Aspects of Higher Spin Symmetry and Its Breaking

Alexander Zhiboedov

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Physics
Adviser: Juan Maldacena

September 2014

Abstract

This thesis explores different aspects of higher spin symmetry and its breaking in the context of Quantum Field Theory, AdS/CFT and String Theory. In chapter 2, we study the constraints imposed by the existence of a single higher spin conserved current on a three-dimensional conformal field theory (CFT). A single higher spin conserved current implies the existence of an infinite number of higher spin conserved currents. The correlation functions of the stress tensor and the conserved currents are then shown to be equal to those of a free field theory. Namely a theory of N free bosons or free fermions. This is an extension of the Coleman-Mandula theorem to CFT's, which do not have a conventional S-matrix. In chapter 3, we consider three-dimensional conformal field theories that have a higher spin symmetry that is slightly broken. The theories have a large N limit, in the sense that the operators separate into single-trace and multi-trace and obey the usual large N factorization properties. We assume that the only single trace operators are the higher spin currents plus an additional scalar. Using the slightly broken higher spin symmetry we constrain the three-point functions of the theories to leading order in N. We show that there are two families of solutions. One family can be realized as a theory of N fermions with an $O(N)$ Chern-Simons gauge field, the other as a N bosons plus the Chern-Simons gauge field.

In chapter 4, we consider several aspects of unitary higher-dimensional conformal field theories. We investigate the dimensions of spinning operators via the crossing equations in the light-cone limit. We find that, in a sense, CFTs become free at large spin and $1/s$ is a weak coupling parameter. The spectrum of CFTs enjoys additivity: if two twists τ_1, τ_2 appear in the spectrum, there are operators whose twists are arbitrarily close to $\tau_1 + \tau_2$. We characterize how $\tau_1 + \tau_2$ is approached at large spin by solving the crossing equations analytically. Applications include the 3d Ising model, theories with a gravity dual, SCFTs, and patterns of higher spin symmetry breaking.

In chapter 5, we consider higher derivative corrections to the graviton three-point coupling within a weakly coupled theory of gravity. We devise a thought experiment involving a high energy scattering process which leads to causality violation if the graviton three-point vertex contains the additional structures. This violation cannot be fixed by adding conventional particles with spins $J \leq 2$. But, it can be fixed by adding an infinite tower of extra massive particles with higher spins, $J > 2$. In AdS theories this implies a constraint on the conformal anomaly coefficients $\left|\frac{a-c}{c}\right| \lesssim \frac{1}{\Delta_{gap}^2}$ in terms of Δ_{gap}, the dimension of the lightest single particle operator with spin $J > 2$. For inflation, or de Sitter-like solutions, it indicates the existence of massive higher spin particles if the gravity wave non-gaussianity deviates significantly from the one computed in the Einstein theory.

Acknowledgements

I am deeply indebted to my advisor, Juan Maldacena, for sharing his time and incredible ideas, for teaching me so many different things, for his encouragement and support. It was always a special pleasure to discuss physics and collaborate with Juan. His office was like a launch site from which the shuttle of conversation goes high in the space of ideas, with impenetrable walls often becoming small dots far below, new landscapes opening around, and the horizon of ideas being increased in size vastly.

I am very grateful to my collaborators Zohar Komargodski, Xian Camanho, Jose Edelstein together with who I wrote papers presented in this thesis. I also would like to thank Andrei Belitsky, Stefan Hohenegger, Grisha Korchemsky and Emery Sokatchev for collaboration on a series of papers that is not included in the thesis.

I benefited greatly from many conversations with IAS and Princeton faculty, postdocs and students. I would like to thank Igor Klebanov, Daniel Marlow and Lyman Page for serving on thesis and related committees, for their help and support during my PhD. I am very grateful to Simone Giombi for agreeing to be a reader of this thesis. I am also indebted to Alexander Gorsky for introducing me to the field.

I would like to thank the Graduate School for providing me with the Charlotte Elizabeth Procter Fellowship during my fifth year, and the Friends of the International Center for their Excellence in Teaching Award.

I would like to thank my family, friends and dearest M. for their encouragement, care and support.

Contents

Abstract . . . iii

Acknowledgements . . . iv

1 Introduction 1

1.1 Gravity and Quantum . . . 1

1.2 The Search of an Ultimate Symmetry . . . 4

1.3 Overview of the Thesis . . . 7

2 Constraining Conformal Field Theories With a Higher Spin Symmetry 10

2.1 Introduction . . . 10

2.1.1 Organization of the chapter . . . 13

2.2 Generalities about higher spin currents . . . 13

2.3 Removing operators in the twist gap, $\frac{1}{2} < \tau < 1$ . . . 15

2.3.1 Action of the charges on twist one fields . . . 18

2.4 Basic facts about three point functions . . . 18

2.4.1 General expression for three point functions . . . 19

2.5 Argument using bilocal operators . . . 20

2.5.1 Light cone limits of correlators of conserved currents . . . 20

2.5.2 Getting an infinite number of currents . . . 23

2.5.3 Definition of bilocal operators . . . 25

2.5.4 Constraining the action of the higher spin charges . . . 26

2.5.5 Quantization of $\widetilde{N}$: the case of bosons . . . 32

2.5.6 Fermionic-like bilocal operators . . . 33

2.5.7 Quantization of $\widetilde{N}$: the case of fermions . . . 34

2.6 Arguments based on more generic three and four point functions . . . 35

2.6.1 Basic operations in the space of charge conservation identities . . . 35

2.6.2 From spin three to spin four 36
2.6.3 Analysis of three point functions using the spin four current. 38
2.6.4 Constraining the four point function of j_0 39
2.6.5 No parity odd piece 40
2.6.6 Case of fermions 41
2.7 Higher spin symmetries broken at order $\frac{1}{N}$ 43
2.8 Case of two conserved spin two currents 46
2.8.1 Adding higher spin conserved currents 48
2.9 Conclusions and discussion 48
2.10 Appendix A. Argument for $[Q_s, j_2] = \partial j_s + \cdots$ 51
2.11 Appendix B. An integral expression for the odd three point functions 51
2.11.1 Light cone limit for the odd three point function 54
2.12 Appendix C. No odd piece from the one point energy correlator 54
2.13 Appendix D. Half-integer higher spin currents implies even spin higher spin currents 56
2.14 Appendix E. Functions that are annihilated by all charges 57
2.14.1 Proof 57
2.14.2 Functions with Fourier Transform 58
2.15 Appendix F. Current conservation equation in terms of cross ratios 58
2.16 Appendix G. Four point function of scalar operators in the fermion-like theory . . . 59
2.17 Appendix I. Conformal blocks for free scalar 60
2.18 Appendix J. Fixing correlators in the light-like limit 61
2.18.1 Fermions 62
2.18.2 General operators 63
2.19 Appendix K. More explicit solutions to the charge conservation identities of section five 63
2.19.1 Expression for the currents in the free theory 64
2.19.2 Analyzing the charge conservation identity 65

3 Constraining Conformal Field Theories With a Slightly Broken Higher Spin Symmetry 68
3.1 Introduction 68
3.2 Possible divergence of the spin four operator 71
3.2.1 Spectrum of the theory 71

3.2.2 Divergence of the spin four current . . . 72
3.3 Structures for the three point functions . . . 75
3.4 Charge non-conservation identities . . . 78
3.4.1 General story . . . 78
3.4.2 Constraints on three point functions with non-zero even spins . . . 80
3.4.3 Closing the chain. The Quasi-fermion case. . . . 85
3.4.4 Closing the chain. The Quasi-boson case. . . . 87
3.4.5 Three point functions involving scalars. The quasi-fermion case. . . . 88
3.4.6 Three point functions involving scalars. The quasi-boson case. . . . 89
3.4.7 Comments about higher point correlation functions . . . 90
3.5 Final results . . . 91
3.5.1 Quasi-fermion theory . . . 92
3.5.2 Quasi-boson theory . . . 93
3.5.3 The critical point of $O(N)$ . . . 93
3.5.4 The critical $O(N)$ fermion theory . . . 95
3.5.5 Relation to other Chern-Simons matter theories . . . 96
3.5.6 Comments about higher dimensions . . . 96
3.5.7 A comment on higher order parity violating terms . . . 97
3.6 Conclusions and discussion . . . 98
3.7 Appendix A. Why the divergence of the currents should be conformal primaries . . . 100
3.8 Appendix B. Structure of the pseudo-charge conservation identities . . . 101
3.9 Appendix C. Compendium of normalizations . . . 102
3.10 Appendix D. Exploring the scalar sector for the quasi-fermion . . . 104
3.10.1 Fixing $\widetilde{c}_{\widetilde{0},\widetilde{0}}$ and nontrivial consistency check . . . 105
3.10.2 Fixing $\langle s\widetilde{0}\widetilde{0}\rangle$ . . . 106
3.10.3 Fixing $\gamma_{\widetilde{0}\widetilde{0}\widetilde{0}}$ . . . 107
3.11 Appendix E. Exploring the scalar sector for the quasi-boson . . . 107
3.11.1 Fixing $\widetilde{c}_{0,0}$ . . . 108
3.11.2 Fixing $\langle s00\rangle$ . . . 109
3.11.3 Vanishing of the double trace term in the $\langle 000\rangle$ identity . . . 110
3.12 Appendix F. Impossibility to add any further double trace deformations . . . 111
3.12.1 Case of quasi-fermion . . . 112
3.12.2 Case of quasi-boson . . . 113

3.12.3 Triple trace deformation 113
3.13 Appendix G. Motivational introduction 114

4 Convexity and Liberation at Large Spin 117

4.1 Introduction 117
4.2 Deep Inelastic Scattering 124
4.2.1 General Theory 126
4.2.2 Convexity 130
4.3 Crossing Symmetry and the Structure of the OPE 132
4.3.1 The Duals of the Unit Operator 133
4.3.2 An Argument Using the Light-cone OPE 135
4.3.3 Relation to Convexity 138
4.3.4 The Leading Finite Spin Correction 139
4.3.5 An Argument in the Spirit of Alday-Maldacena 143
4.4 Examples and Applications of (3.17) 144
4.4.1 Correction due to the Stress Tensor and Theories with Gravity Duals 145
4.4.2 The large Δ Limit 146
4.4.3 The Case of Parametrically Small $\Delta - \frac{\tau_{\min}}{2}$ 147
4.4.4 The Charge to Mass Ratio is Related to Convexity 148
4.5 Applications of Convexity 151
4.5.1 Free Theories and Weakly Coupled Theories 151
4.5.2 The Critical $O(N)$ Models 153
4.5.3 Strongly Coupled Theories and AdS/CFT 155
4.5.4 Higher Spin Symmetry Breaking 156
4.6 Conclusions and Open Problems 157
4.7 Appendix A. Normalizations and Conventions 159
4.8 Appendix B. Relating Spin to σ 160
4.8.1 A Naive Approach 160
4.8.2 A Systematic Approach 162
4.8.3 The Coefficient $c_{\tau_{\min}}$ in (3.12) 163
4.8.4 The Most General (Scalar) Case 164
4.9 Appendix C. OPE of Chiral Primary Operators in SCFT 166
4.10 Appendix D. DIS for Traceless Symmetric Representations 168

4.11 Appendix E. The $\frac{1}{s^2}$ Correction in the Critical $O(N)$ Model 169

5 Causality Constraints on Corrections to the Graviton Three-Point Coupling 171

5.1 Introduction/Motivation 171

5.2 Flat Space Causality and Shock Waves 174

5.2.1 Statement of Flat Space Causality 175

5.2.2 Scattering Through a Plane Wave in General Relativity 176

5.2.3 The Effect of Higher Derivative Interactions on Particles with Spin 179

5.3 General Constrains on the On-Shell Three-Point Functions 182

5.3.1 The Phase Shift in Impact Parameter Representation From Three-Point Functions 182

5.3.2 The Possible Forms of Three-Point Functions in Various Theories 186

5.3.3 Problems with Higher Derivative Corrections to the Three-Point Functions . 188

5.3.4 Scattering of Gravitons in $D > 4$ Dimensions 192

5.3.5 Scattering of Gravitons in Four Dimensions 193

5.4 Fixing the Causality Problem by Adding Massive Particles 194

5.4.1 Massive Spin Two Particles Do Not Fix the Problem in $D = 4$. 195

5.4.2 Massive Spin Two Particles Do Not Fix the Problem in $D > 4$. 197

5.4.3 Exciting the Graviton Into New Particles 199

5.4.4 Massive Higher Spin Particles Can Solve the Problem 200

5.4.5 Compositeness and the Extra Structures for Graviton Scattering 202

5.5 Anti-de Sitter Discussion 202

5.5.1 Motivation: The Emergence of Bulk Locality Should Happen in the Classical Theory 203

5.5.2 Statement of AdS Causality 203

5.5.3 The Effect of Higher Derivative Interactions on Particles with Spin 205

5.5.4 Implications for a, c in Theories with Large Operator Dimensions 207

5.5.5 Implications for Dimensions of Double Trace Operators 208

5.6 Cosmological Applications 212

5.7 Conclusions 213

5.7.1 Open Problems 215

5.8 Appendix A. Shapiro Time Delay 216

5.9 Appendix B. Three-Point Amplitudes and Their Sums 218

5.9.1 Scattering of a Scalar and a Graviton 220
5.10 Appendix C. The QED case 221
5.11 Appendix D. Causality and Unitarity for a Signal Model 222
5.12 Appendix E. Scattering in String Theory 225
5.13 Appendix F. Properties of the AdS Shock Wave 229
5.14 Appendix G. Time Advances and Time Machines 230
5.15 Appendix H. Representations That Couple to Two Gravitons 231

Chapter 1

Introduction

1.1 Gravity and Quantum

All phenomena observed in Nature so far are neatly described within the framework of existing physical theories. The range of phenomena covered by physical experiments is immense. It covers distances from $\sim 10^{-17}$m probed in the collisions of highly accelerated particles to $\sim 10^{25}$m coming from the deep sky surveys. In between there is a bounty of Biophysics, Chemistry and Condensed Matter Systems. Quite remarkably, on the theoretical side very few fundamental ideas are necessary to describe existing observations. Two monumental pillars of this description are Quantum Mechanics and General Relativity. Let us briefly review them from the viewpoint which is relevant for the ideas covered in this thesis.

Quantum Mechanics (QM) appeared in the beginning of the twentieth century as the theory describing atoms, light and their interaction. Quantum Field Theory (QFT) is a relativistic generalization of QM that describes the dynamics of elementary particles or, more precisely, fields. The basic difference between QM and QFT is that in QFT particles are created and destroyed. This makes the number of available degrees of freedom infinite and leads to different complications, like UV divergences, that were found in the 1930's and were overcome in the course of development of QFT and Particle Physics. The list of known elementary particles and their interactions are encoded in the few lines of the celebrated Standard Model Lagrangian [1]. A glorious march of the Standard Model was recently continued with the discovery of the Higgs boson at Large Hadron Collider at CERN [2, 3]. There are three main building blocks of the Standard model: force carriers or gauge bosons, matter fields and the Higgs responsible for the particle masses. Two characteristic scales of the Standard Model are: $\Lambda_{QCD} \sim 100$MeV, energy scale at which QCD becomes strongly couple; $m_{W,Z,H} \sim 100$GeV, energy scale at which Fermi theory of beta decays breaks down.

General Relativity (GR) is a relativistic generalization of Newtonian gravity put forward by Einstein. It postulates that spacetime geometry is dynamical and describes how matter affects geometry and vice versa. The characteristic gravitational scale is the Planck scale $M_{Pl} \sim 10^{19}$GeV and the gravitational coupling in the process of given energy is $\frac{\text{energy}^2}{M_{Pl}^2}$. This is why at short distances and energies of the modern days colliders (~ 10TeV) gravitational force is totally irrelevant. On the other hand, it is a dominant force at large, cosmological scales. The standard Big Bang cosmological model, the so-called ΛCDM, is nowadays a well-established experimental subject thanks mainly due to the extensive observations over the Cosmic Microwave Background.

From the old perturbative QFT point of view the main difference between the Standard Model and General Relativity is that the latter is non-renormalizable. It means that the theory has a lack of predictive power at Planckian energies when it becomes strongly coupled since the coupling constant becomes of order 1. In modern view, however, both the Standard Model and General Relativity are effective field theories. The notion of an effective field theory goes back to works by Wilson in the 1970's and tells us to think of energy as the parameter in the space of Lagrangians. From this point of view the correct physical theory has an onion-like structure, with different layers that correspond to different energies with potentially different degress of freedom being excited and governed by different Lagrangians.[1] From this point of view the non-renormalizability of gravity is the statement that General Relativity should be substituted by a new, more fundamental theory at energies of order, or above the Planck scale.

The problem of formulating a consistent quantum theory that includes as its low-energy limit both the Standard Model[2] and General Relativity is known as the Quantum Gravity (QG) problem. Because energies at which quantum-gravitational effects become significant are not directly accessible via existing experiments, the development of the subject is driven by thought experiments. The most famous example of this sort is the so-called "information paradox" articulated by Hawking in [4]. Hawking considered a process of black hole evaporation and argued that two things cannot be simultaneously true: a) physics in the low-curvature region is described by GR and QFT on the curved background; b) time evolution is unitary. He chose to abandon b) which is to modify Quantum Mechanics. The detailed understanding of the black hole evaporation process is still lacking, even though there is strong theoretical evidence that, at least, for the type of measurement that Hawking

[1] Of course, it could well be that at certain energies the description in terms of local Lagrangians breaks down. For example, this is what happens in String Theory.

[2] Or similar QFT.

described the usual rules of QM are valid.

Even though the subject of QG is largely of academic interest several months ago the first quantum gravity effect was potentially experimentally observed. Remarkably, the BICEP collaboration announced measurement of the B-mode polarization of the CMB background [5]. According to current understanding of the Universe history it could be an imprint of primordial quantum fluctuations of the spacetime metric and, thus, evidence for gravity being quantum.

A leading research program that addresses the problem of QG is String Theory. The basic assumption of String Theory is trivial to state: instead of relativistic structureless point particles we should consider one-dimensional extended relativistic strings. The consequences of this "minor" shift are breathtaking and still largely unexplored after forty years of intense research. Some of them include extra dimensions, supersymmetry, branes, dualities and holography. In this thesis we will mostly focus on the idea of holography and its celebrated incarnation known as AdS/CFT [6, 7, 8].

AdS/CFT[3] in its most general form is the statement that two systems are equivalent:

$$\text{AdS QG} = \text{CFT}, \tag{1.1}$$

Quantum Gravity in asymptotically anti-de Sitter space AdS_{d+1} and Conformal Field Theory (CFT) living on its Minkowski space boundary $\partial AdS_{d+1} = M^d$.[4] Notice that the AdS part of the equation (1.1) involves gravity, evaporating black holes and, in general, strings. Whereas the CFT part is a non-gravitational QFT in some cases very similar to QCD. There are many basic questions one can ask by looking at (1.1):

- Is (1.1) always correct? How to prove it?
- Are there extra observables in AdS QG which are not captured by CFT and vice versa?
- How does the extra dimension emerge and what is the mechanism for locality in AdS?
- How to describe the evaporation of black holes in AdS and their interior?
- What can be said about de Sitter space and its dual?
- For 74 extra questions check [10].[5]

[3]CFT stands for *Conformal Field Theory*, a QFT without any intrinsic scale.

[4]It was written in a similar form in the opening paragraph of [9] with commentary which in our opinion is the most eloquent expression of the excitement that surrounds AdS/CFT.

[5]Participants of the Strings 2014 conference were asked to pose an important question which they think can be answered within the next 5 years.

It is common in physics to study complicated phenomena by first solving simple toy models with extra symmetries and then analyzing deviations from the symmetric case. QG and AdS/CFT is clearly a very complicated problem and it would be very useful to have a simplified setting which allows one to explore it in a greater depth. Such setting does exist and was inspired by many years of impressive work by Vasiliev et al. that culminated in [24]; put forward in the context of AdS/CFT in the paper by Klebanov and Polyakov [11] and invigorated by a heroic computation of Giombi and Yin [12]. Extra simplifying symmetry is the so-called higher spin symmetry which is the main topic of this thesis and which we describe in detail below.

1.2 The Search of an Ultimate Symmetry

The idea of symmetry is an organizing principle in modern theoretical physics (see e.g. [13]). Different branches of physics can often be distinguished by the type of symmetry being assumed and exploited. A symmetry of fundamental importance shared by the Standard Model, General Relativity and their cousins is known as the Poincare group. The Poincare group is the isometry group of Minkowski space. It is a semidirect product of translations and Lorentz group. If we imagine a scattering experiment involving some ingoing and outgoing particles the fact that the underlying dynamics is Poincare-symmetric guarantees that the scattering matrix takes the following form

$$S = \delta^{(4)}(\sum_i p_i)\mathcal{A}(p_i.p_j). \tag{1.2}$$

One of the questions that people ask early on is: given a nontrivial scattering matrix what is the maximal group of symmetries it can have? Equivalently, how much can the Poincare group be extended without disallowing the interactions. As an illuminating example, imagine instead of energy-momentum conservation considering its following generalization

$$\sum_i p_i^{s-1} = 0 \tag{1.3}$$

with s being some integer $s \geq 2$. If we imagine this conservation to be generated by an integrated conserved current via the Noether theorem then the spin of the current is s.[6] At the intuitive level it is easy to see that such symmetries are not consistent with interactions. To demonstrate this let

[6]This is why symmetries of the $s > 2$ type are called *higher spin symmetries*.

us consider scattering of wave packets [14]

$$\psi_i(x) \sim \int\limits_{-\infty}^{\infty} dp\; e^{-\sigma(p-p_i)^2} e^{ip(x-x_i)}. \tag{1.4}$$

This wavefunction describes a particle with momentum being approximately p_i and position centered around x_i. Now let us act on this wavefunction with the operator $e^{i\alpha P^{s-1}}$. The result is

$$e^{i\alpha P^{s-1}} \psi_i(x) \sim \int\limits_{-\infty}^{\infty} dp\; e^{-\sigma(p-p_i)^2} e^{ip(x-x_i)} e^{i\alpha p^{s-1}}. \tag{1.5}$$

This wave function corresponds to the wave packet being shifted

$$x_i \to x_i + \alpha(s-1) p_i^{s-2}. \tag{1.6}$$

The condition (1.3) implies that the scattering amplitude is invariant when all wave packets are shifted according to (1.6). At this point the difference between $s = 2$ and $s > 2$ becomes very clear. For $s = 2$ we shift all the wave packets by the same amount and it is not surprising that it does not change the scattering amplitude. For $s > 2$ the relative position of the wave packets is changed and in interacting theory we expect the scattering amplitude to be changed as well! Thus, no interacting theories with the S-matrix having support only on configurations of the type (1.3) is expected to exist.

This simple argument was turned into a no-go theorem by Coleman and Mandula in 1967 [15]. They showed that the only extension of the Poincare group which is consistent with interactions is to add some internal symmetries.[7] It is often the case with the no-go theorems that they contain subtle loopholes that lead to new ideas. In the case of the Coleman-Mandula theorem there were several loopholes as well. One was that they did not consider fermionic symmetry charges and, thus, missed supersymmetry. This was corrected in the form of the so-called Haag-Lopuszanski-Sonhius theorem [16] which states that in addition to Poincare group and internal symmetries the interacting S-matrix can be supersymmetric. Another famous loophole is the existence of two-dimensional integrable models [14] which is to say that the no-go theorems mentioned above are valid if the number of spacetime dimensions is larger than two.

We can also imagine an interacting theory with some parameter which we call with a hindsight to be E such that when $E \to \infty$ the theory develops extra symmetries of the type (1.3). Then

[7] Familiar examples of this type are baryon number conservation or flavor symmetry.

according to the Coleman-Mandula theorem we expect that the S-matrix becomes trivial in this limit. Interestingly enough, this is exactly what happens in the high energy limit of scattering of strings [17]. In the high energy limit all modes of the string become effectively massless (the string becomes *tensionless*) and the theory supposedly develop new symmetries of the type we discussed such that

$$\lim_{E \to \infty} \mathcal{A}_i = 0, \tag{1.7}$$

where $\lim_{E\to\infty} \frac{\mathcal{A}_i}{\mathcal{A}_j}$ is finite and well-defined. From this point of view String Theory seems to be a spontaneously broken phase of some theory that has a lot of extra symmetries with the string length l_s being the order parameter for breaking of these extra symmetries. From this point of view massive modes of the string are similar to W, Z-bosons. The question of how to think about these extra symmetries that are developed in the high energy limit of scattering of strings in flat space and how to use them to constrain the S-matrix is still open (though see [18]).

There is yet another loophole in the argument of Coleman and Mandula. Especially after the discovery of AdS/CFT it was understood that there are many nontivial QFTs for which the notion of the S-matrix is ill-defined. Indeed, in a CFT it is not correct to think about excitations as being free when they are far from each other which is a basic premise for construction of the S-matrix. So one can wonder if it is possible to have nontrivial CFTs with higher spin symmetry and what are their AdS duals. Independently of AdS/CFT developments (and before them!) in a series of important works Vasiliev et al. (see e.g. [24]) found nontrivial gravitational theories in AdS that contain massless particles of all spins. Vasiliev theory is a type of theory that is expected to describe tensionless strings or the most symmetric phase of String Theory. According to AdS/CFT these should be dual to CFTs with extra symmetries of the type we discussed. Exploring AdS/CFT correspondence in this limit is the subject of the first half of the thesis.

If we assume that String Theory is the only framework for quantization of gravity and higher spin symmetry and higher spin particles are so ubiquitous in String Theory then what exactly is there role? Is it possible to argue based only on general principles that higher spin particles necessarily exist in our Universe if gravity is quantum? These are largely open and unexplored questions, however, there are few cases where these questions can be addressed and an important role of higher spin particles can be explicitly demonstrated. This is the topic of the second half of the thesis and is more relevant to our Universe since no elementary particles of spin larger than 2 were observed.[8]

[8]Strictly speaking, graviton which has spin 2 also was not observed yet.

Even though we mainly discussed higher spin symmetry and its breaking in the context of Quantum Field Theory, AdS/CFT and String Theory it is very important to emphasize that it has applications in many different areas of physics. In principle even a mundane perturbation theory around the free theory can be thought as a perturbation theory around the point where higher spin symmetry is unbroken. The problem is, of course, that in most of the cases it is not clear if this symmetry-based point of view simplifies the computation or puts some extra constraints on the expected results. There are many examples where broken higher spin symmetry can potentially play an important role like large N QCD, three-dimensional critical phenomena and condensed matter systems, cosmology. In our thesis we describe examples of such applications in the simplest possible setting.

1.3 Overview of the Thesis

As we briefly mentioned in the preceding section, this thesis explores different aspects of higher spin symmetry, its breaking and closely related topics in the context of AdS/CFT and String Theory. In the first part of the thesis we analyze cases when higher spin symmetry is present or is broken only slightly.

In chapter two we explore three-dimensional unitary conformal field theories that admits an additional conserved current of spin higher than two. We demonstrate that this implies existence of an infinite number of conserved currents. Using these currents we can construct charges and analyze Ward identities generated by them. As the result we show that correlation functions of the stress tensor and the conserved currents are then shown to be equal to those of a free field theory, a theory of N free bosons or free fermions. N is the central charge of the theory which is argued to be quantized. This analysis closes the loophole in the Coleman-Mandula theorem mentioned in the previous section by extending it to the theories that do not have an S-matrix.

In chapter three we consider the case when higher spin symmetry is broken quantum-mechanically. More precisely, we consider three-dimensional conformal field theories that have a large N limit, in the sense that the operators separate into single trace and multitrace and obey the usual large N factorization properties. We also assume that the spectrum of single trace operators are the higher spin currents plus an additional scalar. N plays the role of the Planck constant and we assume that the anomalous dimensions of the higher spin currents are of order $\hbar \sim 1/N$. Using the slightly broken higher spin symmetry we constrain the three point functions of the theories to leading order in N. As the result we get two families of solutions. Microscopically, one family can be realized as a theory of N fermions with an $O(N)$ Chern-Simons gauge field, the other as a N bosons plus the Chern-Simons gauge field. The family of solutions is parametrized by the 't Hooft coupling and as we change it the bosons are transformed into fermions which is an example of the 3d bosonization. At special parity preserving points we get the critical $O(N)$ models, both the Wilson-Fisher one and the Gross-Neveu one. For the dual Vasiliev theories our analysis fixes the three-point vertices of Vasiliev theory in AdS_4 or dS_4.

In the second part of the thesis we consider systems where higher spin symmetry is badly broken. We elucidate some aspects of this breaking and also clarify the role of the massive higher spin particles for the physical consistency of the theory.

In chapter four, we consider several aspects of unitary higher-dimensional conformal field theories. We first study massive deformations that trigger a flow to a gapped phase. Deep inelastic scattering in the gapped phase leads to a convexity property of dimensions of spinning operators of the original CFT. Via AdS/CFT it bounds masses of higher spin particles in the higher spin broken phase of the quantum gravity in AdS. We further investigate the dimensions of spinning operators via the crossing equations in the light-cone limit. Crossing equation is a defining nonperturbative equation for an abstract CFT. We manage to solve it analytically in a very particular limit. We find that, in a sense, CFTs become free at large spin and $1/s$ is a weak coupling parameter. The spectrum of CFTs enjoys additivity: if two twists τ_1, τ_2 appear in the spectrum, there are operators whose twists are arbitrarily close to $\tau_1 + \tau_2$. We characterize how $\tau_1 + \tau_2$ is approached at large spin by solving the crossing equations analytically. We find the precise form of the leading correction, including the prefactor. We compare with examples where these observables were computed in perturbation theory, or via gauge-gravity duality, and find complete agreement. The crossing equations show that certain operators have a convex spectrum in twist space. Applications include the 3d Ising model, theories with a gravity dual, SCFTs.

In chapter five, we consider the situation where the three-point coupling of the graviton is corrected compared to the in the Einstein theory. We assume that the theory is weakly coupled and analyzed constraints imposed on the theory by causality. We run a thought experiment that that involves high energy scattering of gravitons at fixed impact parameter. We find that in this experiment causality is violated. To fix the problem we can extra particles to the theory. We find that this violation cannot be fixed by adding conventional particles with spins $J \leq 2$. But, it can be fixed by adding an infinite tower of extra massive particles with higher spins, $J > 2$. This is true both in flat space and in AdS. In the context of inflation, or de Sitter-like solutions, we find that our constraint implies that if the gravity wave non-gaussianity deviates significantly from the one computed in the

Einstein theory it indicates the existence of massive higher spin particles. Notice that in String Theory breaking of higher spin symmetry occurs at the classical level. It means that interpolation between a gravitational theory of the Einstein-type and of the Vasiliev-type is a classical field theory question. Moreover, this suggests that in general higher derivative corrections to the Einstein theory should be related to the spectrum of higher spin particles. Chapter five explores this idea in the case of the simplest possible observable which is the graviton three-point coupling.

Chapter 2

Constraining Conformal Field Theories With a Higher Spin Symmetry

2.1 Introduction

The classic Coleman-Mandula result [15], and its supersymmetric extension [16], states that the maximum spacetime symmetry of a theory with an S-matrix is the super-Poincare group.[1] Interacting conformal field theories are interesting theories that do not have an S-matrix obeying the assumptions of [15, 16]. In this chapter we would like to address the question of whether a CFT can have a spacetime symmetry beyond the conformal group. We show that if a CFT has a conserved higher spin current, $s > 2$, then the theory is essentially free. Namely, all the correlators of the conserved currents are those of a free theory. In particular, this implies that the energy correlation function observables [19] of "conformal collider physics" [20] are the same as those of a free theory.

Let us clearly state the assumptions and the conclusions.

<u>Assumptions :</u>

- a) The theory is conformal and it obeys all the usual CFT axioms/properties, such as the operator product expansion, existence of a stress tensor, cluster decomposition, a finite number

[1] [16] also mentions the conformal group in the case of massless particles. However, [16] assumed that these massless particles are free in the IR, so that an S-matrix exists.

of primaries with dimensions less than some number, etc.

- a') The two point function of the stress tensor is finite.
- b) The theory is unitary.
- c) The theory contains a conserved current j_s of spin higher than two $s > 2$.
- d) We are in three spacetime dimensions.
- e) The theory contains unique conserved current of spin two which is the stress tensor.

The theorem, or conclusion :

There is an infinite number of even spin conserved currents that appear in the operator product expansion of two stress tensor. All correlation function of these currents have two possible structures. One is identical to that obtained in a theory of N free bosons, with currents built as $O(N)$ invariant bilinears of the free bosons. The other is identical to those of a theory of N free fermions, again with currents given by $O(N)$ invariant bilinears in the fermions.

Let us discuss the assumptions and conclusions in more detail.

We spelled out a') explicitly, to rule out theories with an infinite number of degrees of freedom, as in the $N = \infty$ limit of $O(N)$ vector models, for example. Unitarity is a very important assumptions since it allows us to put bounds on the dimensions of operators, etc. We assumed the existence of a higher spin current. One might wonder if one could have a symmetry which is not generated by a local current. Presumably, a continuous symmetry implies the existence of an associated current, but, to our knowledge, this has not been proven.[2] Assumption d) should hopefully be replaced by $d \geq 3$. Some of the methods used in this chapter have a simple extension to higher dimensions, and it should be straightforward to extend the arguments to all dimensions $d \geq 3$. In two dimensions, $d = 2$, we expect a richer structure. In fact, the Coleman-Mandula theorem in two dimensions allows integrable theories [21]. Also there are interesting current algebras in two dimensions which contain higher spin primaries beyond the stress tensor. An example are the W_N symmetry algebras [22]. The assumption of a unique stress tensor can also be relaxed, at the expense of making the conclusions a bit more complicated to state. It is really a technical assumption that simplifies the analysis. In fact, we actually generalize the discussion to the case that we have exactly two spin two conserved currents. In that case we have a kind of factorization into two subsectors and one, or

[2] Of course, if we have a theory given in terms of a Lagrangian, then Noether's theorem implies the existence of a current. Also if we assume we can generalize the action of the symmetry to the case that the parameter has some spacetime dependence, $\epsilon(x)$, then the usual argument implies the existence of an associated current, $j_\mu = \frac{\delta}{\delta \partial_\mu \epsilon}$, where the derivative is acting on the partition function, or the generating functional of correlation functions.

both, could have higher spin currents. We expect something similar for a larger number of spin two currents, but we did not prove it. A simple theory when there are two spin two conserved currents is the product of a free theory with a non-trivial interacting theory. In this case all the higher spin currents live in the free subsector.

Now, regarding the conclusions, note that we did not prove the existence of a free field operator ϕ, in the CFT. The reason why we could not do it is easily appreciated by considering a theory of N scalar fields where we restrict the operators to be $O(N)$ singlets. This is sometimes called "the free $O(N)$ model". This theory obeys all the assumptions of our theorem (as well as the conclusions!) but it does not have a free field in the spectrum. Short of establishing the existence of free fields, we will show that the theory contains "bilocal" operators $B(x_1, x_2)$ and $F_-(x_1, x_2)$ whose correlators are the same as those of the free field operators $\sum_{i=1}^N : \phi^i(x_1)\phi^i(x_2) :$ and $\sum_{i=1}^N : \psi_-^i(x_1)\psi_-^i(x_2) :$ in a theory of free fields. Our statements also concern the infinite number of even spin conserved currents that appear in the operator product expansion of two stress tensors. We are not making any statement regarding other possible odd spin conserved currents, or spin two currents that do not appear the operator product of two stress tensors. Such currents, if present, are probably also highly constrained but we leave that to the future.

Notice that we start from the assumption of certain symmetries. By imposing the charge conservation identities we obtained the explicit form of the correlators. Thus, we can view this as a simple example of a realization of the bootstrap program. Namely, we never used "the Lagrangian", we derived everything from physical correlation functions and physical consistency conditions. Of course, the result turned out to be rather trivial since we get free theory correlation functions. We can also view this as an exercise in current algebra, now for currents of higher spin.

Recently, there have been some studies regarding the duality between various $O(N)$ models in three dimensions and Vasiliev-type theories [23, 24, 25] in AdS_4 [11, 12, 26] (and dS_4 [27]). If one considers a Vasiliev theory with boundary conditions (at the AdS_4 boundary) which preserve the higher spin symmetry, then the theory obeys the assumptions of our theorem. Thus, our theorem implies that the theory is equivalent to a free theory of scalars or fermions. This was conjectured in [11] (see references therein for previous work), and tested in [12, 26, 28], see also [29, 30]. In this context the quantization of N implies that the coupling constant of a unitary Vasiliev's theory is quantized, if we preserve the higher spin symmetry at the AdS boundary.

It is also interesting to consider the same Vasiliev theory but with boundary conditions that do not preserve the higher spin charge. An example was proposed in [11] by imposing a boundary condition for the scalar that produced the "interacting" $O(N)$ theory, see also [28]. In such $O(N)$

models, the higher spin currents acquire an anomalous dimension of order $1/N$. Then the conclusions of our theorem do not apply. However, there are still interesting constraints on the correlators and we discuss some of them, as an example. It is likely that this method would give a way to compute correlators in this theory, but we will leave that for the future.

There is a large literature on higher spin symmetry, and we refer the reader to the review [31], which is also available in the arXiv [32, 33, 34, 35, 36, 37].

2.1.1 Organization of the chapter

In section two we discuss some generalities about higher spin currents.

In order to make the discussion clear, we will first present an argument that rules out theories with operators with low anomalous dimensions. This serves as a simple warm up example to the arguments presented later in the chapter for the more general case. This is done in section three.

In section four we discuss some general properties of three point functions of currents, which will be necessary later.

We then present two slightly different approaches for showing the main conclusion. The first, in section five, requires a slightly more elaborate construction but it is computationally more straightforward. The other, in section six, is conceptually more straightforward, but it required us to use Mathematica a lot. We have presented both methods, since they could be useful for other purposes (useful spinoffs). These two sections can be read almost independently, and the reader should feel free to choose which one to read or skip first.

In section seven we discuss the case of slightly broken higher spin symmetry. We just discuss a couple of simple points, leaving a more general analysis to the future.

In section eight we discuss the case of exactly two spin two conserved currents.

In section nine we present some conclusions and discussion.

We also included several appendices with some more technical results, which could also be useful for other purposes.

2.2 Generalities about higher spin currents

We consider higher spin local operators $J_{\mu_1 \cdots \mu_s}$ that transform in the representation of spin s of the rotation group. More specifically, if we consider the operator inserted at the origin, then it transforms in the spin s representation. It is convenient to define the twist, $\tau = \Delta - s$, where Δ is the scaling dimension of the operator. In three dimensions, the unitarity bound restricts the twist of a primary operator to [38]

$$\tau \geq \frac{1}{2}, \qquad \text{for} \qquad s = 0, \frac{1}{2} \tag{2.1}$$

$$\tau \geq 1, \qquad \text{for} \qquad s \geq 1. \tag{2.2}$$

The equality in the first line corresponds to free fields: bosons and fermions. While in the second line it corresponds to conserved currents $\partial_\mu J^\mu{}_{\mu_2\cdots\mu_s} = 0$.

From these conserved currents we can build conserved charges by contracting the current of spin s with an $s-1$ index conformal Killing tensor $\zeta^{\mu_1\cdots\mu_{s-1}}$, which obeys $\partial_{(\mu}\zeta_{\mu_1\cdots\mu_{s-1})} = 0$, where the parenthesis denote the symmetric and *traceless* part. This condition ensures that $\widehat{J}_\mu = J_{\mu\mu_1\cdots\mu_{s-1}}\zeta^{\mu_1\cdots\mu_{s-1}}$ is a conserved current so that the associated charge $Q = \int_\Sigma *\widehat{J}$ is conserved. A simple way to construct a conformal Killing tensor is to take a product of conformal Killing vectors [39]. In a CFT these charges annihilate the conformal invariant vacuum. These charges are conserved in the sense that their value does not depend on the hypersurface where we integrate the current. However, they might not commute with the Hamiltonian, since the Killing vectors can depend explicitly on the coordinates. This is a familiar phenomenon and it happens already with the ordinary conformal generators. These conserved currents lead to identities when they are inserted in correlation functions of other operators. Namely, imagine we consider an n point function of operators $\mathcal{O}_i$. Then we will get a "charge conservation identity" of the form

$$0 = \sum_i \langle \mathcal{O}_1(x_1) \cdots [Q, \mathcal{O}_i(x_i)] \cdots \mathcal{O}_n(x_n) \rangle \tag{2.3}$$

If we know the action of Q on each operator this leads to certain equations which we call "charge conservation identities".

In this chapter, we will focus exclusively on only one very particular charge that arises from a Killing tensor with only minus components. Namely, only $\zeta^{-\cdots-}$ is nonzero, and the rest of the components are zero. Here x^- is a light cone coordinate. We think of the three coordinates as[3] $ds^2 = dx^+dx^- + dy^2$. We denote the charge associated to the spin s current by

$$Q_s = \int_{x^+=\text{const}} dx^- dy \;\; j_s \;, \qquad\qquad j_s \equiv J_{s,-\cdots-} \tag{2.4}$$

We have also defined j_s to be the current with all minus indices. Also from now on we will denote

[3]Throughout the chapter we consider $x^\pm$ as independent variables.

by $\partial \equiv \partial_{x^-}$.

From now on we will consider the twist operator τ defined so that it is the anomalous dimension minus the spin in the $\pm$ directions (or boost generator in the $\pm$ direction). Note that Q_s has spin $s-1$. One can check that it has twist zero. This is a very important property of this particular Q_s charge which we will heavily use.

One more general property that we will need is the following. The fact that Q_s has a non-trivial dimension $(s-1)$ under the dilatation operator implies that

$$[Q_s, j_2] \propto \partial j_s + \cdots \ , \qquad\qquad s > 1 \tag{2.5}$$

(recall that $\partial = \partial_-$). This is indicating that the current j_s, from which we formed the charge Q_s is appearing in the right hand side of the commutator (2.5) . Here j_2 is the stress tensor of the theory. This is shown in more detail in appendix A. This fact is also related to the following. For any operator $\mathcal{O}$, the three point function $\langle \mathcal{O}\mathcal{O}j_2 \rangle$ is nonzero, where j_2 is the stress tensor. This should be non-zero because the stress tensor generates conformal transformations. In particular, the stress tensor should always be present in the OPE of two identical operators. Here we are using that the two point function of the stress tensor is finite. (The stress tensor comes with a natural normalization so that it makes sense to talk about the coefficient in its two point function).

Note that two point functions of conserved currents are proportional to

$$\langle j_s(x_1) j_s(x_2) \rangle \propto \frac{(x_{12}^+)^{2s}}{|x_{12}|^{4s+2}} \tag{2.6}$$

(Recall that our definition of j_s (2.4) contains only the minus components.)

2.3 Removing operators in the twist gap, $\frac{1}{2} < \tau < 1$

Notice that all operators with spin $s \geq 1$ should have twist bigger or equal to one. However, operators with spin $s = 0, \frac{1}{2}$ can have twists less than one. If $\tau = 1/2$, we have a free field which can be factored out of the theory. In this section we show that it is very easy to eliminate operators with twists in the range $\frac{1}{2} < \tau < 1$. We call this range the "twist gap". This will serve as a warm up exercise for the rest of the chapter.

To keep the discussion simple, let us assume that we have a conserved current of spin four, j_4. (We will show in the next section that a current of spin four is always present as soon as we have *any* higher spin current). Let us see how this current could act on a scalar operator of spin zero ϕ and twist in the twist gap, $\frac{1}{2} < \tau < 1$. The action of the charge Q_4, $[Q_4, \phi]$, preserves the twist. Q_4 cannot annihilate ϕ since the operator product expansion of $\phi\phi$ contains the stress tensor (the unique spin two conserved current), and Q_4 cannot annihilate the stress tensor due to (2.5) . The right hand side of $[Q_4, \phi]$ should be a combination of local operators and derivatives. We cannot have explicit functions of the coordinates since that would imply that Q_4 does not commute with translations. The fact that Q_4 commutes with translations can be seen from its integral expression (2.4) and the conservation of the current. The operator that appears in the right hand side has to have the same twist as ϕ. Thus, it should be a scalar. The only derivative that does not change the twist is $\partial = \partial_-$. This derivative should be present three times due to spin conservation. In conclusion, we have that

$$[Q_4, \phi] = \partial_-^3 \phi \qquad (2.7)$$

If we had many scalar operators of the same twist we would get $[Q_4, \phi_a] = c_{ab}\partial^3\phi_b$, where c_{ab} is symmetric due to the charge conservation identities for Q_4 acting on $\langle \phi_a \phi_b \rangle \propto \delta_{ab}$. Thus, we can diagonalize c_{ab} and we return to (2.7) .

As the next step we can write the charge conservation identity for the action of Q_4 on the four point functions of these fields. It is convenient to do it in momentum space. Then we have

$$[\sum_{i=1}^{4} (k_{i-})^3] \langle \phi(k_1)\phi(k_2)\phi(k_3)\phi(k_4) \rangle = 0 \qquad (2.8)$$

Using momentum conservation, and writing $k_i = k_{i-}$ for the minus component of the momentum, we find

$$\sum_{i=1}^{4} k_i^3 = -3(k_1 + k_2)(k_1 + k_3)(k_2 + k_3) = 0. \qquad (2.9)$$

when it acts on the four point function. Thus, the four point function is proportional to $\delta(k_{i_1} + k_{i_2})\delta(k_{i_3} + k_{i_4})$ and other two terms obtained by permutations. Together with rotational invariance and conformal invariance, this means that the four point function is a sum of two point functions

$$\langle \phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4) \rangle = \langle \phi(x_1)\phi(x_2) \rangle \langle \phi(x_3)\phi(x_4) \rangle + \text{permutations} \qquad (2.10)$$

Now we will show that the only operators that can obey (5.5) are free fields (see [40]). For that purpose we perform the ordinary operator product expansion as $|x_{12}| \to 0$. The first term in (5.5) gives the contribution of the identity operator. The other two terms are analytic around $x_{12} = 0$. Thus, they correspond to operators with twists $\tau_{int} = 2\tau + n$, with $n \geq 0$, where τ is the twist of ϕ. If this twist is $\tau > 1/2$, then the stress tensor, which has twist one, would not appear in this operator product expansion. But the stress tensor should always appear in the OPE of two identical currents. Thus, the only possible value is $\tau = 1/2$ and we have a free field. Such a free field will decouple from the rest of the theory.

Actually, by a similar argument we can eliminate spin 1/2 operators in the twist gap, $\frac{1}{2} < \tau < 1$. We repeat the above arguments for the particular component $\psi = \psi_-$ (the spin 1/2 component as opposed to the spin -1/2). All the arguments go through with no change up to (2.9) . That equation restricts the x^- dependence only. However, conformal symmetry and rotational invariance allow us to compute the four point function for any component ψ_α, and we find an answer that factorizes. Which then implies again $\tau = 1/2$ and a free fermion.

Note also that the analysis leading to (2.7) also constrains how all higher spin charges act on free fields, $[Q_s, \phi] = \partial^{s-1}\phi$, and similarly for fermions (we can show it easily for the ψ_- component and then the Dirac equation ensures that it acts in the same way on both components ψ_α[4]).

At this point we should mention a qualitative argument for the reason that we get free fields. If we form wavepackets for the fields ϕ centered around some momentum and somewhat localized in space, then the action of the charge Q_4 will displace them by an amount which is proportional to the square of their momenta. Thus, if the wavepackets were colliding in some region, then these displacements would make them miss. Here we use that $d \geq 3$.

The results of this section can be very simply extended to $d \geq 3$. There the twist gap exists in the region $\frac{d-2}{2} \leq \tau < d - 2$, with the lower bound corresponding to free fields.

Returning to $d = 3$, in what follows, we will mostly consider the twist one operators that correspond to conserved currents. The idea will be similar, we first constrain the action of the higher spin charges and then determine the correlation functions of the operators.

We will present two independent ways of doing this. The first involves the notion of light cone operator product expansions and it is a bit more direct. It is presented in section five. The second involves a more explicit analysis of three point functions and it is conceptually easier, but computationally more complicated (needs use of Mathematica). It is presented in section six. The reader can choose which one to read and/or skip. But first some further generalities.

[4] Here we again use that $[Q_s, P_\mu] = 0$.

2.3.1 Action of the charges on twist one fields

Before jumping to those sections, let us make some preliminary statements on the action of charges Q_s on twist one operators. Since Q_s has twist zero, the only things we can have in the right hand side are other twist one fields. These include other (or the same) twist one fields. Even if we had fields in the twist gap ($\frac{1}{2} < \tau < 1$), they cannot appear in the right hand side of a twist one current transformation, due to twist conservation and the fact that derivatives can change the twist, at most by integer amounts. Thus, we have the general transformation law

$$[Q_s, j_{s'}] = \sum_{s''=\max[s'-s+1,0]}^{s'+s-1} \alpha_{s,s',s''} \, \partial^{s'+s-1-s''} j_{s''} \tag{2.11}$$

(the limits of the sum are explained below). Here we wrote explicitly arbitrary constants $\alpha_{s,s',s''}$ appearing in front of each term, below we will assume them implicitly. We will also sometimes omit the derivatives. The sum (2.11) can involve twist one and $s = 0, 1/2$ fields, which are not conserved currents, but we still denote them by $j_0, \ j_{1/2}$. Here everything has minus indices which we have not indicated. The derivatives are also minus derivatives. The number of derivatives is easily fixed by matching the spin on both sides.

One easy property we can prove is the following, if current j_y appears in the right hand side of $[Q_s, j_x]$, then we should have that current j_x appears in the right hand side of $[Q_s, j_y]$. This follows by considering the action of Q_s on the $\langle j_x j_y \rangle$ two point function and (2.6) . In summary,

$$[Q_s, j_x] = \partial^{s+x-y-1} j_y + \cdots \quad \Rightarrow \quad [Q_s, j_y] = \partial^{s+y-x-1} j_x + \cdots \tag{2.12}$$

This has a simple consequence which is the fact that the spread of spins s'' in (2.11) is as indicated in (2.11) . The upper limit in (2.11) is obvious. The lower limit in (2.11) arises from the fact that the current $j_{s'}$ should appear in the right hand side of $[Q_s, j_{s''}]$, and the upper limit in this commutator, results in the lower limit in (2.11) .

2.4 Basic facts about three point functions

For our analysis it is important to understand the structure of the three point functions of conserved currents and other operators in $d = 3$.

This problem was analyzed in [41] (see also [42]). Unfortunately, it is not clear to us what was proven and what was the result of a case by case analysis with a later extrapolation (a "physics" proof). Thus, in appendix I we prove the results that will be crucial for section five. These only involve certain light-like limits of the correlators. For section six we need some particular cases with low spin, where the statements in [41, 42] can be explicitly checked.

Nevertheless, let us summarize the results of [41]. The three point function of the conserved currents has, generically, three distinct pieces

$$\langle j_{s_1} j_{s_2} j_{s_3} \rangle = \langle j_{s_1} j_{s_2} j_{s_3} \rangle_{boson} + \langle j_{s_1} j_{s_2} j_{s_3} \rangle_{fermion} + \langle j_{s_1} j_{s_2} j_{s_3} \rangle_{odd}. \tag{2.13}$$

Here the piece $\langle j_{s_1} j_{s_2} j_{s_3} \rangle_{boson}$ is generated by the theory of free bosons. The piece $\langle j_{s_1} j_{s_2} j_{s_3} \rangle_{fermion}$ is generated by the theory of free fermion and the piece $\langle j_{s_1} j_{s_2} j_{s_3} \rangle_{odd}$ is not generated by free theories. The boson and fermion pieces are known in closed form while the odd piece can be computed in each particular case explicitly by imposing the conservation of the currents. By considering particular examples it was observed that the odd piece is non-zero whenever the triangle rule is satisfied

$$s_i \leq s_{i+1} + s_{i+2}. \tag{2.14}$$

In appendix B we derive an integral expression for these odd correlators that naturally incorporates this triangle rule. We start with free higher spin conformal fields in four dimensions, introduce a conformal invariant perturbation, and look at correlators on a three dimensional slice. The correlators automatically obey the conservation condition and are conformal invariant. See appendix B for more details.

Another property of three point functions which we found useful is that any three point function of the two identical[5] currents with the third one that has an odd spin is zero

$$\langle j_s j_s j_{s'} \rangle = 0 \qquad \text{for } \; s' \; \text{ odd} \tag{2.15}$$

This is easy to check for the explicit expressions for the boson and fermion solutions. It is also can be seen from the integral representation for the odd piece.

2.4.1 General expression for three point functions

The space of three point functions of conserved currents is conveniently split into two parts according to their transformation under the parity. Practically, we call "even" ones the three point functions that do not contain the three dimensional epsilon tensor and "odd" the ones that do.

[5] We need literally the same current. It is not enough that they have the same spin.

The parity even correlation functions are generated by

$$\mathcal{F}_{even} = e^{\frac{1}{2}(Q_1+Q_2+Q_3)} e^{P_1+P_2} (b \cosh P_3 + f \sinh P_3) \tag{2.16}$$

where the piece proportional to b stands for correlation functions in the theory of free bosons and the one proportional to f for the ones in the free fermion theory.[6]. The P_i and Q_i are some cross ratios whose form can be found in [41] For currents which have indices only along the minus direction we can find the simple expressions

$$Q_i = \ell_i^2 (\widehat{x}^+_{i,i+1} - \widehat{x}^+_{i,i-1}) \ , \qquad P_i = \ell_{i+1} \ell_{i-1} \widehat{x}^+_{i-1,i+1} \ , \qquad \widehat{x}^+ = \frac{x^+}{x^2} \tag{2.17}$$

where ℓ_i are the parameters in the generating function. In other words the term with $\ell_1^{2s_1}$ gives us the current $j_{s_1}(x_1)$, etc.

When spins of the currents are integer (2.16) reproduces the results given in [41], but (2.16) also gives the answer for half integer spin currents.[7]

Due to the different behavior under $P_3 \to -P_3$ the bosonic and fermionic parts never mix inside the charge conservation identities for three point functions that we will consider. Thus, whenever we can solve the charge conservation identities they are satisfied separately for $f = 0$ and $b = 0$ parts.

2.5 Argument using bilocal operators

Here we present an argument for the main conclusion.

2.5.1 Light cone limits of correlators of conserved currents

We consider the light cone OPE limit of two conserved currents $j_s(x) j_{s'}(0)$. All indices are minus. We will consider the twist one contribution to the OPE.[8] This can be cleanly separated from the lower twist contribution which can only arise from the identity or spin $s = 0, \frac{1}{2}$ operators. This clean separation is possible because there is a finite number of operators in the twist gap, $\frac{1}{2} < \tau < 1$.

[6] An equivalent, but symmetric version of (2.16) is given by

$$\mathcal{F}_{even} = e^{\frac{1}{2}(Q_1+Q_2+Q_3)} e^{(P_1+P_2+P_3)} \times \left(b + f + \frac{b-f}{3} [e^{-2P_1} + e^{-2P_2} + e^{-2P_3}] \right)$$

[7] Note that for half integer spins the labels b and f just give two different possible structures, since, of course we need both bosons and fermions to have half integer spins

[8] Here we consider the contribution of twist *exactly* one. This should not be confused with discussions in weakly coupled theories where one is interested in the whole tower of operators which have twists close to one (or in four dimensions twist close to two). Such operators appear in discussions of deep inelastic scattering and parton distribution functions.

To extract the twist one contribution of a given spin, we can consider the three point functions $\langle j_s(x_1)j_{s'}(x_2)j_{s''}(x_3)\rangle$. We can also consider the contribution of all spins together by taking the following light-like limits.

We take the limit $x_{12}^+ \to 0$ first. Then we take the limit $y_{12} \to 0$. By twist conservation, the correlator could behave only like $1/|y_{12}|$ or like $1/y_{12}$, in this limit. The fermion piece vanishes when we take $x_{12}^+ \to 0$. One can see that the $1/y_{12}$ behavior can be produced only by a parity odd piece. This follows from the fact that we can put the third current at $y_3 = 0$, but at generic $x_3^\pm$. This still allows us to extract any possible twist one current that appears in the OPE limit. However, the parity transformation $y \to -y$ would change the sign of the correlator only if we have a $1/y_{12}$ term in the OPE. Below, the part that contains the $1/|y_{12}|$ behavior is called the "boson" piece. It comes from the boson piece of the three point functions.

Similarly, we can look at the piece that goes like x_{12}^+ in the limit that x_{12}^+ goes to zero. Then we extract the piece going like $\frac{x_{12}^+}{|y_{12}|^3}$.

More explicitly, these limits are defined by

$$\underline{j_s j_{s'}}_b = \left(\lim_{y_{12}\to 0^+} + \lim_{y_{12}\to 0^-}\right)|y_{12}| \lim_{x_{12}^+\to 0} j_s(x_1)j_{s'}(x_2) - (\text{lower}) \tag{2.18}$$

$$\underline{j_s j_{s'}}_f = \left(\lim_{y_{12}\to 0^+} + \lim_{y_{12}\to 0^-}\right) \lim_{x_{12}^+\to 0} \frac{|y_{12}|^3}{x_{12}^+}\left[j_s(x_1)j_{s'}(x_2) - \frac{1}{|y_{12}|}\underline{j_s j_{s'}}_b - (\text{lower})\right] \tag{2.19}$$

where we have indicated by an underline the fact that we take the light-like limit (and the subindex on the line reminds us which of the two limits we took). We will also indicate by an underline points that are light-like separated along x^-. In the fermion like limit (2.18) , we have indicated that we extracted the boson piece. However, in practice we will only use it when the boson piece vanishes. Thus, there is no ambiguity in extracting the limit. We also extract any possible twist less than one operators that could appear. There is a finite list of them. Thus, there is no problem in extracting the lower twist operators. In section three we have eliminated these lower twist operators by making the assumption of the existence of Q_4. Here we have not yet derived the existence of Q_4, so we had to argue that these lower twist operators do not affect the definition of the limits and the extraction of the twist one contribution to the OPE. In fact, from now on, these lower twist contribution play no role. After we show the existence of Q_4, we can then remove them using section three. We could define also a limit that extracts the odd piece, but we will not need it.

The three point functions simplify dramatically in this limit. For the boson part we have

$$\left(\lim_{y_{12}\to 0^+} + \lim_{y_{12}\to 0^-}\right) \lim_{x_{12}^+\to 0} |x_{12}| \langle j_{s_1} j_{s_2} j_{s_3}\rangle_b \propto \partial_1^{s_1}\partial_2^{s_2} \langle \underline{\phi\phi^*} j_{s_3}\rangle_{free} \ ; \tag{2.20}$$

$$\langle \underline{\phi\phi^*} j_{s_3}\rangle_{free} \propto \frac{1}{\sqrt{\widehat{x}_{13}\widehat{x}_{23}}} \left(\frac{1}{\widehat{x}_{13}} - \frac{1}{\widehat{x}_{23}}\right)^{s_3} \tag{2.21}$$

where $\langle \underline{\phi\phi^*} j_{s_3}\rangle_{free}$ stands for the correlation function of a free complex boson. The underline reminds us that the points are light-like separated, but there is no limit involved. The reason we take a *complex* boson is to allow for non-zero values when s is odd. We have displayed only the x_i^- dependence[9] and we have defined a slightly shifted version of the coordinates

$$\widehat{x}_1 = x_1^- \ , \qquad \widehat{x}_2 = x_2^- \ , \qquad \widehat{x}_3 = x_3^- - \frac{y_{13}^2}{x_{13}^+} \tag{2.22}$$

We should emphasize that in (2.20) the expression $\langle \phi\phi^* j_{s_3}\rangle_{free}$ denotes a correlator in a theory of a free complex boson. In particular, j_{s_3} is not the original current in the unknown theory, but the current of the free boson theory.

For a free boson theory, it is easy to see why (2.20) is true. For a free boson the currents are given by expressions like $\sum \partial^i\phi\partial^{s-i}\phi^*$. When we take the limit $x_{12}^+ \to 0$ only one term survives from the two sums associated to j_{s_1} and j_{s_2} (recall that $\partial = \partial_-$). The reason is that we need a contraction of the scalar with no derivatives in order *not* to bring down a factor x_{12}^+ in the numerator. That single term contains all the derivatives on the ϕ fields that are contracted with j_{s_3}.

Actually, without assuming the explicit forms given in [41], one can prove that (2.20) follows, in the light cone limit, from conformal symmetry and current conservation for the third current. Since this is a crucial point for our arguments, we provide an explicit proof of (2.20) in appendix I.

Using the same reasoning for the fermion part we obtain (see appendix I)

$$\left(\lim_{y_{12}\to 0^+} + \lim_{y_{12}\to 0^-}\right) \lim_{x_{12}^+\to 0} \frac{|x_{12}|^3}{x_{12}^+} \langle j_{s_1} j_{s_2} j_{s_3}\rangle_f \propto \partial_1^{s_1-1}\partial_2^{s_2-1} \langle \underline{\psi\psi^*} j_{s_3}\rangle_{free} \ ; \tag{2.23}$$

$$\langle \underline{\psi\psi^*} j_{s_3}\rangle_{free} \propto \frac{1}{(\widehat{x}_{13}\widehat{x}_{23})^{3/2}} \left(\frac{1}{\widehat{x}_{13}} - \frac{1}{\widehat{x}_{23}}\right)^{s_3-1} \tag{2.24}$$

where ψ stands for a free complex fermion. If $s = 0$, then $\langle \underline{\psi\psi^*} j_0\rangle_{free} = 0$, and the limit in the first line of (2.23) vanishes. This is simply the statement that in a free fermion theory we do not have a twist one spin zero operator.

[9] There is also a factor of $\frac{1}{x_{13}^+} = \frac{1}{x_{23}^+}$ in the right hand side of (2.20) .

As an aside, for the odd piece one can obtain a similar expression, given in appendix B. This will not be necessary here, since the higher spin symmetry will eliminate the odd piece and we will not need to know its explicit form.

2.5.2 Getting an infinite number of currents

Imagine we have a current of spin s. From (2.5) we know that $[Q_s, j_2] = \partial j_s + \cdots$. Then (2.12) implies that

$$[Q_s, j_s] = \partial^{2s-3} j_2 + \cdots \tag{2.25}$$

Let us first

$$\text{assume that} \quad \langle j_2 j_2 j_2 \rangle|_b \neq 0 \tag{2.26}$$

We now consider the charge conservation identity for Q_s acting on $\langle \underline{j_2 j_2}_b j_s \rangle$, where we mean a charge conservation identity that results from acting with Q_s on $\langle j_2 j_2 j_s \rangle$ and then taking the light cone limit for the first two variables. In other words, we get

$$0 = \langle \underline{[Q_s, j_2] j_2} j_s \rangle + \langle \underline{j_2 [Q_s, j_2]} j_s \rangle + \langle \underline{j_2 j_2} [Q_s, j_s] \rangle \tag{2.27}$$

In the first two terms we first expand the commutator using (2.11) and then take the light cone limit using (2.20) . We will consider first the case of integer s. Then the first two terms in (2.27) , produce

$$\langle \underline{[Q_s, j_2] j_2} j_s \rangle + \langle \underline{j_2 [Q_s, j_2]} j_s \rangle = \partial_1^2 \partial_2^2 \left[\gamma \partial_1^{s-1} + \delta \partial_2^{s-1} \right] \langle \underline{\phi \phi^*} j_s \rangle_{free} \tag{2.28}$$

Here the symmetry under $x_1 \leftrightarrow x_2$ implies that $\gamma = (-1)^s \delta$, due to the symmetries of (2.20) . In the third term in (2.27) , when Q_s acts on j_s, we generate many possible operators through (2.11) . Each of those currents might or might not have an overlap with $\underline{j_2 j_2}_b$. Thus, combining this with (2.28) we conclude that the charge conservation identity that results from acting with Q_s on $\langle \underline{j_2 j_2}_b j_s \rangle$ is

$$0 = \partial_1^2 \partial_2^2 \left[\gamma (\partial_1^{s-1} + (-1)^s \partial_2^{s-1}) \langle \underline{\phi \phi^*} j_s \rangle_{free} + \sum_{k=1}^{2s-1} \widetilde{\alpha}_k \partial_3^{2s-1-k} \langle \underline{\phi \phi^*} j_k \rangle_{free} \right] \tag{2.29}$$

where $\widetilde{\alpha}_k$ is the product of the constants in (2.11) and the actual value of the boson part of the three point function. The overall derivatives can be removed since the right hand side is a sum of terms of the form $\frac{1}{x_{13}^a x_{23}^b}$ where a and b are half integer, so if a term was non-zero before the derivative acted it would be non-zero afterwards, and terms with different powers cannot cancel each other. Notice that (2.29) involves the explicit functions defined in (2.20) . An important property of these functions is that $\langle \underline{\phi\phi^*} j_k\rangle_{free}$ has a zero of order k when $x_{12}^- \to 0$, and the ∂_3 derivative does not change the order of this zero. This implies that all the terms in the last sum in (2.29) are independent. Thus, all the $\widetilde{\alpha}_k$ are fixed. We know that $\widetilde{\alpha}_2$ is nonzero due to (2.25) and (2.26) . Here we used that the spin two current is unique so that there is no other contribution that could cancel the j_2 contribution. Given that this term is non-zero, then γ is nonzero. The $\widetilde{\alpha}_k$ for odd k are set to zero by the $x_1 \leftrightarrow x_2$ symmetry of the whole original equation, (2.27) . The rest of the $\widetilde{\alpha}_k$ are fixed, and are equal to what we would obtain in a free theory, which is the unique solution of (2.29) . It is then possible to check that all $\widetilde{\alpha}_k$, with $k = 2, 4, \cdots, 2s-2$ are non-zero. The explicit proof of this statement is given in appendix J.

In particular, this implies that the current j_4 is in the spectrum, and it is in the right hand side of the $\underline{j_2 j_2}_b$ OPE. This is true regardless of whether the original s is even or odd. Now we can go back to section three and remove the operators in the twist gap, $\frac{1}{2} < \tau < 1$.

Since $s > 2$, we find that $2s - 2 > s$, so that we are also finding currents with spins bigger than the original one. Thus, repeating the argument we find an infinite number of conserved currents.

We now show that a twist one scalar j_0 is present in the spectrum, and in the right hand side of $\underline{j_2 j_2}_b$. This can be shown by considering the Q_4 charge conservation identity acting on $\langle \underline{j_2 j_2}_b j_2\rangle$. This leads to an expression very similar to (2.29) , but with different limits in the sum (see (2.11)), such that now a non-zero $\widetilde{\alpha}_4$ implies the existence of j_0. Now, considering the ward identity from Q_4 on $\langle \underline{j_2 j_2}_b j_0\rangle$ we show that j_2 is in the right hand side of

$$[Q_4, j_0] = \partial^3 j_0 + \partial j_2 + \cdots \tag{2.30}$$

where the dots denote other twist one operators that have no overlap with $\underline{j_2 j_2}_b$ which could possibly appear. Let us now show that the fermion components are zero. We consider the charge conservation identity from Q_4 on $\langle \underline{j_2 j_2}_f j_0\rangle$. Notice that $\langle \underline{j_2 j_2}_f j_0\rangle = 0$, see (2.23) . Thus, this charge conservation identity implies that $\langle j_2 j_2 j_2\rangle_f = 0$. Here we have used that there is a unique stress tensor since we assumed that the j_2 appearing at various places is always the same. Thus, if $\langle j_2 j_2 j_2\rangle_b \neq 0$, then $\langle j_2 j_2 j_2\rangle_f = 0$. Conversely, if $\langle j_2 j_2 j_2\rangle_f \neq 0$, then $\langle j_2 j_2 j_2\rangle_b = 0$, and the discussion after (2.29) implies that also $\langle \underline{j_2 j_2}_b j_s \rangle = 0$ for all s.

In the case that $\langle j_2 j_2 j_2 \rangle_f$ is nonzero, we can start with a ward identity similar to (2.29) , but for $\langle \underline{j_2 j_2}_f j_s \rangle$. We find an identical conclusion, an infinite set of even spin currents in the right hand side of $\underline{j_2 j_2}_f$. (See appendix J for an explicit demonstration). In conclusion, due to the uniqueness of the stress tensor we have only one of two cases:

$$\langle j_2 j_2 j_2 \rangle_b \neq 0 \ , \quad \Rightarrow \langle j_2 j_2 j_2 \rangle_f = 0 \ , \ \ \underline{j_2 j_2}_b = \sum_{k=0}^{\infty} [j_{2k}] \ , \ \ \underline{j_2 j_2}_f = 0 \tag{2.31}$$

$$\langle j_2 j_2 j_2 \rangle_f \neq 0 \ , \quad \Rightarrow \langle j_2 j_2 j_2 \rangle_b = 0 \ , \ \ \underline{j_2 j_2}_f = \sum_{k=1}^{\infty} [j_{2k}] \ , \ \ \underline{j_2 j_2}_b = 0 \tag{2.32}$$

where the brackets denote currents and their derivatives. In other words, the brackets are the contribution of the conformal block of the current j_s to the twist one part of the OPE.

It is also possible to start with a half integer s, higher spin conserved current j_s and to obtain a charge conservation identity similar to (2.29) (we describe it in detail in appendix D). This again shows that currents with even spins $k = 2, 4, \cdots 2s - 1$ appear in the right hand side of $\underline{j_2 j_2}_b$ or $\underline{j_2 j_2}_f$. Again, if $s \geq 5/2$ we get higher spin conserved currents with even spins and we return to the previous case.[10] In the next two subsections we consider the first case in (2.31) and then the second.

2.5.3 Definition of bilocal operators

In the light-like limit we encountered a correlator which was essentially given by a product of free fields (2.20) . Here we will argue that we can "integrate" the derivatives and define a bilocal operator $B(x_1, x_2)$ with x_1 and x_2 separated only along the minus direction.

Now, let us make a comment on light cone OPE's. The light cone OPE of two currents $\underline{j_s j_{s'}}_b$, contains only twist one fields of various spins. The overlap with each individual current is computed by (2.20) . In particular, all such limits will contain a common factor of $\partial_1^s \partial_2^{s'}$. We can "integrate" such derivatives and define a quasi-bilocal operator $\widehat{B}(\underline{x_1, x_2})$ such that

$$\underline{j_s j_{s'}}_b = \partial_1^s \partial_2^{s'} \widehat{B}(\underline{x_1, x_2}) \tag{2.33}$$

The underline reminds us that x_1 and x_2 are null separated (along the x^- direction). Here the right hand side is simply a superposition of twist one fields and their derivatives. They are defined such

[10] If we have a half integer higher spin current, we will also have a supercurrent, a spin 3/2 current. This then requires that both $\langle j_2 j_s j_2 \rangle_b$ and $\langle j_2 j_2 j_2 \rangle_f$ are nonzero. Since we had the dichotomy (2.31) , we should have more than one spin two conserved current if there is any half integer higher spin current.

that

$$\langle \underline{j_s j_{s'}}_b j_{s''}(x_3)\rangle = \partial_1^s \partial_2^{s'} \langle \widehat{B}(\underline{x_1, x_2}) j_{s''}(x_3)\rangle \tag{2.34}$$

Of course, we also see that $\langle \widehat{B}(\underline{x_1, x_2}) j_s(x_3)\rangle \propto \langle \underline{\phi(x_1)\phi^*(x_2)} j_{s_3}(x_3)\rangle_{free}$. This implies that $\widehat{B}$ transforms as two weight 1/2 fields under conformal transformations.

A particular one we will focus on is the quasi-bilocal that we get from the stress tensor, defined by

$$\underline{j_2 j_2}_b = \partial_1^2 \partial_2^2 B(\underline{x_1, x_2}) \tag{2.35}$$

These are reminiscent of the (Fourier transform along x^- of the) operators whose matrix elements define parton distribution functions. One very important difference is that here we are constructing B *only* from the operators that have twist *exactly* equal to one.

We can define similar operators in the case of fermions

$$\underline{j_s j_{s'}}_f = \partial_1^{s-1} \partial_2^{s'-1} \widehat{F}_-(\underline{x_1, x_2}), \tag{2.36}$$

$$\underline{j_2 j_2}_f = \partial_1 \partial_2 F_-(\underline{x_1, x_2}) \tag{2.37}$$

Here F_- transforms as the product of two free fermions with minus polarizations $:\psi_-(x_1)\psi_-(x_2):$. It is defined again through (2.23) , by "integrating" the derivatives. It is the superposition of all the (even spin) currents that appear in the right hand side of $\underline{j_2 j_2}_f$.

We call these operators "quasi-bilocals" because they transform under conformal transformations as a product of two elementary fields. This does not mean that they are honest local operators, in the sense that their correlators have the properties of products of fields, with singularities only at the insertion of other operators. This is illustrated in appendix H. Our task will be to show that in a theory with higher spin symmetry, they become true bilocal operators, with correlators equal to the free field ones. The first step is to constrain the action of Q_s on B and F_- in (2.35) , (2.36) .

2.5.4 Constraining the action of the higher spin charges

Here we will show that

$$[Q_s, B(\underline{x_1, x_2})] = (\partial_1^{s-1} + \partial_2^{s-1}) B(\underline{x_1, x_2}) \tag{2.38}$$

Then we will show that this implies that their correlators have the free field form. Let us first assume that $\langle j_2 j_2 j_2\rangle_b$ is nonzero. In (2.38) we consider charges Q_s with s even constructed out of the currents j_s that appear in the right hand side of $\underline{j_2 j_2}_b$. As we saw before, there is an infinite number of such currents.

We would like to compute $[Q_s, B(\underline{x_1, x_2})]$. We can compute $[Q_s, \underline{j_2, j_2}_b] = \underline{[Q, j_2]j_2}_b + \underline{j_2[Q, j_2]}_b$. In other words, the action of Q_s commutes with the limit. This follows from the charge conservation identity of Q_s on $\langle j_2 j_2 j_k\rangle$ and taking the light cone limit, which gives

$$\langle\underline{[Q_s, j_2]j_2}j_k\rangle + \langle\underline{j_2[Q_s, j_2]}j_k\rangle = -\langle\underline{j_2 j_2}[Q_s, j_k]\rangle = \langle[Q_s, \underline{j_2 j_2}]j_k\rangle \tag{2.39}$$

So, we can write $[Q, j_2]$ in terms of currents and derivatives (with indices and derivatives all along the minus directions). Thus, in the end we can use the formula (2.34) to write

$$[Q_s, B(\underline{x_1, x_2})] = (\partial_1^{s-1} + \partial_2^{s-1})\widetilde{B}(\underline{x_1, x_2}) + (\partial_1^{s-1} - \partial_2^{s-1})B'(\underline{x_1, x_2}) \tag{2.40}$$

$\widetilde{B}$ contains all the even currents and it is symmetric under the interchange of x_1 and x_2 and B' contains the odd currents and it is antisymmetric under the interchange. Of course, the full expression is symmetric under this interchange.

Let us first show that B' is zero. Since B' is odd, it includes only odd currents $j_{s'}$ with s' odd. Imagine that B' contains a particular odd current $j_{s'}$. Then, if $s' > 1$, we consider the charge conservation identity coming from the action of $Q_{s'}$ on $\langle B' j_2\rangle$

$$0 = \langle[Q_{s'}, B']j_2\rangle + \langle B'[Q_{s'}, j_2]\rangle \tag{2.41}$$

$$0 = \gamma(\partial^{s'-1} - \partial^{s'-1})\langle\underline{\phi\phi^*}j_2\rangle + \sum_{k=0}^{s'+1}\widetilde{\alpha}_k\partial_3^{s'+1-k}\langle\underline{\phi\phi^*}j_k\rangle \tag{2.42}$$

This charge conservation identity is very similar, in structure, to the one we considered in (2.29) . Namely, the last term contains a sum over various currents which give rise to different functional forms. Therefore all $\widetilde{\alpha}_k$'s are fixed. $\widetilde{\alpha}_{s'}$ is nonzero because $[Q_{s'}, j_2]$ contains $j_{s'}$, (2.5) . The rest of the terms are fixed to the same coefficients that we would have if we were considering the same charge conservation identity in a free theory of a complex boson.[11] In particular, as shown in appendix J, $[Q_{s'}, j_2]$ produce a j_1 whose overlap with B' is nonzero. Once we have shown that j_1 appears in B',

[11] In (2.41) all α_k with even k are set to zero.

we can consider the charge conservation identity for Q_s on $\langle B j_1\rangle$.

$$0 = \langle [Q_s, B j_1]\rangle = (\partial_1^{s-1} - \partial_2^{s-1})\langle B' j_1\rangle + \langle B[Q_s, j_1]\rangle \tag{2.43}$$

Here the first term is non-zero. Analyzing this charge conservation identity, one can show that it can be obeyed, for a non-zero first term, only if $[Q_s, j_1]$ contains a j_s in the right hand side. The analysis of this charge conservation identity is somewhat similar to the previous ones and it is discussed in more detail in appendix J. Here we have used that Q_s is a current that appears in the right hand side of $\underline{j_2 j_2}_b$. In other words, Q_s is built out of the same current j_s that appears in the right hand side of $[Q_s, j_1]$. This implies that $\langle j_s j_s j_1\rangle$ would need to be non-zero.[12] However all such three point functions are zero when the two j_s currents are identical (2.15) . Since we reached a contradiction, we conclude that B' is zero.

We would like now to show that $\widetilde{B}$ is the same as B. We know that $\widetilde{B}$ is non-zero because we can consider the charge conservation identity for Q_s on $\langle B j_2\rangle$, and use that j_s appears in the right hand side of $[Q_s, j_2]$. This implies that $\langle \widetilde{B} j_2\rangle$ is nonzero. In fact, we can normalize it in such a way that j_2 appears in the same way on B and $\widetilde{B}$. Then we can consider $B - \widetilde{B}$. This does not contain a j_2. Here we are using that there is a unique j_2 in the theory. Say that $j_{s'}$, with s' even and $s' > 2$, is a candidate current to appear in the right hand side of $B - \widetilde{B}$. We can consider the $Q_{s'}$ charge conservation identity for $\langle (B - \widetilde{B}) j_2\rangle$. This charge conservation identity has a form similar to (2.41) ,

$$0 = \gamma(\partial_1^{s'-1} + \partial_2^{s'-1})\langle \phi\phi^* j_2\rangle + \sum_{k=0}^{s'+1} \widetilde{\alpha}_k \partial_3^{s'+1-k}\langle \phi\phi^* j_k\rangle \tag{2.44}$$

Here we are assuming that $\widetilde{\alpha}_{s'}$ is non-zero. However, this equation would also show that $\widetilde{\alpha}_2$ is non-zero which is now in contradiction with the fact that $B - \widetilde{B}$ does not contain j_2 (see appendix J). So we have shown that $B - \widetilde{B}$ cannot have any current of spin $s' > 0$.

Let us now focus on a possible spin zero operator j_0'. We put a prime to distinguish it from j_0 which is the one that appears in B. j_0' might or might not be equal to j_0. If $\langle j_0 j_0'\rangle \neq 0$, then we can consider the charge conservation identity for Q_4 acting on $\langle (B - \widetilde{B}) j_0\rangle$. Using (2.30) we get a non-zero term when Q_4 acts on j_0 producing j_0. However, this charge conservation identity cannot be obeyed if we are setting the term involving j_2 to zero. Thus, $\langle j_0 j_0'\rangle = 0$. Then there is some even

[12] Let's emphasize again we consider j_s that appears in the OPE of $j_2 j_2 \sim j_s$. Then, in principle, we can have $[Q_s, j_1] = \alpha j_s + \beta \widetilde{j}_s + \ldots$. The charge conservation identity (2.43) implies that α is non-zero. This implies that $\langle j_s j_s j_1\rangle$ must be non-zero.

current in the $\underline{j_2 j_2}_b$, call it s'', such that $[Q_s, j_{s''}] \sim j_0' + \cdots$. From (2.11) this implies that $s'' < s$. Then we consider the action of Q_s on $\langle (B - \widetilde{B}) j_{s''} \rangle$. This action produces (up to derivatives) both $\langle (B - \widetilde{B}) j_0' \rangle$ and $\langle (B - \widetilde{B}) j_2 \rangle$. But given that the second is zero, then the first is also zero. This charge conservation identity is very similar to the others we have been discussing (see appendix J). Here it is important that $s'' < s$, so that we get a constraint on $\langle (B - \widetilde{B}) j_0' \rangle$. So we conclude that there cannot be any j_0' in $B - \widetilde{B}$.

In conclusion, we have shown that (2.38) holds when x_1 and x_2 are null separated along the minus direction.

Now, once we argued that (2.38) is true, then we can consider any n point function of bilinears $\langle B(\underline{x_1, x_2}) \cdots B(\underline{x_{2n-1}, x_{2n}}) \rangle$. We have an infinite number of constraints from all the conserved charges (2.38) . These constraints take the form

$$\sum_{i=1}^{2n} \partial_i^{s-1} \langle B(\underline{x_1, x_2}) \cdots B(\underline{x_{2n-1}, x_{2n}}) \rangle = 0 \ , \qquad s = 2, 4, 6, \cdots \tag{2.45}$$

where $\partial_i = \partial_{x_i^-}$. This constrains the x_i^- dependence of this correlator. We show in appendix E that this constrains the correlator to be a sum of functions of differences $x_i^- - x_j^-$, $x_l^- - x_k^-$. More explicitly, the x_i^- dependence is such that the correlator is a sum of functions of the form $\sum_\sigma g_\sigma(x^-_{\sigma(1)} - x^-_{\sigma(2)}, \cdots, x^-_{\sigma(2n-1)} - x^-_{\sigma(2n)})$, where σ are the various permutations of $2n$ elements. In principle, these functions, g_σ are all different.

In addition, we know that the correlator should respect conformal symmetry and rotational invariance. Thus, it should be a function of conformal dimension $1/2$ at each location and a function of distances d_{ik}. Of course, $d_{ij}^2 = x_{ij}^+ x_{ij}^- + (y_i - y_j)^2$. Thus, we can now write the function in terms of distances. Let us explain this point a bit more. We have defined the bilocals in terms of an operator product expansion. But we have also noticed that we can also view these bilocals as special conformal block like objects written in terms of currents. From that perspective, the transformation properties under conformal transformations are identical to those of a product of bosonic fields. This is what we are using at this point. Thus, for each function g_σ, say $g(x_{13}^-, x_{24}^-, \cdots)$ we can now write $\widehat{g}(d_{13}, d_{24}, \cdots)$. But in addition, under conformal transformations, the function has to have weight one half with respect to each variable. Thus, it can only be a function of the form

$$\widehat{g}(d_{13}, d_{24}, \cdots) = \frac{1}{d_{13}} \frac{1}{d_{24}} \times (\cdots) \tag{2.46}$$

where the product runs over n distinct pairs of distances and each point i appears in one and only one of the terms. A distance between two points in the same bilocal is zero and cannot appear. So such terms are not present. In addition, the permutation symmetry under exchanges of B's and also of the two arguments of B imply that the overall coefficients of all terms are the same, up to disconnected terms, which are given by lower point functions of B's by cluster decomposition. For example, for a two point function of B's we have

$$\langle B(\underline{x_1, x_2})B(\underline{x_3, x_4})\rangle = \widetilde{N}\left(\frac{1}{d_{13}d_{24}} + \frac{1}{d_{14}d_{23}}\right) \tag{2.47}$$

The overall coefficients of the connected n point function of B's can be determined from the one in the $n-1$ point function by expanding one of the B's and looking at the stress tensor contribution, which is fixed by the stress tensor charge conservation identity. Therefore, all n point functions of B's are fixed up to a single constant, the constant in the two point function of the stress tensor, which is (up to a numerical factor) the same as $\widetilde{N}$ in (2.47) .

We can consider a theory of N free bosons, with

$$B(\underline{x_1, x_2}) = \sum_{i=1}^{N} \underline{:\phi_i(x_1)\phi_i(x_2):} \tag{2.48}$$

where the :: imply that we do not allow contractions between these two ϕ fields. This theory has a higher spin symmetry and we also get the same correlators, except that the constant in front of the stress tensor and (2.47) is $\widetilde{N} = N$. Another way to state the result is that all the correlators of the currents are the same as the ones we have in a theory of N bosons with N analytically continued, $N \to \widetilde{N}$. However, as we argue below $\widetilde{N}$ should be an integer.

We did not show that the operators for the elementary fields (or partons) are present as good operators in the theory. It is clear that one cannot show that by looking purely at correlators on the plane, since one can project them out by imposing an $SO(N)$ singlet condition which would leave all remaining correlators obeying the crossing symmetry relations, etc.

Let us discuss further lessons. First, we observe that the expansion of B's contains the currents with minus indices j_s. Thus, the three point function of B's contain the three point functions of currents with minus indices. All of these are the same as those of the free boson theory. Note that the possibility of odd three point functions has disappeared. Thus, we never needed to know much about the nature of the odd structures for the three point functions. From the OPE of two currents with minus indices we can get currents with arbitrary indices. Thus, from the six point function of B's we get the three point function of the currents for arbitrary indices, which coincides with the free boson answer. Similarly from a $2n$ point function of B's we get an n point function of currents with arbitrary indices which is equal to the free boson one for a theory of N bosons. By further performing OPE's we get other operators such as "double trace" or "double sum" operators. In conclusion, we have fixed all the correlation functions of the stress tensor, and also all the correlation function for all the higher spin conserved currents that appear in the $j_2 j_2$ operator product expansion. These are currents that have even spins. In addition, we have also fixed the correlators of all other operators that appear in the operator product expansion of these operators. All such correlators are given by the corresponding correlators in a theory of N free bosons restricted to the $O(N)$ invariant subsector.

We should remark that, though we made statements regarding currents that can be written as $O(N)$ invariant bilinears, this does not mean that the theory is the "free $O(N)$" model. The theory can contain additional conserved currents (and still only one stress tensor). For example, for $N = 2M$, we can consider M complex fields and restrict to the $U(M)$ invariant sector. This theory obeys all the assumptions of our theorem. In particular, it still has a single conserved spin two current. However, we also have additional currents of odd spin. Of course, the currents with even spin that appear in the $j_2 j_2$ OPE are still given by an $O(2M)$ invariant combination of the fields and have the same correlators that we discussed above. Probably a little more work would show that if we had odd spin currents, they should also behave like those of free fields. Presumably one would construct a bilocal operator from the odd currents and argue as we did above. One could also wonder about other even spin currents which do not appear in the $j_2 j_2$ OPE. Probably there cannot be such currents (with a single conserved spin two current), but we did not prove it.

Note that the correlation functions of stress tensors are all equal to the ones in the free theory. In particular, if we created a state with an insertion of the stress tensor at the origin we could compute the energy correlation functions that would be measured by idealized detectors (or calorimeters) at infinity. These are the energy correlation functions considered, for example in [19, 20], and are computed by particular limits and integrals of correlation functions of the stress tensor. The n point energy correlator for a state created by the stress tensor is schematically $\langle 0|T^\dagger(q)\epsilon(\theta_1)\cdots\epsilon(\theta_n)T(q)|0\rangle$, where $T(q)$ is the insertion of a stress tensor operator of four momentum (roughly) q at the origin and $\epsilon(\theta)$ are the energies per unit angle collected at ideal calorimeters sitting at infinity at the angle θ. These will give the same result as in the free theory. Namely, that the energy is deposited in two localized points, corresponding to the two partons hitting the calorimeters. This result is qualitatively similar to the Coleman-Mandula result for the triviality of the S-matrix. These energy correlation functions are infrared safe (or well-defined) observables which are, conceptually, rather close to the S-matrix. Here we see that these energy distributions are trivial.

2.5.5 Quantization of $\widetilde{N}$: the case of bosons

The basic idea for showing that $\widetilde{N}$ is quantized uses the fact that the $N = \infty$ theory and the finite N theory have a different operator spectrum. The finite N spectrum is a truncation of the infinite N spectrum. This point was also emphasized recently in [43].

In order to prove the quantization of $\widetilde{N}$ we argue as follows. In the theory of N bosons we consider the operator

$$\mathcal{O}_q = \delta^{[i_1,\cdots,i_q]}_{[j_1,\cdots,j_q]}(\phi^{i_1}\partial\phi^{i_2}\partial^2\phi^{i_3}\cdots\partial^{q-1}\phi^{i_q})(\phi^{j_1}\partial\phi^{j_2}\partial^2\phi^{j_3}\cdots\partial^{q-1}\phi^{j_q}) \tag{2.49}$$

where the δ function is the totally antisymmetric delta function of q indices. It is the object that arises we consider a contraction of two ϵ tensors of the from

$$\delta^{[i_1,\cdots,i_q]}_{[j_1,\cdots,j_q]} \propto \epsilon^{i_1,\ldots,i_q,i_{q+1},\cdots,i_N}\epsilon_{j_1,\cdots,j_q,i_{q+1}\cdots i_N} \tag{2.50}$$

This operator, (2.49) , can be rewritten as sum of products of q bilinear operators. Once we have written it as a particular combination of bilinear operators we can consider any value of N and we can imagine doing analytic continuation in $N \to \widetilde{N}$, with q fixed.

In particular, we can consider the norm of this state. We are interested in the $\widetilde{N}$ dependence of the norm of this state. We can show that

$$\langle\mathcal{O}_q\mathcal{O}_q\rangle = \widetilde{N}(\widetilde{N}-1)(\widetilde{N}-2)\cdots(\widetilde{N}-(q-1)) \tag{2.51}$$

where we only indicated the $\widetilde{N}$ dependence. Since we have q bilinears, the series expansion in $1/\widetilde{N}$ has only q terms. In addition, the result should vanish for $\widetilde{N} = 1, 2, \cdots, q-1$, and the leading power should be $\widetilde{N}^q$.

Now, imagine that $\widetilde{N}$ was not integer. Then we could consider this operator for $q = [\widetilde{N}] + 2$, where $[\widetilde{N}]$ is the integer part of $\widetilde{N}$. Then we find that (2.51) is

$$\langle\mathcal{O}_{[\widetilde{N}]+2}\mathcal{O}_{[\widetilde{N}]+2}\rangle = (\text{positive})(\widetilde{N}-[\widetilde{N}]-1) \tag{2.52}$$

where we only wrote the last term in (2.51) , which is the only negative one. Thus, we have a negative norm state unless $\widetilde{N}$ is an integer. Therefore, unitarity forces $\widetilde{N}$ to be integer. Here we have phrased the argument in terms of the norm of a particular state, which might be changed by choosing a different normalization constant. However, we can get the same argument by consider the contribution to the OPE of a state like $\mathcal{O}_q$. If the norm is negative, then we will get a negative contribution to the OPE in the channel that is selecting this particular $\mathcal{O}_q$.

2.5.6 Fermionic-like bilocal operators

We now return to the second case in (2.31) , where we have to use (2.36) . We can now repeat the operations we did for the bosonic case in order to argue that

$$[Q_s, F_-(\underline{x_1, x_2})] = (\partial_1^{s-1} + \partial_2^{s-1})F_-(\underline{x_1, x_2}) \tag{2.53}$$

The arguments are completely similar to the case of the boson. One needs to apply the same charge conservation identities. All the arguments are very similar, except that we now take the fermion like limit (2.18) and use the functions in (2.23) . One can run over all the arguments presented for the boson-like limit and one can check that all the charge conservation identities have the same implications for the fermion-like case. This is shown in appendix J.

Once we show (2.53) , we can now constrain the form of any correlator of the form $\langle F_-(\underline{x_1, y_1}) \cdots F_-(\underline{x_n, y_n})\rangle$. Again it involves functions of differences of x_{ij}^- with each i appearing only in one argument of the function. In addition, we know that they should be rotational invariant and conformal covariant. However, as opposed to the bosonic case, in this case the conformal transformations of F_- are those of a product of fermions. We can take into account these transformations by using factors of fermion propagators, x_{ij}^+/d_{ij}^3. Together with permutation (anti) symmetry these constraints imply that the correlators are those of free fermions. For example, for the four point function we get

$$\langle F_-(\underline{x_1, x_2})F_-(\underline{x_3, x_4})\rangle = \widetilde{N}\left(\frac{x_{13}^+}{d_{13}^3}\frac{x_{24}^+}{d_{24}^3} - \frac{x_{14}^+}{d_{14}^3}\frac{x_{23}^+}{d_{23}^3}\right) \tag{2.54}$$

Similarly we can use the symmetries to fix all n point functions of F in terms of the single parameter $\widetilde{N}$ that appears in (2.54) , or in the two point function of the stress tensor. For $\widetilde{N} = N$, these n point functions agree with the ones we would obtain in a theory of N Majorana fermions with

$$F_-(\underline{x_1, x_2}) = \sum_{i=1}^{N} \underline{:\psi_-^i(x_1)\psi_-^i(x_2):} \tag{2.55}$$

The expansion of F_- contains all currents of twist one with minus indices. The further OPE of such currents contains twist one currents with other indices and also a scalar operator of twist two.

In a theory of free fermions this is

$$\widetilde{j}_0 = \sum_{i=1}^{N} \psi_+^i \psi_-^i \qquad (2.56)$$

Thus, all the correlators of currents of even spin, plus the twist two scalar operator are fixed by the higher spin symmetry to be the same as in the theory of N free fermions.

2.5.7 Quantization of $\widetilde{N}$: the case of fermions

One can wonder whether there exists any theory $\widetilde{N}$ is non-integer. Since $\widetilde{N}$ appears in the two point function of the stress tensor, we know that $\widetilde{N} > 0$. We now argue for the quantization of $\widetilde{N}$. The argument follows from considering the operator

$$\mathcal{O}_q \equiv : (\widetilde{j}_0)^q : \qquad (2.57)$$

For general $\widetilde{N}$, this operator is defined by looking at the appropriate term in the operator product expansion of q $\widetilde{j}_0$ operators that are coming together. And $\widetilde{j}_0$ is itself also defined via a suitable operator product expansion of the original bilocals F_-. We can now compute the two point function of such operators, (2.57) . More precisely, since the two point function is arbitrarily defined, we can look at the contribution to the OPE of the exchange of the operator $\mathcal{O}_q$. By the way, notice that the correlation function of any of the current bilinears or any product of such bilinears, is given by making an analytic continuation in $N \to \widetilde{N}$ of the free fermion results. It is important that when we make this analytic continuation, the number of currents in a correlator should be independent of N.

A crucial observation is that in a free fermion theory with N fermions we have that (2.57) is zero for $q = N + 1$. This means that such a contribution should not be present in the OPE.

Now, let us fix q and compute the two point function of $\mathcal{O}_q$ for general $\widetilde{N}$. (More precisely, we are talking about the contribution of this operator to the OPE). The expression has a $1/\widetilde{N}$ expansion with precisely q terms, since that is the number of terms we get in a free fermion theory. We argue below that it should have the following $\widetilde{N}$ dependence

$$\langle \mathcal{O}_q \mathcal{O}_{\widetilde{q}} \rangle = \widetilde{N}(\widetilde{N}-1)(\widetilde{N}-2)\cdots(\widetilde{N}-(q-1)) \qquad (2.58)$$

First, we see that (2.58) has precisely q terms. Second, notice that it should vanish for $\widetilde{N}$ =

$1, 2, \cdots, q-1$, since for such integer values, we have correlators identical to those of the free fermion theory where $\mathcal{O}_q$ vanishes.

If $\widetilde{N}$ was not an integer, then we could set $q = [\widetilde{N}] + 2$, where $[\widetilde{N}]$ is the integer part of $\widetilde{N}$. Then the norm (2.58) would be equal to

$$\langle \mathcal{O}_{[\widetilde{N}]+2} \mathcal{O}_{[\widetilde{N}]+2} \rangle = (\text{positive})(\widetilde{N} - [\widetilde{N}] - 1) \tag{2.59}$$

where we only wrote the last factor in (2.58) . The rest of the factors are positive. However, this last factor is negative if $\widetilde{N}$ is not an integer. Thus, we see that unitarity forces $\widetilde{N}$ to be an integer and we get precisely the same values as in the free fermion theory.

2.6 Arguments based on more generic three and four point functions

In this section we derive some of the above results in a conceptually more straightforward fashion, which ends up being more computationally intensive. We checked the statements below by using Mathematica. Of course, if we were to take light-like limits some of the computations simplify, and we go back to the discussion in the previous section.

2.6.1 Basic operations in the space of charge conservation identities

The presence of higher spin symmetries leads to charge conservation identities for three point functions which relate three point functions of conserved currents of different spins. Namely, we start from a three point function and demand that it is annihilated by Q_s. This imposes interesting constraints for the following reason. The three point functions have a very special form due to conformal symmetry and current conservation. The action of the higher spin charge (2.11) gives a linear combination of these three point functions and their derivatives. These derivatives do not commute with the action of the conformal group, so this single equation is equivalent to larger set of equations constraining the coefficients of the various three point functions. In other words, since conformal symmetry restricts the functional form of the three point functions, the single equation that results from Q_s charge conservation is typically enough to fix all the relative coefficients of the various three point functions that appear after we act with Q_s on each of the currents.

Let us first describe some operations that we can use over and over again to constrain the action of the symmetries.

Imagine that we are trying to constrain the action of Q_s on a current j_x and we would like to know whether current j_y is present or not in the transformation law

$$[Q_4, j_x] = j_y + ... \tag{2.60}$$

(generically, with some derivatives acting on j_y). Recall that through (2.12) , then $[Q_4, j_y] = j_x + \cdots$. Now the basic charge conservation identity we can use is the one resulting from the action of Q_4 on $\langle j_2 j_x j_y \rangle$. First of all, notice that from the variation of j_2 we will necessarily get (see (2.5)) the term $\langle j_4 j_x j_y \rangle$ which must be non-zero if the j_x and j_y appear in each other transformations under Q_4. The simplest possibility would be to find that the only solutions of the charge conservation identity are such that $\langle j_4 j_x j_y \rangle = 0$. Thus, our assumption about the presence of j_y in the variation of j_x was wrong. This is the basic operation of the elimination of j_y from $[Q_4, j_x]$.

If we can find that solutions of the charge conservation identity with $\langle j_4 j_x j_y \rangle \neq 0$ exist, this is consistent with the presence of j_y in the variation of j_x but does not necessarily imply it.

Another basic charge conservation identity operation can be used to the check that j_y is definitely present in the transformation of j_x. This is done via the charge conservation identity resulting from the action of Q_4 on $\langle j_4 j_x j_x \rangle$. [13] Notice that from the transformation of j_4 we will necessarily (see (2.5) , (2.12)) get the term $\langle j_2 j_x j_x \rangle$ which must be non-zero due to the fact that j_2 generates conformal transformations. Now imagine that for the solution of the charge conservation identity to exist the term $\langle j_4 j_x j_y \rangle$ should be necessarily non-zero. This means that j_y is necessarily present in the transformation of j_x.

Very often using these two operations allow us to fix completely which operators do appear in the transformation of the given conserved current.

2.6.2 From spin three to spin four

In this section we show that in any theory that contains a conserved current of spin three then the conserved spin four current is necessarily present.

Let us first consider the most general CFT that has spin three current. We have that the most

[13] In practice, we first check that $[Q_4, j_2] \sim \partial^3 j_2 + \cdots$ and then we use $\langle j_2 j_x j_x \rangle$.

general transformation has the form

$$[Q_3, j_2] = \alpha_0 \partial^4 j_0 + \alpha_1 \partial^3 j_1 + \alpha_2 \partial^2 j_2 + \alpha_3 \partial j_3 + \alpha_4 j'_4, \tag{2.61}$$

$$[Q_3, j_3] = \beta_1 \partial^4 j'_1 + \beta_2 \partial^3 j_2 + \beta_3 \partial^2 j'_3 + \beta_4 \partial j_4 + \beta_5 j_5. \tag{2.62}$$

Let's make several comments on this expression. Primes stand for the fact that the same or a different current of the same spin can, in principle, appear in the right hand side. Notice that $\alpha_4 = 0$ since otherwise Q_3 is not translation invariant. Integrating both sides we would get $[Q_3, P_-] = Q_4$, which would mean that Q_3 is not translation invariant, in contradiction with (2.4) and current conservation. From (2.5) , α_3 should be non-zero, and so is β_2.

As the next step we consider the charge conservation identity obtained by Q_3 acting on $\langle j_2 j_2 j_3 \rangle$

$$0 = \langle [Q_3, j_2(x_1) j_2(x_2) j_3(x_3)] \rangle = 0 \tag{2.63}$$

This is essentially the same charge conservation identity we considered in section five. The solution exists only for $\beta_4 \neq 0$. In other words, the spin four current is necessary present in the theory.

Another feature of this exercise that is worth mentioning is that the general solution of the charge conservation identities involves three distinct pieces

$$\langle jjj \rangle = \langle jjj \rangle_{boson} + \langle jjj \rangle_{fermion} + \langle jjj \rangle_{odd} \tag{2.64}$$

the first two pieces correspond to the free boson and free fermion three point functions respectively. We expect these solutions to be present in the theory of complex boson and fermion. The odd piece does not come from any of free theories and is parity violating. [14]. We will elaborate the nature of the odd piece later. But we should emphasize that at the level of three point functions one can find three independent solutions of higher spin charge conservation identities [15]

After an illustration of this particular example we can formulate the general recipe. Assume that the theory has a conserved current j_s of spin $s \neq 4$ and $s > 2$. We again consider the charge conservation identity obtained by Q_s acting on $\langle j_2 j_2 j_s \rangle$ and arrive at the conclusion that j_4 must be present in the spectrum. This type of charge conservation identity is especially simple to analyze using the light cone limit described in the previous section.

So that we take it as a given that we always have a spin four current.

[14]Since it gives odd contribution to the three point function of stress tensors.

[15]We will show below that the odd solutions appear in the theories with higher spin symmetry broken at $\frac{1}{N}$ order .

2.6.3 Analysis of three point functions using the spin four current.

In this section we look in detail at the action of the Q_4 charge. Again we present only the results of the computations which are straightforward but tedious. We can argue that

$$[Q_4, j_2] = \partial^5 j_0 + \partial^3 j_2 + \partial j_4 \tag{2.65}$$

We have eliminated j_1 and j_3 by considering the charge conservation identity corresponding to the action of Q_4 on $\langle j_2 j_2 j_1 \rangle$ and $\langle j_2 j_2 j_3 \rangle$. We used general transformation laws for j_1, j_3 (2.11) . Then we consider the action of Q_4 at $\langle j_2 j_2 j_2 \rangle$ and use that we have already shown that $\langle j_2 j_2 j_4 \rangle$ is nonzero. The stress tensor three point function can have three different pieces

$$\langle j_2 j_2 j_2 \rangle = \langle j_2 j_2 j_2 \rangle_{boson} + \langle j_2 j_2 j_2 \rangle_{fermion} + \langle j_2 j_2 j_2 \rangle_{odd} \tag{2.66}$$

The charge conservation identity gives three different solutions involving other spins corresponding to these three different pieces. In other words, the charge conservation identity equations for the boson, fermion and odd pieces do not mix. However, if the boson or odd pieces in (2.66) are non-zero then the charge conservation identity implies that the current j_0 exists and appears as in (2.65) . The fermion solution does not require a j_0.

If the stress tensor is unique, then the fact that j_0 exists, implies that $\langle j_2 j_2 j_2 \rangle_f = 0$. This is done by considering the Q_4 charge conservation Identity for $\langle j_0 j_2 j_2 \rangle$. Q_4 on j_0 gives j_2 due to (2.12) and (2.65). The non-zero fermion part of these three point functions is not compatible with this charge conservation identity.

Then the problem separates into two problems. First we consider the case where $\langle j_2 j_2 j_2 \rangle_f = 0$ and then the case when $\langle j_2 j_2 j_2 \rangle_b = 0$. Both cannot be zero due the fact that j_2 generates conformal transformations for j_2, and the fact that the odd piece does not contribute to the action of conserved charges.[16] The odd piece can be eliminated at this point by an energy correlation argument, as shown in appendix C. But the reader not familiar with that technique can wait a little longer until we eliminate it in a more straightforward way.

[16]Consider $\int dx^- dy \langle j_{s_1}(x) j_{s_2}(x_2) j_{s_3}(x_3) \rangle_{odd}$ with $y_2 = y_3 = 0$ then under $y \to -y$ the integrand is odd and, thus, the integral vanishes.

2.6.4 Constraining the four point function of j_0

First we can constrain the action of Q_4 on j_0. By a method very similar to the one we used for the stress tensor we can prove that

$$[Q_4, j_0] = \partial^3 j_0 + \gamma \partial j_2 \tag{2.67}$$

where we wrote the constant γ appearing in the transformation explicitly. Again we eliminate j_1 and j_3 by considering the Q_4 charge conservation identities on $\langle j_2 j_1 j_0 \rangle$ and $\langle j_2 j_3 j_0 \rangle$. The other two terms can be found from the charge conservation identity for Q_4 acting on $\langle j_0 j_2 j_2 \rangle$, and using (2.65) and the fact that $\langle j_0 j_0 j_2 \rangle$ is nonzero.

Now we consider Q_4 acting on the four point function $\langle j_0 j_0 j_0 j_0 \rangle$. This gives

$$\partial_1^3 \langle j_0 j_0 j_0 j_0 \rangle + \gamma \partial_1 \langle j_2 j_0 j_0 j_0 \rangle + [1 \leftrightarrow 2] + [1 \leftrightarrow 3] + [1 \leftrightarrow 4] = 0 \tag{2.68}$$

In order to solve this we first need to write the most general four point function with one insertion of the stress tensor and three scalars, $\langle j_2 j_0 j_0 j_0 \rangle$. This can be done using the techniques described in [**?**, 42]. There are two possible forms, one is parity even and the other is parity odd. The equation (2.68) splits into two one for the parity even and the other for the parity odd piece. The general form for the parity even piece involves certain conformal invariants Q_{ijk} constructed out of the positions and the polarization tensors

$$\langle j_2(x_1) j_0(x_2) j_0(x_3) j_0(x_4) \rangle = \frac{Q_{123}^2 g(u,v) + Q_{124}^2 g(v,u) + Q_{134}^2 \widetilde{g}(u,v)}{x_{12}^2 x_{34}^2} \tag{2.69}$$

$$\widetilde{g}(u,v) = \frac{1}{u} g(\frac{v}{u}, \frac{1}{u}) \tag{2.70}$$

where the relevant conformal invariants simplify to

$$Q_{ijk} = \frac{x_{i,j}^+}{x_{i,j}^2} - \frac{x_{i,k}^+}{x_{i,k}^2} \;, \qquad u = \frac{x_{13}^2 x_{24}^2}{x_{12}^2 x_{34}^2} \;, \qquad v = \frac{x_{14}^2 x_{23}^2}{x_{12}^2 x_{34}^2} \tag{2.71}$$

because we consider only the minus polarization. We define $f(u,v)$ via

$$\langle j_0 j_0 j_0 j_0 \rangle = \frac{f(u,v)}{x_{12}^2 x_{34}^2} \tag{2.72}$$

$$f(u,v) = f(v,u) = \frac{1}{v} f(\frac{u}{v}, \frac{1}{v}) \tag{2.73}$$

Inserting (2.72) and (2.69) into (2.68) , we get an equation which depends both on the cross ratios and the explicit points x_i. By applying conformal transformations, and a lot of algebra, we can get a set of equations purely in terms f, g (and u and v). There is a solution where $g = 0$, which has a factorized dependence of the coordinates. There is also a solution with both g and f. The sum of these two solutions is

$$f(u,v) = \alpha(1+\frac{1}{u}+\frac{1}{v}) + \beta(\frac{1}{\sqrt{u}}+\frac{1}{\sqrt{v}}+\frac{1}{\sqrt{uv}}), \tag{2.74}$$

$$\gamma g(u,v) = \beta\frac{9}{20\sqrt{u}}. \tag{2.75}$$

By taking OPE of $\langle j_0 j_0 j_0 j_0\rangle$ we can extract, for example, the $\langle j_0 j_0 j_0\rangle$ structure constant. We can fix γ then by considering the j_2 charge conservation identity. In other words, integrating j_2 to get the action of P_- on $\langle j_0 j_0 j_0\rangle$, for example.

These are the free field theory correlators. The term proportional to α is the disconnected contraction and the one involving β is the connected one. In a theory of N free scalars we can set $\alpha = 1$ by a choice of normalization for the operators. Then $\beta \sim 1/N$. Notice that we were able to fix two four point functions using just one charge conservation identity.

2.6.5 No parity odd piece

While it is clear that we cannot write an odd piece for the four point function of scalars we can do it for $\langle j_2 j_0 j_0 j_0\rangle$. The unique structure in this case takes the following form in the embedding formalism (see [42] for conventions)

$$\langle j_2(x_1) j_0(x_2) j_0(x_3) j_0(x_4)\rangle \sim \frac{\epsilon(Z_1,P_1,P_2,P_3,P_4)}{(P_1P_2)^2(P_1P_3)^{3/2}(P_1P_4)^{3/2}(P_3P_4)^{1/2}} \tag{2.76}$$
$$[Q_{123}g_1(u,v) + Q_{134}g_2(u,v) + Q_{142}g_3(u,v)]$$

Inserting this in the charge conservation identity (2.68), $\sum \partial\langle j_2 j_0 j_0 j_0\rangle = 0$ we find that there is no solution.

By taking the OPE of $\langle j_2 j_0 j_0 j_0\rangle$ we can concentrate on the twist one channel where the stress tensor is propagating. The relevant three point functions are $\langle j_2 j_0 j_2\rangle$ and $\langle j_2 j_0 j_0\rangle$.

Both of them are non-zero. The function $\langle j_2 j_0 j_0\rangle$ is non-zero because stress tensor generates conformal transformation and $\langle j_2 j_0 j_2\rangle$ is non-zero because of the Q_4 charge conservation identity for $\langle j_2 j_2 j_2\rangle$.

Due to the triangle inequality (2.14) the $\langle j_2 j_0 j_2\rangle$ is the only odd three point function that is non-zero in the twist one sector OPE expansion of $\langle j_2 j_0 j_0 j_0\rangle$.

The fact that the four point function does not have the odd piece forces us to set the odd piece of $\langle j_2 j_0 j_2\rangle$ to zero. Then through the higher spin charge conservation identities we will set the odd part of the whole tower of conserved currents three point functions to zero. Thus, while the odd pieces of three point functions respect higher spin symmetry, at the level of four point functions they are eliminated in the theories where higher spin symmetry is exact.

2.6.6 Case of fermions

The above discussion can be repeated for the case that $\langle j_2 j_2 j_2\rangle_f \neq 0$. In this case we do not expect a twist one, spin zero field. Here we can show that

$$[Q_4, j_2] = \partial^3 j_2 + \partial j_4 \tag{2.77}$$

In principle, we could act with Q_4 on the four point function $\langle j_2 j_2 j_2 j_2\rangle$ and use (2.77) to fix it completely. Though this probably works, we have not managed to do it due to the large number of conformal structures that are possible.

Instead, one can take a longer route by first showing that a certain twist two scalar operator $\widetilde{j}_0$ exists, finding its transformation laws under Q_4 and then showing that its four point function is the same as that of $\widetilde{j}_0 = \epsilon^{\alpha\beta}\psi_\alpha\psi_\beta$ in a theory of free fermions.

In this section we consider not just the minus component of the currents but also the third, or perpendicular component. For the charge Q_4 we continue to focus on the all minus component.

We consider the charge conservation identity from the action of Q_4 on $\langle j_2 j_2 j_{2-\perp}\rangle$.

The new ingredient is the variation $[Q_4, j_{-\perp}]$ which should have twist two operators in the right hand side. We write the most general form of this variation denoting by $\widetilde{j}_s$ a twist two operator of spin s (and without tilde for the twist one ones)

$$[Q_4, j_{2\perp}] = \sum_{i=0}^{4} \partial^{4-i}(\widetilde{j}_i + \partial_\perp j_i + j_{(i+1)\perp}) \tag{2.78}$$

where all implicit tensor indices are minus.

Now as we are having in mind the Q_4 charge conservation identity for $\langle j_2 j_2 j_{2\perp}\rangle$ we can ignore all the odd spin twist one current due to (2.15). Moreover, we can fix $\widetilde{j}_{4-}$ to zero using the translation invariance argument. Also we know that for the fermion solution j_0 term is absent. And here we are interested in the fermion part of the solution for $\langle j_2 j_2 j_2\rangle$. One can be confused by the lack of perpendicular index in the twist two sector. It comes from an ϵ tensor as

$$\epsilon_{-\perp\mu}\partial^{\mu} = \epsilon_{-\perp+}\partial_{-} \tag{2.79}$$

, so that with more general indices we have

$$[Q_4, j_{2\mu\nu}] = \partial_-^2[\partial_\mu\epsilon_{-\nu\sigma} + \partial_\nu\epsilon_{-\mu\sigma}]\partial^\sigma\widetilde{j}_0 + ... \tag{2.80}$$

This epsilon tensor implies that in the charge conservation identity $\langle j_2 j_2 j_{2-\perp}\rangle_{even}$ should cancel the odd contribution $\langle j_2 j_2 \widetilde{j}_0\rangle_{odd}$. The relevant terms that we need for the charge conservation identity are

$$\begin{aligned}[Q_4, j_{2\perp}] = \partial^4\widetilde{j}_0 + \partial^2\widetilde{j}_2 + \partial^2\partial_\perp j_2 + \partial_\perp j_4 \\ + \partial^3 j_{2\perp} + \partial j_{4\perp} + ...\end{aligned} \tag{2.81}$$

The dots denote terms that do not have any overlap with $j_2 j_2$. So we end up with the following charge conservation identity to be checked. Here we think about all even parts as being generated by fermion three point functions

$$\begin{aligned}\langle[\partial^3 j_2] j_2 j_{2\perp}\rangle_{even} + \langle[\partial j_4] j_2 j_{2\perp}\rangle_{even} + [1 \leftrightarrow 2] = \langle j_2 j_2[\partial^4\widetilde{j}_0]\rangle_{odd} \\ + \langle j_2 j_2[\partial^2\widetilde{j}_2]\rangle_{odd} + \langle j_2 j_2[\partial^2\partial_{(-} j_{2\perp)}]\rangle_{even} + \langle j_2 j_2[\partial_{(-} j_{4\perp)}]\rangle_{even}.\end{aligned} \tag{2.82}$$

The only new three point functions involved are $\langle j_2 j_2 \widetilde{j}_0\rangle_{odd}$ and $\langle j_2 j_2 \widetilde{j}_2\rangle_{odd}$. Both are fixed by conformal symmetry, and current conservation for j_2, up to a constant. The charge conservation identity implies that the twist two scalar $\widetilde{j}_0$ needs to exist and it needs to appear in the right hand side of

$$[Q_4, j_{2\perp}] = \partial^4\widetilde{j}_0 + \partial^3 j_{2\perp} + \partial j_{4\perp} + ... \tag{2.83}$$

By performing an analysis similar to what was done for the bosonic case one can show that nothing else can appear in the right hand side of (2.83) . In addition, one can then constrain the action of

Q_4 on $\widetilde{j}_0$ and find

$$[Q_4, \widetilde{j}_0] = \partial^3 \widetilde{j}_0 + \partial \partial_{[-} j_{2\perp]} \tag{2.84}$$

This again requires looking at many charge conservation identities to eliminate all the other possible terms that could appear. Again, notice that the second term involves an ϵ tensor.

After this, one can look at the charge conservation identity for the four point function $\langle \widetilde{j}_0 \widetilde{j}_0 \widetilde{j}_0 \widetilde{j}_0 \rangle$. We will need $\langle J_{2\mu\nu} \widetilde{j}_0 \widetilde{j}_0 \widetilde{j}_0 \rangle$, actually, the parity odd version of this correlator, which has a structure similar to (2.76) except for the different conformal dimension of $\widetilde{j}_0$ (now two). After a lot of algebra one can show that this four point function has the form of free fermions (see Appendix G).

2.7 Higher spin symmetries broken at order $\frac{1}{N}$

So far we considered the case where the symmetries are exactly conserved. There are some interesting theories, mainly large N $O(N), Sp(N), U(N)$ vector models where the symmetries are "almost" conserved. By this we mean that the anomalous dimensions of the currents are of order $1/N$. Furthermore the correlators of the theory have $1/N$ expansion. We will make now some remarks about this case here.

To explain this imagine a situation when the higher spin currents have anomalous dimension $\Delta = s + 1 + O(\frac{1}{N})$. The operators have a single "trace" vs. a multitrace structure, and to leading order in N the correlators of multitrace operators factorize.[17]

Rather than giving a general discussion, here we will focus on one specific case that will illustrate the method, leaving a more general discussion for the future.

We can illustrate the method by considering a large N $SO(N)$ Chern-Simons matter theory with an $SO(N)$ level k Chern-Simons action coupled to N Majorana fermions. This theory was considered recently in [44]. [18] In this theory, in addition, we have a coupling $\lambda \sim N/k$ which is very small when the Chern-Simons level is very large, $k \gg N$. So we now have this second expansion parameter which we will also use.

In the limit that $\lambda = 0$ we have the theory of free fermions. The "single trace" operators in this theory are spanned by the twist one currents j_s and the twist two pseudoscalar operator $\widetilde{j}_0$. Let us normalize them so that their two point functions are one. Let us consider the most general form of

[17] We should rather speak of a "single sum" vs. a "multi-sum" structure.

[18] Actually, they considered a $U(N)$ gauge group, but the story is very similar. For a similar theory with scalars see [45].

the divergence of the J_4 current in this theory

$$\partial_\mu J^\mu{}_{---} = \frac{1}{\sqrt{N}}\left[a_1 \partial_- \tilde{j}_0 j_2 + a_2 \tilde{j}_0 \partial_- j_2\right]. \tag{2.85}$$

The divergence should be twist three operator with spin three. The terms in the right hand side are all the operators we can write down in this theory. Note that there are no single trace primaries that can appear. Notice that at this point we do not use any information about the microscopic theory except the spectrum at $N = \infty$.

We can now consider partially broken charge conservation identities, by integrating an insertion of (2.85) in a correlation function

$$\int d^3x \langle \partial_\mu J^\mu{}_{---} \mathcal{O}_1 \cdots \mathcal{O}_n \rangle = \sum_i \int_{S_i} \langle j^n{}_{---} \mathcal{O}_1 \cdots \mathcal{O}_n \rangle \tag{2.86}$$

Here we have used Stoke's theorem to integrate the divergence on a region which is the full R^3 minus little spheres around the insertion of each operator. The surface terms have the usual expression for the charges in terms of the currents. Except that now they are not conserved. However, since the non-conservation is a small correction, we can act on each of the operators by the charges, up to $1/N$ corrections. In the left hand side we can insert the expression (2.85) and use the leading N, factorized correlator. In our normalization the action of the "charges" , have a $1/\sqrt{N}$, so that left hand side and right hand side of (2.86) are of the same order.

As a first example, consider the case where the operators are simply $\mathcal{O}_1 \mathcal{O}_2 = \tilde{j}_0 j_2$. In this case, the charges, to leading order annihilate the correlator in a trivial way, using (2.84), (2.77) . All that remains is the integral of the right hand side of (2.85) . In order for this to vanish, we need that

$$a_2 = -\frac{2}{5} a_1. \tag{2.87}$$

So this relative coefficient is fixed in this simple way, for all λ, to leading order in $1/N$. This is a somewhat trivial result since it also follows from demanding that the special conformal generator K^- annihilates the right hand side of (2.85) . We have spelled it out in order to illustrate the use of the broken symmetry.

As a less trivial example, consider the insertion of the same broken charge conservation identity in the three point function of the stress tensor. We will do this to leading order in λ. We get

$$\sum_i \int_{S_i} \langle j^n_{---} j_2(x_1) j_2(x_2) j_2(x_3) \rangle \sim \frac{a_1}{\sqrt{N}} \int d^3x \langle [\partial \widetilde{j}_0 j_2 - \frac{2}{5} \widetilde{j}_0 \partial j_2](x) j_2(x_1) j_2(x_2) j_2(x_3) \rangle. \qquad (2.88)$$

Now let's take the large N limit in this equation. In the left hand side we can substitute the action of the charges on each of the operators. This gives

$$\langle [Q_4, j_2 j_2 j_2] \rangle \sim \frac{1}{\sqrt{N}} \left(\langle \partial^3 j_2 j_2 j_2 \rangle + \langle \partial j_4 j_2 j_2 \rangle + \text{action on the other } j_2\text{'s} \right) \qquad (2.89)$$

Notice that this is of order $\frac{1}{N}$.

In the right hand side of (2.88) the order $\frac{1}{N}$ terms come from

$$\int d^3x (\partial \langle j_2(x) j_2(x_1) \rangle) \langle \widetilde{j}_0(x) j_2(x_2) j_2(x_3) \rangle \sim \partial_1^5 \int d^3x \frac{1}{|x_1 - x|} \langle \widetilde{j}_0(x) j_2(x_2) j_2(x_3) \rangle \qquad (2.90)$$

where we have integrated by parts and used that all indices are minus ($\partial_1 = \partial_{x_1^-}$).

Now the final result of the integral in the right hand side of (2.90) is the same as the three point function $\langle j_0 j_2 j_2 \rangle$. Namely, the three point function involving a twist *one* scalar, as opposed to the twist two scalar $\widetilde{j}_0$ that we started with. This can be seen by the fact that the integral in (2.90) (before taking the ∂_1^5 derivative) has the same conformal properties as $\langle j_0 j_2 j_2 \rangle$.

If the current were exactly conserved, we would set (2.89) to zero. In that case, the only solution is the one corresponding to the free fermion structure in the three point functions. The reason is that the free boson, or odd solutions of this charge conservation identity require a twist one operator, j_0. Recall that we had said in section six that these three point function ward identities have three independent solutions involving only the fermion, or only the boson or only the odd structures. However, in (2.88) something remarkable has happened. The right hand side, which comes from the lack of conservation of the current, is mocking up perfectly the contribution we would have had in a theory with a twist one scalar. This allows for more nontrivial solutions which can be the superposition of all three structures for the three point functions.

Notice that if we consider now the charge conservation identity expanded in the 't Hooft parameter λ, then the λ^0 term will satisfy free fermion result. At the λ^1 the integral in (2.90) generates a result equal to the *odd* structure for $\langle j_0 j_2 j_2 \rangle$. Thus, at order λ^1, the odd structure for $\langle j_2 j_2 j_2 \rangle$ is generated.

We expect that further analysis along the lines explained above should fix all the leading N three point functions, exactly in λ.

We expect to be able to apply a similar ideas to the interacting fixed point of $O(N)$ model. In this case the spectrum of leading N theory is the same, except that the twist two operator is parity even.

Finally, let us comment on the case of weakly coupled gauge theories with matter in the adjoint. In this case we will typically have single trace operators that can, and will, appear in the right hand side of the divergence of the currents, once we turn on the coupling. In this case, we also expect to be able to use the broken charge conservation identities to analyze the theory, though it is not clear whether this will be any simpler than applying usual perturbation theory. One feature is that we will be dealing purely in terms of observables, without ever discussing things like "gauge fixing", etc.

2.8 Case of two conserved spin two currents

In this section we relax the assumption of a single stress tensor and we generalize the discussion to the case of two stress tensors. Presumably something similar will hold for more than two, but we have not studied it in detail.

We now consider a CFT that has exactly two symmetric traceless conserved spin two currents. One of them generates conformal transformations and we denote it by j_2. The usual minus generator built out of it is translation along the minus direction $Q_2 = P_-$. Another current we denote as $\widehat{j}_2$ and the correspondent charge would be $\widehat{Q}_2 = \widehat{P}_-$. We assume that these two currents are orthogonal and we normalize their two point functions as

$$\langle j_{2-}(x_1) j_{2-}(x_2)\rangle = \langle \widehat{j}_{2-}(x_1)\widehat{j}_{2-}(x_2)\rangle = c\frac{(x^+)^4}{(x^2)^{d+2}} \ , \qquad \langle j_{2-}(x_1)\widehat{j}_{2-}(x_2)\rangle = 0 \tag{2.91}$$

here we used the freedom to rescale $\widehat{j}_2$.

The most general form of the transformation consistent with the conformal properties of generators, two point function charge conservation identities and non-zero three point functions is

$$[P_-, j_2] = \partial_- j_2, \quad [P_-, \widehat{j}_2] = \partial_- \widehat{j}_2 \tag{2.92}$$

$$[\widehat{P}_-, j_2] = a\,\partial^2 j_1 + \partial \widehat{j}_2 \tag{2.93}$$

$$[\widehat{P}_-, \widehat{j}_2] = \partial j_2 + b\,\partial \widehat{j}_2 \tag{2.94}$$

We then consider the $\langle[\widehat{P}_-, j_2 j_2 \widehat{j}_2]\rangle = 0$ charge conservation identity, which sets $a = 0$.

Now, for any b we can introduce a new basis

$$J_2 = \frac{1}{2}\left(1 + \frac{b}{\sqrt{b^2+4}}\right) j_2 - \frac{1}{\sqrt{b^2+4}} \widehat{j}_2 \tag{2.95}$$
$$\widehat{J}_2 = \frac{1}{2}\left(1 - \frac{b}{\sqrt{b^2+4}}\right) j_2 + \frac{1}{\sqrt{b^2+4}} \widehat{j}_2 \tag{2.96}$$

such that the commutation relations take the form

$$[Q_{J_2}, J_2] = \partial J_2, \quad [Q_{\widehat{J}_2}, \widehat{J}_2] = \partial \widehat{J}_2, \tag{2.97}$$
$$[Q_{\widehat{J}_2}, J_2] = 0, \quad [Q_J, \widehat{J}_2] = 0. \tag{2.98}$$

In this new basis there are two orthogonal $\langle J_2 \widehat{J}_2\rangle = 0$ conserved spin two currents such that each of them generates translation for itself and leaves intact the other one.

We can now consider the correlation function with J_2's, $\widehat{J}_2$'s and an insertion of $e^{iQ_{J_2}a}$

$$\begin{aligned}\langle \widehat{J}_2(y_1)...\widehat{J}_2(y_n) J_2(x_1)...J_2(x_n)\rangle =& \langle e^{iQ_{J_2}a} \widehat{J}_2(y_1)...\widehat{J}_2(y_n) J_2(x_1)...J_2(x_n) e^{-iQ_{J_2}a}\rangle \\ =& \langle \widehat{J}_2(y_1)...\widehat{J}_2(y_n) e^{iQ_{J_2}a} J_2(x_1)...J_2(x_n) e^{-iQ_{J_2}a}\rangle\end{aligned} \tag{2.99}$$

where we used the fact that $[Q_{\widehat{J}_2}, J_2] = 0$. Then using the fact that $[Q_{J_2}, P_-] = 0$ and also the fact that $[Q_{J_2}, J_2] = [P_-, J_2]$ we can rewrite (2.99) as

$$\begin{aligned}\langle \widehat{J}_2(y_1)...\widehat{J}_2(y_n) J_2(x_1)...J_2(x_n)\rangle =& \langle \widehat{J}_2(y_1)...\widehat{J}_2(y_n) e^{iP_- a} J_2(x_1)...J_2(x_n) e^{-iP_- a}\rangle \\ =& \langle \widehat{J}_2(y_1)...\widehat{J}_2(y_n) J_2(x_1 + a^-)...J_2(x_n + a^-)\rangle.\end{aligned} \tag{2.100}$$

where all J_2's are translated along the minus direction. By the cluster decomposition assumption we get[19]

$$\langle \widehat{J}_2(y_1)...\widehat{J}_2(y_n) J_2(x_1)...J_2(x_n)\rangle = \langle \widehat{J}_2(y_1)...\widehat{J}_2(y_n)\rangle \langle J_2(x_1)...J_2(x_n)\rangle. \tag{2.101}$$

Thus, these correlators are completely decoupled. Notice that, in particular, if we consider collider physics observables like energy correlation functions for the "energies" defined from J_2 or $\widehat{J}_2$, we

[19] Let us make this more clear. We first choose the x_i and y_j to be spacelike separated and such that as $a \to +\infty$ the distances between any x_i and any y_j grow. One can check that such a choice of points is possible, and it still allows us to move all the points in a small neighborhood without violating this property. Once we establish the decoupling for such points, we can analytically continue the result for all points.

would obtain completely decoupled answers. However, we cannot conclude at this point that we have two decoupled theories, since we could still consider two theories that share a global symmetry and impose a singlet constraint on the operators. However, it is clear that in some sense, the two theories are dynamically decoupled. All that we have said so far is valid for any theory with two spin two conserved currents, independently of whether we have any higher spin generator.

2.8.1 Adding higher spin conserved currents

Now we consider a theory with a conserved higher spin current, j_s, and exactly two stress tensors. We know that $[Q_s, j_2] \sim j_s$, which implies that $[Q_s, j_s] = j_2 + \cdots$. Thus, the right hand side of $[Q_s, j_s]$ has either J_2 or $\widehat{J}_2$, or both. Let us first assume that it has J_2. Then the Q_s ward identity on $\langle J_2 J_2 j_s \rangle$ implies that there is an infinite number of currents in $\underline{J_2 J_2}_b$ or in $\underline{J_2 J_2}_f$. Now, nothing that appears in the right hand side of these light cone limits can have any non-zero correlator with $\widehat{J}_2$, due to (2.100) . Thus, from this point onwards, the analysis is effectively the same as if we had only one stress tensor. Of course, the same holds if $\widehat{J}_2$ appears in the right hand side of $[Q_s, j_s]$.

One simple example with two conserved spin two currents is a free $\mathcal{N} = 1$ supersymmetric theory with N bosons and N fermions, with an $O(N)$ singlet constraint. The two conserved spin two currents are the stress tensor of the boson and the one of the fermions.

The situation is more subtle when we have more than two spin two conserved currents. For example, in a theory with two free bosons, ϕ_1, ϕ_2, we have three conserved spin two conserved currents, schematically $\partial\phi_1\partial\phi_1$, $\partial\phi_1\partial\phi_2$ and $\partial\phi_2\partial\phi_2$. Clearly in this case the theory decouples into two theories, and not into three!

2.9 Conclusions and discussion

In this article we have studied theories with exactly conserved higher spin currents, with spin $s > 2$. We have shown that all correlators of the currents and the stress tensor are those of a free theory. More precisely, we have made the technical assumption of a single spin two current. Under this assumption, the only two cases are those of a free scalar or a free fermion. More precisely, we proved that the correlators of the currents are the same as in the theory of free bosons or free fermions, but we did not demonstrate the existence of a free scalar or a free fermion operator.

This is in the same spirit as the Coleman-Mandula theorem, [15, 16], extended here to theories without an S-matrix.

It can also be viewed as a simple exercise in the bootstrap approach to field theories. One simply starts with the currents, constructs the symmetries, and one ends up fixing all the correlation functions. One never needs to say what the Lagrangian is. Of course, the answer is very simple, because we end up getting free theories!

In this chapter we considered correlators of stress tensors and currents but not of other operators that the theory can have. In two dimensions, one can find explicitly the correlators of stress tensors, but that does not mean that the theory is free. On the other hand, in three dimensions we expect that the simple form of correlators that we obtained will imply that the theory is indeed free. In fact, one can consider the conformal collider physics observables that come from looking at energies collected by "detectors" at infinity for a state created by the stress tensor [19, 20]. The n point energy correlator for a state created by the stress tensor is schematically $\langle 0|T^\dagger(q)\epsilon(\theta_1)\cdots\epsilon(\theta_n)T(q)|0\rangle$, where $T(q)$ is the stress tensor insertion with four momentum (roughly) q at the origin and $\epsilon(\theta)$ are the energies per unit angle collected at ideal calorimeters sitting at infinity situated at the angle θ. These can be computed by considering $n+2$ correlation functions of stress tensors. Since these agree with the ones in the free theory, we expect that the theory is free. Notice also that these infrared safe observables are conceptually rather close to the S-matrix, since they are measured at infinity in Minkowski space. Nevertheless, one would like to be able to understand directly the constrains of the higher spin symmetry on other operators and their correlators. For example, it is natural to conjecture that their dimensions are integer or half integer (if the stress tensor is unique). As an example of a theory that contains an extra operator, we can consider N fermions with an $SO(N)$ singlet constraint and the operator $\epsilon^{i_1,\cdots,i_N}\psi_-^{i_1}\cdots\psi_-^{i_N}$. This can certainly not be produced from the currents when N is odd. For N odd this has a half integer dimension.

It should be simple to extend these results to four dimensional field theories. In fact, our approach that uses the light-like limits in section five is expected to work with few modifications. We expect that we will need light-like limits that pick out the free boson, free fermion and free Maxwell fields. In higher dimensions we can have other free fields, such as the self dual tensor in six dimensions.

Recently there have been many studies of a four dimensional higher spin gravity theory proposed by Vasiliev ([24] and references therein) . These theories have higher spin *gauge* symmetry in the bulk. They also have an AdS_4 vacuum solution. If one chooses AdS boundary conditions that preserve the higher spin symmetry, then our methods show that the theory is essentially the same as a free theory on the boundary. Thus, we have proven the conjecture in [11] for the free case. More concretely it is the theory of N free bosons or free fermions with an $O(N)$ singlet constraint. Given that the free case works, then [28] showed that the conjecture followed for the interacting case. Other arguments were presented in [29, 30]. If the boundary conditions break the higher symmetry in a slight way, then one can also use the symmetries to constrain the results. Notice also that our results for the quantization of $\widetilde{N}$ show that the coupling constant in a unitary Vasiliev theory with higher spin symmetric boundary condition is quantized.

It is also interesting that theories with almost conserved higher spin currents can be analyzed in this spirit. We discussed some simple computations in the case of large N vector models (recently studied in [44],[45]). Here one can use the slightly broken symmetries to get interesting results about the correlators. Of course, our previous analysis of the exact case is the backbone of the analysis for the slightly broken one. We suspect that correlators in these theories can be completely fixed by these considerations alone, and it would be nice to carry this out explicitly.

Notice that for any weakly coupled theory we have a slightly broken higher spin symmetry. In general, the breaking of the higher spin symmetry can occur at the single trace level. This occurs, for example, in gauge theories with adjoint fields. One might be able to perform perturbative computations using the higher spin currents and the pattern of symmetry breaking. Of course, one would recover the results of standard perturbation theory. However, the fact that we work exclusively with gauge invariant physical observables might lead to important simplifications, particularly in the case of gauge theories. Notice, by the way, that the bilocal operators that we have introduced in (2.33) (2.53) are simply the (Fourier transform) of the operators whose matrix elements are the parton distribution functions. In an interacting theory, these operators also have a Wilson line connecting the two "partons".

More ambitiously, one would like to understand large N limits of theories with adjoints, such as $\mathcal{N} = 4$ super Yang Mills, in the 't Hooft limit. In general, the single trace higher spin currents can have large anomalous dimensions compared to one. However, the anomalous dimensions are still small compared to N. It is natural to wonder whether they continue to impose interesting constraints. This is closely related to the possible emergence of a useful higher spin symmetry in string theory at high energies (see e.g. [46] and references therein).

Interestingly, one expects that the *absence* of single trace higher spin states in the low energy spectrum, namely the fact that all single trace higher spin operators have large anomalous dimensions, should impose an intriguing constraint: The theory has an AdS dual well approximated by Einstein gravity. For more discussion of this point see [47, 48, 40].

2.10 Appendix A. Argument for $[Q_s, j_2] = \partial j_s + \cdots$

We now argue that the action of Q_s on the stress tensor produces j_s in the right hand side, where j_s is the current we used to construct Q_s,

$$[Q_s, j_2] = \partial j_s + \ldots , \qquad\qquad s > 1 \tag{2.102}$$

This follows from the fact that we know how the charge Q_s transforms under the conformal group.

The general form of the commutator is

$$[Q_s, T_{\mu\nu}] = c\partial_- J_{--\ldots-\mu\nu} + \ldots. \tag{2.103}$$

Here we eliminated a possible term like $\partial_\mu J_{\nu-\ldots-}$ by applying ∂_μ to both sides of (2.103) and noticing that the left hand side, as well as the right hand side of (2.103) are zero, but the extra possible term would not be zero. Now, let us prove that c is not zero. Imagine it is zero, then contract (2.103) with dilaton Killing vector and integrate over $d^{d-1}x$ to get

$$[Q_\zeta, \int d^{d-1}x T_{0\nu} x^\nu] = [Q_\zeta, D] = 0 \tag{2.104}$$

which is inconsistent with the fact that $[Q_\zeta, D] = (s-1)Q_\zeta$. Note that for $s = 1$, the current does not appear in the right hand side. As usual the stress tensor should be invariant under global symmetries.

2.11 Appendix B. An integral expression for the odd three point functions

In this appendix we provide a "constructive" way to write it and a nice rationale for its existence. The idea is inspired by the AdS expressions for the odd pieces for spin one and spin two currents discussed in [49]. However, we will use something even simpler. We view the three point function as a correlator on a domain wall defect inside a (non-unitary) four dimensional conformal field theory. We construct it as follows. We start with four dimensional operators with spins $(j_l, j_r) = (s, 0),\ (0, s)$ at the four dimensional unitarity bound. We then introduce a quartic bulk interaction that preserves conformal symmetry. We then consider the correlator of three of these fields on a three dimensional subspace $t = 0$, to leading order in the interaction, see 2.1.

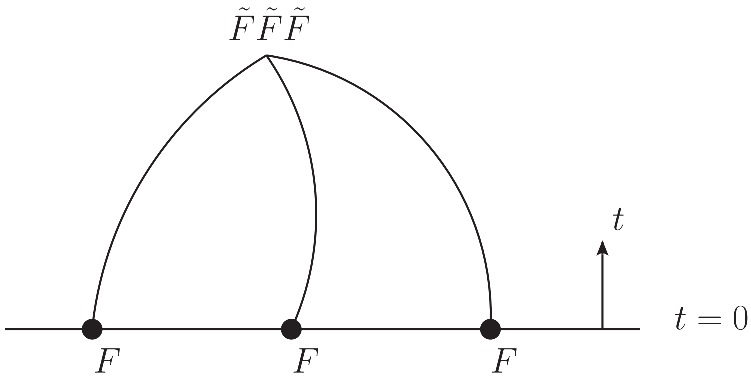


Figure 2.1: Constructing the three dimensional correlator as arising from a four dimensional integral in a CFT. The three dimensional space lies at $t = 0$.

We denote the four dimensional operators with spins $(j_l, j_r) = (s, 0),\ (0, s)$ by $F_{\alpha_1 \ldots \alpha_{2s}},\ \widetilde{F}_{\dot{\alpha}_1 \ldots \dot{\alpha}_{2s}}$. Of course, these operators are at the four dimensional unitarity bound and given in terms of free fields. The conformal dimension is $[F] = s + 1$ which is the same as dimension of conserved currents in $d = 3$ CFT. Their correlators are

$$\langle F_{\alpha_1 \ldots \alpha_{2s}}(x_1) \widetilde{F}_{\dot{\alpha}_1 \ldots \dot{\alpha}_{2s}}(x_2) \rangle = \frac{\prod_{i=1}^{2s} (x_{12})_{\alpha_i \dot{\alpha}_i}}{(x_{12}^2)^{2s+1}} \tag{2.105}$$

we can contract indices with polarization spinors λ^{α} and $\overline{\mu}^{\dot{\alpha}}$ so that we get

$$\langle F(x_1, \lambda) \widetilde{F}(x_2, \overline{\mu}) \rangle = \frac{(\lambda \not{x}_{12} \overline{\mu})^{2s}}{(x_{12}^2)^{2s+1}} = (\lambda \not{\partial}_{x_{12}} \overline{\mu})^{2s} \frac{1}{x_{12}^2}. \tag{2.106}$$

Now let's consider the self dual interaction of the form

$$g \int dt d^3 \vec{z} \chi \widetilde{F}_{s_1} \widetilde{F}_{s_2} \widetilde{F}_{s_3} \tag{2.107}$$

where all indices are contracted using $\epsilon^{\dot{\alpha}\dot{\beta}}$. Here χ is a scalar operator of dimension $[\chi] = 1 - (s_1 + s_2 + s_3)$. The contraction of indices is possible when the triangle inequality is satisfied

$$s_i \leq s_{i+1} + s_{i+2}, \quad i = 1, 2, 3 \mod 3 \tag{2.108}$$

This interaction breaks parity in the bulk so that we can expect that it will generate parity breaking structure for the correlators at $t = 0$, where t is the fourth coordinate.

Now imagine we are computing correlation functions of three $F_{\alpha_1 \cdots \alpha_{2s}}$ inserted at the boundary $t = 0$. There is a subset of four dimensional conformal transformations which map this boundary to itself. They act on this boundary as the three dimensional conformal transformations. From that point of view the $F_{\alpha_1\cdots\alpha_{2s}}$ operators transform as three dimensional twist one operators of spin s. The contraction with the spinor λ_α gives us the same as contracting three dimensional currents with the three dimensional spinors, as in [41]. In addition, one can see that they are conserved currents from the three dimensional point of view.[20] For this purpose it is enough to show that the propagator is annihilated by three dimensional divergence operator $\partial_\lambda \not\partial_x \partial_\lambda$. The structure of (2.106) shows that it is enough to check it for $s = 1$. For $s = 1$ the current is just $j_i = F_{ti} + \epsilon_{ijk}F_{jk}$. Then the divergence is zero due to the Maxwell equation and the Bianchi identity.

We now postulate that χ in (2.107) has expectation value $\langle\chi\rangle = t^{-\Delta_\chi} = t^{s_1+s_2+s_3-1}$. This expectation value is consistent with the three dimensional conformal invariance, and can be viewed as arising from a "domain wall defect" at $t = 0$.

We can now consider the expression that results from considering the three point correlator at first order in the interaction (2.107)

$$\langle F_{s_1}(\vec{x}_1,\lambda_1)F_{s_2}(\vec{x}_2,\lambda_2)F_{s_3}(\vec{x}_3,\lambda_3)\rangle \sim \int_0^\infty dt d^3\vec{x}_0 t^{s_1+s_2+s_3-1} \tag{2.109}$$

$$\frac{(\lambda_1 \not x_{10} \not x_{02}\lambda_2)^{(s_1+s_2-s_3)}(\lambda_1 \not x_{10} \not x_{03}\lambda_3)^{(s_1+s_3-s_2)}(\lambda_2 \not x_{20} \not x_{03}\lambda_3)^{(s_2+s_3-s1)}}{(t^2+(\vec{x}_1-\vec{x}_0)^2)^{2s_1+1}(t^2+(\vec{x}_2-\vec{x}_0)^2)^{2s_2+1}(t^2+(\vec{x}_3-\vec{x}_0)^2)^{2s_3+1}}$$

The operators in the left hand side are inserted at $t = 0$. They can be viewed as three dimensional conserved currents.

This is a conformal invariant expression for three dimensional correlators of conserved currents. It contains both even and odd contributions.

Notice that factors in the numerator of (2.109) have the schematic form

$$(\lambda_1 \not x_{i0} \not x_{0j}\lambda_2) \sim t(\vec{n}\vec{x}_{12}) + \epsilon_{ijk}n^i x_{01}^j x_{02}^k \tag{2.110}$$

where $\vec{n} = \lambda_1^\alpha \vec{\sigma}_{\alpha\beta}\lambda_2^\beta = \lambda_2^\alpha \vec{\sigma}_{\alpha\beta}\lambda_1^\beta$. We see that these two pieces behave differently under three dimensional parity. Thus, to extract the odd piece we can subtract parity conjugated integral, which can also be extracted by changing $t \to -t$. In the end one can show that the parity odd piece

[20] This needs to be checked separately because the interaction (2.107) is not obviously unitary and, thus, we cannot use the unitarity bound argument.

is given by extending the range of integration

$$\langle j_{s_1}(\vec{x}_1,\lambda_1)j_{s_2}(\vec{x}_2,\lambda_2)j_{s_3}(\vec{x}_3,\lambda_3)\rangle_{odd} \sim \int\limits_{-\infty}^{\infty} dt d^3\vec{x}_0 t^{s_1+s_2+s_3-1} \tag{2.111}$$
$$\frac{(\lambda_1 \not{x}_{10} \not{x}_{02}\lambda_2)^{(s_1+s_2-s_3)}(\lambda_1 \not{x}_{10}\not{x}_{03}\lambda_3)^{(s_1+s_3-s_2)}(\lambda_2\not{x}_{20}\not{x}_{03}\lambda_3)^{(s_2+s_3-s_1)}}{(x_{10}^2)^{2s_1+1}(x_{20}^2)^{2s_2+1}(x_{30}^2)^{2s_3+1}}.$$

This formula is valid both for integer and half integer spins. If $s_1 = s_2 = s$, then one can check that under the exchange $(\lambda_1, x_1) \leftrightarrow (\lambda_2, x_2)$ we get a factor of $(-1)^{s_3+2s}$. This implies that the correlator for two identical currents $\langle j_s j_s j_{s_3}\rangle$ is zero if s_3 is odd. Note that this is also zero when s is half integer, since in this case we expect an anticommuting result when we exchange the first two currents.

2.11.1 Light cone limit for the odd three point function

For the odd piece the light cone limit give us

$$\langle \underline{j_{s_1} j_{s_2}}_o j_{s_3}\rangle = \partial_1^{s_1}\partial_2^{s_2}\Upsilon(s_3), \qquad s_3 > 0 \tag{2.112}$$

where

$$\Upsilon(s) = \frac{1}{\widehat{x}_{12}}\left[\frac{\widehat{x}_{12}}{\widehat{x}_{13}\widehat{x}_{23}}\right]^s \sum_{k=0}^{[\frac{s-1}{2}]} C_{s-k-1,2k+1}\frac{1}{2k+1}\left[\frac{\widehat{x}_{13}\widehat{x}_{23}}{\widehat{x}_{12}^2}\right]^k \tag{2.113}$$

here $[\alpha]$ means integer part of α and $\widehat{x}$ were defined in (2.22) . And $C_{n,m}$ is the usual binomial coefficient. Interestingly, in this limit the answer can be written explicitly for any s. Notice that we do not have any square roots of the x_{ij}'s. Essentially for the same reason, we will not have a $|y_{12}|$ as we the limit $y_{12} \to 0$, we only get $1/y_{12}$. In general, it can be checked from the expressions in [41] that all square roots disappear from the odd correlator, even away from the light cone limit.

2.12 Appendix C. No odd piece from the one point energy correlator

Let's consider the case when $\langle j_2 j_2 j_2\rangle$ is the boson piece plus the odd piece. In other words, $\langle j_2 j_2 j_2\rangle_f = 0$.

Consider the one point energy correlator for the state created by a stress tensor at the origin with energy q^0 and zero spatial momentum. It is characterized by a two dimensional polarization tensor $\epsilon^{ij}T_{ij}$. See [20] for more details. The most general $O(2)$ symmetric structure in this case is given by the following expression

$$\langle \mathcal{E}(\vec{n}) \rangle = \frac{q^0}{2\pi}\left[1 + t_4\left(\frac{|\epsilon_{ij}n^j n^j|^2}{\epsilon_{ij}\epsilon^*_{ij}} - \frac{1}{4}\right)\right] - \frac{q^0}{2\pi}d_4\left[\frac{\varepsilon^{ij}(n_i n^m \epsilon_{jm}\epsilon^*_{kp}n^k n^p + n_i n^m \epsilon^*_{jm}\epsilon_{kp}n^k n^p)}{2\epsilon_{ij}\epsilon^*_{ij}}\right], \tag{2.114}$$

it's even part was studied in [50], however, the odd piece was missed in the literature before as far as we know. The positive energy condition $\langle \mathcal{E}(\vec{n}) \rangle \geq 0$ becomes

$$|d_4| \leq \sqrt{16 - t_4^2}. \tag{2.115}$$

For the case of the boson it is easy to check that $t_4 = 4$. Thus, we are forced to set $d_4 = 0$. We got that in the theory of Majorana fermion $t_4 = -4$.

More intuitively consider the state created by $T_{11} - T_{22}$. Then the one point energy correlator is given by

$$\langle \mathcal{E}(\vec{n}) \rangle = \frac{q^0}{2\pi}\left[1 + \frac{t_4}{4}\cos 4\theta + \frac{d_4}{4}\sin 4\theta\right] \tag{2.116}$$

where θ is the angle between x^1 and the detector.

For a free boson, we have that the energy correlator vanishes at $\theta = \pi/4$ which forces $t_4 = 4$.[21] Then near $\theta = \pi/4$ we would have regions with negative energy correlator if d_4 was nonzero.

Thus, we are forced to conclude that the odd piece is not allowed by the assumption of the positivity of the energy flow. In the main text we have also excluded the odd piece in other ways.

Notice also that the energy correlator that we obtain, after setting the fermion and odd parts to zero, is such that it vanishes for one particular angle. This is already a suggestion that the theory is probably free, since in an interacting theory we expect showering so the energy distribution will never be exactly zero. In other words, showering from a neighboring angle would make it non-zero. In fact, it is natural to conjecture that if $\langle j_2 j_2 j_2 \rangle_f = 0$ (or $\langle j_2 j_2 j_2 \rangle_b = 0$), then the theory is free, without assuming that a higher spin current exist.

[21] The operator is odd under the exchange of the 1 and 2 axis, but the state of two massless scalars going back to back along this axis would be even.

2.13 Appendix D. Half-integer higher spin currents implies even spin higher spin currents

Here we assume that we start from a higher, *half*-integer spin current j_s, $s \geq 5/2$. We then argue that we also have even spin currents. Then we can go back to the case treated in the main text.

The analysis is completely parallel to what was done in section five for integer spin currents. We simply need a couple of other formulas. For half integer s and s' we need that

$$\langle \underline{j_{s'} j_{2_b}} j_s \rangle \propto \partial_1^{s'-1/2} \partial_2^2 \langle \underline{\psi \phi} j_s \rangle_{free} \ , \tag{2.117}$$

$$\langle \underline{j_{s'} j_{2_f}} j_s \rangle \propto \partial_1^{s'-1/2} \partial_2 \langle \underline{\phi \psi} j_s \rangle_{free} \ , \tag{2.118}$$

$$\langle \underline{\psi \phi} j_s \rangle_{free} \propto \frac{1}{\widehat{x}_{13}^{3/2} \widehat{x}_{23}^{1/2}} \left(\frac{1}{\widehat{x}_{13}} - \frac{1}{\widehat{x}_{23}} \right)^{s-1/2} \tag{2.119}$$

$$\langle \underline{\phi \psi} j_s \rangle_{free} \propto \frac{1}{\widehat{x}_{13}^{1/2} \widehat{x}_{23}^{3/2}} \left(\frac{1}{\widehat{x}_{13}} - \frac{1}{\widehat{x}_{23}} \right)^{s-1/2} \tag{2.120}$$

It is clear that in a free theory of bosons and fermions these formulas are true. In fact, these formulas follow from conformal invariance and current conservation of the third current. This can be shown directly using the methods of appendix I.

We can now use these expressions to consider the Q_s charge conservation identity on $\langle \underline{j_2 j_{2_b}} j_s \rangle$. $[Q_s, j_s]$ gives a sum of integer spin currents so that we get a contribution to the charge conservation identity similar to the second term in (2.29) . The action of $[Q_s, \underline{j_2 j_{2_b}}]$ requires a bit more words. First $[Q_s, j_2]$ gives some other half integer spin currents. Then we can do the light-like OPE using (2.117) for each term and summing all the terms. After the dust settles, such terms give us

$$\langle [Q_s, \underline{j_2 j_{2_b}}] s \rangle = \partial_1^2 \partial_2^2 \left(\gamma \partial_1^{s-\frac{3}{2}} \frac{1}{\widehat{x}_{13}} + \delta \partial_2^{s-\frac{3}{2}} \frac{1}{\widehat{x}_{23}} \right) \frac{1}{\sqrt{\widehat{x}_{12} \widehat{x}_{13}}} \left(\frac{1}{\widehat{x}_{13}} - \frac{1}{\widehat{x}_{23}} \right)^{s-\frac{1}{2}} \tag{2.121}$$

$$\langle [Q_s, \underline{j_2 j_{2_f}}] s \rangle = \partial_1 \partial_2 \left(\gamma \partial_1^{s-\frac{1}{2}} \frac{1}{\widehat{x}_{23}} + \delta \partial_2^{s-\frac{1}{2}} \frac{1}{\widehat{x}_{13}} \right) \frac{1}{\sqrt{\widehat{x}_{12} \widehat{x}_{13}}} \left(\frac{1}{\widehat{x}_{13}} - \frac{1}{\widehat{x}_{23}} \right)^{s-\frac{1}{2}} \tag{2.122}$$

here permutation symmetry fixes $\delta = (-1)^{s-\frac{1}{2}} \gamma$.

The analysis of these charge conservation identities is very similar to the one we did before. Again we can count the number of independent terms in the right hand side and conclude that we expect a unique solution. This unique solution is the one we get in a free supersymmetric theory. We conclude that currents j_k appear in $\underline{j_2 j_2}$ with even spins $k = 2, 4, \cdots, 2s-1$. If $s \geq 5/2$, then $2s - 1 \geq 4$ and we have a higher spin current with integer spin and we can go back to the analysis done in section five.

2.14 Appendix E. Functions that are annihilated by all charges

Suppose that we have a function $f(x_i)$ of n variables that has a Taylor series expansion and is such that $Q_s = \sum_i (\partial_{x_i})^s$ annihilates it for all odd s. We want to prove that the function can be written as a sum of functions (not all equal)

$$f(x_i) = \sum_{\sigma} g_\sigma(x_{\sigma(1)} - x_{\sigma(2)}, x_{\sigma(3)} - x_{\sigma(4)}, ..., x_{\sigma(n-1)} - x_{\sigma(n)}) \tag{2.123}$$

where the sum is over all permutations. Notice that the each function depends on $[\frac{n}{2}]$ variables, compared to the original n x_i. If n is odd, then we drop the last variable (the function is independent of the last variable). Here we will not use conformal symmetry. Also we assume that the variables x_i are ordinary numbers and not vectors, etc.

2.14.1 Proof

First if f has a Taylor series expansion, we can organize it in terms of the overall degree of each term. Say we have a polynomial in x_i and we can separate the terms according to the overall degree of each term. Then Q_s should annihilate terms of different degrees. In other words, there is no mixing between terms of different degrees. Thus, we have to prove the statement for polynomials of degree k. The statement is obvious for degree $k = 0$. Now assume it is valid up to degree k and we want to prove it for degree $k + 1$.

If we have a polynomial P_{k+1} that is annihilated by the charges, then we can take $\partial_{x_1} P_k$ which is also annihilated and is a polynomial of degree k. Thus, by assumption $\partial_{x_1} P_k$ can be written as in (5.146) . We can now integrate each term in such a way that we preserve its form. For example if we have a function of the form $g(x_1 - x_2, \cdots)$, then we integrate it with respect to x_{12}. After we are done, we would have written $P_{k+1} = \sum_\sigma g_\sigma + h(x_2, \cdots x_n)$. In other words, we are left with an integration "constant" which is a function of $n-1$ variables. Clearly the first term, the one involving the functions g, is annihilated by the Q_s, so h is also annihilated. Now we can repeat the argument for h, viewing it as a function of $n - 1$ variables. In this way we eliminate all variables.

Thus, we have proven what we wanted to prove.

2.14.2 Functions with Fourier Transform

Of course, if the function f has a Fourier transform, then it is even easier to prove the statement. In that case we have $Q_s = \sum_i k_i^s$. Now we take $s \to \infty$, then the largest or lowest value of k dominates (depending on whether there is a $|k|$ bigger than one or not). There must be an even number of largest $|k_i|$ otherwise for large s we would violate the equality. Say there are two, say $k_1 = -k_2$. Thus, the Fourier transform must have a $\delta(k_1 + k_2)$. We can now repeat the same with the second largest, etc, to argue that $f = \sum_\sigma g(k_{\sigma(2)}, k_{\sigma(4)}, \cdots, k_{\sigma(n)})\delta(k_{\sigma(1)} + k_{\sigma(2)}) \cdots \delta(k_{\sigma(n-1)} - k_{\sigma(n)})$. which is the same as saying (5.146) .

The reason we proved it for functions that have a Taylor series expansion, rather than a Fourier expansion, is that our definition of the quasi-bilocal operators, (2.35) and (2.36) , involve the operator product expansion. So these are given in terms of power series expansions. In principle, we do not know how to continue them beyond their convergence radius. For this reason we do not know, in principle, whether a Fourier transform exists, or whether the action of the higher spin charges (2.38) continues to be valid beyond that region. Of course, the result proved in this appendix, together with conformal symmetry and analyticity, allows us to extend the bilocals everywhere.

2.15 Appendix F. Current conservation equation in terms of cross ratios

We consider parity even correlation functions of a conserved current with two spinning operators of general twist

$$\langle O_{\tau_1,s_1}(x_1) O_{\tau_2,s_2}(x_2) j_{s_3}(x_3)\rangle = \frac{1}{|x_{12}|^{\tau_1+\tau_2-1}|x_{13}|^{\tau_1+1-\tau_2}|x_{23}|^{\tau_2+1-\tau_1}} F(Q_i, P_i) \tag{2.124}$$

where Q_i and P_i are the conformal cross ratios introduced in [41], see also [42]. Here $\tau_{2,3}$ are the twists of the second and third operators. The function F is constrained by current conservation. Here we give an expression for the condition for current conservation as a function of the cross ratios. The idea is to write $\partial_{\lambda_3} \not{\partial}_{x_3} \partial_{\lambda_3}$, act on (2.124) , and then rewrite the answer in terms of cross ratios. We end up with an equation for F which can be expressed purely in terms of a differential operator acting on F, $\mathcal{D}_3 F = 0$, with

$$\begin{aligned}\mathcal{D}_3 = & -8(Q_2Q_3 - P_1^2)\partial_{Q_2}\partial_{Q_3}^2 + 8(Q_1Q_3 - P_2^2)\partial_{Q_1}\partial_{Q_3}^2 \\ & -4Q_2\partial_{Q_2}\partial_{Q_3} + 4Q_1\partial_{Q_1}\partial_{Q_3} - Q_1\partial_{P_2}^2 + Q_2\partial_{P_1}^2 + \\ & + 2(-Q_1P_1 + P_2P_3)\partial_{P_1}\partial_{P_2}^2 - 2(-Q_2P_2 + P_1P_3)\partial_{P_2}\partial_{P_1}^2 + \\ & + 2(-Q_1Q_2 + P_3^2)(\partial_{P_2}^2\partial_{Q_2} - \partial_{P_1}^2\partial_{Q_1}) + \\ & + 2(P_2^2 - Q_1Q_3)\partial_{Q_3}\partial_{P_2}^2 - 2(P_1^2 - Q_2Q_3)\partial_{Q_3}\partial_{P_1}^2 + \\ & + 8(Q_1P_1 - P_2P_3)\partial_{Q_1}\partial_{Q_3}\partial_{P_1} - 8(Q_2P_2 - P_1P_3)\partial_{Q_2}\partial_{Q_3}\partial_{P_2} + \\ & + (\tau_2 - \tau_1)\left[-4\partial_{Q_3}Q_3\partial_{Q_3} - 4\partial_{Q_3}(P_1\partial_{P_1} + P_2\partial_{P_2}) - Q_2\partial_{P_1}^2 - Q_1\partial_{P_2}^2 - 2P_3\partial_{P_1}\partial_{P_2}\right]\end{aligned} \tag{2.125}$$

In principle one could also write a similar expression for odd structures. One can write the S_i in [41], by taking square roots in (2.18) of [41], and then being careful about the sign of the square root. Doing this, we can still use (2.125) , for odd structures.

2.16 Appendix G. Four point function of scalar operators in the fermion-like theory

Using the transformation law for the twist two scalar which we derived in section 6.6

$$[Q_4, \tilde{j}_0] = \partial^3\tilde{j}_0 + \gamma(\partial_-^2 j_{2-\perp} - \partial_-\partial_\perp j_{2--}) \tag{2.126}$$

we can solve the charge conservation identity $\langle[Q_4, \tilde{j}_0\tilde{j}_0\tilde{j}_0\tilde{j}_0]\rangle = 0$. The result is

$$\langle\tilde{j}_0\tilde{j}_0\tilde{j}_0\tilde{j}_0\rangle = \frac{f(u,v)}{x_{12}^4x_{34}^4} \tag{2.127}$$

$$f(u,v) = f(v,u) = \frac{1}{v}f(\frac{u}{v},\frac{1}{v}) \tag{2.128}$$

$$f(u,v) = \alpha(1 + \frac{1}{u^2} + \frac{1}{v^2}) + \beta\frac{1 + u^{5/2} + v^{5/2} - u^{3/2}(1+v) - v^{3/2}(1+u) - u - v}{u^{3/2}v^{3/2}} \tag{2.129}$$

and for $\langle J_2\widetilde{j}_0\widetilde{j}_0\widetilde{j}_0\rangle$ we get, using the notation of [42],

$$\langle J_2(x_1)j_0(x_2)j_0(x_3)j_0(x_4)\rangle = \frac{\epsilon(Z_1,P_1,P_2,P_3,P_4)}{(P_1P_2)^3(P_1P_3)(P_1P_4)(P_3P_4)^2} \tag{2.130}$$

$$[Q_{123}f_1(u,v) + Q_{134}f_2(u,v) + Q_{142}f_3(u,v)]\,, \tag{2.131}$$

$$Q_{1ij} = \frac{(Z_1P_i)(P_jP_1) - (Z_1P_j)(P_iP_1)}{(P_iP_j)}, \tag{2.132}$$

$$\gamma f_1(u,v) = \frac{9}{5}\beta\frac{v^{3/2}}{u^{3/2}},\ \gamma f_2(u,v) = \frac{9}{5}\beta\frac{u}{v^{3/2}},\ \gamma f_3(u,v) = \frac{9}{5}\beta\frac{1}{u^{3/2}v^{3/2}}. \tag{2.133}$$

As in the case of bosons β and γ are fixed as soon as we choose the normalization for the two-point function of stress tensors.

2.17 Appendix I. Conformal blocks for free scalar

In the main text we introduced the quasi-bilocal operators by taking a particular light-like limit of stress tensors (2.35) . These are the sum of contributions from several spins. They can be defined in any theory, even in theories that do not have the higher spin symmetry. In a generic interacting theory we only get the contribution of the stress tensor. Here we would like to present some explicit formulas showing that the projection on to a single spin contribution does not define a genuine bilocal operator.

Let us focus on the contribution with a particular spin s. This is again a quasi-bilocal operator which can be defined as follows

$$b_s(x_1,x_2) = \sum c_{i,n}(x_1-x_2)^i\partial^n j_s(\frac{x_1+x_2}{2}), \tag{2.134}$$

$$\langle b_s(x_1,x_2)j_s(x_3)\rangle \sim \langle\phi^*(x_1)\phi(x_2)j_s(x_3)\rangle_{free}. \tag{2.135}$$

The second line fixes the constants in the first line. We can define it even away from the light cone limit by this formula. Defined in this way, $b_s(x_1,x_2)$ transforms as a bi-primary operator with conformal weights ($\Delta_x = \frac{1}{2}, \Delta_y = \frac{1}{2}$) and obeys the Laplace equation for both points. However, as we emphasized, this is not the product of two free fields!. In particular, notice that $b_2(x_1,x_2)$ exists in any CFT.

To see the quasi-bilocal nature of $b_s(x_1,x_2)$ let us consider the two point function of bilocals $\langle b_s(x_1,x_2)b_s(x_3,x_4)\rangle$. In $d=3$ it is given by the contribution of conserved current of spin s, J_s, into the four point function of free fields. This problem was addressed recently in [51]. The answer is given by the formula (6.20) in [51] with $a = b = 0$, $\lambda_1 = s + \frac{1}{2}$, $\lambda_2 = \frac{1}{2}$, $l = s$

$$\langle b_s(x_1, x_2) b_s(x_3, x_4) \rangle = \frac{1}{|x_{12}||x_{34}|} \mathcal{F}_s(u, v) \tag{2.136}$$

$$\mathcal{F}_s(u, v) = \frac{\sqrt{u}}{\pi} \int_0^\pi d\theta \frac{4^s X^s (1 + \sqrt{1 - X})^{-2s}}{\sqrt{1 - X}} \tag{2.137}$$

$$X = z \cos^2 \theta + \overline{z} \sin^2 \theta \tag{2.138}$$

$$u = z\overline{z}, \qquad v = (1 - z)(1 - \overline{z}) \tag{2.139}$$

This integral is not known for general s. Fortunately, for $s = 0$ it is known and we get

$$\langle b_0(x_1, x_2) b_0(x_3, x_4) \rangle = \frac{1}{|x_{13}||x_{24}|} \frac{2}{\pi} \frac{K(-\frac{z - \overline{z}}{1 - z})}{\sqrt{1 - z}} \tag{2.140}$$

where $K(y)$ is the complete elliptic integral of the first kind.

One can check that this solution satisfies the Laplace equation. Also one can check that it has a singularity at the expected locations. However, the behavior near the singularity is not the one that is expected for the correlator of local operators. More precisely, in the limit $z \to 1$, with $\overline{z}$ fixed but close to one, we get a term $\log(1 - z)$. For local operators the singularities are power-like.

2.18 Appendix J. Fixing correlators in the light-like limit

Here we would like to show how the three point correlators are fixed in the light-like limit $\langle \underline{j_{s_1} j_{s_2}} j_{s_3} \rangle$. We use the conventions and notation of Giombi, Prakash and Yin [41]. In the light like limit we have $P_3 = 0$, the rest of the conformal invariants are non-zero.

The general conformal invariant, without any factor of P_3 has the form

$$\langle j_{s_1} j_{s_2} j_{s_3} \rangle = \frac{1}{|x_{12}||x_{13}||x_{23}|} \sum_{a,b} c_{ab} Q_1^a Q_2^b P_2^{2s_1 - 2a} P_1^{2s_2 - 2b} Q_3^{s_3 - s_1 - s_2 + a + b}. \tag{2.141}$$

We can now use the relation the P's and Q's given in equation (2.14) of [41]. Setting $P_3 = 0$ we get

$$P_1^2 Q_1 + P_2^2 Q_2 = Q_1 Q_2 Q_3 \ , \qquad \Rightarrow Q_3 = \frac{P_1^2}{Q_2} + \frac{P_2^2}{Q_1} \tag{2.142}$$

We can use this relation to solve for Q_3 and eliminate it from (2.141) . Then we find a simpler formula

$$\langle j_{s_1} j_{s_2} j_{s_3} \rangle = \frac{1}{|x_{12}||x_{13}||x_{23}|} f(s_1, s_2, s_3) \ , \tag{2.143}$$

$$f(s_1, s_2, s_3) = \sum_{a=0}^{s_3} c(a) Q_1^{s_1 - a} P_2^{2a} P_1^{2(s_3 - a)} Q_2^{s_2 - s_3 + a} \tag{2.144}$$

We would like now to impose current conservation on the third current. The correspondent operator of the divergence can be obtained from (2.125) by setting P_3 to zero. The resulting operator is

$$\mathcal{D}_3 = -(1 + 2P_1 \partial_{P_1} + 2Q_2 \partial_{Q_2}) Q_1 \partial_{P_2}^2 + (1 + 2P_2 \partial_{P_2} + 2Q_1 \partial_{Q_1}) Q_2 \partial_{P_1}^2 . \tag{2.145}$$

By imposing $\mathcal{D}_3 f(s_1, s_2, s_3) = 0$ we find a recursive relation

$$\frac{c(a+1)}{c(a)} = \frac{(s_1 + a + \frac{1}{2})}{(a + \frac{1}{2})} \frac{(s_3 - a - \frac{1}{2})}{(s_3 + s_2 - a - \frac{1}{2})} \frac{(s_3 - a)}{(a+1)} \tag{2.146}$$

This gives a unique solution, which is the one we considered in the main text (2.20) . Once we know the solution is unique, we can simply use the free boson answer to get (2.20) . But one can check explicitly that (2.146) gives (2.20) .

2.18.1 Fermions

For fermions we expect a factor of P_3. In whatever multiplies this factor, we can set $P_3 = 0$ and use (2.142) . This means that the conservation condition will act in the same way on what multiplies P_3 as it acted on the boson case. Thus, we get the following structure

$$f^f(s_1, s_2, s_3) = P_3 P_1 P_2 \sum_{b=0}^{\widetilde{s}_3} \widetilde{c}(b) Q_1^{\widetilde{s}_1 - b} P_2^{2b} P_1^{2(\widetilde{s}_3 - b)} Q_2^{\widetilde{s}_2 - \widetilde{s}_3 + b} \tag{2.147}$$

where $\widetilde{s}_i = s_i - 1$.This looks very similar to (2.143) . In fact, we see that

$$f^f(s_1, s_2, s_3) = P_3 f_h(s_1 - \frac{1}{2}, s_2 - \frac{1}{2}, s_3) \tag{2.148}$$

where f_h is the same sum as in (2.143) but with a running over half integer values $a = \frac{1}{2}, \frac{3}{2}, \cdots, s_3 - \frac{1}{2}$. Thus, the action of $\mathcal{D}_3$, (2.145) , leads to the same recursion relation (2.146) , but now running over half integer values of a.

This again proves that there is a unique structure, which is (2.23) .

In the same way one can obtain (2.117) .

2.18.2 General operators

We now explicitly solve the problem of finding the different parity even structures we can have for the three point function of two operators of the same twist and one conserved current. The problem is very similar to what we solved above. The main observation is that P_3 does not depend on λ_3 or x_3. Then if we have a solution to the current conservation condition, we can generate a new solution by multiplying by P_3. Then the general even structures have the form

$$\langle O_{s_1} O_{s_2} j_{s_3} \rangle \sim \frac{1}{|x_{12}|^{2\tau_0 - 1} |x_{23}||x_{13}|} \sum_l P_3^{2l} \left[\langle j_{s_1-l} j_{s_2-l} j_{s_3} \rangle_b + \langle j_{s_1-l} j_{s_2-l} j_{s_3} \rangle_f \right] \tag{2.149}$$

The sum runs over the bosonic and fermionic three point functions of three conserved currents. We can easily check that these are solutions. The fact that these are all the solutions is obtained by taking light cone limits. To order P_3^0 the solution is constrained as in (2.146) , which is the same as what we get from the $l = 0$ in (3.12) . If we now subtract the $l = 0$ term, which is a solution, we can take the light cone limit and focus on the P_3^1 term, etc.

In (3.12) , the sum over the bosonic structures runs over $1 + \min[s_1, s_2]$ values. The sum over fermion structures runs over $\min[s_1, s_2]$ values. The total number of structures that we have is

$$1 + 2\min[s_1, s_2] \tag{2.150}$$

These formulas are valid for integer spins. Probably, there is a similar story with half integer spins.

It would be interesting to fix also the structures that appear when the twists of the two operators are not the same. For that we would have to include the last term in (2.125) when we analyze the constraints.

2.19 Appendix K. More explicit solutions to the charge conservation identities of section five

Here we discuss further the charge conservation identities of section five and we prove that they have the properties stated there. To find the explicit solution of (2.29) it is convenient to restate the problem in a slightly different language. First remember that it is easy to prove that there is a unique solution of (2.29) which coincides with the free field one. However, we need to check that for all even k $\widetilde{\alpha}_k$ is non-zero. So we focus on the free field theory.

2.19.1 Expression for the currents in the free theory

The expressions for the conserved currents are bilinear in the fields and contain a combination of derivatives that makes the field a primary conformal field. These are computed in multiple places in the literature, (see e.g. [52] [53] and references therein). We would like to start by providing a derivation for the expression for these currents that is particularly simple in three dimensions.

Consider the Fourier space expression for a current contracted with a three dimensional spinor λ, $\lambda^{2s}.J(q)$, where q^μ is the three-momentum. We can consider the matrix element

$$F = \langle p_1, p_2 | (\lambda^{2s}.J(q)) | 0 \rangle \tag{2.151}$$

here p_1 and p_2 are two massless three dimensional momenta of the two bosons or two fermions that make up the current. A massless three dimensional momentum can be written in terms of a three dimensional spinor, π_α. In other words, we write $p_{\alpha\beta} = p^\mu \sigma^\mu_{\alpha\beta} = \pi_\alpha \pi_\beta$. Since we have two massless momenta we have two such spinors π^1 and π^2. Of course, $q = p_1 + p_2$ and it is not an independent variable. We can also view (2.151) as the form factor for the current. It is telling us how the current is made out of the fundamental fields. Clearly, Lorentz invariance implies that (2.151) is a function $F(\lambda.\pi^1, \lambda.\pi^2)$. However, since the current is also conserved, it should obey the equation

$$0 = q^{\alpha\beta} \frac{\partial^2}{\partial \lambda^\alpha \lambda^\beta} F = (\pi_1.\pi_2)^2 \left[\frac{\partial^2}{\partial (\pi^1.\lambda)^2} + \frac{\partial^2}{\partial (\pi^2.\lambda)^2} \right] F \tag{2.152}$$

where we expressed $q_{\alpha\beta} = \pi^1_\alpha \pi^1_\beta + \pi^2_\alpha \pi^2_\beta$ and we acted on the function $F(\lambda.\pi^1, \lambda.\pi^2)$. We can now define the variables

$$z = \pi_1.\lambda - i\pi_2.\lambda \ , \qquad\qquad \overline{z} = \pi_2.\lambda + i\pi_2.\lambda \tag{2.153}$$

Then the equation (2.152) becomes

$$\partial_z \partial_{\overline{z}} F(z, \overline{z}) = 0 \tag{2.154}$$

The solutions are purely holomorphic or purely anti-holomorphic functions. Since the homogeneity degree in λ is fixed to be $2s$, we get the two solutions

$$F_b = z^{2s} + \overline{z}^{2s} \ , \qquad\qquad F_f = z^{2s} - \overline{z}^{2s} \tag{2.155}$$

The first corresponds to the expression for the current in the free boson theory and the second to the expression in the free fermion theory, as we explain below. These expressions contain the same information as the functions $\langle \phi\phi^* j_s \rangle_{free}$ and $\langle \psi\psi^* j_s \rangle_{free}$ that we discussed in section five.

In order to translate to the usual expressions in terms of derivatives acting on fields, without loss of generality, we can choose λ so that it is purely along the direction λ^-. In the boson expression, each $\pi^i.\lambda$ appears to even powers, say $(\pi^1.\lambda)^{2k}(\pi^2.\lambda)^{2s-2k}$ such a term can be replaced by $\partial_-^k \phi^* \partial_-^{2(s-k)} \phi$. The F_b gives the expression

$$j_s^{bos} \sim \sum_{k=0}^{s} (-1)^k \frac{1}{(2k)!(2s-2k)!} \partial^k \phi^* \partial^{s-k} \phi. \tag{2.156}$$

Similarly, for the fermion case we have odd powers and we replace $(\pi^1.\lambda)^{2k+1}(\pi^2.\lambda)^{2s-2k-1}$ by $\partial_-^k \psi_-^* \partial_-^{s-1-k} \psi_-$ and F_f in (2.155) gives

$$j_s^{fer} \sim \sum_{k=0}^{s} (-1)^k \frac{1}{(2k+1)!(2s-2k-1)!} \partial^k \psi^* \partial^{s-k} \psi. \tag{2.157}$$

where all are minus component.

Notice that we do not even need to go to Fourier space to think about the z and $\overline{z}$ variables. We can just think of them as book-keeping devices to summarize the expansions (2.157) , (2.156) via the simple expressions (2.155) .

2.19.2 Analyzing the charge conservation identity

Symmetries act on free field as

$$[Q_s, \phi] = \partial^{s-1}\phi, \quad [Q_s, \phi^*] = (-1)^s \partial^{s-1} \phi^* \tag{2.158}$$

and a similar expression for the fermions. Expressing this in terms of the z and $\overline{z}$ variables we rewrite (2.158) as

$$[Q_s, j_{s'}](z, \overline{z}) = \left[(z+\overline{z})^{2s-2} - (z-\overline{z})^{2s-2} \right] (z^{2s'} \pm \overline{z}^{2s'}). \tag{2.159}$$

where the $\pm$ indicate the results in a boson of fermion theory respectively. Then, using the notation from (2.29), we have

$$[Q_s, j_s](z, \overline{z}) = \sum_{k=1}^{s-1} \widetilde{\alpha}_{2k} z^{2s-2k-1} \overline{z}^{2s-2k-1} (z^{4k} \pm \overline{z}^{4k}) \tag{2.160}$$

where we used the map $\partial_3 \to z\overline{z}$. As in the previous subsection this can be seen by $\partial_3 = \partial_{x_3^-} \to \lambda^2 q$ and $\lambda^2.q = (\lambda.\pi^1)^2 + (\lambda.\pi^2)^2 = z\overline{z}$. Expanding the left hand side of (2.159) for $s' = s$ we get

$$\widetilde{\alpha}_{2k} = \frac{2\,\Gamma(2s-1)}{\Gamma(2k)\Gamma(2s-2k)} \tag{2.161}$$

which are all non-zero. Thus, this solves the problem of showing that $\widetilde{\alpha}_k$ are non-zero for the free theories. This works equally simply for the boson-like and the fermion-like light like limits.

We can also do a similar expansion for the action of $[Q_s, j_{s'}]$ and find which terms appear in the right hand side

$$\begin{aligned}
[Q_s, j_{s'}] =& (2s-2)! \sum_{r=-s'-s+2}^{s+s'-2} \widetilde{\alpha}_r \partial^{s+s'-r-1} j_r \\
\widetilde{\alpha}_r =& \frac{[1+(-1)^{s+s'+r}]}{2} \left(\frac{1}{\Gamma(r+s-s')\Gamma(s+s'-r)} \pm \frac{1}{\Gamma(r+s+s')\Gamma(s-s'-r)} \right).
\end{aligned} \tag{2.162}$$

The $\pm$ indicates the boson/fermion case.

This is how all the other charge conservation identities we discussed in section five can be analyzed. Namely, in section five we argued that all solutions are unique. And by computing for free theories as in this appendix we check which coefficients are non-zero. This works in the same way for the boson-like and fermion-like light like limits.

We can go over the charge conservation identities and the properties we used.

(2.41) : We used that if s is odd, $s > 1$, and $s' = 2$, then both j_1 and j_s appear in the right hand side. This simply amounts to checking that for $s' = 2$, $r = 1, s$, the coefficient in (3.81) is non-zero for any odd s. We can also check that we get a j_s in the right hand side. This is the case both for the boson and fermion case.

(2.44) : We used that if s is even and $s' = 2$, then we get both a j_2 and a j_s in the right hand side of (3.81) .

(2.43) : Here we used a slightly different property. The reason is that (2.43) does not correspond to the transformation of any free field, since an s even charge in free theories acts with + signs in front of the ∂^{s-1}. In order to analyze (2.43) , it is convenient to view the functions in (2.43) in

Fourier space, as we have done here. Then we can rewrite (2.43) as

$$\left[(z+\overline{z})^{2s-2}+(z-\overline{z})^{2s-2}\right](z^2\pm\overline{z}^2)=(2s-2)!\sum_{r=-s}^{s}\widetilde{\beta}_r(z\overline{z})^{s-r}(z^{2r}\pm\overline{z}^{2r}) \tag{2.163}$$

Then one can find that

$$\widetilde{\beta}_r=\frac{[1+(-1)^{s+r}]}{2}\left(\frac{1}{\Gamma(r+s-1)\Gamma(s-r+1)}\pm\frac{1}{\Gamma(r+s+1)\Gamma(s-r-1)}\right) \tag{2.164}$$

which is the same expression as before, in (3.81) , for $s'=1$, except for the (-1) dependent prefactor which now implies that now the sum over r runs from $r=-s,-s+2,\cdots,s$. Now the important point is that $\widetilde{\beta}_s$ is non-zero. We used this in (2.43) .

In the discussion of elimination of j_0', for the boson-like case, we used that $[Q_s, j_{s'}]$ produces both j_0 and j_2 in the boson theory and for s, s' even, with $s>s'$. This can also be checked from (3.81) , by setting $r=0,2$. Note that in (3.81) $\widetilde{\alpha}_0=0$ in fermion theory (the minus sign in (3.81)), as expected.

Chapter 3

Constraining Conformal Field Theories With a Slightly Broken Higher Spin Symmetry

3.1 Introduction

In this paper we study a special class of three dimensional conformal field theories that have a weakly broken higher spin symmetry. The theories have a structure similar to what we expect for the CFT dual to a weakly coupled four dimensional higher spin gravity theory in AdS_4 [24, 54, 55, 56, 57, 58, 11, 59, 12]. We compute the leading order three point functions of the higher spin operators. We use current algebra methods. Our only assumption is that the correlation functions defined on the boundary of AdS_4 obey all the properties that a boundary CFT would obey. But we will not need any details regarding this theory other than some general features which follow from natural expectations for a weakly coupled bulk dual. This seems a reasonable assumption for Vasiliev's theory, since Vasiliev's theory appears to be local on distances much larger than the AdS radius. This would imply that the usual definition of boundary correlators is possible [12, 26, 28]. In order to apply our analysis to Vasiliev's theory we need to make the assumption that these boundary correlators can be defined and that they obey the general properties of a CFT. Thus we are assuming AdS/CFT, but we are not specifying the precise definition of the boundary CFT. Since our assumptions are very general, they also apply to theories involving N scalar or fermions fields coupled to $O(N)$ or $SU(N)$ Chern-Simons gauge fields [44, 45]. So our methods are also useful for computing three point functions in these theories as well.

Our assumptions are the following[1]. We have a CFT with a unique stress tensor and has a large parameter $\widetilde{N}$. In the Chern-Simons examples $\widetilde{N} \sim N$. In the Vasiliev gravity theories, $1/\widetilde{N} \sim \hbar$ sets the bulk coupling constant of the theory. We then assume that the spectrum of operators has the structure of an approximate Fock space, with single particle states and multiparticle states. The dimensions of the multiparticle states are given by the sum of the dimensions of their single particle constituents up to small $1/\widetilde{N}$ corrections. This Fock space should not be confused with the Fock space of a free theory in three dimensions. We should think of this Fock space as the Fock space of the weakly coupled four dimensional gravity theory. In order to avoid this confusion we will call the "single particle states" "single trace" and the multiparticle states "multiple trace". In the Chern-Simons gauge theories this is indeed the case. We also assume that the theory has the following spectrum of single trace states. It has a single spin two conserved current. In addition, it has a sequence of approximately conserved currents J_s, with $s = 4, 6, 8, \cdots$. These currents are approximately conserved, so that their twist differs from one by a small amount of order $1/\widetilde{N}$

$$\tau_s = \Delta_s - s = 1 + O(\frac{1}{\widetilde{N}}). \tag{3.1}$$

In addition, we have one single trace scalar operator. All connected correlators of the single trace operators scale as $\widetilde{N}$. This includes the two point function of the stress tensor. We also assume that the spectrum of single trace operators is such that the higher spin symmetry can be broken only by double trace operators via effects of order $1/\widetilde{N}$. In particular, we assume that there are no twist three single trace operators in the theory. With these assumptions, we will find that the three point functions in these theories, to leading order in $\widetilde{N}$, are constrained to lie on a one parameter family

$$\langle J_{s_1} J_{s_2} J_{s_3} \rangle = \widetilde{N} \left[\frac{\tilde{\lambda}^2}{1+\tilde{\lambda}^2} \langle J_{s_1} J_{s_2} J_{s_3} \rangle_{\rm bos} + \frac{1}{1+\tilde{\lambda}^2} \langle J_{s_1} J_{s_2} J_{s_3} \rangle_{\rm fer} + \frac{\tilde{\lambda}}{1+\tilde{\lambda}^2} \langle J_{s_1} J_{s_2} J_{s_3} \rangle_{\rm odd} \right] \tag{3.2}$$

where the subindices $_{\rm bos}$ and $_{\rm fer}$ indicate the results in the theory of a single real boson or a single Majorana fermion. The subindex $_{\rm odd}$ denotes an odd structure which will be defined more clearly below. Here $\tilde{\lambda}$ is a parameter labeling the family of solutions of the current algebra constraints. More precise statements will be made below, including the precise normalization of the currents.

The class of theories for which our assumptions apply includes Vasiliev higher spin theories in

[1] These assumptions are not all independent from each other, but we will not give the minimal set.

AdS_4 with higher spin symmetry broken by the boundary conditions [24]. It also applies for theories containing N fermions [44] or N bosons [45] interacting with an $SO(N)$ or $U(N)$ Chern-Simons gauge field. We call these theories quasi-fermion and quasi-boson theories respectively. In such theories $1/\widetilde{N} \propto 1/N$ is the small parameter. In addition, $\widetilde{\lambda} \sim \lambda = N/k$ is an effective 't Hooft coupling in these theories. We emphasize that the analysis here is only based on the symmetries and it covers both types of theories, independently of any conjectured dualities between them. Of course, the results we obtain are consistent with the proposed dualities between these Chern-Simons theories and Vasiliev's theories [11, 58, 12, 44, 45]. We can take the limit of large $\widetilde{\lambda}$ in (3.2) and we find that the the correlators of the quasi-fermion theory go over to those of the critical $O(N)$ theory. And we have a similar statement for the quasi-boson theory.

Our analysis is centered on studying the spin four single trace operator J_4. We write the most general form for its divergence, or lack of conservation. With our assumptions this takes the schematic form

$$\partial.J_4 = a_2 J J' + a_3 J J' J'' \qquad (3.3)$$

where on the right hand side we have products of two or three single trace operators, together with derivatives sprinkled on the right hand side. The coefficients a_2 and a_3 are small quantities of order $1/\widetilde{N}$ and $1/\widetilde{N}^2$ respectively. We will be able to use this approximate conservation law in the expression for the three point function in order to get (3.2) . Note that in the case that J_4 is exactly conserved, we simply have free boson or free fermion correlators [60]. We should emphasize that our discussion applies only to the special theories in [44, 45], but not to more general large N Chern-Simons matter theories. The special feature that we are using is the lack of single trace operators of twist three. Such operators can appear in the divergence of the spin four current (3.3) . This can give rise to an anomalous dimension for the higher spin currents already at the level of the classical theory (or large $\widetilde{N}$ approximation). In the theories in [44, 45], we do not have a single trace operator that can appear in the right hand side of (3.3) . In the language of the bulk theory, we have the pure higher spin theory without extra matter[2]. In particular, the bulk theory lacks the matter fields that could give a mass to the higher spin gauge fields via the Higgs mechanism already at the classical level. Thus the higgsing is occurring via quantum effects involving two (or three) particle states [61]. Our analysis can also be viewed as an on shell analysis of the Vasiliev theory with AdS_4 asymptotic boundary conditions. If the higher spin symmetry is unbroken, then we can use [60] to

[2]By "matter" we mean extra multiplets under the higher spin symmetry. The scalar field of the Vasiliev theory is part of the pure higher spin theory.

compute all correlators, just from the symmetry. In this paper we also use an on shell analysis, but for the case that the higher spin symmetry is broken. As it has often been emphasized, on shell results in gauge theories can be simpler than what the fully covariant formalisms would suggest. For a more general motivational introduction, see appendix G. The paper is organized as follows. In section two we discuss the most general form of the divergence of the spin four current. In section three we present several facts about three point functions which are necessary for the later analysis. In section four we explain how one can use the slightly broken higher spin symmetry to fix the three point functions. In section five we present the results and explain their relation to known microscopic theories. In section six we present conclusions and discussions. Several appendices contain technical details used in the main body of the paper.

3.2 Possible divergence of the spin four operator

3.2.1 Spectrum of the theory

We consider theories with a large $\widetilde{N}$ expansion. We do not assume that $\widetilde{N}$ is an integer. We assume that the set of operators develops the structure of a Fock space for large $\widetilde{N}$. Namely, we can talk about single particle operators and multiparticle operators. In the case of the Chern-Simons matter theories discussed in [44, 45], these correspond to single sum and multiple sum operators (sometimes called single trace or multitrace operators). The spectrum of single trace operators includes a conserved spin two current, the stress tensor, $J_{2\,\mu\nu}$. We will often suppress the spacetime indices and denote operators with spin simply by J_s. We have approximately conserved single trace operators J_s, with $s = 4, 6, 8, \cdots$. These operators have twists $\tau = 1 + O(1/\widetilde{N})$. In addition we have a single scalar operator. We will see that the dimension of this operator has to be either one or two. We denote the first possibility as j_0 and the second as $\tilde{j}_0$. The theory that contains $\tilde{j}_0$ is called the quasi-fermion theory. The theory with j_0 is called the quasi-boson theory. The theory might contain also single particle operators with odd spins. For simplicity, let us assume that these are not present, but we will later allow their presence and explain that the correlators of even spin currents are unchanged. An example of a theory that obeys these properties is a theory with N massless Majorana fermions interacting with an $O(N)$ Chern-Simons gauge field at level k [44]. In this theory the scalar is $\tilde{j}_0 = \sum_i \psi_i \psi_i$, which has dimension two, at leading order in N for any $\lambda = N/k$. The name quasi-fermion was inspired by this theory, since we start from fermions and the Chern-Simons interactions turns them into non-abelian anyons, which for large k, are very close to ordinary fermions. Our discussion is valid for any theory whose single particle spectrum was described above. We are just calling "quasi-fermion" the case where the spectrum includes a scalar with dimension two, $\widetilde{j}_0$. A second example is a theory with N massless real scalars, again interacting with an $O(N)$ Chern-Simons gauge field at level k [45]. This theory also allows the presence of a $(\vec{\phi}.\vec{\phi})^3$ potential while preserving conformal symmetry, at least to leading order in N. As higher orders in the $1/N$ expansion are taken into account this coefficient is fixed, if we want to preserve the conformal symmetry [45]. Here we will only do computations to leading order in N, thus, we have two parameters N/k and the coupling of the $(\vec{\phi}.\vec{\phi})^3$ potential. We call this case the quasi-boson theory. Again we will not use any of the microscopic details of its definition. For us the property that defines it is that the scalar has dimension one, j_0. A third example is the critical $O(N)$ theory (as well as interacting UV Gross-Neveu fixed point). Namely, we can have N free scalars, perturbed by a potential of the form $(\vec{\phi}.\vec{\phi})^2$, which flows in the IR to a new conformal field theory (after adjusting the coefficient of the mass term to criticality). This is just the usual large N limit of the Wilson-Fisher fixed point. This theory has no free parameters. Here the scalar operator $\widetilde{j}_0 \sim \phi_i \phi_i$ has dimension two. It starts with dimension one in the UV but it has dimension two in the IR. The UV theory has a higher spin symmetry. In the IR CFT this symmetry is broken by $1/N$ effects. This theory is in the family of what we are calling the quasi-fermion case. A fourth example is a Vasiliev theory in AdS_4 (or dS_4) with general boundary conditions which would generically break the higher spin symmetry. Here the bulk coupling is $\hbar \sim 1/\widetilde{N}$. Depending on whether the scalar has dimensions one or two, we would have a quasi-fermion or quasi-boson case. We should emphasize that the theories we call quasi-fermion or quasi-boson case are not specific microscopic theories. They are *any* theory that obeys our assumptions, where the scalar has dimension two or one respectively.

3.2.2 Divergence of the spin four current

Let us consider the spin four current J_4. We consider the divergence of this current. If it is zero, then we have a conserved higher spin current and all correlators of the currents are as in a theory of either free bosons or free fermions [60]. Here we consider the case that this divergence is nonzero. Our assumptions are that the current is conserved in the large $\widetilde{N}$ limit. This means that in this limit J_4 belongs to a smaller multiplet than at finite $\widetilde{N}$. At finite $\widetilde{N}$, J_4 combines with another operator to form a full massive multiplet. More precisely, it combines with the operator that appears in the right hand side of $\nabla.J_4 = \partial_\mu J^\mu{}_{\nu_1\nu_2\nu_3}$. In other words, $\nabla.J_4$, should be a conformal primary operator in the large N limit [62, 61], see appendix A. $\nabla.J_4$ should be a twist three, spin three primary operator. According to our assumptions there are no single particle operators of this kind. Note that in general matter Chern-Simons theories, such as theories with adjoint fields, we can certainly have single trace operators with twist three and spin three. So, in this respect, the theories we are considering are very special. In our case, we can only have two particle or three particle states with these quantum numbers. Let us choose the metric

$$ds^2 = dx^+ dx^- + dy^2 \tag{3.4}$$

and denote the indices of a vector by $v_\pm$, v_y. Let us see what is the most general expression we can write down for $(\nabla.J)_{---}$. Since the total twist is three, and the total spin is three we can only make this operator out of the stress tensor, J_2, and the scalar field. Any attempt to include a higher spin field would have to raise the twist by more than three. A scalar field can only appear if its twist is one or two. Note that we cannot have two stress tensors. The reason is that we cannot make a twist three, spin three primary out of two stress tensors. Naively, we could imagine an expression like $J_{2\,y-}J_{2\,--}$. But this cannot be promoted into a covariant structure, even if we use the ϵ tensor.[3] Let us consider first the case that the scalar has twist two, $\widetilde{j}_0$. The most general operator that we can write down is

$$\partial_\mu J^\mu{}_{---} = a_2 \left(\partial_- \widetilde{j}_0 j_2 - \frac{2}{5} \widetilde{j}_0 \partial_- j_2 \right). \tag{3.5}$$

Here we are denoting by $j_s = J_{s\,-\cdots-}$, the all minus components of J_s. In (3.5) we have used the fact that the right hand side should be a conformal primary in order to fix the relative coefficient. If we started out from the free fermion theory, then this structure would break parity, since $\widetilde{j}_0$ is parity odd. Then we have $a_2 \propto N/k$, at least for large k. In general $\widetilde{j}_0$ does not have well defined parity and the theory breaks parity. If we had the critical O(N) theory, then (3.5) is perfectly consistent with parity, since in that case $\widetilde{j}_0$ is parity even.

Let us now consider the case where we have a scalar of twist one, j_0. Now there are more

[3] We could have promoted it if we had two different spin two currents: $\epsilon_{-\mu\nu} J^\mu{}_- \widetilde{J}^\nu{}_-$

conformal primaries that we can write down

$$\begin{aligned}\partial_\mu j^\mu{}_{---} &= a_2 \epsilon_{-\mu\nu}\left[8\partial^\mu j_0 \partial^\nu j_{--} - 6\partial^\mu j_0 \partial_- j^\nu{}_- - 5\partial_-\partial^\mu j_0 j^\nu{}_- - j_0 \partial_- \partial^\mu j^\nu{}_-\right] \\ &+ a_3\left[j_0 j_0 \partial^3 j_0 - 9 j_0 \partial j_0 \partial^2 j_0 + 12 \partial j_0 \partial j_0 \partial j_0\right] \\ &+ a_3'\left[\partial j_2 j_0 j_0 - 5 j_2 \partial j_0 j_0\right]. \end{aligned} \tag{3.6}$$

We have only written combinations which are conformal primaries. We have also denoted $\partial \equiv \partial_-$. The analysis of the broken charge conservation identities will relate a_3 and a_3' leaving us with only two parameters (besides $\widetilde{N}$). This agrees with the two parameters in the large N limit in the boson plus Chern-Simons theories of [45].

Here we have concentrated on the case of J_4. Let us briefly discuss the situation for higher spin currents. We focus, as usual, on the all minus component of the current $\partial_\mu J^\mu{}_{--\dots-}$. This operator has twist $\tau = \Delta - S = 3$ and spin $s-1$. Let us examine possible double particle operators that can appear. The minimum twist of a double trace operator is $2 = 1 + 1$. We should make up the twist by considering other components, or derivatives other than ∂_-, which has twist zero. All of these should arise from a rotationally invariant structure involving the flat space metric or the ϵ tensor. The only structure that can raise the twist by one is $\epsilon_{-\mu\nu}$. For the quasi-fermion theory we can also use the scalar operator of twist 2, $\tilde{j}_0$, and one of the twist one currents.

Matching the scaling dimensions in $\partial_\mu J^\mu_s \propto J_{s_1} J_{s_2}$ (with derivatives sprinkled on the right hand side) with all minus indices leads to

$$s + 2 = (s_1 + 1) + (s_2 + 1) + n_\partial \tag{3.7}$$

where $n_\partial \geq 0$ is the number of derivatives which raise the dimension. Thus, we get an inequality

$$s \geq s_1 + s_2, \qquad s > s_1,\ s_2\ , \qquad \text{double trace} \tag{3.8}$$

where we show that that $s > s_1, s_2$ as follows. For $s_1 = s$, the only operator with the right twist would be $j_0 J_{s\,y-\dots-}$, but this is not really a spin $s-1$ operator, namely it does not come from any covariant structure. (3.8) is a constraint on the spins of the operators that can appear in the divergence of a current of spin s. At the level of triple trace operators the product of three operators has already twist three. So the only structure which is allowed is ∂_- by the twist counting. Matching the dimensions in $\partial_\mu J_s^\mu \propto J_{s_1} J_{s_2} J_{s_3}$ we get

$$s + 2 = (s_1 + 1) + (s_2 + 1) + (s_3 + 1) + n_\partial \tag{3.9}$$

and the constraint

$$s \geq s_1 + s_2 + s_3 + 1, \qquad \text{triple trace} \tag{3.10}$$

Twist counting prohibits having the product of more than three operators. Now let us comment on the scaling of the coefficients in (3.3) with $\tilde{N}$. Let us normalize the scaling of single particle operators so that their connected n point functions scale like $\tilde{N}$. Then if we consider a three point correlator of a given current with the two currents that appear in the right hand side of its divergence we get

$$\tilde{N} \sim \partial_\mu \langle J_s^\mu(x) J_{s_1}(x_1) J_{s_2}(x_2)\rangle = a_2 \langle J_{s_1}(x) J_{s_1}(x_1)\rangle\langle J_{s_2}(x) J_{s_2}(x_2)\rangle \sim a_2 \tilde{N}^2 \tag{3.11}$$

with derivatives sprinkled on the right hand side. Thus, we get that $a_2 \propto \frac{1}{\tilde{N}}$. For a_3 the same argument leads to $a_3 \propto \frac{1}{\tilde{N}^2}$. This scaling is the only one that is consistent with the $\tilde{N}$ counting and is such that it leads to non-zero terms in the leading contribution.

3.3 Structures for the three point functions

In this section we constrain the structure of three point functions. When we have exactly conserved currents the possible three point functions were found in [41] (see also [42] [60]). They were found by imposing conformal symmetry and current conservation. The three point functions were given by three possible structures. One structure arises in the free fermion theory and another arises in the free boson theory. We call these the fermion and boson structures respectively. Finally there is a third "odd" structure which does not arise in a free theory. For twist one fields this structure is parity odd (it involves an epsilon tensor). However, for correlators of the form $\langle \tilde{j}_0 J_{s_1} J_{s_2}\rangle$, the fermion structure is parity odd (it involves an epsilon tensor) and the "odd" structure is parity even. This is due to the fact that $\tilde{j}_0$ is parity odd in the free fermion theory. Alternatively, if we assign parity minus to $\tilde{j}_0$ and parity plus to all the twist one operators, then the "odd" structures always violate parity[4]. The reader should think that when we denote a structure as "odd" we simply mean

[4]This is not always the natural parity assignment. For example, in the critical $O(N)$ theory $\tilde{j}_0$ has parity plus and the theory preserves parity. In this theory, we have only the "odd" structure for the correlators of the form $\langle \tilde{j}_0 J_{s_1} J_{s_2}\rangle$.

"strange" in the sense that it does not arise in a theory of a free boson or free fermion.

In our case, the currents are not conserved, so we need to revisit these constraints. For example, the divergence of a current can produce a double trace operator. If these contract with the two remaining operators, as in (3.11) , we get a term that is of the same order in the $1/\widetilde{N}$ expansion as the original three point function. Notice that only double trace operators can contribute in this manner to the current non-conservation of a three point function. We emphasize that we are computing these three point functions to leading order in the $1/\widetilde{N}$ expansion, where we can set their twist to be one. All statements we make in this section are about the structure of correlators to leading order in the $1/\widetilde{N}$ expansion.

We will show below that even correlation functions (fermion and boson) stay the same and all new structures appear in the odd piece. Consider the three point function of twist one operators $\langle J_{s_1} J_{s_2} J_{s_3} \rangle$, with $s_i \geq 2$. Let us say that s_1 is larger or equal than the other two spins. Then the J_{s_2} and J_{s_3} currents are conserved inside this three point function, since the spins appearing in the divergence of a current are always strictly less than those of the current itself (3.8) . On the other hand, in order to get a non-zero contribution we would need to contract J_{s_1} with one of the two currents that appears in the right hand side of the divergence of J_{s_2} or J_{s_3}. Thus, we can impose current conservation on J_{s_2} and J_{s_3} for this three point function. Let us consider the parity even structures first. As we discussed in [60], for two operators of the same twist and one conserved current, say j_{s_3}, we have the most general even structure

$$\langle O_{s_1} O_{s_2} j_{s_3} \rangle \sim \frac{1}{|x_{12}|^{2\tau_0 - 1} |x_{23}||x_{13}|} \sum_{l=0}^{\min[s_1,s_2]} P_3^{2l} \left[\langle j_{s_1-l} j_{s_2-l} j_{s_3} \rangle_{\text{bos}} + \langle j_{s_1-l} j_{s_2-l} j_{s_3} \rangle_{\text{fer}} \right] \tag{3.12}$$

where P_i are as is [41]. We get this result by considering the light cone limit between $\underline{O_{s_1} O_{s_2}}$ and imposing the conservation of j_{s_3} [60]. We then take light cone limit $\underline{j_{s_1} j_{s_2}}$ and impose conservation of j_{s_3}. Then we take the light cone limit $\underline{j_{s_1} j_{s_3}}$ and impose conservation of j_{s_2}. From these two operations we would conclude that

$$\langle j_{s_1} j_{s_2} j_{s_3} \rangle = \frac{1}{|x_{12}||x_{23}||x_{13}|} \sum_{l=0}^{\min[s_1,s_2]} P_3^{2l} \left[b_l \langle j_{s_1-l} j_{s_2-l} j_{s_3} \rangle_{\text{bos}} + f_l \langle j_{s_1-l} j_{s_2-l} j_{s_3} \rangle_{\text{fer}} \right], \tag{3.13}$$

$$\langle j_{s_1} j_{s_2} j_{s_3} \rangle = \frac{1}{|x_{12}||x_{23}||x_{13}|} \sum_{l=0}^{\min[s_1,s_3]} P_2^{2l} \left[\widetilde{b}_l \langle j_{s_1-l} j_{s_2} j_{s_3-l} \rangle_{\text{bos}} + \widetilde{f}_l \langle j_{s_1-l} j_{s_2} j_{s_3-l} \rangle_{\text{fer}} \right].$$

The only consistent solution is $b_0 = \widetilde{b}_0$, $f_0 = \widetilde{f}_0$ and b_l and f_l with $l \neq 0$ are equal to zero. This can be seen by taking the light cone limit in x_{12} first, which sets to zero all terms of the form P_3^{2l}, with $l > 0$, as well as the fermion terms. In the second line only the boson structures survive, but only the $l = 0$ structure is the same as the one surviving in the first line. This shows that all $\widetilde{b}_l = 0$ for $l > 0$. Repeating this argument we can show it for the other cases. We can now consider also the case when one of the particles has spin zero, or is j_0. Then any of the expressions in (3.13) only allows the $l = 0$ term. Thus, the even structures with only one j_0 are the same as in the free boson theory (the free fermion ones are zero).

For the odd structure the situation is more tricky. Inside the triangle, $s_i \leq s_{i+1} + s_{i-1}$ for $i = 1, 2, 3$, we have the structures that we had before, since (3.8) does not allow any of the three currents to have a non-zero divergence. However, outside the triangle we have new structures that obey (3.11) with a non-zero double trace term. Precisely the existence of these new structures make the whole setup consistent. The current non-conservation identity has the form of the current conservation one but with a non-zero term in the right hand side. Since outside the triangle we had no solutions of the homogeneous equations, this guarantees that the solutions are uniquely fixed in terms of the operators that appear in the right hand side of the conservation laws. Thus, we have unique solutions for these structures. One interesting example of this phenomenon is the correlator

$$\langle J_4 J_2 j_0 \rangle_{\text{odd},\, nc} \propto a_2 \frac{S_3 Q_1^2}{|x_{12}||x_{13}||x_{23}|} \left[Q_1 Q_2 + 4P_3^2 \right] . \tag{3.14}$$

where P_i and Q_i are defined in [41]. This odd correlator would be zero if all currents were conserved. However, using the lack of conservation of the J_4 current, (3.6) , we can derive (3.14) . Clearly only a_2 contributes to it.

As an another example, consider $\langle J_4 J_2 \widetilde{j}_0 \rangle$. Here we can have structures that are parity odd and parity even, the fermion and the "odd" structure respectively. Due to the form of the current non-conservation of J_4, (3.5) , we get

$$\langle J_4 J_2 \widetilde{j}_0 \rangle_{\text{odd, nc}} \propto a_2 \frac{Q_1^2}{|x_{12}||x_{13}||x_{23}|} \left[P_3^4 - 10 P_3^2 Q_1 Q_2 - Q_1^2 Q_2^2 \right] \tag{3.15}$$

This "odd" structure would vanish if J_4 were exactly conserved. Of course, we also have the structure that we get in the free fermion theory for these case, which is parity odd (while (3.15) is parity even). In fact, any correlator of two twist one currents and one $\widetilde{j}_0$, which is parity odd will have the same structure as in the free fermion theory since (3.5) (or its higher spin versions), will not modify it. On the other hand, the parity even ones can be modified. Again we find that a structure that was forced to be zero when the current is exactly conserved can become non-zero when the current is not conserved. Finally, we should mention that any correlator that involves a current and two scalars is uniquely determined by conformal symmetry. In this case the current is automatically conserved. Of course, the three point function of three scalars is also unique. In summary, we constrained the possible structures for various three point functions. These are the boson, fermion and odd structures. When the operator $\widetilde{j}_0$ is involved we can only have fermion or odd structures. In the next section we will constrain the relative coefficients of all of these three point functions.

3.4 Charge non-conservation identities

3.4.1 General story

We use the following technique to constrain the three point function. We start from a three point function $\langle O_1 O_2 O_3 \rangle$. We then insert a J_4 current and take its divergence, which gives us an identity of the form

$$\langle \nabla . J_4(x) O_1 O_2 O_3 \rangle = a_2 \langle J J' O_1 O_2 O_3 \rangle + a_3 \langle J J' J'' O_1 O_2 O_3 \rangle \tag{3.16}$$
$$\sim a_2 \langle J O_1 \rangle \langle J' O_2 O_3 \rangle + a_3 \langle J O_1 \rangle \langle J' O_2 \rangle \langle J'' O_3 \rangle + \text{permutations}$$

This equation is schematic since we dropped derivatives that should be sprinkled on the right hand side. The two point functions in the right hand side are non-zero only if J or J' is the same as one of the operators O_i. Thus the the right hand side is non-zero only when any of the operators O_i is the same as one of the currents that appears in the right hand side of the divergence of J_4 (3.3) (3.5) (3.6) . We have only considered disconnected contributions in the right hand side because those are the only ones that survive to leading order in $1/\widetilde{N}$. Here we used the scaling of the coefficients a_2 and a_3 with $\widetilde{N}$ given in (3.11) . Given this equation, we can now integrate over x on the left and right hand sides. We integrate over a region which includes the whole space except for little spheres, S_i, around each of the points x_i where the operators O_i are inserted. The left hand side of (3.16) contributes only with a boundary term of the form

$$\sum_{i=1}^{3} \langle \int_{S_i} n^j J_{j---} O_1 O_2 O_3 \rangle \tag{3.17}$$

where the integral is over the surface of the little spheres or radii r_i around each of the points and n^j is the normal vector to the spheres.

If the current J_4 were exactly conserved, these integrals would give the charge acting on each

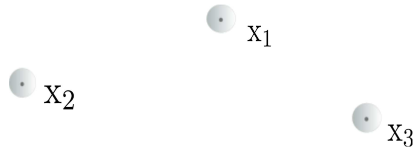


Figure 3.1: The action of the charge is given by the integral of the conserved current over a little sphere around the operator insertion. To derive the pseudo-conservation identity we are integrating the divergence of the current over all space except for these spheres. The boundary terms give the pseudo-charges, which have the same expression in terms of the current as in the conserved case.

of the operators. In our case, the charges are not conserved and the integrals may depend on the radius of the little spheres. This dependence can give rise to divergent terms going like inverse powers of the radius of the spheres. These terms diverge when $r_i \to 0$. These divergent terms should precisely match similar divergent terms that arise in the integral of the right hand side of (3.16) . After matching all the divergent terms we are left with the finite terms in the $r_i \to 0$ limit. These also have to match between the left and right hand sides. Demanding that they match we will get interesting constraints. Notice that, at the order we are working, we do not get any logarithms of r_i since the anomalous dimensions of operators start at higher order of the $1/\widetilde{N}$ expansion. Thus, the separation of the finite and the divergent terms is always unambiguous. Thus, we define a pseudo-charge Q that acts on the operators by selecting the finite part of the above integrals in the small radius limit

$$[Q, O(0)] = \int\limits_{|x|=r} n^j J_{j---} O(0) \Big|_{\text{finite as } r\to 0} \tag{3.18}$$

This action of this pseudo-charge on single trace operators is determined by the three point functions we discussed above. It is also constrained by twist and spin conservation to have a similar structure to the one we had for absolutely conserved currents. For example, on the twist one operators we have

$$[Q, j_s] = \sum_{s'=0}^{s+3} c_{s,s'} \partial^{s-s'+3} j_{s'} \tag{3.19}$$

In concrete computations we found it useful to work in the metric (3.4) and to cut out little "slabs" of width Δx^+ around the operators, instead of cutting out little spheres.

The advantage is that then the integral involves the all minus component of the current. In

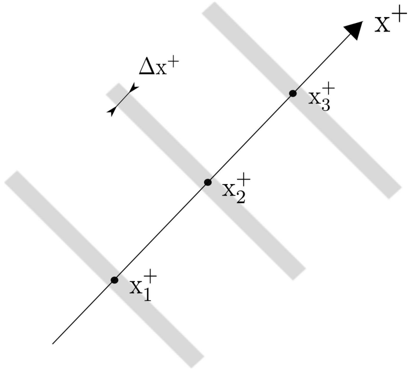


Figure 3.2: Instead of spheres as in 3.1, we can cut out little slabs of width Δx^+ around the insertion point of every operator. The charge is given by integrating current over x^- and y at the edges of these slabs. This simplifies some computations compared to 3.1.

addition, minus and y-derivatives can be integrated by parts or be pulled out of the integral

$$Q_s(x^+) = \int dx^- dy \; j_s(x^-, x^+, y). \tag{3.20}$$

The action of the pseudo-charge on O_i comes from three point functions of the form $\langle J_4 O_i O_k \rangle$. As we mentioned the structures that arise in the boson or fermion theories will continue to produce three point functions where the charges are conserved. The odd correlators can give us something new. The odd structures involving all twist ones fields, such as $\langle j_4 j_{s_2} j_{s_3} \rangle$, are parity odd and do not contribute to the action of Q. This can be seen by setting $y_2 = y_3 = 0$. Then the fact that the three point function is odd under the parity $y_1 \to -y_1$ implies that the integral over y in (3.20) must vanish. This implies that commutator must vanish also for arbitrary $y_2 \neq y_3$ since the two point function structures that could possibly contribute do not vanish for $y_2 = y_3$. When one of the operators O_i has twist two, we can get non-vanishing contributions to the action of the pseudo-charge from odd structures.

In conclusion, after integrating (3.16) we get an expression of the form

$$\begin{aligned}\langle [Q, O_1] O_2 O_3 \rangle + \text{cyclic} = \int d^3x \, [a_2 \langle J(x) O_1 \rangle \langle J'(x) O_2 O_3 \rangle + \\ + \; a_3 \langle J(x) O_1 \rangle \langle J'(x) O_2 \rangle \langle J''(x) O_3 \rangle + \text{permutations}]_{\text{finite}}\end{aligned} \tag{3.21}$$

where the operators O_i are evaluated at x_i. The integral is over the full R^3 after subtracting all the divergent terms that can arise around each of the points x_i. This is the main identity that we will use to relate the various three point functions to each other. We can call it a pseudo-conservation of the pseudo-charges.

3.4.2 Constraints on three point functions with non-zero even spins

In this section we will consider the constraints that arise on three point functions of operators with spins $s_i \geq 2$. We will consider all their indices to be minus. So we take the operators O_i in (3.16) to be j_{s_i}, with $s_i \geq 2$. In this case the action of Q can only produce other twist one fields, which are only single particle states. As we mentioned above the action of Q is determined by three point functions of the form $\langle j_4 j_{s_1} j_{s_2} \rangle$ or $\langle j_4 j_{s_1} j_0 \rangle$. Only the even structures contribute to the charges, and these are the same (up to overall coefficients) as in the free theories, thus the action of the pseudo-charge is well defined and produces

$$[Q, j_s] = c_{s,s-2}\partial^5 j_{s-2} + c_{s,s}\partial^3 j_s + c_{s,s+2}\partial j_{s+2} \tag{3.22}$$

where $\partial = \partial_-$. We can always choose the following normalization conditions

$$c_{2,0} = 1 \ , \qquad c_{2,4} = 1, \quad c_{4,4} = 1, \quad c_{4,6} = 1, \quad c_{s,s+2} = 1, s \geq 2 \tag{3.23}$$

This can be done in all cases, pure fermions, pure bosons, or the interesting theory we are considering. We are not setting the normalization of two point function to one. The two point functions are not used for the time being. The stress tensor is normalized in the canonical way. Notice that, since the normalization of stress tensor is fixed, $c_{2,4}$ does not depend on the normalization of j_4 and is fixed by conformal invariance to 3 (recall $[Q_s, j_2] = (s-1)j_s$ [60]). Thus, we fix it to one by the rescaling freedom in the definition of $Q_4 \to \frac{1}{3}Q_4$. Then we fix $c_{4,4} = 1$ by rescaling j_4 itself. And $c_{s,s+2} = 1$ by rescaling j_{s+2}.

Now, in our case, it is clear that if all $s_i \geq 4$, then the right hand side of (3.21) vanishes. In addition, if one or more, of the s_i is equal to two, then the following happens. In that case the two current terms in (3.5), (3.6) could contribute, since we can contract the J_2 in the divergence of J_4 with the s_i that is equal to two. It is useful to recall the two point functions of various components of the stress tensor

$$\begin{aligned}
\langle J_{--}(x)J_{--}(0)\rangle &\propto \frac{(x^+)^4}{(x^+x^- + y^2)^5} = \frac{1}{4!}\partial_-^4 \frac{1}{x^+x^- + y^2}, \\
\langle J_{--}(x)J_{-y}(0)\rangle &\propto 2\frac{(x^+)^3 y}{(x^+x^- + y^2)^5} = \frac{1}{4!}\partial_-^3 \partial_y \frac{1}{x^+x^- + y^2}, \\
\langle J_{-y}(x)J_{-y}(0)\rangle &\propto -\frac{(x^+)^2(x^+x^- - 3y^2)}{(x^+x^- + y^2)^5} = \frac{1}{4!}\left(\partial_-^2\partial_y^2 \frac{1}{x^+x^- + y^2} - 2\partial_-^2 \frac{1}{(x^+x^- + y^2)^2}\right).
\end{aligned} \tag{3.24}$$

where the equations are true up to a normalization factor common to all three equations. If we look at the first term in the right hand side of (3.21) in the case that $O_1 = j_2$, then we can use the above two point function (3.24) . We can integrate by parts all derivatives so that they act on the two point function. Then we can write them as derivatives acting on x_i and pull them out of the integral[5]. Now the result depends on whether we are dealing with the quasi-fermion or quasi-boson cases. In the quasi-boson case, it is possible to check that the particular combination of currents that appear in the two current term in (3.6) is, after integrating by parts, $\partial_y J_{2\,--} - \partial_- J_{2\,y-}$. When this is contracted with $j_2 = J_{2\,--}$ we get zero after using (3.24) . Thus, the right hand side of (3.21) vanishes in the quasi-boson case. In the quasi-fermion case, we end up having to compute (the ∂_-^5 derivative) of an integral of the form

$$\int d^3x \frac{1}{|x - x_1|^2} \langle \widetilde{j}_0(x) O_2(x_2) O_3(x_3) \rangle \propto \langle j_0^{\text{eff}}(x_1) O_2(x_2) O_3(x_3) \rangle \tag{3.25}$$

The factor $\frac{1}{|x-x_1|^2}$ is exactly the one that makes the integral conformal covariant. It gives a result that effectively transforms as the three point function with the insertion of an operator of dimension $\Delta = 1$ at x_1. We have denoted this in terms of an effective operator, j_0^{eff}, of dimension one. We do not have any real operator of this kind in the theory, this is just a mnemonic to help us keep track of the resulting integral. The net result can be expressed by saying that the action of Q on j_2 in the quasi-fermion theory can also produce a j_0^{eff}. In the free boson theory the action of Q on j_2 produces j_0 (plus other things). This is not true in the free fermion theory, which does not contain a j_0. However, in the quasi-fermion theory, with the pseudo-conservation property (3.5) (3.21) , we get terms in these identities that would be identically to what we would get if we had allowed a j_0^{eff} in the action of Q on j_2. This term is crucial in order to allow a free boson structure in the pseudo-conservation identities. In other words, we define a new effective pseudo-charge $\widehat{Q}$, which has the same action as Q on all currents with spin $s > 2$, but it is

$$[\widehat{Q}, j_2] = c_{2,0}\partial^5 j_0^{eff} + c_{2,2}\partial^3 j_2 + c_{2,4}\partial j_4 \tag{3.26}$$

The net result is that when we consider the pseudo-conservation identities acting on $\langle j_{s_1} j_{s_2} j_{s_3} \rangle$ becomes identical to a charge conservation identity for $\widehat{Q}$. Note also that after adding j_0^{eff} in (3.26) the action of $\widehat{Q}$ is essentially the same as the action of Q in the quasi-boson theory. Thus, we can treat these two cases in parallel, after we remember that j_0^{eff} is not a real operator but just the

[5]If we work with the "slabs" described in figure 3.2, together with the usual $i\epsilon$ prescription, it is clear that these operations do not produce boundary terms since we only have ∂_- or ∂_y derivatives.

integral in (3.25) .

Next we write all twist one three point functions as

$$\langle j_{s_1} j_{s_2} j_{s_3} \rangle = \alpha_{s_1,s_2,s_3} \langle j_{s_1} j_{s_2} j_{s_3} \rangle_{\rm bos} + \beta_{s_1,s_2,s_3} \langle j_{s_1} j_{s_2} j_{s_3} \rangle_{\rm fer} + \gamma_{s_1,s_2,s_3} \langle j_{s_1} j_{s_2} j_{s_3} \rangle_{\rm odd} \tag{3.27}$$

Here the boson and fermion ones are the three point functions for a single real boson and a single Majorana fermion, in the normalization of the currents set by (3.23) . The normalization of the odd piece is fixed so that the identities we describe below are true. In the quasi-fermion theory we are also including in (3.27) the case with $\langle j_0^{eff} j_{s_1} j_{s_2} \rangle$, where only the boson and odd structures are non-zero.

We now consider the charge conservation identities for various cases. An important property of these identities is that each identity separates into three independent equations which relate only the boson structures to each other, the fermion structures to each other and the odd structures to each other. Each equation involves sums of objects of the form $\partial_i^n \langle j_{s_1} j_{s_2} j_{s_3} \rangle_{\rm str}$, where str runs over boson, fermion and odd. These equations are such that the coefficient in front of each of the terms is fixed relative to all the other coefficients. In other words, only an overall constant is left undetermined. See appendix B. This was proven for the boson and fermion terms in [60]. We have not proved this for the odd case. However, given the existence of the Chern-Simons matter theories [44, 45] we know that at least one solution definitely exists! For low values of the spins we have explicitly analyzed the equations and the solution is definitely unique, in the sense that all relative coefficients are fixed. We think that this is likely to be true for all cases. Focusing first on the equations that constrain the boson structures we get that the equations arising from pseudo-charge conservation are

$$\begin{aligned} Q\langle j_2 j_2 j_2 \rangle : &\quad \widetilde{c}_{2,2}\alpha_{222} = \alpha_{224} = \alpha_{022} \\ Q\langle j_2 j_2 j_4 \rangle : &\quad \widetilde{c}_{2,2}\alpha_{224} = \widetilde{c}_{4,4}\alpha_{224} = \alpha_{244} = \widetilde{c}_{4,2}\alpha_{222} = \alpha_{226} = \alpha_{024} \\ Q\langle j_2 j_4 j_4 \rangle : &\quad \widetilde{c}_{2,2}\alpha_{244} = \widetilde{c}_{4,4}\alpha_{244} = \alpha_{444} = \widetilde{c}_{4,2}\alpha_{224} = \alpha_{246} = \alpha_{044} \\ Q\langle j_2 j_2 j_6 \rangle : &\quad \widetilde{c}_{2,2}\alpha_{226} = \widetilde{c}_{6,6}\alpha_{226} = \alpha_{246} = \widetilde{c}_{6,4}\alpha_{224} = \alpha_{228} = \alpha_{026} \end{aligned} \tag{3.28}$$

and we can continue in this way. We are defining $\widetilde{c}_{s,s'}$ to be the ratio

$$\widetilde{c}_{s,s'} = \frac{c_{s,s'}}{c_{s,s'}^{\rm free\ boson}} \tag{3.29}$$

with both in the normalization (3.23) .

We now can start solving (3.28) . We see that the equations (3.28) fix all the α's in terms of α_{222} and the c's. In addition, we get multiple equations for the same α's. This fixes the $\widetilde{c}$'s. For example, we start getting things like

$$\begin{aligned}
&\alpha_{224} = \widetilde{c}_{2,2}\alpha_{222} \\
&\alpha_{244} = \widetilde{c}_{4,2}\alpha_{222} \ , \qquad \widetilde{c}_{2,2} = \widetilde{c}_{4,2} \ , \qquad \alpha_{226} = \widetilde{c}_{4,2}\alpha_{222} \\
&\alpha_{444} = \widetilde{c}_{4,2}\widetilde{c}_{2,2}\alpha_{222} \ , \quad \widetilde{c}_{2,2} = \widetilde{c}_{4,4} \ , \quad \alpha_{244} = \widetilde{c}_{4,2}\alpha_{222} \ , \quad \alpha_{246} = \widetilde{c}_{4,2}\widetilde{c}_{2,2}\alpha_{222}.
\end{aligned} \tag{3.30}$$

where each line in (3.30) comes from the corresponding line in (3.28) . Using the fact that $\widetilde{c}_{4,4} = 1$ we see that $\widetilde{c}_{2,2} = 1$, and also $\widetilde{c}_{4,2} = 1$, and so on. So all the $\widetilde{c}_{s,s'} = 1$. We also find that all the α's are also fixed to be equal to α_{222}. If we did the same for the free fermions we would also obtain the same pattern if we define

$$\widetilde{c}_{s,s'} = \frac{c_{s,s'}}{c_{s,s'}^{\text{free fermion}}} \tag{3.31}$$

Then we get that all $\beta_{s_1,s_2,s_3} = \beta_{222}$ and all $\widetilde{c}_{s,s'} = 1$. One subtlety is that we have defined the $\widetilde{c}'s$ differently for the bosons than for the fermions. So we can only hope to get both structures present only if

$$c_{s,s'}^{\text{free boson}} = c_{s,s'}^{\text{free fermion}} \tag{3.32}$$

This can be checked by using form factors (see Appendix C). Of course, the mere existence of the Chern-Simons matter theories implies that this is true. The conclusion from this analysis is that all the α's are equal to α_{222}. Analogously, all β's are equal to β_{222}. We will also need the coefficients of the two point functions n_s. In a theory of a single free boson or single free fermion, with the normalization conditions (3.23) , these are given by n_s^{bos} and n_s^{fer} respectively. We do not need their explicit forms, but they can be computed in the free theories. We can now determine the two point functions in the full theory by demanding that the stress tensor Ward identities are obeyed. In other words, from the previous discussion we know that $\alpha_{2ss} = \alpha_{222}$ and $\beta_{2ss} = \beta_{222}$. In addition, the Ward identity of the stress tensor relates this to the two point function. More precisely,

$$n_s = \alpha_{222} n_s^{bos} + \beta_{222} n_s^{fer} \tag{3.33}$$

Notice that $n_2^{bos} = n_2^{fer}$ according to formulas (5.1) and (5.7) of [49] [6]. Thus we find

$$\widetilde{n}_2 \equiv \frac{n_2}{n_2^{bos}} = \alpha_{222} + \beta_{222} \tag{3.34}$$

Notice that the analysis so far is equivalently valid for the theories of quasi-bosons and the theory of quasi-fermions. Of course, in the latter case, whenever a j_0 appeared, it should be interpreted as j_0^{eff}. Thus, so far, we have written all correlators in terms of three undetermined coefficients α_{222}, β_{222} and γ_{222}. As an aside we should note that when we solve the equations for the $\gamma_{s_1 s_2 s_3}$ we need to use some of the odd correlators that are outside the triangle. These are possible thanks to the non-conservation of the currents.

3.4.3 Closing the chain. The Quasi-fermion case.

In this section we consider current conservation identities in the case that one of the operators is a twist two fields such as $\widetilde{j}_0$ or $J_{2\,-y}$. Let us start by discussing the possible action of the pseudo-charge Q on J_{-y}. All single trace operators that could appear are already fixed by the correspondent three point functions to be the same as in free fermion theory. The most general expression involving double trace terms takes the form

$$[Q, J_{-y}] = \partial^4 \widetilde{j}_0 + \partial^3 J_{-y} + \partial J_{4\,---y} + \frac{\chi}{\widetilde{N}} j_2 j_2. \tag{3.35}$$

The double trace term comes with an $\frac{1}{\widetilde{N}}$ factor because it enters in the pseudo-conservation identity with the $\langle jj \rangle \langle jj \rangle \sim \widetilde{N}^2$ factor. We have normalized $\widetilde{j}_0$ by setting the first coefficient to one[7]. One can check that χ should be set to zero by considering $\langle j_2 j_2 J_{-y} \rangle$ pseudo-conservation identity [8]. In addition, we write the possible correlators of $\widetilde{j}_0$ in terms of free fermion correlators, introducing a $\beta_{\widetilde{0} s_1 s_2}$. This is the coefficient that multiplies $\langle \widetilde{j}_0 j_{s_1} j_{s_2} \rangle_{\text{fer}}$, which is the correlator in a theory of a single Majorana fermion in the normalization of $\widetilde{j}_0$ in (3.35) . Note, that these free fermion correlators are parity odd. We also introduce a $\gamma_{\widetilde{0} s_1 s_2}$ which multiply the "odd" structures, which are parity even structures involving $\widetilde{j}_0$. These structures are more subtle since they can be affected by the violation of current conservation, as in (3.15) . An additional issue we should discuss is the type of contribution we expect from the right hand side of (3.21) when the operator is J_{-y}. The same reasoning we used around (3.25) , together with (3.24) tells us that we also effectively produce

[6] There the formula for the Dirac fermion is written. Here we consider a Majorana one.

[7] This is not possible for the critical $O(N)$ theory since in that case parity prevents $\widetilde{j}_0$ in the right hand side. We will obtain this case as a limit. It can also be analyzed directly by considering the action of Q on $\widetilde{j}_0$, etc.

[8] This double trace structure can arise in a covariant way as $[Q, J_{2\,\mu\nu}] = \epsilon_{-\mu\rho} J^{\rho}_{\ -} J_{\nu -} + \epsilon_{-\nu\rho} J^{\rho}_{\ -} J_{\mu -}$.

a $j_0^{\rm eff}$. Thus, the net result is that we simply should add terms involving j_0^{eff} in (3.35) , and treat the charge as conserved. This is necessary for getting all the identities to work. We then find that $\beta_{\widetilde{0}s_1s_2} = \beta_{222}$ and $\gamma_{\widetilde{0}s_1s_2} = \gamma_{222}$. Interestingly, in this case, in order to satisfy the pseudo-charge conservation identities for the odd part, we need both $\widetilde{j}_0$ as well as j_0^{eff}. Let us now consider the three point function $\langle j_0^{eff} j_2 j_2\rangle$. We see that its definition via (3.25) involves a three point function given by $\beta_{\widetilde{0}22}$ and $\gamma_{\widetilde{0}22}$. But we have just fixed these coefficients. Thus, going back to the first line (3.28) , we note that α_{022} is really the coefficient of $\langle j_0^{eff} j_2 j_2\rangle_{\rm bos}$. We get this structure from doing the integral (3.25) with the "odd" structure for the three point function. Using that α_{022} was fixed to α_{222} we get the equation

$$\alpha_{222} = x_2 a_2 n_2 \gamma_{\widetilde{0}22} = \widetilde{x}_2 a_2(\alpha_{222} + \beta_{222})\gamma_{222} \tag{3.36}$$

Similarly, the odd structure gives

$$\gamma_{222} = x_1 a_2 n_2 \beta_{\widetilde{0}22} = \widetilde{x}_1 a_2(\alpha_{222} + \beta_{222})\beta_{222} \tag{3.37}$$

where x_1 and x_2 are two calculable numerical coefficients, given by doing the integral in (3.25) , etc. We used (3.34) and $\widetilde{x}_i = x_i/n_2^{\rm bos}$. These are two equations for four unknowns ($\alpha_{222},\ \beta_{222},\ \gamma_{222},\ a_2$). So we have two undetermined coefficients which are $\widetilde{N}$ and $\widetilde{\lambda}$. We can define $\widetilde{N} = \widetilde{n}_2 = \alpha_{222} + \beta_{222}$ and $\widetilde{\lambda} = a_2\widetilde{N}/\sqrt{\widetilde{x}_1\widetilde{x}_2}$. With these definitions we get

$$\alpha_{s_1s_2s_3} = \widetilde{N}\frac{\widetilde{\lambda}^2}{(1+\widetilde{\lambda}^2)}, \quad \beta_{s_1s_2s_3} = \widetilde{N}\frac{1}{(1+\widetilde{\lambda}^2)}\ , \quad \gamma_{s_1s_2s_3} = \widetilde{N}\frac{\widetilde{\lambda}}{(1+\widetilde{\lambda}^2)}. \tag{3.38}$$

Here we are considering three twist one fields. Similar equations are true for correlators involving one $\widetilde{j}_0$ field

$$\beta_{s_1s_2\widetilde{0}} = \widetilde{N}\frac{1}{(1+\widetilde{\lambda}^2)}\ , \quad \gamma_{s_1s_2\widetilde{0}} = \widetilde{N}\frac{\widetilde{\lambda}}{(1+\widetilde{\lambda}^2)}. \tag{3.39}$$

We will discuss the correlation functions that involve more than one $\widetilde{j}_0$ separately. In both equations, (3.38) (3.39), we can choose the ($\widetilde{\lambda}$ independent) numerical normalization of the odd correlators so that we can replace $\sim$ by an equality.

The final conclusion is that we expressed all the correlators of the currents in terms of just two parameters $\widetilde{N}$ and $\widetilde{\lambda}$. Since the fermions plus Chern-Simons theory has precisely two parameters, N and k we conclude that we exhausted all the constraints. Our analysis was based only on general symmetry consideration and does not allow us to find the precise relation between the parameters. However, in the 't Hooft limit we expect[9]

$$\widetilde{N} = N f(\lambda) , \qquad \widetilde{\lambda} = h(\lambda) = d_1 \lambda + d_3 \lambda^3 + \cdots \tag{3.40}$$

where $f(\lambda) = f_0 + f_2\lambda^2 + \cdots$. Where we used the symmetry of the theory under $\lambda \to -\lambda$ (or $k \to -k$), together with parity. Note that the function f encodes how the two point function of the stress tensor depends on λ.

3.4.4 Closing the chain. The Quasi-boson case.

In the quasi-boson theory we will again consider the charge non-conservation identity on $\langle j_2 j_2 J_{y-}\rangle$. Again we need to write the most general action of Q on J_{3-} involving double trace terms. It takes the form

$$\begin{aligned}[Q, J_{-y}] &= \partial^4 \partial_y j_0 + \partial^3 J_{-y} + \partial J_{---y} \\ &+ \frac{\chi_1}{\widetilde{N}} j_2 j_2 + \frac{\chi_2}{\widetilde{N}} \partial^2 j_0 j_2 + \frac{\chi_3}{\widetilde{N}} \partial^4 j_0 j_0\end{aligned} \tag{3.41}$$

We have normalized j_0 by setting the first coefficient to one[10]. Notice also that $\chi_i \sim O(\frac{1}{\widetilde{N}})$ to contribute at leading order in $\widetilde{N}$. We fix $\chi_1 = 0$ as in the case of fermions. The presence or absence of $\chi_{2,3}$ is not important for any of the arguments.

In addition, we need to consider the contribution of J_{-y} to the right hand side of (3.21) . In other words, we will need to consider the right hand side of (3.21) when the operator Q_1 is J_{-y}. Using (3.24) , we find (derivatives of) an integral of the form

$$\int d^3x \frac{1}{|x - x_1|^4} \langle j_0(x) O_2(x_2) O_3(x_3)\rangle \sim \langle \widetilde{j}_0^{\text{eff}}(x_1) O_2(x_2) O_3(x_3)\rangle \tag{3.42}$$

Here we used that, again, this is a conformal integral which behaves as a correlator with a scalar of weight $\Delta = 2$ at x_1. We have denoted this by introducing a ficticious operator $\widetilde{j}_0^{\text{eff}}$. This is not an operator that exists in the theory, but it is appearing in three point functions in the same way as an operator of this form. Thus, the net effect of the action of Q on J_{-y} includes also this operator $\widetilde{j}_0^{eff}$ in the right hand side of (3.35) . With all these features, we now have a situation which is rather

[9] For the theory considered in [44] $f_0 = 2$ and $d_1 = \frac{\pi}{2}$.

[10] Again, this is not possible in the critical $O(N)$ Gross-Neveu theory. But we will obtain this case as a limit. It can also be analyzed directly through a slightly longer route.

similar to the one we had in the quasi-fermion case and we can relate the different coefficients

$$\beta_{222} = y_1 a_2 n_2 \gamma_{022} = \widetilde{y}_1 a_2(\alpha_{222} + \beta_{222})\gamma_{222}. \tag{3.43}$$
$$\gamma_{222} = y_2 a_2 n_2 \alpha_{022} = \widetilde{y}_2 a_2(\alpha_{222} + \beta_{222})\alpha_{222}.$$

Again y_1 and y_2 are some numbers. Defining again $\widetilde{N} = \widetilde{n}_2 = (\alpha_{222} + \beta_{222})$ and $\widetilde{\lambda} = a_2\widetilde{N}/\sqrt{\widetilde{y}_1\widetilde{y}_2}$ we get

$$\alpha_{s_1s_2s_3} = \widetilde{N}\frac{1}{(1+\widetilde{\lambda}^2)}, \quad \beta_{s_1s_2s_3} = \widetilde{N}\frac{\widetilde{\lambda}^2}{(1+\widetilde{\lambda}^2)}\ , \quad \gamma_{s_1s_2s_3} = \widetilde{N}\frac{\widetilde{\lambda}}{(1+\widetilde{\lambda}^2)}. \tag{3.44}$$

Here at least two of the spins should be bigger than zero, $s_i \geq 2$. We will discuss the correlation functions that involve more than one j_0 separately.

3.4.5 Three point functions involving scalars. The quasi-fermion case.

Fixing the three point functions in the scalar sector involves several new subtleties that were absent before. We describe them in detail in the Appendix D and here sketch the method and present the results. We have already discussed how to fix three point functions which include one scalar operator. We used the pseudo-conservation identity on $\langle J_{-y} j_{s_1} j_{s_2}\rangle$ to get (3.39) . To proceed it is necessary to specify how Q acts on $\widetilde{j}_0$. As explained in Appendix D the result is

$$[Q, \widetilde{j}_0] = \partial^3\widetilde{j}_0 + \frac{1}{1+\widetilde{\lambda}^2}\partial\left[\partial_- j_{-y} - \partial_3 j_{--}\right] \tag{3.45}$$

here the interesting new ingredient is a $\frac{1}{1+\widetilde{\lambda}^2}$ factor. This is obtained by inserting arbitrary coefficients and fixing them by analyzing the $\langle j_4 j_2 \widetilde{j}_0\rangle$ three point function. To fix the correlators with two scalars we consider the pseudo-charge conservation identity that is generated by acting with Q on $\langle \widetilde{j}_0 \widetilde{j}_0 j_s\rangle$. This identity involves tricky relations (see Appendix D) between different three point functions. After the dust settles we get that

$$\beta_{s\widetilde{0}\widetilde{0}} = \beta_{222}\ , \qquad\qquad \gamma_{s\widetilde{0}\widetilde{0}} = \gamma_{222} \tag{3.46}$$

The last step is to consider the action of Q on $\langle \widetilde{j}_0 \widetilde{j}_0 \widetilde{j}_0\rangle$ WI to get

$$\gamma_{\widetilde{0}\widetilde{0}\widetilde{0}} = 0. \tag{3.47}$$

Note that $\beta_{\tilde{0}\tilde{0}\tilde{0}} = 0$ by definition, since this correlator vanishes in the free fermion theory (due to parity). From these three point function it is also possible to extract the normalization of the two point function $\langle \tilde{j}_0 \tilde{j}_0 \rangle$. This is two point function is related by a Ward identity to $\langle j_2 \tilde{j}_0 \tilde{j}_0 \rangle$. We then get

$$\widetilde{n}_{\tilde{0}} \equiv \frac{n_{\tilde{0}}}{n_{\tilde{0}}^{\text{free fermion}}} = \beta_{2\tilde{0}\tilde{0}} = \beta_{222} = \widetilde{N}\frac{1}{1+\widetilde{\lambda}^2} \tag{3.48}$$

This can be used, together with (3.5) , to compute the anomalous dimension for the spin four current, as explained in appendix A.

3.4.6 Three point functions involving scalars. The quasi-boson case.

For the quasi-boson sector the story is almost identical. We put details of the analysis in the Appendix E and here again sketch the general idea and present the results. We have already discussed above how to fix three point functions which include one scalar operator. Then the charge conservation identity on $\langle j_2 j_{s_1} j_{s_2} \rangle$ fixes

$$\alpha_{0 s_1 s_2} = \alpha_{222} \ , \qquad \gamma_{0 s_1 s_2} = \gamma_{222}. \tag{3.49}$$

For the action of the charge on the scalar we get

$$[Q, j_0] = \partial^3 j_0 + \frac{1}{1+\widetilde{\lambda}^2}\partial j_2. \tag{3.50}$$

By considering the pseudo-conservation of $\langle j_s j_0 j_0 \rangle$ when $s > 2$ we get

$$\alpha_{s00} = \beta_{222} \ , \qquad \gamma_{s00} = \gamma_{222}. \tag{3.51}$$

We now need to consider the pseudo-conservation identities for $\langle j_2 j_0 j_0 \rangle$ and $\langle j_0 j_0 j_0 \rangle$.

A new feature of these two cases is that the triple trace terms in (3.6) contribute. Analyzing these we obtain that (3.51) is also true for $s = 2$. The triple trace terms contribute as follows. Let us first consider the pseudo-conservation identity for $\langle j_0 j_0 j_0 \rangle$. The triple trace non-conservation term takes the form

$$a_3 n_0^3 \sum_{i=1}^{3} \partial_i^3 \langle j_0(x_1) j_0(x_2) j_0(x_3) \rangle. \tag{3.52}$$

where n_0 is the coefficient in the two point function for the scalar. Thus, we get the equation

$$\alpha_{000} = \frac{1}{1+\tilde{\lambda}^2}\alpha_{222} + z_1 a_3 n_0^3 \tag{3.53}$$

where z_1 is a numerical constant and b_3 is the coefficient in (3.6) . Notice that the double trace deformation does not influence this computation. This fact is established in the Appendix C. The triple trace term in the pseudo-conservation identity for $\langle j_2 j_0 j_0 \rangle$ is

$$a_3' n_2 n_0^2 \partial_1^5 \langle j_0(x_1) j_0(x_2) j_0(x_3) \rangle. \tag{3.54}$$

leading to

$$\alpha_{000} = \frac{1}{1+\tilde{\lambda}^2}\alpha_{222} + z_2 a_3' n_0^2 n_2 + ... \tag{3.55}$$

with z_2 a numerical constant and the dots stands for the contribution of the double trace non-conservation piece whose coefficients are already known. Importantly, we conclude that two triple trace deformations are not independent. In other words, a_3 and a_3' are related by equating (3.55) and (3.53) . Thus, we recover the known counting of marginal deformation of free boson in $d = 3$, namely there are two parameters. Microscopically, one corresponds to the Chern-Simons coupling and the second one to adding a $(\vec{\phi}.\vec{\phi})^3$ operator. On the Vasiliev theory side, the freedom to add this ϕ^6 deformation translates into the fact that we can choose a non-linear boundary conditions for the scalar which preserves the conformal symmetry. These were discussed in a similar situation in [63]. In our context, we have a scalar of mass $(mR_{AdS})^2 = -2)$ which at infinity decays as $\phi = \alpha/r + \beta/r^2$. Then the boundary condition that corresponds to adding the $\lambda_6\phi^6$ deformation is $\beta = \lambda_6\alpha^2$ [63]. Also the whole effect of the presence of the triple trace deformations, at the level of three point functions, is to change $\langle j_0 j_0 j_0 \rangle$. Finally, the two point function of j_0 can be fixed by using the usual stress tensor Ward identity via $\langle j_2 j_0 j_0 \rangle$. We obtain

$$\tilde{n}_0 \equiv \frac{n_0}{n_0^{\text{free boson}}} = \alpha_{222} = \tilde{N}\frac{1}{1+\tilde{\lambda}^2} \tag{3.56}$$

Recall that the normalization of j_0 was given by setting $c_{2,0} = 1$ in (3.23) .

3.4.7 Comments about higher point correlation functions

We can wonder whether we can determine higher point correlation functions. It seems possible to use the same logic. Namely, inserting $\nabla.J_4$ into an n point function and then integrating as in (3.21) . This relates the action of Q on an n point function to integrals of disconnected correlators. These integrals involve also n point functions. (Recall that for three point functions the integrals involved other three point functions). When the charge is conserved this is expected to fix the connected correlation uniquely to that of the free theory. This was done explicitly in [60] for the action of Q on $\langle j_0 j_0 j_0 j_0 \rangle$. Now that the right hand side is non-zero, we still expect this to fix uniquely the correlator, though we have not tried to carry this out explicitly. It is not totally obvious that this will fix the correlators because the integral terms also involve n point functions. But it seems reasonable to conjecture that this procedure would fix the leading order connected correlator for all n point functions of single particle operators. It would be interesting to see whether this is indeed true!

3.5 Final results

In this section we summarize the results for the three point functions. The normalization of the stress tensor is the canonical one. The normalization of the charge is $Q = \frac{1}{3}\int j_4$. This sets $c_{2,4} = 1$ in (3.22) . The normalization of j_4 is fixed by setting $c_{4,4} = 1$. Then all other operators are normalized by setting $c_{s,s+2} = 1$. The operator $\widetilde{j}_0$ is normalized by setting $c_{2,\widetilde{0}} = 1$ in (3.35) . We have the two point functions

$$\langle j_s(x_1) j_s(x_2) \rangle = n_s \frac{(x_{12}^+)^{2s}}{|x_{12}|^{4s+2}}. \tag{3.57}$$

For three point functions

$$\begin{aligned} \langle j_{s_1}(x_1) j_{s_2}(x_2) j_{s_3}(x_3) \rangle = \alpha_{s_1 s_2 s_3} \langle j_{s_1}(x_1) j_{s_2}(x_2) j_{s_3}(x_3) \rangle_{\text{bos}} + \\ + \beta_{s_1 s_2 s_3} \langle j_{s_1}(x_1) j_{s_2}(x_2) j_{s_3}(x_3) \rangle_{\text{fer}} + \gamma_{s_1 s_2 s_3} \langle j_{s_1}(x_1) j_{s_2}(x_2) j_{s_3}(x_3) \rangle_{\text{odd}} \end{aligned} \tag{3.58}$$

where the $_{\text{bos}}$ and $_{\text{fer}}$ denote the three point functions in the theory of free boson and free fermion in the normalization of currents described above. Their functional form can be found in [41]. The odd generating functional for the spins inside the triangle can be found in appendix B of [60]. Outside the triangle the odd correlation functions are the ones that satisfy the double trace deformed non-conservation equations. We do not know their explicit form in general but nevertheless we know that they exist and know how the dependence on the coupling will enter. We fix the numerical normalization of the odd pieces to be such that the pseudo-charge conservation identities are obeyed. In a similar fashion we define the correlators involving a $\tilde{j}_0$ (j_0) operator, except that there is no free boson (fermion) structure.

3.5.1 Quasi-fermion theory

The interacting theory two point functions are given by

$$\begin{aligned} n_s &= \widetilde{N}\frac{\tilde{\lambda}^2}{1+\tilde{\lambda}^2}n_s^{\text{free boson}} + \widetilde{N}\frac{1}{1+\tilde{\lambda}^2}n_s^{\text{free fermion}} = \widetilde{N}n_s^{\text{free boson}}, \quad s \geq 2, \\ n_{\tilde{0}} &= \widetilde{N}\frac{1}{1+\tilde{\lambda}^2}n_{\tilde{0}}^{\text{free fermion}} \end{aligned} \tag{3.59}$$

where $n_s^{\text{free boson}}$ and $n_s^{\text{free fermion}}$ are two point functions computed in the theory of single free boson or single fermion with normalization of operators such that (3.23) (and (3.35)) holds. We also used in the first line the fact that $n_s^{\text{free boson}} = n_s^{\text{free fermion}}$ in the normalization that we adopted. This is explained in the appendix C. The three point functions in the interacting theory are then given by

$$\begin{aligned} \alpha_{s_1s_2s_3} &= \widetilde{N}\frac{\tilde{\lambda}^2}{1+\tilde{\lambda}^2}, \qquad \beta_{s_1s_2s_3} = \widetilde{N}\frac{1}{1+\tilde{\lambda}^2}, \qquad \gamma_{s_1s_2s_3} = \widetilde{N}\frac{\tilde{\lambda}}{1+\tilde{\lambda}^2} \\ \beta_{s_1s_2\tilde{0}} &= \widetilde{N}\frac{1}{1+\tilde{\lambda}^2}, \qquad \gamma_{s_1s_2\tilde{0}} = \widetilde{N}\frac{\tilde{\lambda}}{1+\tilde{\lambda}^2} \\ \beta_{s_1\tilde{0}\tilde{0}} &= \widetilde{N}\frac{1}{1+\tilde{\lambda}^2} \\ \gamma_{\tilde{0}\tilde{0}\tilde{0}} &= 0 \end{aligned} \tag{3.60}$$

All coefficients not explicitly written do not appear because there is no corresponding structure. The two parameters are defined as follows. We take the stress tensor to have a canonical normalization. We then set

$$\begin{aligned} \widetilde{N} &= \frac{n_2}{n_2^{\text{free boson}}}, \\ \tilde{\lambda}^2 &= \frac{\alpha_{222}}{\beta_{222}}. \end{aligned} \tag{3.61}$$

From the bounds on energy correlators discussed in [60] it follows that $\tilde{\lambda}^2 \geq 0$. Note that one would then find that $\tilde{\lambda} \propto a_2\widetilde{N}$, with a_2 defined in (3.5) . We have also computed the anomalous dimension of the spin four current (see appendix A)

$$\tau_4 - 1 = \frac{32}{21\pi^2}\frac{1}{\widetilde{N}}\frac{\widetilde{\lambda}^2}{(1+\widetilde{\lambda}^2)} \tag{3.62}$$

3.5.2 Quasi-boson theory

The interacting theory two point functions are given by

$$\begin{aligned} n_s &= \widetilde{N}\frac{1}{1+\widetilde{\lambda}^2}n_s^{\text{free boson}} + \widetilde{N}\frac{\widetilde{\lambda}^2}{1+\widetilde{\lambda}^2}n_s^{\text{free fermion}} = \widetilde{N}n_s^{\text{free boson}}\,, \qquad s \geq 2, \\ n_0 &= \widetilde{N}\frac{1}{1+\widetilde{\lambda}^2}n_0^{\text{free boson}} \end{aligned} \tag{3.63}$$

and the three point functions are

$$\begin{aligned} \alpha_{s_1s_2s_3} &= \widetilde{N}\frac{1}{1+\widetilde{\lambda}^2}, \qquad \beta_{s_1s_2s_3} = \widetilde{N}\frac{\widetilde{\lambda}^2}{1+\widetilde{\lambda}^2}\,, \qquad \gamma_{s_1s_2s_3} = \widetilde{N}\frac{\widetilde{\lambda}}{1+\widetilde{\lambda}^2} \\ \alpha_{s_1s_20} &= \widetilde{N}\frac{1}{1+\widetilde{\lambda}^2}, \qquad \gamma_{s_1s_20} = \widetilde{N}\frac{\widetilde{\lambda}}{1+\widetilde{\lambda}^2} \\ \alpha_{s_100} &= \widetilde{N}\frac{1}{1+\widetilde{\lambda}^2} \\ \alpha_{000} &= \widetilde{N}\frac{1}{(1+\widetilde{\lambda}^2)^2} + z_1a_3n_0^3 \end{aligned} \tag{3.64}$$

were we have separated out the correlators involving the scalar. Two of the parameters are defined by (taking into account that stress tensor has a canonical normalization)

$$\begin{aligned} \widetilde{N} &= \frac{n_2}{n_2^{\text{free boson}}}, \\ \widetilde{\lambda}^2 &= \frac{\beta_{222}}{\alpha_{222}}. \end{aligned} \tag{3.65}$$

Again one can see that $\widetilde{\lambda}^2 \geq 0$ from the bounds on energy correlators discussed in [60]. A third parameter can be introduced which is the combination involving a_3 that shifts the three scalar correlator in (3.64) . Again we find that $\widetilde{\lambda} \propto a_2\widetilde{N}$ in (3.6) . Thus we find a three parameter family of solutions, as expected to leading order in the large N limit.

3.5.3 The critical point of $O(N)$

The critical point of $O(N)$ is a theory that we obtain from the free boson theory by adding a j_0^2 interaction and flowing to the IR, while tuning the mass of the scalars to criticality. Then the operator j_0 in the IR gets to have dimension two, for large N. Thus, we can view it as a $\widetilde{j}_0$ scalar operator of dimension two. This is a theory that fits into the quasi-fermion case. Namely the divergence of J_4 is given by (3.5) . By the way, if we assume that the IR limit has a scalar operator of dimension different from one, then since (3.5) is the only expression we can write down for the divergence of the current, we conclude that the operator has to have dimension two. Notice that in this case $\widetilde{j}_0$ is parity even and (3.5) is consistent with parity. Thus, we could in principle do the same analysis as above. The only point where something different occurs is at (3.35) where $\widetilde{j}_0$ cannot appear in the right hand side, since it is inconsistent with the parity of the theory. In addition, all parity odd correlators should be set to zero. In this case, parity implies that

$$[Q, \widetilde{j}_0] = \partial^3 \widetilde{j}_0 \tag{3.66}$$

In principle, we have an arbitrary coefficient in this equation but the coefficient is then fixed by considering various charge conservation identities we mention below. When we write the conservation identity for $\langle j_2 j_2 j_2 \rangle$ we will get the correlator $\langle \widetilde{j}_0 j_2 j_2 \rangle$ as the integral term in the right hand side. Again, considering the ward identity on this last one will require $\langle \widetilde{j}_0 \widetilde{j}_0 j_2 \rangle$ in the right hand side. In this fashion one can determine the solution. Note that in this case there is only one parameter which is $\widetilde{N}$.

Interestingly, we can get these correlators by taking the large $\widetilde{\lambda}$ limit of (3.59) (3.60) . At the level of three point functions of currents the limit is simple to take, namely we see that for any $s \geq 2$

$$\langle s_1 s_2 s_3 \rangle \to \langle s_1 s_2 s_3 \rangle_{\text{bos}} \tag{3.67}$$

so that all three point functions become purely boson ones. For $\widetilde{j}_0$, due to (3.59) it is necessary to rescale the operator and define a new operator $\widehat{\widetilde{j}}_0 = \widetilde{\lambda} \widetilde{j}_0$. The two point function of $\widehat{\widetilde{j}}_0$ remains finite. This also has the nice feature of removing $\widehat{\widetilde{j}}_0$ from the right hand side of (3.35) . The three point functions of the form $\langle \widehat{\widetilde{j}}_0 j_{s_1} j_{s_2} \rangle$ loose their β structure and remain only with the parity preserving $\gamma_{\widetilde{0} s_1 s_2}$ structure. The three point structures involving two scalars survive through a β structure, after rescaling the operator. And $\gamma_{\widehat{0}\widehat{0}\widehat{0}}$ stays being equal to zero, which is consistent with the large N limit of the critical $O(N)$ theory. This three point function becomes non-zero at higher orders in the $1/N$ expansion. Notice that at $\widetilde{\lambda} = \infty$ the three point functions become parity invariant. However the parity of the operator $\widetilde{j}_0$ got flipped compared to the one at $\widetilde{\lambda} = 0$. This suggests that the large $\widetilde{\lambda}$ limit of the fermions plus Chern-Simons matter theory should agree with the critical

$O(N)$ theory, at least in the large N limit. This was conjectured in [44]. But this conjecture appears to require a funny relation between the N's and $k's$ of both theories. In other words, if we start with the quasi-fermion theory with N and k and take the limit where λ, defined in [44], goes to one, then the conjecture would say that this should be the same as the critical $O(N')$ theory with some k' in the limit that $k' \to \infty$. But the behavior of the free energy in [44] would require that

$$N' \propto N(\lambda - 1)[-\log(\lambda - 1)]^3 \tag{3.68}$$

as $\lambda \to 1$. Here the proportionality is just a numerical constant. This formula is derived as follows. First notice that [44] derived a formula for the free energy, in the large N limit, for a theory of N fermions with a Chern-Simons coupling k. The conjecture is that this matches a critical bosonic $O(N')$ theory perturbed by a Chern-Simons coupling k'. When $\lambda \to 1$ we expect that $k' \to \infty$. So we can compute the free energy of the critical bosonic $O(N')$ ignoring the Chern-Simons coupling. The free energy of the $O(N')$ theory with no Chern-Simons coupling goes like N', for large N' [64]. Matching the two expressions for the free energy we get the relation (3.68) . At first sight, (3.68) seems incompatible with the fact that N' should be an integer. However, we should recall that (3.68) is only supposed to be true in the large N (and N') limit. Thus, it could be that there is an integer valued function of N and k which reduces to the right hand side of (3.68) in the large N limit. Note that if this were true, we would also find the same function in the two point function of the stress tensor, the function f discussed in (3.40) . Of course, here we are assuming that our $\widetilde{\lambda} \to \infty$ limit is the same as the $\lambda \to 1$ limit in [44].

3.5.4 The critical $O(N)$ fermion theory

Starting from N free fermions, we can add a perturbation $\widetilde{j}_0^2$. This is the three dimensional Gross-Neveu model [65]. This is an irrelevant perturbation. And one can wonder whether there is a UV fixed point that leads in the IR to the free fermion plus this perturbation. In the large N limit, it is easy to see that such a fixed point exists and it is given by a theory were the operator $\widetilde{j}_0$ in the UV has dimension one [66]. Thus it has the properties of the quasi-boson theory. Again, this is a theory that is parity symmetric. In [66] this theory was argued to be renormalizable to all orders in the $1/N$ expansion. We can now take the large g limit of the quasi-boson results (3.63) (3.64) . Again, we need to rescale $\widehat{j}_0 = \widetilde{\lambda} j_0$. Only the β_{s_1,s_2,s_3} survive in this limit. With one scalar operator, we get only the $\gamma_{0 s_1 s_2}$ structure surviving, which is consistent with parity since now $\widehat{j}_0$ is parity odd. Notice that, after the rescaling of the scalar operator, $\alpha_{\widehat{0}\widehat{0}\widehat{0}}$ still goes to zero if we hold a_3 fixed. This

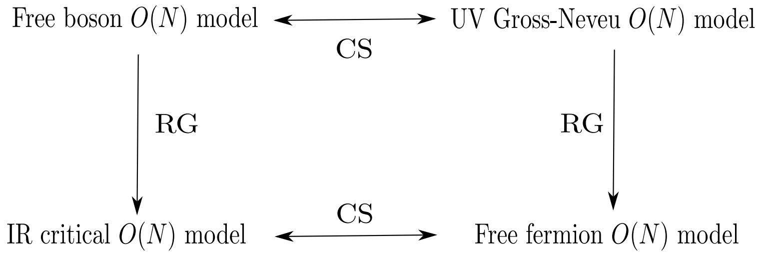


Figure 3.3: The analysis of three point functions can be summarized by this picture. The quasi-fermion theory is the top line and the quasi-boson is the bottom line. At the two end points we have the free boson or fermion theory on one side and the interacting $O(N)$ theories on the other. This is a statement about the three point functions. It would be interesting to understand whether we have a full duality between the two theories. Notice that the duality would relate a theory of fermions with a theory with scalars. This duality would be a form of bosonization in three dimensions. We also expect an RG flow connecting the quasi-boson theory on the top line and the quasi-fermion theory on the bottom line for general values of the Chern-Simons coupling.

is necessary because the three point function of $\langle \widehat{j}_0 \widehat{j}_0 \widehat{j}_0 \rangle$ should be zero by parity.

In principle, we could also introduce, in the large N limit, a parity breaking interaction of the from $\widehat{j}^3$ which would lead to a non-zero three point function. This could be obtained by rescaling a_3 in (3.64) , so that a finite term remains.

3.5.5 Relation to other Chern-Simons matter theories

Thoughout this paper we have focused on theories where the only single trace operators are given by even spin currents, plus a scalar operator. This is definitely the case for theories with N bosons or N fermions coupled to an $O(N)$ Chern-Simons theory. If instead we consider a theory of N complex fermions coupled to a $U(N)$ or $SU(N)$ Chern-Simons gauge field, then we have additional single trace operators. We still have the same spectrum for even spins, but we also have additional odd spin currents. However, the theory has a charge conjugation symmetry under which the odd spin currents get a minus sign. Thus, the odd spin currents cannot appear in the right hand side of $[Q, j_s]$ where s is even. These odd spin currents can (and do) appear in the right hand side of $\nabla.J_4$ in bilinear combinations. However, since we only considered insertions of J_4 into correlators involving operators with even spin, these extra terms do not contribute. Thus, the whole analysis in the paper goes through, for these Chern-Simons theories. Our results give the three point functions of even spin currents[11]. Presumably a similar analysis can be done for correlators of odd spin currents, but we will not do this here. It is also worth mentioning that these theories *do not* have a single trace twist three operator which could appear in the right hand side of $\nabla.J_4$. Thus the higher spin

[11] The result for the anomalous dimension (3.62) would be changed by the presence of the odd spin currents.

symmetry breaking only happens through double trace operators.

3.5.6 Comments about higher dimensions

We can consider the extension of the small breaking ansatz that we explored for $d = 3$ to higher dimensions. For simplicity we limit ourself to the case of almost conserved currents which are symmetric traceless tensors. We assume that the presence of a conserved J_4 will again fix all three point functions[12]. Thus, we are interested in the vector-like scenarios when the conservation of J_4 is broken at $\frac{1}{N}$. To analyze this possibility we consider $\partial_\mu J^\mu{}_{---}$ in an arbitrary number of dimensions. This operator has twist d while the conserved currents has twist $d-2$. Matching also the spin we get that we can write the following equation

$$\partial_\mu J^\mu{}_{---} = a\partial_- \mathcal{O} J_{--} + b\mathcal{O}\partial_- J_{--}. \tag{3.69}$$

where $\frac{b}{a}$ is fixed by the condition that the right hand side is a primary operator. The scalar operator $\mathcal{O}$ has scaling dimension $\Delta = 2$ by matching the quantum numbers. First, notice that the unitarity bound for the scalar operators is $\frac{d-2}{2}$. And, thus, if we restrict our attention to unitary theories, the equation (3.69) can be only valid in $d \leq 6$. We also need to check that there exists a three point function $\langle J_4 J_2 \mathcal{O}\rangle$ that reproduces (3.69) . Imposing the conservation of J_2 leads tho the result (3.69) as long as $\Delta_{\mathcal{O}} \neq d-2$. When $d = 4$, and $\Delta_{\mathcal{O}} = 2$, the correlator $\langle J_4 J_2 \mathcal{O}\rangle$ obeys J_4 current conservation automatically once we impose the J_2 current conservation [13]. Thus, we conclude that the scenario that we considered in $d = 3$ is impossible to realize in $d = 4$[14]. In $d = 5$ it seems possible to realize the scenario via the UV fixed point of a $-(\vec{\phi}.\vec{\phi})^2$ theory. This is a sick theory because the potential is negative, but one would probably not see the problem in $1/N$ perturbation theory. In $d = 6$ we do not know whether there is any example.

3.5.7 A comment on higher order parity violating terms

The parity breaking terms in Vasiliev's theories are characterized by a function $\theta(X) = \theta_0 + \theta_2 X^2 + \cdots$ [24]. The computations we have done are sensitive to θ_0. At tree level, the term proportional to θ_2 would start contributing to a planar five point function, but not to lower point functions [44].

[12] Though we have not proved this, the discussion of [60] looks almost identical in higher dimensions for symmetric traceless tensors.

[13] We thank David Poland for providing us with the Mathematica code to analyze the relevant three point functions in the case of higher dimensions.

[14] Notice that in $d = 4$ we can write also $j_- j_{--}$. However, in this case J_4 is automatically conserved as soon as we impose conservation of J_2 and J_1.

However, if we start with $\theta = 0$ and we tried to add θ_2 (keeping $\theta_0 = 0$), we run into the following issues. For $\theta_0 = 0$ we can choose boundary conditions that preserve the higher spin symmetry. That symmetry fixes all n point functions [60]. Thus, if turning on θ_2 modifies a five point function, then it must be breaking the higher spin symmetry. But if we break the higher spin symmetry we need to modify a three or four point function at the same order. (To see this, we can just consider the correlator of J_4 together with the currents that appear in the right hand side of (3.3)). However, θ_2 would only contribute to the five point function. Thus, our conclusion is that the function $\theta(X)$ is either constrained to be constant, or it's non-constant part can be removed by a field redefinition. In this argument we have assumed that these deformations of Vasiliev's theory in AdS_4 also lead to boundary correlators that obey all the properties of a CFT.

3.6 Conclusions and discussion

In this paper we have computed the three point correlation functions in conformal field theories with a large $\tilde{N}$ approximation, where the higher spin symmetry is broken by $1/\tilde{N}$ effects. Equivalently, we have performed an on shell analysis of Vasiliev theories on AdS_4, constraining the boundary three point correlators. These constraints also apply to the dS_4 case, where they can be viewed as computing possible non-gaussianities in Vasiliev's theory. From the dS_4 point of view we are computing the leading non-gaussianities of the de Sitter wavefunction[15].

We restricted the theories to contain a single trace spectrum with only one spin two current (the stress tensor) and only one scalar. The scalar can only have dimensions either one or two at leading order in $\tilde{N}$. This defines two classes of correlators which we called quasi-fermion (scalar of dimension two) and quasi-bosons, where the scalar has dimension one. These are the only possible dimensions that enable the $1/\tilde{N}$ breaking of the higher spin symmetry. The final three point functions depend on an overall constant, $\tilde{N}$, which is also the two point function of the stress tensor. Thus we can view $1/\tilde{N}$ as the coupling of the bulk theory. In addition, they depend on an extra parameter which selects the relative weights of the three possible structures in the correlators. The final results are given in (3.60) , (3.64) . These results apply, in particular, to theories of N bosons or N fermions coupled to an $SO(N)$ Chern-Simons gauge field. Here $\tilde{N}$ scales with N and the extra parameter $\tilde{\lambda}$ is a function of N/k, where k is the Chern-Simons level, see (3.40) . These results also apply to Vasiliev's theories which have parity breaking terms in the Lagrangian [24]. It also applies to Vasiliev's theories with boundary conditions that break the higher spin symmetry, but preserve

[15] Of course, for gravitons these match the particular structures discussed in [49], but with particular coefficients.

conformal invariance. At strong coupling the quasi-fermion three point functions go into the three point functions of the critical $O(N)$ theory. This suggests that there should be a duality between the small k limit of the Chern-Simons theories and the critical $O(N)$ theory [44]. However, the thermal partition function computed in [44] suggests that if this duality is true, the connection between the values of N and k of the two theories is rather intricate, see (3.68) . It would be nice to further understand this issue, and to find the correct duality, if there is one. Naively, *if* our methods actually fix all n point functions, then all leading order correlators would be consistent with the duality. This duality would be a three dimensional version of bosonization.

Our method was based on starting with the simplest possible even spin higher spin current, J_4, and writing the most general form for its divergence. We then noted that the violation of conservation of the J_4 current leads to constraints on three point functions, which we solved. These equations were mostly identical to the ones one would get in the conserved case, except when acting on correlators with spins two or lower. In these cases we obtained some relations which eliminated some of the free parameters and left only two parameters in the quasi-fermion theory and three parameters in the quasi-boson case. These parameters match with the free parameters that we have in large N Chern-Simons models.

Since we only considered the current J_4 one can wonder whether it is possible to add other deformation parameters which affect current conservation for higher spin currents but not J_4. We argue in appendix F that this is not possible. An interesting extension of these results would be to carry out this procedure for higher point functions. This is in principle conceptually straightforward, but it seems computationally difficult. Note that this analysis amounts to an on-shell study of the Vasiliev theory. We study the physical, gauge invariant, observables of this theory with AdS_4 or dS_4 boundary conditions. As it has often been emphasized, the on shell analysis of gauge theories can be simpler than doing computations in a fully Lorentz invariant formulation.

The methods discussed in this paper could be viewed as on-shell methods to compute correlation functions in certain matter plus Chern-Simons theories. Note that we did not have any gauge fixing issues, since we never considered gauge non-invariant quantities. These methods apply only to the special class of theories that do not contain a twist three, spin three single trace operator that can directly Higgs the J_4 higher spin symmetry already at leading order in N. It would also be interesting to consider cases where this Higgsing can happen already for single trace operators. If the mixing with this other operator is small, which occurs in weakly coupled theories, we can probably generalize the discussion of this paper. We would only need to add the twist three single trace operator in the right hand side of the divergence of the current, $\nabla . J_4$. This would lead to an on shell method for computing correlators. Something in this spirit was discussed in [67], [68]. This Higgsing mechanism, named "La Grande Bouffe" in [69] is also important for understanding the emergence of a more ordinary looking string theory in AdS from the higher spin system.

Our analysis also works for higher spin theories on de-Sitter space. In that case, we are constraining the wavefunction of the universe. It is valid for the proposed examples of dS-CFT [27, 70], as well as further examples that one might propose by looking at the parity violating versions of the Vasiliev theory. To consider these cases one should set $\widetilde{N} < 0$ in our formulas . Of course, the constraints hold whether we know the CFT dual or not! We have restricted our analysis to the case of AdS_4/CFT_3. Recently, examples of AdS_3/CFT_2 theories with higher spin symmetry have been considered. See for example [71, 72, 73, 74, 75, 76]. In lower dimensions the higher spin symmetry appears less restrictive, so one would need a more sophisticated analysis than the one presented in this paper. On the other hand, in higher dimensions the higher spin symmetry is more constraining. The higher dimensional case will hopefully be discussed separately. We have used conformal symmetry in an important way, it would also be interesting to study non-conformal cases. For example, we can expect to constrain the flows between the top and bottom lines of figure 3.

3.7 Appendix A. Why the divergence of the currents should be conformal primaries

In this appendix we show that the divergence of an operator with spin becomes a conformal primary when $\tau - 1 \to 0$. What happens is that the large representation with twist τ and spin s is splitting into a representation with $\tau = 1$ and spin s plus a representation with $\tau = 3$ and spin $s - 1$. This second representation is what appears in the right hand side of the divergence of J_s as $\tau \to 1$. This is well known fact, that was used in the past in [62, 67, 68, 61]. Here we recall its proof for completeness. Let us normalize the current to one, $\langle J_s | J_s \rangle$. Let us define

$$\partial . J_s = \alpha O_{s-1} \tag{3.70}$$

where O_{s-1} is also normalized to one. Then we have

$$\begin{aligned}\langle \partial J | \partial J \rangle &\propto (\tau - 1)\langle J | J \rangle = \tau - 1 \\ \langle \partial J | \partial J \rangle &= \alpha^2\end{aligned} \tag{3.71}$$

Where in the first line we converted the derivative in the bra into a special conformal generator in the ket, and then commuted it through the derivative in the ket using the conformal algebra. We ignored numerical factors. In the second line we used (3.70) . The conclusion is that

$$\tau - 1 \propto \alpha^2 \tag{3.72}$$

We now act on both sides of (3.70) with the special conformal generator K_ν. On the left we use the conformal algebra to evaluate the answer. We then get that $\alpha K_\mu O_{s-1} \propto (\tau - 1)J_\mu \propto \alpha^2 J_\mu$. The conclusion is that $\langle K_\mu O_{s-1}|K_\mu O_{s-1}\rangle \propto \alpha^2$. Thus we see that, to leading order in α, the operator O_{s-1} is a conformal primary. Of course this is true regardless of whether O_{s-1} is a single trace or multitrace operator, as long as $\tau - 1$ is very small. Finally, note that if we know α appearing in (3.70) then we can compute the anomalous dimension $\tau - 1$ of the current via the same formulas. In particular, in the quasi-fermion theory we have argued around (3.38) that $a_2 \sim \widetilde{\lambda}/\widetilde{N}$. This together with the value of the $\widetilde{j}_0$ two point function (3.48) implies that for the spin four current we have

$$\tau_4 - 1 = \frac{32}{21\pi^2}\frac{1}{\widetilde{N}}\frac{\widetilde{\lambda}^2}{1+\widetilde{\lambda}^2} \tag{3.73}$$

This formula should be applied to $O(N)$ theory. For general group and arbitrary spin s we expect the following formula to be true

$$\tau_s - 1 = a_s\frac{1}{\widetilde{N}}\frac{\widetilde{\lambda}^2}{1+\widetilde{\lambda}^2} + b_s\frac{1}{\widetilde{N}}\frac{\widetilde{\lambda}^2}{(1+\widetilde{\lambda}^2)^2} \tag{3.74}$$

where a_s and b_s are some fixed numbers. In principle these arguments allow us to fix the overall numerical coefficient. However, as a shortcut, we have used the formula for the anomalous dimensions for the critical $O(N)$ theory given in eqn. (2.20) of [77]. Thus, we fixed the overall coefficient in (3.73) so that the $\widetilde{\lambda} \to \infty$ limit matches [77]. This also fixes $a_s = \frac{16}{3\pi^2}\frac{s-2}{2s-1}$.

3.8 Appendix B. Structure of the pseudo-charge conservation identities

In this appendix we recall how the various coefficients that appear in the charge conservation identity are fixed. Assume that we consider only twist one three point functions, and that the spins are all non-zero. Then we can use (3.22) , perhaps including a j_0^{eff} operator when it acts on j_2. Then the action of Q on $\langle j_{s_1} j_{s_2} j_{s_3}\rangle$ gives an expression of the form

$$\left[\sum_{i=0,\pm 2} r_{\text{bos},1,i}\partial^{3-i}\langle j_{s_1+i} j_{s_2} j_{s_3}\rangle_{\text{bos}} + \text{cyclic}\right] + [\text{bos}\to\text{fer}] + [\text{bos}\to\text{odd}] \tag{3.75}$$

Here we have the set of coefficients $r_{\text{type},a,i}$. Here type goes over bos, fer, odd. a labels the point and runs over 1,2,3. Finally i runs over $\pm 2, 0$. In total there are up to $3^3 = 27$ coefficients. These coefficients result from multiplying the $c_{s,s'}$ in (3.22) and the α, β and γ in (3.27) . These equations split into three sets of equations, one for each type. In each set of equations the coefficients in front of different three point functions are all fixed up to an overall constant, except in the cases where the corresponding three point structure vanishes automatically, where, obviously, the corresponding coefficient is not fixed. For the boson and fermion types this follows from the discussion in appendix J of [60]. One simply needs to take successive light-cone limits of the three possible pairs of particles to argue that the coefficients are all uniquely fixed. For the odd structure the situation is more subtle since some of the equations do not have any non-zero solutions if we restrict to structures inside the triangle rule $s_i \leq s_{i+1} + s_{i-1}$. However, there are non-zero solutions once we take into account that the current non-conservation allows solutions outside the triangle. We have checked this in some cases, and we think it is likely to be true in general, but we did not prove it. We know that there is at least one solution with non-zero coefficients, which is the one that the Chern-Simons construction would produce. Thus, in order to show that the solution is unique, we would need to show that there is no solution after we set one of the coefficients to zero. We leave this problem for the future.

3.9 Appendix C. Compendium of normalizations

Here we would like to present more details on the normalization convention that we chose for the currents. Noticed that we have set them by the choice (3.23) . Here we will check that with (3.23) the coefficients of all the terms in (3.22) are the same for the single free boson and single free fermion theory. This is related to the fact that the higher spin algebra is the same for bosons and fermions in three dimensions (see, for example, [59]). In this appendix we do computations in the free theories. Then it is convenient to consider the matrix elements of the currents in Fourier space. As explained in the Appendix J of [60] it is convenient to introduce a combination of the two momenta that appear in the on shell matrix elements of a current with two on shell fermions or bosons . The idea is roughly to change $\partial^m \psi_-^1 \to (z+\bar{z})^{2m+1}$, $\partial^k \psi_-^2 \to (-1)^k (z-\bar{z})^{2k+1}$. Where the indices 1 and 2 denote the two fermion fields that make up the current. Then currents take a form

$$j_s = \alpha_s \left[z^{2s} - \overline{z}^{2s}\right] \tag{3.76}$$

where we introduced normalization factor α_s explicitly. The charge generated by j_4 that we consider in the main text is given then by

$$Q = \alpha_Q \left[(z+\overline{z})^6 - (z-\overline{z})^6\right] \propto \left[\partial^3_{x_1^-} + \partial^3_{x_2^-}\right] \tag{3.77}$$

where we again introduced the normalization factor for the charge. We have written the action of the charge on the two free fields that make the current. Also notice that α_2 is fixed in a canonical way. Now we can start fixing the normalizations. We start from $c_{2,4} = 1$. Using $[Q, j_2]$ we get

$$\alpha_Q = \frac{\alpha_4}{12\alpha_2} \tag{3.78}$$

then we consider $[Q, j_4]$ to fix $c_{4,4} = 1$, $c_{4,6} = 1$. It gives

$$\begin{aligned} \alpha_4 &= \frac{3}{10}\alpha_2 \\ \alpha_6 &= \left(\frac{3}{10}\right)^2 \alpha_2 \end{aligned} \tag{3.79}$$

simple analysis further shows that

$$\alpha_s = \left(\frac{3}{10}\right)\alpha_{s-2} \tag{3.80}$$

this follows from the formula (J.12) in [60] which can be written as

$$\begin{aligned} [Q_s, j_{s'}] &= (2s-2)! \sum_{r=0}^{s+s'-2} \widetilde{\alpha}_r(s,s') \partial^{s+s'-r-1} j_r \\ \widetilde{\alpha}_r(s,s') &= \left[1 + (-1)^{s+s'+r}\right] \left(\frac{1}{\Gamma(r+s-s')\Gamma(s+s'-r)} \pm \frac{1}{\Gamma(r+s+s')\Gamma(s-s'-r)}\right). \end{aligned} \tag{3.81}$$

Notice that in this normalization $c_{2,2}$ is also fixed to one automatically. Here the $\pm$ sign corresponds to the boson or fermion case. The case of bosons is almost identical. It is important that that the term with $\pm$ in (3.81) is zero for all terms $\widetilde{\alpha}_r(4, s')$ with $r, s' \geq 2$. These are all the cases where we can compare the normalization between fermions and bosons. This implies that all constants

$c_{s,s}$ and $c_{s,s\pm2}$ that appear in (3.22) , are the same in the free boson or free fermion theories. The case when one of the spins is zero, corresponds to the appearance of the scalar operator of twist one j_0 which is only present in the free boson theory. Moreover, using the formulas above and the results of [78] one can check that in the normalization we are using $n_s^{\text{free boson}} = n_s^{\text{free fermion}}$. More precisely, one can first relate the normalization used in [78] to the normalization here and then used the results of [78] for two point functions to get that

$$n_s^{\text{free boson}} = n_s^{\text{free fermion}} = \alpha_s^2 \Gamma(1+2s) \tag{3.82}$$
$$\alpha_2 = \frac{1}{16\pi}, \quad \alpha_s = (\frac{3}{10})\alpha_{s-2}.$$

We used this result in the main text. After we fix all normalizations of currents in this way we can compute all three point functions in the free fermion theory $\langle j_{s_1} j_{s_2} j_{s_3}\rangle_{\text{fer}}$ (as well as in the theory of free boson to get $\langle j_{s_1} j_{s_2} j_{s_3}\rangle_{\text{bos}}$). These are the solutions that we use in the main text, for example, in (3.27) .

3.10 Appendix D. Exploring the scalar sector for the quasi-fermion

In this appendix we explain how the scalar sector correlation functions could be recovered. We start from writing the general form of the variation of the scalar operator $\widetilde{j}_0$ under the action of Q

$$[Q, \widetilde{j}_0] = \widetilde{c}_{\widetilde{0},\widetilde{0}} c_{\widetilde{0},\widetilde{0}} \partial^3 \widetilde{j}_0 + \widetilde{c}_{\widetilde{0},2} c_{\widetilde{0},2} \partial \left(\partial_- J_{y-} - \partial_y J_{--}\right) \tag{3.83}$$
$$[Q, \widetilde{0}] = \widetilde{c}_{\widetilde{0},\widetilde{0}} c_{\widetilde{0},\widetilde{0}} \partial^3 \widetilde{0} + \widetilde{c}_{\widetilde{0},2} c_{\widetilde{0},2} \partial^{2\perp} 2$$

where remember coefficients c are the ones we would get in the theory of free fermion with the normalization of the operators that we chose. The interesting part is a deviation from the free theory which we denoted by $\widetilde{c}$ following the notations used in the main text. On the second line we introduce a shortened notation where we replaced $j_s \to s$ and in addition, we also introduced the symbol $\partial^{2\perp}$ to denote the combination of derivatives and indices of J_2 that appear in the first line. To fix some of the coefficients we use the $\langle 42\widetilde{0}\rangle$ three point function. It has two different structures: fermion and odd ones and we have already determined their coefficients. It is easy to see that it leads to the identities

$$\gamma_{42\tilde{0}} \propto \tilde{\lambda} n_{\tilde{0}} \qquad (3.84)$$
$$\tilde{c}_{\tilde{0},2} c_{\tilde{0},2} n_2 = \beta_{42\tilde{0}}$$

In the first line we took the divergence of the J_4 current and compared to (3.5) . In the second line we integrated the current to get the charge acting on $\tilde{j}_0$. From the first line we get that $\tilde{n}_{\tilde{0}} = \frac{1}{1+\tilde{\lambda}^2}$ where we used the fact that at $\tilde{\lambda} = 0$ it should be equal to 1. The second equation then fixes

$$\tilde{c}_{\tilde{0},2} = \frac{1}{1+\tilde{\lambda}^2}. \qquad (3.85)$$

3.10.1 Fixing $\tilde{c}_{\tilde{0},\tilde{0}}$ and nontrivial consistency check

Here we fix $\tilde{c}_{\tilde{0},\tilde{0}}$ by considering the pseudo-conservation identities for $\langle \tilde{0} s_1 s_2 \rangle$ where both s_1 and s_2 are larger than 2 for simplicity. We will encounter a rather intricate structure for this identity. We schematically write relevant terms in pseudo-conservation identities with their $\tilde{\lambda}$ scaling

$$\begin{aligned}
&\tilde{c}_{\tilde{0},\tilde{0}} \frac{1}{1+\tilde{\lambda}^2} \langle \partial^3 \tilde{0} s_1 s_2 \rangle_{\text{fer}} + \tilde{c}_{\tilde{0},\tilde{0}} \frac{\tilde{\lambda}}{1+\tilde{\lambda}^2} \langle \partial^3 \tilde{0} s_1 s_2 \rangle_{\text{odd}} + \\
&\frac{1}{1+\tilde{\lambda}^2} \left(\frac{1}{1+\tilde{\lambda}^2} \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{fer}} + \frac{\tilde{\lambda}}{1+\tilde{\lambda}^2} \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{odd}} + \frac{\tilde{\lambda}^2}{1+\tilde{\lambda}^2} \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{bos}} \right) + \\
&\frac{1}{1+\tilde{\lambda}^2} \langle \tilde{0} \text{ standard terms} \rangle_{\text{fer}} + \frac{\tilde{\lambda}}{1+\tilde{\lambda}^2} \langle \tilde{0} \text{ standard terms} \rangle_{\text{odd}} = \\
&= \frac{\tilde{\lambda}}{1+\tilde{\lambda}^2} \partial_{x_1^-} \int \frac{1}{|x-x_1|^4} \left(\frac{1}{1+\tilde{\lambda}^2} \langle 2 s_1 s_2 \rangle_{\text{fer}} + \frac{\tilde{\lambda}}{1+\tilde{\lambda}^2} \langle 2 s_1 s_2 \rangle_{\text{odd}} + \frac{\tilde{\lambda}^2}{1+\tilde{\lambda}^2} \langle 2 s_1 s_2 \rangle_{\text{bos}} \right)
\end{aligned} \qquad (3.86)$$

Here, except for the terms involving $\partial^{2\perp} 2$ all components of the currents are minus, as are all derivatives. Now let's start looking for a solution. All $\tilde{\lambda}$ dependence is explicit in this equation. Note that the equations contain terms with various $\tilde{\lambda}$ dependence. First we match the double poles at $\tilde{\lambda}^2 = -1$. This requires the following two identities

$$\langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{odd}} = \partial_{x_1^-} \int \frac{1}{|x-x_1|^4} \left(\langle 2 s_1 s_2 \rangle_{\text{fer}} - \langle 2 s_1 s_2 \rangle_{\text{bos}} \right), \qquad (3.87)$$

$$\partial_{x_1^-} \int \frac{1}{|x-x_1|^4} \langle 2 s_1 s_2 \rangle_{\text{odd}} = -\langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{fer}} + \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{bos}}. \qquad (3.88)$$

The second line can also be used to make all terms of the $_{\text{fer}}$ terms work. This would work nicely if we set $\tilde{c}_{\tilde{0},\tilde{0}} = 1$, which is what we wanted to argue. In fact, replacing the second line, we get all the parity odd pieces in the full equation work out. By parity odd we mean the terms that are odd under $y \to -y$.

To summarize, in this subsectiontion we fixed $\widetilde{c}_{\widetilde{0},\widetilde{0}} = 1$ and presented the self-consistency relations (3.87) (**??**) that are necessary for the whole construction to work. We checked one of the (3.87) in the light cone limit and it indeed works. It would be nice to check these identities more fully.

3.10.2 Fixing $\langle s\widetilde{0}\widetilde{0}\rangle$

To fix these three point functions we consider pseudo-conservation identities that we get from $\langle s\widetilde{0}\widetilde{0}\rangle$. Notice that we already know that $\langle 2\widetilde{0}\widetilde{0}\rangle$ and $\langle 4\widetilde{0}\widetilde{0}\rangle$ functions are given by β_{222}. The first follows from the stress tensor Ward identity and the two point function. The second follows from the action of the charge Q on $\widetilde{j}_0$, with $\widetilde{c}_{\widetilde{0},\widetilde{0}} = 1$ that we derived above. Starting from this we can build the induction. Consider first pseudo-conservation identity for $\langle 4\widetilde{0}\widetilde{0}\rangle$. We get schematically

$$\begin{aligned}
&\beta_{6\widetilde{0}\widetilde{0}}\langle \partial 6\widetilde{0}\widetilde{0}\rangle_{\text{fer}} + \frac{1}{1+\widetilde{\lambda}^2}\left(\langle \partial^5 2\widetilde{0}\widetilde{0}\rangle_{\text{fer}} + \langle \partial^3 4\widetilde{0}\widetilde{0}\rangle_{\text{fer}}\right) + \\
&\frac{1}{1+\widetilde{\lambda}^2}\left(\frac{1}{1+\widetilde{\lambda}^2}\langle 4\partial^{2\perp}2\widetilde{0}\rangle_{\text{fer}} + \frac{\widetilde{\lambda}}{1+\widetilde{\lambda}^2}\langle 4\partial^{2\perp}2\widetilde{0}\rangle_{\text{odd}} + \langle 4\partial^3\widetilde{0}\widetilde{0}\rangle_{\text{fer}} + \widetilde{\lambda}\langle 4\partial^3\widetilde{0}\widetilde{0}\rangle_{\text{odd}}\right) + x_2 \leftrightarrow x_3 \\
&= \frac{\widetilde{\lambda}}{1+\widetilde{\lambda}^2}\partial_{x_2^-}\int \frac{1}{|x-x_1|^4}\left(\frac{1}{1+\widetilde{\lambda}^2}\langle 42\widetilde{0}\rangle_{\text{fer}} + \frac{\widetilde{\lambda}}{1+\widetilde{\lambda}^2}\langle 42\widetilde{0}\rangle_{\text{odd}} + [x_2 \leftrightarrow x_3]\right),
\end{aligned} \tag{3.89}$$

Again, matching the double poles at $\widetilde{\lambda}^2 = -1$ requires

$$\langle 4\partial^{2\perp}2\widetilde{0}\rangle_{\text{odd}} = \partial_{x_2^-}\int \frac{1}{|x-x_1|^4}\langle 42\widetilde{0}\rangle_{\text{fer}}, \tag{3.90}$$

$$\langle 4\partial^{2\perp}2\widetilde{0}\rangle_{\text{fer}} = -\partial_{x_2^-}\int \frac{1}{|x-x_1|^4}\langle 42\widetilde{0}\rangle_{\text{odd}}. \tag{3.91}$$

Then replacing these identities in the equations implies that the odd piece cancels and the fermion piece works if

$$\beta_{6\widetilde{0}\widetilde{0}} = \beta_{222} = \frac{\widetilde{N}}{1+\widetilde{\lambda}^2}. \tag{3.92}$$

Then we proceed by induction. Considering pseudo-charge conservation identities for $\langle s\widetilde{0}\widetilde{0}\rangle$ we get a similar story. Double pole matching leads to

$$\langle s\partial^{2\perp}\widetilde{2}\widetilde{0}\rangle_{\rm odd} = \partial_{x_2^-}\int \frac{1}{|x-x_1|^4}\langle s\widetilde{2}\widetilde{0}\rangle_{\rm fer}, \tag{3.93}$$

$$\langle s\partial^{2\perp}\widetilde{2}\widetilde{0}\rangle_{\rm fer} = -\partial_{x_2^-}\int \frac{1}{|x-x_1|^4}\langle s\widetilde{2}\widetilde{0}\rangle_{\rm odd}, \tag{3.94}$$

and the fermion piece fixes

$$\beta_{s\widetilde{0}\widetilde{0}} = \beta_{222} = \frac{\widetilde{N}}{1+\widetilde{\lambda}^2}. \tag{3.95}$$

This fixes the three point function involving two $\widetilde{j}_0$ operators.

3.10.3 Fixing $\gamma_{\widetilde{0}\widetilde{0}\widetilde{0}}$

To fix the last three point function we consider the WI $\langle\widetilde{0}\widetilde{0}\widetilde{0}\rangle$. We get schematically the equation

$$\gamma_{\widetilde{0}\widetilde{0}\widetilde{0}}(\partial_1^3+\partial_2^3+\partial_3^3)\langle\widetilde{0}\widetilde{0}\widetilde{0}\rangle_{\rm odd} = \frac{\widetilde{\lambda}n_0}{\widetilde{N}}\beta_{\widetilde{2}\widetilde{0}\widetilde{0}}\sum_{perm}\partial_{x_1^-}\int\frac{1}{|x-x_1|^4}\langle\widetilde{2}\widetilde{0}\widetilde{0}\rangle_{\rm fer}. \tag{3.96}$$

The right hand side of this equation can be computed explicitly using the star-triangle identity (see, for example, [79]). More precisely, each term in the right hand side of (3.96) can be rewritten (thanks to our sandwich geometry as we can pull out the derivatives) as

$$\frac{1}{x_{23}^{3/2}}\partial_{x_1^-}\left(\partial^2_{x_2^-}+\partial^2_{x_3^-}-6\partial_{x_2^-}\partial_{x_3^-}\right)\int d^3x\frac{1}{|x-x_1|^2|x-x_2|^3|x-x_3|}+\text{perm}. \tag{3.97}$$

The integral is finite and is given by $\frac{x_{23}}{x_{12}^2x_{13}^2}$. By taking the derivatives and summing over permutations which come from different contractions in (3.96) one can check that the sum is actually zero. Thus we are forced to set

$$\gamma_{\widetilde{0}\widetilde{0}\widetilde{0}} = 0. \tag{3.98}$$

This is nicely consisted with the critical $O(N)$ limit.

3.11 Appendix E. Exploring the scalar sector for the quasi-boson

Now we would like to go through the same analysis but for the quasi-boson theory. We start by writing the general form of the variation of the scalar operator j_0 under the action of Q

$$[Q, j_0] = \widetilde{c}_{0,0} c_{0,\widetilde{0}} \partial^3 j_0 + \widetilde{c}_{0,2} c_{0,2} \partial j_2 \tag{3.99}$$

Recall that coefficients c are the ones we would get in the theory of free boson with the normalization of the operators that we chose. The interesting part is a deviation from the free theory which we denoted by $\widetilde{c}$ following the notations used in the main text. First we consider the $\langle 420 \rangle$ three point function. It has two different structures: boson and odd ones. This leads to the identities

$$\begin{aligned} \gamma_{420} &\propto \widetilde{\lambda} n_0 \\ \widetilde{c}_{0,2} c_{0,2} n_2 &= \alpha_{420} = \alpha_{222} \end{aligned} \tag{3.100}$$

In the first line we simply have taken the divergence of J_4 and used (3.6) . In the second line we have integrated J_4 around j_0. From the first line we get that $\widetilde{n}_0 = \frac{1}{1+\widetilde{\lambda}^2}$ where we used the fact that at $\widetilde{\lambda} = 0$ it should be equal to 1. The second equation then fixes

$$\widetilde{c}_{0,2} = \frac{1}{1 + \widetilde{\lambda}^2}. \tag{3.101}$$

3.11.1 Fixing $\widetilde{c}_{0,0}$

To fix $\widetilde{c}_{0,0}$ we are analogously considering the pseudo-conservation identities for $\langle 0 s_1 s_2 \rangle$ where both s_1 and s_2 are larger than 2 for simplicity. We schematically write relevant terms in pseudo-conservation identities with their $\widetilde{\lambda}$ scaling

$$\begin{aligned} &\widetilde{c}_{0,0} \frac{1}{1+\widetilde{\lambda}^2} \langle \partial^3 0 s_1 s_2 \rangle_{\text{bos}} + \widetilde{c}_{0,0} \frac{\widetilde{\lambda}}{1+\widetilde{\lambda}^2} \langle \partial^3 0 s_1 s_2 \rangle_{\text{odd}} + \\ &\frac{1}{1+\widetilde{\lambda}^2} \left(\frac{1}{1+\widetilde{\lambda}^2} \langle \partial 2 s_1 s_2 \rangle_{\text{bos}} + \frac{\widetilde{\lambda}}{1+\widetilde{\lambda}^2} \langle \partial 2 s_1 s_2 \rangle_{\text{odd}} + \frac{\widetilde{\lambda}^2}{1+\widetilde{\lambda}^2} \langle \partial 2 s_1 s_2 \rangle_{\text{fer}} \right) + \\ &\frac{1}{1+\widetilde{\lambda}^2} \langle 0 \text{ standard terms} \rangle_{\text{bos}} + \frac{\widetilde{\lambda}}{1+\widetilde{\lambda}^2} \langle 0 \text{ standard terms} \rangle_{\text{odd}} = \\ &= \frac{\widetilde{\lambda}}{(1+\widetilde{\lambda}^2)^2} \int \frac{1}{|x - x_1|^2} \Big(\langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{bos}} + \widetilde{\lambda} \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{odd}} + \widetilde{\lambda}^2 \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{fer}} \Big) \end{aligned} \tag{3.102}$$

Again we match the double poles at $\widetilde{\lambda}^2 = -1$ to obtain

$$\langle \partial 2 s_1 s_2 \rangle_{\text{odd}} = \int \frac{1}{|x - x_1|^2} \left(\langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{bos}} - \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{fer}} \right) \tag{3.103}$$

$$- \langle \partial 2 s_1 s_2 \rangle_{\text{bos}} + \langle \partial 2 s_1 s_2 \rangle_{\text{fer}} = \int \frac{1}{|x - x_1|^2} \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{odd}}$$

Now using these equations, we find that the all the even pieces under $y \to -y$ work properly only if $\widetilde{c}_{0,0} = 1$. Then the odd piece reduces to the condition

$$\langle \partial^3 0 s_1 s_2 \rangle_{\text{odd}} + \langle 0 \text{ standard terms} \rangle_{\text{odd}} = \int \frac{1}{|x - x_1|^2} \langle \partial^{2\perp} 2 s_1 s_2 \rangle_{\text{fer}} \tag{3.104}$$

We have not checked explicitly whether (3.104) and (3.103) are true, but they should be for consistency.

3.11.2 Fixing $\langle s00 \rangle$

To fix these three point functions we consider pseudo-conservation identities that we get from $\langle s00 \rangle$. Notice that we already know that $\langle 200 \rangle$ and $\langle 400 \rangle$ functions are given by α_{222}. This follows from the stress tensor Ward identity and from the action of Q on j_0, together with the normalization of j_0.

Starting from this we can build the induction. Consider first pseudo-conservation identity for $\langle 400 \rangle$. We get schematically

$$\begin{aligned}
&\alpha_{600} \langle \partial 600 \rangle_{\text{bos}} + \frac{1}{1 + \widetilde{\lambda}^2} \left(\langle \partial^5 200 \rangle_{\text{bos}} + \langle \partial^3 400 \rangle_{\text{bos}} \right) + \\
&+ \frac{1}{1 + \widetilde{\lambda}^2} \left(\frac{1}{1 + \widetilde{\lambda}^2} \langle 4 \partial 20 \rangle_{\text{bos}} + \frac{\widetilde{\lambda}}{1 + \widetilde{\lambda}^2} \langle 4 \partial 20 \rangle_{\text{odd}} \right) + \\
&+ \frac{1}{1 + \widetilde{\lambda}^2} \langle 4 \partial^3 00 \rangle_{\text{bos}} + \frac{\widetilde{\lambda}}{1 + \widetilde{\lambda}^2} \langle 4 \partial^3 00 \rangle_{\text{odd}} + [x_2 \leftrightarrow x_3] = \\
&= \frac{\widetilde{\lambda}}{1 + \widetilde{\lambda}^2} \int \frac{1}{|x - x_1|^2} \left(\frac{1}{1 + \widetilde{\lambda}^2} \langle 4 \partial^{2\perp} 20 \rangle_{\text{bos}} + \frac{\widetilde{\lambda}}{1 + \widetilde{\lambda}^2} \langle 4 \partial^{2\perp} 20 \rangle_{\text{odd}} + [x_2 \leftrightarrow x_3] \right),
\end{aligned} \tag{3.105}$$

Again looking at the double pole at $\widetilde{\lambda}^2 = -1$ we get the equations

$$\langle 4 \partial 20 \rangle_{\text{bos}} = - \int \frac{1}{|x - x_1|^2} \langle 4 \partial^{2\perp} 20 \rangle_{\text{odd}}, \tag{3.106}$$

$$\langle 4 \partial 20 \rangle_{\text{odd}} = \int \frac{1}{|x - x_1|^2} \langle 4 \partial^{2\perp} 20 \rangle_{\text{bos}} \tag{3.107}$$

This implies that $\alpha_{600} = \alpha_{222}$.

By induction through the identities

$$\langle s\partial 20\rangle_{\text{bos}} = -\int \frac{1}{|x-x_1|^2}\langle s\partial^{2\perp} 20\rangle_{\text{odd}}, \tag{3.108}$$

$$\langle s\partial 20\rangle_{\text{odd}} = \int \frac{1}{|x-x_1|^2}\langle s\partial^{2\perp} 20\rangle_{\text{bos}} \tag{3.109}$$

we get that

$$\alpha_{s00} = \alpha_{222}. \tag{3.110}$$

Again, we did not verify the identities (3.108) but they are necessary for consistency.

3.11.3 Vanishing of the double trace term in the $\langle 000\rangle$ identity

When we studied the pseudo-charge conservation identity for the triple scalar correlator $\langle j_0 j_0 j_0\rangle$ in the main text we said that the double trace term in the divergence of J^4, (3.6) , vanishes. Here we prove that assertion. We start from the following integral

$$I(x_1,x_2,x_3) = a_2 \int d^3x \langle j_0(x) j_0(x_1)\rangle\langle \partial^{2\perp} J_2(x) j_0(x_2) j_0(x_3)\rangle \tag{3.111}$$

where operator $j_0(x)\partial^{2\perp}J_2(x)$ comes from an insertion of the double trace term in (3.6) . After some algebra one can re-express this as the following integral

$$I(x_1,x_2,x_3) \sim \partial_{x_1^-}\left(\partial_{x_2^-} - \partial_{x_3^-}\right)\left(\partial_{x_2^-}\partial_{y_3} - \partial_{x_3^-}\partial_{y_2}\right)\int d^3x \frac{1}{|x-x_1|^2|x-x_2||x-x_3|} \tag{3.112}$$

written in this way it is manifestly finite. However, we rewrite it again as follows

$$I(x_1,x_2,x_3) \sim \partial_{x_1^-}\left(\partial_{x_2^-} - \partial_{x_3^-}\right)\left[y_{23}\partial_{x_3^-}J(x_1,x_2,x_3) - \frac{1}{2}x_{23}^+\partial_{y_3}J(x_1,x_2,x_3)\right] \tag{3.113}$$

where

$$J(x_1,x_2,x_3) = \int d^3x \frac{1}{|x-x_1|^2|x-x_2|^3|x-x_3|} \tag{3.114}$$

is the conformally invariant integral. It is divergent, however, the difference in (3.113) is, of course, finite. We can take this integral using the well-known star-triangle identity. We regularize the integral as

$$J_\delta(x_1, x_2, x_3) = \int d^{3+\delta}x \frac{1}{|x - x_1|^{2+\delta}|x - x_2|^3|x - x_3|} \tag{3.115}$$

using the star-triangle formula, expanding in δ and plugging into (3.112) we see that the divergent piece cancels and the finite piece is given by

$$I(x_1, x_2, x_3) \sim \frac{1}{|x_{23}|} \left(\partial_{x_1^-}\partial_{x_2^-} - \partial_{x_1^-}\partial_{x_3^-}\right) \frac{x_1^+ y_{23} + x_2^+ y_{31} + x_3^+ y_{12}}{|x_1 - x_2|^2|x_1 - x_3|^2|x_2 - x_3|}. \tag{3.116}$$

The total contribution to the pseudo-charge conservation identity is given by the sum of three terms which is zero

$$I(x_1, x_2, x_3) + I(x_2, x_3, x_1) + I(x_3, x_1, x_2) = 0. \tag{3.117}$$

Thus, double trace non-conservation does not contribute to the $\langle 000 \rangle$ pseudo-charge conservation identity.

3.12 Appendix F. Impossibility to add any further double trace deformations

Our whole analysis was based on studying the possible terms that appear in the divergence of the spin four current J_4. We could wonder whether we can add further double terms to the divergence of higher spin currents which are not fixed by the analysis we have already done. In other words, these would be terms that appear for higher spin currents but not for the spin four current. This would only be possible via the odd terms, which are the only ones that could have a non-conserved current. We suspect that all odd terms are fixed by the J_4 pseudo-conservation identities, but we did not prove it. Therefore we will do a separate analysis to argue that we cannot continuously deform the divergence of the higher spin currents once we have fixed the J_4 one.

Let us imagine that it is possible to introduce an additional parameter for the breaking of some higher spin current. We will focus first on possible double trace terms. Consider the *lowest* spin s

at which the new term enters[16]

$$\nabla.J_s = qJ_{s_1}J_{s_2} + \text{rest} \tag{3.118}$$

By assumption $s > 4$. The rest denotes terms that are required by J_4 non-conservation. This extra term contributes to the $\langle J_s J_{s_1} J_{s_2}\rangle_{\text{odd}}$ three point function. Consider then pseudo-conservation identity for the J_4 current on $\langle j_{s-2} j_{s_1} j_{s_2}\rangle$. For $s_1, s_2 < s-2$ we get

$$0 = q\langle j_s j_{s_1} j_{s_2}\rangle_{\text{odd}} + \text{rest} = q\langle j_s j_{s_1} j_{s_2}\rangle_{\text{odd}} \tag{3.119}$$

where by "rest" we denote the terms that were present when $q = 0$. These sum up to zero by construction so we have to conclude that $q = 0$. The argument slightly changes when one of the spins is equal to $s-2$. Below we discuss this case separately for the quasi-boson and quasi-fermion.

3.12.1 Case of quasi-fermion

In this case if, say, $s_1 = s-2$ we have two possibilities for s_2: 2 or $\widetilde{0}$. So we write schematically

$$\nabla.J_s = q(\epsilon J_{s-2}J_2 + \partial J_{s-2}\widetilde{J}_0) + \text{rest} \tag{3.120}$$

where ϵ is the three dimension Levi-Civita tensor. This modifies the correlators $\langle J_s J_{s-2} J_2\rangle_{\text{odd}}$ and $\langle J_s J_{s-2}\widetilde{j}_0\rangle_{\text{odd}}$. Consider now pseudo-conservation of the J_4 current for $\langle j_{s-2} j_{s-2} j_2\rangle$ we get

$$q\left(\partial_1\langle j_s j_{s-2} j_2\rangle_{\text{odd}} + \partial_2\langle j_{s-2} j_s j_2\rangle_{\text{odd}}\right) = 0. \tag{3.121}$$

To check this identity it is convenient to introduce in the formula above the dependence on the insertion of $j_2(x)$ and consider the integral

$$\partial_{x_3^-}\int d^3x \frac{1}{|x-x_3|^4}\left(\partial_1\langle s\ s-2\ 2(x)\rangle_{\text{odd}} + \partial_2\langle s-2\ s\ 2(x)\rangle_{\text{odd}}\right) = 0 \tag{3.122}$$

[16]This expression is schematic. We require and ϵ tensor in the right hand side if the two operators J_{s_1} and J_{s_2} have twist one.

Using formula (3.87) we get that it is equivalent to the identity

$$\partial_1 \langle s\ s-2\ \partial^{2\perp} 2\rangle_{\rm fer} + \partial_2 \langle s-2\ s\ \partial^{2\perp} 2\rangle_{\rm fer} \tag{3.123}$$
$$-\partial_1 \langle s\ s-2\ \partial^{2\perp} 2\rangle_{\rm bos} - \partial_2 \langle s-2\ s\ \partial^{2\perp} 2\rangle_{\rm bos} = 0$$

We now take a light cone limit $\langle s_1 \underline{s_2 \partial^{2\perp} 2}\rangle$. More specifically we take the limits $\lim_{y\to 0} y|y| \lim_{x^+\to 0}$ and $\lim_{y\to 0} y^2 \lim_{x^+\to 0}$, which pick out the boson and fermion pieces respectively. Then one can check that (3.123) does not have a solution. The next case to consider is $\langle s-2\ s-2\ \widetilde{0}\rangle$. In this case we have

$$q\left(\partial_1 \langle s\ s-2\ \widetilde{0}\rangle_{\rm odd} + \partial_2 \langle s-2\ s\ \widetilde{0}\rangle_{\rm odd}\right) = 0. \tag{3.124}$$

Again we introduce the explicit dependence on the insertion point $\widetilde{0}(x)$ and integrate to get

$$q\int d^3x \frac{1}{|x-x_3|^2}\left(\partial_1 \langle s\ s-2\ \widetilde{0}(x)\rangle_{\rm odd} + \partial_2 \langle s-2\ s\ \widetilde{0}(x)\rangle_{\rm odd}\right) = 0 \tag{3.125}$$

Using (3.25) this becomes

$$q\left[\partial_1 \langle s\ s-2\ 0\rangle_{\rm bos} + \partial_2 \langle s-2\ s\ 0\rangle_{\rm bos}\right] = 0. \tag{3.126}$$

This identity does not hold and we conclude that $q = 0$.

3.12.2 Case of quasi-boson

In this case if, say, $s_1 = s-2$ we have two possibilities for s_2: 2 or 0. So we write schematically

$$\nabla . J_s = q(\epsilon J_{s-2} J_2 + \epsilon \partial^2 J_{s-2} J_0) + \text{rest} \tag{3.127}$$

We modify the correlators $\langle J_s J_{s-2} J_2\rangle_{\rm odd}$ and $\langle J_s J_{s-2} j_0\rangle_{\rm odd}$. Consider now the J_4 pseudo-conservation identity for $\langle j_{s-2} j_{s-2} j_2\rangle$. The argument is then identical to the fermion one. We conclude that the first term in (3.127) is not possible. Then we consider the J_4 pseudo-conservation on $\langle j_{s-2} j_{s-2} j_0\rangle$. In this case we have

$$q\left(\partial_1 \langle j_s j_{s-2} j_0\rangle_{\rm odd} + \partial_2 \langle j_{s-2} j_s j_0\rangle_{\rm odd}\right) = 0. \tag{3.128}$$

we introduce the explicit dependence on the insertion point $0(x)$ and integrate to get

$$\int d^3x \frac{1}{|x-x_3|^4} \left(\partial_1 \langle s\ s-2\ 0(x)\rangle_{\rm odd} + \partial_2 \langle s-2\ s\ 0(x)\rangle_{\rm odd}\right) = 0 \qquad (3.129)$$

Using (3.42) we find

$$\partial_1 \langle s\ s-2\ \widetilde{0}\rangle_{\rm fer} + \partial_2 \langle s-2\ s\ \widetilde{0}\rangle_{\rm fer} = 0. \qquad (3.130)$$

This identity does not hold and we conclude that $q=0$.

3.12.3 Triple trace deformation

Here we would like to analyze the new triple trace deformations.

$$\nabla.J_s = qJ_{s_1}J_{s_2}J_{s_3} + \text{rest}. \qquad (3.131)$$

First notice that this deformation does not affect any of three point functions. Thus, it does not affect any of the identities that we got using J_4. Thus, if our assumption that the identities for the three point functions fix the odd pieces completely is correct then any of these terms will not affect any of our conclusions. Applying arguments similar to the above ones for pseudo-conservation of J_4 on $\langle j_{s-2}j_{s_1}j_{s_2}j_{s_3}\rangle$ we expect to find that we cannot add the q deformation. However, we did not perform a complete analysis.

3.13 Appendix G. Motivational introduction

In this appendix we give a longer motivational introduction for the study of the constraints imposed by the higher spin symmetry. As it is well known, the structure of theories with massless particles with spin is highly constrained. For example, massless particles of spin one lead to the Yang Mills theory, at leading order in derivatives. Similarly massless particles of spin two lead to general relativity. We also need the assumption that the leading order interaction at low energies is such that the particles are charged under the gauge symmetries or that gravitons couple to energy in the usual way. Now for *massive* spin one particles, what can we say? If we assume that we have a weakly interacting theory for energies much bigger than the mass of the particle, and we assume a local bulk lagrangian without other higher spin particles, then we find that the theory should contain at least a Yang-Mills field plus a Higgs particle. We can then view the theory as having a spontaneously broken gauge symmetry. Notice that we are assuming that the theory is weakly coupled. Now consider a theory that has massless particles with higher spin, $s > 2$. If we are in flat space and the S-matrix can be defined, then we expect that the Coleman-Mandula theorem should forbid any interaction [15, 16][17]. Here we are assuming that the couplings of the higher spin fields are such that particles are charged under the higher spin transformations. This is the case if we have a graviton in the theory, the same vertex that makes sure that the graviton couples to the energy of the higher spin particle also implies that the particles transforms under a higher spin charge.

Now we can consider theories that contain massive particles with spin $s > 2$. We assume that we have a weakly coupled theory at all energies, even at very high energies. In other words, we assume a suitable decay of the amplitudes at high energy (such as the one we have in string theory). Then it is likely that the theory in question is a full string theory. In other words, we can propose the following conjecture: *A Lorentz invariant theory in three or more dimensions, which is weakly coupled*[18]*, contains a massive particle with spin $s > 2$, and has amplitudes with suitably bounded behavior at high energies, should be a string theory.*[19] Of course, there are many string compactifications, so we do not expect this to fix the theory uniquely. This is precisely the problem that the old dual models literature tried to solve, and string theory was found as a particular solution. The above conjecture is just that it is the *only* solution. As an analogy, in gauge theories with a spontaneously broken symmetry we can have many realizations of the Higgs mechanism. However, the paradigm of spontaneously broken gauge symmetry constrains the theories in an important way. In string theories, or theories with higher spin massive particles, we also expect that their structure is constrained by the spontaneous breaking of the higher spin symmetry. For example, we expect that once we have a higher spin particle, we have infinitely many of them. It is notoriously difficult to study infinite dimensional symmetries. It is even more difficult if there is no unbroken phase that is easy to study. These ideas were discussed in [80, 81] (see [82] for a recent discussion containing many further references).

Here is where the AdS case appears a bit simpler. In AdS space, as opposed to flat space, it is possible to have a theory which is interacting (in AdS) but that nevertheless realizes the unbroken higher spin symmetry [24, 54, 55]. The fact that these theories exist can be understood from

[17] The Coleman-Mandula theorem assumes that we have a finite number of states below a certain mass shell. For the purpose of this discussion we assume that it still applies...

[18] The weak coupling assumption is important. In fact, the reader might think of the following apparent counter example. Consider the scattering of higher spin excited states of hydrogen atoms. These are higher spin states, but they are not strings! However, these states are not weakly coupled to each other in the particle physics sense.

[19] One probably needs to assume some interactions between the massive spin particles which makes sure that the particles are "charged" under the higher spin symmetry. This would be automatically true if we also include an $m = 0$, $s = 2$ graviton. This should also apply to theories like large N QCD, which have a conserved stress tensor leading to massless graviton in the bulk.

AdS/CFT, they are duals to free theories [56, 57, 58, 11, 59, 12]. In any example of AdS/CFT that has a coupling constant on the boundary we can take the zero coupling limit. In this limit, the boundary theory has single trace operators sitting at the twist bound. These are all states that are given by bilinears in the fields. Furthermore, this forms a closed subsector under the OPE [52]. Thus we can always find a Vasiliev-type theory as a consistent truncation of the zero coupling limit of the full theory. The Vasiliev-type theories contain only the higher spin fields (perhaps plus a scalar). They are analogous to the Yang Mills theory without an elementary Higgs particle. For example, if we take the zero λ limit of the $\mathcal{N} = 4$ super Yang-Mills theory, we get a certain theory with higher spin symmetry. Besides the higher spin currents, this theory contains lots of other states that are given by single trace operators which contain more than two field insertions, say four, six, fields in the trace. Here the Vasiliev-like theory is the restriction to the bilinears. In AdS_4 we can realize the dream of constraining the theory from its symmetries. If the higher spin symmetry is unbroken, we can determine all the correlation functions on the AdS boundary. This can be done not only at tree level, but for all values of the bulk coupling constant $(1/N)$. This is one way of reading the results of [60]. Here we are assuming that the AdS theory is such that we can define boundary correlators obeying the usual axioms of a CFT. In other words, we are assuming the AdS/CFT correspondence. In Vasiliev-type theories the higher spin symmetry can be broken only by two (or three) particle states, which is what we study in this article. Finally, one would like to study theories that, besides the higher spin fields, also contain enough extra fields that can Higgs the higher spin symmetry at the classical level. In such theories the mass of the higher spin particles would be nonzero and finite even for very small bulk coupling. First one needs to understand what kind of "matter" can be added to the Vasiliev theory and still preserve the higher spin symmetry. Since the higher spin symmetry is highly constraining, it is likely that the only "matter" that we can add is what results from large N gauge theories. As far as we know, this is an unproven speculation, and it is closely related to the conjecture above. Note that such gauge theories have a Hagedorn density of states, so that the bulk theories would look more like ordinary string theories. Once, this problem is understood, one could consider a case where the Higgs mechanism can be introduced with a small parameter (as in $\mathcal{N} = 4$ SYM, for example). Here we will have single trace terms in the divergence of the higher spin currents. This breaking mechanism will be constrained by the higher spin symmetry. Understanding how it is constrained when the coupling is small will probably give us clues for how the mechanism works for larger values of the coupling. In addition, it could give us a way to do perturbation theory for correlators in gauge theories in a completely on shell fashion. Whether this idea is feasible or not, it remains to be seen...

Chapter 4

Convexity and Liberation at Large Spin

4.1 Introduction

Even though conformal field theories (CFTs) have been studied for many years, very little is known about the general structure of their operator spectrum and three-point functions. The interest in conformal field theories stems from their role in the description of phase transitions [83] and from their relation to renormalization group flows.[1] In addition, significant motivation to study conformal field theories is derived from the connection between quantum gravity in AdS and conformal field theories [6, 7, 8].

A classic result is that, in unitarity CFTs, operators obey the so-called unitarity bounds [91] (see also [92]). For example, the dimension of primary symmetric traceless tensors of spin $s \geq 1$ satisfies the following inequality

$$\Delta \geq d - 2 + s \ . \tag{4.1}$$

(Here and below d is the dimension of space-time.)

Primary operators that saturate the bound are known to be conserved, namely, they satisfy the equation $\partial_\mu J^\mu_{\cdots} = 0$ inside any correlation function. Similarly, scalar operators ($s = 0$) are known

[1] A priori there may also be renormalization group flows where one finds scale invariant theories at the end points without the special conformal generators. The feasibility of this scenario has been discussed recently, for example, in [84, 85, 86, 87, 88, 89, 90]. See also references therein. To date, scale invariant but non-conformal theories have not been constructed.

to satisfy $\Delta \geq \frac{d-2}{2}$. If this inequality is saturated, a second order differential equation holds true $\Box\mathcal{O} = 0$.

The unitarity bounds above are traditionally derived from reflection positivity of Euclidean correlation functions. It has been known for a while that considering processes in Lorentz signature can lead to further constraints on quantum field theory, and often such constraints do not seem to admit a straightforward derivation in Euclidean space. Simple examples based on an application of the optical theorem are discussed in [93, 94, 95, 96].

In this note we will use some properties of quantum field theory that are particularly easy to understand in Minkowski space in order to derive various constraints. In addition, we will find it useful to embed our conformal field theories in renormalization group flows that lead to a gapped phase.

Before we plunge into a detailed summary of our results, we would like to review quickly some basic notions in conformal field theory, mostly in order to set the terminology we will be using throughout this note. Given scalar operators $\mathcal{O}_1(x)$, $\mathcal{O}_2(x)$ we can consider their OPE

$$\mathcal{O}_1(x)\mathcal{O}_2(0) \sim \sum_s \sum_{k=1}^{\alpha(s)} C_s^k x^{-\Delta_1-\Delta_2+\Delta_s^k-s} x^{\mu_1} x^{\mu_2} \cdots x^{\mu_s} \mathcal{O}^k_{\mu_1,\mu_2,\ldots,\mu_s}(0) \ . \tag{4.2}$$

Clearly, only symmetric traceless representations of the Lorentz group can appear. The sum over s is the sum over spin and the sum over k is the sum over all the operators that have spin s. The sum over k is a finite sum. One can further re-package the sum above in terms of representations of the conformal group by combining the contributions of all the operators that are derivatives of some primary operator. (The coefficients C_s^k are fixed by conformal symmetry in terms of the coefficients of primary operators.)

In Euclidean space it is natural to consider $x^2 \to 0$ which means that all the components of x^μ go to zero. In this case the terms that dominate the sum (4.3) are the operators $\mathcal{O}^k_{\mu_1,\ldots,\mu_s}$ with the smallest dimensions Δ_s^k.

In Minkowski signature we can send $x^2 \to 0$ while being on the light cone. For convenience, we can take light-cone coordinates on a two-dimensional plane x^-, x^+, and set $x^i = 0$. Then we consider the light-cone limit $x^+ \to 0$ while x^- is finite. The operators that dominate then are $\mathcal{O}_{-,-,\ldots,-}$ and their contribution to the OPE goes like $x^{-\Delta_1-\Delta_2+\Delta_s^k-s}$, hence the strength with which they contribute is dictated by the twist $\tau_s^k \equiv \Delta_s^k - s$. We denote $\tau_s^* \equiv \min_k(\tau_s^k)$, i.e. τ_s^* is the minimal twist that appears in the OPE for a given spin. Another useful notation is $\tau_{\min} \equiv \min_s \tau_s^*$, that is, the overall minimal twist that appears in the OPE, *excluding the unit operator*, which has twist zero.

The twists of conformal primaries are constrained by (4.1). Traceless symmetric conformal primaries of spin $s \geq 1$ must have twist at least $d-2$ and a scalar conformal primary must have twist at least $\frac{d}{2}-1$. The unit operator has twist zero. The energy momentum tensor and conserved currents have twist $d-2$.

When two points are light-like separated, the group $SL(2,R)$ preserves the corresponding light-ray. It is therefore natural to consider representations of it. The representations are classified by collinear primaries of dimension Δ and spin s_{+-} in the two dimensional plane. We can act on this collinear primary any number of times with ∂_- in order to generate the complete collinear representation. The twist of a collinear primary is defined as $\tau^{\rm coll} \equiv \Delta - s_{+-}$.

A conformal primary $\mathcal{O}_{\mu_1,\ldots,\mu_n}$ can be decomposed into (generally infinitely many) collinear primaries. In this decomposition, the collinear primary with the smallest collinear twist is clearly $\mathcal{O}_{-,\ldots,-}$, and this minimal collinear twist coincides with the conformal twist.

It would be useful for our purposes to rewrite the OPE (4.3) in terms of collinear representations. From the discussion above we see that it takes the form as $x^+ \to 0$

$$\mathcal{O}_1(x^+,x^-)\mathcal{O}_2(0,0) \sim (x^+)^{-\frac{1}{2}(\Delta_1+\Delta_2)} \times \sum_{\tau^{\rm coll},s_{+-},\ldots} \widetilde{C}_s^k \, (x^+)^{\frac{1}{2}\tau^{\rm coll}} \mathcal{F}(x^-,\partial_-) \mathcal{O}^k_{\underbrace{-,-,\ldots,-}_{\frac{1}{2}(s+s_{+-})},\underbrace{+,+\ldots,+}_{\frac{1}{2}(s-s_{+-})}}(0) \; . \tag{4.3}$$

The functions $\mathcal{F}$ are determined from representation theory, see [97] for a review. The important message in (4.3) is that the approach to the light cone is controlled by the collinear twists.

Having introduced the basic terminology that we will employ, let us now summarize the constraints we discuss in this note.

First, we merely extend the observations of Nachtmann [98] about the QCD sum rules to an argument pertaining to arbitrary CFTs. In [98] it was pointed out that the Deep Inelastic Scattering (DIS) QCD sum rules imply certain convexity properties of the high energy limit of QCD. We discuss such constraints on the operators in the OPE in the context of a generic CFTs. We emphasize the assumptions involved in this argument.

To this end, we consider an RG flow where the CFT of interest is the UV theory and in the IR we have a gapped phase. We consider an experiment with the operator of interest, $\mathcal{O}(x)$, playing the role of the usual electromagnetic current in DIS. Assuming Regge asymptotics of the amplitude in the gapped phase (namely, the amplitude is bounded by some power of s)

$$\lim_{s\to\infty} A(s,t=0) \leq s^{N-1} \tag{4.4}$$

one is led to a convexity property of the minimal twists τ_s^* of operators appearing in the $\mathcal{O}(x)\mathcal{O}^\dagger(0)$ OPE

$$\frac{\tau_{s_3}^* - \tau_{s_1}^*}{s_3 - s_1} \leq \frac{\tau_{s_2}^* - \tau_{s_1}^*}{s_2 - s_1} \ , \qquad s_3 > s_2 > s_1 \geq s_c \ . \tag{4.5}$$

Here s_c is the minimal spin above which the convexity property starts. From (4.4) we know that s_c is finite and certainly does not exceed N. The argument leading to the convexity property (4.5) depends on the following three assumptions:

- a) unitarity;
- b) the CFT can flow to a gapped phase;
- c) polynomial boundedness of the DIS cross section in the gapped phase (4.4).

Assumption c) may seem as the least innocuous. However, there are formal arguments connecting it to causality and other basic properties of QFT. See for example [99], and a more recent discussion in [100]. In addition, we do not know of field theories violating assumption b) (such theories are known to exist in the realm of critical phenomena, but, to our knowledge, there are no concrete examples of unitary Lorentz invariant theories that cannot be deformed to a gapped phase). Hence, convexity of the function τ_s^* is established under very general assumptions. Moreover, in all the examples we checked, the convexity above holds true starting from $s_c = 2$. We do not prove here that this is the case in most generality.

A second topic we study here is the crossing symmetry (i.e. the bootstrap equations) in Lorentzian signature. The bootstrap equations were introduced long ago by [101, 102, 103]. It has been demonstrated recently, starting from [104], that one can go about and systematically bound numerically solutions to these formidable equations. Sometimes these numerical bounds seem to be saturated by actual CFTs, as has been demonstrated, for example, in [105]. These methods of bounding solutions to the bootstrap equations are easily described in Euclidean space, and the consistency of the procedure follows from the fact that the contribution of the operators with high dimension to the OPE in Euclidean space is exponentially suppressed [106].

By contrast, as we recalled above, the structure of the operator product expansion in Minkowski space is such that we can probe directly operators with very high spin and small twist. We therefore use the bootstrap methods in Lorentzian signature to constrain operators with high dimension and high spin, but with low twist.[2]

Studying the light-cone OPE and bootstrap equations in the appropriate limit, we conclude that starting from any two primaries with twists τ_1 and τ_2, operators with twists arbitrarily close to $\tau_1 + \tau_2$ are necessarily present. This can be regarded as an additivity property of the twist spectrum of general CFT. In the derivation of this result, it is important to assume that $d > 2$. Indeed, this additivity property is definitely violated in two-dimensional models, such as the minimal models. The condition $d > 2$ comes from the fact that we need to separate the contribution of the unit operator from everything else in the light-cone OPE.

In particular, this means that the minimal twists appearing in the OPE of an operator $\mathcal{O}_1$ with any other operator $\mathcal{O}_2$ are bounded from above as follows (in $d > 2$):

$$\lim_{s\to\infty} \tau_s^* \leq \tau_{\mathcal{O}_1} + \tau_{\mathcal{O}_2} \ . \tag{4.6}$$

We conclude that there should not be unitary solutions of the crossing equations with $\tau_s^* > \tau_{\mathcal{O}_1} + \tau_{\mathcal{O}_2}$. This could be useful to bear in mind when trying to bound solutions to the crossing equations along the lines of [104].

Additivity implies that for large enough spin there are operators whose twists are arbitrary close to $\tau_{\mathcal{O}_1} + \tau_{\mathcal{O}_2}$. We refer to these operators as double-twist operators. We also denote these operators symbolically as $\mathcal{O}_1 \partial^s \mathcal{O}_2$. Similar argument shows that operators with twists asymptoting to $\tau_{\mathcal{O}_1} + \tau_{\mathcal{O}_2} + 2n$ for any integer n are present as well. Those can be denoted symbolically as $\mathcal{O}_1 \partial^s \Box^n \mathcal{O}_2$. We do not discuss the case $n \neq 0$ in as much detail as $n = 0$. In addition, it also makes sense to talk about multi-twist operators.

Given that their twists approach $\tau_{\mathcal{O}_1} + \tau_{\mathcal{O}_2}$ for large enough spin, it is interesting to ask how precisely this limit of $\tau_{\mathcal{O}_1} + \tau_{\mathcal{O}_2}$ is reached. It has been argued in [110] that for sufficiently large spin we can parameterize the twists of these operators as follows

$$\tau(s) = \tau_{\mathcal{O}_1} + \tau_{\mathcal{O}_2} - \frac{c_{\tau_{\min}}}{s^{\tau_{\min}}} + \cdots \ . \tag{4.7}$$

[2]Different applications of the bootstrap equations in Lorentzian signature were discussed in [107, 108, 109]. In these papers the properties of operators with both high spin and high twist were analyzed.

Here $\tau_{\rm min}$ is the twist of $\mathcal{O}_{\rm min}$ which is the operator of smallest twist exchanged by $\mathcal{O}_1$ and $\mathcal{O}_2$.[3] (As in the definition before, we exclude from this the unit operator.) In many cases we would expect this to be the energy momentum tensor, thus $\tau_{\rm min} = d-2$.

The essence of the argument of [110] for (4.7) is as follows. One maps conformally the CFT from flat space to an $AdS_3 \times S^{d-3}$ background. By choosing special coordinates in AdS_3 one can think about the four-point functions above as being described by a two-dimensional gapped theory embedded in AdS_3. The energy in this theory is the twist, and (4.7) corresponds to the leading interaction between separated localized excitations where the separation is governed by the spin.

Here we arrive at the same conclusion (4.7) by studying the crossing equation in flat Minkowski space. Furthermore, by using the bootstrap equations, we compute the coefficient $c_{\tau_{\rm min}}$ in (4.7) analytically in every $d>2$ CFT. The simplest case is when we consider the OPE of an operator with its Hermitian conjugate $\mathcal{O}(x)\mathcal{O}^\dagger(0)$. Then the coefficient governing the asymptotic twists of double-twist operators is

$$c_{\tau_{\rm min}} \quad = \frac{\Gamma(\tau_{\rm min}+2s_{\rm min})}{2^{s_{\rm min}-1}\Gamma\left(\frac{\tau_{\rm min}+2s_{\rm min}}{2}\right)^2}\frac{\Gamma(\Delta_{\mathcal{O}})^2}{\Gamma(\Delta_{\mathcal{O}}-\frac{\tau_{\rm min}}{2})^2}f^2 \ , \ f^2 \quad = \frac{C^2_{\mathcal{O}\mathcal{O}^\dagger\mathcal{O}_{\tau_{\rm min}}}}{C_{\mathcal{O}\mathcal{O}}C_{\mathcal{O}^\dagger\mathcal{O}^\dagger}C_{\mathcal{O}_{\tau_{\rm min}}\mathcal{O}_{\tau_{\rm min}}}} \ , \qquad (4.8)$$

where $s_{\rm min}$ and $\tau_{\rm min}$ are the spin and the twist of the minimal twist operator (other than the unit operator) appearing in the OPE of $\mathcal{O}$ with its conjugate. Hence, $c_{\tau_{\rm min}}$ is fixed by some universal factors and the two- and three-point functions of the operators $\mathcal{O}$ and $\mathcal{O}_{\tau_{\rm min}}$. As we remarked, in many models one would expect $\mathcal{O}_{\tau_{\rm min}}$ to be the stress tenors, in which case the three-point function above is fixed in terms of $\Delta_{\mathcal{O}}$ and the two-point function of the stress tensor (which is some measure of the number of degrees of freedom of the theory).

Note that in unitary CFTs we have $\tau_{\rm min} \le 2\Delta_{\mathcal{O}}$, and an equality can only be reached for free fields. In general, there could be several operators of the same minimal twist $\tau_{\rm min}$. In this case c is the sum over all of these operators. It is worth mentioning that (4.8) provides another point of view on convexity. Since all the leading contributions to c are manifestly positive, the spectrum of double-twist operators at large enough spins is convex. This statement is slightly different from the one outlined in (4.5), which is about the minimal twists in reflection positive OPEs.[4] Here the statement is about double-twist operators in reflection positive OPEs.

We also consider non-reflection positive OPEs, such as those of a charged operator (with charge q under some $U(1)$ global symmetry, with the generalization to non-Abelian symmetries being straight-

[3] By this we mean it appears both in the OPE of $\mathcal{O}_1\mathcal{O}_1^\dagger$ and $\mathcal{O}_2\mathcal{O}_2^\dagger$.

[4] In this note when we say "reflection positive OPE" we mean an OPE of the type $\mathcal{O}(x)\mathcal{O}(0)^\dagger$.

forward) $\mathcal{O}$ with itself, $\mathcal{O}(x)\mathcal{O}(0)$. The double-twist operators $\mathcal{O}\partial^s\mathcal{O}$ with twists that approach $2\Delta_{\mathcal{O}}$ are necessarily present. Here we can compute the corrections to the twists at large spin $\tau_s - 2\Delta_{\mathcal{O}}$ and we find a formula very similar to (4.8) (we will describe it in detail in the text) with one small but important difference, a factor of $(-1)^s$. Hence, the conserved current and the energy momentum tensor, both of which have twist $d-2$ contribute with opposite signs. We assume those are the minimal twist operators in the OPE $\mathcal{O}\mathcal{O}^\dagger$. The spectrum is non-concave at large enough spin if and only if

$$\Delta \geq \frac{d-1}{2}|q| \ , \tag{4.9}$$

in some normalization for the two-point function of the $U(1)$ current that we define carefully in the main text. Hence convexity at large spin of such operators is related to whether or not the operator $\mathcal{O}$ satisfies a BPS-like bound.

In SUSY theories, for chiral primaries, (4.9) is saturated (the charge q is that of the superconformal $U(1)_R$ symmetry) and the leading corrections to the anomalous twists vanish. For chiral operators that are not superconformal primaries, (4.9) is satisfied in unitary theories. Hence, the spectrum of twists of the double-twist operators $\mathcal{O}\partial^s\mathcal{O}$ is convex. It would be interesting to understand more generally whether convexity at large spin is a general phenomenon or not. The arguments in this note show that a large class of OPEs is consistent with being convex.[5]

If we combine (4.5) and (4.6) (i.e. convexity and additivity) we are led to rather stringent predictions concerning interesting models. In particular, we predict that in the 3d critical $O(N)$ models (which are highly relevant for 3d phase transitions) there is a set of almost conserved currents of spin $s \geq 4$ with the dimension $\Delta_s = s + \tau_s^*$ with

$$1 \leq \tau_s^* \leq 2\Delta_\sigma \ , \tag{4.10}$$

where Δ_σ is the dimension of the spin field σ_i. Furthermore, we can argue using additivity that

$$\lim_{s\to\infty} \tau_s^* = 2\Delta_\sigma \ , \tag{4.11}$$

and we can characterize in detail how this limit is approached via (4.7) and (4.8). One can use the numerical results from the literature [111] to arrive at concrete predictions. Furthermore, the function τ_s^* is monotonically rising and convex as a function of the spin.

[5]Throughout the chapter, convexity is to be understood as non-concavity.

In the case of the 3d Ising model, it is known that there is a spin 4 operator with the twist $\tau_4^* \sim 1.02$, which agrees with our expectations. Using convexity, this allows us to further narrow down the possible window for higher spin currents and predict that in the 3d Ising model (which describes, among others, boiling water)

$$1.02 \leq \tau_s^* \leq 1.037, \quad s \geq 6 \ , \qquad\qquad \lim_{s\to\infty} \tau_s^* = 2\Delta_\sigma \approx 1.037. \tag{4.12}$$

Recall that τ_s^* is again monotonic and convex as a function of the spin. In addition, the upper bound above is in fact saturated for $s \to \infty$ (this follows from additivity and the fact that the set of operators of the Ising model is continuously connected to free field theory). We will present several other examples in the text.

We also discuss the relation of some of our results to small breaking of higher spin symmetry and, through AdS/CFT, to quantum gravity in AdS.

For the parity preserving Vasiliev theory in AdS_4 [24] with higher spin breaking boundary conditions, (4.11) implies for the masses of the higher spin bosons in the bulk

$$m_s^2 \approx 4\gamma_\sigma s + O(1), \quad s \gg 1 \tag{4.13}$$

here γ_σ is the anomalous dimension of the spin field in a non-singlet $O(N)$ model and the result is exact in N.

The outline of the chapter is as follows. In section 2 we discuss DIS and establish (4.5). In section 3 we consider the bootstrap equations in Lorentz signature and derive (4.6),(4.7),(4.8). We also review the argument of [110] in this section. In sections 4,5 we consider examples and applications. In section 6 we describe some open problems and conclude. Five appendices complement the main text with technical details.

4.2 Deep Inelastic Scattering

In this section we explain how one can derive inequalities about operator dimensions in CFT by considering RG flows that start in the ultraviolet from the CFT and end up in the gapped phase. The basic ingredients are as follows.

Take any CFT (in any number of space-time dimensions, d) and assume it can be perturbed by a relevant deformation to a gapped phase (see 4.1). Then we take the lightest particle in the Hilbert

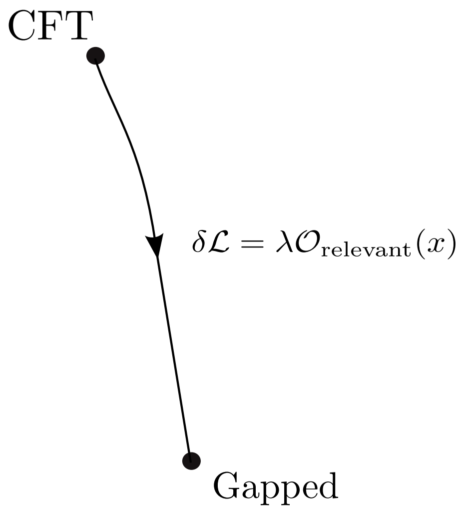


Figure 4.1: A relevant perturbation of a CFT that leads to a gapped phase.

space, and denote this particle by $|P\rangle$. Its mass is denoted by M.

There may exist CFTs which have no relevant operators whatsoever ("self-organizing" CFTs), or CFTs which have relevant operators but can never flow to a gapped phase. We are not aware of such examples, and we will henceforth assume that flowing to a gapped phase is possible. In the critical $O(N)$ models one can make the following heuristic argument to support this scenario.

The critical $O(N)$ models can be reached if one tunes appropriately the relevant perturbations of free N bosons. On the other hand, if the mass parameter in the UV is large enough, the model clearly flows to a gapped phase. Now we can decrease the mass gradually until the RG flow hovers very close to the critical $O(N)$ model. In this case we can describe the late part of the flow as a perturbation of the $O(N)$ model by the energy operator, $\int d^3x\epsilon(x)$. Indeed, it is a primary relevant operator in the critical $O(N)$ models.[6] If there are no phase transitions as a function of this mass, then it means that the critical $O(N)$ models, perturbed by the energy operator, flow to a gapped phase. This is also consistent with the fact that the energy operator corresponds to dialing the temperature away from the critical temperature.

This heuristic argument is summarized in fig. 4.2. In the following we discuss general CFTs, assuming they can be deformed to a gapped phase by adding relevant operators (and maintaining unitarity).

[6]This is known to be certainly true at large N and at some small values of N. It seems natural that this would be indeed the case for any N.

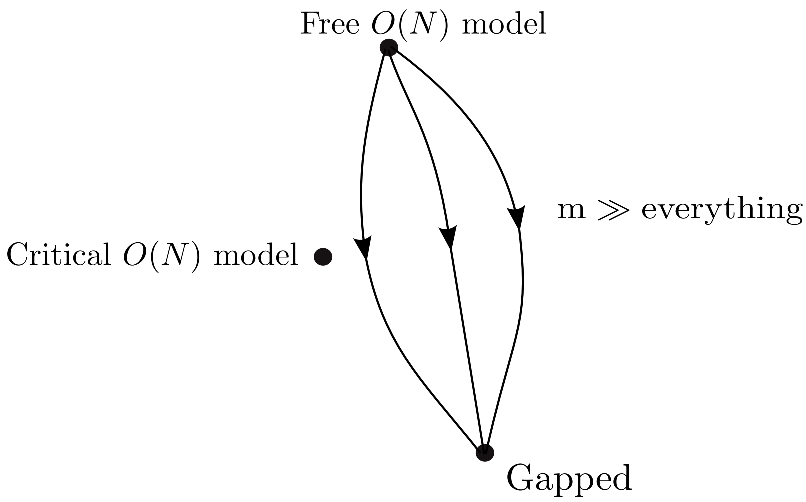


Figure 4.2: When we add a large mass m in the ultraviolet, the model clearly flows to a gapped phase. We can gradually reduce the mass and tune the flow such that it passes close to the critical model. The flow after this stage can be interpreted as a perturbation of the critical $O(N)$ model by the energy operator, $\epsilon(x)$.

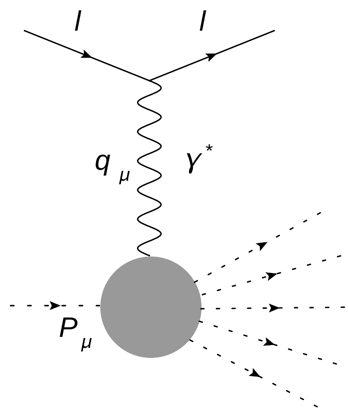


Figure 4.3: A lepton emits a virtual, space-like, photon that strikes a hadron. As a result, the hadron generally breaks up into a complicated final state.

4.2.1 General Theory

Deep Inelastic Scattering (DIS) allows to probe the internal structure of matter. One bombards some state with very energetic (virtual) particles and examines the debris. Of particular interest is the total cross section for this process, as a function of various kinematical variables.

Traditionally, the most common setup to consider is that of fig. 4.3. A lepton emits a virtual, space-like, photon that strikes a hadron.

To study the inner structure of $|P\rangle$ we can try to shoot various particles at it. If the theory (i.e. the CFT and the relevant deformations) preserves a global symmetry, we can couple the conserved currents to external gauge fields and repeat the experiment of fig. 4.3. (We can introduce arbitrarily

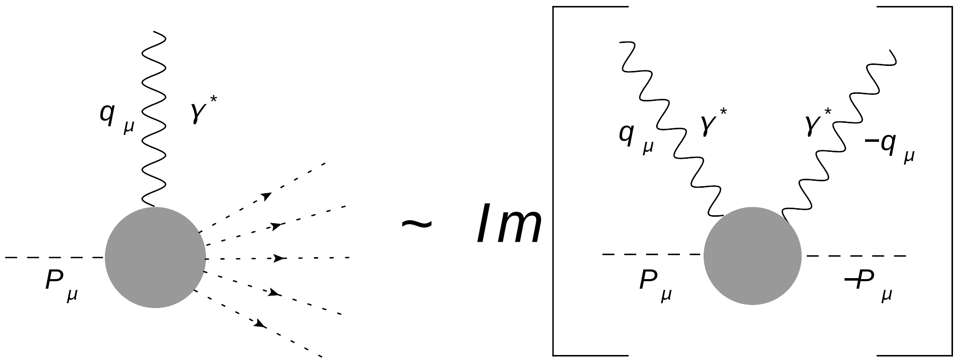


Figure 4.4: The total inclusive cross section can be extracted from the imaginary part of a Compton-like scattering where the momenta of the outgoing states are identical to the momenta of the incoming states.

weakly interacting charged matter particles and make the gauge field dynamical in order to have a perfect analog of fig. 4.3.) A more universal probe that exists in any local QFT is the EM tensor. We can couple it to a background graviton and consider the same experiment. We can also perform DIS with any other operator in the theory. Let us start by reviewing the kinematics of DIS in the case of a scalar background field $J(x)$. In other words, a background field that is coupled to some scalar operator in the theory $\mathcal{O}(x)$ via $\int d^dxJ(x)\mathcal{O}(x)$. ($J(x)$ and $\mathcal{O}(x)$ are taken to be real.)

The optical theorem allows to extract the inclusive amplitude for DIS via the analog of Compton scattering (see fig. 4.4). The amplitude in our "scalar DIS" setup is

$$A(q_\mu, P_\mu) \equiv \int d^dye^{iqy}\langle P|T\left(\mathcal{O}(y)\mathcal{O}(0)\right)|P\rangle \ , \tag{4.14}$$

where we have denoted the momentum of the target $|P\rangle$ by P_μ. We will only discuss the case of $q^2 < 0$, i.e. space-like momentum for the virtual particle. The above amplitude obviously depends on the mass scales of the theory (we set $P^2 = 1$ for convenience), and the two invariants $q^2, \nu \equiv 2q \cdot P$.

We can promote ν to a complex variable. Since we have assumed the particle $|P\rangle$ is the lightest in the theory, the above amplitude has a branch cut for $\nu \geq -q^2$ and $\nu \leq q^2$. The ν plane is depicted in fig. 4.5. The optical theorem connects the discontinuity along the branch cut of (4.14) to the square of the DIS amplitude. (There are no other cuts since q is space-like.)

To get a handle on (4.14), we can invoke the OPE expansion. In the case that $\mathcal{O}(x)$ is real, only even spins contribute. (Odd spins would be allowed if $\mathcal{O}^+ \neq \mathcal{O}$.)

$$\mathcal{O}(y)\mathcal{O}(0) = \sum_{s=0,2,4,\ldots}\ \sum_{\alpha\in\mathcal{I}_s} f_s^{(\alpha)}(y)y^{\mu_1}\ldots y^{\mu_s}\mathcal{O}^{(\alpha)}_{\mu_1\ldots\mu_s}(0) \ . \tag{4.15}$$

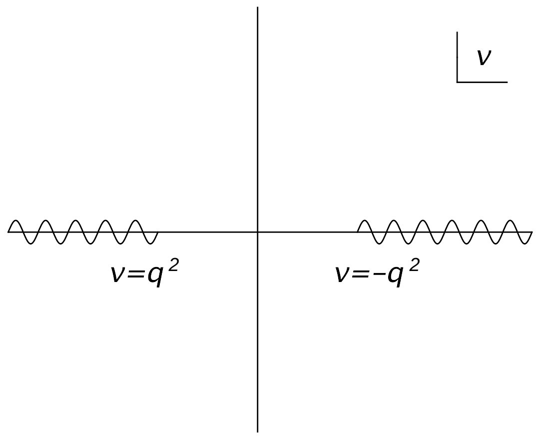


Figure 4.5: The analytic structure in the ν plane.

Here s labels the spins, and we only write primary fields and ignore descendants since they do not contribute to the correlation function under consideration. The set $\mathcal{I}_s$ corresponds to the collection of operators of spin s.

Conformal symmetry at small y fixes the OPE coefficients in (4.15) to be

$$f_s^{(\alpha)}(y) = C_s^{(\alpha)} y^{\tau_s^{(\alpha)} - 2\Delta_{\mathcal{O}}} \left(1 + \cdots\right) \ , \tag{4.16}$$

where $\tau_s^{(\alpha)} = \Delta_{\mathcal{O}^{(\alpha)}_{\mu_1 \ldots \mu_s}} - s$ is the twist of the operator. The $\cdots$ denote corrections suppressed by positive powers of y^2. These corrections depend on the relevant operators. Parameterizing these corrections would be unnecessary for our purposes.

The expectation values of $\mathcal{O}^{\alpha}_{\mu_1 \ldots \mu_s}$ in the state P are parameterized as follows

$$\langle P | \mathcal{O}^{\alpha}_{\mu_1 \ldots \mu_s}(0) | P \rangle = \mathcal{A}_s^{(\alpha)} \left(P_{\mu_1} P_{\mu_2} \cdots P_{\mu_s} - \text{traces} \right) \ , \tag{4.17}$$

with the $\mathcal{A}_s^{(\alpha)}$ some dimensionful coefficients.

We can now insert the expansion (4.15) into (4.14) and find

$$A(q_\mu, P_\mu) \equiv \sum_{s=0,2,4,\ldots} \sum_{\alpha \in \mathcal{I}_s} \mathcal{A}_s^{(\alpha)} \left(\left(P \cdot \frac{\partial}{\partial q} \right)^s - \text{traces} \right) \widetilde{f}_s^{(\alpha)}(q) \ . \tag{4.18}$$

Here "traces" stands for terms of the form $P^2 \left(P \cdot \frac{\partial}{\partial q} \right)^{s-2} \left(\frac{\partial}{\partial q} \cdot \frac{\partial}{\partial q} \right)$ etc.

At this point it is natural to switch to the variable $x = -q^2/\nu$, and write the answer in terms of x, q^2. The OPE allows to control easily the leading terms as $-q^2 \to \infty$. Hence, for any given power of x we keep only those terms that are leading in the limit of large $-q^2$.

The leading terms take the form

$$A(x,q^2) \equiv \sum_{s=0,2,4,\ldots} \mathcal{A}_s^* C_n^* x^{-s} (q^2)^{-\frac{1}{2}\tau_s^* + \Delta_{\mathcal{O}} - d/2} \ . \tag{4.19}$$

Among the set of operators of given spin s, $\mathcal{I}_s$, we select the one that has the smallest twist and denote the corresponding coefficients and twists with an asterisk. In the limit of large $-q^2$ the "traces" appearing in (4.18) can be dropped.[7]

We see that the OPE expansion is useful for finite ν and large q^2 (i.e. large x and large q^2). The physically relevant configuration is, however, large q^2 and $x \in [0,1]$. These two regimes can be related by the usual contour deformations in the complex plane.

To this end we promote x to a complex variable and recall that the branch cut extends over $x \in [-1,1]$. Being careless for a moment about the behavior near $x \sim 0$, we write sum rules by following the usual trick of pulling the contour from infinite x to the branch cut. We define $\mu_s(q^2)$ as the s-th moment

$$\mu_s(q^2) = \int_0^1 dx x^{s-1} \mathrm{Im} A(x,q^2) \ , \tag{4.20}$$

and from (4.18) we infer the sum rule for large $-q^2$

$$\mu_s(q^2) \to (q^2)^{-\frac{1}{2}\tau_s^* + \Delta_{\mathcal{O}} - d/2} \mathcal{A}_s^* C_s^* \ . \tag{4.21}$$

The finiteness of these moments depends on the small x behavior. Equivalently, it depends on the behavior of the amplitude for fixed q^2 and large ν. This is the Regge limit. We would like to take a conservative approach to the problem of determining the asymptotics in the Regge limit. We simply

[7] The reason is as follows. The nature of the expansion is that we keep all powers of x and for each power we only retain the most dominant power of q^2. The trace terms are negligible, because, for instance, the first "trace term" scales as $x^{-s+2}(q^2)^{-\frac{1}{2}\tau_s^* + \Delta_{\mathcal{O}} - d/2 - 2}$. This has to be compared to the contribution from the spin $s-2$ operator, $x^{-s+2}(q^2)^{-\frac{1}{2}\tau_{s-2}^* + \Delta_{\mathcal{O}} - d/2}$. We see that the "trace terms" can be indeed consistently neglected, for example, if the smallest twist of a spin n operator, τ_s^*, is monotonically increasing as a function of the spin. We will justify this assumption self-consistently later. This subtlety is apparently overlooked in various places. Another way to avoid this issue is to use the spin projection trick [98].

assume the amplitude is bounded polynomially[8]

$$\lim_{x\to 0} A(x,q^2) \leq x^{-N+1} \ , \tag{4.22}$$

for some integer N. Polynomial boundedness is discussed in [99]. A recent discussion and more references to the original literature where polynomial boundedness is discussed can be found in [100]. In this case the contour manipulation leading to (4.20) can be justified only for $s \geq N$. This has to be borne in mind in the following, where we derive some consequences of (4.20) using convexity properties. The first place we are aware of where convexity properties appear in this context is [98].

4.2.2 Convexity

We will discuss various simple inequalities that the moments $\mu_s(q^2)$ have to satisfy. These properties follow simply from unitarity

$$\mathrm{Im}A(x,q^2) \geq 0 \ , \tag{4.23}$$

which is nothing but saying that the cross section for the process of DIS is nonnegative. In fact, $\mathrm{Im}A(x,q^2)$ must be nonzero at least around some points, otherwise, the scattering is trivial.

The general inequalities stated below can be proven with the method we are using only for $s \geq N$ with some finite N. However, they may or may not be true for smaller spins as well. We will denote the spin from which these inequalities become true by s_c. We see that s_c is finite and it is at most N.

From (4.20) it is clear (because of (4.23)) that $\mu_s > \mu_{s+1}$, which together with the sum rule (4.21) leads to

$$\tau_s^* \leq \tau_{s+1}^* \ . \tag{4.24}$$

Hence, the smallest twist as a function of the spin is a nondecreasing function. There is a stronger conclusion that follows from (4.20),(4.21),(4.23). That is, take three ordered even integers $s_1 < s_2 <$

[8]For on-shell scattering processes, we have the bound of Froissart-Martin [112, 113] $\sigma \leq C(\log(s))^2$, where C is some dimensionful coefficient. In deep inelastic scattering we are dealing with an amplitude involving an off-shell particle (for example, a virtual photon), so the argument of Froissart-Martin does not carry over. In the context of QCD, one can use various phenomenological approaches to model the small x behavior. For example, see the review [114]. The result of these phenomenological models is that the amplitude grows only as a power of a logarithm with $1/x$. This may be more general than QCD, for example, it would be interesting to understand the small x asymptotics in $\mathcal{N}=4$ [115]. We thank M. Lublinsky and A. Schwimmer for discussions about the Regge limit.

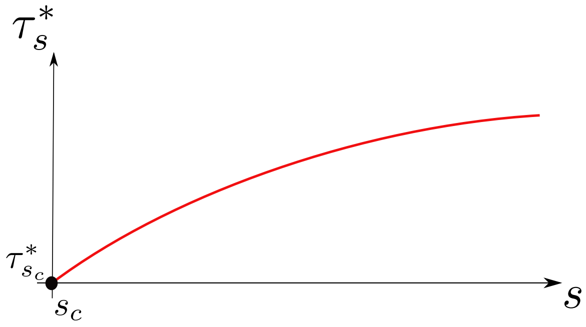


Figure 4.6: The generic structure of the spectrum of minimal twists in a CFT.

s_3, then

$$\left(\frac{\mu_{s_3}}{\mu_{s_1}}\right)^{s_2-s_1} \geq \left(\frac{\mu_{s_2}}{\mu_{s_1}}\right)^{s_3-s_1} . \tag{4.25}$$

This implies for the twists

$$\frac{\tau_{s_3}^* - \tau_{s_1}^*}{s_3 - s_1} \leq \frac{\tau_{s_2}^* - \tau_{s_1}^*}{s_2 - s_1} . \tag{4.26}$$

Hence, the twist is a monotonic, convex (more precisely, non-concave) function of the spin.[9] We remind again that these inequalities start from some finite spin, s_c. In all the examples we checked they are in fact true for all spins above two (at spin two we have the energy momentum tensor, which has the minimal possible twist allowed by unitarity, $d-2$).

The typical resulting spectrum of minimal twists is depicted in fig. 4.6. This is consistent with the known examples of CFTs in $d > 2$. The situation is slightly subtler in $d = 2$. In two-dimensional CFTs there are infinitely many operators with vanishing twist (these are found in the Verma module of the unit operator). Hence, the minimal twist is zero for all spins. This is however consistent with our analysis here, as we have only shown that τ_s^* is non-concave. (When all the minimal twists are identical, unitarily (4.23) simply translates to inequalities between the coefficients appearing in (4.21).)

[9]Let us first prove this in the case $s_3 = s_2 + 2 = s_1 + 4$, and we denote $s \equiv s_1$. In this case, (4.25) is true due to the simple identity

$$0 < \frac{1}{2} \int\limits_{(x,y)\in[0,1]^2} dxdy x^{s-1} y^{s-1} (x^2 - y^2)^2 \mathrm{Im}A(x,q^2)\mathrm{Im}A(y,q^2) = \mu_s \mu_{s+4} - \mu_{s+2}^2 .$$

This establishes local convexity, namely,

$$\tau_{s+4}^* - \tau_s^* \leq 2\left(\tau_{s+2}^* - \tau_s^*\right) .$$

From this, global convexity (4.26) follows trivially. One can also show that without further assumptions on $\mathrm{Im}A(x,q^2)$ (or some extra input) there is no stronger inequality that follows from (4.20),(4.21),(4.23).

So far we have established convexity and monotonicity (starting from some finite spin). It is also natural to ask what is the asymptotic limit of the minimal twist as the spin goes to infinity. We will see that τ_s^* has a finite and determined limit as $s \to \infty$. This is the topic of the next section.

4.3 Crossing Symmetry and the Structure of the OPE

In this section we analyze the crossing equations in Minkowski space. Using a conformal transformation we can always put any four points on a plane. On this plane it is natural to introduce light-like coordinates, such that the metric on $R^{d-1,1}$ takes the form $ds^2 = dx^+dx^- + (dx^i)^2$. We study the crossing equations in the limit when one of the points approaches the light cone of another point. We have explained in the introduction that one should distinguish between the collinear twist $\Delta - s_{+-}$ and the conformal twist $\Delta - s$.

Our analysis consists of two parts. First, we analyze how the unit operator constrains general CFTs via the crossing equations. We conclude that in any unitary CFT in $d > 2$, given any two operators with collinear twists τ_1^{coll} and τ_2^{coll}, we either have in the theory operators with collinear twist precisely $\tau_1^{\text{coll}} + \tau_2^{\text{coll}}$, or there exist operators whose twists are arbitrarily close to $\tau_1^{\text{coll}} + \tau_2^{\text{coll}}$.

Combined with convexity (which holds in reflection positive OPEs), this implies that in any unitary CFT the minimal twist obeys the following inequality[10]

$$\lim_{s\to\infty} \tau_s^* \leq 2\tau_{\mathcal{O}} \ . \tag{4.27}$$

So far we have discussed additivity in the space of collinear twists. We then give an argument of why in fact additivity holds in the space of the ordinary conformal twists. Namely, *given any two operators of twists τ_1 and τ_2, there will be an operator with twist arbitrary close to $\tau_1 + \tau_2$.* This sharper statement motivates the symbolic notation $\mathcal{O}_1\partial^s\mathcal{O}_2$ introduced above.

We present several arguments for this stronger statement, and we also emphasize that it can be understood in the analysis of Alday and Maldacena [110]. The arguments we give and the discussion in [110] provide very compelling reasons for making this stronger claim. We leave the task of constructing a detailed proof of this statement to the future.[11]

Thus, the spectrum of any CFT has an additivity property in twist space. It makes sense to talk

[10] In some sense, a predecessor of this statement is the so-called Callan-Gross theorem [116]. It was derived for Lagrangian scalar theories.

[11] In the setup of [110] this would amount to understanding better some analytic continuations. In our setup this presumably amounts to understanding in more detail the structure of the decomposition of a conformal primary into collinear primaries.

about double- and multi-twist operators. We denote the double-twist operators appearing in the OPE of $\mathcal{O}_1$ and $\mathcal{O}_2$ symbolically as $\mathcal{O}_1\partial^s\mathcal{O}_2$. The three-point functions $\langle\mathcal{O}_1\mathcal{O}_2\,(\mathcal{O}_1\partial^s\mathcal{O}_2)\rangle$ are given by the generalized free fields values to leading order in $\frac{1}{s}$. By analyzing terms that are sub-leading in the small z limit we conclude that operators which can be symbolically denoted as $\mathcal{O}_1\partial^s\Box^n\mathcal{O}_2$ appear as well, and for large spin their twist asymptotes to $\tau_1+\tau_2+2n$. Again the three-point functions $\langle\mathcal{O}_1\mathcal{O}_2\,(\mathcal{O}_1\partial^s\Box^n\mathcal{O}_2)\rangle$ approach the generalized free fields values to leading order in $\frac{1}{s}$. In the current analysis we do not discuss the case $n\neq 0$ in detail.

We then include in the crossing equations the operator with the minimal twist after the unit operator. Using the crossing equations we characterize how the limiting twist $\tau_1+\tau_2$ is approached as the spin is taken to infinity. The corrections to the twist, $\tau_s-\tau_1-\tau_2$, go to zero at large s as some power of s that we characterize in great detail below. For simplicity, we mostly concentrate on the case where $\mathcal{O}_1$ and $\mathcal{O}_2$ are scalar operators.

4.3.1 The Duals of the Unit Operator

As a warm-up, consider any large N vector model, and a four-point function of some flavor-neutral operators. (With minor differences the same holds for adjoint large N theories.) We decompose the four-point function into its disconnected and connected contributions

$$\begin{aligned}\langle\mathcal{O}(x_1)\mathcal{O}(x_2)\mathcal{O}(x_3)\mathcal{O}(x_4)\rangle &= \langle\mathcal{O}(x_1)\mathcal{O}(x_2)\rangle\langle\mathcal{O}(x_3)\mathcal{O}(x_4)\rangle+\langle\mathcal{O}(x_1)\mathcal{O}(x_3)\rangle\langle\mathcal{O}(x_2)\mathcal{O}(x_4)\rangle\\ &+\langle\mathcal{O}(x_1)\mathcal{O}(x_4)\rangle\langle\mathcal{O}(x_2)\mathcal{O}(x_3)\rangle+\langle\mathcal{O}(x_1)\mathcal{O}(x_2)\mathcal{O}(x_3)\mathcal{O}(x_4)\rangle_{conn}.\end{aligned} \tag{4.28}$$

For any configuration of points and large enough N, this correlator is dominated by the disconnected pieces. The connected contribution is suppressed by $\frac{1}{N}$ compared to the disconnected pieces. This fact by itself guarantees the existence of the double trace operators, which, as a matter of definition, we can denote $\mathcal{O}\partial^s\mathcal{O}$, with the scaling dimensions $\Delta_s=2\Delta_{\mathcal{O}}+s+O(\frac{1}{N})$.

To see that these operators are present and that their dimensions are as above, we simply expand the disconnected pieces in a specific channel. This exercise is reviewed in appendix B. Of course, it is well known that such operators indeed exist in all the large N theories, and the notation $\mathcal{O}\partial^s\mathcal{O}$ makes sense because the dimensions of these operators are very close to the sum of dimensions of the "constituents."

We would like to prove that a similar structure exists in arbitrary $d>2$ CFT, where for asymptotically large spins we find operators whose twists approach $\tau_1+\tau_2$.

This universal property of higher-dimensional CFTs can be seen from crossing symmetry and the

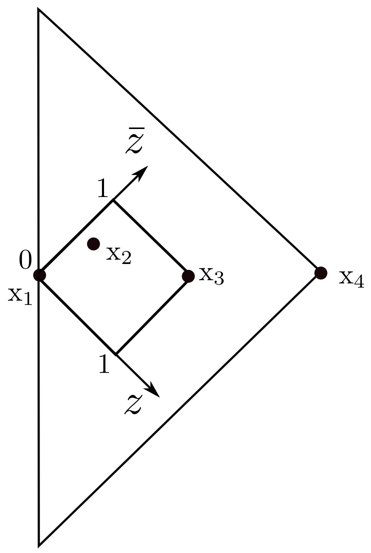


Figure 4.7: We consider all four points to lie on a plane, with the following light cone $(z, \overline{z})$ coordinates: $x_1 = (0, 0)$, $x_2 = (z, \overline{z})$, $x_3 = (1, 1)$, $x_4 = (\infty, \infty)$. We consider the light-cone OPE in the small z fixed $\overline{z}$ channel and then explore its asymptotic as $\overline{z} \to 1$.

presence of a twist gap. By the twist gap we mean the non-zero twist difference between the unit operator that is always present in reflection positive OPEs and any other operator in the theory. This fact follows from unitarity.

To establish this we consider a four-point correlation function in the Lorentzian domain.[12] The main idea is to consider the consequences in the s-channel of the existence of the unit operator in the t-channel. In Euclidean space, the consequences of the unit operator in the dual channel were recently discussed in [106]. It turns out to constrain the high energy asymptotics of the integrated spectral density. Here we would like to exploit the presence of the unit operator in the Lorentzian domain.

For simplicity we consider the case of four identical real scalar operators $\mathcal{O}$ of dimension Δ and then generalize the discussion. The argument in the most general case goes through essentially verbatim.

Consider the four-point function

$$\langle \mathcal{O}(x_1)\mathcal{O}(x_2)\mathcal{O}(x_3)\mathcal{O}(x_4)\rangle = \frac{\mathcal{F}(z,\overline{z})}{(x_{12}^2 x_{34}^2)^{\Delta}} \ , \tag{4.29}$$

$$u = \frac{x_{12}^2 x_{34}^2}{x_{13}^2 x_{24}^2} = z\overline{z} \ , \quad v = \frac{x_{14}^2 x_{23}^2}{x_{13}^2 x_{24}^2} = (1-z)(1-\overline{z}) \ . \tag{4.30}$$

Let us introduce two variables which will parameterize the approach to different light cones $\sigma, \beta \in$

[12]We thank Juan Maldacena for many useful discussions on this topic.

$(0, \infty)$. In this domain all points are space-like separated. The cross ratios are related to σ and β as follows

$$z = e^{-2\beta}, \quad \overline{z} = 1 - e^{-2\sigma}. \tag{4.31}$$

Note that z and $\overline{z}$ are two independent real numbers. The reason for the notation we are using is that in Euclidean signature z is a complex number and $\overline{z}$ is its conjugate. A convenient choice of a four-point function realizing the notation above is depicted in fig. 4.7.

Below we present two versions of the argument. In the first version we exploit crossing symmetry in Minkowski space. We use some basic properties of the light-cone OPE and show how crossing symmetry can lead to certain constraints on the operators spectrum of the theory. We find that crossing symmetry leads to a certain "duality" between fast spinning operators with very large dimensions and the low-lying operators of the theory. Here we compute the first two orders in the large spin expansion that we introduce. The second version of the argument relies on the picture of [110] for fast spinning operators. In this nice, intuitive, picture many facts that we derive from the bootstrap equations have a straightforward interpretation in terms of particles that interact weakly with each other in some gapped theory. We begin by presenting the approach based on the bootstrap equations, and then re-interpret some of the facts we find using the ideas of [110].

4.3.2 An Argument Using the Light-cone OPE

We would like to consider the s-channel light-cone OPE expansion $z \to 0^+$ and $\overline{z}$ fixed. As we have reviewed in the introduction, this expansion is governed by the twists of the operators. More precisely, it is governed by the collinear twist $\Delta - s_{-+}$. Any primary operator can be decomposed into irreducible representations of the collinear conformal group and one can easily see that, for any operator, the minimal collinear twist coincides with the usual conformal twist $\Delta - s$. An operator with collinear twist τ^{coll} contributes $\mathcal{F}(z, \overline{z}) \sim z^{\frac{\tau^{\text{coll}}}{2}} \mathcal{F}(\overline{z})$. The function $\mathcal{F}(\overline{z})$ is a partial wave of the collinear conformal group. We quote some properties of it when needed. As usual, by re-summing the light-cone OPE expansion we can recover the correlation function at any fixed z and $\overline{z}$.

In the notation we defined after (4.29), the light-cone OPE expansion in the s-channel corresponds to a large β, fixed σ, expansion

$$\mathcal{F}(\beta, \sigma) = \sum_i e^{-\tau_i^{coll}\beta} f_i(\sigma) = \int_0^\infty d\tau e^{-\tau^{\text{coll}}\beta} f(\sigma, \tau). \tag{4.32}$$

We assume this OPE to be convergent in the region where all points are spacelike separated, namely $\sigma, \beta \in (0, \infty)$.

The s-channel unit operator gives the most important contribution for finite σ and large β. Let us now consider a different regime. We keep β fixed and start increasing σ. In this way our coordinate $\bar{z}$ approaches 1 and hence gradually becomes light-like separated from the insertion x_3. If we take $\sigma \gg \beta$ the four-point function is dominated by the t-channel OPE, and the most dominant term comes from the unit operator in the t-channel. We can estimate the behavior of (4.32) in this limit:

$$\mathcal{F}(\beta, \sigma) \sim e^{-2\tau_{\mathcal{O}}\beta} e^{2\tau_{\mathcal{O}}\sigma} \left(1 + O(e^{-\sigma})\right) . \tag{4.33}$$

Operators other than the unit operator in the t-channel contribute to the exponentially suppressed terms. Here we are using the fact that there is a twist gap in $d > 2$ CFTs.

One can naively conclude that (4.33) implies that the spectral density is non-zero at $\tau^{\text{coll}} = 2\tau_{\mathcal{O}}$. This is, however, not necessarily true. Imagine we have a series of operators with twists $2\tau_{\mathcal{O}} + a_i$ such that $a_i \to 0$ when $i \to \infty$. Then for arbitrary large σ there will be $a_i < e^{-\tau_{\min}\sigma}$ such that their twist will be effectively $2\tau_{\mathcal{O}}$ to the precision we are probing the correlation function. Thus, equation (4.33) only implies that the spectral density $f(\sigma, \tau)$ is non-zero in any neighborhood of $\tau^{\text{coll}} = 2\tau_{\mathcal{O}}$. This shows that there are operators with collinear twist in any neighborhood of $\tau^{\text{coll}} = 2\tau_{\mathcal{O}}$.

The discussion above concerned with the collinear twists. As mentioned in the introduction to this section, we claim a slightly stronger result – that there are operators whose conformal twists are arbitrarily close to $\tau = 2\tau_{\mathcal{O}}$.

Let us explain the motivations for this stronger claim. First of all, as we will review later, this follows from the construction of [110]. Second, if $\tau = 2\tau_{\mathcal{O}}$ had not been the minimal collinear twist in the decomposition of a conformal primary into collinear primaries, then the conformal twist of the corresponding conformal primary would have been $2\tau_{\mathcal{O}} - n$ for some nonzero integer n. That conflicts with unitarity for some choices of $\mathcal{O}$. Finally, a heuristic argument: different collinear primaries that follow from a single conformal primary are obtained by applying ∂_z to the collinear primary with the lowest collinear twist. However, since we are interested in the behavior for small $1 - \bar{z}$ (this is the $e^{2\tau_{\mathcal{O}}\sigma}$ in (4.33)), higher collinear twists are expected to produce the same (or weaker) singularity as they roughly transport $\mathcal{O}$ in the z direction.

We conclude that there must be operators with twists arbitrarily close to $2\tau_{\mathcal{O}}$. This follows from the presence of the unit operator in the t-channel. We can say more about these operators. An easy lemma to prove is that we must have infinitely many operators with twists $2\tau_{\mathcal{O}}$ (or arbitrarily close to it). The argument is simply that conformal blocks behave logarithmically as $\bar{z} \to 1$ and hence any finite number of them cannot reproduce the exponential growth of (4.33).

Let us also comment about the three-point functions of these operators. To fix them we expand the leading term $\frac{z^\Delta}{(1-\bar{z})^\Delta}$ in terms of s-channel collinear conformal blocks. This can be done using the generalized free fields solution of the crossing equations. The $\bar{z} \to 1$ limit is dominated by the large spin operators. To reproduce the leading $\bar{z} \to 1$ asymptotics correctly, the three-point functions $\langle \mathcal{O}_1 \mathcal{O}_2 \left(\mathcal{O}_1 \partial^s \mathcal{O}_2\right) \rangle$ should coincide with the ones of generalized free fields to the leading order in $\frac{1}{s}$.

So far we analyzed the consequences of matching the $\frac{z^\Delta}{(1-\bar{z})^\Delta}$ piece from the unit operator contribution to the s-channel OPE. However, we can consider the complete dependence on z coming from the unit operator in the t-channel, $\frac{z^\Delta}{(1-z)^\Delta (1-\bar{z})^\Delta}$. Clearly, the primary operators discussed so far are not enough. Again the theory of generalized free fields reproduces the necessary function by means of additional primary operators of the form $\mathcal{O} \partial^s \Box^n \mathcal{O}$. By taking the large spin limit of this solution (equivalently, $\bar{z} \to 1$) we conclude that operators with twists $2\tau_{\mathcal{O}} + 2n$ and three-point functions approaching those of generalized free fields are always present in the spectrum, of any CFT.

All these facts naturally combine into the following picture (that we will henceforth use): CFTs are free at large spin – they are given by generalized free fields. More precisely, in the OPE of two arbitrary operators $\mathcal{O}_1(x)$ and $\mathcal{O}_2(x)$, at large enough spin, there are operators whose conformal twists are arbitrarily close to $\tau_1 + \tau_2$. We denote these operators symbolically by $\mathcal{O}_1 \partial^s \mathcal{O}_2$. Their twists behave as $\tau_1 + \tau_2 + O(\frac{1}{s^\alpha})$ where the power correction will be discussed in the next subsectiontion (α is some positive number). These operators, $\mathcal{O}_1 \partial^s \mathcal{O}_2$, are referred to as *double-twist operators.* Their three-point functions asymptote the ones of the theory of generalized free fields. Similar remarks apply to $\mathcal{O} \partial^s \Box^n \mathcal{O}$.

In principle, one can have either of the following two scenarios:

- a) Operators with the twist $\tau = \tau_1 + \tau_2$ are present in the spectrum.
- b) The point $\tau = \tau_1 + \tau_2$ is a limiting point of the spectrum at infinite spin.

By option b) we mean that in a general CFT there is a set of operators with large enough spin such that their twist is $\tau_1 + \tau_2 - O(\frac{1}{s^\alpha})$ with $\alpha > 0$. In this way, for arbitrary $\epsilon > 0$, there will be operators X with twist $\tau_1 + \tau_2 - \tau_X < \epsilon$. They will be responsible for the behavior (4.33).[13]

Strictly speaking, so far we only discussed scalar external operators. However, the generalization to operators with spin is trivial in this case. Indeed, the usual complication of having many different

[13] At one loop this was observed, for example, in [117],[118]. Here we see this is a general property of any CFT above two dimensions.

structures in three point functions for operators with spin is irrelevant for us. The reason is that our problem is effectively two dimensional, and by considering, for example, operators with all the indices along the z direction $\langle \mathcal{O}_{z\dots z}\widetilde{\mathcal{O}}_{z\dots z}\widetilde{\mathcal{O}}_{z\dots z}\mathcal{O}_{z\dots z}\rangle$ the argument above goes through.

Indeed, in the case of four identical real operators, by keeping only unit operators in the s- and t-channel we get

$$\begin{aligned}\langle \mathcal{O}_{z\dots z}\mathcal{O}_{z\dots z}\mathcal{O}_{z\dots z}\mathcal{O}_{z\dots z}\rangle &= \frac{z^{2s}}{(z\overline{z})^{\Delta+s}} + \frac{(1-z)^{2s}}{[(1-z)(1-\overline{z})]^{\Delta+s}} \\ &= \frac{z^{2s}}{(z\overline{z})^{\Delta+s}}\left[1+\left(\frac{z}{1-z}\right)^{\Delta-s}\left(\frac{\overline{z}}{1-\overline{z}}\right)^{\Delta+s}\right].\end{aligned} \tag{4.34}$$

Again we see that $z^{\Delta-s}$ corresponds to the $\tau^{\rm coll} = 2\tau_{\mathcal{O}} = 2(\Delta - s)$ operators in the s-channel. The crucial simplification is that our perturbation theory is effectively 2d and, thus, all the usual complications of three-point functions of operators with spin are absent.

Notice that our conclusion are obviously incorrect in $d = 2$. There are many $2d$ theories where the double-twist operators are absent. As we have mentioned above, this is due to the fact that in two dimensions there is no twist gap above the unit operator.

This is also related to why some theories in $2d$ can be very simple: the absence of the twist gap provides a way out from this additivity property. Then we can have a very simple twist spectrum. For example, in minimal models the twist spectrum takes the form $\tau_i + n$, where n is an integer that corresponds to the presence of the Virasoro descendants, and τ_i is some finite collection of real numbers. Such a spectrum is generally inconsistent in higher dimensions.

As a simple illustration of how this result is evaded in two dimensions, we can consider the four-point function of spin fields in the 2d Ising model [103]. In this case the only Virasoro primaries have quantum numbers $(0, 0)$, $(\frac{1}{2}, \frac{1}{2})$, and $(\frac{1}{16}, \frac{1}{16})$ for the spin field. One can check explicitly that the leading piece in the $\overline{z} \to 1$ limit has the the leading twist contribution $z^{\frac{\tau}{2}}$ with $\tau = 0$, consistently with the known spectrum. The spectral density around $2\tau_\sigma$ vanishes.

4.3.3 Relation to Convexity

The discussion above allows us to put an upper bound on the minimal twist of the large spin operators appearing in the OPE of any operator with its Hermitian conjugate in $d > 2$ CFTs. The upper bound is, thus, $2\tau_{\mathcal{O}}$. This leads to a compelling picture of the spectrum of general CFTs if we combine this observation with the main claims of section 2, (4.24),(4.26). τ_s^* is thus a nondecreasing convex function that asymptotes to a finite number, not exceeding $2\tau_{\mathcal{O}}$. See fig. 4.8. In addition,

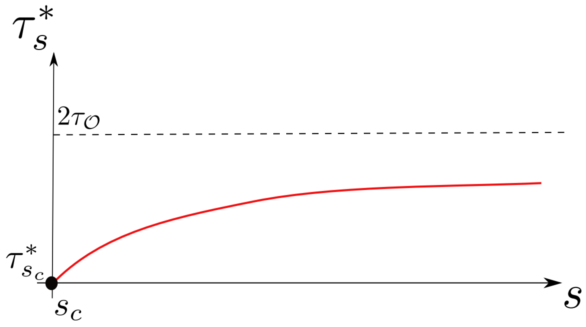


Figure 4.8: The general form of the minimal twist spectrum in an arbitrary CFT in $d > 2$.

as we emphasized above, there must be some operators populating the region of twists around $2\tau_{\mathcal{O}}$. These may be operators of non-minimal twist in the OPE.

An interesting direct application of these results is for the Euclidean bootstrap algorithm of [104]. If we assume that $s_c = 2$ (which is consistent with all the examples we know), then we see that *imposing the gap condition for operators with $s > 2$, $\tau_s^* > \tau_1 + \tau_2$ should not have any solutions (in unitary conformal field theories)* .

4.3.4 The Leading Finite Spin Correction

Let us, for simplicity, focus on the case of four scalar operators of the type $\langle \mathcal{O}\mathcal{O}^\dagger\mathcal{O}\mathcal{O}^\dagger \rangle$ where the dimension of $\mathcal{O}$ is Δ. (The case of non-identical scalar operators is considered in appendix B.) The crossing equation in this case takes the form

$$\mathcal{F}(z,\bar{z}) = \left(\frac{z\bar{z}}{(1-z)(1-\bar{z})}\right)^{\Delta} \mathcal{F}(1-z,1-\bar{z}) \ . \tag{4.35}$$

So far we have focused on the limit where we probe the t-channel OPE, $\bar{z} \to 1$ with z finite. We have taken into account the contribution of the unit operator in this limit. Let us now consider the next operator that contributes in this limit. This operator has the smallest twist among the operators appearing in the OPE (excluding the unit operator). We denote this smallest twist by $\tau_{\min}$. We denote the dimension and spin of this operator by Δ_{min}, $s_{\min}$ such that $\Delta_{\min} - s_{\min} = \tau_{\min}$. The twist gap $\tau_{\min} > 0$ in theories that live in more than two space-time dimensions guarantees that the contribution of such operators in (4.33) is exponentially suppressed by $e^{-\sigma\tau_{\min}}$ compared to the unit operator contribution.

It would be very important for our purpose to determine the contribution of this operator in more detail. For this we need the conformal block corresponding to this operator. The general conformal block entering the s-channel is denoted by $g_{\Delta,s}(z,\bar{z})$. Since in this case we are doing it for the t-channel we need to evaluate it with the arguments $g_{\Delta,s}(1-z,1-\overline{z})$.

One useful property of conformal blocks that we need is their behavior near the light-cone. In the limit $\overline{z}\to 0$ with fixed z the conformal block is given by $\overline{z}^{\frac{\tau}{2}}\mathcal{F}(z)$, where $\mathcal{F}$ is simply related to the hypergeometric function ${}_2F_1$. This is true in any number of dimensions. We can consider $\mathcal{F}(z)$ as z approaches 1. We find the following leading asymptotic form [51]

$$g_{\Delta,s}(z,\overline{z})\to -\frac{\Gamma(\tau+2s)}{(-2)^s\Gamma\left(\frac{\tau+2s}{2}\right)^2}\overline{z}^{\frac{\tau}{2}}\left(\log\left(1-z\right)+\mathcal{O}(1)\right). \tag{4.36}$$

In (4.36) we have taken the limit $\overline{z}\to 0$ first and subsequently we let z approach 1.

Combining all the factors we arrive at the following result for (4.32) in the limit $\sigma\gg\beta$

$$\mathcal{F}(z,\overline{z})=\left(\frac{z\overline{z}}{(1-z)(1-\overline{z})}\right)^{\Delta_{\mathcal{O}}}\left(1-\widetilde{f}^2\frac{\Gamma(\tau_{\rm min}+2s_{\rm min})}{(-2)^{s_{\rm min}}\Gamma\left(\frac{\tau_{\rm min}+2s_{\rm min}}{2}\right)^2}\left(1-\overline{z}\right)^{\frac{\tau_{\rm min}}{2}}\log\left(z\right)\right)+\cdots\,. \tag{4.37}$$

Here $\cdots$ stand for operators with higher twist than $\tau_{\rm min}$ and for various subleading contributions from the $\mathcal{O}_{min}$ conformal block. The contributions from operators with higher twist than $\tau_{\rm min}$ are further suppressed by powers of $1-\overline{z}$ (this translates to exponential suppression in σ in the limit $\sigma\gg 1$ and β fixed). f^2 is the usual coefficient appearing in the conformal block decomposition. It is fixed by the three point function $\langle\mathcal{O}_{min}\mathcal{O}\mathcal{O}^\dagger\rangle$ and the two-point functions of these operators.

We rewrite (4.37) in terms of β and σ. We get

$$\mathcal{F}(\beta,\sigma)=\left(\frac{e^{-2\beta}(1-e^{-2\sigma})}{(1-e^{-2\beta})e^{-2\sigma}}\right)^{\Delta_{\mathcal{O}}}\left(1+2\beta\widetilde{f}^2\frac{\Gamma(\tau_{\rm min}+2s_{\rm min})}{(-2)^{s_{\rm min}}\Gamma\left(\frac{\tau_{\rm min}+2s_{\rm min}}{2}\right)^2}e^{-\tau_{\rm min}\sigma}\right)+\cdots\,. \tag{4.38}$$

Now we need to interpret this result in the s-channel. Consider first the case of unit operator. Its contribution can be accounted for in terms of the s-channel conformal blocks of operators that have twist 2Δ and large spin (as in the theory of generalized free fields). In fact, as we show in appendix B, by doing a saddle point analysis one can see that the leading contribution comes from operators with spin

$$\log s_{\rm dom}=\sigma+\log\frac{(2\Delta-1)^{\frac{3}{2}}(2\Delta+1)^{\frac{1}{2}}}{4\Delta}+O(e^{-\sigma})\,. \tag{4.39}$$

The second derivative at the saddle point is parametrically large only when $\Delta\to\infty$. (In this case the formula simplifies to $\log s_{\rm dom}=\sigma+\log\Delta+O(e^{-\sigma})$.) Otherwise, the saddle point is not localized and its precise location is not very meaningful. It is however meaningful that when σ becomes large the dominant spins are large approximately as (4.39) dictates, $s_{\rm dom}\sim e^{\sigma}$. Below we will "re-sum" all the contributions around the saddle point in order to get various detailed predictions.

Thus, we are building a perturbation theory around the point $s = \infty$, which is dominated by the unit operator in the t-channel and double-twist operators in the s-channel. Note that this perturbation theory is very different from the one in [47], where the expansion is for large N theories. The expansion we are considering is valid in any CFT in $d > 2$.

In (4.38) we have included the contribution of the first leading operator in the t-channel, after the unit operator. Unitarity guarantees that $\tau_{\text{min}} < 2\Delta$. Notice that $\tau_{\text{min}} \leq d-2$ (the upper bound comes from the energy momentum tensor, which is always present), while $\Delta \geq \frac{d-2}{2}$. Thus, equality is only possible for free fields. This is important because it means that this contribution to $\mathcal{F}(\beta,\sigma)$ from τ_{min} is rising exponentially. Hence, it cannot be accounted for by adding finitely many operators in the s-channel, or changing finitely many three-point functions. The saddle point (4.39) suggests that this contribution from τ_{min} should be accounted for by introducing small corrections to operators with very large spins.

Taking into account the exponentially suppressed correction proportional to $e^{-\tau_{\text{min}}\sigma}$ in (4.38), the twists of the operators propagating in the s-channel are now modified. The modification is very small for large enough spin because of the relation (4.39) between the spin and σ. Treating this term proportional to $e^{-\tau_{\text{min}}\sigma}$ as a perturbation, to leading order we can thus replace $e^{-\tau_{\text{min}}\sigma}$ by $1/s^{\tau_{\text{min}}}$, where s is the spin. Then we can interpret this perturbation as a correction to the twist, $\delta\tau_s$, of such high spin operators by matching the dependence on β.[14] We find that for large s

$$\delta\tau_s = -\frac{c_{\tau_{\text{min}}}}{s^{\tau_{\text{min}}}} + \mathcal{O}\left(\frac{1}{s^{\tau_{\text{min}}+\epsilon}}\right) , \tag{4.40}$$

where c is some constant independent of the spin and $\epsilon > 0$. If these double-twist operators are also the minimal twist operators in the OPE (as we will discuss in some examples), then we know from monotonicity of the twist that $c_{\tau_{\text{min}}} > 0$. We see that the limiting twist $2\tau_{\mathcal{O}}$ is approached in a manner that is completely fixed by the operator with the lowest lying nontrivial twist in the problem. For example, it could be some low dimension scalar operator, or the energy momentum tensor (for which $\tau_{\text{min}} = d-2$).

It is natural to ask what can be said about $c_{\tau_{\text{min}}}$. Here we compute it for the case of external scalar operators of dimension Δ. As explained in appendix B, the relevant contribution from the

[14]In other words, we use $e^{-\beta\tau-\beta\delta\tau} = e^{-\beta\tau}(1-\beta\delta\tau+\cdots)$.

s-channel takes the form[15]

$$-\frac{1}{2}\log z\ z^{\Delta} \times \lim_{\overline{z}\to 1}\sum_{s=\Lambda}^{\infty}\frac{c_{\tau_{\min}}}{s^{\tau_{\min}}}c_s\overline{z}^{\Delta+s}\ {}_2F_1(\Delta+s,\Delta+s,2s+2\Delta,\overline{z})\ . \tag{4.41}$$

First we switch from the sum to the integral $\sum_{s=0}^{\infty} \to \int ds$ and also make a change of variables motivated by the saddle point analysis of appendix B, $s \to \frac{s}{\sqrt{1-\overline{z}}}$. Denoting $\epsilon = 1-\overline{z}$ and using for the $\epsilon \to 0$ limit

$${}_2F_1(\frac{h}{\sqrt{\epsilon}},\frac{h}{\sqrt{\epsilon}},2\frac{h}{\sqrt{\epsilon}},1-\epsilon) \sim 4^{\frac{h}{\sqrt{\epsilon}}}\frac{\sqrt{h}K_0(2h)}{\sqrt{\pi}\epsilon^{\frac{1}{4}}}\ , \tag{4.42}$$

(which can be derived using the integral representation for the hypergeometric function) we get (in the small ϵ limit) that (4.41) is equal to

$$\begin{aligned}&\frac{c_{\tau_{\min}}}{(1-\overline{z})^{\Delta-\frac{\tau_{\min}}{2}}}\frac{4}{\Gamma(\Delta)^2}\int_0^{\infty} ds\ s^{2\Delta-\tau_{\min}-1}K_0(2s)\\ &=\frac{c_{\tau_{\min}}}{(1-\overline{z})^{\Delta-\frac{\tau_{\min}}{2}}}\frac{\Gamma(\Delta-\frac{\tau_{\min}}{2})^2}{\Gamma(\Delta)^2}\ .\end{aligned} \tag{4.43}$$

Thus, the crossing equation, i.e. the requirement that (4.41) is equal to (4.37), takes the form

$$c_{\tau_{\min}}\frac{\Gamma(\Delta-\frac{\tau_{\min}}{2})^2}{\Gamma(\Delta)^2} = 2\widetilde{f}^2\frac{\Gamma(\tau_{\min}+2s_{\min})}{(-2)^{s_{\min}}\Gamma\left(\frac{\tau_{\min}+2s_{\min}}{2}\right)^2}\ , \tag{4.44}$$

which can be solved for $c_{\tau_{\min}}$ to get

$$\begin{aligned}c_{\tau_{\min}} &= \frac{\Gamma(\tau_{\min}+2s_{\min})}{2^{s_{\min}-1}\Gamma\left(\frac{\tau_{\min}+2s_{\min}}{2}\right)^2}\frac{\Gamma(\Delta)^2}{\Gamma(\Delta-\frac{\tau_{\min}}{2})^2}f^2,\\ f^2 &= \frac{C^2_{\mathcal{O}\mathcal{O}^\dagger\mathcal{O}_{\tau_{\min}}}}{C_{\mathcal{O}\mathcal{O}}C_{\mathcal{O}^\dagger\mathcal{O}^\dagger}C_{\mathcal{O}_{\tau_{\min}}\mathcal{O}_{\tau_{\min}}}} = (-1)^{s_{\min}}\widetilde{f}^2\ .\end{aligned} \tag{4.45}$$

The additional factor $(-1)^{s_{\min}}$ in the second line of (4.45) comes about because $C_{\mathcal{O}\mathcal{O}^\dagger\mathcal{O}_{\tau_{\min}}}$ differs from $C_{\mathcal{O}^\dagger\mathcal{O}\mathcal{O}_{\tau_{\min}}}$ by a minus sign for odd spins. This will be important below.

For convenience, above we have included the expression connecting f^2 with various two- and three-point functions (our conventions for two- and three-point functions are summarized in appendix A). Thus, we have computed the coefficient that controls the leading correction to the anomalous dimension of double-twist operators in any CFT $\tau_{\mathcal{O}\partial^s\mathcal{O}^\dagger} = 2\tau_{\mathcal{O}} - \frac{c_{\tau_{\min}}}{s^{\tau_{\min}}} + \cdots$. Let us emphasize

[15]For the small spin cut-off we chose Λ. Since the sum is dominated by large spin operators in the limit we are considering, nothing will depend on Λ in an important way.

several features of this formula:

- - It is manifestly invariant under redefinitions of the normalizations of operators;
- - The formula is valid for any $d > 2$ CFT and does not depend on the existence of any perturbative expansion parameters. Thus, it is an exact result.
- - The sign of all the corrections is positive, both for odd and even spins. In the next section we will consider more general situations where different contributions come with different signs. Here, since all the contributions are positive, the spectrum of twists of $\mathcal{O}\partial^s\mathcal{O}^\dagger$ is clearly convex for large enough spin, in any CFT.
- - If the minimal twist operator $\mathcal{O}_{min}$ is not unique, we just have to sum the different contributions $c = \sum c_i$;
- - From unitarity we know that $\tau_{\text{min}} < 2\Delta$, and also $\tau_{\text{min}} \leq d-2$. In section 5 we will see an example where all the contributions with $\tau_{\text{min}} \leq 2\Delta$ cancel out. Then the leading contribution to the asymptotic anomalous twist will come from $\tau_{\text{min}} > 2\Delta$, and our formula reproduces it correctly.
- - The formula should be modified to incorporate the case when the external operators are not scalars. We expect that this should be manageable because the problem is effectively two dimensional and all conformal blocks are known [119]. We do not pursue this here.

This formula has several interesting consequences and applications that we study in the next section. The following subsectiontion is dedicated to a brief review of the picture of [110] and its connection to the discussion above.

4.3.5 An Argument in the Spirit of Alday-Maldacena

The general results discussed in the previous subsectiontions can be naturally understood at the qualitative level using the ideas of [110]. Such ideas were also applied in [120].

One can think about the four-point correlation function in a d-dimensional CFT as a four-point function in a two-dimensional gapped theory. Denote the coordinates on this two-dimensional space by (u, v). The gap in the spectrum is the twist gap of the underlying d-dimensional CFT. More generally, the evolution in the u direction is dictated by the twist of the underlying d-dimensional CFT. The operators could be thought as being inserted at $(\pm u_0, \pm v_0)$.

When operators are largely separated in the v direction the interaction is weak and corrections to the free propagation $e^{-2\tau_{\mathcal{O}} u_0}$ are small. The corrections are governed by the separation v_0. Since the theory is gapped, the correction to the energy due to the exchange of a particle of "mass" τ_{exch} takes the form $e^{-\tau_{\text{exch}} v_0}$, which could be made arbitrarily small by taking large enough v_0.

Since ∂_u measures the twist we see that in this limit of large v_0 we have a propagating state with the twist arbitrarily close to $2\tau_{\mathcal{O}}$. By the state-operator correspondence we have to identify it with some operators in the theory. This is the additivity property alluded to above.

Moreover, [110] identified how the v_0 separation is related to the spin of the operators. They suggested the relation $\log s_{\text{dom}} \sim v_0$. That allows to interpret the correction potential energy due to the leading interaction, $e^{-\tau_{\text{exch}} v_0}$, in terms of a correction to the twist of high spin operators, $\delta\tau \sim \frac{c}{s^{\tau_{exch}}}$. For large v_0 we the most important exchanged particle is the one with the smallest twist, hence τ_{exch} is identified with $\tau_{\min}$. This is the statement of (4.40). The fact that at large spin we approach the generalized free field picture corresponds to the fact that at large separations excitations do not interact and propagate freely. (In other words, locality in the v coordinate.)

We see that the free propagation in the Alday-Maldacena picture corresponds to the unit operator dominance in the OPE picture. Corrections due to the exchange of massive particles in the Alday-Maldacena picture correspond to the inclusion of the next terms in the light-cone OPE. Also in this picture the state-operator correspondence plays an important role.

The relation $\log s_{\text{dom}} \sim v_0$ of [110] is qualitatively correct, but there are very important corrections to it that we could determine precisely in our formalism (4.39). (And as we explained, the saddle point itself is wide unless $\Delta \to \infty$, so one needs to re-sum the corrections around it. This is what we have essentially done in (4.42),(4.43).) Those corrections play a central role in the determination of the precise shift in the twists of fast spinning operators. In the picture of [110] those corrections should result from taking into account the finite width of the wave function in the v coordinate. This would be interesting to understand better, perhaps along the lines of [120].

4.4 Examples and Applications of (3.17)

In this section we study the correction formula (4.45) in different regimes and circumstances, comparing it to known results when possible.

4.4.1 Correction due to the Stress Tensor and Theories with Gravity Duals

It is well known that in the case of the stress tensor the coupling of it to other operators is universal and depends only on the dimension of the operator and two-point function of the stress tensor. More precisely, in (4.45) $f^2 = \frac{d^2\Delta^2}{(d-1)^2 c_T^2}$ so that we get

$$c_{stress} = \frac{d^2\Gamma(d+2)}{2c_T(d-1)^2\Gamma(\frac{d+2}{2})^2} \frac{\Delta^2\Gamma(\Delta)^2}{\Gamma(\Delta - \frac{d-2}{2})^2}. \tag{4.46}$$

Let us consider a CFT where in the OPE of certain operators the minimal twist (after the unit operator) is realized by the stress tensor. Simple examples of such CFTs arise in the context of holography. In this case the two-point function is given by (see for example [121])

$$c_T = \frac{d+1}{d-1} \frac{L^{d-1}}{2\pi G_N^{(d+1)}} \frac{\Gamma(d+1)\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})^3} , \tag{4.47}$$

so that we get for the leading correction

$$c_{stress} = \frac{4\Gamma(\frac{d}{2})}{\pi^{\frac{d-2}{2}}(d-1)} \frac{G_N^{(d+1)}}{L^{d-1}} \frac{\Gamma(\Delta+1)^2}{\Gamma(\Delta - \frac{d-2}{2})^2} . \tag{4.48}$$

We can apply this formula, for example, in the case of $\mathcal{N}=4$ SYM theory. We substitute $d=4$ and $\frac{G_N^{(d+1)}}{L^{d-1}} = \frac{\pi}{2N^2}$ so that $c_T = 40N^2$ and we get that the correction due to the stress tensor exchange is given by

$$c_{stress}^{\mathcal{N}=4} = \frac{2\Delta^2(\Delta-1)^2}{3N^2} . \tag{4.49}$$

One of the first computations of a four-point function at strong coupling was 2-2 dilaton scattering. in this case the anomalous dimensions of double trace operators are known [122]

$$\tau_{:\mathcal{O}\partial^s\mathcal{O}:} = 8 - \frac{96}{(s+1)(s+6)} \frac{1}{N^2} . \tag{4.50}$$

In the formula above $\mathcal{O}$ stands for the operator dual to the dilaton. The only twist 2 operator in the t-channel is the stress tensor. By plugging $\Delta = 4$ into our formula we correctly reproduce $\frac{96}{N^2}$!

For other computations of the anomalous dimensions of double-trace operators in theories with gravity duals see [123, 124]. In all those cases we found that the sign of the correction is consistent with convexity, and the leading power of s is as predicted. However, often more than one operator of twist 2 is exchanged in the t-channel and (4.49) will not be the complete answer. To get the complete answer we just have to sum over finitely many contributions.

4.4.2 The large Δ Limit

It is curious to consider large Δ limit of (4.45) . The factor $\frac{\Gamma(\Delta-\frac{\tau_{\rm min}}{2})^2}{\Gamma(\Delta)^2}$ becomes $\Delta^{\tau_{\rm min}}$ and, thus, $c_{\tau_{\rm min}}$ has a simple dependence on Δ, $c_{\tau_{\rm min}} \sim \Delta^{\tau_{\rm min}} f^2$ where f^2 is the combination of three- and two-point functions written in (4.45). The limit of large Δ is also nice in that the saddle point described in appendix B becomes exact and simplifies as described after (4.39).

The point we would like to make in this subsectiontion is that since for large Δ the saddle point (4.39) becomes exact, the idea of [110], which we outlined in subsectiontion 3.5, can be made precise.

According to the picture we reviewed in subsectiontion 3.5, the scaling dimensions of the double-twist operators are related to the total energy of two static charges in some two-dimensional space interacting through a Yukawa-like potential

$$\Delta_1 + \Delta_2 + g q_1 q_2 e^{-m v_0} \ , \tag{4.51}$$

where v_0 is the separation between the particles. We assume for simplicity $\Delta_1 = \Delta_2 \equiv \Delta$. Now we use the fact that the saddle point is exact if Δ is large and the relation at the saddle point is $v_0 = \log \frac{s}{\Delta}$. Additionally, from our formula for the leading correction to the dimensions of double-twist operators we read $g q_1 q_2 \sim \frac{C^2_{\mathcal{O}\mathcal{O}^\dagger \mathcal{O}_{\tau_{\rm min}}}}{C_{\mathcal{O}\mathcal{O}} C_{\mathcal{O}^\dagger\mathcal{O}^\dagger} C_{\mathcal{O}_{\tau_{\rm min}}\mathcal{O}_{\tau_{\rm min}}}}$. The mass m appearing in (4.51) corresponds to the minimal twist that we exchange (this is directly related to the fact that the lightest particle that we exchange leads to the longest range interactions).

These identifications are natural, for example, the sign of the correction has a very simple interpretation. Gravity is always an attractive force, thus, the correction from the potential energy is always negative. This is reflected by the fact that for the energy momentum tensor we have $g \sim -\frac{1}{c_T}$ and $q \sim \Delta$. However, if we have a $U(1)$ current the force could be either attractive or repulsive depending on the signs of the charges.

The identification $v_0 = \log \frac{s}{\Delta}$ and the relation between $g q_1 q_2$ and various two- and three-point functions all receive nontrivial corrections in Δ^{-1}. We have accounted for all of them precisely in the previous section.

4.4.3 The Case of Parametrically Small $\Delta - \frac{\tau_{\rm min}}{2}$

The formula (4.45) goes to zero when $\Delta = \frac{\tau_{\rm min}}{2}$. This is consistent because such an equality could only be realized in free field theory (or, if one abandons unitarity it could be realized in models such as generalized free fields). And indeed, in free field theory (or generalized free fields) the answer is zero; there are no corrections to the twists of double-twist operators.

However, the case of small $\Delta - \frac{\tau_{\rm min}}{2}$ is in fact more subtle. For example, consider some weakly coupled CFT. Then if we choose the external state appropriately, the leading twist operators in the OPE would have $\tau_{\rm min}$ very close to 2Δ. But there would be infinitely many such operators. In order to get the right answer for the anomalous dimensions of fast spinning operators we would generally have to sum them all and be careful about performing the perturbation theory consistently.

An example is the ϕ^4 theory in $4-\epsilon$ dimensions. We consider the $\phi(x)\phi(0)$ OPE, which includes many operators whose twists are very close to 2. In order to obtain the right anomalous twists of fast spinning operators of the type $\phi\partial^s\phi$ as discussed in section 3, we have to sum all of those operators up. In this case ϵ breaks a higher spin symmetry and so the infinitely many operators with twists close to 2 can be regarded as due to a slightly broken higher spin symmetry.

More generally, imagine we have a CFT which contains a small parameter ϵ. Imagine we also know the twist $\tau_{\mathcal{O}\partial^s\mathcal{O}}(s,\epsilon)$ of some double-twist operators exactly as a function of Δ and s. From the discussion above we know that at *sufficiently large spin* we expect it to have the form

$$\tau_{\mathcal{O}\partial^s\mathcal{O}}(s,\epsilon) = 2\Delta - \frac{c_2(\epsilon)}{s^{d-2}} - \frac{c_4(\epsilon)}{s^{d-2+\gamma_4(\epsilon)}} - ... \tag{4.52}$$

where $\gamma_4(\epsilon)$ is the anomalous dimension of the minimal twist spin-four operator. In the above we have assumed that the minimal twists are realized by the energy-momentum tensor, spin-four operators etc. Of course it could also be that $\mathcal{O}^2$ has a twist smaller than $d-2$ (and hence dominates over the energy momentum tensor) but that won't change the point we would like to make. Clearly, if ϵ is small, the expansion (4.52) is only useful for spins much larger than any other parameter in the theory (including $e^{\frac{1}{\gamma_4(\epsilon)}}$ that would naturally arise in a situation like (4.52)).

If one decides to fix the spin and take ϵ arbitrarily small, then all the operators in (4.52) become important. In this regime, the problem we are discussing is essentially that of solving perturbative CFTs via the bootstrap program. This is not our goal here, and we will not comment on it further. The interested reader should consult [47, 125] for interesting recent progress in this direction.

In any given CFT, the results (4.40),(4.45) always apply for sufficiently large spin. For example, in the Ising model

$$\tau_{\sigma\partial^s\sigma}^{3d\ Ising} \sim 1.037 - \frac{0.0028}{s} + ... \tag{4.53}$$

where we have used here the numerical value for the central charge found in [105]. This is an analytic result about the Ising model, but in practice its power is limited because one has to go to quite large spin s in order for the next terms to be suppressed (the next term is controlled by the anomalous twist of the spin 4 operator, which is 0.02, and thus the formula (4.53) is useful only for exponentially large spins). In the next section we will discuss some results about the Ising model that are more powerful from the practical point of view.

4.4.4 The Charge to Mass Ratio is Related to Convexity

As we explained above, the correction $c_{\tau_{\min}}$ in (4.40) could have either sign depending on which operators propagates in the t-channel and depending on the precise form of the four-point function we study. We demonstrate this point here by considering four point functions of the form $\langle \mathcal{O}_i(x_1)\mathcal{O}_j(x_2)\mathcal{O}^\dagger_{\bar{i}}(x_3)\mathcal{O}^\dagger_{\bar{j}}(x_4)\rangle$ where $\mathcal{O}_i$ and $\mathcal{O}_{\bar{i}}$ are operators charged under global symmetries. In the expansion in the s-channel we necessarily encounter double-twist operators of the type $\mathcal{O}\partial^s\mathcal{O}$ which are charged under some global symmetries. We cannot directly apply the results of section 2 for such operators since we do not have positivity of the cross section argument. It is thus interesting to ask whether such operators approach the limiting twist 2Δ in a convex, flat, or concave manner. Here we will explain that under some very general circumstances this is, roughly speaking, fixed by the charge to dimension ratio of $\mathcal{O}$. Let us set up the framework more precisely.

Imagine we have N_f flavor currents J^μ_K. And imagine also we have a set of scalar operators $\mathcal{O}_i$ charged under this symmetry so that

$$[Q, \mathcal{O}_i] = -(T_K)^j_i \mathcal{O}_j \tag{4.54}$$

This means that the OPE includes

$$J^\mu_K(x)\mathcal{O}_i(0) \sim \frac{-i}{S_d}\frac{x^\mu}{x^d}(T_K)^j_i\mathcal{O}_j(0). \tag{4.55}$$

where S_d is the surface area of a $d-1$ sphere.

Let us define the two-point functions

$$\langle \mathcal{O}_i(x)\mathcal{O}^\dagger_{\bar{i}}(0)\rangle = \frac{g_{i\bar{i}}}{x^{2\Delta}} \ , \ \langle J^\mu_I(x)J^\mu_J(0)\rangle \qquad = \frac{\tau_{IJ}}{S_d^2}\frac{I^{\mu\nu}}{x^{2(d-1)}}. \tag{4.56}$$

Using these and the OPE (4.55) we can fix the three-point functions to be

$$\langle \mathcal{O}_i(x_1)\mathcal{O}^\dagger_{\bar{j}}(x_2)J^\mu_I(x_3)\rangle = \frac{-i}{S_d}(T_I)^j_i g_{j\bar{j}}\frac{Z^\mu}{x_{12}^{2\Delta-(d-2)}x_{13}^{d-2}x_{23}^{d-2}} \tag{4.57}$$

where $Z^\mu = \frac{x^\mu_{31}}{x^2_{31}} - \frac{x^\mu_{32}}{x^2_{32}}$.

Now we consider the four-point function $\langle \mathcal{O}_i(x_1)\mathcal{O}_j(x_2)\mathcal{O}^\dagger_{\bar{i}}(x_3)\mathcal{O}^\dagger_{\bar{j}}(x_4)\rangle$. Using (4.45) we can account for the contributions of the conserved currents in the t-channel to $\delta\tau_s$. We find

$$(\delta\tau_s)_{\text{currents}} \sim -\sum_{I,J}\frac{C_{\mathcal{O}_i\mathcal{O}^\dagger_{\bar{i}}J^I}C_{\mathcal{O}_j\mathcal{O}^\dagger_{\bar{j}}J^J}}{C_{\mathcal{O}_i\mathcal{O}^\dagger_{\bar{i}}}C_{\mathcal{O}_j\mathcal{O}^\dagger_{\bar{j}}}C_{J^IJ^J}} = -\sum_{I,J}\frac{(T_I)^k_i g_{k\bar{i}}(T_J)^m_j g_{m\bar{j}}}{\tau_{IJ}g_{i\bar{i}}g_{j\bar{j}}}. \tag{4.58}$$

Restoring the pre-factors from (4.45) (for a configuration of four insertions like we consider here there is an additional factor of $(-1)^s$) we get for the sum of the stress-tensor and repulsive flavor corrections (here we have assumed that the minimal twists that appear in the t-channel are the EM tensor and conserved currents)

$$\delta\tau_s \sim -\frac{1}{s^{d-2}}\left[\frac{d^2\Delta^2}{4(d-1)^2 c_T}\frac{\Gamma(d+2)}{\Gamma(\frac{d+2}{2})^2} - \frac{\Gamma(d)}{2\Gamma(\frac{d}{2})^2}\sum_{I,J}\frac{(T_I)^k_i g_{k\bar{i}}(T_J)^m_j g_{m\bar{j}}}{\tau_{IJ}g_{i\bar{i}}g_{j\bar{j}}}\right] . \tag{4.59}$$

In general, the odd spin contributions are always repulsive in the context considered in this subsectiontion and the even spin contributions (including scalar and stress tensor exchange) are attractive. Therefore, whether $\delta\tau_s$ is convex, flat, or concave for large spin is determined by the ratio of Δ/c_T compared to, roughly speaking, q^2/τ_{IJ} where q is the charge of the operators (more generally, the representation). *Therefore, the question of convexity at large spin is determined by whether $\mathcal{O}$ satisfies a BPS-like bound.* The BPS-like bound that controls convexity for the simple case of stress tensor and a single $U(1)$ symmetry is

$$\frac{\Delta}{\sqrt{c_T}} \geq \frac{1}{\sqrt{2}}\frac{d-1}{\sqrt{d(d+1)}}\frac{|q|}{\sqrt{\tau}} \ . \tag{4.60}$$

Let us now mention that if we consider instead the correlation function in the ordering $\langle \mathcal{O}_i(x_1)\mathcal{O}^\dagger_{\bar{i}}(x_2)\mathcal{O}_j(x_3)\mathcal{O}^\dagger_{\bar{j}}$

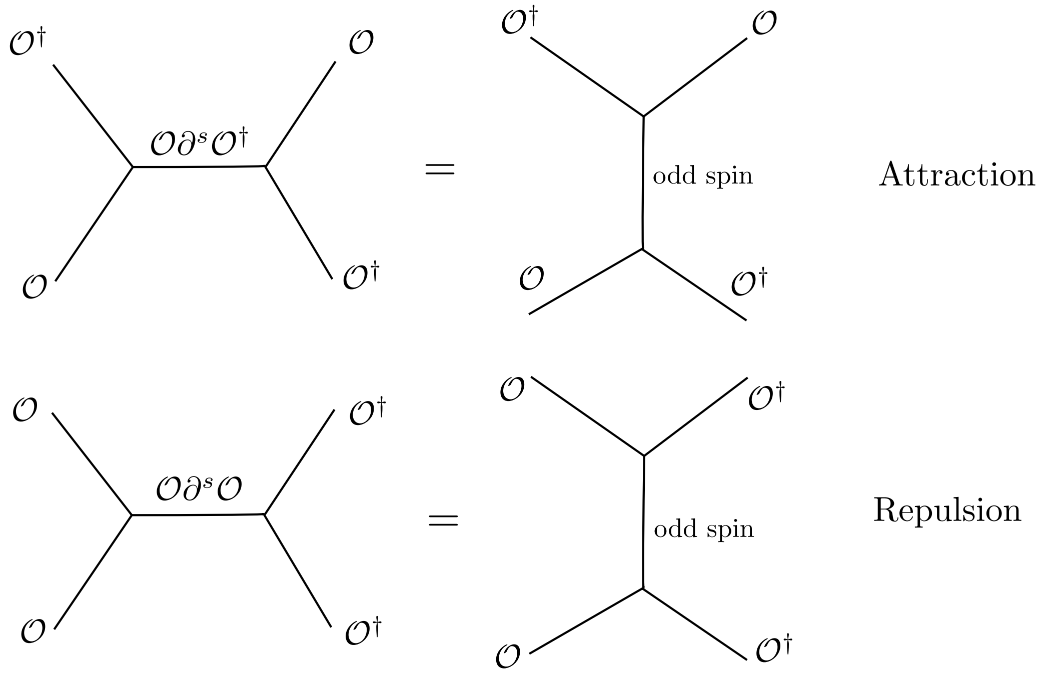


Figure 4.9: The correction due to the $U(1)$ current exchange has the opposite sign for the anomalous twists of $\mathcal{O}\partial^s\mathcal{O}$ and $\mathcal{O}\partial^s\mathcal{O}^\dagger$. At the level of the bootstrap equations, it corresponds to the fact that the three-point function $\langle \mathcal{O}\mathcal{O}^\dagger J_s\rangle$ is odd under a permutation of $\mathcal{O}$ and $\mathcal{O}^\dagger$ when s is odd. (Here J_s is an operator of spin s.)

and study the s-channel expansion, then we encounter double-twist operators of the type $\mathcal{O}\partial^s\mathcal{O}^\dagger$. In this case, the t-channel expansions gives a different results from before. Because we flipped two operators (see fig. 4.9), all the even spins still contribute positively to c but now the odd spins contribute positively as well. Hence, this essentially proves convexity in a new way (without alluding to an RG flow) for large enough spins! In fact, in some sense, it proves a slightly different result than the one in section 2, since here we do not need to assume that the double-twist operators are the minimal twist operators to get convexity – the convexity of section 2 is always about the minimal twist operators.

As an example of the discussion above let us consider an $\mathcal{N}=1$ SCFT in four dimensions. In the case of SCFT we have a $U(1)_R$ global symmetry. The two-point function of the R-current is related via supersymmetry to the two-point function of the stress tensor (see, for example, formula (1.11) in [126] [16]). The relation in our conventions is

$$\frac{c_T}{\tau} = \frac{d(d+1)}{2} \ . \tag{4.61}$$

[16] Our conventions are related to theirs by $\frac{\tau_{here}}{S_d^2} = \frac{2(d-1)(d-2)}{(2\pi)^d}\tau_{there}$ and $\frac{c_{T;here}}{S_d^2} = c_{T;there}$.

Now consider the correlation function $\langle \mathcal{O}(x_1)\mathcal{O}(x_2)\mathcal{O}^\dagger(x_3)\mathcal{O}^\dagger(x_4)\rangle$ where $\mathcal{O}$ is a chiral primary and $\mathcal{O}^\dagger$ is an anti-chiral primary. We can use the formula (4.59) to evaluate the asymptotic corrections to the operators $\mathcal{O}\partial^s\mathcal{O}$ appearing in the chiral×chiral OPE. Using the relation (4.61) and the fact that for chiral primaries $\Delta = \frac{3r}{2}$ where r is the R-symmetry charge of $\mathcal{O}$, we find that the correction (4.59) precisely vanishes. Furthermore, if $\mathcal{O}$ carries some global symmetry quantum numbers other than the R-symmetry, we should include the contributions to $\delta\tau_s$ from the exchanges in the t-channel of the flavor supermultiplet (this is a linear multiplet of $\mathcal{N} = 1$). We find that this contribution again vanishes identically.

One could ask why these corrections to $\delta\tau_s$ vanish identically for chiral primaries. In the OPE of two chiral primaries one finds short representations with twists precisely 2Δ and long representations with twists larger than 2Δ. (There are no smaller twists.) The cancelation above may be related to the presence of infinitely many short representations in the OPE. It would be nice to understand whether the OPE of two chiral primaries contains infinitely many protected operators, as the computations above hint. For completeness, the structure of the OPE of two chiral primaries is reviewed in appendix D.

Another curiosity that we would like to note is that if $\mathcal{O}$ is chiral but not a chiral primary, then $\Delta > \frac{3r}{2}$ and the exchange of the stress tensor dominates over the R-current. Thus, we have convexity again. We see that the general connection between the BPS-like bound (4.60) and convexity is realized very naturally in SUSY.

4.5 Applications of Convexity

In this section we consider different known examples and verify that the convexity and asymptotic form of the twists hold true. Moreover, we will see that in all known examples the convexity starts from spin 2 (thus $s_c = 2$).

4.5.1 Free Theories and Weakly Coupled Theories

It is easy to see that the propositions made in the previous sections are true in free theories. In this case we have an infinite set of conserved currents j_s with the minimal possible twist $\tau = d - 2$.

For a generic operator $\mathcal{O}$ we expect the three-point functions $\langle \mathcal{O}\mathcal{O}^\dagger j_s\rangle$ to be non-zero. This is suggested by the way higher spin symmetry is realized in the sector of conserved currents as well as from the way it acts on the fundamental field.

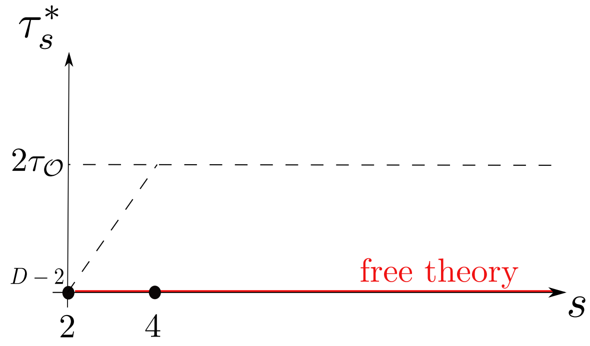


Figure 4.10: The form of the minimal twist spectrum in free theories.

The conserved currents will be the leading twist operators. In this way convexity with $s_c = 2$ is realized in the case of free theories. See fig. 4.10.

It is also trivial to see that there exist operators with twist arbitrarily close to $2\tau_{\mathcal{O}}$. Indeed, for any composite operator $\mathcal{O}$ in a free theory, in the OPE of this operator with itself there will appear the operators $\mathcal{O}\partial^s\mathcal{O}^\dagger$ (written schematically). Those have twist precisely $2\tau_{\mathcal{O}}$. If the operator $\mathcal{O}$ is the elementary free field itself, then these operators are the higher spin currents. Their twist is precisely $d-2$, which coincides with twice the dimension (and hence twice the twist) of the elementary scalar field.

Weakly coupled CFTs inherit some of their OPE structure from their higher spin symmetric parents. (The OPE coefficients and the dimensions change a little. Various short multiplets could combine, for example, the exactly conserved higher-spin currents undergo a "Higgs mechanism" and acquire a small anomalous twist. Also new operators that were not present in the free field theory could appear.) If the CFT is sufficiently weakly coupled, the almost conserved currents will still be the minimal twist operators in the OPE of any operator with itself. This means that in this case our statement from section 2 boils down to the convexity of twists of almost conserved currents. The examples below are of this type.

Notice that in the case when we have a flux (e.g. in gauge theories) the situation is different. The almost conserved currents universally receive a correction to the twist of the form $\log s$. Thus, for large enough spin, their twist could be arbitrary large so that they cease to be the minimal twist operators. In this case some other operators (the double twist operators) ensure additivity and convexity.

4.5.2 The Critical $O(N)$ Models

Let us consider the $O(N)$ critical point in $4-\epsilon$ dimensions. We can use the results of [127] for the anomalous dimensions. We have

$$\delta\tau_{\sigma_i\partial^s\sigma_i} - 2\gamma_\sigma = -\epsilon^2 \frac{N+2}{2(N+8)^2}\frac{6}{s(s+1)} \ , \tag{4.62}$$
$$\delta\tau_{\sigma_{(i}\partial^s\sigma_{j)}} - 2\gamma_\sigma = -\epsilon^2 \frac{N+2}{2(N+8)^2}\frac{6(N+6)}{(N+2)s(s+1)} \ , \tag{4.63}$$

where γ_σ is the anomalous dimension of the spin field. The operators $\sigma_i\partial^s\sigma_i$ are singlets under the global symmetry $O(N)$, and $\sigma_{(i}\partial^s\sigma_{j)}$ are symmetric traceless representations of $O(N)$.

Let us make several comments about the formulae above. First of all, both formulae have a well-defined large $s\to\infty$ limit that implies $\delta\tau_s - 2\gamma_\sigma \to 0$.[17] This means that the prediction (4.27) is saturated, as it *must* in weakly coupled CFTs where we consider operators that are bilinear in the fundamental field. Second, the leading correction around infinite spin in both cases is of the form $\frac{1}{s^2}$. This is what we expect due to the presence of the stress tensor and other currents which are conserved at this order. (To this order also σ^2 has twist $d-2$.) Third, the convexity starts from the energy momentum tensor, $s=2$. At higher orders in ϵ, the higher spin-currents will lead to different typical exponents at large spin, $\frac{1}{s^{\tau_{\min}(\epsilon)}}$. The dominant exponent at very large spin will still be due to the stress tensor, $\frac{1}{s^{2-\epsilon}}$. (The reason for that is that higher spin currents have a larger anomalous twist and the operator σ^2 appears to have dimension, and thus twist, larger than $d-2$ as well. This information about σ^2 is suggested from the epsilon expansion, from the known results about low N critical $O(N)$ models, and the results of large N.)

Alternatively, we can consider the case of large N with the number of space-time dimensions being d. In this case we can use the results of the large N expansion [128],[129]

$$\delta\tau_{\sigma_i\partial^s\sigma_i} = 2\gamma_\sigma \frac{4}{(d+2s-4)(d+2s-2)}\left[(d+s-2)(s-1) - \frac{1}{2}\frac{\Gamma[d+1]\Gamma[s+1]}{2(d-1)\Gamma[d+s-3]}\right] \ , \tag{4.64}$$
$$\delta\tau_{\sigma_{(i}\partial^s\sigma_{j)}} = 2\gamma_\sigma \frac{4(d+s-2)(s-1)}{(d+2s-4)(d+2s-2)} \ , \tag{4.65}$$
$$\gamma_\sigma = \frac{2}{N}\frac{\sin\pi\frac{d}{2}}{\pi}\frac{\Gamma(d-2)}{\Gamma(\frac{d}{2}-2)\Gamma(\frac{d}{2}+1)} \ . \tag{4.66}$$

[17] To leading order $2\gamma_\sigma = \frac{N+2}{2(N+8)^2}\epsilon^2$.

In the large s limit these formulae become

$$\frac{\delta\tau_{\sigma_i\partial^s\sigma_i} - 2\gamma_\sigma}{2\gamma_\sigma} \sim -\frac{\Gamma(d+1)}{2(d-1)}\frac{1}{s^{d-2}} - \frac{d(d-2)}{4}\frac{1}{s^2} \ , \tag{4.67}$$

$$\frac{\delta\tau_{\sigma_{(i}\partial^s\sigma_{j)}} - 2\gamma_\sigma}{2\gamma_\sigma} \sim -\frac{d(d-2)}{4}\frac{1}{s^2} \ . \tag{4.68}$$

These results are consistent with the fact that at leading order in N we have conserved currents operators with the twist $\tau = d-2$, and we also see $1/s^2$ coming from the operator σ^2.

The difference in the coefficients of $\frac{1}{s^{d-2}}$ for the operators $\sigma_i\partial^s\sigma_i$ and $\sigma_{(i}\partial^s\sigma_{j)}$ comes from the fact that different conserved higher spin currents contribute. We see that the sum over all the higher spin currents (each contributing to $1/s^{d-2}$) actually vanishes for the symmetric combination. However, the $\frac{1}{s^2}$ contribution is governed solely by the three point function $\langle\sigma_i\sigma_i\sigma^2\rangle$ and, thus, the coefficient of $1/s^2$ is the same for both operators $\sigma_i\partial^s\sigma_i$ and $\sigma_{(i}\partial^s\sigma_{j)}$.[18] We can easily reproduce the coefficient of $1/s^2$ using our formula (4.45) and the known two- and three-point functions in this model (at large N). We find a precise match between the perturbative calculation (4.67) and our general methods using the bootstrap equations. The necessary details are collected in appendix E.

Now we can turn to predictions using convexity. We consider arbitrary N and ϵ. Of particular interest are three-dimensional small N models (for which there is no small parameter in the problem). They describe many second order phase transitions. For example, in the case of $N = 1$ we have the 3d Ising model which describes the vapor-liquid critical point of water.

All these theories are known to contain an almost free scalar operator (the spin field) $\Delta_\sigma \sim 0.5$, thus, the prediction is that all these theories contain operators of spin $4, 6, ..., \infty$ with twists

$$1 < \tau_s < 2\Delta_\sigma \ . \tag{4.69}$$

This happens to be a rather stringent constraint on the spectrum of the leading twist higher spin operators. For example, in the case of the 3d Ising model, it is known that $\Delta_\sigma = 0.5182(3)$ (see e.g. [105]). Thus, we would predict that there is an infinite set of operators with $1 < \tau_s < 1.037$. Moreover, based on other examples and on the known twist of the spin four operator that nicely satisfies (4.69) ($\tau_4 \sim 1.02$), we expect that these operators are present starting from $s = 2$.

Using monotonicity and the known twist of the spin four operator, we can further make the prediction that there is an infinite set of almost conserved currents in the 3D Ising model starting

[18] The cancelation of $1/s^{d-2}$ in the second line of (4.67) is easily understood in our language. Consider the four-point function of spin fields $\langle\sigma_1\sigma_2\sigma_2\sigma_1\rangle$. The only operator in the t-channel that can propagate is the interacting field σ^2. The higher spin currents cannot propagate because they behave essentially like free fields and so contribute zero to the t-channel diagrams. This shows that the operators $\sigma_1\partial^s\sigma_2$ do not have a piece that goes like $1/s^{d-2}$ at large spin.

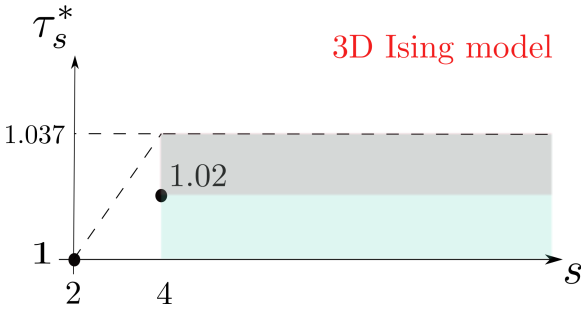


Figure 4.11: The form of the minimal twist spectrum implied by convexity in the 3D Ising model. The shaded region describes the allowed values for almost conserved currents of spin $s \geq 6$.

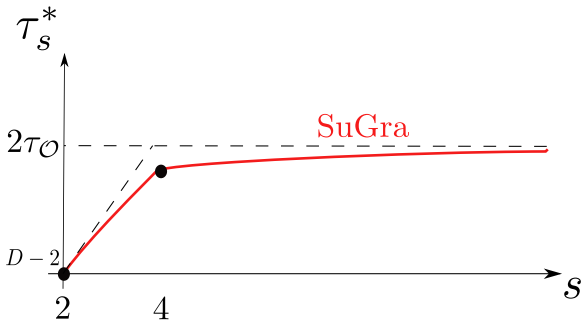


Figure 4.12: The form of the minimal twist spectrum implied by convexity in a theory described by supergravity.

from $s = 6$ with the twists

$$1.02 < \tau_s < 1.037 \ , \qquad s = 6, 8, ..., \infty \ . \tag{4.70}$$

This is illustrated in fig. 4.11.

4.5.3 Strongly Coupled Theories and AdS/CFT

For theories with gravity duals, convexity could be tested via AdS/CFT. One may be able to think about convexity as some hidden constraint on the low-energy effective actions in AdS in the spirit of [94]. We leave the exploration of this question for the future. A typical characteristic of theories with gravity duals is that these theories possess a large gap in the spectrum. See [47] for a systematic approach to analyzing such theories.

In this context, convexity is at statement about the anomalous dimensions of various double-trace operators, which are the double-twist operators in this case. For very large spins convexity was proved using the crossing equations in section 3 while for spins of order $\mathcal{O}(1)$ we can use the argument of section 2. In the bulk gravity duals, this corresponds to the energy of various spinning two-particle states. Hence, convexity implies that the binding energy of fast spinning operators in the bulk is always negative (i.e. the particles have net attraction) and the binding energy approaches zero in a power law fashion. This is depicted schematically in fig. 4.12, where $2\tau_{\mathcal{O}}$ is the twist of the two-particle state at very large spin (hence, the binding energy goes to zero). We emphasize that all these binding energies are of order $1/N$ or less. Hence, the properties delineated above are about various $1/N$ corrections that come from tree-level diagrams in supergravity.

4.5.4 Higher Spin Symmetry Breaking

Recently there was some progress in understanding the higher spin symmetric phase of the AdS/CFT correspondence [11, 25, 12]. On the CFT side one can notice that such symmetries fix correlations functions up to a number [60].[19]

However, in the case when higher spin symmetry is broken the situation is much less clear. On the boundary sometimes it is possible to use the constraints to get some general predictions [44, 45, 131]. But they are very limited and are not applicable in the general situation. In the bulk the situation is even less clear. It would be very interesting to understand higher spin symmetry breaking in the bulk better.

Convexity is a generic constraint on any possible higher spin symmetry breaking. On the boundary, slightly broken higher spin symmetry currents are the minimal twist operators discussed before. In the bulk, this maps to a statement about the masses of Higgsed gauge bosons.

The scaling dimensions of operators with spin s are mapped to the masses of the dual states in the bulk [132]

$$m^2 = C_2(\Delta, s) - C_2(d-2+s, s) = \delta\tau_s(\delta\tau_s + 2s + d - 4). \tag{4.71}$$

here C_2 is the quadratic Casimir of $SO(d,2)$ and $\delta\tau_s$ is the anomalous twist (when it is zero the particle is massless).

Thus, convexity is some general constraint on the higher spin symmetry breaking in quantum gravities in AdS. It could be applied both to Vasiliev theories [24] and tensionless limits of string theories [56].

Consider as an example case of parity preserving Vasiliev theory in AdS_4 with higher spin breaking boundary conditions. We know from additivity that for large spins $\tau_s = d-2+2\gamma_\sigma+O(s^{2-d})$

[19] The expression for n-point correlation functions of currents was recently found in [130].

so we can approximate the masses of the higher spin bosons in the bulk as

$$m_s^2 \approx 4\gamma_\sigma s + O(1) \; . \tag{4.72}$$

Here γ_σ is the anomalous dimension of the spin field in the $O(N)$ model (the singlet $O(N)$ model was first considered in [11]).

In the singlet $O(N)$ model, σ is not part of the spectrum, but we can nevertheless define its anomalous dimension:

$$\gamma_\sigma = \lim_{s\to\infty} \frac{1}{4}\frac{\partial m_s^2}{\partial s} = \lim_{s\to\infty} \frac{1}{2}(\tau_s - 1) \; . \tag{4.73}$$

Hence, it is defined from the limit of anomalous dimensions of almost conserved currents (alternatively, from the masses of bulk gauge bosons).

4.6 Conclusions and Open Problems

In this chapter we investigated unitary CFTs in the Lorentzian domain. By flowing to the gapped phase and using positivity of a certain cross section, we established a convexity property of the minimal twists in the $\mathcal{O}(x)\mathcal{O}^\dagger(0)$ OPE (4.26). Studying the crossing equations in the light-cone limit we found that there is an additivity property in the twist space and that the inequality (4.27) has to hold. Additivity means that in $d > 2$ it is meaningful to talk about double- and multi-twist operators, in any CFT. We also computed explicitly the leading large spin corrections to the twist of double-twist operators in the case of scalar operators and explained how to compute it for arbitrary operators. We discussed various double-twist operators and the leading correction to their twists at large spin. This led to a purely CFT-based proof of a certain convexity property of the twists of double-twist operators.

Our general findings are related naturally to the picture of [110], where 4d CFT correlators look like correlators in a 2d gapped local theory. In the original CFT this locality is realized in $\log s$. We reformulated their picture in terms of the OPE and that allowed us to extend the analysis of [110] and compute the leading spin corrections exactly.

We proceeded by studying various examples and applications of these general properties. We have found that convexity is a recurring theme at large spin. We checked that the general picture that we obtained is indeed realized in known theories (mostly weakly coupled theories or theories with gravity duals).

We concluded that there should not be solutions of the crossing equations with $\tau_s^* > \tau_1 + \tau_2$ and we checked our picture against the available numerical results [105]. We emphasized some predictions for the 3d critical $O(N)$ models that are relevant for phase transitions. One can set bounds on an infinite set of operators in the critical $O(N)$ models (4.69) . In the case of the 3d Ising model one can do better (4.70) .

Another example we considered is the correction to the twists of chiral operators in $\mathcal{N} = 1$ SCFTs. We found that to the order we computed the twist spectrum of chiral primaries remained flat at large spin (in principle it could have been concave). If we consider chiral operators that are not chiral primaries, then, as we have explained in section 4, the usual superconformal unitarity bound leads to convexity at large spin.

For a generic CFT we found that there is a relation between the convexity/concavity of the large spin twists of double-twist operators and the mass/charge ratio. More precisely, assuming that the minimal twist operators are the stress tensor and a $U(1)$ current we found that

$$\mathcal{O}\partial^s\mathcal{O} \text{ convexity} \leftrightarrow \Delta \geq \frac{(d-1)}{2}|q| \ , \tag{4.74}$$

where we used the freedom to normalize the current to choose $\frac{c_T}{\tau} = \frac{d(d+1)}{2}$ (we can always do this for Abelian currents. For concavity the inequality works in the opposite direction.)

Hence, by studying the dimensions and charges of arbitrary primary operators in various CFTs one can conclude about convexity/concavity of the large spin anomalous dimensions.

Our results are also relevant for the high spin symmetry breaking in CFTs and through AdS/CFT to theories of quantum gravity in the bulk. As higher spin symmetry breaks, higher spin gauge bosons become massive through the Higgs mechanism. The minimal twists discussed above correspond to almost conserved higher spin currents. For the case of parity preserving Vasiliev theory, from the property of additivity of the twists we infer that the masses of these gauge bosons at high spin follow a Regge-like formula, with the prefactor fixed by the dimension of the spin field.

There are several concrete open questions which would be interesting to address.

One obvious generalization would be to compute the next, subleading, orders in the $\frac{1}{s}$ expansion. Another concrete problem that would be nice to address is the computation of the coefficients controlling the power law corrections for the operators $\mathcal{O}_1\partial^s\Box^n\mathcal{O}_2$. Finally, it would be interesting to derive convexity using the ideas of [20]. It was noticed in [20] that the Regge limit of the DIS amplitude we considered in section 2 is related through a conformal transformation to the problem of measuring energy distributions in the final state created by some operator. Moreover, the fact that the energy distributions are integrable may constrain the Regge asymptotics of graviton deep inelastic scattering. It would be interesting to understand whether this is indeed the case.

4.7 Appendix A. Normalizations and Conventions

Here we collect our normalization and conventions for two- and three-point functions as well as the normalizations of conformal blocks. We mostly follow the conventions of [133]

$$\begin{aligned} r_{ij} = (x_i - x_j)^2,\ I^{\mu\nu}(x) &= \delta^{\mu\nu} - 2\frac{x^\mu x^\nu}{x^2}, \\ Z^\mu = \frac{x_{13}^\mu}{r_{13}} - \frac{x_{12}^\mu}{r_{12}}. \end{aligned} \tag{4.75}$$

Then for the two- and three-point functions contracted with polarization tensors C^μ we have

$$\langle O_l \cdot C_1(x_1) O_l \cdot C_2(x_2)\rangle = c_{O_l O_l} \frac{(C_1^\mu I_{\mu\nu} C_2^\nu)^l}{r_{12}^\Delta}, \tag{4.76}$$

$$\langle \mathcal{O}_{\Delta_1}(x_1)\mathcal{O}_{\Delta_2}(x_2) O_l \cdot C_3(x_3)\rangle = c_{\mathcal{O}\mathcal{O}O_l} \frac{(Z\cdot C)^l}{r_{12}^{\frac{1}{2}(\Delta_1+\Delta_2-\Delta)} r_{13}^{(\Delta_1+\Delta-\Delta_2)} r_{23}^{(\Delta_2+\Delta-\Delta_1)}}. \tag{4.77}$$

For the conformal block we have the following asymptotic behavior [134]

$$G_{d,\Delta,s}(z,\bar{z}) \sim_{z\to 0} (-\frac{1}{2})^s z^{\frac{\Delta-s}{2}} \bar{z}^{\frac{\Delta+s}{2}}\ {}_2F_1(\Delta+s, \Delta+s, 2s+2\Delta, \bar{z}) \ . \tag{4.78}$$

This leading behavior is independent of the number of space-time dimensions.

In the text various notions of twist, minimal twist etc. appeared. For convenience, we summarize here some of the notation and terminology we used.

- \- Conformal twist. It is given by $\tau = \Delta - s$ and governs the DIS experiment we considered. Unless stated otherwise, by twist we always mean the conformal twist.
- \- Collinear twist. It is given by $\tau^{\text{coll}} = \Delta - s_{+-}$ and the light-cone OPE is arranged according to collinear twists of operators.
- \- By τ_s^* we denote the conformal twist of the spin s *minimal* conformal twist operator that appears in the OPE of some operator with its conjugate $\mathcal{O}(x)\mathcal{O}^\dagger(0)$. This is again relevant for DIS.
- \- By $\tau_{\min}$ we denote the minimal twist among the operators different from the unit operator

that appear in the t-channel of various four point functions. The spin of the corresponding operator could be arbitrary.

- - By τ_s we denote the twists of double-twist operators that appear in the s-channel and correspond via the bootstrap equations to low twist operators in the t-channel.

4.8 Appendix B. Relating Spin to σ

4.8.1 A Naive Approach

Consider the disconnected part of the four-point function of real scalar operators with dimension Δ. In our notation for the cross ratios, the four point function takes the form

$$\mathcal{F}(z,\overline{z}) = 1 + (z\overline{z})^{\Delta} + \left[\frac{z\overline{z}}{(1-z)(1-\overline{z})}\right]^{\Delta} . \tag{4.79}$$

As in the main text, we take z to be small and consider the piece z^{Δ} only. This corresponds to focusing on the contributions of the leading collinear twist "double-trace" operators. This leading piece in the four-point function takes the form $z^{\Delta} f(\overline{z})$ where

$$f(\overline{z}) = \overline{z}^{\Delta}\left[1 + \frac{1}{(1-\overline{z})^{\Delta}}\right] . \tag{4.80}$$

We can decompose this function in terms of collinear conformal blocks

$$f(\overline{z}) = \sum_{s=0}^{\infty} c_s \overline{z}^{\Delta+s} \; {}_2F_1(\Delta+s, \Delta+s, 2s+2\Delta, \overline{z}) \; , \tag{4.81}$$

where the c_s are known to be (see e.g. [47])

$$c_s = (1+(-1)^s)\frac{\Gamma(\Delta+s)^2\Gamma(2\Delta+s-1)}{\Gamma(s+1)\Gamma(\Delta)^2\Gamma(2\Delta+2s-1)} . \tag{4.82}$$

This expansion is valid in any number of dimensions even though the full conformal blocks for generic d are not known. The simplification is due to the fact we consider small z and thus collinear conformal blocks, which are much simpler than the full conformal blocks.

Let us now substitute $1-\overline{z} = e^{-2\sigma}$. We would like to study where does the contribution come from in the sum (4.81) for some given σ. In other words, we would like to find the typical spin of operators that dominate the contribution for some given σ. It is helpful to replace the sum over spins by an integral and also utilize the integral representation of the hypergeometric function. We get for $f(\overline{z})$

$$\int_0^1 dt \int_\Delta^\infty d\widehat{s}\, \frac{2(2\widehat{s}-1)\Gamma(\Delta+\widehat{s}-1)}{\Gamma(\Delta)^2\Gamma(1+\widehat{s}-\Delta)} \frac{\left[\frac{(1-e^{-2\sigma})t(1-t)}{1-(1-e^{-2\sigma})t}\right]^{\widehat{s}}}{t(1-t)} \ , \tag{4.83}$$

We expect the large $\sigma \gg 1$ limit to be dominated by the high spin operators. Thus, we expand to the leading order in $\widehat{s}$

$$\frac{2(2\widehat{s}-1)\Gamma(\Delta+\widehat{s}-1)}{\Gamma(\Delta)^2\Gamma(1+\widehat{s}-\Delta)} = \frac{4\widehat{s}^{2\Delta-1}}{\Gamma(\Delta)^2} + O(\frac{1}{\widehat{s}}) \ , \tag{4.84}$$

so that we get for the integral

$$\frac{4}{\Gamma(\Delta)^2} \int_0^1 dt \int_\Delta^\infty d\widehat{s} \frac{e^{(2\Delta-1)\log\widehat{s}+\widehat{s}\log\left[\frac{(1-e^{-2\sigma})t(1-t)}{1-(1-e^{-2\sigma})t}\right]}}{t(1-t)} \ . \tag{4.85}$$

We perform the integral over $\widehat{s}$ by the saddle point approximation. We get for the extremum value $\widehat{s}_0$

$$\widehat{s}_0 = -\frac{2\Delta-1}{\log\left[\frac{(1-e^{-2\sigma})t(1-t)}{1-(1-e^{-2\sigma})t}\right]}. \tag{4.86}$$

After evaluating the integral we get

$$\frac{4e^{1-2\Delta}\sqrt{\pi}(2\Delta-1)^{2\Delta-\frac{1}{2}}}{\sqrt{2\Delta-1}\Gamma(\Delta)^2} \int_0^1 dt \frac{\left(-\log\frac{(e^{2\sigma}-1)(1-t)t}{e^{2\sigma}(1-t)+t}\right)^{-2\Delta}}{t(1-t)}. \tag{4.87}$$

Next we use the saddle point approximation to evaluate the t integral. The extremal value for t is

$$t_0 = 1 - \left(\frac{2\Delta-1}{2\Delta+1}\right)^{\frac{1}{2}} e^{-\sigma} + O(e^{-2\sigma}) \ . \tag{4.88}$$

Plugging it back to (4.86) we see that the dominant spin, s_{dom}, is

$$\log s_{\text{dom}} = \sigma + \log\frac{(2\Delta-1)^{\frac{3}{2}}(2\Delta+1)^{\frac{1}{2}}}{4\Delta} + O(e^{-\sigma}) \ . \tag{4.89}$$

The leading term in the relation (4.89) coincides with the identification of [110]. Notice that for

$\Delta \gg 1$ the expression (4.89) further reduces to

$$\log s_{\text{dom}} \approx \sigma + \log \Delta \ . \tag{4.90}$$

It is important to remark that the saddle point above is only localized when $\Delta \gg 1$, hence, one cannot perform reliable computations (for finite Δ) by the usual saddle point approximation method (one needs to re-sum various corrections). The purpose of this discussion is to motivate the claim that for $\bar{z} \to 1$ the important terms in the sum (4.81) are those with large spin, according to the relation $\log s_{\text{dom}} \approx \sigma$.

In the next subsectiontion we will present a more rigorous approach to the problem. This approach allows us to control the sum (4.81) efficiently. The following discussion is crucial for reproducing the main results of section 3.

4.8.2 A Systematic Approach

In the previous subsectiontion we learned that the spins dominating the s-channel expansion which reproduces the unit operator in the t-channel are given by $s \sim \frac{1}{(1-\bar{z})^{\frac{1}{2}}}$. We denote $\epsilon = 1 - \bar{z}$ and plug this into the hypergeometric function. Below we will discuss the leading terms are $\epsilon \to 0$. From the integral representation of the hypergeometric function we see that it behaves as follows in this scaling limit (A is an arbitrary coefficient that stays finite as $\epsilon \to 0$)

$${}_2F_1\left(\frac{A}{\sqrt{\epsilon}}, \frac{A}{\sqrt{\epsilon}}, 2\frac{A}{\sqrt{\epsilon}}, 1-\epsilon\right) = \frac{\Gamma(\frac{2A}{\sqrt{\epsilon}})}{\Gamma^2(\frac{A}{\sqrt{\epsilon}})} \int_0^1 dt (1-t)^{\frac{A}{\sqrt{\epsilon}}-1} t^{\frac{A}{\sqrt{\epsilon}}-1} (1-t+\epsilon t)^{-\frac{A}{\sqrt{\epsilon}}} \ . \tag{4.91}$$

We approximate the pre-factor using the Stirling formula as $\sqrt{\frac{A}{4\pi}} \frac{1}{\epsilon^{1/4}} e^{\frac{2A\log(2)}{\sqrt{\epsilon}}}$ and so

$${}_2F_1\left(\frac{A}{\sqrt{\epsilon}}, \frac{A}{\sqrt{\epsilon}}, 2\frac{A}{\sqrt{\epsilon}}, 1-\epsilon\right) = \sqrt{\frac{A}{4\pi}} \frac{1}{\epsilon^{1/4}} 4^{\frac{A}{\sqrt{\epsilon}}} \int_0^1 dt (1-t)^{\frac{A}{\sqrt{\epsilon}}-1} t^{\frac{A}{\sqrt{\epsilon}}-1} (1-t+\epsilon t)^{-\frac{A}{\sqrt{\epsilon}}} \ . \tag{4.92}$$

Now we need to take the $\epsilon \to 0$ limit of this integral.

From the analysis of the previous subsectiontion we are inspired to make a change of variables $1 - t = \lambda\sqrt{\epsilon}$ and find

$${}_2F_1\left(\frac{A}{\sqrt{\epsilon}}, \frac{A}{\sqrt{\epsilon}}, 2\frac{A}{\sqrt{\epsilon}}, 1-\epsilon\right) = \sqrt{\frac{A}{4\pi}} \frac{1}{\epsilon^{1/4}} 4^{\frac{A}{\sqrt{\epsilon}}} \int_0^{1/\sqrt{\epsilon}} \frac{d\lambda}{\lambda} (1-\lambda\sqrt{\epsilon})^{\frac{A}{\sqrt{\epsilon}}-1} (1+\frac{\sqrt{\epsilon}}{\lambda})^{-\frac{A}{\sqrt{\epsilon}}} \ . \tag{4.93}$$

We take the $\epsilon \to 0$ limit and get

$$
{}_2F_1\left(\frac{A}{\sqrt{\epsilon}}, \frac{A}{\sqrt{\epsilon}}, 2\frac{A}{\sqrt{\epsilon}}, 1-\epsilon\right) = \sqrt{\frac{A}{4\pi}}\frac{1}{\epsilon^{1/4}}4^{\frac{A}{\sqrt{\epsilon}}}\int_0^\infty \frac{d\lambda}{\lambda}e^{-\lambda A-\frac{A}{\lambda}} \ . \tag{4.94}
$$

This is a nicely convergent integral, given by the modified Bessel function of the second kind. Thus

$$
{}_2F_1\left(\frac{A}{\sqrt{\epsilon}}, \frac{A}{\sqrt{\epsilon}}, 2\frac{A}{\sqrt{\epsilon}}, 1-\epsilon\right) \to \sqrt{\frac{A}{\pi}}\frac{4^{\frac{A}{\sqrt{\epsilon}}}}{\epsilon^{1/4}}K_0(2A) \ . \tag{4.95}
$$

We can now use this limiting expression and reconsider (4.81) using this approximation for the hypergeometric function. We find

$$
f(\bar{z}) = \frac{4}{\Gamma^2(\Delta)\epsilon^\Delta}\int_0^\infty dA A^{2\Delta-1}K_0(2A) \ . \tag{4.96}
$$

It is encouraging to see that the dependence on ϵ is correct in the small ϵ limit. Now let us consider the pre-factor. This would be another test for the fact that the procedure above indeed captures the important contributions in the small ϵ limit. We use the integral $\int_0^\infty dAA^{2\Delta-1}K_0(2A) = \frac{1}{4}\Gamma^2(\Delta)$. Plugging this back into (4.96) we find

$$
f(\bar{z}) = \frac{1}{\epsilon^\Delta}\left(1+\mathcal{O}(\epsilon)\right) \ . \tag{4.97}
$$

This is precisely the correct asymptotics for small ϵ, including the pre-factor.

4.8.3 The Coefficient $c_{\tau_{\min}}$ in (3.12)

In the main text in section 3 we have seen that since the correction to the anomalous dimension is of the form $e^{-\tau_{\min}}$ and $\log s_{\text{dom}} = \sigma$, we expect the anomalous twist at large s to take the form $\frac{c_{\tau_{\min}}}{s^{\tau_{\min}}}$. Here we would like to compute $c_{\tau_{\min}}$ exactly. The idea is that the small corrections to the anomalous twists in the s-channel reproduce the contribution of the operator $\mathcal{O}_{min}$ in the t-channel. This contribution in the t-channel is given by

$$
-\frac{\Gamma(\tau_{\min}+2s_{\min})}{(-2)^{s_{\min}}\Gamma\left(\frac{\tau_{\min}+2s_{\min}}{2}\right)^2}\frac{C^2_{\mathcal{O}\mathcal{O}\mathcal{O}_{\tau_{\min}}}}{C^2_{\mathcal{O}\mathcal{O}}C_{\mathcal{O}_{\tau_{\min}}\mathcal{O}_{\tau_{\min}}}}\log z \ \ e^{(2\Delta-\tau_{\min})\sigma} \ . \tag{4.98}
$$

It is easy to match this in the s-channel because of the $\log z$. We reproduce this logarithm by summing the corrections to the twists in the s-channel $z^{\Delta+\frac{\delta\tau_s}{2}} \sim \frac{\delta\tau_s}{2}z^\Delta \log z$. Setting $\delta\tau_s = \frac{c_{\tau_{\min}}}{s^{\tau_{\min}}}$, we get the basic equation that determines $c_{\tau_{\min}}$

$$
\begin{aligned}
&-\frac{2}{c_{\tau_{\min}}}\frac{\Gamma(\tau_{\min}+2s_{\min})}{(-2)^{s_{\min}}\Gamma\left(\frac{\tau_{\min}+2s_{\min}}{2}\right)^2}\frac{C^2_{\mathcal{O}\mathcal{O}\mathcal{O}_{\tau_{\min}}}}{C^2_{\mathcal{O}\mathcal{O}}C_{\mathcal{O}_{\tau_{\min}}\mathcal{O}_{\tau_{\min}}}}e^{(2\Delta-\tau_{\min})\sigma} \\
&=\sum_{s=0}^{\infty}\frac{c_s}{s^{\tau_{\min}}}\bar{z}^{\Delta+s}\ {}_2F_1(\Delta+s,\Delta+s,2s+2\Delta,\bar{z}),
\end{aligned}
\tag{4.99}
$$

where $\bar{z}=1-e^{-2\sigma}$ and the equality should be understood in terms of *leading pieces in* σ.

Solving for $c_{\tau_{\min}}$ we get

$$
c_{\tau_{\min}}=-2\frac{\Gamma(\tau_{\min}+2s_{\tau_{\min}})}{(-2)^{s_{\tau_{\min}}}\Gamma\left(\frac{\tau_{\min}+2s_{\tau_{\min}}}{2}\right)^2}\frac{C^2_{\mathcal{O}\mathcal{O}\mathcal{O}_{\tau_{\min}}}}{C^2_{\mathcal{O}\mathcal{O}}C_{\mathcal{O}_{\tau_{\min}}\mathcal{O}_{\tau_{\min}}}}f(\tau_{\min},\Delta)\ , \tag{4.100}
$$

$$
f(\tau_{\min},\Delta)=\lim_{\sigma\to\infty}\frac{e^{(2\Delta-\tau_{\min})\sigma}}{\sum_{s=\Lambda}^{\infty}\frac{c_s}{s^{\tau_{\min}}}\bar{z}^{\Delta+s}\ {}_2F_1(\Delta+s,\Delta+s,2s+2\Delta,\bar{z})}\ . \tag{4.101}
$$

and since the sum is dominated by the large spins we expect the result to be independent of Λ.

We compute $f(\tau_{\min},\Delta)$ by performing the sum in the denominator carefully. We first switch to an integral and repeat the procedure of the previous subsectiontion, including the double scaling limit for the hypergeometric function (4.95). Doing the integral we finally find

$$
f(\tau_{\min},\Delta)=\frac{\Gamma(\Delta)^2}{\Gamma(\Delta-\frac{\tau_{\min}}{2})^2}\ . \tag{4.102}
$$

4.8.4 The Most General (Scalar) Case

We are interested in generalizing the computation of the correction $c_{\tau_{\min}}$ to the case when not all the operators are identical and $\Delta_1\neq\Delta_2$. This is relevant for a variety of applications in the main body of the text. More precisely, we are considering the following correlation function $\langle\mathcal{O}_1(x_1)\mathcal{O}_2(x_2)\mathcal{O}_2(x_3)\mathcal{O}_1(x_4)\rangle$. What we call the s-channel is the OPE 12 $\to$ 34. What we call the t-channel is the OPE 23 $\to$ 14. Let us write the OPE expansion in each of the channels.

For the s-channel OPE we have

$$
\langle\mathcal{O}_1(x_1)\mathcal{O}_2(x_2)\mathcal{O}_2^\dagger(x_3)\mathcal{O}_1^\dagger(x_4)\rangle=\left(\frac{x_{24}x_{13}}{x_{14}^2}\right)^{\Delta_1-\Delta_2}\sum_{\Delta,s}\frac{p_{\Delta,s}g_{\Delta,s}(z,\bar{z})}{x_{12}^{\Delta_1+\Delta_2}x_{34}^{\Delta_1+\Delta_2}}\ . \tag{4.103}
$$

And for the t-channel OPE we have

$$
\langle\mathcal{O}_2^\dagger(x_3)\mathcal{O}_2(x_2)\mathcal{O}_1(x_1)\mathcal{O}_1^\dagger(x_4)\rangle=\sum_{\Delta,s}\frac{\widetilde{p}_{\Delta,s}g_{\Delta,s}(1-z,1-\bar{z})}{x_{23}^{2\Delta_2}x_{14}^{2\Delta_1}}\ . \tag{4.104}
$$

The crossing equation takes the form

$$\sum_{\Delta,s} p_{\Delta,s} g_{\Delta,s}(z,\overline{z}) = \frac{z^{\frac{\Delta_1+\Delta_2}{2}} \overline{z}^{\frac{\Delta_1+\Delta_2}{2}}}{(1-z)^{\Delta_2}(1-\overline{z})^{\Delta_2}} \sum_{\Delta,s} \widetilde{p}_{\Delta,s} g_{\Delta,s}(1-z,1-\overline{z}) \ . \tag{4.105}$$

As before we focus on the unit operator in the t-channel. It is dominant when $\overline{z} \to 1$ so that the RHS at leading order takes the form $\frac{z^{\frac{\Delta_1+\Delta_2}{2}}}{(1-\overline{z})^{\Delta_2}}$ where we again keep only the leading piece in the z expansion. This is so that we only have to deal with the collinear conformal blocks instead of the full ones.

The collinear conformal blocks in this case are slightly different from the case $\Delta_1 = \Delta_2$, so that we get the following equation

$$\lim_{\overline{z}\to 1} \sum_{s=\Lambda}^{\infty} c_s \overline{z}^s \ {}_2F_1(s+\Delta_2, s+\Delta_2, 2s+\Delta_1+\Delta_2) = \frac{1}{(1-\overline{z})^{\Delta_2}} \ . \tag{4.106}$$

We know that the coefficients of the theory of generalized free fields solve this problem. They could be easily found using the results of [135]

$$c_s = \frac{\Gamma(\Delta_1+s)\Gamma(\Delta_2+s)\Gamma(\Delta_1+\Delta_2+s-1)}{\Gamma(s+1)\Gamma(\Delta_1)\Gamma(\Delta_2)\Gamma(\Delta_1+\Delta_2+2s-1)} \ . \tag{4.107}$$

The limit $\overline{z} \to 1$ in (4.106) is again dominated by large spins with the scaling as before. We have a similar double scaling limit of the hypergeometric function

$${}_2F_1\left(\frac{s}{\sqrt{\epsilon}}+\Delta_2, \frac{s}{\sqrt{\epsilon}}+\Delta_2, 2\frac{s}{\sqrt{\epsilon}}+\Delta_1+\Delta_2\right) \sim 2^{\Delta_1+\Delta_2} \epsilon^{\frac{\Delta_1-\Delta_2}{2}} 4^{\frac{s}{\sqrt{\epsilon}}} \frac{\sqrt{s} K_{\Delta_2-\Delta_1}(2s)}{\sqrt{\pi}\epsilon^{\frac{1}{4}}} \ . \tag{4.108}$$

Switching from the sum (4.106) to an integral we get

$$\frac{4(1-\overline{z})^{-\Delta_2}}{\Gamma(\Delta_1)\Gamma(\Delta_2)} \int_0^{\infty} ds \ s^{\Delta_1+\Delta_2-1} K_{\Delta_2-\Delta_1}(2s) = \frac{1}{(1-\overline{z})^{\Delta_2}} \ . \tag{4.109}$$

This is analogous to (4.96) and (4.97). It serves to check that we have understood correctly how to resum the hypergeomtric functions on the LHS of (4.106) in this subtle limit.

To compute the correction to anomalous dimension we need a slightly different integral

$$\frac{4}{\Gamma(\Delta_1)\Gamma(\Delta_2)} \int_0^{\infty} ds \ s^{\Delta_1+\Delta_2-\tau_{\min}-1} K_{\Delta_2-\Delta_1}(2s) = \frac{\Gamma(\Delta_1 - \frac{\tau_{\min}}{2})}{\Gamma(\Delta_1)} \frac{\Gamma(\Delta_2 - \frac{\tau_{\min}}{2})}{\Gamma(\Delta_2)} \ . \tag{4.110}$$

So that the crossing equation becomes

$$c_{\tau_{\min}} \frac{\Gamma(\Delta_1 - \frac{\tau_{\min}}{2})}{\Gamma(\Delta_1)} \frac{\Gamma(\Delta_2 - \frac{\tau_{\min}}{2})}{\Gamma(\Delta_2)} = 2f^2 \frac{\Gamma(\tau_{\min} + 2s_{\min})}{(-2)^{s_{\min}} \Gamma\left(\frac{\tau_{\min}+2s_{\min}}{2}\right)^2} \ , \tag{4.111}$$

and we have for the correction to the twist of $\mathcal{O}_1 \partial^s \mathcal{O}_2$ in this case

$$c_{\tau_{\min}} = 2f^2 \frac{\Gamma(\tau_{\min} + 2s_{\min})}{(-2)^{s_{\min}} \Gamma\left(\frac{\tau_{\min}+2s_{\min}}{2}\right)^2} \frac{\Gamma(\Delta_1)}{\Gamma(\Delta_1 - \frac{\tau_{\min}}{2})} \frac{\Gamma(\Delta_2)}{\Gamma(\Delta_2 - \frac{\tau_{\min}}{2})} \ , \tag{4.112}$$

$$f^2 = \frac{C_{\mathcal{O}_1 \mathcal{O}_1^\dagger \mathcal{O}_{\tau_{\min}}} C_{\mathcal{O}_2^\dagger \mathcal{O}_2 \mathcal{O}_{\tau_{\min}}}}{C_{\mathcal{O}_1 \mathcal{O}_1^\dagger} C_{\mathcal{O}_2 \mathcal{O}_2^\dagger} C_{\mathcal{O}_{\tau_{\min}} \mathcal{O}_{\tau_{\min}}}} \ . \tag{4.113}$$

If we take $\mathcal{O}_2 = \mathcal{O}_1^\dagger$, the coefficient $c_{\tau_{\min}}$ is positive for any intermediate operator $\mathcal{O}_{\min}$.

4.9 Appendix C. OPE of Chiral Primary Operators in SCFT

Here we discuss $d = 4$ $\mathcal{N} = 1$ SCFTs (see also [136] and references therein). The superconformal algebra contains as a bosonic sub-algebra the conformal algebra times $U(1)_R$. Performing radial quantization, we label representations by the quantum numbers (Δ, j_1, j_2, R), where Δ is the dimension of the operator.

Superconformal primaries are annihilated by S_α and $\overline{S}_{\dot\alpha}$. This implies that they are annihilated by special conformal transformations as well, hence, they are primaries in the usual sense. The complete representation can be constructed by acting with Q and $\overline{Q}$ on the superconformal primary.

The quantum numbers of Q_α are $(1/2, 1/2, 0, -1)$ and of $\overline{Q}_{\dot\alpha}$ are $(1/2, 0, 1/2, 1)$. Given a superconformal primary in the representation (Δ, j_1, j_2, R), we can act on it with Q_α and find states in the representation $(\Delta + \frac{1}{2}, j_1 - \frac{1}{2}, j_2, R-1) \oplus (\Delta + \frac{1}{2}, j_1 + \frac{1}{2}, j_2, R-1)$. Similarly, we can act with $\overline{Q}_{\dot\alpha}$ and find states in $(\Delta + \frac{1}{2}, j_1, j_2 - \frac{1}{2}, R+1) \oplus (\Delta + \frac{1}{2}, j_1, j_2 + \frac{1}{2}, R+1)$.

In both cases the irreducible representation with the smaller norm is then one with the smaller spins. Demanding therefore that the states above have non-negative norm we find two inequalities

$$\Delta + 2\delta_{j_1,0} - \left(2 + 2j_1 - \frac{3r}{2}\right) \geq 0 \ , \Delta + 2\delta_{j_2,0} - \left(2 + 2j_2 + \frac{3r}{2}\right) \geq 0 \ . \tag{4.114}$$

If one of the inequalities above is saturated, the superconformal primary is annihilated by some supercharges.

In the special case that either $j_1 = 0$ or $j_2 = 0$ (or both), there is more information that can be extracted by considering states at level 2. (This is similar to the case $j_1 = j_2 = 0$ in the ordinary conformal group.) Let us take for example $j_1 = 0$. Then states at level 2 transform in $(\Delta+1, 0, j_2, R-2)$. The norm of these states is such that the window

$$-\frac{3r}{2} < \Delta < 2 - \frac{3r}{2} \tag{4.115}$$

is *excluded*. The norm of this state at level 2 precisely vanishes when $\Delta = -3r/2$ or $\Delta = 2 - 3r/2$. Similarly, if $j_2 = 0$, the disallowed range is $\frac{3r}{2} < \Delta < 2 + \frac{3r}{2}$.

Chiral primaries are annihilated by all the $\overline{Q}_{\dot\alpha}$, which means they carry $j_2 = 0$, and thus satisfy $\Delta = 3r/2$.

We will now use these facts to analyze the operator product expansion of two chiral primaries $\Phi(x)\Phi(0)$. It is well known that there are no singularities in this OPE. Let us examine all the possible primaries that can appear in this OPE. (The operators that appear on the right hand side are not necessarily superconformal primaries.) We denote $R(\Phi) \equiv R_\Phi = 2\Delta_\Phi/3$.

- **A** Chiral primaries, i.e. states with $j_2 = 0$. Those must have $R = 2R_\Phi$ and thus $\Delta = 2\Delta_\Phi$. So this is just the operator Φ^2 in the OPE. It has $j_1 = j_2 = 0$.

  The other operators in the OPE must still be chiral (because the LHS is chiral) but they cannot be chiral primaries. So we have to write general supeconformal descendants that are chiral (annihilated by $\overline{Q}$) and are primaries of the usual conformal group. So we can have

- **B** $\overline{Q}^{(\dot\alpha_1}\mathcal{O}^{\dot\alpha_2\ldots\dot\alpha_l),(\alpha_1,\ldots,\alpha_l)}$ with $\mathcal{O}$ a superconfromal primary. This is a conformal primary, and to make sure it is chiral we need the superconformal primary to obey $\overline{Q}_{\dot\alpha}\mathcal{O}^{(\dot\alpha\ldots\dot\alpha_l),(\alpha_1,\ldots,\alpha_l)} = 0$. This first order condition guarantees that $\overline{Q}^{(\dot\alpha_1}\mathcal{O}^{\dot\alpha_2\ldots\dot\alpha_l),(\alpha_1,\ldots,\alpha_l)}$ is a chiral field. By matching the R-charge we also find $R_\mathcal{O} = 2R_\Phi - 1$. The shortening condition and the R-charge imply immediately from (4.114) that $\Delta_\mathcal{O} = 2 + 2j_2 + 3R_\Phi - 3/2$. We dropped the term $\delta_{j_2,0}$ because we know that $\mathcal{O}$ has to have half-integer spin in the $\dot\alpha$ indices. (In other words, l is even.) Finally, we substitute $j_2 = l/2 - 1/2$ and find $\Delta_\mathcal{O} = l + 2\Delta_\Phi - 1/2$, and hence the dimension of $\overline{Q}^{(\dot\alpha_1}\mathcal{O}^{\dot\alpha_2\ldots\dot\alpha_l),(\alpha_1,\ldots,\alpha_l)}$ is $l + 2\Delta_\Phi$. This means that the twist of these operators is precisely 2Δ.[20]

- **C** We can have $\overline{Q}^2\mathcal{O}^{(\dot\alpha_1\ldots\dot\alpha_l),(\alpha_1,\ldots,\alpha_l)}$ for $\mathcal{O}$ superconformal primary. We immediately find that $R_\mathcal{O} = 2R_\Phi - 2$ and plugging this into the unitarity bounds we obtain $\Delta(\overline{Q}^2\mathcal{O}) \geq 1 + 2 + l +$

[20]Another closely related class of operators one could think of would be obtained by antisymmetrizing the spinor index of $\overline{Q}$ with some superconformal primary $\mathcal{O}$ and imposing a first order equation that the symmetrization gives zero. (This equation is necessary for chirality.) This is clearly inconsistent, since, as we explained, among the descendants, the state of lowest norm is always the state of smallest spin.

$3R_\Phi - 3 = l + 2\Delta$ and $\Delta(\overline{Q}^2\mathcal{O}) \geq 1 + 2 + l - 3R_\Phi + 3 = l + 6 - 2\Delta$. Either way, the twists of these operators are clearly not smaller than 2Δ.[21]

4.10 Appendix D. DIS for Traceless Symmetric Representations

Consider a symmetric traceless operator of spin s, $\mathcal{O}_{\mu_1\ldots\mu_s}(y)$. We can contract it with a light like complex polarization vector ζ^μ, $\zeta^2 = 0$. We denote the result as $\mathcal{O}(\zeta, y)$. The object we are interested in is

$$A(\nu, q^2, \zeta) = \frac{i}{\pi}\int d^d y e^{iqy}\langle P|\mathrm{T}\left(\mathcal{O}(\zeta^*, y)\mathcal{O}(\zeta, 0)\right)|P\rangle \ , \qquad \nu = 2q\cdot P \ . \tag{4.116}$$

The imaginary part of $\mathcal{A}$ is related to the total cross section for DIS.

Let us consider the most general possible form for the operator product expansion

$$\mathcal{O}(\zeta^*, y)\mathcal{O}(\zeta, 0) = \sum_{s=0,2,4\ldots}\ \sum_{\alpha\in\mathcal{I}_s} f_s^{(\alpha)}(y, \zeta, \zeta^*)^{\mu_1\ldots\mu_s}\mathcal{O}^{(\alpha)}_{\mu_1\ldots\mu_s}(0) \ , \tag{4.117}$$

where the symbols $s, \mathcal{I}_s$ represent the same objects as in section 2. If the operator $\mathcal{O}_{\mu_1\ldots\mu_s}(x)$ is conserved that would constrain the allowed structures in (4.117). In the OPE (4.117) we retain only primary symmetric traceless operators on the right hand side. The other operators are discarded because they do not contribute to DIS as long as the target particle $|P\rangle$ is a scalar. (We will henceforth assume the target is a scalar particle for simplicity.)

We substitute the OPE expansion (4.117) in (4.116) and evaluate the expectation values of the operators according to

$$\langle P|\mathcal{O}^\alpha_{\mu_1\ldots\mu_s}(0)|P\rangle = \mathcal{A}_n^{(\alpha)}\left(P_{\mu_1}P_{\mu_2}\cdots P_{\mu_s} - traces\right) \ . \tag{4.118}$$

The kinematics is a little more complicated than in the scalar case of section 2 because of the polarization vector ζ. However, there is a simple choice for ζ which makes the problem virtually

[21] One can in fact argue the twist must be strictly larger than 2Δ. Suppose the twist were 2Δ. Then the second inequality in (4.114) would have been saturated. In this case $\overline{Q}_{\dot\alpha}\mathcal{O}^{(\dot\alpha\ldots\dot\alpha_l),(\alpha_1,\ldots,\alpha_l)} = 0$. But in this case $\overline{Q}^2\mathcal{O}^{(\dot\alpha_1\ldots\dot\alpha_l),(\alpha_1,\ldots,\alpha_l)} = 0$ and thus there is no such operator in the OPE.

isomorphic to the scalar case. We can always pick

$$\zeta \cdot P = 0 \ . \tag{4.119}$$

In this case, the contribution from the $P_{\mu_1} P_{\mu_2} \cdots P_{\mu_s}$ term in (4.118) automatically selects the same kinematic structure as in the scalar case, and all the arguments from section 2 go through.[22] One finds the same sum rules and obtains convexity under the same assumptions as in section 2.

4.11 Appendix E. The $\frac{1}{s^2}$ Correction in the Critical $O(N)$ Model

Here we would like to show how one can use the general formula for the coefficient $c_{\tau_{\min}}$ in order to compute the $\frac{1}{s^2}$ tem in (4.67). The correction $1/s^2$ arises from the exchange of the σ^2 operator in the t-channel.

To contrast (4.67) with our formula (4.45) we need to compute the three-point function $\langle \sigma_i(x_1)\sigma_j(x_2)\sigma^2(x_3)\rangle$. For the two-point functions we have [12]

$$\begin{aligned} \langle \sigma_i(x)\sigma_j(0)\rangle &= \frac{\delta_{ij}}{N}\frac{\Gamma(\frac{d}{2}-1)}{4\pi^{\frac{d}{2}}}\frac{1}{x^{d-2}} = \frac{\gamma_s}{N}\frac{1}{x^{d-2}} \ , \\ \langle \sigma^2(x)\sigma^2(0)\rangle &= \frac{\gamma_{\phi^2}}{N}\frac{1}{x^4} \ , \quad \gamma_{\phi^2} = \frac{2^{d+2}\sin\left(\frac{\pi d}{2}\right)\ \Gamma(\frac{d-1}{2})}{\pi^{\frac{3}{2}}\Gamma(\frac{d}{2}-2)} \ . \end{aligned} \tag{4.120}$$

The three-point function $\langle \sigma_i(x_1)\sigma_i(x_2)\sigma^2(x_3)\rangle$ at leading order is given by the diagram fig. 4.13, where the interaction vertex is $-\frac{N}{2}\sigma_i^2$. So that we get

$$\begin{aligned} \langle \sigma_i(x_1)\sigma_i(x_2)\sigma^2(x_3)\rangle &= -\frac{\gamma_{\sigma^2}\gamma_s^2}{N^2}\int d^dx_0 \frac{1}{x_{10}^{d-2}x_{20}^{d-2}x_{30}^4} \\ &= -\frac{\gamma_{\sigma^2}\gamma_s^2}{N^2}\frac{\pi^{\frac{d}{2}}\Gamma(\frac{d}{2}-2)}{\Gamma(\frac{d-2}{2})^2}\frac{1}{x_{12}^{d-4}x_{13}^2x_{23}^2} \ . \end{aligned} \tag{4.121}$$

[22] Again, we need to be careful that the "trace" terms in (4.118) do not overwhelm contributions coming from lower spins (with a different power of ν). In section 2 we have already analyzed this for the case $f_s^{(\alpha)}(y,\zeta,\zeta^*)^{\mu_1\cdots\mu_s} \sim y^{\mu_1}\cdots y^{\mu_s}$. Here we could obtain new structures, such as $f_s^{(\alpha)}(y,\zeta,\zeta^*)^{\mu_1\cdots\mu_s} \sim \zeta^{\mu_1}(\zeta^*)^{\mu_2}y^{\mu_3}\cdots y^{\mu_s}$. Due to (4.119), this could contribute only when dotted into the "trace" terms, e.g. $g_{\mu_1\mu_2}P_{\mu_3}\cdots P_{\mu_s}$. A simple calculation shows that this would scale like (at large $-q^2$) $x^{-s+2}(q^2)^{-\frac{1}{2}\tau_s^*+\Delta_\mathcal{O}-d/2}$. Comparing this to the already existing contribution from spin $s-2$, $x^{-s+2}(q^2)^{-\frac{1}{2}\tau_{s-2}^*+\Delta_\mathcal{O}-d/2}$, we see that again it is sufficient that the twists τ_s^* are nondecreasing, as found in (4.24).

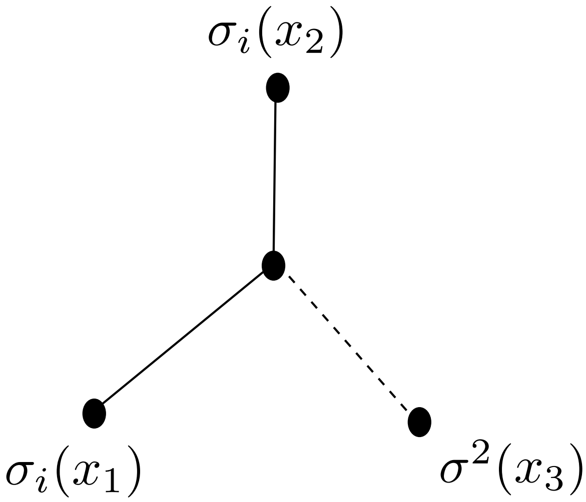


Figure 4.13: The diagram for the three-point function $\langle\sigma_i(x_1)\sigma_i(x_2)\sigma^2(x_3)\rangle$.

(We used the standard formula [79], and there is no summation over i). Thus, we get

$$\frac{C^2_{\sigma\sigma\sigma^2}}{C_{\sigma^2\sigma^2}C^2_{\sigma\sigma}} = \frac{1}{N}\gamma_{\sigma^2}\gamma_s^2\frac{\pi^d\Gamma(\frac{d}{2}-2)^2}{\Gamma(\frac{d-2}{2})^4} \ . \tag{4.122}$$

Now we have everything we need to apply (4.45). (Notice that in this case $\tau_{\rm min} > 2\Delta$, but we can think about our formula in terms of an analytic continuation.) Thus we set $s_{\rm min} = 0$, $\tau_{\rm min} = 2$ and $\Delta = \frac{d-2}{2}$ to get

$$c_{\sigma^2} = \frac{2^{d-1}\Gamma(\frac{d-1}{2})\sin(\frac{\pi d}{2})}{N\pi^{3/2}\Gamma(\frac{d}{2}-2)} \ , \tag{4.123}$$

which can be easily checked to reproduce the result (4.67) precisely.

Chapter 5

Causality Constraints on Corrections to the Graviton Three-Point Coupling

5.1 Introduction/Motivation

In this chapter we consider weakly coupled gravity theories in the tree-level approximation. It is well-known that at long distances such theories should reduce to the Einstein gravity theory. However, at intermediate energies we can have higher derivative corrections. By intermediate energies we mean those that are low enough that the theory is still weakly coupled but high enough that we are sensitive to possible higher derivative corrections. An example of such a theory is weakly coupled string theory where the corrections appear at a length scale $\sqrt{\alpha'}$, which is much larger than the Planck length. The theory at energies comparable to $1/\sqrt{\alpha'}$ is still weakly coupled. In this case, the higher derivative corrections are accompanied by extra massive higher spin particles which appear at the same scale. For higher energies the description is via a string theory which departs significantly from ordinary local quantum field theory. It is reasonable to expect that this is a generic feature. Namely, that higher derivative corrections only arise due to the presence of extra states with masses comparable to the scales where the higher derivative corrections become important. The objective of this chapter is to sharpen this link for the simplest possible correction, that of the graviton three-point coupling. Due to the fact that the graviton has spin, the flat space on-shell three-point function is not uniquely specified. In general, it has three different possible structures. The most familiar is the one we get in the Einstein theory. The others can be viewed as arising from higher derivative terms in the gravitational action. The first new structure has two more derivatives. Relative to the size of the Einstein-Hilbert term, it scales like αp^2, where α is a new quantity with dimensions of length squared which characterizes the relative importance of the new term. Here p is the typical scale of the momenta, it is not a Mandelstam invariant, since they all vanish for on-shell three-point functions. We work in a regime where both of the three-point vertices are small, so that gravity is weakly coupled. We find that new three-point vertices lead to a potential causality violation unless we get contributions from extra particles. This causality violation is occurring when the theory is still weakly coupled. It occurs in a high-energy, fixed impact parameter, scattering process at a center of mass (energy)^2, s, which is large compared to $1/\alpha$ but still small enough for the coupling to be weak. In general relativity this scattering process leads to the well-known Shapiro time delay [137], which is one of the classical tests of general relativity [138]. See also [139]. When the graviton three-point vertex is corrected, the new term can lead to a time advance, depending on the spin of the scattered graviton. At short enough impact parameter this time advance can overwhelm Shapiro's time delay and lead to a causality problem. This troublesome feature arises at an impact parameter of order $b^2 \sim \alpha$. At tree-level, this problem can only be fixed by introducing an infinite number of new massive particles with spin $J > 2$ and m^2 comparable to α^{-1}.[1] In other words, it cannot be fixed by adding particles with spins $J \leq 2$, or by considering the existence of extra dimensions. These causality constraints are similar in spirit to those considered in [140, 94] but they differ in two ways. First of all, here the problem will be shown to arise for small t/s, but large s. Second, the fact that the graviton has spin is crucial. On the other hand, in both cases we have locally Lorentz-invariant Lagrangians that nevertheless can lead to causality violations in non-trivial backgrounds. An example of a theory that is constrained by these considerations is given by the action

$$S = l_p^{2-D} \int d^Dx\sqrt{g}\left[R + \alpha\left(R^2_{\mu\nu\rho\sigma} - 4R^2_{\mu\nu} + R^2\right)\right] \ , \tag{5.1}$$

where the second term is the Lanczos-Gauss-Bonnet term.[2] The constant α has dimensions of length squared. For $\alpha \gg l_p^2$, we will show that the theory is not causal. Furthermore, there is no way to

[1] In $D > 4$ dimensions by spin $J > 2$ particles we mean particles in the representations of little group $SO(D-1)$ with both of the following two properties (see also appendix H): a) their maximal spin projection $J_{+-} \geq 2$; b) their representations are labeled by Young tableaux with three or more boxes.

[2] It is the dimensional continuation of the four-dimensional Euler density. In four dimensions it is a topological term, while in higher dimensions it is not topological, in fact it contributes to the three-point coupling of the graviton.

make it causal by adding local higher curvature terms. In fact, our discussion refers to on-shell data, namely the three-point function, which reflects the real physical information and does not depend on the particular way that we write the Lagrangian. In other words, the discussion is invariant under field redefinitions. Note that if we view gravity as a low-energy effective theory with a UV cutoff of order M_{pl} and we add higher derivative terms with dimensionless coefficients which are of order one, then we have nothing to say. The remarks in this chapter *only apply* for theories where the coefficients of the higher derivative terms are much larger. The "natural" value for the coefficient α if we view (5.1) as an effective gravity theory is $\alpha \sim l_p^2$. We will only constrain larger values of α. Of course, we are discussing this problem because it *is indeed possible* to have theories with $\alpha \gg l_p^2$, for example a weakly coupled string theory. Another example where this discussion is relevant is the following. Imagine that we consider a large N gauge theory. Such a gauge theory is expected to have a weakly coupled string dual. This theory will have a weakly coupled graviton corresponding to the stress tensor operator [6, 8, 7]. However, we are not guaranteed that the dual will be an ordinary Einstein gravity theory. It might even be a Vasiliev-like gravity theory [54, 37]. The field theory three-point functions of the stress tensor [141, 142] determine the three-point functions of the gravity theory [143, 20]. One can imagine a theory where the only light single trace operator is the stress tensor. It is natural to expect that this theory will have an Einstein gravity dual. It would be nice to prove that. For scalar interactions, [47] argued that the solutions of the crossing equation have solutions that correspond to local vertices in the bulk, see also [48] . However, one can worry that the size of the vertices with higher derivatives might be comparable to the one in the Einstein theory. As a case in point, we can think about the graviton three-point coupling. This is constrained to be a linear combination of three structures, only one of which is the Einstein one. Here we will argue that the other ones lead to a causality problem unless we introduce new higher spin particles at the scale that appears in these new three-point functions. Thus we link the three-point function of the stress tensor to the operator spectrum of the theory. These three-point functions were constrained by causality in [20] . Here we get stronger constraints because we are making further assumptions about the operator spectrum of the theory. As another application we consider the possibility that gravity waves during inflation were generated by a theory that indeed had these higher derivative corrections with a size comparable to the Hubble scale. This is a possibility which is allowed by conformal invariance and would be realized if the dual description to inflation (in the spirit of dS/CFT [144, 145, 146]) was a weakly coupled theory or if inflation occurs in a string theory where the string scale is close to the Hubble scale. The theory is still weakly coupled, so that scalar and tensor fluctuations are small. In this case the gravity wave non-gaussianities would be different from the ones in the Einstein theory [49, 147]. The observation of such gravity wave signals, combined with the arguments in this chapter, would imply the existence of extra particles with spin $J > 2$ during inflation. Finally, as a further motivation we should mention the grand dream of deriving the most general weakly coupled consistent theory of gravity. It is quite likely that the only such theory is a string-like theory, broadly defined. We are certainly very far away from this dream, but hopefully our simple observation about three-point functions could be useful. In particular, this observation highlights the importance of spin. Spin is likely to grow in importance as we consider constraints on the four-point function, given what we got for the much simpler three-point function. In [**?**, 149, 150] various interesting constraints were derived by using crossing symmetry and the correct factorization on the pole singularities. The constraints discussed in this chapter are additional constraints, not covered by their analysis. This chapter is organized as follows. In section two we discuss the notion of asymptotic causality, both for flat space and AdS space. We also discuss the propagation of particles and fields through a shock wave. The purpose of this discussion is to set the stage for the more general argument in the following section so that it becomes more intuitively clear. In section three, we present the main thought experiment which involves only on-shell amplitudes and does not refer explicitly to the the shock waves. In section four we discuss the effects of adding extra massive particles. We show that particles with spins two or less cannot solve the problem. We also discuss the appearance of these massive particles among the final states of the scattering process. In order to argue that the problem persists we present an alternative presentation of the problem where we consider the analyticity properties of the S matrix in impact parameter representation. In section five we discuss various aspects of the AdS version of the thought experiment and its implications for properties of the dual theory. In section 6 we briefly mention the implications of a possible time advance in the context of wormholes. In section 7 we discuss a cosmological application, where we link the possible existence of new structures for the gravity wave non-gaussianity to the presence of higher spin massive particles during inflation.

5.2 Flat Space Causality and Shock Waves

In this section we consider the problem in asymptotically flat space. We start by discussing the causal structure in asymptotically flat space. As a motivation for our later discussion we analyze the scattering of a probe graviton from the shock wave. We then present the argument for causality violation in purely on-shell terms.

5.2.1 Statement of Flat Space Causality

When we consider a gravity theory in asymptotically flat space we expect to be able to define scattering amplitudes. In particular, this presupposes that one can fix the asymptotic structure of the spacetime so that we can compare times between the past and the future asymptotic regions. In dimensions $D > 4$ we expect to be able to do this. Let us give a simple argument. Imagine we have a series of observers that sit at a large distance L from each other, and also from the center. For simplicity imagine them at the vertices of a large spatial hypercube and moving along the time direction. We would like to argue that they can synchronize their clocks. If they were in flat space, they can just send signals to each other. On the other hand if there is an object of mass m in the interior, then the metric components decay as $1/r^{D-3}$. This gives rise to a redshift of order $\frac{Gm}{L^{D-3}}$ at the position of the detectors. More importantly, as a signal travels from one detector to the other, staying at a distance of order L from the center, it will get delayed by an amount $\delta t = \delta L \sim \frac{Gm}{L^{D-4}}$ (see appendix A). In $D \geq 5$, we can make this as small as we want by moving out to large enough L. For $D = 4$ this delay does not go to zero and we have a problem. This problem has been discussed in [151] , and it seems closely related to the soft graviton issues that arise in perturbative attempts to define the gravitational S-matrix. This issue is not present in AdS_4. In order to deal with the $D = 4$ case, Gao and Wald [152] have introduced another notion of causality saying that *we cannot send signals faster than what is allowed by the asymptotic causal structure of the spacetime.* In general relativity, with the null energy condition, they argued that this notion of causality is respected, see [152] for a precise statement of the theorem. This holds in any number of dimensions. For $D > 4$ this becomes the more naive notion of causality introduced above. We expect that this notion of causality is actually a *requirement* for any theory of quantum gravity. For asymptotically AdS theories we expect that we should not be able to send signals through the bulk faster than through the boundary. For theories of gravity that are dual to a quantum field theory in the boundary, this is implied by causality in the boundary theory. Also, since we expect that quantum gravity in asymptotically flat space is Lorentz invariant, then these time delays can lead to acausality or closed time-like curves (see appendix G). Another reason to require this notion is to ensure that Lorentzian wormholes, such as the one obtained form the eternal Schwarzschild black hole and discussed in [153] , do not lead to causality violations in the ambient spacetime. This is required by the ER=EPR interpretation of such geometries [154, 155, 156]. In the rest of the chapter we will *assume* this notion of causality and derive constraints on some higher derivative corrections.

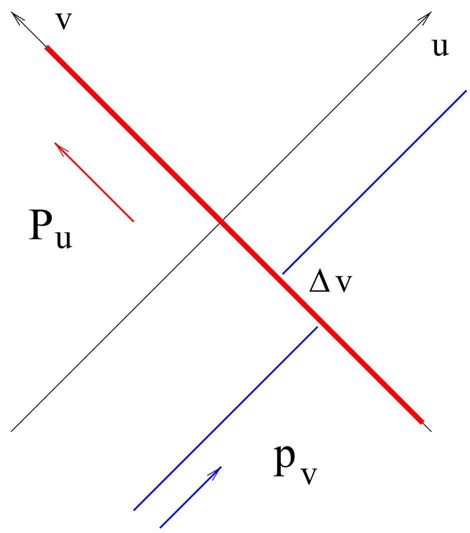


Figure 5.1: A particle creates a shock wave localized at $u = 0$. A second probe particle propagates on the geometry and experiences a time delay, Δv. The two particles are separated along the transverse directions, which are suppressed in this diagram.

5.2.2 Scattering Through a Plane Wave in General Relativity

A well-known general relativistic effect is that light going near a massive body (e.g. the Sun) would suffer a time delay relative to the same propagation in flat space. This is known as the Shapiro time delay [137] and constitutes one of the classical tests of general relativity (see appendix A). We will recall here the derivation of this time delay in the shock wave approximation, which will be all we need for our purposes. First we review the basic properties of shock wave solutions in gravitational theories [157]. This solution describes the gravitational field of an ultrarelativistic particle in a generic theory of gravity [158, 159] and is directly relevant for the high-energy scattering.

A generic shock wave solution in flat space can be written in the following form

$$ds^2 = -dudv + h(u, x_i)(du)^2 + \sum_{i=1}^{D-2}(dx_i)^2. \tag{5.2}$$

This geometry admits a covariantly constant null Killing vector $l^\mu \partial_\mu = \partial_v$. See (5.1).

We will be interested in the case when this geometry is sourced by a particle that moves very fast in the v direction. Classically, we can model such particle via the stress tensor [158]

$$T_{uu} = -P_u \delta(u)\delta^{D-2}(\vec{x}) \tag{5.3}$$

where $r = \sqrt{\sum_{i=1}^{D-2} x_i^2}$ and $P_u < 0$ is momentum of a particle. The Einstein equations then take the

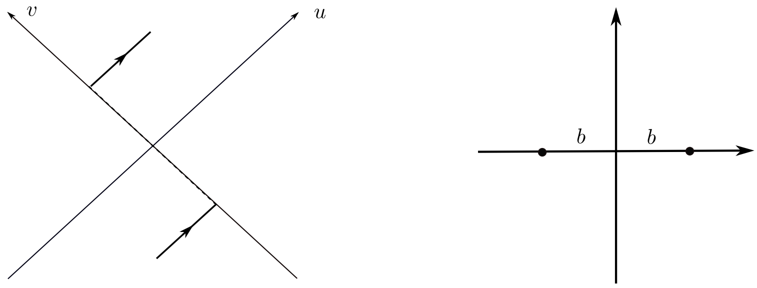


Figure 5.2: Left: we consider a particle propagating through the superposition of two left moving shock waves localized at $u = 0$. The particle trajectory is given by the arrows. Right: in the transverse plane we separate shock waves by distance $2b$ and send the particle between them so that the net deflection angle is zero. The time delay is the sum of the time delays due to each shock wave.

form

$$\partial_\perp^2 h(u, x_i) = -16\pi G|P_u|\delta(u)\delta^{D-2}(\vec{x}). \tag{5.4}$$

The solution is

$$h(u, x_i) = \frac{4\Gamma(\frac{D-4}{2})}{\pi^{\frac{D-4}{2}}}\delta(u)\frac{G|P_u|}{r^{D-4}}. \tag{5.5}$$

Now we consider a probe particle that moves in the other light cone direction with momentum p_v and is such that it crosses the shock with impact parameter b. In other words, the displacement in the transverse dimensions is $r = b$. The metric (5.2) is a bit peculiar because of the delta function $\delta(u)$ in $h(u, x_i)$. We can remove this delta function at the transverse position b by defining a new coordinate

$$v = v_{\text{new}} + \frac{4\Gamma(\frac{D-4}{2})}{\pi^{\frac{D-4}{2}}}\frac{G|P_u|}{b^{D-4}}\theta(u) \tag{5.6}$$

This then cancels the $\delta(u)$ term in the metric at $r = b$. Thus, the geodesic that goes through this point is continuous in the v_{new} coordinates. This means that in the original coordinates it suffers a shift

$$\Delta v = v_{\text{After crossing}} - v_{\text{Before crossing}} = \frac{4\Gamma(\frac{D-4}{2})}{\pi^{\frac{D-4}{2}}}\frac{G|P_u|}{b^{D-4}} > 0 \tag{5.7}$$

This represents the Shapiro time delay, see (5.1).

In addition to this time delay, the trajectory also is subject to a deflection angle. We might worry that the deflection angle would hamper our ability to see a possible time delay or time advance from far away. In fact, we could consider two shocks in succession, but separated in the transverse direction. We set one at $r = 0$ and another at $r = 2b$. See (5.2). In this case, the probe particle coming at impact parameter b, or $r = b$, does not get a net deflection, but the time delays add. It is instructive to reproduce this formula for the time delay when we treat the probe as a quantum mechanical particle. The wave equation for a scalar field takes the form

$$\nabla^2\phi = 0 \quad \longrightarrow \quad \partial_u\partial_v\phi + h\partial_v^2\phi - \frac{1}{4}\partial_i^2\phi = 0, \tag{5.8}$$

Let us now consider the change in the value of ϕ from $u = 0^-$ to $u = 0^+$. Since the variation of the h term is much faster than the variation in the other variables, we neglect the ∂_i derivatives and write

$$\begin{aligned} \phi(u = 0^+, v, x^i) &= e^{-\int_{0^-}^{0^+} du h \partial_v} \phi(u = 0^-, v, x^i) \\ &= e^{-\Delta v \partial_v}\phi(u = 0^-, v, x^i) = e^{-i\Delta v p_v}\phi(u = 0^-, v, x^i) \end{aligned} \tag{5.9}$$

where Δv is given in (5.7) . Thus we see that we reproduce the answer we had with the geodesic analysis. Note that $p_v = -i\partial_v$ is the generator of translations in v. In quantum field theory we would end the discussion here. A time advance in v, relative to the background Minkowski metric would be a problem. In gravity, the situation is more subtle, in principle, we need to make observations from asymptotically far away. However, in that case, the p_u energy also has a p_v dependence and it contributes positively to the time delay. Since $p_u = \vec{q}^2/4p_v$, this is a small effect for large p_v. However, if we multiply by the total $u-$time elapsed, it can add to a very big time delay. Here we will assume that we can sit far enough from the shock to be able to neglect the dynamics of spacetime, but close enough that we can neglect the $v-$time delay produced by the p_u energy. This seems possible for small G.

subsectionConnection with the Scattering Amplitude Computation Let us reproduce the computation above using scattering amplitudes [139] . It is well-known that the shock wave computation can be reproduced using the so-called eikonal approximation [160]. Consider the scattering amplitude for gravitating scalar particles. It is given by

$$\mathcal{A}_{tree}(s,t) = -8\pi G\frac{s^2}{t}. \tag{5.10}$$

The eikonal approximation resums a particular set of diagrams (horizontal ladders) in the deflection-less limit when $\frac{t}{s} \to 0$. Under favorable circumstances,[3] the amplitude exponentiates in the impact parameter space (see e.g. [161, 162])

$$i\mathcal{A}_{eik} = 2s \int d^{D-2}\vec{b}\; e^{-i\vec{q}.\vec{b}} \left[e^{i\delta(b,s)} - 1 \right] \tag{5.11}$$

where the phase is given by

$$\delta(b,s) = \frac{1}{2s} \int \frac{d^{D-2}\vec{q}}{(2\pi)^{D-2}} e^{\vec{q}.\vec{b}} \mathcal{A}_{tree}(s, -\vec{q}^2) = \frac{\Gamma(\frac{D-4}{2})}{\pi^{\frac{D-4}{2}}} \frac{Gs}{b^{D-4}}. \tag{5.12}$$

This result matches the shock wave computation

$$\delta(b,s) = -p_v \Delta v_{|r=b}, \tag{5.13}$$

where we used $s = 4P_u p_v$. As we will see below this picture can be naturally generalized to the scattering of particles with spin.

5.2.3 The Effect of Higher Derivative Interactions on Particles with Spin

If we had a photon propagating through the plane wave, it will have the same time delay we computed above (5.7) . However, if the Lagrangian contains certain higher order interactions, such as

$$\S = \int d^D x \sqrt{-g} \frac{1}{4} \left[F_{\mu\nu} F^{\mu\nu} + \widehat{\alpha}_2 R^{\mu\nu}{}_{\sigma\delta} F_{\mu\nu} F^{\sigma\delta} \right] \tag{5.14}$$

then the second term gives rise to a time delay that depends on the polarization of the electromagnetic wave. The equations of motion take the form

$$\nabla^\mu F_{\mu\nu} - \widehat{\alpha}_2\, R_\nu{}^{\mu\alpha\beta} \nabla_\mu F_{\alpha\beta} = 0 \tag{5.15}$$

In the shock wave background we have $R_{uiju} = \frac{1}{2}\partial_i \partial_j h$. In the high energy limit (5.15) is reduced to

$$\partial_u F_{vi} + (\delta_{ij} h + \widehat{\alpha}_2 \partial_i \partial_j h)\, \partial_v F_{vj} = 0. \tag{5.16}$$

[3] This particular point will be discussed later in more detail.

The total time delay then has the form

$$\Delta v = \left[1 + \widehat{\alpha}_2 \frac{\epsilon^i \epsilon^j \partial_{b^i} \partial_{b^j}}{\epsilon^i \epsilon^i}\right] \frac{4\Gamma(\frac{D-4}{2})}{\pi^{\frac{D-4}{2}}} \frac{G|P_u|}{b^{D-4}} \tag{5.17}$$
$$= \frac{4\Gamma(\frac{D-4}{2})}{\pi^{\frac{D-4}{2}}} \frac{G|P_u|}{b^{D-4}} \left[1 + \frac{\widehat{\alpha}_2 (D-4)(D-2)}{b^2} \left(\frac{(\epsilon.n)^2}{\epsilon.\epsilon} - \frac{1}{D-2}\right)\right]$$

where ϵ is the (real) transverse linear polarization direction of the electromagnetic wave. We also introduced $\vec{n} \equiv \frac{\vec{b}}{b}$. The derivatives in (5.17) come from the derivatives present in the Riemann tensor. The second term in (5.14) can be viewed as a spin dependent gravitational force. We see that as b becomes small, $b^2 < \widehat{\alpha}_2$, the second term in (5.17) can overwhelm the first and will lead to time advance instead of time delay. The crucial feature of this new term is that it can be positive or negative, depending on the choice of polarization. If the polarization is along $\vec{b}$, it has one sign, while if the polarization vector is orthogonal to $\vec{b}$, $\vec{\epsilon}.\vec{n} = 0$, then it has the other sign. If the polarization is a linear combination of these two, we first decompose the wave in linear polarizations along and perpendicular to $\vec{b}$ and then exponentiate (5.17) for each of these two cases separately, and then add the two results. In other words, the expression for the time delay is now a matrix that can be diagonalized by choosing the polarizations to be along $\vec{b}$ or perpendicular to $\vec{b}$. In conclusion, for either choice of the sign of $\widehat{\alpha}_2$, there is a choice of polarization that can lead to time advance. Thus, if we require the theory to be causal, we see that $\widehat{\alpha}_2$ should be set to zero. More precisely, it should be small enough so that the computation we did above breaks down for some reason. An example of a theory where the coupling in (5.14) arises at tree level is bosonic string theory [163, 164]. We will see later that in string theory the potential causality problem is fixed by the presence of extra massive states.

As another example, let us consider the Gauss-Bonnet theory. This consists of the usual gravity action plus a specific R^2 interaction of the form

$$S = \frac{1}{16\pi G} \int d^D x \sqrt{-g} \left(R + \lambda_{GB} \left[R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - 4 R_{\mu\nu} R^{\mu\nu} + R^2\right]\right). \tag{5.18}$$

The term in brackets is a topological invariant in $D = 4$, but it is not topological for $D > 4$. This theory has been extensively studied because it has the nice feature that the equations for small fluctuations around any background are second order [165] .

As explained in [159] the shock wave solution (5.2) is also an exact solution in the Gauss-Bonnet theory as well. We can consider propagation of a gravitational perturbation through the shock wave background. Before and after the shock the graviton moves as in flat space. All we need to know is what happens when it crosses the shock.

We consider a high-energy graviton δh_{ij} that propagates in the v direction with momentum p_v and traceless polarization in the transverse plane. Near the shock we approximate the equations as

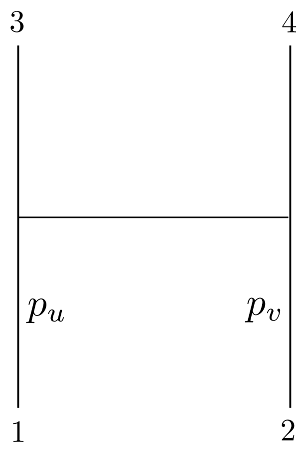
$$\partial_u \partial_v \delta h_{ij} + (\delta_{ik} + 4\lambda_{GB} \partial_i \partial_k h)\, \partial_v^2 \delta h_{kj} = 0 \tag{5.19}$$

Using (5.19) we can find the time delay which takes the following form

$$\begin{aligned}\Delta v &= [1 + 4\lambda_{GB} \frac{\epsilon^{ik}\epsilon^{jk}\partial_{b^i}\partial_{b^j}}{\epsilon.\epsilon}] \frac{4\Gamma(\frac{D-4}{2})}{\pi^{\frac{D-4}{2}}} \frac{G|P_u|}{b^{D-4}} \\ &= \frac{4\Gamma(\frac{D-4}{2})}{\pi^{\frac{D-4}{2}}} \frac{G|P_u|}{b^{D-4}} \left[1 + \frac{4\lambda_{GB}(D-4)(D-2)}{b^2} \left(\frac{(\epsilon.n)^2}{\epsilon.\epsilon} - \frac{1}{D-2} \right) \right]\end{aligned} \tag{5.20}$$

Again, by choosing different polarizations we can get time advance for $b^2 \sim |\lambda_{GB}|$ for any sign of λ_{GB}. Notice that the formula for the time delay is very similar to the ones discussed in the context of energy correlators in AdS/CFT with the parameters depending on the impact parameter of scattering b (see e.g. [50]). We will see below that in the case of AdS the causality constraint interpolates between the usual energy correlator bounds and the flat space bounds obtained above. So we see that imposing positivity in all channels exclude λ_{GB} completely unless new physics kicks in. In other words, the purely Gauss-Bonnet theory (5.18) is acausal. As an aside, the Gauss-Bonnet theory was found in [166] to violate the second law of black hole thermodynamics in certain processes.[4] We could also consider fermions coupled to gravity. At the level of the on-shell scattering amplitudes there is an additional spin-flipping structure that leads to time advance. This additional structure does not come from any known local Lagrangian and is known to be ruled out by considering consistency conditions imposed on the four-point amplitude [150] , at least, in four dimensions. If the same is true in all dimensions then the effects that we are discussing could not be observed for fermions. We leave exploration of this point for the future. Note that having an explicit local Lagrangian which leads to second order equations of motions guarantees that all consistency conditions for the four-point amplitude imposed in [150] are obeyed. Thus, none of the problems discussed in [150] arise in the case of photons or gravitons with the modified three-point functions that we are discussing.

[4] More precisely, [166] considered the compactification of the Gauss-Bonnet theory to four dimensions on a circle. Then the Gauss-Bonnet term becomes topological. This leads to a constant contribution to the black hole entropy. The sign of this constant contribution depends on the coefficient λ. When this contribution is negative, a small enough black hole can have negative entropy. When the constant is positive, a merger of two black holes can violate the second law. The constraints arise when the Schwarzschild radius $r_s^2 \sim \lambda_{GB}$. In this sense they are similar to the ones we have. In both cases one needs $\lambda_{GB} \ll l_{Planck}$ in order to have a meaningful statement. [166] has the nice feature of also constraining purely topological terms in four dimensions. On the other hand, they might be modified by higher derivative corrections to the action, which could also modify black hole thermodynamics.



$$\sum_{i=1}^{4} p_i^\mu = 0$$

$$s = -(p_1 + p_2)^2$$

$$t = -(p_1 + p_3)^2$$

Figure 5.3: Kinematics of the two-to-two scattering that we are interested in. Particle 1 has a very large momentum p_u and particle two has a very large momentum p_v.

5.3 General Constrains on the On-Shell Three-Point Functions

The examples we have discussed so far have shown that there are causality issues with specific theories. In this section we will isolate the crucial elements that produce the problem. It turns out that these causality problems are produced purely and exclusively by the form of the *on-shell* three-point functions of the theory. Therefore they are insensitive to higher order contact terms. First we will explain why the three-point functions determine the time delay. Then we will recall some of the properties of three-point functions. Finally, we will present a thought experiment that makes the causality violation more manifest.

5.3.1 The Phase Shift in Impact Parameter Representation From Three-Point Functions

Let us consider the tree-level four-point amplitude $\mathcal{A}_4$. It depends on the kinematic invariants that we can produce with the four on-shell conserved momenta and polarization tensors of the external particles. Since it is a tree-level amplitude its only singularities are poles in the s, t and u Mandelstam variables. We can now consider external momenta such that $s \gg t$, but s small enough that the theory is still weakly coupled. We can take the first incoming particle to have very large momentum along p_u and the second with large momentum p_v.

One can show that, in impact parameter space the amplitude is given by (for review see appendix

C)

$$\delta(\vec{b}, s) = \frac{1}{2s} \int \frac{d^{D-2}\vec{q}}{(2\pi)^{D-2}} e^{i\vec{q}.\vec{b}} \mathcal{A}_4(\vec{q}). \tag{5.21}$$

where $\mathcal{A}_4(\vec{q})$ is really a short-hand for the four-point amplitude evaluated in the following momentum configuration

$$\begin{aligned} p_{1\,\mu} &= \left(p_u, \frac{q^2}{16p_u}, \frac{\vec{q}}{2}\right) \quad , \qquad p_{2\,\mu} = \left(\frac{q^2}{16p_v}, p_v, -\frac{\vec{q}}{2}\right) \quad , \\ p_{3\,\mu} &= -\left(p_u, \frac{q^2}{16p_u}, -\frac{\vec{q}}{2}\right) \quad , \qquad p_{4\,\mu} = -\left(\frac{q^2}{16p_v}, p_v, \frac{\vec{q}}{2}\right) \quad , \\ s &\simeq 4p_u p_v \ , \qquad t \simeq -(\vec{q})^2 \end{aligned} \tag{5.22}$$

where in all cases we indicated only the leading order term in the t/s expansion, assuming $t/s \ll 1$. If we had scalar particles, the amplitude would depend only on the Mandelstam invariants and we can write $\mathcal{A}_4(\vec{q}) = \mathcal{A}_4(s, t = -\vec{q}^2)$. However, in the case of particles with spin, the amplitude depends also on the polarization vectors contracted with the various momenta. Let us assume that $\vec{b}$ in (5.123) is chosen along $\vec{b} = (b, 0, 0, \cdots)$, $b = |\vec{b}|$. Then let us consider the integral over the first component of $\vec{q}$, call it q_1. Due to the exponential factor in (5.123) this integrand is suppressed in if we give q_1 a positive imaginary part. Here we assumed that the amplitude does not grow exponentially. This is true if we consider particles with polynomial interactions. Setting $q_1 = i\kappa$+real, we see that the exponential in (5.123) is suppressed in this region as $e^{-\kappa b}$, $\kappa > 0$. Thus we would get a vanishing result (in the large κ limit) unless we cross poles under this contour shift. In fact, we do cross poles. For example the pole at $t = m^2$ coming from the exchange of a particle of mass m in the t-channel leads to a pole at

$$\kappa_*^2 = m^2 + (\vec{q}_{\rm rest})^2 \tag{5.23}$$

where $\vec{q}_{\rm rest}$ are the rest of the components of $\vec{q}$ except the first one. These are still real. The residue of the pole is given by a product of on-shell three-point functions. These three-point functions are non-vanishing because the intermediate momentum has one imaginary and one real component. Thus, we can think of the whole on-shell three-point function as being in two time directions. Note that this is a particular computation where the on-shell three-point function is relevant in ordinary signature. Notice that somewhat abstract notion of complex factorization in mixed signature [167, 149] spacetimes has a direct physical meaning in the context of computing scattering amplitudes in

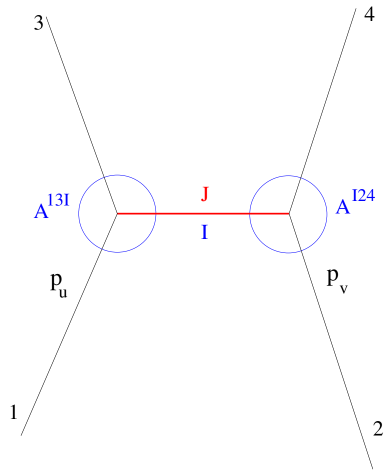


Figure 5.4: At fixed impact parameter we find a contribution to the amplitude given by the on-shell propagator in the intermediate channel, labeled by the letter I. The contribution involves a product of two on-shell intermediate amplitudes, $\mathcal{A}^{13I}$ and $\mathcal{A}^{I24}$, which are circled in the figure. Particles 1 and 3 carry a large momentum p_u and 2 and 4 have a large p_v. We get a factor of s^J from contracting p_u on the left side with p_v on the right side through the sum over polarizations of the intermediate particle.

the impact parameter representation. The pole (5.23) gives a contribution to the amplitude of the form $e^{-\kappa_* b}$. For a massive particle this gives something going like e^{-mb} for large bm, a Yukawa-like potential.[5] For a massless particle, the integral over the rest of the components of $\vec{q}$ produces an inverse power of b, $1/b^{D-4}$. In addition, factors of $\vec{q}$ which are contracted with the polarization tensors give derivatives with respect to $\vec{b}$. In a theory which only contains a massless graviton we just get the massless graviton pole. It should be noted that in impact parameter representation, for non-zero b, we only get a contribution from the diagram that contains an on-shell particle in the t-channel. In particular, s-channel exchanges do not contribute. The reason is that the two incoming particles have to actually overlap in order to give rise to the intermediate particle in the s-channel. Similarly, a four point contact interaction does not contribute for the same reasons. In both of these remarks we used that we are looking at the tree level diagrams to leading order in the weak coupling expansion. At higher orders there can be other contributions. However, since we are at weak coupling, we can ignore them. In these paragraph, we have used that the interactions are local. Any non-local effect has to come from the propagation of a physical particle in the t-channel.[6]

We are interested in the part of the amplitude that increases fastest with the Mandelstam variable s. Since there are no kinematic invariants in the on-shell three-point amplitudes, a factor of s can

[5]More precisely, after we integrate over the rest of the components of $\vec{q}$ we get $(2\pi)^{\frac{2-D}{2}} \left(\frac{m}{b}\right)^{\frac{D-4}{2}} K_{\frac{D-4}{2}}(mb)$.

[6]If our theory contains extended objects, such as strings, we can actually create particles in the s-channel by going to small enough (but non-zero) impact parameter. We discuss this below.

only come from factors of p_u or p_v that are in each of the three-point function contracted, via the sum of the polarizations of the intermediate particle, see figure (5.4). With a particle of spin J, the maximum power we can get in the amplitude is $\mathcal{A}_4 \sim s^J$. This leads to a factor of s^{J-1} in the phase shift (5.123). In fact, this sum over polarizations over the intermediate state receives a large contribution only from one specific polarization, when the intermediate state has all u indices. In other words, the polarization tensor in the intermediate state appearing in the left three-point function has the form $\epsilon^{I\, u\cdots u}$ and the one in the right three-point function is $\epsilon^{I\, v\cdots v}$. These are contracted with the corresponding factors of the large momentum to give $(p_u p_v)^J \sim s^J$. The polarization tensors of the external particles can be written as products of vectors $\epsilon_{\mu\nu} = \epsilon_\mu \epsilon_\nu$ where for particles one and three we can write

$$\epsilon^{1\,\mu} = (-\frac{\vec{q}.\vec{e}_1}{2p_u}, 0, \vec{e}_1)\ , \qquad \epsilon^{3\,\mu} = (\frac{\vec{q}.\vec{e}_3}{2p_u}, 0, \vec{e}_3) \tag{5.24}$$

where $\vec{e}_{1,3}$ are vectors in the purely transverse directions. The on-shell three-point functions will contain factors of the form

$$\epsilon^1.\epsilon^3 = \vec{e}_1.\vec{e}_3\ , \quad \epsilon^1.p^3 = \vec{q}.\vec{e}_1\ , \quad \epsilon^3.p^1 = \vec{q}.\vec{e}_3 \tag{5.25}$$

The conclusion is that we can think of the polarizations of the external particles as contained effectively in the transverse space. In addition, when we contract them with the external states we get factors of $\vec{q}$. These translate into derivatives with respect to $\partial_{\vec{b}}$.

The final result from a massless pole is

$$\begin{aligned} \delta(\vec{b}, s) &= \frac{\mathcal{A}_3^{13I}(-i\partial_{\vec{b}})\mathcal{A}_3^{I24}(-i\partial_{\vec{b}})}{2s} \int \frac{d^{D-2}\vec{q}}{(2\pi)^{D-2}} \frac{e^{i\vec{q}.\vec{b}}}{\vec{q}^2} \\ &= \frac{\Gamma(\frac{D-4}{2})}{4\pi^{\frac{D-2}{2}}} \frac{\mathcal{A}_3^{13I}(-i\partial_{\vec{b}})\mathcal{A}_3^{I24}(-i\partial_{\vec{b}})}{2s} \frac{1}{|\vec{b}|^{D-4}}. \end{aligned} \tag{5.26}$$

where the three-point functions are evaluated on three momenta (5.22) . As explained above, in each three-point function there is only one relevant polarization for the intermediate state that produces the factor of s^J. The polarizations of the external particles can be viewed as living purely in the transverse space, after we use (5.24) (5.25) . In this way we can compute $\delta(\vec{b}, s)$ using on-shell methods. For more details see appendix B. This gives the phase shift to leading order in perturbation theory. The answers agree with with the infinitesimal form of the phase shift computed using the shock wave computation. In fact, the shock wave itself can be viewed as the on-shell graviton intermediate state we discussed above. The fact that the metric contains only the h_{uu} component is related to the fact that only one polarization of the intermediate state produces the most factors of s.

5.3.2 The Possible Forms of Three-Point Functions in Various Theories

Given that the answer for the time delay depends on the on-shell three-point functions, it is useful to recall the possible three-point functions. In four dimensions we can use the usual helicity basis. The usual Einstein gravity action gives rise to the $++-$ and $--+$ three-point functions. We can also have $+++$ and $---$ structures which could come from $(\text{Riemann})^3$ terms. There are two combinations, one of them is parity violating. In higher dimensions, $D > 4$ we have three possible structures for the graviton three-point functions, all parity preserving. They have the schematic structure, writing $\epsilon_{\mu\nu} = \epsilon_\mu \epsilon_\nu$,

$$\mathcal{A}_R = (\epsilon_1.\epsilon_2\epsilon_3.p_1 + \epsilon_1.\epsilon_3\epsilon_2.p_3 + \epsilon_2.\epsilon_3\epsilon_1.p_2)^2 \tag{5.27}$$

$$\mathcal{A}_{R^2} = (\epsilon_1.\epsilon_2\epsilon_3.p_1 + \epsilon_1.\epsilon_3\epsilon_2.p_3 + \epsilon_2.\epsilon_3\epsilon_1.p_2)\epsilon_1.p_2\epsilon_2.p_3\epsilon_3.p_1 \tag{5.28}$$

$$\mathcal{A}_{R^3} = (\epsilon_1.p_2\epsilon_2.p_3\epsilon_3.p_1)^2. \tag{5.29}$$

The first is the usual one from Einstein gravity, the second can arise from the Gauss-Bonnet term (5.18) . The third one can arise from a $(\text{Riemann})^3$ term.[7] They can be viewed as products of the two possible structures that we can have for on-shell spin one particles. In a given theory, the total three-point function is given by a linear combination of these three answers

$$\mathcal{A}_{ggg} = \sqrt{32\pi G}\,[\mathcal{A}_R + \alpha_2\,\mathcal{A}_{R^2} + \alpha_4\mathcal{A}_{R^3}] \tag{5.30}$$

where α_2 and α_4 are two parameters with dimension of $(\text{length})^2$ and $(\text{length})^4$ respectively. Notice that we have an overall coupling, given by G, which we take to be parametrically smaller than the other two parameters. In the high energy limit the three-point functions appearing in (5.26) simplify

[7] $(\text{Riemann})^3 = R_{\mu\nu\sigma\delta}R^{\sigma\delta\rho\gamma}R_{\rho\gamma}{}^{\mu\nu}$. In four dimensions we can replace one of the curvatures by $R_{\mu\nu\sigma\delta} \to \widetilde{R}_{\mu\nu\sigma\delta} = \epsilon_{\mu\nu\rho\gamma}R^{\rho\gamma}{}_{\sigma\delta}$ to obtain the parity violating term. This parity violating term gives rise to the three-point vertex $\epsilon_{\mu\nu\delta\sigma}p_1^\delta p_2^\sigma \epsilon_{1,\mu\mu'}\epsilon_{2,\nu\nu'}p_2^{\mu'}p_1^{\nu'}\epsilon_{3,\gamma\rho}p_1^\rho p_2^\gamma$, which, despite appearances, is properly symmetric under exchange of any of the three particles. When this term is present the coefficient of the $+++$ amplitude is complex and the coefficient of the $---$ amplitude is the complex conjugate.

further and become

$$\mathcal{A}_{R}^{13I} = 2p_u^2(\vec{e}_1.\vec{e}_3)^2 \tag{5.31}$$

$$\mathcal{A}_{R^2}^{13I} = 2p_u^2(\vec{e}_1.\vec{e}_3)(\vec{q}.\vec{e}_1)(\vec{q}.\vec{e}_3) \tag{5.32}$$

$$\mathcal{A}_{R^3}^{13I} = 2p_u^2(\vec{q}.\vec{e}_1)^2(\vec{q}.\vec{e}_3)^2 \tag{5.33}$$

for the three terms in (5.27) . We used (5.24) . The amplitude $\mathcal{A}^{I24}$ has a similar expression with $p_u \to p_v$ and $1,3 \to 2,4$. The reader can express these in terms of transverse traceless two index tensors by using the replacement rule $e_i e_j \to e_{ij}$ in (5.31) , where i,j are indices in the $D-2$ transverse directions. Note that the third structure in (5.27) or (5.31) is not allowed in a supersymmetric theory. This can be seen as follows. In $D = 4$ this structure gives rise to $+++$ and $---$ amplitudes. However, by the methods described in [168] it is possible to show that supersymmetry implies that this structure should be set to zero. This is actually true in all dimensions. The reason is that we can start with a theory in D dimensions and consider external graviton three-point functions with four-dimensional kinematics. If we had a non-zero contribution from the third structure in D dimensions, then it would lead to a contribution to the $+++$ or $---$ amplitudes in four-dimensional kinematics. Since the arguments in [168] are purely kinematical, based on the symmetries of the theory, then they also force the amplitudes to vanish. In conclusion, supersymmetry implies that $\alpha_4 = 0$. In the heterotic string we have a non-zero α_2. And this is also true for compactifications of the string to D dimensions. Thus α_2 is compatible with half maximal supersymmetry. With maximal supersymmetry, e.g. $\mathcal{N} = 8$ in $D = 4$, we should also have $\alpha_2 = 0$. This can be understood from the fact that the four point amplitude for the full supergravity multiplet is determined up to a unique function. Thus, there is no freedom to introduce the polarization dependent terms that would arise if we had the freedom to switch on α_2. Of course, in standard maximal supergravity α_2 is not present, therefore the only value consistent with maximal supersymmetry is $\alpha_2 = 0$. As an aside, notice that this is also related to the fact that in four-dimensional $\mathcal{N} = 4$ superconformal theories, the three-point functions of the stress tensor are completely fixed by supersymmetry in terms of the two-point functions. There are two possible three-point functions between a graviton and two photons.

$$\mathcal{A}_{F^2} = \epsilon_{\mu\nu}\left[p_1^\mu p_3^\nu(\epsilon_1.\epsilon_3) - \epsilon_1^\mu p_3^\nu(\epsilon_3.p_1) - \epsilon_3^\mu p_1^\nu(\epsilon_1.p_3)\right], \tag{5.34}$$

$$\mathcal{A}_{RFF} = \epsilon_{\mu\nu}p_1^\mu p_3^\nu(\epsilon_1.p_3)(\epsilon_3.p_1), \tag{5.35}$$

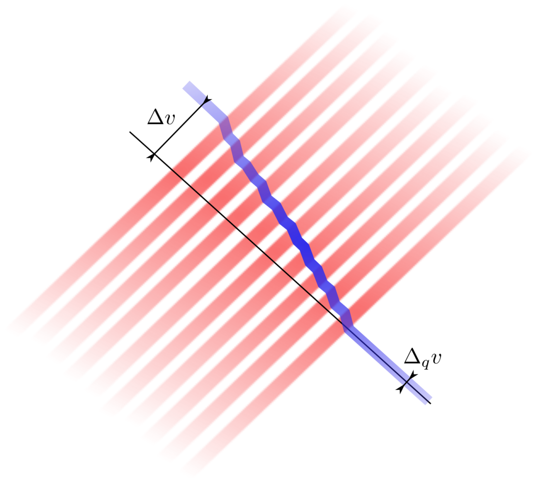


Figure 5.5: We imagine a particle going through a set of successive scattering events. The intrinsic quantum uncertainty in v is $\Delta_q v$. We have drawn a situation where there is a final time advance after going through all the shocks that is larger than the quantum uncertainty. In this figure we have neglected the delay of the u-localized particles.

so that $\mathcal{A}_{g\gamma\gamma} = \sqrt{32\pi G}\,[\mathcal{A}_{F^2} + \widehat{\alpha}_2\,\mathcal{A}_{RFF}]$. One arising from pure gravity and the other from the second term in (5.14) . Again this second structure is forbidden in a supersymmetric theory. As we did in (5.31) , in the high energy limit we can write them as

$$\mathcal{A}^{13I}_{F^2} = 2p_u^2(\vec{e}_1.\vec{e}_3), \tag{5.36}$$

$$\mathcal{A}^{13I}_{RFF} = 2p_u^2(\vec{q}.\vec{e}_1)(\vec{q}.\vec{e}_3). \tag{5.37}$$

We emphasize that we work here with the on-shell three-point functions, independently of the precise way we write the Lagrangian. This discussion depends only on on-shell three-point functions and not on other contact terms. Any contact four point interaction does not give rise (at tree level) to the long range force at a non-zero value of the impact parameter.

5.3.3 Problems with Higher Derivative Corrections to the Three-Point Functions

Above we discussed how the three-point functions lead to the leading order expression for the phase shift $\delta(\vec{b}, s) = sF(\vec{b})$. If this result were exponentiated, as $e^{i\delta}$, then we could get a time advance problem similar to what we found for the shock waves. Here we would like to explain how to get a time advance problem without using the particular non-linear structure of shock waves. The goal is to present the problem in a way that depends only on very general principles. First note that in order for time delay to be a problem we would like to find that the time delay $\Delta v = \partial_{p_{2,v}}\delta$ is larger than the quantum mechanical uncertainty that is implicit in the definition of the wave packet for a particle of of momentum $p_{2,v}$. This uncertainty is of the order of $\Delta_q v \sim 1/p_{2,v}$. Thus the figure of merit is

$$\frac{\Delta v}{\Delta_q v} = p_{2,v}\partial_{p_{2,v}}\delta = \delta \tag{5.38}$$

here we used that δ is linear in s, and therefore linear in $p_{2,v}$ ($s \propto p_{1,u}p_{2,v}$). Thus in order to see a problem, we expect that δ should be greater than one. On the other hand, the validity of perturbation theory suggests that δ should be much less than one. In order to amplify the effect, we imagine that particle number two undergoes N successive instances of particle number one, see (5.5). Through each instance it gets a small phase shift which leads to a factor in the out state of the form $(1 + i\delta)$ with small δ. If we repeat this N times we get

$$(1 + i\delta)^N \sim e^{iN\delta} \qquad \delta \ll 1 \ , \quad N \gg 1. \tag{5.39}$$

This is the total phase shift, and as explained in (5.38), we want $N\delta$ to be of order one. This can be achieved by taking $N \sim 1/\delta$. In addition, we would like to make sure that the approximations that we used remain valid. In particular, we have said that particle 2 remains localized at some distance b through the whole process. We can choose light-cone coordinates for the evolution of particle two so that u is time, then we have a non-relativistic problem with mass $m \sim p_v$. The spreading of the wavefunction during the time U that the whole process takes is $\Delta b \sim \sqrt{U/p_v}$. The time U that the process takes is determined as follows. We want to separate the N instances of the scattering process from each other so that we can view them as independent and we can use (5.39) . The best we can localize each of the N particles of type 1 is by an amount $\Delta_q u \sim 1/p_u$. Thus $U = N/p_u$ and then $\Delta b \sim \sqrt{\frac{N}{s}}$. We want $\Delta b \ll b$. This translates into the condition

$$N \ll sb^2 \quad \longrightarrow \quad \frac{1}{sb^2} \ll \delta \tag{5.40}$$

where we used that $N \sim 1/\delta$ in order to have a problem. We see that the simultaneous validity of (5.40) and (5.39) can be achieved if $sb^2 \gg 1$ and if s can also be increased so as to achieve

$$\frac{1}{sb^2} \ll \delta \ll 1 \tag{5.41}$$

which can be done if δ grows with s. An additional issue is that we wanted to neglect the deflection angle. Then we can replace each of the 1 particles by a pair of particles localized in the transverse dimension at the origin and at $2\vec{b}$, with particle number 2 still localized at $\vec{b}$ in the transverse plane, as shown in (5.2) . Notice that if we had a $g\phi^3$ vertex, then the exchange of the spin zero particle leads to $\delta \sim \frac{1}{s}\frac{g^2}{b^{D-4}}$. In this case we cannot obey the conditions (5.41) to obtain a possible causality problem. In fact, (5.40) implies that $g^2/b^{D-6} > 1$, which means that we are at strong coupling. Thus we see that a crucial feature that we used is that the phase shift increases as a function of s. In the case that we exchange a spin one field, δ is independent of s and there is also no causality issue. The reason is simple, the spin two field is effectively changing the metric and causal structure while the spin zero or one fields are not.

The conclusion is that we have justified the exponentiation (5.39) and thus the derivation of the time delay problem in a way that does not depend on the details of the shock wave solution. Let us emphasize that the preceding shock wave discussion was motivational, but the thought experiment that we have set up in this section does not require us to rely on the non-linear structure of the shock wave. It was all derived from the on-shell three-point functions plus certain assumptions about the locality of the theory that allowed us to view each shock as an independent event. In fact, there are some cases where the shock wave computation does not give the right answer. For example, consider the pure gravity case, where $\delta = Gs/b^{D-4}$. If the energy is large enough to form a black hole, then the time delay computed from the shock wave is not the correct description for the physics. We expect to form a black hole when the center-of-mass energy is such that the associated Schwarzschild radius $r_s^{D-3} = \sqrt{s}G$ is larger than b [169, 170, 171]. It is easy to check that this never happens if (5.41) is obeyed. It is not obeyed even if we consider the energy of the N particles that we used for the argument (with $N \sim 1/\delta$). There is still one more complication that we need to deal with in order to make the argument clearer. In the shock wave discussion of section two, the spin of the particle creating the shock did not matter. However, with the modified three-point functions, the spins of the scattered particles can change. The full interaction is a spin dependent force which acts of both spins. It acts on both the spin of the left and the right moving particles (particle one and two, see (5.3)). I will be necessary for us to be able to fix the polarizations of particles one and three, see (5.3) . This can be achieved by replacing particle one, by a coherent state of particles. In this case, due to the usual Bose enhancement factor, particle three will have a larger probability of remaining with the same polarization. In this set up we can set the spins, or polarization vectors, of particles one and three to be the same. Or, more precisely, $\epsilon_3 = \epsilon_1^*$. Since we are at weak coupling we can consider a coherent state with a mean occupation number which is large enough for us to be able to neglect the spin flips but small enough that the total scattering amplitude is still small. In other words, the use of coherent states allows us to effectively select the final state for particle 3. More explicitly, say that we form a coherent state for the oscillator mode created by $a^\dagger$, $e^{\lambda a^\dagger}|0\rangle$. We could then have terms in the interaction Hamiltonian that leave this oscillator the same $H_{1,int} = h_1 a^\dagger a$ or that mix it with a second oscillator $H_{2,int} = h_2 b^\dagger a + h.c.$. Here h_1 and h_2 can act on other degrees of freedom. Then, for a coherent state, the matrix elements where there is no change are enhanced, due to the usual Bose enhancement factors, relative to the ones where there is a change

$$\langle 0|e^{\overline{\lambda}a}\, H_{1,int}\, e^{\lambda a^\dagger}|0\rangle \propto |\lambda|^2 \langle h_1\rangle \tag{5.42}$$

$$\langle 0|e^{\overline{\lambda}a} b\, H_{2,int}\, e^{\lambda a^\dagger}|0\rangle \propto \lambda \langle h_2\rangle \tag{5.43}$$

We see that for large λ the first term is enhanced. We have not indicated explicitly the initial and final states on which h_1 and h_2 act, since they involve other oscillators. Since we are at weak coupling, we can choose λ large enough so that the first term dominates relative to the second term, while still the whole process is in the weakly coupled approximation, or the total effect of the interaction Hamiltonian is small. Of course an alternative way to say this is that we are creating a classical background with the particles of type 1 in (5.3) the terms that have a non-zero expectation value in this classical background dominate over the others. We want a classical background with small enough amplitude that we can still trust the leading order perturbative expansion of the interaction Hamiltonian. The use of coherent states also allows us to select final states for particle 3 in (5.3) with a a small momentum p_v. Namely, we form a coherent state out of a superposition of particles with large momenta $p_{1,\mu}$. We need a superposition since we need to localize this particles within the transverse plane to a location smaller than b. Thus we have some dispersion in the transverse momenta $\vec{q}$. With a large p_u component of the momentum, we then get a small momentum along the $p_v = \frac{\vec{q}^2}{4p_u}$ direction. Since we have a coherent state, the particle 3 is also taken out of this superposition and has the same range of values for the momentum. Therefore the total momentum transfer in the t-channel along the $v$direction is very small. Then the kinematics chosen in (5.22) is representative for the process in question. This still allows a possibly large amount of momentum transfer along the u direction. We will discuss this in subsection 4.3 . Another minor point, is that the phase shift δ represents a time delay for both particles, it affects both particle 1 and particle 2. So far we have been focusing only on the effects on particle 2. These coherent states also allow us to effectively select a final state for particles 3, so that we can focus more clearly on the time delay with which particle four emerges, see (5.3) .

5.3.4 Scattering of Gravitons in $D > 4$ Dimensions

It is also easy to argue that α_2 and α_4 structures lead to causality problem in $D > 4$. To show this we consider the probe graviton that scatters off the coherent state. The phase shift takes in this case the following form (see appendix B for details)

$$\delta \sim Gs(e_1^{ij}e_3^{ij} + \alpha_2 e_1^{ij}e_3^{ik}\partial_{b^j}\partial_{b^k} + \alpha_4 e_1^{ij}e_3^{kl}\partial_{b^i}\partial_{b^j}\partial_{b^k}\partial_{b^l})\times \tag{5.44}$$
$$(e_2^{ij}e_4^{ij} + \alpha_2 e_2^{ij}e_4^{ik}\partial_{b^j}\partial_{b^k} + \alpha_4 e_2^{ij}e_4^{kl}\partial_{b^i}\partial_{b^j}\partial_{b^k}\partial_{b^l})\frac{1}{|\vec{b}|^{D-4}}.$$

In order to find problems we will be choosing various polarizations for the particles. For example, let us firt consider particle 1 with polarization $e^1_{xx} = -e^1_{yy}$ and the other components equal to zero. Here x, y represent to two directions in the transverse plane. We call this the $\oplus$ polarization. We also choose $e^3 = e^1$. We enforce this by sending a coherent state with this polarization. Now for particle 2 we can choose the same polarization or the crossed polarization, called $\otimes$, given by $e^2_{xy} = e^2_{yx} = 1/\sqrt{2}$ and all other components equal to zero. Then we find the following. If $\vec{b}$ is along the $\widehat{x}$ direction, then these two different polarizations for particle two do not mix as they go through the shock. They diagonalize the phase shift matrix. For small enough b the α_4^2 terms dominate since they are the most singular in the small b expansion. These terms have the form

$$\delta_{\oplus\oplus} \sim Gs\alpha_4^2 O_\oplus O_\oplus \frac{1}{b^{D-4}} = \frac{Gs\alpha_4^2}{b^{D+4}}(\text{positive}) \tag{5.45}$$
$$\delta_{\oplus\otimes} \sim Gs\alpha_4^2 O_\oplus O_\otimes \frac{1}{b^{D-4}} = \frac{Gs\alpha_4^2}{b^{D+4}}(\text{negative}) \tag{5.46}$$
$$O_\oplus = (\partial_{b_x}^2 - \partial_{b_y}^2)^2\ , \qquad O_\otimes = 4\partial_{b_x}^2\partial_{b_y}^2 \tag{5.47}$$

where we take the derivatives first and then set $\vec{b} = (b_x, 0, \cdots, 0)$. Where the terms in parenthesis are polynomials in D which are positive or negative definite.[8] Notice that the positivity of the first case is due to the following argument. If the polarizations of particle 2 and 4 are the same as those of 1 and 3, then the configuration is constrained by unitarity along the t-channel.[9] Therefore in this case we should get a strictly positive answer for the time delay for the contribution of any particle with a non-zero coupling. Since we obtained a negative time delay for the second case in (5.54) , we conclude that α_4 should be set to zero unless new particles are present. Once α_4 has been set to zero, we can discuss α_2. In that case we can still choose the $\otimes_{xy}$ polarization for particles 1 and 3

[8] For $D > 4$, (positive)= $(D-4)(D^2-4)D(336 + 128D + 20D^2 + 4D^3 + D^4)$ and (negative)= $-4(D-4)(D^2-4)D(D+3)(20 + 6D + D^2)$. For $D = 4$ the derivatives act on $-\log b$ and we get (positive) = -(negative) =80640.

[9] t-channel unitarity is the following statement. When the polarizations of 2 and 4 are related by conjugation and reflection along the $\vec{b}$ axis to the polarizations of 1 and 3 respectively, then the residue of the t-channel pole is positive.

and the $\otimes_{yz}$ polarizations for particles 2 and four. We then focus on the terms proportional to α_2^2 since they are the dominant terms at small b (once we have set $\alpha_4 = 0$). We then get

$$\delta_{\otimes_{xy},\otimes_{yz}} \sim G s \alpha_2^2 \widehat{O}_{xy} \widehat{O}_{yz} \frac{1}{b^{D-4}} = -\frac{G s \alpha_2^2}{b^D} 2(D-4)(D-3)(D-2) \tag{5.48}$$

$$\widehat{O}_{xy} = -\partial_{b_x}^2 - \partial_{b_y}^2 \, , \qquad \widehat{O}_{yz} = -\partial_{b_y}^2 - \partial_{b_z}^2 \tag{5.49}$$

Then we conclude that α_2 should also be set to zero unless new massive particles appear.

5.3.5 Scattering of Gravitons in Four Dimensions

Let us now discuss in more detail the four-dimensional case, $D = 4$, which is a bit special. First, need to take into account that the Einstein term produces a $\log(L/b)$ time delay, where L is an IR cutoff. Second we need to take into account the parity violating structure. The logarithm can be taken into account by modifying the causality criterion in the form suggested by Gao and Wald, who define it by comparing to the behavior of the same metric far away. In this way the $\log L$ term is eliminated and it is easy for a power law behavior produced by α_4, which goes as $1/b^4$ to overwhelm the logarithm. Also we will later repeat the computations for AdS_4 space and we will see that $L \to R_{AdS_4}$. In conclusion, this is not a real issue.

Note that in four dimensions the α_2 structure is identically zero. This is related to the fact that the Gauss-Bonnet term becomes topological in four dimensions. Thus, we have only the Einstein term structure and the α_4 one. With four-dimensional kinematics, we can consider the situation with coherent states of particles of type 1. Let us choose the spin of these particles to be plus. Due to the coherent state considerations, the spin of particle 3 also needs to be plus (in the outgoing notation, or minus in the incoming notation). In other words the amplitude does not have a spin flip. In fact, without a spin flip the α_4 structure does not contribute in four dimensions. Thus in the vertex involving particles one, three, and the intermediate one, the only structure that contributes is the Einstein one. This contribution is effectively the same as the spin zero one. Then we can run an argument similar to the one above. Let us now discuss the parity violating structure, together with the parity preserving one. By considering the coherent state for particle one (see (5.3)), with definite helicity (positive or negative) we get that only the Einstein structure contributes to the $\mathcal{A}^{13I}$ three-point amplitude (see (5.4)). The we get the following matrix form for the phase shift for particle two

$$
\begin{aligned}
\delta =& 2Gs\left(\mathbf{1} \ + \ \gamma\partial_\beta^4|-\rangle\langle+| \ + \ \gamma^*\partial_{\beta^*}^4|+\rangle\langle-|\right)(-\log|\beta|) \\
=& 2Gs\left(-\mathbf{1}\log|\beta| \ + \ \frac{3\gamma}{\beta^4}|-\rangle\langle+| \ + \ \frac{3\gamma^*}{{\beta^*}^4}|+\rangle\langle-|\right)
\end{aligned} \tag{5.50}
$$

where we introduced a complex variable $\beta = b_1 + ib_2$ in the two dimensional transverse space. We also used $\partial_\beta = \frac{1}{2}(\partial_{b^1} - i\partial_{b^2})$ and $e_\pm \propto e_x \mp ie_y$. Here γ is the coefficient of the $+++$ amplitude and γ^* the coefficient of the $---$ one. Notice that these physically imply that the particle two undergoes a spin flip. We see that δ is a two by two matrix in the space of helicities. In (5.50) , $\mathbf{1}$ represents the identity matrix in this two dimensional space. The matrix in (5.50) can be diagonalized by choosing the polarization directions

$$
p_{1,2} \propto |+\rangle \pm \sqrt{\frac{\gamma\,{\beta^*}^4}{\gamma^*\,\beta^4}}|-\rangle \tag{5.51}
$$

Then we find that

$$
\delta_{1,2} = 2Gs(-\log|\beta| \pm 3\frac{|\gamma|}{|\beta|^4}) \tag{5.52}
$$

Thus we see that we have a causality problem at small enough $b = |\beta|$.

5.4 Fixing the Causality Problem by Adding Massive Particles

Let us discuss how to evade the causality problem that we found above. This problem can be evaded by adding new particles at the scale α_2 or α_4. We will discuss the case of a weakly coupled theory where the problem should be fixed at tree level. This is indeed what happens in string theory, see appendix D. For a case involving loops see appendix C. Let us first consider the corrections to the time delay due to new particles being exchanged in the t-channel. The new particles have to lead to a phase shift growing like s, or a higher power of s. *Thus, we should add massive particles with spin $J \geq 2$.*

We will now argue that massive spin two particles do not help and that we need particles of higher spin. In particular, this will then rule out a solution involving Kaluza-Klein gravity, which would be a special example of the addition of massive spin two particles. For this reason we will

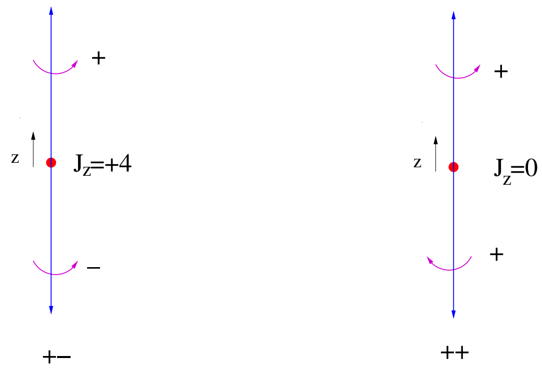


Figure 5.6: Consider the coupling of a massive spin two particle to two massless gravitons. Let us choose the kinematic configuration so that the massive particle decays into two massless gravitons along the $\hat{z}$ axis. The $+-$ helicity configuration is impossible since the angular momentum along the z axis would be $+4$. The $++$ configuration is allowed.

analyze it in detail.

5.4.1 Massive Spin Two Particles Do Not Fix the Problem in $D = 4$.

Let us first discuss the four-dimensional case. Since the external states are massless spin two particles, the on-shell three-point vertices involve two massless particles and a massive spin two particle.

In four dimensions, we can label the massless particles by their helicities. An important result is that, in all incoming notation, the only non-zero amplitudes involve $++$ or $--$ helicities for the massless particles. In particular the $+-$ combinations are zero. The argument is essentially the same as in the Weinberg Witten theorem [172] , or the statement that gravity does not have a local stress tensor operator.[10] Imagine that we have the massive spin two particle in its rest frame. We let it decay into two massless spin two particles. Let us suppose that the two decay products move in opposite directions along the $\hat{z}$ axis, see (5.6) . In the $+-$ configuration the total sum of the spins of the decay products along the $\hat{z}$ axis is $+4$ or -4. However, the initial massive particle had spin at most ± 2. Therefore a $+-$ configuration is impossible. With a $++$ or $--$ configuration there is not problem because the sum of the spins is zero. One can also write down explicitly the corresponding three-point amplitude

$$\begin{aligned} &\widetilde{\alpha}_4 \epsilon^I_{\mu\nu} p_1^\nu p_3^\mu \left[(\epsilon_1.p_3)(\epsilon_3.p_1) - (\epsilon_1.\epsilon_3)(p_1.p_3)\right]^2 \\ &\to 2\widetilde{\alpha}_4 p_u^2 \left[(\vec{e}_1.\vec{q})(\vec{e}_3.\vec{q}) + \frac{m_I^2}{2}(\vec{e}_1.\vec{e}_3)\right]^2 \end{aligned} \tag{5.53}$$

[10] The matrix elements of the stress tensor operator between two on-shell graviton states is like the coupling to a massive spin two particle, where the square of the momentum of the stress tensor, q^2 corresponds to the mass of the massive spin two particle.

where $\epsilon^1_{\mu\nu} = \epsilon^1_\mu \epsilon^1_\nu$, and we used that the component of $\epsilon^{I\mu\nu}$ that contributes the largest factor of s in the sum over intermediate states is $\epsilon^{Iuu} = 2$. We have denoted the coupling by $\widetilde{\alpha}_4$ since it reduces to the α_4 structure in the massless limit. Here 1 and 3 are the massless particles. Of course, $p_1.p_3$ is given by the mass of the massive particle. We see that this result is invariant under $\epsilon^1 \to \epsilon^1 + p^1$, and so on. In the second line of (5.53) we have written the three-point amplitude including the leading terms in the high energy limit. This is written in terms of the purely transverse polarization vectors (or tensors) introduced in (5.24) . In $D = 4$ there is also a parity violating structure which we will not need to write explicitly. If we now consider particles 1 and 3 in (5.3) to be associated to a coherent state with definite spin, then we have no spin flip allowed and this coherent state does not couple to the massive spin two particles. Therefore in four dimensions the massive spin two particles cannot solve the problem, they simply do not couple to the type of source that we are considering. Note that it is important that the massless intermediate gravitons are still coupling to the 1-3 coherent state through the Einstein three-point function, and as discussed in section 3.5, it leads to a causality problem for particle two in (5.3) . We can further show that the massive spin two particle with a coupling (5.53) by itself also leads to a causality problem and should therefore not be present. In fact, it will be useful for our later argument to understand this in more detail. For simplicity let us set to zero the parity violating massive structure. For the coherent state that involves particles one and three in (5.3) we choose the $\oplus$ polarization with $e^1_{xx} = -e^1_{yy}$ and the other components equal to zero, and the same for e^3. Here x, y represent the two directions in the transverse plane. Now for particle 2 we can choose the same polarization or the crossed polarization, called $\otimes$, given by $e^2_{xy} = e^2_{yx} = 1/\sqrt{2}$ and all other components equal to zero. Then we find the following. If $\vec{b}$ is along the $\widehat{x}$ direction, then these two different polarizations for particle two do not mix as they go through the shock. They diagonalize the phase shift matrix. If the polarizations of particle 2 and 4 are the conjugate to those of 1 and 3, and reflected along $\vec{b}$, then the configuration is constrained by unitarity along the t-channel to give a strictly positive answer for the contribution to the time delay of any particle with a non-zero coupling. On the other hand, if we average over all polarizations for particle 2, it is possible to see that the terms involving α_4 or $\widetilde{\alpha}_4$ (the massive particle contributions) all vanish. Thus, the contribution from the crossed polarization has to have the opposite sign. In other words, unitarity fixes a plus sign for the time delay for one polarization and this implies a negative sign for the other. Indeed, it is possible to see this explicitly by computing the massive particle contribution to both answers, which are

$$\delta_{\oplus,\oplus} =4Gs(\sum_m \widetilde{\alpha}_4^2 O_m O_m)K_0(mb) \tag{5.54}$$

$$\delta_{\oplus,\otimes} =-4Gs(\sum_m \widetilde{\alpha}_4^2 O_m O_m)K_0(mb) \tag{5.55}$$

$$O_m \equiv\partial_{b_x}^2\partial_{b_y}^2 - \frac{m^4}{8} \tag{5.56}$$

where the first subindex of δ is the polarization of particles 1 and 3 and the second that of particles 2 and 4 in (5.3) . By acting with this operator explicitly one can see that it gives a negative answer in the second case. This is independent of the sign of $\widetilde{\alpha}_4$. In fact, it is also negative for the contribution of the massless case when we have the α_4 structure on both sides. The full phase shift also has the general relativity contribution. Once we have a single massive particle, it is possible to go to a small enough b so that we overwhelm the positive contributions from the General Relativity vertices. This shows that in a theory with up to spin two particles we cannot solve the causality problem that arises when α_4 is nonzero. In addition, we see any massive spin two particles, even if present, they should have $\widetilde{\alpha}_4 = 0$ in order not to cause further causality problems.

5.4.2 Massive Spin Two Particles Do Not Fix the Problem in $D > 4$.

Now we now move on to a higher dimensional gravity theory, $D > 4$. The three-point amplitudes for two gravitons and a massive spin two particle now have two possible structures, first the one in (5.53) , which can be multiplied by a coefficient which we will still call $\widetilde{\alpha}_4$. And a second one of the form

$$\begin{aligned}
&\widetilde{\alpha}_2\epsilon^I_{\mu\nu}\left[\epsilon_1^\mu p_3^\nu(\epsilon_3.p_1) + \epsilon_3^\mu p_1^\nu(\epsilon_1.p_3) - p_1^\mu p_3^\nu(\epsilon_1.\epsilon_3) - \epsilon_1^\mu\epsilon_3^\nu(p_1.p_3)\right]\times \\
&\qquad\qquad [(\epsilon_1.p_3)(\epsilon_3.p_1) - (\epsilon_1.\epsilon_3)(p_1.p_3)] \\
&\to 2\widetilde{\alpha}_2 p_u^2\left[e^1_{ki}q^i e^3_{kj}q^j + \frac{m^2}{2}e^1_{ij}e^3_{ij}\right]
\end{aligned} \tag{5.57}$$

where again $\epsilon^I_{\mu\nu}$ is the intermediate state polarization vector and we used that we only care about its $\epsilon^{Iuu} = 2$ component. We have introduced a new coefficient $\widetilde{\alpha}_2$. In the second line we have indicated the form that it takes in the high energy limit. In the last line the polarization tensors are purely in the transverse directions and q is the momentum transfer. We can first consider a setup with four-dimensional kinematics. Namely, we can consider particles 1 and 3 to be associated to a coherent state which is uniformly distributed along $D - 4$ of the original dimensions. In this case the problem is essentially four-dimensional and the three-point amplitudes involving α_2 and $\widetilde{\alpha}_2$ (both massless and massive) do not contribute. If we want to avoid causality problems, and without spin > 2 particles, we conclude that both α_4 and $\widetilde{\alpha}_4$ should be zero. The argument is the same as the one we presented in the four-dimensional discussion. Note that since we are getting to four dimensions by effectively dimensionally reducing the higher dimensional theory, then the parity violating four-dimensional structure does not arise. We would now like to rule out the possibility of having contributions with non-zero α_2. For this we will assume that α_4 and $\widetilde{\alpha}_4$ are zero, as shown by the previous argument. Let us first consider the case $D = 5$. Now we have three transverse directions, let us call them $x,\ y,\ z$. We choose the polarizations of 1 and 3 to be of the $\oplus_{xy}$ type. Inserting this into (5.57) we see that this produces a factor of $O_{xy} = m^2 - \partial^2_{b_x} - \partial^2_{b_y}$ acting on the massive propagator, which is simply $\frac{1}{b}e^{-mb}$. For the particles 2 and 4 we choose the polarization $\otimes_{yz}$, which also produces a similar operator O_{yz}. The final result has the form

$$
\begin{aligned}
\delta_{\oplus_{xy},\otimes_{yz}} &\sim Gs\left(\sum_m (\widetilde{\alpha}_2)^2 O_{xy}O_{yz}\frac{1}{b}e^{-mb}\right) \\
&= -Gs\sum_m (\widetilde{\alpha}_2)^2\frac{(b^3m^3 + 5b^2m^2 + 12bm + 12)}{b^5}e^{-mb}
\end{aligned}
\tag{5.58}
$$

where we have set $\vec{b} = (b, 0, 0)$ after taking the derivatives. We see that for any sign of $\widetilde{\alpha}_2$ this produces a negative result. Furthermore, in the massless case, $m = 0$, this also gives the part of the graviton contribution proportional to α_2^2. The graviton also contains other contributions involving the ordinary Einstein piece on both three-point functions, as well as a mixed term. These contributions behave like $1/b$ and α_2/b^3 respectively. Thus, for small b, the graviton contribution involving α_2^2 dominates, since it goes as $\alpha_2^2 b^{-5}$. Therefore, we conclude that if α_2 is non-zero, then by going to small enough b we get a causality problem. Furthermore, this problem cannot be fixed by adding massive spin two particles. For $D > 5$ one can run a similar argument. But of course, we could also set up the problem with five-dimensional kinematics. In other words, we choose a coherent state spread over $D-5$ of the dimensions and we get the same as what we discussed above. The final conclusion is that if we have extra structures in the graviton three-point function (if α_2 or α_4 are nonzero), they lead to a causality problem *which cannot be fixed by adding massive particles with spins* $J \leq 2$.

With a risk of being repetitive, let us summarize the argument that rules out massive spin two particles. First we go to four-dimensional kinematics where the massless or massive couplings proportional to α_2 or α_2 do not contribute and we rule out both α_4 and the similar coupling $\widetilde{\alpha}_4$

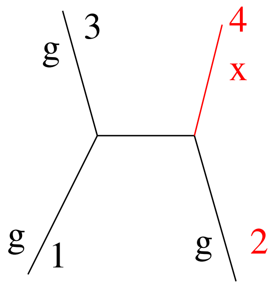


Figure 5.7: When the particles scatter, the graviton can become another massive particle, here labeled by X.

to massive intermediate gravitons. Then we go to five (or higher) dimensional kinematics and we rule out both α_2 and $\widetilde{\alpha}_2$. In particular, notice that, even in the case of ordinary Einstein gravity, with $\alpha_2 = \alpha_4 = 0$, we have rule out tree level couplings to massive gravitons (or massive spin two particles).

5.4.3 Exciting the Graviton Into New Particles

In the above discussion we ignored the possibility of exciting a graviton when it passes through the shock and transforming it into a new state.[11] In this section we discuss this possibility and conclude that it cannot fix the problem. Since the available energy is large, compared to $1/b$, it is possible to turn the incoming graviton into an outgoing massive particle, let us call it X. If we use coherent states for particles 1 and 3 in (5.3) , then we suppress the processes where particle 3 becomes a new massive particle and we only have to worry about the possibility of particle 2 turning into this new massive particle. This can happen even if the mass of the new particle, call it X, is much larger than b^{-1}, but smaller than $\sqrt{s}$. The reason is that there can be some p_u energy transfer from particles 1 and 3 to particle 4 in (5.3) . Let us view the process of the graviton 2 passing through the shock as a signal transmission problem. Focusing on the v dependence of the signal we can say that the out-signal, given an in-signal must be causal. Namely, if the in-signal vanishes for $v < 0$, then the out-signal must vanish for $v < 0$. In Fourier space these signals are related by $f_{out}(\omega) = S(\omega) f_{in}(\omega)$, with $\omega = -p_v$. Here we are using that particles 1 and 3 carry a negligible amount of p_v, since we are using $v-$translation invariance. Of course, if we have physical particles, we cannot localize them sharply in v because they only have positive frequencies. In order to obtain a sharp causality bound we need to invoke the vanishing of the field commutators, $[\phi_{out}(v), \partial_v \phi_{in}(v')] = 0$ for $v < v'$ (we put the derivative to remove possible zero mode issues).[12] As reviewed in appendix D , causality

[11] In string theory these are called tidal excitations of the string.

[12] In a theory of gravity we do not have local field operators. However, we can imagine defining such operators in the asymptotic past and future. More precisely, in order to run into the causality problems we need to put them far

implies that $S(\omega)$ is analytic in the upper half plane. In addition, the fact that we can produce other particles can only make the strength of the graviton signal in the future smaller. This, in turn, implies that the S matrix element for graviton going into graviton, call it $S_{gg}(\omega)$ should be be smaller than one in the upper half plane,

$$|S_{gg}(\omega)| \leq 1 \ , \qquad \text{for} \qquad \text{Im}(\omega) > 0 \tag{5.59}$$

See appendix D for further discussion. In situations where we have some time advance for the graviton, we are getting an infinitesimal matrix element of the form

$$S_{gg} = 1 - i\Delta v p_v = 1 + i\Delta v\omega \tag{5.60}$$

with $\Delta v < 0$. Then if we set $\omega \to i\gamma$, with $\gamma > 0$ we get $S_{gg} = 1 - \gamma\Delta v$ which is bigger than one in the upper half plane. Note that we do not need to go to very large values of γ to obtain a violation, we only need p_v or γ, to be large enough so that this impact parameter description is good enough. In this presentation of the argument, it is clear that adding extra particles as possible extra final states does not help. We need to modify S_{gg}. In other words, the transformation of the graviton to X is irrelevant for this argument because we are considering the graviton-graviton S matrix element. The transformation to X and then back to the graviton can contribute to this matrix element at higher orders in G. But, this cannot fix the problem we run into (5.60), which is of first order in G.

5.4.4 Massive Higher Spin Particles Can Solve the Problem

Now we consider the exchange of a massive spin $J > 2$ particle in the t-channel. Its contribution to the phase shift will rise with energy like Gs^{J-1} and at high energies it will dominate over the graviton contribution. This can happen even in the regime that the theory is weakly coupled. If we have a single contribution of this type, we also run into a problem. The problem is the following. We can think of the propagation of the particle number 2 as signal transmission problem where time is v. In other words, we start with a signal $f_{in}(v)$ which vanishes for $v < 0$, then the out-signal $f_{out}(v)$ should be zero for $v < 0$. In addition we want that the total norm of the out wave should not grow. From these two conditions we can deduce that the $S(\omega)$ matrix as a function of the "energy", $\omega = -p_v$, should be analytic in the upper half complex ω-plane and, in addition, it should

enough to the past and future that we can neglect the change in the spacetime metric but close enough so that the quantum mechanically dictated momentum p_u of the signal particles does not wash out the possible time advance. Here we will assume that it is possible to do this.

be bounded $|S(\omega)| \leq 1$ in the upper half plane. See appendix D for a review of these properties. However, a particle of spin $J > 2$ leads to a contribution $S \propto 1 + iGs^{J-1} + \cdots$ which becomes bigger than one in some regions of the upper half complex s plane. Notice that the problem arises already at weak coupling, for a small value for Gs^{J-1}.

Thus, a *finite* number of higher spin particles does not fix the problem. In fact, it generates problems of its own. On the other hand, an *infinite* number of particles with higher spin can solve the problem. An example is string theory. This problem has been discussed extensively in the classic papers by Amati, Ciafaloni and Veneziano [173, 174, 175, 176, 177] (see also [178]). In fact, the amplitude Reggeizes

$$\delta \propto \frac{Gs}{t} s^{\frac{\alpha'}{2}t} e^{-i\pi\frac{\alpha'}{4}t} \tag{5.61}$$

for large s small t ($s\alpha' \gg 1$, $t\alpha' \ll 1$). This expression has a cut in the s-channel, due to the creation of physical states along the s-channel. These are simply the massive closed string that are present along the s-channel. For spacelike t, $t < 0$, we see that this effectively leads to a phase shift that decreases faster than s at large s. Taking (5.61) and transforming to the impact parameter representation we find that for $b^2 \ll \alpha' \log s$ we get a behavior [175]

$$\delta \sim (\text{Pol}) \left[\frac{Gs}{(D-4)(\log s)^{\frac{D-4}{2}}} \left(1 - \frac{b^2}{\log s} + \cdots \right) + i\pi^2 \frac{Gs}{(\log s)^{\frac{D-4}{2}}} + \cdots \right] \tag{5.62}$$

where $\text{Pol} = 1 + \alpha_2 \epsilon.\partial_b \epsilon.\partial_b + \cdots$ is the part coming from the polarization tensors, which includes the new structures in the three-point functions.[13] This is indeed compatible with causality. We also get a large imaginary part that is reflecting the fact that we are creating strings along the s-channel. Notice that we had argued before that in a local theory we expect that by going to impact parameter space we can suppress tree level s-channel processes. This is not true in string theory, which contains extended objects. Furthermore, since their size increases with mass logarithmically, we see that at high energies their effects appear at $b^2 \sim \alpha' \log(s\alpha')$ rather than the more naive expectation of $b^2 \sim \alpha'$. This justifies the small t expansion used in (5.61) . Further aspects of the string case are discussed in appendix E.

The conclusion is that an infinite number of higher spin particles can solve the problem. We need a tower of particles with increasing spins and intricate relations between them so that the expansion can be resummed into an amplitude that does not have a problem. Besides string theory, we do not

[13] For type II superstrings Pol= 1.

know if there are other ways of doing this.

5.4.5 Compositeness and the Extra Structures for Graviton Scattering

In this subsection, with the string theory case as an inspiration, we make some remarks about the extra structures for the graviton scattering. Imagine that the graviton has a composite structure.[14] Let us imagine that the graviton is given by a pair of particles which are in a bound state given by a wavefunction $\psi(r)$ where r is the relative distance. We further assume that these particles scatter through the shock via the usual general relativity three-point functions. However, since the two particles feel slightly different forces we find that the total scattering amplitude will have the form

$$\delta_0(b_0) + \partial_i\partial_j\delta_0(b)\langle\psi_\epsilon|r_ir_j|\psi_\epsilon\rangle + \cdots \tag{5.63}$$

where δ_0 is the general relativistic expression (5.12) , and the subindex ϵ indicates the spin of the graviton. Notice that since the Laplacian $\nabla_b^2\delta_0 = 0$, the only terms that can contribute to the second part are those which are not rotationally invariant in the relative coordinates. These are possible because the graviton spin or polarization ϵ. Notice that this is a simple argument for the presence of extra structures in the graviton three-point function. Even though we motivated this with a graviton composed with two particles, the same final formula works if the graviton is made out of many more elementary constituents as it happens in string theory, when it is a string. In any case, the size of the new structure,α_2, due to compositeness, is of order $\alpha_2 \sim r_s^2$ where r_s is the typical size of the graviton. Given that the bound state has this typical size, then we also expect that it can be excited to other states with masses $m^2 \sim 1/r_s^2$. Indeed, by imposing the causality constraint, we found that there should be new particles with masses of this order of magnitude.

5.5 Anti-de Sitter Discussion

The case of asymptotically AdS space is very similar to the asymptotically flat space one. The causal structure is defined by the causal structure of the boundary. We then require that signals that go through the bulk cannot go faster than signals that remain on the boundary. As argued in [179, 152], general relativity with the null energy condition implies that this is obeyed. In terms of the dual CFT, this is just the statement that CFT observers cannot exchange information faster than light. Equivalently boundary CFT operators commute outside the boundary light-cone.

[14] Note that we cannot make the graviton as a zero energy bound state in a local relativistic quantum field theory that contains a stress tensor operator [172] .

5.5.1 Motivation: The Emergence of Bulk Locality Should Happen in the Classical Theory

In this subsection we discuss in more detail the AdS considerations that motivated the present chapter. We expect that the dual of a large N gauge theory should be a weakly coupled string theory with coupling $g_s \sim 1/N$. This should be true both at weak and strong 't Hooft coupling. As we increase the 't Hooft coupling we are supposed to interpolate between a Vasiliev-like theory and an ordinary Einstein-like theory at strong 't Hooft coupling. This whole interpolation happens within the classical string theory in the bulk. Of course, in the ordinary Einstein description we see a local theory in the bulk. Thus the emergence of bulk locality is something that should be contained within *classical* string theory. It is for this reason that it is interesting to understand the constraints of tree level interactions of gravitons and the link between the masses of the higher spin particles and the size of the corrections to Einstein gravity. Here we attempted to link them via the causality considerations for the simplest gravitational interactions. Given the interest of the AdS case, we will discuss in more detail some of its features.

5.5.2 Statement of AdS Causality

In asymptotically AdS gravitational theories the causal structure given by the Minkowski light-cone on the boundary of AdS. This allows us to formulate the causality criterion in a very simple manner. None of the subtleties that existed in flat space regarding the definition of causality appear for gravitational theories in AdS. The basic condition we would like to impose is that

$$\langle\Psi|[T_{\mu\nu}(y), T_{\rho\sigma}(0)]|\Psi\rangle = 0, \quad y^2 > 0, \tag{5.64}$$

where Ψ is some nontrivial state in the theory. We will be interested in the commutator computed on the shock wave background [107, 180, 181, 182, 183]. The Aichelburg-Sexl shockwave in AdS can be created by inserting a pair of operators creating a coherent state bulk wavefunctions that localize the bulk stress tensor on the light ray [107, 108, 109]. As before, instead of computing the commutator, we study propagation of an energetic graviton through the bulk and impose positivity of the time delay. The whole discussion is similar to the flat space one. The only difference is that the t-channel propagator is now in AdS. Instead to computing it directly, we consider plane wave solutions which encode intermediate state gravitons with the properties we need. Plane wave solutions in AdS_D have the form

$$ds^2 = \frac{-dudv + h(u, y_i, z)(du)^2 + \sum_{i=1}^{D-3} dy_i^2 + dz^2}{z^2} \tag{5.65}$$

Here the function h is only constrained by the Laplace equation in the transverse hyperbolic space

$$z^{D-2}\partial_z(z^{-(D-2)}\partial_z h) + \partial_{y_i}^2 h = 0 \tag{5.66}$$

which can be equivalently written as

$$\nabla_{D-2}^2 f - (D-2)f = 0 \tag{5.67}$$

where $f = \frac{h}{z}$. In appendix F we give an argument that this is a solution, to all orders in the derivative expansion. We can also a plane wave with a delta function source so that instead of (5.66) we write

$$z^{D-2}\partial_z(z^{-(D-2)}\partial_z h) + \partial_{y_i}^2 h = -16\pi G|P_u|\delta(u) z_0^{D-2}\delta^{D-3}(\vec{y} - \vec{y}_0)\delta(z - z_0). \tag{5.68}$$

where the RHS corresponds to an insertion of a delta-function source in the hyperbolic space in (5.67) . The Green's function in the hyperbolic space is well-known, so that we get

$$h = \frac{z\varpi(\rho)}{1-\rho^2}\delta(u) = 16\pi G|P_u|\delta(u)\frac{z(4\pi)^{\frac{2-D}{2}}\Gamma(\frac{D}{2})}{(D-1)(D-2)}\left(\frac{\rho^2}{1-\rho^2}\right)^{2-D} {}_2F_1(D-2, \frac{D}{2}, D; -\frac{1-\rho^2}{\rho^2}), \tag{5.69}$$

$$\rho = \sqrt{\frac{(z-z_0)^2 + |\vec{y} - \vec{y}_0|^2}{(z+z_0)^2 + |\vec{y} - \vec{y}_0|^2}}. \tag{5.70}$$

where $\vec{x}_0$ and z_0 are the coordinates of the source in the bulk. It can be checked that P_u is also the total momentum from the boundary point of view by integrating the boundary stress tensor (read off from the small z expansion of this metric) on the boundary. For example, in AdS_4 and AdS_5 (5.69) takes the following form

$$\varpi_4(\rho)/(G|P_u|) = -8\left(1 - \rho^2 + (1+\rho^2)\log\rho\right), \tag{5.71}$$

$$\varpi_5(\rho)/(G|P_u|) = 2\frac{(1-\rho)^4}{\rho}. \tag{5.72}$$

More generally, we have

$$\lim_{\rho\to 0}\varpi(\rho)\sim\frac{1}{\rho^{D-4}}, \tag{5.73}$$

$$\lim_{\rho\to 1}\varpi(\rho)\sim(1-\rho)^{D-1}. \tag{5.74}$$

Let us understand better the symmetries of the problem. The shock wave is localized around $u=0$ and is probed by a particle which is localized in v. The role of the transverse plane in flat space is played here by the transverse H_{D-2}. It is convenient to think of the probe crossing this hyperbolic space at the center.[15] The shock centered at $(\vec{y}_0, z_0)$ has a number of Killing vectors that depend on $f(u)$. For arbitrary $f(u)$ the background has an obvious $SO(D-3)$ symmetry that rotates $\vec{y}$ and a translation in v. For $f(u)=\delta(u)$ the geometry has extra Killing vectors which enhance the rotational symmetry to $SO(D-2)$, as we had in flat space. In the original coordinates (5.65) , some of the extra symmetries involve special conformal generators. Since we are working in the high energy limit, we effectively have a delta function in u, so that we also expect to have this extra $SO(D-2)$ symmetry. See [182] for a coordinate system that makes this manifest. General properties of the AdS shock wave and its different limits are considered in appendix F. In particular, when the center of the shock goes to the boundary $z_0\to 0$ the problem becomes very similar to the one arising in the computation of energy correlators [20], whereas in the limit $z_0\to\infty$ it reduces to the setup used in [183] to study causality. Our formulas will reduce to the ones considered before in those limits.

5.5.3 The Effect of Higher Derivative Interactions on Particles with Spin

Now we would like to consider different type of probes and compute the time delay for them. We start with a simple example of a scalar probe and then move to the case of particles of spin one and two. If we consider a minimally coupled scalar in the shock wave background its equation of motion takes the form

$$\nabla^2\phi=0. \tag{5.75}$$

[15] From the CFT point of view we can create such a probe by acting with the operator of given energy and zero momentum in the AdS Poincare coordinates for which $u=0$ is the future null infinity as explained in [20] . In [20] terms we are working here in the y-coordinates, while the operator with given momentum is inserted in the x-coordinates. In pure AdS case isometries of H_{D-2} at fixed $u=0$ correspond to the usual Lorentz symmetry group in the x-coordinates of [20].

In our setup we are interested in corrections to this equation which are second order in derivatives.[16] Considering terms yielding two derivative equations of motion for ϕ we have to consider terms like

$$\mathcal{H}^{\mu\nu}\nabla_\mu\nabla_\nu\phi \tag{5.76}$$

where the tensor $\mathcal{H}$ is made from the background metric, Riemann and covariant derivatives. However one can check that there is no two index symmetric tensor that is not vanishing on-shell [182] . Of course, this statement is equivalent to the uniqueness of the scalar-scalar-graviton three-point vertex. In the high energy limit we get similar to flat space

$$\partial_u\partial_v\phi + f(u)h(z,\vec{y})\partial_v^2\phi = 0 \tag{5.77}$$

which produces the time delay

$$\Delta v = \frac{\varpi(\rho)}{1-\rho^2} \tag{5.78}$$

reproducing the flat space computation for small ρ. We assumed that the perturbation crosses the shock at $z = 1$ and $\vec{y} = 0$. For the gauge boson we imagine at the level of two derivative the following equation

$$\nabla^\mu F_{\mu\nu} + \mathcal{H}_\nu{}^{\mu\alpha\beta}\nabla_\mu F_{\alpha\beta} = 0 \tag{5.79}$$

where $\mathcal{H}$ is built out of the Riemann tensor and its covariant derivatives. Using the properties of the background discussed above (we defer the details to appendix G) one can check that the only term that we can have is the correction that appeared in the case of flat space

$$\nabla^\mu F_{\mu\nu} - \widehat{\alpha}_2\,\check{R}_\nu{}^{\mu\alpha\beta}\nabla_\mu F_{\alpha\beta} = 0. \tag{5.80}$$

If we compute the time delay using the same action (5.80), considering that each mode corresponds to a different constant polarization, ϵ_i, we get

$$\Delta v = \frac{\varpi(\rho)}{1-\rho^2}\left(1 - \widehat{\alpha}_2\left(1-\rho^2\right)^2\frac{\varpi'(\rho)-\rho\varpi''(\rho)}{4\rho\varpi(\rho)}\left(\frac{\epsilon.n^2}{\epsilon.\epsilon} - \frac{1}{D-2}\right)\right). \tag{5.81}$$

[16] Since we need derivatives to bring down large factors of momentum.

The final result is very similar to the one obtained in flat space. The only different is that the polarization dependent delay is slightly more complicated. The flat space limit is reproduced by considering $\rho \to 0$ limit, whereas the energy correlator constrained is recovered in the limit $\rho \to 1$. Similarly, in the case of gravity we are interested in the most general form of the second order equations. We choose to parameterize the equations of motions for perturbations as follows

$$\delta R_{\mu\nu} + \alpha_2 \, \check{R}_{(\mu}{}^{\rho\alpha\beta} \delta R_{\nu)\rho\alpha\beta} + \frac{\alpha_4}{2} \, [\nabla_{(\mu} \nabla_{\nu)} \check{R}^{\alpha\beta\rho\sigma}] \delta R_{\alpha\beta\rho\sigma} = 0 \tag{5.82}$$

where the parameters α_i are in units of the AdS radius R_{AdS} that we set to one. The time delay for these equation of motion is then given by

$$\Delta v = \frac{\varpi(\rho)}{1-\rho^2} \left(1 + t_2(\rho) \left(\frac{(\epsilon.n)^2}{\epsilon.\epsilon} - \frac{1}{D-2} \right) + t_4(\rho) \left(\frac{(\epsilon.n)^4}{(\epsilon.\epsilon)^2} - \frac{2}{D(D-2)} \right) \right), \tag{5.83}$$

$$t_2(\rho) = \left(1-\rho^2\right)^2 \frac{\varpi' - \rho\varpi''}{4\rho \, \varpi(\rho)} \left(-\alpha_2 + \alpha_4 \frac{D(1+\rho^2)^2 - 2\rho^2}{\rho^2} \right), \tag{5.84}$$

$$t_4(\rho) = -D\alpha_4 \left(1-\rho^2\right)^2 \frac{\varpi' - \rho\varpi''}{4\rho \, \varpi(\rho)} \frac{D(1+\rho^2)^2 + 2(1+\rho^4)}{4\rho^2}. \tag{5.85}$$

where $\vec{n}$ is a vector pointing from the center of the shock to the probe particle, and ϵ is the polarization of the probe particle. These time delays can become negative for small enough ρ if α_2 or α_4 are non-zero. These results can be also reproduced using slightly different method of evaluating the on-shell in an explicit gravitational theories in the shock wave background like Lovelock or quasi-topological theories (see e.g. [184, 185]). In the limit $\rho \to 0$ these constraints reproduce the flat space analysis with $\rho = \frac{b}{2}$. In the $\rho \to 1$ limit the above result reproduces constraints discussed in the past. If we take the limit $\rho \to 1$ by taking the shock center to the boundary $z_0 \to 0$ we recover energy correlator computation [20]. If we on the other hand consider $\rho \sim 1$ by taking the shock center to the horizon $z_0 \to \infty$ we recover the shock wave discussed by Hofman [183].

5.5.4 Implications for a, c in Theories with Large Operator Dimensions

Imagine an abstract CFT_4 with large $N \gg 1$ and large gap $\Delta_{gap} \gg 1$, where Δ_{gap} is the dimension of the lightest higher spin single trace operator. This theory is described by a gravitational theory in the bulk with potentially some higher derivative corrections. String theory inspired intuition suggests that higher derivative corrections should be suppressed by the $\frac{1}{\Delta_{gap}}$ factor. Our argument shows that this is indeed the case for the simplest corrections, which are the ones affecting the three-point function of stress tensor. Indeed, as we showed in (5.83) , we run into a potential problems with causality at impact parameters $\rho_c \sim (\alpha_2)^{1/2}, (\alpha_4)^{1/4}$. Since this can only be fixed by higher spin particles, we conclude that Δ_{gap} has to be small enough so that it can start correcting the amplitude before we run into this problem.

For $\Delta_{gap} \gg 1$ then the relevant impact parameters are such that we can approximate the formulas by the flat space limit. Thus, we get the bound of the type

$$(\alpha_2)^{1/2} \lesssim \frac{1}{\Delta_{gap}}, \quad (\alpha_4)^{1/4} \lesssim \frac{1}{\Delta_{gap}}, \tag{5.86}$$

where $\lesssim$ stands for some numerical coefficient that we cannot fix using our simple analysis. In the case of $\mathcal{N}=1$ superconformal theories, $\alpha_4 = 0$ by supersymmetry and $\alpha_2 \propto \frac{a-c}{c}$. Then we get

$$\left|\frac{a-c}{c}\right| \lesssim \frac{1}{\Delta_{gap}^2}. \tag{5.87}$$

In the case of $\mathcal{N}=4$ SYM (5.86) is satisfied trivially since $\Delta_{gap} \sim \lambda^{1/4}$ and $a=c$. It will be very interesting to find an independent field theoretic argument that leads to the bounds of the type (5.86), (5.87) . It would also be nice to find the precise numerical coefficients in (5.86) , (5.87) .

5.5.5 Implications for Dimensions of Double Trace Operators

This computation of the time delay can be viewed as a special four-point correlation function with particular wave functions for external operators. We can use the OPE to expand the four-point function computation in terms of the time evolution eigenstates (or, equivalently, local operators). One can ask then the following question: what is the relation between the time delay and the OPE data? This question was addressed in a nice series of papers [107, 108, 109] where the time delay was shown to be equal to anomalous dimension of double trace operators of the type $\mathcal{O}(\partial^2)^n\partial_{\mu_1}...\partial_{\mu_j}\mathcal{O}$ for large n and large j

$$\delta(s,\rho) = -\pi\gamma(n,j),\ \rho = \frac{j}{j+2n},\ s = 4n(j+n). \tag{5.88}$$

where the relations between parameters on both sides of the equation are reviewed below. Intuitively, (5.88) follows from the fact that the phase shift of the correlator $e^{i\delta}$ is given by $e^{-i(\Delta_* - 2\Delta_{\mathcal{O}})\Delta\tau}$ where Δ_* is the dimension of the operator that dominates the OPE, and $\Delta\tau$ is the global AdS time that passed from the beginning to the end of the process. The time delay that we computed corresponds to particles starting at the boundary, getting close to each other at the center and then reaching

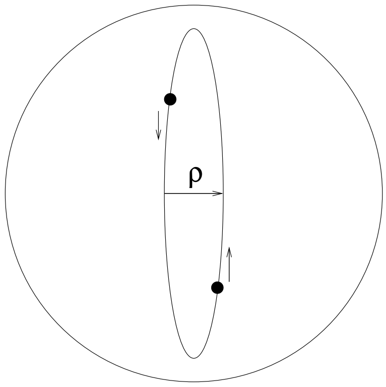


Figure 5.8: We consider two energetic particles in AdS that oscillate back and forth with energy E and angular momentum J. This models the high twist, high spin double trace operator in the dual CFT.

the boundary again, see (5.8) . It takes $\Delta\tau = \pi$ for this process to occur. From this fact (5.88) follows. Even though (5.88) was derived in general relativity in the limit when $\delta \ll 1$ it follows simply from the AdS graviton diagram exchanged. In the impact parameter representation, only the on-shell t-channel exchange diagram constributes. This diagram is fixed in terms of three-point function $\langle \mathcal{O}\mathcal{O}T_{\mu\nu}\rangle$ in a generic gravitational theory with generic three-point couplings similar to our flat space analysis. Let us briefly review the results in [107, 108, 109]. The basic idea is the following: the state created by $\mathcal{O}(\partial^2)^n\partial_{\mu_1}...\partial_{\mu_j}\mathcal{O}$ for large n and large j can be thought of as two highly energetic particles that follow null geodesics in AdS, see (5.8) . For two geodesics that are characterized by total energy E and spin J the minimal separation is achieved in global coordinates at

$$\rho = \frac{J}{E} \simeq \frac{j}{j+2n}. \tag{5.89}$$

where we matched energy and spin of the pair of particles to the ones of the double trace operator $J = j$, $E = 2\Delta + 2n + j$ and used that $n, s \gg 1$, with Δ also of order one. Thus, we see that probing distances much smaller than AdS radius ($\rho \ll 1$) corresponds to considering operators with $n \gg j$. On the other hand $n \ll j$ corresponds to scattering at very large impact parameters. The Mandelstam variable s is given by

$$s = E^2 - J^2 \simeq 4n(j+n) \tag{5.90}$$

and the relation to the anomalous dimensions is that

$$\pi\gamma(n,j) = p_v \Delta v = -Gs\frac{\varpi(\rho)}{1-\rho^2}. \tag{5.91}$$

Let us consider different limits of this formula. First, consider very large impact parameters $\rho \sim 1$ or $j \gg n$. In this limit we get

$$\gamma(n,j) \sim -G\frac{n^{D-1}}{j^{D-3}} \tag{5.92}$$

which is in agreement with the general results derived using the crossing equation [110, 186, 187]. In the opposite limitof small impact parameters scattering, $n \gg j$ and $\frac{j}{n} > \frac{1}{\Delta_{gap}}$, we have

$$\gamma(n,j) \sim -Gn^2\left(\frac{n}{j}\right)^{D-4}. \tag{5.93}$$

We have several comments to add to this story. First, these results should be universal and applicable to generic CFTs with large N and large gap. To write the answer in an abstract form we have to use the relation of G to the two-point function of stress tensors which is well-known and is roughly $G \sim \frac{1}{c_T}$. It means that it should be possible to derive them using crossing equation which still be dominated by the stress tensor exchange. Probably the relevant limit is $z \to 0$, $\bar{z} \to 1$ with $\frac{z}{1-\bar{z}}$ being fixed. It would be nice to reproduce the formulas above using crossing equations. Second, we see that causality, or positivity of the time delay, implies the constraint $\gamma(n,s) < 0$, which generalizes the ones that were previously known [98, 186, 187] for asymptotically large $s \gg n$. Again it would be very interesting to understand how to prove these constraints purely from the field theory point of view. Third, we see that considering double trace operators of the type $T_{\mu\nu}(\partial^2)^n\partial_{\mu_1}...\partial_{\mu_j}T_{\rho\sigma}$ we get new structures due to the dependence on polarization which potentially lead to causality violations and bounds (5.86), (5.87) . Of course, the same is true about the double trace operators that involve the conserved current J_μ. It will be very interesting to understand them from the purely CFT viewpoint. Note that in the scattering process the polarizations of particles 3 and 4 in (5.3) can change relative to those of particles 2 and 4, so that the phase shift is an operator that acts on this space of polarization tensors. Both t-channel unitarity, as well as our considerations, constrain only some matrix elements of this general matrix. While we leave the general case for the future, we note that, in some cases, we can ensure that the polarizations of 3 and 4 do not change by using conservation of angular momentum along the directions orthogonal to the impact parameter

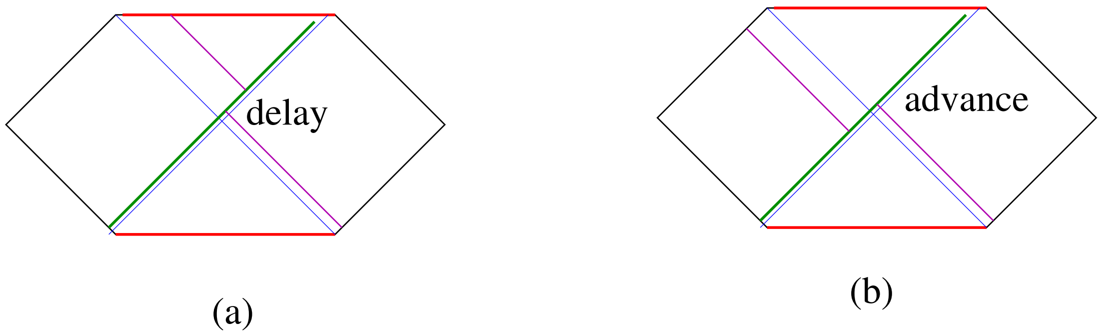


Figure 5.9: We consider a Lorentzian wormhole configuration that is described, near each wormhole, by the maximally extended Schwarzschild solution. We send a (green) particle from the left very close to the past horizon. We then send a (purple) particle from the right. (a) If this particle gets a time delay, it will fall into the singularity. (b) If the particle gets a time advance, then it can make it out of the other black hole and we would have a way of sending signals through the wormhole. Here the blue lines represent the average position of the horizon of the black hole, defined by null lines that are very far away from the two particles we send in. We assume that the impact parameter is much smaller than the Schwarzschild radius.

direction. In these cases our considerations apply and we can say that the anomalous dimensions of the corresponding double trace operators should be negative. The positivity statement applies to the part of the phase shift that grows with the Mandelstam invariant s, which translate into the growth with n via (5.90) . However, in our case, this positivity requirement is not obvious from the CFT point of view. It would be nice to see whether this is a general requirement or is one that is present only in theories with a local bulk dual.

sectionWormholes and Time Advances General relativity has Lorentzian wormhole solutions that join far away points by short Einstein-Rosen bridges. The simplest configuration is the maximally extended Schwarzschild solution interpreted as an approximation to the metric of two distant black holes which share a single interior. As discussed in [153] , these solutions do not lead to a violation of causality in the ambient space because it is not possible to send signals through the wormhole [188] .

The inability to send signals through the wormhole depends crucially on the fact that we have a Shapiro time *delay* as opposed to a time advance. For example, if one sends a fast moving particle from the left side, then a particle send from the right will suffer a time delay that will make it go into the singularity, see e.g. [158, 189] . However, if that particle were to suffer a time advance, as opposed to a time delay, then it would be able to go through the wormhole and we would have a violation of causality, see (5.9) . Note that we can make the wormhole big enough that we can neglect the higher derivative corrections in the description of the background metric. It can also be big enough that we can neglect the backreaction of the two particles. So we are considering a situation where the impact parameter $b \ll$(Schwarzschild radius). In the $D = 4$ case, the Schwarzschild radius acts at the IR cutoff of the logarithm. This impossibility of sending signals is crucial for interpreting the wormhole as an EPR state of two disconnected systems [154, 155, 156].

5.6 Cosmological Applications

The gravity wave non-gaussianities produced by inflation are a direct measure of the graviton three-point vertex during inflation [144, 49]. To leading order in the slow roll approximation we can do the computation in de Sitter space. The symmetries of de Sitter imply that only two different parity preserving structures are possible. These correspond to the two parity preserving structures that we have in four-dimensional flat space. One is the one produced by the Einstein action and the other can be produced by a term in the action of the form $M_{pl}^2 \alpha_4 R^3$, where R is a Riemann tensor (not the Ricci tensor). The relative size between the two types of gravity wave non-gaussianity is proportional to

$$\frac{\langle hhh \rangle_{R^3}}{\langle hhh \rangle_{Einstein}} \propto \alpha_4 H^4 \tag{5.94}$$

where H is the Hubble scale during inflation. Of course, both are small compared to the two point function, $\frac{\langle 3\text{point} \rangle}{\langle 2\text{point} \rangle^{3/2}} \sim \frac{H}{M_{pl}}$. See [49] for the explicit expressions. Thus, if the gravity wave three-point function was measured *and it was found that this exotic new structure is present at a level comparable to the Einstein one*, then one concludes that α_4 is of the order of the Hubble scale. The considerations in this chapter imply that there should also be new particles with spins $J > 2$ with masses comparable to the Hubble scale during inflation. Thus, this would be an indirect evidence for string theory during inflation. Note that $\alpha_4 H^4 \sim 1$ implies that supersymmetry had to be broken at the Planck scale and not at a lower scale, since the $+++$ and $---$ structures are forbidden by supersymmetry.[17] Let us be a bit more explicit about this point. If the short distance theory is supersymmetric, then the field theory Lagrangian does not contain the couplings giving rise to α_4. Now, since supersymmetry is broken, this three-point function could arise from integrating out massive particles. These are expected to contribute to α_4 as

$$\alpha_4 \sim \frac{1}{M_{pl}^2} \left(\frac{1}{m_B^2} - \frac{1}{m_F^2} \right) , \quad \longrightarrow \quad \alpha_4 H^4 \sim \frac{H^2}{M_{pl}^2} \ll 1 \tag{5.95}$$

[17] We thank Nima Arkani-Hamed for emphasizing this to us.

which is very small. Where, to maximize the effect, we assumed that the masses of the bosons and fermions, as well as their differences, are of order H. We then see from (5.94) and (5.95) that in this supersymmetric scenario the contributions are very small. Thus, in order for the right hand side of (5.94) to be of order unity (or say a few percent) the supersymmetry should be broken at the Planck scale (during inflation) so that the three-point vertex is present in the original classical theory. Notice that most of the string inflation models *do not* predict a large α_4 since they are based on compactifications of the ten dimensional superstring. It should be a model where the string length is comparable to the Hubble radius and with a very weak coupling to account for the small experimental upper bound for H/M_{pl} [190] . Note that if one imagines that inflation is given a dual description in the spirit of dS/CFT and the dual field theory is weakly coupled, then one expects that $\alpha_4 H^4 \sim 1$. This is what happens in the Vasiliev theory [27] . Of course, this theory also contains massless higher spin particles. It is also not suitable for building an inflationary model because the scalar does not appear to obey the slow roll conditions. If the gravity waves produced by inflation are as large as to explain the signal seen by BICEP2 [191] , then probing the gravity wave three-point functions might be possible (with a lot of optimism!).[18]

5.7 Conclusions

In this chapter, we studied causality constraints on higher derivative corrections to the graviton three-point function. We considered a weakly coupled theory and studied higher derivative corrections which are important before the theory becomes strongly coupled. These are higher derivative corrections that arise in the classical regime of the theory. The constraints arise from a thought experiment where we scatter two gravitons at relatively high energy and fixed impact parameter. The energy is high compared to the inverse of the impact parameter but low compared to the scale where the theory becomes strongly coupled. More explicitly, we have the very small overall coupling G and we consider corrections to the three-point functions which scale as powers of αp^2 relative to the Einstein one. The three-point amplitudes are very small because G is very small. But α is a fixed quantity and we look at impact parameters of the order of $b^2 \sim \alpha$. We found that when the impact parameter $b^2 \sim \alpha$, then we see a causality violation. In this impact parameter representation, and in the field theory regime (without higher spin particles), the time delay comes from the singularities in the t-channel, which is simply a pole at $t = 0$ for the massless theory. More precisely it comes from the part that is quadratic in s of its residue. In other words, terms going like s^2/t at $t = 0$.

[18] At this time, there are alternative explanations for this signal [192, 193], so we might have to wait till the dust settles.

It is important that while $t = 0$, the momentum transfer itself is non-zero.[19] The overall sign of the time delay, then depends on the contractions of the polarization tensors of the external particles with the momentum transfer in the t-channel.

We have argued that this type of tree level causality violation can only be fixed, at tree level, by higher spin particles at a mass scale $m^2 \sim 1/\alpha$. In string theory, this issue is fixed because the amplitude Reggeizes. Namely, it has a behavior $s^{2+\frac{\alpha' t}{2}}$ for large s and fixed t. This is due to extended strings being exchanged in the s-channel. If the amplitude Reggeizes, then corrections appear at a scale $b^2 \sim \alpha' \log(s\alpha')$. Due to the presence of the logarithm we did not find a sharp bound between the corrections to the graviton three point amplitude, α, and the Regge slope α'.

We should stress that in this discussion we have assumed that we have a weakly coupled gravitational theory. We have also assumed the notion of asymptotic causality which says that the causal structured determined by the far away regions of spacetime cannot be violated by its interior regions. The analysis in this chapter was also motivated by trying to understand better the AdS/CFT correspondence. In particular, if we consider large N gauge theories we know that we have a weakly coupled theory in the bulk. However, we do not know under what conditions that weakly coupled theory is a local in the bulk. It is clear that the absence of light massive higher spin states is a requirement. Here we have tried to address the question of whether it is sufficient. We have only studied the simplest possible correction to gravity, its three-point function. We have argued that, as long as higher spin particles are very massive, there cannot be higher derivative corrections to the three-point functions. Previous discussions argued against such corrections by saying that they would make the theory strongly coupled at energies that are lower than the Planck scale [47, 48], but still parametrically larger than the scale of the corrections.[20] Here we have strengthened the bound by linking the scale of corrections to the appearance of new particles at the same scale. As a more concrete statement, we are linking the values of the constants appearing in the stress tensor three-point functions to the dimensions of the lightest particles with higher spins, $J > 2$. In other words, $\frac{a-c}{c} \lesssim \frac{1}{\Delta_{gap}^2}$. Unfortunately, we could not determine the precise numerical constant in this inequality. Using the results in [107, 108, 109], we can link the time delay for a high energy scattering process in the bulk to the anomalous dimensions of certain double trace operators. These double trace operators have the rough form $T\partial_+^j(\partial^2)^n T$. They have both relative spin and relative radial excitations. The anomalous dimensions are $\gamma(n,j) \sim -\delta(s,b)/\pi$, where the values of b and s are

[19] It is a null, non-zero momentum.

[20] In [47, 48], or in talks referring to those papers, they impose the bound $\alpha \lesssim \frac{1}{G(\Delta_{gap})^4}$ (with $G \sim 1/N^2$). This bound comes from demanding that the theory remains perturbatively unitary at the scale Δ_{gap} where new particles appear. Here we argued for the stronger bound $\alpha \lesssim \frac{1}{\Delta_{gap}^2}$.

given in terms of j, n in (5.89) (5.90) . The requirement that the time delay is positive leads to the statement that the anomalous dimensions for some of these operators should be negative.

For the de Sitter case, this analysis has *potentially* interesting phenomenological applications. *If* the gravity wave three-point functions were measured, we expect to see the structure predicted by Einstein theory. However a new structure is also possible. This new structure is the only one allowed by the approximate scale and conformal invariance of the inflationary phase. *If* such new structure was found with a strength comparable to the Einstein one, then it would be a direct signal of dramatically new physics at the Hubble scale: a tower of higher spin particles. Which is a rather drastic departure from ordinary field theory at the inflationary scales. It is not clear how likely this inflationary scenario is in the space of possible inflationary theories.

5.7.1 Open Problems

It would be nice to derive these constraints in a more direct way. If one understood directly the constraints of unitarity and causality at the level of the four point function, then one would not need to resort to the exponentiation argument we discussed in section 3.3. Furthermore, it might lead to sharper bounds that include numerical factors.

Our discussion of massive intermediate particles in mixed representations was not complete.[21] These are representations that have maximal spin two in the uv-plane, but have additional indices in the other directions (see appendix H). We suspect that a finite number of these cannot solve the causality problem, but we did not prove it.

In the AdS case, it would be nice to derive the constraints from the conformal bootstrap point of view. This is in the spirit of [47] , but it would involve the stress tensor as an external state. One of the main messages from this chapter is the importance of spin in order to derive constraints. The structure constants (or three-point functions) of operators with spin are very numerous but, as we have shown, there can be powerful constraints on them. These constraints are not so easily seen when we scatter external operators with no spin. Notice that, even the simplest bounds for a and c discussed in [20] , which should be valid for arbitrary CFTs, have not been derived from the conformal bootstrap approach. It would be nice to further constrain the interactions of all the higher spin particles and derive the general structure of the tree level theory. This is the program that was pursued in the sixties and that led to string theory. However, we would like to know how

[21] These are not present in $D = 4$, so that the existence of an infinite tower of higher spin states is clear in $D = 4$, or if $\alpha_4 \neq 0$ in higher dimensions. In higher dimensional theories with only $\alpha_4 = 0$ but non-zero α_2, in order to establish that the tower is really infinite we need to rule out the possibility that the causality problem is fixed with a finite number of mixed representations. We leave this to the future.

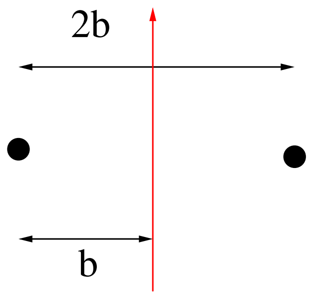


Figure 5.10: We consider the metric produced by two massive sources at distance $2b$. We then send a particle between them and measure the time delay. There is no deflection angle, but there is a non-zero time delay. We do the computation to leading order in the r_S/b expansion, by simply consider a linear superposition of the two metrics.

unique string theory is. Methods developed to tackle this problem might also be useful for analyzing large N gauge theories such as large N QCD. It would also be nice to see whether in the de Sitter context there is a sharp bound for the gravity wave three-point correlators analogous to the one in [20] for AdS correlators. Our de Sitter discussion assumed the local gravity description and a locally flat space discussion in the bulk, so it applies most clearly when $\alpha_4 H^4$ is somewhat smaller than one, but still of order one compared with H/M_{pl}.

5.8 Appendix A. Shapiro Time Delay

The physical effect discussed in the chapter is known in general relativity as the Shapiro time delay [137] . The usual setup to discuss the Shapiro time delay is to consider propagation of light near a massive body (a star or a planet). Here we would like to consider a slight variation of it by considering propagation of light between two massive bodies. We consider the masses to be equal and consider the geodesic that is equally separated from each of them, such that the deflection is absent (each of the masses bends the trajectory in the opposite direction such that the net deflection is zero). The time delay on the other hand accumulates. Recall that the Schwarzschild metric takes the form

$$ds^2 = -\left(1 - \frac{r_s^{D-3}}{r^{D-3}}\right) dt^2 f + \frac{dr^2}{1 - \frac{r_s^{D-3}}{r^{D-3}}} + r^2 d\Omega^2 \tag{5.96}$$

We are interested in the metric produced by the superposition of two equal masses. In the

Schwarzschild coordinates to first order in the mass, or r_S, we get

$$ds^2 = ds^2_{Mink} + \sum_i \frac{r_s^{D-3}}{r_i^{D-3}}(dt^2 + dr_i^2) \tag{5.97}$$

with $r_i = |\vec{x} - \vec{x}_i|$. Let us now consider two masses separated along the direction x^1, one at $x^1 = -b$ and one at $x^1 = b$. Then we consider a probe particle moving along the direction $x^{D-1} = z$. By symmetry, it will stay at $x^1 = ... = x^{D-2} = 0$ if it starts there with velocity along the z direction. We have that $r = \sqrt{b^2 + z^2}$ for both masses. We also find that $dr = dzz/r$. We get

$$dt^2(1 - 2\frac{r_s^{D-3}}{r^{D-3}}) = dz^2(1 + \frac{r_s^{D-3}}{r^{D-3}}z^2/r^2) \tag{5.98}$$

Then the time delay is

$$\Delta t = \int_{-\infty}^{\infty} dz r_s^{D-3} \frac{1}{(b^2+z^2)^{\frac{D-3}{2}}}(1 + \frac{z^2}{z^2+b^2}) = \frac{r_s^{D-3}}{b^{D-4}} \frac{(D-2)\sqrt{\pi}\Gamma(\frac{d}{2}-2)}{2\Gamma(\frac{D-1}{2})} \tag{5.99}$$

We now do the same for a particle moving at a velocity v. We find

$$dt^2(1 - 2\frac{r_s^{D-3}}{r^{D_3}}) = dz^2\left(1 + \frac{r_s^{D-3}}{r^{D_3}}\frac{z^2}{z^2+b^2}\right) + d\tau^2 \tag{5.100}$$

$$\frac{dt}{d\tau}(1 - 2\frac{r_s^{D-3}}{r^{D_3}}) = \gamma\ , \qquad \gamma^{-2} = 1 - v^2 \tag{5.101}$$

$$dt = \frac{dz}{v}\left(1 + (2 + \frac{z^2}{z^2+b^2} - \frac{1}{v^2})\frac{r_s^{D-3}}{r^{D_3}}\right) \tag{5.102}$$

The result of doing the integral over z is the same as in (5.99) multiplied by a factor of

$$\Delta t = \frac{1}{v}\frac{(2 + \frac{1}{D-3} - \frac{1}{v^2})}{(1 + \frac{1}{D-3})}\Delta t_{\text{relativistic}} \tag{5.103}$$

We see that for slow velocities, we indeed find a negative time delay (time advance) proportional to $1/v^3$. This is time advance relative to the particle moving with the same velocity v in flat space (not relative to a particle moving at the speed of light!). We could have repeated the computation in different coordinate system. For example, we can consider the computation in the so-called isotropic coordinates given by

$$ds^2 = -\left(\frac{1-\frac{r_s^{D-3}}{4r^{D-3}}}{1+\frac{r_s^{D-3}}{4r^{D-3}}}\right)^2 dt^2 + \left(1+\frac{r_s^{D-3}}{4r^{D-3}}\right)^{\frac{4}{D-3}} \sum_{i=1}^{D-1} dx_i^2, \tag{5.104}$$

$$r^2 = \sum_{i=1}^{D-1} x_i^2. \tag{5.105}$$

One can easily check that the time delay for the geodesic in this coordinates coincides with the one computed in Schwarzschild coordinates to leading order in r_S.

5.9 Appendix B. Three-Point Amplitudes and Their Sums

In this appendix we recollect different three-point amplitudes that involve a graviton and two other particles and present different polarization sums that appear in the time delay. In particular cases we reproduce the shock wave computations but the results obtained using on-shell amplitudes are much more general and are valid in any theory with given on-shell three-point amplitudes. Let us be more explicit on the kinematics we are interested in. As defined in the bulk of the chapter we consider the following kinematics

$$\begin{aligned} p_{1\,\mu} &= \left(p_u, \frac{q^2}{16p_u}, \frac{\vec{q}}{2}\right) \ , \qquad p_{2\,\mu} = \left(\frac{q^2}{16p_v}, p_v, -\frac{\vec{q}}{2}\right) \ , \\ p_{3\,\mu} &= -\left(p_u, \frac{q^2}{16p_u}, -\frac{\vec{q}}{2}\right) \ , \qquad p_{4\,\mu} = -\left(\frac{q^2}{16p_v}, p_v, \frac{\vec{q}}{2}\right) \ , \\ s &\simeq 4p_u p_v \ , \qquad t \simeq -(\vec{q})^2 \end{aligned} \tag{5.106}$$

We define the polarization tensors as follows

$$\begin{aligned} \epsilon^{1\,\mu} &= (-\frac{\vec{q}.\vec{e}_1}{2p_u}, 0, \vec{e}_1) \ , \qquad \epsilon^{3\,\mu} = (\frac{\vec{q}.\vec{e}_3}{2p_u}, 0, \vec{e}_3), \\ \epsilon^{2\,\mu} &= (0, \frac{\vec{q}.\vec{e}_2}{2p_v}, \vec{e}_2) \ , \qquad \epsilon^{4\,\mu} = (0, -\frac{\vec{q}.\vec{e}_4}{2p_v}, \vec{e}_4). \end{aligned} \tag{5.107}$$

An important point is that all $\epsilon_i.p_j$ are of order 1 and, thus, are sub-leading in the high energy limit compared to powers of s. Thus, in all our computations we can think of polarization tensors as being purely transverse, see (5.25) . We focus on the massless scalar, vector and graviton three-point amplitudes which are parameterized as follows

$$\mathcal{A}_{hhg} = \sqrt{32\pi G}\epsilon_{\mu\nu}p_1^\mu p_3^\nu \tag{5.108}$$

$$\mathcal{A}_{\gamma\gamma g} = \sqrt{32\pi G}\epsilon_{\mu\nu}\left([p_1^\mu p_3^\nu(\epsilon_1.\epsilon_3) - \epsilon_1^\mu p_3^\nu(\epsilon_3.p_1) - \epsilon_3^\mu p_1^\nu(\epsilon_1.p_3)] + \widehat{\alpha}_2 p_1^\mu p_3^\nu(\epsilon_1.p_3)(\epsilon_3.p_1)\right), \tag{5.109}$$

$$\begin{aligned}\mathcal{A}_{ggg} &= \sqrt{32\pi G}\left[(\epsilon_1.\epsilon_2\epsilon_3.p_1 + \epsilon_1.\epsilon_3\epsilon_2.p_3 + \epsilon_2.\epsilon_3\epsilon_1.p_2)^2\right. \\ &+\alpha_2(\epsilon_1.\epsilon_2\epsilon_3.p_1 + \epsilon_1.\epsilon_3\epsilon_2.p_3 + \epsilon_2.\epsilon_3\epsilon_1.p_2)\epsilon_1.p_2\epsilon_2.p_3\epsilon_3.p_1 \\ &\left.+\alpha_4(\epsilon_1.p_2\epsilon_2.p_3\epsilon_3.p_1)^2\right].\end{aligned} \tag{5.110}$$

where we used the $\epsilon^{\mu\nu} \to \epsilon^\mu\epsilon^\nu$ form of the polarization tensor for the graviton . Using these three-point amplitudes we can compute the time delay using (5.123) in the high energy limit. In the high energy limit the relevant part of the sum over graviton polarization tensors takes a very simple form

$$\sum_i \epsilon^i_{\mu\nu}(q)(\epsilon^i_{\rho\sigma}(q))^* \sim \frac{1}{2}\left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho}\right) \tag{5.111}$$

so that it leads to factors of s^2 when we contract with the p_1 or p_3 momenta from the left side and with p_2 or p_4 momenta on the right side, see (5.4) . Note that the large components of the external momenta are transverse to the momentum $\vec{q}$ of the intermediate particle, $p_1 + p_3 = (0, 0, \vec{q})$, see (5.106) . The results are the following

$$\sum \mathcal{A}_{hhg}(-i\partial_{\vec{b}})\mathcal{A}_{hhg}(-i\partial_{\vec{b}}) = 8\pi G s^2, \tag{5.112}$$

$$\sum \mathcal{A}_{hhg}(-i\partial_{\vec{b}})\mathcal{A}_{\gamma\gamma g}(-i\partial_{\vec{b}}) = 8\pi G s^2\left(e_2.e_4 + \widehat{\alpha}_2 e_2.\partial_b e_4.\partial_b\right), \tag{5.113}$$

$$\sum \mathcal{A}_{hhg}(-i\partial_{\vec{b}})\mathcal{A}_{ggg}(-i\partial_{\vec{b}}) = 8\pi G s^2\left(e_2^{ij}e_4^{ij} + \alpha_2 e_2^{ij}e_4^{ik}\partial_{b^j}\partial_{b^k} + \alpha_4 e_2^{ij}e_4^{kl}\partial_{b^i}\partial_{b^j}\partial_{b^k}\partial_{b^l}\right), \tag{5.114}$$

$$\sum \mathcal{A}_{\gamma\gamma g}(-i\partial_{\vec{b}})\mathcal{A}_{\gamma\gamma g}(-i\partial_{\vec{b}}) = 8\pi G s^2\left(e_1.e_3 + \widehat{\alpha}_2 e_1.\partial_b e_3.\partial_b\right)\left(\epsilon_2.e_4 + d_2 e_2.\partial_b\epsilon_4.\partial_b\right), \tag{5.115}$$

$$\begin{aligned}\sum \mathcal{A}_{\gamma\gamma g}(-i\partial_{\vec{b}})\mathcal{A}_{ggg}(-i\partial_{\vec{b}}) &= 8\pi G s^2\left(e_1.e_3 + \widehat{\alpha}_2 e_1.\partial_b e_3.\partial_b\right) \\ &(e_2^{ij}e_4^{ij} + \alpha_2 e_2^{ij}e_4^{ik}\partial_{b^j}\partial_{b^k} + \alpha_4 e_2^{ij}e_4^{kl}\partial_{b^i}\partial_{b^j}\partial_{b^k}\partial_{b^l}),\end{aligned} \tag{5.116}$$

$$\begin{aligned}\sum \mathcal{A}_{ggg}(-i\partial_{\vec{b}})\mathcal{A}_{ggg}(-i\partial_{\vec{b}}) &= 8\pi G s^2(e_1^{ij}e_3^{ij} + \alpha_2 e_1^{ij}e_3^{ik}\partial_{b^j}\partial_{b^k} + \alpha_4 e_1^{ij}e_3^{kl}\partial_{b^i}\partial_{b^j}\partial_{b^k}\partial_{b^l}) \\ &(e_2^{ij}e_4^{ij} + \alpha_2 e_2^{ij}e_4^{ik}\partial_{b^j}\partial_{b^k} + \alpha_4 e_2^{ij}e_4^{kl}\partial_{b^i}\partial_{b^j}\partial_{b^k}\partial_{b^l}).\end{aligned} \tag{5.117}$$

where all of the operators are acting on the propagator $1/b^{D-4}$. To reproduce the usual shock wave computations only the first three formulas are relevant. In the case of electrodynamics matching with (5.17) is manifest. In the case of Gauss-Bonnet theory we have from (5.20) $\alpha_2 = \frac{\lambda_{GB}}{4}$, $\alpha_4 = 0$. Consider now the coupling of the graviton to a massive spin two particle which is a relevant amplitude for the discussion in the bulk of the chapter. We get

$$\mathcal{A}_{g g \widetilde{g}} = \widetilde{\alpha}_2 \epsilon_{\mu\nu} \left[\epsilon_1^\mu p_3^\nu (\epsilon_3.p_1) + \epsilon_3^\mu p_1^\nu (\epsilon_1.p_3) - p_1^\mu p_3^\nu (\epsilon_1.\epsilon_3) - \epsilon_1^\mu \epsilon_3^\nu (p_1.p_3)\right] \\ \left[(\epsilon_1.\epsilon_3)(p_1.p_3) - (\epsilon_3.p_1)(\epsilon_1.p_3)\right]] \\ + \widetilde{\alpha}_4 \epsilon_{\mu\nu} p_1^\mu p_3^\nu \left[(\epsilon_1.p_3)(\epsilon_3.p_1) - (\epsilon_1.\epsilon_3)(p_1.p_3)\right]^2 . \tag{5.118}$$

Notice that instead of three structures that are present in the massless amplitude we have only two. For the coupling to spin four particle or higher we can have three structures [42] . They take the same form as above with extra indices of particle of spin J being contracted with momenta. An additional structure takes the form

$$\widetilde{c}_{extra} \epsilon_{\mu\nu\rho\sigma}.p_1...p_1 \left[\epsilon_1^\mu p_3^\nu (\epsilon_3.p_1) + \epsilon_3^\mu p_1^\nu (\epsilon_1.p_3) - p_1^\mu p_3^\nu (\epsilon_1.\epsilon_3) - \epsilon_1^\mu \epsilon_3^\nu (p_1.p_3)\right] \\ \left[\epsilon_1^\rho p_3^\sigma (\epsilon_3.p_1) + \epsilon_3^\rho p_1^\sigma (\epsilon_1.p_3) - p_1^\rho p_3^\sigma (\epsilon_1.\epsilon_3) - \epsilon_1^\rho \epsilon_3^\sigma (p_1.p_3)\right] . \tag{5.119}$$

We are interested in the contribution of new particles to the time delay. The computation is almost identical to the one we did for the graviton exchange. The first difference is that we have to compute the Fourier transform of the massive propagator in $D-2$ dimensions. The result is given by $(\frac{m}{b})^{\frac{D-4}{2}} K_{\frac{D-4}{2}}(mb)$ which decays exponentially fast for $mb \gg 1$. The second difference comes from slightly different structure of the three-point functions. Similarly, we can write down the contribution to the time delay of some higher spin particle the difference being that the sum of three-point amplitudes give s^J and we have an extra structure in the three-point amplitude for external gravitons which is an analogous to the Einstein one. The formulas above are valid in $D > 4$, whereas in $D = 4$ α_2 and $\widetilde{\alpha}_2$ structures above are absent[22] but instead we have a parity odd structure. In $D \geq 5$ parity odd structures are absent in the class of amplitudes considered above.

5.9.1 Scattering of a Scalar and a Graviton

As an example let us reproduce the computation for the graviton that we did using the shock wave. It corresponds to scattering from the energetic scalar particle (we could have considered graviton-graviton scattering as well if we average of all polarizations of gravitions 1 and 3).[23]

As an example let us reproduce the computation for the graviton that we did using the shock

[22] They are identically zero.

[23] More precisely we set polarizations one and three to be equal and then we sum over all of them. Physically we are considering a sequence of coherent state with the various alternative polarizations. This makes the α_2 and α_4 contributions vanish in the $\mathcal{A}^{13I}$ part of the amplitude.

wave. It corresponds to scattering from the energetic scalar particle (we could have considered graviton-graviton scattering as well).

$$\sum_{\text{states}} \mathcal{A}^{hhg}(-i\partial_{\vec{b}})\mathcal{A}^{ggg}(-i\partial_{\vec{b}}) \propto Gs^2(e_2^{ij}e_4^{ij} + \alpha_2 e_2^{ij}e_4^{ik}\partial_{b^j}\partial_{b^k} + \alpha_4 e_2^{ij}e_4^{kl}\partial_{b^i}\partial_{b^j}\partial_{b^k}\partial_{b^l})\frac{1}{b^{D-4}} \quad (5.120)$$

Setting $\alpha_4 = 0$, $\alpha_2 = -\frac{\lambda_{GB}}{4}$ we reproduces the computation from the previous section (5.20) . Let us now study the constraints that follow from the positivity of the time delay. Introducing new variables

$$t_2 = \frac{(D-2)(D-4)\alpha_2}{b^2} - \frac{4(D-4)(D-2)D\alpha_4}{b^4} \quad (5.121)$$

$$t_4 = \frac{(D-4)(D-2)D(D+2)\alpha_4}{b^4} \quad (5.122)$$

we can write the phase shift in the form familiar from the study of the energy correlators

$$\delta(s,\vec{b}) \sim 1 + t_2\left(\frac{(e.n)^2}{e.e} - \frac{1}{D-2}\right) + t_4\left(\frac{(e.n)^4}{(e.e)^2} - \frac{2}{D(D-2)}\right) \quad (5.123)$$

which coincides with the formula (3.6) in [194, 50].[24] Here $\vec{n} = \frac{\vec{b}}{|\vec{b}|}$ and $\vec{e}$ is the graviton polarization of particles two and four. Thus, positivity constraints from causality are identical to the ones obtained in their analysis with identification of parameters as above (5.121), namely

$$1 - \frac{t_2}{D-2} - \frac{2t_4}{D(D-2)} \geq 0 \quad (5.124)$$

$$\left(1 - \frac{t_2}{D-2} - \frac{2t_4}{D(D-2)}\right) + \frac{t_2}{2} \geq 0 \quad (5.125)$$

$$\left(1 - \frac{t_2}{D-2} - \frac{2t_4}{D(D-2)}\right) + \frac{D-3}{D-2}(t_2 + t_4) \geq 0 \quad (5.126)$$

So we get the bounds on α_2 and α_4 depending for how small a b we can trust the computation. If the computation is trustworthy for arbitrarily small b we are forced to set α_2 and α_4 to zero.

5.10 Appendix C. The QED case

Note that the action (5.14) also arises in QED (quantum electrodynamics) after we integrate out the massive electron [195] . In that case $\widehat{\alpha} \propto \frac{e^2}{m^2}$. This is a one loop effect, suppressed by the coupling e^2. In this chapter we have taken the coupling to be very small, so that we would have treated this

[24] In [50] $d = D - 1$.

$\widehat{\alpha}$ as being essentially zero. The discussion of this chapter concentrated on the case that the higher curvature corrections were present at tree level, so that the causality problem had to be solved by tree level physics. In this appendix, we consider this loop generated term in QED and we will show that the potential causality problem is solved by one loop effects.

In QED, when we get to an impact parameter of order m^{-1} we cannot be satisfied with the low energy action (5.14) . Fortunately the necessary diagrams were computed in [196] . Using their results it is possible to go to the impact parameter representation (doing the Fourier transform) and check that for $b > 1/m$ we get a result that agrees with the the simple Lagrangian (5.14) , but for $b < 1/m$ we get a different result which displays no causality problem. In other words, the potential causality problems arising from (5.14) appear at $b \sim 1/\sqrt{|\alpha|} \sim e/m$, but at a larger distance, $b \sim 1/m$, the computation should be already modified and we obtain results consistent with causality.[25] We can view this modification as arising from the propagation of electron positron pairs along the t-channel. Notice, that, in addition, when the photon goes through the shock we can have electron positron pair creation. We can view this an another example of tidal excitations. Indeed in QED, the photon has a non-zero probability of being an electron positron pair.

5.11 Appendix D. Causality and Unitarity for a Signal Model

In this appendix we review the constraints from causality and unitarity in the context of a simple signal model. We imagine a signal propagating along one dimension. We have an initial signal which is a function of time $f_{\rm in}(t)$ and an out-signal $f_{\rm out}(t)$ which, in Fourier space is given by $f_{\rm out}(\omega) = S(\omega) f_{\rm in}(\omega)$ or

$$f_{\rm out}(t) = \int dt' \int d\omega S(\omega) e^{-i\omega(t-t')} f_{\rm in}(t') \tag{5.127}$$

Causality implies that if $f_{\rm in}(t') = 0$ for $t' < 0$, then $f_{\rm out}(t) = 0$ for $t < 0$. By unitarity we mean that the L_2 norm of the out-signal should be smaller than that of the in-signal $\int dt |f_{\rm out}(t)|^2 \leq \int dt |f_{\rm in}(t)|^2$. Now it is well known that the Fourier transform of a function which vanishes for $t < 0$ is analytic in the upper half ω plane. This follows directly from the explicit integral expression for the Fourier transform. Then if $f_{\rm in} = 0$ for $t < 0$ we find that both $f_{\rm in}(\omega)$ and $f_{\rm out}(\omega)$ are analytic in the upper half plane. This also implies that $S(\omega)$, which is given their ratio, is also analytic. One might worry that $S(\omega)$ could have poles at zeros of $f_{\rm in}(\omega)$. However, we can change the location of the zeros of

[25] See [197] and references therein for a related discussion of this problem.

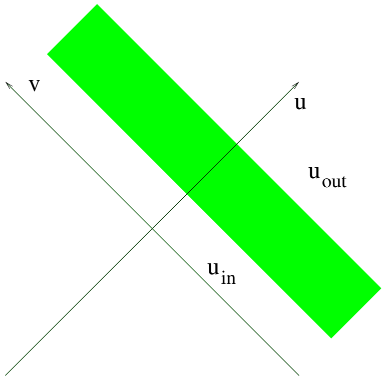


Figure 5.11: We consider a v independent perturbation that is localized in the u direction, given here by the shaded region. We then consider signal propagating along the u direction, which are v dependent and demand causality. We can consider an S matrix that connects the region before the perturbation to the region after the perturbation.

$f_{\text{in}}(\omega)$ by choosing different functions. Therefore $S(\omega)$ is analytic in the upper half plane. We will now prove that unitarity implies that $|S(\omega)| \leq 1$ in the upper half plane. With some foresight, we pick a particular $f_{\text{in}}(t)$ of the form

$$f_{\text{in}}(t) = e^{-\gamma t} e^{-i\omega_0 t} \theta(t) \sqrt{2\gamma} \tag{5.128}$$

with $\gamma > 0$ and ω_0 real. Note that $||f_{\text{in}}||^2 = \int dt |f_{\text{in}}(t)|^2 = 1$. For $Im(\omega) > 0$ we can now write

$$|f_{\text{out}}(\omega)|^2 \leq \left| \int_0^\infty dt e^{i\omega t} f_{\text{out}}(t) \right| \leq \int_0^\infty dt |e^{i\omega t}|^2 \int_0^\infty dt |f_{\text{out}}(t)|^2 = \frac{1}{2Im(\omega)} ||f_{\text{out}}||^2 \tag{5.129}$$

$$|f_{\text{out}}(\omega)|^2 \leq \frac{1}{2Im(\omega)} \tag{5.130}$$

$$|S(\omega)|^2 = \frac{|f_{\text{out}}(\omega)|^2}{|f_{\text{in}}(\omega)|^2} \leq \frac{1}{2Im(\omega)} \frac{1}{|f_{in}(\omega)|^2} \tag{5.131}$$

Here we used the Cauchy-Schwartz inequality. Note that $|e^{i\omega t}|^2 = e^{-2Im(\omega)t}$. We also used that $||f_{\text{out}}||^2 \leq ||f_{\text{in}}||^2 = 1$. We can now set $\omega = \omega_0 + i\gamma$ and find that $f_{\text{in}}(\omega_0 + i\gamma) = 1/\sqrt{2\gamma}$ for the specific function (5.128) . Inserting this into (5.129) we then find that $|S(\omega_0 + i\gamma)| \leq 1$, which is what we wanted to prove, since ω_0 and γ are arbitrary. In conclusion, we find that $S(\omega)$ should be analytic and bounded $|S(\omega)| \leq 1$ in the upper half plane. These are necessary and sufficient conditions.[26] Let us now briefly mention how this is connected to the field theory situation. We consider light cone coordinates u and v. We consider a perturbation that is translation invariant in v but is localized in the u coordinate. We call this "the shock". We expand the fields in the

[26]Note that some functions which are analytic in the upper half plane, such as $S(\omega) = e^{i\omega^3}$ are actually not causal.

v coordinate at some u_{in} and then we expand them again at some u_{out} after the shock. To make contact with the above discussion we call $t = v$ and $p_v = -\omega$. We can expand the field as

$$\phi(t) = \int_0^\infty \frac{d\omega}{\sqrt{\omega}} (a_\omega e^{-i\omega t} + a_\omega^\dagger e^{i\omega t}) \tag{5.132}$$

We can do this for ϕ_{in} and ϕ_{out} in terms of a_{in} and a_{out}. These oscillators then are related by

$$a_{\omega,\text{out}} = S(\omega) a_{\omega,\text{in}} \ , \qquad a^\dagger_{\omega,\text{out}} = S(\omega)^* a^\dagger_{\omega,\text{in}} \tag{5.133}$$

This defines $S(\omega)$ for positive ω. For negative ω we can define $S(-\omega) = S(\omega)^*$. Alternatively, we can define

$$S(\omega) = -\int_{-\infty}^{\infty} dt e^{i\omega t} [\phi_{\text{out}}(t), i\partial_t \phi_{\text{in}}(0)] = -\int_0^\infty dt e^{i\omega t} [\phi_{\text{out}}(t), i\partial_t \phi_{\text{in}}(0)] \tag{5.134}$$

The commutation relations for the in and out oscillators require that $|S(\omega)|^2 = 1$. In a case where there is particle mixing, but no particle creation, then the fields have indices ϕ^i_{in} and ϕ^j_{out}. Now S is a matrix which obeys $S_{ij}(-\omega) = S_{ij}(\omega)^*$ and $S_{ij}(\omega) S_{kj}(\omega)^* = \delta_{ik}$ due to the commutation relations of the oscillators before and after the shock. In the preceding discussion we have neglected the transverse dimensions. We can now remedy that by including the momentum in the transverse dimensions as part of the indices we are discussing here. If we consider a signal that is made out of physical particles, one might correctly worry that the fact that $\omega > 0$ will preclude us from localizing the signal in time. In order to avoid this issue we can consider a coherent state of the form

$$|\Psi\rangle = e^{i \int dt f_{\text{in}}(t) \phi_{\text{in}}(t)} |0\rangle \tag{5.135}$$

with a real function f_{in}. This is a state that could be produced by adding a hermitian term to the Hamiltonian at some early time u_{in}. On this state we have the expectation values

$$\langle \Psi | \partial_t \phi_{\text{in}}(t) | \Psi \rangle = f_{\text{in}}(t) \ , \qquad \langle \Psi | \partial_t \phi_{\text{out}}(t) | \Psi \rangle = f_{\text{out}}(t) \tag{5.136}$$

where the functions are related as in the signal model.[27] Here we assumed a linear relation between the in- and out-signals. Furthermore, we can also consider the expectation values of the normal ordered product $T_{vv} = T_{tt} =: \partial_t \phi(t) \partial_t \phi(t) :$. When this is evaluated on the state (5.135) , and

[27] f_{out} is real if f_{in} is real when $S(-\omega) = S(\omega)^*$.

integrated over t we find that the answer is given by

$$-P_v^{\text{in,out}} = \int dv T_{vv} = \int dt (f_{\text{in,out}}(t))^2 = ||f_{\text{in,out}}||^2 \tag{5.137}$$

Thus, the condition that the total light-cone momentum P_v should not increase implies the norm condition $||f_{\text{out}}||^2 \leq ||f_{\text{in}}||^2$. More precisely, we can consider the signal f_{in} exciting a mode involving a graviton with a given polarization. The signal f_{out} is the same mode of the graviton. In addition, the initial graviton could go into other massive particles. Then the condition that the total P_v in the out-graviton mode should be no bigger than the initial P_v, which was all contained in the graviton mode, leads to the norm condition (or unitarity condition) for the signal model. In conclusion, the graviton-graviton matrix element $S_{gg}(\omega)$ obeys all the assumptions of the signal model. Therefore, it should be analytic and $|S_{gg}(\omega)| \leq 1$ in the upper half plane. Note that we have assumed here a perfect v-translation symmetry for the perturbation that creates the shock. In our scattering problem, see (5.3) , particles 1 and 3 have small p_v momentum. Thus, in this discussion, we have neglected this small momentum. This is reasonable for $sb^2 \gg 1$.

5.12 Appendix E. Scattering in String Theory

String theory in flat space is the simplest example of a theory that follows into the category of weakly coupled gravitational theories with higher derivative corrections that are subject of our analysis. As explained in the introduction, a motivation for this work was actually to argue that string theory is inevitable, at least, under certain assumptions. It is well known that effective gravitational actions in string theory contain higher derivative corrections at the string scale [198, 199]. In particular graviton three-point amplitudes can contain the higher derivative terms that we constrained in this chapter using causality. The potential problem is fixed by extra particles with string scale masses. Here we would like to understand how this is happening in detail. Let us first recall the form of the three-point gravity amplitudes in bosonic, heterotic and type II string theories [200]

$$\mathcal{A}^{ggg} = \sqrt{32\pi G}\epsilon_1^{\mu_1\mu_1'}\epsilon_2^{\mu_2\mu_2'}\epsilon_3^{\mu_3\mu_3'}N_{\mu_1\mu_2\mu_3}\overline{N}_{\mu_1'\mu_2'\mu_3'} \tag{5.138}$$

$$N^{\mu_1\mu_2\mu_3} = k_2^{\mu_1}\eta^{\mu_2\mu_3} + k_3^{\mu_2}\eta^{\mu_1\mu_3} + k_1^{\mu_3}\eta^{\mu_1\mu_2} + \frac{\alpha'}{2}\epsilon k_2^{\mu_1}k_3^{\mu_2}k_1^{\mu_3} \tag{5.139}$$

and we have $\epsilon_{bos} = \bar{\epsilon}_{bos} = \epsilon_{het} = 1$ and $\epsilon_{II} = \bar{\epsilon}_{II} = \bar{\epsilon}_{het} = 0$. Translating it to our notations we get

$$\alpha_2^{bos} = 2\alpha_2^{het} = \alpha', \qquad \alpha_2^{II} = 0, \tag{5.140}$$

$$\alpha_4^{bos} = \frac{(\alpha')^2}{4}, \qquad \alpha_4^{het} = \alpha_4^{II} = 0. \tag{5.141}$$

The vanishing of some of these corrections can be understood from supersymmetry, as explained in section 3.2. For the purposes of this chapter, the type II case is not interesting since there are not corrections at all for the graviton three-point function. The high energy scattering problem in string theory was studied in a nice series of papers by Amati, Ciafaloni and Veneziano [173, 174, 175, 176, 177], see also e.g. [**?**, 201]. Let us first review their picture. The scattering can be described in terms of a phase shift defined as

$$\delta(\vec{b}, s) = [\mathrm{POL}]\delta^{ACV}(\vec{b}, s), \tag{5.142}$$

$$\delta^{ACV}(\vec{b}, s) = \int \frac{d^{D-2}\vec{q}}{(2\pi)^{D-2}} e^{i\vec{q}.\vec{b}} C(s,t,u), \tag{5.143}$$

$$C(s,t,u) = \frac{\Gamma(-\frac{\alpha' s}{4})\Gamma(-\frac{\alpha' t}{4})\Gamma(-\frac{\alpha' u}{4})}{\Gamma(1+\frac{\alpha' s}{4})\Gamma(1+\frac{\alpha' t}{4})\Gamma(1+\frac{\alpha' u}{4})}, \tag{5.144}$$

where [POL] represents a factor that depends on the polarizations and is polynomial in the momenta. We will only need its form in a specific limit. In the high energy limit $C(s,t,u)$ has the celebrated Regge behavior

$$C(s,t,u) \sim \frac{\Gamma(-\frac{\alpha' t}{4})}{\Gamma(1+\frac{\alpha' t}{4})} \left(-i\frac{s\alpha'}{4}\right)^{-2+\frac{t\alpha'}{2}}. \tag{5.145}$$

This Regge form is reflecting the creation of particles in the s-channel. The infinite sequence of s-channel poles is becoming a cut, the cut arising when $s \to se^{2\pi i}$. The creation of the massive s-channel states is also related to the fact that we get an imaginary part from the $(-i)^{\frac{t\alpha'}{2}}$ factor in (5.145) is saying that the most likely process is to create a massive closed string, rather than scattering the gravitons.

For large s, only small q will contribute and we can approximate the prefactor in (5.145) by $1/t$. Then the integral becomes

$$\delta^{[ACV]} \propto \int \frac{d^{D-2}\vec{q}}{2(2\pi)^{D-2}} \frac{e^{i\vec{q}.\vec{b}}\left(-i\frac{s\alpha'}{4}\right)^{-\frac{\vec{q}^2\alpha'}{2}}}{\vec{q}^2} = \frac{1}{(\frac{\alpha' Y}{2})^{\frac{D-4}{2}}} \int_0^1 d\rho \rho^{\frac{D-6}{2}} e^{-\frac{b^2}{2\alpha' Y}\rho} \tag{5.146}$$

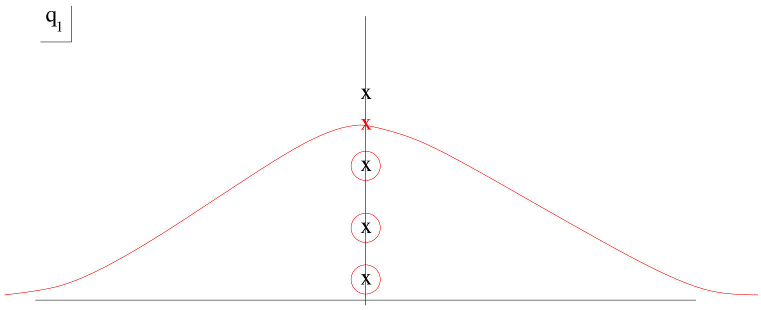


Figure 5.12: We display the complex q_1 plane. We have displayed the poles in the t-channel by black crosses. The saddle point (5.148) of the Gaussian integral has been denoted by a red cross. We have shifted the original integration contour for q_1, which was along the real axis, to the complex plane so that it passes through the saddle point (5.148) . In the process we have picked up some of the poles in the t-channel.

where $Y = \log(-is\alpha'/4)$. We have two characteristic behaviors, depending on whether b^2 is larger or smaller than $\alpha' \log(s\alpha')$. For large b we get the usual $1/b^{D-4}$ behavior. For small b we get the result in brackets in (5.62) . Note that since the transition region occurs for a b^2 which is larger than α' by a $\log s$ factor, we can, a posteriori, justify the fact that we have approximated the prefactor in (5.145) by $1/t$.

In our field theory discussion we had represented the phase shift as a sum over poles. We can wonder how this applies to the string theory discussion. Notice that from (5.145) we get a Gaussian integrand factor of the form

$$e^{i\vec{q}.\vec{b}} e^{-(\vec{q})^2 \frac{\alpha'}{2} \log(-is\alpha'/4)} \tag{5.147}$$

Let us assume that $\vec{b} = (b, 0, \cdots, 0)$, so that it points along the first coordinates. When we do the integral over the first component of $\vec{q}$, call it q_1, we get a saddle point for (5.147) at

$$q_s = i \frac{b}{\alpha' \log(-is\alpha'/4)} \tag{5.148}$$

It is thus convenient to shift the contour to this location, (5.148) , where this saddle point contribution gives something of the order of

$$e^{-\frac{b^2}{2\alpha' \log(-is\alpha'/4)}} \tag{5.149}$$

This is not the whole answer, since by shifting the contour to this location, we can pick up some poles from the prefactor (5.145) , see (5.12) . We always pick up the pole at $t = 0$[28], which was the center of our gravity discussion, but we can also pick up the poles at $t = \frac{4}{\alpha'} n$ for $n < -q_s^2$, where q_s is given in (5.148) . Notice that at these saddles the t dependent part of the exponent gives us factors of s^{2n} as we expect for the corresponding spins[29], once we take into account that [POL] contains a factor of s^4. When b is large $b^2 \gg \alpha' \log(s\alpha')$, we pick up many poles, but when b is small we only pick up the $t = 0$ pole, but, even then, the integral is more accurately computed using (5.146) . Note that the factor [POL] in string theory phase simplifies at $t = 0$ and becomes the product of three-point amplitudes we discussed in the body of the chapter and that we added as Pol in (5.62) . In other words [POL]$\to$Pol as $t \to 0$. In particular, note that the residues of poles associated to the massive states go as $\frac{1}{(n!)^2} e^{-b\sqrt{\frac{4n}{\alpha'}}} s^{2n}$. As a function of n, these contributions decrease and then start increasing again with a transition at a value of n corresponding to the saddle point (5.148) [30]. We can ask: Why don't we include all poles in the t-channel?. If we were to include all poles in the t-channel, we would obtain the wrong answer. The reason it is wrong is that in string theory the argument that we can shift the contour is not correct because of contributions for large values of q. Such large values of q were never meant to be included in the integral, since the kinematics of the process we consider restricts the real values of $\vec{q}^2$ to be much smaller than s. In deriving the physical picture, we certainly assumed that $\vec{q}^2 \ll s$.[31] Indeed, if we look at (5.145) we get a very small contribution from large real values of $\vec{q}^2$. On the other hand, if we were to keep s fixed and we formally look at large real values of $\vec{q}^2$ in the original expression, (5.142) , then we would encounter the u channel poles. The conclusion is that shifting the contour for the q_1 integral, while it can be done formally, it does not represent the real physical computation we want to do. Approximating the integrand using (5.145) , and then integrating gives the physically correct answer. One can qualitatively say that, for large $b^2/(\alpha' \log s\alpha')$, we get a contribution of some of the t-channel poles, as in (5.12) , and then the rest of the poles are completely resummed via the saddle point integral in (5.12) . Their contribution should be better thought of as coming from the creation of extended objects in the s-channel. Another remark we want to make is the following. It was shown in [159] that the plane wave solution is a solution to all orders in the α' expansion. This gravitational plane wave encodes the contribution from the $t = 0$ pole. However, we have seen that sometimes we get a subleading contribution from the other poles, due to massive states along the t-channel. These

[28] The location of this pole in the q_1 plane depends on $\vec{q}_{\text{rest}}$ which we take to be real.

[29] For a closed string with $N_L = N_R = n$, the maximum spin is $J = 2 + 2n$.

[30] Here we are assuming that both b^2 and $\log s\alpha'$ are large with a ratio larger than one but fixed, so that we can neglect the $1/(n!)^2$ in this discussion. This last factor makes the sum converge for large n, but this convergent answer, as discussed below, is not the correct one.

[31] Momentum conservation and on-shell conditions impose $q^2 < \frac{s}{4}$.

mean that the physical scattering process contains extra contributions not captured by the plane wave[32].

In string theory we can also take into account the tidal excitations. In this case the phase shift can be viewed as an operator that maps the two initial gravitons to two final generic string states. [173, 174, 175, 176, 177] have shown that this operator has the remarkably simple expression $\widehat{\delta} \propto \int d\sigma d\sigma' \delta_{\rm grav}(\widehat{X}_L(\sigma) - \widehat{X}_R(\sigma'))$ where X_L and X_R are the transverse space positions of the string on the worldsheet and $\delta_{\rm grav}(b)$ is the ordinary gravity phase shift. This is valid for distances $b^2 \gg \alpha' \log(s\alpha')$. The effects of these tidal excitations do not help in resolving the causality issues discussed here and are unrelated to the appearance of closed strings in the s-channel discussed above. See [173, 174, 175, 176, 177] for further discussion.

5.13 Appendix F. Properties of the AdS Shock Wave

Let us first examine the problem of higher derivative corrections for the AdS shock wave. As in the case of flat space the shock wave at hand continue to be an exact solution when arbitrary higher derivative corrections are included. The argument for this is identical to the one in section 5 of [182] . The vector $l_\mu = \partial_\mu u = \{1, 0, 0, 0..\}$ in coordinates u, v, y_i, z. We can now compute the vector V_μ that the argument talks about. We find $V_\mu = \{0, \cdots, 0, -2/z\}$. In other words $V_z = -2/z$ and the rest of the components are zero. This obeys that $V_\mu l^\mu = 0$ are required in their argument. In addition, one can also compute the first order Riemann tensor $\check{R}_{\mu\nu\delta\sigma}$. It is indeed of the form stated in [182] with the symmetric tensor K given by

$$K_{y^i,y^j} = \frac{1}{2}(\delta_{ij}\partial_z h/z^3 - \partial_i\partial_j h/z^2)\ , \quad K_{z,y^i} = K_{y^i,z} = -\frac{\partial_{y^i}\partial_z h}{z^2}\ , \quad K_{zz} = \frac{1}{2}(\partial_z h/z^3 - \partial_z^2 h/z^2) \tag{5.150}$$

with the rest of the components equal to zero. This K obeys that $K_{\mu\nu} l^\nu = 0$ as required by [182] . Notice that this form of K also leads to the equation of the form (5.66), when the Riemann tensor is used in Einstein's equations, namely $K^\mu_{\ \mu} = 0$. Using this shock wave we can once again compute the time delay for different theories. It is convenient to evaluate Riemann tensor in the coordinates that make rotation symmetry manifest. The result is

$$\widehat{R}_{uiju}|_{\vec{y}=0;z=1} = K_{ij} = -f(u)\left(1-\rho^2\right)\frac{\varpi'(\rho) - \rho\varpi''(\rho)}{8\rho}\left(n^i n^j - \frac{1}{D-2}\delta^{ij}\right) \tag{5.151}$$

[32]This is not in contradiction with [159] , since these extra terms can be viewed as a non-perturbative contribution in α'.

where $i = 1, .., D-2$ so that we span $\vec{y}$ and z components. Below we consider propagation of different perturbations in the shock wave background described above. We first write the most general form of the second order equations of motion and then compute the time delay in the high energy limit. Let us consider several limits of the shock wave to make contact with previous investigations of similar type. First, we expect to recover the flat space shock wave for probes that come close enough to the center of the shock or, equivalently, for $\rho \to 0$. Indeed, it is easy to check that this asymptotic is correctly recovered (5.73) . Let us consider couple of other limits. We can fix $\vec{y}_0$ and send $z_0 \to 0$ in which case we have a source at the boundary and the shock wave takes the form

$$h \sim \frac{z^{D-3}}{(z^2 + |\vec{y} - \vec{y}_0|^2)^{D-2}}, \tag{5.152}$$

which is exactly the shock wave considered in the energy correlator problem [20]. In the opposite limit $z_0 \to \infty$ we get

$$h \sim z^{D-3}, \tag{5.153}$$

and the time delay in this background was computed, for example, in [183] .

5.14 Appendix G. Time Advances and Time Machines

In this appendix we would like to argue that a negative time delay enables us to build a time machine which leads to closed time-like curves. This is a standard argument [202] and the only thing we check is that the long range gravitational forces do not prevent us from setting it up. The setup we would like to consider is the following. We have two shock waves that correspond to energetic particles with momenta $q_{1,u} = \frac{\sqrt{s}}{2}$ and $q_{2,v} = \frac{\sqrt{s}}{2}$ separated by distance r in the transverse plane. The first shock is localized around $u = 0$ and the second one around $v = 0$. We would like the separation to be such that $r \gg r_S$, namely we are in the regime where the black hole formation does not occur

$$r^{D-3} \gg r_S^{D-3} = G\sqrt{s}. \tag{5.154}$$

Next we would like to consider a test particle that propagates through both shocks in such a way that it ends up at the same position where it started, thus, forming a closed time-like curve. The situation is depicted on (5.13) . When the test particle crosses each of the shocks it gets shifted by

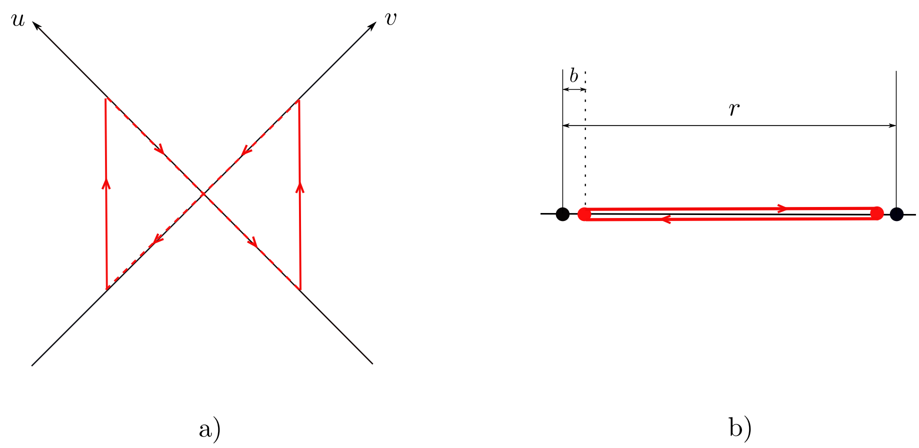


Figure 5.13: a) We imagine a background that consists of two shock waves located at $u = 0$ and $v = 0$ widely separated in the transverse directions which is not presented on the figure.The arrows show the motion of the probe massless particle projected on to the u, v plane. b) Same motion but projected on the transverse plane. The two background shocks are separated by r and the probe passes at a short distance b from each of them. The vertical region of the path in (a) corresponds to the horizontal motion in (b). We can build a closed time-like curve as depicted on the picture by crossing this pair of shocks if time advances are allowed. We need mirrors to reverse the motion in the transverse plane as we pass through the shocks.

$\Delta v = \Delta u \sim \frac{G\sqrt{s}}{b^{D-4}}$. Between the shock the particle travels the distance of order r. Thus, we want the time shift to be $\frac{G\sqrt{s}}{b^{D-4}} = \left(\frac{r_S}{b}\right)^{D-3} b \sim r \gg r_S$ which becomes

$$\left(\frac{r_S}{b}\right)^{D-4} \gg 1. \tag{5.155}$$

We also want $b \gg \widetilde{r}_S$ where $\widetilde{r}_S$ is the Schwarzschild radius for the shock wave-test particle pair, $\widetilde{r}_S^{D-3} \sim G\sqrt{p\sqrt{s}}$, where p is the energy of the probe particle. Together with (5.155) it implies $\sqrt{s} \gg p$ where p is the energy of the probe particle. We also need that $pb \gg 1$. The conclusion is that in $D > 4$ we can construct closed time-like curve using negative time delays by choosing b,r and s appropriately. For example, if we have a causality problem that appears at a scale $b^2 \sim \alpha_2$, then we take this value for b. Since we are at weak coupling, we know that the Planck length $l_p \ll b$. We can then pick $\sqrt{s}l_p \sim X^{1+a}$,$pl_p \sim X^{1-a'}$, with $X = (b/l_p)^{D-3}$ and we can choose $a > 0$, $a - a' < 0$, $1 - a' + 1/(D-3) > 0$ to ensure that $r_S \gg b$, $\widetilde{r}_S \ll b$ and $pb \gg 1$. We can achieve this with $a' = 1/2$ and $a = 1/4$, for example.

5.15 Appendix H. Representations That Couple to Two Gravitons

Here we would like to understand better what are the representations of little group $SO(D-1)$ of massive particles that can couple to two gravitons. We are interested only in bosonic fields since a single fermion does not couple to two gravitons.

Let us consider the decay of a massive particle in its rest frame so that $p_2 = (M, \vec{0})$. Gravitons produced have $\vec{p}_1 = -\vec{p}_3 = \vec{p}$. We characterize the original particle by some polarization tensor $e_{i_1 \dots i_k}$ which has only spatial components and is traceless with respect to any pair of indices. We do not specify the symmetry properties of this tensor yet. We characterize gravitons by polarization tensors e_1 and e_3 such that $\vec{e}_1.\vec{p} = \vec{e}_3.\vec{p} = 0$. We have three type of contractions (see also [42])

$$\mathcal{A}_1 = e_{i_1 \dots i_k} p^{i_1} \dots p^{i_k} (e_1.e_3)^2, \tag{5.156}$$

$$\mathcal{A}_2 = e_{i_1 \dots i_k} e_1^{i_1} e_3^{i_2} p^{i_3} \dots p^{i_k} (e_1.e_3), \tag{5.157}$$

$$\mathcal{A}_3 = e_{i_1 \dots i_k} e_1^{i_1} e_1^{i_2} e_3^{i_3} e_3^{i_4} p^{i_5} \dots p^{i_k}. \tag{5.158}$$

The first amplitude $\mathcal{A}_1$ exists only for particles in the symmetric traceless representations (Young tableau that consists of single horizontal row with k boxes). Actually all three amplitudes are are allowed for symmetric representation and we discussed them in the bulk of the chapter in detail. For the second and third amplitude we can add more rows to the Young diagram. Properties of these, so-called, mixed-symmetry tensors are nicely reviewed, for example, in appendix E of [203]. By thinking about the representation in terms of tensors which are manifestly anti-symmetric with respect to indices in a given column we can read off possible representations. In particular the fact that we have only three different vectors means that we can have at most three rows. Let us write all the amplitudes in a covariant manner. The general prescription is the following

$$e_1^i e_3^j \to E_{13}^{\mu\nu} \equiv \epsilon_1^\mu p_3^\nu (\epsilon_3.p_1) + \epsilon_3^\mu p_1^\nu (\epsilon_1.p_3) - p_1^\mu p_3^\nu (\epsilon_1.\epsilon_3) - \epsilon_1^\mu \epsilon_3^\nu (p_1.p_3). \tag{5.159}$$

We then have for the amplitudes

$$\mathcal{A}_1 = \epsilon_{\mu_1 \ldots \mu_k} p_1^{\mu_1} \ldots p_1^{\mu_k} \left[(\epsilon_1.\epsilon_3)(p_1.p_3) - (\epsilon_1.p_3)(\epsilon_3.p_1)\right]^2 , \tag{5.160}$$

$$\mathcal{A}_2 = \epsilon_{\mu_1 \ldots \mu_k} E_{13}^{\mu_1 \mu_2} p_1^{\mu_3} \ldots p_1^{\mu_k} \left[(\epsilon_1.\epsilon_3)(p_1.p_3) - (\epsilon_1.p_3)(\epsilon_3.p_1)\right] , \tag{5.161}$$

$$\mathcal{A}_3 = \epsilon_{\mu_1 \ldots \mu_k} E_{13}^{\mu_1 \mu_3} E_{13}^{\mu_2 \mu_4} p_1^{\mu_5} \ldots p_1^{\mu_k} . \tag{5.162}$$

To compute the time delay we need to use the completeness relation

$$\sum_i \epsilon^*_{\nu_1 \ldots \nu_k} \epsilon_{\mu_1 \ldots \mu_k} = \Pi_{\nu_1 \ldots \nu_k | \mu_1 \ldots \mu_k} \tag{5.163}$$

and contract both sides of the amplitude, where Π is a projector on to the space orthogonal to the intermediate momentum. As the next step we would like to focus on those diagrams that produce s^a with $a \geq 2$ for the amplitude. Everything that is less is irrelevant for causality violation. In the language of Young tableaux it corresponds to having $a \geq 2$ boxes in the first row. It will be curious to understand if mixed symmetry fields with $a = 2$ can resolve the causality problem we observed in the bulk of the chapter. As an example of such field would be $(2, 2)$ (a square with four boxes) or $(2, 2, 2)$ fields (a vertical rectangle of 2×3 boxes). There is a short list of representations of this kind. One could wonder whether these particles alone (without the infinite tower of higher spin particles) could solve the causality problem. We think that the answer is no. We leave a full exploration of this question for the future.

Bibliography

[1] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[2] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1 (2012) [arXiv:1207.7214 [hep-ex]].

[3] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716**, 30 (2012) [arXiv:1207.7235 [hep-ex]].

[4] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975) [Erratum-ibid. **46**, 206 (1976)].

[5] P. A. R. Ade *et al.* [BICEP2 Collaboration], Phys. Rev. Lett. **112**, 241101 (2014) [arXiv:1403.3985 [astro-ph.CO]].

[6] J. M. Maldacena, Adv. Theor. Math. Phys. **2**, 231 (1998). [hep-th/9711200].

[7] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, Phys. Lett. B **428**, 105 (1998). [hep-th/9802109].

[8] E. Witten, Adv. Theor. Math. Phys. **2**, 253 (1998). [hep-th/9802150].

[9] J. Polchinski, arXiv:1010.6134 [hep-th].

[10] A. Strominger, Talk at Strings 2014, http://physics.princeton.edu/strings2014/slides/Strominger.pdf

[11] I. R. Klebanov and A. M. Polyakov, Phys. Lett. B **550**, 213 (2002). [hep-th/0210114].

[12] S. Giombi and X. Yin, JHEP **1009**, 115 (2010). [arXiv:0912.3462 [hep-th]].

[13] P. W. Anderson, Science **177**, no. 4047, 393 (1972).

[14] P. Dorey, hep-th/9810026.

[15] S. R. Coleman, J. Mandula, Phys. Rev. **159**, 1251-1256 (1967).

[16] R. Haag, J. T. Lopuszanski, M. Sohnius, Nucl. Phys. **B88**, 257 (1975).

[17] D. J. Gross, Phys. Rev. Lett. **60**, 1229 (1988).

[18] G. W. Moore, hep-th/9310026.

[19] C. L. Basham, L. S. Brown, S. D. Ellis and S. T. Love, Phys. Rev. D **17**, 2298 (1978).

[20] D. M. Hofman and J. Maldacena, JHEP **0805**, 012 (2008). [arXiv:0803.1467 [hep-th]].

[21] A. B. Zamolodchikov, A. B. Zamolodchikov, Annals Phys. **120**, 253-291 (1979).

[22] V. A. Fateev, S. L. Lukyanov, Int. J. Mod. Phys. **A3**, 507 (1988).

[23] E. S. Fradkin and M. A. Vasiliev, Nucl. Phys. B **291**, 141 (1987).

[24] M. A. Vasiliev, In *Shifman, M.A. (ed.): The many faces of the superworld* 533-610. [hep-th/9910096].

[25] E. Sezgin and P. Sundell, Nucl. Phys. B **644**, 303 (2002), [Erratum-ibid. B **660**, 403 (2003)]. [hep-th/0205131].

[26] S. Giombi and X. Yin, JHEP **1104**, 086 (2011). [arXiv:1004.3736 [hep-th]].

[27] D. Anninos, T. Hartman and A. Strominger, [arXiv:1108.5735 [hep-th]].

[28] S. Giombi and X. Yin, [arXiv:1105.4011 [hep-th]].

[29] R. d. M. Koch, A. Jevicki, K. Jin and J. P. Rodrigues, Phys. Rev. D **83**, 025006 (2011). [arXiv:1008.0633 [hep-th]].

[30] M. R. Douglas, L. Mazzucato and S. S. Razamat, Phys. Rev. D **83**, 071701 (2011). [arXiv:1011.4926 [hep-th]].

[31] Solvay workshop proceedings , http://www.ulb.ac.be/sciences/ptm/pmif/HSGT.htm

[32] M. Bianchi and V. Didenko, [hep-th/0502220].

[33] D. Francia and C. M. Hull, [hep-th/0501236].

[34] S. de Buyl and A. Kleinschmidt, [hep-th/0410274].

[35] A. C. Petkou, [hep-th/0410116].

[36] N. Bouatta, G. Compere and A. Sagnotti, [hep-th/0409068].

[37] X. Bekaert, S. Cnockaert, C. Iazeolla and M. A. Vasiliev, [hep-th/0503128].

[38] N.T.Evans, J.Math.Phys. 8 (1967) 170-185.

[39] M. G. Eastwood, Annals Math. **161**, 1645 (2005). [hep-th/0206233].

[40] S. El-Showk, K. Papadodimas, [arXiv:1101.4163 [hep-th]].

[41] S. Giombi, S. Prakash, X. Yin, [arXiv:1104.4317 [hep-th]].

[42] M. S. Costa, J. Penedones, D. Poland and S. Rychkov, JHEP **1111**, 071 (2011). [arXiv:1107.3554 [hep-th]].

[43] S. H. Shenker and X. Yin, [arXiv:1109.3519 [hep-th]].

[44] S. Giombi, S. Minwalla, S. Prakash, S. P. Trivedi, S. R. Wadia and X. Yin, Eur. Phys. J. C **72**, 2112 (2012). [arXiv:1110.4386 [hep-th]].

[45] O. Aharony, G. Gur-Ari and R. Yacoby, JHEP **1203**, 037 (2012). [arXiv:1110.4382 [hep-th]].

[46] A. Sagnotti and M. Taronna, Nucl. Phys. B **842**, 299 (2011). [arXiv:1006.5242 [hep-th]].

[47] I. Heemskerk, J. Penedones, J. Polchinski and J. Sully, JHEP **0910**, 079 (2009). [arXiv:0907.0151 [hep-th]].

[48] A. L. Fitzpatrick, E. Katz, D. Poland and D. Simmons-Duffin, JHEP **1107**, 023 (2011). [arXiv:1007.2412 [hep-th]].

[49] J. M. Maldacena and G. L. Pimentel, JHEP **1109**, 045 (2011). [arXiv:1104.2846 [hep-th]].

[50] A. Buchel, J. Escobedo, R. C. Myers, M. F. Paulos, A. Sinha and M. Smolkin, JHEP **1003**, 111 (2010). [arXiv:0911.4257 [hep-th]].

[51] F. A. Dolan and H. Osborn, [arXiv:1108.6194 [hep-th]].

[52] A. Mikhailov, [hep-th/0201019].

[53] S. E. Konstein, M. A. Vasiliev and V. N. Zaikin, JHEP **0012**, 018 (2000). [hep-th/0010239].

[54] M. A. Vasiliev, Phys. Lett. B **567**, 139 (2003). [hep-th/0304049].

[55] M. A. Vasiliev, [arXiv:1203.5554 [hep-th]].

[56] B. Sundborg, Nucl. Phys. Proc. Suppl. **102**, 113 (2001). [hep-th/0103247].

[57] E. Witten, Talk at the John Schwarz 60-th Birthday Symposium, http://theory.caltech.edu/jhs60/witten/1.html.

[58] E. Sezgin and P. Sundell, Nucl. Phys. B **644**, 303 (2002), [Erratum-ibid. B **660**, 403 (2003)]. [hep-th/0205131]

[59] E. Sezgin and P. Sundell, JHEP **0507**, 044 (2005). [hep-th/0305040].

[60] J. Maldacena and A. Zhiboedov, [arXiv:1112.1016 [hep-th]].

[61] L. Girardello, M. Porrati and A. Zaffaroni, Phys. Lett. B **561**, 289 (2003). [hep-th/0212181].

[62] W. Heidenreich, J. Math. Phys. **22**, 1566 (1981).

[63] T. Hertog and K. Maeda, JHEP **0407**, 051 (2004). [hep-th/0404261].

[64] S. Sachdev, Phys. Lett. B **309**, 285 (1993). [hep-th/9305131].

[65] D. J. Gross and A. Neveu, Phys. Rev. D **10**, 3235 (1974)..

[66] B. Rosenstein, B. J. Warr and S. H. Park, Phys. Rev. Lett. **62**, 1433 (1989).

[67] D. Anselmi, Nucl. Phys. B **541**, 369 (1999). [hep-th/9809192].

[68] J. Henn, C. Jarczak and E. Sokatchev, Nucl. Phys. B **730**, 191 (2005). [hep-th/0507241].

[69] M. Bianchi, J. F. Morales and H. Samtleben, JHEP **0307**, 062 (2003). [hep-th/0305052].

[70] G. S. Ng and A. Strominger, [arXiv:1204.1057 [hep-th]].

[71] M. R. Gaberdiel and R. Gopakumar, Phys. Rev. D **83**, 066007 (2011). [arXiv:1011.2986 [hep-th]].

[72] M. R. Gaberdiel and T. Hartman, JHEP **1105**, 031 (2011). [arXiv:1101.2910 [hep-th]].

[73] M. R. Gaberdiel, R. Gopakumar, T. Hartman and S. Raju, JHEP **1108**, 077 (2011). [arXiv:1106.1897 [hep-th]].

[74] C. -M. Chang and X. Yin, [arXiv:1106.2580 [hep-th]].

[75] K. Papadodimas and S. Raju, Nucl. Phys. B **856**, 607 (2012). [arXiv:1108.3077 [hep-th]].

[76] C. -M. Chang and X. Yin, [arXiv:1112.5459 [hep-th]].

[77] W. Ruhl, [hep-th/0607197].

[78] D. Anselmi, Class. Quant. Grav. **17**, 1383 (2000). [hep-th/9906167].

[79] D. I. Kazakov, Phys. Lett. B **133**, 406 (1983)..

[80] D. J. Gross, Phys. Rev. Lett. **60**, 1229 (1988)..

[81] G. W. Moore, In *Boston 1993, Proceedings, Supersymmetry and unification of fundamental interactions, SUSY 93* 540-550. [hep-th/9308052].

[82] A. Sagnotti, [arXiv:1112.4285 [hep-th]].

[83] A. M. Polyakov, JETP Lett. **12**, 381 (1970), [Pisma Zh. Eksp. Teor. Fiz. **12**, 538 (1970)].

[84] I. Antoniadis and M. Buican, Phys. Rev. D **83**, 105011 (2011). [arXiv:1102.2294 [hep-th]].

[85] M. A. Luty, J. Polchinski and R. Rattazzi, [arXiv:1204.5221 [hep-th]].

[86] J. -F. Fortin, B. Grinstein and A. Stergiou, [arXiv:1206.2921 [hep-th]].

[87] J. -F. Fortin, B. Grinstein and A. Stergiou, [arXiv:1208.3674 [hep-th]].

[88] Y. Nakayama, [arXiv:1208.4726 [hep-th]].

[89] J. -F. Fortin, B. Grinstein, C. W. Murphy and A. Stergiou, [arXiv:1210.2718 [hep-th]].

[90] Y. Nakayama, [arXiv:1210.6439 [hep-th]].

[91] G. Mack, Commun. Math. Phys. **55**, 1 (1977)..

[92] B. Grinstein, K. A. Intriligator and I. Z. Rothstein, Phys. Lett. B **662**, 367 (2008). [arXiv:0801.1140 [hep-ph]].

[93] T. N. Pham and T. N. Truong, Phys. Rev. D **31**, 3027 (1985)..

[94] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis and R. Rattazzi, JHEP **0610**, 014 (2006). [hep-th/0602178].

[95] Z. Komargodski and A. Schwimmer, JHEP **1112**, 099 (2011). [arXiv:1107.3987 [hep-th]].

[96] Z. Komargodski, JHEP **1207**, 069 (2012). [arXiv:1112.4538 [hep-th]].

[97] V. M. Braun, G. P. Korchemsky and D. Mueller, Prog. Part. Nucl. Phys. **51**, 311 (2003). [hep-ph/0306057].

[98] O. Nachtmann, Nucl. Phys. B **63**, 237 (1973)..

[99] H. Epstein, V. Glaser and A. Martin, Commun. Math. Phys. **13**, 257 (1969)..

[100] S. B. Giddings and R. A. Porto, Phys. Rev. D **81**, 025002 (2010). [arXiv:0908.0004 [hep-th]].

[101] S. Ferrara, A. F. Grillo and R. Gatto, Annals Phys. **76**, 161 (1973).

[102] A. M. Polyakov, Zh. Eksp. Teor. Fiz. **66**, 23 (1974).

[103] A. A. Belavin, A. M. Polyakov and A. B. Zamolodchikov, Nucl. Phys. B **241**, 333 (1984).

[104] R. Rattazzi, V. S. Rychkov, E. Tonni and A. Vichi, JHEP **0812**, 031 (2008). [arXiv:0807.0004 [hep-th]].

[105] S. El-Showk, M. F. Paulos, D. Poland, S. Rychkov, D. Simmons-Duffin and A. Vichi, [arXiv:1203.6064 [hep-th]].

[106] D. Pappadopulo, S. Rychkov, J. Espin and R. Rattazzi, [arXiv:1208.6449 [hep-th]].

[107] L. Cornalba, M. S. Costa, J. Penedones and R. Schiappa, JHEP **0708**, 019 (2007). [hep-th/0611122].

[108] L. Cornalba, M. S. Costa, J. Penedones and R. Schiappa, Nucl. Phys. B **767**, 327 (2007). [hep-th/0611123].

[109] L. Cornalba, M. S. Costa and J. Penedones, JHEP **0709**, 037 (2007). [arXiv:0707.0120 [hep-th]].

[110] L. F. Alday and J. M. Maldacena, JHEP **0711**, 019 (2007). [arXiv:0708.0672 [hep-th]].

[111] A. Pelissetto and E. Vicari, Phys. Rept. **368**, 549 (2002). [cond-mat/0012164].

[112] M. Froissart, Phys. Rev. **123**, 1053 (1961)..

[113] A. Martin, Phys. Rev. **129**, 1432 (1963)..

[114] L. Frankfurt, M. Strikman and C. Weiss, Ann. Rev. Nucl. Part. Sci. **55**, 403 (2005). [hep-ph/0507286].

[115] J. Polchinski and M. J. Strassler, JHEP **0305**, 012 (2003). [hep-th/0209211].

[116] C. G. Callan, Jr. and D. J. Gross, Phys. Rev. D **8**, 4383 (1973)..

[117] S. E. Derkachov and A. N. Manashov, J. Phys. A **29**, 8011 (1996). [hep-th/9604173].

[118] S. K. Kehrein, Nucl. Phys. B **453**, 777 (1995). [hep-th/9507044].

[119] H. Osborn, Phys. Lett. B **718**, 169 (2012). [arXiv:1205.1941 [hep-th]].

[120] L. F. Alday, B. Eden, G. P. Korchemsky, J. Maldacena and E. Sokatchev, JHEP **1109**, 123 (2011). [arXiv:1007.3243 [hep-th]].

[121] P. Kovtun and A. Ritz, Phys. Rev. Lett. **100**, 171606 (2008). [arXiv:0801.2785 [hep-th]].

[122] L. Hoffmann, L. Mesref and W. Ruhl, Nucl. Phys. B **608**, 177 (2001). [hep-th/0012153].

[123] F. A. Dolan and H. Osborn, Nucl. Phys. B **629**, 3 (2002). [hep-th/0112251].

[124] P. J. Heslop, JHEP **0407**, 056 (2004). [hep-th/0405245].

[125] P. Liendo, L. Rastelli and B. C. van Rees, [arXiv:1210.4258 [hep-th]].

[126] E. Barnes, E. Gorbatov, K. A. Intriligator, M. Sudano and J. Wright, Nucl. Phys. B **730**, 210 (2005). [hep-th/0507137].

[127] K. G. Wilson and J. B. Kogut, Phys. Rept. **12**, 75 (1974)..

[128] K. Lang and W. Ruhl, [hep-th/9311046].

[129] K. Lang and W. Ruhl, Nucl. Phys. B **400**, 597 (1993)..

[130] V. E. Didenko and E. D. Skvortsov, [arXiv:1210.7963 [hep-th]].

[131] J. Maldacena and A. Zhiboedov, [arXiv:1204.3882 [hep-th]].

[132] S. Ferrara, C. Fronsdal and A. Zaffaroni, Nucl. Phys. B **532**, 153 (1998). [hep-th/9802203].

[133] F. A. Dolan and H. Osborn, Nucl. Phys. B **599**, 459 (2001). [hep-th/0011040].

[134] M. S. Costa, J. Penedones, D. Poland and S. Rychkov, JHEP **1111**, 154 (2011). [arXiv:1109.6321 [hep-th]].

[135] A. L. Fitzpatrick and J. Kaplan, JHEP **1210**, 032 (2012). [arXiv:1112.4845 [hep-th]].

[136] D. Poland and D. Simmons-Duffin, JHEP **1105**, 017 (2011). [arXiv:1009.2087 [hep-th]].

[137] I. I. Shapiro, Phys. Rev. Lett. **13**, 789 (1964).

[138] C. M. Will, [arXiv:1403.7377 [gr-qc]].

[139] G. 't Hooft, Phys. Lett. B **198**, 61 (1987).

[140] Y. Aharonov, A. Komar and L. Susskind, Phys. Rev. **182**, 1400 (1969)..

[141] H. Osborn and A. C. Petkou, Annals Phys. **231**, 311 (1994). [hep-th/9307010].

[142] J. Erdmenger and H. Osborn, Nucl. Phys. B **483**, 431 (1997). [hep-th/9605009].

[143] G. Arutyunov and S. Frolov, Phys. Rev. D **60**, 026004 (1999). [hep-th/9901121].

[144] J. M. Maldacena, JHEP **0305**, 013 (2003). [astro-ph/0210603].

[145] A. Strominger, JHEP **0110**, 034 (2001). [hep-th/0106113].

[146] E. Witten, [hep-th/0106109].

[147] P. McFadden and K. Skenderis, JCAP **1106**, 030 (2011). [arXiv:1104.3894 [hep-th]].

[148] P. Benincasa and F. Cachazo, [arXiv:0705.4305 [hep-th]].

[149] P. C. Schuster and N. Toro, JHEP **0906**, 079 (2009). [arXiv:0811.3207 [hep-th]].

[150] D. A. McGady and L. Rodina, [arXiv:1311.2938 [hep-th]].

[151] R. Penrose, "On Schwarzschild Causality- A problem for "Lorentz Covariant" General Relativity", in "Essays in General Relativity", Edited by F. Tipler, 1980, Elsevier.

[152] S. Gao and R. M. Wald, "Theorems on gravitational time delay and related issues," Class. Quant. Grav. **17**, 4999 (2000). [gr-qc/0007021].

[153] R. W. Fuller and J. A. Wheeler, Phys. Rev. **128**, 919 (1962)..

[154] W. Israel, Phys. Lett. A **57**, 107 (1976)..

[155] J. M. Maldacena, JHEP **0304**, 021 (2003). [hep-th/0106112].

[156] J. Maldacena and L. Susskind, Fortsch. Phys. **61**, 781 (2013). [arXiv:1306.0533 [hep-th]].

[157] P. C. Aichelburg and R. U. Sexl, Gen. Rel. Grav. **2**, 303 (1971).

[158] T. Dray and G. 't Hooft, Nucl. Phys. B **253**, 173 (1985)..

[159] G. T. Horowitz and A. R. Steif, Phys. Rev. Lett. **64**, 260 (1990).

[160] D. N. Kabat and M. Ortiz, Nucl. Phys. B **388**, 570 (1992). [hep-th/9203082].

[161] S. B. Giddings, [arXiv:1105.2036 [hep-th]].

[162] M. Ciafaloni and D. Colferai, [arXiv:1406.6540 [hep-th]].

[163] S. Stieberger, [arXiv:0907.2211 [hep-th]].

[164] H. Kawai, D. C. Lewellen and S. H. H. Tye, Nucl. Phys. B **269**, 1 (1986)..

[165] B. Zwiebach, Phys. Lett. B **156**, 315 (1985)..

[166] S. Sarkar and A. C. Wall, Phys. Rev. D **83**, 124048 (2011). [arXiv:1011.4988 [gr-qc]].

[167] R. Britto, F. Cachazo, B. Feng and E. Witten, Phys. Rev. Lett. **94**, 181602 (2005). [hep-th/0501052].

[168] M. T. Grisaru, H. N. Pendleton and P. van Nieuwenhuizen, Phys. Rev. D **15**, 996 (1977)..

[169] D. M. Eardley and S. B. Giddings, Phys. Rev. D **66**, 044011 (2002). [gr-qc/0201034].

[170] V. S. Rychkov, Phys. Rev. D **70**, 044003 (2004). [hep-ph/0401116].

[171] S. B. Giddings and V. S. Rychkov, Phys. Rev. D **70**, 104026 (2004). [hep-th/0409131].

[172] S. Weinberg and E. Witten, Phys. Lett. B **96**, 59 (1980)..

[173] D. Amati, M. Ciafaloni and G. Veneziano, Phys. Lett. B **197**, 81 (1987).

[174] D. Amati, M. Ciafaloni and G. Veneziano, Int. J. Mod. Phys. A **3**, 1615 (1988).

[175] D. Amati, M. Ciafaloni and G. Veneziano, Phys. Lett. B **216**, 41 (1989).

[176] D. Amati, M. Ciafaloni and G. Veneziano, Phys. Lett. B **289**, 87 (1992).

[177] D. Amati, M. Ciafaloni and G. Veneziano, Nucl. Phys. B **403**, 707 (1993).

[178] L. Susskind, Phys. Rev. D **1**, 1182 (1970)..

[179] D. N. Page, S. Surya and E. Woolgar, Phys. Rev. Lett. **89**, 121301 (2002). [hep-th/0204198].

[180] M. Brigante, H. Liu, R. C. Myers, S. Shenker and S. Yaida, Phys. Rev. Lett. **100**, 191601 (2008). [arXiv:0802.3318 [hep-th]].

[181] M. Hotta and M. Tanaka, Class. Quant. Grav. **10**, 307 (1993)..

[182] G. T. Horowitz and N. Itzhaki, JHEP **9902**, 010 (1999). [hep-th/9901012].

[183] D. M. Hofman, Nucl. Phys. B **823**, 174 (2009). [arXiv:0907.1625 [hep-th]].

[184] R. C. Myers, M. F. Paulos and A. Sinha, JHEP **1008**, 035 (2010). [arXiv:1004.2055 [hep-th]].

[185] K. Sen and A. Sinha, [arXiv:1405.7862 [hep-th]].

[186] A. L. Fitzpatrick, J. Kaplan, D. Poland and D. Simmons-Duffin, JHEP **1312**, 004 (2013). [arXiv:1212.3616 [hep-th]].

[187] Z. Komargodski and A. Zhiboedov, JHEP **1311**, 140 (2013). [arXiv:1212.4103 [hep-th]].

[188] G. J. Galloway, K. Schleich, D. Witt and E. Woolgar, Phys. Lett. B **505**, 255 (2001). [hep-th/9912119]. .

[189] S. H. Shenker and D. Stanford, JHEP **1403**, 067 (2014). [arXiv:1306.0622 [hep-th]].

[190] P. A. R. Ade *et al.* [Planck Collaboration], [arXiv:1303.5082 [astro-ph.CO]].

[191] P. A. R. Ade *et al.* [BICEP2 Collaboration], [arXiv:1403.3985 [astro-ph.CO]].

[192] M. J. Mortonson and U. Seljak, [arXiv:1405.5857 [astro-ph.CO]].

[193] R. Flauger, J. C. Hill and D. N. Spergel, [arXiv:1405.7351 [astro-ph.CO]].

[194] X. O. Camanho and J. D. Edelstein, JHEP **1004**, 007 (2010). [arXiv:0911.3160 [hep-th]].

[195] I. T. Drummond and S. J. Hathrell, Phys. Rev. D **22**, 343 (1980)..

[196] F. A. Berends and R. Gastmans, Annals Phys. **98**, 225 (1976)..

[197] S. Dubovsky, A. Nicolis, E. Trincherini and G. Villadoro, Phys. Rev. D **77**, 084016 (2008). [arXiv:0709.1483 [hep-th]].

[198] D. J. Gross and E. Witten, Nucl. Phys. B **277**, 1 (1986).

[199] R. R. Metsaev and A. A. Tseytlin, Nucl. Phys. B **293**, 385 (1987)..

[200] R. R. Metsaev and A. A. Tseytlin, Phys. Lett. B **185**, 52 (1987).

[201] H. J. de Vega and N. G. Sanchez, Nucl. Phys. B **317**, 706 (1989).

[202] R.C. Tolman, “The theory of relativity of motion”, Berkely Univ. Press, 1917. See also S. Coleman, “Acausality”, Erice School 1969, New York (1970) page 282.

[203] V. E. Didenko and E. D. Skvortsov, [arXiv:1401.2975 [hep-th]].