Holography, Locality and Symmetries of The Universe

Xiao Xiao

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2015

Abstract

Holography, Locality and Symmetries of The Universe

Xiao Xiao

It is an interesting question that, with a well tested duality between the quantum gravity in anti de Sitter space and a quantum field theory in one lower dimension, whether quantum gravity in a cosmological background has a well defined dual description. In large N limit, this duality could be a correspondence between an approximately local gravity theory describing cosmology and a quantum field theory. In dS/CFT, the quantum field theory is a Euclidean CFT living at the conformal boundary of de Sitter space, in large N limit, we should expect the local observables in de Sitter cosmology be recovered from the CFT. We explicitly develop this construction for scalar fields and derive the operator map at lowest order of $\frac{1}{N}$ expansion.

Having addressed the fundamental question of how local fields in de Sitter cosmology arise via holography, we focus on the theory of cosmological perturbations that is described in terms of local field theory. The curvature perturbations during inflation, which originated from quantum fluctuations of inflaton and which induced the CMB inhomogeneity we see today, as well as the large scale structure, can be described as Goldstone boson fields which nonlinearly realize a subset of general coordinate transformations as residual symmetries. This fact puts strong constraints on the behavior of the cosmological correlation functions, and a series of consistency relations constraining the soft limits of these correlation functions can be derived as Ward identities.

Contents

List of Figures

Acknowledgements

First I thank my family, my Dad, my Mom and my wife Huijun, for their care and understanding through these years. It is never easy for my Dad and Mom having me thousands of miles away from them, nor is it easy for my wife spending three years in a different time zone from me. I don't know how my life would've been without my family taking care of me.

I thank my advisor, professor Erick Weinberg, who guided me through my whole Ph.D. life as a great mentor. Working with a physicist at Erick's level has been such a pleasure and also a great opportunity to learn. As an outstanding physicist, Erick shaped the way I view and think about physics, and as a great mentor, Erick gave me support when time was difficult and showed me what the highest standard of personality is like.

I thank Ali Masoumi and Bart Horn, with whom I went through the joy and pain of tackling research problems and with whom I finished research works that I'll be proud of through my life. Without these two gentlemen I could've never accomplished what I've done.

I thank Kurt Hinterbichler for reviewing my first paper with Ali and encouraging us to publish, which gave me a good start for later research.

I thank Andrej Ficnar, Ali Masoumi and Junpu Wang, for countless discussions in both physics and non-physics topics, with you guys the office had been such an enjoyable place to stay and an inspiring place where I can get stimulating ideas.

I thank professor Frederik Denef, Gregory Gabadadze, Brian Greene, Lam Hui, Daniel Kabat and Alberto Nicolis, with whom I worked and from whom I learned a great deal, it is great luck for me that I have the opportunity to talk to some of the best theoretical physicists I've ever seen.

I thank all my friends who are with me here in New York: folks in and had been in the lattice QCD group at Columbia, Ziyuan Bai, Luchang Jin, Zhongjie Lin, Qi Liu, Hantao Yin, Jianglei Yu and Daiqian Zhang, it has been really enjoyble to drop by your office, having random chats, getting lunch together, watching Game of Thrones together and play doodle jump together; my classmates at Columbia, especially Xinyuan Ai, Alex Chen, Lian Liu, Zhe Liu, Shuo Zhang and Lei Zhou, I feel really lucky to know all of you, to sit in the same classrooms with you and to have awesome parties with you; my friends from high school and college, especially Shanchao Fu, Bo Zhao, Tianqiang Liu, Wen Ma and Qian Peng, it is so heartwarming to have old friends around.

My thank goes to everyone who cares about me and thinks of me as a friend or a good kid, I will keep on being awesome, for all of you.

Introduction

To understand the cosmic microwave background (CMB) as well as the rich large scale structure (LSS) in our universe, it is necessary to go beyond the background geometry and look into the fluctuations of various fields, which inevitably introduces the methodology of quantum field theory (QFT) into the study of cosmology. When the quantum field theory on an expanding background is formulated, the tasks of computing the fluctuations of CMB photons as well as the density of dark matter become standard applications of Feynman integral techniques, and meaningful information about CMB as well as LSS can be extracted from the correlation functions of field theories.

The universe we observe has a flat, expanding geometry. Thus we can slice the universe with flat slices and regard the cosmological perturbations as degrees of freedom living on these slices and study their correlation functions on these slices. A certain formulation of quantum field theory which is dubbed the "in-in formalism" is applied. We will come to this kind of formulation later in this thesis. As an example, the CMB perturbation is observed on the last scattering surface which lies at 14 billion years ago in the history of our universe, and correlation functions are computed with assumptions regarding the initial state of the universe as well as the dynamics of its evolution. As another example, through the large scale structure survey we look at the cosmological correlation functions on the slice that we are living now. Information regarding the distribution of galaxies and dark matter is inferred from the observation and statistics of galaxies, and theoretical predictions are given by coupling fluid dynamics with gravity or other modified gravity models. Theoretically this procedure can be carried out for any cosmological background. For the background we are living in now — a flat FRW universe dominated by a small cosmological constant — observables such as the power spectrum and bi-spectrum are computed from theories and compared to the results from various observations.

Among all kinds of cosmologies, de Sitter space is especially interesting: It has the maximum number of isometries a spacetime can have and thus has various interesting coordinate patches. In flat slicing it appears to be a flat universe which is exponentially expanding, which makes it a good approximation to the inflation phase of our universe as well as its current dark energy-dominated stage. In the static patch it has a observer-dependent horizon, and thus has temperature and finite entropy, which bring interesting puzzles regarding the nature of quantum gravity in de Sitter space. One especially interesting property of de Sitter space is that it has future and past conformal boundaries. This inspires the idea of formulating a holographic duality between quantum gravity in de Sitter space and a quantum field theory, the so called dS/CFT duality. The quantum field theory dual to gravity in de Sitter space lives on a special slice — the future/past conformal boundary. The cosmological correlation functions of scalar and tensor perturbations, when extrapolated to the boundary, become the correlation functions of the scalar part and the tensor part of the stress tensor of the quantum field theory. The isometries of de Sitter space naturally introduce the whole set of conformal symmetries to the degrees of freedom on the boundary and the dual quantum field theory is a conformal field theory (CFT). Ths correspondence between quantum gravity in de Sitter space and a Euclidean CFT was originally proposed by Strominger [1] and the realization of this idea on high spin fields was discovered in [2]. Since de Sitter space in flat patch is linked to anti-de Sitter space via a simple analytic continuation, the holography in this particular patch strongly resembles the AdS/CFT correspondence and the operator dictionary is formulated. There is a subtlety regarding the operator dictionary when the microcausality of the correlation functions is considered. We will come to this issue later in this thesis and address it. The holographic description of de Sitter space in static patch is relatively less understood.

Surprisingly, conformal symmetries are shown to be important not only for quantum field theories in de Sitter space, but also for cosmological perturbations in a generic FRW universe. From the point of view of a field theory, the conformal symmetries are non-linearly realized by the Goldstone bosons — the scalar field for curvature perturbations and the graviton field for tensor perturbations. Rather than originating from the isometries of the background spacetime, the conformal symmetries come from a subset of general coordinate transformations which remain after gauge fixing. These conformal symmetries put strong constraints on the soft limits of cosmological correlation functions, in the form of consistency relations which can be understood as the Ward identities associated with these symmetries. The existence of non-linearly realized conformal symmetries does not rely on any assumption regarding the specific dynamics except for general covariance. Thus, consistency relations can be found for both CMB correlation functions and LSS correlation functions. The existence of consistency relations for the LSS correlation functions in especially interesting since these observables are highly non-linear at small scales and thus difficult to compute. However, consistency relations for LSS make non-trivial constraints on these observables and thus provide a probe into the highly non-linear regime. The first such consistency relations was discovered in Maldacena's work [3], from a simple argument which was later formalized into the so-called "background wave argument", which is still a very efficient and intuitive procedure for deriving consistency relations. Maldacena's consistency relation is derived from a rescaling of the spatial coordinates, which corresponds to the scale transformation in the conformal group. In work [4] and [5] the consistency relations corresponding to the full conformal group were discovered along with an infinite set of consistency relations corresponding to infinitely many "residual transformations". This infinite set of consistency relations are organized into a power series in the soft momentum. It was later shown that this can be resummed into a master relation originating from gauge invariance [6],[7]. The first consistency relation for large scale structure was discussed in [8] and [9]. The Galilean transformation, which was later clarified to be the time-dependent translation of coordinates, was shown to lead to interesting consistency relations constraining the correlation functions of the dark matter over-density. It was soon discovered by [10] and [11] that there is an infinite set of such relations in LSS.

The structure of this thesis is as follows: in Chapter One we briefly discuss the basics of conformal field theory, the AdS/CFT correspondence, and quantum field theory in a cosmological background. In Chapter Two, we discuss the operator dictionary of the dS/CFT correspondence, and address a subtlety regarding the construction of local operators in de Sitter space from the CFT. This chapter is based on papers [13], [14]. In Chapter Three, we introduce the conformal symmetries in the theoretical description of large scale structure and derive an infinite set of consistency relations constraining the soft limits of correlation functions. Further we look into Lagrangian space and derive the Newtonian consistency relation therein. The chapter is based on work [11], [12]. Following Chapter Three is the Conclusion for the thesis and the Appendices.

Chapter 1

Quantum Field Theory in AdS, and in Our Universe

The purpose of this chapter is to outline the necessary background for introducing the thesis projects in the following chapters. The concept of conformal field theory and the geometry of anti de Sitter space are introduced. Then we introduce field theory on and AdS background and the operator dictionary of the AdS/CFT correspondence. We live in a universe with a positive cosmological constant, which is approximately described by de Sitter geometry. It is then an interesting question whether a holographic correspondence can be realized in such a background. We introduce the geometries of de Sitter space and a generic FRW universe. Then we describe the idea of a dS/CFT correspondence, which is one of the generalizations of AdS/CFT to de Sitter space. Then observables in an expanding universe as well as cosmological perturbation theory are introduced.

1.1 Local quantum field theory in AdS and AdS/CFT duality

1.1.1 A bit of conformal field theory

Of all the quantum field theories, conformal field theories form an interesting subset; they play a central role in the physics of phase transitions as the fixed points of the renormalization group flow. The worldsheet field theories of strings are also conformal field theories, to ensure the consistency of the string theories.

To see what conformal field theory is, we need to know about the conformal group. The conformal group in D-dimensional Minkowski spacetime is composed of all Lorentz transformations supplemented by the following scale and special conformal transformations (SCT):

$$\begin{aligned} x^\mu &\to x'^\mu = \lambda x^\mu \\ x^\mu &\to x'^\mu = \frac{x^\mu + x^2 b^\mu}{1 + 2b \cdot x + b^2 x^2} \end{aligned} \tag{1.1}$$

It is also possible to defined a conformal field theory as a theory which is invariant under Lorentz transformations, dilation and inversion:

$$x'^\mu = \frac{x^\mu}{x^2} \tag{1.2}$$

The reason is that the special conformal transformations can be reproduced by a translation sandwiched by two times of inversions:

$$\begin{aligned} x^\mu &\to \frac{x^\mu}{x^2} \\ x^\mu &\to x^\mu - b^\mu \\ x^\mu &\to \frac{x^\mu}{x^2} \end{aligned} \tag{1.3}$$

The conformal group has a set of generators forming the conformal algebra. In terms of differential operators on spacetime fields, the generators are:

$$\begin{aligned} P_\mu &= -i\partial_\mu \\ D &= -ix^\mu\partial_\mu \\ L_\mu\nu &= i\left(x_\mu\partial_\nu - x_\nu\partial_\mu\right) \\ K_\mu &= -i\left(2x_\mu x^\nu\partial_\nu - x^2\partial_\mu\right) \end{aligned} \tag{1.4}$$

These generators form the conformal algebra, which is defined by the following commutation relations:

$$\begin{aligned} [D, P_\mu] &= iP_\mu \\ [D, K_\mu] &= -iK_\mu \\ [k_\mu, P_\nu] &= 2i\left(\eta_{\mu\nu}D - L_{\mu\nu}\right) \\ [K_\rho, L_{\mu\nu}] &= i\left(\eta_{\rho\mu}K_\nu - \eta_{\rho\nu}K_\mu\right) \\ [P_\rho, L_{\mu\nu}] &= i\left(\eta_{\rho\mu}P_\nu - \eta_{\rho\nu}P_\mu\right) \\ [L_{\mu\nu}, L_{\rho\sigma}] &= i\left(\eta_{\nu\rho}L_{\mu\sigma} + \eta_{\mu\sigma}L_{\nu\rho} - \eta_{\mu\rho}L_{\nu\sigma} - \eta_{\nu\sigma}L_{\mu\rho}\right) \end{aligned} \tag{1.5}$$

This algebra, when written in terms of a set of linear combinations of the generators, takes a much simpler form and can be connected to a more familiar algebra. Define:

$$\begin{aligned} J_{\mu\nu} &\equiv L_{\mu\nu} \\ J_{D+1,\mu} &\equiv \frac{1}{2}\left(P_\mu - K_\mu\right) \\ J_{D+1,0} &\equiv D \\ J_{0,\mu} &\equiv \frac{1}{2}\left(P_\mu + K_\mu\right) \end{aligned} \tag{1.6}$$

Then the conformal algebra (1.5) takes the form:

$$\left[J_{MN}, J_{PQ}\right] = i\left(\eta_{MQ}J_{NP} + \eta_{NP}J_{MQ} - \eta_{MP}J_{NQ} - \eta_{NQ}J_{MP}\right) \tag{1.7}$$

which is the algebra for $SO\left(D+1,1\right)$. This means that the conformal group in D dimensions is isomorphic to $SO\left(D+1,1\right)$

A conformal field theory (CFT) is a field theory invariant under the conformal group. Roughly speaking, it is a system in which the physics is independent of the spacetime scale or energy scale, i.e. invariant when all the coordinates are rescaled by the same amount[1]. There is a subset of dynamical variables in a CFT characterized by their scaling dimensions in addition to their Lorentz transformation properties, called quasi-primary operators. These operators are of special interests since their scaling dimensions reflect important knowledge about the dynamics of the CFT. For instance, under a scale transformation, a quasi-primary operator $\mathcal{O}$ in the CFT with scaling dimension Δ transforms like:

$$\mathcal{O}\left(x\right) \to \lambda^{\Delta}\mathcal{O}\left(x'\right) = \lambda^{\Delta}\mathcal{O}\left(\lambda x\right) \tag{1.8}$$

[1] The intuitive statement here is not completely rigorous. Whether scale invariance necessarily leads to conformal invariance is itself a deep and interesting question. People are trying to establish the statement rigorously in various dimensions. Given certain reasonable constraints on viable field theories (unitarity, Lorentz invariance, etc.), no widely-accepted counterexample is known.

In general, under a conformal transformation, it transforms like:

$$\mathcal{O}\left(x\right) \to \det\left(\frac{\partial x'}{\partial x}\right)^{\frac{\Delta}{d}} \mathcal{O}\left(x'\right) \tag{1.9}$$

We will constrain our discussion to quasi-primary operators when we study operators in a CFT.

The form of the two-point and three-point correlation functions of quasi-primary operators in a CFT is completely fixed by conformal invariance. The two-point function between operators with scaling dimension Δ is:

$$\left\langle \mathcal{O}\left(x_1\right) \mathcal{O}\left(x_2\right) \right\rangle = \frac{c_{12}}{x_{12}^{2\Delta}} \tag{1.10}$$

where

$$x_{12}^2 \equiv \left(x_1 - x_2\right)^2 \tag{1.11}$$

Notice that it is only when the two operators in the two-point function have the same scaling dimension that the two-point function is not vanishing; otherwise, the operators are not correlated at this level.

The three-point function between operators with dimensions Δ_1, Δ_2 and Δ_3 is:

$$\left\langle \mathcal{O}_1\left(x_1\right) \mathcal{O}_2\left(x_2\right) \mathcal{O}_3\left(x_3\right) \right\rangle = \frac{C_{123}}{x_{12}^{\Delta_1+\Delta_2-\Delta_3} x_{23}^{\Delta_2+\Delta_3-\Delta_1} x_{31}^{\Delta_3+\Delta_1-\Delta_2}} \tag{1.12}$$

It is always possible to normalize any quasi-primary operator $\mathcal{O}$ so that the coefficient of the two-point function is one (or any certain value). After this is done, one is no longer free to choose the coefficients of the three-point functions. Thus, these coefficients are intrinsic characteristics of the CFT, in addition to the scaling dimensions. The coefficient C_{abc} is called an "OPE coefficient", since it is the coefficient for the dimension Δ_c operator in the operator product expansion between operators with dimensions Δ_a and Δ_b:

$$\mathcal{O}_{\Delta_a}(x \to 0)\,\mathcal{O}_{\Delta_b}(0) \sim \sum_c C_{abc}\mathcal{O}_{\Delta_c}(0) \tag{1.13}$$

This collection of numbers, the scaling dimensions of all the operators and the OPE coefficients, form the whole set of data defining a CFT: a conformal field theory is uniquely specified when the scaling dimensions of its operators and the OPE coefficients are determined. No more information is needed.

Coming to four-point functions, we see when there are four spacetime points considered, the correlation functions are no longer uniquely fixed by conformal symmetries. The reason is the existence of invariant cross ratios. One can verify that the cross ratios

$$\frac{x_{12}x_{34}}{x_{13}x_{24}} \equiv \frac{|x_1-x_2||x_3-x_4|}{|x_1-x_3||x_2-x_4|}\,,\quad \frac{x_{12}x_{34}}{x_{23}x_{14}} \equiv \frac{|x_1-x_2||x_3-x_4|}{|x_2-x_3||x_1-x_4|} \tag{1.14}$$

are invariant under conformal transformations, which means that any functions of these two ratios should be invariant and can be freely multiplied on a function with the right dimensions while keeping the transformation properties under the conformal group. The generic form of the four-point functions in a CFT is

$$G_4(x_1,x_2,x_3,x_4) = F\left(\frac{x_{12}x_{34}}{x_{13}x_{24}},\frac{x_{12}x_{34}}{x_{23}x_{14}}\right)\prod_{i<j}^{4} x_{ij}^{\frac{\Delta}{3}-\Delta_i-\Delta_j} \tag{1.15}$$

where $\Delta \equiv \sum_{i=1}^{4}\Delta_i$

1.1.2 Anti-de Sitter space

Anti-de Sitter space is a solution of Einstein's equations with negative cosmological constant. The solution has the maximal number of isometries and has been studied extensively. There are several choices of coordinate patches to describe the metric of Anti de Sitter space. For instance, in the Poincaré patch the spacetime metric is

$$ds^2_{AdS} = \frac{dz^2 + dx^2_D}{z^2} \tag{1.16}$$

where dx^2_D is the spacetime line element of D-dimensional Minkowski space, and z runs from 0 to ∞. The boundary of AdS space, in the Poincaré patch, is located at $z = 0$ and $z = \infty$ is the Poincaré horizon.

In the Poincaré patch it is easy to write down the isometries of anti-de Sitter space AdS_{D+1}. The most obvious ones are the translations and (pseudo)rotations in D-dimensions:

$$x^\mu \to x^\mu + a^\mu \ , \ x^\mu \to \Lambda^\mu{}_\nu x^\nu \tag{1.17}$$

which correspond to all the Lorentz transformations on the boundary $z = 0$.

Another obvious isometry to see is the rescaling of all the coordinates:

$$z \to \lambda z \ , \ x^\mu \to \lambda x^\mu \tag{1.18}$$

which induces a rescaling of coordinates on the boundary.

Less transparent are the isometries that induce the special conformal transformations on the boundary:

$$\begin{aligned} x^\mu &\to \frac{x^\mu - b^\mu \left(x^2 + z^2\right)}{1 - 2b \cdot x + b^2 \left(x^2 + z^2\right)} \\ z &\to \frac{z}{1 - 2b \cdot x + b^2 \left(x^2 + z^2\right)} \end{aligned} \tag{1.19}$$

Here we see that the AdS isometries are in one-one correspondence with the conformal transformations that acts on the boundary degrees of freedom.

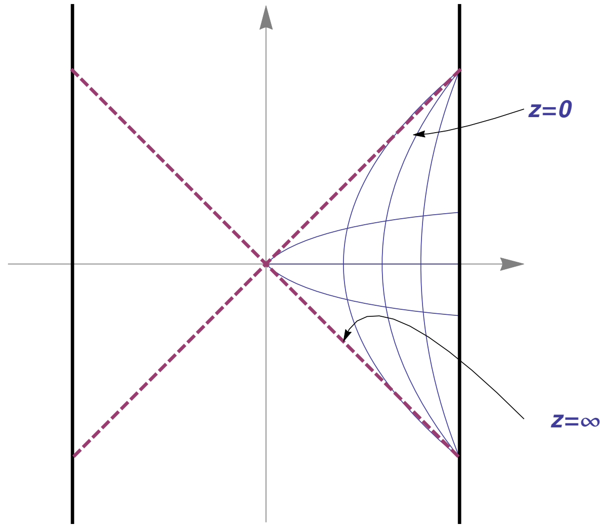


Figure 1.1: *AdS_2 and Poincaré patch*

In global coordinates the metric is instead:

$$ds^2_{AdS} = \frac{1}{\cos^2\rho}\left(-d\tau^2 + d\rho^2 + \sin^2\rho d\Omega^2_{D-1}\right) \tag{1.20}$$

with ρ ranging from 0 to $\frac{\pi}{2}$

There is another parametrization for the global patch:

$$ds^2 = -\left(r^2+1\right)dt^2 + \frac{1}{r^2+1}dr^2 + r^2 d\Omega^2 \tag{1.21}$$

We see that at large r the volume of the space grows along with the surface area:

$$Volume \sim 4\pi \times \int^R \frac{r^2 dr}{\sqrt{1+r^2}} \propto R^2, R \to \infty \tag{1.22}$$

If we work in the coordinate system (1.21) and put a massive test particle near $r = 0$, we see that the geodesic motion of this particle look like a harmonic oscillator. This is

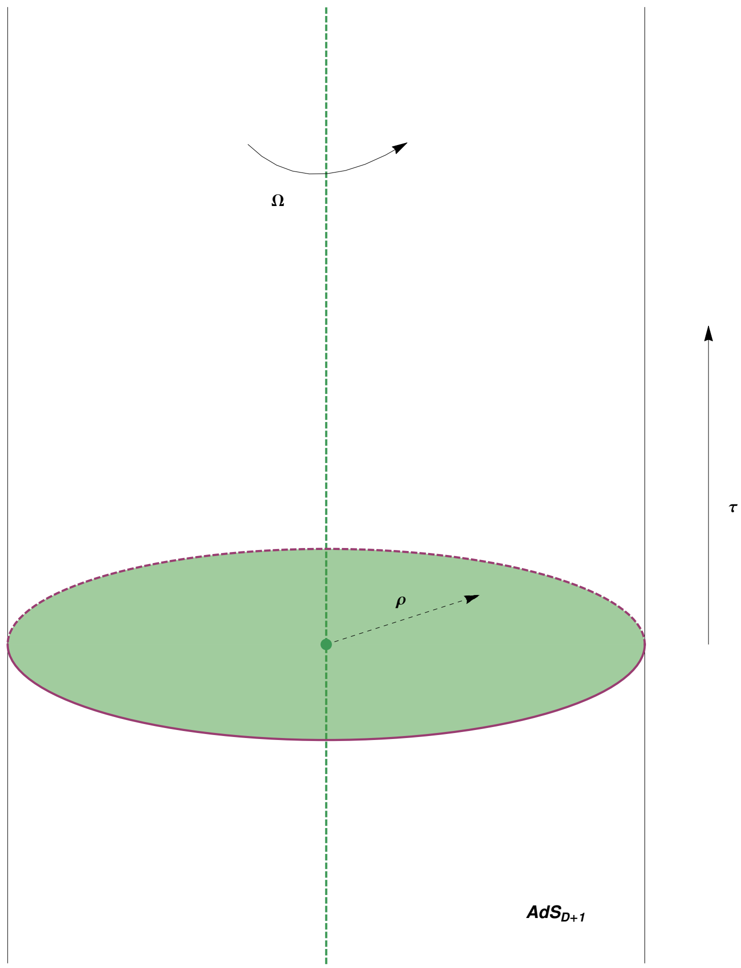


Figure 1.2: *Global AdS*

because we have a gravitational potential which grows with r:

$$V_{gravity} \propto \sqrt{-g_{00}} \sim \sqrt{1+r^2} \tag{1.23}$$

Thus giving a kick to the particle will result in a motion around the center of the coordinate $r = 0$ and the particle will never reach the boundary of the space because of the infinite gravitational potential at $r = 0$.

1.1.3 Field theory on AdS and AdS/CFT

The conjecture of AdS/CFT correspondence is that supergravity or string theory in anti-de Sitter space is dual to a conformal field theory living on the boundary of the space. The original paper of Maldacena [15] gave a brane construction which consists of N D3 branes. The Type-IIB string theory in the near horizon geometry is shown to have a duality with $\mathcal{N} = 4$ super Yang-Mills theory, which is the world volume theory of the D3 branes. The near horizon geometry of the D3 branes is $AdS_5 \times S^5$, which establishes a correspondence between a certain type of string theory and a certain superconformal field theory.

We will not delve into the brane construction of the correspondence, rather we focus on the operator dictionary relating the AdS (bulk) degrees of freedom and the boundary CFT degrees of freedom. In the low energy theory of the bulk string theory there are various degrees of freedom, including the moduli, gauge fields and graviton. For simplicity, let's look at a scalar field theory defined on the Poincaré patch of AdS_{D+1}, with a mass term the only potential term.

$$S = \int d^{D+1}x\sqrt{-g}\left\{-\frac{1}{2}\left(\nabla\phi\right)^2 - \frac{1}{2}m^2\phi^2\right\} \tag{1.24}$$

The equation of motion for the scalar field on the AdS background is

$$\partial_\alpha \partial^\alpha \phi + z^{D-1}\partial_z \left(z^{1-D}\partial_z \phi\right) - \frac{m^2}{z^2}\phi = 0 \tag{1.25}$$

In order to establish the operator dictionary, we focus on the behavior of the solutions to this equation near the boundary of AdS.

Suppose we have a simple solution that is trivial on x^μ directions and has the form

$$\phi \sim z^\delta \beta \tag{1.26}$$

near the boundary $z = 0$, where β is the boundary condition we fix.

The equation of motion implies that the possible values δ can take are:

$$\delta_\pm = \frac{D}{2} \pm \sqrt{\left(\frac{D}{2}\right)^2 + m^2} \tag{1.27}$$

The two values correspond to two different boundary conditions that we can impose when we solve the classical equation as well as when we define the functional integral of the theory in AdS.

We can see that if we take the positive value of δ then β can be interpreted as an operator in a D-dimensional CFT. Suppose we rescale all the coordinates:

$$z \to \lambda z, x^\mu \to \lambda x^\mu \tag{1.28}$$

As a scalar field ϕ should be invariant:

$$\phi\left(z, x\right) = z^\delta \beta\left(x\right) \to \phi\left(\lambda z, \lambda x\right) = \lambda^\delta z^\delta \beta\left(\lambda x\right) \tag{1.29}$$

Thus we see that under the scale transformation induced by the AdS isometry, β transforms as

$$\beta(x) \to \beta(\lambda x) = \lambda^{-\delta}\beta(x) \tag{1.30}$$

which is exactly the behavior for a quasi-primary field with scaling dimension δ.

The alternative boundary condition with $\delta_- = D - \delta_+ = \frac{D}{2} - \sqrt{\left(\frac{D}{2}\right)^2 + m^2} < 0$ seems to define a β field with negative scaling dimension. However in this case we treat β as a source field coupled to the CFT operator and it is not itself in the specturm of the theory.

AdS/CFT duality claims that the partition function of quantum gravity in AdS_{D+1}, with the boundary condition $\phi(z \to 0, x) \to z^{\delta_-}\beta(x)$, is equal to the partition function of a conformal field theory deformed by a coupling between β and a dimension-δ_+ operator:

$$Z_{AdS}(\beta) = \left\langle e^{\beta\mathcal{O}} \right\rangle_{CFT} \tag{1.31}$$

The CFT correlation functions can be obtained by computing the AdS partition function of gravity in terms of boundary condition β and then taking the derivatives:

$$\langle \mathcal{O}(x_1)\,\mathcal{O}(x_2)\ldots\mathcal{O}(x_n)\rangle = \frac{\delta^n}{\delta\beta(x_1)\ldots\delta\beta(x_n)} Z_{AdS}(\beta) \tag{1.32}$$

Since corresponding to a strongly-coupled conformal field theory on the boundary is a weakly coupled supergravity which can be solved in semi-classical approximation, the partition function $Z_{AdS}(\beta)$ is thus often computable for strongly-coupled CFT. This gives us a powerful way of handling such systems: in order to study a strongly coupled CFT — in which lots of quantities seem impossible to compute — we go to AdS space in one higher dimension and solve a weakly coupled supergravity, which is doable.

1.2 Local quantum field theory in de Sitter space and generic FRW universe

1.2.1 Geometry of de Sitter space and FRW universe

Observations show that our universe has a flat, expanding geometry with a positive cosmological constant. Such a geometry in general can be described by an FRW anzatz, which is a 3-dimensional flat spatial slice with a certain scale factor

$$ds^2_{FRW} = -dt^2 + a^2(t)\, d\boldsymbol{x}^2 \tag{1.33}$$

The scale factor $a(t)$ satisfies the Friedmann equations

$$\begin{aligned} H^2 &= \frac{8\pi G}{3}\rho \\ \dot{H} + H^2 &= -\frac{4\pi G}{3}(\rho + 3p) \end{aligned} \tag{1.34}$$

where the Hubble parameter $H(t) \equiv \frac{\dot{a}}{a}$. How the scale factor evolves depends on the matter and energy content of the universe. For instance, in a matter dominated universe, $p = 0$ and $a(t) \propto t^{\frac{2}{3}}$; in a radiation dominated universe $\rho = 3p$ and $a(t) \propto t^{\frac{1}{2}}$.

De Sitter space corresponds to a certain evolution of the scale factor, with

$$a(t) = e^{Ht} \tag{1.35}$$

This is the solution for (1.34) with $\rho = -p$, which is a universe dominated by a cosmological constant. The Hubble parameter is a constant $H(t) = H$ in de Sitter space.

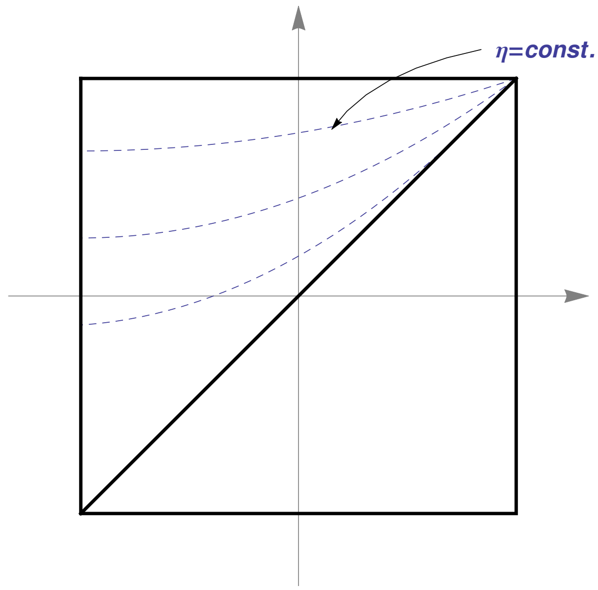


Figure 1.3: *de Sitter space and the flat patch*

Defining conformal time $\eta \equiv e^{-Ht}$ we have

$$ds^2_{dS} = \frac{H^2}{\eta^2}\left(-d\eta^2 + d\boldsymbol{x}^2\right) \tag{1.36}$$

We see that there is a singularity of the metric at $\eta = 0$. The Penrose diagram of de Sitter space is shown in Fig. 1.3. It looks like a square with each point in the diagram corresponding to a two-sphere. The infinite future which corresponds to $t \to \infty$ is pulled back to the top of the square $\eta = 0$ via defining the conformal time. The metric (1.36) covers either the upper left wedge or the lower right wedge of de Sitter space and describes flat spatial slices that evolve in time. This is the flat patch of de Sitter space, and $\eta = 0$ is called the conformal boundary of de Sitter.

Equation (1.36) describes either an expanding universe with η ranging from $-\infty$ to 0, or a shinking universe with η ranging from 0 to ∞. The global coordinates cover both:

$$ds^2 = \frac{1}{H^2\cos^2\tau}\left(-d\tau^2 + d\Omega^2_{d-1}\right) \tag{1.37}$$

where τ is the global time with range $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

These coordinates show that global de Sitter space is a closed universe with an expanding phase following a contracting phase. In these coordinates the universe looks very different from what we see in the flat patch, while both share the same degree of symmetry.

There is a simple analytic continuation relating the flat patch of de Sitter space and the Poincaré patch of anti-de Sitter space. Take a de Sitter geometry with Hubble parameter H and an anti-de Sitter space with AdS radius R_{AdS}. Taking $z \to i\eta$, $x_D \to it$ and $R_{AdS} \to H$, we have

$$ds^2_{AdS_{D+1}} = \frac{R^2_{AdS}}{z^2}\left(dz^2 + dx^2_D\right) \to ds^2_{dS_{D+1}} = \frac{H^2}{\eta^2}\left(-d\eta^2 + d\boldsymbol{x}^2_D\right) \tag{1.38}$$

In this analytic continuation, the spatial boundary $z = 0$ of AdS space become the timelike boundary $\eta = 0$ of de Sitter space. Functional integrals that are defined respectively by certain boundary conditions are also related to each other via the analytic continuation. Thus the dS/CFT correspondence as the analytic continuation of AdS/CFT correspondence seems to be a natural candidate for a holographic description of quantum gravity in de Sitter space. However, since in the analytic continuation, spacelike separated points become timelike separated, issues regarding microcausality become subtle and require further thinking. We will see that such issues appear when we try to construct local bulk fields in de Sitter space from the dS/CFT operator dictionary. The construction formula is not the analytic continuation from the version in anti-de Sitter space.

1.2.2 dS/CFT correspondence

Quantum gravity in de Sitter space is conjectured to be dual to a conformal field theory living at the timelike boundary $\mathcal{I}^+$. The CFT has no time evolution since the space in which it is defined is Euclidean, thus there is not directly a notion of unitarity. In the future wedge of the flat patch of de Sitter, the asymptotic boundary $\mathcal{I}^+$ is a flat Euclidean space. The wave equation as well as the solutions can be obtained directly by analytically continuing from the corresponding parts in AdS. For a scalar field with mass m in de Sitter space, what we have in the CFT are two operators with dimensions

$$\delta_\pm = \frac{D}{2} \pm \sqrt{\left(\frac{D}{2}\right)^2 - m^2} \tag{1.39}$$

We see that for a heavy scalar with

$$m^2 > \left(\frac{D}{2}\right)^2 \tag{1.40}$$

we have imaginary scaling dimensions. In co-moving time coordinates, the imaginary parts are in the exponential:

$$e^{\pm i\sqrt{m^2 - \left(\frac{D}{2}\right)^2} t} \tag{1.41}$$

and they have the right form for the evolution of the positive energy and negative energy components of a local field in de Sitter space. This suggests that unlike the case in anti de Sitter space, for a local field with both positive and negative energy components in de Sitter space the corresponding degrees of freedom in the CFT also have two components with complementary scaling dimensions. We will see this explicitly in Chapter Two.

1.3 Observables and perturbation theory for the early universe

1.3.1 Observables in the universe and Schwinger-Keldysh formalism

The observables in the cosmic microwave background radiation and the large scale structure survey are the statistical correlation functions, which have their origin in the quantum fluctuations during inflation. These statistical correlation functions characterize expectation values of the dynamical variables and their products under certain probability measures, that are determined by quantum processes during inflation and the subsequent evolutions.

Computing such expectation values is different from computing the off-shell time-ordered correlation functions encountered in quantum field theories in Minkowski spacetime, which we use to compute the scattering amplitudes. Those correlation functions are the matrix elements of operators between in and out vacua, instead of the expectation values of operators in certain quantum states, therefore the path-integration procedure for such correlation functions cannot be directly applied to computing the cosmological correlation functions.

Here we briefly introduce the Schwinger-Keldysh formalism for computing expectation values of observables. The formalism is also often called the “in-in formalism” since the path integral has a folded integration path which starts from the in-vacuum and goes back to the in-vacuum at the end.

Here we are interested in values of observables in a general spacetime, especially an expanding universe. These observables can be chosen to sit on the same spatial slice if they are all space-like separated from each other, or we are free to leave them at different times.

Let us start with the simple case where we have already aligned all the observables on the same spatial slice at time t, and underlying these observables we have a fundamental variable ϕ which we integrate over when performing path integration. Then at time t the expectation value of these observables is given by a probability measure which is determined by the wave function:

$$\langle \mathcal{O}\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots \boldsymbol{x}_n; t\right)\rangle = \int \mathcal{D}\phi\left(t\right) \Psi^*\left(\phi, t\right) \mathcal{O}\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots \boldsymbol{x}_n; t\right) \Psi\left(\phi, t\right) \tag{1.42}$$

The Hamiltonian of the field theory can be split into the free part and the interaction part,

$$H = H_0 + H_{int} \tag{1.43}$$

In the interaction picture, the expectation value is taken on the the initial wave function evolving with the free Hamiltonian H_0, while we insert the exponential of the interaction Hamiltonian into the expectation value:

$$\begin{aligned} &\langle \mathcal{O}\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots \boldsymbol{x}_n; t\right)\rangle \\ =& \int \mathcal{D}\phi\left(t\right) \Psi_0^*\left(\phi, t\right) e^{i\tilde{\mathcal{T}}\int_{t_0}^t H_{int}(t')dt'} \mathcal{O}\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots \boldsymbol{x}_n; t\right) e^{-i\mathcal{T}\int_{t_0}^t H_{int}(t')dt'} \Psi_0\left(\phi, t\right) \end{aligned} \tag{1.44}$$

Therefore the expectation value of the observables evaluated at time t is exactly the expectation value of the observables evolved back to the initial time and evaluated under the initial probability distribution:

$$\langle \mathcal{O}\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots \boldsymbol{x}_n; t\right)\rangle = \int \mathcal{D}\phi\left(t\right) e^{i\tilde{\mathcal{T}}\int_{t_0}^t H_{int}(t')dt'} \mathcal{O}\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots \boldsymbol{x}_n; t\right) e^{-i\mathcal{T}\int_{t_0}^t H_{int}(t')dt'} \mathcal{P}_0\left(\phi, t\right) \tag{1.45}$$

In actual calculations, the initial time t_0 is often taken to be $-\infty$ and we usually want the wave function $\Psi_0\left(\phi\right)$ to be the Bunch-Davies vacuum, which is the ground state of the free Hamiltonian in a approximate de Sitter background. We can achieve this by giving a small imaginary part to the Hamiltonian. This selects the vacuum state from a general initial wave function:

$$e^{-i\mathcal{T}\int_{-\infty}^{t} H_{int}(t')dt'} \to e^{-i\mathcal{T}\int_{-\infty(1+i\varepsilon)}^{t} H_{int}(t')dt'} \tag{1.46}$$

The expectation value thus specifies a closed time contour which goes from $-\infty\,(1+i\varepsilon)$ to the time t of the operator insertions, and then goes back to $-\infty\,(1-i\varepsilon)$. Notice that in evaluating Green's functions for scattering amplitudes, it is not the initial wave function but the asymptotic boundary conditions in time that are fixed, thus the time integration contour which selects the vacuum state goes from $-\infty\,(1+i\varepsilon)$ straight to $\infty\,(1+i\varepsilon)$.

1.3.2 Cosmological perturbation theories for primordial fluctuations and large scale structure

Here we briefly introduce the theories for fluctuations during inflation as well as the subsequent evolution which is responsible for the large scale structure we see today. We discuss briefly the physical quantities and the observables we compute, based on our introduction of the FRW universe and the in-in formalism in the previous sections.

1.3.2.1 Primordial fluctuations

Primordial fluctuations are the quantum fluctuations of gravity coupled to the inflaton during inflation. The simplest model describing this system is general relativity minimally coupled to a single scalar field which interacts with itself via a potential term:

$$S = \int d^4x\sqrt{-g}\left(M_p^2 R - \frac{1}{2}\left(\nabla\phi\right)^2 - V\left(\phi\right)\right) \tag{1.47}$$

Theories with non-minimal couplings and derivative couplings are extensive in the literature, but this type of models is a good starting point. Also, some of the models that have the best fits with the observed data lie within this type.

There are three dynamical degrees of freedom in the system: two from the metric fluctuations and one from the inflaton. The slow-roll inflation paradigm requires that the inflaton potential $V(\phi)$ has a near-flat region which is able to support the exponential growth of the scale factor for enough e-folds. Thus the background solution of the spacetime looks like

$$\begin{aligned} &\phi = \phi_0(t) \\ &ds^2 = -dt^2 + a(t)^2 d\boldsymbol{x}^2 \end{aligned} \tag{1.48}$$

The solution should be homogeneous in the spatial coordinates because the universe we see now is homogeneous and isotropic. Upon the background solution, the dynamical degrees of freedom manifest themselves as fluctuations on the background. There are three independent fluctuations, one for the scalar part and two for the tensor part. The scalar fluctuation can be chosen by a gauge choice as either the scalar part of the metric or as the inflaton fluctuation. We take the ζ-gauge in which the scalar fluctuation is treated as a component of the metric fluctuations.

The primordial fluctuations are treated in the ADM formalism, which splits the metric into dynamical degrees of freedom and auxilliary fields:

$$ds^2 = -N^2 dt^2 + h_{ij}\left(dx^i + N^i dt\right)\left(dx^j + N^j dt\right) \tag{1.49}$$

The dynamical fluctuations are encoded in h_{ij}. The lapse N and shifts N_i are auxiliary fields which can be integrated out by solving the Hamiltonian constraints.

In ζ-gauge the perturbations look like:

$$\begin{aligned}
&\phi = \phi\left(t\right) \\
&h_{ij} = a^2\left(t\right) e^{2\zeta}\left(e^{\gamma}\right)_{ij} \\
&\gamma^i{}_i = 0 \\
&\partial_i \gamma^i{}_j = 0
\end{aligned} \tag{1.50}$$

where the ζ field encodes the scalar fluctuations and the transverse traceless tensor γ_{ij} is the graviton fluctuation.

The Lagrangian for the scalar perturbation ζ has the quadratic part

$$S_\zeta = M_p^2 \int dt d^3x \frac{\dot{\phi}^2}{H(t)^2} a\left(t\right)\left(a^2\dot{\zeta}^2 - \left(\partial_i\zeta\right)^2\right) \tag{1.51}$$

Using the conformal time $d\eta \equiv e^{-Ht}dt$, the massless scalar field in de Sitter space has the quadratic lagrangian

$$S_\zeta = \int d^4x \frac{M_p^2}{H^2\eta^2}\left(\left(\partial_\eta\zeta\right)^2 - \left(\partial_i\zeta\right)^2\right) \tag{1.52}$$

which gives the two-point functions for ζ

$$\langle \zeta_{\boldsymbol{k}}\zeta_{-\boldsymbol{k}}\rangle = \left(2\pi\right)^3 \frac{H^2}{2M_p^2\boldsymbol{k}^3}\left(1 + \boldsymbol{k}^2\eta^2\right) \tag{1.53}$$

At late time $\eta \to 0$, only the contribution from the first term remains.

The higher correlation functions can be computed with the interaction Hamiltonian. If in H_{int} there is a three-point interaction, then there is a non-trivial three-point function:

$$\langle \zeta^3\left(t\right)\rangle = -i\int_{-\infty(1+i\varepsilon)}^{t} dt' \langle\left[\zeta^3\left(t\right), H_{int}\right]\rangle \tag{1.54}$$

Similar calculations apply to arbitrary higher point functions.

1.3.2.2 Large scale structure as dark matter fluid

After the inflation phase, the universe went through the stages of radiation domination and matter domination, during which matter lumps formed under gravity. The system of dark matter lumps is described by a nearly perfect fluid coupled to gravity, with varying fluid velocity and density. In comoving coordinates the background configuration for the dark matter fluid is just a homogeneous fluid which moves with the expanding background:

$$\rho\left(\boldsymbol{x}, t\right) = \overline{\rho}(t) \ , \ v^i = 0 \tag{1.55}$$

The primordial fluctuations during inflation serve as the seeds for structure formation. Gravity amplifies the primordial inhomogeneity. The density fluctuations grow and exit the linear regime. To describe the density fluctuation, we introduce the overdensity δ,

$$\delta \equiv \frac{\rho}{\overline{\rho}} - 1 \tag{1.56}$$

The fluid velocity field itself can be treated as a fluctuation since the background value is zero everywhere. From the conservation of the stress-energy tensor we can derive the equations satisfied by the overdensity and the fluid velocity:

$$\begin{aligned} &\partial_\eta \delta + \partial_i \left(\left(1+\delta\right) v^i\right) = 0 \\ &\partial_\eta v^i + \mathcal{H} v^i + v^j \partial_j v^i = -\partial_i \Phi \end{aligned} \tag{1.57}$$

where η is the conformal time coordinate in which the metric of the FRW universe takes the form

$$ds^2 = a^2\left(\eta\right)\left(-d\eta^2 + d\boldsymbol{x}^2\right) \tag{1.58}$$

and $\mathcal{H}$ is the Hubble parameter defined with respect to conformal time

$$\mathcal{H} \equiv \frac{\partial_\eta a}{a} \tag{1.59}$$

This set of equations is supplemented by Poisson's equation from one of the constraints in Einstein's equations

$$\partial_i^2 \Phi = 4\pi G \overline{\rho} a^2 \delta \tag{1.60}$$

In the regime that δ and v^i are small, we can group the terms in the equations into linear and non-linear parts:

$$\begin{aligned} \partial_\eta \delta + \partial_i v^i &= -\partial_i \left(\delta v^i\right) \\ \partial_\eta v^i + \mathcal{H} v^i + \partial_i \Phi &= -v^j \partial_j v^i \end{aligned} \tag{1.61}$$

The dark matter fluid has no curl; i.e. the dark matter velociity field is a pure gradient with only one independent component. Thus we can defined a scalar quantity θ for the divergence of the velocity field

$$\theta \equiv \partial_i v^i \tag{1.62}$$

Then considering Poisson's equation and taking the divergence of the second equation of (1.61), the linear part of the set of equations become:

$$\begin{aligned} &\dot{\delta} + \partial^2 \theta = 0 \\ &\dot{\theta} + \mathcal{H}\theta + 4\pi G \overline{\rho} a^2 \delta = 0 \end{aligned} \tag{1.63}$$

It is of the general form:

$$\dot{\Theta} + M \cdot \Theta = 0 \tag{1.64}$$

where

$$\Theta \equiv (\delta, \theta) \tag{1.65}$$

The equations can be solved in perturbation theory by first solving the linear part, obtaining the linear propagator for the doublet Θ, and then adding in the non-linear terms as vertices. The correlation functions can be computed with diagrammatics. For instance, suppose we have a three-point vertex $V_{k,p,q}$in the theory and we have solved the linear theory so in the linear level we have

$$\Theta_k(\eta) = G(\eta, \eta_0)\,\Theta_k(\eta_0) + \int d\eta' G(\eta, \eta')\,V_{k,p,q}(\eta')\,G(\eta', \eta_0)\,G(\eta', \eta_0)\,\Theta_p(\eta_0)\,\Theta_q(\eta_0) \tag{1.66}$$

Then the correction to the power spectrum for Θ is computed by connecting the three-point diagrams with the power spectrum at initial time t_0

All other diagrams can be computed in a similar fashion. Described above is the standard perturbation theory (SPT) which treats both δ and θ as perturbations. However when the dark matter lumps form, δ can eqsily exceed order ~ 1, and in sufficiently short scales it is far larger than one. Thus we need ways to improve the SPT and develop a more powerful theory which is more systematic and can be applied to short scales where non-linearity becomes much more significant. Renormalized perturbation theory (RPT) and the effective field theory of large scale structure are two directions that are being developed.

Chapter 2

Bulk Microcausality from Boundary

2.1 Holographic Representation of Local Operators in de Sitter Space

In this chapter we discuss how the local operators that satisfy microcausality arise in the dual conformal field theory in AdS and dS space. At first sight, it seems impossible to get local observables in a $D+1$-dimensional curved spacetime from a local field theory in D-dimensions since there are simply not enough degrees of freedom to achieve this, and indeed the dual theory of the boundary CFT is a string theory, which is not a local field theory in the strict sense. However in the large-N limit, it is possible to construct approximately local operators in the bulk that satisfy microcausality with $\frac{1}{N}$ corrections. The large number of degrees of freedom makes up for the number of degrees of freedom, in a rough sense. Below we see explicitly how this works in AdS as well as in dS, and how the dS case differs from the AdS case in an essential way.

2.1.1 Introduction

Gauge/gravity duality[15], which equates a theory of quantum gravity to a quantum field theory in one lower dimension, has provided a deeper understanding of both non–perturbative string theories and conformal field theories, and also finds applications in different areas such as nuclear physics and condensed matter physics.

Despite the progress in the area of holographic duality, some basic questions regarding bulk locality remain to be clarified. Recently attention has been focused on sub–AdS locality [16][17]—locality of physics within the AdS radius, which might help understanding the recent puzzles regarding black holes [18]. It is well–known that in order to be dual to weakly–coupled gravity in the form of a local field theory in AdS, a conformal field theory must have a large number of degrees of freedom as well as being strongly coupled. The operator dictionary of AdS/CFT [19][20][21] can be understood as a series of claims about locality in the near–boundary region of AdS. There are two kinds of operator dictionaries in AdS/CFT. One of them is the GKPW dictionary[19][20] which identifies the boundary condition for a non–normalizable mode in AdS space as the coupling of a deformation to the boundary CFT, and the boundary correlation functions are obtained by differentiating this coupling to the partition function of bulk gravity. On the other hand, the BDHM dictionary [21] identifies the boundary condition for a normalizable mode as an operator in the un–deformed CFT, and then CFT correlation functions are recovered by extrapolating the bulk quantum gravity correlation functions to the boundary. In both cases, there is a one–one correspondence between a local operator in the bulk and a local operator on the boundary.

While the dictionary is well–defined in the limit that the bulk operator approaches the boundary, the story for an operator probing deeper inside the space is less transparent—such an operator corresponds to non–local operators on the boundary and the property of microcausality is not manifest. There are several approaches towards understanding this "sub–AdS locality" issue including the conformal bootstrap [16] and the use of Mellin representation of CFT correlation functions [17]. Constraints on operator dimensions and the behavior of Mellin amplitudes are conjectured. In [16], the authors count the constraints arising from the OPE, conformal invariance and the bootstrap conditions for large$-N$ conformal field theories in $d = 2$ and $d = 4$, and match the number of solutions to the constraints to the counting of quartic bulk local interactions. In [17], CFT correlation functions are formulated as scattering amplitudes in AdS space, with the help of a Mellin transform. It is demonstrated that to have local interactions in the AdS bulk, the Mellin amplitudes of the CFT should grow no faster than a power of the Mellin space coordinate δ, in the limit that δ is large. In this chapter, we focus on another approach which starts from microcausality and explicitly construct local operators from CFT data. The particular construction we are describing was developed in anti–de Sitter space by several authors [28][29][31], and recently further developed to describe the interior of eternal black holes in AdS space[32], in order to explore the "firewall" problem[18]. In this thesis, we parallelly develop the construction to local operators in de Sitter space, at the level of $\left(\frac{1}{N}\right)^0$ (two-point functions), in the context of the de Sitter/CFT correspondence [1].

It is still not completely clear whether quantum gravity in de Sitter space can be described holographically. There are several proposals for such a correspondence, including dS/CFT[1], dS/dS[33] and static patch solipsism[34]. Among these dS/CFT seems to be the simplest extension of AdS/CFT to de Sitter space in the sense that quantities like CFT correlation functions and bulk wavefunctions can be related to AdS case via analytic continuation, and the bulk de Sitter isometries match nicely to the conformal symmetry of the Euclidean theory at the future or past conformal boundary. Recently a realization of dS/CFT in the context of higher spin gravity was proposed[2]; namely the Vasiliev theory[35] in de Sitter space is conjectured to be dual to the critical or the free Euclidean

Sp(N) model with anti–commuting scalars. Despite a nice analogy to AdS/CFT as well as a certain proposed realization, the idea of dS/CFT suffers from several problems [36][37]: The CFT correlators are not observables for any observer in de Sitter space, but rather are "meta–observables"[1]; gravity is not decoupled; the dual field theory is non–unitary; and it is hard to see how bulk unitarity arises. Also the de Sitter nature of future infinity may be spoiled by bubble nucleation, and a boundary CFT may not exist at all.

In this thesis we will not get into any of these subtleties and will just assume the existence of the dS/CFT correspondence and try to construct local bulk observables from the boundary CFT data. We work in both the flat patch and the global patch of de Sitter space. In contrast to what happens in the case of AdS/CFT, at the level of two–point functions, a local operator in de Sitter space is shown to be constructed from two sets of single–trace operators in the boundary CFT, with dimensions Δ and $d-\Delta$ respectively. The observation that an operator in de Sitter is dual to two operators in CFT is not itself new, having already been pointed out by the original dS/CFT paper [1] by looking at the bulk correlation function in the limit that the bulk operators approach the boundary. In the paper by Harlow and Stanford[22], the GKPW (differentiating) and BDHM (extrapolating) dictionaries in de Sitter space are shown to be inequivalent—while the "differentiating" dictionary gives correlators with a single scaling dimension, the "extrapolating" dictionary gives correlators with two different near–boundary behaviors. To the knowledge of the author, the construction of de Sitter local operators is new, and helps clarify the understanding of how the bulk observables of de Sitter space emerge from a lower dimensional space, as well as clarifying the difference between dS/CFT and the analytic continuation of AdS/CFT. We also generalize the construction to gauge fields

[1] Actually we are meta–observers for the nearly de Sitter geometry during inflation [27], and the CMB correlation functions are "meta–observables" for the observers in an inflating universe. CMB is observable to us because after inflation the universe exits the near–de Sitter phase and the CMB photons fall into causal contact with us.

with integer spin s. It is shown that the construction for a spin$-s$ gauge field in de Sitter space can be identified with a construction of scalar fields with $m^2 < \left(\frac{d}{2}\right)^2$ and a bulk operator that matches with a spin$-s$ boundary current is explicitly constructed.

2.1.2 Construction of a Local Scalar Field in Anti–de Sitter Space

In this section I describe how a bulk scalar field in an empty anti–de Sitter space emerges from a conformal field theory. I briefly review the construction in AdS space following[28][29][31]. I work in the Poincaré patch, which has a direct analytic continuation to the flat slicing of de Sitter space.

There are two approaches leading to the same result. One is based on solving a space–like Cauchy problem and uses Green's function to express the local field, while the other starts from summing the normalizable modes in the bulk. In the Green's function approach, one first solves for the Green's function in AdS space

$$\left(\Box' - m^2\right) G(z,x|z',x') = \frac{1}{\sqrt{-g}}\delta^d(x-x')\delta(z-z') \tag{2.1}$$

Then from Green's theorem we have a bulk field expressed as

$$\Phi(z,x) = \int_{z'\to 0} d^dx' \sqrt{-g'}\left(G(z,x;z',x')\partial_{z'}\Phi(z',x') - \Phi(z',x')\partial_{z'}G(z,x;z',x')\right) \tag{2.2}$$

where for $\Phi(z,x)$ we just choose a single fall–off behavior near the boundary, as $z \to 0$

$$\Phi(z,x) \sim z^{\Delta}\mathcal{O}(x) \tag{2.3}$$

which corresponds to a normalizable solution to the bulk equation.

Pushing z' to the boundary and using the Green's function, finally we get

$$\Phi(z,x) = \int d^d x' K\left(z, x|x'\right) \mathcal{O}(x') \tag{2.4}$$

with $K(z, x|x')$ being a function which behaves like $z^{d-\Delta}$ when z approaches zero. We call it the "smearing function", for it smears over the operators in a certain region in CFT, defining a non–local operator in CFT as a local operator in the bulk. In the Poincaré patch it is

$$\boxed{K(z,x|x') = c_{d,\Delta}\left(\frac{z^2+(x-x')^2}{z}\right)^{\Delta-d}\Theta\left(z-|x-x'|\right)} \tag{2.5}$$

The domain of integration on the boundary is finite and within the intersection between the boundary and the bulk lightcone originated from the bulk operator, as shown in figure (2.1). Though it looks like an unconventional Cauchy problem—the "initial data" are spacelike separated from the bulk point, the result is causal: the commutator between two bulk operators constructed in this way vanishes when they are spacelike separated, to order N^0 in the large–N expansion. When considering interactions, the commutator turns on in three–point functions, which is order N^{-1}, but this can be cured by including multi–trace operators in the smearing prescription [29]. Schematically we have:

$$\Phi(z,x) = \int d^d x' K(z,x|x')\mathcal{O}(x') + \sum_i \int d^d x' K_{\Delta_i}\left(z,x|x'\right)\mathcal{O}_{\Delta_i}(x') \tag{2.6}$$

with multi–trace operators $\mathcal{O}_{\Delta_i}$ such as $\mathcal{O}^2$. The commutator at order N^{-1} is then cancelled by the contributions from the multi–trace operators. With a concrete demonstration at order N^{-1} [29], the procedure is conjectured to work order–by–order in the large–N expansion, and by adding multi–trace operators one can construct local operators in AdS to any order of $\frac{1}{N}$. The construction is believed to break down away from the large–N

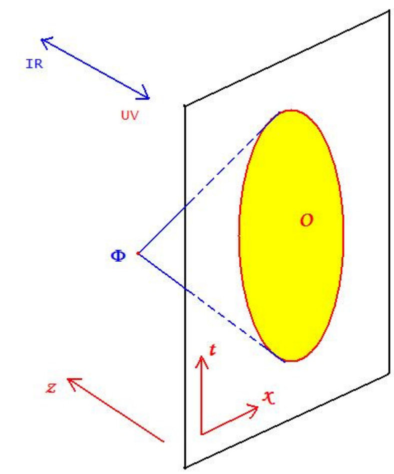


Figure 2.1: *Construction of A Local Observable in Anti–de Sitter space*

limit, where the bulk gravity fluctuates and the notion of microcausality as well as the notion of the background spacetime itself break down; one would not expect to define local observables in a full–fledged gravity theory [30]

The second approach—summing over modes—is more transparent for seeing why the result is causal. One starts by solving the free field equation in AdS and ends up with two independent solutions corresponding to each wave number. In the simple example of AdS_2 these modes are

$$\Phi_\omega(z) \sim C_1\sqrt{z}J_{\Delta-\frac{1}{2}}(z\omega) + C_2\sqrt{z}Y_{\Delta-\frac{1}{2}}(\omega z) \tag{2.7}$$

where J_ν and Y_ν are Bessel functions. Only the part proportional to J_ν is normalizable, so we just keep this branch of the solutions and sum over it.

Summing over the normalizable modes, including both positive frequencies and negative frequencies, we get

$$\Phi(z,t) = \int_0^\infty d\omega \left(a_\omega e^{-i\omega t} + a_\omega^\dagger e^{i\omega t}\right)\sqrt{z}J_{\Delta-\frac{1}{2}}(\omega z) \tag{2.8}$$

which will recover the same result obtained by the Green's function approach.

Here we can easily see why the result should be causal: although we just choose one of the fall–off behaviors in the z–direction, we are keeping both positive frequencies and negative frequencies in the time direction. This is crucial for ensuring microcausality, just as in flat spacetime. Thus we can express a local operator in AdS_{d+1} space in terms of the CFT_d operators inside its spatial lightcone, and these bulk operators satisfy microcausality.

We see that a local operator inside AdS space emerges as a non–local operator on the boundary. From the boundary point of view, the AdS coordinates t and z are just parameters defining the non–local operator in the CFT. The duality between bulk and boundary physics ensures that this particular non–local operator in CFT satisfies a free field equation in a higher dimension as well as being local in the sense of a higher dimensional microcausality. We should emphasize that the map between boundary and bulk here depends on the state of the boundary CFT [32] which maps to a certain bulk background geometry. The smearing function—and thus the construction—is made referring to a certain background. In this case it is empty AdS space, which is dual to the vacuum of a zero–temperature CFT. In a more general background, for instance, with a black hole sitting in the bulk, the construction would be different.

2.1.3 Analytic Continuation and Operator Dictionaries

We will see that though de Sitter space and anti–de Sitter space are related to each other via analytic continuation, the analytic continuation of the AdS smearing prescription above to de Sitter space does not give causal correlation functions. As a starting point one can analytically continue the AdS Poincaré patch to de Sitter flat slicing via

$$z \to \eta \ , \ t \to t \ , \ x^i \to ix^i \ , \ R_{AdS} = iR_{dS} \tag{2.9}$$

and get

$$ds^2 = \frac{-d\eta^2 + d\mathbf{x}^2}{\eta^2} \tag{2.10}$$

with t treated as one of the spatial coordinates in de Sitter space [2] .

Thus if one does the analytic continuation to the prescription introduced in the section above, a field operator in de Sitter space is then expressed as an integral defined on the past or future boundary. The domain of integration for smearing is in the past/future light cone of the bulk point, and it is now a standard Cauchy problem to express the bulk point in terms of boundary operators as evolving the initial conditions using the retarded Green's function in de Sitter space.

However this cannot be what we aim for. First, it violates microcausality: after the analytic continuation, the spatial lightcone in AdS becomes time–like, and the bulk operator now commutes with the operators inside its own time–like lightcone and fails to commute with the ones outside, which is not the right behavior for being causal. One can see the reason why this happens—in AdS smearing procedure, to construct a local operator, we just need to sum over the set of normalizable modes. Continuing to de Sitter space the z-direction becomes the time direction and keeping only one set of modes in this direction turns into keeping either positive or negative frequency modes, which spoils microcausality. In AdS we do not have this problem because we go from the bulk to the boundary in a spatial direction, and we can still keep both positive frequency and negative frequency modes in the time direction while sticking to just normalizable modes in the z-direction.

Second, from the smearing prescription above one recovers the correct AdS bulk correlation functions, but as has been pointed out by several authors [22][26][1], the analytic continuation of AdS correlators would not give the correlation function in any de Sit-

[2]Here we just set R_{dS} to one.

ter invariant vacuum. In paper [22] the authors used the language of the holographic renormalization group [25] to clarify this point. They claim that in de Sitter space the "GKPW dictionary" and the "BDHM dictionary" are not equivalent, though the vacuum wavefunctions in AdS and dS are related by analytic continuation. The reason why this happens is the definitions of correlation functions in dS and AdS are not related to each other by analytic continuation. In the language of the holographic RG, one can define the correlation functions in AdS in the following way [22]: split the bulk path integral with a plane at $z = l$ and the path integrals in the UV side and IR side give UV and IR wavefunctions separately, and then one can insert operators on the plane and thus obtain a bulk correlation function

$$\langle\tilde{\phi}(x_1, l)\ldots\tilde{\phi}(x_n, l)\rangle_{AdS} = \int_{z=l} \mathcal{D}\tilde{\phi}\Psi_{IR}[\tilde{\phi}]\tilde{\phi}(x_1, l)\ldots\tilde{\phi}(x_n, l)\Psi_{UV}[\tilde{\phi}, \phi_0] \tag{2.11}$$

where ϕ_0 is the boundary condition for the path integral in Ψ_{UV}. One recovers the boundary correlation function by taking the limit $l \to 0$, and it agrees with the result one gets by differentiating with the boundary condition ϕ_0 [22]. The Ψ_{IR} is shown to be related to the Hartle–Hawking vacuum in de Sitter space Ψ_{HH} via analytic continuation [22]; however if we analytically continue the definition of the correlation function to de Sitter space one gets something peculiar: taking the future wedge of flat slicing, the Ψ_{UV} is now a wavefunction in the later stage of the universe and we call it Ψ_L, and Ψ_{IR} is defined in an earlier period and we call it Ψ_E. Then the analytically continued correlation function is defined as

$$\langle\tilde{\phi}(x_1, \eta)\ldots\tilde{\phi}(x_n, \eta)\rangle_{dS} = \int_{\eta} \mathcal{D}\tilde{\phi}\Psi_E[\tilde{\phi}]\tilde{\phi}(x_1, \eta)\ldots\tilde{\phi}(x_n, \eta)\Psi_L[\tilde{\phi}, \phi_0] \tag{2.12}$$

This is different from how one computes correlation functions in de Sitter space, or in a more generic FRW cosmology. With this definition, in order to compute the correlation functions at a certain time η it is not enough to know the earlier stage evolution of the wavefunction, but one needs as well the later stage. This is not what one would do in cosmology, since to compute the correlation functions of temperature fluctuations in the cosmic microwave background we don't have to know the wavefunction of the universe during the subsequent structure formation. Also, fixing a certain boundary condition at future infinity is manifestly acausal [24]. The radiation fails to pass through future infinity and will be reflected back into the past. This acausal behavior will manifest itself as the breakdown of microcausality: operators on a single spatial slice will fail to commute. The way to define the correlation function in de Sitter and in more generic cosmology should just involve the Hartle–Hawking wavefunction and its complex conjugate, and corresponds to an in–in path integral:

$$\langle\Psi|\tilde{\phi}(x_1,\eta)\ldots\tilde{\phi}(x_n,\eta)|\Psi\rangle_{dS,FRW} = \int_\eta \mathcal{D}\tilde{\phi}\,\Psi_E^*[\tilde{\phi}]\tilde{\phi}(x_1,\eta)\ldots\tilde{\phi}(x_n,\eta)\Psi_E[\tilde{\phi}] \tag{2.13}$$

where η is a certain spatial slice such as the last scattering surface of CMB photons in our universe, on which we compute correlation functions and compare with data, and Ψ_E refers to both "a wavefunction at early time" and "a wavefunction of the universe in the Euclidean (Hartle–Hawking) vacuum". Here one no longer specifies the boundary condition at the future boundary. This is a natural definition of expectation values under the Born rule, and it is clearly different from the analytic continuation from AdS. Also this definition obeys microcausality—the spacelike separated operators commute inside the correlation functions and timelike separated ones do not commute. The simplest one of this type of correlation function is the Wightman function for a free scalar field in de

Sitter space[3]. Thus a construction of a de Sitter bulk operator that computes de Sitter cosmology should reproduce the Wightman function, and it should also contain both positive and negative frequency modes in de Sitter space in order to ensure causality.

2.1.4 de Sitter Construction

In this subsection, we construct the local scalar operators in de Sitter space explicitly from operators in an Euclidean CFT. We perform the construction in both the flat patch and the global patch. We also explore the issues associated with building up local gauge fields in de Sitter space, and try to rewrite the construction in terms of an embedding formalism which is explicitly dS(AdS) covariant.

2.1.4.1 Flat Slicing

Now we look at how a local scalar operator with mass $m^2 > \left(\frac{d}{2}\right)^2$ in de Sitter space is constructed from a CFT located at the boundary. In the AdS construction the boundary is timelike, and the extra direction emerges from the boundary as a spatial direction. In de Sitter space, the boundaries are located at future and past conformal infinity, which are spacelike boundaries, so what emerges from the CFT is the bulk time. From the boundary point of view the bulk time η appears as a parameter in the definition of non–local CFT operators. As we will see, a local bulk operator that is far from the boundary will be highly non–local from the CFT point of view.

In this subsection we work in the flat patch of de Sitter space, which covers only half of the global geometry. One can either choose the past wedge to work on, or the future wedge, and the boundary CFT will live on $\mathcal{I}^-$ or $\mathcal{I}^+$ respectively. Here for the moment we choose the past wedge. The construction in the global patch of de Sitter space is left

[3]Wightman function is the expectation value of the product of two field operators for the same field inserted at different points, in the vacuum state of the field theory.

to the next subsection.

We have seen that a construction prescription for local operators in de Sitter space should involve modes with both positive and negative frequencies, corresponding to "normalizable" and "non-normalizable" behaviors in AdS. Here we define

$$\Delta = \frac{d}{2} + i\sqrt{m^2 - \left(\frac{d}{2}\right)^2} \tag{2.14}$$

and near the boundary a positive/negative frequency mode has behavior

$$\begin{aligned} \Phi_+ \left(\eta \to 0\right) &\sim \eta^{\Delta} \mathcal{O}_+ \\ \Phi_- \left(\eta \to 0\right) &\sim \eta^{d-\Delta} \mathcal{O}_- \end{aligned} \tag{2.15}$$

where $\mathcal{O}_\pm$ are single–trace operators in the boundary CFT, with scaling dimensions Δ and $d - \Delta$ respectively.

For the case of interest here, since $m^2 - \left(\frac{d}{2}\right)^2$ is positive, near the boundary both Φ_+ and Φ_- are damped by the same factor $\eta^{\frac{d}{2}}$ and oscillate with frequency $\sqrt{m^2 - \left(\frac{d}{2}\right)^2}$. If $m^2 < \left(\frac{d}{2}\right)^2$ then the two modes fall at different rates near the boundary and do not oscillate.

According to the reasoning in the section above, a causal operator should have both components, schematically:

$$\Phi(\eta \to 0) \sim A\eta^{\Delta}\mathcal{O}_+ + B\eta^{d-\Delta}\mathcal{O}_- \tag{2.16}$$

With a certain linear combination, one can reproduce the Wightman function in the Euclidean vacuum.

To construct the bulk operator, we evolve the initial data at $\mathcal{I}^-$ with the retarded

Green's function, which is

$$G_{ret}|_{\eta'\to 0} \approx c_{\Delta,d}(-\sigma - i\epsilon)^{\Delta-d} + c^*_{\Delta,d}(-\sigma - i\epsilon)^{-\Delta} - c.c. \tag{2.17}$$

in the limit that $\eta' \to 0$. Here

$$\sigma = \frac{\eta^2 + \eta'^2 - (\mathbf{x} - \mathbf{x}')^2}{2\eta\eta'} \tag{2.18}$$

is a de Sitter invariant distance and

$$c_{\Delta,d} = \frac{\Gamma(2\Delta - d)\Gamma(d - \Delta)}{2^{\Delta-d}(4\pi)^{\frac{d+1}{2}}\Gamma(\Delta - \frac{d-1}{2})} \tag{2.19}$$

The bulk operator is constructed by evolving an operator near the boundary:

$$\Phi(\eta, \mathbf{x}) = \lim_{\eta'\to 0}\int_{|\mathbf{x}'|<\eta} d^d x' \left(\frac{1}{\eta'}\right)^{d-1} \left(G_{ret}(\eta, \mathbf{x}; \eta', \mathbf{x}')\partial_{\eta'}\Phi(\eta', \mathbf{x}') - \Phi(\eta', \mathbf{x}')\partial_{\eta'}G_{ret}(\eta, \mathbf{x}; \eta', \mathbf{x}')\right) \tag{2.20}$$

where

$$\Phi(\eta', \mathbf{x}') \sim A\,(\eta')^{\Delta}\,\mathcal{O}_+(\mathbf{x}') + B\,(\eta')^{d-\Delta}\,\mathcal{O}_-\,(\mathbf{x}') \tag{2.21}$$

By keeping both sets of operators, we are keeping both the positive and negative frequency parts of the solution [4].

Evaluating the integrand, we have the local operator in de Sitter space expressed as

$$\begin{aligned}\Phi(\eta, \mathbf{x}) =& A_{\Delta,d}\int_{|\mathbf{x}'|<\eta} d^d x' \left(\frac{\eta^2 - \mathbf{x}'^2}{\eta}\right)^{\Delta-d} \mathcal{O}_+(\mathbf{x} + \mathbf{x}') + \\ & B_{\Delta,d}\int_{|\mathbf{x}'|<\eta} d^d x' \left(\frac{\eta^2 - \mathbf{x}^2}{\eta}\right)^{-\Delta} \mathcal{O}_-(\mathbf{x} + \mathbf{x}')\end{aligned} \tag{2.22}$$

[4] Before we take the limit $\eta' \to 0$, there should be four kinds of components present: $\sigma^{\Delta-d}\mathcal{O}_\pm$ and $\sigma^{-\Delta}\mathcal{O}_\pm$, but in the limit $\eta' \to 0$ only $\sigma^{\Delta-d}\mathcal{O}_+$ and $\sigma^{-\Delta}\mathcal{O}_-$ survive because the other contributions oscillate quickly and go to zero as η' approaches zero. The details are presented in Appendix B.

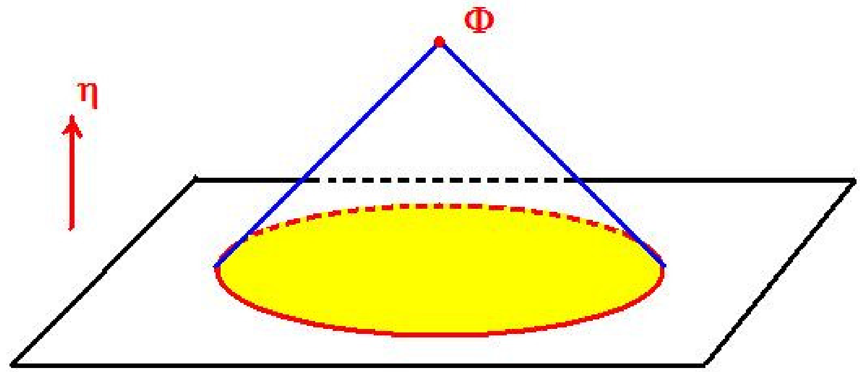


Figure 2.2: *Construction in the Flat Patch of De Sitter Space*

where $A_{\Delta,d}$ and $B_{\Delta,d}$ are certain coefficients. Here for a moment we keep them free, since in principle we can rescale the boundary operators and change the coefficients of the two–point functions, as a marginal deformation to the boundary CFT. This freedom of rescaling the operators, as well as the freedom of choosing a certain linear combination of two modes with different fall–off behaviors, enables us to keep $A_{\Delta,d}$ and $B_{\Delta,d}$ free for a moment. We will finally fix them by demanding that the correlation function of Φ recover the Wightman function in the Euclidean vacuum. This means that the choice of the coefficients is state–dependent: for other de Sitter invariant vacua such as "α–vacua", we should have different prescriptions in order to recover the Wightman functions.

Having a prescription, we would like to calculate the bulk two–point function and compare it with the two–point bulk Wightman function in the limit that one of the bulk points approaches the boundary. The Euclidean vacuum Wightman function in de Sitter space is [26]

$$W\left(x,x'\right)=\frac{\Gamma(\Delta)\Gamma(d-\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\frac{d+1}{2})}F\left(\Delta,d-\Delta,\frac{d+1}{2},\frac{1+\sigma}{2}\right) \tag{2.23}$$

where $F(\alpha,\beta,\gamma,x)$ is the hypergeometric function ${}_2F_1$, and σ is the de Sitter invariant distance defined beforehand.

When one of the bulk points x' approaches the boundary $\eta' \to 0$, the fourth argument of the hypergeometric function grows large and is dominated by σ

$$\frac{1+\sigma}{2} \sim \frac{\sigma}{2} \sim \frac{\eta^2 - (\mathbf{x}-\mathbf{x}')^2}{4\eta\eta'} \tag{2.24}$$

For convenience we can set $\mathbf{x}'$ to zero. In this limit we have

$$\begin{aligned} W(\eta, \mathbf{x}; \eta' \sim 0, \mathbf{x}' = 0) \to & \frac{\Gamma(\Delta)\Gamma(d-2\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\frac{d+1}{2}-\Delta)} \left(-\frac{4\eta\eta'}{\eta^2 - \mathbf{x}^2}\right)^{\Delta} + \\ & \frac{\Gamma(2\Delta-d)\Gamma(d-\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\Delta-\frac{d-1}{2})} \left(-\frac{4\eta\eta'}{\eta^2 - \mathbf{x}^2}\right)^{d-\Delta} \end{aligned} \tag{2.25}$$

As expected, the Wightman function has two components with dimensions Δ and $d-\Delta$. Next we want to reproduce it from the smearing formula (2.22).

Here we would like to normalize the boundary two–point functions so that we have

$$W(\eta \to 0, \mathbf{x}; \eta' \to 0, \mathbf{x}' = 0) \to (\eta\eta')^{\Delta} D_+(\mathbf{x}) + (\eta\eta')^{d-\Delta} D_-(\mathbf{x}) \tag{2.26}$$

where $D_\pm$ are the boundary CFT correlation functions which we take to be

$$D_+(\mathbf{x}) = \frac{2^{2\Delta}\Gamma(\Delta)\Gamma(d-2\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\frac{d+1}{2}-\Delta)} \left(\frac{1}{\mathbf{x}^2}\right)^{\Delta} \tag{2.27}$$

$$D_-(\mathbf{x}) = \frac{2^{2(d-\Delta)}\Gamma(2\Delta-d)\Gamma(d-\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\Delta-\frac{d-1}{2})} \left(\frac{1}{\mathbf{x}^2}\right)^{d-\Delta} \tag{2.28}$$

Taking the smearing formula (2.22) and computing the correlation function between the bulk operator and an operator near the boundary, we have

$$\begin{aligned}\langle\Phi(\eta,\mathbf{x})\Phi(\eta'\to 0,0)\rangle &= A_{\Delta,d}\int_{|\mathbf{x}'|<\eta} d^dx'\left(\frac{\eta^2-\mathbf{x}'^2}{\eta}\right)^{\Delta-d}\eta'^{\Delta}\langle\mathcal{O}_+(\mathbf{x}+\mathbf{x}')\mathcal{O}_+(0)\rangle\\ &+ B_{\Delta,d}\int_{|\mathbf{x}'|<\eta} d^dx'\left(\frac{\eta^2-\mathbf{x}'^2}{\eta}\right)^{-\Delta}\eta'^{d-\Delta}\langle\mathcal{O}_-(\mathbf{x}+\mathbf{x}')\mathcal{O}_-(0)\rangle\end{aligned} \tag{2.29}$$

With the boundary correlator of the operators $\mathcal{O}_\pm$:

$$\langle\mathcal{O}_+(\mathbf{x})\mathcal{O}_+(0)\rangle = D_+(\mathbf{x})\ ,\ \langle\mathcal{O}_-(\mathbf{x})\mathcal{O}_-(0)\rangle = D_-(\mathbf{x})\ ,\ \langle\mathcal{O}_+(\mathbf{x})\mathcal{O}_-(0)\rangle = 0 \tag{2.30}$$

we obtain

$$\begin{aligned}\langle\Phi(\eta,\mathbf{x})\Phi(\eta'\to 0,0)\rangle &= A_{\Delta,d}\frac{2^{2\Delta}\Gamma(\Delta)\Gamma(d-2\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\frac{d+1}{2}-\Delta)}\int_{|\mathbf{x}'|<\eta} d^dx'\left(\frac{\eta^2-\mathbf{x}'^2}{\eta}\right)^{\Delta-d}\eta'^{\Delta}\frac{1}{(\mathbf{x}+\mathbf{x}')^{2\Delta}}\\ &+ B_{\Delta,d}\frac{2^{2(d-\Delta)}\Gamma(2\Delta-d)\Gamma(d-\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\Delta-\frac{d-1}{2})}\int_{|\mathbf{x}'|<\eta} d^dx'\left(\frac{\eta^2-\mathbf{x}'^2}{\eta}\right)^{-\Delta}\eta'^{d-\Delta}\frac{1}{(\mathbf{x}+\mathbf{x}')^{2(d-\Delta)}}\end{aligned} \tag{2.31}$$

After evaluating the integrals we end up with the result

$$\begin{aligned}&\langle\Phi(\eta,\mathbf{x})\Phi(\eta'\to 0,0)\rangle =\\ &A_{\Delta,d}\frac{\pi^{\frac{d}{2}}\Gamma(\Delta-d+1)}{\Gamma(\Delta-\frac{d}{2}+1)}\frac{2^{2\Delta}\Gamma(\Delta)\Gamma(d-2\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\frac{d+1}{2}-\Delta)}(-1)^{\Delta}\left(\frac{\eta\eta'}{\eta^2-\mathbf{x}^2}\right)^{\Delta}+\\ &B_{\Delta,d}\frac{\pi^{\frac{d}{2}}\Gamma(1-\Delta)}{\Gamma(\frac{d}{2}-\Delta+1)}\frac{2^{2(d-\Delta)}\Gamma(2\Delta-d)\Gamma(d-\Delta)}{(4\pi)^{\frac{d+1}{2}}\Gamma(\Delta-\frac{d-1}{2})}(-1)^{d-\Delta}\left(\frac{\eta\eta'}{\eta^2-\mathbf{x}^2}\right)^{d-\Delta}\end{aligned} \tag{2.32}$$

We see that the coefficients

$$A_{\Delta,d} = \frac{\Gamma(\Delta - \frac{d}{2} + 1)}{\pi^{\frac{d}{2}}\Gamma(\Delta - d + 1)}$$
$$B_{\Delta,d} = A_{d-\Delta,d} = \frac{\Gamma(\frac{d}{2} - \Delta + 1)}{\pi^{\frac{d}{2}}\Gamma(1 - \Delta)}$$

give the two–point Wightman function in the Euclidean vacuum, in the limit that one of the bulk points approaches the boundary. It is not hard to show that the prescription also gives Wightman function away from this limit. Notice that each of the two terms in the smearing prescription can actually be obtained by analytically continuing from AdS space. Thus the Wightman function breaks into two pieces. One is for a scalar with $\Delta_+ = \Delta = \frac{d}{2} + i\sqrt{m_{dS}^2 - \left(\frac{d}{2}\right)^2}$ and the other for a scalar with $\Delta_- = d - \Delta_+$. The two–point functions for the Δ_+ and Δ_- components are in the form:

$$G_\Delta(x, x') = \left(\frac{2}{1+\sigma}\right)^{\Delta} F\left(\Delta, \Delta - \frac{d-1}{2}, 2\Delta - d + 1, \frac{2}{1+\sigma}\right) \tag{2.33}$$

with $\Delta = \Delta_\pm$.

Summing the two pieces using the property of hypergeometric function introduced in Appendix A, with the coefficients obtained in this section, we recover the Wightman function in de Sitter space with both points deep in the bulk.

Now we can write the expression for a local operator in de Sitter space explicitly, with coefficients set by the Euclidean vacuum state:

$$\boxed{\begin{aligned}\Phi(\eta, \mathbf{x}) = \frac{\Gamma(\Delta - \frac{d}{2} + 1)}{\pi^{\frac{d}{2}}\Gamma(\Delta - d + 1)} \int_{|\mathbf{x}'|<\eta} d^d x' \left(\frac{\eta^2 - \mathbf{x}'^2}{\eta}\right)^{\Delta - d} \mathcal{O}_+(\mathbf{x} + \mathbf{x}') + \\ \frac{\Gamma(\frac{d}{2} - \Delta + 1)}{\pi^{\frac{d}{2}}\Gamma(1 - \Delta)} \int_{|\mathbf{x}'|<\eta} d^d x' \left(\frac{\eta^2 - \mathbf{x}^2}{\eta}\right)^{-\Delta} \mathcal{O}_-(\mathbf{x} + \mathbf{x}')\end{aligned}} \tag{2.34}$$

This is our main result in this section: to construct local operators in de Sitter space that probe and create particles in the Euclidean vacuum state, we start from the Wightman function in the bulk Euclidean vacuum $W(x, x')$, and construct the retarded propagator by taking the expectation value of the commutator $G_{ret} \equiv W(x, x') - W(x', x)$ which has support only inside the bulk time–like lightcone. This retarded propagator gives the smearing functions for CFT operators with coefficients $A_{\Delta,d}$ and $B_{\Delta,d}$, and using the constructed smearing function we can recover the Wightman function we started with. What we get is a representation of local bulk operator in terms of boundary CFT operators, in a certain vacuum state. The bulk operator is constructed with CFT data inside the past lightcone of the bulk point, as shown in figure (2.2).

Here one can also see that, as opposed to the AdS case, to check microcausality one is no longer supposed to just compute the correlation function between a bulk operator and a single boundary operator. Considering for example $\mathcal{O}_+$, the result one gets in this way is the same as the one continued from AdS, and thus acausal. The reason why we shouldn't do this is clear: unlike the case for AdS, $\mathcal{O}_+$ or $\mathcal{O}_-$ alone no longer match smoothly onto any local bulk operator that approaches the boundary.

The construction above in terms of CFT operators at the past boundary $\mathcal{I}^-$ is not directly relevant to cosmology. In cosmology it is the flat FRW slicing of de Sitter space defined on the future wedge which is relevant. It describes the expansion phase of the universe. In the future wedge the bulk operators are constructed with CFT operators on $\mathcal{I}^+$, which seems unappealing because the "retarded propagator" is now propagating the boundary operators back in time. However in terms of the physical observables everything is causal: by rerunning the calculation we obtain the Wightman function in the future wedge, and the operators satisfy microcausality. From the point of view of the evolution of wavefunctions, it is more appealing to phrase the construction in the future wedge: to compute correlation functions at a late time $\eta \to 0$ one starts with the vacuum state defined on some spatial slice at earlier time $\eta \to -\infty$, evolves forward in time, and then computes the expectation value. In the above construction we chose $\mathcal{I}^-$ simply because it is more appealing from the perspective of the retarded propagator. One can rerun everything we formulated above in the future wedge and get a local operator in the future wedge in terms of operators at $\mathcal{I}^+$.

One may also question about the operator content: in a CFT, it seems we don't necessarily have enough operator content for the construction. Take the example of a scalar. It is totally possible that the theory only contains a dynamical scalar current $\mathcal{O}_+$ with dimension Δ but not one with dimension $d - \Delta$. However if we couple the operator $\mathcal{O}_+$ to the Lagrangian via a coupling term $\beta\mathcal{O}_+$, then the coupling β naturally has dimension $d - \Delta$. This means that when doing the path integral for the CFT, we are not only integrating over the constituents that forming $\mathcal{O}_+$, but also the source [5]. This is legitimate as one can always treat the coupling as a multiplier and integrate over it. Furthermore, this is consistent with the fact that when computing expectation values in de Sitter space, we have to integrate over sources also. An example is computing correlation functions at future infinity with the wavefunction of the universe $\Psi[g_{ij}]$. To obtain this wavefunction we do a path integral in de Sitter space using the Hartle–Hawking prescription, which is equivalent to coupling g_{ij} as a source to the boundary stress tensor and integrating over the CFT field contents. When we want to compute correlation functions on the boundary, according to Born's rule, we also have to integrate over g_{ij}, which is the degrees of freedom that corresponds to a certain classical configuration of gravitational field. A CFT enlarged to include operators corresponding to sources then becomes a "doubled CFT" as discussed in [22][23]. In the next subsection, we can see that the second set of operators $\mathcal{O}_-$ also has a natural interpretation as "shadow operators" in the CFT.

[5] I thank Frederik Denef for pointing this out.

Further, one may wonder if by writing down $\Phi \to A\eta^{\Delta}\mathcal{O}_+ + B\eta^{d-\Delta}\mathcal{O}_-$ near the boundary, we are imposing boundary conditions similar to what we do in anti–de Sitter space. From the perspective of path integration, we have seen that what produces the Wightman function in de Sitter space is an in–in type path integral which does not fix any boundary condition at the future or past boundary. Rather it specifies a particular vacuum state in which we calculate the expectation values. Indeed we are not fixing the boundary condition even we write down a schematic form of the operator near the boundary. The reason is that we are not fixing the coefficients: Φ can be any linear combination of the two components. What fixes a particular linear combination is the vacuum we choose: here we chose a particular A and B to recover the Wightman function in the Euclidean vacuum. Therefore there is no contradiction with the fact that the correlation functions are computed by an in–in path integral.

One further comment about the case of $m^2 \leq \left(\frac{d}{2}\right)^2$ here. Our construction is done for the case $m^2 > \left(\frac{d}{2}\right)^2$ and we have seen the positive and negative frequency modes join nicely into the Wightman function in the Euclidean vacuum. For $m^2 < \left(\frac{d}{2}\right)^2$, apart from some specific values, one can continue our result trivially and obtain Wightman function for the light scalar. This corresponds to summing over bulk modes with near boundary behavior η^{Δ} and $\eta^{d-\Delta}$ with $\Delta = \frac{d}{2} + \sqrt{\left(\frac{d}{2}\right)^2 - m^2}$. However in this case the modes have no oscillatory behavior—they just fall at different rates, thus the interpretation of positive–negative frequencies is not a good one despite the fact that we can construct the local bulk operator in the same way. For some specific values of m^2, the construction fails, as we will see in section 2.1.4.3. The reason is that when $2\Delta - d$ is an integer which happens when we have a light scalar with $m^2 = -(s-2)(s+d-2) \leq \left(\frac{d}{2}\right)^2$, with s a positive integer which turns out to be the spin of a gauge field, the Wightman function no longer split into two parts corresponding to complementary dimensions and there is a logarithmic term. Our construction formula has explicit singularities at such mass parameters.

2.1.4.2 Global Slicing

The construction in the global patch is similar to the flat patch, but with new elements from having two boundaries. Now we can define conformal field theory operators separately at $\mathcal{I}^+$ and $\mathcal{I}^-$; for any local operator in the bulk, the CFT operators at $\mathcal{I}^+$ and $\mathcal{I}^-$ can be regarded as two different bases, which should be related to each other via a Bogoliubov transformation [26]. In order to see the relation, one may construct a local field with CFT operators in $\mathcal{I}^-$ and push it to $\mathcal{I}^+$, or vice versa, thus getting the expression of an operator on $\mathcal{I}^+$ in terms of operators on $\mathcal{I}^-$. As a starting point, we first formulate the global patch smearing function.

In the global patch we work in conformal time, with the metric

$$ds^2 = \frac{1}{\cos^2 \tau} \left(-d\tau^2 + d\Omega_d^2\right). \tag{2.35}$$

Here the topology of the spacetime is $R \times S^d$ with the conformal time τ running from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$.

In these coordinates, the de Sitter invariant distance is expressed as

$$\sigma(x, x') = \frac{\cos(\Omega - \Omega') - \sin\tau \sin\tau'}{\cos\tau \cos\tau'}. \tag{2.36}$$

As x' goes to the future boundary $\mathcal{I}^+$, τ' goes to $\frac{\pi}{2}$, and the regularized distance from a bulk point to the boundary point is

$$\sigma(x, x') \cos\tau' \sim \frac{\cos(\Omega - \Omega') - \sin\tau}{\cos\tau} \tag{2.37}$$

Therefore the smearing functions that evolve future boundary operators back into the bulk will be proportional to

$$K_{+}^{\mathcal{I}^{+}}(\tau,\Omega|\Omega') \sim \left(\frac{\cos(\Omega-\Omega')-\sin\tau}{\cos\tau}\right)^{\Delta-d}$$
$$K_{-}^{\mathcal{I}^{+}}(\tau,\Omega|\Omega') \sim \left(\frac{\cos(\Omega-\Omega')-\sin\tau}{\cos\tau}\right)^{-\Delta}$$

with the support to be the region on the boundary inside the bulk lightcone. A simple example is dS_{1+1}, for which the support for the smearing function is

$$|\rho-\rho'| < \frac{\pi}{2} - \tau \tag{2.38}$$

on $\mathcal{I}^+$, where $\rho \in [-\pi, \pi]$ is the spatial coordinate of dS_{1+1}.

For the smearing functions evolving operators from the past boundary $\mathcal{I}^-$, we have

$$K_{+}^{\mathcal{I}^{-}}(\tau,\Omega|\Omega') \sim \left(\frac{\cos(\Omega-\Omega')+\sin\tau}{\cos\tau}\right)^{\Delta-d}$$
$$K_{-}^{\mathcal{I}^{-}}(\tau,\Omega|\Omega') \sim \left(\frac{\cos(\Omega-\Omega')+\sin\tau}{\cos\tau}\right)^{-\Delta}$$

and the support for the case of dS_{1+1}, is

$$|\rho-\rho'| < \tau + \frac{\pi}{2} \tag{2.39}$$

The supports for the $\mathcal{I}^+$ and $\mathcal{I}^-$ smearing functions are each defined within a single lightcone originating from the bulk point and extending to both past and future, as shown in figure (2.3). The smearing prescription for a dimension Δ operator which reduces to

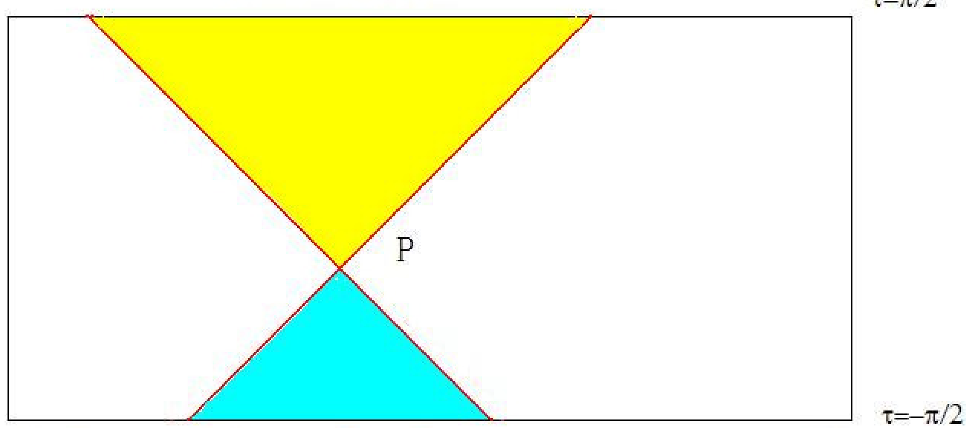


Figure 2.3: *Local Operator In the Global 1+1 dimensional de Sitter space*

the flat de Sitter space smearing function is

$$\boxed{\begin{aligned}\Phi(\tau,\Omega) = &\frac{2^{\Delta-d}\Gamma(\Delta-\frac{d}{2}+1)}{\pi^{\frac{d}{2}}\Gamma(\Delta-d+1)}\int d\Omega'\left(\frac{\cos(\Omega-\Omega')\mp\sin\tau}{\cos\tau}\right)^{\Delta-d}\mathcal{O}_+(\Omega')\\ &+\frac{2^{-\Delta}\Gamma(\frac{d}{2}-\Delta+1)}{\pi^{\frac{d}{2}}\Gamma(1-\Delta)}\int d\Omega'\left(\frac{\cos(\Omega-\Omega')\mp\sin\tau}{\cos\tau}\right)^{-\Delta}\mathcal{O}_-(\Omega')\end{aligned}} \tag{2.40}$$

respectively for $\mathcal{I}^{\pm}$ smearing, with the integration inside the bulk lightcone region. This will reproduce the Wightman function in the Euclidean vacuum, expressed in global coordinates.

One can construct a bulk operator from $\mathcal{I}^-$ and push it to $\mathcal{I}^+$, in which limit

$$\begin{aligned}\Phi(\tau\to\frac{\pi}{2},\Omega)\to(\cos\tau)^{d-\Delta}&\frac{2^{\Delta-d}\Gamma(\Delta-\frac{d}{2}+1)}{\pi^{\frac{d}{2}}\Gamma(\Delta-d+1)}\int d\Omega'\left(\cos(\Omega-\Omega')+1\right)^{\Delta-d}\mathcal{O}_+(\Omega')\\ &+(\cos\tau)^{\Delta}\frac{2^{-\Delta}\Gamma(\frac{d}{2}-\Delta+1)}{\pi^{\frac{d}{2}}\Gamma(1-\Delta)}\int d\Omega'\left(\cos(\Omega-\Omega')+1\right)^{-\Delta}\mathcal{O}_-(\Omega')\end{aligned} \tag{2.41}$$

This has precisely the form we would expect as the boundary limit of a local bulk operator, but now with the "−" component expressed by $\mathcal{O}_+$ on $\mathcal{I}^-$ and vice versa:

$$\begin{aligned}
\mathcal{O}_+(\Omega,\mathcal{I}^+) &= \frac{2^{-\Delta}\Gamma(\frac{d}{2}-\Delta+1)}{\pi^{\frac{d}{2}}\Gamma(1-\Delta)}\int d\Omega'\,(\cos(\Omega-\Omega')+1)^{-\Delta}\,\mathcal{O}_-(\Omega',\mathcal{I}^-)\\
\mathcal{O}_-(\Omega,\mathcal{I}^+) &= \frac{2^{\Delta-d}\Gamma(\Delta-\frac{d}{2}+1)}{\pi^{\frac{d}{2}}\Gamma(\Delta-d+1)}\int d\Omega'\,(\cos(\Omega-\Omega')+1)^{\Delta-d}\,\mathcal{O}_+(\Omega',\mathcal{I}^-)
\end{aligned} \tag{2.42}$$

Here the integration is over the whole past boundary—the past lightcone covers the whole d–sphere. The equations above can be regarded as a Bogoliubov transformation in coordinate space, as well as an operator dictionary relating two copies of CFT on $\mathcal{I}^\pm$

The form of this boundary–boundary map can be better formulated by expressing operators not at angular position Ω, but its antipodal point $\tilde{\Omega}$ instead, since $\cos(\tilde{\Omega}-\Omega') = -\cos(\Omega-\Omega')$ we have

$$\boxed{\begin{aligned}
\mathcal{O}_+(\tilde{\Omega},\mathcal{I}^+) &= \alpha_{\Delta,d}\int d\Omega'\langle\mathcal{O}_+(\Omega)\mathcal{O}_+(\Omega')\rangle\mathcal{O}_-(\Omega',\mathcal{I}^-)\\
\mathcal{O}_-(\tilde{\Omega},\mathcal{I}^+) &= \beta_{\Delta,d}\int d\Omega'\langle\mathcal{O}_-(\Omega)\mathcal{O}_-(\Omega')\rangle\mathcal{O}_+(\Omega',\mathcal{I}^-)
\end{aligned}} \tag{2.43}$$

with α and β being some coefficients that depend on d and Δ, where in global coordinates the CFT two–point function with dimension Δ is proportional to $\left(\frac{1}{\sin^2\left(\frac{\Omega-\Omega'}{2}\right)}\right)^{\Delta}$

The operator relations above between the CFT's at $\mathcal{I}^\pm$ are the dictionary relating two equivalent holographic descriptions of the same bulk. With the state–operator map one can regard them as the transformations relating different bases for a quantum state. Here the collection of operators at either copy of the CFT are a complete description for de Sitter space in its Euclidean vacuum state. The total Hilbert space for the two CFTs on both boundaries forms a redundant description, with the constraints above necessary to relate half of the space to the other half.

An interesting way of looking at these operator relations is provided by techniques developed for calculating conformal blocks, in which the relations in Eq. (2.42) are actually the definition of shadow operators[6] [38]. For a given primary operator $\mathcal{O}(x)$ with dimension Δ in a CFT, its shadow operator $\tilde{O}(x)$ is a non–local operator with dimension $d-\Delta$. To get the part involving a certain primary operator $\mathcal{O}(x)$ in the conformal block decomposition of a CFT four–point function $\langle\varphi_1(x_1)\varphi_2(x_2)\varphi_3(x_3)\varphi_4(x_4)\rangle$, one can insert a projection operator with dimension zero which is defined by both $\mathcal{O}$ and its shadow $\tilde{O}$ [38][39]. Then

$$\int d^d x \langle\varphi_1(x_1)\varphi_2(x_2)\mathcal{O}(x)\rangle\langle\tilde{O}(x)\varphi_3(x_3)\varphi_4(x_4)\rangle \tag{2.44}$$

gives the conformal block for exchanging the $\mathcal{O}$ operator after projecting out the shadow blocks. In a CFT, one can always construct the shadow for a local primary operator. The shadow operators are non–local operators in the CFT, but they transform like local primary operators under conformal transformations. An explicit relation between a primary operator and its shadow is given in [40]:

$$\tilde{O}(x) = \int d^d y \frac{1}{(x-y)^{2(d-\Delta)}}\mathcal{O}(y) \propto \int d^d y D_{d-\Delta}(x-y)\mathcal{O}(y) \tag{2.45}$$

here $D_{d-\Delta}(x-y)$ is the two–point function of a primary operator with dimension $d-\Delta$.

One can immediately notice that this is a generalization of Eq. (2.43), and Eq. (2.43) gives a physical interpretation of shadow operators in the special case of Euclidean CFTs on spheres. The shadow for an operator $\mathcal{O}_+$ defined at one of the boundaries of dS, is a local operator $\mathcal{O}_-$ in the CFT defined at the other boundary, with an antipodal map on the sphere. Thus instead of phrasing the construction of de Sitter local operators in terms of two sets of CFT local operators defined at the same boundary, we can also phrase it as a construction with a single copy of operators defined at both boundaries—we use the operator $\mathcal{O}$ defined on one of the boundaries and use its shadow $\tilde{O}$ defined on the other.

[6] I thank Daliang Li for pointing this out.

This also ensures that we are always able to come up with the operators required for the construction: we are always able to construct the shadow operator $\tilde{O}$ from an operator $\mathcal{O}$. Even in the flat slicing, with only a single boundary, the shadow operator for an operator $\mathcal{O}_+$ in the same CFT fits in the properties we need for the corresponding $\mathcal{O}_-$. Thus in the flat patch, the construction can be made by a local operator $\mathcal{O}_+$ and its non–local shadow $\mathcal{O}_- = \int D_-\mathcal{O}_+$. Schematically we have

$$\begin{aligned}\Phi(\eta, \mathbf{x}) =& A_{d,\Delta} \int d^3x' K_+ \left(\eta, \mathbf{x} = 0|\mathbf{x}'\right) \mathcal{O}_+(\mathbf{x} + \mathbf{x}') + \\ & B_{d,\Delta} \int d^3x' d^3y K_- \left(\eta, \mathbf{x} = 0|\mathbf{x}'\right) D_- \left(\mathbf{x} + \mathbf{x}' - \mathbf{y}\right) \mathcal{O}_+(\mathbf{y})\end{aligned} \tag{2.46}$$

The second part is from the shadow operator. After integrating over $\mathbf{x}'$ inside the lightcone, we have a contribution proportional to

$$\int d^d y \left(\frac{\eta^2 - \mathbf{y}^2}{\eta}\right)^{\Delta - d} \mathcal{O}_+(\mathbf{x} + \mathbf{y}) \tag{2.47}$$

The integrand is the same as the contribution from the operator $\mathcal{O}_+$, which is the first term in (2.22), but the support for the integration is non–compact. The non–compact support is from the definition of the shadow operator, which is the price for expressing a de Sitter local operator with a single CFT operator. In this representation, a bulk operator which is close to the boundary maps to a highly non–local operator: a mixture of a local boundary operator and its shadow.

2.1.4.3 Comments on Gauge Fields in de Sitter Space

In the discussion above the attention was on scalar operators in de Sitter space. Here we make some comments on the construction for local fields with integer spins. We focus on gauge fields propagating in de Sitter space. In [2] Vasiliev theory [35] in de Sitter space was proposed as a higher spin realization of dS/CFT correspondence, with the dual conformal field theory being an $Sp(N)$ model with anti–commuting scalars. In this theory there are infinitely many higher spin conserved currents that are bilinear operators constructed from the scalar multiplet:

$$\mathcal{O}_{i_1\ldots i_s} = \Omega_{ab}\chi^a\partial_{(i_1}\ldots\partial_{i_s)}\chi^b \tag{2.48}$$

where Ω_{ab} is symplectic tensor.

In [41][42] the holographic constructions for a massless vector field and a graviton field in anti–de Sitter space were established and were generalized to gauge field $\Phi_{M_1\ldots M_s}$ with generic integer spin in [14]. With the choice of holographic gauge,

$$\Phi_{z\ldots z} = \Phi_{\mu_1 z\ldots z} = \cdots = \Phi_{\mu_1\mu_2\ldots\mu_{s-1}z} = 0, \tag{2.49}$$

in the Poincaré patch it was shown that for a generic gauge field with spin $s > 1$,

$$\boxed{\Phi_{\mu_1\ldots\mu_s} = \frac{\Gamma\left(s+\frac{d}{2}-1\right)}{\pi^{\frac{d}{2}}\Gamma\left(s-1\right)}\frac{1}{z^s}\int_{t'^2+|\mathbf{y}'|^2<z^2} dt'd^{d-1}y'\left(\frac{z^2-t'^2-|\mathbf{y}'|^2}{z}\right)^{s-2}\mathcal{O}_{\mu_1\ldots\mu_s}(t+t',\mathbf{x}+i\mathbf{y}')} \tag{2.50}$$

where the μ_i are d–dimensional indices. The operator $\mathcal{O}_{\mu_1\ldots\mu_s}$ is a symmetric traceless conserved current on the AdS boundary. In $d-$dimensional CFT such an operator has dimension $\Delta = s+d-2$. Thus the twist $\Delta - d$ is always $s-2$, as indicated by the smearing function above.

It is shown in [14] that it is very convenient to convert $\Phi_{\mu_1\ldots\mu_s}$ to a scalar muitiplet with vierbeins in AdS

$$e_a{}^{\mu} = z\delta_a{}^{\mu} \tag{2.51}$$

Let us define

$$Y_{a_1\dots a_s} \equiv e_{a_1}{}^{\mu_1}\dots e_{a_s}{}^{\mu_s}\Phi_{\mu_1\dots\mu_s} = z^s\Phi_{a_1\dots a_s} \tag{2.52}$$

Here $\Phi_{a_1\dots a_s}$ is written in the sense of components; Φ itself is still defined as a tensor under diffeomorphism.

One can show that $Y_{a_1\dots a_s}$ obeys a free scalar equation in AdS with mass parameter and scaling dimension

$$\begin{aligned} m^2R^2_{AdS} &= (s-2)(s+d-2) \\ \Delta &= s+d-2 \end{aligned} \tag{2.53}$$

and therefore near the boundary

$$Y \to z^{\Delta}\mathcal{O} \tag{2.54}$$

The near boundary behavior of the gauge field is given by:

$$\Phi_{\mu_1\dots\mu_s} = \frac{1}{z^s}Y_{\mu_1\dots\mu_s} \to z^{\Delta-s}\mathcal{O}_{\mu_1\dots\mu_s} = z^{d-2}\mathcal{O}_{\mu_1\dots\mu_s} \tag{2.55}$$

Therefore we are able to relate a spin$-s$ bulk gauge field propagating in AdS_{d+1} with a scalar with mass parameter above the Breitenlohner–Freedman bound.

Now we look at the case of gauge fields in de Sitter space. The Poincaré patch of AdS can be analytically continued to the flat patch of dS with double analytic continuation:

$$\begin{aligned} &R^2_{AdS} \to -R^2_{dS} \\ &z \to \eta \\ &x^i_{AdS} \to ix^i_{ds} \end{aligned} \tag{2.56}$$

With the analytic continuation the mass parameter in de Sitter space turns into

$$m^2 R_{dS}^2 = -(s-2)(s+d-2) \tag{2.57}$$

The map between the scaling dimension and the mass parameter in de Sitter space is

$$\Delta = \frac{d}{2} + \sqrt{\frac{d^2}{4} - m^2 R_{dS}^2} \tag{2.58}$$

and thus gives real dimensions

$$\Delta = s + d - 2 \tag{2.59}$$

We see that for scalars the scaling dimensions can in general be imaginary, but that for conserved currents, the dimensions are still real integers.

Therefore for a spin$-s > 1$ gauge field in de Sitter space, the construction is equivalent to the construction for a massless scalar ($s = 2$) or tachyons ($s > 2$). For spin-1, the mass of the scalar is $m^2 = d - 1$ which is positive, but in general for a $d-$dimensional space[7], $d-1 < \frac{d^2}{4}$ and thus does not satisfy $m^2 > \left(\frac{d}{2}\right)^2$. Therefore we see that a local observable for a gauge field in de Sitter space behaves quite differently from the heavy scalar operators we have constructed. These gauge field operators have real dimensions and when approaching the boundaries, components with $(\sigma\eta')^{\Delta}$ and $(\sigma\eta')^{d-\Delta}$ fall at different rates, and have no oscillatory behaviors.

One can see this difference explicitly when directly extending the construction for the scalar field to gauge fields. One might naively expect that for gauge fields the construction involves two sets of single$-$trace operators with dimensions Δ and $d - \Delta$ and gives the

[7] With the exception of $d = 2$, which satuates the bound, which suggests that maybe the bulk Chern-Simons field has something special in this story.

construction equation:

$$\Phi_{i_1\ldots i_s} = \frac{\Gamma\left(s+\frac{d}{2}-1\right)}{\pi^{\frac{d}{2}}\Gamma\left(s-1\right)}\frac{1}{\eta^s}\int_{\mathbf{x}'^2<\eta^2} d^dx'\left(\frac{\eta^2-\mathbf{x}'^2}{\eta}\right)^{s-2}\mathcal{O}^{(+)}_{i_1\ldots i_s}(\mathbf{x}+\mathbf{x}')$$
$$+\frac{\Gamma(3-s-\frac{d}{2})}{\pi^{\frac{d}{2}}\Gamma(3-s-d)}\frac{1}{\eta^s}\int_{\mathbf{x}'^2<\eta^2} d^dx'\left(\frac{\eta^2-\mathbf{x}'^2}{\eta}\right)^{2-s-d}\mathcal{O}^{(-)}_{i_1\ldots i_s}(\mathbf{x}+\mathbf{x}')$$

which is obtained by simply substituting the dimension $\Delta = s + d - 2$ into the scalar expression and identifying $Y_{i_1\ldots i_s} = \eta^s\Phi_{i_1\ldots i_s}$as a bulk scalar.

One would then notice that this proposed solution has problems: The first term has a diverging denominator when $s = 1$ and the second term has a diverging denominator when $d + s$ is an integer larger than two. Therefore, for all the cases of interest, s and d taking values on positive integers, the construction equation above is not well–defined.

The root of the problem is the starting point of the construction for these fields. They are constructed by demanding $\eta^s\Phi_{i_1\ldots i_s}$ as scalars recover the Wightman function for the Euclidean vacuum, Eq (2.23). Here in general the $\eta^s\Phi_{i_1\ldots i_s}$ are scalars with mass parameters that go below $\left(\frac{d}{2}\right)^2$. Also, for the case of interest s and d are integers and so the Wightman function is ill–defined for $s > 1$ due to the factor involving $\Gamma(1-\Delta = 3-s-d)$, and for spin $s = 1$ it doesn't exhibit the nice property of spliting into two parts with fall–off behaviors η^Δ and $\eta^{d-\Delta}$ because the hypergeometric function behaves in a different way when its arguments are integers.

Interestingly enough, these operators are exactly the ones relevant for the proposal of duality between the $Sp(N)$ model in 3 dimensions and dS_4 [2]. There we have currents $J_{i_1\ldots i_s} = \Omega_{ab}\chi^a\partial_{(i_1}\ldots\partial_{i_s)}\chi^b$ with dimension $s + 1$, which corresponds to mass parameter $m^2 = (2-s)(s + 1)$. Due to these values for the mass parameter, for these fields the approach fails and we no longer have a nice picture of a bulk operator being constructed from a pair of CFT operators, and recovering bulk Wightman functions in a de Sitter invariant vacuum.

Despite the remarks above, we can still get a bulk operator that is dual to a boundary current, in the following sense. Given a boundary spin$-s$ current there is an operator in the bulk that matches it smoothly when approaching the boundary. To get such an operator, one can just keep the first part of the expression which approaches the boundary higher spin current:

$$\boxed{\Phi^{(+)}_{i_1\ldots i_s} = \frac{\Gamma\left(s+\frac{d}{2}-1\right)}{\pi^{\frac{d}{2}}\Gamma\left(s-1\right)}\frac{1}{\eta^s}\int_{\mathbf{x}'^2<\eta^2} d^d x'\left(\frac{\eta^2-\mathbf{x}'^2}{\eta}\right)^{s-2}\mathcal{O}^{(+)}_{i_1\ldots i_s}(\mathbf{x}+\mathbf{x}')} \tag{2.60}$$

Here we still have the singularity from the coefficient for a vector field $s = 1$. A more careful treatment following [42][14] gives the expression

$$\boxed{A^{(+)}_i = \frac{1}{V(S^{d-1})}\frac{1}{\eta}\int_{|\mathbf{x}'|=\eta} d^{d-1}x'\mathcal{J}^{(+)}_i\left(\mathbf{x}+\mathbf{x}'\right)} \tag{2.61}$$

where A_i is the bulk vector field and $\mathcal{J}_i$ is a boundary conserved current.

Here the integration is over the intersection of the bulk lightcone and the boundary which is a sphere, and $V\left(S^{d-1}\right) = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}$ is the surface area of a $(d-1)$–sphere. This is just the analytically continued version of the construction in anti–de Sitter space. Here we are actually imposing Dirichlet type boundary condition at $\mathcal{I}^{\pm}$ by demanding that $\Phi_{\mu_1\ldots\mu_s} \to \eta^{\Delta-s}\mathcal{O}_{\mu_1\ldots\mu_s}$. As has been discussed in [24], such Dirichlet boundary conditions are acausal—they force the radiation that hits the future boundary to reflect back into the past. As was discussed in the section 2.1.3, the boundary conditions kill the bulk positive or negative mode, and thus spoil microcausality. The correlation functions computed with such operators are the ones we can obtain by analytically continuing the AdS correlation functions, and thus do not correspond to any de Sitter invariant vacuum. They are not the operators for computing the correlation functions if one would like to look at the cosmology in de Sitter space.

Looking back at gauge fields in AdS space, one notices that the situation is much simpler: the mass parameter for the scalar which corresponds to a spin$-s$ gauge field is always within the Breitenlohner–Freedman bound $m^2 > -\frac{d^2}{4}$. The constructions can be carried out in a standard way [14].

2.1.4.4 Implementation in the Embedding Formalism

We can ask if it is possible to recast the constructions above in the language of the embedding formalism, which was developed in [43][44][45], and recently generalized for superconformal field theories in [46][47]. In this language, the conformal group in d dimensions, which is $SO(d, 2)$, is realized as the Lorentz group in a $d + 2$ dimensional Minkowski space with two time directions. The transformations on the $d+2$ coordinates are linear, and the conformally invariant quantities in d dimensions can be built as Lorentz invariant quantities in $d + 2$ dimensional embedding space, which shows the conformal invariance manifestly.

The set–up of the embedding formalism starts with a $(d + 2)$-dimensional Minkowski space with two time–like directions:

$$ds^2 = \eta_{IJ} dX^I dX^J = -dX^+ dX^- + \eta_{\mu\nu} dX^\mu dX^\nu \tag{2.62}$$

where the indices I, J run over $d+2$ coordinates and μ, ν run over d coordinates including one of the timelike directions.

The anti–de Sitter space and de Sitter space are realized as the hypersurfaces defined by:

$$
\begin{aligned}
X_{AdS} \cdot X_{AdS} &= -R_{AdS}^2 \\
X_{dS} \cdot X_{dS} &= R_{dS}^2
\end{aligned} \tag{2.63}
$$

At large X both the hypersurfaces for de Sitter and for anti–de Sitter space approach the $(d+2)$–dimensional embedding space lightcone, which is

$$
X \cdot X = 0 \tag{2.64}
$$

We can define a d–dimensional Minkowski space by turning the $(d+2)$–dimensional embedding space lightcone into a projective space, denoting the points on the embedding space lightcone as P^I. We then demand that P^I satisfy:

$$
\begin{aligned}
&P \cdot P = 0 \\
&P^I \sim \lambda P^I
\end{aligned} \tag{2.65}
$$

Here we identify the points on the embedding space lightcone that are on the same ray from the origin, thus forming a d dimensional space.

One can parametrize the embedding space in the following ways. To recover the d–dimensional Minkowski space, we define coordinates on the projective lightcone as

$$
P^I = \left(1, y^2, y^\mu\right) \tag{2.66}
$$

where y^μ are d–dimensional coordinates, and y^2 here denotes $y^\mu y_\mu$. The distance between two points on the projective lightcone is then $-2P_1 \cdot P_2 = (y_1 - y_2)^2$. We see that it recovers the distance between two points in the Minkowski spacetime.

For AdS_{d+1} one can define:

$$X^I_{AdS} = \frac{1}{z}\left(1, z^2 + x^2, x^\mu\right) \tag{2.67}$$

We then have the distance between two points in AdS space as

$$-2X_{AdS,1} \cdot X_{AdS,2} = \frac{z_1^2 + z_2^2 + (x_1 - x_2)^2}{z_1 z_2} \tag{2.68}$$

which is proportional to the distance σ we used throughout the section.

Also the regularized distance between a boundary and a bulk point is

$$-2P \cdot X = \frac{z^2 + (x - y)^2}{z} \tag{2.69}$$

Now we see that in anti–de Sitter space, with the embedding coordinates, we can write down the smearing function in a very simple and manifestly AdS–invariant way:

$$\boxed{\Phi(X) = A_{\Delta,d} \int_{\partial AdS} dP \left(-2P \cdot X\right)^{\Delta - d} \Theta\left(-P \cdot X\right) \mathcal{O}(P)} \tag{2.70}$$

Here we integrate over the boundary points denoted by P to get a bulk operator sitting at point X in the embedding coordinates. The domain of integration is over the region with $P \cdot X < 0$, where the boundary points are spacelike separated from the bulk point in the z direction.

In the construction of causal three–point functions in AdS [29], there is an AdS–invariant cross–ratio which is particularly interesting:

$$\chi\left(z, x; x_1; x_2\right) = \frac{\left(z^2 + (x - x_1)^2\right)\left(z^2 + (x - x_2)^2\right)}{z^2 \left(x_2 - x_1\right)^2} \tag{2.71}$$

This cross–ratio is an AdS–invariant quantity built from a single bulk point and two boundary points. It turns out that in AdS space the towers of multi–trace operators to be added into the smearing prescription for recovering bulk microcausality at the level of N^{-1} are organized by powers of this cross–ratio [29]. In the embedding formalism, this cross–ratio is simple:

$$\chi\left(z,x;x_1;x_2\right)=4\frac{\left(P_1\cdot X\right)\left(P_2\cdot X\right)}{P_1\cdot P_2} \tag{2.72}$$

where X denotes the bulk point and P_i the boundary points.

We can also describe the smearing prescription in de Sitter space with the embedding formalism, starting from embedding de Sitter space into the higher dimensional Minkowski space:

$$X^I_{dS}=\frac{1}{\eta}\left(1,-\eta^2+\mathbf{x}^2,x^i\right) \tag{2.73}$$

We then have the distance between two points in de Sitter space:

$$-2X_1\cdot X_2=\frac{-\eta_1^2-\eta_2^2+\left(\mathbf{x}_1-\mathbf{x}_2\right)^2}{\eta_1\eta_2} \tag{2.74}$$

and the regularized distance between a bulk point and a boundary point is:

$$-2P\cdot X=\frac{-\eta^2+\left(\mathbf{x}-\mathbf{y}\right)^2}{\eta} \tag{2.75}$$

Therefore for a scalar with $m^2>\left(\frac{d}{2}\right)^2$ we have the smearing prescription in the embedding space:

$$\boxed{\Phi(X)=A_{\Delta,d}\int_{\partial dS}dP\left(2P\cdot X\right)^{\Delta-d}\Theta\left(P\cdot X\right)\mathcal{O}_+(P)+B_{\Delta,d}\int_{\partial dS}dP\left(2P\cdot X\right)^{-\Delta}\Theta\left(P\cdot X\right)\mathcal{O}_-(P)} \tag{2.76}$$

The dS–invariant cross–ratio

$$\chi(\eta, \mathbf{x}; \mathbf{x}_1; \mathbf{x}_2) = 4\frac{(P_1 \cdot X)(P_2 \cdot X)}{P_1 \cdot P_2} = \frac{(-\eta^2 + (\mathbf{x} - \mathbf{x}_1)^2)(-\eta^2 + (\mathbf{x} - \mathbf{x}_2)^2)}{\eta^2(\mathbf{x}_2 - \mathbf{x}_1)^2} \tag{2.77}$$

could be useful when one considers microcausality for three–point functions in de Sitter space.

It could be interesting to perform the construction for gauge fields in the embedding space, which could potentially make the AdS invariance manifest in the construction. In [42][14] the construction is done in AdS space by imposing the holographic gauge. The construction is not done in a manifestly AdS covariant way, therefore one has to check the AdS covariance of the constructions afterwards. The embedding formalism could be helpful in that direction. Here we didn't derive the construction starting from the embedding formalism, but rather just wrote down the final results in the embedding space. It also could be interesting if we can solve the Cauchy problem or sum over the modes starting with the embedding formalism and derive the equations above.

2.2 Holographic Representation of Higher Spin Gauge Fields

In this section we switch to the topic of how to build a local field operator with integer spin in anti de Sitter space. All the discussions are about theories in AdS space with a general number of dimensions.

2.2.1 Introduction

It is well-known through the AdS/CFT correspondence [15] that a strongly-coupled conformal field theory with a large number of degrees of freedom in d dimensions is dual to a semiclassical gravity theory in $d+1$ dimensions which is a local field theory. In [19] and [20], explicit operator dictionaries were constructed that relate near-boundary bulk fields to operators in the boundary CFT. To probe deeper into anti-de Sitter space, one has to identify the local operators deep in AdS with non-local operators in the boundary CFT. For scalar field this is done scalar field at leading order in large-N expansion in [21][56][57], and further refined in [28]. The construction is carried out to order $\frac{1}{N}$ with interactions in [29].

In [41] and [42], local operators with integer spins, especially gauge fields with zero mass such as a bulk photon field and a graviton field were constructed. The construction was shown to be AdS covariant, with two-point functions between bulk gauge fields and boundary currents having bulk and boundary light-cone singularities. Causality was shown to be respected by gauge-invariant operators such as the electro-magnetic field strength and the Weyl tensor.

In this section, we extend this construction to higher spin gauge fields with $s > 2$. Higher spin gauge field have recently attracted attention since a certain consistent higher spin gauge theory with interactions in AdS space [35] is conjectured [58] to be dual to a free SO(N) vector model. Furthermore, the analytic continuation of this duality was proposed as a realization of the idea of dS/CFT in [1], [2].

To construct gauge fields with higher spin, we work in holographic gauge, in which the calculation is simplified. We show that as in the cases for massless spin-1 and spin-2 fields, one can construct local spin-s field in AdS bulk as a non-local operator smeared over certain region on the boundary; for $s > 1$ the smearing function has support inside the bulk lightcone while for $s = 1$ the support is on the intersection between the bulk lightcone and the boundary (i.e. over a boundary ball and a spherical region respectively). Applying AdS isometries on the bulk operators generically brings the field out of the holographic gauge, but it is shown that one can always do a gauge transformation to bring the field back to holographic gauge, thus establishing the AdS covariance of the construction. Two-point functions of higher-spin fields and currents are calculated and shown to possess the singularity structure compatible with microcausality.

2.2.2 Holographic Representation of Scalar, Vector and Tensor Fields

In this subsection we briefly review the construction of local spin-0 and massless spin-1 and spin-2 fields in AdS_{d+1}/CFT_d. We work in the Poincaré patch with metric (taking the AdS radius R_{AdS} to be 1)

$$ds^2 = \frac{1}{z^2}\left(-dt^2 + dz^2 + d\mathbf{x}^2\right) \tag{2.78}$$

A bulk scalar with dimension $\Delta = \frac{d}{2} + \sqrt{m^2 + \left(\frac{d}{2}\right)^2} \equiv \nu + \frac{d}{2}$ can be constructed by summing over all normalizable modes in the bulk [28]:

$$\Phi\left(t, z, \mathbf{x}\right) = \int_{|\omega|>|\mathbf{k}|} d\omega d^{d-1}\mathbf{k} a_{\omega\mathbf{k}} e^{-i\omega t} e^{i\mathbf{k}\cdot\mathbf{x}} z^{\frac{d}{2}} J_\nu\left(z\sqrt{\omega^2 - \mathbf{k}^2}\right) \tag{2.79}$$

with the mode related to a boundary local operator by

$$a_{\omega\mathbf{k}} = \frac{2^\nu\Gamma(\nu+1)}{(2\pi)^d\left(\omega^2 - \mathbf{k}^2\right)^{\frac{\nu}{2}}} \int dt d^{d-1}\mathbf{x} e^{i\omega t} e^{-i\mathbf{k}\cdot\mathbf{x}} \mathcal{O}(x) \tag{2.80}$$

Putting $a_{\omega\mathbf{k}}$ into the mode sum one obtains the representation of a local scalar field in the AdS bulk as an integral on the boundary with compact support, by making the boundary coordinates complex.

$$\Phi(t,z,\mathbf{x}) = \frac{\Gamma(\Delta - \frac{d}{2}+1)}{\pi^{\frac{d}{2}}\Gamma(\Delta - d+1)} \int_{t'^2+\mathbf{y}'^2<z^2} dt' d^{d-1}\mathbf{y}' \left(\frac{z^2-t'^2-\mathbf{y}'^2}{z}\right)^{\Delta-d} \mathcal{O}(t+t',\mathbf{x}+i\mathbf{y}') \tag{2.81}$$

For $\Delta > d-1$ the integral is well-defined, but it diverges for a field with $\Delta = d-1$, which is a tachyon with mass $m^2 = 1-d$. The construction in this case is carried out in [42], where it turns out that the integration domain is the sphere S^{d-1} on which the bulk lightcone intersects with the boundary:

$$\Phi(t,z,\mathbf{x}) = \frac{1}{Vol(S^{d-1})} \int_{t'^2+\mathbf{y}'^2=z^2} dt' d^{d-1}\mathbf{y}' \mathcal{O}(t+t',\mathbf{x}+i\mathbf{y}') \tag{2.82}$$

Here $Vol(S^{d-1}) = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}$ is the surface area of a $(d-1)$-sphere.

The construction for scalars is very useful for the construction for gauge fields; the case for gauge fields (in holographic gauge) can always be reduced to scalars with a certain mass and then directly constructed with the help of the results above for scalars.

Let's start with a massless spin-one field. In this case, the source-free bulk Maxwell equation

$$\nabla_M F^M{}_N = 0 \tag{2.83}$$

can be solved in holographic gauge

$$A_z = 0 \tag{2.84}$$

One can further impose that

$$\partial_\mu A^\mu = 0 \tag{2.85}$$

which can be thought of as the conservation of the boundary current $\partial_\mu A^\mu \sim \partial_\mu j^\mu = 0$.

Here the Greek index μ runs from 0 to $d-1$ of the boundary coordinates, while the capital letters M and N run from 0 to d, including the radial or holographic coordinate z. The bulk equation for A_μ becomes

$$\partial_\nu \partial^\nu A_\mu + z^{d-3} \partial_z \left(z^{3-d} \partial_z A_\mu \right) = 0 \tag{2.86}$$

If we define

$$\Phi_\mu \equiv z A_\mu \tag{2.87}$$

the equation for A_μ can be written as the free wave equation for a multiplet of scalars with mass squared $m^2 = 1 - d$

$$\partial_\alpha \partial^\alpha \Phi_\mu + z^{d-1} \partial_z \left(z^{1-d} \partial_z \Phi_\mu \right) + \frac{d-1}{z^2} \Phi_\mu = 0 \tag{2.88}$$

Thus the gauge field A_μ can be constructed via the construction of scalars Φ_μ:

$$A_\mu(t, z, \mathbf{x}) = \frac{1}{z} \Phi_\mu(t, z, \mathbf{x}) = \frac{1}{Vol(S^{d-1})} \frac{1}{z} \int_{t'^2 + \mathbf{y}'^2 = z^2} dt' d^{d-1} \mathbf{y}' j_\mu(t + t', \mathbf{x} + i\mathbf{y}') \tag{2.89}$$

The construction for the graviton is similar. Working in holographic gauge

$$h_{zz} = h_{z\mu} = 0 \tag{2.90}$$

one solves the bulk equation for linearized graviton propagation in AdS space

$$\nabla_Q \nabla^Q h_{MN} - 2 \nabla_Q \nabla_M h^Q_{\ N} + \nabla_M \nabla_N h^Q_{\ Q} - 2d h_{MN} = 0 \tag{2.91}$$

One can further impose the conditions

$$h^{\alpha}{}_{\alpha} = 0 \ , \ \partial_{\mu} h^{\mu\nu} = 0 \tag{2.92}$$

because of the conservation of the boundary currents

$$\partial_{\mu} T^{\mu\nu} = 0 \tag{2.93}$$

$$\partial_{\mu} \left(x_{\nu} T^{\mu\nu} \right) = T^{\nu}_{\ \nu} = 0 \tag{2.94}$$

The $\mu\nu$ component of the bulk equation then gives

$$\partial_{\alpha}\partial^{\alpha} h_{\mu\nu} + \partial_z^2 h_{\mu\nu} + \frac{5-d}{z}\partial_z h_{\mu\nu} - \frac{2(d-2)}{z^2} h_{\mu\nu} = 0 \tag{2.95}$$

Defining a multiplet of scalars

$$\Phi_{\mu\nu} \equiv z^2 h_{\mu\nu} \tag{2.96}$$

we have the equation for $\Phi_{\mu\nu}$ as an equation for massless scalars

$$\partial_{\alpha}\partial^{\alpha}\Phi_{\mu\nu} + z^{d-1}\partial_z \left(z^{1-d}\partial_z \Phi_{\mu\nu} \right) = 0 \tag{2.97}$$

Therefore the bulk graviton can be represented as

$$h_{\mu\nu}(t,z,\mathbf{x}) = \frac{\Gamma(\Delta - \frac{d}{2} + 1)}{\pi^{\frac{d}{2}}\Gamma(\Delta - d + 1)} \frac{1}{z^2} \int_{t'^2 + \mathbf{y}'^2 < z^2} dt' d^{d-1}\mathbf{y}' \left(\frac{z^2 - t'^2 - \mathbf{y}'^2}{z} \right)^{\Delta - d} T_{\mu\nu}(t+t', \mathbf{x}+i\mathbf{y}') \tag{2.98}$$

2.2.3 Holographic Representation of Massless Spin-s field

In this section we carry out the construction for a general integer-spin gauge field in AdS_{d+1} in terms of smeared local operators in the field theory.

As is well known, a massless gauge field in AdS is represented by a totally symmetric rank-s tensor $\Phi_{M_1\ldots M_s}$ satisfying double-tracelessness conditions,

$$\Phi^{MN}{}_{MNM_5\ldots M_s} = 0 \tag{2.99}$$

The linearized equation for a spin-s gauge field on AdS$_{d+1}$ is [62], [63], [35]

$$\nabla_N\nabla^N\Phi_{M_1\ldots M_s} - s\nabla_N\nabla_{M_1}\Phi^N{}_{M_2\ldots M_s} + \frac{1}{2}s(s-1)\nabla_{M_1}\nabla_{M_2}\Phi^N{}_{N\ldots M_s} - 2(s-1)(s+d-2)\Phi_{M_1\ldots M_s} = 0 \tag{2.100}$$

This equation is invariant under the gauge transformation

$$\begin{aligned} &\Phi_{M_1\ldots M_s} \to \Phi_{M_1\ldots M_s} + \nabla_{M_1}\Lambda_{M_2\ldots M_s} \\ &\Lambda^N{}_{NM_3\ldots M_s} = 0 \end{aligned} \tag{2.101}$$

We would like to work in holographic gauge, in which the z-components of the gauge field all vanish, so we shall generalize the holographic gauge from a vector and a rank-two tensor to a rank-s tensor

$$\Phi_{z\ldots z} = \Phi_{\mu_1 z\ldots z} = \cdots = \Phi_{\mu_1\ldots\mu_{s-1}z} = 0 \tag{2.102}$$

One can always make this gauge choice because given a general field $\Phi^{(0)}_{M_1\ldots M_s}$, one can perform a gauge transformation $\Phi \to \Phi + \nabla\Lambda$ so that

$$\Phi^{(0)}_{z\dots z} + \nabla_z \Lambda_{z\dots z} = 0 \tag{2.103}$$

$$\Phi^{(0)}_{\mu_1 z\dots z} + \nabla_z \Lambda_{\mu_1 z\dots z} = 0 \tag{2.104}$$

$$\dots \tag{2.105}$$

$$\Phi^{(0)}_{\mu_1\dots\mu_{s-1}z} + \nabla_z \Lambda_{\mu_1\dots\mu_{s-1}} = 0 \tag{2.106}$$

The number of gauge parameters is just right to satisfy the set of equations and fix the holographic gauge for generic spin- s gauge field.

The bulk gauge field is dual to a totally symmetric, traceless, conserved rank-s tensor on the boundary:

$$\begin{aligned} \mathcal{O}^{\nu}{}_{\nu\mu_3\dots\mu_s} &= 0 \\ \partial_\nu \mathcal{O}^{\nu}{}_{\mu_2\dots\mu_s} &= 0 \end{aligned} \tag{2.107}$$

In holographic gauge, only three types of components of the bulk equation are non-trivial, they are

$zz\mu_3\dots\mu_s$:

$$\begin{aligned} &\frac{s(s-1)}{2}\partial_z^2 \Phi^{\alpha}{}_{\alpha\mu_3\dots\mu_s} + \left(\frac{s(s-1)(2s-3)}{2} - s\right)\frac{1}{z}\partial_z \Phi^{\alpha}{}_{\alpha\mu_3\dots\mu_s} + \\ &\left(2 - s(s-1) + \frac{s!}{2(s-4)!}\right)\frac{1}{z^2}\Phi^{\alpha}{}_{\alpha\mu_3\dots\mu_s} = 0 \end{aligned} \tag{2.108}$$

$z\mu_2\dots\mu_s$:

$$(2 - s(s-1))\frac{1}{z}\partial_\alpha \Phi^{\alpha}{}_{\mu_2\dots\mu_s} - s\partial_z\partial_\alpha \Phi^{\alpha}{}_{\mu_2\dots\mu_s} + \frac{s(s-1)}{2}\left(\partial_z\partial_{\mu_1}\Phi^{\alpha}{}_{\alpha\mu_2\dots\mu_s} + \frac{s-1}{z}\partial_{\mu_1}\Phi^{\alpha}{}_{\alpha\mu_2\dots\mu_s}\right) = 0 \tag{2.109}$$

$\mu_1 \dots \mu_s$:

$$z^2\left(\partial_z^2+\partial_\alpha\partial^\alpha\right)\Phi_{\mu_1\dots\mu_s}+(2s+1-d)z\partial_z\Phi_{\mu_1\dots\mu_s}+2(s-1)(2-d)\Phi_{\mu_1\dots\mu_s}-s\partial_{\mu_1}\partial_\alpha\Phi^\alpha{}_{\mu_2\dots\mu_s}+ \tag{2.110}$$

$$\frac{1}{2}s(s-1)\left(\partial_{\mu_1}\partial_{\mu_2}\Phi^\alpha{}_{\alpha\mu_3\dots\mu_s}-g_{\mu_1\mu_2}z\partial_z\Phi^\alpha{}_{\alpha\mu_3\dots\mu_s}-(s-2)g_{\mu_1\mu_2}\Phi^\alpha{}_{\alpha\mu_3\dots\mu_s}-\sum_{i=3}^{s}g_{\mu_1\mu_i}\Phi^\alpha{}_{\alpha\mu_2\dots\mu_{i-1}\mu_{i+1}\dots\mu_s}\right) \tag{2.111}$$

$$=0 \tag{2.112}$$

the components with more than two z's vanish trivially.

Since the boundary currents are conserved and traceless, we can consistently set

$$\Phi^\alpha{}_{\alpha\mu_3\dots\mu_s}=0 \tag{2.113}$$

$$\partial_\alpha\Phi^\alpha{}_{\mu_2\dots\mu_s}=0 \tag{2.114}$$

and we have the $\mu_1 \dots \mu_s$ component of the equation as

$$\left(\partial_z^2+\partial_\alpha\partial^\alpha\right)\Phi_{\mu_1\dots\mu_s}+\frac{2s+1-d}{z}\partial_z\Phi_{\mu_1\dots\mu_s}+\frac{2(s-1)(2-d)}{z^2}\Phi_{\mu_1\dots\mu_s}=0 \tag{2.115}$$

To turn the problem of constructing a spin-s gauge field into constructing scalars, we define[8]

$$Y_{\mu_1\dots\mu_s}=z^s\Phi_{\mu_1\dots\mu_s} \tag{2.116}$$

[8] Just as in the spin-1 and spin-2 cases, this amounts to expressing the higher spin fields in the vielbein basis. $Y_{a_1\dots a_s}=e^{\mu_1}_{a_1}\dots e^{\mu_s}_{a_s}Y_{\mu_1\dots\mu_s}$ with $e^{\mu_i}_{a_i}=\frac{z}{R_{AdS}}\delta^{\mu_i}_{a_i}$.

as a multiplet of scalars. The equation for $Y_{\mu_1\ldots\mu_s}$ is

$$\partial_\alpha\partial^\alpha Y_{\mu_1\ldots\mu_s} + z^{d-1}\partial_z\left(z^{1-d}\partial_z Y_{\mu_1\ldots\mu_s}\right) - \frac{(s-2)(s+d-2)}{z^2}Y_{\mu_1\ldots\mu_s} = 0 \tag{2.117}$$

which is just free scalar equations with mass parameter

$$m^2 = (s-2)(s+d-2) \tag{2.118}$$

corresponding to scaling dimension

$$\Delta = s + d - 2 \tag{2.119}$$

The near-boundary behavior of $Y_{\mu_1\ldots\mu_s}$ is

$$Y_{\mu_1\ldots\mu_s} \sim z^\Delta \mathcal{O}_{\mu_1\ldots\mu_s} \tag{2.120}$$

So one can directly construct the bulk spin-s field:

$$\Phi_{\mu_1\ldots\mu_s} = \frac{\Gamma\left(s+\frac{d}{2}-1\right)}{\pi^{\frac{d}{2}}\Gamma\left(s-1\right)}\frac{1}{z^s}\int_{t'^2+|\mathbf{y}'|^2<z^2} dt' d^{d-1}\mathbf{y}'\left(\frac{z^2-t'^2-|\mathbf{y}'|^2}{z}\right)^{s-2}\mathcal{O}_{\mu_1\ldots\mu_s}(t+t',\mathbf{x}+i\mathbf{y}') \tag{2.121}$$

for a field with integer spin $s > 1$. We see the field behaves like $z^{\Delta-s} = z^{d-2}$ near the boundary. Also, it gives rise to the expected scaling dimension for the conserved boundary currents.

2.2.4 AdS Covariance

In this section we check the covariance of the construction. We apply the anti-de Sitter group transformations to the local operator constructed in last section, and see if it is covariant up to gauge transformations.

The covariance under dilation is pretty straightforward, since both sides of (2.121) have the same scaling dimension and thus rescale in the same way when a dilation is applied. The special conformal transformations are less trivial. The bulk AdS isometries that correspond to special conformal transformations are

$$x^\mu \to x^\mu + 2b \cdot x x^\mu - b^\mu \left(x^2 + z^2\right) \tag{2.122}$$

$$z \to z + 2b \cdot x z \tag{2.123}$$

Also, for a higher spin field we have

$$\Phi_{M_1 \ldots M_s} \to \Phi'_{M_1 \ldots M_s} = \frac{\partial x^{N_1}}{\partial x^{M_1}} \cdots \frac{\partial x^{N_s}}{\partial x^{M_s}} \Phi_{N_1 \ldots N_s} \tag{2.124}$$

which gives the transformation laws for components:

$$\delta\Phi_{zz\ldots z} = \delta\Phi_{zz\ldots\mu_s} = \cdots = \delta\Phi_{zz\mu_3\ldots\mu_s} = 0 \tag{2.125}$$

$$\delta\Phi_{z\mu_2\ldots\mu_s} = 2zb^\alpha\Phi_{\alpha\mu_2\ldots\mu_s} \tag{2.126}$$

$$\delta\Phi_{\mu_1\ldots\mu_s} = 2b^\alpha \sum_{j=1}^{s} x_{\mu_j}\Phi_{\alpha\mu_1\ldots\mu_{j-1}\mu_{j+1}\ldots\mu_s} - 2x^\alpha \sum_{j=1}^{s} b_{\mu_j}\Phi_{\alpha\mu_1\ldots\mu_{j-1}\mu_{j+1}\ldots\mu_s} - 2s\left(b\cdot x\right)\Phi_{\mu_1\ldots\mu_s} \tag{2.127}$$

The transformations bring the gauge field out of holographic gauge, which requires a compensating gauge transformation to recover the holographic gauge. Such a gauge transformation takes the form

$$\delta\Phi_{z\mu_2\ldots\mu_s} = \frac{1}{z^{2s-2}}\partial_z\epsilon_{\mu_2\ldots\mu_s} \tag{2.128}$$

$$\delta\Phi_{\mu_1\ldots\mu_s} = \frac{1}{z^{2s-2}}\sum_{j=1}^{s}\partial_{\mu_j}\epsilon_{\mu_1\ldots\mu_{j-1}\mu_{j+1}\ldots\mu_s} \tag{2.129}$$

with the gauge parameters

$$\epsilon_{\mu_2\ldots\mu_s} = -\frac{\Gamma(s+\frac{d}{2}-1)}{2^{2-s}\pi^{\frac{d}{2}}2\Gamma(s)}\int d^dx'\Theta(\sigma z')(\sigma zz')^{s-1}2b^\alpha\mathcal{O}_{\alpha\mu_2\ldots\mu_s} \tag{2.130}$$

Here we have defined the AdS invariant length

$$\sigma = \frac{z^2+z'^2+(x-x')^2}{2zz'} \tag{2.131}$$

and in the gauge parameters above, $\sigma z'$ and $\sigma zz'$ should be understood as in the limit $z' \to 0$. Then we have

$$\delta\Phi_{z\mu_2\ldots\mu_s} = -2zb^\alpha\Phi_{\alpha\mu_2\ldots\mu_s} \tag{2.132}$$

and

$$\delta\Phi_{\mu_1\ldots\mu_s} = -\frac{\Gamma(s+\frac{d}{2}-1)}{2^{2-s}\pi^{\frac{d}{2}}\Gamma(s-1)}\frac{1}{z^{2s-2}}\int d^dx\Theta(\sigma z')(\sigma zz')^{s-2}\sum_{j=1}^{s}(x-x')_{\mu_j}2b^\alpha\mathcal{O}_{\alpha\mu_1\ldots\mu_{j-1}\mu_{j+1}\ldots\mu_s} \tag{2.133}$$

We see the $\Phi_{z\mu_2\ldots\mu_s}$ components are brought back to zero and holographic gauge restored, while the $\Phi_{\mu_1\ldots\mu_s}$ components get an extra piece from the gauge transformation which combines with the AdS transformation and gives

$$\begin{aligned} z^s\delta\Phi_{\mu_1\ldots\mu_s} =& \frac{\Gamma(s+\frac{d}{2}-1)}{\pi^{\frac{d}{2}}\Gamma(s-1)}\int d^dx\Theta(\sigma z')(2\sigma z')^{s-2}\sum_{j=1}^{s}\Big(2x'_{\mu_j}b^\alpha\mathcal{O}_{\alpha\ldots\mu_s}-2x^\alpha b_{\mu_j}\mathcal{O}_{\alpha\ldots\mu_s}\Big) \\ & -2s\,(b\cdot x)\,z^s\Phi_{\mu_1\ldots\mu_s} \end{aligned} \tag{2.134}$$

This can be further simplified if we consider current conservation, which gives

$$\int d^dx'\Theta(\sigma z')(\sigma zz')^{s-1}\partial^\alpha\Phi_{\alpha\mu_2\ldots\mu_s}=0 \tag{2.135}$$

Integrating by parts we get

$$\int d^dx\Theta(\sigma z')\,(\sigma z')^{s-2}\,(x-x')^\alpha\mathcal{O}_{\alpha\ldots\mu_s}=0 \tag{2.136}$$

This suggests that we can replace the x' with x in (2.134) above:

$$\begin{aligned} z^s\delta\Phi_{\mu_1\ldots\mu_s} &= \frac{\Gamma(s+\frac{d}{2}-1)}{\pi^{\frac{d}{2}}\Gamma(s-1)}\int d^dx\Theta(\sigma z')(2\sigma z')^{s-2}\sum_{j=1}^{s}\big(2x_{\mu_j}b^\alpha\mathcal{O}_{\alpha\ldots\mu_s}-2x^\alpha b_{\mu_j}\mathcal{O}_{\alpha\ldots\mu_s}\big) \\ & -2s\,(b\cdot x)\,z^s\Phi_{\mu_1\ldots\mu_s} \\ &= z^s\sum_{j=1}^{s}\big(2x_{\mu_j}b^\alpha\Phi_{\alpha\ldots\mu_s}-2x^\alpha b_{\mu_j}\Phi_{\alpha\ldots\mu_s}\big)-2sz^s\,(b\cdot x)\,\Phi_{\mu_1\ldots\mu_s} \end{aligned} \tag{2.137}$$

which matches with the transformation of the boundary current under special conformal transformations when one combines with the transformation of the integration measure:

$$\delta\mathcal{O}_{\mu_1\ldots\mu_s}=\sum_{j=1}^{s}\big(2x_{\mu_j}b^\alpha\mathcal{O}_{\alpha\ldots\mu_s}-2x^\alpha b_{\mu_j}\mathcal{O}_{\alpha\ldots\mu_s}\big)-2(d+s-2)(b\cdot x)\mathcal{O}_{\mu_1\ldots\mu_s} \tag{2.138}$$

Chapter 3

Consistency Relations As Ward Identities

In the previous chapter we saw how the degrees of freedom in a CFT, which are organized by conformal symmetries, give rise to degrees of freedom in AdS and dS space which are organized by the isometries of the spacetime. The transition kernels for the constructions act like transformation matrices between different representations — they map representations of the conformal group to representations of the AdS and dS group. Knowing how a local field theory arises in a curved spacetime, in this chapter we further explore the implications of symmetries for local field theories in a cosmology, especially the field theories describing the primordial fluctuations and the large scale structure. We will see that, very similar to the soft pion theorems in strong interaction dynamics, soft mode relations can be derived for cosmology from non-linearly realized conformal symmetries. These consistency relations put special constraints on correlation functions and serve as smoking guns for violations of fundamental principles.

3.1 Symmetries and Ward Identities for Large Scale Structure

3.1.1 Introduction

In this chapter the main topic we will be concerned with is the implications of nonlinearly realized symmetries in cosmology. The physics of nonlinearly realized symmtries and Goldstone bosons has a long history. Before the discovery of QCD, people could already do calculations involving pions, for instance the emission and absorption of soft pions. The guiding principle for such calculations is an approximate $SU(2)\times SU(2)$ chiral symmetry—which is postulated to be the origin of pions and whose soft breaking gives the pion masses. The chiral symmetry dictates the existence of current algebra, with which the amplitude involving arbitrary number of pions can be calculated. Chiral symmetry has another kind of prediction, as shown by the generation of physicists in the 60's, that relates the amplitudes with soft pions to those that without pions. One example is Adler's rule for nucleon amplitudes with a single soft pion.

Pions are pseudo-scalars that carry isospin. They can be generated by the axial current[1]

$$\langle 0|J_A^\mu|\pi^-\rangle \sim p^\mu \frac{F_\pi}{\sqrt{2}} e^{-ip\cdot x} \tag{3.1}$$

Taking the divergence we see that this equation gives:

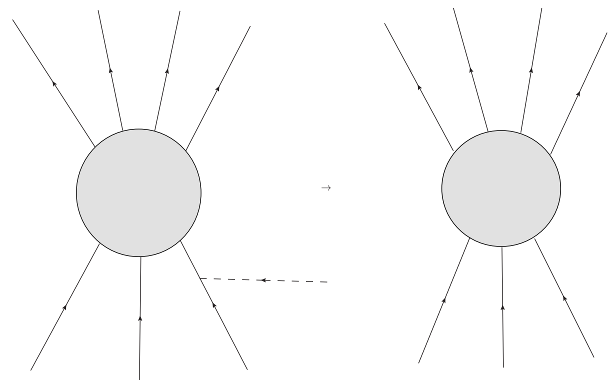
$$\langle 0|\partial_\mu J_A^\mu|\pi^-\rangle \sim m_\pi^2 F_\pi e^{-ip\cdot x} \tag{3.2}$$

Thus with suitable normalization, the operator $\partial_\mu J_A^\mu$ is a good pion field operator.

Now we look at the scattering between nucleons and a pion. Adler's rule relates a

[1] Here the current J_A is defined as $J_{A1} + iJ_{A2}$ so that it creates π^-, here the number denotes SU(2) isospin number

process involving a single soft pion with the process without the soft pion, as depicted in the graph:



which expressed in matrix elements would be

$$\langle \pi(q), f|i\rangle_{q\to 0} \sim \mathcal{M}_{fi} \tag{3.3}$$

where we use i and f to denote the initial and final states of the nuleons.

The argument that leads to the relation is as follows:

Look at the amplitude involving one pion. From the reduction formula, it is proportional to the matrix element

$$\langle f, \pi(q)|i\rangle \sim \langle f|\partial_\mu J^\mu_A|i\rangle = q_\mu \langle f|J^\mu_A|i\rangle \tag{3.4}$$

This might lead one to think that in the soft pion limit, $q \to 0$, the matrix element vanishes. However this is only true if the matrix element $\langle f|J^\mu_A|i\rangle$ is regular in this limit. But indeed it has a pole which cancels the q^μ factor.

The origin of the pole is from the graphs in which the pion (axial current, or soft momentum) is attached to an external nucleon line, as shown in the figure. Between the pion insertion and the blob of hard interactions there is then a section carrying a propagator

$$\frac{1}{(p+q)^2+m^2} \sim \frac{1}{2p\cdot q} \tag{3.5}$$

where p and m are the momentum and the mass of the corresponding external nucleon.

The zero (which is named "Adler's zero") in the denominator cancels with the prefactor and gives a finite result that has nothing to do with the soft pion. Therefore in order to calculate the amplitude with a soft pion, we can just calculate all the diagrams without any pions, and then attach a pion to the external legs.

Adler's rule, as an example of a soft pion theorem in strong interaction dynamics, resulted from the underlying chiral symmetry. In terms of spacetime correlators, we can see it is very similar to a Ward identity:

$$\langle\pi(x)\mathcal{N}(x_1)\cdots\mathcal{N}(x_n)\rangle = \langle\partial J_A(x)\mathcal{N}(x_1)\cdots\mathcal{N}(x_n)\rangle \sim \sum_{a=1}^{n}\delta^4(x-x_a)\langle\mathcal{N}(x_1)\cdots\mathcal{N}(x_n)\rangle \tag{3.6}$$

We will see that nonlinearly realized symmetries have important consequences in the study of cosmological perturbations, and equations similar to Eq. (3.6) can be derived. These constrain the correlation functions of cosmological perturbations.

In section (1.3.2.2) we briefly introduced the relevant physical observables in the study of large scale structure. The velocity field $\vec{v}$ is of special interest to us since it is related to the overdensity field δ at the linear level, and transforms nonlinearly under some of the diffeomorphism transformations:

$$\Delta\vec{v} = \Delta_{\rm lin.}\vec{v} + \Delta_{\rm nl.}\vec{v}\,, \tag{3.7}$$

where $\Delta_{\rm lin.}\vec{v}$ depends linearly on the fluctuating variable $\vec{v}$, while $\Delta_{\rm nl.}\vec{v}$ does not depend on the fluctuating field, or any fluctuating field. An example is a *time-dependent* spatial translation:

$$\vec{x} \to \vec{x} + \vec{n}(\eta)\,, \tag{3.8}$$

where $\vec{n}$ depends on (conformal) time η but not space. Under such a transformation, in addition to the usual linear transformation $\Delta_{\rm lin.}\vec{v} = -\vec{n}\cdot\vec{\nabla}\vec{v}$, the velocity experiences a nonlinear shift:

$$\Delta_{\rm nl.}\vec{v} = \vec{n}'\,, \tag{3.9}$$

where $'$ denotes a time derivative. It was pointed out recently by Kehagias & Riotto and Peloso & Pietroni [77, 78] (KRPP) that just such a time-dependent spatial translation is in fact a symmetry of the familiar system of equations for a pressureless fluid coupled to gravity in the Newtonian limit: the continuity, Euler and Poisson equations. As pointed out by the same authors, the consequence of such a nonlinearly realized symmetry is *not* the simple invariance of a general correlation function, but rather a relation between an $(N{+}1)$-point function and an N-point function:

$$\lim_{\vec{q}\to 0}\frac{\langle \vec{v}(\vec{q})\mathcal{O}_{\vec{k}_1}\mathcal{O}_{\vec{k}_2}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}}{P_v(q)} \sim \langle \mathcal{O}_{\vec{k}_1}\mathcal{O}_{\vec{k}_2}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}\,, \tag{3.10}$$

where P_v is the velocity power spectrum, the superscript c' denotes a *connected* correlation function with the overall delta function removed, and $\mathcal{O}$ denotes some observable, which could be different for each $\vec{k}$.

These *consistency relations*, which relates a squeezed (N+1)-point function to an N-point function, are well known in the context of single field inflation. The first example was pointed out by Maldacena [3] (see also [79]):

$$\lim_{\vec{q}\to 0} \frac{\langle \zeta_{\vec{q}}\zeta_{\vec{k}_1}...\zeta_{\vec{k}_N}\rangle^{c'}}{P_\zeta(q)} = -\left(3(N-1) + \sum_{a=1}^{N} \vec{k}_a \cdot \frac{\partial}{\partial \vec{k}_a}\right) \langle \zeta_{\vec{k}_1}...\zeta_{\vec{k}_N}\rangle^{c'} , \tag{3.11}$$

where ζ is the curvature perturbation. It arises from a spatial dilation symmetry, which is non-linearly realized on ζ. Recently, more non-linearly realized symmetries have been discovered, including the special conformal symmetry [4, 80] and in fact a whole infinite tower of symmetries [5] (H2K hereafter). Recent work has emphasized the non-perturbative nature of these consistency relations as Ward/Slavnov-Taylor identities [5–8, 81–83]. They can be viewed as the cosmological analog of the classic soft-pion theorems introduced above. In the examples above, ζ and the velocity play the role of the pion – both shift nonlinearly under the respective symmetries.

The consistency relations are symmetry statements that are valid beyond the perturbative regime. This is especially interesting for the study of large scale structures since at short scales dark matter forms lumps and galaxies form. Perturbation theory breaks down, while the short wavelength modes in the consistency relations can be highly nonlinear.

In this chapter, we take a close look at the symmetries of the large scale structure, and derive the corresponding consistency relations. The origin of these symmetries is shown to be residual gauge transformations, and we discuss the relations between different symmetries which are established via the adiabatic mode conditions. Starting with the symmetries, we derive an infinite tower of consistency relations constraining the correlation functions involving scalar modes and tensor modes. The robustness and limitations of the consistency relations in LSS are extensively discussed. We examine the underlying assumptions behind the consistency relations by looking into models of galaxy dynamics. A Lagrangian describing large scale structure dynamics is introduced.

Further, we extend our discussion to Lagrangian space, and try to look at the symmetries and consistency relations therein. The consistency relation for time-dependent translation is derived in Lagrangian space, where it takes a very simple form. The equivalence between the Eulerian space consistency relation and the Lagrangian space consistency relation is proved to all orders of perturbation theory.

A few words are in order on our notation and terminology. We use the symbol π to represent the Nambu-Goldstone boson of a non-linearly realized symmetry (the pion), in accordance with standard practice. For our LSS application, π is the velocity potential. The same symbol is also used to denote the numerical value 3.14.... Which is meant should be obvious from the context. Essentially, the numerical π always precedes the Newtonian constant G in the combination $4\pi G$. In cases where both could potentially appear, we use $M_P^2 \equiv 1/(8\pi G)$ to avoid confusion. Also, we use the term *nonlinear* to refer to quantities that are not linear in the LSS observables (fields such as density or velocity). Sometimes, this has the usual meaning that such quantities go like the fields raised to higher powers: quadratic and so on. But, sometimes, this means the quantities of interest do not depend on the field variables at all, such as the nonlinear part of certain symmetry transformations e.g. Eq. (3.9). We rely on the context to differentiate between the two.

3.1.2 Consistency Relation from Time-dependent Translation – a Newtonian Symmetry

In this subsection, we focus on the Newtonian symmetry uncovered by KRPP. Subsection 3.1.3 is a review of the symmetry and its implied consistency relation. In section (3.1.3.1) we discuss the robustness and limitations of the consistency relation, and go over what assumptions can or cannot be relaxed, especially concerning the nonlinear, astrophysically messy, galaxy observables. We discuss what kind of galaxy dynamics, and what sort of galaxy selection, could lead to violations of the consistency relation. As a by-product of our investigation, we describe a simple Lagrangian for LSS in 3.1.3.2.

3.1.3 Time-dependent Translation Symmetry and the Background Wave Argument – a Review

We begin with a review of the Newtonian symmetry discovered by KRPP. We go over the background wave derivation of the consistency relation in some detail, emphasizing the underlying assumptions, and making the derivation easily generalizable to the general relativistic case. Two fundamental concepts are: (1) the existence of a non-linearly realized symmetry (one that shifts at least some of the LSS observables by an amount that is independent of the observables), and (2) an adiabatic mode condition, which is an additional condition that dictates the time-dependence of the symmetry.

Time-dependent Translation Symmetry. The set of Newtonian equations of motion for LSS is: [2]

$$\begin{aligned}
&\delta' + \partial_i[(1+\delta)v^i] = 0 \\
&v^{i\prime} + v^j\partial_j v^i + \mathcal{H}v^i = -\partial_i\Phi \\
&\nabla^2\Phi = 4\pi G\bar{\rho}a^2\delta
\end{aligned} \tag{3.12}$$

where δ is the mass overdensity, $' \equiv \partial/\partial\eta$ denotes the derivative with respect to conformal time η, ∂_i denotes the derivative with respect to the comoving coordinate x^i, v^i is the peculiar velocity $dx^i/d\eta$, Φ is the gravitational potential, G is Newton's constant, $\bar{\rho}$ is the mean mass density at the time of interest, a is the scale factor, and $\mathcal{H} \equiv a'/a$ is the

[2] Throughout this chapter, we will be cavalier about the placement of indices for objects with Latin indices, e.g. v^i, x^i are the same as v_i, x_i.

comoving Hubble parameter. The first equation expresses continuity or mass conservation. The second equation is the Euler equation or momentum conservation for a pressureless fluid. The third equation is the Poisson equation. Let us start with this basic set. We will later consider generalizations to include pressure, relativistic corrections, and even complex galaxy formation processes.

KRPP pointed out that this system of equations admits the following symmetry:

$$\begin{aligned}
&\eta \to \tilde{\eta} = \eta \\
&x^i \to \tilde{x}^i = x^i + n^i \\
&v^i \to \tilde{v}^i = v^i + n^i \\
&\Phi \to \tilde{\Phi} = \Phi - (\mathcal{H} n^{i\prime} + n^{i\prime\prime}) x^i, \\
&\delta \to \tilde{\delta} = \delta
\end{aligned} \tag{3.13}$$

where n^i is a function of time but not space. It can be shown that under this set of transformations, Eq. (3.12) takes on exactly the same form, with all the variables replaced by ones with a $\tilde{}$ on top. To see that this is true, it is important to keep in mind that:

$$\frac{\partial}{\partial \tilde{\eta}} = -n^{i\prime} \frac{\partial}{\partial x^i} + \frac{\partial}{\partial \eta}\,, \tag{3.14}$$

where on the left, $\tilde{x}^i$ is held fixed, and on the right, $\partial/\partial x^i$ is at a fixed η, and $\partial/\partial\eta$ is at a fixed x^i. On the other hand $\partial/\partial\tilde{x}^i = \partial/\partial x^i$. The symmetry transformation described by Eq. (3.13) is a time-dependent spatial translation. (Henceforth, we will occasionally refer to this somewhat sloppily as simply translation.) Under this translation, the velocity gets shifted in the expected manner. The gravitational potential needs to be shifted correspondingly to preserve the form of the Euler equation. The density δ, on the other hand, does not change at all, in the sense that $\tilde{\delta}(\tilde{x}) = \delta(x)$.

Eq. (3.13) is a symmetry regardless of the time-dependence of n^i. For the purpose of deriving the consistency relations, however, we need to impose an additional condition. Suppose we start with $v^i = 0$ in Eq. (3.13); we would like the velocity generated by the transformation, i.e. $\tilde{v}^i = n^{i\prime}$, to be the long wavelength limit of an actual physical mode that satisfies the equations of motion. We say long wavelength because $n^{i\prime}$ has no spatial dependence and so is strictly speaking a $q = 0$ mode (q being the wavenumber/momentum in Fourier space). What we want to impose is this: if we take a physical velocity mode at a finite q, and make its q smaller and smaller, we would like $n^{i\prime}$ to match its time-dependence. Following the terminology used in general relativity, we refer to this as the *adiabatic mode condition* [84]. This condition ensures that the symmetry transformation generates a velocity mode that evolves in a physical way. It is easy to see that at long wavelength, where the equations can be linearized, Eq. (3.12) can be combined into a single equation:

$$\partial_i v^{i\prime} + \mathcal{H}\partial_i v^i = 4\pi G\bar{\rho}a^2\partial_i \int d\eta\, v^i \,, \tag{3.15}$$

or written in a more familiar form:

$$\delta'' + \mathcal{H}\delta' = 4\pi G\bar{\rho}a^2\delta \quad \text{with} \quad \delta' = -\partial_i v^i \,. \tag{3.16}$$

The linear evolution of the overdensity field δ is encoded in a time dependent linear growth factor $D(\eta)$:

$$\delta(\eta, \mathbf{x}) = D(\eta)\, \delta_0(\mathbf{x}) \tag{3.17}$$

where $\delta_0(\mathbf{x})$ is the initial condition of overdensity evolution. This means that the linear growth factor obeys the equation

$$D''(\eta) + \mathcal{H}D'(\eta) = 4\pi G\bar{\rho}a^2(\eta)\,D(\eta) \tag{3.18}$$

One might be tempted to say a velocity going like $n^{i\prime}$ satisfies Eq. (3.15) trivially for an $n^{i\prime}$ of arbitrary time-dependence, since $n^{i\prime}$ has no spatial dependence. What we want, however, is for $n^{i\prime}$ to have the same time-dependence as that of a velocity mode at a low, but finite momentum. In other words, we impose the *adiabatic mode condition*:

$$n^{i\prime\prime} + \mathcal{H}n^{i\prime} = 4\pi G\bar{\rho}a^2 n^i \tag{3.19}$$

i.e. $n^i(\eta)$ has the same time-dependence as the linear growth factor $D(\eta)$, assuming growing mode initial conditions. Effectively, we demand that our symmetry-generated velocity-shift (or more precisely, the nonlinear part thereof) satisfy Eq. (3.15) with the spatial gradient removed.

The Background Wave Argument. Next, we give the background wave derivation of the consistency relation. More sophisticated and rigorous derivations exist [5–7, 81, 83], but the background wave argument has the virtue of being fairly intuitive. Our goal here is to go over the underlying assumptions, and formulate the argument in such a way to ease later generalizations. The form of our expressions follow closely those in H2K [5].

Before we carry out the argument, it is convenient (especially for later discussions) to introduce the velocity potential π, assuming potential flow on large scales,[3] in which case,

[3] Assuming potential flow is not strictly necessary, as the argument can be made using the velocity v^i itself in place of $\partial_i\pi$. The reason we make this assumption is that the velocity enters into our derivation mainly as a large scale or low momentum mode. Assuming the growing mode initial condition, the large scale velocity does take the form of a potential flow. Indeed, vorticity remains zero until orbit crossing. We do *not* assume potential flow on small scales.

the symmetry transformation of Eq. (3.13) tells us:

$$\pi \to \tilde{\pi} = \pi + n^{i\prime} x^i \quad \text{with} \quad v^i = \partial_i \pi \,. \tag{3.20}$$

The velocity potential π has the hallmark of a pion or Nambu-Goldstone boson: it experiences a *nonlinear* shift under the symmetry transformation.[4] Note that π also implicitly has a *linear* shift:

$$\tilde{\pi}(\tilde{x}) = \pi(x) + n^{i\prime} x^i \sim \pi(\tilde{x}) - n^i \partial_i \pi + n^{i\prime} x^i \tag{3.21}$$

where we have Taylor expanded $\pi(x)$ to first order in $\tilde{x}^i - x^i$, assuming a small n^i. Here, linear and nonlinear refer to whether or not the transformation is linear in π (or any other LSS fields). The total shift in π, i.e. $\tilde{\pi}(\tilde{x}) - \pi(\tilde{x})$, thus has both linear and nonlinear pieces:

$$\Delta_{\text{lin.}} \pi = -n^i \partial_i \pi \quad , \quad \Delta_{\text{nl.}} \pi = n^{i\prime} x^i \,. \tag{3.22}$$

Similarly, the overdensity changes by the amount $\tilde{\delta}(\tilde{x}) - \delta(\tilde{x})$:

$$\Delta_{\text{lin.}} \delta = -n^i \partial_i \delta \quad , \quad \Delta_{\text{nl.}} \delta = 0 \,, \tag{3.23}$$

i.e. δ experiences no nonlinear shift.

Consider an N-point function involving a product of N LSS observables. Let us denote each as $\mathcal{O}$, labeled by momentum, so that the N-point function is $\langle \mathcal{O}_{\vec{k}_1} \mathcal{O}_{\vec{k}_2} ... \mathcal{O}_{\vec{k}_N} \rangle$. Here, the $\mathcal{O}$'s at different momenta need not be the same observable. For instance, one can be δ, the other can be the gravitational potential, et cetera. They need not even be

[4] Note that Φ also experiences a nonlinear shift (Eq. 3.13), and thus can also be used as the pion in this derivation. The two give the same result, see section 3.1.4.3.

evaluated at the same time. We are interested in this N-point function in the presence of some long wavelength (soft) π. Let us imagine splitting all fluctuations into hard and soft modes, with $\vec{k}_1, ..., \vec{k}_N$ falling into the hard category. The N-point function obtained by integrating over the hard modes, but leaving the soft modes of π unintegrated, can be Taylor expanded as:

$$\langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle_{\pi_{\rm soft}} \approx \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle_0 + \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{\delta \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle_{\pi_{\rm soft}}}{\delta \pi^*_{\vec{p}}} \Big|_0 \pi^*_{\vec{p}} \,, \tag{3.24}$$

where we have taken the functional derivative with respect to, and summed over the Fourier modes of, π with *soft* momenta $\vec{p}$. Multiplying both sides by $\pi_{\vec{q}}$ (where $\vec{q}$ is also soft) and ensemble averaging over the soft modes, one finds

$$\frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle}{P_\pi(q)} = \frac{\delta \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle_{\pi_{\rm soft}}}{\delta \pi^*_{\vec{q}}} \Big|_0 \,. \tag{3.25}$$

We have used the definition of the power spectrum: $\langle \pi_{\vec{q}} \pi^*_{\vec{p}} \rangle = (2\pi)^3 \delta_D(\vec{q} - \vec{p}) P_\pi(q)$, with δ_D being the Dirac delta function. We can on the other hand compute the derivative on the right hand side this way:

$$\int \frac{d^3\vec{p}}{(2\pi)^3} \frac{\delta \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle_{\pi_{\rm soft}}}{\delta \pi^*_{\vec{p}}} \Big|_0 \Delta_{\rm nl.} \pi^*_{\vec{p}} = \Delta_{\rm lin.} \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle \,. \tag{3.26}$$

This statement says that the change to the N-point function induced by the symmetry transformation (the right hand side) is equivalent to the change to the N-point function by adding a long-wavelength background π induced by the same symmetry (the left hand side). We will unpack it a bit more in section 3.1.3.1. A careful reader might note that there is no reason why one should include on the right hand side only the *linear* part of the transformation of the N-point function. That is true: by including only the linear transformation, we are effectively dealing with the *connected* N-point function. For a proof, see H2K. [5] Combining Eqs. (3.25) and (3.26), and adding the superscript c for connected N-point function, we have:

$$\boxed{\int \frac{d^3q}{(2\pi)^3} \frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\pi(q)} \Delta_{\rm nl.} \pi^*_{\vec{q}} = \Delta_{\rm lin.} \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^c \, .} \tag{3.27}$$

It is important to note that the connected correlation functions on both sides contain delta functions. This way of writing the consistency relation follows the Ward identity treatment of H2K, and is applicable to any symmetries with π as the Nambu-Goldstone boson. The background wave argument has the advantage of being intuitive, but is a bit heuristic. Readers interested in subtleties can consult e.g. H2K. Our final result here matches theirs.

Let us apply Eq. (3.27) to the translation symmetry. To be specific, let us take our observable $\mathcal{O}$ to be the mass overdensity δ. We use the following convention for the Fourier transform of some function f:

$$f(\vec{q}) = \int d^3x f(\vec{x}) e^{i\vec{q}\cdot\vec{x}} \, , \; f(\vec{x}) = \int \frac{d^3q}{(2\pi)^3} f(\vec{q}) e^{-i\vec{q}\cdot\vec{x}} \, . \tag{3.28}$$

Eqs. (3.22) and (3.23) thus imply:

$$\Delta_{\rm nl.} \pi^*_{\vec{q}} = i \, n^{j\prime} \frac{\partial}{\partial q^j} [(2\pi)^3 \delta_D(\vec{q})] \quad , \quad \Delta_{\rm lin.} \delta_{\vec{k}} = i n^j k^j \delta_{\vec{k}} \, . \tag{3.29}$$

[5] The restriction to the *connected* N-point function will not be so relevant for the Newtonian (KRPP) consistency relation or the general relativistic dilation consistency relation, i.e. they take the same form whether the consistency relation is phrased in terms of *connected* or *general* correlation functions. The restriction is relevant for those consistency relations that involve more than one derivative on the right hand side. For them, it is important to keep in mind that the N momenta on the right hand side sum to zero (see [83]); deriving *general* consistency relations from *connected* ones necessarily introduce extra terms.

Substituting this into Eq. (3.27), we find

$$\lim_{\vec{q}\to 0} n^{j\prime}(\eta)\frac{\partial}{\partial q^j}\left[\frac{\langle \pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^c}{P_\pi(q)}\right] = -\sum_{a=1}^{N} n^j(\eta_a)k_a^j\,\langle \delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^c\,, \tag{3.30}$$

where η is the time implicitly assumed for $\pi_{\vec{q}}$, and η_a is the time for $\delta_{\vec{k}_a}$. Note that k_a^j refers to the j-component of the vector $\vec{k}_a$. Since $n^j(\eta)$ has the time-dependence of the linear growth factor $D(\eta)$ (from the adiabatic mode condition), but can otherwise point in an arbitrary direction, we conclude:

$$\lim_{\vec{q}\to 0} \frac{\partial}{\partial q^j}\left[\frac{\langle \pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^c}{P_\pi(q)}\right] = -\sum_{a=1}^{N} \frac{D(\eta_a)}{D'(\eta)}k_a^j\,\langle \delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^c\,. \tag{3.31}$$

We have yet to remove the delta functions from both sides. To do so, we use the following, pure shift, symmetry:

$$\pi \to \tilde{\pi} = \pi + \text{const.}\,, \tag{3.32}$$

This symmetry does not involve transforming space-time at all, and so none of the observables receive a *linear* shift. The argument leading to Eq. (3.27) thus tells us

$$\lim_{\vec{q}\to 0} \frac{\langle \pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)} = 0\,. \tag{3.33}$$

We use the superscript c' to denote the connected correlation function with the overall delta function removed:

$$\langle \mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^c = (2\pi)^3\delta_D(\vec{k}_1+...+\vec{k}_N)\langle \mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'} \tag{3.34}$$

Combining Eqs. (3.31) and (3.33), we have the Newtonian *translation consistency relation*:

$$\boxed{\lim_{\vec{q}\to 0}\frac{\partial}{\partial q^j}\left[\frac{\langle\pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\right] = -\sum_{a=1}^{N}\frac{D(\eta_a)}{D'(\eta)}k_a^j\,\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\,,} \tag{3.35}$$

where η is the time for the soft mode $\pi_{\vec{q}}$, and each η_a is the time for the corresponding hard mode $\delta_{\vec{k}_a}$. This turns out to be a common feature for all consistency relations as we will see: Eqs. (3.33) and (3.31) are two consistency relations that differ by one derivative with respect to q. The former allows us to remove the delta function from the latter in a straightforward way – the result is Eq. (3.35). The form adopted by KRPP is to integrate Eq.(3.35) over q, and using Eq. (3.33), to obtain:

$$\lim_{\vec{q}\to 0}\left[\frac{\langle\pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\right] = -\sum_{a=1}^{N}\frac{D(\eta_a)}{D'(\eta)}\vec{q}\cdot\vec{k}_a\,\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\,. \tag{3.36}$$

The soft mode $\pi_{\vec{q}}$, by the linearized continuity equation, is related to $\delta_{\vec{q}}$ by[6]

$$\delta_{\vec{q}}(\eta) = q^2\frac{D(\eta)}{D'(\eta)}\pi_{\vec{q}}(\eta)\,. \tag{3.37}$$

One can therefore rewrite Eq. (3.36) as[7]

$$\lim_{\vec{q}\to 0}\left[\frac{\langle\delta_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\delta(q)}\right] = -\sum_{a=1}^{N}\frac{D(\eta_a)}{D(\eta)}\frac{\vec{q}\cdot\vec{k}_a}{q^2}\,\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\,. \tag{3.38}$$

In the context of this Newtonian derivation, we would like to think of the form expressed in Eq. (3.35) and Eq. (3.33) as more fundamental, since it is π that experiences a nonlinear

[6] In the context of a consistency relation which is purported to be non-perturbative, one might wonder if using the linear relation between δ and π (for the soft mode only) is justified. It can be shown that including nonlinear corrections to this relation leads to terms subdominant in the squeezed limit of the correlation function. See [91].

[7] The use of the equation of motion within the *connected* correlation function does not lead to contact terms. See e.g. H2K and [83] on the role of contact terms in Ward identity arguments.

shift in the symmetry transformation, acting as the pion, and since the expression in Eq. (3.38) contains two non-relativistic consistency relations, one trivial and one nontrivial. It is also worth stressing that Eq. (3.35) does not constrain $\partial^2/\partial q^2$ of $[\langle\pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}/P_\pi(q)]$ in the soft limit, i.e. Eq. (3.36) can in principle contain $O(q^2)$ corrections, and Eq. (3.38) can contain $O(q^0)$ corrections.

Let us close this derivation by observing that the only assumption made about the hard modes is how they transform under the symmetry (in particular, the linear part of their transformation; see Eqs. (3.23, 3.29). Thus, suppose we have some observable $\mathcal{O}$ whose linear transformation under the spatial translation is:

$$\Delta_{\rm lin.}\mathcal{O}_{\vec{k}} = in^j g^j \mathcal{O}_{\vec{k}}\,. \tag{3.39}$$

For instance, if $\mathcal{O}$ is the galaxy overdensity, we expect $g^j = k^j$, but g^j could take other forms for other observables. Exactly the same derivation then gives the Newtonian translation consistency relation in a more general form:

$$\boxed{\lim_{\vec{q}\to 0}\frac{\partial}{\partial q^j}\left[\frac{\langle\pi_{\vec{q}}\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\right] = -\sum_{a=1}^{N}\frac{D(\eta_a)}{D'(\eta)}g_a^j\,\langle\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}\,,} \tag{3.40}$$

where we have allowed the possibility that the N hard modes correspond to different observables, thus potentially a different g_a^j for each $a = 1, ..., N$. Corollaries – analogs of Eqs. (3.36, 3.38) – follow in the same way:

$$\lim_{\vec{q}\to 0}\left[\frac{\langle\pi_{\vec{q}}\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\right] = -\sum_{a=1}^{N}\frac{D(\eta_a)}{D'(\eta)}\vec{q}\cdot\vec{g}_a\,\langle\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}\,, \tag{3.41}$$

and

$$\lim_{\vec{q}\to 0}\left[\frac{\langle\delta_{\vec{q}}\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\right] = -\sum_{a=1}^{N}\frac{D(\eta_a)}{D(\eta)}\frac{\vec{q}\cdot\vec{g}_a}{q^2}\langle\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}\,, \tag{3.42}$$

where it should be understood that Eq. (3.41) and Eq. (3.42) contain $O(q^2)$ and $O(q^0)$ corrections respectively. Phrased as such, the consistency relation is fairly robust. The detailed dynamics of the hard modes has no relevance; it matters not whether the corresponding observables are astrophysically messy or highly nonlinear. All we need to know is how they transform under a spatial translation. To understand this robustness better, it is helpful to study concrete examples, which is the subject of the next section.

3.1.3.1 Robustness and Limitations of the Consistency Relation(s)

To understand better the robustness of the consistency relation, it is instructive to ask the question: when does it fail? As we will see, the consistency relation stands on three legs: the existence of the time-dependent translation symmetry, the single-field initial condition, and the adiabatic mode condition. All three are necessary in order for the consistency relation to hold. Here, we focus on the KRPP consistency relation as a specific example, but the points we raise are general, pertaining to other consistency relations sec. (3.1.4) as well.

1. Initial condition: the single-field assumption. A crucial step in the derivation is Eq. (3.26): that linear transformations of a collection of hard modes, represented by $\mathcal{O}_{\vec{k}}$, can be considered equivalent to placing the same hard modes in the presence of a soft mode – the pion $\pi_{\vec{q}}$. That this single soft mode is sufficient to account for all the transformations of the hard modes is an assumption about initial conditions. In the context of inflation, the assumption is often phrased as that of a single field or a single clock. In our Newtonian LSS context, in addition to keeping only the growing modes, essentially the assumption is that of Gaussian initial conditions.[8] More precisely, one demands that the initial condition does not contain a coupling between soft and hard modes beyond that captured by Eq. (3.26). We will follow the inflation terminology and call this the *single-field assumption*.

2. Adiabatic mode condition. Another crucial ingredient in the derivation is the *adiabatic mode condition*, that the symmetry transformation have the correct time-dependence so that the nonlinear shift $\Delta_{\rm nl.}\pi$ may be the long wavelength limit of an actual physical mode. We stress that this is an additional requirement with non-trivial implications on top of demanding a symmetry. To spell them out, it is useful to have a concrete example. Since the consistency relation is purported to be robust, in the sense that the hard modes can be highly nonlinear and even astrophysically complex, let us write down a system of equations that allow for these complexities:

$$\delta'_{(a)} + \vec{\nabla}\cdot[(1+\delta_{(a)})\vec{v}_{(a)}] = R_{(a)} \quad , \quad {\vec{v}_{(a)}}' + \vec{v}_{(a)}\cdot\vec{\nabla}\vec{v}_{(a)} + \mathcal{H}\vec{v}_{(a)} = -\vec{\nabla}\Phi + \vec{F}_{(a)}\,. \tag{3.43}$$

Here a labels the species: for instance, it can be dark matter, populations of galaxies, baryons and so on. $R_{(a)}$ represents a source term for the density evolution. For dark matter, we expect $R_{(a)} = 0$ (barring significant annihilation or decay). For galaxies, $R_{(a)}$ quantifies the effect of galaxy formation and mergers. All particles are subjected to the same gravitational force plus a species-dependent force $\vec{F}_{(a)}$. The gravitational potential Φ is sourced by the total mass fluctuation:

$$\nabla^2\Phi = 4\pi G a^2\bar{\rho}\delta_T\,, \tag{3.44}$$

where δ_T represents the effective total mass density fluctuation from all particles.

[8]Note that in the inflationary context, Gaussianity is also assumed – for the wave-function in the far past.

A natural generalization of the (time-dependent) translational symmetry from the previous section would be:

$$
\begin{aligned}
&\eta \to \tilde{\eta} = \eta \\
&x^i \to \tilde{x}^i = x^i + n^i \\
&{v_{(a)}}^i \to {\tilde{v}_{(a)}}^i = {v_{(a)}}^i + n^i \\
&\Phi \to \tilde{\Phi} = \Phi - (\mathcal{H} n^{i\prime} + n^{i\prime\prime}) x^i \\
&\delta_{(a)} \to \tilde{\delta}_{(a)} = \delta_{(a)} \\
&\delta_T \to \tilde{\delta}_T = \delta_T \\
&R_{(a)} \to \tilde{R}_{(a)} = R_{(a)} \\
&{F_{(a)}}^i \to {\tilde{F}_{(a)}}^i = {F_{(a)}}^i
\end{aligned}
\tag{3.45}
$$

This means that $R_{(a)}$ and $\vec{F}_{(a)}$ remain invariant under this symmetry. For instance, they are invariant if R and $\vec{F}$ depend only on δ, on the spatial gradient of $\vec{v}$, or on the second gradients of Φ. R and $\vec{F}$ could even have explicit dependence on time η. If R or $\vec{F}$ depends on $\vec{v}$ with no gradients, unless the dependence is of the form $\vec{v}_{(a)} - \vec{v}_{(b)}$ (so that the shifts in the velocities of the two different species (a) and (b) cancel out), the invariance is violated. Specializing to the case of galaxies, what this means is that the number density evolution and dynamics of the galaxies do not care about the absolute size of the velocity, but only about the velocity difference (either between neighbors, or between species). The only context in which the absolute size of velocity plays a role is through Hubble friction – this is the origin of the $\mathcal{H}$ dependent term in the nonlinear shift of Φ. In other words, Hubble friction aside, *galaxy formation and dynamics is frame invariant*, which seems a fairly safe assumption. For instance, dynamical friction, which no doubt exerts an influence on galaxies, should depend on velocity difference.

Thus, let us assume Eq. (3.45) is a symmetry of our system – note that this is a symmetry regardless of the time-dependence of n^i. As emphasized earlier, this is not enough to guarantee the validity of the consistency relation. To derive the consistency relation, n^i must have the correct time-dependence: $n^{i\prime}$ (the nonlinear shift in v^i) must match the time-dependence of a physical long-wavelength velocity perturbation.[9] This has to hold for all species, meaning the *same* $n^{i\prime}$ matches the long-wavelength velocity perturbation of each and everyone of the species. In other words, *all species should move with the same velocity on large scales.* This leads to two subtleties, which are best illustrated by assuming an explicit form for $\vec{F}$. Consider:

$$\begin{aligned}&\vec{F}_{(a)} = -{c_s}_{(a)}^2 \vec{\nabla} \ln(1+\delta_{(a)}) - \beta_{(a,b)}(\vec{v}_{(a)} - \vec{v}_{(b)}) - \alpha_{(a)}\vec{\nabla}\varphi \\ &\nabla^2\varphi = 8\pi G \sum_a \alpha_{(a)} a^2 \bar{\rho}_a \delta_a\,. \end{aligned} \tag{3.46}$$

The first term on the right of the expression for $\vec{F}$ represents some sort of pressure – c_s is the sound speed – this would be relevant if the subscript (a) represents baryons at finite temperature. The second term represents some sort of friction that depends on the velocity difference between two species, with a coefficient β. The third term represents an additional fifth force, mediated by the scalar φ, with a coupling α. The scalar φ obeys a Poisson-like equation. In scalar-tensor theories, the tensor part of the theory mediates a universal gravitational force (described by the gravitational potential Φ), but the scalar need not be universally coupled: hence we allow the coupling $\alpha_{(a)}$ to depend on the species (see e.g. [85, 86]). The form for the additional force $\vec{F}$ proposed in Eq. (3.46) is fairly generic: counting derivatives, we can see that the pressure term goes like $\partial\delta$, whereas the other two terms go like $\partial^{-1}\delta$(due to the equations of motion). In terms of the symmetry

[9] This typically means n^i should satisfy Eq. (3.19), i.e. n^i must evolve like the linear growth factor D. This holds if all particles, on large scales, evolve like dark matter and if gravity is the only long-range force. See further discussion below.

transformation, one can see that

$$\varphi \to \tilde{\varphi} = \varphi \tag{3.47}$$

is compatible with Eq. (3.45). Thus, we have a system (Eq. 3.43) that respects the translational symmetry spelled out in Eq. (3.45), even if many different kinds of forces are present, including non-gravitational or modified gravitational ones such as in Eq. (3.46). We wish to see how, despite the presence of the (time-dependent) translational symmetry, there can still be a breakdown of the consistency relation, due to obstructions in satisfying the adiabatic mode condition – that the velocity perturbations of all species should be equal on large scales.

2a. Soft dynamics constraint. In the long wavelength limit, one can ignore the pressure term compared to the other two terms in the expression for $\vec{F}_{(a)}$. Let us first focus on the fifth-force term. This term is at the same level in derivatives as the normal gravitational force $-\vec{\nabla}\Phi$, and thus both have to be taken into account on large scales. The problem with a long-range fifth-force is the non-universal coupling: if there is a different coupling $\alpha_{(a)}$ for each kind of particles, the different species will move with different velocities even on large scales. This means no *single* $n^{i\prime}$ can possibly generate long-wavelength velocity perturbations for all species. In other words, unless the soft (large-scale) dynamics obeys the equivalence principle, the consistency relation would be violated, as emphasized by [92, 100]. We stress that, in our example, the violation of the equivalence principle occurs *without* the violation of the translation symmetry described by Eq. (3.45). The fact that the consistency relation is not obeyed is entirely because of the failure to satisfy the adiabatic mode condition when the equivalence principle is violated. The friction term (the second term on the right of Eq. 3.46), on the other hand, is compatible with the adiabatic mode condition – it simply vanishes if the velocities of different species are equal, and is therefore consistent with the large scale requirement that all species flow with the same velocity. To sum up, the *soft dynamics constraint* is: for the consistency relation to be valid, the dynamics on large scales must be consistent with all species moving with the same velocity.

2b. Squeezing constraint. Let us next turn to the pressure term (the first term on the right of Eq. (3.46)). Since different species have different sound speeds, this also leads to differences in velocity flows. This is relatively harmless though, since the pressure term becomes subdominant on large scales. Thus, there is no problem with the adiabatic mode condition, which is really a condition on motions in the soft limit $q \to 0$. The presence of pressure does lead to a practical limitation on the application of the consistency relation, however. The consistency relation is a statement about an $(N+1)$-point function in the squeezed limit $q \ll k_1, ..., k_N$. There is the practical question of how small q has to be. An important requirement is that q must be sufficiently small that the velocity perturbations of different species have the same time-dependence as that generated by a single $n^{i\prime}$. In the present context, it means $q < \mathcal{H}/c_s$, i.e. the length scale must be above the Jeans scale.[10]We refer to this as the *squeezing constraint*: the soft leg of the consistency relation must be sufficiently soft that any difference in force on the different species becomes negligible. This is worth emphasizing, because clearly dark matter and baryons are subject to different forces: while that does not by itself lead to the breakdown of the consistency relations, one has to be careful to make sure that the squeezed correlation function is sufficiently squeezed.

3. Galaxy-biasing. It is also instructive to approach the subject of consistency relation violation from the viewpoint of galaxy-biasing. What kind of galaxy-biasing would lead to the violation of consistency relation? We can only address this question in the perturbative regime, but it nonetheless provides some useful insights. Suppose the galaxy overdensity

$\delta_{(a)}$ (of type a) and matter density δ are related by:

$$\delta_{(a)\vec{k}} = b_{(a)} \left(\delta_{\vec{k}} + \int \frac{d^3k'}{(2\pi)^3} \frac{d^3k''}{(2\pi)^3} (2\pi)^3 \delta_D(\vec{k} - \vec{k}' - \vec{k}'') W_{(a)}(\vec{k}', \vec{k}'') \left[\delta_{\vec{k}'} \delta_{\vec{k}''} - \langle \delta_{\vec{k}'} \delta_{\vec{k}''} \rangle \right] \right) \tag{3.48}$$

where $b_{(a)}$ is a linear bias factor (independent of momentum) and $W_{(a)}$ is a kernel that describes a general quadratic bias. To the lowest order in perturbation theory, it can be shown that the bispectrum between three types of galaxies a, b and c, at momenta $\vec{q}, \vec{k}_1, \vec{k}_2$ and times η, η_1, η_2 respectively, is

$$\begin{aligned}
&B_{(abc)}(\vec{q}, \eta; \vec{k}_1, \eta_1; \vec{k}_2, \eta_2) = \\
&2b_{(a)}b_{(b)}b_{(c)}P_\delta(k_1, \eta, \eta_1)P_\delta(k_2, \eta, \eta_2)\Big[\left(\frac{5}{7} + \frac{1}{2}\hat{k}_1 \cdot \hat{k}_2 \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7}(\hat{k}_1 \cdot \hat{k}_2)^2 \right) + W_{(a)}(-\vec{k}_1, -\vec{k}_2) \Big] \\
&+ 2b_{(a)}b_{(b)}b_{(c)}P_\delta(q, \eta, \eta_1)P_\delta(k_2, \eta_1, \eta_2)\Big[\left(\frac{5}{7} + \frac{1}{2}\hat{q} \cdot \hat{k}_2 \left(\frac{q}{k_2} + \frac{k_2}{q} \right) + \frac{2}{7}(\hat{q} \cdot \hat{k}_2)^2 \right) + W_{(b)}(-\vec{q}, -\vec{k}_2) \Big] \\
&+ 2b_{(a)}b_{(b)}b_{(c)}P_\delta(q, \eta, \eta_2)P_\delta(k_1, \eta_1, \eta_2)\Big[\left(\frac{5}{7} + \frac{1}{2}\hat{q} \cdot \hat{k}_1 \left(\frac{q}{k_1} + \frac{k_1}{q} \right) + \frac{2}{7}(\hat{q} \cdot \hat{k}_1)^2 \right) + W_{(c)}(-\vec{q}, -\vec{k}_1) \Big]
\end{aligned} \tag{3.49}$$

where P_δ is the linear mass power spectrum – its two time-arguments signify the fact that the two δ's involved can be at different times. We are interested in the $\vec{q}\to 0$ limit:

$$\begin{aligned}
&\lim_{\vec{q}\to 0} b_{(a)}\frac{B_{(abc)}(\vec{q},\eta;\vec{k}_1,\eta_1;\vec{k}_2,\eta_2)}{P_{(aa)}(q,\eta,\eta)} \\
&= E + \frac{D(\eta_1)}{D(\eta)}\left[\frac{\vec{q}\cdot\vec{k}_2}{q^2} + 2W_{(b)}(-\vec{q},-\vec{k}_2) + O(q^0)\right] P_{(bc)}(k_2,\eta_1,\eta_2) \\
&\quad + \frac{D(\eta_2)}{D(\eta)}\left[\frac{\vec{q}\cdot\vec{k}_1}{q^2} + 2W_{(c)}(-\vec{q},-\vec{k}_1) + O(q^0)\right] P_{(bc)}(k_1,\eta_1,\eta_2) \\
= E - &\left[\frac{D(\eta_1)}{D(\eta)}\left(\frac{\vec{q}\cdot\vec{k}_1}{q^2} - 2W_{(b)}(-\vec{q},\vec{k}_1)\right) + \frac{D(\eta_2)}{D(\eta)}\left(\frac{\vec{q}\cdot\vec{k}_2}{q^2} - 2W_{(c)}(-\vec{q},\vec{k}_2)\right) + O(q^0)\right] \\
&P_{(bc)}(k_1,\eta_1,\eta_2)
\end{aligned} \tag{3.50}$$

where

$$E \equiv 2\frac{D(\eta)^2}{D(\eta_1)D(\eta_2)}\frac{b_{(a)}^2}{b_{(b)}b_{(c)}}\frac{P_{(bc)}(k_1,\eta_1,\eta_2)P_{(bc)}(k_2,\eta_1,\eta_2)}{P_{(aa)}(q,\eta,\eta)}\left[O(q^2) + W_{(a)}(-\vec{k}_1,-\vec{k}_2)\right]. \tag{3.51}$$

We have used

$$\begin{aligned}
P_{(aa)}(q,\eta,\eta) &= (D(\eta)/D(\eta_1))b_{(a)}^2 P_\delta(q,\eta,\eta_1) \\
P_{(bc)}(k_1,\eta_1,\eta_2) &= b_{(b)}b_{(c)}P_\delta(k_1,\eta_1,\eta_2)
\end{aligned} \tag{3.52}$$

and so on (appropriate only perturbatively). Comparing this expression with the consistency relation expressed in Eq. (3.42) (identifying $\vec{g}_a$ with $\vec{k}_a$), we see that the two agree if E, $W_{(b)}$ and $W_{(c)}$ can be ignored, and $b_{(a)} = 1$. A number of comments are in order.

[10]Since the Jeans scale changes with time in general, the formal requirement is therefore that the soft-mode be longer than the Jeans scale at all times: $q < (\mathcal{H}/c_s)_{rec.}$ where the maximum size of the sound horizon may be conservatively estimated to be the size at recombination. In practice, since most of the present day non-Gaussianity is generated at late times, it is typically sufficient to require that the contribution from early times be subdominant. Using second order perturbation theory, we estimate that $q < \mathcal{H}/c_{s_{rec.}}(D(\eta_{obs.})/D(\eta_{rec.}))$ is a sufficient parametric condition for Eq. 3.35 to be valid (at least away from the equal time limit).

First, let us focus on the case with no galaxy biasing, so that Eq. (3.50) simply constitutes a perturbative check of the consistency relation for the mass overdensity i.e. Eq. (3.38). We see that the term E can be ignored compared to the terms that are kept only if the soft power spectrum is not too blue: assuming $P_{(aa)}(q) \sim q^n$ for small q, the validity of the consistency relation requires $n < 3$. Note that $P_{(aa)}$ is the mass spectrum in the absence of galaxy biasing. This is a limitation on the consistency relation that is not often emphasized. In practice though, the realistic power spectrum has no problem satisfying this requirement.

Let us next consider the effects of galaxy biasing. The second point we would like to raise is that the soft mode must be kept unbiased. There are two reasons for this, one trivial, the other less so. The trivial reason is that, if the soft mode is biased, the left hand side of the consistency relation then has to be corrected by a factor of $b_{(a)}$, the linear bias factor for the soft mode. This is not a big problem: one can obtain an estimate of the linear bias and correct the consistency relation when comparing against observations of the galaxy bispectrum. The more non-trivial problem is the presence of the quadratic bias kernel $W_{(a)}$ in E. Consider for instance a local biasing model of the form: $\delta_{(a)} = b_{(a)}\delta + b_{(a),2}\delta^2/2$ in real space, where $b_{(a)}$ and $b_{(a),2}$ are constants, typically referred to as the linear and quadratic bias factors. In this case $W_{(a)} = b_{(a),2}/(2b_{(a)})$ has no momentum dependence, and so E contains a contribution that goes like q^{-n} for $P_{(aa)} \sim q^n$. This means one needs $n < 1$ for E to be negligible compared to the terms we keep in the consistency relation. On the largest scales, n approaches 1, though observations suggest it is slightly less than 1. On smaller scales (but still keeping $q \ll k_1, ..., k_N$), the relevant n is on the safe side. Nonetheless, this perturbative check suggests that one should be careful in using the a biased observable for the soft-mode.

Henceforth, let us assume the soft-mode is unbiased but the hard modes are biased, in which case E is safely negligible in the squeezed limit as long as $n < 3$. The third point we wish to raise is that the validity of the consistency relation requires the hard modes be biased in a way that is not too infrared-divergent: $W_{(b)}(-\vec{q}, \vec{k}_1)$ and $W_{(c)}(-\vec{q}, \vec{k}_2)$ cannot contain terms that go like k_1/q or k_2/q i.e.

$$\boxed{W_{(b)}(-\vec{q}, \vec{k}), W_{(c)}(-\vec{q}, \vec{k}) \; < \; 1/q} \tag{3.53}$$

in the $q \to 0$ limit. As mentioned above, the local biasing model typically assumed in LSS studies implies that the kernels $W_{(b)}$ and $W_{(c)}$ are momentum independent, and is thus consistent with the consistency relation. It is worth emphasizing that the word *local* in local biasing is a bit misleading: it merely states that the galaxy density at a given point in real space is related to the mass density at the same point. In reality, galaxies form out of the collapse of larger regions, influenced by the tidal field of the environment: there are therefore good reasons to believe that galaxy biasing is at some level *non-local*, i.e. the galaxy density at a given point is affected by the mass density at other points. This non-locality is not non-locality in the field theory sense, in that there is nothing non-local in the dynamics, and the so-called non-local galaxy bias arises completely out of local processes. A violation of the consistency relation requires more than a non-local galaxy bias though. It requires the non-local biasing kernel to be infrared divergent. This does not appear to be easily obtained, even if tidal effects are taken into account [95]. One way it arises is in a model in which galaxies are born with a velocity bias, as pointed out by [96]. The quadratic kernel $W_{(c)}$ (or $W_{(b)}$), for some galaxy population with a velocity bias of b_v^* at birth (i.e. the galaxy velocity equals b_v^* times the dark matter velocity when the galaxy forms) is

$$W_{(c)}(\vec{q}, \vec{k}) \sim 2b_{(c)}^{-1}(b_v^* - 1) \left(\frac{D_*}{D}\right)^{3/2} \hat{q} \cdot \hat{k} \left(\frac{q}{k} + \frac{k}{q}\right) , \tag{3.54}$$

where D_* is the linear growth factor at the time of birth, and D is growth factor at the time of interest. We display only the term that has a dipolar dependence on the angle between $\vec{q}$ and $\vec{k}$, and have taken the late time limit. This has precisely the kind of infrared divergence in the $q \ll k$ limit which would invalidate the consistency relation. It is interesting that this is also an example where we should have expected a violation of the consistency relation based on earlier arguments – the existence of a scale-independent velocity bias b_v^* means dark matter and galaxies do not flow in the same way, even on large scales. This violates the adiabatic mode condition, and so it is not a surprise that the consistency relation fails. Realistically, velocity bias is present at some level of course, but is expected to approach unity on sufficiently large scales, unless of course the equivalence principle is violated [92]. As a general statement, we can say that a non-local galaxy bias that is more infrared-divergent than Eq. (3.53) is what one needs to violate the consistency relation. It is interesting to ask whether there are other ways to physically generate such a galaxy bias besides through equivalence principle violations. This naturally brings us to the issue of selection.

It is worth emphasizing that the galaxy bias is also partly a selection bias: one chooses to study galaxies of a certain luminosity, color, morphology or some other property of interest. The question is then: can one choose the galaxy sample in such a way as to violate Eq. (3.53)? What if one chooses galaxies based on their motions, for instance, selecting galaxies that have systematically large velocities? It would seem that by hand we have introduced a velocity bias, and thus a violation of the consistency relation. This is actually not a violation in the technical sense. Choosing galaxies based on their motions can be thought of as weighing the galaxies by velocities, i.e. $\delta_g \to \delta_g(1+\pi)$. From the point of view of violating the consistency relation, it is most relevant to consider weighing by the large scale velocity. In that case, it is not surprising one finds additional terms that diverge in the squeezed limit of the correlation function – this is because we have included in the correlation function additional soft modes that carry with them additional powers $1/q$.

4. Robustness. Let us close this section by reiterating how robust the consistency relation is. As long as the underlying assumptions – the existence of the time-dependent translation symmetry, single field initial condition and adiabatic mode condition – are satisfied, the consistency relation is robust. The hard or high momentum modes can be those of any LSS observable (or even mixtures of observables), referred to as $\mathcal{O}_{\vec{k}}$ in Eq. (3.42). No assumption is made about the size of $\mathcal{O}_{\vec{k}}$: it matters not at all how nonlinear or non-perturbative these high momentum observables are. Indeed, we do not even need to know their detailed dynamics: all we need to know is how they transform under the symmetry in Eq. (3.39).[11] Their evolution can be a lot more complicated than that of dark matter (e.g. Eq. 3.43). In other words, $\mathcal{O}_{\vec{k}}$ can be astrophysically messy observables, such as those associated with galaxies. The presence of pressure effects, multiple components, multiple-streaming,[12] star formation, supernova explosions, etc. does not lead to violations of the consistency relations, as long as the adiabatic mode condition – i.e. the soft dynamics constraint and the squeezing constraint – is satisfied. This is why the LSS consistency relations are interesting: they provide a reliable window into the non-perturbative, astrophysically complex regime.

3.1.3.2 A Simple Fluid Lagrangian for LSS

The time-dependent translation symmetry laid out above was justified at the level of the equations of motion. It would be useful to see the same at the level of the action. In this section, we provide the action that describes the dark (i.e. pressureless) matter

[11] If the observable turns out to transform differently from Eq. (3.39), one can go back to the more fundamental Eq. (3.27) to figure out the correct consistency relation.

[12] The soft-mode, by the assumption of growing mode initial condition, is potential flow, but the hard modes need not be, and can even involve multiple streams. See footnote 3.

dynamics under gravity. We should stress that, for our discussion of the consistency relation, the action is not strictly necessary; the equations of motion are as good a guide to the symmetry. Moreover, the action we will write down concerns only dark matter; it does not cover realistic observables such as galaxies, while the consistency relation applies regardless of the complex astrophysics that might be present in such observables. Nonetheless, the dark matter action is useful for conceptual understanding. We provide it here for completeness, and connect it with a more well known fluid action in Appendix B.1. For simplicity, we assume potential flow; an extension to allow for vorticity should be straightforward, along the lines of [97]. Readers not interested in the action perspective can skip to sec. 3.1.4 – the rest of the chapter does not depend on this section.

Let us motivate the construction of the action by reducing the standard pressureless LSS equations (3.12) into a single equation for the velocity potential π. The Euler equation can be integrated once to give:

$$\Phi = -\frac{1}{a}\left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2\right] , \tag{3.55}$$

where $(\nabla\pi)^2$ stands for $\partial_i\pi\partial_i\pi$ and a stands for the scale factor.[13] The Poisson equation then gives us:

$$\delta = -\frac{2M_P^2}{\bar{\rho}a^3}\nabla^2\left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2\right] . \tag{3.56}$$

The continuity equation can thus be turned into a single equation for π:

$$-2M_P^2\nabla^2\left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2\right]' + \bar{\rho}a^3\nabla^2\pi - 2M_P^2\nabla_i\left(\nabla_i\pi\nabla^2\left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2\right]\right) = 0 \tag{3.57}$$

[13] We use ∇_i and ∂_i interchangeably, preferring the former where there is the danger of confusing ∂ with the space-time derivative.

This is a complicated looking equation, but it is not too difficult to guess the form of the associated action:

$$\boxed{S = -\int d^4x \left[\frac{1}{2}\bar{\rho}a^4(\nabla\pi)^2 + M_P^2\left(\nabla\left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2\right]\right)^2\right]\,.} \tag{3.58}$$

The overall normalization (and sign) is arbitrary from the point of view of reproducing the desired equation of motion, but is chosen to conform to a more general action discussed in Appendix B.1. It is straightforward to check that this action is invariant under the time-dependent translation symmetry discussed earlier, namely:

$$\eta \to \tilde{\eta} = \eta \quad , \quad x^i \to \tilde{x}^i = x^i + n^i \quad , \quad \pi \to \tilde{\pi} = \pi + n^{i\prime}x^i\,. \tag{3.59}$$

The dynamics of the velocity potential π is completely fixed by this action. From this point of view, the π equation of motion has the interpretation of the continuity equation, if δ is *defined* by Eq. (3.56) and if the gravitational potential Φ is *defined* by Eq. (3.55) so as to reproduce the Poisson equation. With this understanding, the action takes a fairly simple form:

$$S = \int d^4x\, \frac{\bar{\rho}a^4}{2}\left(\Phi\delta - \vec{v}^2\right)\,, \tag{3.60}$$

i.e. the Lagrangian is the difference between what resembles potential energy and kinetic energy, though with an unexpected overall sign, which can be understood from the larger context of a fluid with pressure (see Appendix B.1).

3.1.4 Consistency Relations from Diffeomorphisms – General Relativistic Symmetries

The time-dependent translation symmetry noted by KRPP Eq. (3.13) appears to be a global symmetry of the Newtonian LSS equations. (Or, more generally, the time-dependent translation as described by Eq. (3.45) is a symmetry of the equations of motion for dark matter and galaxies.) Our goal in this section is to place it in a larger context: the claim is that this symmetry is actually part of a diffeomorphism in the context of general relativity. This perspective is useful for two reasons: first, it helps us make contact with the earlier work on consistency relations in inflation, which are based on diffeomorphism invariance; second, diffeomorphism invariance allows us to systematically write down further consistency relations. The earlier work generally uses the ζ-gauge, alternatively referred to as the unitary or comoving gauge. On the other hand, in LSS studies, the Newtonian gauge is the more natural one to use. Here, we take advantage of the fact that the full list of consistency relations are already known in the unitary or ζ-gauge [5], and transform each known symmetry in ζ-gauge into a symmetry in the Newtonian gauge. This way, we will obtain an infinite tower of consistency relations in the Newtonian gauge. We emphasize that we could equally well proceed by directly working in the Newtonian gauge, and obtain the same results (see [10] on the dilation and special conformal consistency relations obtained this way). One might wonder why writing down consistency relations in the Newtonian gauge is useful if we already know what they are in the unitary gauge. It has to do with the taking of the Newtonian limit, a subject we will discuss later in this section.

In the interest of generality, we allow the presence of multiple components of which pressureless matter/dust is one. We assume adiabatic initial conditions in the sense that all components fluctuate in the same way in the long wavelength limit: in particular their velocity potentials coincide in this limit. We give in sec. 3.1.4.1 the general prescription for transforming symmetries known in the unitary gauge to symmetries in Newtonian gauge. In sec. 3.1.4.2 we focus on the dilation and the special conformal symmetries, which are the symmetries that generate only scalar modes, and we show how the KRPP Newtonian consistency relation arises as the sub-Hubble limit of the latter. We comment on the robustness and limitations of the consistency relations in sec. 3.1.4.3, adding a relativistic twist to some of the comments made earlier. We also discuss the taking of the Newtonian/sub-Hubble limit. We close with sec. 3.1.4.4 on further consistency relations that form an infinite tower – they generally involve the tensor modes. We comment on why there is no useful sub-Hubble limit in these cases.

3.1.4.1 Symmetry Transformations from Diffeomorphisms

Here we are interested in symmetry transformations coming from residual gauge/coordinate transformations (i.e. diffeomorphisms)

$$x'^{\mu} = x^{\mu} + \xi^{\mu}\,, \tag{3.61}$$

that are allowed even after we have applied the usual gauge-fixing. In the context of inflation, a common gauge is the unitary or $\zeta-$gauge:

$$ds^2 = ... + a^2 e^{2\zeta}(e^{\gamma})_{ij}dx^i dx^j\,, \delta\phi = 0, \tag{3.62}$$

where we have omitted the time-time and time-space components of the metric which are obtainable from the given space-space parts by solving the Hamiltonian and momentum constraints. Here, ζ represents the scalar perturbation and the transverse traceless γ_{ij} represents the tensor perturbation. Vector perturbations are ignored because they are not generated by single field models (a brief discussion of vector modes can be found in Appendix B.4). The equal time surface is chosen so that the matter field, which we have called ϕ, has no spatial fluctuation. For our application, there can in general be multiple components, in which case $\delta\phi$ is chosen to vanish for one of them. To be concrete, let us choose this to be the dark matter fluid, i.e. we model it as a fluid described by a Lagrangian of the form $P(X)$, where P is some function of $X \equiv -(\partial\phi)^2$. The velocity potential π is related to ϕ by $\delta\phi = \phi - \bar{\phi} = -\bar{\phi}'\pi$, where $\bar{\phi}$ is the background, and $\bar{\phi}'$ is its conformal time derivative (see Appendix B.1).[14]

The full list of residual diffeomorphisms that respect the unitary gauge is worked out in H2K. Since the unitary gauge is a complete gauge-fixing for diffeomorphisms that vanish at spatial infinity, the residual diffeomorphisms must be those that do not vanish at infinity. They take the form:

$$\xi^0 = 0 \quad , \quad \xi^i = \xi^i_{\rm unit.} \sim x^n \, . \tag{3.63}$$

No time-diffeomorphism is allowed since that would violate the $\delta\phi = 0$ (or $\pi = 0$) unitary gauge condition, and the allowed spatial diffeomorphism, (which we refer to as $\xi^i_{\rm unit.}$) goes like x^n, where $n = 1, 2, ...$ We will give explicit expressions for $\xi^i_{\rm unit.}$ later. They satisfy:[15]

$$\text{scalar} \ + \ \text{tensor symmetries}: \quad \nabla^2 \xi^i_{\rm unit.} + \frac{1}{3}\partial_i(\partial_k \xi^k_{\rm unit.}) = 0 \, . \tag{3.64}$$

This set of symmetries contains subsets that only generate (nonlinearly) scalar modes,

[14] By describing the dark matter using a single fluid field ϕ, we are ignoring orbit-crossing and also vorticity. As is clear in the Newtonian discussion, neither one of these assumptions is strictly necessary. We make them only to simplify the general relativistic discussion.

[15] This holds only to the lowest order in tensor modes. See H2K, and discussion in footnote 30.

and subsets that only generate tensor modes:

$$\begin{aligned} \text{scalar symmetries}: \quad & \partial_i \xi^j_{\rm unit.} + \partial_j \xi^i_{\rm unit.} - \frac{2}{3}\delta_{ij}\partial_k \xi^k_{\rm unit.} = 0 \\ \text{tensor symmetries}: \quad & \partial_i \xi^i_{\rm unit.} = 0 \ , \ \nabla^2 \xi^i_{\rm unit.} = 0 \,. \end{aligned} \tag{3.65}$$

The spatial diffeomorphism $\xi^i_{\rm unit.}$ can be considered to be time-independent.[16]

In LSS studies, it is more common to employ the Newtonian gauge instead:

$$ds^2 = a^2 \left[-(1+2\Phi)d\eta^2 + 2S_i dx^i d\eta + ((1-2\Psi)\delta_{ij} + \gamma_{ij})dx^i dx^j \right] \,, \tag{3.66}$$

where we no longer impose $\pi = 0$; Φ and Ψ are the scalar modes, the transverse traceless γ_{ij} denotes the tensor modes as before, and the divergence-free S_i represents the vector modes (which is set to zero here). Here, we work perturbatively in the metric perturbations, since the Newtonian-gauge metric perturbations are expected to be small even in the highly nonlinear regime where the density fluctuation δ is large, and including higher order metric perturbations corrects the consistency relations by negligible amounts. [17] Under a small diffeomorphism ξ^μ, the nonlinear transformations of the metric fluctuations

[16] Adiabatic mode conditions in the unitary gauge actually make $\xi^i_{\rm unit.}$ time-dependent in general. As shown in H2K, its time-independent part alone is sufficient to deduce the consistency relations. We will implement the adiabatic mode conditions separately in the Newtonian gauge computation.

[17] See footnote 30 for a more detailed discussion of this point.

are:[18]

$$
\begin{aligned}
&\Delta_{\rm nl.}\Phi = -\xi^{0\prime} - \mathcal{H}\xi^0 \\
&\Delta_{\rm nl.}\Psi = \mathcal{H}\xi^0 + \frac{1}{3}\partial_k\xi^k \\
&\Delta_{\rm nl.}g_{0i} = a^2(\partial_i\xi^0 - \partial_0\xi^i) \\
&\Delta_{\rm nl.}g_{ij} - \frac{1}{3}\delta_{ij}\Delta_{\rm nl.}g_{kk} = -a^2\left(\partial_i\xi^j + \partial_j\xi^i - \frac{2}{3}\delta_{ij}\partial_k\xi^k\right)
\end{aligned}
\tag{3.67}
$$

Given each symmetry in the unitary gauge, it is straightforward to deduce the corresponding symmetry in the Newtonian gauge. Let us break it down into a number of steps. First, we begin with the metric in Newtonian gauge, where $\pi = \pi_0 \neq 0$. We assume $\Psi = \Phi$, in the absence of anisotropic stress.[19] To convert to the unitary gauge, we apply a time-diffeomorphism $\xi^0 = -\pi_0$ to make the scalar field ϕ spatially homogeneous. Second, we apply the known unitary-gauge symmetry transformation $\xi^i = \xi^i_{\rm unit.}$. Third, we wish to return to Newtonian gauge. The first and second steps in general make $\Psi \neq \Phi$. To restore equality, we apply an additional time-diffeomorphism $\xi^0 = \pi_0 + \xi^0_{\rm add.}$. We also need to ensure $g_{0i} = 0$ (no vector modes[20]), and thus an additional spatial diffeomorphism $\xi^i_{\rm add.}$ may be necessary. It is shown in Appendix B.2 that the requisite additional time- and space-diffeomorphisms are:

$$
\xi^0_{\rm add.} = -\frac{1}{3c}D'\partial_i\xi^i_{\rm unit.} \quad , \quad \xi^i_{\rm add.} = \frac{1}{c}D\nabla^2\xi^i_{\rm unit.}\,. \tag{3.68}
$$

[18]The net (linear + nonlinear) transformation of the metric is given by $\Delta g_{\mu\nu} = -\xi^\alpha\partial_\alpha g_{\mu\nu} - g_{\alpha\mu}\partial_\nu\xi^\alpha - g_{\alpha\nu}\partial_\mu\xi^\alpha$.

[19]This is an adiabatic mode condition in the Newtonian gauge. See discussions in Appendix B.2.

[20] The absence of vector modes is assumed in two places. Assuming $\nabla^2\xi^i_{\rm unit.} + \partial_i(\partial\cdot\xi)/3 = 0$ means there is no vector mode in the spatial part of the metric. In addition, our choice of $\xi^\mu_{\rm add.}$ ensures there is no vector mode in the space-time part of the metric either.

Here, D is the linear growth factor satisfying the following equation:

$$D'' + 2\mathcal{H}D' - c = 0\,, \tag{3.69}$$

where c is a constant (independent of time and space). In other words, the following diffeomorphism is a symmetry of Newtonian gauge:

$$\xi^0 = \xi^0_{\text{add.}} \quad , \quad \xi^i = \xi^i_{\text{unit.}} + \xi^i_{\text{add.}}\,, \tag{3.70}$$

where $\xi^i_{\text{unit.}}$ is the residual (time-independent) diffeomorphism allowed by the unitary gauge Eqs. (3.64) and (3.65). Furthermore, it can be shown that this diffeomorphism satisfies the adiabatic mode conditions, i.e. the perturbations that are nonlinearly generated match the time-dependence of very soft (growing) physical modes. This is why the linear growth factor D appears in the diffeomorphism. The derivation is given in Appendix B.2. (The attentive reader might wonder why the linear growth factor D – a quantity that shows up in the Newtonian discussion of sub-Hubble perturbations – appears also in a general relativistic discussion, and how Eq. (3.69) is related to the more familiar growth equation (Eq. 3.19). This is discussed in Appendix B.3). An important underlying assumption is that all fluid components move with the same velocity in the soft limit. Under this assumption, it is shown in Appendix B.3 that the velocity, or velocity potential π, evolves as:

$$\lim_{\vec{q}\to 0} \pi_{\vec{q}} \propto D'\,. \tag{3.71}$$

In the context of a general relativistic discussion, this statement (strictly speaking) holds in the super-Hubble limit $q \ll \mathcal{H}$. What is interesting is that for the $\pi_{\vec{q}}$ of pressureless matter, this statement holds also for sub-Hubble (but linear) scales . It is this fact that makes an interesting Newtonian consistency relation possible.[21]

For the purpose of deducing the consistency relations, we also need to know how other LSS observables transform under a diffeomorphism. From the way a scalar should transform, one can see the velocity potential $\pi \equiv -\delta\phi/\bar{\phi}'$ should transform by

$$\begin{aligned} \Delta\pi &= \Delta_{\rm lin.}\pi + \Delta_{\rm nl.}\pi \\ \Delta_{\rm lin.}\pi &= -\xi^0 \frac{(\bar{\phi}'\pi)'}{\bar{\phi}'} \\ &-\xi^i \partial_i \pi \\ \Delta_{\rm nl.}\pi &= \xi^0 \end{aligned} \tag{3.72}$$

We will mostly need only the nonlinear part of the π transformation. As emphasized above, the assumption of potential flow is not strictly necessary. The nonlinear transformation of the velocity can also be deduced by transforming the 4−velocity U^μ:[22]

$$\Delta_{\rm nl.} v^i = \xi^{i\,\prime} \tag{3.73}$$

Another LSS observable of interest is the mass density fluctuation δ. Its transformation is:

$$\begin{aligned} \Delta\delta &= \Delta_{\rm nl.}\delta + \Delta_{\rm lin.}\delta \\ \Delta_{\rm nl.}\delta &= -\xi^0 \frac{\bar{\rho}'}{\bar{\rho}} \\ \Delta_{\rm lin.}\delta &= -\xi^\mu \partial_\mu \delta - \xi^0 \\ &\frac{\bar{\rho}'}{\bar{\rho}}\delta \end{aligned} \tag{3.74}$$

[21] The fact that the soft π is proportional to D' is nicely consistent with $\xi^0 \propto D'$, since $\Delta_{\rm nl.}\pi = \xi^0$ (see Eq. 3.72).

[22] One can use $U^\mu = (1 - \Phi, v^i)/a$, valid to the lowest order in velocity and perturbations, with the understanding that $v^i = dx^i/d\eta$. In this chapter, by relativistic effects, we are generally interested in effects on super-Hubble scales as opposed to effects associated with high peculiar velocities.

One could set $-\bar{\rho}'/\bar{\rho} = 3\mathcal{H}$ for $\bar{\rho}$ that redshifts like pressureless matter, but we will keep the discussion general. The generalization to the galaxy density fluctuation δ_g (or the fluctuation of any component) is immediate:

$$\Delta\delta_g = \Delta_{\text{nl.}}\delta_g + \Delta_{\text{lin.}}\delta_g \quad , \quad \Delta_{\text{nl.}}\delta_g = -\xi^0 \frac{\bar{\rho}_g'}{\bar{\rho}_g} \ , \ \Delta_{\text{lin.}}\delta_g = -\xi^\mu \partial_\mu \delta_g - \xi^0 \frac{\bar{\rho}_g'}{\bar{\rho}_g}\delta_g \, , \tag{3.75}$$

where $\bar{\rho}_g$ is the mean galaxy number density. In both cases, the linear part of the transformations would resemble more what one expects for a scalar if we consider $\delta\rho = \bar{\rho}\delta$ instead of δ:

$$\Delta_{\text{nl.}}\delta\rho = -\xi^0\bar{\rho}' \ , \ \Delta_{\text{lin.}}\delta\rho = -\xi^\mu\partial_\mu\delta\rho \, . \tag{3.76}$$

3.1.4.2 Scalar Consistency Relations

Let us first derive the consistency relations that involve only scalar modes, i.e. where only scalar modes are nonlinearly generated. Recall from sec. 3.1.4.1 that the scalar symmetries take the form:

$$\xi^0 = \xi^0_{\text{add.}} \quad , \quad \xi^i = \xi^i_{\text{unit.}} + \xi^i_{\text{add.}} \, , \tag{3.77}$$

with $\xi^i_{\text{unit.}}$ and $\xi^\mu_{\text{add.}}$ satisfying:

$$\partial_i\xi^j_{\text{unit.}} + \partial_j\xi^i_{\text{unit.}} - \frac{2}{3}\delta_{ij}\partial_k\xi^k_{\text{unit.}} = 0 \, , \tag{3.78}$$

$$\xi^0_{\text{add.}} = -\frac{1}{3c}D'\partial_i\xi^i_{\text{unit.}} \quad , \quad \xi^i_{\text{add.}} = \frac{1}{c}D\nabla^2\xi^i_{\text{unit.}} \, , \tag{3.79}$$

where D is the linear growth factor obeying $D'' + 2\mathcal{H}D' - c = 0$, with c being a constant.

As discussed before, since the unitary gauge is a complete gauge-fixing for diffeomorphisms that vanish at spatial infinity, the residual diffeomorphism of interests must be one where $\xi^i_{\rm unit.}$ does not vanish at infinity. Following H2K, we can express $\xi^i_{\rm unit.}$ as a power series:

$$\xi^i_{\rm unit.} = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} M_{i\ell_0\ldots\ell_n} x^{\ell_0}\ldots x^{\ell_n} \,, \tag{3.80}$$

where each $M_{i\ell_0\ldots\ell_n}$ represents a constant coefficient, symmetric in its last $n+1$ indices. As pointed out by [4, 80], the only scalar symmetries are those associated with $n = 0$: $\xi^i_{\rm unit.} \sim x$ (dilation) and $n = 1$: $\xi^i_{\rm unit.} \sim x^2$ (special conformal transformation).

The Dilation Consistency Relation Dilation is described by $\xi^i_{\rm unit.} = \lambda x^i$ where λ is a constant. Plugging this into Eq. (3.79) tells us that $\xi^0_{\rm add.} = -(\lambda/c)D'$ and $\xi^i_{\rm add.} = 0$. In other words, the net residual diffeomorphism in Newtonian gauge is

$$\xi^0 = \epsilon \quad , \quad \xi^i = \lambda x^i \quad \text{with} \quad \epsilon \equiv -\frac{\lambda}{c} D' \,, \tag{3.81}$$

where λ is a constant. This symmetry involves a spatial dilation plus an accompanying time translation, with the two related by a differential equation: $\epsilon' + 2\mathcal{H}\epsilon + \lambda = 0$. We will refer to the resulting consistency relation simply as the dilation consistency relation, even though the symmetry involves more than spatial dilation. To deduce the associated consistency relation, we employ Eq. (3.27). Two pieces of information are needed to use it. One is the nonlinear shift of π in Fourier space, obtained by taking the Fourier transform of Eq. (3.72):

$$\Delta_{\rm nl.} \pi^*_{\vec{q}} = (2\pi)^3 \delta_D(\vec{q})\, \epsilon \,. \tag{3.82}$$

The other piece of information we need is the linear transformation of the high momentum observable(s). Here, let us use the density fluctuation $\delta_{\vec{k}}$ as the observable at high momentum. By a Fourier transform of Eq. (3.74), we find

$$\Delta_{\rm lin.}\delta_{\vec{k}} = \left[-\epsilon\left(\frac{\bar{\rho}'}{\bar{\rho}} + \partial_\eta\right) + \lambda(3 + \vec{k}\cdot\partial_{\vec{k}})\right]\delta_{\vec{k}}\,. \tag{3.83}$$

Plugging these two pieces into the master equation (3.27), we see that

$$\begin{aligned}&\lim_{\vec{q}\to 0}\frac{\langle\pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^c}{P_\pi(q)}\epsilon(\eta)\\ =&\sum_{a=1}^{N}\left[-\epsilon(\eta_a)\left(\frac{\bar{\rho}'}{\bar{\rho}}\Big|_{\eta_a} + \partial_{\eta_a}\right) + \lambda\left(3 + \vec{k}_a\cdot\partial_{\vec{k}_a}\right)\right]\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^c\,,\end{aligned} \tag{3.84}$$

where the time-dependence should be understood as follows: the soft $\vec{q}$ mode is evaluated at time η, while the hard mode $\vec{k}_a$ is evaluated at time η_a, meaning that each hard mode can be at a different time. This is why the ϵ on the left is at time η – it is associated with the nonlinear shift in π and therefore the soft mode – and the ϵ's on the right are evaluated at the respective η_a, since each is associated with the linear transformation of the corresponding hard mode.

The connected N- and $(N+1)$-point functions on both sides contain the momentum conserving delta function. Its removal requires some care since the derivatives with respect to momentum on the right hand side act on the delta function:

$$\sum_{a=1}^{N}\vec{k}_a\cdot\partial_{\vec{k}_a}\delta_D(\vec{k}_1 + ... + \vec{k}_N) = -3\delta_D(\vec{k}_1 + ... + \vec{k}_N). \tag{3.85}$$

This can be established by rewriting the delta function as $(2\pi)^{-3}\int d^3x\, e^{i(\vec{k}_1+...+\vec{k}_N)\cdot\vec{x}}$, and integrating by parts. Thus, removing the delta function on both sides, with $\langle...\rangle^{c'}$ representing the connected correlation function without δ_D, we have[23]

$$\begin{aligned}&\lim_{\vec{q}\to 0}\frac{\langle \pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\epsilon(\eta)\\ &=\left(3\lambda(N-1)+\sum_{a=1}^{N}\left[-\epsilon(\eta_a)\left(\frac{\bar{\rho}'}{\bar{\rho}}\Big|_{\eta_a}+\partial_{\eta_a}\right)+\lambda\vec{k}_a\cdot\partial_{\vec{k}_a}\right]\right)\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\,,\end{aligned}\tag{3.86}$$

with the understanding that ϵ and (the constant) λ are related by Eq. (3.81), and where the N-point function depends on the time associated with each of the N modes. Using the relation, we can rewrite the *dilation consistency relation* as

$$\boxed{\begin{aligned}&\lim_{\vec{q}\to 0}\frac{\langle \pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\\ &=-\left(\frac{3c}{D'(\eta)}(N-1)+\sum_{a=1}^{N}\left[\frac{D'(\eta_a)}{D'(\eta)}\left(\frac{\bar{\rho}'}{\bar{\rho}}\Big|_{\eta_a}+\partial_{\eta_a}\right)+\frac{c}{D'(\eta)}\vec{k}_a\cdot\partial_{\vec{k}_a}\right]\right)\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\end{aligned}}\tag{3.87}$$

with the understanding that $c = D'' + 2\mathcal{H}D'$ is a constant. It is trivial to generalize the consistency relation by changing the hard modes from δ for the mass density to δ_g for the galaxy density: simply change the mean mass density $\bar{\rho}$ on the right hand side to the mean galaxy number density $\bar{\rho}_g$. This consistency relation can be further rewritten in different forms. We will postpone this discussion until after we discuss the special conformal consistency relation.

The Special Conformal Consistency Relation – Containing the Newtonian Translation Consistency Relation Next, we consider the special conformal transformation: $\xi^i_{\rm unit.} = 2\vec{b}\cdot\vec{x}x^i - b^i\vec{x}\cdot\vec{x}$, where $\vec{b}$ is a constant vector. Plugging this into Eq. (3.79), we see that the requisite accompanying diffeomorphism is: $\xi^0_{\rm add.} = -(2/c)D'\vec{b}\cdot\vec{x}$

[23] Note that $\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}$ should be understood to be a function of only $N-1$ momenta. For instance, we can think of $\vec{k}_N = -\vec{k}_1 - \vec{k}_2... - \vec{k}_N$. Thus the derivative $\partial_{\vec{k}_N}$, keeping $\vec{k}_1, ..., \vec{k}_{N-1}$ fixed, vanishes. This point is particularly important for the higher consistency relations. See [83].

and $\xi^i_{\rm add.} = -(2/c)Db^i$. Putting everything together, we see that the symmetry is:

$$\xi^0 = n^{i\prime}\, x^i \quad , \quad \xi^i = n^i + 2\vec{b}\cdot\vec{x}\, x^i - b^i\, \vec{x}\cdot\vec{x} \quad \text{with} \quad n^i \equiv -\frac{2}{c} Db^i \tag{3.88}$$

where $\vec{b}$ is a constant vector. We refer to the implied consistency relation as the special conformal consistency relation, even though the full symmetry transformation involves a time diffeomorphism and a spatial translation in addition to the special conformal transformation – these transformations are related via $n^{i\prime\prime} + 2\mathcal{H}n^{i\prime} + 2b^i = 0$. As we will see, the Newtonian translation consistency relation is contained in here.

Once again, we employ the master equation (Eq. 3.27), for which we need the nonlinear transformation of the velocity potential π and the linear transformation of the hard modes – as in the case of dilation, we choose the observable to be the density fluctuation δ for the hard modes. Under the current symmetry transformation, we have:

$$\Delta_{\rm nl.}\pi^*_{\vec{q}} = i\vec{n}^{\prime}\cdot\partial_{\vec{q}}[(2\pi)^3\,\delta_D(\vec{q})]\,, \tag{3.89}$$

$$\Delta_{\rm lin.}\delta_{\vec{k}} = i\left[\vec{n}^{\prime}\cdot\partial_{\vec{k}}\left(\frac{\bar{\rho}^{\prime}}{\bar{\rho}} + \partial_\eta\right) + \vec{n}\cdot\vec{k} - (6\vec{b}\cdot\partial_{\vec{k}} + 2b^j k^i \partial_{k^j}\partial_{k^i} - \vec{b}\cdot\vec{k}\,\nabla^2_k)\right]\delta_{\vec{k}}\,. \tag{3.90}$$

Substituting the above into Eq. (3.27), we see that the left hand side (LHS) is

$$\begin{aligned}\text{LHS} \,&= \lim_{\vec{q}\to 0}(-i\vec{n}^{\prime}(\eta))\cdot\partial_{\vec{q}}\left[(2\pi)^3\delta_D(\vec{q}+\vec{k}_1+...+\vec{k}_N)\right]\frac{\langle\pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c\prime}}{P_\pi(q)} \\ &\quad -i(2\pi)^3\delta_D(\vec{q}+\vec{k}_1+...+\vec{k}_N)\vec{n}^{\prime}(\eta)\cdot\partial_{\vec{q}}\left[\frac{\langle\pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c\prime}}{P_\pi(q)}\right]\,. \end{aligned} \tag{3.91}$$

The right hand side (RHS) is

$$
\begin{aligned}
&\text{RHS} \\
=&\sum_{a=1}^{N} i\left[\vec{n}'(\eta_a)\cdot\partial_{\vec{k}_a}\left(\frac{\bar{\rho}'}{\bar{\rho}}\Big|_{\eta_a}+\partial_{\eta_a}\right)+\vec{n}(\eta_a)\cdot\vec{k}_a-(6\vec{b}\cdot\partial_{\vec{k}_a}+2b^i k_a^j\partial_{k_a^j}\partial_{k_a^i}-\vec{b}\cdot\vec{k}_a\,\nabla^2_{k_a})\right] \\
&\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^c\,.
\end{aligned}
\tag{3.92}
$$

The connected N-point function $\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^c$ contains an overall momentum conserving delta function. The momentum derivative acts non-trivially on it. To simplify, it is useful to know that:

$$
\sum_{a=1}^{N}\left(2b^j k_a^i\partial_{k_a^j}\partial_{k_a^i}-\vec{b}\cdot\vec{k}_a\,\nabla^2_{k_a}\right)\delta_D(\vec{k}_{\text{tot.}})=-6\,\vec{b}\cdot\partial_{\vec{k}_{\text{tot.}}}\delta_D(\vec{k}_{\text{tot.}})\,,
\tag{3.93}
$$

where $\vec{k}_{\text{tot}}\equiv\vec{k}_1+...+\vec{k}_N$. This which can be proved by rewriting the delta function as the spatial integral of a plane wave. The term $\sum_a 6\vec{b}\cdot\partial_{\vec{k}_a}\delta_D(\vec{k}_{\text{tot.}})$ can be rewritten as $6N\vec{b}\cdot\partial_{\vec{k}_{\text{tot.}}}\delta_D(\vec{k}_{\text{tot.}})$; indeed, any $\partial_{k_a^i}\delta_D(\vec{k}_{\text{tot.}})$ can be written as $\partial_{k^i_{\text{tot.}}}\delta_D(\vec{k}_{\text{tot.}})$. Then there are terms that involve one momentum derivative on the delta function and one momentum derivative on the N-point function:

$$
\begin{aligned}
&\sum_{a=1}^{N}2b^i k_a^j\,\partial_{k_a^j}\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\,\partial_{k_a^i}\delta_D(\vec{k}_{\text{tot.}}) \\
&+2b^i\partial_{k_a^j}\delta_D(\vec{k}_{\text{tot.}})\left(k_a^j\partial_{k_a^i}\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}-k_a^i\partial_{k_a^j}\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\right) \\
&=\left[2\vec{b}\cdot\partial_{\vec{k}_{\text{tot.}}}\delta_D(\vec{k}_{\text{tot.}})\right]\sum_{a=1}^{N}\vec{k}_a\cdot\partial_{\vec{k}_a}\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\,,
\end{aligned}
\tag{3.94}
$$

where we have used rotational invariance of the N-point function to remove the last two terms on the first line. With this understanding, Eq. (3.92) can be expressed as

$$\text{RHS} = -i\Big[\partial_{k^i_{\text{tot.}}}(2\pi)^3\delta_D(\vec{k}_{\text{tot.}})\Big]\Big[6b^i(N-1) + \sum_{a=1}^{N}\left(-n^{i\prime}(\eta_a)\left(\frac{\bar{\rho}'}{\bar{\rho}}\Big|_{\eta_a} + \partial_{\eta_a}\right) + 2b^i\vec{k}_a\cdot\partial_{\vec{k}_a}\right)\Big]$$
$$\times\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'} - i(2\pi)^3\delta_D(\vec{k}_{\text{tot.}})\sum_{a=1}^{N}\Big[-\left(\frac{\bar{\rho}'}{\bar{\rho}}\Big|_{\eta_a} + \partial_{\eta_a}\right)\vec{n}'(\eta_a)\cdot\partial_{\vec{k}_a} - \vec{n}(\eta_a)\cdot\vec{k}_a$$
$$+ (6\vec{b}\cdot\partial_{\vec{k}_a} + 2b^ik^j_a\partial_{k^j_a}\partial_{k^i_a} - \vec{b}\cdot\vec{k}_a\,\nabla^2_{k_a})\Big]\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}, \tag{3.95}$$

The first term of the above can be equated with the first line of Eq. (3.91), since what multiplies the derivative of the delta function on both sides replicates the dilation consistency relation Eq. (3.86). Note that n^i and b^i are related by Eq. (3.88). Using this, eliminating the dilation consistency relation from both sides,[24] and removing the delta function, we obtain the *special conformal consistency relation*:

$$\boxed{\begin{aligned}\lim_{\vec{q}\to 0}\partial_{q^i}\left[\frac{\langle\pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\right] = -\sum_{a=1}^{N}\Big[\left(\frac{\bar{\rho}'}{\bar{\rho}}\Big|_{\eta_a} + \partial_{\eta_a}\right)\frac{D'(\eta_a)}{D'(\eta)}\partial_{k^i_a} + \frac{D(\eta_a)}{D'(\eta)}k^i_a \\ + \frac{c}{D'(\eta)}\left(3\partial_{k^i_a} + k^j_a\partial_{k^j_a}\partial_{k^i_a} - \frac{1}{2}k^i_a\nabla^2_{k_a}\right)\Big]\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'},\end{aligned}} \tag{3.96}$$

with the understanding that $c = D'' + 2\mathcal{H}D'$ is a constant (Eq. B.50). Just as in the case of the dilation consistency relation, this consistency relation can be easily generalized to the hard modes being the galaxy overdensity – changing δ to δ_g, and changing $\bar{\rho}$ to $\bar{\rho}_g$. Examining the terms on the right hand side, we see that in the sub-Hubble limit, i.e. $k \gg \mathcal{H}$, the term that dominates on the right hand side is $-\sum_a(D(\eta_a)/D'(\eta))k^i_a\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}$, reproducing the translation consistency relation (Eq. 3.35) derived from the Newtonian

[24]This is a general pattern: one can obtain the correct consistency relation for a given symmetry by using the master equation (3.27), *and* simply ignoring the delta functions on both sides. At first sight, this might appear dangerous as there are derivatives acting on the delta functions, but they invariably multiply consistency relations from symmetries at the lower levels, and so can be removed.

equations. We will have more to say about the non-relativistic limit in sec. 3.1.4.3.

At the level of the spatial diffeomorphisms, dilation and special conformal transformations exhaust the list of purely scalar symmetries, since it is only dilation and special conformal transformations that respect the second expression of Eq. (3.78) and do not generate vector or tensor modes. It is also worth noting that the special conformal transformation consistency relation strictly speaking receives (small) corrections on the right hand side, a point to which we will return.

3.1.4.3 Robustness and Limitations of the Consistency Relations - a Relativistic Perspective and the Newtonian Limit

It is useful to pause and reflect on the fully relativistic consistency relations derived so far. Some of our discussions here mirror the earlier ones in the Newtonian context (sec. 3.1.3.1), but with a relativistic twist. We also discuss the issue of taking the Newtonian, i.e. sub-Hubble, limit.

1. Newtonian limit. The special conformal consistency relation Eq. (3.96) is the relativistic analog of the Newtonian translation consistency relation Eq. (3.35). The former reduces to the latter in the sense that:

$$\lim_{\vec{q}\to 0}\frac{\partial}{\partial q_j}\left[\frac{\langle \pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\right] = -\sum_{a=1}^{N}\frac{D(\eta_a)}{D'(\eta)}k_{aj}\,\langle \delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\times\left(1+O(\mathcal{H}^2/k^2)\right)\,, \tag{3.97}$$

where the $\mathcal{H}^2/k^2$-suppressed terms can be ignored in the sub-Hubble limit. Note that the unsuppressed (Newtonian) terms are of the order of $\frac{k}{\mathcal{H}}\langle \delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}$. Similarly, we can think of the dilation consistency relation Eq. (3.87) as the relativistic analog of Eq. (3.33). The dilation consistency relation takes the form:

$$\begin{aligned}\lim_{\vec{q}\to 0}\frac{\langle \pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)} &= O\left(\frac{k}{\mathcal{H}}\langle \delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\right)\times O\left(\mathcal{H}^2/k\right)\\ =&O\left(q\frac{k}{\mathcal{H}}\langle \delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\right)\times O(\mathcal{H}/k)(\mathcal{H}/q)\end{aligned} \tag{3.98}$$

which can be compared against $q\times$ Eq. (3.97). We can see that the right hand side of the above expression is $O(\mathcal{H}/k)O(\mathcal{H}/q)$ times $q\times$ Eq. (3.97). In the sub-Hubble limit where $\mathcal{H}$ is small compared to both q and k, it is therefore consistent to think of Eq. (3.98) as vanishing – reducing to Eq. (3.33).[25]

2. Combining consistency relations. It is worth pointing out that, just as in the Newtonian case where Eqs. (3.33) and (3.35) can be combined into a single equation (3.36), the general relativistic dilation and special conformal consistency relations can be combined into:

$$\boxed{\begin{aligned}&\lim_{\vec{q}\to 0}\frac{\langle \pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\\ &=-\left(\frac{3c}{D'(\eta)}(N-1)+\sum_{a=1}^{N}\left[\frac{D'(\eta_a)}{D'(\eta)}\left(\frac{\bar\rho'}{\bar\rho}\Big|_{\eta_a}+\partial_{\eta_a}\right)\left(1+\vec{q}\cdot\partial_{\vec{k}_a}\right)+\frac{D(\eta_a)}{D'(\eta)}\vec{q}\cdot\vec{k}_a\right.\right.\\ &\left.\left.+\frac{c}{D'(\eta)}\left(\vec{k}_a\cdot\partial_{\vec{k}_a}+3\vec{q}\cdot\partial_{\vec{k}_a}+q^i k_a^j\partial_{k_a^j}\partial_{k_a^i}-\frac{1}{2}\vec{q}\cdot\vec{k}_a\nabla^2_{k_a}\right)\right]\right)\langle \delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'},\end{aligned}} \tag{3.99}$$

where the constant $c = D'' + 2\mathcal{H}D'$.

3. Alternative pions. Recall that π, δ and Φ all shift nonlinearly under the symmetries

[25] Eq. (3.33) was derived using the shift symmetry $\pi\to\pi+b$, where b is a constant. The reader might wonder how that argument breaks down in the relativistic context. The point is that a constant shift in π has to be accompanied by a time-dependent shift in Φ (see e.g. Eq. 3.55). Such a time-dependent shift is not a symmetry of the kinetic term for the metric once time-derivatives are taken into account, unless the coordinates change too. It is interesting to note that $\pi\to\pi+b/\bar\phi'$ is a symmetry (see Appendix B.1) because of the shift symmetry in ϕ; however, this symmetry does not correspond to the growing mode vacuum and therefore does not lead to a consistency relation. See Appendix B.4 for a further discussion.

of interest.[26] One might wonder whether we could have derived the consistency relation with δ or Φ playing the role of the pion instead. The answer is affirmative. Let us compare these nonlinear shifts: $\Delta_{\rm nl.}\pi = \xi^0$, $\Delta_{\rm nl.}\delta = -\xi^0\bar{\rho}'/\bar{\rho}$, and $\Delta_{\rm nl.}\Phi = -\xi^{0\prime} - \mathcal{H}\xi^0$. Recalling that $\xi^0 \propto D'$, we see that $\Delta_{\rm nl.}\delta = -\bar{\rho}'/\bar{\rho} \times \Delta_{\rm nl.}\pi$, and $\Delta_{\rm nl.}\Phi = -(D'' + \mathcal{H}D')/D' \times \Delta_{\rm nl.}\pi$. One can thus run the same arguments as before, and arrive at essentially the same consistency relation Eq. (3.99), with the right hand side unaltered, but the left hand side replaced by

$$\text{LHS} \to -\frac{\bar{\rho}'}{\bar{\rho}}\bigg|_\eta \times \lim_{\vec{q}\to 0} \frac{\langle \delta_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\delta(q)}\,, \tag{3.100}$$

or

$$\text{LHS} \to -\frac{D'' + \mathcal{H}D'}{D'}\bigg|_\eta \times \lim_{\vec{q}\to 0} \frac{\langle \Phi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\Phi(q)}\,. \tag{3.101}$$

The consistency relations expressed using π, δ or Φ as the soft pion are all equivalent – with one important caveat, which is related to the squeezing constraint.

4. Squeezing constraint. The reader might want to consult sec. 3.1.3.1 for a parallel discussion of the squeezing constraint in the Newtonian context. The consistency relation, whether expressed in terms of π as in Eq. (3.99), or expressed in terms of δ or Φ as in Eq. (3.100) or (3.101), is a statement about the $\vec{q} \to 0$ limit. In the relativistic context, this means that, in addition to $q \ll k_1, ..., k_N$, the soft mode should strictly speaking be super-Hubble, i.e. $q < \mathcal{H}$. On the other hand, in LSS we are typically interested in sub-Hubble modes, so the question arises: under what conditions does the consistency relation remain valid when all modes, including the soft one, are sub-Hubble (while maintaining the hierarchy $q \ll k_1, ..., k_N$)?

[26] We could also discuss the nonlinear shift of δ_n, which coincides with δ for pressureless matter. (See Appendix B.3.)

We show in Appendix B.3 a special fact about the velocity potential π for pressureless matter: it has the same time dependence $\propto D'$ (Eqs. B.49 and B.50) regardless of whether the mode of interest is inside or outside the Hubble radius. (When it is outside the Hubble radius, this statement is true throughout the entire history of the universe; when it is inside the Hubble radius, this statement is strictly true only after radiation domination.) The consistency relation written in terms of the matter π Eq. (3.99) can therefore be safely taken inside the Hubble radius, even for the soft mode, since the same diffeomorphism is capable of generating the correct π regardless of whether it is on super- and sub-Hubble scales. In this limit, we recover the Newtonian consistency relation: the term $-\sum_a [D(\eta_a)/D'(\eta)]\vec{q}\cdot\vec{k}_a \langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}$ dominates on the right hand side. [27] Similar statements hold for Φ as the soft pion.

The same is not true for δ (here, we focus on the *matter* δ as the soft mode): from the continuity equation (B.28), it is evident that the time dependence of δ (which is the same as δ_n for pressureless matter) depends on whether the mode is inside or outside the Hubble radius. The consistency relation written using δ as the soft mode takes the form of Eq. (3.100) only for $q < \mathcal{H}$. If the soft δ mode is within the horizon, the continuity equation tells us that $\delta_{\vec{q}} = q^2(D/D')\pi_{\vec{q}}$, and so the left hand side of the consistency relation should read:

$$\text{LHS} \to \lim_{\vec{q}\to 0} q^2 \frac{D(\eta)}{D'(\eta)} \times \frac{\langle\delta_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\delta(q)}\,, \tag{3.102}$$

[27] There is one subtlety though: for a mode that enters the Hubble radius during radiation domination, the time evolution deviates from D' during part of its history, and so strictly speaking the consistency relation does not apply if the soft mode belongs to this category. An alternative way to put it is this: when the mode is within the Hubble radius (or more precisely, within the sound horizon) during radiation domination, neither the matter nor the radiation moves with a velocity that agrees with D'. A diffeomorphism that obeys the adiabatic mode conditions (e.g. Eq. (3.81) for dilation, or Eq. (3.88) for special conformal transformation) cannot generate the correct velocity for either component. Even in this case, we expect the consistency relation to still be a good approximation in the late universe, to the extent that most of the late-time non-Gaussianity is generated after radiation domination. We thank Paolo Creminelli for discussions on this point.

while, as discussed above, the right hand side reduces to $-\sum_a [D(\eta_a)/D'(\eta)]\vec{q}\cdot\vec{k}_a \langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}$. This reproduces the Newtonian translation consistency relation written in terms of $\delta_{\vec{q}}$ (Eq. (3.38)). To conclude: the consistency relation expressed in terms of a soft $\delta_{\vec{q}}$ takes a different form outside versus inside the Hubble radius; i.e. Eq. (3.100) versus Eq. (3.102). The consistency relation expressed using the matter π or Φ as the soft pion maintains the same form regardless.[28]

5. The existence of an interesting Newtonian limit. From the discussion above, we see that the special conformal consistency relation has a non-trivial Newtonian limit (i.e. the right hand side is non-vanishing), whereas the dilation one does not. What is the underlying reason? From Eq. (3.79), we see that for a given unitary-gauge transformation $\xi^i_{\rm unit.}$, the corresponding residual diffeomorphism in Newtonian gauge is

$$\xi^0 \sim \mathcal{H}^{-1}\partial_i \xi^i_{\rm unit.} \quad , \quad \xi^i \sim \xi^i_{\rm unit.} + \mathcal{H}^{-2}\nabla^2\xi^i_{\rm unit.} \,. \tag{3.103}$$

The associated consistency relation, making use of the relation $\delta_{\vec{q}} \sim q^2\pi_{\vec{q}}/\mathcal{H}$ in the sub-Hubble limit, can be written schematically as:

$$\lim_{\vec{q}\to 0} \frac{q^2}{\mathcal{H}^2}\frac{\langle\delta_{\vec{q}}\delta_{\vec{k}}...\rangle^{c'}}{P_\delta(q)}\left[\partial_i\xi^i_{\rm unit.}\right]_q \sim \left(\left[\partial_i\xi^i_{\rm unit.}\right]_k + k_i\left[\xi^i_{\rm unit.}\right]_k + k_i\mathcal{H}^{-2}\left[\nabla^2\xi^i_{\rm unit.}\right]_k\right)\langle\delta_{\vec{k}}...\rangle^{c'} \tag{3.104}$$

Here, $[\ ]_k$ denotes the Fourier transform of the quantity of interest at momentum k, with the delta function removed. For instance, for $\xi^i_{\rm unit.} \sim x^{n+1}$, we have $[\xi^i_{\rm unit.}]_k \sim k^{-n-1}$, $[\partial_i\xi^i_{\rm unit.}]_k \sim k^{-n}$, $[\partial_i\xi^i_{\rm unit.}]_q \sim q^{-n}$, and $[\nabla^2\xi^i_{\rm unit.}]_k \sim k^{-n+1}$. This gives

$$\lim_{\vec{q}\to 0}\frac{\langle\delta_{\vec{q}}\delta_{\vec{k}}...\rangle^{c'}}{P_\delta(q)} \sim \left(\frac{q}{k}\right)^n\left(\frac{\mathcal{H}^2}{q^2}+\frac{k^2}{q^2}\right)\langle\delta_{\vec{k}}...\rangle^{c'} \,. \tag{3.105}$$

In the sub-Hubble (and squeezed) limit where $\mathcal{H} \ll q \ll k$, this suggests we have the

[28]Using the baryon π as the soft pion is permissible too, as long as one stays above the Jeans scale.

dimensionless ratio $\langle\delta_{\vec{q}}\delta_{\vec{k}}...\rangle^{c'}/(P_\delta(q)\langle\delta_{\vec{k}}...\rangle^{c'}) \sim (q/k)^n(k/q)^2$. For $n = 1$, the special conformal case, this reproduces correctly the Newtonian translation consistency relation. For $n = 0$, the dilation case, this does not work. The reason is that the k^2/q^2 term in Eq. (3.105), which is the dominant term in the sub-Hubble limit, originates from $\nabla^2\xi^i_{\rm unit.}$, which vanishes for the dilation $\xi^i_{\rm unit.} = \lambda x^i$. Our naïve power-counting argument also suggests there could be additional $n > 1$ consistency relations that are non-trivial in the Newtonian limit. As we will see in the next section, the $n > 1$ consistency relations generally involve tensors, which complicates taking the squeezed mode to within the Hubble radius.

Let us close this section by emphasizing the robustness of the consistency relations. As in the Newtonian derivation, the general relativistic derivation makes no assumptions about the dynamics of the hard modes – all we need to know is how they transform under diffeomorphisms. Thus, we expect the consistency relations to hold even for nonlinear, or astrophysically messy, hard modes (though the right hand side of the consistency relations might need to be modified depending on exactly how the modes of interest transform; (see the comments after Eq. (3.87) and in footnote 30). Besides the existence of symmetries, which according to the general relativistic perspective are nothing but residual diffeomorphisms, the two key assumptions are the same as in the Newtonian derivation: single field initial condition and adiabatic mode conditions, in particular that all species move with the same velocity in the soft limit.

3.1.4.4 Consistency Relations Involving Tensor Modes

In this section, we move beyond dilation and special conformal transformation to discuss residual diffeomorphisms that generate tensor modes (with or without accompanying scalar modes). We apply the same strategy as the one used for the pure scalar symmetries:

use the full set of symmetries derived in the unitary gauge by H2K, and map each to a symmetry in the Newtonian gauge by Eq. (3.70).

The unitary gauge residual diffeomorphisms can be written as (Eq. 3.80):

$$\xi^i_{\rm unit.} = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} M_{i\ell_0\ldots\ell_n} x^{\ell_0} \ldots x^{\ell_n} \,, \tag{3.106}$$

where the constant coefficients M satisfy:

$$M_{i\ell\ell\ell_2\cdots\ell_n} = -\frac{1}{3} M_{\ell i\ell\ell_2\cdots\ell_n} \,. \tag{3.107}$$

This condition is derived by substituting the power series into Eq. (3.64): $\nabla^2\xi^i_{\rm unit.} + \partial_i(\partial_k\xi^k_{\rm unit.})/3 = 0$. Note that M by definition is symmetric in its last $n+1$ indices. Since we are interested in M that generates tensor modes (in addition to possibly scalar modes), we should impose an additional *adiabatic transversality condition*:

$$\hat{q}^i \left[M_{ij\ldots}(\hat{q}) + M_{ji\ldots}(\hat{q}) - \frac{2}{3} M_{kk\ldots}(\hat{q})\delta_{ij} \right] = 0 \,. \tag{3.108}$$

This condition can be understood as requiring that the tensor generated by our diffeomorphism be extensible to the $\vec{q} \to 0$ limit of a transverse physical tensor mode. Imagine an M that is nearly constant but tapers off to zero at sufficiently large x. While a constant M yields (derivatives of) a delta function peaked at $\vec{q} = 0$ in Fourier space, a tapering M yields a smoothed out version thereof. A tensor mode at a small but finite momentum should be transverse to its own momentum. We demand that even as we take the $\vec{q} = 0$ limit (allowing the tapering of M to occur at larger and larger distances), transversality continues to hold, keeping the direction $\hat{q}$ fixed. This is the content of Eq. (3.108). The choice of $\hat{q}$ is arbitrary; one could for instance choose it to point in the z direction. In addition, if the diffeomorphism generates *only* tensor modes, further conditions on M

come from Eq. (3.65): $\partial_i \xi^i_{\rm unit.} = 0$ and $\nabla^2 \xi^i_{\rm unit.} = 0$. This implies

$$M_{\ell\ell\ell_1\ell_2\cdots\ell_n} = 0 \quad \textit{and} \quad M_{i\ell\ell\ell_2\cdots\ell_n} = 0\,, \tag{3.109}$$

i.e. that M is traceless over any pairs of indices, which also trivially satisfies Eq. (3.107).

As discussed in sec. 3.1.4.1, for each unitary gauge residual diffeomorphism, there is a corresponding one in Newtonian gauge:

$$\xi^\mu = \xi^\mu_{\rm unit.} + \xi^\mu_{\rm add.} \quad \text{where} \quad \xi^0_{\rm add.} = -\frac{1}{3c} D' \partial_i \xi^i_{\rm unit.} \quad , \quad \xi^i_{\rm add.} = \frac{1}{c} D \nabla^2 \xi^i_{\rm unit.}\,, \tag{3.110}$$

with D being the linear growth factor and where c is a constant satisfying $D'' + 2\mathcal{H}D' - c = 0$. Thus, at level n, the Newtonian gauge diffeomorphism is:

$$\boxed{\begin{aligned} \xi^0 &= -\frac{D'}{3cn!} M_{\ell\ell\ell_1\ldots\ell_n} x^{\ell_1} \ldots x^{\ell_n} \\ \xi^i &= \frac{1}{(n+1)!} M_{i\ell_0\ell_1\ldots\ell_n} x^{\ell_0} \ldots x^{\ell_n} + \frac{D}{c(n-1)!} M_{i\ell\ell\ell_2\ldots\ell_n} x^{\ell_2} \ldots x^{\ell_n} \end{aligned}}\ . \tag{3.111}$$

This expression holds for all n, with the exception of $n = 0$, in which case the last term on the right for ξ^i is absent. One thing which is immediately clear is that for purely tensor symmetries — diffeomorphisms that generate only tensor modes — we have $\xi^0_{\rm add.} = \xi^i_{\rm add.} = 0$, and so they are identical in the unitary gauge and the Newtonian gauge, as expected. It is worth emphasizing that Eqs. (3.111) and (3.107) are general: they apply to symmetries that generate only scalar modes, or only tensor modes, or both. For symmetries that generate tensor modes, the adiabatic transversality expressed in Eq. (3.108) is an additional requirement, and Eq. (3.109) applies if *only* tensor modes are generated.

To derive the corresponding consistency relations, we need a master equation analogous to Eq. (3.27) but generalized to allow for the possibility of tensor modes:

$$\int \frac{d^3q}{(2\pi)^3} \left[\frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\pi(q)} \Delta_{\rm nl.} \pi^*_{\vec{q}} + \sum_s \frac{\langle \gamma^s_{\vec{q}} \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\gamma(q)} \Delta_{\rm nl.} \gamma^{s*}_{\vec{q}} \right]$$
$$= \Delta_{\rm lin.} \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^c \,. \tag{3.112}$$

Here, the label s denotes one of the two possible tensor polarization states; a given tensor perturbation $\gamma_{ij}(\vec{q})$ can be decomposed as $\gamma_{ij}(\vec{q}) = \sum_s \epsilon^s_{ij}(\hat{q}) \gamma^s_{\vec{q}}$, where the symmetric traceless polarization tensor $\epsilon^s_{ij}(\hat{q})$ obeys $\hat{q}^i \epsilon^s_{ij}(\hat{q}) = 0$ and $\epsilon^s_{ij}(\hat{q}) \epsilon^{s'}_{ij}(\hat{q})^* = 2\delta^{ss'}$. The tensor power spectrum is defined by $\langle \gamma^s_{\vec{q}} \gamma^{s'}_{\vec{q}'} \rangle = (2\pi)^3 \delta_D(\vec{q} + \vec{q}') \delta^{ss'} P_\gamma(q)$.

$$\Delta_{\rm nl.} \gamma_{ij} = -\left(\partial_i \xi^j + \partial_j \xi^i - \frac{2}{3} \delta_{ij} \partial_k \xi^k \right) \,. \tag{3.113}$$

Eq. (3.112) can alternatively be written as:

$$\boxed{\begin{aligned} &\int \frac{d^3q}{(2\pi)^3} \left[\frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\pi(q)} \Delta_{\rm nl.} \pi^*_{\vec{q}} + \frac{1}{2} \frac{\langle \gamma_{ij}(\vec{q}) \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\gamma(q)} \Delta_{\rm nl.} \gamma_{ij}(\vec{q})^* \right] \\ &= \Delta_{\rm lin.} \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^c \end{aligned}} \tag{3.114}$$

using the fact that $\Delta_{\rm nl.} \gamma^{s*}_{\vec{q}} = \Delta_{\rm nl.} \gamma_{ij}(\vec{q})^* \epsilon^s_{ij}(\hat{q})/2$. Let us step through a few low n examples to get a feel for the kind of consistency relations that arise from these diffeomorphisms. The discussion follows that of H2K, with suitable deformations to the Newtonian gauge.

For $n = 0$, $M_{i\ell_0}$ can be written as the sum of a trace (dilation), an antisymmetric part (which does not generate a nonlinear shift in the metric[29]) and a symmetric traceless part (anisotropic rescaling, which generates tensor perturbations). Note that $n = 0$ is a special case, in the sense that the last term of Eq. (3.111) does not exist (because $\nabla^2 \xi^i_{\rm unit.} = 0$).

[29] Such a transformation would correspond to a time-independent rotation. See Appendix B.4 though for the corresponding decaying mode, which although visible corresponds to a nonstandard choice of initial conditions.

Focusing on a symmetric traceless $M_{i\ell_0}$, there are five independent components. Imposing the adiabatic transversality condition $\hat{q}^i M_{i\ell_0}(\hat{q}) = 0$ reduces the number of independent tensor modes to two. Thus, at the level of $n = 0$, we have one pure scalar and two pure tensor symmetries. It is straightforward to infer the corresponding symmetries in Newtonian gauge: Dilation gets deformed as discussed in sec. 3.1.4.2; and the purely tensor symmetries take exactly the same form in the two gauges. The $n = 0$ anisotropic rescaling tensor consistency relation reads:

$$\lim_{\vec{q}\to 0} \frac{\langle \gamma^s_{\vec{q}} \delta_{\vec{k}_1} ... \delta_{\vec{k}_N} \rangle^{c'}}{P_\gamma(q)} = -\frac{1}{2} \epsilon^s_{ij}(\hat{q})^* \sum_{a=1}^{N} k^i_a \partial_{k^j_a} \langle \delta_{\vec{k}_1} ... \delta_{\vec{k}_N} \rangle^{c'} \,, \tag{3.115}$$

a relation that was first pointed out by Maldacena [3]. To derive this, we use $\Delta_{\rm nl.} \gamma_{i\ell_0}(\vec{q})^* = -2M_{i\ell_0}(2\pi)^3 \delta_D(\vec{q})$, which gives $\Delta_{\rm nl.} \gamma^{s*}_{\vec{q}} = -2(2\pi)^3 \delta_D(\vec{q})$ if we choose $M_{i\ell_0} = \epsilon^s_{i\ell_0}(\hat{q})^*$. The linear transformation of δ under the diffeomorphism $\xi^i = M_{i\ell_0} x^{\ell_0}$ is $\Delta_{\rm lin.} \delta_{\vec{k}} = M_{i\ell_0} k^i \partial_{k^{\ell_0}} \delta_{\vec{k}}$.

For $n = 1$, there are three purely scalar symmetries (the special conformation transformations) and four purely tensor ones. The special conformal transformations correspond to

$$M_{i\ell_0\ell_1} = 2(b^{\ell_1} \delta_{i\ell_0} + b^{\ell_0} \delta_{i\ell_1} - b^i \delta_{\ell_0\ell_1}) \,. \tag{3.116}$$

This is manifestly symmetric between ℓ_0 and ℓ_1, and satisfies Eq. (3.107). We see that plugging this into Eq. (3.111) reproduces Eq. (3.88), and thus the special conformal consistency relation of Eq. 3.96 follows. Each of the four tensor symmetries come from an $M_{i\ell_0\ell_1}$ that is symmetric between ℓ_0 and ℓ_1, fully traceless over any pair of indices, and transverse in the sense of Eq. (3.108). The corresponding $n = 1$ tensor consistency relation reads:

$$\lim_{\vec{q}\to 0} M_{i\ell_0\ell_1}\frac{\partial}{\partial q^{\ell_1}}\left[\frac{\langle\gamma_{i\ell_0}(\vec{q})\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\gamma(q)}\right] = -M_{i\ell_0\ell_1}\sum_{a=1}^{N}\left[\frac{1}{2}k_a^i\frac{\partial^2}{\partial k_a^{\ell_0}\partial k_a^{\ell_1}}\right]\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\,, \quad (3.117)$$

where the dependence on $M_{i\ell_0\ell_1}$ can be removed by applying suitable projectors (see H2K).

For each $n \geq 2$, there are four purely tensor symmetries and two mixed symmetries where both π and γ transform nonlinearly. In general, any $n \geq 0$ consistency relation reads:

$$\begin{aligned}
&\lim_{\vec{q}\to 0} M_{i\ell_0\ell_1...\ell_n}\frac{\partial^n}{\partial q^{\ell_1}...\partial q^{\ell_n}}\left[\frac{\langle\gamma_{i\ell_0}(\vec{q})\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\gamma(q)} + \delta_{i\ell_0}\frac{D'(\eta)}{3c}\frac{\langle\pi_{\vec{q}}\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}}{P_\pi(q)}\right] \\
&= -M_{i\ell_0\ell_1...\ell_n}\sum_{a=1}^{N}\Bigg[\delta_{i\ell_0}\frac{\partial^n}{\partial k_a^{\ell_1}...\partial k_a^{\ell_n}} - \delta_{i\ell_0}\frac{\delta_{n0}}{N} + \frac{k_a^i}{n+1}\frac{\partial^{n+1}}{\partial k_a^{\ell_0}...\partial k_a^{\ell_n}} \\
&\quad + \delta_{i\ell_0}\frac{D(\eta_a)'}{3c}\left(\frac{\bar{\rho}'}{\bar{\rho}}\Big|_{\eta_a} + \partial_{\eta_a}\right)\frac{\partial^n}{\partial k_a^{\ell_1}...\partial k_a^{\ell_n}} \\
&\quad - \frac{n}{c}\delta_{\ell_0\ell_1}\left(D(\eta_a) - D(\eta)(1-\delta_{n1})\right)k_a^i\frac{\partial^{n-1}}{\partial k_a^{\ell_2}...\partial k_a^{\ell_n}}\Bigg]\langle\delta_{\vec{k}_1}...\delta_{\vec{k}_N}\rangle^{c'}\,.
\end{aligned} \quad (3.118)$$

This is our most general result: each Newtonian gauge residual diffeomorphism described by Eq. (3.111) gives rise to a consistency relation given by Eq. (3.118).[30] The consistency relations can also be written in a form in which the matrix M is projected out (see H2K). The familiar dilation and special conformal consistency relations are contained here and the M's in those cases take a form that projects out the tensor term on the left hand side. Note that the soft modes are assumed to be at time η, while the hard modes are

[30] With the exception of the dilation consistency relation ($n = 0$ with $M_{i\ell_0} \propto \delta_{i\ell_0}$), these consistency relations in general receive corrections on the right hand side which either involve replacing one of the hard modes by a hard (scalar or tensor) metric perturbation, or involve higher powers of the metric perturbations. These corrections arise because the associated diffeomorphisms generally need to be corrected order by order in metric perturbations (H2K, [10]). What we focus on in this chapter are the lowest order terms in the diffeomorphisms (i.e. metric-independent contributions). Even in the nonlinear regime where density perturbations are large, the metric perturbations are in general small. Thus, the corrections to the consistency relations are negligible in applications where the hard modes are density (as opposed to metric) perturbations on sub-Hubble scales.

at time η_a for each momentum $\vec{k}_a$. Purely tensor consistency relations follow from those M's that are fully traceless. Hence we set the scalar contributions on the left hand side to be zero (and zeroing out terms proportional to the Kronecker delta on the right hand side as well). For $n \geq 2$, there are choices of M (two for each $n \geq 2$) that have a structure that gives both non-vanishing tensor and scalar contributions on the left hand side. It is worth pointing out that on the right hand side, the first set of terms (the second line) are time-independent; they originate from the unitary gauge diffeomorphisms. [31] The second set of terms (the third line) originate from $\xi^0_{\rm add.}$, the additional time diffeomorphism that is necessary to keep us in Newtonian gauge. Likewise, the last set of terms (the fourth line) come from $\xi^i_{\rm add.}$, and we have used the pure tensor consistency relation at level $(n-2)$ to move the terms proportional to $D(\eta)$ to the right hand side.

Let us study the taking of the Newtonian, i.e. sub-Hubble, limit. As explained in sec. 3.1.4.3, it is helpful to rewrite the consistency relations using $\delta_{\vec{q}} \sim q^2\pi_{\vec{q}}/\mathcal{H}$ (the precise relation is $\delta_{\vec{q}} = q^2\pi_{\vec{q}} D/D'$ for the δ and π of pressureless matter in the sub-Hubble limit). Recalling that $c = D'' + 2\mathcal{H}D' \sim \mathcal{H}^2$, we see that Eq. (3.118) naïvely has a sub-Hubble limit of the schematic form:

$$\lim_{\vec{q}\to 0}\left\{\frac{\mathcal{H}^2}{q^2}\frac{\langle\gamma_{\vec{q}}\delta_{\vec{k}}...\rangle^{c'}}{P_\gamma(q)} + \frac{\langle\delta_{\vec{q}}\delta_{\vec{k}_1}...\rangle^{c'}}{P_\delta(q)}\right\} \sim \left(\frac{q}{k}\right)^n \left(\frac{\mathcal{H}^2}{q^2} + n\frac{k^2}{q^2}\right)\langle\delta_{\vec{k}}...\rangle^{c'}, \tag{3.119}$$

where we have equated $\partial_q \sim 1/q$ and $\partial_k \sim 1/k$. Of the terms on the right hand side, the term suppressed in the sub-Hubble limit by $(\mathcal{H}^2/q^2)$ arises from the second and third lines of Eq. (3.118), and the unsuppressed term comes from the last line of Eq. (3.118). It is also worth noting that the unsuppressed term is in general non-vanishing even if all the hard modes are at the same time, as long as the soft mode is at a different time; the $n = 1$ case (that gives rise to the KRPP consistency relation) is an exception rather than

[31] The term $-\delta_{i\ell_0}\delta_{n0}/N$ arises from the removal of delta functions. See H2K for discussion.

the rule.

At first sight, this suggests that there is a non-trivial Newtonian limit for each $n > 0$, with the $n = 1$ case (KRPP) being one example. This is not the case because of the presence of tensor modes. In all $n \geq 2$ cases where the diffeomorphism generates a soft scalar, the same diffeomorphism generates a soft tensor as well. The tensor equation of motion $\gamma''_{ij} + 2\mathcal{H}\gamma'_{ij} - \nabla^2\gamma_{ij} = 0$ (Eq. B.36) tells us that (1) ignoring the $\nabla^2\gamma_{ij}$ term, then $\gamma_{ij} =$ const. is the growing mode solution (or more properly, the dominant mode solution; the other mode decays); (2) allowing for a small ∇^2, the growing mode tensor solution gets corrected by a term proportional to D (see Appendix B.2); (3) when ∇^2 is important, the tensor mode oscillates with an amplitude that decays as $1/a$. Cases (1) and (2) pertain to super-Hubble modes while case (3) has to do with sub-Hubble ones. The purely tensor consistency relations follow from diffeomorphisms that are time-independent and generate tensor modes of type (1). The mixed scalar-tensor consistency relations follow from diffeomorphisms that generate tensor modes of type (2) (see footnote 4). In neither case are we allowed to take the soft tensor mode to within the Hubble radius. This is in contrast with the purely scalar consistency relations (such as dilation and special conformal transformation), where the time-dependence of the soft $\pi_{\vec{q}}$ remains the same whether it is outside or inside the Hubble radius. One might be tempted to say: within the Hubble radius, the tensor mode decays anyway, so why not just drop the tensor term from the consistency relations? This is not allowed because the tensor mode enters in both the numerator and denominator of $\langle\gamma_{\vec{q}}...\rangle^{c'}/P_\gamma(q)$. In general this ratio is independent of the amplitude of the tensor mode; the consistency relations express precisely this fact.

3.2 Lagrangian space consistency relations

3.2.1 Introduction

Consistency relations are statements which relate the squeezed limit of an (N+1)-point correlation function to an N-point function of cosmological perturbations; i.e., they take the following schematic form in momentum space:

$$\lim_{\mathbf{k}\to 0} \frac{\langle \pi_{\mathbf{k}} \mathcal{O}_{\mathbf{k_1}} \mathcal{O}_{\mathbf{k_2}} ... \mathcal{O}_{\mathbf{k_N}} \rangle^{c'}}{P_\pi(\mathrm{k})} \sim \langle \mathcal{O}_{\mathbf{k_1}} \mathcal{O}_{\mathbf{k_2}} ... \mathcal{O}_{\mathbf{k_N}} \rangle^{c'} \,, \tag{3.120}$$

where $\pi_{\mathbf{k}}$ represents a squeezedmode (long wavelength) of what turns out to be a Goldstone boson or pion, $P_\pi(\mathrm{k})$ is the power spectrum of the pion (k represents the magnitude of the vector $\mathbf{k}$), and $\mathcal{O}$ represents observables at high momenta $\mathbf{k_1}, ..., \mathbf{k_N}$. The symbol $\langle ... \rangle^{c'}$ denotes the connected correlation function with the overall delta function removed. Consistency relations can be understood as analogues of 'soft-pion' theorems in particle physics, which arise generally when a symmetry is spontaneously broken/nonlinearly realized. In the case of cosmology, the symmetries in question are diffeomorphisms (i.e. coordinate transformations), and consistency relations arise from a particular set of residual symmetries of a given gauge where the transformation does not fall off at infinity. The first example of a consistency relation was pointed out by Maldacena [3] in the context of a computation of the three-point correlation function from inflation. The utility of this as a test of single field/clock inflation was emphasized by Creminelli and Zaldarriaga [79]. Recent work pointed out new symmetries and therefore further consistency relations [4, 80], indeed an infinite tower of them [5], and explicated their non-perturbative nature [6–8, 81–83].

These consistency relations are extremely robust: they remain valid when the high momentum modes ($\mathcal{O}$ in Eq. 3.120) are deep in the nonlinear regime, and even when the observables are astrophysically complex (such as galaxy density). This point might appear academic when applied to (small) perturbations in the early universe, such as are revealed in the cosmic microwave background. When applied to large scale structure (LSS) in the late universe, however, the robustness of the consistency relations becomes very interesting. It thus came as welcome news when Kehagias/Riotto [8] and Peloso/Pietroni [9] (KRPP) pointed out that non-trivial consistency relations exist even if all modes (including the squeezed one) are within the Hubble radius, within the Newtonian regime which is the realm of LSS (see also [9–11, 91, 92, 94, 100, 101]).

The KRPP consistency relation can be stated in the following form:

$$\lim_{\mathbf{k}\to 0} \frac{\langle v^j_{\mathbf{k}}(\eta)\, \mathcal{O}_{\mathbf{k}_1}(\eta_1) \ldots \mathcal{O}_{\mathbf{k}_N}(\eta_N)\, \rangle^{c'}}{P_v(\mathrm{k},\eta)} = i\mathrm{k}^j \sum_{a=1}^{N} \frac{D(\eta_a)}{D'(\eta)} \frac{\mathbf{k}\cdot\mathbf{k}_a}{\mathrm{k}^2} \langle \mathcal{O}_{\mathbf{k}_1}(\eta_1) \ldots . \mathcal{O}_{\mathbf{k}_N}(\eta_N)\, \rangle^{c'} , \tag{3.121}$$

where $v^j_{\mathbf{k}}$ is the j-th component of the peculiar velocity in momentum space, and P_v is the velocity power spectrum defined by $\langle v^i_{\mathbf{k}}(\eta)\, v^{j}_{\mathbf{k}'}{}^*(\eta)\rangle = (2\pi)^3 \delta_D(\mathbf{k}-\mathbf{k}')(\mathrm{k}^i\mathrm{k}^j/\mathrm{k}^2) P_v(\mathrm{k},\eta)$.[32] The observables can be thought of as mass or galaxy overdensity at different momenta and times, and D and D' represent the linear growth factor and its conformal time derivative. The fluctuation variables will in general depend on time, although we will often suppress the time dependence to simplify the notation: $v^j_{\mathbf{k}}$ (and its power spectrum) is at conformal time η, $\mathcal{O}_{\mathbf{k}_1}$ is at time η_1, and so on. The times need not be equal. The symbol k^2 denotes $\mathbf{k}\cdot\mathbf{k}$.

We wish to show that the KRPP consistency relation takes a particularly simple form in Lagrangian space:

$$\boxed{\lim_{\mathbf{p}\to 0} \frac{\langle \boldsymbol{v}_{\mathbf{p}}(\eta)\, \mathcal{O}_{\mathbf{p}_1}(\eta_1) \ldots \mathcal{O}_{\mathbf{p}_N}(\eta_N)\, \rangle^{c'}}{P_v(\mathrm{p},\eta)} = 0\,.} \tag{3.122}$$

[32] This form of the velocity power spectrum assumes no vorticity. This is acceptable since $P_v(\mathrm{k})$ is used only for small k, or large scales, where the growing mode initial condition ensures gradient flow.

Unless otherwise stated, we use $\mathbf{p}$ to denote momentum in Lagrangian space and $\mathbf{k}$ to denote momentum in Eulerian space. In other words:

$$\mathcal{O}_{\mathbf{k}} = \int d^3\mathbf{x}\, \mathcal{O}(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}} \quad , \quad \mathcal{O}_{\mathbf{p}} = \int d^3\mathbf{q}\, \mathcal{O}(\mathbf{x}(\mathbf{q})) e^{i\mathbf{p}\cdot\mathbf{q}} \,, \tag{3.123}$$

where $\mathbf{x}$ and $\mathbf{q}$ are the Eulerian space and Lagrangian space coordinates respectively.[33] In both cases, we rely on context to distinguish between $\mathcal{O}$ in Fourier space and $\mathcal{O}$ in configuration space.

Since the velocity $\boldsymbol{v}$ (whose j-th component is v^j) is nothing other than the time derivative of the displacement $\boldsymbol{\Delta}$ in Lagrangian space, we can also rewrite the Lagrangian space consistency relation as:

$$\boxed{\lim_{\mathbf{p}\to 0} \frac{\langle \boldsymbol{\Delta}_{\mathbf{p}}(\eta)\, \mathcal{O}_{\mathbf{p}_1}(\eta_1) \ldots \mathcal{O}_{\mathbf{p}_N}(\eta_N) \rangle^{c'}}{P_\Delta(\mathrm{p}, \eta)} = 0 \,,} \tag{3.124}$$

where the power spectrum of displacement is defined by

$$\langle \Delta^i_{\mathbf{p}} \Delta^j_{\mathbf{p}'}{}^* \rangle = (2\pi)^3 \delta_D(\mathbf{p} - \mathbf{p}') (\mathrm{p}^i \mathrm{p}^j / \mathrm{p}^2) P_\Delta(\mathrm{p}) \tag{3.125}$$

It is important to emphasize that the Eulerian space consistency relation (Eq. 3.121) already yields a vanishing right hand side if $\eta_1 = \eta_2 ... = \eta_N$. The Lagrangian space consistency relation (Eq. 3.122 or 3.124), on the other hand, has a vanishing right hand side *regardless* of what the times $\eta_1, ..., \eta_N$ happen to be. The consistency relation can also be viewed as a statement about how the squeezed correlation function (normalized by the soft power spectrum) scales with the soft momentum: the Eulerian space consistency relation states that such a squeezed correlation function goes like k^0 ($\mathbf{k}$ is the soft momentum); the Lagrangian space consistency relation states that there is no p^0 term,

[33] The definitions given apply even in the presence of multiple streaming. See discussion in sec. 3.2.2.1.

and at best there is a p^ϵ contribution with $\epsilon > 0$.

The simplest way to derive Eq. (3.124) is to work out the implications of the KRPP symmetry entirely within Lagrangian space. This is done in Sec. 3.2.2. We perform a perturbative check of this Lagrangian space consistency relation using Lagrangian perturbation theory in Sec. 3.2.3.1. Because the Eulerian space and the Lagrangian space relations look so different, as a further check, we show how one can be obtained from the other in Sec. 3.2.3.2. Since observations are performed in Eulerian, not Lagrangian, space, the fact that the consistency relation takes a particularly simple form in Lagrangian space is mainly of theoretical interest. The simplicity of the Lagrangian space consistency relation should not be interpreted as the lack of physical content, however – in the Lagrangian as well as in the Eulerian picture, the consistency relation can be viewed as a test of the single-field initial condition and of the equivalence principle. Rather, the simplicity suggests that an analytical understanding of nonlinear clustering might be most promising in Lagrangian space. This will be discussed in Chapter Four.

3.2.2 The Lagrangian space consistency relation: derivation

After a brief review of notation, we derive our main result – the Lagrangian space consistency relation – using the background wave argument phrased entirely in Lagrangian space.

3.2.2.1 Notation

We use $\mathbf{q}$ to denote the Lagrangian space coordinate of a particle, which coincides with its initial position, and $\mathbf{x}$ to denote the Eulerian space coordinate which is its position at a later time. To be definite, in cases where multiple components are present, the Lagrangian space coordinate $\mathbf{q}$ refers to that of the dark matter particle, which has only gravitational interactions.[34] Both coordinates are defined in comoving space where the expansion of the universe is scaled out. The (dark matter) displacement $\mathbf{\Delta}$ is the difference:

$$\mathbf{x}(\mathbf{q},\eta) = \mathbf{q} + \mathbf{\Delta}(\mathbf{q},\eta) \ . \tag{3.126}$$

The (dark matter) velocity is given by the conformal time derivative of $\mathbf{\Delta}$ at a fixed Lagrangian coordinate:

$$\boldsymbol{v}(\mathbf{q},\eta) = \frac{\partial \mathbf{\Delta}}{\partial \eta}\Big|_{\mathbf{q}} . \tag{3.127}$$

The (dark matter) overdensity δ can be obtained by mass conservation, assuming the initial overdensity is negligible:

$$1 + \delta(\mathbf{x},\eta) = |J(\mathbf{q},\eta)|^{-1} \tag{3.128}$$

with $J(\mathbf{q},\eta)$ being the Jacobian relating the volume elements in Eulerian and Lagrangian space:

$$J(\mathbf{q},\eta) \equiv \det\left[\frac{\partial x^i(\mathbf{q},\eta)}{\partial q^j}\right] . \tag{3.129}$$

The Jacobian J as a function of $\mathbf{q}$ is well-defined even in the presence of multiple-streaming – where a single $\mathbf{x}$ corresponds to multiple $\mathbf{q}$'s – but Eq. (3.128) requires modification in that case:

$$1 + \delta(\mathbf{x},\eta) = \sum_{\mathbf{x}=\mathbf{q}+\mathbf{\Delta}(\mathbf{q},\eta)} |J(\mathbf{q},\eta)|^{-1} , \tag{3.130}$$

where the sum is over all $\mathbf{q}$'s that reach the same $\mathbf{x}$.

Suppose we have some LSS observable $\mathcal{O}$. This could represent many different quantities, such as mass overdensity or galaxy number overdensity.[35] What we typically observe

[34] Our derivation of the Lagrangian space consistency relation would go through even if we chose the Lagrangian coordinate to track other constituents of the universe.

[35] Unless otherwise stated, whenever we discuss mass or galaxy density, we mean the mass or galaxy

is $\mathcal{O}$ as a function of $\mathbf{x}$ (and possibly time, which we suppress). Given this function $\mathcal{O}(\mathbf{x})$, one can define unambiguously a corresponding function of $\mathbf{q}$: $\mathcal{O}(\mathbf{x}(\mathbf{q}))$. In other words, suppose we are interested in the value of $\mathcal{O}$ at a Lagrangian location $\mathbf{q}$: we can define it by working out the $\mathbf{x}$ that $\mathbf{q}$ maps to, and then evaluating $\mathcal{O}(\mathbf{x})$. This procedure is well defined even if multiple $\mathbf{q}$'s map to the same $\mathbf{x}$, which is expected to happen for dark matter in the nonlinear regime.

Some quantities defined in Lagrangian space, on the other hand, might not have an unambiguous meaning in Eulerian space. For instance, the velocity $\boldsymbol{v}$ given in Eq. (3.127) is defined for a dark matter particle labeled by the Lagrangian coordinate $\mathbf{q}$. At an Eulerian position $\mathbf{x}$ where multiple Lagrangian streams cross, additional inputs are required to define a velocity; a reasonable definition is:

$$\text{average } \boldsymbol{v} = \frac{\sum_{\mathbf{x}=\mathbf{q}+\boldsymbol{\Delta}} |J(\mathbf{q})|^{-1} \boldsymbol{v}(\mathbf{q})}{\sum_{\mathbf{x}=\mathbf{q}+\boldsymbol{\Delta}} |J(\mathbf{q})|^{-1}} \,, \tag{3.131}$$

where the sum is over all $\mathbf{q}$'s that map to the same $\mathbf{x}$. This gives a mass weighted velocity.

It is interesting to contrast the Fourier transform in Lagrangian versus Eulerian space, as described by Eq. (3.123). In particular, the Eulerian space Fourier transform can be rewritten as (suppressing time dependence):

$$\mathcal{O}_{\mathbf{k}} = \int d^3\mathbf{x}\, \mathcal{O}(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}} = \int d^3\mathbf{q}\, J(\mathbf{q})\, \mathcal{O}(\mathbf{x}(\mathbf{q})) e^{i\mathbf{k}\cdot(\mathbf{q}+\boldsymbol{\Delta})} \,, \tag{3.132}$$

where J comes without absolute value; this expression remains valid in the presence of multiple streaming. Note how an Eulerian space Fourier transform of $\mathcal{O}(\mathbf{x})$ can be interpreted as a Lagrangian space Fourier transform of $J(\mathbf{q})\mathcal{O}(\mathbf{x}(\mathbf{q}))e^{i\mathbf{k}\cdot\boldsymbol{\Delta}}$.

count per unit *Eulerian* space volume. Such a quantity can of course be expressed as a function of either Eulerian space coordinate $\mathbf{x}$ or Lagrangian space coordinate $\mathbf{q}$.

3.2.2.2 Derivation from the displacement symmetry

We now deduce our main result, making use of a master formula derived in an earlier paper [11]. At the heart of the consistency relation is the existence of a nonlinearly realized symmetry, under which some field – the Goldstone boson or pion π – transforms as $\pi \to \pi + \Delta_{\rm lin.}\pi + \Delta_{\rm nl.}\pi$. Here, $\Delta_{\rm lin.}\pi$ is the part of the transformation that is linear in π, and $\Delta_{\rm nl.}\pi$ is the part of the transformation that is independent of π (i.e., nonlinear in π, though 'sub-linear' or 'inhomogeneous' would be a better description). The fact that $\Delta_{\rm nl.}\pi \neq 0$ is the sign of a nonlinearly realized, or spontaneously broken, symmetry. At the same time, there are other fields or observables $\mathcal{O}$ that could have their own linear and/or nonlinear transformations. The master formula (in momentum space) reads [11]:

$$\int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{\langle \pi_{\mathbf{p}} \mathcal{O}_{\mathbf{p}_1} \cdots \mathcal{O}_{\mathbf{p}_N} \rangle^c}{P_\pi(\mathrm{p})} \Delta_{\rm nl.}\pi^*_{\mathbf{p}} = \Delta_{\rm lin.} \langle \mathcal{O}_{\mathbf{p}_1} \cdots \mathcal{O}_{\mathbf{p}_N} \rangle^c \,, \tag{3.133}$$

where $\langle ... \rangle^c$ refers to the connected correlation function *without* removing the overall delta function (as opposed to $\langle ... \rangle^{c'}$ which has the delta function removed). Note how it is the nonlinear transformation of π and the linear transformation of $\mathcal{O}$ that show up on the left and the right respectively. Note also that the $\mathcal{O}$'s need not even be the same observable. Nor do π and the $\mathcal{O}$'s need be at the same time: they can be at *arbitrary, potentially different, times.* The derivation of this master formula made no assumption about whether the quantities (or the Fourier transform thereof) are defined in Eulerian or Lagrangian space. We are thus free to use it in either. This master relation can be used to derive the large scale structure analog of Ward identities or soft-pion theorems in particle physics.

As a warm-up, let us first apply this formula to a simple system that involves the dark matter only. The dynamics is described by: (1) $\mathbf{x} = \mathbf{q} + \boldsymbol{\Delta}$ as in Eq. (3.126); (2) the dark mater overdensity δ determined by the Jacobian as in Eq. (3.130); (3) the displacement

$\mathbf{\Delta}$ which evolves according to:

$$\left.\frac{\partial^2 \mathbf{\Delta}}{\partial \eta^2}\right|_{\mathbf{q}} + \frac{a'}{a}\left.\frac{\partial \mathbf{\Delta}}{\partial \eta}\right|_{\mathbf{q}} = -\mathbf{\nabla}_x \Phi\,, \tag{3.134}$$

where a is the scale factor, a' is its derivative with respect to conformal time η, Φ is the gravitational potential and $\mathbf{\nabla}_x$ is the partial derivative with respect to $\mathbf{x}$; lastly (4) the Poisson equation:

$$\nabla_x^2 \Phi = 4\pi G a^2 \bar{\rho}\delta\,, \tag{3.135}$$

where G is Newton's constant and $\bar{\rho}$ is the mean mass density.

This system has the following symmetry:

$$\boxed{\mathbf{q} \to \mathbf{q} \quad , \quad \mathbf{\Delta} \to \mathbf{\Delta} + \mathbf{n}(\eta) \quad , \quad \Phi \to \Phi - \left(\mathbf{n}'' + \frac{a'}{a}\mathbf{n}'\right)\cdot \mathbf{x}\,,} \tag{3.136}$$

where $\mathbf{n}(\eta)$ is a function of time alone. We will refer to this as the *displacement symmetry*. Note how $\mathbf{\Delta}$ shifts by a nonlinear (or sub-linear) amount and can be thought of as our Goldstone boson. The same is true for Φ. The interesting point is that the mass overdensity δ does not transform at all under this symmetry. Nor are $\mathbf{q}$ or η transformed. Applying the master formula, choosing the observable $\mathcal{O} = \delta$, we thus find:

$$\lim_{\mathbf{p}\to 0} \frac{\langle \mathbf{\Delta}_{\mathbf{p}} \delta_{\mathbf{p_1}} ... \delta_{\mathbf{p_N}} \rangle^{c'}}{P_\Delta(\mathrm{p})} = 0\,. \tag{3.137}$$

Here, we have used the fact that the nonlinear transformation of $\mathbf{\Delta}$ in Fourier space is $\Delta_{\mathrm{nl.}}\mathbf{\Delta}_{\mathbf{p}} = \mathbf{n}(\eta)(2\pi)^3\delta_D(\mathbf{p})$, where $\delta_D(\mathbf{p})$ is the Dirac delta function, $\Delta_{nl.}$ denotes the sublinear change of quantities and $\mathbf{\Delta}_{\mathbf{p}}$ denotes the displacement mode carrying momentum $\mathbf{p}$. We have also removed the overall momentum-conserving delta function. The power spectrum of displacement P_Δ is as defined in sec. 3.2.1.

Two comments are in order before we proceed to generalize this derivation to more realistic, astrophysically complex observables. First, while the Lagrangian coordinate $\mathbf{q}$ does not transform under the symmetry of interest, the Eulerian coordinate $\mathbf{x} = \mathbf{q} + \mathbf{\Delta}$ does, because the displacement $\mathbf{\Delta}$ shifts. This implies that an observable like δ, when expressed as a function of $\mathbf{x}$, transforms as: $\delta \to \delta + \Delta_{\text{lin.}}\delta$ with $\Delta_{\text{lin.}}\delta = \mathbf{\Delta} \cdot \nabla\delta$. Plugging this into the master formula Eq. (3.133), we see that there is a non-vanishing right hand side, unlike the situation in Lagrangian space where δ expressed as a function of $\mathbf{q}$ does not shift at all. The is the fundamental reason why the KRPP consistency relation takes a more complicated form in Eulerian space (Eq. 3.121) than in Lagrangian space (Eq. 3.137).

Second, the reader might wonder about the validity of our application of the master formula: on the one hand, the master relation is phrased in terms of a scalar pion; on the other, our application effectively uses the vector displacement $\mathbf{\Delta}$ as the pion. The short answer is that the master formula is applicable to any field π that shifts nonlinearly under the symmetry of interest; one can use it for each component of $\mathbf{\Delta}$ for instance. The long answer is that since $\mathbf{\Delta}$ is used in the consistency relation only as a soft (long wavelength) mode, one is justified in treating it as a gradient mode (assuming the growing mode initial condition) with $\mathbf{\Delta} = \nabla_q \pi$ and π playing the role of the displacement potential. The master formula can then be applied with the displacement potential as the pion. The resulting consistency relation can be shown to be equivalent to the one we have derived.[36]

Let us turn to the derivation of a stronger form of the Lagrangian space consistency relation. So far, we have focused on a simple system of dark matter particles that interact

[36] There are actually two different nonlinear realized symmetries associated with the displacement potential. One is shifting it by a constant or a function of time (but not space). The other is shifting it by a linear gradient, i.e., $\pi \to \pi + \mathbf{n} \cdot \mathbf{q}$ where $\mathbf{n}$ is the same as that in Eq. (3.136). There are as a result two consistency relations which can be succinctly combined into one, Eq. (3.137). See our earlier paper [11] for further discussions.

only gravitationally, as embodied in Eqs. (3.134) and (3.135). Let us consider the addition of galaxies into the mix. They have their own number overdensity δ_g, displacement $\mathbf{\Delta}_g$ and velocity $\mathbf{v}_g = \mathbf{\Delta}'_g$. Their number density is not necessarily conserved by evolution, since galaxies can form and merge:

$$\delta'_g + (1+\delta_g)\boldsymbol{\nabla}_x \cdot \mathbf{v}_g = R_g \,, \tag{3.138}$$

where $'$ refers to conformal time derivative at a fixed Lagrangian coordinate and R_g is a source term that incorporates the formation and merger rates. The equation of motion for the galaxies is:

$$\mathbf{\Delta}''_g + \frac{a'}{a}\mathbf{\Delta}'_g = -\boldsymbol{\nabla}_x \Phi + \mathbf{F}_g \,, \tag{3.139}$$

where $\mathbf{F}_g$ encodes additional forces that might act on galaxies, such as gas pressure, dynamical friction et cetera. The gravitational potential Φ is determined of course by the Poisson equation (3.135) as before.

The displacement symmetry of Eq. (3.136) can be extended to include also:

$$\mathbf{\Delta}_g \to \mathbf{\Delta}_g + \mathbf{n}(\eta) \,, \tag{3.140}$$

which also implies $\mathbf{v}_g \to \mathbf{v}_g + \mathbf{n}'$. The galaxy overdensity δ_g, like its dark matter counterpart, does not transform under this symmetry. Eqs. (3.136) and (3.140) represent the displacement symmetry of the combined dark-matter-galaxies system, as long as R_g and $\mathbf{F}_g$ depend only on (dark matter/galaxy) densities and gradients of (dark matter/galaxy) velocities – recall that neither shifts under our symmetry. What happens if R_g and/or $\mathbf{F}_g$ depends on velocities as opposed to gradients of velocities? In that case, shifting velocities by a spatially constant amount would affect the galaxy formation and dynamics

– this is a violation of the equivalence principle which states that local physical processes (such as galaxy formation, mergers and motion) should not be dependent on the absolute state of motion. Note that a dependence on the dark-matter-galaxy velocity *difference* $\mathbf{v} - \mathbf{v}_g$, on the other hand, is consistent with the equivalence principle, and the velocity difference is indeed unchanged under our symmetry. Thus, as long as the equivalence principle is respected, whether R_g and $\mathbf{F}_g$ depend on densities, gradients of velocities or velocity differences, the displacement symmetry holds. Furthermore, the same statement is expected to be valid in a system with many different species, such as baryons, galaxies or even dark matter of different kinds. The argument that leads to Eq. (3.137) can be rerun to give the more general Lagrangian space consistency relation:

$$\boxed{\lim_{\mathbf{p}\to 0} \frac{\langle \boldsymbol{\Delta}_{\mathbf{p}}(\eta)\, \mathcal{O}_{\mathbf{p}_1}(\eta_1) \ldots \mathcal{O}_{\mathbf{p}_N}(\eta_N)\, \rangle^{c'}}{P_\Delta(\mathrm{p}, \eta)} = 0\,,} \tag{3.141}$$

where $\mathcal{O}$ is any observable that has no linear shift under the displacement symmetry – this includes for instance the densities, displacements and velocities of the galaxies and of dark matter.[37] Note that the $\mathcal{O}$'s need not be the same observables. We have restored the explicit time-dependence of each fluctuation variable to emphasize the fact that the times need not be equal. Note also that we have chosen the dark matter displacement to be the pion. We could have chosen the galaxy displacement instead. Assuming that gravity is the dominant interaction on large scales and that adiabatic initial conditions hold, the two displacements are expected to coincide in any case in the soft limit. Furthermore, we could have chosen the velocity instead of the displacement as the soft-pion, in which case Eq. (3.122) follows.

[37] The reader might wonder: given that the Lagrangian coordinate $\mathbf{q}$ does not get transformed at all under the displacement symmetry, is there any observable that has a linear shift? The answer is yes. For instance, the combination $\mathcal{O} = \mathbf{v}\delta$ transforms to $(\mathbf{v} + \mathbf{n}')\delta$ giving a shift that is linear in the fluctuation variable δ.

3.2.3 The Lagrangian space consistency relation: checks

The above derivation of the Lagrangian space consistency relation is a bit terse, and the form the relation takes is surprisingly simple. It is thus worth performing some non-trivial checks of the relation. We will first do this using second order Lagrangian perturbation theory (sec. 3.2.3.1). Then, in sec. 3.2.3.2, we demonstrate how the Eulerian space consistency relation can be derived from its counterpart in Lagrangian space.

3.2.3.1 Perturbative check

Let us perform an explicit check of Eq. (3.141) using second-order Lagrangian space perturbation theory. For simplicity, we will focus on the case where the only species present is dark matter and the observable $\mathcal{O} = \delta$. We will confine the discussion to the squeezed three-point function; extension to a general (N+1)-point function is straightforward. Expanding Eq. (3.128) to second order, we have

$$\delta(\mathbf{x}(\mathbf{q},\eta),\eta) = -\nabla_{\mathbf{q}} \cdot \mathbf{\Delta} + \frac{1}{2}(\nabla_{\mathbf{q}} \cdot \mathbf{\Delta})^2 + \frac{1}{2}\nabla_{\mathbf{q}^i}\mathbf{\Delta}^j\nabla_{\mathbf{q}^j}\mathbf{\Delta}^i \,. \tag{3.142}$$

Expanding out $\delta_{\mathbf{p}} = \delta^{(1)}_{\mathbf{p}} + \delta^{(2)}_{\mathbf{p}} + \cdots$, $\mathbf{\Delta}_{\mathbf{p}} = \mathbf{\Delta}^{(1)}_{\mathbf{p}} + \mathbf{\Delta}^{(2)}_{\mathbf{p}} + \cdots$, and plugging into Eqs. (3.134) and (3.135), we have [102]:

$$\begin{aligned}
\Delta^{j\,(1)}_{\mathbf{p}}(\eta) &= \frac{-i\,\mathrm{p}^j}{\mathrm{p}^2}\,\delta^{(1)}_{\mathbf{p}}(\eta)\,,\\
\Delta^{j\,(2)}_{\mathbf{p}}(\eta) &= \frac{1}{2}\frac{D_2(\eta)}{D(\eta)^2}\frac{i\,\mathrm{p}^j}{\mathrm{p}^2}\int\frac{d^3\mathbf{p}_A d^3\mathbf{p}_B}{(2\pi)^3}\delta_D(\mathbf{p}_A+\mathbf{p}_B-\mathbf{p})\left(1-\frac{(\mathbf{p}_A\cdot\mathbf{p}_B)^2}{\mathrm{p}_A^2\mathrm{p}_B^2}\right)\delta^{(1)}_{\mathbf{p}_A}(\eta)\delta^{(1)}_{\mathbf{p}_B}(\eta)\,,\\
\delta^{(2)}_{\mathbf{p}}(\eta) &= \frac{1}{2}\int\frac{d^3\mathbf{p}_A d^3\mathbf{p}_B}{(2\pi)^3}\delta_D(\mathbf{p}_A+\mathbf{p}_B-\mathbf{p})\left(1-\frac{D_2(\eta)}{D(\eta)^2}+\frac{(\mathbf{p}_A\cdot\mathbf{p}_B)^2}{\mathrm{p}_A^2\mathrm{p}_B^2}\left[1+\frac{D_2(\eta)}{D(\eta)^2}\right]\right)\delta^{(1)}_{\mathbf{p}_A}(\eta)\delta^{(1)}_{\mathbf{p}_B}(\eta)\,,
\end{aligned} \tag{3.143}$$

where D is the linear growth factor determining the time-dependence of the first order displacement (and density), and D_2 is the second order growth factor determining that of the second order displacement. They satisfy the equations:

$$\begin{aligned} &D'' + \frac{a'}{a}D' - 4\pi G a^2 \bar{\rho} D = 0\,, \\ &D_2'' + \frac{a'}{a}D_2' - 4\pi G a^2 \bar{\rho} D_2 = -4\pi G a^2 \bar{\rho} D^2\,. \end{aligned} \tag{3.144}$$

For instance, in a flat universe with $\Omega_m = 1$, $D_2 = -3D^2/7$. Using these expressions, we can work out the lowest order contributions to the relevant squeezed bispectrum:

$$\begin{aligned} &\langle \Delta^j_{\mathbf{p}}(\eta)\delta_{\mathbf{p}_1}(\eta_1)\delta_{\mathbf{p}_2}(\eta_2)\rangle \\ =&\langle \Delta^{j\,(2)}_{\mathbf{p}}(\eta)\delta^{(1)}_{\mathbf{p}_1}(\eta_1)\delta^{(1)}_{\mathbf{p}_2}(\eta_2)\rangle + \langle \Delta^{j\,(1)}_{\mathbf{p}}(\eta)\delta^{(2)}_{\mathbf{p}_1}(\eta_1)\delta^{(1)}_{\mathbf{p}_2}(\eta_2)\rangle + \langle \Delta^{j\,(1)}_{\mathbf{p}}(\eta)\delta^{(1)}_{\mathbf{p}_1}(\eta_1)\delta^{(2)}_{\mathbf{p}_2}(\eta_2)\rangle \\ =&O(\mathrm{p}^j) + O(\mathrm{p}^j P_\Delta(\mathrm{p}))\,, \end{aligned} \tag{3.145}$$

where we spell out the dependence on the soft momentum $\mathbf{p}$: the $O(\mathrm{p}^j)$ piece comes from the first term on the right in the first line, and the $O(\mathrm{p}^j P_\Delta(\mathrm{p}))$ piece comes from the other two terms. We have used the fact that $1-[(\mathbf{p}_A\cdot\mathbf{p}_B)^2/\mathrm{p}_A^2\mathrm{p}_B^2] = O(\mathrm{p}^2)$ for $\mathbf{p}_A+\mathbf{p}_B = \mathbf{p}$. The $O(\mathrm{p}^j P_\Delta(\mathrm{p}))$ piece is obviously compatible with the Lagrangian space consistency relation; i.e., it gives $\langle \Delta^j_{\mathbf{p}}\delta_{\mathbf{p}_1}\delta_{\mathbf{p}_2}\rangle^{c'}/P_\Delta(\mathrm{p}) = 0$ in the $\mathbf{p}\to 0$ limit. The $O(\mathrm{p}^j)$ piece does the same, provided the power spectrum $P_\Delta(\mathrm{p})$ is not too blue. Parametrizing the power spectrum $P_\Delta(\mathrm{p}) \propto \mathrm{p}^{n-2}$ in the low momentum limit, the consistency relation holds as long as $n < 3$. Exactly the same condition is needed for the Eulerian space consistency relation (see e.g. [11]).[38]

[38] The Eulerian space consistency relation is often given in a form where δ is used as the soft mode. One might be tempted to do the same for the Lagrangian space consistency relation. However, one can check using perturbation theory that such a consistency relation would have required $n < 1$, a condition considerably stronger than expected. This is related to the fact that $\delta^{(2)}_{\mathbf{p}}$ does not vanish in the $\mathbf{p}\to 0$

3.2.3.2 Recovering the Eulerian space consistency relation from Lagrangian space

The consistency relation takes such a different form in Lagrangian versus Eulerian space that it is worth considering how one can be derived from the other. Let us compute the following:

$$\begin{aligned} &E \equiv E_L + E_R \\ &E_L \equiv \lim_{\mathbf{k}\to 0} \frac{\langle v^j_{\mathbf{k}}(\eta)\, \delta_{\mathbf{k}_1}(\eta_1)\, \delta_{\mathbf{k}_2}(\eta_2)\,\rangle^{c'}}{P_v(\mathrm{k},\eta)} \quad , \quad E_R \equiv -i\mathrm{k}^j \sum_{a=1}^{2} \frac{D(\eta_a)}{D'(\eta)} \frac{\mathbf{k}\cdot\mathbf{k}_a}{\mathrm{k}^2} \langle \delta_{\mathbf{k}_1}(\eta_1)\, \delta_{\mathbf{k}_2}(\eta_2)\,\rangle^{c'} \,. \end{aligned} \tag{3.146}$$

The Eulerian space consistency condition is the statement that $E = 0$. We will content ourselves with deriving this – a special case of the more general Eulerian space consistency relation (3.121) – from the Lagrangian space consistency relation.

To relate E to quantities in Lagrangian space, we will slightly abuse our notation. So far, we have been using $\mathbf{k}$ for the Eulerian space momentum and $\mathbf{p}$ for the Lagrangian space momentum. For instance, $\delta_{\mathbf{k}_1}$ is defined as

$$\delta_{\mathbf{k}_1} = \int d^3\mathbf{x}\, \delta(\mathbf{x})\, e^{i\mathbf{k}_1\cdot\mathbf{x}} \,. \tag{3.147}$$

Let us rewrite this as

$$\begin{aligned} \delta_{\mathbf{k}_1=\mathbf{p}_1} &= \int d^3\mathbf{q}\, J(\mathbf{q})\, \delta(\mathbf{x}(\mathbf{q}))\, e^{i\mathbf{p}_1\cdot(\mathbf{q}+\boldsymbol{\Delta}(\mathbf{q}))} \\ &= \int d^3\mathbf{q}\, J(\mathbf{q})\, \delta(\mathbf{x}(\mathbf{q}))\, e^{i\mathbf{p}_1\cdot\mathbf{q}} + i\mathrm{p}_1^m \int d^3\mathbf{q}\, J(\mathbf{q})\, \delta(\mathbf{x}(\mathbf{q}))\, \Delta^m(\mathbf{q})\, e^{i\mathbf{p}_1\cdot\mathbf{q}} \\ &\quad - \frac{1}{2}\mathrm{p}_1^m\mathrm{p}_1^n \int d^3\mathbf{q}\, J(\mathbf{q})\, \delta(\mathbf{x}(\mathbf{q}))\, \Delta^m(\mathbf{q})\Delta^n(\mathbf{q})\, e^{i\mathbf{p}_1\cdot\mathbf{q}} + \ldots\,, \end{aligned} \tag{3.148}$$

limit, unlike $\Delta^{j\,(2)}_{\mathbf{p}}$.

where $J(\mathbf{q})$ is defined by Eq. (3.129) (with no absolute value). Defining

$$\tilde{\delta}(\mathbf{q}) \equiv J(\mathbf{q})\delta(\mathbf{x}(\mathbf{q}))\,, \tag{3.149}$$

we see that the first term on the right of Eq. (3.148) is $\tilde{\delta}_{\mathbf{p}_1}$, i.e. the Fourier transform of $\tilde{\delta}$ in *Lagrangian space*. The other terms on the right can likewise be thought of as the Fourier transform of some quantity in Lagrangian space. This is why we introduce $\mathbf{p}_1$ as the momentum label for these Fourier components. On the other hand, upon summation, they give the quantity on the left $\delta_{\mathbf{k}_1=\mathbf{p}_1}$ which is the Fourier transform of density in *Eulerian space* – this is why we use $\mathbf{k}_1$ as its momentum label; it just happens to take on the numerical value $\mathbf{p}_1$ which conveniently gives us the appropriate momentum label for quantities on the right. It is worth emphasizing that our definitions are general, in that they are valid even in the presence of multiple-streaming (see Sec. 3.2.2.1).

The expansion in terms of Δ in Eq. (3.148) is purely formal. In the nonlinear regime, there is no sense in which Δ is small. The expansion provides a convenient way to relate the Fourier transform in Eulerian space to the Fourier transform in Lagrangian space. We will argue $E = 0$ holds to arbitrary order in a power series expansion.

For the soft mode, we have

$$v^j_{\mathbf{k}=\mathbf{p}} = v^j_{\mathbf{p}} + \ldots\,, \tag{3.150}$$

This is where our abuse of notation is the most egregious: on the left is the velocity Fourier transformed in Eulerian space; on the right is the velocity Fourier transformed in Lagrangian space. They agree only to lowest order in perturbations. For the soft-mode, ignoring the higher order corrections is permissible: the higher order corrections will give higher powers of the soft momentum $\mathbf{p}$ compared to what is kept in the consistency relation, provided that the soft power spectrum $P_v(\mathrm{p})$ or $P_\Delta(\mathrm{p})$ is not too blue (see sec. 3.2.3.1). Similarly, it can be shown that in the soft limit, there is no need to distinguish between P_v in Lagrangian versus Eulerian space.[39]

Let us substitute Eq. (3.148) for the hard modes, and Eq. (3.150) for the soft mode, into the expression for E in Eq. (3.146). Consider first what contributes to E_L:

$$\begin{aligned}\langle v^j_{\mathbf{k}=\mathbf{p}}\delta_{\mathbf{k}_1=\mathbf{p}_1}\delta_{\mathbf{k}_2=\mathbf{p}_2}\rangle =&\langle v^j_{\mathbf{p}}\tilde{\delta}_{\mathbf{p}_1}\tilde{\delta}_{\mathbf{p}_2}\rangle \\ &+\left[i\mathrm{p}_1^m\int\frac{d^3\mathbf{p}_A}{(2\pi)^3}\langle v^j_{\mathbf{p}}\Delta^m_{\mathbf{p}_A}(\eta_1)\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}\tilde{\delta}_{\mathbf{p}_2}\rangle + 1\leftrightarrow 2\right] \\ &-\left[\mathrm{p}_1^m\mathrm{p}_2^n\int\frac{d^3\mathbf{p}_A}{(2\pi)^3}\frac{d^3\mathbf{p}_B}{(2\pi)^3}\langle v^j_{\mathbf{p}}\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}\Delta^m_{\mathbf{p}_A}(\eta_1)\tilde{\delta}_{\mathbf{p}_2-\mathbf{p}_B}\Delta^n_{\mathbf{p}_B}(\eta_2)\rangle\right. \\ &\quad\left.+\left[\frac{1}{2}\mathrm{p}_1^m\mathrm{p}_1^n\int\frac{d^3\mathbf{p}_A}{(2\pi)^3}\frac{d^3\mathbf{p}_B}{(2\pi)^3}\langle v^j_{\mathbf{p}}\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A-\mathbf{p}_B}\Delta^m_{\mathbf{p}_A}(\eta_1)\Delta^n_{\mathbf{p}_B}(\eta_1)\tilde{\delta}_{\mathbf{p}_2}\rangle + (1\leftrightarrow 2)\right]\right] \\ &+O(\Delta^3)+...\end{aligned} \tag{3.151}$$

where we have largely suppressed the time-dependence to minimize clutter (η for the soft mode, and η_1 and η_2 respectively for the hard modes), except for variables with internal momenta. We emphasize that the expansion in Δ is purely formal, and comes entirely from expanding $e^{i\mathbf{p}_1\cdot\boldsymbol{\Delta}}$ or $e^{i\mathbf{p}_2\cdot\boldsymbol{\Delta}}$. The first term on the right can be set to zero by virtue of the Lagrangian space consistency condition (keeping in mind that this term is divided by $P_v(\mathrm{p})$ as part of the quantity E_L). We will be assuming the Lagrangian space consistency relation in its general form (Eq. 3.122):

$$\lim_{\mathbf{p}\to 0}\frac{\langle \boldsymbol{v}_{\mathbf{p}}(\eta)\,\mathcal{O}_{\mathbf{p}_1}(\eta_1)\,...\,\mathcal{O}_{\mathbf{p}_N}(\eta_N)\,\rangle^{c'}}{P_v(\mathrm{p},\eta)}=0\,, \tag{3.152}$$

[39] It is also worth emphasizing that the notion of a well-defined velocity in Eulerian space is valid only when multiple-streaming is ignored. This is acceptable for the soft-mode. We do not assume single-streaming for the hard modes.

where the observables at hard momenta need not be the same observable. Finally, note that while we are interested in the connected part of the correlator on the left hand side of Eq. (3.151), the correlators on the right hand side are the full correlators, minus the contributions where some proper subset of the original hard and soft momenta sum to zero. In particular, the correlators on the right hand side of Eq. (3.151) contain both connected and disconnected pieces.

The second term on the right of Eq. (3.151), formally $O(\Delta)$, equals

$$i\mathrm{p}_1^m \int \frac{d^3\mathbf{p}_A}{(2\pi)^3}\Bigg[\langle v^j_{\mathbf{p}}\Delta^m_{\mathbf{p}_A}(\eta_1)\rangle\langle\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}\tilde{\delta}_{\mathbf{p}_2}\rangle + \langle v^j_{\mathbf{p}}\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}\rangle\langle\Delta^m_{\mathbf{p}_A}(\eta_1)\tilde{\delta}_{\mathbf{p}_2}\rangle + \langle v^j_{\mathbf{p}}\tilde{\delta}_{\mathbf{p}_2}\rangle\langle\Delta^m_{\mathbf{p}_A}(\eta_1)\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}\rangle + \langle v^j_{\mathbf{p}}\Delta^m_{\mathbf{p}_A}(\eta_1)\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}\tilde{\delta}_{\mathbf{p}_2}\rangle^c\Bigg] + (1\leftrightarrow 2)\,, \tag{3.153}$$

where the connected trispectrum term $\langle v^j_{\mathrm{p}}...\rangle^c$ (anticipating division by $P_v(\mathrm{p})$) can be set to zero using the Lagrangian space consistency relation. The terms involving $\langle v^j_{\mathbf{p}}\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}\rangle$, $\langle v^j_{\mathbf{p}}\tilde{\delta}_{\mathbf{p}_2}\rangle$ and the like have one more power of the soft momentum $\mathbf{p}$ (and are thus subdominant) compared to terms involving $\langle v^j_{\mathrm{p}}\Delta^m_{\mathbf{p}_A}\rangle$, which give:

$$(2\pi)^3\delta_D(\mathbf{p}_1+\mathbf{p_2}+\mathbf{p})\,P_v(\mathrm{p},\eta)\,i\mathrm{p}^j\,\frac{D(\eta_1)}{D'(\eta)}\frac{\mathbf{p}_1\cdot\mathbf{p}}{\mathrm{p}^2}\langle\tilde{\delta}_{\mathbf{p}_1+\mathbf{p}}(\eta_1)\tilde{\delta}_{\mathbf{p}_2}(\eta_2)\rangle^{c'} + (1\leftrightarrow 2)\,. \tag{3.154}$$

The third term on the right of Eq. (3.151), formally $O(\Delta^2)$, can be treated in a similar way: some can be ignored by assuming the Lagrangian space consistency relation, some are subdominant in the soft-limit (i.e., they vanish upon division by $P_v(p)$ and sending

$\mathbf{p} \to 0$), and the dominant terms are those that involve $\langle v^j_{\mathbf{p}} \Delta \rangle$ which give:

$$-(2\pi)^3 \delta_D(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}) P_v(\mathrm{p}, \eta) \int \frac{d^3\mathbf{p}_A}{(2\pi)^3} \Bigg[\mathrm{p}^j \mathrm{p}_2^m \frac{D(\eta_1)}{D'(\eta)} \frac{\mathbf{p}_1 \cdot \mathbf{p}}{\mathrm{p}^2} \langle \tilde{\delta}_{\mathbf{p}_1+\mathbf{p}}(\eta_1) \tilde{\delta}_{\mathbf{p}_2-\mathbf{p}_A}(\eta_2) \Delta^m_{\mathbf{p}_A}(\eta_2) \rangle^{c'} + \mathrm{p}^j \mathrm{p}_1^m \frac{D(\eta_1)}{D'(\eta)} \frac{\mathbf{p}_1 \cdot \mathbf{p}}{\mathrm{p}^2} \langle \tilde{\delta}_{\mathbf{p}_1+\mathbf{p}-\mathbf{p}_A}(\eta_1) \tilde{\delta}_{\mathbf{p}_2}(\eta_2) \Delta^m_{\mathbf{p}_A}(\eta_1) \rangle^{c'} + (1 \leftrightarrow 2) \Bigg] . \tag{3.155}$$

Thus, combining Eqs. (3.154) and (3.155), E_L of Eq. (3.146) can be rewritten as:

$$\begin{aligned} E_L \equiv& \lim_{\mathbf{p}\to 0} \frac{\langle v^j_{\mathbf{k}=\mathbf{p}}(\eta)\, \delta_{\mathbf{k}_1=\mathbf{p}_1}(\eta_1)\, \delta_{\mathbf{k}_2=\mathbf{p}_2}(\eta_2)\, \rangle^{c'}}{P_v(\mathrm{p},\eta)} \\ =& i\mathrm{p}^j \frac{D(\eta_1)}{D'(\eta)} \frac{\mathbf{p}_1 \cdot \mathbf{p}}{\mathrm{p}^2} \langle \tilde{\delta}_{\mathbf{p}_1}(\eta_1) \tilde{\delta}_{\mathbf{p}_2}(\eta_2) \rangle^{c'} - \\ & \int \frac{d^3\mathbf{p}_A}{(2\pi)^3} \Bigg[\mathrm{p}^j \mathrm{p}_2^m \frac{D(\eta_1)}{D'(\eta)} \frac{\mathbf{p}_1 \cdot \mathbf{p}}{\mathrm{p}^2} \langle \tilde{\delta}_{\mathbf{p}_1}(\eta_1) \tilde{\delta}_{\mathbf{p}_2-\mathbf{p}_A}(\eta_2) \Delta^m_{\mathbf{p}_A}(\eta_2) \rangle^{c'} \\ & + \mathrm{p}^j \mathrm{p}_1^m \frac{D(\eta_1)}{D'(\eta)} \frac{\mathbf{p}_1 \cdot \mathbf{p}}{\mathrm{p}^2} \langle \tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}(\eta_1) \tilde{\delta}_{\mathbf{p}_2}(\eta_2) \Delta^m_{\mathbf{p}_A}(\eta_1) \rangle^{c'} \Bigg] + (1 \leftrightarrow 2) + \dots . \end{aligned} \tag{3.156}$$

Next, let us rewrite E_R using the same strategy:

$$\begin{aligned} E_R \equiv& - i\mathrm{p}^j \sum_{a=1}^{2} \frac{D(\eta_a)}{D'(\eta)} \frac{\mathbf{p} \cdot \mathbf{p}_a}{\mathrm{p}^2} \langle \delta_{\mathbf{k}_1=\mathbf{p}_1}(\eta_1)\, \delta_{\mathbf{k}_2=\mathbf{p}_2}(\eta_2)\, \rangle^{c'} \\ =& - i\mathrm{p}^j \frac{D(\eta_1)}{D'(\eta)} \frac{\mathbf{p} \cdot \mathbf{p}_1}{\mathrm{p}^2} \langle \tilde{\delta}_{\mathbf{p}_1}(\eta_1)\, \tilde{\delta}_{\mathbf{p}_2}(\eta_2)\, \rangle^{c'} \\ & + \int \frac{d^3\mathbf{p}_A}{(2\pi)^3} \Bigg[\mathrm{p}^j \mathrm{p}_2^m \frac{D(\eta_1)}{D'(\eta)} \frac{\mathbf{p}_1 \cdot \mathbf{p}}{\mathrm{p}^2} \langle \tilde{\delta}_{\mathbf{p}_1}(\eta_1) \tilde{\delta}_{\mathbf{p}_2-\mathbf{p}_A}(\eta_2) \Delta^m_{\mathbf{p}_A}(\eta_2) \rangle^{c'} \\ & + \mathrm{p}^j \mathrm{p}_1^m \frac{D(\eta_1)}{D'(\eta)} \frac{\mathbf{p}_1 \cdot \mathbf{p}}{\mathrm{p}^2} \langle \tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_A}(\eta_1) \tilde{\delta}_{\mathbf{p}_2}(\eta_2) \Delta^m_{\mathbf{p}_A}(\eta_1) \rangle^{c'} \Bigg] + (1 \leftrightarrow 2) + \dots . \end{aligned} \tag{3.157}$$

Thus, we see that $E \equiv E_L + E_R = 0$, at least to the two lowest non-trivial orders in

Δ. The cancelation works like this: expanding $e^{i\mathbf{p}_1\cdot\boldsymbol{\Delta}}$ and $e^{i\mathbf{p}_2\cdot\boldsymbol{\Delta}}$ as a formal power series in Δ, a given order for E_L is canceled by one lower order for E_R. It can be shown that this pattern continues to arbitrarily high orders. The proof is given in the Appendix. This completes our derivation of the Eulerian space consistency relation, embodied in the statement $E = 0$ (Eq. 3.146), from the Lagrangian space consistency relation (Eq. 3.152).

Chapter 4

Conclusions

In this thesis we discussed extensively how holography and locality are related in the quantum field theory description of cosmology, and the implications of symmetries, especially conformal symmetries. There is a one-to-one map between the conformal group and the isometry group of dS and AdS, which makes it possible to establish the AdS/CFT and dS/CFT dualities. In the large-N limit, this duality between bulk and boundary local observables holds. The AdS and dS isometries are linearly realized by the fields defined on the space, but in the actual cosmology, the background solution breaks the de Sitter group. For instance the inflationary geometry is only approximately de Sitter, with the deviation from de Sitter space controlled by the slow-roll parameter. Even in these cases conformal symmetries play unexpected while important roles. As residual gauge transformations they are non-linearly realized by the cosmological fluctuations and thus constraints on the soft limits of the correlation functions as Ward identities can be derived. Below we make a brief overview and discussions of the chapters.

In Chapter 2 we made progress in the construction of de Sitter bulk operators in terms of non–local boundary operators. For heavy scalars with $m^2 > \left(\frac{d}{2}\right)^2$ the construction recovers the bulk Wightman function in the Euclidean vacuum state. The construction is state–dependent: for different vacuum states in de Sitter space, we have different de Sitter and CFT correlation functions, and thus different construction prescriptions relating bulk observables and boundary operators.

We have performed the construction at the level of two point function, which is at order N^0 in the large–N expansion. To go beyond this one would like to think about three–point functions with each of the bulk operators corresponding to two boundary operators. There can be more subtleties than in AdS case. For instance, the three–point function of scalars in de Sitter space, with one scalar deep inside de Sitter and two others close to the boundary, is schematically

$$\begin{aligned}&\langle\Phi(\eta_1,\mathbf{x}_1)\Phi(\eta_2\sim 0,\mathbf{x}_2)\Phi(\eta_3\sim 0,\mathbf{x}_3)\rangle\sim\\&(\eta_2\eta_3)^{\Delta}\int d^3\mathbf{x}_1' K_+(\eta_1,\mathbf{x}_1|\mathbf{x}_1')\langle\mathcal{O}_+(\mathbf{x}_1')\mathcal{O}_+(\mathbf{x}_2)\mathcal{O}_+(\mathbf{x}_3)\rangle+\\&\eta_2^{\Delta}\eta_3^{d-\Delta}\int d^3\mathbf{x}_1' K_-(\eta_1,\mathbf{x}_1'|\mathbf{x}_1')\langle\mathcal{O}_+(\mathbf{x}_1')\mathcal{O}_+(\mathbf{x}_2)\mathcal{O}_-(\mathbf{x}_3)\rangle+\dots\end{aligned}$$

which involves all the boundary three–point functions that are constructed from $\mathcal{O}_\pm$. As in the AdS case, one would expect that the bulk lightcone singularities will show up in each of these terms, thus breaking microcausality. The hope is to include towers of multi–trace operators into the construction and recover microcausality.

One can also consider bulk operators in de Sitter space in other coordinates, apart from the flat and the global patch. The static patch would be a very interesting case. Once we go into the static patch, we have no asymptotic boundaries anymore—they are behind the horizon, so one can ask in that case what data should be used to construct the local operators. Possibly the approaches people take for the problem of constructing local observables behind the horizon of an eternal black hole in AdS could shed some light on it. There, although the interior is separated from the boundary by a horizon, still the construction from the boundary is shown to be possible with extra degrees of freedom involved—either from the CFT's thermofield double [55] or from a fine–grained sector in the CFT [32]. In de Sitter space, it is not clear if a similar construction will work. There are proposals for static patch holography such as dS/dS[33][48] and static patch solipsism[34], in which the dual theory to the static patch lives on the central slice of dS slicing and the central worldline[1] respectively. It would be interesting to explore whether and how the discussion of operator dictionaries and bulk locality can be extended to these proposals. A related question is how to understand local fields when bubble nucleation is considered. We have already mentioned the possible subtleties which may potentially falsify the existence of a dS/CFT correspondence. One of them is the metastability of de Sitter space. As eternal inflation populates the landscape, what appears in the asymptotic future may be a fractal of FRW universes in nucleated bubbles, instead of a flat Euclidean space on which we can define a Euclidean CFT. Since our construction is state–dependent, it refers to a certain background spacetime. Thus the nucleation of bubbles could potentially modify the construction prescription dramatically. However from the point of view of a FRW universe as a semi–classical background, the notion of microcausality should be well–defined. A proposal for a holographic description of an FRW universe in the Coleman–de Luccia bubble is described in [49], aiming to giving a holographic description to eternal inflation. Thus one may think about how to formulate local bulk physics in terms of the dual data in such a proposal.

All the constructions that we presented in Chapter Two are about empty de Sitter space in certain vacuum states. One can also think about black holes in de Sitter space. The construction of local operators behind the event horizon of an eternal black hole in anti de Sitter space has been performed by several authors, such as in [55], where the construction is established for the BTZ black hole in AdS_{2+1} as a special case, and a local

[1] The de Sitter slicing of de Sitter static patch refers to the coordinate system that covers the static patch of dS_{d+1} with slices of dS_d, the central slice is the slice that passes through the center of the static patch. The central worldline is the worldline with respect to which one defines the static patch at hand.

operator inside a black hole is shown to involve operators in both a conformal field theory and its thermo–field double. In [32] the construction is generalized to an eternal black hole in AdS of generic dimensions. The construction data one uses are from a single copy of CFT, established with the help of "mirror operators". It would be interesting to think about black holes in de Sitter space, how the structure of a CFT with its thermo–field double is realized in a de Sitter black hole, and how observables behind the black hole horizon can be constructed.

In Chapter three we discussed the existence of residual gauge transformations in the theory of large scale structure, and their implications for the LSS correlation functions, especially correlation functions involving overdensities. Here is a summary of the main results in this chapter:

Consistency relations connect the soft limit of $N+1$-point correlation functions with N-point functions; they are the analogs of soft-pion theorems for cosmology. The physics of soft pions is governed by chiral symmetry, and similarly consistency relations are derived as Ward identities for a set of spacetime coordinate transformations. These transformations are diffeomorphisms that are not nailed by the gauge conditions. In ζ-gauge, schematically the consistency relations take the form:

$$\lim_{\vec{q}\to 0}\partial_q^n\left[\frac{\langle\zeta_{\vec{q}}\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}}{P_\zeta(q)}+\Theta(n-1)\frac{\langle\gamma_{\vec{q}}\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}}{P_\gamma(q)}\right]\sim\partial_k^n\langle\mathcal{O}_{\vec{k}_1}...\mathcal{O}_{\vec{k}_N}\rangle^{c'}\,, \tag{4.1}$$

where ζ and γ are the (soft) curvature and tensor perturbations, P_ζ and P_γ are their respective power spectra, and $\mathcal{O}$ represents some observables of interest at hard momenta $\vec{k}_1,...\vec{k}_N$. Higher n corresponds to high powers of x in the corresponding symmetry transformation and higher powers of soft momentum in the consistency relation.

The fact that consistency relations arise from symmetries makes them robust against some of the complex dynamics due to gravitational lumping, which makes these relations interesting in the study of large scale structure. We work in the sub-Hubble scale in Newtonian gauge and find a tower of symmetries and corresponding consistency relations, schematically these relations are:

$$\begin{aligned} &\lim_{\vec{q}\to 0} \partial_q^n \left[\mathcal{H}^{-1} \frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} + \Theta(n-1) \frac{\langle \gamma_{\vec{q}} \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\gamma(q)} \right] \\ &\sim \left[\partial_k^n + n\mathcal{H}^{-2} \partial_k^{n-2} \right] \langle \mathcal{O}_{\vec{k}_1} ... \mathcal{O}_{\vec{k}_N} \rangle^{c'} , \end{aligned} \tag{4.2}$$

where $\pi_{\vec{q}}$ is the soft velocity potential. The precise form is given in Eq. (3.118).

These consistency relations are not based on assumptions on specific dynamics of the hard modes, but they are not completely free of assumptions. There are three underlying assumptions backing the consistency relations: the existence of nonlinearly realized symmetries, the single field initial condition, and the adiabatic mode condition. The adiabatic mode condition states that a zero-momentum mode generated by a symmetry transformation is physical only when it satisfies the equations of the system at a small but finite momentum. This means that in a universe with multiple particle species, all the particles must move with the same velocity on large scales, implying that the equivalence principle is obeyed in the large scale, while in the short scale different matter contents can move differently. Only in this way can the large scale motion of the matter fluid be reproduced by a single soft mode which obeys the equations of motion.

Another subtlety is that in consistency relations the soft momentum q is taken to zero and thus should be super-Hubble in the strict sense. We are interested in the sub-Hubble dynamics and would like to take both hard and soft modes within the Hubble scale: $\mathcal{H} < q \ll \vec{k}_a$. This is not always possible since it is not always true that the soft mode has the same time dependence in and out of the Hubble scale. If the soft mode is the velocity potential π it is true that sub-Hubble and super-Hubble modes have the same time dependence and thus the consistency relations involving soft π can be pushed inside the Hubble scale. For soft overdensity mode δ and soft tensor γ this is not in general true.

Though the LSS correlation functions involve hard modes that can be well inside the nonlinear scale and the physics at short scale can be highly non-perturbative and astrophysically complex, to derive the symmetries and consistency relations, the only thing we need to know is how the observables transform. Though independent of the short scale physics of galaxy formation and mergers, the consistency relations do require that in the large scale all objects fall in the same way. Thus a violation of consistency relations, if observed, may lead to the violation of equivalence principle in the large scale.

Further we showed that the consistency relation takes a particularly simple form in Lagrangian space: the squeezed correlation function with a soft displacement field vanishes (Eqs. 3.124):

$$\lim_{\mathbf{p}\to 0} \frac{\langle \mathbf{\Delta}_{\mathbf{p}}(\eta)\, \mathcal{O}_{\mathbf{p}_1}(\eta_1) \ldots \mathcal{O}_{\mathbf{p}_N}(\eta_N)\, \rangle^{c'}}{P_{\Delta}(\mathrm{p}, \eta)} = 0\,, \tag{4.3}$$

where Δ is the displacement field in Lagrangian space, and $\mathcal{O}$ can be different observables such as mass or galaxy density; the quantities can be at different times, and $\mathbf{p}, \mathbf{p}_1, \mathbf{p}_2, \ldots$ label the momenta with $\mathbf{p}$ being the soft one.[2] The derivation given in 3.2.2.2 is fully non-perturbative and is valid even in the presence of multiple-streaming. It follows the derivation in an earlier paper [11], which relates an $(N+1)-$point function to the linear transformation of an $N-$point function, for a general nonlinearly-realized symmetry (Eq. 3.133). The reason why the right-hand side of the consistency relation vanishes and we have a particularly simple form is that the symmetry shifts the displacement field without transforming the Lagrangian coordinate $\mathbf{q}$ (Eq. 3.136):

$$\mathbf{q} \to \mathbf{q} \quad , \quad \mathbf{\Delta} \to \mathbf{\Delta} + \mathbf{n}(\eta) \quad , \quad \Phi \to \Phi - \left(\mathbf{n}'' + \frac{a'}{a}\mathbf{n}'\right) \cdot \mathbf{x}\,, \tag{4.4}$$

[2]See also Eq. 3.122 with velocity $\boldsymbol{v}$ as the soft mode.

where $\mathbf{\Delta}$ is the displacement, $\mathbf{n}$ is some function of time, Φ is the gravitational potential, a is the scale factor and $\mathbf{x}$ is the Eulerian coordinate.[3] Thus, observables $\mathcal{O}$ such as the mass density or the galaxy density[4] , when expressed as functions of the Lagrangian coordinate $\mathbf{q}$, do not receive linear transformations, and so the right-hand side of the master formula vanishes. [5] This contrasts with what happens when these observables are thought of as functions of the Eulerian coordinate $\mathbf{x}$. They then receive linear transformations because under the symmetry transformation, the Eulerian coordinate shifts:

$$\mathbf{x} \to \mathbf{x} + \mathbf{n}(\eta)\,. \tag{4.5}$$

Though the consistency relation we study can be written in a very simple, even trivial form in Lagrangian space, it does not mean that it is free of physical content . The consistency relation can be violated if certain physical conditions are not met, such as if the initial condition were not of the single-field/clock type. Rather, the simplicity suggests an analytic understanding of nonlinear clustering is perhaps more promising in Lagrangian space. This view has a long history, starting from Zeldovich [103]. What is interesting is that the consistency relation, by virtue of its being a symmetry statement, is non-perturbative, and thus goes beyond perturbation treatments such as the Zeldovich approximation.

In reality, observations are performed in Eulerian space, not Lagrangian space. At the nonlinear level, the relation between the two descriptions is complex. Our derivation of the consistency relation in Eulerian space from its counterpart in Lagrangian space requires a formal series expansion in the displacement Δ. In relating the two descriptions,

[3] If there are multiple species present such as dark matter and galaxies, the same transformation applies to the displacement of all species. See Eq. (3.140).

[4] See footnote 35 in Chapter 3.

[5] Note that even quantities such as Δ or Φ have no linear shift (a shift that is linear in fluctuation variables). More complicated observables could have a linear shift; see footnote 37 in Chapter 3.

the expansion in Δ is done in an uneven manner: only phase factors such as $e^{i\mathbf{p}_1\cdot\boldsymbol{\Delta}}$ are expanded even though other variables, such as the density δ, also depend on the displacement. This is not unexpected in relations that are supposed to be non-perturbative – *partial* resummation of perturbations is often a useful technique. It would be interesting if this particular example we concern can lead to a more general resummation scheme.

A natural question is whether there are relativistic generalizations of statements like Eq. (4.3) – consistency relations with a vanishing right hand side. The Lagrangian coordinate (attached to dark matter particles) is essentially the freely-falling coordinate. Indeed [106] showed that using the freely-falling coordinate, the dilation consistency relation [3] can be rewritten in a similarly simple form (see also [107]). Their derivation is perturbative. It should be possible to extend their proof using the non-perturbative arguments presented here. More generally, it would be interesting to see if further general relativistic consistency relations, such as those found by [5], can also be recast in this fashion. The hope is to find the infinite tower of symmetries and consistency relations in Lagrangian space and further to see if it is possible to resum these relations into a Ward-Takahashi identity for gauge symmetries.

Bibliography

[1] Strominger, Andrew. "The dS/CFT correspondence." Journal of High Energy Physics 2001.10 (2001): 034.

[2] Anninos, Dionysios, Thomas Hartman, and Andrew Strominger. "Higher spin realization of the dS/CFT correspondence." arXiv preprint arXiv:1108.5735 (2011).

[3] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," JHEP **0305**, 013 (2003) [astro-ph/0210603].

[4] K. Hinterbichler, L. Hui and J. Khoury, "Conformal Symmetries of Adiabatic Modes in Cosmology," JCAP **1208**, 017 (2012) [arXiv:1203.6351 [hep-th]].

[5] (H2K) K. Hinterbichler, L. Hui and J. Khoury, "An Infinite Set of Ward Identities for Adiabatic Modes in Cosmology," arXiv:1304.5527 [hep-th].

[6] G. L. Pimentel, "Inflationary Consistency Conditions from a Wavefunctional Perspective," JHEP **1402**, 124 (2014) [arXiv:1309.1793 [hep-th]].

[7] L. Berezhiani and J. Khoury, "Slavnov-Taylor Identities for Primordial Perturbations," JCAP **1402**, 003 (2014) [arXiv:1309.4461 [hep-th]].

[8] A. Kehagias and A. Riotto, Nucl. Phys. B **864**, 492 (2012) [arXiv:1205.1523 [hep-th]].

[9] M. Peloso and M. Pietroni, "Ward identities and consistency relations for the large scale structure with multiple species," arXiv:1310.7915 [astro-ph.CO].

[10] P. Creminelli, J. Noreña, M. Simonović and F. Vernizzi, "Single-Field Consistency Relations of Large Scale Structure," JCAP **1312**, 025 (2013) [arXiv:1309.3557 [astro-ph.CO]].

[11] B. Horn, L. Hui and X. Xiao, JCAP **1409**, no. 09, 044 (2014) [arXiv:1406.0842 [hep-th]].

[12] B. Horn, L. Hui and X. Xiao, arXiv:1502.06980 [hep-th].

[13] X. Xiao, Phys. Rev. D **90**, no. 2, 024061 (2014) [arXiv:1402.7080 [hep-th]].

[14] D. Sarkar and X. Xiao, Phys. Rev. D **91**, no. 8, 086004 (2015) [arXiv:1411.4657 [hep-th]].

[15] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2**, 231 (1998) [hep-th/9711200].

[16] I. Heemskerk, J. Penedones, J. Polchinski and J. Sully, "Holography from Conformal Field Theory," JHEP **0910**, 079 (2009) [arXiv:0907.0151 [hep-th]].

[17] A. L. Fitzpatrick and J. Kaplan, "AdS Field Theory from Conformal Field Theory," JHEP **1302**, 054 (2013) [arXiv:1208.0337 [hep-th]].

[18] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, "Black Holes: Complementarity or Firewalls?," JHEP **1302**, 062 (2013) [arXiv:1207.3123 [hep-th]].

[19] E. Witten, "Anti-de Sitter space and holography," Adv. Theor. Math. Phys. **2**, 253 (1998) [hep-th/9802150].

[20] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," Phys. Lett. B **428**, 105 (1998) [hep-th/9802109].

[21] T. Banks, M. R. Douglas, G. T. Horowitz and E. J. Martinec, "AdS dynamics from conformal field theory," hep-th/9808016.

[22] D. Harlow and D. Stanford, "Operator Dictionaries and Wave Functions in AdS/CFT and dS/CFT," arXiv:1104.2621 [hep-th].

[23] D. Harlow and L. Susskind, "Crunches, Hats, and a Conjecture," arXiv:1012.5302 [hep-th].

[24] D. Anninos, G. S. Ng and A. Strominger, "Future Boundary Conditions in De Sitter Space," JHEP **1202**, 032 (2012) [arXiv:1106.1175 [hep-th]].

[25] I. Heemskerk and J. Polchinski, "Holographic and Wilsonian Renormalization Groups," JHEP **1106**, 031 (2011) [arXiv:1010.1264 [hep-th]].

[26] R. Bousso, A. Maloney and A. Strominger, "Conformal vacua and entropy in de Sitter space," Phys. Rev. D **65**, 104039 (2002) [hep-th/0112218].

[27] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," JHEP **0305**, 013 (2003) [astro-ph/0210603].

[28] A. Hamilton, D. N. Kabat, G. Lifschytz and D. A. Lowe, "Local bulk operators in AdS/CFT: A Boundary view of horizons and locality," Phys. Rev. D **73**, 086003 (2006) [hep-th/0506118].

[29] D. Kabat, G. Lifschytz and D. A. Lowe, "Constructing local bulk observables in interacting AdS/CFT," Phys. Rev. D **83**, 106009 (2011) [arXiv:1102.2910 [hep-th]]. Alishahiha, M., Karch, A., Silverstein, E., Tong, D. (2004). The dS/dS correspondence. arXiv preprint hep-th/0407125.

[30] B. S. DeWitt, "Quantum Theory of Gravity. 2. The Manifestly Covariant Theory," Phys. Rev. **162**, 1195 (1967).

[31] I. Heemskerk, D. Marolf, J. Polchinski and J. Sully, "Bulk and Transhorizon Measurements in AdS/CFT," JHEP **1210**, 165 (2012) [arXiv:1201.3664 [hep-th]].

[32] K. Papadodimas and S. Raju, "State-Dependent Bulk-Boundary Maps and Black Hole Complementarity," arXiv:1310.6335 [hep-th].

[33] M. Alishahiha, A. Karch, E. Silverstein and D. Tong, "The dS/dS correspondence," AIP Conf. Proc. **743**, 393 (2005) [hep-th/0407125].

[34] D. Anninos, S. A. Hartnoll and D. M. Hofman, "Static Patch Solipsism: Conformal Symmetry of the de Sitter Worldline," Class. Quant. Grav. **29**, 075002 (2012) [arXiv:1109.4942 [hep-th]].

[35] M. A. Vasiliev, "Higher spin gauge theories in four-dimensions, three-dimensions, and two-dimensions," Int. J. Mod. Phys. D **5**, 763 (1996) [hep-th/9611024].

[36] N. Goheer, M. Kleban and L. Susskind, "The Trouble with de Sitter space," JHEP **0307**, 056 (2003) [hep-th/0212209].

[37] D. Anninos, "De Sitter Musings," Int. J. Mod. Phys. A **27**, 1230013 (2012) [arXiv:1205.3855 [hep-th]].

[38] S. Ferrara, A. F. Grillo, G. Parisi and R. Gatto, "The shadow operator formalism for conformal algebra. vacuum expectation values and operator products," Lett. Nuovo Cim. **4S2**, 115 (1972) [Lett. Nuovo Cim. **4**, 115 (1972)].

[39] D. Simmons-Duffin, "Projectors, Shadows, and Conformal Blocks," arXiv:1204.3894 [hep-th].

[40] A. L. Fitzpatrick, J. Kaplan, Z. U. Khandker, D. Li, D. Poland and D. Simmons-Duffin, "Covariant Approaches to Superconformal Blocks," arXiv:1402.1167 [hep-th].

[41] I. Heemskerk, "Construction of Bulk Fields with Gauge Redundancy," JHEP **1209**, 106 (2012) [arXiv:1201.3666 [hep-th]].

[42] D. Kabat, G. Lifschytz, S. Roy and D. Sarkar, "Holographic representation of bulk fields with spin in AdS/CFT," Phys. Rev. D **86**, 026004 (2012) [arXiv:1204.0126 [hep-th]].

[43] P. A. M. Dirac, "Wave equations in conformal space," Annals Math. **37**, 429 (1936).

[44] G. Mack and A. Salam, "Finite component field representations of the conformal group," Annals Phys. **53**, 174 (1969).

[45] S. Weinberg, "Six-dimensional Methods for Four-dimensional Conformal Field Theories," Phys. Rev. D **82**, 045031 (2010) [arXiv:1006.3480 [hep-th]].

[46] W. D. Goldberger, W. Skiba and M. Son, "Superembedding Methods for 4d N=1 SCFTs," Phys. Rev. D **86**, 025019 (2012) [arXiv:1112.0325 [hep-th]].

[47] W. D. Goldberger, Z. U. Khandker, D. Li and W. Skiba, "Superembedding Methods for Current Superfields," Phys. Rev. D **88**, 125010 (2013) [arXiv:1211.3713 [hep-th]].

[48] X. Dong, B. Horn, E. Silverstein and G. Torroba, "Micromanaging de Sitter holography," Class. Quant. Grav. **27**, 245020 (2010) [arXiv:1005.5403 [hep-th]].

[49] B. Freivogel, Y. Sekino, L. Susskind and C. -P. Yeh, "A Holographic framework for eternal inflation," Phys. Rev. D **74**, 086003 (2006) [hep-th/0606204].

[50] H. Bondi, M. G. J. van der Burg and A. W. K. Metzner, “Gravitational waves in general relativity. 7. Waves from axisymmetric isolated systems,” Proc. Roy. Soc. Lond. A **269**, 21 (1962).

[51] R. K. Sachs, “Gravitational waves in general relativity. 8. Waves in asymptotically flat space-times,” Proc. Roy. Soc. Lond. A **270**, 103 (1962).

[52] A. Strominger, “Asymptotic Symmetries of Yang-Mills Theory,” arXiv:1308.0589 [hep-th].

[53] T. He, V. Lysov, P. Mitra and A. Strominger, “BMS supertranslations and Weinberg’s soft graviton theorem,” arXiv:1401.7026 [hep-th].

[54] E. Witten, “Quantum gravity in de Sitter space,” hep-th/0106109.

[55] A. Hamilton, D. N. Kabat, G. Lifschytz and D. A. Lowe, “Local bulk operators in AdS/CFT: A Holographic description of the black hole interior,” Phys. Rev. D **75**, 106001 (2007) [Erratum-ibid. D **75**, 129902 (2007)] [hep-th/0612053].

[56] Balasubramanian, Vijay, Per Kraus, and Albion Lawrence. “Bulk versus boundary dynamics in antide Sitter spacetime.” Physical Review D 59.4 (1999): 046003.

[57] Bena, Iosif. “Construction of local fields in the bulk of AdS_5 and other spaces.” Physical Review D 62.6 (2000): 066007.

[58] Klebanov, Igor R., and A. M. Polyakov. “AdS dual of the critical O(N) vector model.” Physics Letters B 550.3 (2002): 213-219.

[59] D. Sarkar, “(A)dS Holography with a Cut-off,” Phys. Rev. D **90**, 086005 (2014) [arXiv:1408.0415 [hep-th]].

[60] I. Heemskerk, J. Penedones, J. Polchinski and J. Sully, "Holography from Conformal Field Theory," JHEP **0910**, 079 (2009) [arXiv:0907.0151 [hep-th]].

[61] A. L. Fitzpatrick and J. Kaplan, "AdS Field Theory from Conformal Field Theory," JHEP **1302**, 054 (2013) [arXiv:1208.0337 [hep-th]].

[62] Mikhailov, Andrei. "Notes on higher spin symmetries." arXiv preprint hep-th/0201019 (2002).

[63] Giombi, Simone, and Xi Yin. "Higher spin gauge theory and holography: the three-point functions." Journal of High Energy Physics 2010.9 (2010): 1-80.

[64] Y. S. Stanev, "Correlation Functions of Conserved Currents in Four Dimensional Conformal Field Theory," Nucl. Phys. B **865**, 200 (2012) [arXiv:1206.5639 [hep-th]].

[65] Kabat, Daniel, and Gilad Lifschytz. "CFT representation of interacting bulk gauge fields in AdS." Physical Review D 87.8 (2013): 086004.

[66] D. Kabat and G. Lifschytz, "Decoding the hologram: Scalar fields interacting with gravity," arXiv:1311.3020 [hep-th].

[67] A. Zhiboedov, "A note on three-point functions of conserved currents," arXiv:1206.6370 [hep-th].

[68] M. S. Costa, J. Penedones, D. Poland and S. Rychkov, "Spinning Conformal Correlators," JHEP **1111**, 071 (2011) [arXiv:1107.3554 [hep-th]].

[69] V. Balasubramanian, P. Kraus and A. E. Lawrence, "Bulk versus boundary dynamics in anti-de Sitter space-time," Phys. Rev. D **59**, 046003 (1999) [hep-th/9805171].

[70] I. Heemskerk, D. Marolf, J. Polchinski and J. Sully, "Bulk and Transhorizon Measurements in AdS/CFT," JHEP **1210**, 165 (2012) [arXiv:1201.3664 [hep-th]].

[71] B. de Wit and D. Z. Freedman, "Systematics of Higher Spin Gauge Fields," Phys. Rev. D **21**, 358 (1980).

[72] S. Giombi, S. Prakash and X. Yin, "A Note on CFT Correlators in Three Dimensions," JHEP **1307**, 105 (2013) [arXiv:1104.4317 [hep-th]].

[73] J. Maldacena and A. Zhiboedov, "Constraining Conformal Field Theories with A Higher Spin Symmetry," J. Phys. A **46**, 214011 (2013) [arXiv:1112.1016 [hep-th]].

[74] C. M. Chang, S. Minwalla, T. Sharma and X. Yin, "ABJ Triality: from Higher Spin Fields to Strings," J. Phys. A **46**, 214009 (2013) [arXiv:1207.4485 [hep-th]].

[75] M. R. Gaberdiel and R. Gopakumar, "Higher Spins & Strings," arXiv:1406.6103 [hep-th].

[76] Coleman, Sidney. Aspects of symmetry: selected Erice lectures. Cambridge University Press, 1988.

[77] A. Kehagias and A. Riotto, "Symmetries and Consistency Relations in the Large Scale Structure of the Universe," Nucl. Phys. B **873**, 514 (2013) [arXiv:1302.0130 [astro-ph.CO]].

[78] M. Peloso and M. Pietroni, "Galilean invariance and the consistency relation for the nonlinear squeezed bispectrum of large scale structure," JCAP **1305**, 031 (2013) [arXiv:1302.0223 [astro-ph.CO]].

[79] P. Creminelli and M. Zaldarriaga, "Single field consistency relation for the 3-point function," JCAP **0410**, 006 (2004) [astro-ph/0407059].

[80] P. Creminelli, J. Norena and M. Simonović, "Conformal consistency relations for single-field inflation," JCAP **1207**, 052 (2012) [arXiv:1203.4595 [hep-th]].

[81] V. Assassi, D. Baumann and D. Green, "On Soft Limits of Inflationary Correlation Functions," JCAP **1211**, 047 (2012) [arXiv:1204.4207 [hep-th]].

[82] V. Assassi, D. Baumann and D. Green, "Symmetries and Loops in Inflation," JHEP **1302**, 151 (2013) [arXiv:1210.7792 [hep-th]].

[83] W. D. Goldberger, L. Hui and A. Nicolis, "One-particle-irreducible consistency relations for cosmological perturbations," Phys. Rev. D **87**, 103520 (2013) [arXiv:1303.1193 [hep-th]].

[84] S. Weinberg, "Adiabatic modes in cosmology," Phys. Rev. D **67**, 123504 (2003) [astro-ph/0302326].

[85] L. Hui, A. Nicolis and C. Stubbs, "Equivalence principle implications of modified gravity models," Phys. Rev. D **80** 104002 (2009) [arXiv:0905.2966 [astro-ph.CO]].

[86] L. Hui and A. Nicolis, "An Equivalence principle for scalar forces," Phys. Rev. Lett. **105**, 231101 (2010) [arXiv:1009.2520 [hep-th]].

[87] L. Boubekeur, P. Creminelli, J. Norena and F. Vernizzi, "Action approach to cosmological perturbations: the 2nd order metric in matter dominance," JCAP **0808**, 028 (2008) [arXiv:0806.1016 [astro-ph]].

[88] P. Creminelli, G. D'Amico, J. Norena and F. Vernizzi, "The Effective Theory of Quintessence: the w¡-1 Side Unveiled," JCAP **0902**, 018 (2009) [arXiv:0811.0827 [astro-ph]].

[89] A. Kehagias, J. Norea, H. Perrier and A. Riotto, "Consequences of Symmetries and Consistency Relations in the Large-Scale Structure of the Universe for Non-local bias and Modified Gravity," arXiv:1311.0786 [astro-ph.CO].

[90] A. Kehagias, H. Perrier and A. Riotto, “Equal-time Consistency Relations in the Large-Scale Structure of the Universe,” arXiv:1311.5524 [astro-ph.CO].

[91] P. Creminelli, Jrm. Gleyzes, M. Simonović and F. Vernizzi, “Single-Field Consistency Relations of Large Scale Structure. Part II: Resummation and Redshift Space,” arXiv:1311.0290 [astro-ph.CO].

[92] P. Creminelli, Jrm. Gleyzes, L. Hui, M. Simonović and F. Vernizzi, “Single-Field Consistency Relations of Large Scale Structure. Part III: Test of the Equivalence Principle,” arXiv:1312.6074 [astro-ph.CO].

[93] P. Valageas, “Angular averaged consistency relations of large-scale structures,” arXiv:1311.4286 [astro-ph.CO].

[94] P. Valageas, “Consistency relations of large-scale structures,” arXiv:1311.1236 [astro-ph.CO].

[95] R. K. Sheth, K. C. Chan and R. Scoccimarro, “Non-local Lagrangian bias,” Phys. Rev. D **87**, 083002 (2013) [arXiv:1207.7117 [astro-ph.CO]].

[96] K. C. Chan, R. Scoccimarro and R. K. Sheth, “Gravity and Large-Scale Non-local Bias,” Phys. Rev. D **85**, 083509 (2012) [arXiv:1201.3614 [astro-ph.CO]].

[97] S. Dubovsky, T. Gregoire, A. Nicolis and R. Rattazzi, “Null energy condition and superluminal propagation,” JHEP **0603**, 025 (2006) [hep-th/0512260].

[98] M. Crocce and R. Scoccimarro, “Renormalized cosmological perturbation theory,” Phys. Rev. D **73**, 063519 (2006) [astro-ph/0509418].

[99] S. Dodelson, “Modern cosmology,” Amsterdam, Netherlands: Academic Pr. (2003) 440 p

[100] A. Kehagias, J. Noreña, H. Perrier and A. Riotto, "Consequences of Symmetries and Consistency Relations in the Large-Scale Structure of the Universe for Non-local bias and Modified Gravity," Nucl. Phys. B **883**, 83 (2014) [arXiv:1311.0786 [astro-ph.CO]].

[101] P. Berger, A. Kehagias and A. Riotto, "Testing the Origin of Cosmological Magnetic Fields through the Large-Scale Structure Consistency Relations," JCAP **1405**, 025 (2014) [arXiv:1402.1044 [astro-ph.CO]].

[102] F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, "Large scale structure of the universe and cosmological perturbation theory," Phys. Rept. **367**, 1 (2002) [astro-ph/0112551].

[103] Y. B. Zeldovich, "Gravitational instability: An Approximate theory for large density perturbations," Astron. Astrophys. **5**, 84 (1970).

[104] R. A. Porto, L. Senatore and M. Zaldarriaga, "The Lagrangian-space Effective Field Theory of Large Scale Structures," JCAP **1405**, 022 (2014) [arXiv:1311.2168 [astro-ph.CO]].

[105] L. Senatore and M. Zaldarriaga, "The IR-resummed Effective Field Theory of Large Scale Structures," arXiv:1404.5954 [astro-ph.CO].

[106] E. Pajer, F. Schmidt and M. Zaldarriaga, "The Observed Squeezed Limit of Cosmological Three-Point Functions," Phys. Rev. D **88**, no. 8, 083502 (2013) [arXiv:1305.0824 [astro-ph.CO]].

[107] T. Tanaka and Y. Urakawa, "Dominance of gauge artifact in the consistency relation for the primordial bispectrum," JCAP **1105**, 014 (2011) [arXiv:1103.1251 [astro-ph.CO]].

Appendix A

Appendix for holographic representation of operators

A.1 Integration

Here we evaluate the integral

$$f(\alpha,\beta) = \int_{|\mathbf{x}'|<\eta} d^d x' \left(\frac{\eta^2 - \mathbf{x}'^2}{\eta}\right)^\alpha \frac{1}{(\mathbf{x}+\mathbf{x}')^{2\beta}} \tag{A.1}$$

For convenience we make the choice $x^1 = |\mathbf{x}| \equiv R$, $x^2 = \cdots = x^d = 0$, thus we have

$$f(\alpha,\beta) = Vol(S^{d-2}) \int_0^\eta dr r^{d-1} \left(\frac{\eta^2 - r^2}{\eta}\right)^\alpha \int_0^\pi \frac{\sin^{d-2}\theta d\theta}{(R^2 + 2Rr\cos\theta + r^2)^\beta} \tag{A.2}$$

where

$$Vol(S^{d-2}) = \frac{2\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2})} \tag{A.3}$$

Using the formulae:

$$\int_0^\pi \frac{\sin^{2\mu-1}\theta}{(1+2a\cos\theta+a^2)^\nu}d\theta = \frac{\Gamma(\mu)\Gamma(\frac{1}{2})}{\Gamma(\mu+\frac{1}{2})}F\left(\nu,\nu-\mu+\frac{1}{2},\mu+\frac{1}{2},a^2\right)$$
$$\int_0^1 (1-x)^{\mu-1}x^{\gamma-1}F(\alpha,\beta,\gamma,ax)dx = \frac{\Gamma(\mu)\Gamma(\gamma)}{\Gamma(\mu+\gamma)}F(\alpha,\beta,\gamma+\mu,a)$$

we then have

$$f(\alpha,\beta) = \frac{\pi^{\frac{d}{2}}\Gamma(\alpha+1)}{\Gamma(\alpha+\frac{d}{2}+1)}\frac{\eta^{\alpha+d}}{|\mathbf{x}|^{2\beta}}F\left(\beta,\beta-\frac{d}{2}+1,\alpha+\frac{d}{2}+1,\frac{\eta^2}{\mathbf{x}^2}\right) \tag{A.4}$$

Also to get the near–boundary two–point Wightman function we need the following property of the hypergeometric function:

$$F(\alpha,\beta,\beta,z) = (1-z)^{-\alpha} \tag{A.5}$$

A.2 Smearing Function from Green's Function

In de Sitter space, an operator near the past boundary can be expressed by operators in the CFT:

$$\Phi(\eta\to 0,\mathbf{x}) \sim \eta^\Delta\mathcal{O}_+ + \eta^{d-\Delta}\mathcal{O}_- \tag{A.6}$$

Then if we want to probe deeper into de Sitter space, we need something like

$$\Phi(\eta,\mathbf{x}) = \int K_+(\eta,\mathbf{x}|\mathbf{x}')\mathcal{O}_+(\mathbf{x}') + \int K_-(\eta,\mathbf{x}|\mathbf{x}')\mathcal{O}_-(\mathbf{x}') \tag{A.7}$$

Here it is important that we have both components, which means that the correct construction of a local operator in de Sitter space is not an analytic continuation from anti de Sitter space; otherwise, we would spoil microcausality.

Here the bulk operator is linked to the boundary CFT operators by a retarded Green's function defined as

$$G_{ret}(x,x') \equiv G_E(x,x') - G_E(x',x) \tag{A.8}$$

with G_E being the Wightman function in Euclidean vacuum.

Here we just need the asymptotic form of G_{ret} in the limit $\eta' \to 0$:

$$G_{ret}|_{\eta'\to 0} \sim c_{\Delta,d}(-\sigma - i\epsilon)^{\Delta-d} + c^*_{\Delta,d}(-\sigma - i\epsilon)^{-\Delta} - c.c \tag{A.9}$$

The bulk operator is evaluated by:

$$\Phi(\eta,\mathbf{x}) = \int_{|\mathbf{x}'|<\eta} d^d x' \left(\frac{1}{\eta'}\right)^{d-1} \left(G_{ret}(\eta,\mathbf{x};\eta',\mathbf{x}')\partial_{\eta'}\Phi(\eta',\mathbf{x}') - \Phi(\eta',\mathbf{x}')\partial_{\eta'}G_{ret}(\eta,\mathbf{x};\eta',\mathbf{x}')\right) \tag{A.10}$$

Taking a partial derivative on the retarded Green's function and working in the small η' limit, we have

$$\partial_{\eta'}G_{ret} \sim \frac{1}{\eta'}\left(c(\Delta-d)(-\sigma-i\epsilon)^{\Delta-d} - c^*\Delta(-\sigma-i\epsilon)^{-\Delta} + c^*\Delta(-\sigma+i\epsilon)^{-\Delta} - c(\Delta-d)(-\sigma+i\epsilon)^{\Delta-d}\right) \tag{A.11}$$

Therefore for $\Phi_+ \sim \eta^\Delta \mathcal{O}_+$ we have

$$\left(\frac{1}{\eta'}\right)^{d-1} \left[\Phi_+(\eta',\mathbf{x}')\partial_{\eta'}G_{ret}(\eta,\mathbf{x}|\eta',\mathbf{x}') - G_{ret}(\eta,\mathbf{x}|\eta',\mathbf{x}')\partial_{\eta'}\Phi_+(\eta',\mathbf{x}')\right]$$
$$= (\eta')^{\Delta-d}\left[cd\left((-\sigma+i\epsilon)^{\Delta-d} - (-\sigma-i\epsilon)^{\Delta-d}\right) + 2c^*\Delta\left((-\sigma+i\epsilon)^{-\Delta} - (-\sigma-i\epsilon)^{-\Delta}\right)\right]\mathcal{O}_+$$

Here the factor $(\eta')^{\Delta-d}$ cancels with the factor of η' with the inverse power from $\sigma^{\Delta-d}$ and gives a well-defined limit when $\eta' \to 0$, but it doesn't cancel with the factor in $\sigma^{-\Delta}$, leading to a fast oscillation when $\eta' \to 0$ so the term proportional to $\sigma^{-\Delta}$ vanishes.

For $\Phi_- \sim \eta^{d-\Delta}\mathcal{O}_-$ we have

$$\left(\frac{1}{\eta'}\right)^{d-1}\left[\Phi_-(\eta',\mathbf{x}')\partial_{\eta'}G_{ret}(\eta,\mathbf{x}|\eta',\mathbf{x}') - G_{ret}(\eta,\mathbf{x}|\eta',\mathbf{x}')\partial_{\eta'}\Phi_-(\eta',\mathbf{x}')\right]$$
$$= \eta'^{-\Delta}\left[2c(\Delta-d)\left((-\sigma-i\epsilon)^{\Delta-d}-(-\sigma+i\epsilon)^{\Delta-d}\right)+c^*d\left((-\sigma+i\epsilon)^{-\Delta}-(-\sigma-i\epsilon)^{-\Delta}\right)\right]\mathcal{O}_-$$

Similarly we only have the contribution from the terms proportional to $\sigma^{-\Delta}$

To evaluate the integration kernel, we notice that outside the bulk lightcone $\sigma \propto \eta^2 - (\mathbf{x}-\mathbf{x}')^2 < 0$ so the ϵ prescription can be dropped and the integral gives a vanishing result. When we analytically continue the result into the bulk lightcone, the ϵ prescription will give a phase shift proportional to $Im\,(\Delta - d)$ and $Im\,(-\Delta)$ respectively. For instance, in

$$(-\sigma + i\epsilon)^{\Delta-d} \tag{A.12}$$

the cut starts from $\sigma = i\epsilon$. To analytically continue we go under the branch point and thus get a phase $e^{-i\pi\left(i\sqrt{m^2-\frac{d^2}{4}}\right)}$ and therefore

$$(-\sigma+i\epsilon)^{\Delta-d} - (-\sigma-i\epsilon)^{\Delta-d} = -2ie^{i\pi(\Delta-d)}\sin\left(\pi i\left(\Delta-d\right)\right)\sigma^{\Delta-d} \tag{A.13}$$

In this way for $\mathcal{O}_+$ we have a smearing function proportional to $(\sigma\eta')^{\delta-d}$ and for $\mathcal{O}_-$ we have a smearing function proportional to $(\sigma\eta')^{-\Delta}$

A.3 A Property of Hypergeometric Function

In finding the asymptotic form of the scalar two–point Wightman function, we find the relation below very useful:

$$\begin{aligned}F(\alpha,\beta,\gamma;z) = &\frac{\Gamma(\gamma)\Gamma(\beta-\alpha)}{\Gamma(\gamma-\alpha)\Gamma(\beta)}(-z)^{-\alpha}F\left(\alpha,\alpha-\gamma+1,\alpha-\beta+1;\frac{1}{z}\right)\\ &+\frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\alpha)\Gamma(\gamma-\beta)}(-z)^{-\beta}F\left(\beta,\beta-\gamma+1,\beta-\alpha+1;\frac{1}{z}\right)\end{aligned}$$

We see here that when $z \to \infty$ the function nicely splits into two parts with behaviors $z^{-\alpha}$ and $z^{-\beta}$. They give two components with different scaling dimensions in the Wightman function.

This relation is true when neither $\alpha - \beta$ nor $\gamma - \alpha - \beta$ is an integer, and is thus applicable to the case when a de Sitter scalar has mass parameter

$$m^2 > \left(\frac{d}{2}\right)^2 \tag{A.14}$$

as well as to light particles with non–integer dimensions. However, for gauge fields the dimensions are integers determined by the spin and spatial dimension, so this property is not applicable.

A.4 Higher Spin Fields in AdS and dS

Here we briefly review some results about general integer spin gauge field in AdS_{d+1}, following [14].

Massless gauge fields in AdS are represented by totally symmetric rank–s tensors

$\Phi_{M_1\dots M_s}$ satisfying double–tracelessness conditions;

$$\Phi^{MN}{}_{MNM_5\dots M_s} = 0 \tag{A.15}$$

The linear equation for a spin–s gauge field on AdS is [35]

$$\nabla_N\nabla^N\Phi_{M_1\dots M_s} - s\nabla_N\nabla_{M_1}\Phi^N{}_{M_2 i\dots M_s} + \frac{1}{2}s(s-1)\nabla_{M_1}\nabla_{M_2}\Phi^N{}_{N\dots M_s} - 2(s-1)(s+d-2)\Phi_{M_1\dots M_3} = 0 \tag{A.16}$$

This equation is invariant under the gauge transformation

$$\Phi_{M_1\dots M_s} \to \Phi_{M_1\dots M_s} + \nabla_{M_1}\Lambda_{M_2\dots M_s}\ ,\ \Lambda^N{}_{NM_3\dots M_s} = 0 \tag{A.17}$$

We can choose the holographic gauge in which all the z–components of the gauge field vanish [14]

$$\Phi_{z\dots z} = \Phi_{\mu_1 z\dots z} = \cdots = \Phi_{\mu_1\dots\mu_{s-1}z} = 0 \tag{A.18}$$

The bulk gauge field is dual to a totally symmetric, traceless, conserved rank–s tensor on the boundary:

$$\mathcal{O}^\nu{}_{\nu\mu_3\dots\mu_s} = 0\ ,\ \partial_\nu\mathcal{O}^\nu{}_{\mu_3\dots\mu_s} = 0 \tag{A.19}$$

Therefore, to be consistent we have to set

$$\Phi^\nu{}_{\nu\mu_3\dots\mu_s} = 0\ ,\ \partial_\nu\Phi^\nu{}_{\mu_3\dots\mu_s} = 0 \tag{A.20}$$

thus we get:

$$\left(\partial_z^2 + \partial_\alpha\partial^\alpha\right)\Phi_{\mu_1\dots\mu_s} + \frac{2s+1-d}{z}\partial_z\Phi_{\mu_1\dots\mu_s} + \frac{2(s-1)(2-d)}{z^2}\Phi_{\mu_1\dots\mu_s} = 0 \tag{A.21}$$

We define

$$Y_{\mu_1 \ldots \mu_s} = z^s \Phi_{\mu_1 \ldots \mu_s} \tag{A.22}$$

as a multiplet of scalars. The equation for $Y_{\mu_1 \ldots \mu_s}$ is

$$\partial_\alpha \partial^\alpha Y_{\mu_1 \ldots \mu_s} + z^{d-1} \partial_z \left(z^{1-d} \partial_z Y_{\mu_1 \ldots \mu_s} \right) - \frac{(s-2)(s+d-2)}{z^2} Y_{\mu_1 \ldots \mu_s} = 0 \tag{A.23}$$

which is just the free scalar equation with mass parameter

$$m^2 = (s-2)(s+d-2) \tag{A.24}$$

corresponding to scaling dimension

$$\Delta = s + d - 2 \tag{A.25}$$

The near–boundary behavior of $Y_{\mu_1 \ldots \mu_s}$ is

$$Y_{\mu_1 \ldots \mu_s} \sim z^\Delta \mathcal{O}_{\mu_1 \ldots \mu_s} \tag{A.26}$$

So one can directly construct the bulk spin–s field:

$$\Phi_{\mu_1 \ldots \mu_s} = \frac{\Gamma\left(s + \frac{d}{2} - 1\right)}{\pi^{\frac{d}{2}} \Gamma\left(s-1\right)} \frac{1}{z^s} \int_{t'^2 + |\mathbf{y}'|^2 < z^2} dt' d^{d-1} y' \left(\frac{z^2 - t'^2 - |\mathbf{y}'|^2}{z} \right)^{s-2} \mathcal{O}_{\mu_1 \ldots \mu_s}(t + t', \mathbf{x} + i \mathbf{y}') \tag{A.27}$$

for fields with integer spin $s > 1$

We see that the field behaves like $z^{\Delta - s} = z^{d-2}$ near the boundary.

The reason why $Y_{\mu_1 \ldots \mu_s}$ turns out to be a scalar is that it is actually the components of the gauge field in a vierbein basis. In AdS we have that

$$e_a{}^{\mu} = z\delta_a{}^{\mu} \tag{A.28}$$

and

$$e_{a_1}{}^{\mu_1} \dots e_{a_s}{}^{\mu_s} \Phi_{\mu_1 \dots \mu_s} \equiv Y_{a_1 \dots a_s} \tag{A.29}$$

are scalars because they don't actually carry any spacetime indices—they are defined with respect to a certain vierbein basis at each point in the spacetime. It is a bit of abuse of the notation not to distinguish $Y_{a_1 \dots a_s}$ and $Y_{\mu_1 \dots \mu_s}$, but at the end of the day we multiply the inverse vierbeins and recover $\Phi_{\mu_1 \dots \mu_s}$ and it does not matter whether we make the vierbeins explicit.

In de Sitter space, the free field equation for $Y_{i_1 \dots i_s}$ is obtained by direct analytic continuation:

$$\ddot{Y}_{i_1 \dots i_s} + \frac{1-d}{\eta} \dot{Y}_{i_1 \dots i_s} + \left(\frac{(2-s)(s+d-2)}{\eta^2} - \partial_j^2 \right) Y_{i_1 \dots i_s} = 0 \tag{A.30}$$

which matches with the generic form of scalar equations in dS:

$$\left(\Box - m^2\right) \phi = 0 \to \ddot{\phi} + \frac{1-d}{\eta} \dot{\phi} + \left(\frac{m^2}{\eta^2} - \partial_j^2 \right) \phi = 0 \tag{A.31}$$

Thus $Y_{i_1 \dots i_s}$ is a free scalar muitiplet in de Sitter space with mass $m^2 = (2-s)(s+d-2)$, with the "dot" meaning a derivative with respect to conformal time η.

Appendix B

Appendix on Consistency Relations

B.1 A Lagrangian for Fluid with Pressure

In this Appendix, we connect the LSS Lagrangian discussed in Sec. 3.1.3.2 with a more commonly used fluid Lagrangian. We will continue to work within the zero vorticity regime. For generalizations to include a non-vanishing vorticity, see [97]. Let us consider the following action:

$$S = \int d^4x \sqrt{-g} \left[\frac{1}{2} M_P^2 R + \mathcal{P}(X) + ... \right] , \tag{B.1}$$

where the first term is the Einstein-Hilbert action, and $\mathcal{P}(X)$ is the fluid action, where $\mathcal{P}$ is some function of $X \equiv \sqrt{-g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi}$ with ϕ describing the single degree of freedom of an irrotational fluid.[1] The ... stands for other possible matter or energy content in the universe, i.e. the background expansion need not be determined solely by the $\mathcal{P}(X)$ fluid in question. This is a completely relativistic action, and has been used by many authors [87, 88]. Our goal here is to take the non-relativistic limit, and connect the result with

[1] The vanishing of vorticity can be expressed covariantly as $\epsilon^{\mu\nu\rho\sigma} u_\nu \partial_\rho u_\sigma = 0$.

the action in sec. 3.1.3.2 (Eq. 3.58).

We assume a metric of the form:

$$ds^2 = a^2\left[-(1+2\Phi)d\eta^2 + (1-2\Psi)d\vec{x}^2\right] . \tag{B.2}$$

The fluid energy-momentum $T_{\mu\nu}$ can be obtained from the fluid action by $\sqrt{-g}T_{\mu\nu} = -2\delta S_{\rm fluid}/\delta g^{\mu\nu}$:

$$T_{\mu\nu} = 2\mathcal{P}_{,X}\partial_\mu\phi\partial_\nu\phi + g_{\mu\nu}\mathcal{P} \tag{B.3}$$

which is the energy-momentum of a perfect fluid, with the 4-velocity U^μ, energy density ρ and pressure P given by:

$$U^\mu = \frac{-\partial^\mu\phi}{\sqrt{X}} \quad , \quad \rho = 2X\mathcal{P}_{,X} - \mathcal{P} \quad , \quad P = \mathcal{P} . \tag{B.4}$$

A fluid with an equation of state $P = w\rho$ can be modeled by a $\mathcal{P}(X)$ of the form:

$$\mathcal{P}(X) \propto X^{\frac{1+w}{2w}} . \tag{B.5}$$

We are interested in the case of a small w. Let us split ϕ into a background $\bar{\phi}(\eta)$ and a perturbation:

$$\phi = \bar{\phi} + \delta\phi \quad , \quad \pi \equiv -\delta\phi/\bar{\phi}' , \tag{B.6}$$

where we have defined π in terms of the field fluctuation $\delta\phi$. This definition is consistent with the interpretation of π as the velocity potential, as can be seen by working out U^μ in terms of ϕ and equating $U^\mu = a^{-1}(1,\vec{v})$; i.e. $v_i = \nabla_i\pi$ to the lowest order in the perturbation. The background $\bar{\phi}$ obeys:

$$\partial_\eta(a^4\bar{\rho}/\bar{\phi}') = 0\,, \tag{B.7}$$

which implies $\bar{\phi}' \propto a^{1-3w}$, using the fact that $\bar{\rho} \propto a^{-3(1+w)}$. We denote by $\bar{X}$ the value of X evaluated at $\phi = \bar{\phi}$. Using the fact that $P = w\rho$, we find that $1+\delta = (1+[\delta X/\bar{X}])^{(1+w)/2w}$, which implies

$$\frac{2w}{1+w}\ln(1+\delta) = \ln\left(1+\frac{\delta X}{\bar{X}}\right) \tag{B.8}$$

Assuming both w and $\delta X/\bar{X}$ are small, but without assuming δ is small, we can approximate this by

$$2w\ln(1+\delta) \sim \frac{\delta X}{\bar{X}}\,. \tag{B.9}$$

Let us write out $\delta X/\bar{X}$ explicitly in terms of the metric and ϕ fluctuations:

$$\frac{\delta X}{\bar{X}} = -2\left[\Phi + \frac{1}{a}(a\pi)' + \frac{1}{2}(\nabla\pi)^2\right]\,, \tag{B.10}$$

Here we have approximated $\bar{\phi}' \propto a$ (for small w), assumed $\Phi \sim v^2 \lesssim \pi' \sim \mathcal{H}\pi \ll 1$, and ignored higher order terms. (we regard Φ^2, Φ' and $w\Phi$ as both higher order). Eqs. (B.9) and (B.10) combine to give:

$$-w\ln(1+\delta) = \Phi + \frac{1}{a}(a\pi)' + \frac{1}{2}(\nabla\pi)^2\,, \tag{B.11}$$

Applying the spatial gradient on this equation reproduces the Euler equation in the presence of pressure (Eq. 3.46), upon identifying w with c_s^2, the sound speed squared.

We are interested in rewriting the action in Eq. (B.1) in terms of the fluctuations. In other words, we are not so much interested in the background as in the dynamics of the fluctuations. Thus, we ignore the background term in $\sqrt{-g}[M_P^2 R/2 + \mathcal{P}(X)]$. We also remove (tadpole) terms that are linear in fluctuations – they only serve to multiply the background equation of motion. Thus, we have (in the sub-Hubble, non-relativistic limit):

$$
\begin{aligned}
S &= S_{\rm EH} + S_{\rm fluid} \\
S_{\rm EH} &= \int d^4x\, a^2 M_P^2 (\nabla_i\Psi\nabla_i\Psi - 2\nabla_i\Psi\nabla_i\Phi) \\
S_{\rm fluid} &= \int d^4x\, wa^4\bar{\rho}\left[\frac{1+w}{2w}\frac{\delta X}{\bar{X}} + \left(\delta - \frac{1+w}{2w}\frac{\delta X}{\bar{X}}\right)\right] ,
\end{aligned}
\tag{B.12}
$$

where $S_{\rm EH}$ comes from the Einstein-Hilbert action, and $S_{\rm fluid}$ comes from the fluid part of the action. For the latter, we have used the fact that $\mathcal{P}(X) = w\rho = w\bar{\rho}(1+\delta)$, and removed the background piece $w\bar{\rho}$. We add and subtract $(1+w)\delta X/(2w\bar{X})$ to facilitate the removal of tadpole terms that arise from expanding out $(1+\delta) = (1+\delta X/\bar{X})^{1+w/2w}$. We are to understand the last line as follows: the first $(1+w)\delta X/(2w\bar{X})$ term should be understood to have the linear fluctuations removed, while the second $(1+w)\delta X/(2w\bar{X})$ term has all terms in it.[2] The fluid part of S is therefore

$$
S_{\rm fluid} = \int d^4x\, a^4\bar{\rho}\left[-\frac{1}{2}(\nabla\pi)^2 + F\right] , \tag{B.13}
$$

where F is

$$
F \equiv w\left(\delta - \frac{1}{2w}\frac{\delta X}{\bar{X}}\right) , \tag{B.14}
$$

[2]The determinant $\sqrt{-g}$ contains terms of order Φ and Φ^2. Terms of order Φ multiplying the background are removed as tadpoles. Surviving terms can be seen to multiply at least one factor of w or of $v^2 \sim \Phi$ (the latter with no compensating $1/w$), and so are small compared to what we keep (which are of order v^2 or $v^2(v^2/w)$).

with δ and δX understood to be expressible in terms of Φ and π using Eqs. (B.9) and (B.10). We have already verified that the Euler equation with pressure holds (Eq. (B.11)). From the point of view of the action $S_{\rm fluid}$, this merely serves as a definition for δ. Let us verify that we obtain the Poisson and continuity equations by varying the action. First, we see that Ψ can be integrated out by setting $\Psi = \Phi$. In other words, let us work with the action:

$$\boxed{S = \int d^4x \left\{ -a^2 M_P^2 (\nabla\Phi)^2 + a^4 \bar{\rho} \left[-\frac{1}{2}(\nabla\pi)^2 + F \right] \right\} .} \tag{B.15}$$

The variation of F when we vary Φ, using Eqs. (B.9) and (B.10), is

$$\Delta F = \frac{1}{2}\frac{\Delta\delta X}{\bar{X}}\delta = -\delta\Delta\Phi \,, \tag{B.16}$$

giving us

$$\Delta S = \int d^4x \, [2a^2 M_P^2 \nabla^2\Phi - a^4\bar{\rho}\delta]\Delta\Phi \,, \tag{B.17}$$

and therefore the Poisson equation. This assumes that the only fluctuations sourcing Φ are from the $\mathcal{P}(X)$ fluid. This can of course be relaxed. The π equation of motion, on the other hand, follows from

$$\Delta F = \frac{1}{2}\frac{\Delta\delta X}{\bar{X}}\delta = -\left[\frac{1}{a}(a\Delta\pi)' + \nabla_i\pi\nabla\Delta\pi\right]\delta \,, \tag{B.18}$$

which together with the variation $\Delta\left(-\frac{1}{2}(\nabla\pi)^2\right)$ gives us the continuity equation $\delta' + \nabla_i[(1+\delta)\nabla_i\pi] = 0$.

The action in Eq. (B.13) is a bit hard to use, because F involves a fairly nonlinear function of the fields. There are two possible simplifications.

We are interested in the $w \to 0$ limit. Keeping δ finite, Eq. (B.11) tells us

$$\Phi = -\frac{1}{a}\left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2\right] \tag{B.19}$$

which is just the pressureless Euler equation again. Sending $F \to 0$, and substituting the above into Eq. (B.15), we obtain:

$$S = -\int d^4x \left[\frac{1}{2}\bar{\rho}a^4(\nabla\pi)^2 + M_P^2\left(\nabla\left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2\right]\right)^2\right] \tag{B.20}$$

reproducing Eq. (3.58) that we wrote down in Sec. 3.1.3.2. This justifies the normalization and sign that was adopted there.

The other possible simplification is to expand out $(1+\delta) = (1+\delta X/\bar{X})^{1+w/2w}$ to second order in $\delta X/\bar{X}$. We have resisted doing so earlier, because doing so effectively assumes $\delta X/\bar{X}$ is parametrically smaller than w (which is itself small). This is equivalent to assuming small δ, something we might not want to impose. It is nonetheless instructive to see what results:

$$\boxed{S = \int d^4x \left\{-a^2M_P^2(\nabla\Phi)^2 + \frac{1}{2}a^4\bar{\rho}\left[-(\nabla\pi)^2 + c_s^{-2}\left(\frac{1}{a}(a\pi)' + \Phi + \frac{1}{2}(\nabla\pi)^2\right)^2\right]\right\}} \tag{B.21}$$

where we have set $w = c_s^2$. In the context of this action, we treat δ as *defined* by:

$$-c_s^2\delta = \Phi + \frac{1}{a}(a\pi)' + \frac{1}{2}(\nabla\pi)^2\,. \tag{B.22}$$

Notice how this differs from Eq. (B.11) in replacing $\ln(1+\delta)$ by δ on the left hand side. The reason for this definition is so that the Φ equation of motion gives the Poisson equation as usual. The π equation of motion can be seen to give the continuity equation.

In other words, the full set of equations in this system are:

$$\delta' + \nabla_i[(1+\delta)v_i] = 0 \quad , \quad v_i' + v_j\nabla_j v_i + \mathcal{H}v_i = -\nabla_i\Phi - c_s^2\nabla_i\delta \quad , \quad \nabla^2\Phi = 4\pi G\bar{\rho}a^2\delta \tag{B.23}$$

This is the set of equations one expects for a fluid with pressure, except the pressure term in the Euler equation is slightly modified from the non-perturbative one displayed in Eq. (3.46). Aside from this modification, this system of equations has the correct nonlinear structure. In particular, on length scales above the Jeans scale i.e. $k < k_J$ where $k_J^2 \equiv a^2\bar{\rho}/(2M_P^2c_s^2)$, one can ignore the pressure term, and the system reduces exactly to the standard pressureless LSS equations (Eq. 3.12). A useful feature of the action in Eq. (B.21) is that it shows clearly that π has a kinetic term of the correct sign.

To summarize, the fluid action Eq. (B.15) gives the exact nonlinear equations for the perturbations of a fluid with pressure in the Newtonian limit. It simplifies in the zero-pressure limit to the action in Eq. (3.58), which gives the exact nonlinear equations for a pressureless fluid. It can be approximated by the action in Eq. (B.21) which gives a linearized pressure term for the Euler equation, but otherwise retains the full nonlinear structure of the exact theory.

B.2 Derivation of the General Relativistic Adiabatic Mode Conditions in Newtonian Gauge

In this Appendix, we derive the adiabatic mode conditions appropriate for the Newtonian gauge, and derive the additional diffeomorphism laid out in Eq. (3.68).

For the purpose of deriving consistency relations, it is important that the modes generated nonlinearly by the symmetries be the low momentum limit of actual physical modes, i.e. they must obey adiabatic mode conditions (see sec. 3.1.3.1). For the low momentum modes (and for them only), it is sufficient to consider the linearized Einstein equations, and study the time-dependence they imply for the perturbations. The linearized Einstein equations in Newtonian gauge are:

$$-\frac{1}{2}\delta G^0{}_0 = -4\pi G\delta T^0{}_0 \to \nabla^2\Psi - 3\mathcal{H}(\Psi' + \mathcal{H}\Phi) = 4\pi G a^2 \sum(\rho - \bar{\rho}) \,, \tag{B.24}$$

$$\frac{1}{2}\delta G^0{}_i = 4\pi G\delta T^0{}_i \to -\partial_i\left(\Psi' + \mathcal{H}\Phi\right) + \frac{1}{4}\nabla^2 S_i = 4\pi G a^2 \sum(\bar{\rho} + \bar{P})(v_i + S_i) \tag{B.25}$$

$$\frac{1}{6}\delta G^k{}_k = \frac{4\pi}{3}G\delta T^k{}_k \to \Psi'' + \mathcal{H}(\Phi' + 2\Psi') + (2\mathcal{H}' + \mathcal{H}^2)\Phi - \frac{1}{3}\nabla^2(\Psi - \Phi) = 4\pi G a^2 \sum(P - \bar{P}) \tag{B.26}$$

$$\begin{aligned} &\delta G^i{}_j - \frac{1}{3}\delta^i{}_j\delta G^k{}_k = 8\pi G\left(\delta T^i{}_j - \frac{1}{3}\delta^i{}_j\delta T^k{}_k\right) \to \\ &\left(\partial_i\partial_j - \frac{\delta_{ij}}{3}\nabla^2\right)(\Psi - \Phi) - (\partial_0 + 2\mathcal{H})\partial_{(i}S_{j)} + (\partial_0^2 + 2\mathcal{H}\partial_0 - \nabla^2)\frac{\gamma_{ij}}{2} \\ &= 8\pi G\left(\delta T^i{}_j - \frac{1}{3}\delta_{ij}\delta T^k{}_k\right) . \end{aligned} \tag{B.27}$$

We have allowed the possibility that there might be multiple fluid components present (for instance dark matter, baryons, radiation, etc.), hence the summation on the right hand side, though we suppress the label for each component. Also useful are the linearized conservation equations, assuming each fluid is individually conserved. The continuity equation for each fluid component is

$$\delta_n' + \partial_i v_i = 3\Psi' \,, \tag{B.28}$$

where δ_n is related to the density fluctuation $\delta \equiv (\rho-\bar{\rho})/\bar{\rho}$ by $(1+w)\delta_n = \delta$, with $w = P/\rho$ being the equation of state parameter of the fluid component of interest. This definition of δ_n is motivated by the fact that $\bar{\rho} \propto a^{-3(1+w)}$, and so it is $\bar{\rho}^{[1/(1+w)]}$ that redshifts like a^{-3}, i.e. one can think of $n \equiv \rho^{[1/(1+w)]}$ as the "number" density, and of δ_n as its fractional (small) fluctuation (for instance, for $w = 1/3$, n would be the number density of photons). In deriving Eq. (B.28), it is useful to know that $\mathcal{H}^2 - \mathcal{H}' = 4\pi G a^2 \sum(\bar{\rho} + \bar{P})$. Note also that, in an analogous manner to Eq. (3.74):

$$\Delta_{\rm nl.}\delta_n = 3\mathcal{H}\xi^0 \,. \tag{B.29}$$

The relativistic Euler equation for each component is:

$$(v_i + S_i)' + (1-3w)\mathcal{H}(v_i + S_i) = -\partial_i\Phi - w\partial_i\delta_n \,, \tag{B.30}$$

where we have assumed the fourth Einstein equation has a vanishing source.[3] Decomposing this last equation into scalar, vector and tensor parts, we have

$$\left(\partial_i\partial_j - \frac{\delta_{ij}}{3}\nabla^2\right)(\Psi - \Phi) = 0 \,, \tag{B.31}$$

$$(\partial_0 + 2\mathcal{H})\partial_{(i}S_{j)} = 0 \,, \tag{B.32}$$

$$(\partial_0^2 + 2\mathcal{H}\partial_0 - \nabla^2)\gamma_{ij} = 0 \,. \tag{B.33}$$

Following Weinberg [84], we demand that the (nonlinear part of the) symmetry-generated perturbations (as described in Sec. 3.1.4.1), solve the Einstein equations in a non-trivial way, that is, in a way that works even if we deform those perturbations slightly away from the zero momentum $q = 0$ limit. (See Eq. 3.19 for the Newtonian

[3]This does not strictly hold if for instance the fluid is made out of a collection of relativistic particles, but it is a reasonably good approximation in LSS. In essence, we assume $\Phi = \Psi$, and $(a^2 S_i)' = 0$.

analog of this statement.) For scalar fluctuations, we therefore insist:

$$\text{scalar adiabatic mode condition } : \Psi = \Phi \quad , \quad -(\Psi' + \mathcal{H}\Phi) = 4\pi G a^2 \sum (\bar{\rho} + \bar{P})\pi \tag{B.34}$$

which comes from removing the spatial gradients from Eq. (B.31) and the scalar part of Eq. (B.25). Similarly, (if vector modes are present) the vector adiabatic mode conditions are:

$$\text{vector adiabatic mode condition } : (\partial_0 + 2\mathcal{H})S_i = 0 \quad , \quad (v_i^\perp + S_i)' + (1-3w)\mathcal{H}(v_i^\perp + S_i) = 0 \tag{B.35}$$

which comes from Eq. (B.32) and the transverse component of Eq. (B.30) ($v_i^\perp$ is the transverse part of the velocity v_i, i.e. the vorticity component). Note how the vector modes have only a single solution, which decays. Since single field inflation cannot generate vector modes, we will not consider them further here. The tensor adiabatic mode condition is simply the tensor equation of motion (B.33):

$$\text{tensor adiabatic mode condition } : \gamma_{ij}'' + 2\mathcal{H}\gamma_{ij}' - \nabla^2\gamma_{ij} = 0\,. \tag{B.36}$$

Note that we do not wish to simply set the gradient to zero, because we are interested in diffeomorphisms generating a γ_{ij} that is the soft limit of a finite momentum physical mode.

Applying the above conditions to the (nonlinear part of the) symmetry-generated perturbations (Eqs. 3.67 and 3.72), we obtain:

$$\xi^{0\prime} + 2\mathcal{H}\xi^0 + \frac{1}{3}\partial_i\xi^i = 0 \quad , \quad \partial_i\xi^{i\prime} = 0 \quad , \quad \gamma_{ij} = -\left(\partial_i\xi^j + \partial_j\xi^i - \frac{2}{3}\delta_{ij}\partial_k\xi^k\right)\,. \tag{B.37}$$

The first equality enforces $\Phi = \Psi$. The second equality enforces the second part of the scalar adiabatic mode condition, with the understanding that in the soft limit, all fluid components share the same velocity perturbation π. The third equality equates the tensor mode with the traceless part of the spatial metric generated by the diffeomorphism. This holds only if a certain gauge condition is satisfied such that the scalar contribution to the spatial metric resides entirely in its trace (see below). As far as the adiabatic mode condition is concerned, the important point is that γ_{ij} defined this way satisfies the tensor equation of motion (B.36). These three expressions constitute the adiabatic mode conditions on residual diffeomorphisms in Newtonian gauge. For a diffeomorphism to respect the Newtonian gauge, it must satisfy

$$\partial_i \xi^0 = \partial_0 \xi^i \quad , \quad \nabla^2 \xi^i + \frac{1}{3}\partial_i(\partial_k \xi^k) = 0 \tag{B.38}$$

such that $\Delta_{\rm nl} g_{0i} = 0$ and the traceless part of $\Delta_{\rm nl} g_{ij}$ is transverse (see Eq. 3.67).

As discussed in Sec. 3.1.4.1, one way to organize the set of Newtonian-gauge diffeomorphisms that satisfy Eqs. (B.37) and (B.38) is to relate each such diffeomorphism to a corresponding known residual diffeomorphism in the unitary gauge $\xi_{\rm unit.}$ (Eq. 3.70):

$$\xi^0 = \xi^0_{\rm add.} \quad , \quad \xi^i = \xi^i_{\rm unit.} + \xi^i_{\rm add.} \, , \tag{B.39}$$

where the time-independent $\xi^i_{\rm unit.}$ is supplemented by a time- and space- diffeomorphism $\xi^0_{\rm add.}$, $\xi^i_{\rm add.}$. The time-independent unitary-gauge diffeomorphism $\xi^i_{\rm unit.}$ satisfies Eq. (3.64). Comparing this with Eq. (B.38), we see that $\xi^i_{\rm add.}$ itself must satisfy the same: $\nabla^2 \xi^i_{\rm add.} + \partial_i(\partial_k \xi^k_{\rm add.})/3 = 0$. For this reason, we might as well absorb any time-independent part of $\xi^i_{\rm add.}$ into the definition of $\xi^i_{\rm unit.}$. From the second condition in Eq. (B.37), we see that $\partial_i \xi^i_{\rm add}$ must be independent of time. Suppose it is equal to some function $f(x)$. One can express $\xi^i_{\rm add.}$ as a gradient and a curl (plus possibly some function that depends only on time). The divergence of the gradient is what matches up with $f(x)$, i.e. the gradient part is time-independent, and so by definition, it should have been absorbed into $\xi_{\text{unit.}}$ already. Thus, we can set $f(x) = 0$ and we can assume $\partial_i \xi^i_{\text{add.}} = 0$ without loss of generality. The first condition of Eq. (B.37) thus tells us

$$\xi^0_{\text{add.}}{}' + 2\mathcal{H}\xi^0_{\text{add.}} + \frac{1}{3}\partial_i \xi^i_{\text{unit.}} = 0 \tag{B.40}$$

Recall from Eq. (3.72) that $\Delta_{\text{nl.}}\pi = \xi^0 = \xi^0_{\text{add.}}$. From Appendix B.3, we see that π in the soft limit has the time dependence of D', where $D(\eta)$ is the linear growth factor satisfying:

$$D'' + 2\mathcal{H}D' - c = 0 \tag{B.41}$$

where c is a constant. Comparing this against Eq. (B.40) and keeping only the growing solution, we see that

$$\xi^0_{\text{add.}} = -\frac{1}{3c} D' \partial_i \xi^i_{\text{unit.}} \, , \tag{B.42}$$

confirming the time-diffeomorphism of Eq. (3.68). We can then solve for $\xi^i_{\text{add.}}$ from the first expression of Eq. (B.38) which tells us that $\partial_i \xi^0_{\text{add.}} = \partial_0 \xi^i_{\text{add.}}$, i.e.

$$\xi^i_{\text{add.}} = -\frac{1}{3c} D \partial_i (\partial_k \xi^k_{\text{unit.}}) = \frac{1}{c} D \nabla^2 \xi^i_{\text{unit.}} \, , \tag{B.43}$$

where the second equality follows from Eq. (3.64). This confirms the space-diffeomorphism of Eq. (3.68). As a self-consistency check, one can see that Eq. (3.64) also implies that $\partial_i \xi^i_{\text{add.}} = 0$. Lastly, it can also be checked that the tensor mode created by this diffeomorphism (the third expression of Eq. B.37) obeys the tensor equation of motion. To see this, it is useful to note that because $\xi^i_{\text{unit.}}$ satisfies $\nabla^2 \xi^i_{\text{unit.}} + \partial_i (\partial_k \xi^k_{\text{unit.}})/3 = 0$, we also know that $\nabla^2\partial_k\xi^k_{\rm unit.} = 0$, $\nabla^2\partial_i\xi^j_{\rm unit.} = -\partial_i\partial_j\partial_k\xi^k_{\rm unit.}/3$, and $\nabla^2\nabla^2\partial_i\xi^j_{\rm unit.} = 0$.[4] It is worth noting that for pure tensor symmetries, where $\partial_i\xi^i_{\rm unit} = 0$, both $\xi^0_{\rm add.}$ and $\xi^i_{\rm add.}$ vanish, and so the pure tensor symmetries coincide in the Newtonian gauge and unitary gauge, as they should.

B.3 Derivation of the General Relativistic Velocity Equation

Our goal in this Appendix is to derive the following equation for the velocity potential π:

$$(\pi' + 2\mathcal{H}\pi - C)' - 3w\mathcal{H}(\pi' + 2\mathcal{H}\pi - C) = w(g' + \mathcal{H}g) - (1+3w)\left[(\mathcal{H}^2 - \mathcal{H}')\pi - 4\pi G a^2 \sum(\bar\rho + \bar P)\pi\right] \quad \text{(B.44)}$$

where $g \equiv \int d\eta \nabla^2\pi$, and C, determined by initial conditions, is a constant in time but not space . Here, π refers to the velocity potential of some particular fluid component of interest with an equation of state parameter w – except in the very last term where $\sum(\bar\rho + \bar P)\pi$ refers to a sum over all fluid components. We will use this equation to deduce useful statements about the time-dependence of π in the soft limit.

The continuity equation (B.28) can be integrated once to obtain:

$$\delta_n = 3(\Psi + C) - g \quad , \quad g \equiv \int d\eta \nabla^2\pi \,, \quad \text{(B.45)}$$

where C denotes an integration constant that is independent of time, but in general dependent on space. This can be substituted into (the scalar part of) the relativistic

[4] In other words, the combined action of $\xi^i_{\rm unit.} + \xi^i_{\rm add.}$ generates a tensor mode of the form $\gamma_{ij} = [1 + (D/c)\nabla^2]\gamma_{ij\,\rm unit.}$ where $\gamma_{ij\,\rm unit.}$ is the tensor mode generated by the time-independent unitary diffeomorphism alone. That the constant tensor (growing) mode gets corrected at finite momentum by a term proportional to momentum squared should not be surprising. The time dependence can also be checked explicitly by solving the tensor equation of motion in the small but finite momentum limit.

Euler equation (B.30) and integrated once to give

$$\pi' + \mathcal{H}(1-3w)\pi = -(1+3w)\Psi - 3wC + wg\,. \tag{B.46}$$

On the other hand, (the scalar part of) the $\delta G^0{}_i$ equation (B.25) can be integrated once to obtain

$$-(\Psi' + \mathcal{H}\Psi) = 4\pi G a^2 \sum(\bar{\rho} + \bar{P})\pi\,, \tag{B.47}$$

where we have assumed $\Psi = \Phi$. One can solve for Ψ from Eq. (B.46), substitute the result into Eq. (B.47), and subtract from both sides $(\mathcal{H}^2 - \mathcal{H}')\pi$. This gives Eq. (B.44). Note that in this derivation, we have not thrown away any gradient terms; i.e., we have not made any super-Hubble approximation.

Equation (B.44) simplifies if π happens to be the same for all fluid components, in which case what appears within the square brackets sums to zero, by virtue of $\mathcal{H}^2 - \mathcal{H}' = 4\pi G a^2 \sum(\bar{\rho} + \bar{P})$. This happens, for instance, if we work on super-Hubble scales and assume adiabatic initial conditions. One can check that this is a self-consistent solution on super-Hubble scales, and assuming all fluid components move with the same π, the entire right hand side of Eq. (B.44) vanishes, implying:

$$\pi' + 2\mathcal{H}\pi - C \propto a^{3w}\,. \tag{B.48}$$

This suggests different fluid components (with different w's) evolve differently, *unless* the proportionality constant is in fact zero, i.e.

$$\pi' + 2\mathcal{H}\pi - C = 0\,. \tag{B.49}$$

With this choice of the initial condition, it is thus consistent to have the same π for all fluid components on super-Hubble scales. *Interestingly, for pressureless matter ($w = 0$), Eq. (B.49) holds even on sub-Hubble (but linear) scales, after radiation domination.* This can be seen by setting $w = 0$ in Eq. (B.44), and noting that during matter or cosmological constant domination, the terms within the square brackets still sum to zero. This means that for a mode (of pressureless matter) that enters the Hubble radius after radiation domination, Eq. (B.49) holds for its entire history. For a mode that enters the Hubble radius before matter domination, however, Eq. (B.49) does not hold in the intermediate period when the mode is within the Hubble radius during the radiation dominated phase.[5]

As we see in sec. 3.1.4.2, the fact that Eq. (B.49) holds for pressureless matter both inside and outside the Hubble radius (as long as the mode of interest crosses the Hubble radius after radiation domination) enables us to have interesting consistency relations in the Newtonian limit. It is also worth noting that since $\partial_i \pi$ describes the dark matter velocity on all (linear) scales, including sub-Hubble ones, where we know the velocity scales with time as D' (D being the linear growth factor), we expect

$$D'' + 2\mathcal{H}D' - c = 0\,, \tag{B.50}$$

where c is some constant whose normalization is arbitrary – its normalization is tied to the normalization of the growth factor D. That this relation holds for the Newtonian growth factor in a matter dominated universe is easy to check: $D \propto a$. That this is true for more general cases is less familiar. Let us check this for a universe with a cosmological constant.

For a flat universe with pressureless matter and a cosmological constant, the linear

[5] It is worth pointing out that Eq. (B.49), when substituted into Eq. (B.46) gives $\Psi = -(\pi' + \mathcal{H}\pi)$ – this holds as long as the wg term can be ignored, which can be justified either for super-Hubble scales, or for $w = 0$.

growth factor can be written in closed form [99]:

$$D = \frac{5\Omega_m^0}{2}\frac{H(a)}{H_0}\int_0^a \frac{d\tilde{a}}{(\tilde{a}H(\tilde{a})/H_0)^3}\,, \tag{B.51}$$

where Ω_m^0 is the matter density today, H_0 is the Hubble constant today, and the normalization is chosen such that D equals the scale factor a in the early universe. Note that $H = a'/a^2$ whereas $\mathcal{H} = a'/a$.

Let us rewrite what we want to show, Eq. (B.50), as follows:

$$H\frac{d}{da}\left(a^4 H\frac{dD}{da}\right) = c\,. \tag{B.52}$$

First, note that

$$\frac{d}{da}\left(\frac{H}{H_0}\right) = -\frac{3\Omega_m^0}{2a^4}\frac{H_0}{H}\,. \tag{B.53}$$

We can therefore work out:

$$\frac{dD}{da} = -\frac{3\Omega_m^0}{2a^4}\frac{H_0}{H}\frac{5\Omega_m^0}{2}\int_0^a \frac{d\tilde{a}}{(\tilde{a}H(\tilde{a})/H_0)^3} + \frac{5\Omega_m^0}{2}\frac{H_0^2}{a^3H^2}\,. \tag{B.54}$$

Therefore,

$$\frac{d}{da}\left(a^4 H\frac{dD}{da}\right) = \frac{5\Omega_m^0}{2}\frac{H_0^2}{H}\,, \tag{B.55}$$

implying the desired result Eq. (B.52).

B.4 Decaying Modes and Vector Symmetries

At several places in our discussion we have omitted the second solution to the scalar and tensor equations of motion, which decays at late times. It is worth discussing the decaying solution in a little more detail, both for completeness and because the reasons for ignoring it are slightly subtle.

The decaying solution arises at the linearized level because both the scalar and tensor equations of motion for the linearized modes (Eq. (3.16) or Eq. (B.50), and Eq. (B.36)) are second order. In particular, we note that Eq. (B.50) for the linear growth factor $D(\eta)$ can allow for a decaying piece $D' \propto 1/a^2$ which is independent of c as well as the strictly growing piece (which does depend on having $c \neq 0$, and which is what is usually meant by the linear growth factor). The decaying piece corresponds to varying the lower limit of the integral in Eq. (B.51), which is arbitrary. It is straightforward to check that this solution gives the correct decaying solution $\delta \propto \mathcal{H}/a$ for Eq. (3.16) in the Newtonian limit, using the helpful lemma

$$\mathcal{H}' - \mathcal{H}^2 = -4\pi G_N a^2 \Sigma(\bar{\rho} + \bar{P}) = -4\pi G_N \bar{\rho}_{\text{matter}} \propto \frac{1}{a^2} \tag{B.56}$$

in a ΛCDM universe. From Eq. (B.40), the most general diffeomorphism $\xi^0_{\text{add.}} and \xi^i_{\text{unit.}} + \xi^i_{\text{add.}}$ involving scalar and tensor modes allowed by the adiabatic mode conditions includes the decaying piece

$$\begin{aligned} \xi^0_{\text{add.,decay}} &= \frac{d(x)}{a^2} \\ \xi^i_{\text{add.,decay}} &= \int^\eta \frac{\partial^i d(x)}{a^2} \end{aligned} \tag{B.57}$$

where $d(x)$ is harmonic because of Eqs. (B.37) and (B.38). It is straightforward to check that this satisfies the tensor equation B.36 in the super-Hubble limit as well.

Note that the decaying mode is independent of the time-independent (growing mode) unitary transformation, and so the diffeomorphism corresponding to a decaying mode is a separate symmetry. We can Taylor expand as in Sec. 3.1.4.4 to write the most general decaying symmetries as

$$\xi^0_{\rm decay} = \frac{1}{n!}\frac{1}{a^2}M_{\ell\ell_1\cdots\ell_n}x^{\ell_1}\cdots x^{\ell_n},\ \xi^i_{\rm decay} = \frac{1}{(n-1)!}\int^\eta \frac{1}{a^2}M^i_{\ell\ell_2\cdots\ell_n}x^{\ell_2}\cdots x^{\ell_n}. \tag{B.58}$$

for $(n \geq 1)$, where the M's are constant and obey the usual transversality and adiabatic transversality conditions.

Can we derive consistency relations for the decaying modes using these symmetries, using Eq. (3.27) or its generalization Eq. (3.112)? We argue that the answer is no, though it is not enough to say that these simply decay away. Rather, keeping the decaying modes would correspond to a nonstandard choice of the initial vacuum state in the far past: if the decaying mode is not set to zero the energy associated with these modes (the scalar part of the action scales like $\rho a^4 \sim a^2(\nabla\Phi)^2 \sim \mathcal{H}^2/a^2$) becomes divergent at early times. Had we chosen to ignore this problem and work within the putative vacuum containing only decaying modes, we could have, in which case the lack of the time-dependent piece would make our consistency relations look slightly different from Eq. (3.118). Note that it is not true that the consistency relations should vanish in this case because the modes decay at late times: this is because the consistency relation is a ratio between the (N+1)-point function and the power spectrum, both of which decay, but the ratio on the right hand side does not.

For the sake of completeness, we discuss the case where the symmetries may involve vector modes – unlike the scalars and tensors, these have only a decaying solution. The condition

$$\nabla^2\xi^i + \frac{1}{3}\partial^i(\partial_j\xi^j) = 0. \tag{B.59}$$

on the spatial part of the diffeomorphism will still be obeyed, but the condition $\partial_0\xi^i = \partial_i\xi^0$ will be violated and replaced by the weaker condition

$$\nabla^2\xi^0 = \partial^0\partial_i\xi^i = 0. \tag{B.60}$$

where in the second equality we have made use of the second equation in Eq. (B.37). Using the vector adiabatic mode conditions (Eq. (B.35)) we have

$$\begin{aligned}(\partial_i\xi^0 - \partial_0\xi^i)_{\text{vec.}} &= \frac{\bar{\xi}^i}{a^2}\\ (\partial_i\xi^0)_{\text{vec.}} &\propto \frac{1}{a}\end{aligned} \tag{B.61}$$

where $\bar{\xi}^i$ is transverse and time-independent. The second of these conditions is clearly incompatible with the first condition in Eq. B.37:

$$\xi^{0\prime} + 2\mathcal{H}\xi^0 + \frac{1}{3}\partial_i\xi^i = 0, \tag{B.62}$$

unless $(\partial_i\xi^0)_{\text{vec.}}$ vanishes, and so the vector part of the symmetry will be $\bar{\xi}^i \int \frac{d\eta}{a^2} \subset \xi^i$. We can Taylor expand

$$\bar{\xi}^i = \sum \frac{1}{(n+1)!}\bar{M}^i{}_{\ell_0\ell_1\cdots\ell_n}x^{\ell_0}\cdots x^{\ell_n} \tag{B.63}$$

The $\bar{M}$'s are completely traceless, and they obey the usual tensor transversality conditions Eqs. (3.107), (3.108) as well, so these are vector-tensor symmetries. Note that they obey the tensor equation of motion Eq. (B.36) on superhorizon scales, though they correspond to the decaying mode solution. For $n \geq 1$ there will be four such symmetries at each level. For $n = 0$, there are additional symmetries with $\bar{M}_{i\ell_0}$ antisymmetric in the indices; these correspond to time-dependent rotations.[6] They will obey the further adiabatic transversality condition

$$\hat{q}^i(\bar{M}_{i\ell_0}(\hat{q}) - \bar{M}_{\ell_0 i}) = 0 \tag{B.64}$$

which will reduce the number of allowed polarizations from 3 to 2. Since a localized rotation necessarily involves shearing, we need this condition to enforce transversality in addition to the antisymmetric tensor structure.

To summarize, at $n = 0$ there are two purely vector symmetries, and for $n \geq 1$ there are four vector plus tensor symmetries. Since vector modes always decay, for our choice of vacuum there are no consistency relations that involve vector modes.

B.5 Recovering the Eulerian space consistency relation from Lagrangian space – to arbitrary orders in displacement

In Sec. 3.2.3.2, we argue that the Eulerian space consistency relation follows from its Lagrangian space counterpart, at least to the two lowest non-trivial orders in a formal expansion in displacement. In this Appendix, we show that this works to arbitrary orders.

[6] Note that for $n \geq 1$, the tensor structure that is antisymmetric in the first two indices and symmetric in the last $n + 1$ will vanish identically.

We begin by expanding Eq. (3.148) to all orders in Δ:

$$\begin{aligned} \delta_{\mathbf{k}_1=\mathbf{p}_1} &= \int d^3\mathbf{q}\, J(\mathbf{q})\,\delta(\mathbf{x}(\mathbf{q}))\, e^{i\mathbf{p}_1\cdot(\mathbf{q}+\boldsymbol{\Delta}(\mathbf{q}))} \\ &= \sum_{n=0}^{\infty}\frac{i^n}{n!}\mathrm{p}_1^{i_1}\cdots\mathrm{p}_1^{i_n}\int_{\mathbf{p}_{j_1},\cdots,\mathbf{p}_{j_n}}\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_{j_1}-\cdots-\mathbf{p}_{j_n}}\Delta^{i_1}_{\mathbf{p}_{j_1}}\cdots\Delta^{i_n}_{\mathbf{p}_{j_n}} \end{aligned} \tag{B.65}$$

where, as before, $\tilde{\delta}(\mathbf{q}) = J(\mathbf{q})\delta(\mathbf{x}(\mathbf{q}))$. Collecting together the terms of order Δ^n in the three-point correlator $\langle v^j_{\mathbf{p}}(\eta)\delta_{\mathbf{p}_1}(\eta_1)\delta_{\mathbf{p}_2}(\eta_2)\rangle$ in E_L, we have

$$\begin{aligned} \sum_{m=0}^{n}\frac{i^m}{m!}\frac{i^{n-m}}{(n-m)!}\mathrm{p}_1^{i_1}\cdots\mathrm{p}_1^{i_m}\mathrm{p}_2^{i_{m+1}}\cdots\mathrm{p}_2^{i_n}\int_{\mathbf{p}_{j_1},\cdots,\mathbf{p}_{j_n}}\langle v^j_{\mathbf{p}}(\eta)\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_{j_1}-\cdots-\mathbf{p}_{j_m}}(\eta_1)\Delta^{i_1}_{\mathbf{p}_{j_1}}(\eta_1)\cdots\Delta^{i_m}_{\mathbf{p}_{j_m}}(\eta_1) \\ \cdot\,\tilde{\delta}_{\mathbf{p}_2-\mathbf{p}_{j_{m+1}}-\cdots-\mathbf{p}_{j_n}}(\eta_2)\Delta^{i_{m+1}}_{\mathbf{p}_{j_{m+1}}}(\eta_2)\cdots\Delta^{i_n}_{\mathbf{p}_{j_n}}(\eta_2)\rangle \end{aligned} \tag{B.66}$$

where the correlator in the integral is the full correlator containing both connected and disconnected pieces, but where no proper subset of the original momenta $\mathbf{p}, \mathbf{p}_1, \mathbf{p}_2$ sums to zero, since it is the connected correlator that appears in E_L. The correlator can be split into a sum over products of connected blocks. Anticipating division by $P_v(\mathbf{p})$, we see that the Lagrangian space consistency relation implies that all contributions where the soft velocity $v^j(\mathbf{p})$ is part of a connected correlator with two or more other fields will vanish. All the remaining terms contain factors such as

$$\langle v^j(\mathbf{p})\tilde{\delta}_{\mathbf{p}_1-\cdots}\rangle\,,\ \langle v^j(\mathbf{p})\tilde{\delta}_{\mathbf{p}_2-\cdots}\rangle\,,\ \langle v^j(\mathbf{p})\Delta_{\mathbf{p}_j}(\eta_1)\rangle\,,\ \langle v^j(\mathbf{p})\Delta_{\mathbf{p}_j}(\eta_2)\rangle\,. \tag{B.67}$$

The first two types of terms will be suppressed by an additional power of the soft momentum, and are subdominant in the squeezed limit. The final set of terms are

$$\sum_{m=0}^{n}\frac{i^m}{m!}\frac{i^{n-m}}{(n-m)!}\mathrm{p}_1^{i_1}\cdots\mathrm{p}_1^{i_m}\mathrm{p}_2^{i_{m+1}}\cdots\mathrm{p}_2^{i_n}\int_{\mathbf{p}_{j_1},\cdots,\mathbf{p}_{j_n}}\Big[m\langle v^j_{\mathbf{p}}(\eta)\Delta^{i_1}_{\mathbf{p}_{j_1}}(\eta_1)\rangle\langle\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_{j_1}-\cdots-\mathbf{p}_{j_m}}(\eta_1)$$
$$\cdot\,\Delta^{i_1}_{\mathbf{p}_{j_1}}(\eta_1)\cdots\Delta^{i_m}_{\mathbf{p}_{j_m}}(\eta_1)\tilde{\delta}_{\mathbf{p}_2-\mathbf{p}_{j_{m+1}}-\cdots-\mathbf{p}_{j_n}}(\eta_2)\Delta^{i_{m+1}}_{\mathbf{p}_{j_{m+1}}}(\eta_2)\cdots\Delta^{i_n}_{\mathbf{p}_{j_n}}(\eta_2)\rangle\Big]$$
$$+\,(1\leftrightarrow 2) \tag{B.68}$$

and they give (relabelling m as $m+1$)

$$\Big[i\mathrm{p}^j\frac{\mathbf{p}\cdot\mathbf{p}_1}{\mathbf{p}^2}\frac{D(\eta_1)}{D'(\eta)}\sum_{m=0}^{n-1}\frac{i^{n-1}}{m!(n-1-m)!}\mathrm{p}_1^{i_1}\cdots\mathrm{p}_1^{i_m}\mathrm{p}_2^{i_{m+1}}\cdots\mathrm{p}_2^{i_n}\int_{\mathbf{p}_{j_1},\cdots,\mathbf{p}_{j_n}}\langle\tilde{\delta}_{\mathbf{p}_1+\mathbf{p}-\mathbf{p}_{j_1}-\cdots-\mathbf{p}_{j_m}}(\eta_1)$$
$$\cdot\,\Delta^{i_1}_{\mathbf{p}_{j_1}}(\eta_1)\cdots\Delta^{i_m}_{\mathbf{p}_{j_m}}(\eta_1)\tilde{\delta}_{\mathbf{p}_2-\mathbf{p}_{j_{m+1}}-\cdots-\mathbf{p}_{j_n}}(\eta_2)\Delta^{i_{m+1}}_{\mathbf{p}_{j_{m+1}}}(\eta_2)\cdots\Delta^{i_n}_{\mathbf{p}_{j_n}}(\eta_2)\rangle\Big]$$
$$+\,(1\leftrightarrow 2) \tag{B.69}$$

Comparing this to the order Δ^{n-1} terms in the expansion of E_R,

$$\Big[-i\mathrm{p}^j\frac{\mathbf{p}\cdot\mathbf{p}_1}{\mathbf{p}^2}\frac{D(\eta_1)}{D'(\eta)}\sum_{m=0}^{n-1}\frac{i^{n-1}}{m!(n-1-m)!}\mathrm{p}_1^{i_1}\cdots\mathrm{p}_1^{i_m}\mathrm{p}_2^{i_{m+1}}\cdots\mathrm{p}_2^{i_n}\int_{\mathbf{p}_{j_1},\cdots,\mathbf{p}_{j_n}}\langle\tilde{\delta}_{\mathbf{p}_1-\mathbf{p}_{j_1}-\cdots-\mathbf{p}_{j_m}}(\eta_1)$$
$$\cdot\,\Delta^{i_1}_{\mathbf{p}_{j_1}}(\eta_1)\cdots\Delta^{i_m}_{\mathbf{p}_{j_m}}(\eta_1)\tilde{\delta}_{\mathbf{p}_2-\mathbf{p}_{j_{m+1}}-\cdots-\mathbf{p}_{j_n}}(\eta_2)\Delta^{i_{m+1}}_{\mathbf{p}_{j_{m+1}}}(\eta_2)\cdots\Delta^{i_n}_{\mathbf{p}_{j_n}}(\eta_2)\rangle\Big]$$
$$+\,(1\leftrightarrow 2) \tag{B.70}$$

the terms cancel in the $\mathbf{p}\to 0$ limit to give $E_L + E_R = 0$ order by order in the formal expansion in Δ, Q.E.D.