field. Moreover these poles do sometimes go tachyonic (i.e. their squared mass sometimes goes below zero). The condition for this to happen can be worked out by following discussion in Section 4.5 of [59] and Appendix C of [60]. The particle-antiparticle scattering S-matrix has a pole with positive squared mass when

$$4\lambda_B \leq \lambda_B^2(3x_6^B + 4) - 4\frac{b_4\lambda_B}{c_B} \leq 4 \ . \tag{3.89}$$

When

$$4\lambda_B = \lambda_B^2(3x_6^B + 4) - 4\frac{b_4\lambda_B}{c_B} \ ,$$

the pole is at threshold, i.e. $\sqrt{s} = 2c_B$. On the other hand when

$$\lambda_B^2(3x_6^B + 4) - 4\frac{b_4\lambda_B}{c_B} = 4 \ , \tag{3.90}$$

the pole lies at $\sqrt{s} = 0$. Using the definitions (3.74) for A_u and $B_{4,u}$ it is easy to see that the condition (3.90) can be rewritten as

$$B_{4,u} = -A_u c_B \ . \tag{3.91}$$

(recall that the quantity c_B is positive by definition). When

$$\lambda_B^2(3x_6^B + 4) - 4\frac{b_4\lambda_B}{c_B} > 4 \ , \tag{3.92}$$

we have a bound state with negative squared mass, i.e. a tachyonic bound state. The condition for the existence of this tachyonic pole is

$$A_u c_B \leq -B_{4,u} \ . \tag{3.93}$$

Of course the quantity c_B is not independent of A_u and $B_{4,u}$ but is determined in terms of these quantities by the gap equation. The solutions to the gap equation are given by

$$c_B = \frac{-B_{4,u} \pm \sqrt{B_{4,u}^2 + A_u m_B^2}}{A_u} \tag{3.94}$$

Inserting these solutions into the condition (3.93), we find that the condition (3.93) is met whenever

$$-B_{4,u} \pm \sqrt{B_{4,u}^2 + A_u m_B^2} \leq -B_{4,u} \tag{3.95}$$

This condition is obeyed by the 'minus' branch of solutions (3.86) but not by the 'plus' branch of solutions (3.88). But we have seen above that this is precisely the split between the local maxima (solutions (3.86)) and local minima (solutions (3.88)) of the effective action (3.83). It is thus natural to identify the tachyonic bound states as the explanation for the instability of the 'minus' branch of solutions (3.86).

It follows, in other words, that the instabilities in the unHiggsed phase occur for the same reason as the instabilities in the Higgsed phase, but for a different field. Unstable Higgsed 'phases' occurred when our solution to the gap equations was at a maximum of the potential for the field ϕ. We propose that instabilities in the unHiggsed phase occur for solutions to the gap equation around maxima for the field $(\overline{\phi}\phi)_{\rm cl}$ that is very ostensibly related to the bound state of $\overline{\phi}$ and ϕ.

In the case of the Higgsed theory in the $\lambda_B \to 0$ limit (3.11) we obtained a classical theory (with an overall factor of λ_B^{-1} outside the action) in terms of the variable φ. In the current unHiggsed context the effective potential does not have a clear classical limit as $\lambda_B \to 0$. On solutions to the gap equation that follows from varying (3.83) w.r.t. $|\sigma_B|$, it turns out that σ_B and $\overline{\phi}\phi$ (rather than $\lambda_B\sigma_B$ and $\overline{\varphi}\varphi$ as in the Higgsed phase) are finite as $\lambda_B \to 0$. The field φ which was the natural classical variable at weak coupling in the Higgsed phase does not seem to be useful in the analysis of the unHiggsed branch (this is probably a reflection of the fact that dynamics is always quantum on this branch).

3.5.4 Landau-Ginzburg Analysis of the zero temperature phase diagram[17]

In subsections 3.5.2 and 3.5.3 we have already explored the qualitative structure of the Landau-Ginzburg potential (3.69), plotted as a function of σ_B, separately for $\sigma_B > 0$ and $\sigma_B < 0$. In this section we will simply put the analyses of subsections 3.5.2 and 3.5.3 together to obtain a global picture of the Landau-Ginzburg potential as a function of σ_B over all possible ranges of parameters x_6, $\lambda_B b_4$ and m_B^2. We reproduce the potential below. Recall that our exact Landau-Ginzburg potential as a function of σ_B (or equivalently, using (3.57), a function of $(\overline{\phi}\phi)_{\rm cl}$) is given by

$$U_{\rm eff}(\sigma_B) = \begin{cases} \frac{N_B}{2\pi}\left[(x_6-\phi_2)\lambda_B^2\sigma_B^3 + 2\lambda_B b_4\sigma_B^2 + m_B^2\sigma_B\right] & \text{for } \sigma_B < 0 \ , \\ \frac{N_B}{2\pi}\left[(x_6-\phi_1)\lambda_B^2\sigma_B^3 + 2\lambda_B b_4\sigma_B^2 + m_B^2\sigma_B\right] & \text{for } \sigma_B > 0 \ . \end{cases} \tag{3.96}$$

Recall from (3.70) that $\phi_1 < \phi_2$. As we have explored in detail above, the plots of the effective potential are qualitatively different when $x_6 < \phi_1$, $\phi_1 < x_6 < \phi_2$ and $x_6 > \phi_2$. For this reason we analyze these three ranges of parameters separately. It is also useful to recall the formulae for the discriminants (3.80) and (3.85) of the gap equations (3.75) in either phase of the theory:

$$\begin{aligned} D_h &= \frac{4(2-|\lambda_B|)^2}{\lambda_B^2}\left[(\lambda_B b_4)^2 - \tfrac{3\lambda_B^2}{4}(x_6-\phi_1)m_B^2\right] \ , \\ D_u &= 4\left[(\lambda_B b_4)^2 - \tfrac{3\lambda_B^2}{4}(x_6-\phi_2)m_B^2\right] \ . \end{aligned} \tag{3.97}$$

Case I: $x_6 > \phi_2$

In this case the coefficient of σ_B^3 in the effective potential (3.96) is positive both when $\sigma_B > 0$ and when $\sigma_B < 0$. It follows that $U_{\rm eff}(\sigma_B)$ is an increasing function in the limits $\sigma_B \to \pm\infty$. Note, in particular, that $U_{\rm eff}(\sigma_B)$ is unbounded from below at large negative σ_B presumably indicating a runaway instability of the theory. In other words, the theory has no truly stable phase in this range of x_6. In this subsection we will sketch the 'phase diagram' of the theory, defined as the diagram that tracks the dominant metastable phase as a function of the relevant parameters. [18] In order to do this we simply plot $U_{\rm eff}(\sigma_B)$ as a function of σ_B. The detailed behaviour of the curve $U_{\rm eff}(\sigma_B)$ at finite values of σ_B depends on the signs and values of m_B^2 and $\lambda_B b_4$. We have the following sub cases.

1. *m_B^2 positive.*

[17] This subsection was worked out in collaboration with O. Aharony.

[18] We emphasize that this phase diagram is formal; no phase - not even the dominant one - is stable. The theory always has a run away instability to tunnel to large negative values of σ_B.

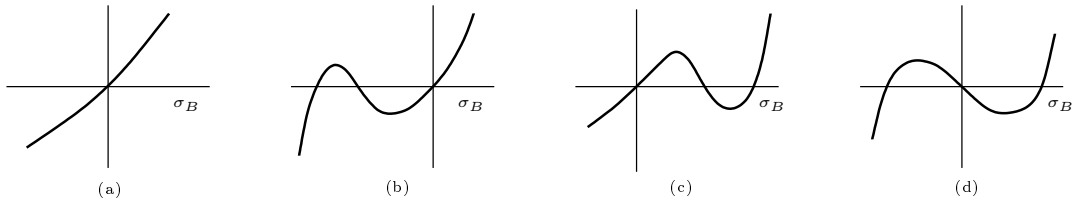


Figure 3.5: Effective potential for $x_6 > \phi_2$.

(a) $\lambda_B b_4$ *positive with* D_u *negative* or $\lambda_B b_4$ *negative with* D_h *negative*: In this case $U_{\text{eff}}(\sigma_B)$ is a monotonically increasing function of σ_B as depicted in Fig 3.5(a). $U_{\text{eff}}(\sigma_B)$ has no extrema and so the gap equation has no solutions.

(b) $\lambda_B b_4$ *positive with* D_u *positive*: In this case the curve of $U_{\text{eff}}(\sigma_B)$ takes the schematic form depicted in Fig 3.5(b). $U_{\text{eff}}(\sigma_B)$ has two extrema; a local minimum and a local maximum both for $\sigma_B < 0$, so both in the unHiggsed branch. The local minimum is the only metastable phase of the theory (the maximum is unstable) and so is the dominant 'phase'.

(c) $\lambda_B b_4$ *negative with* D_h *positive:* The graph of $U_{\text{eff}}(\sigma_B)$ takes the schematic form depicted in Fig 3.5(c). $U_{\text{eff}}(\sigma_B)$ has two extrema; a local minimum and a local maximum both for positive σ_B so in the Higgsed branch. The local minimum is the only metastable phase of the theory (the maximum is unstable) and so is the dominant 'phase'.

2. m_B^2 *negative,* $\lambda_B b_4$ *arbitrary:* In this case the graph of $U_{\text{eff}}(\sigma_B)$ versus σ_B takes the schematic form depicted in Fig 3.5(d). We have a local maximum at negative σ_B (so in the unHiggsed phase) and a local minimum - so a metastable phase - at positive σ_B, so in the Higgsed branch. This local minimum is the dominant (metastable) phase.

Putting all this together we conclude that our theory has the (metastable) phase structure depicted in Fig. 8 of [57] and redrawn here for convenience in Fig. 3.6 of this chapter. Notice that whenever a metastable phase exists, a subdominant unstable local maximum of $U_{\text{eff}}(\sigma_B)$ also exists in the vicinity. These are the subdominant 'phases' that appear in Fig. 22(a) of [57].

To end this subsubsection let us study what happens in the limit in which $x_6 \to \phi_2$ from above. First, nothing special happens to $U_{\text{eff}}(\sigma_B)$ for positive σ_B. At negative σ_B, however, the coefficient of the σ_B^3 term tends to zero when $\sigma_B < 0$. In this limit D_u is always positive, so the top half of the red curve in Fig. 3.6 tends to a horizontal line (the $m_B^2 > 0$ axis). Moreover, when m_B^2 and $\lambda_B b_4$ are both positive (i.e. the case of Fig. 3.5(b)) the local minimum (which can be thought of as arising due to a competition between the linear and quadratic terms in the action) continues to occur at a fixed value of σ_B and the $U_{\text{eff}}(\sigma_B)$ evaluated at this minimum also remains fixed. But the local maximum of this

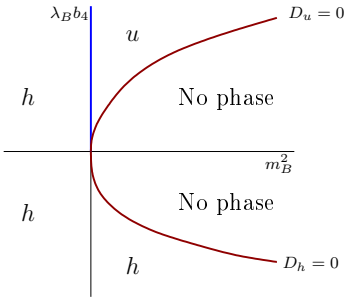


Figure 3.6: Phase diagram for $x_6 > \phi_2$. The positive $\lambda_B b_4$ axis (shown in blue) corresponds to a second-order phase transition.

diagram (which is a result of the competition between the cubic and quadratic terms in the action) now occurs at a value of σ_B that tends to $-\infty$. Moreover the value of $U_{\text{eff}}(\sigma_B)$ at this maximum also tends to ∞. For $x_6 \leq \phi_2$ this local maximum simply does not exist any more.

Case II: $\phi_1 < x_6 < \phi_2$

In this case, the coefficient of σ_B^3 is positive for $\sigma_B > 0$ and negative for $\sigma_B < 0$ which implies that the potential is bounded below for all values of σ_B, so the theory is stable. $U_{\text{eff}}(\sigma_B)$ is a decreasing function of σ_B for large negative σ_B, but is an increasing function of σ_B for large positive σ_B.

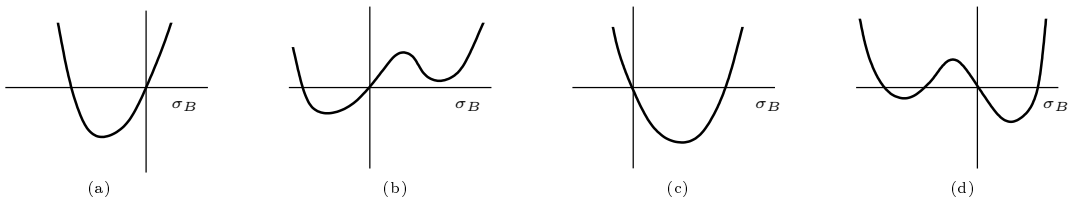


Figure 3.7: Effective potential for $\phi_1 < x_6 < \phi_2$.

1. m_B^2 *positive*: $\lambda_B b_4$ *positive or* $\lambda_B b_4$ *negative with* D_h *negative:* In this case the graph of $U_{\text{eff}}(\sigma_B)$ versus σ_B takes the form depicted in Fig 3.7(a). The global minimum in the unHiggsed phase (negative σ_B) is the only extremum of $U_{\text{eff}}(\sigma_B)$; this phase dominates the phase diagram.

2. m_B^2 *positive and* $\lambda_B b_4$ *negative with* D_h *positive* or
m_B^2 *negative and* $\lambda_B b_4$ *negative with* D_u *positive*: In this case the graph of $U_{\text{eff}}(\sigma_B)$ versus σ_B takes the form depicted in Fig 3.7(b) when m_B^2 is positive and of the form depicted in Fig 3.7(d) when m_B^2 is negative. In either case the graph has a

local minimum in the unHiggsed branch (negative σ_B) and a local minimum in the Higgsed branch (positive σ_B) separated by a local maximum. The maximum occurs in the Higgsed branch when $m_B^2 > 0$ but in the unHiggsed branch when $m_B^2 < 0$. [19] The dominant phase is the local minimum with the smaller free energy. Which phase dominates depends on the precise values of m_B^2, $\lambda_B b_4$ and x_6. A detailed analysis has been performed in [57] and we summarize the results here. When x_6 is strictly between ϕ_1 and ϕ_2 the theory has a first order phase transition line along the curve

$$D_\nu = m_B^2 - \nu_c(x_6)(\lambda_B b_4)^2 = 0 \ . \tag{3.98}$$

The function $\nu_c(x_6)$ was studied in detail in [57]; see around Fig. 25 and Fig. 26. The function $\nu_c(x_6)$ is monotonically decreasing as a function of x_6 with $x_6 \in (\phi_1, \phi_2)$. The function $\nu_c(x_6)$ is negative when x_6 is near ϕ_2 and hence the first order transition line is in the third quadrant (corresponding to Fig. 3.8(a)). When x_6 is near ϕ_1, the function $\nu_c(x_6)$ is positive and hence the first order transition line is in the fourth quadrant (Fig. 3.8(b)). The phase transition line crosses over to the fourth quadrant from the third quadrant (equivalently, $\nu_c(x_6)$ goes from being negative to positive) at some intermediate value of x_6. This intermediate value occurs at $x_6 = \frac{1}{2}(\phi_1 + \phi_2)$ and the phase transition line coincides with the negative $\lambda_B b_4$ axis. We plot the phase diagram for this case and also the corresponding Landau-Ginzburg potential on the phase transition line in Figure 3.9. Clearly, we have two exactly equal minima and hence the onset of a first order phase transition.

Let us study the behaviour in the limits $x_6 \to \phi_1$ and $x_6 \to \phi_2$. In the limit $x_6 \to \phi_2$ from below, the unHiggsed branch minimum occurs at $\sigma_B \to -\infty$ and the potential $U_{\text{eff}}(\sigma_B)$ evaluated on this solution tends to $-\infty$. In this limit the unHiggsed branch local minimum is the dominant phase for every value of $\lambda_B b_4 < 0$ and m_B^2. In the opposite limit $x_6 \to \phi_1$ the Higgsed branch local minimum occurs at very large values of σ_B and $U_{\text{eff}}(\sigma_B)$ evaluated on this solution tends to $-\infty$. In this limit the Higgsed branch local minimum is the dominant phase for every value of $\lambda_B b_4 < 0$ and m_B^2.

3. *m_B^2 negative: $\lambda_B b_4$ positive or $\lambda_B b_4$ negative with D_u negative:* The graph of $U_{\text{eff}}(\sigma_B)$ versus σ_B takes the form depicted in Fig. 3.7(c). The global minimum in the Higgsed branch (positive σ_B) is the only extremum of $U_{\text{eff}}(\sigma_B)$; this phase dominates the phase diagram.

Putting all this together we arrive at the phase diagram presented in Fig. 7 of [57]. This phase diagram is resketched in Fig 3.8 for convenience.

Case III: $x_6 < \phi_1$

In this case, the coefficient of σ_B^3 is negative for both $\sigma_B > 0$ and $\sigma_B < 0$. It follows that $U_{\text{eff}}(\sigma_B)$ is a increasing function in the limits $\sigma_B \to \pm\infty$. Note, in particular, that $U_{\text{eff}}(\sigma_B)$

[19] In fact at $m_B^2 = 0$ the local maximum goes through $\sigma_B = 0$; this maximum undergoes a 'second order phase transition' at this point from the Higgsed to the unHiggsed phase. This point is depicted in Figure 3.9.

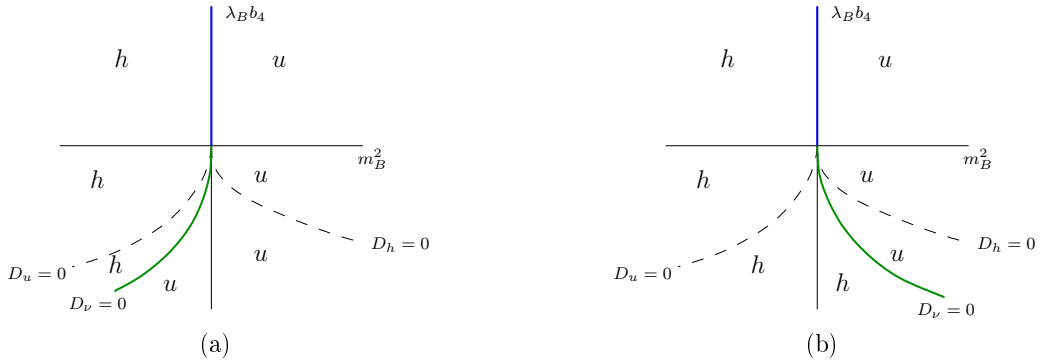


Figure 3.8: The phase diagram for $\phi_1 < x_6 < \phi_2$. There is a second order phase transition (shown in blue) along the positive $\lambda_B b_4$ axis. The first order phase transition line is the curve (shown in green) between the two dashed curves. The precise location of this phase transition curve varies as we change x_6. Two possible locations of this curve have been sketched in the two figures above. The first figure corresponds to x_6 near ϕ_2 and the second figure corresponds to x_6 near ϕ_1.

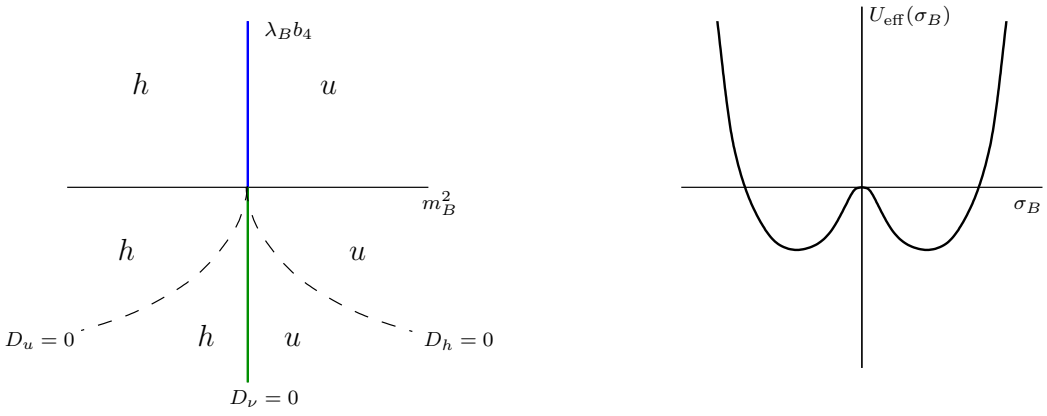


Figure 3.9: The first figure is the phase diagram for $x_6 = \frac{1}{2}(\phi_1 + \phi_2)$. The second figure is the Landau-Ginzburg potential at the same value of x_6 for a point on the first order phase transition line (green) corresponding to $m_B^2 = 0$ and some $\lambda_B b_4 < 0$.

is unbounded from below at large positive σ_B presumably implying a runaway instability of the theory. In this case the instability is easy to understand as it is present even in the classical theory at sufficiently negative values of x_6. Just as in Section 3.5.4, in range of parameters the RB the theory has no truly stable phases. As in Section 3.5.4, in this subsubsection we will sketch the 'phase diagram' of the theory, defined as the diagram that tracks the dominant metastable phase as a function of relevant parameters. As in Section 3.5.4 we read off our results from plots of $U_{\rm eff}(\sigma_B)$ as a function of σ_B. We have the following subcases.

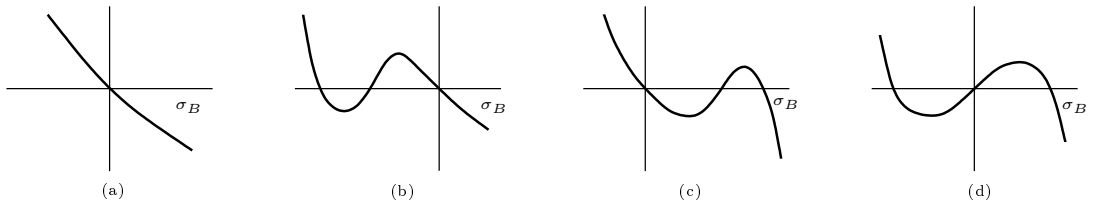


Figure 3.10: Effective potential for $x_6 < \phi_1$.

1. m_B^2 *negative.*

   (a) $\lambda_B b_4$ *negative with* D_u *negative or* $\lambda_B b_4$ *positive with* D_h *negative:* $U_{\rm eff}(\sigma_B)$ is a monotonically decreasing function of σ_B as depicted in Fig 3.10(a). $U_{\rm eff}(\sigma_B)$ has no extrema, and so the gap equation has no solutions.

   (b) $\lambda_B b_4$ *negative with* D_u *positive:* In this case the curve of $U_{\rm eff}(\sigma_B)$ takes the schematic form depicted in Fig. 3.10(b). $U_{\rm eff}(\sigma_B)$ has two extrema; a local minimum and a local maximum both for $\sigma_B < 0$, so both in the unHiggsed phase. The local minimum is the only metastable phase of the theory (the maximum is unstable) and so is the dominant 'phase'.

   (c) $\lambda_B b_4$ *positive with* D_h *positive:* The graph of $U_{\rm eff}(\sigma_B)$ takes the schematic form depicted in Fig 3.10(c). $U_{\rm eff}(\sigma_B)$ has two extrema; a local minimum and a local maximum both for positive σ_B so in the Higgsed phase. The local minimum is the only metastable phase of the theory (the maximum is unstable) and so is the dominant 'phase'.

2. m_B^2 *positive,* $\lambda_B b_4$ *arbitrary:* In this case the graph of $U_{\rm eff}(\sigma_B)$ versus σ_B takes the form depicted in Fig 3.10(d). We have a local minimum in at negative σ_B (so in the unHiggsed phase) and a local maximum at positive σ_B, so in the Higgsed phase. This local minimum is the dominant (metastable) phase.

Putting all this together we conclude that our theory has the (metastable) phase structure depicted in Fig. 9 of [57] and redrawn here for convenience in Fig. 3.11 of this chapter.

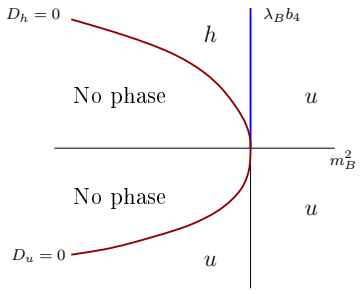


Figure 3.11: Phase diagram for $x_6 < \phi_1$. The positive $\lambda_B b_4$ axis (shown in blue) corresponds to a second-order phase transition.

Notice that whenever a metastable phase exists, a subdominant unstable local maximum of $U_{\text{eff}}(\sigma_B)$ also exists in the vicinity. These are the subdominant 'phases' that appear in Fig. 27(b) of [57].

To end this subsubsection let us study what happens in the limit in which $x_6 \to \phi_1$ from above. For negative values of σ_B, nothing special happens to $U_{\text{eff}}(\sigma_B)$. At positive σ_B, however, the coefficient of the σ_B^3 term tends to zero. In this limit D_h is always positive, so the top half of the red curve in Fig. 3.11 tends to a horizontal line (the negative m_B^2 axis). Moreover, when m_B^2 and $\lambda_B b_4$ are both positive (i.e. the case of sub Fig. 3.10(d)) the local minimum (which can be thought of as arising due to a competition between the linear and quadratic terms in the action) continues to occur at a fixed value of σ_B and the $U_{\text{eff}}(\sigma_B)$ evaluated at this minimum also remains fixed. But the local maximum of this diagram (which is a result of the competition between the cubic and quadratic terms in the action) now occurs at a value of σ_B that tends to ∞. Moreover the value of $U_{\text{eff}}(\sigma_B)$ at this maximum also tends to ∞. For $x_6 \geq \phi_1$ this local maximum simply does not exist any more.

3.6 Discussion

The results of this chapter suggest several questions for future work. First, it would be interesting to generalize the computation of S-matrices presented in [59, 60] to the Higgsed phase of the Regular Boson theory. The fact that this (and related) computations may throw light on the dual fermionic interpretation of the Z boson - as discussed in detail in [54] - make it particularly interesting.

One of the most interesting results of this chapter is the off-shell effective action (3.69). It would be interesting to generalize this result to finite values of temperature and chemical potential and explicitly observe the smoothing-out of the non-analyticity which was present at zero temperature. It would be also interesting to compute a similar action for the theory of one fundamental boson and one fundamental fermion studied in [62] and to use this action to unravel the phase structure of the deformed $\mathcal{N} = 2$ supersymmetric matter

Chern-Simons theory with a single chiral multiplet in the fundamental representation. It is possible that such an investigation will have interesting interplays with supersymmetry: for example it may be possible to find a superspace version of (3.69).

In Section 3.4 we have presented a three-variable off-shell free energy that reproduces the gap equation and thermal free energy of the regular boson theory. We have also presented a physical interpretation of the variable σ_B that enters this action. It would be interesting to investigate whether there are interesting off-shell interpretations of the other dynamical variables - c_B and $\widetilde{\mathcal{S}}$ - that appear in Section 3.4.

Above, we have found a preferred value for the cosmological constant counterterm Λ of the CB theory - one that correctly reproduces the tadpole condition for the regular boson theory (see (3.39)). It would be interesting to derive (3.39) from a more fundamental physical principle. It would also be interesting to investigate if this result makes any physical predictions for the critical boson theory: is it correct, for instance, to interpret the Legendre transform of (3.33) w.r.t $m_B^{\rm cri}$ (with Λ chosen to have the value (3.39)) as the Coleman-Weinberg potential of the CB theory w.r.t its dimension two scalar operator J_0? It would be interesting to further investigate this and similar questions, and their implications.

Finally it would be interesting to generalize the considerations of this chapter - even qualitatively - to finite values of N_B. We leave all these questions for future work.

Chapter 4

Classifying and Constraining 4 Photon and 4 Graviton S-Matrices

4.1 Introduction

The study of string theory suggests a surprising rigidity in the structure of quantum theories of gravity. For instance, there are only 5 known Lorentz Invariant theories of gravity in flat 10 dimensional space, namely Type I, Type IIA, Type IIB, $SO(32)$ Heterotic, $E8 \times E8$ Heterotic. Is it possible that these 5 are the only 10 dimensional stable Lorentz Invariant quantum theories of gravity? But how can we establish that there is no other Lorentz Invariant theory of gravity? Atleast with our current state of understanding of quantum gravity the only practical way of tackling such a question is to employ simple general low energy consistency considerations. This is the strategy we will employ in this chapter.

Consider all consistent Lorentz invariant d dimensional theories that admit a classical limit. We conjecture: The classical gravitational S matrix in every such theory is necessarily one of either the Einstein S matrix, or the Type II string theory S-matrix on $R^d \times M$ or the Heterotic string S-matrix on $R^d \times M$, where M is any 'compact space'.

Note the S-matrices above are independent of M. The conjecture of the last paragraph asserts that the gravitational part of the classical limit of any consistent theory of flat space gravity admits a consistent truncation to one of the three universal theories described above. Perhaps low energy consistency is enough to establish this result? To proceed towards proving this conjecture, we first tackle an easier sub-problem. Recall that the Type II and Heterotic S-matrices have intermediate massive poles corresponding to the exchange of higher spin massive particles. The conjecture of the previous paragraph - if true - implies a simpler result as a special case. Namely that Einstein gravity is the only consistent local (i.e. finite number of derivatives) classical theory of gravity that admits a consistent truncation involving no other fields. We wish to investigate this special case first. To complete this special case we need to prove this special case for scattering of $n = 3, 4, \ldots$ gravitons. The case of 3 gravitons is simple and there is already an interesting result about this in the literature - the 3 graviton S-matrix and constraint thereof [131].

This simplicity has its root in the fact that 3 graviton S matrices are highly kinematically constrained. The most general 3 graviton S matrix is kinematically forced to be a linear combination of three structures.

$$\begin{aligned} T_1 &= (\epsilon_1.\epsilon_2\epsilon_3.p_1 + \text{perm})^2 \qquad \text{2 der: Einstein} \\ T^2 &= (\epsilon_1 \wedge \epsilon_2 \wedge \epsilon_3 \wedge p_1 \wedge p_2)^2 \qquad \text{4 der : Gauss-Bonnet} \\ T^3 &= (\epsilon_1.p_2\epsilon_2.p_3\epsilon_3.p_1)^2 \qquad \text{6 der : Riemann}^3 \end{aligned} \tag{4.1}$$

Therefore, the most general 3 graviton S matrix takes the form

$$aT_1 + bT_2 + cT_3 \tag{4.2}$$

where, a, b, and c are numbers. [131] demonstrated that any theory in which either b or c is nonzero is necessarily acausal unless it couples to higher spin particles of arbitrarily high spin. In other words, in a causal gravitational theory with a local S matrix, $b = c = 0$. This already established our conjecture for 3 graviton scattering. However note that 3 graviton scattering is special as it is parameterized by finite data. We encounter qualitatively greater complexity when scattering 4 (or more) gravitons.

In the work that is described in this chapter, we have systematically listed all the possible $2 - 2$ S matrices for a classical theory of electrodynamics and gravity. Now, to prove the above stated conjecture, we wish to use causality and constrain the possible theories. This remains to be done in the future work.

As a warm-up for counting the independent function data, consider the scattering of 4 identical scalars. The most general S matrix is a fully symmetric function $\mathcal{S}$ of s, t, u with $s + t + u = 0$. For the case of local S matrices, $\mathcal{S}$ is a polynomial. The number of such polynomials at degree m is given by coefficient of x^m in

$$\begin{aligned} Z_{sym} &= \frac{1}{(1-x^2)(1-x^3)} \\ d_{sym}(m) &\sim \frac{m+1}{6} \qquad \text{asymptotically} \end{aligned} \tag{4.3}$$

This d_{sym} also counts the number of field redefinition inequivalent m derivative 4ϕ terms one can add to the free boson Lagrangian. There is a simple map from S matrices and Lagrangian structures.

Next we do this counting for the case when we have objects with indices. Such S matrices are labelled by polarization tensor in addition to Mandelstam invariants s, t, u. The full S matrix has to be S_4 invariant. Now it is easy to check that the $\mathbb{Z}_2 \times \mathbb{Z}_2$ subgroup of S_4 consisting of $I, P_{12}P_{34}, P_{13}P_{24}$ and $P_{14}P_{23}$ leaves s, t, and u unchanged. S_4 invariance thus requires that index structure that appears in the S matrix is $\mathbb{Z}_2 \times \mathbb{Z}_2$ invariant. The conditions above just on index structure ensure the S matrix is invariant under $\mathbb{Z}_2 \times \mathbb{Z}_2$ permutations. To ensure invariance under all of S_4 we must now also ensure invariance of the S matrix under $S_4/(\mathbb{Z}_2 \times \mathbb{Z}_2) = S_3$.

Consider an index structure that happens to be invariant under a subgroup G of S_3 . The coefficient function of s, t and u that multiplies this structure must also be invariant

under this subgroup - which can vary from nothing to all of S_3. We decompose polynomials of s, t and u into the 3 irreps of S_3 , namely the 1 dimensional completely symmetrical representation, the one dimensional completely antisymmetric representation and the 2 dimensional irrep (in which it turns out that every permutation operator (e.g.P_{12}) has eigenvalues ± 1. We find

$$\begin{aligned}
Z_{no-sym} &= \frac{1}{(1-x)^2}, \qquad d_{no-sym}(m) = m+1 \\
Z_{sym} &= \frac{1}{(1-x^2)(1-x^3)}, \qquad d_{sym}(m) \sim \frac{m+1}{6} \\
Z_{as} &= \frac{x^3}{(1-x^2)(1-x^3)}, \qquad d_{as}(m) \sim \frac{m+1}{6} \\
Z_{mixed} &= \frac{2x}{(1-x)(1-x^3)}, \qquad d_{mixed}(m) = \frac{2(m+1)}{3} \\
Z_{\mathbb{Z}_2-sym} &= \frac{1+x}{(1-x)^2}, \qquad d_{\mathbb{Z}_2-sym}(m) = \frac{m+1}{2}
\end{aligned} \tag{4.4}$$

Its clear from the large m values of d above that a structure with no symmetry will give rise to 6 functions worth degrees of freedom because fully symmetric d corresponds to one function worth degree of freedom and d_{sym} grows like $\frac{m+1}{6}$ for large enough m. Similarly, d_{mixed} suggests that a mixed symmetry structure would correspond to 4 function worth degrees of freedom, and $\mathbb{Z}_2 - sym$ structures have 3 function worth degrees of freedom. This is verified in our explicit listing of structures.

As a first exercise we present result for the most general local parity invariant S matrix for 4 photons. For $d \geq 5$ this function is parameterized by 2 $\mathbb{Z}_2$ invariant functions ($A^{0,1}(s,t) = A^{0,1}(t,s)$) and a single S_3 invariant function $A^{2,1}(s,t)$; a total of 7 degrees of freedom. We say a Lagrangian structure A is a descendent of a structure B if first A has more derivatives than B, but all the extra derivatives that are in A but not in B have indices that contract with each other.

We have three primary structures given in terms of Lagrangians. Let $A^{0,1}$ and $A^{0,2}$ parameterize descendents of the four derivative structures $(\mathrm{Tr}F^2)^2$ and $\mathrm{Tr}(F^4)$ respectively while $A^{1,2}$ parameterizes descendents of the six derivative term

$$F_{ab}\mathrm{Tr}(\partial_a F \partial_b F F) \tag{4.5}$$

Note, that in $d = 4$ we have special relation given by

$$F_{ab}F_{c[d}F_{ef}F_{gh]} = 0 \tag{4.6}$$

Because any product of 5 dimensional antisymmetric tensors is trivially 0 in 4 dimensions, as there are not enough indices in 4 dimensions to construct such an object. One might be tempted to write further relations involving $6, 7, 8$ indices for $5, 6, 7$ dimensions respectively. But antisymmetrization on our 4 F structures for 6 or more indices vanishes identically for any dimensions because if we chose any 3 Fs from our 4 Fs the product would necessarily

be symmetric in exchange of two of the Fs. We go on to parametrize the photon S matrix explicitly written in terms of polarization vector and momenta of the photons.

As a check to the claim that these three structures form a basis for any $2-2$ scattering, we write the tree level scattering of 4 photons in Type-I theory (or Type II theory on D branes). It has a single index structure given in our basis as:

$$L_{4V}^{SS} \propto \frac{1}{16}\left(\mathrm{Tr}(F^4) - \frac{1}{4}(\mathrm{Tr}(F^2))^2\right) \tag{4.7}$$

which itself can be obtained by expanding the Born Infeld action to quartic order in F . Consequently this scattering amplitude can be cast into our general form with $A^{2,1} = 0$ and $A^{0,2} = \frac{1}{4}A^{0,1}$. The expression for $A^{0,1}$ is a well known Veneziano type function.

We have also recast the formula for tree level scattering in the open bosonic string into our general form. It involves all three of our structures.

Now, we present the result for the most general parity invariant S matrix for 4 graviton scattering.

For $d \geq 7$ the most general S matrix turns out to be parameterized by 7 $\mathbb{Z}_2$ invariant functions, one function that enjoys no permutation symmetry and two functions that are completely permutation symmetric. A total of 29 degrees of freedom.

Listing these primary terms with derivatives orders, we have one term at six derivative order. This multiplies a fully symmetric function of s, t, u. Going up in derivatives, at eight derivative order we have 5 $\mathbb{Z}_2$ invariant terms and one term with no additional symmetry. This gives a total of $15 + 6 = 21$ degrees of freedom parameterizing a general descendent coming from these structures. At degree ten we have 2 $\mathbb{Z}_2$ invariant structures giving 6 degrees of freedom for parameterizing descendants from these primary structures. Finally, we have a single completely symmetric function (one degree of freedom) parameterizing descendents of a 12 derivative terms (4 derivatives of 4 Reimanns). We explicitly list this parameterization in the later sections of this chapter.

As a check of our parametrization we write the 4 graviton S matrix from the Einstein Lagrangian, in terms of our 10 basis structures of momenta and polarization as mentioned above. Then we construct the 4 Riemann term which, when multiplied by $\frac{1}{stu}$, generates the same S matrix at tree level as the Einstein action.

Next, the 4-graviton amplitude in Type II superstring theory is proportional (in the sense of index structure) to the S matrix for Einstein gravity, and so can also be easily written in our basis.

$$A_{4h}^{SS} = h(s, t, u, \alpha')A_{4h}^{EG} \tag{4.8}$$

The tree level S matrices for the Heterotic string and the bosonic string are more complicated, but also can each be written as a linear combination of the last 9 structures we discussed above. The first structure - descendents of the 6 derivative term - never appears in tree level string amplitudes. It would be interesting to check whether this structure appears in string loop amplitudes. We have not yet tried this.

The final goal of this work is to use the principle of causality to constrain all the possible classical effective theories of gravity (and electrodynamics) upto 4 Riemann (and field strength, $F_{\mu\nu}$) level. This is a work for the future.

Now, to systematically proceed to do our computation, we first get the answer from the CFT side. Using group theory arguments first we constrain the number of possible structures and their appropriate symmetries. The counting from the CFT side is already done in the [132]. We present it in the next section.

4.2 CFT four point structures

In this section we compute the number of index structures for the four point function of identical spin 1 and spin 2 currents. This section is a review of the relevant parts, for our counting, from [132] and [136].

As the spin-1 operator is conserved, effectively it can be thought of as a vector of $SO(d-1)$ obtained after subtracting the scalar degree of freedom of the conservation equation [132]. The index structures of the four point function are the same as number of $SO(d-2)$ singlets in the tensor product of four $SO(d-1)$ vectors. This is not exactly true, as one needs to take into account the permutation symmetry of the external operators. We expect the four point function of identical operators to be invariant under S_4, the group of permutations of four objects. The four point function is written as a function of cross-ratios, the invariance of the four point function under permutations that do not leave cross-ratio fixed is known as the crossing symmetry. This symmetry imposed via the crossing equation on the CFT side or by summing over s,t and u channels on the scattering side. Only the left-over symmetry, $S_4/S_3 = \mathbb{Z}_2 \times \mathbb{Z}_2$ is to be imposed at the outset. It is important to understand that these $\mathbb{Z}_2$'s act as a simultaneous exchange of pairs of external operators and they leave the cross-ratios invariant. So the total number of four-point structures is

$$N = \Big(\mathrm{Res}^{SO(d-1)}_{SO(d-2)}\Big(\otimes^4_{i=1} \rho\Big)^{\mathbb{Z}_2\times\mathbb{Z}_2}\Big)^{SO(d-2)}. \tag{4.9}$$

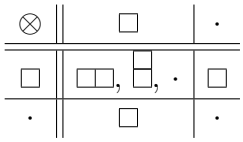
Here we have used the notations in [132]. This is computed by the representation ring identity

$$\Big(\otimes^4_{i=1} \rho\Big)^{\mathbb{Z}_2\times\mathbb{Z}_2} = \rho^4 \ominus 3\Big(S^2\rho \otimes \wedge^2\rho\Big). \tag{4.10}$$

Let us compute this for a vector representation of $SO(d-1)$.

$$\mathrm{Res}^{SO(d-1)}_{SO(d-2)} \square = \square \oplus \bullet \tag{4.11}$$

We will compute the representation appearing in $S^2\rho$ and $\wedge^2\rho$.



	$\#S^2$	$\#\wedge^2$
$\square\square$	1	0
$\begin{smallmatrix}\square\\\square\end{smallmatrix}$	0	1
$\square$	1	1
$\cdot$	2	0

$$N = 1^2 + 1^2 + 2^2 + 2^2$$
$$-3(1\cdot 0 + 0\cdot 1 + 1\cdot 1 + 2\cdot 0) = 7. \tag{4.12}$$

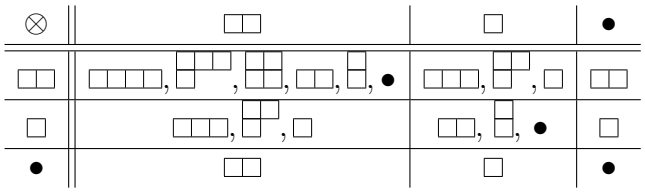
The counting above assumes large enough dimensions, CFT dimension $d \geq 4$. The Lorentz group in lower dimensions is special and there are identities between these 7 structures causing these orbits to shrink to a total of 5 terms in $d = 3$. This is described in [132] and Appendix A in [136].

In 3-d the restriction of (4.11) of the spin-1 traceless symmetric tensor of $O(2)$ to $O(1)$ is just

$$\text{Res}^{O(2)}_{O(1)} \square = \bullet^+ \oplus \bullet^- \tag{4.13}$$

Where superscript $+$ is for parity even and $-$ is for parity odd. Plugging this in (4.10) we immediately see that there are 5 parity even structures and 2 parity odd structures. Its interesting to note that this result is independent of spin l of the current, i.e. on the l.h.s. of (4.13) we could have taken any spin l representation the r.h.s. would remain the same in this special case. Therefore, even for the case of spin-2 current, we will have the same counting of structures in $d = 3$ (bulk dimensions 4).

Let us do a similar counting for the spin-2 conserved current. It decomposes into $SO(d-2)$ as $\square\square \oplus \square \oplus \bullet$.



	$\#S^2$	$\#\wedge^2$
$\square\square\square\square$	1	0
$\begin{smallmatrix}\square\square\square\\\square\end{smallmatrix}$	0	1
$\begin{smallmatrix}\square\square\\\square\square\end{smallmatrix}$	1	0
$\square\square\square$	1	1
$\begin{smallmatrix}\square\square\\\square\end{smallmatrix}$	1	1
$\square\square$	3	1
$\begin{smallmatrix}\square\\\square\end{smallmatrix}$	0	2
$\square$	2	2
$\cdot$	3	0

(4.14)

From here it follows that the number of four point structure for spin-2 current is

$$\begin{aligned} N &= 1^2 + 1^2 + 1^2 + 2^2 + 2^2 + 4^2 + 2^2 + 4^2 + 3^2 \\ &\quad -3(1\cdot 0 + 0\cdot 1 + 1\cdot 0 + 1\cdot 1 + 1\cdot 1 + 3\cdot 1 + 0\cdot 2 + 2\cdot 2 + 3\cdot 0) \\ &= 29. \end{aligned} \tag{4.15}$$

This counting also assumes large enough dimensions, namely CFT dimensions $d \geq 6$.

In the case of spin 2 conserved current, there are two more special cases apart from $d = 3$. First, there is the case of $d = 5$. There is a degeneracy condition mentioned in section 2.7 and Appendix A of [136], which results in the reduction from 29 function worth degrees of freedom to 28 degrees of freedom.

Secondly, the case of $d = 4$. [132]. In this case we have to consider the restriction of parity even spin 2 representation of $O(3)$ to $O(2)$, $\square\!\square \oplus \square \oplus \bullet^+$. This representation has the character

$$\chi_{(\square\!\square\oplus\square\oplus\bullet^+)}(x,s) = \frac{1+s}{2}\left(\frac{x^{\frac{5}{2}} - x^{-\frac{5}{2}}}{x^{\frac{1}{2}} - x^{-\frac{1}{2}}}\right) + \frac{1-s}{2} \tag{4.16}$$

Where,

$$O(2) = U(1) \times \mathbb{Z}_2 = \{(x,s) : x \in U(1), s = \pm 1\} \tag{4.17}$$

Plugging this in the (4.10) we find that the total number of parity even structures is 22 and parity odd structures is 3. Parity odd structures count is just the coefficient of s in the character of product of 4 of the above mentioned representations, i.e. (4.9).

4.3 S-matrices

In this section we review the relevant parts of [134, 135]. We also explicitly list all possible photon and graviton S-matrices and find them to be in one-to-one correspondence with the CFT four point structures. The S-matrix is a scalar function of four polarizations ϵ_i and four momenta p_i with the following properties

$$\begin{aligned} f(\lambda_i \epsilon_i, p_i) &= \Big(\prod_i^4 \lambda_i^{\ell_i}\Big) f(\epsilon_i, p_i) \\ f(\epsilon_i + \alpha_i p_i, p_i) &= f(\epsilon_i, p_i) \\ \text{with} \qquad & p_i^2 = 0, \quad \epsilon_i \cdot p_i = 0. \end{aligned} \tag{4.18}$$

Here ℓ_i is the spin of the i-th operator. This is exactly the same algebraic problem as the classification of conformal four point function of spinning operators in embedding space. It has been studied in [134, 135]. The function with above properties can be constructed using only two types of structures,

$$h_{ij} = (\epsilon_i \cdot \epsilon_j)(p_i \cdot p_j) - (\epsilon_i \cdot p_j)(\epsilon_j \cdot p_i), \tag{4.19}$$

$$v_{i,jk} = (\epsilon_i \cdot p_j)(p_i \cdot p_k) - (\epsilon_i \cdot p_k)(p_i \cdot p_j). \tag{4.20}$$

The structure h_{ij} contributes homogeneity 1 to ϵ_i and ϵ_j each, while $v_{i,jk}$ contributes homogeneity 1 only to ϵ_i. The S-matrix is constructed with h's and v's such that the required homogeneity in all the ϵ's is obtained. For a given i, j, the structure h_{ij} is unique while a priory, there are multiple $v_{i,jk}$ structures for a given i. For example a single homogeneity in ϵ_1 is obtained by $v_{1,23}, v_{1,34}$ and $v_{1,24}$. For CFT four point functions, where the role of p_i is played by the embedding space coordinate X_i, it turns out there are

only two of these structures that are linearly independent. In the context of S-matrices however, these structures are even more constrained thanks for momentum conservation. After defining $s = (p_1 + p_2)^2/2, t = (p_1 + p_3)^2/2, u = (p_1 + p_4)^2/2$,

$$\begin{aligned} v_{1,23} &= (\epsilon_1 \cdot p_2)(p_1 \cdot p_3) - (\epsilon_1 \cdot p_3)(p_1 \cdot p_2) \\ &= (\epsilon_1 \cdot p_2)t - (\epsilon_1 \cdot p_3)s \\ &= -(\epsilon_1 \cdot (p_1 + p_3 + p_4))t - (\epsilon_1 \cdot p_3)(-t - u) \\ &= -(\epsilon_1 \cdot p_4)t + (\epsilon_1.p_3)u \\ &= -v_{1,43} \end{aligned} \tag{4.21}$$

Similarly, $v_{1,23} = -v_{1,24}$. This implies that for a given i, $v_{i,jk}$ structure is unique, we can change either j or k in $v_{i,jk}$ at the cost of a minus sign. Let us define a convention $v_i \equiv v_{i,(i+1,i+2)}$. Now the S-matrix is to be written only using h_{ij} and v_i, both these structures are unique for the given value of their indices. We write the S-matrix structures using a standard graphical notation, h_{ij} are denoted as a line between the vertices i and j after drawing the lines, all remaining homogeneity is saturated by v's. This is elaborated in figure 4.3

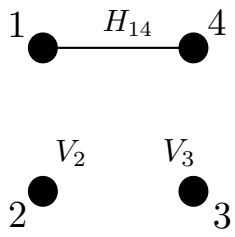


Figure 4.1: This figure denotes the S-matrix $h_14v_2v_3$ for photons. The same diagram also has an interpretation for graviton S-matrix, in that case it represents the S-matrix $h_{14}v_1v_4v_2^2v_3^2$.

With this notation, we have computed the possible gauge invariant 4-photon S-matrix structures as well as 4-graviton S-matrix structures in figure 4.2 and 4.3 respectively. In both cases, the number of structures agrees with the group theory counting done in the previous section.

These results are only true in $d \geq 4$ for the case of photon and $d \geq 6$ for the case of graviton. The identities between h_{ij} and $v_{i,jk}$ for $d = 3$ are are discussed in Appendix C of [133]. In the case of $d = 3$ where there is a reduction from 7 degrees of freedom to 5, but number of structures remain the same, as was done in the counting in the previous section. Similar identities between h_{ij} and $v_{i,jk}$ for the case of $d = 4, 5$ for the case of graviton are not listed but from the counting of the previous section we know how many structures reduce in these dimensions for the case of gravitons.

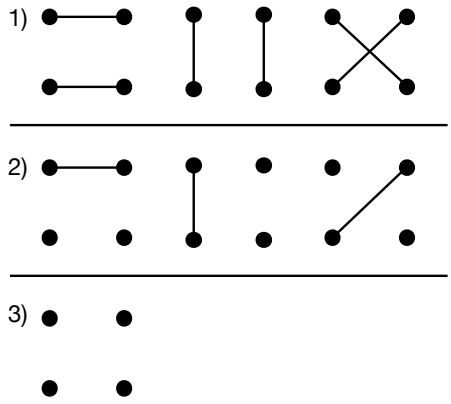


Figure 4.2: Linearly gauge invariant photon S-matrices

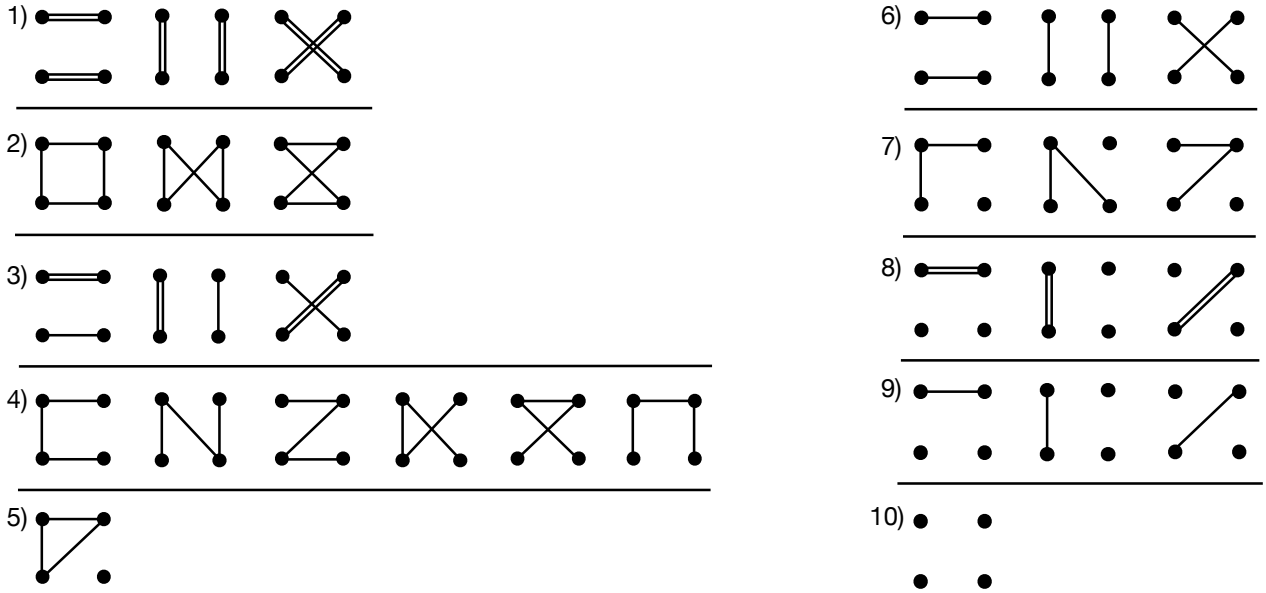


Figure 4.3: Linearly gauge invariant graviton S-matrices

Now, with the counting of previous section in mind, we start to write the Lagrangian structures for the bulk theories.

4.4 S matrices for 4 scalars - enumeration and counting

Consider the S matrix of a theory of four identical massless scalars. Clearly the most general S matrix in such a theory is specified by a single function

$$A(s,t)$$

which enjoys the following symmetry properties

$$A(s,t) = A(t,s) = A(s,u), \quad u = -s - t \tag{4.22}$$

(4.22) reflects the fact that the S matrix enjoys invariance under the full permutation group, and so in particular interchanges $2 \leftrightarrow 3$ and $2 \leftrightarrow 4$ respectively. Of course s, t and u are the usual Mandlestam invariants that obey

$$s + t + u = 0 \tag{4.23}$$

If we focus attention on S matrices that arise from local Lagrangians, the functions $A(s,t)$ are necessarily polynomials of s and t and u. These polynomials can be sorted by their degree. At degree zero we have a constant - corresponding to the Lagrangian term ϕ^4. The only completely symmetric polynomial at degree one is $s+t+u$ which vanishes by (4.23). At degree two we have two symmetric polynomials $s^2+t^2+u^2$ and $st+tu+us$. However the equation $(s+t+u)^2 = 0$ gives us one relation between these two polynomials. The independent polynomial can be taken to be $st + tu + us$ and is generated by the Lagrangian term

$$\partial_\mu \partial \nu \phi \partial_\mu \phi \partial_\nu \phi \phi \tag{4.24}$$

At degree three we have 3 independent symmetric polynomials, $s^3+t^3+u^3$, $s^2(t+u)+t^2(s+u)+u^2(s+t)$ and stu. The equations $(s+t+u)^3 = 0$ and $(s+t+u)(st+tu+us) = 0$ give us three relations between these three independent quantities. The independent polynomial can be taken to be stu and is generated by the Lagrangian structure

$$\partial_\mu \partial \nu \partial_\alpha \phi \partial_\mu \phi \partial_\nu \phi \partial_\alpha \phi \tag{4.25}$$

This enumeration can be continued indefinitely. In order to see how the counting works let us compute two partition function

$$Z_n(x) = Trx^d = \sum_{m=0}^{\infty} d_n(m) x^m, \quad Z_{sym}(x) = Trx^d = \sum_{m=0}^{\infty} d_{sym}(m) x^m \tag{4.26}$$

(the subscript n stands for naive) where the 'trace' in Z_n i taken over all symmetric polynomials, the trance in Z is taken over all equivalence classes of symmetric polynomials (polynomials that agree when $s+t+u=0$ are regarded as identical). $Z_n(x)$ is simply the partition function of three identical particles in a harmonic oscillator and is given by

$$Z_n(x) = \frac{1}{3(1-x^3)} + \frac{1}{6(1-x)^3} + \frac{1}{2(1-x)(1-x^2)} \tag{4.27}$$

As we have seen in the examples above, the relationship between $d_n(m)$ and $d_{sym}(m)$ is

$$d_{sym}(m) = d_n(m) - \sum_{p=0}^{m-1} d_{sym}(p) \tag{4.28}$$

(the last term on the RHS reflects the fact that any lower order nontrivial polynomial can be multiplied by the appropriate power of $(s+t+u)$ to generate a symmetric polynomial at degree m, as we have seen in examples). It follows that

$$d_n(m) = \sum_{p=0}^{m} d_{sym}(p) \tag{4.29}$$

so that

$$d(p) = d_n(p) - d_n(p-1) \tag{4.30}$$

It follows that

$$\begin{aligned} Z_{sym}(x) = (1-x)Z_n(x) =& \frac{1}{3(1+x+x^2)} + \frac{1}{6(1-x)^2} + \frac{1}{2(1-x^2)} \\ &= \frac{1}{(1-x^2)(1-x^3)} \\ &= 1 + x^2 + x^3 + x^4 + x^5 + 2x^6 + x^7 + 2x^8 + 2x^9 + 2x^{10} \dots \end{aligned} \tag{4.31}$$

It is straightforward but unilluminating to expand (4.31) to obtain an analytic expression for $d(m)$. At large m

$$d_{sym}(m) \sim \frac{m+1}{6} \tag{4.32}$$

Recall that $d(m)$ counts the number of independent n Lagrangian structures that contribute to the scattering of 4 identical scalars.

4.5 S matrices for 4 photon scattering

In section 4.2 above we have already counted the number of independent index structures that can appear in 4 photon S matrices. For $D \geq 5$ we found that there was a total of 7 such structures, which transformed in under the S^3 'residual permutation symmetry' group as 3 structures in the (singlet) completely symmetric representation and two structures in the unique 2 dimensional representation of S^3. In this section we explicitly construct the corresponding S matrix index structures. As our classification follows just from simple symmetry considerations, it follows that the full 4 photon S matrix of any theory - classical or quantum - is given by a linear combination of the index structures listed in this section. The coefficients of this linear combination are, in general, arbitrary functions (subject to certain symmetry requirements, see below) of the Mandelstam variables s, t, u.

In this chapter we are especially interested in S matrices generated by the local part of the 4 photon quantum effective action of our theory, i.e. the S matrices that we can obtain from local (i.e with finite number of derivatives) gauge invariant bulk interaction terms. As a special case of the construction of the most general 4 photon S matrix (described in the previous paragraph) we also present an explicit parameterization of all inequivalent

local four photon effective actions [1]. Our final answer is given in terms of 3 infinite classes of Lagrangian terms - terms in the same class generate 4 photon S matrices with the same index structure but with different dependences on s, t, u. Our final answers are very explicit.

4.5.1 Direct Enumeration

Our method to construct 4 photon S matrices is completely straightforward (if slightly tedious). We systematically list all gauge invariant local Lagrangian structures that can contribute to four photon scattering subject to the following equivalences

- Terms in the Lagrangian that differ by terms that vanish when we impose the linearized equations of motion $\partial^\mu F_{\mu\nu}$ can be converted to each other by the an appropriate field redefinition. Such terms generate the same 4 photon S matrix and so are regarded as equivalent.

- Terms in the Lagrangian that differ by a total derivative generate the same S matrix and so are regarded as equivalent.

- Terms in the Lagrangian that involve the dot product - i.e. an index contraction - of two separate derivatives (e.g.

  $$B = \mathrm{Tr}(\partial_\mu F_{ab} \partial^\mu F_{bc} F_{cd} F_{da})$$

  generate S matrices with the same index structures. Consequently, if we are interested only in the index structures of S matrices, all such terms may be regarded as equivalent.

- Let us now specialize - as we do in much of this section - to the study of local Lagrangians. It is useful to define the notion of a 'descendent' in the space of local Lagrangians. A term B is said to be a descendent of a term C if B can be obtained from C by acting on the fields that appear in C by derivatives whose indices contract in pairs. For instance the term B above is a descendent of

  $$C = \mathrm{Tr}(F_{ab} F_{bc} F_{cd} F_{da})$$

  It is also useful to define the notion of a 'primary' in the space of Lagrangians. A primary is a Lagrangian structure that cannot itself be written as the 'descendent' of some other another local Lagrangian. In order to classify local Lagrangian terms it is useful to first construct all inequivalent primary structures. With these in hand it is straightforward to construct all descendents, and so all S matrices that arise out of local Lagrangian structures. In the first part of our analysis - where we seek to classify all inequivalent primaries - we simply set all descendents to zero; two terms that differ by a descendent are regarded as equivalent.

[1] We regard local effective actions as equivalent when they generate the same S matrix, or, equivalently, when they can be related to each other by a filed redefinition of a specified sort, see below.

4.5.2 Triviality of 3 photon scattering

As a warm up to the study of 4 photon S matrices, it is useful to first rederive a well known fact, namely that there are no nontrivial index structures for 3 photon scattering; 3 photon scattering simply vanishes.

Consider a term in the Lagrangian with three explicit factors of the field strength. We can always use integration by parts to transfer all derivatives that act on the third field strength to the first and second ones. So the most general Lagrangian term of this sort takes the schematic form

$$\partial\partial...\partial F\ \partial\ldots\partial F\ F \tag{4.33}$$

Now let us examine the index structure of any such term. Focusing on primaries, none of the indices of the derivative operators ∂_μ are allowed to contract with each other. Also, when ∂_μ acts on a given F, the μ index is not allowed to contract with an index of that F, because when that happens the corresponding term vanishes because of the equation of motion. It follows that when ∂_μ derivative acts on the first F must contract with a free index on either the second F or the third F. Now suppose we have two derivatives, say ∂_μ and ∂_ν, acting on the first F. Then one of the two indices (say μ) must contract with a free index of the second F while the other one - ν contracts with a free index of the third F (if both indices contract with the free indices of - say - the second F, the corresponding term vanishes because of antisymmetry of $F_{\mu\nu}$).

In summary primaries formed out of 3 Fs cannot have more than four derivatives, and the unique candidate for a four derivative primary is

$$T_4 = \partial_\mu\partial_\theta F_{a\nu}\partial_\nu\partial_\phi F_{a\mu}F_{\theta\phi} \tag{4.34}$$

The reader can, however, easily verify that the swap of dummy variables

$$\mu \leftrightarrow \nu, \quad \theta \leftrightarrow \phi, \quad a \leftrightarrow b \tag{4.35}$$

allows us to prove that $T_4 = -T_4$, so that T_4 vanishes. It follows that there are no three F primaries with four derivatives.

Let us now turn to two derivative primaries. The reader can easily convince herself that any candidate two derivative primary can, by an integration by parts (and by use of the equation of motion), be manipulated into one of the following two terms

$$\begin{aligned} T_2^a &= \partial_\mu F_{\nu a}\partial_\nu F_{\mu b}F_{ab} \\ T_2^b &= \partial_\mu F_{ab}\partial_\nu F_{ab}F_{\mu\nu} \end{aligned} \tag{4.36}$$

[2] However the change of dummy variables (4.35) can, once again, be used to demonstrate that $T_2^a = -T_2^a$ and $T_2^b = -T_2^b$. It thus follows that each of T_2^a and T_2^b vanish identically.

Finally, the unique operator with 3 Fs and no derivatives is simply $\mathrm{Tr}F^3$ (where the trace is taken over the Lorentz indices of the antisymmetric $D \times D$ matrix F). Using

$$\mathrm{Tr}F^3 = \mathrm{Tr}(F^3)^T = \mathrm{Tr}(F^T)^3 = -\mathrm{Tr}F^3$$

[2] In particular, integration by parts with ∂_μ takes $\partial_\mu F_{\theta\phi}\partial_\nu F_{\mu\theta}F_{\nu\phi}$ to T_2^a.

it follows that this term also vanishes identically. We conclude that there exist no primaries with 3 F fields.

4.5.3 4 photon scattering

Clearly all gauge invariant contributions to the Lagrangian that contribute to 4 photon scattering consist of products of derivatives (of arbitrary number) multiplying 4 $F_{\mu\nu}$ fields in such a way that all indices contract so that the Lagrangian is a scalar. It is very easy to see that no term involving 6 or more derivatives (distributed and contracted in any manner among the 4 $F_{\mu\nu}$ operators) can yield a 'primary' contribution to the 4 photon S matrix. In order to see why this is the case, suppose it were not true. Then there must be a scalar expression built out of 4 $F_{\mu\nu}s$ and 6 derivatives in which none of the derivatives contract with each other. It follows that both indices of atleast two F_{ab} operators must contract with derivatives. A candidate term of this term might be

$$\partial_a F_{\mu\nu} \partial_\mu F_{ab} \partial_b \partial_\nu \partial^p F_{mn} \partial^m F_{pn} \tag{4.37}$$

in which the indices of the first two field strength operators are both contracted with derivatives. In order to see that the term above is trivial (according to our rules of counting as spelt out above) we use the the Bianchi identity

$$\partial_a F_{\mu\nu} = -\partial_\mu F_{\nu a} - \partial_\nu F_{a\mu}$$

to re express the first field field strength in (4.37) as a sum of two other terms. This gives us a sum of terms, each of which is a product of four field strengths. Note, however, that both of these terms have a pair of derivatives with contracted indices, and so both terms are descendents, and so are trivial for the purposes of the current section.

The reader can easily convince herself that exactly the same argument can be made whenever two separate field strength operators have both their indices contracted with derivatives. Let the two field strengths of this form be the 'first' and the second $F_{\mu\nu}$ operators in the expression. The two derivatives that contract with the second $F_{\mu\nu}$ must act on distinct $F_{\alpha\beta}$ fields (else the expression would vanish by the antisymmetry of F_{ab}). Moreover neither of these derivatives can act on the second field itself (else the expression would vanish by the equations of motion). An integration by parts can be used to ensure that none of the derivatives act on the 'fourth' $F_{\mu\nu}$. With this convention it follows that one of the two derivatives that contracts with the second $F_{\mu\nu}$ must act on the 'third' $F_{\mu\nu}$ while the second derivative must act on the first $F_{\mu\nu}$. We can now replace the expression involving the derivative acting on the first $F_{\mu\nu}$ by two different terms via the Bianchi identity. A moment's consideration will convince the reader that both these terms are descendents and so can be dropped.

Let us now turn to terms involving four derivatives acting on the four $F_{\mu\nu}$ operators. The reader can quickly convince herself that there are five terms of this sort that are not

obviously trivial. These terms are

$$\begin{aligned}
T_1^4 &= \partial_\delta \partial_\nu F_{\alpha a} F_{\beta a} \partial_\alpha \partial_\beta F_{\nu\beta} F_{\delta b} \\
T_2^4 &= \partial_\nu F_{\alpha a} \partial_\delta F_{\beta a} \partial_\alpha F_{\nu b} \partial_\beta F_{\delta b} \\
T_3^4 &= \partial_\delta \partial_\gamma F_{a\alpha} F_{a\beta} \partial_\alpha F_{b\gamma} \partial_\beta F_{b\delta} \\
T_4^4 &= \partial_\beta \partial_d F_{ab} F_{bc} \partial_\alpha \partial_a F_{cd} F_{\alpha\beta} \\
T_5^4 &= \partial_\gamma \partial_\alpha F_{ab} \partial_\delta \partial_\beta F_{ba} F_{\alpha\beta} F_{\gamma\delta}
\end{aligned} \tag{4.38}$$

(Any other expression that the reader may care to write down can be manipulated into one of the five forms above upto total derivatives - without, at this stage, the use of Bianchi identities).

In Appendix C.1 we employ Bianchi identities to find relations between the structures T_i^4 $(i = 1 \to 5)$. The relations we obtain turn out to be strong enough to allow us to deduce that each of the terms listed in (4.38) actually are actually trivial. It follows that there is no Lagrangian structure built out of four field strengths and four derivatives that generates a 'primary' S matrix.

The situation is a bit more complicated with terms involving two derivatives of the four field strengths. By using the equivalence of terms that differ by total derivatives, the reader can convince herself that there there are fourteen naively inequivalent structures at this order. They are

$$\begin{aligned}
T_1^2 &= \partial_b F_{\beta a} \partial_a F_{\alpha b} F_{\theta\alpha} F_{\theta\beta} \\
T_2^2 &= \partial_b F_{\beta a} \partial_a F_{\beta b} F_{\mu\nu} F_{\mu\nu} \\
T_3^2 &= \partial_b F_{\beta a} \partial_a F_{\alpha b} F_{\theta\beta} F_{\theta\alpha} \\
T_4^2 &= F_{\beta a} \partial_a F_{\alpha b} \partial_b F_{\theta\alpha} F_{\theta\beta} \\
T_5^2 &= F_{\beta a} \partial_a F_{\beta b} \partial_b F_{\mu\nu} F_{\mu\nu} \\
T_6^2 &= F_{\beta a} \partial_a F_{\alpha b} \partial_b F_{\theta\beta} F_{\theta\alpha} \\
T_7^2 &= \partial_\beta F_{ba} F_{\theta\beta} \partial_a F_{b\alpha} F_{\theta\alpha} \\
T_8^2 &= F_{\alpha a} F_{\beta b} \partial_a F_{\alpha\theta} \partial_b F_{\beta\theta} \\
T_9^2 &= F_{\alpha a} F_{\beta b} \partial_b \partial_a F_{\alpha\theta} F_{\beta\theta} \\
T_{10}^2 &= F_{b\alpha} F_{a\alpha} \partial_b F_{\mu\nu} \partial_a F_{\mu\nu} \\
T_{11}^2 &= F_{b\alpha} F_{a\alpha} \partial_a \partial_b F_{\mu\nu} F_{\mu\nu} \\
T_{12}^2 &= F_{\alpha a} F_{\beta b} \partial_a F_{\beta\theta} \partial_b F_{\alpha\theta} \\
T_{13}^2 &= F_{\alpha a} F_{\beta b} \partial_a \partial_b F_{\beta\theta} F_{\alpha\theta} \\
T_{14}^2 &= F_{ab} \partial_a F_{\mu\nu} \partial_b F_{\nu\rho} F_{\rho\mu}
\end{aligned} \tag{4.39}$$

T_1^2 and T_3^2 are equivalent upto relabeling.

Once again these naively independent structures are not really all distinct; once again Bianchi identities may be used to relate these 14 structures. In Appendix C.2 we demonstrate that the use of Bianchi identities generates 13 nontrivial identities between the

structures listed in (4.39). These identities can use used to relate each of these structures to a single independent term which we choose to be

$$T_I^2 = F_{ab}\partial_a F_{\mu\nu}\partial_b F_{\nu\rho}F_{\rho\mu} \tag{4.40}$$

4.5.4 Most General Local S matrix for 4 photon scattering

We have demonstrated in the last subsection that the three primaries for for photon scattering are

$$S^1 : \ (\mathrm{Tr}F^2)^2, \quad S^2 : \ \mathrm{Tr}F^4, \quad S^3 : \ F_{ab}\partial_a F_{\mu\nu}\partial_b F_{\nu\rho}F_{\rho\mu} \tag{4.41}$$

In this subsection we will explicitly parameterize the most general 'descendent' S matrix built out of these primaries.

Let us first consider descendents of $(\mathrm{Tr}F^2)^2$. Every such descendent is a linear combination of the terms

$$A^{0,1}_{m,n} = \prod_{i=1}^{m}\prod_{j=1}^{n} \mathrm{Tr}\left(\partial_{\nu_j}\partial_{\mu_i}FF\right)\mathrm{Tr}\left(\partial_{\mu_i}F\partial_{\nu_j}F\right) \tag{4.42}$$

In order to see this let us first note that we can always choose, by an integration by parts , to have no derivative acting on the second appearance of F in (4.42). Once we adopt this convention, it follows that all descendent derivatives that act on the first F necessarily have indices that contract with a derivative that acts either on the third or on the fourth F as we have displayed in (4.42). In addition to the terms displayed in (4.42) we could have descendent derivatives acting on the third and fourth F contracting with each other. However these terms are easily manipulated into terms of the form (4.42) using integration by parts and the equations of motion. [3]

It is manifest from the expression (4.42) that

$$A^{0,1}_{m,n} = A^{0,1}_{n.m} \tag{4.43}$$

[3]For instance

$$\mathrm{Tr}\left(F\partial_\mu F\right)\left(\partial_\mu FF\right) = -\mathrm{Tr}\left(\partial_\mu F\partial_\mu F\right)\left(FF\right) - \mathrm{Tr}\left(F\partial_\mu F\right)\left(F\partial_\mu F\right)$$

Using the equation of motion $\partial^2 F = 0$, the second term in the expression above can be written as

$$-\frac{1}{2}(\partial_\mu^2+\partial_\mu^4)^2\mathrm{Tr}\left(FF\right)\left(FF\right) = -\frac{1}{2}(\partial_\mu^1+\partial_\mu^3)^2\mathrm{Tr}\left(FF\right)\left(FF\right) = -\mathrm{Tr}\left(\partial_\mu FF\right)\left(\partial_\mu FF\right)$$

Where the notation ∂_μ^a denotes a derivative that acts only on the a^{th} appearance of F. In the manipulation in the equation above we have first integrated by parts twice and then used the equation of motion. In summary we conclude that

$$\mathrm{Tr}\left(F\partial_\mu F\right)\left(\partial_\mu FF\right) = -A^{0,1}_{1,0} + -A^{0,1}_{0,1}.$$

While the manipulation in this footnote might seem complicated in position space, it is completely standard in momentum space. The manipulation simply demonstrates the familiar fact that $p_2.p_4 = p_1.p_3$ using

$$p_2.p_4 = \frac{1}{2}(p_2+p_4)^2 = \frac{1}{2}(p_1+p_3)^2 = p_1.p_3.$$

It is convenient to package the information of the parameters $A^{0,1}_{m,n}$ into a generating function

$$A^{0,1}(t,u) = \sum_{m,n=0}^{\infty} A^{0,1}_{m.n} t^m u^n \tag{4.44}$$

It follows from (4.43) that

$$A^{0,1}(t,u) = A^{0,1}(u,t) \tag{4.45}$$

The most general local descendent of $(\mathrm{Tr}F^2)^2$ is parameterized by the most general polynomial function $A^{0,1}(t,u)$ that is symmetric under interchange of t and u. [4] Explicitly, the S matrix generated by the descendent parameterized by $A^{0,1}(t,u)$ is given by

$$\begin{aligned} &A^{0,1}(t,u)\left(p^1_\mu\epsilon^1_\nu - p^1_\nu\epsilon^1_\mu\right)\left(p^2_\mu\epsilon^2_\nu - p^2_\nu\epsilon^2_\mu\right)\left(p^3_\alpha\epsilon^3_\beta - p^3_\beta\epsilon^3_\alpha\right)\left(p^4_\alpha\epsilon^4_\beta - p^4_\beta\epsilon^4_\alpha\right) \\ &+A^{0,1}(s,u)\left(p^1_\mu\epsilon^1_\nu - p^1_\nu\epsilon^1_\mu\right)\left(p^3_\mu\epsilon^3_\nu - p^3_\nu\epsilon^3_\mu\right)\left(p^2_\alpha\epsilon^2_\beta - p^2_\beta\epsilon^2_\alpha\right)\left(p^4_\alpha\epsilon^4_\beta - p^4_\beta\epsilon^4_\alpha\right) \\ &+A^{0,1}(t,s)\left(p^1_\mu\epsilon^1_\nu - p^1_\nu\epsilon^1_\mu\right)\left(p^4_\mu\epsilon^4_\nu - p^4_\nu\epsilon^4_\mu\right)\left(p^3_\alpha\epsilon^3_\beta - p^3_\beta\epsilon^3_\alpha\right)\left(p^2_\alpha\epsilon^2_\beta - p^2_\beta\epsilon^2_\alpha\right) \end{aligned} \tag{4.46}$$

In a similar manner, the most general descendent of the primary $\mathrm{Tr}(F^4)$ is a linear combination of the quantities

$$A^{0,2}_{m,n} = \prod_{i=1}^{m}\prod_{j=1}^{n} \mathrm{Tr}\left(\partial_{\nu_j}\partial_{\mu_i}F\partial_{\mu_i}FF\partial_{\nu_j}F\right) \tag{4.47}$$

Using

$$\mathrm{Tr}\left(\partial_{\nu_j}\partial_{\mu_i}F\partial_{\mu_i}FF\partial_{\nu_j}F\right) = \mathrm{Tr}\left(\partial_{\nu_j}FF\partial_{\mu_i}F\partial_{\nu_j}\partial_{\mu_i}F\right) = \mathrm{Tr}\left(\partial_{\nu_j}\partial_{\mu_i}F\partial_{\nu_j}FF\partial_{\mu_i}F\right)$$

(where we have taken the transpose in the first step and used cyclicity of the trace in the second step) it follows that

$$A^{0,2}_{m,n} = A^{0,2}_{n.m} \tag{4.48}$$

As above, the information in the coefficients $A^{0,2}_{m,n}$ is most conveniently packaged into the generating function

$$A^{0,2}(t,u) = \sum_{m,n=0}^{\infty} A^{0,2}_{m.n} t^m u^n \tag{4.49}$$

It follows from (4.48) that

$$A^{0,2}(t,u) = A^{0,2}(u,t) \tag{4.50}$$

Note that [5]

[4] More generally, the most general S matrix with the same index structure as that generated by $(\mathrm{Tr}F^2)^2$ - whether local or not - is parameterized by the most general function $A^{0,1}(t,u)$ that is symmetric under the interchange of t and u.

[5] We have used the symbol t rather than the symbol s for contractions between derivatives acting on the first and second appearances of F in (4.42) because we are using conventions in which the field strengths that appear in (4.42) correspond, in order, to the first, third, second and fourth particles respectively (and then cyclically sum. With this convention the S matrix has a $1 \leftrightarrow 2$ symmetry (rather than, for instance, a $1 \leftrightarrow 3$ symmetry), as was the case with (4.46). While this choice is convenient it carries no physical content. 1, 2, 3,4 are dummy variables in the S matrix (4.46).

As above the most general descendent of $\mathrm{Tr}(F^4)$ is parameterized by the most general polynomial function $A^{0,2}(t,u)$ that is symmetric in t and u. The Lagrangian term $A^{0,2}(t,u)$ generates the following S matrix

$$\begin{aligned} &A^{0,2}(t,u)\left(p^1_\mu\epsilon^1_\nu - p^1_\nu\epsilon^1_\mu\right)\left(p^3_\nu\epsilon^3_\alpha - p^3_\alpha\epsilon^3_\nu\right)\left(p^2_\alpha\epsilon^2_\beta - p^2_\beta\epsilon^2_\alpha\right)\left(p^4_\beta\epsilon^4_\mu - p^4_\mu\epsilon^4_\beta\right)\\ &+A^{0,2}(s,u)\left(p^1_\mu\epsilon^1_\nu - p^1_\nu\epsilon^1_\mu\right)\left(p^2_\nu\epsilon^2_\alpha - p^2_\alpha\epsilon^2_\nu\right)\left(p^3_\alpha\epsilon^3_\beta - p^3_\beta\epsilon^3_\alpha\right)\left(p^4_\beta\epsilon^4_\mu - p^4_\mu\epsilon^4_\beta\right)\\ &+A^{0,2}(t,s)\left(p^1_\mu\epsilon^1_\nu - p^1_\nu\epsilon^1_\mu\right)\left(p^3_\nu\epsilon^3_\alpha - p^3_\alpha\epsilon^3_\nu\right)\left(p^4_\alpha\epsilon^4_\beta - p^4_\beta\epsilon^4_\alpha\right)\left(p^2_\beta\epsilon^2_\mu - p^2_\mu\epsilon^2_\beta\right)\end{aligned} \tag{4.51}$$

Now let us consider the third structure which is

$$F_{ab}\mathrm{Tr}(\partial_a F\partial_b F F) \tag{4.52}$$

Note that

$$F_{ab}\mathrm{Tr}(\partial_a F\partial_b F F) = F_{ab}\mathrm{Tr}(F\partial_b F\partial_a F) \tag{4.53}$$

(where we have taken the transpose within the trace and used the antisymmetry of F_{ab}). It follows that the three Fs that enter in the trace in (4.52) all appear on equal footing.

The most general descendent of this structure can be written as linear combinations of

$$A^{2,1}_{m,n} = \prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}F_{ab}\mathrm{Tr}\left(\partial_{\mu_i}\partial_a F\partial_{\nu_j}\partial_b F F\right) \tag{4.54}$$

Clearly

$$A^{2,1}_{m,n} = A^{2,1}_{n.m} \tag{4.55}$$

(once again this follows by taking transpose within the trace in (4.47) and using the antisymmetry of F_{ab}). The coefficients $A^{2,1}_{m,n}$ also obey a slightly less obvious symmetry property[6]

$$A^{2,1}_{m,n} = (-1)^m\sum_{p=0}^{m}\frac{m!}{(m-p)!p!}A^{2,1}_{m-p,n+p} \tag{4.57}$$

[6] A nice way to see this is to consider the following manipulations

$$\begin{aligned} A^{2,1}_{m,n} &= \prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}F_{ab}\mathrm{Tr}\left(\partial_{\mu_i}\partial_a F\partial_{\nu_j}\partial_b F F\right)\\ &= -\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}F_{ab}\mathrm{Tr}\left(\partial_{\mu_i}F\partial_{\nu_j}\partial_b F\partial_a F\right)\\ &= \prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}F_{ab}\mathrm{Tr}\left(\partial_a F\partial_{\nu_j}\partial_b F\partial_{\mu_i}F\right)\\ &= \prod_{i=1}^{m}\prod_{j=1}^{n}(-1)^m\partial_{\mu_i}\partial_{\nu_j}F_{ab}\mathrm{Tr}\left(\partial_{\mu_i}\left(\partial_a F\partial_{\nu_j}\partial_b F\right)F\right)\\ &= \prod_{i=1}^{m}\prod_{j=1}^{n}(-1)^m\partial_{\mu_i}\partial_{\nu_j}F_{ab}\sum_{p=0}^{m}\frac{m!}{(m-p)!p!}\mathrm{Tr}\left(\partial_{\mu_{m-p}}\partial_a F\partial_{\nu_j}\partial_{\mu_{m+p}}\partial_b F F\right)\end{aligned} \tag{4.56}$$

In the second line we have used integration by parts. In the third line we have used the transpose property and cyclicity of trace. In the fourth line we again use integration by parts

The meaning of this symmetry is clearer when expressed in terms of the generating function

$$A^{2,1}(s,t) = \sum_{m,n=0}^{\infty} A^{2,1}_{m.n} s^m t^n \tag{4.58}$$

(4.57) is equivalent to the symmetry property

$$A^{2,1}(s,t) = A^{2,1}(u,t), \quad \text{where} \quad u \equiv -(s+t) \tag{4.59}$$

It also follows from (4.55) that

$$A^{2,1}(s,t) = A^{2,1}(t,s) \tag{4.60}$$

Combining (4.59) and (4.60) it follows that $A^{2,1}$ is a completely symmetric function of s, t, u, i.e.

$$A^{2,1}(s,t) = A^{2,1}(t,s) = A^{2,1}(u,t) = A^{2,1}(t,u) = A^{2,1}(s,u) = A^{2,1}(u,s), \quad \text{where} \quad u \equiv -(s+t) \tag{4.61}$$

The S matrix is given by

$$\begin{aligned} A^{2,1}(s,t) \big[& \left(p_a^1 \epsilon_b^1 - p_b^1 \epsilon_a^1\right) p_a^2 \left(p_\mu^2 \epsilon_\nu^2 - p_\nu^2 \epsilon_\mu^2\right) p_b^3 \left(p_\nu^3 \epsilon_\alpha^3 - p_\alpha^3 \epsilon_\nu^3\right) \left(p_\alpha^4 \epsilon_\mu^4 - p_\mu^4 \epsilon_\alpha^4\right) \\ & + \left(p_a^2 \epsilon_b^2 - p_b^2 \epsilon_a^2\right) p_a^1 \left(p_\mu^1 \epsilon_\nu^1 - p_\nu^1 \epsilon_\mu^1\right) p_b^4 \left(p_\nu^4 \epsilon_\alpha^4 - p_\alpha^4 \epsilon_\nu^4\right) \left(p_\alpha^3 \epsilon_\mu^3 - p_\mu^3 \epsilon_\alpha^3\right) \\ & + \left(p_a^3 \epsilon_b^3 - p_b^3 \epsilon_a^3\right) p_a^4 \left(p_\mu^4 \epsilon_\nu^4 - p_\nu^4 \epsilon_\mu^4\right) p_b^1 \left(p_\nu^1 \epsilon_\alpha^1 - p_\alpha^1 \epsilon_\nu^1\right) \left(p_\alpha^2 \epsilon_\mu^2 - p_\mu^2 \epsilon_\alpha^2\right) \\ & + \left(p_a^4 \epsilon_b^4 - p_b^4 \epsilon_a^4\right) p_a^3 \left(p_\mu^3 \epsilon_\nu^3 - p_\nu^3 \epsilon_\mu^3\right) p_b^2 \left(p_\nu^2 \epsilon_\alpha^2 - p_\alpha^2 \epsilon_\nu^2\right) \left(p_\alpha^1 \epsilon_\mu^1 - p_\mu^1 \epsilon_\alpha^1\right) \big] \end{aligned} \tag{4.62}$$

4.5.5 S matrix written in terms of unconstrained polarizations

Our explicit expressions for the S matrices are all presented in terms of polarization vectors that obey $\epsilon_i.k_i = 0$ and $\epsilon_i = \epsilon_i \sim k_i$. In other words the polarization tensors that enter our expressions are constrained rather than free data. For some purposes it is useful to have expressions for the S matrix expressed entirely in terms of unconstrained data. We turn to this problem in this subsection.

Throughout our discussion we will always choose all polarization vectors to be normalized, i.e. we chose $\epsilon_i.\epsilon_i^* = 1$. The polarization vector ϵ_i, for the i^{th} particle can be written as

$$\epsilon_i = \cos\zeta_i \epsilon_i^p + \sin\zeta_i \epsilon_i^o \tag{4.63}$$

where ϵ_i^p is a normalized polarization vector that lies within the three dimensional manifold spanned by the four momenta $p_1 \ldots p_4$ and ϵ_i^o is another normalized polarization vector that is completely orthogonal to this plane. Now the equation $p_i.\epsilon_i^p = 0$ forces ϵ_i^p to lie in a two dimensional subspace of the momentum plane. Recall also that ϵ_i^p occurs in equivalence classes; polarizations in the same equivalence class are related by $\epsilon_i^p \sim \epsilon_i^p + \alpha p_i$. All vectors in the same equivalence have the same norm and define the same S matrices and are thought of as physically identical. Modding out by this ambiguity, ϵ_i^p is uniquely determined (upto a phase that we choose arbitrarily and will play no essential role below.)

It follows that the 'in momentum plane' part of ϵ_i is is completely determined by the angle ζ_i. On the other hand the 'out of momentum plane' part of ϵ_i is fixed by ζ_i together with the choice of a unit vector in the $D-3$ dimensional transverse space.

The full S matrix - which is separately linear in each of the $\epsilon_i s$ - is consequently the sum of three distinct kinds of terms. Term 1 is proportional to $\sin\zeta_1 \sin\zeta_2 \sin\zeta_3 \sin\zeta_4$. Term 2 is proportional to $\sin\zeta_1 \sin\zeta_2 \cos\zeta_3 \cos\zeta_4 +$ permutations. Term 3 is proportional to $\cos\zeta_1 \cos\zeta_2 \cos\zeta_3 \cos\zeta_4$. In order that the whole S matrix grow no faster than s^2, the same must be true of each of Term 1, Term 2 and Term 3 separately [7] We now explicitly list the 3 different terms in the S matrix. In the rest of this section we will never again use the full polarization ϵ_i but will make extensive use of ϵ_i^o. In order to lighten the notation in the formulae below we will drop the superscript o above. Every occurrence of ϵ_i in the rest of this subsection actually denotes ϵ_i^0.

Term 1 of the S matrix is given by

$$\begin{aligned}
S_1 &= -8(A^{0,1}(t,u)(s^2\epsilon_1\cdot\epsilon_2\epsilon_3\cdot\epsilon_4) + A^{0,1}(s,u)(t^2\epsilon_1\cdot\epsilon_3\epsilon_2\cdot\epsilon_4) \\
&\quad + A^{0,1}(t,s)((s+t)^2\epsilon_1\cdot\epsilon_4\epsilon_2\cdot\epsilon_3)) \qquad (4.64)\\
S_2 &= -2(A^{0,2}(t,u)(t^2\epsilon_1\cdot\epsilon_4\epsilon_2\cdot\epsilon_3 + (s+t)^2\epsilon_1\cdot\epsilon_3\epsilon_2\cdot\epsilon_4) + A^{0,2}(s,u)(s^2\epsilon_1\cdot\epsilon_4\epsilon_2\cdot\epsilon_3 \\
&\quad + (s+t)^2\epsilon_1\cdot\epsilon_2\epsilon_3\cdot\epsilon_4) + A^{0,2}(t,s)(s^2\epsilon_1\cdot\epsilon_3\epsilon_2\cdot\epsilon_4 + t^2\epsilon_1\cdot\epsilon_2\epsilon_3\cdot\epsilon_4)) \\
S_3 &= 0 \\
S_1+S_2+S_3 &= -2\epsilon_1\cdot\epsilon_2\epsilon_3\cdot\epsilon_4\left(4s^2A^{0,1}(t,u) + u^2A^{0,2}(s,u) + t^2A^{0,2}(t,s)\right) \\
&\quad -2\epsilon_1\cdot\epsilon_3\epsilon_2\cdot\epsilon_4\left(4t^2A^{0,1}(s,u) + u^2A^{0,2}(t,u) + s^2A^{0,2}(t,s)\right) \\
&\quad -2\epsilon_1\cdot\epsilon_4\epsilon_3\cdot\epsilon_2\left(4u^2A^{0,1}(t,s) + s^2A^{0,2}(s,u) + t^2A^{0,2}(t,u)\right) \qquad (4.65)
\end{aligned}$$

Term 3 of the S matrix is given by

$$\begin{aligned}
S_1 &= s^2(-A^{0,1}(t,u)) - t^2A^{0,1}(s,u) - (s+t)^2A^{0,1}(t,s) \\
S_2 &= -\frac{1}{4}\left(\left(2s^2+2st+t^2\right)A^{0,2}(s,u) + \left(s^2+t^2\right)A^{0,2}(t,s)\right. \qquad (4.66)\\
&\quad \left. + \left(s^2+2st+2t^2\right)A^{0,2}(t,u)\right) \\
S_3 &= st(s+t)(A^{2,1}(t,s) + A^{2,1}(u,s) + A^{2,1}(t,u)) \\
S_1+S_2+S_3 &= -s^2A^{0,1}(t,u) - t^2A^{0,1}(s,u) - u^2A^{0,1}(t,s) \\
&\quad +\frac{1}{4}\left(-\left(s^2+u^2\right)A^{0,2}(s,u) - \left(s^2+t^2\right)A^{0,2}(t,s) - \left(u^2+t^2\right)A^{0,2}(t,u)\right) \\
&\quad -stu(A^{2,1}(t,s) + A^{2,1}(u,s) + A^{2,1}(t,u)) \qquad (4.67)
\end{aligned}$$

[7] For instance, if each of the $\zeta_i = \frac{\pi}{2}$ the full S matrix is given by Term 1. If each of the $\zeta_i = 0$ then the full S matrix equals Term 3. If two of the $\zeta_i = 0$ and the other two equals $\frac{\pi}{2}$ we get one of the six permutations of Term 2.

Finally, the part of term 2 proportional to $\cos\zeta_1 \sin\zeta_2 \cos\zeta_3 \sin\zeta_4$ is given by [8]

$$\begin{aligned} S_1 &= t^2 \epsilon_2^\perp \cdot \epsilon_4^\perp A^{0,1}(s,u) \\ S_2 &= \frac{1}{4}\epsilon_2^\perp \cdot \epsilon_4^\perp (s(u-t)A^{0,2}(t,s) + u(s-t)A^{0,2}(t,u)) \\ S_3 &= \epsilon_2^\perp \cdot \epsilon_4^\perp \frac{stu}{2}(A^{2,1}(t,s) + A^{2,1}(u,s) + A^{2,1}(t,u)) \\ S_1+S_2+S_3 &= \frac{1}{4}\epsilon_2^\perp \cdot \epsilon_4^\perp \left[4t^2 A^{0,1}(s,u) + u(s-t)A^{0,2}(t,u) + s(u-t)A^{0,2}(t,s)\right. \\ &\quad \left. +2stu\left\{A^{2,1}(t,s) + A^{2,1}(u,s) + A^{2,1}(t,u)\right\}\right] \end{aligned} \tag{4.68}$$

where $\epsilon_i^\perp$ refers to the transverse part of the polarization.

In 4 dimensions, at the level of descendants, the full S matrix= 0 has a solution. The equation (4.46)+(4.51) = 0 is equivalent to the following two constraints

$$\begin{aligned} &s^2 A^{0,2}(-s-t,t) + t^2 A^{0,2}(s,-s-t) + (s+t)^2 A^{0,2}(s,t) = 0 \\ &-\frac{1}{2}s^2 A^{0,1}(t,-s-t) - \frac{1}{8}s^2 A^{0,2}(s,-s-t) - \frac{1}{8}s^2 A^{0,2}(t,s) - \frac{1}{8}s^2 A^{0,2}(t,-s-t) = 0 \end{aligned} \tag{4.69}$$

has the following solution

$$\begin{aligned} A^{0,2}(s,t) &= st \\ A^{0,1}(s,t) &= \frac{1}{4}(s^2+t^2+st) \end{aligned} \tag{4.70}$$

The s-matrix of the structure which becomes fully symmetric is

$$\begin{aligned} S^{1,2,3,4} =& \frac{1}{2}\left(u\mathrm{Tr}(F^1F^2F^3F^4) + t\mathrm{Tr}(F^1F^3F^4F^2)\right) \\ &+ \frac{s}{8}\left(\mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4) + \mathrm{Tr}(F^1F^3)\mathrm{Tr}(F^4F^2) + \mathrm{Tr}(F^1F^4)\mathrm{Tr}(F^2F^3)\right) \end{aligned} \tag{4.71}$$

So, the counting for the case of 4-dimensions is $3+1+1=5$ degrees of freedom, as is expected from the CFT side counting in 3-dimensions. One of the orbits with 3 elements ($S^3/\mathbb{Z}_2$) becomes an orbit of 1 element, i.e. fully symmetric.

4.5.6 Counting Data in S matrices

We have seen above that the most general S matrix for 4 photons is specified by functions $A^{(0,1)}$, $A^{0,2}$) and $A^{2,1}$. These three functions are not quite on the same footing - in particular they don't carry equivalent amounts of data - because they have different symmetries. The function $A^{2,1}$ is completely symmetric in s t and u. The number of polynomials of degree m of this sort is given by $d_{sym}(m)$ where $d_{sym}(m)$ is the coefficient of x^m in (4.73). At large m, in particular, $d_{sym}(m) \sim \frac{m+1}{6}$.

[8] Of course Term 2 has 5 additional terms related to this one by permutations.

In contrast, the most general degree m function of s, t, u, with no symmetry constraints (this is what we would find if we were scattering, for instance, 4 distinguishable scalars) is simply the most general degree m polynomial built out of s and t and has $m+1$ terms. The partition function that counts these polynomials is

$$Z_{no-sym} = \frac{1}{(1-x)^2}$$

Somewhere in the middle of these two extremes, the most general degree m polynomial that preserves a Z_2 subgroup - say, of interchange between s and t - is simply the most general symmetric degree m polynomial of two variables, and so has

$$\left[\frac{m}{2}\right] + 1$$

independent coefficients. The partition function that counts these polynomials is given by

$$Z_{Z_2-sym}(x) = \frac{1+x}{(1-x^2)^2}$$

We can say all this more systematically in terms of the permutation group of three objects (which, in this case, are s, t and u). The irreducible representations of the group S_3 are labelled by Young Tableaux made out of 3 boxes. There are three such Young Tableaux - the completely symmetric one (with all three boxes in a single row), the completely antisymmetrical one (with all three boxes in a single column) and the mixed representation with two boxes in the first row and one in the second row (equivalently with two boxes in the first column and one in the second column). The completely symmetrical and completely antisymmetrical representations are one dimensional, while the mixed representation is two dimensional.

There is only one polynomial of degree zero; it transforms in the symmetric representation. At degree 1 we have two polynomials which can be taken to be $2s-t-u = 3s$ and $2t-s-u = 3t$. Note that these polynomials all vanish upon complete symmetrization. Moreover they are each annihilated by a Z_2 permutation symmetry. They are mapped to each other under permutation and consequently transform in the mixed representation. Note also that exactly one of these expressions is symmetric under the interchange of, say, s and t while the other one can be chosen to be antisymmetric under the same exchange ($s+t$) and $s-t$ are the two choices here).

At degree 2 we have three polynomials which can be taken to be $(s^2+t^2+u^2)$, and $(s^2+t^2-2u^2)$ and $t^2+u^2-2s^2)$. The first of these is in the symmetric representation. The next two transform in the mixed representation. A convenient basis for these two elements is $s^2+t^2-2u^2$ and s^2-t^2. Clearly the first element is symmetric under the interchange of s and t, while the second element is antisymmetric under the same interchange.

At degree three we have four polynomials. These can be taken to be $s^3+t^3+u^3$, $s^2t-t^2s-s^2u+u^2-u^2t-t^2u$, $s^3+t^3-2u^3$, and $s^3+u^3-2t^3$. The first of these expressions is completely symmetric. The second is completely antisymmetric. The remaining two belong

to the mixed representation. Once again we can choose a basis in this set of elements that is symmetric and antisymmetric under interchange of s and t (the symmetric expression is $s^3+t^3-2u^3$, whereas the antisymmetric expression is s^3-t^3)

The full set of polynomials of stu of degree m can be decomposed into these representations. The number of completely symmetrical representations is given by $d_{sym}(m)$ as described above. The number of completely antisymmetrical polynomials is easily computed along the lines of the discussion in (4.26). The analogue of Z_n computed there is the partition function of three identical particles in a fermionic harmonic oscillator and is given by

$$Z_n(x) = \frac{1}{3(1-x^3)} + \frac{1}{6(1-x)^3} - \frac{1}{2(1-x)(1-x^2)} \tag{4.72}$$

Correspondingly, the number of degree m completely antisymmetric polynomials of s, t and u is $d_{as}(m)$, the coefficient of x^m in

$$\begin{aligned} Z_{as}(x) = (1-x)Z_n(x) &= \frac{1}{3(1+x+x^2)} + \frac{1}{6(1-x)^2} - \frac{1}{2(1-x^2)} \\ &= \frac{x^3}{(1-x^2)(1-x^3)} \\ &= x^3(1+x^2+x^3+x^4+x^5+2x^6+x^7+2x^8+2x^9+2x^{10}\dots \end{aligned} \tag{4.73}$$

Finally, there are

$$\left[\frac{m}{2}\right]+1$$

representations of the mixed type (and so twice as many polynomials that transform in these representations because the representations are two dimensional). The partition function that counts the polynomials that transform in this representation is given by

$$\begin{aligned} Z_{mix-sym} =& 2\left(Z_{Z_2-sym} - Z_{sym}\right) \\ =& \frac{x}{(1-x)^2\left(x^2+x+1\right)} \end{aligned} \tag{4.74}$$

As a check, the sum of the partition functions of the polynomials that transform in various representations is given by

$$2Z_{Z_2-sym} - Z_{sym} + Z_{as} = \frac{1}{(1-x)^2} = Z_{no-sym} \tag{4.75}$$

4.5.7 Growth of S matrices with s at fixed t

In this subsection we will classify all local 4 photon S matrices that grow no faster than s^2 in the Regge limit for every choice of polarization vectors.

Let us first study term 1 of the S matrix. Provided that $D \geq 4$ the condition that the S matrix grow no faster than s^2 is only met provided the same condition holds independently

for the coefficients of $\epsilon_1.\epsilon_2\epsilon_3.\epsilon_4$ and the two other similar terms. Our condition is thus simply that

$$\left(4u^2A^{0,1}(t,s)+s^2A^{0,2}(s,u)+t^2A^{0,2}(t,u)\right) \tag{4.76}$$

(together with the two crossing related expressions) grow no faster than s^2 in the Regge limit.

If we assume that $A^{0,1}$ and $A^{0,2}$ are polynomials then the expression in (4.76) is a polynomial of degree 2 or greater. Moreover it is symmetric under interchange of s and t. The only polynomials that meet these conditions and still do not grow faster than s^2 are s^2+t^2, st, s^2t+t^2s and s^2t^2. Now if the polynomial in (4.76) were to evaluate to s^2t^2 then the (permutation related) polynomial that occurs in the bracket of the second-last line of (4.64) would evaluate to s^2u^2. As this expression grows faster than s^2 and so is disallowed. We conclude that the expression in (4.76) must be a linear combination of s^2+t^2, st and s^2t+t^2s.

Let us now turn to the result of term 3 above. Notice that the part of this answer that depends on $A^{(0,1)}$ and $A^{(0,2)}$ is proportional to the term in (4.76) completely symmetrized (i.e. is proportional to the sum of the three brackets in the last three lines of (4.64)). Given the conditions of the last paragraph, this term automatically grows no faster than s^2 in the Regge limit. It follows that Term 3 grows no faster than s^2 in the Regge limit provided the same is s true of

$$-stu(A^{2,1}(t,s)+A^{2,1}(u,s)+A^{2,1}(t,u)).$$

This condition immediately forces $A^{2,1}(s,t)$ to be a constant.

Finally, let us turn to the expression in term 2 above. The coefficient of $\epsilon_2.\epsilon_3$ in that expression is given by

$$\left[4u^2A^{0,1}(s,t)+t(s-u)A^{0,2}(t,u)+s(t-u)A^{0,2}(s,u)+2stu\left\{A^{2,1}(t,s)+A^{2,1}(u,s)+A^{2,1}(u,s)\right\}\right] \tag{4.77}$$

As we now know that $A^{(2,1)}$ is a constant, the term proportional to $A^{(2,1)}$ in this expression is proportional to stu and so automatically grows no faster than s^2 in the Regge limit. In order that our S matrix grow no faster than s^2 at fixed t, it must be that the same is true of the expression

$$4u^2A^{0,1}(s,t)+t(s-u)A^{0,2}(t,u)+s(t-u)A^{0,2}(t,s) \tag{4.78}$$

By repeating the reasoning in the paragraph under (4.76) it must be that (4.78), like (4.76), is a linear combination of the polynomials s^2+t^2, st, and stu.

In summary we require that the expressions in (4.76) and (4.78) must simultaneously be (possibly different) linear combinations of the three polynomials listed above. If $A^{0,1}$ and $A^{0,2}$ are constants, then both (4.76) and (4.78) are automatically linear combinations of s^2+t^2 and st. The only other possibility for $A^{0,1}$ and $A^{0,2}$ is that they are proportional to the (unique symmetric degree one) polynomial $s+t$. If we suppose that $A^{0,1}(s,t)=a(s+t)$ and that $A^{0,2}(s,t)=b(s+t)$ then (4.76) evaluates to

$$-4au^3-b\left(s^2t+t^2s\right)$$

The condition that this expression grow no faster than s^2 in the Regge limit sets $a = 0$. With this condition (4.78) evaluates to

$$-st(s-u) - st(t-u) = -3stu$$

So b is allowed to be nonzero.

4.5.8 String tree level 4 photon S matrices

Superstring

In this section we express the type 1 superstring amplitude for photons using our basis. The amplitude is given by [137]

$$\begin{aligned} A_{4V}^{ss} &= f(s,t,u)\left(\frac{1}{2}\epsilon_2.\epsilon_3\left(s\epsilon_1.k_3\epsilon_4.k_2 + t\epsilon_1.k_2\epsilon_4.k_3\right) + \frac{1}{2}\epsilon_1.\epsilon_4\left(s\epsilon_2.k_4\epsilon_3.k_1 + t\epsilon_2.k_1\epsilon_3.k_4\right)\right. \\ &+\frac{1}{2}\epsilon_2.\epsilon_4\left(s\epsilon_1.k_4\epsilon_3.k_2 + u\epsilon_1.k_2\epsilon_3.k_4\right) + \frac{1}{2}\epsilon_1.\epsilon_3\left(s\epsilon_2.k_3\epsilon_4.k_1 + u\epsilon_2.k_1\epsilon_4.k_3\right) \\ &+\frac{1}{2}\epsilon_3.\epsilon_4\left(t\epsilon_1.k_4\epsilon_2.k_3 + u\epsilon_1.k_3\epsilon_2.k_4\right) + \frac{1}{2}\epsilon_1.\epsilon_2\left(t\epsilon_3.k_2\epsilon_4.k_1 + u\epsilon_3.k_1\epsilon_4.k_2\right) \\ &\left.-\frac{1}{4}st\epsilon_1.\epsilon_4\epsilon_2.\epsilon_3 - \frac{1}{4}su\epsilon_1.\epsilon_3\epsilon_2.\epsilon_4 - \frac{1}{4}tu\epsilon_1.\epsilon_2\epsilon_3.\epsilon_4\right) \end{aligned} \tag{4.79}$$

where $f(s,t,u)$ is a function of the Mandelstam variables with no polarization dependence. This is reproduced by the s-matrix coming from the following Lagrangian term

$$L_{4V}^{ss} \propto \frac{1}{16}\left(\mathrm{Tr}(F^4) - \frac{1}{4}(\mathrm{Tr}(F^2))^2\right) \tag{4.80}$$

Bosonic String

The open bosonic string amplitude is given by [138],

$$\begin{aligned} A_{4V}^{\mathrm{Open}} &= \left(-A_{4V}^{ss} + \frac{st\left(\epsilon_2.\epsilon_3 - \epsilon_2.k_3\epsilon_3.k_2\right)\left(\epsilon_1.\epsilon_4 - \epsilon_1.k_4\epsilon_4.k_1\right)}{4\left(\frac{u}{2}+1\right)} + \frac{su\left(\epsilon_1.\epsilon_3 - \epsilon_1.k_3\epsilon_3.k_1\right)\left(\epsilon_2.\epsilon_4 - \epsilon_2.k_4\epsilon_4.k_2\right)}{4\left(\frac{t}{2}+1\right)}\right. \\ &+\frac{tu\left(\epsilon_1.\epsilon_2 - \epsilon_1.k_2\epsilon_2.k_1\right)\left(\epsilon_3.\epsilon_4 - \epsilon_3.k_4\epsilon_4.k_3\right)}{4\left(\frac{s}{2}+1\right)} + \frac{1}{2}s\left(\frac{1}{3}\left(\epsilon_1.k_2\epsilon_2.k_3\epsilon_3.k_1 - \epsilon_1.k_3\epsilon_2.k_1\epsilon_3.k_2\right)\right. \\ &\left(\epsilon_4.k_1 - \epsilon_4.k_2\right) + \epsilon_1.k_3\epsilon_2.k_3\left(\epsilon_3.k_1\epsilon_4.k_1 + \epsilon_3.k_2\epsilon_4.k_2\right)\Big) + \frac{1}{2}t\left(\frac{1}{3}\left(\epsilon_1.k_3\epsilon_2.k_1\epsilon_3.k_2 - \epsilon_1.k_2\epsilon_2.k_3\epsilon_3.k_1\right)\right. \\ &\left(\epsilon_4.k_1 - \epsilon_4.k_3\right) + \epsilon_1.k_2\epsilon_3.k_2\left(\epsilon_2.k_1\epsilon_4.k_1 + \epsilon_2.k_3\epsilon_4.k_3\right)\Big) + \frac{1}{2}u\left(\frac{1}{3}\left(\epsilon_1.k_3\epsilon_2.k_1\epsilon_3.k_2 - \epsilon_1.k_2\epsilon_2.k_3\epsilon_3.k_1\right)\right. \\ &\left(\epsilon_4.k_3 - \epsilon_4.k_2\right) + \epsilon_2.k_1\epsilon_3.k_1\left(\epsilon_1.k_2\epsilon_4.k_2 + \epsilon_1.k_3\epsilon_4.k_3\right)\Big) - \frac{1}{4}st\epsilon_1.\epsilon_4\epsilon_2.\epsilon_3 - \frac{1}{4}su\epsilon_1.\epsilon_3\epsilon_2.\epsilon_4 \\ &\left.-\frac{1}{4}tu\epsilon_1.\epsilon_2\epsilon_3.\epsilon_4\right)\left(\frac{s}{2}+1\right)\left(\frac{t}{2}+1\right)\left(\frac{u}{2}+1\right)g(s,t,u) \end{aligned} \tag{4.81}$$

This is reproduced by the following Lagrangian term

$$\begin{aligned}
L_{4V}^{\rm Open} \propto\ & \left(\frac{s}{2}+1\right)\left(\frac{t}{2}+1\right)\left(\frac{u}{2}+1\right)\frac{1}{16}\left(-{\rm Tr}(F^4)+\frac{1}{4}({\rm Tr}(F^2))^2\right)+\frac{1}{32}{\rm Tr}\left(\partial_{\mu_1}\partial_{\nu_1}FF\right){\rm Tr}\left(\partial_{\mu_1}F\partial_{\nu_1}F\right) \\
& \frac{1}{32}{\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}\partial_{\nu_1}\partial_{\nu_2}FF\right){\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}F\partial_{\nu_1}\partial_{\nu_2}F\right)+\frac{1}{16}{\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}\partial_{\nu_1}FF\right){\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}F\partial_{\nu_1}F\right) \\
& -\frac{1}{4}{\rm Tr}\left(\partial_{\mu_1}F\partial_{\mu_1}FFF\right)+\frac{1}{2}{\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}\partial_{\nu_1}F\partial_{\mu_1}\partial_{\mu_2}FF\partial_{\nu_1}F\right)+\frac{1}{4}{\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}\partial_{\mu_3}F\partial_{\mu_1}\partial_{\mu_2}\partial_{\mu_3}FFF\right) \\
& +\frac{1}{4}{\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}\partial_{\nu_1}\partial_{\nu_2}F\partial_{\mu_1}\partial_{\mu_2}FF\partial_{\nu_1}\partial_{\nu_2}F\right)+\frac{1}{4}{\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}\partial_{\mu_3}\partial_{\nu_1}F\partial_{\mu_1}\partial_{\mu_2}\partial_{\mu_3}FF\partial_{\nu_1}F\right) \\
& -\frac{1}{12}F_{ab}{\rm Tr}\left(\partial_a F\partial_b FF\right)+\frac{1}{6}\partial_{\mu_1}\partial_{\mu_2}\partial_{\nu_1}F_{ab}{\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}\partial_a F\partial_b F\partial_{\nu_1}F\right) \\
& +\frac{1}{8}\partial_{\mu_1}\partial_{\mu_2}F_{ab}{\rm Tr}\left(\partial_{\mu_1}\partial_{\mu_2}\partial_a F\partial_b FF\right)
\end{aligned} \tag{4.82}$$

4.6 Lagrangian terms for 4 point functions

In this section we study diffeomorphically invariant Lagrangians for pure gravity theories. The only field in our theory is the graviton. We will attempt to parameterize all Lagrangians that lead to inequivalent gravitational 3 and 4 point S matrices. In this section we work in an expansion about flat space - we will generalize our results to gravitational theories in asymptotically AdS spaces in later sections.

Let us start with some generalities. Firstly, any diffeomorphically invariant Lagrangian can be built out of products of covariant derivatives of the Riemann tensor. [9] As the Riemann tensor vanishes in flat space, terms in the gravitational action that contribute to three and four point scattering are necessarily built out of products of at most four copies of the Riemann tensor; of course these four copies can each appear with arbitrary numbers of derivatives.

4.6.1 Triviality of terms proportional to $R_{\mu\nu}$

Consider a Lagrangian that takes the form

$$\int \sqrt{g}\left(R+\epsilon R_{\mu\nu}A^{\mu\nu}\right) \tag{4.83}$$

(where $A^{\mu\nu}$ is any two index tensor built out of the Riemann tensor and its derivatives). We will now explain that at first order in ϵ, the S matrix the classical S matrix that follows from this Lagrangian is identical to that for the pure Einstein theory. In order to see this we first note that under the field redefinition

$$g_{\mu\nu}=\widetilde{g}_{\mu\nu}+\epsilon\left(R_{\mu\nu}+g_{\mu\nu}R\right) \tag{4.84}$$

[9] In this chapter we restrict our attention to gravitational theories that have a Lagrangian description, and ignore the possibility of gravitational Chern Simons terms.

the Lagrangian (4.85) turns into

$$\int \sqrt{\widetilde{g}} \left(\widetilde{R} + n\mathcal{O}(\epsilon^2) \right) \tag{4.85}$$

where $\widetilde{R}$ is the Ricci scalar built out of the metric $\widetilde{g}_{\mu\nu}$. Returning to the original variables for a moment, let us substitute the expansion

$$g_{\mu\nu} = \eta_{\mu\nu} + \sum_{i=1}^{n} h^i_{\mu\nu}(k_i) \tag{4.86}$$

into the Lagrangian (4.85) (here $h^i_{\mu\nu}$ are the linearized solutions of Einstein's equations corresponding to gravitons of particular polarization and momenta. Once we make this substitution, we can read off the n graviton S matrix for the Lagrangian (4.85) by evaluating the coefficient of the term linear separately in each field h^i. Schematically

$$S = \sum_{n=1}^{\infty} S_n h^1 h^2 \dots h^n$$

where S is the action and S_n is the n particle S. Working to linearized order in ϵ we have

$$S_n = S_n^E + \epsilon S_n^{new} \tag{4.87}$$

where S_n is the n particle S matrix that follows from the Einstein Lagrangian and S_n^{new} is the correction to this S matrix at first order in ϵ due to the correction term in (4.85).

After making the field redefinition, however, it also follows that

$$S = \sum_{n=1}^{\infty} \widetilde{S}_n \widetilde{h}^1 \widetilde{h}^2 \dots \widetilde{h}^n \tag{4.88}$$

where

$$\widetilde{S}_n = S_n^E$$

(this follows from the fact that the Lagrangian in the new variables is the Einstein action: all formulas are correct only upto order ϵ)

The relationship between $h^i_{\mu\nu}$ and $\widetilde{h}^i_{\mu\nu}$ may be obtained by linearizing (4.84). The really important point is that, at linear order, this relationship is trivial

$$h^i_{\mu\nu} = \widetilde{h}^i_{\mu\nu} \tag{4.89}$$

This follows because it follows by index matching that $A_{\mu\nu}$ is a linear combination of $R_{\mu\nu}$, $g_{\mu\nu}R$, and terms that are of second or higher order in the Riemann tensor. Linearizing and using the fact that both R and $R_{\mu\nu}$ vanish onshell, (4.89) follows.

Comparing (4.87), (4.88) and (4.89) we immediately conclude that S_n^{new} all vanish, as we wanted to show.

4.6.2 Identities of the Riemann Tensor

In this section we list all the identities and symmetry properties of the Riemann tensor that we will be using in the subsequent analysis to determine the diffeomorphism invariant Lagrangian structures.

1. The symmetry properties of the Riemann tensor,

$$R_{abcd} = -R_{bacd} = -R_{abdc}, \qquad R_{abcd} = R_{cdab} \tag{4.90}$$

2. The algebraic Bianchi identity

$$R_{abcd} + R_{acdb} + R_{adbc} = 0 \tag{4.91}$$

3. The differential Bianchi identity

$$\nabla_a R_{bcde} + \nabla_b R_{cade} + \nabla_c R_{abde} = 0 \tag{4.92}$$

4. The contracted Bianchi identity, which we get from the differential Bianchi identity by contracting the Riemann tensor appropriately.

$$\nabla_a R_{ce} + \nabla^b R_{cabe} - \nabla_c R_{ae} = 0 \tag{4.93}$$

5. Commutator of derivatives

$$[\nabla_f, \nabla_e] R_{abcd} = R_{pbcd} R^p_{afe} + R_{apcd} R^p_{bfe} + R_{abpd} R^p_{cfe} + R_{abcp} R^p_{dfe} \tag{4.94}$$

We will be using these identities and symmetry properties of the Riemann tensor along with integration by parts judiciously to fix the independent Lagrangian terms.

4.6.3 Classification of terms that contribute to 3 graviton scattering

We will now classify all terms in the Lagrangian that can contribute to 3 point gravitational scattering. In this section we work in asymptotically high dimensions and so do not allow ourselves to use dimension dependent structures like ϵ tensors or special identities for the Riemann tensor that work only in special dimensions.

All terms that contribute to three point functions are built out of at most 3 copies of the Riemann tensor. The simplest such term in the Einstein Lagrangian itself; it is a two derivative term built out of a single copy of the Riemann tensor and results in the following onshell 3 point gravitational S matrix

(4.95)

This is the unique term in the Lagrangian that we can build out of a single copy of the Riemann tensor.

Let us now turn to terms built out of two copies of the Riemann tensor. The most general term of this form is given by

$$S = \int \sqrt{g} R_{abcd} R^{abcd} \tag{4.96}$$

every other term two Riemann term is either equivalent to the term listed in (4.96) or to a term built out of three copies of the Riemann tensor. ..... To be precise, let us consider the following possible Lagrangian term

$$S = \int \sqrt{g} R_{abcd} R^{acbd} \tag{4.97}$$

Using (4.6.2), we can systematically relate it to (4.96) as follows

$$\begin{aligned} R_{abcd}R^{acbd} &= R_{abcd}\left(-R^{abdc} - R^{adcb}\right) \\ &= R_{abcd}R^{abcd} - R_{abcd}R^{acbd} \\ \therefore R_{abcd}R^{acbd} &= \frac{1}{2} R_{abcd}R^{abcd} \end{aligned} \tag{4.98}$$

where in the first line we have used the algebraic Bianchi identity and in the second line we have used the symmetry properties of the Riemann tensor. We will be mostly concerning ourselves, for now, with Lagrangian terms upto order eight in derivatives. Let us now examine two derivatives acting on two Riemann tensors, which we show to be equivalent to either three Riemann terms or proportional to Ricci tensor, which according to arguments in section 4.6.1, can be discarded due to field redefinition.

$$\begin{aligned} S &= \int \sqrt{g} \nabla_\mu R_{abcd} \nabla^\mu R^{acbd} \\ &= \int \sqrt{g} \left(-\nabla_a R_{b\mu cd} - \nabla_b R_{\mu acd}\right) \nabla^\mu R^{acbd} \\ &= \int \sqrt{g} \nabla^\mu \left(-\nabla_a R_{b\mu cd} - \nabla_b R_{\mu acd}\right) R^{acbd} + C_{\partial\mathcal{M}} \\ &= \int \sqrt{g} \left(-\nabla_a \nabla^\mu R_{b\mu cd} - \nabla_b \nabla^\mu R_{\mu acd}\right) R^{acbd} + C_{\partial\mathcal{M}} + \widetilde{C}_{R^3} \\ &= \widehat{C}_{R_{\mu\nu}} + C_{\partial\mathcal{M}} + \widetilde{C}_{R^3} \end{aligned} \tag{4.99}$$

where in the second line we have used the differential Bianchi identity. In the third line we have used integration by parts to get the total derivative term $C_{\partial\mathcal{M}}$. In the fourth line we have used commutator of derivatives to get the $\widetilde{C}_{R^3}$ term which is of higher order in Riemann tensor and hence will be dealt with when we classify terms with three Riemann tensors. Finally we use the contracted Bianchi identity to relate this to the Ricci tensor term $\widehat{C}_{R_{\mu\nu}}$. Hence for all practical purposes such terms will be discarded. Detailed analysis of other possible two derivative contractions and higher order derivatives on two Riemann tensors are considered in Appendix C.3.

Finally let us consider terms built out of three copies of the Riemann tensor. Let us first study terms with no derivatives. It is easy to convince oneself that there are two unique terms of this form ([139],[140]).

$$R^{pqrs}R_{pq}{}^{tu}R_{rstu}, \qquad R^{pqrs}R_{p}{}^{t}{}_{r}{}^{u}R_{qtsu} \tag{4.100}$$

All possible non-zero index contractions of three Riemann tensors to yield scalars can be obtained from this two by repeated application of the symmetry properties and the algebraic Bianchi identities. Let us consider an explicit example: $R_{pq}{}^{rs}R_{rtsu}R^{tu}{}_{pq}$

$$\begin{aligned} R_{pq}{}^{rs}R_{rtsu}R^{tu}{}_{pq} &= R_{pq}{}^{rs}(-R_{rsut}-R_{ruts})R^{tu}{}_{pq} \\ &= R_{pq}{}^{rs}R_{rstu}R^{tu}{}_{pq} - R_{pq}{}^{rs}R_{rtsu}R^{tu}{}_{pq} \\ \therefore R_{pq}{}^{rs}R_{rtsu}R^{tu}{}_{pq} &= \frac{1}{2}R_{pq}{}^{rs}R_{rstu}R^{tu}{}_{pq} \end{aligned} \tag{4.101}$$

Where in the second line we have used the algebraic Bianchi identity and subsequently used the symmetry properties of the Riemann tensor. Similar manipulations can be done for any non-zero contraction of three Riemann tensors to reduce it to one of the two forms or a linear combination of both. It is to be noted that the contribution of the both the three Riemann terms to the linearized three point function is the same so we can consider a linear combination of the two which contribute to the three point function and another one which does not. To be precise, let us consider the linearized Riemann tensor ($g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$),

$$\begin{aligned} R_{abcd} &= \frac{1}{2}F_{ab}F_{cd} \\ F_{ab} &= (k_a e_b - k_b e_a), \qquad h_{ab} = e_a e_b \end{aligned} \tag{4.102}$$

This satisfies the necessary symmetry properties of the Riemann tensor while being gauge invariant. This satisfies both the algebraic Bianchi identity and the differential Bianchi identity. Finally let us comment on the fact that only one of the two three Riemann structures contribute to three point functions. Explicit computations show that the linear combination

$$R^{pqrs}R_{pq}{}^{tu}R_{rstu} + 2R^{pqrs}R_{p}{}^{t}{}_{r}{}^{u}R_{qtsu} \tag{4.103}$$

contributes to the three point function but the combination

$$R^{pqrs}R_{pq}{}^{tu}R_{rstu} - 2R^{pqrs}R_{p}{}^{t}{}_{r}{}^{u}R_{qtsu} \tag{4.104}$$

does not. We can consider the non-linear completion of this term such that the three point S matrix coming from this term is proportional to a total derivative even without imposing equations of motion. The reason why we consider a non linear completion is because only after this completion the four point function coming from the new term is gauge invariant.

$$\begin{aligned} \chi_6 &= \frac{1}{8}\epsilon_{abcdef}\epsilon^{ghijkl}R_{ab}{}^{gh}R_{cd}{}^{ij}R_{ef}{}^{kl} \\ &= 4R_{ab}{}^{cd}R_{cd}{}^{ef}R_{ef}{}^{ab} - 8R_{a}{}^{c}{}_{b}{}^{d}R_{c}{}^{e}{}_{d}{}^{f}R_{e}{}^{a}{}_{f}{}^{b} - 24R_{abcd}R^{abc}{}_{e}R^{de} + 3R_{abcd}R^{abcd}R \\ &\quad + 24R_{abcd}R^{ac}R^{bd} + 16R_{a}{}^{b}R_{b}{}^{c}R_{c}{}^{a} - 12R_{a}{}^{b}R_{b}{}^{a}R + R^3 \end{aligned} \tag{4.105}$$

This term is in-fact the six dimensional Euler density.

Now let us turn to the terms with derivatives on three copies of the Riemann tensor. The basis for structures with two derivatives on three Riemann tensors have been discussed in [139].

$$\begin{aligned} &R^{pqrs}R_{p}{}^{tuv}\nabla_v\nabla_s R_{qtru}, \qquad R^{pqrs}\nabla_q R^{tuv}{}_{p}\nabla_s R_{tuvr} \\ &R^{pqrs}\nabla_r R^{tuv}{}_{p}\nabla_s R_{tuvq}, \qquad R^{pqrs}\nabla^v R^{t\ u}_{p\ r}\nabla_v R_{tqus} \end{aligned} \tag{4.106}$$

We claim that upto total derivatives, there are no structures with two derivatives acting on three Riemann tensors, which contribute to the three point function. All of them are four Riemann structures or polynomials involving Ricci tensors.

We now consider four derivatives acting on three Riemann tensors.

4.6.4 Classification of terms that contribute to 4 graviton scattering

In this section we list all the terms that can contribute to the analytic part of the four point scattering amplitude. We begin by listing the terms without any derivatives acting on them.

$$\begin{aligned} &(R_{pqrs}R^{pqrs})^2 \\ &R^{pqrs}R_{pqr}{}^{t}R^{uvw}{}_{s}R_{uvwt} \\ &R^{pqrs}R_{pq}{}^{tu}R_{tu}{}^{vw}R_{rsvw} \\ &R^{pqrs}R_{pq}{}^{tu}R_{rt}{}^{vw}R_{suvw} \\ &R^{pqrs}R_{p\ r}^{t\ u}R_{tvws}R_{q\ u}^{v\ w} \\ &R^{pqrs}R_{p\ r}^{t\ u}R_{t\ u}^{v\ w}R_{qvsw} \\ &R^{pqrs}R_{p\ r}^{t\ u}R_{t\ q}^{v\ w}R_{uvsw} \end{aligned} \tag{4.107}$$

From these structures explicit computation shows that at the linearized level out of these seven structures only six are independent. Note that the linearized Riemann tensor can be expressed as product of two electromagnetic field strengths, then its immediately obvious that these structures are just inequivalent contractions of the two electromagnetic Lagrangian structures, namely the first two structures in (4.41). As an explicit example we show the equivalence for the structure $(R_{pqrs}R^{pqrs})^2 \sim \left(\mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4)\right)^2$. In addition to these six independent structures above, we can also write three more structures which are given by the contraction of the first two structures with the third one in (4.41).

$$\begin{aligned} S^2S^3: &\ \mathrm{Tr}(F^1F^2F^3F^4)F^1_{ab}\mathrm{Tr}(p_a^2F^2p_b^3F^3F^4) \\ S^1S^3: &\ \mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4)F^1_{ab}\mathrm{Tr}(p_a^2F^2p_b^3F^3F^4) \\ S^3S^3: &\ F^1_{pq}\mathrm{Tr}(p_p^2F^2p_q^3F^3F^4)F^1_{ab}\mathrm{Tr}(p_a^2F^2p_b^3F^3F^4) \end{aligned} \tag{4.108}$$

4.6.5 Explicit Listing of Kinematically Allowed Gravitational S matrices

In this subsection we list all the possible tree level graviton $2-2$ S matrices in spacetime dimension greater than or equal to 7. First structure has a natural interpretation, and in the rest of 9 structures Riemann tensor can be thought of as product two $F_{\mu\nu}$ tensors, both at same momenta.[10]

- The combination which vanish at 3 point level

$$T_1 : R^{pqrs}R_{pq}{}^{tu}R_{rstu} + 2R^{pqrs}R_{p\ r}^{\ t\ u}R_{qtsu} \tag{4.109}$$

has a completion, which makes it a topological term in 6 dimensions, so its non-trivial only form 7 dimensions onward. The S matrix coming from this has the following form

$$3B^{0,0}(s,t,u)\epsilon^{ijklmnp}\epsilon^{asdfghj}\epsilon^1_i\epsilon^2_j\epsilon^3_k\epsilon^4_l p^1_m p^2_n p^3_p \epsilon^1_a\epsilon^2_s\epsilon^3_d\epsilon^4_f p^1_g p^2_h p^3_j \tag{4.110}$$

ϵ^{ijklmp} is Levi-Civita tensor of rank 7. $B^{0,0}(s,t,u)$ is a fully symmetric function of s,t,u. Note that since the expression involves the product of Levi-Civita tensors, it is gauge invariant in dimensions greater than 7 also.

- T_2 : $\mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4)\mathrm{Tr}(F^1F^2F^3F^4)$, where superscript labels particles. This structure has only $(\mathbb{Z}_2\times\mathbb{Z}_2)$ symmetry. The corresponding orbit from the CFT side contains 6 elements. The general descendant is of the following form

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}\left(F^1_{ab}F^1_{pq}\right)\partial_{\mu_i}\left(F^2_{ba}F^2_{qr}\right)\partial_{\nu_i}\left(F^3_{cd}F^3_{rs}\right)F^4_{dc}F^4_{sp} \tag{4.111}$$

The S matrix has the following form

$$\begin{aligned}
&B^{0,1}(s,t)\left[\left(p^1_p\epsilon^1_q - p^1_q\epsilon^1_p\right)\left(p^2_p\epsilon^2_q - p^2_q\epsilon^2_p\right)\left(p^3_r\epsilon^3_s - p^3_s\epsilon^3_r\right)\left(p^4_r\epsilon^4_s - p^4_s\epsilon^4_r\right)\right.\\
&\left.\left(p^1_a\epsilon^1_b - p^1_b\epsilon^1_a\right)\left(p^2_b\epsilon^2_c - p^2_c\epsilon^2_b\right)\left(p^3_c\epsilon^3_d - p^3_d\epsilon^3_c\right)\left(p^4_d\epsilon^4_a - p^4_a\epsilon^4_d\right)\right]\\
&+ B^{0,1}(s,u)\left[3\leftrightarrow 4\right] + B^{0,1}(t,s)\left[2\leftrightarrow 3\right] + B^{0,1}(t,u)\left[2\leftrightarrow 3 \ \text{ then } \ 2\leftrightarrow 4\right]\\
&+ B^{0,1}(u,t)\left[2\leftrightarrow 4\right] + B^{0,1}(u,s)\left[2\leftrightarrow 4 \ \text{ then } \ 2\leftrightarrow 3\right]
\end{aligned} \tag{4.112}$$

The function $B^{0,1}$ has no special symmetry property. The notation has been shortened to avoid clutter. $B^{0,1}$ dependence on s,t,u has been written showing S^3 permutations. $[A\leftrightarrow B...]$ means to flip A and B particle labels from the term shown in the first two lines. Notice that there are 6 terms.

[10] It is best thought of in the momentum space.

- $T_3 : \mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4)\mathrm{Tr}(F^1F^3F^2F^4)$. This is a very similar structure to the previous one, but now it has extra $\mathbb{Z}_2$ symmetry, which can be clearly seen as label 3 and 4 swap. Therefore, the orbit of S^3 is cut by half. General descendant

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\nu_j}\partial_{\mu_i}\left(F^1_{ab}F^1_{pq}\right)F^2_{ba}F^2_{rs}\partial_{\mu_i}\left(F^3_{cd}F^3_{qr}\right)\partial_{\nu_i}\left(F^4_{dc}F^4_{sp}\right) \tag{4.113}$$

$m \leftrightarrow n$ is manifest in the above expression. S matrix, where $B^{0,2}(s,t)$ is symmetric in s,t

$$\begin{aligned}&B^{0,2}(t,u)\left[\left(p^1_p\epsilon^1_q-p^1_q\epsilon^1_p\right)\left(p^2_p\epsilon^2_q-p^2_q\epsilon^2_p\right)\left(p^3_r\epsilon^3_s-p^3_s\epsilon^3_r\right)\left(p^4_r\epsilon^4_s-p^4_s\epsilon^4_r\right)\right.\\&\left.\left(p^1_a\epsilon^1_b-p^1_b\epsilon^1_a\right)\left(p^3_b\epsilon^3_c-p^3_c\epsilon^3_b\right)\left(p^2_c\epsilon^2_d-p^2_d\epsilon^2_c\right)\left(p^4_d\epsilon^4_a-p^4_a\epsilon^4_d\right)\right]\\&+B^{0,2}(s,u)\left[3\leftrightarrow 2\right]+B^{0,2}(s,t)\left[2\leftrightarrow 4\right]\end{aligned} \tag{4.114}$$

With the following symmetry property.

$$B^{0,2}(t,u)=B^{0,2}(u,t) \tag{4.115}$$

- $T_4 : \mathrm{Tr}(F^1F^2F^3F^4)\mathrm{Tr}(F^1F^2F^3F^4)$. This structure, like $\mathrm{Tr}(F^4)$ again has $\mathbb{Z}_2$ symmetry of $1 \leftrightarrow 3$. Most general descendant

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\nu_j}\partial_{\mu_i}(F^1_{ab}F^1_{pq})\partial_{\mu_i}(F^2_{bc}F^2_{qr})F^3_{cd}F^3_{rs}\partial_{\nu_j}(F^4_{da}F^4_{sp}) \tag{4.116}$$

$m \leftrightarrow n$ is the symmetry. The S matrix is

$$\begin{aligned}&B^{0,3}(s,u)\left[\left(p^1_a\epsilon^1_b-p^1_b\epsilon^1_a\right)\left(p^2_b\epsilon^2_c-p^2_c\epsilon^2_b\right)\left(p^3_c\epsilon^3_d-p^3_d\epsilon^3_c\right)\left(p^4_d\epsilon^4_a-p^4_a\epsilon^4_d\right)\right.\\&\left.\left(p^1_p\epsilon^1_q-p^1_q\epsilon^1_p\right)\left(p^2_q\epsilon^2_r-p^2_r\epsilon^2_q\right)\left(p^3_r\epsilon^3_s-p^3_s\epsilon^3_r\right)\left(p^4_s\epsilon^4_p-p^4_p\epsilon^4_s\right)\right]\\&+B^{0,3}(t,u)\left[3\leftrightarrow 2\right]+B^{0,3}(s,t)\left[3\leftrightarrow 4\right]\end{aligned} \tag{4.117}$$

With the following symmetry property.

$$B^{0,3}(s,u)=B^{0,3}(u,s) \tag{4.118}$$

- $T_5 : \mathrm{Tr}(F^1F^2F^3F^4)\mathrm{Tr}(F^1F^3F^2F^4)$. This structure has $\mathbb{Z}_2$ symmetry of $2 \leftrightarrow 3$. Most general descendant

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\nu_j}\partial_{\mu_i}(F^1_{ab}F^1_{pq})\partial_{\mu_i}(F^2_{bc}F^2_{rs})\partial_{\nu_j}(F^3_{cd}F^3_{qr})F^4_{da}F^4_{sp} \tag{4.119}$$

$m \leftrightarrow n$ is the symmetry. The S matrix is

$$\begin{aligned}&B^{0,4}(s,t)\left[\left(p^1_a\epsilon^1_b-p^1_b\epsilon^1_a\right)\left(p^2_b\epsilon^2_c-p^2_c\epsilon^2_b\right)\left(p^3_c\epsilon^3_d-p^3_d\epsilon^3_c\right)\left(p^4_d\epsilon^4_a-p^4_a\epsilon^4_d\right)\right.\\&\left.\left(p^1_p\epsilon^1_q-p^1_q\epsilon^1_p\right)\left(p^3_q\epsilon^3_r-p^3_r\epsilon^3_q\right)\left(p^2_r\epsilon^2_s-p^2_s\epsilon^2_r\right)\left(p^4_s\epsilon^4_p-p^4_p\epsilon^4_s\right)\right]\\&+B^{0,4}(s,u)\left[3\leftrightarrow 4\right]+B^{0,4}(u,t)\left[2\leftrightarrow 4\right]\end{aligned} \tag{4.120}$$

With the following symmetry property.

$$B^{0,4}(s,t)=B^{0,4}(t,s) \tag{4.121}$$

- T_6 : $\mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4)\mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4)$. This structure has extra $\mathbb{Z}_2$ symmetry of 1 to 2 exchange. The most general descendant is

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}\left(F^1_{ab}F^1_{pq}\right)F^2_{ba}F^2_{qp}\partial_{\mu_i}\left(F^3_{cd}F^3_{rs}\right)\partial_{\nu_j}\left(F^4_{dc}F^4_{sr}\right) \tag{4.122}$$

Again with $m \leftrightarrow n$ symmetry. S matrix

$$\begin{aligned}&B^{0,5}(t,u)\left[\left(p^1_p\epsilon^1_q-p^1_q\epsilon^1_p\right)\left(p^2_p\epsilon^2_q-p^2_q\epsilon^2_p\right)\left(p^3_r\epsilon^3_s-p^3_s\epsilon^3_r\right)\left(p^4_r\epsilon^4_s-p^4_s\epsilon^4_r\right)\right.\\&\left.\left(p^1_a\epsilon^1_b-p^1_b\epsilon^1_a\right)\left(p^2_a\epsilon^2_b-p^2_b\epsilon^2_a\right)\left(p^3_c\epsilon^3_d-p^3_d\epsilon^3_c\right)\left(p^4_c\epsilon^4_d-p^4_d\epsilon^4_c\right)\right]\\&+B^{0,5}(s,u)\left[3\leftrightarrow 2\right]+B^{0,5}(s,t)\left[2\leftrightarrow 4\right]\end{aligned} \tag{4.123}$$

With the following symmetry property.

$$B^{0,5}(t,u)=B^{0,5}(u,t) \tag{4.124}$$

- T_7 : $\mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4)\mathrm{Tr}(F^1F^4)\mathrm{Tr}(F^2F^3)$. This structure has extra $\mathbb{Z}_2$ symmetry of 1 to 3 exchange. The most general descendant is

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}\left(F^1_{ab}F^1_{pq}\right)\partial_{\mu_i}\left(F^2_{ba}F^2_{rs}\right)F^3_{cd}F^3_{sr}\partial_{\nu_j}\left(F^4_{dc}F^4_{qp}\right) \tag{4.125}$$

Again with $m \leftrightarrow n$ symmetry. S matrix

$$\begin{aligned}&B^{0,6}(s,u)\left[\left(p^1_p\epsilon^1_q-p^1_q\epsilon^1_p\right)\left(p^4_p\epsilon^4_q-p^4_q\epsilon^4_p\right)\left(p^2_r\epsilon^2_s-p^2_s\epsilon^2_r\right)\left(p^3_r\epsilon^3_s-p^3_s\epsilon^3_r\right)\right.\\&\left.\left(p^1_a\epsilon^1_b-p^1_b\epsilon^1_a\right)\left(p^2_a\epsilon^2_b-p^2_b\epsilon^2_a\right)\left(p^3_c\epsilon^3_d-p^3_d\epsilon^3_c\right)\left(p^4_c\epsilon^4_d-p^4_d\epsilon^4_c\right)\right]\\&+B^{0,6}(t,u)\left[3\leftrightarrow 2\right]+B^{0,6}(s,t)\left[3\leftrightarrow 4\right]\end{aligned} \tag{4.126}$$

With the following symmetry property.

$$B^{0,6}(s,u)=B^{0,6}(u,s) \tag{4.127}$$

- T_8 : $\mathrm{Tr}(F^1F^2F^3F^4)F^1_{ab}\mathrm{Tr}(p^2_aF^2p^3_bF^3F^4)$. Where, p^2_a from the Lagrangian perspective is a derivative acting on R^2_{abcd} This structure has neither $\mathbb{Z}_2\times\mathbb{Z}_2$, which although is preserved by the first trace but broken by the $F\mathrm{Tr}(...)$ part, nor it has S^3 which is preserved by the $F\mathrm{Tr}(...)$ part but broken by the $\mathrm{Tr}(F^4)$ part. Only $\mathbb{Z}_2$ is preserved, that is just $2\leftrightarrow 4$ flip symmetry. The descendant is of the general form

$$\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}\left(F^1_{pq}F^1_{ab}\right)\partial_a\partial_{\mu_i}\left(F^2_{qr}F^2_{\mu\nu}\right)\partial_b\left(F^3_{rs}F^3_{\nu\alpha}\right)\partial_{\nu_j}\left(F^4_{sp}F^4_{\alpha\mu}\right) \tag{4.128}$$

The S matrix for this structure is

$$
\begin{aligned}
&+\left(B^{2,1}(s,u)\left(p_p^1\epsilon_q^1-p_q^1\epsilon_p^1\right)\left(p_q^2\epsilon_r^2-p_r^2\epsilon_q^2\right)\left(p_r^3\epsilon_s^3-p_s^3\epsilon_r^3\right)\left(p_s^4\epsilon_p^4-p_p^4\epsilon_s^4\right)\right.\\
&\quad B^{2,1}(t,u)\left(p_p^1\epsilon_q^1-p_q^1\epsilon_p^1\right)\left(p_q^3\epsilon_r^3-p_r^3\epsilon_q^3\right)\left(p_r^2\epsilon_s^2-p_s^2\epsilon_r^2\right)\left(p_s^4\epsilon_p^4-p_p^4\epsilon_s^4\right)\\
&+\left.B^{2,1}(t,s)\left(p_p^1\epsilon_q^1-p_q^1\epsilon_p^1\right)\left(p_q^3\epsilon_r^3-p_r^3\epsilon_q^3\right)\left(p_r^4\epsilon_s^4-p_s^4\epsilon_r^4\right)\left(p_s^2\epsilon_p^2-p_p^2\epsilon_s^2\right)\right)\\
&\quad\left(\left(p_a^1\epsilon_b^1-p_b^1\epsilon_a^1\right)p_a^2\left(p_\mu^2\epsilon_\nu^2-p_\nu^2\epsilon_\mu^2\right)p_b^3\left(p_\nu^3\epsilon_\alpha^3-p_\alpha^3\epsilon_\nu^3\right)\left(p_\alpha^4\epsilon_\mu^4-p_\mu^4\epsilon_\alpha^4\right)\right.\\
&\quad+\left(p_a^2\epsilon_b^2-p_b^2\epsilon_a^2\right)p_a^1\left(p_\mu^1\epsilon_\nu^1-p_\nu^1\epsilon_\mu^1\right)p_b^4\left(p_\nu^4\epsilon_\alpha^4-p_\alpha^4\epsilon_\nu^4\right)\left(p_\alpha^3\epsilon_\mu^3-p_\mu^3\epsilon_\alpha^3\right)\\
&\quad+\left(p_a^3\epsilon_b^3-p_b^3\epsilon_a^3\right)p_a^4\left(p_\mu^4\epsilon_\nu^4-p_\nu^4\epsilon_\mu^4\right)p_b^1\left(p_\nu^1\epsilon_\alpha^1-p_\alpha^1\epsilon_\nu^1\right)\left(p_\alpha^2\epsilon_\mu^2-p_\mu^2\epsilon_\alpha^2\right)\\
&\quad\left.+\left(p_a^4\epsilon_b^4-p_b^4\epsilon_a^4\right)p_a^3\left(p_\mu^3\epsilon_\nu^3-p_\nu^3\epsilon_\mu^3\right)p_b^2\left(p_\nu^2\epsilon_\alpha^2-p_\alpha^2\epsilon_\nu^2\right)\left(p_\alpha^1\epsilon_\mu^1-p_\mu^1\epsilon_\alpha^1\right)\right)
\end{aligned}
\tag{4.129}
$$

With the following symmetry property.

$$
B^{2,1}(s,u)=B^{2,1}(u,s) \tag{4.130}
$$

With total of 12 terms, which is $S^4/\mathbb{Z}_2$.

- T_9 : $\mathrm{Tr}(F^1F^2)\mathrm{Tr}(F^3F^4)F^1_{ab}\mathrm{Tr}(p_a^2F^2p_b^3F^3F^4)$. This structure again has only $\mathbb{Z}_2$ symmetry of $3\leftrightarrow 4$.

  Descendants

$$
\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}\left(F^1_{pq}F^1_{ab}\right)\partial_aF^2_{qp}F^2_{\mu\nu}\partial_b\partial_{\mu_i}\left(F^3_{rs}F^3_{\nu\alpha}\right)\partial_{\nu_j}\left(F^4_{sr}F^4_{\alpha\mu}\right) \tag{4.131}
$$

  This has $m\leftrightarrow n$ symmetry. S matrix

$$
\begin{aligned}
&\left(B^{2,2}(t,u)\left(p_p^1\epsilon_q^1-p_q^1\epsilon_p^1\right)\left(p_q^2\epsilon_q^2-p_q^2\epsilon_p^2\right)\left(p_r^3\epsilon_s^3-p_s^3\epsilon_r^3\right)\left(p_r^4\epsilon_s^4-p_s^4\epsilon_r^4\right)\right.\\
&+B^{2,2}(s,u)\left(p_p^1\epsilon_q^1-p_q^1\epsilon_p^1\right)\left(p_p^3\epsilon_q^3-p_q^3\epsilon_p^3\right)\left(p_r^2\epsilon_s^2-p_s^2\epsilon_r^2\right)\left(p_r^4\epsilon_s^4-p_s^4\epsilon_r^4\right)\\
&+\left.B^{2,2}(t,s)\left(p_p^1\epsilon_q^1-p_q^1\epsilon_p^1\right)\left(p_p^4\epsilon_q^4-p_q^4\epsilon_p^4\right)\left(p_r^3\epsilon_s^3-p_s^3\epsilon_r^3\right)\left(p_r^2\epsilon_s^2-p_s^2\epsilon_r^2\right)\right)\\
&\quad\left(\left(p_a^1\epsilon_b^1-p_b^1\epsilon_a^1\right)p_a^2\left(p_\mu^2\epsilon_\nu^2-p_\nu^2\epsilon_\mu^2\right)p_b^3\left(p_\nu^3\epsilon_\alpha^3-p_\alpha^3\epsilon_\nu^3\right)\left(p_\alpha^4\epsilon_\mu^4-p_\mu^4\epsilon_\alpha^4\right)\right.\\
&\quad+\left(p_a^2\epsilon_b^2-p_b^2\epsilon_a^2\right)p_a^1\left(p_\mu^1\epsilon_\nu^1-p_\nu^1\epsilon_\mu^1\right)p_b^4\left(p_\nu^4\epsilon_\alpha^4-p_\alpha^4\epsilon_\nu^4\right)\left(p_\alpha^3\epsilon_\mu^3-p_\mu^3\epsilon_\alpha^3\right)\\
&\quad+\left(p_a^3\epsilon_b^3-p_b^3\epsilon_a^3\right)p_a^4\left(p_\mu^4\epsilon_\nu^4-p_\nu^4\epsilon_\mu^4\right)p_b^1\left(p_\nu^1\epsilon_\alpha^1-p_\alpha^1\epsilon_\nu^1\right)\left(p_\alpha^2\epsilon_\mu^2-p_\mu^2\epsilon_\alpha^2\right)\\
&\quad\left.+\left(p_a^4\epsilon_b^4-p_b^4\epsilon_a^4\right)p_a^3\left(p_\mu^3\epsilon_\nu^3-p_\nu^3\epsilon_\mu^3\right)p_b^2\left(p_\nu^2\epsilon_\alpha^2-p_\alpha^2\epsilon_\nu^2\right)\left(p_\alpha^1\epsilon_\mu^1-p_\mu^1\epsilon_\alpha^1\right)\right)
\end{aligned}
\tag{4.132}
$$

  With the following symmetry property.

$$
B^{2,2}(t,u)=B^{2,2}(u,t) \tag{4.133}
$$

- T_{10} : $F^1_{pq}\mathrm{Tr}(p_p^2F^2p_q^3F^3F^4)F^1_{ab}\mathrm{Tr}(p_a^2F^2p_b^3F^3F^4)$. This structure has S^3 symmetry, because $2,3,4$ can be permuted and the structure remains invariant. The most general descendant of this structure is

$$
\prod_{i=1}^{m}\prod_{j=1}^{n}\partial_{\mu_i}\partial_{\nu_j}(F^1_{ab}F^1_{pq})\partial_p\partial_a(F^2_{\mu\nu}F^2_{\beta\gamma})\partial_q\partial_b(F^3_{\nu\alpha}F^3_{\gamma\delta})F^4_{\alpha\mu}F^4_{\delta\beta} \tag{4.134}
$$

The S matrix contribution from this structure is

$$\begin{aligned}&\left(B^{4,1}(s,t)+B^{4,1}(t,u)+B^{4,1}(u,s)\right)\times\\&\left[\left(p_a^1\epsilon_b^1-p_b^1\epsilon_a^1\right)p_a^2\left(p_\mu^2\epsilon_\nu^2-p_\nu^2\epsilon_\mu^2\right)p_b^3\left(p_\nu^3\epsilon_\alpha^3-p_\alpha^3\epsilon_\nu^3\right)\left(p_\alpha^4\epsilon_\mu^4-p_\mu^4\epsilon_\alpha^4\right)\right.\\&\left.\left(p_p^1\epsilon_q^1-p_q^1\epsilon_p^1\right)p_p^2\left(p_\beta^2\epsilon_\gamma^2-p_\gamma^2\epsilon_\beta^2\right)p_q^3\left(p_\gamma^3\epsilon_\delta^3-p_\delta^3\epsilon_\gamma^3\right)\left(p_\delta^4\epsilon_\beta^4-p_\beta^4\epsilon_\delta^4\right)\right.\\&\left.+(1\leftrightarrow 2)+(1\leftrightarrow 3)+(1\leftrightarrow 4)\right]\end{aligned}\tag{4.135}$$

With the following symmetry property.

$$B^{4,1}(s,t)=B^{4,1}(u,t)=B^{4,1}(t,s)=B^{4,1}(u,s)=B^{4,1}(s,u)=B^{4,1}(t,u)\tag{4.136}$$

A priori there are many 'structures' one can write down for a particular kind of term, but with the use of various identities the above list is the minimal independent set.

4.6.6 Reduction of Structures in lower dimensions

Clearly, the first structure, T_1 starts to contribute to the s-matrix from $D\geq 7$, because in $D=6$ it becomes identically zero. So, in 6 dimensions we have 28 degrees of freedom.

In $D=5$ there are two relations between T_3,T_4,T_5,T_7, removing two of the structures entirely. So, we are left with 7 structures with $28-6=22$ degrees of freedom.

$$\begin{aligned}T_7&=\frac{1}{8}\left(T_4-4T_5\right)\\T_3&=-\frac{T_5}{4}\end{aligned}\tag{4.137}$$

In $D=4$, T_2 vanishes identically. Both T_8,T_9 are descendants. Also, there is one more relationship between T_3,T_4,T_5,T_7, resulting in the left over structures, T_6, T_{10} and one of T_3,T_4,T_5,T_7 (say T_3).

$$T_7=\frac{T_3}{2}\tag{4.138}$$

So, far count for the degrees of freedom is $3+3+1=7$. But there is again a linear combination of T_3 and T_6 which becomes fully symmetric, thereby reducing the count of degrees of freedom to 5 as is expected from the CFT counting. The two derivative descendant s matrix from this fully symmetric structure must vanish, we have found the s matrix from the combinations of the two structures T_3 and T_6 that vanishes, but the fully symmetric combination in has yet to be computed explicitly, but it will be written in the paper which is yet to appear.

4.6.7 String tree level 4 graviton S matrices

Einstein gravity

In this section we reproduce the s-matrix coming from Einstein gravity in our basis. The scattering amplitude for Einstein gravity is given by [141],

$$
\begin{aligned}
A_{4h}^{EG} &= \frac{-4\kappa^2}{stu}\left(\frac{1}{2}\epsilon_2.\epsilon_3\left(s\epsilon_1.k_3\epsilon_4.k_2+t\epsilon_1.k_2\epsilon_4.k_3\right)+\frac{1}{2}\epsilon_1.\epsilon_4\left(s\epsilon_2.k_4\epsilon_3.k_1+t\epsilon_2.k_1\epsilon_3.k_4\right)\right.\\
&+\frac{1}{2}\epsilon_2.\epsilon_4\left(s\epsilon_1.k_4\epsilon_3.k_2+u\epsilon_1.k_2\epsilon_3.k_4\right)+\frac{1}{2}\epsilon_1.\epsilon_3\left(s\epsilon_2.k_3\epsilon_4.k_1+u\epsilon_2.k_1\epsilon_4.k_3\right)\\
&+\frac{1}{2}\epsilon_3.\epsilon_4\left(t\epsilon_1.k_4\epsilon_2.k_3+u\epsilon_1.k_3\epsilon_2.k_4\right)+\frac{1}{2}\epsilon_1.\epsilon_2\left(t\epsilon_3.k_2\epsilon_4.k_1+u\epsilon_3.k_1\epsilon_4.k_2\right)\\
&\left.-\frac{1}{4}st\epsilon_1.\epsilon_4\epsilon_2.\epsilon_3-\frac{1}{4}su\epsilon_1.\epsilon_3\epsilon_2.\epsilon_4-\frac{1}{4}tu\epsilon_1.\epsilon_2\epsilon_3.\epsilon_4\right)^2 \qquad (4.139)
\end{aligned}
$$

This can be written in term of our basis structures, and from there we can read off the Lagrangian which produces the same $2-2$ s-matrix at tree level (just the tensor structure, without the $\frac{1}{stu}$).

$$
\begin{aligned}
L_{4h}^{EG} &\propto \frac{1}{32}(R_{pqrs}R^{pqrs})^2-\frac{1}{2}R^{pqrs}R_{pqr}{}^{t}R^{uvw}{}_{s}R_{uvwt}+\frac{1}{16}R^{pqrs}R_{pq}{}^{tu}R_{tu}{}^{vw}R_{rsvw}\\
&-\frac{1}{4}R^{pqrs}R_{pq}{}^{tu}R_{rt}{}^{vw}R_{suvw}-R^{pqrs}R_{p}{}^{t}{}_{r}{}^{u}R_{tvws}R_{q}{}^{v}{}_{u}{}^{w}+\frac{1}{2}R^{pqrs}R_{p}{}^{t}{}_{r}{}^{u}R_{t}{}^{v}{}_{u}{}^{w}R_{qvsw} \qquad (4.140)
\end{aligned}
$$

Type II

The 4-graviton amplitude in Type II superstring theory is given by [141]

$$
A_{4h}^{ss} = h(s,t,u,\alpha')A_{4h}^{EG} \qquad (4.141)
$$

Hence it is reproduced by the same Lagrangian term which produces Einstein gravity upto momentum factors.

4.6.8 Heterotic String

The Heterotic string tree level amplitude can be written in our basis as

$$
A_2T_2+A_3T_3+A_4T_4+A_5T_5+A_6T_6+A_7T_7+A_8T_8+A_9T_9 \qquad (4.142)
$$

Where,

$$
\begin{aligned}
A_2 &= \frac{1}{8}\left(s^2t + st^2 + \frac{t^3}{2}\right) + \frac{1}{16}\left(s^2t^2 + \frac{3st^3}{2} + \frac{t^4}{2}\right) + \frac{1}{4}\left(-\frac{st}{2} - \frac{t^2}{2}\right) - \frac{t}{4} \\
A_3 &= \frac{1}{8}\left(\frac{1}{2}\left(-t^3 - u^3\right) + \frac{1}{2}\left(t\left(-u^2\right) - t^2u\right)\right) + \frac{1}{16}\left(\frac{1}{2}\left(t\left(-u^3\right) - t^3u\right) - \frac{t^2u^2}{2}\right) + \frac{tu}{8} + \frac{t+u}{4} \\
A_4 &= \frac{1}{8}\left(2\left(s^3 + u^3\right) + 4\left(s^2u + su^2\right)\right) + \frac{1}{16}\left(2\left(s^3u + su^3\right) + 4s^2u^2\right) - s - u \\
A_5 &= \frac{1}{8}\left(2\left(s^3 + t^3\right) + 4\left(s^2t + st^2\right)\right) + \frac{1}{16}\left(2\left(s^3t + st^3\right) + 4s^2t^2\right) - s - t \\
A_6 &= -\frac{1}{128}t^2u^2 + \frac{1}{64}\left(t\left(-u^2\right) - t^2u\right) - \frac{tu}{32} \\
A_7 &= \frac{1}{32}\left(s\left(-u^2\right) - s^2u\right) + \frac{1}{4}\left(\frac{1}{8}\left(s^2 + u^2\right) + \frac{su}{4}\right) + \frac{1}{16}\left(\frac{1}{8}\left(-s^4 - u^4\right) + \frac{1}{4}\left(s\left(-u^3\right) - s^3u\right) - \frac{1}{4}s^2u^2\right) \\
A_8 &= \frac{1}{32}\left(s\left(-u^2\right) - s^2u\right) + \frac{1}{4}\left(\frac{1}{4}\left(-s^2 - u^2\right) - \frac{su}{4}\right) + \frac{1}{4} \\
A_9 &= \frac{1}{128}\left(t^2u + tu^2\right) + \frac{1}{4}\left(\frac{1}{16}\left(t^2 + u^2\right) + \frac{tu}{16}\right) - \frac{1}{16}
\end{aligned}
\tag{4.143}
$$

4.6.9 Closed Bosonic String

Previously in 4.5.8, we wrote the open bosonic string theory amplitude using our F structures. Now, closed bosonic string theory amplitude is just the square of the open string amplitude. All, the possible ways to square our F structures, including the cross terms, are included in the listing of 10 basis structures for the case of gravity. Therefore, it is obvious that we would be able to write the closed bosonic string tree level amplitude in our basis, thereby writing an effective Lagrangian which produces the same s-matrix at the tree level.

4.7 Conclusions

We conclude by classifying all the possible S matrix structures[11] for $2-2$ scattering for an effective theory of pure electrodynamics as well as a classification for an effective theory of pure gravity. We find that for electrodynamics, this scattering is completely specified by 3 polynomials of s, t, u, i.e. any general S matrix can be written down as a linear combination of these three functions dressed with a known dependence on polarization and momenta. In the case of gravity there are 10 such polynomials corresponding to 10 structures listed in section 4.6.5. The dependence on polarization is completely fixed. In the future extensions

[11] For electrodynamics our conclusions are valid in dimension ≥ 5, and for gravity they are valid in dimension ≥ 7.

of this work, we want to impose the physical input of causality and constrain the possible physical theories of electrodynamics and gravity. The expectation is that for the case of gravity, only Einstein theory survives. Next we plan to study the $2-2$ exchange diagrams involving massive spin particles and try to understand how String theory continues to stay causal. Our conclusions are valid if a finite number of terms are involved in the Lagrangian but String theory involves an infinite tower of higher spin massive particles. We also wish to understand if there is a fine-tuning happening in the case of string theory, also how constrained it is.

Appendix A

Appendices for Chapter 1

A.1 Conformal Kernel

In this appendix following main result is proved

$$\int dt_3\ dt_4\ \widetilde{K_c}(t_1,t_2;t_3,t_4)(g_c)_a{}^b(t_3,t_4) = \frac{1}{|J|^2}\ (g_c)_a{}^b(t_1,t_2), \tag{A.1}$$

where relevant quantities are defined by

$$\begin{aligned} \widetilde{K_c}(t_1,t_2;t_3,t_4) =& -|G_c(t_1,t_2)|^{\frac{q-2}{2}} G_c(t_1,t_3)G_c(t_2,t_4)|G_c(t_3,t_4)|^{\frac{q-2}{2}}, \\ g_c(t_1,t_2) =& |G_c(t_1,t_2)|^{\frac{q-2}{2}} G_c(t_1,t_2)[H(t_1)-H(t_2)]. \end{aligned} \tag{A.2}$$

Important part of the integration is given by:

$$\begin{aligned} Q(t_1,t_2) \equiv& \int dt_3\ dt_4\ G_c(t_1,t_3)G_c(t_2,t_4)G_c(t_3,t_4)^{q-1}[H(t_3)-H(t_4)] \\ =& -\frac{1}{|J|^2}\int dt_3\ G_c(t_1,t_3)H(t_3)\int dt_4\ G_c(t_2,t_4)\ |J|^2 G_c(t_4,t_3)^{q-1} \\ & -\frac{1}{|J|^2}\int dt_4\ G_c(t_2,t_4)H(t_4)\int dt_3\ G_c(t_1,t_3)\ |J|^2 G_c(t_3,t_4)^{q-1} \\ =& -\frac{1}{|J|^2}\int dt_3\ G_c(t_1,t_3)H(t_3)(-\delta(t_2-t_3)) - \frac{1}{|J|^2}\int dt_4\ G_c(t_2,t_4)H(t_4)(-\delta(t_1-t_4)) \\ =& \frac{1}{|J|^2}\left[G_c(t_1,t_2)H(t_2)+G_c(t_2,t_1)H(t_1)\right] \\ =& -\frac{1}{|J|^2}G_c(t_1,t_2)\left[H(t_1)-H(t_2)\right]. \end{aligned} \tag{A.3}$$

This proves claimed result when multiplied with $-|G_c(t_1,t_2)|^{\frac{q-2}{2}}$.

A.2 Details of the perturbative computations

A.2.1 Leading Power of β

Two melon graphs

In this subsection we consider the contribution to the free energy given by fig. A.1. First non-trivial effect of winding is seen at this level as explained below. The term whose Wick contraction is calculated is $\frac{1}{4!}(J\psi^4+h.c.)^4$ - where each of ${}^4C_{4/2}$ terms contribute the same. The symmetry factor is calculated as follows. Any one of q number of ψ's of first ψ-vortex

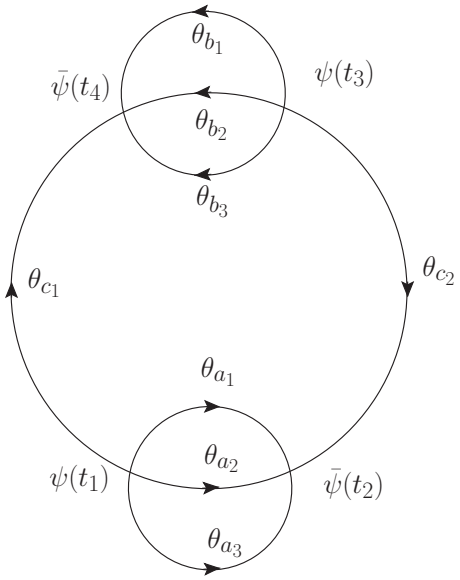


Figure A.1: Direction of arrow is from ψ to $\overline{\psi}$. The diagram is drawn for $q = 4$.

contracts with any one of q number of $\overline{\psi}$'s of any one of two $\overline{\psi}$-vortex to give a factor of $2q^2$. Any one of q number of ψ's of second ψ-vortex contracts with any one of q number of $\overline{\psi}$'s of remaining $\overline{\psi}$-vortex to give a factor of q^2. In large-N only non-suppressed diagram is obtained by joining ψ to $\overline{\psi}$ (of same vortex) of same common colour. Choice of external propagator gives $q-1$ possibilities at each blob. Sign of the symmetry factor comes from noticing as there are two identical 'blobs' sign of contraction of each blob cancel and overall sign is just because of contraction between two 'blobs', it turns out to be -1. Contribution of symmetry factor at this order becomes

$$F_4 = \frac{1}{4!}\ {}^4C_{4/2}\ (-1)2[q^2(q-1)]^2 I^{(4)}, \tag{A.4}$$

where

$$I^{(4)} = \int \prod_{i=1}^{4} dt_i \ \left(\prod_{i=1}^{q-1} G_0(t_{12}, \theta_{a_i})\right) \left(\prod_{i=1}^{q-1} G_0(t_{34}, \theta_{b_i})\right) G_0(t_{32}, \theta_{c_2}) G_0(t_{14}, \theta_{c_1}). \tag{A.5}$$

Where θs are holonomies on different propagators. Here time differences are not necessarily single valued and to satisfy the constraint

$$t_{12} + t_{23} + t_{34} + t_{41} = w\beta,$$

where $w = 0, \pm 1, \pm 2$ (note that each t_{ik} is in $(-\frac{\beta}{2}, \frac{\beta}{2})$, and this restricts allowed values of n) we introduce dimensionless Lagrange multiplier integration

$$P \equiv \beta \int_{-\infty}^{+\infty} \frac{ds}{2\pi} e^{is(t_{12}+t_{23}+t_{34}+t_{41}-w\beta)} = \delta\left(\frac{t_{12}+t_{23}+t_{34}+t_{41}-w\beta}{\beta}\right). \tag{A.6}$$

In the scaling limit (assuming $m > 0$), the propagator becomes

$$G_0(t) = e^{-(m+i\theta_a)t}\theta(t) - xe^{-i\theta_a\beta}e^{-(m+i\theta_a)t}\theta(t) - x^{1/2}e^{-m\beta/2}e^{-i\theta_a\beta}e^{-(m+i\theta_a)t}\theta(-t). \tag{A.7}$$

This way of writing ensures in each of three parts of G_0 excluding explicit x dependence integration over $-\frac{\beta}{2}$ to $\frac{\beta}{2}$ gives only positive powers of x. We will refer to these three parts of G_0 as $x^0, x, x^{1/2}$ contributions.

In the scaling limit of interest $I^{(4)}$ can receive contribution from 5 different types of integration

$$\begin{aligned} I^{(4)} =& x^0 \textit{ everywhere} + x^{1/2} \textit{ on one of the outer } (\theta_{c_1}, \theta_{c_2}) \textit{ lines} + x^{1/2} \textit{ on both of the outer lines} \\ &+ x \textit{ on one of the outer lines} + x \textit{ on one of the inner lines } (\theta_{a_1}, \theta_{a_2} \ldots \theta_{a_{q-1}}, \theta_{b_1}, \theta_{b_2} \ldots \theta_{b_{q-1}}). \end{aligned} \tag{A.8}$$

Note that choosing $x^{1/2}$ on one of the inner propagators will force choosing all the inner propagators in the same blob to be $x^{1/2}$ term due to unit step function. Therefore this choice is ignored in scaling limit calculation. Here we'll present the calculation corresponding to the first one and mention results for others.

Consider $x^{1/2}$ on θ_{c_1} say, and on all others we choose x independent part of G_0. This ensures following time ordering for non-zero integrand $t_{12} > 0, t_{32} > 0, t_{34} > 0, t_{41} > 0$, with which only consistent values of n are $0, 1$. Contribution to $I^{(4)}$ becomes, omitting $\beta(-x^{1/2})e^{-i\theta_{c_1}\beta}e^{+i\theta_{c_1}w\beta}$ (for a contribution like F_0 we must have $n = 0$ which is shown to be true below)

$$\begin{aligned} I^{(4)} \sim \beta(-x^{1/2})e^{-i\theta_{c_1}\beta}e^{+i\theta_{c_1}w\beta} \int \frac{ds}{2\pi}\ dt_{12}\ dt_{32}\ dt_{34}\ dt_{41}\ e^{-isw\beta}\ e^{-(m(q-1)-is)t_{12}} \times \\ e^{-(m(q-1)-is)t_{34}}\ e^{+(m+is)t_{41}-m\beta/2}e^{-(m+is)t_{32}} \\ = -\beta(-x^{1/2})e^{-i\theta_{c_1}\beta}e^{+i\theta_{c_1}w\beta} \int \frac{ds}{2\pi}\ \frac{(e^{is\beta/2}-x^{1/2})(x^{1/2}e^{-is\beta/2}-1)}{(s+i(q-1)m)^2(s-im)^2}e^{-isw\beta} + \mathcal{O}(x^{3/2}), \end{aligned} \tag{A.9}$$

where we ignored higher order contributions in x. Simplifying the numerator gives 3 terms: x independent piece that comes with a non-zero phase factor $e^{is\beta/2}$ (which will give a factor of β upon integration because only $w=0$ will contribute), $x^{1/2}$ term that comes with no non-trivial phase (cannot give a β upon integration), x term drops out in scaling limit. Rest of the integration can be done easily choosing proper contour (semi-circle on upper or lower half plane as required by convergence) to ensure only $w=0$ term contributes to give the following result

$$\delta_{w,0} x^{1/2} \frac{2}{(qm)^3} \left(-1 + \frac{q}{4} m\beta\right). \tag{A.10}$$

All other integrations can be performed similarly to give leading order contribution to free energy

$$F_4 = \frac{1}{4!}\ {}^4C_{4/2}\ 2q^4 \frac{(q-1)^2}{q^2} m^2\beta^2\ N^{(q-1)^2} x \prod_{m=1}^{q-1} \rho_m^1 + \mathcal{O}(\beta). \tag{A.11}$$

n melon graphs

Here a circle diagram with $n \geq 2$ blobs is considered and leading term in β is calculated using methods demonstrated in previous sub-section.

Symmetry factor for the diagram in large N limit is [1]

$$\frac{(-1)^{\frac{nq}{2}+n+1}}{(n!)^2}\ \ n!\ (q^2)^n\ (n-1)!. \tag{A.12}$$

The leading order contribution in β comes from two distinct choices - i) considering $x^{1/2}$ in any one of the n external propagators (with holonomy θ_a say) with x^0 part of the free propagator in all others and ii) x^0 part of the free propagator in all propagators.

Contribution from the integral due to choice (i) is easily seen to be

$$\begin{aligned} &- x^{1/2}|g|^n\ e^{-i\theta_a\beta + iw\theta_a\beta}\ \beta \int \frac{ds}{2\pi}\ e^{-i(w-\frac{1}{2})s\beta}\ \frac{1}{(-is+m(q-1))^n (is+m)^n} \\ &= -2x\ |g|^n\ e^{-i\theta_a\beta} \frac{1}{(n-1)!} \left(\frac{\beta}{2mq}\right)^n\ \delta_{w,0}, \end{aligned} \tag{A.13}$$

where we have kept only highest power of β. Note that extra powers of beta β^{n-1} came from the integration because of evaluation of residue around a pole of order n. This contribution is to be multiplied with a factor of n due to freedom in choosing one external propagator on which $x^{1/2}$ is considered.

Now we turn to the choice (ii). In this case contribution to the integral is

$$\begin{aligned} &|g|^n\ e^{iw\theta_a\beta}\ \beta \int \frac{ds}{2\pi}\ e^{-iws\beta}\ \frac{(1-x^{1/2}e^{-is\frac{\beta}{2}})^n}{(-is+m(q-1))^n (is+m)^n} \\ &= 2x\ |g|^n\ e^{-i\theta_a\beta} \frac{1}{(n-1)!} \left(\frac{\beta}{2mq}\right)^n \left(\frac{2^{n-1}}{n} - 1\right)\ \delta_{w,1}. \end{aligned} \tag{A.14}$$

[1] Here an extra factor of $(n-1)!$ comes as compared to $n=2$ case because of freedom of joining n blobs with one another.

As before we have kept only highest power of β. Note that this contribution vanishes for $n = 2$.

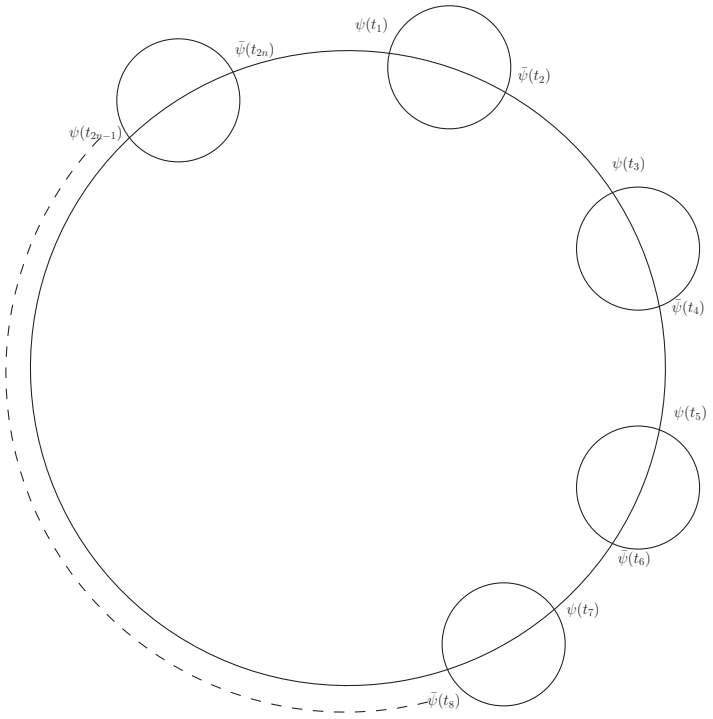
After summing over the holonomies, and canceling loop N's with that of scaling of g, contribution to free energy becomes

$$F_{2n} = 2x\ N^{q-1} \left(\prod_{m=1}^{q-1} \rho_m^1 \right)\ \frac{1}{(n-1)!} \left[\gamma(q)\ \frac{(-\beta)}{m} |J|^2 \right]^n \left(2 - \frac{2^{n-1}}{n} \right) + \mathcal{O}(\beta^{n-1}), \quad \text{(A.15)}$$

where

$$\gamma(q) = (-1)^{\frac{q}{2}(q-1)} \frac{q}{2}. \quad \text{(A.16)}$$

A.2.2 All powers of β in a circle diagram



In this subsection we shall compute explicitly the integral involved in computing the contribution to the free energy in the scaling limit linear in $x = e^{-m\beta}$.

The free fermionic Green's function at any finite temperature is given by,

$$\begin{aligned} \langle \psi(t)\overline{\psi}(0) \rangle &\equiv G_0(t) \\ &= \frac{1}{2} e^{-(m+i\alpha_j)t} \Big[\mathrm{sgn}(t) + \tanh\left(\frac{\beta}{2}(m + i\alpha_j) \right) \Big] \\ &= e^{-(m+i\alpha_j)t} \Big[\theta(t) - x e^{-i\alpha_j \beta} \Big], \end{aligned} \quad \text{(A.17)}$$

where, $x = e^{-m\beta} << 1$ (scaling limit). Hence, one can also write the 'reversed' Green's function at finite temperature as,

$$\langle\overline{\psi}(0)\psi(t)\rangle = G_0^*(-m) \;\; = \;\; \frac{1}{2}e^{(m+i\alpha_j)t}\Big[\mathrm{sgn}(t) - \tanh\Big(\frac{\beta}{2}(m+i\alpha_j)\Big)\Big]. \qquad \text{(A.18)}$$

Here α_j are holonomies, satisfying the following constraint

$$\sum_{j=1}^{q}\alpha_j = 0\,. \qquad \text{(A.19)}$$

Now in the computation we use discrete representation of the delta function

$$\begin{aligned}&\delta(t_{21}+t_{32}+t_{43}+t_{54}+t_{65}+\;\ldots - t_{\overline{2n-1}\;\overline{2n}}+t_{1\;\overline{2n}})\\ &= \frac{1}{2\pi\beta}\sum_{\omega=-\infty}^{\infty} e^{-2\pi i\frac{\omega}{\beta}(t_{21}+t_{32}+t_{43}+t_{54}+t_{65}+\;\ldots\;-t_{\overline{2n-1}\;\overline{2n}}+t_{1\;\overline{2n}}\;)}. \end{aligned}\qquad \text{(A.20)}$$

Evaluating the integral

Let us focus on the diagram which can be computed as using the integral,

$$\begin{aligned}I^{(2n)} \;\; = \;\; \tfrac{1}{2\pi\beta}\left(\tfrac{J}{4}\right)^{2n}\sum_{\omega=-\infty}^{\infty}\Bigg[&\int_{-\beta/2}^{\beta/2}dt_1\; e^{-t_1(m+i\alpha_q)}e^{-2\pi i\frac{\omega}{\beta}t_1}\left(\mathrm{sgn}(t_1)+\tanh(\tfrac{m\beta+i\alpha_q\beta}{2})\right)\\ &\int_{-\beta/2}^{\beta/2}dt_q\; e^{t_q((q-1)m-i\alpha_q)}e^{-2\pi i\frac{\omega}{\beta}t_q}\left(A\;\mathrm{sgn}(t_q)-B\right)\Bigg]^n\,. \end{aligned}\qquad \text{(A.21)}$$

Here the first integral inside the sum is a single propagator while the second one represents the melon with $q-1$ propagators, where A and B are defined as

$$\prod_{j=1}^{q-1}\Big[\mathrm{sgn}(t_q) - \tanh(\tfrac{m\beta+i\alpha_j\beta}{2})\Big] = (-1)^q\left(A\;\mathrm{sgn}(t_q) - B\right)\,. \qquad \text{(A.22)}$$

We integrate over the time intervals of these propagators in (A.21) and since there are n of them we raise it to the power n. However, we would also have to implement the constraint that the times add up to an integral of β. This is achieved by representing the delta function on a circle of length β as an infinite sum. This contributes a factor of $e^{2\pi i\frac{\omega}{\beta}t_i}$ in each of the propagators as shown in (A.21).

Now we would like to focus on the integrals within the box brackets in (A.21)

$$F^{(2n)} \;\; = \;\; \sum_{\omega=-\infty}^{\infty}\left[I_\omega^{(q)}\right]^n. \qquad \text{(A.23)}$$

Upon integrating over t_1 and t_q one finds that

$$I_\omega^{(q)} = \frac{f_1+f_2}{((q-1)m-i\alpha_q+z)(m+i\alpha_q-z)},\ z=-\frac{2\pi i\omega}{\beta} \tag{A.24}$$

Here f_1 consists of terms with $e^{\pm kz\beta}$ where $k\in\mathbb{Z}_{\rm even}$ while f_2 consists of terms $e^{\pm kz\beta/2}$ where $k\in\mathbb{Z}_{\rm odd}$. Its is evident that upon raising $I_\omega^{(q)}$ to n one would have to evaluate sums in z of the form

$$\begin{aligned} S_1 &= \sum_{\omega=-\infty}^{\infty}\frac{e^{\pm kz\beta}}{(((q-1)m-i\alpha_q+z)(m+i\alpha_q-z))^n},\ k\in\mathbb{Z}_{\rm even}\\ S_2 &= \sum_{\omega=-\infty}^{\infty}\frac{e^{\pm kz\beta/2}}{(((q-1)m-i\alpha_q+z)(m+i\alpha_q-z))^n},\ k\in\mathbb{Z}_{\rm odd}. \end{aligned} \tag{A.25}$$

Since $z=-\frac{2\pi i\omega}{\beta}$ we see that these reduce to

$$\begin{aligned} S_1 &= \sum_{\omega=-\infty}^{\infty}\frac{1}{(((q-1)m-i\alpha_q+z)(m+i\alpha_q-z))^n},\\ S_2 &= \sum_{\omega=-\infty}^{\infty}\frac{e^{z\beta/2}}{(((q-1)m-i\alpha_q+z)(m+i\alpha_q-z))^n}. \end{aligned} \tag{A.26}$$

We will use the technique of Matsubara summation to evaluate the above, where a weighting function is included to replace the sum by a contour integral. So, first let us evaluate S_1. With a weighting function $f(z)=\dfrac{1}{1-e^{z\beta}}$, one can replace the above summation with the following contour integral,

$$S_1=\oint\frac{dz}{(1-e^{z\beta})(((q-1)m-i\alpha_q+z)(m+i\alpha_q-z))^n} \tag{A.27}$$

Notice that the integrand has two poles at $z\equiv z_a=-(q-1)m+i\alpha_q$ and $z\equiv z_b=m+i\alpha_q$ and both are of n-th order. Using the residue theorem, one can evaluate the above integral as,

$$S_1=\lim_{z\to z_a}\frac{1}{(n-1)!}\partial_z^{(n-1)}\frac{1}{(1-e^{z\beta})(z-z_b)^n}+\lim_{z\to z_b}\frac{1}{(n-1)!}\partial_z^{(n-1)}\frac{1}{(1-e^{z\beta})(z-z_a)^n} \tag{A.28}$$

Now, it is very easy to verify that for any function $f(z)$,

$$\partial_z^{(n-1)}\Big[f(z)\frac{1}{(z-z_a)^n}\Big]=\sum_{k=0}^{n-1}(-1)^k\ {}^{(n-1)}C_k\ \frac{(n+k-1)!}{(n-1)!}\ \frac{\partial_z^{(n-k-1)}f(z)}{(z-z_a)^{n+k}} \tag{A.29}$$

In the present case, taking $f(z)=\dfrac{1}{(1-e^{\beta z})}$, one can evaluate

$$\partial_z^{(n)}\Big[\frac{1}{1-e^{\beta z}}\Big]=\frac{\beta^n e^{-\beta z}}{(e^{-\beta z}-1)^{n+1}}A(n) \tag{A.30}$$

where, $A(n)$ is the Eulerian polynomial in $e^{-\beta z}$, given by,

$$A(n) = \sum_{m=0}^{n-1} \sum_{k=0}^{m+1} (-1)^k \ {}^{n+1}C_k (m+1-k)^n \ e^{-\beta z m} \tag{A.31}$$

Using equation (A.30) and (A.31), one can easily obtain,

$$\begin{aligned} \partial_z^{(n-k-1)} f(z) &= \partial_z^{(n-k-1)} \Big[\frac{1}{1-e^{\beta z}} \Big] \\ &= \frac{\beta^{n-k-1}}{(e^{-\beta z}-1)^{n-k}} \sum_{m=0}^{n-k-2} \sum_{l=0}^{m+1} (-1)^l \ {}^{n-k}C_l \ (m+1-l)^{n-k-1} e^{-\beta(m+1)z} \\ &+ \frac{1}{1-e^{\beta z}} \delta_{n-k-1,0} \end{aligned} \tag{A.32}$$

Finally Substituting equation (A.32) into equation (A.29), we have,

$$\begin{aligned} \partial_z^{(n-1)} \Big[\frac{f(z)}{(z-z_a)^n} \Big] &= \partial_z^{(n-1)} \Big[\frac{1}{(1-e^{\beta z})(z-z_a)^n} \Big] \\ &= \sum_{k=0}^{n-1} \frac{(-1)^k}{(z-z_a)^{n+k}} \ {}^{(n-1)}C_k \ \frac{(n+k-1)!}{(n-1)!} \Big[\frac{\beta^{n-k-1}}{(e^{-\beta z}-1)^{n-k}} \sum_{m=0}^{n-k-2} \sum_{l=0}^{m+1} (-1)^l \\ &\quad {}^{n-k}C_l \ (m+1-l)^{n-k-1} e^{-\beta(m+1)z} + \frac{1}{1-e^{\beta z}} \delta_{n-k-1,0} \Big] \end{aligned} \tag{A.33}$$

Evaluating the above expression at both the poles $z = z_a$ and z_b, one can compute S_1 as expressed in equation (A.28).

Now let us discuss about evaluating the summation S_2 as given in equation (A.26). With a weighting function $f(z) = \dfrac{e^{\beta z/2}}{1-e^{\beta z}}$, one can replace the above summation with the following contour integral,

$$S_2 = \oint \frac{e^{\beta z/2} dz}{(1-e^{\beta z})(((q-1)m - i\alpha_q + z)(m + i\alpha_q - z))^n} \tag{A.34}$$

Notice that we encounter the same n-th order poles in the contour integral as we had with S_1. The residue computation for evaluating this contour integral needs to evaluate the following term as before,

$$\partial_z^{(n)} f(z) = \partial_z^{(n)} \Big[\frac{e^{\beta z/2}}{1-e^{\beta z}} \Big] = \frac{\beta^n e^{-\beta z/2}}{2^n (e^{-\beta z}-1)^{n+1}} B(n) \tag{A.35}$$

where, $B(n)$ is the Eulerian polynomial of type-B in $e^{-\beta z}$, given by,

$$B(n) = \sum_{m=0}^{n} \sum_{k=0}^{m} (-1)^{m-k} \ {}^{n+1}C_{m-k} (2k+1)^n \ e^{-\beta z m} \tag{A.36}$$

Finally, using equation (A.29), (A.35) and (A.36), one can obtain,

$$\begin{aligned}\partial_z^{(n-1)}\Big[\frac{f(z)}{(z-z_a)^n}\Big] &= \partial_z^{(n-1)}\Big[\frac{e^{\beta z/2}}{(1-e^{\beta z})(z-z_a)^n}\Big] \\ &= \sum_{k=0}^{n-1}\frac{(-1)^k}{(z-z_a)^{n+k}}\ {}^{(n-1)}C_k\ \frac{(n+k-1)!}{(n-1)!}\frac{\beta^{n-k-1}}{2^{n-k-1}(e^{-\beta z}-1)^{n-k}} \\ &\quad \sum_{m=0}^{n-k-1}\sum_{l=0}^{m}(-1)^{m-l}\ {}^{n-k}C_{m-l}\ (2l+1)^{n-k-1}e^{-(2m+1)\beta z/2} \qquad \text{(A.37)}\end{aligned}$$

Now using the above equation one can compute the residue and hence the integral (A.34). This finishes the computation of S_2 as given in equation (A.26).

One finds that S_1 depends only linearly on $x = e^{-m\beta}$ while S_2 depends as $\sqrt{x}$. Further noting that the difference in A and B in (A.22) behaves as $A - B = \mathcal{O}(x^{q-1})$ we find that $f_1 = (A-B)\mathcal{O}(x^{\frac{-q}{2}+1}) = \mathcal{O}(x^{q/2})$ and $f_2 = (A-B)\mathcal{O}(x^{\frac{-q+1}{2}}) = \mathcal{O}(x^{\frac{q}{2}-1})$. Therefore in the scaling limit one can take $A = B = 2^{q-2}\left(1 - x\sum_{j=1}^{q-1}e^{-i\beta\alpha_j}\right)$.

Therefore evaluating $(f_1+f_2)^N \approx F_1 + F_2$ in the scaling limit- where once again F_1 consists of terms with $e^{\pm kz\beta}$ where $k \in \mathbb{Z}$ while F_2 consists of terms $e^{\pm kz\beta/2}$ where $k \in \mathbb{Z}_{\text{odd}}$, $F^{(n)} = S_1\overline{F}_1 + S_2\overline{F}_2$. Here $\overline{F}_{1,2} = F_{1,2}(z=0)$[2].

The fact that only these two type of summations contribute for any integer value of k, makes it easier to evaluate equation (A.21) in the scaling limit as,

$$I^{(2n)} = J^{2n}\sum_{k=0}^{n-2}\frac{2^{(q-1)n}x\beta^{n-k}}{(mq)^{n+k}\Gamma(n)^2}(2^n - 2^{2+k}n)({}^{n-1}C_k)\Gamma(n+k)\prod_{m=1}^{q-1}\rho_m^1 \qquad \text{(A.38)}$$

which can be re-written as

$$I^{(2n)} = \left(\frac{J^2\beta}{mq}\right)^n\sum_{k=0}^{n-2}\left(\frac{J^2}{m^2}\right)^k\left(\frac{m}{qJ^2\beta}\right)^k\frac{2^{(q-1)n}x}{\Gamma(n)^2}(2^n - 2^{2+k}n)({}^{n-1}C_k)\Gamma(n+k)\prod_{m=1}^{q-1}\rho_m^1 + \mathcal{O}(\beta) \qquad \text{(A.39)}$$

or equivalently keep all orders in β as

$$I^{(2n)} = \frac{J^{2n}}{m^{2n}q^n}\frac{(-1)^{q(n-1)}}{(n-1)!}\Bigg[\quad \frac{(2n-2)!}{(n-1)!\ q^{n-1}}m\beta\Big(1 - n\Big(q-2n+3\Big)\Big) + \sum_{k=0}^{n-2}\frac{(n+k-1)!}{k!(n-k-1)!\ q^k}(1-2^{k+2-n}n)(m\beta)^{n-k}\ \Bigg]\ x\left(\prod_{m=1}^{q-1}\rho_m^1\right). \qquad \text{(A.40)}$$

This multiplied with (A.12)$\times N^{q-1}$ gives contribution of a circle diagram with n melons.

[2]Since their z dependences were where taken into account in evaluating S_1 and S_2

A.2.3 Evaluating the subleading correction

We end this appendix by presenting a technical result which we do not use in the main text of the paper, but record here anyway, just in case this result finds application subsequent work.

The technical result we report here is the evaluation of the Feynman integral for diagram Fig. (A.2) (the figure is drawn for $q = 4$ but we present the evaluation in general), which is one of the diagrams that would contribute to the generalization of the results presented in this chapter to subleading orders in $\frac{1}{\beta}$. We present the result for the Feynman diagram ignoring the symmetry factor (which can easily be independently evaluated). We evaluate the diagram of Fig. (A.2) as follows. In order to get the integrand of the diagram we first multiply together all the propagators that make it up, keeping careful track of holonomy factors and making use of the fact that holonomies at any interaction vertex sum to zero. The integrand is the term in the big square bracket in (A.41) with ϵ_1 and ϵ_2 temporarily set to zero. The first two lines on the RHS of (A.41) are the $n-2$ factors on the in the diagram Fig.(A.3). [3]

The next four lines on the RHS of (A.41) represent the second factor in Fig.(A.3). Lines 3-6 on the RHS of (A.41) are the remaining factors (the propagators outside the square bracket) in Fig.(A.3). [4]

[3] t_1 in this term is the length of the straight line in these factors, while t_2 is the length of the 3 (or more generally $q-1$) melonic lines in the part Fig.(A.3) that is enclosed in the square bracket. Really there are $n-1$ different t_1s and $n-2$ different t_2. As t_1 and t_2 are dummy variables that we integrate over, we have used the same symbol for all of them.

[4] The third line in (A.41) is the straight line in this part of Fig.(A.3). The last and second-last lines in (A.41) are, respectively, the blobs of $q-1$ and $q-2$ propagators in this part of Fig.(A.3). Finally the fourth line in (A.41) is the product of the two propagators that run between the '$q-1$ blob' and the '$q-2$ blob'. The times in all these terms represent the lengths of the corresponding propagators.

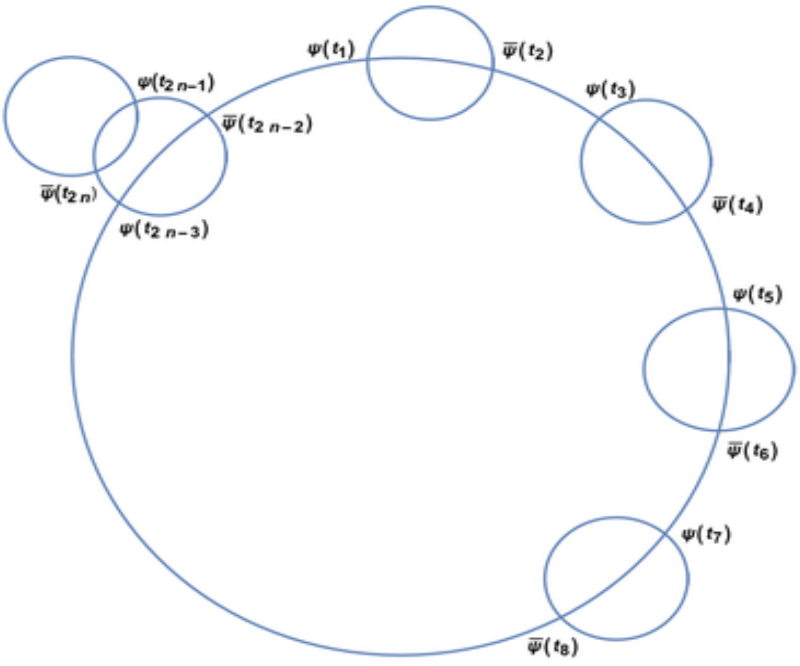


Figure A.2: subleading diagram

$$
\begin{aligned}
I^{(2n-2)} \;=\; \left(\tfrac{1}{2\pi\beta}\right)^2 \sum_{\omega_1,\omega_2=-\infty}^{\infty} \Bigg[& \left(\int_{-\beta/2}^{\beta/2} dt_1 e^{-(m+i\alpha_1+i\frac{\epsilon_1}{\beta})t_1} \left(\mathrm{sgn}(t_1) + \tanh(\tfrac{m\beta+i\alpha_1\beta}{2})\right) \right. \\
& \left. \int_{-\beta/2}^{\beta/2} dt_2 e^{((q-1)m-i\alpha_1-i\frac{\epsilon_1}{\beta})t_2} \left(\mathrm{sgn}(t_2)A_1 - B_1\right) \right)^{n-2} \\
& \int_{-\beta/2}^{\beta/2} dt_1 e^{-(m+i\alpha_1+i\frac{\epsilon_1}{\beta})t_1} \left(\mathrm{sgn}(t_1) + \tanh(\tfrac{m\beta+i\alpha_1\beta}{2})\right) \\
& \left(\int_{-\beta/2}^{\beta/2} dt_3 e^{(m+i\alpha_2+i\frac{\epsilon_2}{\beta})t_3} \left(\mathrm{sgn}(t_3) - \tanh(\tfrac{m\beta+i\alpha_2\beta}{2})\right) \right)^2 \\
& \int_{-\beta/2}^{\beta/2} dt_4 e^{((q-2)m-i(\alpha_1+\alpha_2)-i\frac{(\epsilon_1+\epsilon_2)}{\beta})t_4} \left(\mathrm{sgn}(t_4)A_{1,2} - B_{1,2}\right) \\
& \int_{-\beta/2}^{\beta/2} dt_5 e^{(-(q-1)m+i\alpha_2+i\frac{\epsilon_2}{\beta})t_5} \left(\mathrm{sgn}(t_5)A_2 + B_2\right) \Bigg] \qquad \text{(A.41)}
\end{aligned}
$$

After evaluating the integrand we need to perform the integrals. Roughly speaking we must integrate all propagator lengths in the integrand above from $-\dfrac{\beta}{2}$ to $\dfrac{\beta}{2}$. However we

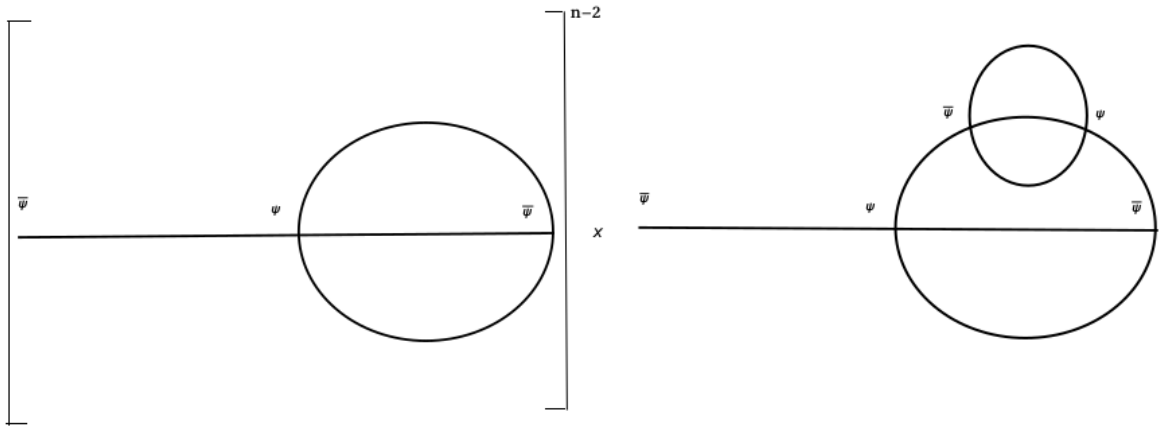


Figure A.3: parts of subleading diagram

need to do this subject to the constraint that as we go round either of the two circles in the diagram Fig.(A.2) we come back to the same time as we started out, modulo β. This is where the parameters ϵ_1 and ϵ_2 in (A.41) come in. ϵ_1 couples to the sum of lengths of propagators in units of β around the big circle in Fig.(A.2), while ϵ_2 multiplies the sum of the lengths of all the propagators as we go around the small circle - again in units of β in Fig.(A.2). The constraint that these lengths evaluate to an integral multiple of β can then be implemented by setting $\epsilon_{1,2} = 2\pi\omega_{1,2}$ and then summing ω_i over all integral values, as we have done in (A.41).

In order to proceed we perform the time integrals in an unconstrained manner. The result can be rearranged (according to its ω_i dependence) as a sum of four types of terms.

1. Terms containing $e^{k(z_1+z_2)\beta}$ where $k \in \mathbb{Z}$
2. those with $e^{kz_1\beta/2}$ where $k \in \mathbb{Z}_{\rm odd}$
3. with $e^{kz_2\beta/2}$ where $k \in \mathbb{Z}_{\rm odd}$
4. and $e^{k(z_1+z_2)\beta/2}$ where $k \in \mathbb{Z}_{\rm odd}$;

where $z_i = -\dfrac{2\pi i}{\beta}\omega_i$.

We deal with these four classes of terms separately; for each class we explicitly perform the sum over ω_i (by reducing it to a contour integral as in the previous subsection) and expand the resultant expression in a Taylor series in x (again as in the previous subsection), keep only the terms that are linear in x. Combining together the results from each of the four classes we obtain our final result

$$I^{(2n-2)} \;=\; -\left(\frac{J^2\beta}{mq}\right)^n \sum_{k=0}^{n-4} \frac{x(q-1)}{(mq\beta)^{k+1}} \frac{2^{(q-1)n}(2^n+(n-1)2^{3+k})(2n+k-2)\Gamma(n+k-1)}{\Gamma(n-k-1)\Gamma(n)\Gamma(1+k)} \prod_{m=1}^{q-1} \rho_m^1$$

$$+\mathcal{O}(\beta^2) \qquad \text{(A.42)}$$

(the terms $\mathcal{O}(\beta^2)$ that we have omitted to list in (A.42) are the terms with $k = n-3$ and $k = n-2$ which exist in the final answer but the values of whose coefficients do not follow the uniform rule of the other terms).

Note that (A.42) scales like $\frac{1}{\beta}$ in coordinated large β small J limit in which $J^2\beta$ is held fixed.

A.3 The holonomy effective action from the sigma model

In this section we ask the following question: what is the contribution to $S_{\text{eff}}(U)$ - the effective action for holonomies - resulting from integrating out the new light degrees of freedom discovered in the massless tensor model in early sections in this chapter? In the bulk of this section we address this question at the technical level. At the end of the section we turn to a quick discussion of its physical import.

Turning on holonomy is equivalent to putting appropriate boundary condition on fermion fields. This translates into boundary condition on V_l, given by $V_l(-\frac{\beta}{2}) = UV_l(+\frac{\beta}{2})$, $U \in O(N)$. This boundary condition is equivalent to the computation of the partition function

$$Z = e^{-S_{\text{eff}}(U)} = \text{Tr} e^{-\beta H}\widehat{U}, \qquad \text{(A.43)}$$

where H is the Hamiltonian of the quantum mechanical system (2.3) and U is the quantum mechanical operator that implements left rotations on the sigma model by the $O(N)^{q-1}$ group rotation U. The partition function (2.2) is the product of $q-1$ factors, associated with the sigma models on the $q-1$ gauge groups. It follows that the effective action $S_{\text{eff}}(U)$ that follows from this computation takes the form

$$S_{\text{eff}}(U) = \sum_i S(U_i). \qquad \text{(A.44)}$$

In the rest of this section we compute the functions $S(U_i)$

Let us first note that the Hilbert H space on which any one of the factors of $q-1$ distinct factors the sigma model (2.3) acts is given as follows. The Hamiltonian acts on the Hilbert space H

$$H = \sum_{R_i} \widetilde{R}_i \otimes \widetilde{R}_i. \qquad \text{(A.45)}$$

The sum R_i runs over all genuine (as opposed to spinorial) representations of $O(N)$. $\widetilde{R}_i$ denotes the vector space on which $O(N)$ acts in the i^{th} representation. The space $\widetilde{R}_i \otimes \widetilde{R}_i$ transforms in the representation $R_i \times R_i$ under $O(N)_L \times O(N)_R$; the operator $\widehat{U}$ acts as an $O(N)$ rotation on the first $\widetilde{R}_i$ but as identity on the second $\widetilde{R}_i$. The Hamiltonian corresponding to action (2.3) is diagonal under the decomposition (A.45); the energy of the i^{th} factor of the Hilbert space is $\frac{JC_2(R_i)}{2\mathcal{A}N^{q-2}}$.

Representations of $O(N)$ are conveniently labeled by the highest weights $(h_1, h_2, h_3 \ldots)$, the charges under rotations in mutually orthogonal two planes. Let $h = \sum_i h_i$. At leading order in the large N limit the dimensionality of the representation R_i depends only on h and is given by

$$d(R_i) = \frac{N^h}{h!}.$$

Moreover the Casimir $C_2(R_i)$ of representations of $O(N)$ also depends only on h at leading order in the large N limit and is given by

$$C_2(R_i) = Nh.$$

Let $\chi_{R_i}(U)$ denote the character in the R_i representation of $O(N)$ and let

$$\chi_n(U) = \sum_{R_i \in \widehat{n}} \chi_{R_i}(U), \tag{A.46}$$

where $\widehat{n}$ denotes the collection of all representations of $O(N)$ with $h = n$. In other words $\chi_n(U)$ is the sum over the characters of all representations with $h = n$.

Note that all representations with $h = n$ can be constructed - and can be constructed exactly once - from the direct products of n vectors of $O(N)$ (this is true when $N \gg n$ as we assume). [5] Let P_n denote the projector onto representations with $h = n$

$$P_n\left[f(U))\right] = \int dU' \sum_{R_i \in \widehat{n}} \chi_{R_i}(U)\chi^*_{R_i}(U')f(U'). \tag{A.47}$$

It follows that

$$\chi_n(U) = P_n\left[(\mathrm{Tr}U)^n\right]. \tag{A.48}$$

where U on the RHS of (A.48) represents the group element in the vector representation of $O(N)$.

Finally we define

$$z = e^{-\frac{J}{2AN^{q-3}}}. \tag{A.49}$$

It follows immediately from all the facts and definitions presented above that

$$e^{-S(U_i)} = \sum_{n=0}^{\infty} \frac{(zN)^n}{n!}\chi_n(U_i). \tag{A.50}$$

Using (A.48), (A.50) can be rewritten in the (perhaps deceptively) elegant form

$$e^{-S(U_i)} = P_{z\partial_z} e^{Nz\mathrm{Tr}(U_i)}. \tag{A.51}$$

[5] Note, however, that not every representation of n vectors has $h = n$; the product space includes representations (formed by contracting 2 vector indices) with $h = n - 2$, and representations (formed by contracting 4 vector indices) with $h = n - 4 \ldots$.

Note that

$$\int dU e^{-S(U)} = 1. \tag{A.52}$$

This is an immediate consequence of the fact that the vacuum is the only representation in the spectrum of the group sigma model that is a singlet under $O(N)_L$. It follows that the partition function generated by $S(U)$ by itself is trivial. However $S(U)$ is only one piece of the effective action for U in the massless tensor model (2.1); we get other contributions to the effective action by integrating out the fermionic fields themselves (as was explicitly done earlier in this chapter for the case of massive fermions). When put together with other contributions the effective action (A.51) could have a significant impact on the partition function, especially at temperatures scaled to ensure that the matter contribution to the effective action - like the contribution of the sigma model considered in this section - is of order N^2.

Appendix B

Appendices for Chapter 2

B.1 Previously known results for the large N free energy

Below we will encounter several equations that involve the quantities

$$\begin{aligned}\mathcal{C} =&\frac{1}{2}\int_{-\pi}^{\pi} d\alpha\, \rho_F(\alpha)\left(\log(2\cosh(\tfrac{\widehat{c}_F+i\alpha}{2}))+\log(2\cosh(\tfrac{\widehat{c}_F-i\alpha}{2}))\right),\\ \mathcal{S} =&\frac{1}{2}\int_{-\pi}^{\pi} d\alpha\, \rho_B(\alpha)\left(\log(2\sinh(\tfrac{\widehat{c}_B+i\alpha}{2}))+\log(2\sinh(\tfrac{\widehat{c}_B-i\alpha}{2}))\right),\end{aligned} \tag{B.1}$$

where $\widehat{c}_B$ and $\widehat{c}_F$ are (dimensionless versions of) the thermal masses in the boson and fermion theory respectively.

Using (3.15), it is not difficult to verify the following identities:

$$\lambda_B\mathcal{S} = -\frac{\text{sgn}(\lambda_F)}{2}c_F + \lambda_F\mathcal{C}\ , \quad \lambda_F\mathcal{C} = -\frac{\text{sgn}(\lambda_B)}{2}c_B + \lambda_B\mathcal{S}\ . \tag{B.2}$$

B.1.1 Results for the critical fermion theory

In [56] the 'fixed holonomy' $\mathbb{R}^2\times S^1$ partition function - $v_F[\rho_F]$ - of the fermionic theory has been evaluated in both fermionic phases. The final result of this calculation is most conveniently given in terms of an auxiliary off-shell free energy

$$\begin{aligned}&F_F(c_F,\zeta_F,\widetilde{\mathcal{C}})\\ &=\frac{N_F}{6\pi}\Bigg[\widehat{c}_F^3-2\lambda_F^2\widetilde{\mathcal{C}}^3-\frac{3}{2}\left(\widehat{c}_F^2-\frac{16\pi^2}{\kappa_F^2}\widehat{\zeta}_F^2\right)\widetilde{\mathcal{C}}+\frac{6\pi\widehat{y}_2^2}{\kappa_F\lambda_F}\widehat{\zeta}_F-\frac{24\pi^2\widehat{y}_4}{\kappa_F^2\lambda_F}\widehat{\zeta}_F^2+\frac{24\pi^3x_6^F}{\kappa_F^3\lambda_F}\widehat{\zeta}_F^3\\ &\qquad-\frac{3}{2}\widetilde{\mathcal{C}}\left(\widehat{c}_F^2-\left(2\lambda_F\widetilde{\mathcal{C}}-\frac{4\pi}{\kappa_F}\widehat{\zeta}_F\right)^2\right)\\ &\qquad-3\int_{-\pi}^{\pi}d\alpha\rho_F(\alpha)\int_{\widehat{c}_F}^{\infty}dy\ y\ \left(\log\left(1+e^{-y-i\alpha}\right)+\log\left(1+e^{-y+i\alpha}\right)\right)\Bigg].\end{aligned} \tag{B.3}$$

The auxiliary off-shell free energy (B.3) is a function of three variables - c_F, ζ_F and $\widetilde{\mathcal{C}}$ - in addition to the temperature and the holonomies. The free energy $v_F[\rho_F]$ defined in (3.12) is obtained from $F_F[\rho_F]$ in (B.3) by extremizing the latter quantity w.r.t. these three 'dynamical' variables. Extremizing the free energy (B.3) w.r.t. the variable $\widetilde{\mathcal{C}}$ yields the equation of motion

$$\widehat{c}_F^2 = \left(2\lambda_F\widetilde{\mathcal{C}} - \frac{4\pi}{\kappa_F}\widehat{\zeta}_F\right)^2 \ . \tag{B.4}$$

Varying w.r.t. c_F yields

$$\widetilde{\mathcal{C}} = \mathcal{C} \ , \tag{B.5}$$

with $\mathcal{C}$ given in (B.1), while the stationarity of variation w.r.t ζ_F yields

$$-\frac{3}{4}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^2 x_6^F - \frac{16\pi\widehat{\zeta}_F}{\kappa_F}\lambda_F\widetilde{\mathcal{C}} + \frac{8\pi\widehat{\zeta}_F}{\kappa_F}\widehat{y}_4 + 4\lambda_F^2\widetilde{\mathcal{C}}^2 - \widehat{y}_2^2 = 0 \ . \tag{B.6}$$

Plugging (B.5) into (B.4) and (B.6) respectively yields the simplified gap equations

$$\widehat{c}_F^2 = \left(2\lambda_F\mathcal{C} - \frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^2 \ , \tag{B.7}$$

and

$$-\frac{3}{4}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^2 x_6^F - \frac{16\pi\widehat{\zeta}_F}{\kappa_F}\lambda_F\mathcal{C} + \frac{8\pi\widehat{\zeta}_F}{\kappa_F}\widehat{y}_4 + 4\lambda_F^2\mathcal{C}^2 - \widehat{y}_2^2 = 0 \ , \tag{B.8}$$

for the quantities c_F and ζ_F.

Below we will find it useful to work with a reduced off-shell free energy, obtained by integrating $\widetilde{\mathcal{C}}$ out of (B.3). In order to do this we note that (B.4) has two solutions

$$2\lambda_F\widetilde{\mathcal{C}} = \frac{4\pi\widehat{\zeta}_F}{\kappa_F} \pm \widehat{c}_F \ , \tag{B.9}$$

The undetermined sign in (B.9) is completely free. Clearly this sign is (tautologically) given by $\mathrm{sgn}(\widetilde{X}_F)$ where

$$\widetilde{X}_F \equiv 2\lambda_F\widetilde{\mathcal{C}} - \frac{4\pi\widehat{\zeta}_F}{\kappa_F} \ . \tag{B.10}$$

It follows that (B.9) may formally be rewritten as

$$2\lambda_F\widetilde{\mathcal{C}} = \frac{4\pi\widehat{\zeta}_F}{\kappa_F} + \mathrm{sgn}(\widetilde{X}_F)\widehat{c}_F \ . \tag{B.11}$$

The reduced free energy - which is now a function only of two variables c_F and ζ_F - is obtained by plugging either of these two solutions into (B.3). Note that when we do this

the second line of (B.3) vanishes as a consequence of (B.4).[1] Inserting the solutions (B.11) into the free energy (B.3), we have the following explicit expression for the reduced off-shell free energy as a function of c_F and ζ_F

$$F_F(c_F,\zeta_F) = \frac{N_F}{6\pi}\Bigg[\frac{|\lambda_F| - \mathrm{sgn}(\lambda_F)\mathrm{sgn}(\widetilde{X}_F)}{|\lambda_F|}\widehat{c}_F^3 - \frac{3}{2\lambda_F}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)(\widehat{c}_F^2 - \widehat{y}_2^2) \\ - \frac{3\widehat{y}_4}{2\lambda_F}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^2 + \frac{(3x_6^F+4)}{8\lambda_F}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^3 \\ - 3\int_{-\pi}^{\pi} d\alpha\,\rho_F(\alpha)\int_{\widehat{c}_F}^{\infty} dy\; y\;\left(\log\left(1+e^{-y-i\alpha}\right)+\log\left(1+e^{-y+i\alpha}\right)\right)\Bigg]. \tag{B.12}$$

Note that the above reduced off-shell free energy function has two branches depending on the sign $\mathrm{sgn}(\widetilde{X}_F)$. We refer to the branch in which $\mathrm{sgn}(\widetilde{X}_F)\mathrm{sgn}(\lambda_F) > 0$ as the unHiggsed branch, and the branch in which $\mathrm{sgn}(\widetilde{X}_F)\mathrm{sgn}(\lambda_F) < 0$ as the Higgsed branch. It is easily verified that the variation of (B.12) with respect to $\widehat{c}_F$ yields (B.7) while the variation of (B.12) w.r.t ζ_F yields (B.8).

Note that that the gap equations (B.4), (B.5) (equivalently (B.7)) and (B.6) - unlike the free energy (B.12) - have no explicit dependence on $\mathrm{sgn}(\widetilde{X}_F)$. Nonetheless the same gap equations hold for both 'phases' of the theory, i.e. for both choices of $\mathrm{sgn}(\widetilde{X}_F\lambda_F)$.[2] It follows from this observation that the solutions to the finite temperature gap equations of the fermionic theory vary analytically as we pass from one 'phase' to another. In fact more is true; the finite temperature free energy of the fermionic theory is itself analytic as one passes from the unHiggsed to the Higgsed phase. At the physical level, the sharp zero temperature distinction between the Higgsed and unHiggsed phases gets blurred out by finite temperature effects.

[1] This reduced form of the off-shell free energy - rather than the fully off-shell free energy (B.3) - was presented in [56]. The off-shell free energy (B.3) is a new formula that has not previously been presented in the literature.

[2] The fact that $\mathrm{sgn}(\widetilde{X}_F)$ disappears from the gap equations is obvious when we obtain the equations from (B.3), even though this fact might appear mysterious when derived starting from (B.12).

B.1.2 Results for regular bosons in the unHiggsed Phase

The off-shell free energy for the RB theory was computed in [56] and is given by

$$F_B(c_B,\widetilde{\mathcal{S}}) = \frac{N_B}{6\pi}\Bigg[-\widehat{c}_B^{\,3} + 2\left(\widehat{c}_B^{\,2}-\widehat{m}_B^2\right)\widetilde{\mathcal{S}} + 2\lambda_B\widehat{b}_4\widetilde{\mathcal{S}}^2 \\ + \widetilde{\mathcal{S}}\left(\widehat{c}_B^{\,2}-\widehat{m}_B^2-(4+3x_6^B)\lambda_B^2\widetilde{\mathcal{S}}^2+4\lambda_B\widehat{b}_4\widetilde{\mathcal{S}}\right) \\ + 3\int_{-\pi}^{\pi}d\alpha\rho_B(\alpha)\int_{\widehat{c}_B}^{\infty}dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right)+\log\left(1-e^{-y+i\alpha}\right)\right)\Bigg], \tag{B.13}$$

or equivalently by

$$F_B(c_B,\widetilde{\mathcal{S}}) = \frac{N_B}{6\pi}\Bigg[-\widehat{c}_B^{\,3} + 3\widetilde{\mathcal{S}}\left(\widehat{c}_B^{\,2}-\widehat{m}_B^2\right) + 6\widehat{b}_4\lambda_B\widetilde{\mathcal{S}}^2 - (4+3x_6^B)\lambda_B^2\widetilde{\mathcal{S}}^3 \\ + 3\int_{-\pi}^{\pi}d\alpha\rho_B(\alpha)\int_{\widehat{c}_B}^{\infty}dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right)+\log\left(1-e^{-y+i\alpha}\right)\right)\Bigg]. \tag{B.14}$$

As in the previous subsection, the free energy $v_B[\rho_B]$ defined in (3.12) is obtained by extremizing the action (B.13) w.r.t. the dynamical variables $\widetilde{\mathcal{S}}$ and c_B. The equations of motion that follow by varying (B.13) w.r.t $\widetilde{\mathcal{S}}$ and c_B respectively are

$$\widehat{c}_B^{\,2} = (4+3x_6^B)\lambda_B^2\widetilde{\mathcal{S}}^2 - 4\lambda_B\widehat{b}_4\widetilde{\mathcal{S}} + \widehat{m}_B^2\ , \tag{B.15}$$

and

$$\widetilde{\mathcal{S}} = \mathcal{S}\ , \tag{B.16}$$

where $\mathcal{S}$ was defined in (B.1). Inserting (B.16) into (B.15) yields the gap equation for the single variable c_B

$$\widehat{c}_B^{\,2} = (4+3x_6^B)\lambda_B^2\mathcal{S}^2 - 4\lambda_B\widehat{b}_4\mathcal{S} + \widehat{m}_B^2\ . \tag{B.17}$$

As in the previous subsection it is possible to obtain a reduced free energy by integrating $\widetilde{\mathcal{S}}$ out of (B.13). This may be achieved by using (B.15) to solve for $\widetilde{\mathcal{S}}$ as a function of c_B and plugging this solution into (B.13). [3] This is the form in which the off-shell free energy for the scalar theory in the unHiggsed phase was presented in [56].

[3] Note that the second line of (B.13) vanishes when we do this as a this line is proportional to (B.15).

B.1.3 Duality in the unHiggsed phase and a prediction for the Higgsed phase

We first list the duality map between various quantities in the RB and CF theories. Recall that the 't Hooft parameters λ_B and λ_F were defined as

$$\lambda_B = \frac{N_B}{\kappa_B} \ , \quad \lambda_F = \frac{N_F}{\kappa_F} \ . \tag{B.18}$$

The duality maps different parameters as follows:

$$\begin{aligned} &N_F = |\kappa_B| - N_B \ , \quad \kappa_F = -\kappa_B \ , \quad \lambda_F = \lambda_B - \mathrm{sgn}(\lambda_B) \ , \\ &x_6^F = x_6^B \ , \quad y_4 = b_4 \ , \quad y_2^2 = m_B^2 \ , \quad |\lambda_B|\rho_B(\alpha) + |\lambda_F|\rho_F(\pi-\alpha) = \frac{1}{2\pi} \ . \end{aligned} \tag{B.19}$$

The last relation gives rise to

$$N_F\rho_F(\alpha) = \frac{|\kappa_B|}{2\pi} - N_B\rho_B(\pi-\alpha) \ . \tag{B.20}$$

Consider the off-shell free energies for the critical fermion theory in terms of the two 'fields' c_F and ζ_F given in (B.12):

$$\begin{aligned} F_F(c_F,\zeta_F) = \frac{N_F}{6\pi}\Bigg[&\frac{\lambda_F - \mathrm{sgn}(\widetilde{X}_F)}{\lambda_F}\widehat{c}_F^3 - \frac{3}{2\lambda_F}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)(\widehat{c}_F^2 - \widehat{y}_2^2) \\ &- \frac{3\widehat{y}_4}{2\lambda_F}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^2 + \frac{(3x_6^F+4)}{8\lambda_F}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^3 + \\ &- 3\int_{-\pi}^{\pi} d\alpha\rho_F(\alpha)\int_{\widehat{c}_B}^{\infty} dy\ y\ \left(\log\left(1+e^{-y-i\alpha}\right) + \log\left(1+e^{-y+i\alpha}\right)\right)\Bigg]. \end{aligned} \tag{B.21}$$

Using the relation (B.20), the last line in (B.21) can be rewritten as [4]

$$\begin{aligned} &-\frac{3N_F}{6\pi}\int_{-\pi}^{\pi} d\alpha\rho_F(\alpha)\int_{\widehat{c}_F}^{\infty} dy\ y\ \left(\log\left(1+e^{-y-i\alpha}\right) + \log\left(1+e^{-y+i\alpha}\right)\right) = \\ &\qquad = \frac{3N_B}{6\pi}\int_{-\pi}^{\pi} d\alpha\rho_B(\alpha)\int_{\widehat{c}_F}^{\infty} dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right) + \log\left(1-e^{-y+i\alpha}\right)\right) \ . \end{aligned} \tag{B.22}$$

[4] The two terms in the RHS of (B.20) simplify as follows. The integral over α in the term proportional to $|\kappa_B|$ can be performed by Taylor-expanding the logarithms and gives zero since the integrals are of the form $\int_{-\pi}^{\pi} d\alpha\, e^{in\alpha} = 0$ for non-zero integers n. In the term proportional to $\rho_B(\pi-\alpha)$, we have performed the variable change $\alpha \to \pi - \alpha$ resulting in an additional minus sign in the argument of the logarithms.

Substituting the various fermionic parameters with their bosonic counterparts in (B.19), we get the following expression for the dual of the fermionic off-shell free energy in terms of two 'fields' c_F and ζ_F:

$$\begin{aligned}\mathcal{F}(c_F,\zeta_F) =&\frac{N_B}{6\pi}\Bigg[-\frac{\lambda_B-\text{sgn}(\lambda_B)-\text{sgn}(\widetilde{X}_F)}{\lambda_B}\widehat{c}_F^3+\frac{3}{2\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)(\widehat{c}_F^2-\widehat{m}_B^2)+\\&+\frac{3\widehat{b}_4}{2\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^2-\frac{(3x_6^B+4)}{8\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^3+\\&+3\int_{-\pi}^{\pi}d\alpha\rho_B(\alpha)\int_{\widehat{c}_F}^{\infty}dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right)+\log\left(1-e^{-y+i\alpha}\right)\right)\Bigg].\end{aligned} \tag{B.23}$$

unHiggsed phase: $\text{sgn}(X_F)=-\text{sgn}(\lambda_B)$

In this case $\mathcal{F}(c_F,\zeta_F)$ in (B.23) simplifies to

$$\begin{aligned}\mathcal{F}(c_F,\zeta_F) =& \frac{N_B}{6\pi}\Bigg[-\widehat{c}_F^3+\frac{3}{2\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)(\widehat{c}_F^2-\widehat{m}_B^2)+\\&+\frac{3\widehat{b}_4}{2\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^2-\frac{(3x_6^B+4)}{8\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^3+\\&+3\int_{-\pi}^{\pi}d\alpha\rho_B(\alpha)\int_{\widehat{c}_F}^{\infty}dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right)+\log\left(1-e^{-y+i\alpha}\right)\right)\Bigg].\end{aligned} \tag{B.24}$$

If we now perform the field redefinitions

$$\widehat{c}_F=\widehat{c}_B\ ,\quad \frac{4\pi\widehat{\zeta}_F}{\kappa_F}=2\lambda_B\widetilde{\mathcal{S}}\ . \tag{B.25}$$

we see that the fermionic off-shell free energy reduces exactly to the regular boson off-shell free energy (B.14), establishing the duality of the CF and RB theories in their unHiggsed phases. The matching of off-shell free energies between the two theories automatically guarantees the matching of gap equations, as the latter are obtained by extremizing the off-shell free energies w.r.t. their 'fields'.

Higgsed phase: $\mathrm{sgn}(X_F) = \mathrm{sgn}(\lambda_B)$

In this case $\mathcal{F}(c_F, \zeta_F)$ simplifies to

$$\mathcal{F}(c_F,\zeta_F) = \frac{N_B}{6\pi}\Bigg[-\frac{\lambda_B - 2\mathrm{sgn}(\lambda_B)}{\lambda_B}\widehat{c}_F^3 + \frac{3}{2\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)(\widehat{c}_F^2 - \widehat{m}_B^2) + \\ + \frac{3\widehat{b}_4}{2\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^2 - \frac{(3x_6^B+4)}{8\lambda_B}\left(\frac{4\pi\widehat{\zeta}_F}{\kappa_F}\right)^3 + \\ + 3\int_{-\pi}^{\pi} d\alpha\rho_B(\alpha)\int_{\widehat{c}_F}^{\infty} dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right) + \log\left(1-e^{-y+i\alpha}\right)\right)\Bigg]. \tag{B.26}$$

If we now make the field redefinitions

$$\widehat{c}_F = \widehat{c}_B\ , \quad \frac{4\pi\widehat{\zeta}_F}{\kappa_F} = -2\lambda_B\widehat{\sigma}_B\ , \tag{B.27}$$

we find that (B.26) reduces to

$$F_B(c_B,\sigma_B) = \frac{N_B}{6\pi}\Bigg[-\frac{\lambda_B - 2\mathrm{sgn}(\lambda_B)}{\lambda_B}\widehat{c}_B^3 - 3\widehat{\sigma}_B(\widehat{c}_F^2 - \widehat{m}_B^2) + 6\widehat{b}_4\lambda_B\widehat{\sigma}_B^2 + (3x_6^B+4)\lambda_B^2\widehat{\sigma}_B^3 \\ + 3\int_{-\pi}^{\pi} d\alpha\rho_B(\alpha)\int_{\widehat{c}_B}^{\infty} dy\ y\ \left(\log\left(1-e^{-y-i\alpha}\right) + \log\left(1-e^{-y+i\alpha}\right)\right)\Bigg]. \tag{B.28}$$

(B.28) may be regarded as the prediction of duality for the off-shell free energy of the RB theory in the Higgsed phase.

B.2 The tadpole from W boson loops

The exact all-orders propagator $G_{\mu\nu}(q)$ in (3.44) is the saddle point value in the large N_B limit of a gauge-singlet field $\alpha_{\mu\nu}(q)$ that appears in [54] as one of two gauge-singlet fields $\alpha_{\mu\nu}$ and $\Sigma^{\mu\nu}$ that describe the effective dynamics of the W boson:

$$G_{\mu\nu}(q) = -\frac{1}{\lambda_B}\alpha_{\mu\nu}(q)\ . \tag{B.29}$$

The first term of the gap equation (3.45) is then given by the tadpole contribution

$$\lambda_B N_B T(\sigma_B) = \frac{\lambda_B N_B}{2\pi}\int\frac{\mathcal{D}^3q}{(2\pi)^3}g^{\mu\rho}G(q)_{\rho\mu} = -\frac{N_B}{2\pi}\int\frac{\mathcal{D}^3q}{(2\pi)^3}g^{\mu\rho}\alpha(q)_{\rho\mu}\ . \tag{B.30}$$

For brevity, we work with the parameter $m = \lambda_B \sigma_B$ in what follows. The field $\alpha_{\mu\nu}(q)$ is given in terms of four known functions F_1, F_2, F_3 and F_4 of $w = 2q_+q_-$ and the all-loop exact kernel $Q(q) = G^{-1}(q)$.

$$\begin{aligned}
\alpha_{++}(q) &= \frac{\lambda_B}{(2\pi)^2 \det Q} \frac{1}{q_-^2} \left(imF_1 + (F_3 + \tfrac{w}{2})^2 \right) \ , \\
\alpha_{-+}(q) &= \frac{\lambda_B}{(2\pi)^2 \det Q} \left((1 - F_4)(F_3 + \tfrac{w}{2}) - im(F_2 + im - q_3) \right) = \alpha_{+-}(-q) \ , \\
\alpha_{--}(q) &= \frac{\lambda_B}{(2\pi)^2 \det Q} q_-^2 (1 - F_4)^2 \ , \\
\alpha_{-3}(q) &= -\frac{\lambda_B}{(2\pi)^2 \det Q} q_- (1 - F_4)(F_2 + im - q_3) = \alpha_{3-}(-q) \ , \\
\alpha_{3+}(q) &= -\frac{\lambda_B}{(2\pi)^2 \det Q} \frac{1}{q_-} \left(F_1(1 - F_4) + (F_2 + im - q_3)(F_3 + \tfrac{w}{2}) \right) = \alpha_{+3}(-q) \ , \\
\alpha_{33}(q) &= -\frac{\lambda_B}{(2\pi)^2 \det Q} \left((F_2 + im)^2 - q_3^2 \right) \ .
\end{aligned} \tag{B.31}$$

We give explicit expressions for the functions $F_{1\ldots4}$:

$$\begin{aligned}
&F_2(w) = im(g(w) - 1) \ , \quad F_4(w) = 1 - \frac{1}{g(w)} \ , \\
&F_3(w) = -\frac{w}{2} + \frac{1}{g(w)} \left(\frac{1}{2} \mathcal{I}(w) - \frac{m^2}{3} (g(w)^3 - g(0)^3) \right) \ , \\
&F_1(w) = img(w) \left(c_B^2 (g(w) - g(0)) - \frac{m^2}{3} (g(w)^3 - g(0)^3) + wg(w) - \mathcal{I}(w) \right) \ ,
\end{aligned} \tag{B.32}$$

where the functions $g(w)$ and $\mathcal{I}(w)$ are given by

$$g(w) = 1 + \lambda_B \xi(w) \ , \quad \mathcal{I}(w) = \int_0^w dz \, g(z) = w + \lambda_B \mathcal{I}_\xi(w) \ , \tag{B.33}$$

and the function $\xi(w)$ in the definition of $g(w)$ above is given by

$$\begin{aligned}
\xi(w) = \frac{1}{2m\beta} \int_{-\pi}^{\pi} d\alpha \, \rho_B(\alpha) \Bigg[&\log 2 \sinh \left(\tfrac{\beta}{2} \sqrt{w + c_B^2} + \tfrac{i}{2} \alpha \right) \\
&+ \log 2 \sinh \left(\tfrac{\beta}{2} \sqrt{w + c_B^2} - \tfrac{i}{2} \alpha \right) \Bigg] \ .
\end{aligned} \tag{B.34}$$

We recognize the quantity $\mathcal{S}$ defined in (B.1) to be the value of $\beta m \, \xi(w)$ at $w = 0$:

$$\mathcal{S} = \beta m \, \xi(0) = \frac{1}{2} \int_{-\pi}^{\pi} d\alpha \, \rho_B(\alpha) \left[\log 2 \sinh \left(\tfrac{\widehat{c}_B + i\alpha}{2} \right) + \log 2 \sinh \left(\tfrac{\widehat{c}_B - i\alpha}{2} \right) \right] \ . \tag{B.35}$$

In the above expressions, the constant c_B is the pole mass of the W boson which occurs in the determinant of the all-loop kernel $Q(q) = G^{-1}(q)$:

$$\det Q(q) = -\frac{m}{(2\pi)^3}(q^2 + c_B^2) \ , \tag{B.36}$$

and is given in terms of the parameter m and the function $\xi(w)$ above by

$$c_B^2 = m^2(1 + \lambda_B \xi(0))^2 \quad \text{or equivalently} \quad \beta^2 c_B^2 = \widehat{c}_B^2 = (\widehat{m} + \lambda_B \mathcal{S})^2 \ . \tag{B.37}$$

Substituting the expressions (B.31), (B.32) for $\alpha_{\mu\nu}(q)$ in (B.30), we have

$$\begin{aligned} T(m) &= -\frac{1}{2\pi\lambda_B}\int \frac{\mathcal{D}^3 q}{(2\pi)^3}(\alpha_{+-}(q) + \alpha_{-+}(q) + \alpha_{33}(q)) \ , \\ &= \frac{1}{m}\int \frac{\mathcal{D}^3 q}{(2\pi)^3}\frac{1}{q^2 + c_B^2}\left(\frac{\mathcal{I}(w)}{g(w)^2} + \frac{4m^2}{3}g(w) + m^2 g(w)^2 + \frac{2m^2}{3}\frac{g(0)^3}{g(w)^2} + q_3^2\right) \ , \\ &= \frac{1}{m}\int \frac{\mathcal{D}^3 q}{(2\pi)^3}\frac{1}{q^2 + c_B^2}\left(\mathcal{L}(w) - (w + c_B^2)\right) + \frac{1}{m}\int \frac{\mathcal{D}^3 q}{(2\pi)^3} \ , \end{aligned} \tag{B.38}$$

where we have added and subtracted the term $w + c_B^2$ inside the integrand to complete the quantity q_3^2 to $q^2 + c_B^2$. The quantity $\mathcal{L}(w)$ is given by

$$\mathcal{L}(w) = \frac{\mathcal{I}(w)}{g(w)^2} + \frac{4}{3}m^2 g(w) + m^2 g(w)^2 + \frac{2}{3}m^2\frac{g(0)^3}{g(w)^2} \ . \tag{B.39}$$

The discrete sum over q_3 in the first term is given in terms of the function $\chi(w)$

$$\begin{aligned} \chi(w) &= -\frac{(2\pi)^3}{m}\int \frac{\mathcal{D}q_3}{2\pi}\frac{1}{q_3^2 + w + c_B^2} = -\frac{(2\pi)^3}{m\beta}\int_{-\pi}^{\pi} d\alpha \rho_B(\alpha) \sum_{n\,\in\,\mathbb{Z}} \frac{1}{(2\pi\frac{n}{\beta} + \frac{\alpha}{\beta})^2 + w + c_B^2} \ , \\ &= -\frac{2\pi^3}{m}\int_{-\pi}^{\pi} d\alpha \rho_B(\alpha)\frac{1}{\sqrt{w + c_B^2}}\times \\ &\qquad \times \left(\coth\left(\tfrac{\beta}{2}\sqrt{w + c_B^2} + \tfrac{i}{2}\alpha\right) + \coth\left(\tfrac{\beta}{2}\sqrt{w + c_B^2} - \tfrac{i}{2}\alpha\right)\right) \ . \end{aligned} \tag{B.40}$$

The q_3 sum in the last term in (B.38) is given by $c_0 \equiv \sum_n 1$ and is hence divergent. We regularize the divergent sum using ζ-function regularization in which case we have $c_0(\text{reg.}) = 1 + 2\zeta(0) = 0$. Thus, equation (B.38) becomes

$$T(m) = -\frac{1}{(2\pi)^3}\int_0^\infty \frac{dw}{4\pi}\ \chi(w)\left(\mathcal{L}(w) - (w + c_B^2)\right) = 2\int_0^\infty \frac{dw}{4\pi}\,\xi'(w)(\mathcal{L}(w) - (w + c_B^2)) \ , \tag{B.41}$$

where we have used $\chi(w) = -2(2\pi)^3\xi'(w)$. Next, recall the expressions (B.33) and (B.37):

$$g(w) = 1 + \lambda_B \xi(w) \ , \quad \mathcal{I}(w) = \int_0^w dz g(z) = w + \lambda_B \mathcal{I}_\xi(w) \ , \quad c_B^2 = m^2(1 + \lambda_B\xi(0))^2 \ . \tag{B.42}$$

Inserting (B.42) into (B.41) and Taylor-expanding $\mathcal{L}(w) - (w + c_B^2)$ in all explicit factors of λ_B (around $\lambda_B = 0$) we find

$$\begin{aligned}\mathcal{L}(w) - (w + c_B^2) &= \frac{\mathcal{I}(w)}{g(w)^2} + \frac{4}{3}m^2 g(w) + m^2 g(w)^2 + \frac{2}{3}m^2\frac{g(0)^3}{g(w)^2} - (w + m^2 g(0)^2) \ , \\ &= \sum_{n=0}^{\infty}(-\lambda_B)^n \mathcal{L}_n(w) \ , \end{aligned} \tag{B.43}$$

with

$$\begin{aligned}&\mathcal{L}_0(w) = 2m^2 \ , \quad \mathcal{L}_1(w) = -\left(\mathcal{I}_\xi(w) - 2w\xi(w)\right) - 2m^2\xi(w) \ , \\ &\mathcal{L}_2(w) = \left(-2\mathcal{I}_\xi(w)\xi(w) + 3w\xi(w)^2\right) + m^2\left(3\xi(w)^2 - 4\xi(w)\xi(0) + \xi(0)^2\right) \ , \\ &\mathcal{L}_n(w) = \left((n+1)w\xi(w)^n - n\xi(w)^{n-1}\mathcal{I}_\xi(w)\right) + \\ &\frac{2m^2}{3}\Big((n+1)\xi(w)^n - (n-2)\xi(w)^{n-3}\xi(0)^3 + 3(n-1)\xi(w)^{n-2}\xi(0)^2 - 3n\ \xi(w)^{n-1}\xi(0)\Big) \quad \text{for} \quad n \geq 3 \ . \end{aligned} \tag{B.44}$$

The integral over w in (B.41) becomes

$$\sum_{n=0}^{\infty}(-\lambda_B)^n \int_0^\infty \frac{dw}{4\pi}\xi'(w)\mathcal{L}_n(w) \ . \tag{B.45}$$

The integral over the first two terms in the expressions for $\mathcal{L}_{n\geq 1}$ in (B.44) can be simplified by writing this in a total derivative form

$$dw\ \xi'(w)\left((n+1)\xi(w)^n w - n\xi(w)^{n-1}\mathcal{I}_\xi(w)\right) = d(\xi(w)^{n+1}w - \xi(w)^n\mathcal{I}_\xi(w)) \tag{B.46}$$

In the dimensional regularization scheme used in our previous paper [54] we have $\xi(\infty) = 0$. Also, by definition we have $\mathcal{I}_\xi(0) = 0$. This implies that

$$\int_0^\infty dw\ \xi'(w)\left((n+1)\xi(w)^n w - n\xi(w)^{n-1}\mathcal{I}_\xi(w)\right) = 0 \ . \tag{B.47}$$

The remaining terms in $\mathcal{L}$ are simple polynomials in ξ and the integrations can be easily performed. Only $\mathcal{L}_0$ and $\mathcal{L}_1$ give non-zero contributions:

$$\int_0^\infty dw\ \xi'(w)\mathcal{L}_0(w) = -2m^2\xi(0) \ , \quad \int_0^\infty dw\ \xi'(w)\mathcal{L}_1(w) = +m^2\xi(0)^2 \ . \tag{B.48}$$

Substituting the above results into (B.41), we get

$$T(m) = -\frac{m^2}{2\pi}\Big(2\xi(0) + \lambda_B\xi(0)^2\Big) \ . \tag{B.49}$$

Recalling the equation (B.37) for $\widehat{c}_B$ and $m = \lambda_B\sigma_B$, we have the final expression for the tadpole contribution from the W boson propagator (B.30):

$$\lambda_B N_B T(\sigma_B) = -\frac{N_B}{2\pi}\left(c_B^2 - \lambda_B^2\sigma_B^2\right) \ . \tag{B.50}$$

B.3 The critical boson scaling limit

Recall that the RB theory reduces to the critical boson or CB theory in the scaling limit

$$m_B^2 \to \infty, \quad \lambda_B b_4 \to \infty, \quad \frac{m_B^2}{2\lambda_B b_4} = m_B^{\rm cri} = \text{fixed} \ . \tag{B.51}$$

In this subsection we study the reduction of the off-shell free energy under this scaling limit. The off-shell free energy (3.52) simplifies in the limit (B.51) as follows. The second term in the second line of (3.52) reduces to

$$6\lambda_B \widehat{b}_4 \widehat{\sigma}_B \left(\widehat{m}_B^{\rm cri} + \widehat{\sigma}_B\right) = 6\lambda_B \widehat{b}_4 \left[\left(\widehat{\sigma}_B + \frac{\widehat{m}_B^{\rm cri}}{2}\right)^2 - \frac{(\widehat{m}_B^{\rm cri})^2}{4}\right] \tag{B.52}$$

Note that confining potential (B.52) is infinitely stiff in the CB scaling limit. It follows that $\widehat{\sigma}_B$ is frozen at the minimum of (B.52) i.e. at $\sigma_B = -\frac{m_B^{\rm cri}}{2}$ in the CB scaling limit. It follows that in this limit (3.52) simplifies to

$$\begin{aligned} F(c_B, \widetilde{\mathcal{S}}) = \frac{N_B}{6\pi}\Bigg[& -\widehat{c}_B^3 - 4\widetilde{\mathcal{S}}^3\lambda_B^2 + \frac{3}{2}\widehat{c}_B^2 \widehat{m}_B^{\rm cri} \\ & + 6\widetilde{\mathcal{S}}^2\lambda_B^2 \widehat{m}_B^{\rm cri} - 3\widetilde{\mathcal{S}}\lambda_B^2 (\widehat{m}_B^{\rm cri})^2 + 6\widehat{c}_B|\lambda_B|(\widetilde{\mathcal{S}} - \tfrac{\widehat{m}_B^{\rm cri}}{2})^2 \\ & + 3\int_{-\pi}^{\pi} d\alpha \rho_B(\alpha) \int_{\widehat{c}_B}^{\infty} dy\ y\ \left(\log\left(1 - e^{-y-i\alpha}\right) + \log\left(1 - e^{-y+i\alpha}\right)\right)\Bigg]. \end{aligned} \tag{B.53}$$

(we have omitted a divergent constant proportional to b_4 that can be cancelled by a cosmological constant counterterm.) Extremizing (B.53) w.r.t. $\widetilde{\mathcal{S}}$ we recover the first of the gap equations in (3.53) under the replacement

$$\sigma_B \to -\frac{m_B^{\rm cri}}{2} \ .$$

The two inequivalent solutions of this equation are (3.54) and (3.55) under the same replacement for σ_B. These solutions correspond to the unHiggsed and Higgsed branches respectively.

On the Higgsed branch we can plug the solution of (3.55) back into (B.53) to find a free energy as a function of the single off-shell variable c_B; the final result of this exercise is given by the critical boson free energy given in (3.33)[5]. In a similar manner, on the unHiggsed

[5]This procedure automatically produces a particular choice of the cosmological constant counterterm. It would be interesting to investigate if this particular value has physical significance.

branch we can plug the solution of (3.54) into (B.53) to find off-shell free energy as a function of c_B, given by

$$F_{\rm CB}(c_B) = \frac{N_B}{6\pi}\Bigg[-\widehat{c}_B^3 + \frac{3}{2}\widehat{m}_B^{\rm cri}\widehat{c}_B^2 - \frac{\lambda_B^2(\widehat{m}_B^{\rm cri})^3}{2} + 3\int_{-\pi}^{\pi}\rho(\alpha)d\alpha\int_{\widehat{c}_B}^{\infty}dyy\left(\log\left(1-e^{-y-i\alpha}\right)+\log\left(1-e^{-y+i\alpha}\right)\right)\Bigg]. \tag{B.54}$$

Appendix C

Appendices for Chapter 3

C.1 Four derivatives on four F structures

We show that all possible Lagrangian structures built out of four derivatives and four field strengths, are essentially descendants of structures that can be constructed out of two derivatives on four F_{ab}s.For convenience, we will work in momentum space with appropriate particle labels.

$$\begin{aligned}
T_1^4 &= k_\alpha^3 F_{\alpha a}^1 k_\beta^3 F_{\beta a}^2 k_\gamma^1 F_{\gamma b}^3 k_\delta^1 F_{\delta b}^4 \\
&= -k_\alpha^3 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{\gamma b}^3 k_\delta^1 F_{\delta b}^4 - k_\alpha^3 F_{a\gamma}^1 k_\beta^3 F_{\beta a}^2 k_\alpha^1 F_{\gamma b}^3 k_\delta^1 F_{\delta b}^4 \\
&= -k_\alpha^3 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{\gamma b}^3 k_\delta^1 F_{\delta b}^4 + I_{\text{desc}} \\
&= k_\gamma^3 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{b\alpha}^3 k_\delta^1 F_{\delta b}^4 + k_b^3 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{\alpha\gamma}^3 k_\delta^1 F_{\delta b}^4 + I_{\text{desc}} \\
&= k_\alpha^3 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{\gamma b}^3 k_\delta^1 F_{\delta b}^4 + k_b^3 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{\alpha\gamma}^3 k_\delta^1 F_{\delta b}^4 + I_{\text{desc}} \\
&= k_\alpha^3 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{\gamma b}^3 k_\delta^1 F_{\delta b}^4 - k_b^2 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{\alpha\gamma}^3 k_\delta^1 F_{\delta b}^4 + I_{\text{desc}} \\
\therefore T_1^4 &\sim -k_b^2 F_{\gamma\alpha}^1 k_\beta^3 F_{\beta a}^2 k_a^1 F_{\alpha\gamma}^3 k_\delta^1 F_{\delta b}^4 + I_{\text{desc}} \sim \widetilde{I}_{\text{desc}}
\end{aligned} \tag{C.1}$$

In deriving this we have used Bianchi identity between the first k^1 and F^1 in the second line. In the fourth line we use Bianchi identity between first k^3 and F^3. The structure in the sixth line is due to momentum conservation and antisymmetry of F_{ab}. In the final step, to go from I_{desc} to $\widetilde{I}_{\text{desc}}$, we have used Bianchi identity between k^2 and F^2. Similar manipulations can be used to prove that all of the T_i^4s are descendants of the lower structures and do not constitute independent lagrangian terms.

C.2 Two derivatives on four F structures

We classify all the possible independent lagrangian structures that can be constructed out of two derivatives acting on four field strengths. Essentially we show that out of all the

structures listed in (4.39), only one is independent. We show some of the manipulations in the hope that it is clear to the reader that rest follow from similar manipulations. Of the possible T_i^2s, we show that each one of them can be related to T_{10}^2.

$$\begin{aligned}
T_1^2 &= k_a^2 F_{\beta a}^1 k_b^1 F_{\alpha b}^2 F_{\theta\alpha}^3 F_{\theta\beta}^4 \\
&\equiv k_a^1 F_{\beta a}^2 k_b^2 F_{\alpha b}^1 F_{\theta\alpha}^3 F_{\theta\beta}^4 \\
&= k_b^2 F_{\alpha b}^1 k_a^1 F_{\beta a}^2 F_{\theta\alpha}^3 F_{\theta\beta}^4 \\
&= T_3^2 \\
&= -k_a^2 F_{ab}^1 k_\beta^1 F_{\alpha b}^2 F_{\theta\beta}^3 F_{\theta\alpha}^4 - k_a^2 F_{b\beta}^1 k_a^1 F_{\alpha b}^2 F_{\theta\beta}^3 F_{\theta\alpha}^4 \\
&= -k_a^2 F_{ba}^1 k_\beta^1 F_{\theta\beta}^3 F_{b\alpha}^2 F_{\theta\alpha}^4 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right) \\
&= -k_a^2 F_{ba}^1 k_\beta^1 F_{\theta\beta}^3 F_{b\alpha}^2 F_{\theta\alpha}^4 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right) \\
&= k_b^2 F_{ba}^1 k_\beta^1 F_{\theta\beta}^3 F_{\alpha a}^2 F_{\theta\alpha}^4 + k_\alpha^2 F_{ba}^1 k_\beta^1 F_{\theta\beta}^3 F_{ab}^2 F_{\theta\alpha}^4 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right) \\
&= k_a^2 F_{ab}^1 k_\beta^1 F_{\theta\beta}^3 F_{\alpha b}^2 F_{\theta\alpha}^4 + k_\alpha^2 F_{\theta\alpha}^4 k_\beta^1 F_{\theta\beta}^3 F_{ba}^1 F_{ab}^2 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right) \\
&= k_a^2 F_{ab}^1 k_\beta^1 F_{\theta\beta}^3 F_{\alpha b}^2 F_{\theta\alpha}^4 + T_{10}^2 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right) \\
\therefore T_1^2 &= T_3^2 \sim T_{10}^2 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right)
\end{aligned} \tag{C.2}$$

The steps in the manipulation are as follows. In the second line we relabel $(1 \leftrightarrow 2)$ to establish the fact that $T_1^2 \sim T_3^2$ in momentum space. In the fifth line, we use Bianchi identity corresponding to particle 1. In the eighth line we use Bianchi identity corresponding to particle 2. Equating the seventh and tenth line we obtain the final identity. Hence we have the lagrangian term T_1^2 is identical to T_{10}^2 upto descendants of four photon Lagrangians of derivative order 4. Let us look at a second example

$$\begin{aligned}
T_6^2 &= k_a^2 F_{\beta a}^1 k_b^3 F_{\alpha b}^2 F_{\theta\beta}^3 F_{\theta\alpha}^4 \\
&= -k_\alpha^2 F_{\beta a}^1 k_b^3 F_{ba}^2 F_{\theta\beta}^3 F_{\theta\alpha}^4 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right) \\
&= -k_\alpha^2 F_{\theta\alpha}^1 k_b^3 F_{ba}^2 F_{\beta a}^4 F_{\theta\beta}^3 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right) \\
&= -k_\alpha^2 F_{\theta\alpha}^1 k_b^3 F_{ab}^2 F_{\beta\theta}^3 F_{\beta a}^4 + \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right) \\
\therefore T_6^2 &\sim \mathcal{O}\left(\mathrm{desc}(\mathrm{Tr}F^4)\right)
\end{aligned} \tag{C.3}$$

where we have used Bianchi identity corresponding to the second particle and relabelling of $(1 \leftrightarrow 4)$. In this way all the lagrangian structures can be represented in terms of T_{10}^2 and descendants of $\mathrm{Tr}F^4$ and $(\mathrm{Tr}F^2)^2$.

C.3 Derivatives on two Riemann tensors

We look at possible two and four derivative contractions on two Riemann tensors and systematically reduce them to higher Riemann polynomials and Ricci tensor polynomials

1

$$\begin{aligned} S &= \int \sqrt{g}\nabla_\mu R_{abcd}\nabla^\mu R^{acbd} \\ &= \int \sqrt{g}\left(-\nabla_a R_{b\mu cd} - \nabla_b R_{\mu acd}\right)\nabla^\mu R^{acbd} \\ &= \int \sqrt{g}\nabla^\mu\left(-\nabla_a R_{b\mu cd} - \nabla_b R_{\mu acd}\right) R^{acbd} + C_{\partial\mathcal{M}} \\ &= \int \sqrt{g}\left(-\nabla_a\nabla^\mu R_{b\mu cd} - \nabla_b\nabla^\mu R_{\mu acd}\right) R^{acbd} + C_{\partial\mathcal{M}} + \widetilde{C}_{R^3} \\ &= \widehat{C}_{R_{\mu\nu}} + C_{\partial\mathcal{M}} + \widetilde{C}_{R^3} \end{aligned} \tag{C.4}$$

2

$$\begin{aligned} S &= \int \nabla_a R_{bcfe}\nabla^b R^{acfe} \\ &= -\int \left(\nabla^b\nabla_a R_{bcfe}\right) R^{acfe} + C_{\partial\mathcal{M}} \\ &= \widehat{C}_{R_{\mu\nu}} + C_{\partial\mathcal{M}} + \widetilde{C}_{R^3} \end{aligned} \tag{C.5}$$

3

$$\begin{aligned} S &= \int \sqrt{g}\nabla_\mu\nabla_\nu R_{abcd}\nabla^\mu\nabla^\nu R^{acbd} \\ &= \int \sqrt{g}\nabla_\nu\nabla^\mu\nabla_\mu R_{abcd}R^{acbd} + C_{\partial\mathcal{M}} \\ &= \int \sqrt{g}\nabla_\nu\nabla^\mu\nabla_\mu\left(-\nabla_a R_{b\nu cd} - \nabla_b R_{\nu acd}\right) R^{acbd} + C_{\partial\mathcal{M}} \\ &= \widehat{C}_{R_{\mu\nu}} + C_{\partial\mathcal{M}} + \widetilde{C}_{\nabla R\nabla RR} \end{aligned} \tag{C.6}$$

4

$$\begin{aligned} S &= \int \sqrt{g}\nabla_a\nabla_b R_{ecfd}\nabla^e\nabla^f R^{acbd} \\ &= \int \sqrt{g}\nabla^e\nabla^f\nabla_a\nabla_b R_{ecfd}R^{acbd} + C_{\partial\mathcal{M}} \\ &= \widehat{C}_{R_{\mu\nu}} + C_{\partial\mathcal{M}} + \widetilde{C}_{\nabla R\nabla RR} \end{aligned} \tag{C.7}$$

Similar manipulations can be used to get rid of all higher order derivatives at the two Riemann tensor level.

C.4 Derivatives on three Riemann tensors

We look at the possible scalar polynomials constructed out of two derivatives and three Riemann tensors

1

$$\begin{aligned}\int \sqrt{g} R^{pqrs} R_p{}^{tuv} \nabla_v \nabla_s R_{qtru} &= \int \sqrt{g} R^{pqrs} R_p{}^{tuv} \nabla_v (-\nabla_r R_{qtus} - \nabla_u R_{qtsr}) \\ &= -\int \sqrt{g} R^{pqsr} R_p{}^{tuv} \nabla_v \nabla_r R_{qtsu} + C'_{R^4} \\ \therefore \int \sqrt{g} R^{pqrs} R_p{}^{tuv} \nabla_v \nabla_s R_{qtru} &\sim \frac{1}{2} C'_{R^4} \end{aligned} \tag{C.8}$$

2

$$\begin{aligned}\int \sqrt{g} R^{pqrs} \nabla_r R^{tuv}{}_p \nabla_s R_{tuvq} &= -\int \sqrt{g} R^{pqrs} \nabla_s \nabla_r R^{tuv}{}_p R_{tuvq} + C_{\partial\mathcal{M}} \\ &\sim C'_{R^4} + C_{\partial\mathcal{M}} \end{aligned} \tag{C.9}$$

3

$$\begin{aligned}\int \sqrt{g} R^{pqrs} \nabla^v R^t{}_p{}^u{}_r \nabla_v R_{tqus} &= -\int \sqrt{g} (\nabla^v R^{pqrs} R^t{}_p{}^u{}_r \nabla_v R_{tqus} + R^{pqrs} R^t{}_p{}^u{}_r \nabla^v \nabla_v R_{tqus}) + C_{\partial\mathcal{M}} \\ \int \sqrt{g} R^{pqrs} \nabla^v R^t{}_p{}^u{}_r \nabla_v R_{tqus} &= \frac{-1}{2} \int \sqrt{g} R^{pqrs} R^t{}_p{}^u{}_r \nabla^v \nabla_v R_{tqus} + C_{\partial\mathcal{M}} \\ &\sim \widehat{C}_{R_{\mu\nu}} + C'_{R^4} + C_{\partial\mathcal{M}} \end{aligned} \tag{C.10}$$

4

$$\begin{aligned}\int \sqrt{g} R^{pqrs} \nabla_q R_{tuvp} \nabla_s R^{tuv}{}_r &= -\int \sqrt{g} R^{pqrs} (\nabla_v R_{tupq} + \nabla_p R_{tuqv}) \nabla_s R^{tuv}{}_r \\ \int \sqrt{g} R^{pqrs} \nabla_q R_{tuvp} \nabla_s R^{tuv}{}_r &= \frac{-1}{2} \int \sqrt{g} R^{pqrs} \nabla_v R_{tupq} \nabla_s R^{tuv}{}_r \\ &= \frac{-1}{2} \int \sqrt{g} R^{pqrs} \nabla_v R_{tupq} (-\nabla^v R^{tu}{}_{rs} - \nabla_r R^{tu}{}_s{}^v) \\ \int \sqrt{g} R^{pqrs} \nabla_q R_{tuvp} \nabla_s R^{tuv}{}_r &= \frac{1}{4} \int \sqrt{g} R^{pqrs} \nabla_v R_{tupq} \nabla^v R^{tu}{}_{rs} \\ &\sim \widehat{C}_{R_{\mu\nu}} + C'_{R^4} \end{aligned} \tag{C.11}$$

C.5 Counting Index Structures

C.5.1 3 point functions

Vectors (gauge fields)

According to the analysis presented in Section 4.2, the distinct 3 point functions of 3 vectors in a d dimensional CFT are in one to one correspondence with the permutation invariant $SO(d-1)$ singlets one can form out of 3 $SO(d-1)$ vectors [1]

Consider the $3 \times (d-1)$ dimensional vector space spanned by the indices of these 3 vectors. According to a classic theorem, this 3 $(d-1)$ dimensional vector space decomposes into representations of $U(d-1) \times S_3$ as

$$\sum_Y U_Y P_Y$$

where Y are Young Tableaux with 3 boxes, the sum over Y runs over all these Tableaux, U_Y is the $U(d-1)$ representation with the Young Tableaux Y, and P_Y is the S_3 representation with the same Young Tableaux.

As there are no nontrivial cross ratios for 3 particles, the only way for a 3 point function to be invariant under S_3 is for the index structure itself to be invariant under S_3 (see below for the contrast with 4 point functions). Consequently we are forced to restrict our attention to the completely symmetric representation of $U(d-1)$. This single representation decomposes into the vector and the traceless symmetric 3 tensor of $SO(d-1)$. As there are no singlets in this decomposition, it follows that there are no structures for the 3 point functions of 3 vectors.

Traceless symmetric 2 tensors (gravitons)

A traceless symmetric tensor in d dimensions has $d(d+1)/2-1$ components. Subtracting the d components (due to conservation) leave us with $r = d(d-1)/2-1$ components of - presumably - a traceless symmetric tensor of $SO(d-1)$. As above the $3r$ dimensional vector space decomposes into

$$\sum_Y U_Y P_Y$$

where Y are Young Tableaux with 3 boxes, the sum over Y runs over all these Tableaux, U_Y is the $U(r)$ representation with the Young Tableaux Y, and P_Y is the S_3 representation with the same Young Tableaux. As above we are only interested in Young Tableaux Y corresponding to the 3 box completely symmetric representation. We appear to be left with a 6 index $SO(d-1)$ tensor with indices split into 3 groups of 2. The indices are traceless and symmetric in each pair, and we have complete symmetry under interchange

[1] We deal with $SO(d-1)$ singlets rather than $SO(d-2)$ singlets, as in the next subsection, as any three operators can be put in a straight line using conformal transformations, and $SO(d-1)$ is the part of the rotation group that leaves a line in d dimensions invariant.

of pairs. We should be counting the inequivalent $SO(d-1)$ singlets in such a tensor. As far as I can tell there is exactly one way of creating such a singlet - by contracting indices in a chain, i.e. a contraction of the form

$$(a,b)\ (b,c)\ (c,a)$$

This is not the expected answer of 3 inequivalent 3 point functions for the stress tensor. What am i doing wrong?

C.5.2 Four Point Functions

Consider the symmetric group S_4. Consider the $Z_2 \times Z_2$ subgroup of S_4 generated by the two elements $A = P_{12}P_{34}$ and $B = P_{13}P_{24}$ where P_{ab} is the permutation of the elements a and b. The four elements of this group are

$$I, \quad A = P_{12}P_{34}, \quad B = P_{13}P_{24}, \quad AB = BA = P_{14}P_{23}$$

In other words the nontrivial elements of this subgroup are obtained by double pairing 1, 2, 3, 4 in each of the 3 possible ways and by permuting in pairs.

Note that this $Z_2 \times Z_2$ subgroup leaves two conformal cross ratios

$$u = \frac{r_{12}r_{34}}{r_{13}r_{24}}, \quad v = \frac{r_{14}r_{23}}{r_{13}r_{24}}$$

invariant. [2] If a correlator has to be invariant under S_4, $Z_2 \times Z_2$ has to leave its index structure invariant.

Now consider the action of the symmetric group S_4 on 'states' that are annihilated by $Z_2 \times Z_2$. This is equivalent to studying S_4 modded out by $Z_2 \times Z_2$ from the right. [3] It is not difficult to convince oneself that under this equivalence rule, every element of S_4 can be put in the form $(1, a, b, c)$ where (a, b, c) represent any any permutation of the numbers of $2, 3, 4$. It follows, in other words, that $\frac{S_4}{Z_2 \times Z_2} = S_3$.

Representations R of S_4 are labelled by Young Tableaux with 4 boxes. The R^{th} representation acts on a vector space V_R. Let us denote the $Z_2 \times Z_2$ invariant subspace of V_R by $\widetilde{V}_R$. The states in $\widetilde{V}_R$ transform in representations of some representation $\widetilde{R}$ of $\frac{S_4}{Z_2 \times Z_2} = S_3$. As $\widetilde{R}$ is a representation of S_3 it is specified by a (sum of) Young Tableaux with 3 boxes.

S_4 has 5 irreducible representations. These representations fall into 2 classes. In the first class we have 3 representations; the one dimensional representation with four boxes in the first row, the one dimensional representation with four boxes in the first column and the 2 dimensional representation with two boxes respectively in each of the first and

[2] In the context of scattering, the same subgroup leaves s, t and u unchanged.

[3] We use the following labelling of elements of S_4; (2314), for instance, denotes the permutation that takes 1 to 2, 2 to 3, 3 to 1 and 4 to 4.

second rows (this also means that the Young Tableaux has two boxes in each column). In the second class we have two 3 dimensional representations - the first is labelled by a Tableaux that contain three boxes in the first row and one in the second, and the second representation is labelled by a Tableaux that contains 3 boxes in the first column and one box in the second column.

The two classes of representations in the previous paragraph differ in the following way. Representations in class 2 have no states that are left invariant by *any* of the nontrivial elements of $Z_2 \times Z_2$. The reason for this is intuitively clear from the nature of the Tableaux for these representations; that any double paring of boxes in these Tableaux has one pair 'symmetric' but the other pair 'antisymmetric'. On the other hand *every* state in the representations that occur in the first class is left invariant by every element of $Z_2 \times Z_2$. Intuitively, this is because every double pairing of boxes in such Tableaux is such that the paired elements are both symmetric or both antisymmetric. It follows that every state in each of the representations in class 1 is automatically invariant under $Z_2 \times Z_2$.

Focusing on representations of class 1, is not too difficult to convince oneself that the one dimensional representation of S_4 with 4 boxes descends to the one dimensional representation with 3 boxes in a single row of S_3. Similarly the one dimensional representation of S_4 with four boxes in a single column descends to the one dimensional representation of S_3 with 3 boxes in a single column. Finally the 2 dimensional representation (two boxes in the first and second row) of S_4 descends to the two dimensional representation of S_3 that is labelled by a Tableaux with two boxes in the first row and one box in the second row.

Vectors (gauge fields)

In this case the state space above is given by the three class one representations of $U(d-1)$. Lets first consider the completely symmetric tableaux. This gives us a completely symmetric - but not traceless tensor. This decomposes into 3 $SO(d-1)$ representations: (symmetric and traceless) 4 tensors, 2 tensors and the singlet. Each of these representations has a single SO(d-2) singlet (obtained by setting all indices in the $SO(d_1)$ singlet to $d-1$). So we have 3 $SO(d-1)$ singlets from this representation, each of which transform in the 1 dimensional (completely symmetric) representation of S_3.

Now let us turn to the $U(d-1)$ representation with two boxes in each of the first two rows of the Young Tableaux. In this case the symmetries of the tensor in question are precisely those of the Riemann tensor. In terms of $SO(d-1)$ we have the representations with $h_1 = 2, h_2 = 2$, $h_1 = 2, h_2 = 0$, and the singlet. The first of these representations has no $SO(d-2)$ singlets. The second and third of these each have exactly one $SO(d-2)$ singlets. It follows that from this representation we get 2 $SO(d-2)$ singlets, each of which transform in the 2 dimensional representation of S_3 with two boxes in the first row and one in the second row. In summary we get a total of 4 singlets from this representation.

Finally, the representation with a single columns of four boxes is a single $SO(d-1)$ representation (the completely antisymmetric 4 index tensor).

C.6 Functions of s t and u as representations of S^3

In this section we study polynomials of s, t and u graded by their degree. We will be particularly interested in decomposing the space of such functions into distinct representations of the permutation group S_3 which permutes the three variables. We will perform our study both for unconstrained functions of s, t and u, as well as for 'constrained' functions, i.e. functions that are regarded as identical if they agree when $s+t+u=0$. We start with a brief discussion of the permutation group S_3 and its representations.

C.6.1 S_3 and its irreducible representations

Let us first recall that the group S_3 has six elements. An element of the permutation group is said to be odd or even depending on whether it is built out of an odd or even number of permutations. We label an element of S_3 by the result of the action of that element on $(1,2,3)$, Thus $(1,2,3)$, the identity element I is even. The other two even elements are the cyclical permutations $C=(2,3,1)$ and $C^{-1}=(3,1,2)$. The set of even elements of S_3 form the abelian subgroup Z_3. The odd elements of this group are the three permutations $P_{12}=(2,1,3)$, $P_{13}=(3,2,1)$ and $P_{23}=(1,3,2)$.

If we think of 1, 2 and 3 as basis elements of a three dimensional vector space then the action above yields a representation of S_3 in terms of 3×3 matrices. The representation is clearly reducible: all permutation elements act as identity on the basis vector $(1+2+3)$. This one dimensional representation is the completely symmetrical representation of S_3; this is the representation labelled by three boxes in the first row of the Young Tableaux.

On the other hand the two dimensional set of vectors with $1+2+3=0$ mix only among themselves under the permutation group, and so transform under a 2 dimensional representation of this group. A convenient basis for this space is found by diagonalizing C. Let

$$B_1 = e^{-\frac{2\pi i}{3}}|1\rangle + |2\rangle + e^{\frac{2\pi i}{3}}|3\rangle, \quad B_2 = e^{\frac{2\pi i}{3}}|1\rangle + |2\rangle + e^{-\frac{2\pi i}{3}}|3\rangle \tag{C.12}$$

Then

$$C\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} e^{-\frac{2\pi i}{3}} & 0 \\ 0 & e^{\frac{2\pi i}{3}} \end{pmatrix}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \tag{C.13}$$

The action of the permutations on the same basis is given by

$$\begin{aligned} P_{12}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} &= \begin{pmatrix} 0 & e^{-\frac{2\pi i}{3}} \\ e^{\frac{2\pi i}{3}} & 0 \end{pmatrix}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \\ P_{13}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} &= \begin{pmatrix} 0 & e^{\frac{2\pi i}{3}} \\ e^{\frac{-2\pi i}{3}} & 0 \end{pmatrix}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \\ P_{23}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \end{aligned} \tag{C.14}$$

Note that the phases that appear in the top right corner of the three matrices (C.14) are, respectively, $e^{-\frac{2\pi i}{3}}$, 1 and $e^{\frac{2\pi i}{3}}$. Of course the redefinition $B^1 \to \alpha B^1$ changes each of these

phases by α. It follows that while the actual value of each phase is convention dependent, the ratios of the phases are physical (convention independent). We will encounter this fact below.

The equations (C.13) and (C.14) give a complete characterization of this two dimensional irreducible representation of S_3 (this is the representation labelled by the Young Tableaux with two boxes in its first row and one in its second row, or equivalently two boxes in the first column and one in the second column). From the fact that $P_{ij}^2 = 1$ it follows that the eigenvalues of the operator $P_{ij} = \pm 1$ in every representation of S_3. In this particular 2 dimensional representation it is easily verified that the two eigenvalues of P_{ij} are plus one and minus one for every choice of i and j.

Though it does not show up in the decomposition described above, there is a third irreducible representation of S_3. This is the completely antisymmetric representation labelled by a Young tableaux with three boxes in the first column. In this one dimensional representation, every even element of the permutation element acts as unity (identity) while every odd element as -1 (negative identity).

C.6.2 Left action of S_3 on itself

Now consider the left action of S_3 on itself. This clearly generates a 6 dimensional reducible representation of the permutation group. Similarly the basis vector

$$\begin{aligned} &(123) + (231) + (312) + (213) + (321) + (132) \\ &I + S + S^{-1} + P_{12} + P_{13} + P_{23} \end{aligned} \tag{C.15}$$

transforms in the one dimensional symmetric representation. Similarly the basis vector

$$\begin{aligned} &(123) + (231) + (312) - (213) - (321) - (132) \\ &I + S + S^{-1} - P_{12} - P_{13} - P_{23} \end{aligned} \tag{C.16}$$

transforms in the one dimensional antisymmetric representation. What remains is the four dimensional vector space of elements

$$A(123) + B(231) + C(312) + p(213) + q(132) + r(321)$$

with $A + B + C = 0$ and $p + q + r = 0$. It is not difficult to decompose this four dimensional vector space into a direct sum of two copies of the two dimensional irreducible representation defined in the previous subsection. Define

$$\begin{aligned} b_1 &= e^{-\frac{2\pi i}{3}}(123) + (231) + e^{\frac{2\pi i}{3}}(312) \\ b_2 &= e^{\frac{2\pi i}{3}}(123) + (231) + e^{-\frac{2\pi i}{3}}(312) \\ \beta_1 &= e^{-\frac{2\pi i}{3}}(213) + (132) + e^{\frac{2\pi i}{3}}(321) \\ \beta_2 &= e^{\frac{2\pi i}{3}}(213) + (132) + e^{-\frac{2\pi i}{3}}(321) \end{aligned} \tag{C.17}$$

With this definition it is easy to see that b_1 and β_1 are eigenstates of C with eigenvalue $e^{-\frac{2\pi i}{3}}$ while b_2 and β_2 are eigenstates of C with eigenvalue $e^{\frac{2\pi i}{3}}$. Moreover

$$
\begin{aligned}
P_{12}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} &= \begin{pmatrix} 0 & e^{\frac{2\pi i}{3}} \\ e^{-\frac{2\pi i}{3}} & 0 \end{pmatrix}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \qquad P_{12}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 & e^{\frac{2\pi i}{3}} \\ e^{-\frac{2\pi i}{3}} & 0 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\
P_{13}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \qquad P_{13}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\
P_{23}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} &= \begin{pmatrix} 0 & e^{-\frac{2\pi i}{3}} \\ e^{\frac{2\pi i}{3}} & 0 \end{pmatrix}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \qquad P_{23}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 & e^{-\frac{2\pi i}{3}} \\ e^{\frac{2\pi i}{3}} & 0 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}
\end{aligned}
\tag{C.18}
$$

from which it immediately follows that the vectors $(b_1+\beta_1), (b_2+\beta_2)$ transform in one irreducible 2 dimensional representation of S_3, while the vectors $(b_1-\beta_1), (b_2-\beta_2)$ transform in a second copy of the same representation [4]

C.6.3 Functions of 3 variables and the permutation group.

Consider a function of three variables s, t and u. Let the permutation group S_3 act on these three variables. Given any particular function $f(s,t,u)$, the action of the permutation group generates upto 5 new functions.

If the original function was invariant under a subgroup of the permutation group then we would obtain fewer than 5 new functions. Let us first suppose that this is not the case. In this case the six dimensional linear vector space of the obtained functions transforms in a six dimensional representation of S_3. Infact the representation we find is identical to that of the previous subsection (representation of S_3 by left action on itself). So we obtain one copy of the symmetric representation, one copy of the antisymmetric representation and two copies of the two dimensional representation.

Consider a general function $f(s,t,u)$. Given any such function it is easy to break it up into a part that is completely symmetric, a part that is completely antisymmetric and a part that lies somewhere in the (generically 4 dimensional) representation vector space of

[4] In the basis that we have arbitrarily adopted, the phases in the to right corner of the 2×2 matrices that represent P_{12}, P_{13} an P_{23} respectively are $e^{\frac{2\pi i}{3}}$, 1 and $e^{-\frac{2\pi i}{3}}$ in the first representation (plus sign) , while the same phases are $-e^{\frac{2\pi i}{3}}$, -1 and $-e^{-\frac{2\pi i}{3}}$ in the second representation. In neither case are these the same phases that appear in (C.14), but in both cases the ratio of phases is the same as in (C.14). This is enough to ensure that both choices transform in the representation (C.14), see the discussion under (C.14).

the two dimensional representations. We have

$$
\begin{aligned}
f(s,t,u) &= f^{sym}(s,t,u) + f^{as}(s,t,u) + f^{mixed}(s,t,u) \\
f^{sym}(s,t,u) &= P^{sym} f = \frac{f(s,t,u) + f(t,u,s) + f(u,s,t) + f(t,s,u) + f(u,t,s) + f(s,u,t)}{6} \\
f^{as}(s,t,u) &= P^{as} f = \frac{f(s,t,u) + f(t,u,s) + f(u,s,t) - f(t,s,u) - f(u,t,s) - f(s,u,t)}{6} \\
f^{mixed}(s,t,u) &= P^{mixed} f = \frac{2f(s,t,u) - f(t,u,s) - f(u,s,t)}{3}
\end{aligned}
\tag{C.19}
$$

It is easy to verify that P^{sym}, P^{as} and P^{mixed} all square to themselves and so are projectors. Moreover they project onto orthogonal subspaces, so that the product of two non equal projectors vanishes. Finally, these projectors commute with the action of the permutation group. The last equation in (C.19) asserts that the polynomials that transform in the mixed representations vanish under Z_3 symmetrization as well as under complete symmetrization (these two facts imply these functions also vanish under complete antisymmetrization).

C.6.4 Counting S_3 representations in polynomials

Let us consider the partition function

$$\widetilde{Z}_{no-sym} = \sum_m \widetilde{d}_{ns}(m) x^m \tag{C.20}$$

where $\widetilde{d}_{ns}(m)$ is the number of polynomials of s, t u of degree m. Clearly $\widetilde{Z}_{no-sym}$ is the product of a partition function for s, a partition function for t and a partition function for u. Each of these partition functions is $\frac{1}{1-x}$ and so we find

$$\widetilde{Z}_{no-sym} = \frac{1}{(1-x)^3} \tag{C.21}$$

from which it follows that

$$\widetilde{d}_{ns}(m) = \frac{(m+2)(m+1)}{2} \tag{C.22}$$

The correctness of (C.22) (note the RHS of this equation is the number of ways of arranging m circles and two crosses on a line) can also be verified directly by simple combinatorial counting.

Now the set of polynomials of any given degree - counted by (C.21) - transform in a representation of the symmetric group S_3. The representation in question can be decomposed into the three irreducible representations of S^3 that we have discussed above - the completely symmetric, completely antisymmetric and 2 dimensional. We would like to find separate partition functions for polynomials in the completely symmetric, completely antisymmetric and mixed representations of S_3.

The partition function for completely symmetric polynomials of 3 variables is the same as the partition functions for 3 identical bosons in a harmonic oscillator potential (with zero point energy ignored). Using the formulas for Bose exponentiation, it is given by the power of p^3 in the expression

$$\exp\left[\frac{p}{1-x}+\frac{p^2}{2(1-x^2)}+\frac{p^3}{3(1-x^3)}\right]$$

and so is given by

$$\widetilde{Z}_{sym}(x)=\frac{1}{3(1-x^3)}+\frac{1}{6(1-x)^3}+\frac{1}{2(1-x)(1-x^2)}=\frac{1}{(1-x)(1-x^2(1-x^3)} \quad \text{(C.23)}$$

The partition function for completely antisymmetric polynomials of 3 variables is the same as the partition functions for 3 identical fermions a harmonic oscillator potential (with zero point energy ignored). Using the formulas for Bose exponentiation, it is given by the power of p^3 in the expression

$$\exp\left[\frac{p}{1-x}-\frac{p^2}{2(1-x^2)}+\frac{p^3}{3(1-x^3)}\right]$$

and so is given by

$$\widetilde{Z}_{as}(x)=\frac{1}{3(1-x^3)}+\frac{1}{6(1-x)^3}-\frac{1}{2(1-x)(1-x^2)}=\frac{x^3}{(1-x)(1-x^2)(1-x^3)} \quad \text{(C.24)}$$

The partition function for polynomials that transform in the nontrivial two dimensional representations is given by (C.21)-(C.23) -(C.24) and is given

$$\widetilde{Z}_{mixed}=\frac{2x}{(1-x)^2(1-x^3)} \quad \text{(C.25)}$$

(we emphasized that $\widetilde{Z}_{mixed}$ is the partition function for polynomials in the two dimensional representation and not for the representations themselves; as each representation has two polynomials, the partition function that counts the number of mixed representations is half of (C.27). As a satisfying consistency check, note that even the partition function for representations clearly has an expansion in positive integers).

Note of course that - by construction

$$\widetilde{Z}_{no-sym}(x)=\widetilde{Z}_{sym}(x)+\widetilde{Z}_{as}(x)+\widetilde{Z}_{mixed}(x) \quad \text{(C.26)}$$

As an application of these ideas, let us compute the partition function over those polynomials that happen to be symmetric under an interchange of s and t, but are otherwise unconstrained. This is the same as counting the partition function over all polynomials that have eigenvalue unity under the permutation P_{12}. We know that this is true of every polynomial in a symmetric representation and exactly half of the polynomials in the two dimensional representations (because P_{12} has eigenvalues ± 1 in such representations. It follows that

$$\widetilde{Z}_{Z_2-sym}=\frac{x}{(1-x)^2(1-x^3)}+\frac{1}{3(1-x^3)}+\frac{1}{6(1-x)^3}+\frac{1}{2(1-x)(1-x^2)}=\frac{1+x}{(1-x^2)^2(1-x)} \quad \text{(C.27)}$$

C.6.5 Counting Polynomials with $s+t+u=0$

When dealing with S matrices, we are not interested in all all polynomials of s, t and u, merely those that are nonzero when $s+u+t=0$. As this condition - namely that $s+t+u=0$ - is S_3 invariant, the resultant polynomials still transform in representations of S_3, even though they are effectively functions of only two variables. We denote the partition functions over polynomials with $s+t+u=0$ that transform in representation r as $Z_r(x)$ (note that Z_r (note the lack of the tilde on Z). The relationship between Z_r and $\widetilde{Z}_r$ is very simple. For every representation r

$$\widetilde{Z}_r = Z_r \times \frac{1}{1-x} \tag{C.28}$$

The equation (C.28) expresses the fact that if we have a polynomial that transforms in the representation r and also are nonvanishing when $s+t+u=0$ we can find another polynomial that also transforms in the representation r - but this time vanishes when $s+t+u=0$ - by multiplying the first polynomial with $(s+t+u)^m$ for any possible value of m (the factor of $\frac{1}{1-x}$ in (C.28) is the partition function for polynomials in $(s+t+u)$.). [5]

It follows immediately that

$$\begin{aligned}
Z_{no-sym} &= \frac{1}{(1-x)^2} = \sum_{m=0}^{\infty}(m+1)x^m \\
Z_{sym} &= \frac{1}{(1-x^2)(1-x^3)} = 1+x^2+x^3+x^4+x^5+2x^6+x^7+2x^8+2x^9+2x^{10}\ldots \\
Z_{as} &= \frac{x^3}{(1-x^2)(1-x^3)} = x^3\left(1+x^2+x^3+x^4+x^5+2x^6+x^7+2x^8+2x^9+2x^{10}\ldots\right) \\
Z_{mixed} &= \frac{2x}{(1-x)(1-x^3)} \\
Z_{Z_2-sym} &== \frac{1+x}{(1-x^2)^2} = \sum_{m=0}^{\infty}\left(\left[\frac{m}{2}\right]+1\right)x^m
\end{aligned} \tag{C.29}$$

As a check, it is a simple matter to directly count the number of polynomials with $s+t+u=0$ that are symmetric under an s and t interchange. Using $u=-s-t$ we can write all such polynomials as symmetric polynomials in the 2 variables s and t. It is clear that the number of such polynomials of degree m is given by $[\frac{m}{2}]+1$ in agreement with the last of (C.29). Of course the first of (C.29) is also obviously true.

[5] One way to think of this is to use s, t, and $s+t+u$ as our variables in the original full space of polynomials and to note that - as $s+t+u$ is permutation invariant - the transformation properties of the polynomial lie entirely in its s and t dependence.

For each r we perform the expansion

$$Z_r(x) = \sum_{m=0}^{\infty} d_r(m) x^m \tag{C.30}$$

At large m we have

$$\begin{aligned} d_{no-sym}(m) &= m+1 \\ d_{sym}(m) &\sim \frac{m+1}{6} \\ d_{as}(m) &=\sim \frac{m+1}{6} \\ d_{mixed}(m) &= \frac{2(m+1)}{3} \\ Z_{Z_2-sym} &== \frac{m+1}{2} = \end{aligned} \tag{C.31}$$

When we need to count degrees of freedom, we will assign one unit of 'degrees of freedom' to functions whose number grows like $\frac{m+1}{6}$ at large m. With this counting a symmetry-less polynomial of s, t u has six degrees of freedom. Completely symmetric/antisymmetric polynomials have one degree of freedom. The set of all mixed polynomials have 4 degrees of freedom. The set of mixed polynomials that are also Z_2 symmetric have 2 degrees of freedom. The set of all Z_2 invariant polynomials have 3 degrees of freedom.

C.6.6 Examples at low orders

There is only one polynomial of degree zero; it transforms in the symmetric representation. At degree 1 we have two polynomials which can be taken to be $2s - t - u = 3s$ and $2t - s - u = 3t$. Note that these polynomials all vanish upon complete symmetrization. Moreover they are each annihilated by a Z_2 permutation symmetry. They are mapped to each other under permutation and consequently transform in the mixed representation. Note also that exactly one of these expressions is symmetric under the interchange of, say, s and t while the other one can be chosen to be antisymmetric under the same exchange $(s + t)$ and $s - t$ are the two choices here).

At degree 2 we have three polynomials which can be taken to be $(s^2 + t^2 + u^2)$, and $(s^2 + t^2 - 2u^2)$ and $t^2 + u^2 - 2s^2)$. The first of these is in the symmetric representation. The next two transform in the mixed representation. A convenient basis for these two elements is $s^2 + t^2 - 2u^2$ and $s^2 - t^2$. Clearly the first element is symmetric under the interchange of s and t, while the second element is antisymmetric under the same interchange.

At degree three we have four polynomials. These can be taken to be $s^3 + t^3 + u^3$, $s^2t - t^2s - s^2u + u^2 - u^2t - t^2u$, $s^3 + t^3 - 2u^3$, and $s^3 + u^3 - 2t^3$. The first of these expressions is completely symmetric. The second is completely antisymmetric. The remaining two belong

to the mixed representation. Once again we can choose a basis in this set of elements that is symmetric and antisymmetric under interchange of s and t (the symmetric expression is $s^3 + t^3 - 2u^3$, whereas the antisymmetric expression is $s^3 - t^3$)

Bibliography

[1] S. Sachdev, *Holographic metals and the fractionalized Fermi liquid*, *Phys. Rev. Lett.* **105** (2010) 151602, [arXiv:1006.3794].

[2] Juan Maldacena, Stephen H. Shenker, and Douglas Stanford. A bound on chaos. *JHEP*, 08:106, 2016.

[3] Sayantan Choudhury, Anshuman Dey, Indranil Halder, Lavneet Janagal, Shiraz Minwalla, and Rohan Poojary. Notes on melonic $O(N)^{q-1}$ tensor models. *JHEP*, 06:094, 2018.

[4] A. Kitaev, "A simple model of quantum holography", http://online.kitp.ucsb.edu/online/entangled15/kitaev/ http://online.kitp.ucsb.edu/online/entangled15/kitaev2/ "Talks at KITP, April 7, 2015 and May 27, 2015."

[5] J. Maldacena and D. Stanford, *Remarks on the Sachdev-Ye-Kitaev model*, *Phys. Rev.* **D94** (2016), no. 10 106002, [arXiv:1604.07818].

[6] A. Almheiri and J. Polchinski, *Models of AdS_2 backreaction and holography*, *JHEP* **11** (2015) 014, [arXiv:1402.6334].

[7] K. Jensen, *Chaos in AdS_2 Holography*, *Phys. Rev. Lett.* **117** (2016), no. 11 111601, [arXiv:1605.06098].

[8] J. Maldacena, D. Stanford, and Z. Yang, *Conformal symmetry and its breaking in two dimensional Nearly Anti-de-Sitter space*, *PTEP* **2016** (2016), no. 12 12C104, [arXiv:1606.01857].

[9] Joseph Polchinski and Vladimir Rosenhaus. The Spectrum in the Sachdev-Ye-Kitaev Model. *JHEP*, 04:001, 2016.

[10] G. Mandal, P. Nayak, and S. R. Wadia, *Coadjoint orbit action of Virasoro group and two-dimensional quantum gravity dual to SYK/tensor models*, arXiv:1702.04266.

[11] D. J. Gross and V. Rosenhaus, *The Bulk Dual of SYK: Cubic Couplings*, *JHEP* **05** (2017) 092, [arXiv:1702.08016].

[12] S. Forste and I. Golla, *Nearly AdS_2 sugra and the super-Schwarzian*, *Phys. Lett.* **B771** (2017) 157–161, [arXiv:1703.10969].

[13] S. R. Das, A. Jevicki, and K. Suzuki, *Three Dimensional View of the SYK/AdS Duality*, arXiv:1704.07208.

[14] A. M. García-García and J. J. M. Verbaarschot, *Spectral and thermodynamic properties of the Sachdev-Ye-Kitaev model*, *Phys. Rev.* **D94** (2016), no. 12 126010, [arXiv:1610.03816].

[15] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, *Black Holes and Random Matrices*, *JHEP* **05** (2017) 118, [arXiv:1611.04650].

[16] A. M. García-García and J. J. M. Verbaarschot, *Analytical Spectral Density of the Sachdev-Ye-Kitaev Model at finite N*, arXiv:1701.06593.

[17] D. Stanford and E. Witten, *Fermionic Localization of the Schwarzian Theory*, arXiv:1703.04612.

[18] V. V. Belokurov and E. T. Shavgulidze, *The Exact Solution of the Schwarzian Theory*, arXiv:1705.02405.

[19] E. Witten, *An SYK-Like Model Without Disorder*, arXiv:1610.09758.

[20] I. R. Klebanov and G. Tarnopolsky, *Uncolored Random Tensors, Melon Diagrams, and the SYK Models*, arXiv:1611.08915.

[21] I. R. Klebanov and G. Tarnopolsky, *On Large N Limit of Symmetric Traceless Tensor Models*, arXiv:1706.00839.

[22] R. Gurau, *Colored Group Field Theory*, *Commun. Math. Phys.* **304** (2011) 69–93, [arXiv:0907.2582].

[23] R. Gurau, *The 1/N expansion of colored tensor models*, *Annales Henri Poincare* **12** (2011) 829–847, [arXiv:1011.2726].

[24] R. Gurau and V. Rivasseau, *The 1/N expansion of colored tensor models in arbitrary dimension*, *Europhys. Lett.* **95** (2011) 50004, [arXiv:1101.4182].

[25] R. Gurau, *The complete 1/N expansion of colored tensor models in arbitrary dimension*, *Annales Henri Poincare* **13** (2012) 399–423, [arXiv:1102.5759].

[26] V. Bonzom, R. Gurau, A. Riello, and V. Rivasseau, *Critical behavior of colored tensor models in the large N limit*, *Nucl. Phys.* **B853** (2011) 174–195, [arXiv:1105.3122].

[27] R. Gurau and J. P. Ryan, *Colored Tensor Models - a review*, *SIGMA* **8** (2012) 020, [arXiv:1109.4812].

[28] P. Narayan and J. Yoon, *SYK-like Tensor Models on the Lattice*, arXiv:1705.01554.

[29] J. Yoon, *SYK Models and SYK-like Tensor Models with Global Symmetry*, arXiv:1707.01740.

[30] S. Dartois, H. Erbin, and S. Mondal, *Conformality of $1/N$ corrections in SYK-like models*, arXiv:1706.00412.

[31] T. Nishinaka and S. Terashima, *A Note on Sachdev-Ye-Kitaev Like Model without Random Coupling*, arXiv:1611.10290.

[32] C. Peng, M. Spradlin, and A. Volovich, *A Supersymmetric SYK-like Tensor Model*, *JHEP* **05** (2017) 062, [arXiv:1612.03851].

[33] C. Krishnan, S. Sanyal, and P. N. Bala Subramanian, *Quantum and Holographic Tensor Models*, *JHEP* **03** (2017) 056, [arXiv:1612.06330].

[34] F. Ferrari, *The Large D Limit of Planar Diagrams*, arXiv:1701.01171.

[35] R. Gurau, *Quenched equals annealed at leading order in the colored SYK model*, arXiv:1702.04228.

[36] V. Bonzom, L. Lionni, and A. Tanasa, *Diagrammatics of a colored SYK model and of an SYK-like tensor model, leading and next-to-leading orders*, *J. Math. Phys.* **58** (2017), no. 5 052301, [arXiv:1702.06944].

[37] C. Krishnan, K. V. P. Kumar, and S. Sanyal, *Random Matrices and Holographic Tensor Models*, *JHEP* **06** (2017) 036, [arXiv:1703.08155].

[38] S. Chaudhuri, V. I. Giraldo-Rivera, A. Joseph, R. Loganayagam, and J. , *Abelian Tensor Models on the Lattice*, arXiv:1705.01930.

[39] T. Azeyanagi, F. Ferrari, and F. I. Schaposnik Massolo, *Phase Diagram of Planar Matrix Quantum Mechanics, Tensor and SYK Models*, arXiv:1707.03431.

[40] S. Giombi, I. R. Klebanov, and G. Tarnopolsky, *Bosonic Tensor Models at Large N and Small ϵ*, arXiv:1707.03866.

[41] S. Sachdev, K. Jensen, Y. Gu, A. Georges, W. Fu, and R. A. Davison, *Thermoelectric transport in disordered metals without quasiparticles: the Sachdev-Ye-Kitaev models and holography*, arXiv:1612.00849.

[42] O. Aharony, J. Marsano, S. Minwalla, K. Papadodimas, and M. Van Raamsdonk, *The Hagedorn - deconfinement phase transition in weakly coupled large N gauge theories*, *Adv. Theor. Math. Phys.* **8** (2004) 603–696, [hep-th/0310285]. [,161(2003)].

[43] Anshuman Dey, Indranil Halder, Sachin Jain, Lavneet Janagal, Shiraz Minwalla, and Naveen Prabhakar. Duality and an exact Landau-Ginzburg potential for quasi-bosonic Chern-Simons-Matter theories. *JHEP*, 11:020, 2018.

[44] E. Witten, *Anti-de Sitter Space, Thermal Phase Transition, And Confinement In Gauge Theories*, 9803131v2.

[45] O. Aharony, J. Marsano, S. Minwalla, K. Papadodimas, and M. Van Raamsdonk, *A First order deconfinement transition in large N Yang-Mills theory on a small 3-sphere*, *Phys. Rev.* **D71** (2005) 125018, [hep-th/0502149].

[46] J. M. Maldacena, *Eternal black holes in anti-de Sitter*, *JHEP* **04** (2003) 021, [hep-th/0106112].

[47] K. Bulycheva, I. R. Klebanov, A. Milekhin, and G. Tarnopolsky, *Spectra of Operators in Large N Tensor Models*, arXiv:1707.09347.

[48] C. Krishnan and K. V. P. Kumar, *Towards a Finite-N Hologram*, arXiv:1706.05364.

[49] D. J. Gross and E. Witten, *Possible Third Order Phase Transition in the Large N Lattice Gauge Theory*, *Phys. Rev.* **D21** (1980) 446–453.

[50] S. R. Wadia, *A Study of U(N) Lattice Gauge Theory in 2-dimensions*, arXiv:1212.2906.

[51] S. R. Wadia, *N = Infinity Phase Transition in a Class of Exactly Soluble Model Lattice Gauge Theories*, *Phys. Lett.* **B93** (1980) 403–410.

[52] M. R. Gaberdiel and R. Gopakumar, *Minimal Model Holography*, *J. Phys.* **A46** (2013) 214002, [arXiv:1207.6697].

[53] C.-M. Chang, S. Minwalla, T. Sharma, and X. Yin, *ABJ Triality: from Higher Spin Fields to Strings*, *J. Phys.* **A46** (2013) 214009, [arXiv:1207.4485].

[54] S. Choudhury, A. Dey, I. Halder, S. Jain, L. Janagal, S. Minwalla, and N. Prabhakar, *Bose-Fermi Chern-Simons Dualities in the Higgsed Phase*, *JHEP* **11** (2018) 177, [arXiv:1804.08635].

[55] C. Cordova, P.-S. Hsin, and N. Seiberg, *Global Symmetries, Counterterms, and Duality in Chern-Simons Matter Theories with Orthogonal Gauge Groups*, *SciPost Phys.* **4** (2018), no. 4 021, [arXiv:1711.10008].

[56] S. Minwalla and S. Yokoyama, *Chern Simons Bosonization along RG Flows*, *JHEP* **02** (2016) 103, [arXiv:1507.04546].

[57] O. Aharony, S. Jain, and S. Minwalla, *Flows, Fixed Points and Duality in Chern-Simons-matter theories*, *JHEP* **12** (2018) 058, [arXiv:1808.03317].

[58] O. Aharony, G. Gur-Ari, and R. Yacoby, *Correlation Functions of Large N Chern-Simons-Matter Theories and Bosonization in Three Dimensions*, *JHEP* **12** (2012) 028, [arXiv:1207.4593].

[59] S. Jain, M. Mandlik, S. Minwalla, T. Takimi, S. R. Wadia, and S. Yokoyama, *Unitarity, Crossing Symmetry and Duality of the S-matrix in large N Chern-Simons theories with fundamental matter*, *JHEP* **04** (2015) 129, [arXiv:1404.6373].

[60] S. Yokoyama, *Scattering Amplitude and Bosonization Duality in General Chern-Simons Vector Models*, *JHEP* **09** (2016) 105, [arXiv:1604.01897].

[61] S. Jain, S. Minwalla, T. Sharma, T. Takimi, S. R. Wadia, and S. Yokoyama, *Phases of large N vector Chern-Simons theories on $S^2 \times S^1$*, *JHEP* **09** (2013) 009, [arXiv:1301.6169].

[62] S. Jain, S. Minwalla, and S. Yokoyama, *Chern Simons duality with a fundamental boson and fermion*, *JHEP* **11** (2013) 037, [arXiv:1305.7235].

[63] S. Sachdev and J. Ye, *Gapless spin fluid ground state in a random, quantum Heisenberg magnet*, *Phys. Rev. Lett.* **70** (1993) 3339, [cond-mat/9212030].

[64] A. M. Perelomov, *Hypergeometric Solutions of Some Algebraic Equations*, *Theoretical and Mathematical Physics* **140** (July, 2004) 895–904, [math-ph/0303016].

[65] O. Aharony, O. Bergman, D. L. Jafferis, and J. Maldacena, *N=6 superconformal Chern-Simons-matter theories, M2-branes and their gravity duals*, *JHEP* **10** (2008) 091, [arXiv:0806.1218].

[66] R. Gurau, *The complete $1/N$ expansion of a SYK–like tensor model*, *Nucl. Phys.* **B916** (2017) 386–401, [arXiv:1611.04032].

[67] G. Turiaci and H. Verlinde, *Towards a 2d QFT Analog of the SYK Model*, arXiv:1701.00528.

[68] T. Li, J. Liu, Y. Xin, and Y. Zhou, *Supersymmetric SYK model and random matrix theory*, *JHEP* **06** (2017) 111, [arXiv:1702.01738].

[69] C. Peng, *Vector models and generalized SYK models*, *JHEP* **05** (2017) 129, [arXiv:1704.04223].

[70] R. Gurau, *The $\iota\epsilon$ prescription in the SYK model*, arXiv:1705.08581.

[71] A. M. G. T. K. Bulycheva, I. Klebanov, *Spectra of Operators in Large N tensor Models. (2017)*, .

[72] S. Giombi, V. Kirilin, and E. Skvortsov, *Notes on Spinning Operators in Fermionic CFT*, *JHEP* **05** (2017) 041, [arXiv:1701.06997].

[73] Z. Komargodski and N. Seiberg, *A symmetry breaking scenario for QCD_3*, *JHEP* **01** (2018) 109, [arXiv:1706.08755].

[74] E. Sezgin, E. D. Skvortsov, and Y. Zhu, *Chern-Simons Matter Theories and Higher Spin Gravity*, *JHEP* **07** (2017) 133, [arXiv:1705.03197].

[75] K. Aitken, A. Baumgartner, A. Karch, and B. Robinson, *3d Abelian Dualities with Boundaries*, *JHEP* **03** (2018) 053, [arXiv:1712.02801].

[76] A. Karch, D. Tong, and C. Turner, *Mirror Symmetry and Bosonization in 2d and 3d*, *JHEP* **07** (2018) 059, [arXiv:1805.00941].

[77] O. Aharony, L. F. Alday, A. Bissi, and R. Yacoby, *The Analytic Bootstrap for Large N Chern-Simons Vector Models*, *JHEP* **08** (2018) 166, [arXiv:1805.04377].

[78] R. Yacoby, *Scalar Correlators in Bosonic Chern-Simons Vector Models*, arXiv:1805.11627.

[79] K. Aitken, A. Baumgartner, and A. Karch, *Novel 3d bosonic dualities from bosonization and holography*, *JHEP* **09** (2018) 003, [arXiv:1807.01321].

[80] I. R. Klebanov and A. M. Polyakov, *AdS dual of the critical O(N) vector model*, *Phys. Lett.* **B550** (2002) 213–219, [hep-th/0210114].

[81] E. Sezgin and P. Sundell, *Massless higher spins and holography*, *Nucl. Phys.* **B644** (2002) 303–370, [hep-th/0205131]. [Erratum: Nucl. Phys.B660,403(2003)].

[82] S. Giombi and X. Yin, *Higher Spin Gauge Theory and Holography: The Three-Point Functions*, *JHEP* **09** (2010) 115, [arXiv:0912.3462].

[83] S. Giombi, S. Minwalla, S. Prakash, S. P. Trivedi, S. R. Wadia, and X. Yin, *Chern-Simons Theory with Vector Fermion Matter*, *Eur. Phys. J.* **C72** (2012) 2112, [arXiv:1110.4386].

[84] J. Maldacena and A. Zhiboedov, *Constraining Conformal Field Theories with A Higher Spin Symmetry*, *J. Phys.* **A46** (2013) 214011, [arXiv:1112.1016].

[85] J. Maldacena and A. Zhiboedov, *Constraining conformal field theories with a slightly broken higher spin symmetry*, *Class. Quant. Grav.* **30** (2013) 104003, [arXiv:1204.3882].

[86] G. Gur-Ari and R. Yacoby, *Correlators of Large N Fermionic Chern-Simons Vector Models*, *JHEP* **02** (2013) 150, [arXiv:1211.1866].

[87] A. Bedhotiya and S. Prakash, *A test of bosonization at the level of four-point functions in Chern-Simons vector models*, *JHEP* **12** (2015) 032, [arXiv:1506.05412].

[88] G. J. Turiaci and A. Zhiboedov, *Veneziano Amplitude of Vasiliev Theory*, *JHEP* **10** (2018) 034, [arXiv:1802.04390].

[89] N. Seiberg, T. Senthil, C. Wang, and E. Witten, *A Duality Web in 2+1 Dimensions and Condensed Matter Physics*, *Annals Phys.* **374** (2016) 395–433, [arXiv:1606.01989].

[90] J. Murugan and H. Nastase, *Particle-vortex duality in topological insulators and superconductors*, *JHEP* **05** (2017) 159, [arXiv:1606.01912].

[91] S. Yokoyama, *Chern-Simons-Fermion Vector Model with Chemical Potential*, *JHEP* **01** (2013) 052, [arXiv:1210.4109].

[92] S. Jain, S. P. Trivedi, S. R. Wadia, and S. Yokoyama, *Supersymmetric Chern-Simons Theories with Vector Matter*, *JHEP* **10** (2012) 194, [arXiv:1207.4750].

[93] O. Aharony, S. Giombi, G. Gur-Ari, J. Maldacena, and R. Yacoby, *The Thermal Free Energy in Large N Chern-Simons-Matter Theories*, *JHEP* **03** (2013) 121, [arXiv:1211.4843].

[94] T. Takimi, *Duality and higher temperature phases of large N Chern-Simons matter theories on $S^2 \times S^1$*, *JHEP* **07** (2013) 177, [arXiv:1304.3725].

[95] S. Yokoyama, *A Note on Large N Thermal Free Energy in Supersymmetric Chern-Simons Vector Models*, *JHEP* **01** (2014) 148, [arXiv:1310.0902].

[96] G. Gur-Ari and R. Yacoby, *Three Dimensional Bosonization From Supersymmetry*, *JHEP* **11** (2015) 013, [arXiv:1507.04378].

[97] M. Geracie, M. Goykhman, and D. T. Son, *Dense Chern-Simons Matter with Fermions at Large N*, *JHEP* **04** (2016) 103, [arXiv:1511.04772].

[98] K. Inbasekar, S. Jain, S. Mazumdar, S. Minwalla, V. Umesh, and S. Yokoyama, *Unitarity, crossing symmetry and duality in the scattering of $\mathcal{N} = 1$ susy matter Chern-Simons theories*, *JHEP* **10** (2015) 176, [arXiv:1505.06571].

[99] Y. Dandekar, M. Mandlik, and S. Minwalla, *Poles in the S-Matrix of Relativistic Chern-Simons Matter theories from Quantum Mechanics*, *JHEP* **04** (2015) 102, [arXiv:1407.1322].

[100] K. Inbasekar, S. Jain, P. Nayak, and V. Umesh, *All tree level scattering amplitudes in Chern-Simons theories with fundamental matter*, *Phys. Rev. Lett.* **121** (2018), no. 16 161601, [arXiv:1710.04227].

[101] K. Inbasekar, S. Jain, S. Majumdar, P. Nayak, T. Neogi, T. Sharma, R. Sinha, and V. Umesh, *Dual Superconformal Symmetry of $\mathcal{N} = 2$ Chern-Simons theory with Fundamental Matter and Non-Renormalization at Large N*, arXiv:1711.02672.

[102] F. Benini, C. Closset, and S. Cremonesi, *Comments on 3d Seiberg-like dualities*, *JHEP* **10** (2011) 075, [arXiv:1108.5373].

[103] G. Gur-Ari, S. A. Hartnoll, and R. Mahajan, *Transport in Chern-Simons-Matter Theories*, *JHEP* **07** (2016) 090, [arXiv:1605.01122].

[104] O. Aharony, G. Gur-Ari, and R. Yacoby, *d=3 Bosonic Vector Models Coupled to Chern-Simons Gauge Theories*, *JHEP* **03** (2012) 037, [arXiv:1110.4382].

[105] T. Nosaka and S. Yokoyama, *Complete factorization in minimal $\mathcal{N} = 4$ Chern-Simons-matter theory*, *JHEP* **01** (2018) 001, [arXiv:1706.07234].

[106] S. Giombi, *Higher Spin -CFT Duality*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1-26, 2015*, pp. 137–214, 2017. arXiv:1607.02967.

[107] S. R. Wadia, *ChernÃćÂĂÂŞSimons theories with fundamental matter: A brief review of large N results including FermiÃćÂĂÂŞBose duality and the S-matrix*, *Int. J. Mod. Phys.* **A31** (2016), no. 32 1630052.

[108] Y. Frishman and J. Sonnenschein, *Large N Chern-Simons with massive fundamental fermions - A model with no bound states*, *JHEP* **12** (2014) 165, [arXiv:1409.6083].

[109] V. Gurucharan and S. Prakash, *Anomalous dimensions in non-supersymmetric bifundamental Chern-Simons theories*, *JHEP* **09** (2014) 009, [arXiv:1404.7849]. [Erratum: JHEP11,045(2017)].

[110] S. Giombi, *Testing the Boson/Fermion Duality on the Three-Sphere*, arXiv:1707.06604.

[111] W. A. Bardeen, *The Massive Fermion Phase for the U(N) Chern-Simons Gauge Theory in D=3 at Large N*, *JHEP* **10** (2014) 39, [arXiv:1404.7477].

[112] W. A. Bardeen and M. Moshe, *Spontaneous breaking of scale invariance in a D=3 U(N) model with Chern-Simons gauge fields*, *JHEP* **06** (2014) 113, [arXiv:1402.4196].

[113] M. Moshe and J. Zinn-Justin, *3D Field Theories with Chern–Simons Term for Large N in the Weyl Gauge*, *JHEP* **01** (2015) 054, [arXiv:1410.0558].

[114] S. Giombi, V. Gurucharan, V. Kirilin, S. Prakash, and E. Skvortsov, *On the Higher-Spin Spectrum in Large N Chern-Simons Vector Models*, *JHEP* **01** (2017) 058, [arXiv:1610.08472].

[115] V. Gurucharan and S. Prakash, *On the Higher Spin Spectrum of Chern-Simons Theory coupled to Fermions in the Large Flavour Limit*, *JHEP* **02** (2018) 094, [arXiv:1711.11300].

[116] D. Radicevic, *Disorder Operators in Chern-Simons-Fermion Theories*, *JHEP* **03** (2016) 131, [arXiv:1511.01902].

[117] O. Aharony, *Baryons, monopoles and dualities in Chern-Simons-matter theories*, *JHEP* **02** (2016) 093, [arXiv:1512.00161].

[118] O. Aharony, P. Narayan, and T. Sharma, *On monopole operators in supersymmetric Chern-Simons-matter theories*, *JHEP* **05** (2015) 117, [arXiv:1502.00945].

[119] D. Radicevic, D. Tong, and C. Turner, *Non-Abelian 3d Bosonization and Quantum Hall States*, *JHEP* **12** (2016) 067, [arXiv:1608.04732].

[120] O. Aharony, F. Benini, P.-S. Hsin, and N. Seiberg, *Chern-Simons-matter dualities with SO and USp gauge groups*, *JHEP* **02** (2017) 072, [arXiv:1611.07874].

[121] A. Karch and D. Tong, *Particle-Vortex Duality from 3d Bosonization*, *Phys. Rev.* **X6** (2016), no. 3 031043, [arXiv:1606.01893].

[122] A. Karch, B. Robinson, and D. Tong, *More Abelian Dualities in 2+1 Dimensions*, *JHEP* **01** (2017) 017, [arXiv:1609.04012].

[123] P.-S. Hsin and N. Seiberg, *Level/rank Duality and Chern-Simons-Matter Theories*, *JHEP* **09** (2016) 095, [arXiv:1607.07457].

[124] J. Gomis, Z. Komargodski, and N. Seiberg, *Phases Of Adjoint QCD_3 And Dualities*, *SciPost Phys.* **5** (2018), no. 1 007, [arXiv:1710.03258].

[125] K. Jensen, *A master bosonization duality*, *JHEP* **01** (2018) 031, [arXiv:1712.04933].

[126] D. Gaiotto, Z. Komargodski, and N. Seiberg, *Time-reversal breaking in QCD_4, walls, and dualities in 2 + 1 dimensions*, *JHEP* **01** (2018) 110, [arXiv:1708.06806].

[127] F. Benini, P.-S. Hsin, and N. Seiberg, *Comments on global symmetries, anomalies, and duality in (2 + 1)d*, *JHEP* **04** (2017) 135, [arXiv:1702.07035].

[128] K. Jensen and A. Karch, *Embedding three-dimensional bosonization dualities into string theory*, *JHEP* **12** (2017) 031, [arXiv:1709.07872].

[129] K. Jensen and A. Karch, *Bosonizing three-dimensional quiver gauge theories*, *JHEP* **11** (2017) 018, [arXiv:1709.01083].

[130] F. Benini, *Three-dimensional dualities with bosons and fermions*, *JHEP* **02** (2018) 068, [arXiv:1712.00020].

[131] X. O. Camanho, J. D. Edelstein, J. Maldacena, and A. Zhiboedov, *Causality Constraints on Corrections to the Graviton Three-Point Coupling*, *JHEP* **02** (2016) 020, [arXiv:1407.5597].

[132] P. Kravchuk and D. Simmons-Duffin, *Counting Conformal Correlators*, *JHEP* **02** (2018) 096, [arXiv:1612.08987].

[133] Daliang Li, David Meltzer, and David Poland. Conformal Collider Physics from the Lightcone Bootstrap. *JHEP*, 02:143, 2016.

[134] M. S. Costa, J. Penedones, D. Poland, and S. Rychkov, *Spinning Conformal Correlators*, *JHEP* **11** (2011) 071, [arXiv:1107.3554].

[135] Miguel S. Costa, Joao Penedones, David Poland, and Slava Rychkov. *Spinning Conformal Blocks*, *JHEP*, 11:154, (2011).

[136] Anatoly Dymarsky. On the four-point function of the stress-energy tensors in a CFT. *JHEP*, 10:075, 2015.

[137] J. H. Schwarz, *Superstring Theory*, *Phys. Rept.* **89** (1982) 223–322.

[138] H. Kawai, D. C. Lewellen, and S. H. H. Tye, *A Relation Between Tree Amplitudes of Closed and Open Strings*, *Nucl. Phys.* **B269** (1986) 1–23.

[139] S. A. Fulling, R. C. King, B. G. Wybourne, and C. J. Cummins, *Normal forms for tensor polynomials. 1: The Riemann tensor*, *Class. Quant. Grav.* **9** (1992) 1151–1197.

[140] R. C. Myers and B. Robinson, *Black Holes in Quasi-topological Gravity*, *JHEP* **08** (2010) 067, [arXiv:1003.5357].

[141] S. Sannan, *Gravity as the Limit of the Type II Superstring Theory*, *Phys. Rev.* **D34** (1986) 1749.