國立臺灣大學理學院物理學研究所
碩士論文
Department of Physics
College of Science
National Taiwan University
Master Thesis

$\mathcal{N} = 7$ 超重力中之優化漸進表現以及 BCFW
Bonus Scaling and BCFW in $\mathcal{N} = 7$ Supergravity

石恩
En Shih

指導教授：黃宇廷 博士
Advisor: Yu-tin Huang, Ph.D.

中華民國 104 年 8 月
August 2015

國立臺灣大學碩士學位論文

口試委員會審定書

N=7 超重力中之優化表現與 BCFW

Bonus Scaling and BCFW in N=7 Supergravity

本論文係 石恩 君（R01222055）在國立臺灣大學物理學系、所完成之碩士學位論文，於民國 104 年 7 月 24 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

（簽名）

（指導教授）

Acknowledgements

Foremost of all, I would like to thank my brilliant collaborator Jin-Yu Liu (柳君諭) and academic advisor Yu-tin Huang (黃宇廷). Thanks for your continuous encouragement, the patience, and the successful collaboration that we have been through. This thesis would never have been if not for both of you.

I also express gratitude to those who have taught me during the time that I have stayed in National Taiwan University: Heng-Yu Chen (陳恆榆), Jiuun-Wei Chen (陳俊瑋), Pisin Chen (陳丕燊), Pei-Ming Ho (賀培銘), Yeong-Chuan Kao (高涌泉).

Special thanks to those who have accompanied me on the path: Hsun-Chi Chan (詹勳奇), Shun-Hsiang Chan (詹舜翔), Chia-Hsun Chang (張家勳), Chih-Kai Chang (張智凱), Po-Wen Chang (張柏文), Wen-Hsuan Chang (張文軒), Ching-Han Chen (陳璟翰), Chin-Jen Chen (陳勁任), Hsiao-Yi Chen (陳曉邑), Wei-Ming Chen (陳韋名), Yen-Chi Chen (陳彥奇), Ta-Yu Chiang (江大右), Dah-Wei Chiou (邱大維), Shih-Kai Chou (周士凱), Shu-Kuan Chu (朱書寬), Yi-Hsien Du (杜怡嫻), Chih-Shan Fang (方之珊), Yao-Chieh Hu (胡耀傑), Tzu-Hsuan Hung (洪子軒), Chao-Jung Lee (李昭融), Norton Lee (李瑞安), Ka Hou Leong (梁家豪), Hung-Hsi Li (李鴻禧), Yen-Hsien Li (李彥賢), Yen-Ting Li (李彥頲), Chia-Hao Lin (林家豪), Kuan-Ting Lin (林冠廷), Tzu-Lu Lin (林子路), Yao-Yu Lin (林曜宇), Kota Naito (內藤航太), Yen-Chin Ong (王元君), Jason Payne (潘傑森), Bryan Teo (張業祥), Tsung-Hsuan Tsai (蔡宗軒), Chih-Yin Tzeng (曾芝寅), Wei-Han Tzeng (曾威翰), Wayne Wang (翁偉恩), Yu-Jie Wei (魏瑀潔), Chih-Wei Yang (楊智偉), Twang-Hao Zhang (詹敦皓) and many others.

I would also like to thank Donal O'Connell, Nima Arkani-Hamed, Johannes Henn, Henrik Johansson, and Kasper Larsen for the interesting discussions shared during the IAS Focused Program on Scattering Amplitudes in Hong Kong. Thanks to Shiu Fu Hong and Sam Wong for their hospitality during the visit.

Finally, I want to thank my family: Mom, Dad and Lei, for your never-ending support along the way.

摘要

本論文概覽重力理論中計算散射振幅的新發展，並且針對一種新的計算方法做探討。我們首先回顧旋量-螺度理論、在殼超重力以及在殼遞迴式的理論基礎。之後我們重點討論使用 BCFW 遞迴式對於超重力振幅的計算。特別是我們呈現一種基於 $\mathcal{N}=7$ 超重力中 BCFW 遞迴式的散射振幅展開式。這個表示式能夠顯現重力振幅在特定高能量極限下優化的表現。這是尋找重力振幅的自然構件所踏出的初步研究，其終極目標是揭露重力的結構，以其建構能顯現其隱藏對稱性的描述。

關鍵詞：散射振幅、超重力、$\mathcal{N}=8$、$\mathcal{N}=7$、BCFW 遞迴、大 z、壞位移

Abstract

This thesis reviews some aspects of the modern developments in calculation methods and assesses a new expression for scattering amplitudes in gravity. We first revisit the basics of spinor helicity formalism, on-shell superspace, and on-shell recursion relations. Special focus is then given to calculating supergravity amplitudes using BCFW recursion relations. In particular, we present an expansion in the form of a BCFW representation in $\mathcal{N} = 7$ supergravity which can manifest bonus behavior of gravity amplitudes under certain high energy limits. This is a initial step in search of natural building blocks for supergravity amplitudes, taken with the eventual goal of uncovering the structure of gravity and providing a description that can manifest its hidden symmetries.

Keywords: scattering amplitudes, supergravity, $\mathcal{N} = 8$, $\mathcal{N} = 7$, BCFW recursion, large-z, bad shift

Contents

1. Introduction **7**

2. On-shell amplitude methods **12**

2.1. Spacetime symmetry and spinor-helicity 12
2.1.1. The Lorentz and Poincaré groups 13
2.1.2. Spinor helicity formalism 19
2.2. Supersymmetry 23
2.2.1. Global supersymmetry 23
2.2.2. On-shell superspace 24
2.2.3. Supersymmetry Ward identities 25
2.3. Scattering amplitudes 26
2.3.1. The MHV classification 26
2.3.2. 3-point amplitudes 27
2.4. On-shell recursion relations 28
2.4.1. On-shell recursion relations: a general formulation 28
2.4.2. BCFW recursion relations 30
2.4.3. Super-BCFW recursion relations 31
2.4.4. Large z behavior under BCFW shifts 32

3. Supergravity amplitudes **34**

3.1. Perturbative gravity 34
3.2. Supergravity 37
3.3. $\mathcal{N}=8$ supergravity 39
3.3.1. $\mathcal{N}=7$ formalism of $\mathcal{N}=8$ supergravity 39
3.3.2. $\mathcal{N}=7$ BCFW recursion 41
3.4. $\mathcal{N}=7$ "bad shift" BCFW representation 42
3.4.1. A particular $[-,+\rangle$ test shift: $\mathrm{N^kMHV}$ amplitudes 42
3.4.2. Large z behavior under general test shifts 46
3.4.3. Comparison to other formulas for supergravity amplitudes 48
3.4.4. $\mathcal{N}=8$ bonus relations and $\mathcal{N}=7$ bonus scaling: the MHV case . 49
3.5. Bonus scaling of "bad shift" BCFW for string amplitudes 50

4. Conclusion and Future directions **54**

A. Amplitudes of Yang-Mills **56**

A.1. Yang-Mills and super-Yang-Mills 56
A.2. Color structure of Yang-Mills amplitudes 58
A.3. $\mathcal{N}=4$ super-Yang-Mills 60

1. Introduction

In search of a geometrical picture for gravity amplitudes

One of the fascinating themes in the study of scattering amplitudes, is that the amplitude is often a solution to a geometric question. A notable example arises from the Britto, Cachazo, Feng, and Witten (BCFW) recursion relations [1], which serve as a method to construct higher-point tree-level amplitudes from lower-point amplitudes iteratively,

$$\mathcal{A}_n \overset{\text{BCFW recursion}}{=} \sum \underbrace{\left(\text{Input from lower-point amplitudes } \mathcal{A}_{m,m<n}\right)}_{\text{BCFW term}} . \tag{1.1}$$

A geometric picture realized in the context of $\mathcal{N} = 4$ super Yang-Mills (SYM), is that the building blocks for the n-point amplitude with k negative helicity gluons constructed via BCFW recursion are actually associated with the Grassmannian $\mathrm{G}(k, n)$ [2, 3], which is the moduli space of k-planes in n-dimensional space.

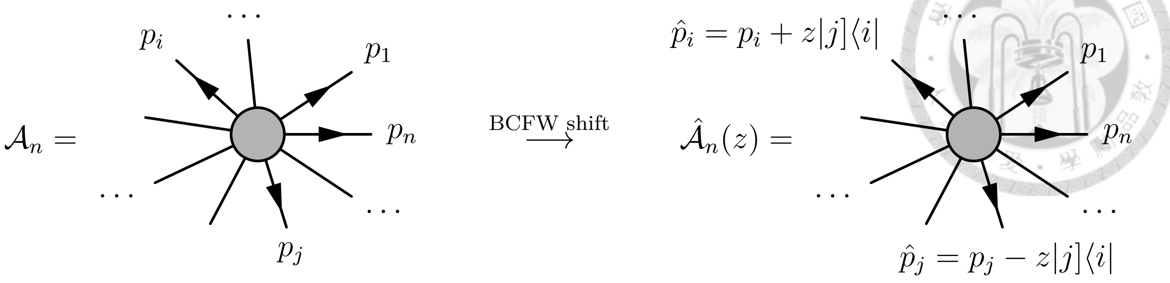
A natural question is whether such geometric structure exists outside of $\mathcal{N} = 4$ SYM. While certain progress along a similar track has been made for $\mathcal{N} = 6$ super Chern-Simons matter theory (CSM) [5, 6, 7, 8, 9], progress has been more limited for gravity. The common property between $\mathcal{N} = 4$ SYM and $\mathcal{N} = 6$ CSM theory is that both allow for color decomposition such that color ordered amplitudes can be defined, and that the theories enjoy an infinite dimensional Yangian symmetry [10]. In fact the building blocks that arise from BCFW recursion are individually Yangian invariant,

$$\underbrace{\mathcal{A}_n^{\mathcal{N}=4 \text{ SYM}}}_{\text{Yangian invariant}} \overset{\text{BCFW recursion}}{=} \sum \underbrace{(\text{BCFW term})}_{\text{Yangian invariant}} . \tag{1.2}$$

For gravity, both of the above properties are absent and thus it may be unclear how to proceed. Instead a question we can ask is that, if there are natural building blocks for gravity amplitudes what would be the nice property one can ask from it, similar to Yangian invariance for the gauge theories.

Large-z behavior under BCFW shifts: a probe of the high energy limit

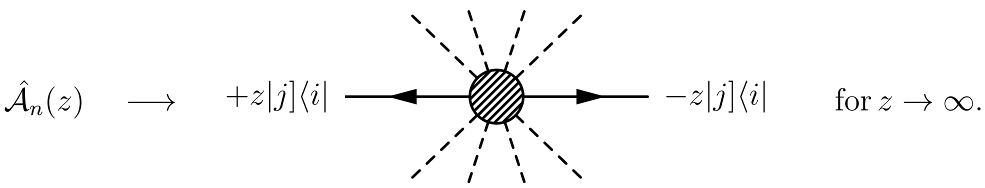
One special property of gravity amplitudes is the asymptotic behavior in the large energy limit. This can be probed in the context of a BCFW shift, where we deform the amplitude by choosing two of the external momenta, say i and j, and shifting them,



(1.3)

This BCFW shift is called an $[i^{h_i}, j^{h_j}\rangle$ shift, where h_i, h_j are the helicities of legs i and j respectively. Note that a BCFW recursion relation is constructed out of a certain choice of BCFW shift. Therefore, the BCFW shift plays a double role in our discussion. First, as the foundation for building amplitudes in terms of BCFW recursion relations. Second, as a test to probe the high energy behavior of the amplitude (or partial expressions of it.)

In order to probe the high energy behavior of the amplitude, we perform a test shift and look at the limit where z is taken to infinity. The large z behavior of the amplitude depends on the helicity configurations of the shifted legs, among which $[-,+\rangle$ shifts are especially of physical interest. Physically, this limit corresponds to a hard particle shooting through a soft background,



(1.4)

Under this probe, gravity amplitudes express their tamed high energy behavior as faster large z fall-off. More precisely, if one performs a test $[-,+\rangle$ shift, as the deformation parameter z is taken to infinity, gravity amplitudes fall off as $1/z^2$ [11][1], compared with $1/z$ of Yang-Mills,

$$\begin{aligned} \hat{A}(z) &\longrightarrow \mathcal{O}\left(\frac{1}{z}\right) \\ \hat{M}(z) &\longrightarrow \mathcal{O}\left(\frac{1}{z^2}\right) \end{aligned} \qquad \text{for } z \to \infty. \tag{1.5}$$

(We will use A_n to denote a Yang-Mills amplitude, and M_n to denote a gravity amplitude.)

Note that individual terms in a generic representation of the amplitude are allowed to cancel only in the sum to express the $1/z^2$ fall-off of the whole amplitude. Thus in the same light that "good" building blocks manifest the Yangian symmetry of $\mathcal{N} = 4$ SYM

[1] Recently, it has been shown that this asymptotic behavior can be attributed to the permutation invariance of gravity amplitudes [12].

amplitudes, we propose that a criteria for a "good" building block for gravity amplitudes is the manifestation of the $1/z^2$ scaling of the full amplitude under any pair of test shift momenta, which we write schematically as

$$\underbrace{\hat{M}_n(z)}_{\text{shifted amplitude}} = \mathcal{O}\left(\frac{1}{z^2}\right) = \sum \underbrace{\mathcal{O}\left(\frac{1}{z^2}\right)}_{\text{shifted individual term}} \quad \text{for a "good" formula.} \tag{1.6}$$

To begin with, let us relax our criteria and ask if one chooses two particular legs to deform, whether there is a formula such that individual terms scale as $1/z^2$ under large deformation. There are already known representations with term-by-term $1/z^2$ fall off for MHV amplitudes under particular choices of test shift legs [13]. However, expressions with such properties have not been previously obtained beyond the MHV sector.

$\mathcal{N} = 8$ supergravity and its $\mathcal{N} = 7$ formulation

As the unique maximally supersymmetric gravity theory in 4d, $\mathcal{N} = 8$ supergravity is chosen as our theory of interest. The individual terms in representations of $\mathcal{N} = 8$ supergravity amplitudes constructed via usual BCFW recursion do not exhibit the desired $1/z^2$ fall-off of the amplitude under BCFW test shifts,

$$\hat{\mathcal{M}}_n^{\mathcal{N}=8}(z) \overset{\text{BCFW recursion}}{=} \sum \left[\mathcal{O}\left(\frac{1}{z}\right) \text{ or worse}\right]. \tag{1.7}$$

In search of a formula that manifest $1/z^2$ fall-off of the amplitude term-by-term, we reformulate $\mathcal{N} = 8$ supergravity in an alternative $\mathcal{N} = 7$ formalism, as suggested by the work of Hodges [14]. There it was noted that BCFW recursion based on a $\mathcal{N} = 7$ formalism gives a more efficient evaluation of MHV amplitudes, and leads to formulas with greater symmetry. The $\mathcal{N} = 7$ formalism can be obtained from the usual $\mathcal{N} = 8$ by SUSY reduction.

$\mathcal{N} = 7$ supergravity $[+, -\rangle$ "bad-shift" BCFW representation

The final ingredient that leads us to term-by-term $1/z^2$ fall off of is the choice of using the $[+, -\rangle$ "bad shift" to construct recursion. This special property of terms arising from the $\mathcal{N} = 7$ bad-shift was first noted for the six-point MHV amplitude by Hodges.

In this thesis we present a representation manifesting the $1/z^2$ fall off that extends to general tree-level amplitudes, in the form of a BCFW recursion in $\mathcal{N} = 7$ supergravity, constructed out of a $[+, -\rangle$ "bad shift". More precisely we claim that if one constructs the representation of a $\mathcal{N} = 7$ amplitude based on the $[j^+, i^-\rangle$ "bad shift", then the individual terms in the BCFW expansion scale at large z as $1/z^2$ under the dual $[i^-, j^+\rangle$ test shift of the same primary shifted legs i, j,

$$\underbrace{\hat{\mathcal{M}}_n^{\mathcal{N}=7}(z)}_{[i^-, j^+\rangle \text{ test shifted amplitude}} \overset{[j^+, i^-\rangle \text{ recursion}}{=} \sum \underbrace{\mathcal{O}\left(\frac{1}{z^2}\right)}_{[i^-, j^+\rangle \text{ test shifted BCFW term}} . \tag{1.8}$$

As we will argue, the reason why $\mathcal{N} = 7$ bad shift recursion allows for term-by-term $1/z^2$ fall-off is because it secretly uses the $1/z^2$ fall-off of the full amplitude. The presence of $1/z^2$ fall-off for the $\mathcal{N} = 8$ amplitude implies extra "bonus relations" [15]. As we will show, for MHV amplitudes, it is precisely due to the incorporation of these bonus relations that the $\mathcal{N} = 7$ bad shift recursion exhibit improved term-by-term fall-off relative to $\mathcal{N} = 8$.

Layout of the thesis

Conventional formalisms for calculating scattering amplitudes suffers from unmanageable large number of Feynman diagrams and redundant information that obscures the hidden structure and symmetries that have become more apparent due to the developments in recent years. It is due to modern amplitude methods such as the spinor-helicity formalism, on-shell superspace, and on-shell recursion relations that have made the discussions in this thesis possible. Therefore before we discuss the main topic of this thesis, we will first review the necessary modern amplitude methods that the discussion is built on. We list here the topics discussed in this thesis in order, which will be further expanded in later chapters. More information on these topics can be found in reviews such as [4][16][17][18].

- **Spinor helicity formalism**
  Weyl spinor brackets are identified as the fundamental invariants for massless kinematics, in terms of which scattering amplitudes are expressed. On-shell amplitudes are specified using only on-shell information: the momenta and helicities of external particles, removing the reference to polarization vectors.

- **Supersymmetry**
  Supersymmetry unify particles of different helicites. The existence of supersymmetry can largely constrain the theory and lead to nicer properties. Maximally supersymmetric theories such as $\mathcal{N} = 4$ SYM and $\mathcal{N} = 8$ supergravity are often studied as toy models which allow more simplified descriptions in the on-shell formalism. In our current discussion, the theory of interest is $\mathcal{N} = 8$ supergravity.

- **On-shell superspace formalism**
  On-shell superspace combines the ideas of spinor-helicty and supersymmetry, such that supercharges are also described in terms of on-shell variables. Consequently a supermultiplet can be efficiently described by a single superfield, and several different amplitudes can be grouped into a single superamplitude.

- **3-point amplitudes**
  3-point amplitudes of massless particles can be determined uniquely by helicity structure under the assumption of locality. Furthermore, complex-momentum 3-point amplitudes serve as the basic building blocks to construct higher-point amplitudes.

- **Analytical structure of tree level amplitudes**
  From the knowledge of the analytical structure of an amplitude, we are able to reconstruct the functional form. The singularities of tree level amplitudes consist

of only propagator poles. On these poles, the amplitude factorizes into two on-shell subamplitudes. To probe the singularities in a way that enable us to utilize the powerful tools of complex analysis and reconstruct the amplitude, we introduce a complex parameter z to deform chosen momenta.

- **On-shell recursion relations**
  On-shell recursion relations allows up to build higher-point on-shell amplitudes form lower-point on-shell amplitudes recursively, ultimately reducing to 3-point building blocks. The derivation of these relations is presented based on complex analysis combined with the knowledge of the analytical structure of amplitudes. What we then obtain is an on-shell formalism which enables us to calculate amplitudes without requiring the Lagrangian or other off-shell information. Different schemes of deforming momenta lead to different recursion relations.

- **BCFW deformations and recursion relations**
  Among the on-shell recursion relations, the most famous are the BCFW recursion relations, which we will use extensively in this thesis. Furthermore, the BCFW shift which recursion is based on can be utilized to characterize high energy behavior through the fall-off of the deformation parameter z. The supersymmetric versions of the BCFW relations are called super-BCFW relations.

- $\mathcal{N}=8$ **supergravity and its** $\mathcal{N}=7$ **formalism**
  The unique maximally supersymmetric gravity theory in 4d is $\mathcal{N}=8$ supergravity. We can SUSY reduce a theory of greater supersymmetry to obtain a theory of less supersymmetry. When the on-shell superspace of $\mathcal{N}=8$ supergravity is SUSY reduced to $\mathcal{N}=7$, the $\mathcal{N}=8$ superfield splits into two superfields Φ and Ψ. The $\mathcal{N}=7$ theory, written in the $\Phi-\Psi$ formalism, has the same on-shell degrees of freedom as with $\mathcal{N}=8$ supergravity, only with a reduced set of supersymmetry being manifest. High energy behavior represented by the large z fall-off under BCFW deformation depend on the supersymmetry $\mathcal{N}$.

- $\mathcal{N}=7$ **"bad shift" BCFW and term-by-term bonus fall-off**
  We will show that by utilizing the $\mathcal{N}=7$ BCFW recursion relations, which are derived by SUSY truncation from $\mathcal{N}=8$ BCFW, we are able to obtain an expression which manifest the $1/z^2$ large z fall-off of the $\mathcal{N}=8$ amplitude. Furthermore, the improved fall-off of $\mathcal{N}=7$ BCFW terms can be linked to the incorporation of bonus relations implied by the improved fall-off of the $\mathcal{N}=8$ amplitude.

- **Bonus fall-off of "bad shift" BCFW for string amplitudes**
  We extend the discussion of bonus term-by-term fall-off in "bad shift" BCFW recursion to string amplitudes. The improved high-energy behavior of string amplitudes is represented by enhanced large z scaling, such that we can expose better term-by-term large z fall-off compared to field theory when supersymmetry is reduced to less than $\mathcal{N}=7$.

2. On-shell amplitude methods

2.1. Spacetime symmetry and spinor-helicity

The conventional formalism of expressing scattering amplitudes regards them as functions of the momenta p_i and polarization vectors ϵ_i (or spin states, polarization tensors, ..., depending on the spin of the particle to represent the external line state ϕ_i) of each external particle i,

$$A(p_i; \epsilon_i). \tag{2.1}$$

Since the amplitude is invariant under spacetime symmetries, the actual expression is given in terms of the invariants formed by p_i and ϵ_i:

$$p_i \cdot p_j, \quad \epsilon_i \cdot \epsilon_j, \quad p_i \cdot \epsilon_j. \tag{2.2}$$

This formulation manifests Lorentz covariance but suffers from much redundancy. For example, given a momentum p_i for an external gluon, we know that the polarization vectors ϵ_i must be transverse and satisfy $p_i \cdot \epsilon_i = 0$, and that the gauge transformation $\epsilon_i \to \epsilon_i + w\, p_i$ leaves the amplitude invariant. This generates many different but physically and numerically equivalent expressions of the same amplitude. Moreover, the on-shell condition $p_i^2 = 0$ of massless kinematics needs to be additionally imposed on the momenta. In light of this, it would be much more economical if we could find a suitable set of invariant variables that can embody the information of the external particles without the redundancy. This leads us to the spinor-helicity formalism, which replaces the reference to vector momenta p_i by Weyl spinor variables $\lambda_i, \tilde{\lambda}_i$ to incorporate the on-shell condition, and removes the necessity of polarization vectors (tensors) by only referencing the helicities h_i of the external particles. We can write the amplitude as,

$$A(\lambda_i, \tilde{\lambda}_i; h_i). \tag{2.3}$$

Instead of referring to invariants formed by p_i and ϵ_i, the amplitude is given in terms of fundamental kinematic invariants formed by the spinors $\lambda_i, \tilde{\lambda}_i$, also written as square and angle brackets:

$$\lambda_i \cdot \lambda_j \equiv [ij], \quad \tilde{\lambda}_i \cdot \tilde{\lambda}_j \equiv \langle ij \rangle. \tag{2.4}$$

Spinor-helicity does more than providing the amplitude a nicer expression. In adopting the formalism, calculation of amplitudes can be vastly reduced in length due to the reduction of redundant information. Moreover, the powerful on-shell superspace and on-shell recursion relies much on the spinor-helicity language.

The foundations of spinor-helicity methods are inherently tied to the symmetries of spacetime. In this section we will first review the Lorentz and Poincaré groups and their representations, and then introduce the basics of the spinor-helicity formalism. We will work exclusively in 1+3 dimensions. The discussion follows [19][20][21][22].

2.1.1. The Lorentz and Poincaré groups

Spacetime symmetry is described by the Poincaré group. Note that this is only an approximate symmetry, broken by the non-flat spacetime due to the presence of gravity. We will however work with perturbative gravity, which is formulated as standard quantum field theory in flat spacetime with a fluctuating graviton field. We consider four-dimensional flat Minkowski spacetime endowed with a "mostly minus" signature metric, $\eta = \mathrm{diag}(1, -1, -1, -1)$, which can be used to raise and lower indices,

$$x_\mu = \eta_{\mu\nu} x^\nu, \quad x^\mu = \eta^{\mu\nu} x_\nu. \tag{2.5}$$

Consider the set of Minkowski spacetime isometries (transformations which preserve the metric). Under such a transformation acting on the coordinates, $x \to x'$, the inverse metric is invariant,

$$\eta^{\mu\nu} \to \eta'^{\mu\nu} = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} \eta^{\alpha\beta} = \eta^{\mu\nu}. \tag{2.6}$$

We therefore infer they are transformations of the form

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \tag{2.7}$$

with the condition that

$$\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \eta^{\alpha\beta} = \eta^{\mu\nu}. \tag{2.8}$$

The Minkowski spacetime isometries form a Lie group, called the Poincaré group. The group elements are parameterized by $(\Lambda^\mu{}_\nu, a^\mu)$, and have the group multiplication

$$(\Lambda_2, a_2) \cdot (\Lambda_1, a_1) = (\Lambda_2 \Lambda_1, \Lambda_2 a_1 + a_2). \tag{2.9}$$

Poincaré group elements of the form $(\Lambda^\mu{}_\nu, 0)$ correspond to pure Lorentz transformations, while $(\delta^\mu{}_\nu, a^\mu)$ correspond to pure spacetime translations. Lorentz transformations and spacetime translations form a group among themselves respectively.

The Lorentz and Poincaré algebras

Due to the correspondence between Lie groups and Lie algebras, we can study the Lorentz and Poincaré groups in terms of their algebra, and write the group elements in terms of generators.

Let us reparametrize the Lorentz group elements by defining $\Lambda^\mu{}_\nu \equiv (e^\omega)^\mu{}_\nu$. Now consider an infinitesimal transformation $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu$, we see that (2.8) translates to the condition that $\omega^{\mu\nu}$ is antisymmetric,

$$\omega^{\mu\nu} + \omega^{\nu\mu} = 0. \tag{2.10}$$

Then let us write an element of the Lorentz group connected to the identity as an exponentiation of the generator $M_{\mu\nu}$, defined with respect to the group parameter $\omega^{\mu\nu}$,

$$\Lambda = e^{i\omega^{[\mu\nu]} M_{[\mu\nu]}} \tag{2.11}$$

Considering its infinitesimal transformation on x^α near the identity, we get

$$i\omega^{[\mu\nu]} M_{[\mu\nu]} x^\alpha = \omega^\alpha{}_\beta x^\beta. \tag{2.12}$$

We are therefore able to identify following the field representation for the generators of the Poincaré group,

$$M_{[\mu\nu]} = i(x_\mu \partial_\nu - x_\nu \partial_\mu) \tag{2.13}$$

$$P_\mu = -i\partial_\mu. \tag{2.14}$$

Under this field representation, the commutation relations of the Poincaré algebra are easily derived:

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(M^{\mu\rho}\eta^{\nu\sigma} - M^{\mu\sigma}\eta^{\nu\rho} - M^{\nu\rho}\eta^{\mu\sigma} + M^{\nu\sigma}\eta^{\mu\rho}) \tag{2.15}$$

$$[P_\mu, P_\nu] = 0 \tag{2.16}$$

$$[M_{\mu\nu}, P_\rho] = i(P_\mu \eta_{\nu\rho} - P_\nu \eta_{\mu\rho}), \tag{2.17}$$

From the commutation relations, we see that Lorentz transformations and spacetime translations do not commute. The Poincaré group is a semi-direct product of the Lorentz group and the translation group,

$$\mathrm{ISO}(1,3) = T^{(4)} \ltimes \mathrm{SO}(1,3). \tag{2.18}$$

The Poincaré group is also referred to as the inhomogeneous Lorentz group, $\mathrm{ISO}(1,3)$.

From the rules of matrix multiplication, we can also identify the defining matrix representation for the Lorentz generators:

$$(M_{[\mu\nu]})^{\alpha\beta} = i(\delta^\alpha{}_\mu \delta^\beta{}_\nu - \delta^\alpha{}_\nu \delta^\beta{}_\mu). \tag{2.19}$$

Under this representation, the infinitesimal Lorentz transformation acts as the matrix

$$(i\omega^{[\mu\nu]} M_{[\mu\nu]})^\alpha{}_\beta = \omega^\alpha{}_\beta, \tag{2.20}$$

and a generic group element is just the Lorentz transformation matrix:

$$\begin{aligned}(\Lambda)^{\alpha}{}_{\beta} &= (e^{i\omega^{[\mu\nu]}M_{[\mu\nu]}})^{\alpha}{}_{\beta} \\ &= (e^{\omega})^{\alpha}{}_{\beta} \\ &= \Lambda^{\alpha}{}_{\beta}.\end{aligned} \tag{2.21}$$

More generally, any set of matrices satisfying the commutation relations (2.15) furnish a linear representation of the Lorentz algebra.

Fields and particles

We now turn our focus to the actors on our spacetime stage: fields and particles, which correspond to particular representations of the Lorentz and Poincaré groups.

To classify all irreducible representations of the Lorentz algebra, we are going to build up on the fundamental spin representations of $\mathfrak{su}(2)$. On the other hand, the classification of nonnegative energy unitary irreducible representations of the Poincaré algebra is a classic result by Wigner, namely the Wigner classification. Both constructions follow the standard procedure of characterizing representations by the eigenvalues of Casimir operators. The main result for massless kinematics is that we can use the helicity h to characterize different kinds of particles.

Representations of the Lorentz group

To classify representations of the Lorentz group, we would like to rewrite the Lorentz commutation relations into a more convenient form. To do so, first we separate the Lorentz generators into rotations and boosts,

$$L^i = \frac{1}{2}\epsilon^{ijk}M_{jk}, \quad K^i = M_{0i}, \tag{2.22}$$

and define

$$J_L^i = \frac{1}{2}(L^i + K^i), \quad J_R^i = \frac{1}{2}(L^i + K^i) \tag{2.23}$$

Then the Lorentz commutation relations become

$$[J_L^i, J_R^j] = 0 \tag{2.24}$$

$$[J_L^i, J_L^j] = i\epsilon^{ijk}J_L^k, \quad [J_R^i, J_R^j] = i\epsilon^{ijk}J_R^k. \tag{2.25}$$

We see that the algebra of these generators decouple into two SU(2) algebras. The representations of the Lorentz algebra are therefore characterized by the two independent Casimir operators J_L^2 and J_R^2. We have to be careful though that $\mathrm{SO}(1,3)$ is not isomorphic to $\mathrm{SU}(2) \times \mathrm{SU}(2)$. The actual relation is that the universal covering (double covering in this case) of the $\mathrm{SO}(1,3)$, which is the spin group $\mathrm{Spin}(1,3)$, is isomorphic to the special linear group $\mathrm{SL}(2,\mathbb{C})$, which is in turn isomorphic to the complexification (complex linear combinations) of SU(2). We can write $\mathrm{SO}(1,3) = \mathrm{SL}(2,\mathbb{C})/\mathbb{Z}_2$. Since the mapping from $\mathrm{SL}(2,\mathbb{C})$ to $\mathrm{SO}(1,3)$ is two-to-one, we will actually miss the spinor representation if we only look at representations of $\mathrm{SO}(1,3)$ but not the representations of $\mathrm{SL}(2,\mathbb{C})$. Therefore from now on, we will regard $\mathrm{SL}(2,\mathbb{C})$ as the "Lorentz group" and study its representations.

Aside these subtleties, the main result is that we can label each irreducible representation of the Lorentz group by a pair of half-integer SU(2) "spins" (j_L, j_R),

The eigenvalues (j_L, j_R) of the Casimir operators J_L^2 and J_R^2 characterize the irreducible representations of the Lorentz group.

For example, the $(\frac{1}{2}, 0)$ representation is the left-handed(chiral) Weyl spinor, the $(0, \frac{1}{2})$ representation is the right-handed(anti-chiral) Weyl spinor, while the $(\frac{1}{2}, \frac{1}{2})$ representation is the Lorentz vector. The quantum fields are realizations of various representations of the Lorentz group.

Weyl spinor formalism

As we have seen in the previous section, we can build all representations of the Lorentz group in terms of the two-component chiral and antichiral Weyl spinor representations. Therefore we can discard $\mathrm{SO}(1,3)$ vector indices and rewrite representations in terms of $\mathrm{SL}(2,\mathbb{C})$ spinor indices. Such a notation is sometimes called the van der Waerden notation.

Chiral spinors, transforming as $(\frac{1}{2}, 0)$, are given undotted spinor indices,

$$\psi_a, \tag{2.26}$$

while antichiral spinors, transforming as $(0, \frac{1}{2})$, are given dotted spinor indices,

$$\tilde{\psi}^{\dot{a}}. \tag{2.27}$$

Analogous to the invariant metric under $\mathrm{SO}(1,3)$, we can build the antisymmetric tensors ϵ^{ab} and $\epsilon^{\dot{a}\dot{b}}$,

$$\epsilon^{ab} = \epsilon^{\dot{a}\dot{b}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \epsilon_{ab} = \epsilon_{\dot{a}\dot{b}}. \tag{2.28}$$

Spinor indices can be raised and lowered using ϵ^{ab} and $\epsilon^{\dot{a}\dot{b}}$,

$$\psi^a = \epsilon^{ab}\psi_b, \quad \tilde{\psi}_{\dot{a}} = \epsilon_{\dot{a}\dot{b}}\tilde{\psi}^{\dot{b}}. \tag{2.29}$$

Lorentz invariant spinor products can be formed by contracting pairs of spinors indices,

$$\psi\chi \equiv \psi^a\chi_a, \quad \tilde{\psi}\tilde{\chi} \equiv \tilde{\psi}_{\dot{a}}\tilde{\chi}^{\dot{a}}. \tag{2.30}$$

Notice that the position of the contracted index in the antisymmetric metric matters, and that raising one while lowering the other contracted index will incur a negative sign.

Other Lorentz representations can be built up from various tensor products of the spinor representations, and can be reduced to irreducible representations by Clebsch–Gordan decomposition.

The transcription between vector and spinor indices are provided by the Pauli matrices. The defining representation of the Lorentz group $\mathrm{SL}(2,\mathbb{C})$ are the complex 2×2 matrices with unit determinant, spanned by the Pauli matrices:

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.31}$$

and their contravariant basis

$$\bar{\sigma}_\mu = \eta_{\mu\nu}\sigma^\nu. \tag{2.32}$$

A Lorentz vector A^μ can be represented by a $\mathrm{SL}(2,\mathbb{C})$ matrix using the Pauli matrices as the basis,

$$A = A_\mu \sigma^\mu = A^\mu \bar{\sigma}_\mu = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 + x^3. \end{pmatrix} \tag{2.33}$$

The four-vector components can be recovered by (see also section A.1)

$$A^\mu = \frac{1}{2}\operatorname{tr}\sigma^\mu A. \tag{2.34}$$

The determinant is the Lorentz invariant vector norm squared,

$$\det A = A_\mu A^\mu. \tag{2.35}$$

Writing out the $\mathrm{SL}(2,\mathbb{C})$ spinor indices on the matrices explicitly, we see that

$$A_{a\dot{b}} = A_\mu \sigma^\mu_{a\dot{b}}. \tag{2.36}$$

The Poincaré symmetry generators can be written in spinor index notation by contracting the Lorentz indices using the Pauli sigma matrices,

$$M_{a\dot{a}b\dot{b}} = M_{\mu\nu}\sigma^\mu_{a\dot{a}}\sigma^\nu_{b\dot{b}} \tag{2.37}$$

$$P_{a\dot{a}} = P_\mu \sigma^\mu_{a\dot{a}}. \tag{2.38}$$

The Lorentz generator can be separated into two parts acting on the chiral and antichiral indices respectively,

$$M_{a\dot{a}b\dot{b}} = \epsilon_{ab}\ell_{(\dot{a}\dot{b})} + \epsilon_{\dot{a}\dot{b}}\tilde{\ell}_{(ab)}. \tag{2.39}$$

Representations of the Poincaré group and the little group

The classification of the representations of the Poincaré group was first introduced by Wigner. The representations can be characterized by the eigenvalues of the following independent Casimir operators of the Poincaré group: the square of the momentum operator P_μ and the square of the Pauli-Lubanski pseudovector W_μ. The Pauli-Lubanski pseudovector is defined by

$$W_\mu \equiv \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}M^{\nu\rho}P^\sigma. \tag{2.40}$$

The Casimir operators of the Poincaré group are

$$P^2 = P_\mu P^\mu, \quad W^2 = W_\mu W^\mu. \tag{2.41}$$

The eigenvalue of P^2 is the squared mass m^2. We will restrict our discussion to representations with $m^2 \geq 0$, and distinguish between the cases $m^2 > 0$ (massive particles) and $m^2 = 0$ (massless particles). Since W^2 is a Lorentz invariant, we can choose a particular frame in which P_μ is reduced to a standard normal form to evaluate W_μ.

- $m^2 > 0$:
  For massive particles, we can go to the rest frame of the particle where $P_\mu = (m, 0, 0, 0)$ to evaluate W_μ,

  $$W_\mu = 0, \quad W_i = mS_i, \tag{2.42}$$

  with the spin vector S_i is defined by

  $$S_i \equiv \frac{1}{2}\epsilon_{ijk}M^{jk}. \tag{2.43}$$

  so we can classify representations of the Poincaré group by the Casimir operator S^2 since

  $$W^2 = m^2 S^2. \tag{2.44}$$

  The S_i are the generators of the little group, the transformations which leave the normal form of P_μ invariant. For the massive case, the little group is SO(3) and the eigenvalues of S^2 are $s(s+1)$ with $s \in \frac{1}{2}\mathbb{N}$. Therefore we learn that massive representations of the Poincaré group can be labeled by the quantum number s, called the spin.

The mass m and the spin $s \in \frac{1}{2}\mathbb{N}$ characterize the massive irreducible representations of the Poincaré group.

- $m^2 = 0$:
  For the massless case, we choose the lightcone frame $P_\mu = (E, 0, 0, E)$ as the normal form. The Pauli-Lubanski pseudovector becomes

  $$W_0 = EL, \quad W_3 = EL, \quad W_1 = ET_1, \quad W_2 = ET_2, \tag{2.45}$$

where

$$L \equiv M_{12}, \quad T_1 \equiv M_{02} + M_{23}, \quad T_2 \equiv M_{01} + M_{13} \tag{2.46}$$

are the rotation and translation generators of the little group ISO(2). We classify the representations of the little group (and the massless Poincaré group) by the Casimir operator

$$T_1^2 + T_2^2 = \mu^2. \tag{2.47}$$

The case for μ^2 is analogous to the discussion for m^2. We distinguish between the cases $\mu^2 > 0$ and $\mu^2 = 0$. The physical representations correspond to the case where $\mu^2 = 0$, such that T_1 and T_2 act trivially, so the group action is generated by the single SO(2) generator L. Therefore the representations of the little group are labeled by $h \in \frac{1}{2}\mathbb{N}$, called the helicity.

> The helicity $h \in \frac{1}{2}\mathbb{N}$ characterizes the massless irreducible representations of the Poincaré group.

Since under parity transformations $(x_0, \vec{x}) \to (x_0, -\vec{x})$ the helicity changes sign, the spectrum of a CPT invariant theory containing a state of helicity $+h$ must therefore also contain a state of the opposite helicity $-h$. For instance, Weyl spinors comes in left-handed and right-handed pairs ($h = \pm\frac{1}{2}$), and there are two polarizations for gluons ($h = \pm 1$) and gravitons ($h = \pm 2$).

2.1.2. Spinor helicity formalism

As implied by its name, the main idea of the spinor helicity formalism is to capture all information of scattering amplitudes in terms of the representations of the physical states, and perform calculations in that context. The physical representations we are going to work with are labeled by helicity h_i and built up on the chiral and antichiral Weyl spinors. We will discuss only the spinor helicity formalism for massless particles in 1+3 dimensional spacetime, although the formalism can also be extended to massive particles and other dimensions. The discussion follows [4].

In $\mathrm{SL}(2, \mathbb{C})$ notation, the massless on-shell condition $p^2 = 0$ makes the momentum $p_{a\dot{a}}$ a rank-1 matrix, so we can factorize the momentum into the product of some chiral and antichiral Weyl spinors,

$$p_{a\dot{b}} = \lambda_a \tilde{\lambda}_{\dot{b}}. \tag{2.48}$$

Therefore instead of considering the amplitude as the function of momenta, we can consider it as a function of the spinors λ and $\tilde{\lambda}$. The introduction of these variables makes the on-shell condition automatically implied and built-in. The amplitude can then be expressed by Lorentz invariant inner products of the spinor variables.

We can write these spinor variables in a convenient Dirac bra-ket notation, where chiral spinors are written as square bra-kets, and antichiral spinors as angle bra-kets,

$$\lambda_a = |p]_a, \quad \tilde{\lambda}_{\dot{a}} = \langle p|_{\dot{a}} \tag{2.49}$$

$$p_{a\dot{b}} = \lambda_a \tilde{\lambda}_{\dot{b}} = |p]_a \langle p|_{\dot{b}} \tag{2.50}$$

Bras are converted to kets and vice-versa by raising and lowering indices,

$$|p]^a = \epsilon^{ab}[p|_b, \quad \langle p|^{\dot{a}} = \epsilon^{\dot{a}\dot{b}}|p\rangle_{\dot{b}}, \tag{2.51}$$

and the Lorentz invariant spinor inner products are written as brackets,

$$[pq] = [p|^a|q]_a, \quad \langle pq\rangle = \langle p|_{\dot{a}}|q\rangle^{\dot{a}}. \tag{2.52}$$

These products are antisymmetric due to the antisymmetry of ϵ^{ab} and $\epsilon^{\dot{a}\dot{b}}$,

$$[pq] = -[qp], \quad \langle pq\rangle = -\langle qp\rangle. \tag{2.53}$$

The massless Weyl equation can be written as

$$p|p] = 0, \quad p|p\rangle = 0, \tag{2.54}$$

which is now just the trivial consequence of $[pp] = \langle pp\rangle = 0$ due to antisymmetry.

We will also use angle-square brackets $\langle q|p|r]$ and square-angle brackets $[q|p|r\rangle$, where p may not be necessary lightlike, which are evaluated by projecting p onto the matching sigma-matrix,

$$[q|p|r\rangle = p_\mu[q|\sigma^\mu|r\rangle = p_\mu[q|^a\sigma^\mu_{a\dot{b}}|r\rangle^{\dot{b}} \tag{2.55}$$

$$\langle q|p|r] = p^\mu\langle q|\bar{\sigma}_\mu|r] = p^\mu\langle q|_{\dot{a}}\bar{\sigma}^{\dot{a}b}_\mu|r]_b. \tag{2.56}$$

Note that a angle-square bracket can be rewritten as a square-angle bracket,

$$\langle q|p|r] = [r|p|q\rangle. \tag{2.57}$$

If p is also lightlike, then the square-angle bracket reduces to square and angle brackets,

$$[q|p|r\rangle = [qp]\langle pr\rangle. \tag{2.58}$$

Using the completeness relation (see also section A.1), we can derive

$$\begin{aligned} p\cdot q &= p_\mu q^\mu \\ &= \frac{1}{2}\operatorname{tr}(p\sigma_\mu)q^\mu \\ &= \frac{1}{2}\operatorname{tr}(pq) \\ &= \frac{1}{2}[pq]\langle qp\rangle, \end{aligned} \tag{2.59}$$

so

$$(p+q)^2 = 2p\cdot q = [pq]\langle qp\rangle. \tag{2.60}$$

Note also that

$$\begin{aligned}[p|\gamma^\mu|p\rangle &= [p|\sigma^\mu|p\rangle \\ &= \mathrm{tr}(\sigma^\mu p) \\ &= 2p^\mu \end{aligned} \tag{2.61}$$

Related is the Fierz identity,

$$\langle p|\bar\sigma_\mu|q][r|\sigma^\mu|s\rangle = 2\langle ps\rangle[rq]. \tag{2.62}$$

Another useful relation is the Schouten identity, which results from the fact that any two-component spinor can be written as a linear combination of two others. Let us write a spinor $|k]$ in terms of two other spinors $|i]$ and $|j]$ with some undetermined coefficients,

$$|k] = a|i] + b|j]. \tag{2.63}$$

We can then dot in spinors to solve for the coefficients, then we can obtain an identity

$$|i][jk] + |j][ki] + |k][ij] = 0. \tag{2.64}$$

A similar identity holds for angle spinors.

The $\mathrm{SL}(2,\mathbb{C})$ representations $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ can be regarded as complex conjugate representations, as can be seen from the relation $J_{(L/R)} = L \pm iK$. The representations are sent into one another by Dirac conjugation. For left-handed and right-handed spinors ψ_a and $\bar\chi_{\dot a}$, their Dirac conjugates are

$$(\psi_a)^* = \bar\psi_{\dot a}, \quad (\bar\chi_{\dot a})^* = \chi_a. \tag{2.65}$$

For real momenta p, the reality condition relate the associated square and angle spinors $|p], |p\rangle$ by complex conjugation,

$$[p|^a = |p\rangle^{\dot a *}, \quad \langle p|_{\dot a} = |p]_a{}^*. \tag{2.66}$$

For complex-valued momenta, the associated square and angle spinors are independent.

In amplitudes with many different momenta $p_1, p_2, \dots, p_i, \dots$, we will use the notation $|i] = |p_i], |i\rangle = |p_i\rangle$ etc. Momenta can be written as

$$p_i = |i]\langle i|. \tag{2.67}$$

The total momentum is denoted by

$$P = \sum_i p_i. \tag{2.68}$$

The Mandelstam variables are defined as

$$s_{ij} = (p_i + p_j)^2, \quad s_{ijk} = (p_i + p_j + p_k)^2, \quad \text{etc.} \tag{2.69}$$

Specifically for the two particles i and j,

$$s_{ij} = (p_i + p_j)^2 = 2p_i \cdot p_j = [ij]\langle ji\rangle. \tag{2.70}$$

The power of the spinor helicity formalism is its ability to trivialize the kinematic constraints by introducing a suitable set of variables. It can also be seen as the natural consequence of using the representation theory of the Poincaré group to express physical quantities. Although the spinor-helicity method presented here is formulated in $D = 4$ and restricted to massless particles, a natural generalization to dimensions other than four and the massive case is provided by [23][24].

Little group scaling

In the spinor helicity formalism, the expression for momentum by the Weyl spinors $p = |p]\langle p|$ is invariant under the transformation

$$|p\rangle \to t|p\rangle, \quad |p] \to t^{-1}|p]. \tag{2.71}$$

This is termed little group scaling, as the little group is the transformations which leave the momentum invariant. For a massless particle in 4d, we can go to the frame where $p^\mu = (E, 0, 0, E)$. For real momenta, the little group consists of the rotation in the $x - y$ plane, which can be identified with $\mathrm{SO}(2) = \mathrm{U}(1)$. Correspondingly, t has to be a complex phase such that the reality condition $|p]^* = \langle p|$ is preserved. For complex momenta, t can be any non-zero complex number.

We can deduce how an amplitude scale under little group transformations by noting that internal particles and vertices are invariant, so the transformation is determined by the external line states. Therefore under little group scaling of each external particle i with helicity h_i, the amplitude transforms homogeneously as

$$\mathcal{A}_n(\ldots, \phi_i^{h_i}, \ldots) \to \mathcal{A}_n(\ldots, t^{-2h_i}\phi_i^{h_i}, \ldots) = t^{-2h_i}\mathcal{A}_n(\ldots, \phi_i^{h_i}, \ldots). \tag{2.72}$$

Polarization vectors

For spin-1 fields, polarization vectors are used to specify the state. Writing the momentum as $p^\mu = (E, E\sin\theta\cos\phi, E\sin\theta\sin\phi, E\cos\theta)$, the transverse polarization vectors are given by

$$\epsilon_\pm^\mu(p) = \pm\frac{e^{\mp i\phi}}{\sqrt{2}}(0, \cos\theta\cos\phi \pm i\sin\phi, \cos\theta\sin\phi \mp i\cos\phi, -\sin\theta). \tag{2.73}$$

For $\theta = \phi = 0$, we have $p^\mu = (E, 0, 0, E)$, and the polarization vectors reduce to $\epsilon_\pm^\mu(p) = \pm\frac{1}{\sqrt{2}}(0, 1, \mp i, 0)$. They satisfy $p \cdot \epsilon_\pm(p) = 0$ and the normalization $\epsilon_\pm^* \cdot \epsilon_\pm = -1$. Furthermore, our choice of the helicity basis for polarization have the property of being null: $\epsilon_\pm(p)^2 = 0$. This allows us to rewrite the polarization vectors in the spinor helicity language,

$$\epsilon_+ = \frac{\sqrt{2}}{\langle qp\rangle}|p]\langle q|, \quad \epsilon_- = \frac{\sqrt{2}}{[qp]}|q]\langle p|. \tag{2.74}$$

for arbitrary spinors $|q], |q\rangle$. This is a reflection of the equivalence of choosing any set the polarization vectors under gauge transformation: $\epsilon_\pm \to \epsilon_\pm + c\,|p]\langle p|$.

2.2. Supersymmetry

In discussing scattering amplitudes and symmetry raises the question of the most general symmetry group of the S-matrix that is allowed. This is answered by the Coleman-Mandula theorem which states that the most general symmetry Lie group of the S-matrix in four dimensions is the direct product of the Poincaré group with an internal symmetry group. Supersymmetry, whereas, is an extension of spacetime symmetry that is able to evade the Coleman-Mandula theorem by generalizing the symmetry from a Lie algebra to a graded Lie algebra. Fermionic generators are introduced and the algebra is given by anticommutators between two fermionic generators, and commutators between one fermionic and one bosonic generator.

2.2.1. Global supersymmetry

Specifically, we introduce $\mathcal{N}$ supercharges Q^A_a and their conjugates $\tilde{Q}^A_{\dot{a}}$, satisfying

$$\{Q^A_a, Q^B_b\} = \{\tilde{Q}^A_{\dot{a}}, \tilde{Q}^B_{\dot{b}}\} = 0 \tag{2.75}$$

$$\{Q^A_a, \tilde{Q}^B_{\dot{b}}\} = \delta^{AB} p_{a\dot{b}}. \tag{2.76}$$

The indices $A, B, \dots$ are labels of the $\mathrm{SU}(\mathcal{N})$ R-symmetry that rotate the supersymmetry generators. A way of viewing the relations is to say that the supersymmetry generators are some kind of square root of the momentum generators which are fermionic analogs of momenta, therefore the term "supermomentum". We also have the supermomenta of individual particles, q_i and $\tilde{q}_i$, which sum up to total supermomenta,

$$\sum_i q_i = Q, \quad \sum_i \tilde{q}_i = \tilde{Q}. \tag{2.77}$$

The supercharges can relate states of different helicities. Consider the action of the supercharges on the annihilation operators a associated with the fields. The supercharge $\tilde{Q}^A_{\dot{a}}$ decreases the helicity by $\frac{1}{2}$ and removes the index A, while Q^A_a does the opposite. States related by the supercharges group together to form a supermultiplet. For example for an $\mathcal{N}=2$ supermultiplet,

$$\begin{aligned}
[\tilde{Q}^A_{\dot{a}}, a] &= 0 \\
[\tilde{Q}^A_{\dot{a}}, a^B] &= \tilde{\lambda}_{\dot{a}}\delta^{AB} a \\
[\tilde{Q}^A_{\dot{a}}, a^{BC}] &= \tilde{\lambda}_{\dot{a}}\delta^{A[B} a^{C]}
\end{aligned} \tag{2.78}$$

and

$$\begin{aligned}
[Q^A_a, a] &= \lambda_a a^A \\
[Q^A_a, a^B] &= \lambda_a a^{AB} \\
[Q^A_a, a^{BC}] &= 0.
\end{aligned} \tag{2.79}$$

2.2.2. On-shell superspace

Similar to the way we introduced the spinors $\lambda, \tilde{\lambda}$ to parameterize on-shell massless kinematics in the spinor-helicity formalism, we can extend the on-shell formalism to the supersymmetric case by introducing the Grassmann variables η^A, such that

$$Q^A_a = \lambda_a \frac{\partial}{\partial \eta_A}, \quad \tilde{Q}^A_{\dot{a}} = \tilde{\lambda}_{\dot{a}} \eta^A. \tag{2.80}$$

The super-Poincaré generators then read as

$$\ell_{ab} = \lambda_a \frac{\partial}{\partial \lambda^b}, \quad \tilde{\ell}_{\dot{a}\dot{b}} = \tilde{\lambda}_{\dot{a}} \frac{\partial}{\partial \tilde{\lambda}^{\dot{b}}} \tag{2.81}$$

$$P_{a\dot{b}} = \lambda_a \tilde{\lambda}_{\dot{b}} \tag{2.82}$$

$$Q^A_a = \lambda_a \frac{\partial}{\partial \eta_A}, \quad \tilde{Q}^A_{\dot{a}} = \tilde{\lambda}_{\dot{a}} \eta^A \tag{2.83}$$

$$R^{AB} = \eta^A \frac{\partial}{\partial \eta_B}. \tag{2.84}$$

The on-shell variables can be alternaively defined in terms of the conjugate of η_A

$$\tilde{\eta}^A = \frac{\partial}{\partial \eta_A}, \quad \frac{\partial}{\partial \tilde{\eta}_A} = \eta^A \tag{2.85}$$

In this formalism, we can group states of various helicities of a supermultiplet into a single superfield, by acting the superspace displacement operator $e^{\eta Q}$ on the top helicity state $|+\rangle$,

$$\Phi = e^{\eta Q}|+\rangle. \tag{2.86}$$

For example for an $\mathcal{N}=4$ theory, denoting the top state as Φ_{+1},

$$
\begin{aligned}
\Phi &= e^{\eta^A Q_A}\Phi^{+1} \\
&= \sum_n \frac{1}{n!}(\eta^A Q_A)^n \Phi^{+1} \\
&= \Phi^{+1} + Q_A\Phi^{+1}\eta^A + \frac{1}{2!}Q_AQ_B\Phi^{+1}\eta^A\eta^B + \frac{1}{3!}Q_AQ_BQ_C\Phi^{+1}\eta^A\eta^B\eta^C + \frac{1}{4!}Q_AQ_BQ_CQ_D\Phi^{+1}\eta^A\eta^B\eta^C\eta^D \\
&= \Phi_{+1} + \Phi_A^{+\frac{1}{2}}\eta^A + \frac{1}{2!}\Phi^0_{AB}\eta^A\eta^B + \frac{1}{3!}\Phi^{-\frac{1}{2}}_{ABC}\eta^A\eta^B\eta^C + \frac{1}{4!}\Phi^{-1}_{ABCD}\eta^A\eta^B\eta^C\eta^D
\end{aligned}
\tag{2.87}
$$

The superfield terminates at η^4 due to the antisymmetric nature of the Grassmann variables. Under the on-shell superspace formalism, amplitudes are grouped into superamplitudes which are now given in terms of the superfields: $\mathcal{A}_n[\Phi_1,\ldots,\Phi_n]$. Component amplitudes can be recovered by integrating out the desired external states. For example,

$$
A_n[\Phi_1^-,\ldots,\Phi_i^+,\ldots,\Phi_j^+,\ldots,\Phi_n^-] = \int \mathrm{d}^4\eta_1\ldots\mathrm{d}^4\eta_n\,\eta_i^4\eta_j^4\mathcal{A}_n[\Phi_1,\ldots,\Phi_n]. \tag{2.88}
$$

2.2.3. Supersymmetry Ward identities

Supersymmetry Ward identities were first studied by Grisaru, Pendleton, and van Nieuwenhuizen [25]. See [26] for a review of supersymmetry Ward identities and their applications.

In a supersymmetric model, the value of the scalar potential V at the vacuum is an order parameter of supersymmetry breaking [27][28]. $V=0$ preserves supersymmetry while $V>0$ breaks supersymmetry. Considering a supersymmetry preserving vacuum, the supercharges annihilates the vacuum state:

$$
Q^A|0\rangle = \tilde{Q}^A|0\rangle = 0. \tag{2.89}
$$

Writing an n-point superamplitude $\mathcal{A}_n[\Phi_1(p_1)\cdots\Phi_n(p_n)]$ with all outgoing particles as an S-matrix element in which the corresponding annihilation operators $a_1(p_1)\cdots a_n(p_n)$ create the particles by acting on the out-vacuum $\langle 0|$,

$$
\mathcal{A}_n[\Phi_1(p_1)\cdots\Phi_n(p_n)] = \langle 0|a_1(p_1)\cdots a_n(p_n)|0\rangle. \tag{2.90}
$$

From this follows the supersymmetry Ward identities,

$$
\begin{aligned}
\langle 0|[Q^A, a_1(p_1)\cdots a_n(p_n)]|0\rangle &= 0 \\
\langle 0|[\tilde{Q}^A, a_1(p_1)\cdots a_n(p_n)]|0\rangle &= 0,
\end{aligned}
\tag{2.91}
$$

which relates component amplitudes in a web of linear relations. The supersymmetry Ward identities are equivalent to the statement that the supercharges annihilate the superamplitude,

$$
Q^A\mathcal{A}_n = 0, \quad \tilde{Q}^A\mathcal{A}_n = 0, \tag{2.92}
$$

which is also equivalent to the statement of supermomentum conservation.

2.3. Scattering amplitudes

Scattering amplitudes are complex numbers that describe transition from an initial asymptotic free "in" momentum eigenstate to a final asymptotic free "out" momentum eigenstate, which are each characterized by the momenta and particle types. The amplitudes can also be viewed as elements of the scattering matrix, also called the "S-matrix", which relates all in and out states by mapping each initial state $|i\rangle$ to a final state $|f\rangle$ and encompasses all information of the perturbative theory. The probability of an initial state $|i\rangle$ changing into a final state $|f\rangle$ is given by $|\langle f|S|i\rangle|^2$. Separating out the trivial part of the process where no scattering occurs, we write

$$S = 1 + iT, \tag{2.93}$$

where T is called the transfer matrix. Then the amplitude for n particles is

$$\mathcal{A}_n[\Phi_1^{out}(p_1)\cdots;\cdots\Phi_n^{in}(p_n)] = \langle\Phi_1^{out}(p_1)\cdots|T|\cdots\Phi_n^{in}(p_n)\rangle. \tag{2.94}$$

Solving the S-matrix then generates all scattering amplitudes at any order in perturbation theory. The calculation of scattering amplitudes is typically performed using Feynman diagrams.

Due to crossing symmetry, we can consider all massless particles as outgoing, and not distinguish between particles and antiparticles but only note their helicities. More explicity, we can consider the amplitude with all outgoing particles as an matrix element in which the corresponding annihilation operators $a_1(p_1)\cdots a_n(p_n)$ create the particles by acting on the out-vacuum $\langle 0|$:

$$\mathcal{A}_n[\Phi_1(p_1)\cdots\Phi_n(p_n)] = \langle 0|a_1(p_1)\cdots a_n(p_n)|0\rangle \tag{2.95}$$

At tree level, amplitudes only develop poles as physical singularities, corresponding to propagators. By deforming the momenta into the complex plane we can probe the amplitude as a meromorphic function that enables us to use to powerful tools of complex analysis. For example, by using the residue theorem and Cauchy's theorem we obtain on-shell recursion relations, which will be the foundations of the methods employed in our discussions.

2.3.1. The MHV classification

Let us consider special cases of the supersymmetry Ward identities. For example in super Yang-Mills,

$$\langle 0|[\tilde{Q}^A, a_1^B a_2 \dots a_n]|0\rangle = 0 \tag{2.96}$$

gives

$$\delta^{AB} A_n[g^+ \dots g^+] = 0, \tag{2.97}$$

which says that the all-plus gluon amplitudes vanish. Similar relations hold for all-negative gluon amplitudes. Amplitudes of the form $A_n[g^- g^+ \dots g^+]$ also vanish except for the 3-point case. The first non-vanishing amplitudes are $A_n[g^- g^- g^+ \dots g^+]$, which are termed Maximally Helicity Violating (MHV) amplitudes. The next class of amplitudes are called Next-to-MHV (NMHV), and onto $\mathrm{N^k}$MHV with $k+2$ negative-helicity gluons. Similar notation and relations hold for gravity amplitudes.

2.3.2. 3-point amplitudes

3-point amplitudes are special in that they can be completely determined by little group scaling and mass dimension, without invoking the Lagrangian. We arrive at this result from *3-point special kinematics*, which states that on-shell 3-point amplitudes of all massless particles only depend on either square or angle brackets of the external momenta. The following discussion follows [4]. We start by noting that momentum conservation of the 3 particles, $p_1 + p_2 + p_3 = 0$ gives

$$p_3^2 = (p_1 + p_2)^2 = [12]\langle 12\rangle, \tag{2.98}$$

so either [12] or $\langle 12\rangle$ must vanish. Suppose $[12] = 0$ and $\langle 12\rangle \neq 0$, then we have

$$\langle 12\rangle[23] = \langle 1|p_2|3] = -\langle 1|(p_1 + p_3)|3] = 0, \tag{2.99}$$

so $[23] = 0$, and similarly $[31] = 0$. Thus all square brackets vanish and the amplitude only depends on angle brackets. An equivalent statement is that all the square spinors are proportional. Therefore, 3-point special kinematics can be stated as

$$|1] \propto |2] \propto |3] \quad \text{or} \quad |1\rangle \propto |2\rangle \propto |3\rangle. \tag{2.100}$$

Again supposing that the amplitude depends on angle brackets only, we can write an general Ansatz

$$A_3(1^{h_1} 2^{h_2} 3^{h_3}) = c\,\langle 12\rangle^{x_{12}} \langle 23\rangle^{x_{23}} \langle 31\rangle^{x_{31}} \tag{2.101}$$

for undetermined powers x_{12}, x_{23}, x_{31} and constant c. Little group scaling fixes

$$h_1 = -\frac{1}{2}(x_{12} + x_{31}), \quad h_2 = -\frac{1}{2}(x_{12} + x_{23}), \quad h_3 = -\frac{1}{2}(x_{31} + x_{23}), \tag{2.102}$$

which we can solve to find $x_{12} = h_3 - h_1 - h_2$ etc. so the amplitude is determined up to a constant,

$$A_3(1^{h_1} 2^{h_2} 3^{h_3}) = c\,\langle 12\rangle^{h_3 - h_1 - h_2} \langle 23\rangle^{h_1 - h_2 - h_3} \langle 31\rangle^{h_2 - h_3 - h_1}. \tag{2.103}$$

In deriving the preceding result, we assumed that the amplitude depended only on angle brackets. However we could have made the alternative assumption that the amplitude depended only on square brackets, which would result in

$$A_3(1^{h_1}2^{h_2}3^{h_3}) = c\,[12]^{h_1+h_2-h_3}[23]^{h_2+h_3-h_1}[31]^{h_3+h_1-h_2}. \tag{2.104}$$

To determine the correct formula, we have to consider mass dimension. Both square and angle brackets have mass dimension 1.

For example, the 3-point graviton MHV amplitude is

$$M_3[1^-2^-3^+] = \frac{\langle 12\rangle^8}{\langle 12\rangle^2\langle 23\rangle^2\langle 31\rangle^2}. \tag{2.105}$$

As we will see in the next section, massless complex-momentum 3-point amplitudes serve as building blocks for determining higher-point amplitudes recursively using on-shell recursion relations.

2.4. On-shell recursion relations

On-shell recursion relations are a powerful method that enable us to build up higher-point amplitudes recursively from lower-point amplitudes. The starting point is to study the amplitude as a function of complex on-shell momenta by deforming external momenta into the complex plane while maintaining on-shell and momentum conservation. Since tree-level amplitudes are meromorphic functions, they are completely characterized by their poles, which correspond to propagators going on-shell. The recursion relations then arise as amplitudes factorize into lower-point subamplitudes on the poles.

Utilizing on-shell recursion relations, we may systematically reduce higher-point amplitudes to basic building blocks – the 3-point amplitudes, which are fixed by little group scaling and locality. As such, we are able to construct all tree-level on-shell amplitudes of the theory without ever having to reference the Lagrangian or any off-shell information.

2.4.1. On-shell recursion relations: a general formulation

Consider a generic n-point amplitude $A_n[p_i^{h_i}]$ as a function of the momenta p_i and helicities h_i, where $i = 1 \ldots n$ labels the particle number. In the following, the dependence on h_i will be omitted unless needed to be stated explicitly. Since the amplitude is an analytic function of its momenta, we can reconstruct its functional form from the information of its analytical structure (for tree level, the poles and residues.) In practice, it is difficult to deal with a function of n momenta, instead a useful method is to introduce a single complex parameter z.

Let us introduce a deformation in the momenta

$$p_i \to \hat{p}_i = p_i + z r_i \tag{2.106}$$

depending on a complex parameter z for some r_i. By considering the amplitude as a function of the parameter z and varying it, we care able probe the behavior of the amplitude in the resulting complex plane. The deformed momenta $\hat{p}_i$ must satisfy momentum conservation $\sum_i p_i = 0$ and remain on-shell $\hat{p}_i^2 = 0$ in order for $A_n[\hat{p}_i(z)]$ to remain being an on-shell amplitude. We also impose the condition $(\sum r_i)^2 = 0$, in order for $(\sum_i \hat{p}_i)^2$ to be linear in z such that each propagator develops only one pole. In terms of the r_i, these conditions become

$$\sum_i r_i = 0 \tag{2.107}$$

$$r_i \cdot r_j = 0 \tag{2.108}$$

$$p_i \cdot r_i = 0 \tag{2.109}$$

Using the residue theorem, we can consider the original amplitude as a contour integral of the function $\frac{A_n[\hat{p}_i(z)]}{z}$ around the origin,

$$A_n[p_i] = A_n[\hat{p}_i(z=0)] = \text{Res}_{z=0} \frac{A_n[\hat{p}_i(z)]}{z} = \oint_C dz \frac{A[\hat{p}_i(z)]}{z}. \tag{2.110}$$

By Cauchy's theorem, we can consider the same integral as surrounding the poles outside of the loop around the origin. Then the amplitude become a sum over residues,

$$A_n[p_i] = -\sum_I \text{Res}_{z=z_I} \frac{A_n[\hat{p}_i(z)]}{z} + B_n, \tag{2.111}$$

with a possible boundary term B_n arising from $z \to \infty$. The validity of the recursion relation derived in the following requires the boundary term to vanish. Assuming no contribution of the boundary term, let us specialize to the case of tree-level amplitudes, then the places $z = z_I$ where $\frac{A_n[\hat{p}_i(z)]}{z}$ develop singularities are where a certain internal propagator $\hat{p}_I$ goes on shell, $\hat{p}_I^2(z = z_I) = 0$. On such a propagator pole, the deformed amplitude factorizes into two on-shell amplitudes on each side,

$$A_n^{tree}[\hat{p}_i(z)] = A_L[\hat{p}_i, i \in L] \frac{1}{\hat{p}_I^2} A_R[\hat{p}_i, i \in R] \tag{2.112}$$

The locations where poles develop are at $\hat{p}_I^2 = 0$, which we can solve for. Noting that the internal momentum is the sum of the momenta on the right hand side, $p_I = \sum_{i \in R} p_i$,

$$\begin{aligned}\hat{p}_I^2 &= (p_I + z r_I)^2 \\ &= 2\, p_I \cdot r_I\, z,\end{aligned} \tag{2.113}$$

we solve for $\hat{p}_I^2 = 0$,

$$z_I = \frac{-p_I^2}{2 p_I \cdot r_I}, \tag{2.114}$$

so

$$\frac{1}{\hat{p}_I^2} = \frac{1}{2 p_I \cdot r_I (z - z_I)} = \frac{-z_I}{p_I^2 (z - z_I)}. \tag{2.115}$$

The residue is

$$\mathrm{Res}_{z=z_I}\frac{1}{\hat{p}_I^2}=\frac{1}{2p_I\cdot r_I}=\frac{-z_I}{p_I^2}. \tag{2.116}$$

Therefore we conclude the original amplitude is

$$\begin{aligned}
A_n[p_i] &= -\sum_I \mathrm{Res}_{z=z_I}\frac{\mathcal{M}_n[\hat{p}_i(z)]}{z} \\
&= -\sum_I\sum_{h_I}\frac{1}{z}A_L[\hat{p}_i, i\in L]\frac{1}{\hat{p}_I^2}A_R[\hat{p}_i, i\in R] \\
&= \sum_I\sum_{h_I}A_L[\hat{p}_i(z_I), i\in L]\frac{1}{p_I^2}A_R[\hat{p}_i(z_I), i\in R],
\end{aligned} \tag{2.117}$$

which is given in terms of two lower-point on-shell amplitudes A_L and A_R of deformed momenta evaluated at $z=z_I$. We have to sum over all possible helicity configurations h_I corresponding to the same propagator momentum pole. Writing $A_{L/R}[\hat{p}_i(z)]=\hat{A}_{L/R}(z)$, the on-shell recursion relation is

$$A_n[p_i]=\sum_I\sum_{h_I}\hat{A}_L(z_I)\frac{1}{p_I^2}\hat{A}_R(z_I), \tag{2.118}$$

with I summed over all factorization channels. In the following, we will choose certain schemes of deforming the momenta to obtain specific recursion relations.

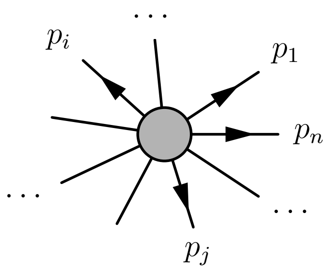
Note that the boundary term B_n arising from the pole at infinity is not readily known in general, so recursion is preferably constructed using a shift under which the boundary term vanishes. In most applications the recursion relations are justified by proving that

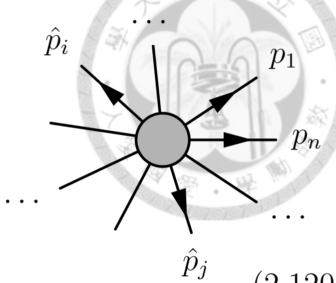
$$\hat{A}_n(z)\to 0 \quad \text{for} \quad z\to\infty \tag{2.119}$$

Such a shift is called a valid (or good) shift, while shifts without this property are called bad shifts.

2.4.2. BCFW recursion relations

The most famous of recursion relations are the Britto, Cachazo, Feng, and Witten (BCFW) recursion relations, which we will use extensively in the later sections. In this scheme, only two momenta, which we label as i and j, are selected to be shifted,

$$\mathcal{A}_n(p_1, \ldots, p_n) =$$ 

$$\longrightarrow \quad \hat{\mathcal{A}}_n(z) =$$ 

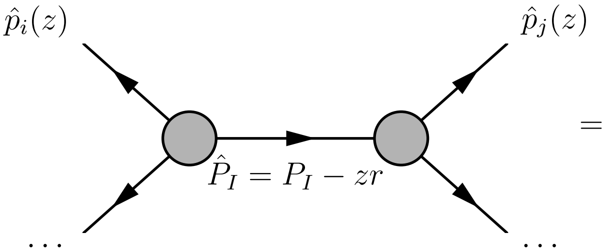
(2.120)

$$\hat{p}_i = p_i + z|j]\langle i|, \quad \hat{p}_j = p_j - z|j]\langle i|. \tag{2.121}$$

In terms of the spinors, the shift reads as

$$|\hat{i}] = |i] + z|j], \quad |\hat{j}\rangle = |j\rangle - z|i\rangle. \tag{2.122}$$

Such a shift is called the $[i, j\rangle$ shift. The brackets $[\hat{i}k]$ and $\langle\hat{j}k\rangle$ are linear in z while all other brackets are unshifted. The only propagators detectable by the BCFW shift are those which separate the shifted legs i and j onto opposite sides, since if the two shifted momenta are on the same side, the internal propagator will remain unshifted due to momentum conservation,

$$\hat{\mathcal{A}}_n(z) =$$ 

$$= \mathcal{A}_L(z)\, \frac{1}{\hat{P}_I^2(z)}\, \mathcal{A}_R(z) \tag{2.123}$$

Using BCFW recursion, we can obtain the classic Parke-Taylor formula for MHV gluon amplitudes iteratively,

$$A[1^-2^-3^+ \ldots n^+] = \frac{\langle 12\rangle^4}{\langle 12\rangle \ldots \langle n1\rangle}. \tag{2.124}$$

2.4.3. Super-BCFW recursion relations

In accordance with the on-shell superspace formalism, we can extend BCFW to include a shift in the Grassmann variable and preserve supermomentum conservation. The $[i, j\rangle$ supershift is

$$|\hat{i}] = |i] + z|j], \quad |\hat{j}\rangle = |j\rangle - z|i\rangle, \quad \hat{\eta}_i = \eta_i + z\eta_j. \tag{2.125}$$

In evaluating the BCFW recursion relations, we have to sum over all possible states that can be exchanged over the internal propagator, as in the non-supersymmetric case. Due to the on-shell superspace formalism, we can represent the sum as a Grassmann integral,

$$\mathcal{A}_n[p_i] = \sum_I \int \mathrm{d}^{\mathcal{N}}\hat{\eta}_I \, \hat{\mathcal{A}}_L(z_I)\frac{1}{p_I^2}\hat{\mathcal{A}}_R(z_I). \tag{2.126}$$

The development of the Grassmannian picture for $\mathcal{N} = 4$ SYM is then related to the realization that the terms arising from super-BCFW recursion actually manifest the Yangian symmetry of the full amplitude.

2.4.4. Large z behavior under BCFW shifts

The validity of BCFW recursion relations requires the boundary term B_n in (2.111) be absent[1]. This is typical approach is to show that the shifted amplitude vanish as $z \to \infty$,

$$\hat{A}_n(z) \to 0 \quad \text{for} \quad z \to \infty \tag{2.127}$$

The large z behavior of the amplitude, and hence the validity of BCFW recursion, depends on the helicities of the shifted legs. For pure Yang-Mills theory, an argument based on the background field method establishes the following large z behavior of color-ordered gluon tree amplitudes under shifts of adjacent gluons with indicated helicites:

$$\hat{A}_n(z) \sim \begin{array}{cccc} [-,+\rangle & [-,-\rangle & [+,+\rangle & [+,-\rangle \\ \frac{1}{z} & \frac{1}{z} & \frac{1}{z} & z^3 \end{array} \tag{2.128}$$

For non-adjacent shifted legs, an extra power $1/z$ is gained in each case. Thus the $[-,+\rangle, [-,-\rangle, [+,+\rangle$ shifts give valid recursion relations, while $[+,-\rangle$ shifts do not and is termed the bad shift as such.

For amplitudes of pure gravity, the large z behavior were shown by [31][32][33] to be

$$\hat{M}_n(z) \sim \begin{array}{cccc} [-,+\rangle & [-,-\rangle & [+,+\rangle & [+,-\rangle \\ \frac{1}{z^2} & \frac{1}{z^2} & \frac{1}{z^2} & z^6 \end{array} \tag{2.129}$$

Since gravity amplitudes are not color ordered, there is no notion of adjacency. That the large z behavior of gravity amplitudes is the square of Yang-Mills can be anticipated via the KLT relations which express graviton amplitudes as products of Yang-Mills amplitudes: $M_n \sim A_n \times A_n$.

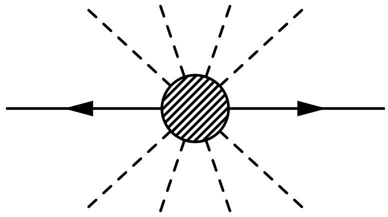
Naively, the large z behavior of YM and gravity amplitudes should have been much worse from the result of power counting looking at Feynman diagrams [34]. The better than expected large z behavior signals that YM and gravity amplitudes are actually tamed in the high energy limit. We can use BCFW shifts as a tool to probe the high energy behavior of amplitudes. The large z limit under opposite helicity shifts is especially noteworthy, since it has the physical interpretation of a hard light-like particle shooting

[1]Recently, a new algorithm known as multi-step BCFW recursion relations has been developed that can systematically deal with situations with the boundary term being present [29][30].

through a soft background [34]. This can be understood by noting that for an $[i, j\rangle$ shift, as $z \to \infty$, $\hat{p}_i$ and $\hat{p}_j$ become opposite, while other moment become irrelevant,

$$\hat{p}_i \to z|j]\langle i|, \quad \hat{p}_j \to -z|j]\langle i| \quad \text{for } z \to \infty \tag{2.130}$$

Since we have chosen to represent all momenta as outgoing, for this to be interpreted as a single particle moving through a background, the helicites of i and j must be opposite (such that the helcities are actually the same when we reverse p_j to be incoming.) This is drawn schematically below, where the dashed lines indicate other momenta, which are suppressed in the large z limit.

$\hat{\mathcal{A}}(z) \quad \longrightarrow \quad \hat{p}_i$  $\hat{p}_j \quad$ for $z \to \infty$ (2.131)

For maximally supersymmetric theories, all particles are of the same superfield, so there is no distinction between different kinds of super-BCFW shifts. Specifically, $\mathcal{N} = 4$ SYM amplitudes behave as $1/z$ and $\mathcal{N} = 8$ supergravity amplitudes as $1/z^2$. The large z behavior of component amplitudes can be recovered from the maximally supersymmetric superamplitude by integrating away some Grassmann variables.

3. Supergravity amplitudes

To lay ground for the discussion of BCFW in $\mathcal{N} = 7$ supergravity, we first review perturbative gravity and the calculation of amplitudes using on-shell methods following the discussion in [4]. We introduce supergravity in general and the maximally supersymmetric gravity theory, $\mathcal{N} = 8$ supergravity. Calculation of scattering amplitudes using BCFW recursion is discussed.

3.1. Perturbative gravity

The Lagrangian formulation of pure gravity is given in terms of the Einstein-Hilbert action,

$$S_{\text{EH}} = S_{\text{pure gravity}} + S_{\text{matter}}. \tag{3.1}$$

The pure gravitational part is

$$S_{\text{pure gravity}} = \frac{1}{2\kappa^2} \int \mathrm{d}^D x \sqrt{-g} R \tag{3.2}$$

where R is the Ricci scalar and $2\kappa^2 = 16\pi G_N$ with G_N being Newton's constant. The gravitational field is realized by the metric $g_{\mu\nu}$. The equation of motion given by the variation of S_{EH} with respect to the metric $g_{\mu\nu}$ is the Einstein equation,

$$G_{\mu\nu} = 8\pi T_{\mu\nu}, \tag{3.3}$$

where the Einstein tensor part $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$ is derived from the variation of $S_{\text{pure gravity}}$ and the variation of S_{matter} gives the stress-tensor $T_{\mu\nu} = \frac{-2}{\sqrt{-g}} \frac{\delta S_{matter}}{\delta g^{\mu\nu}}$. The matter action S_{matter} can be obtained from a theory without gravity by promoting $\eta_{\mu\nu}$ into $g_{\mu\nu}$ and $\mathrm{d}^D x$ into $\mathrm{d}^D x \sqrt{-g}$ to introduce gravitational coupling with matter. For example, the matter action for a massless free scalar and for a photon put in a gravitational background respectively are

$$S_{\text{scalar}} = \int \mathrm{d}^D x \sqrt{-g}\, \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi \tag{3.4}$$

$$S_{\text{photon}} = \int \mathrm{d}^D x \sqrt{-g}\, \left[-\frac{1}{4} g^{\mu\alpha} g^{\nu\beta} \nabla_{[\mu} A_{\nu]} \nabla_{[\alpha} A_{\beta]} \right]. \tag{3.5}$$

The property that gravity interacts with all fields can be explained by the universal appearance of $\sqrt{-g}$ as well as metric contractions upon the introduction of gravity. Our interest of discussion in the following is pure gravity, which is given by $S_{\text{matter}} = 0$.

In our discussion of gravity amplitudes, the focus is not the application of quantum field theory in curved spacetime, but rather perturbative gravity in a flat spacetime background. More precisely, we expand the gravitation field $g_{\mu\nu}$ around the flat metric, $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$ and regard the fluctuation $h_{\mu\nu}$ as the graviton field. From the view of perturbative gravity, black holes are considered non-perturbative states which are suppressed by powers e^{-1/κ^2} in the weak coupling $\kappa << 1$ scattering processes. Suppressing the intricate index structure, we can write the connection as $\Gamma \sim \kappa\partial h + \kappa^2 h\partial h$, the Ricii scalar as $R \sim (\Gamma\Gamma + \partial\Gamma)^2$, and $\sqrt{-g} \sim c + \kappa h + \kappa^2 h^2 + \cdots$, which involves an infinite series. Therefore, the expansion can be written schematically as

$$S_{\text{EH}} = \int \mathrm{d}^D x \left[h\partial^2 h + \kappa h^2\partial^2 h + \kappa^2 h^3\partial^2 h + \kappa^3 h^4\partial^2 h + \cdots\right] \tag{3.6}$$

The infinite set of interaction terms generate Feynman rules for n-graviton vertices for any $n = 3, 4, 5, \ldots$.

In order to extract Feynman rules from the Lagrangian and do explicit calculation of scattering amplitudes, we have to first choose a gauge. A common choice is the de Donder gauge, $\partial^\mu h_{\mu\nu} = \frac{1}{2}\partial_\nu h^\mu_\mu$, which results in the propagator

$$P_{\mu_1\nu_1,\mu_2\nu_2} = -\frac{i}{2}\left(\eta_{\mu_1\mu_2}\eta_{\nu_1\nu_2} + \eta_{\mu_1\nu_2}\eta_{\nu_1\mu_2} - \frac{2}{D-2}\eta_{\mu_1\nu_1}\eta_{\mu_2\nu_2}\right)\frac{1}{k^2}, \tag{3.7}$$

and the 3-point vertex

$$V_3(p_1, p_2, p_3) = p_1^{\mu_3} p_2^{\nu_3}\eta^{\mu_1\nu_2}\eta^{\mu_2\nu_1} + (\text{many other terms}), \tag{3.8}$$

and so on. Each graviton leg has two Lorentz indices, which are to be contracted by graviton polarization tensors. The polarizations can be constructed as products of spin-1 polarization vectors,

$$\epsilon_+^{\mu\nu}(p) = \epsilon_+^\mu(p)\epsilon_+^\nu(p), \quad \epsilon_-^{\mu\nu}(p) = \epsilon_-^\mu(p)\epsilon_-^\nu(p). \tag{3.9}$$

The complicated terms derived from the Lagrangian makes calculation of even the simplest tree-level amplitudes a formidable task. However, final results for on-shell amplitudes turn out to have expressions that are much more simple than the process. In accordance to the central theme of the thesis, instead of starting with the Lagrangian, we can construct gravity amplitudes much more efficiently using on-shell methods. In 4d, little group scaling and dimensional analysis fix the gravitational 3-point amplitudes to be

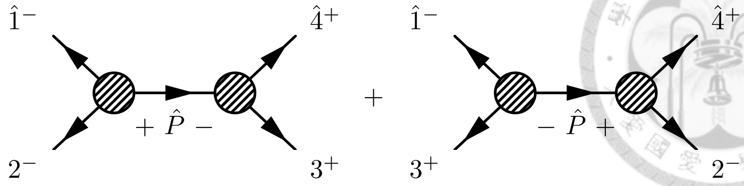
$$M_3(1^-2^-3^+) = \frac{\langle 12\rangle^6}{\langle 23\rangle^2\langle 31\rangle^2} = A_3(1^-2^-3^+)^2 \tag{3.10}$$

$$M_3(1^+2^+3^-) = \frac{[12]^6}{[23]^2[31]^2} = A_3(1^+2^+3^-)^2. \tag{3.11}$$

They serve as the building blocks from which all higher-point tree-level amplitudes can then be calculated using on-shell recursion relations. For example, we calculate the 4-

point tree-level MHV graviton amplitude using BCFW recursion by a $[1^-, 4^+\rangle$ shift,

$$M_4^{\text{tree}}(1^-2^-3^+4^+) =$$



$$
\begin{aligned}
&= M_3(\hat{1}^-2^-\hat{P}^+)\frac{1}{P^2}M_3(3^+\hat{4}^+\text{-}\hat{P}^-) + M_3(\hat{1}^-\hat{P}^-3^+)\frac{1}{P^2}M_3(\text{-}\hat{P}^+\hat{4}^+2^-) \\
&= \frac{\langle 12\rangle^6}{\langle 2\hat{P}\rangle^2\langle \hat{P}1\rangle^2}\frac{1}{\langle 12\rangle[12]}\frac{[34]^6}{[4\hat{P}]^2[\hat{P}3]^2} + \frac{\langle 1\hat{P}\rangle^6}{\langle \hat{P}3\rangle^2\langle 13\rangle^2}\frac{1}{\langle 13\rangle[13]}\frac{[\hat{P}4]^6}{[24]^2[2\hat{P}]^2} \\
&= \frac{\langle 12\rangle^5[34]^6}{[12]\langle 1|\hat{P}|3]^2\langle 2|\hat{P}|1]^4} + \frac{\langle 1|\hat{P}|4]^6}{\langle 13\rangle^3[13][24]^2\langle 3|\hat{P}|2]^2} \\
&= \frac{\langle 12\rangle^5[34]^6}{[12]\langle 12\rangle^2[23]^2\langle 23\rangle^2[34]^2} + \frac{[13]^6\langle 34\rangle^6}{\langle 13\rangle^3[13][24]^2\langle 13\rangle^2[\hat{1}2]^2} \\
&= \frac{\langle 12\rangle^3[34]^4}{[12]\langle 23\rangle^2[23]^2} + \frac{\langle 13\rangle[34]^6}{[13][24]^2[\hat{1}2]^2}.
\end{aligned}
$$

We can solve for $[\hat{1}2]$ by noting that for the diagram on the right,

$$
\begin{aligned}
&\hat{P}^2 = (\hat{p}_1 + p_3)^2 = \langle 13\rangle[\hat{1}3] = 0 \\
\Rightarrow\ &[\hat{1}3] = [13] + z[43] = 0 \\
\Rightarrow\ &z = \frac{[13]}{[34]} \\
\Rightarrow\ &|\hat{1}] = |1] + \frac{[13]}{[34]}|4] \\
\Rightarrow\ &[\hat{1}2] = \frac{[12][34] + [13][42]}{[34]} = -\frac{[14][23]}{[34]},
\end{aligned}
$$

where we have used the Schouten identity in the last line. Therefore,

$$
\begin{aligned}
M_4^{\text{tree}}(1^-2^-3^+4^+) &= \frac{\langle 12\rangle^4[34]^4}{[12]\langle 23\rangle^2[23]^2} + \frac{\langle 13\rangle[34]^8}{[13][24]^2[14]^2[23]^2} \\
&= \frac{\langle 12\rangle^4[34]^4}{[12]\langle 23\rangle^2[23]^2} + \frac{\langle 13\rangle[34]^8\langle 12\rangle^4}{[13]\langle 1|2|4]^2\langle 2|1|4]^2[23]^2} \\
&= \frac{\langle 12\rangle^4[34]^4}{[12]\langle 23\rangle^2[23]^2} + \frac{\langle 13\rangle[34]^8\langle 12\rangle^4}{[13]\langle 1|3|4]^2\langle 2|3|4]^2[23]^2} \\
&= -\frac{\langle 12\rangle^4[34]^4}{s_{14}^2}\left(\frac{1}{s_{12}} + \frac{1}{s_{13}}\right) \\
&= \frac{\langle 12\rangle^4[34]^4}{s_{12}s_{13}s_{14}},
\end{aligned}
$$

where we have used momentum conservation $p_1 + p_2 + p_3 + p_4 = 0$ in the process.

In terms of the usual Mandelstam variables, we can write the compact expression for the

4-point amplitude:

$$M_4^{\text{tree}}(1^-2^-3^+4^+) = \frac{\langle 12\rangle^4[34]^4}{stu}. \tag{3.12}$$

This derivation for the 4-point graviton amplitude is drastically simpler than brute force calculation using Feynman diagrams, which can be looked up in [35], clearing demonstrating the power of BCFW recursion. Beyond tree level, loop amplitudes can be constructed with unitary techniques. From the point of on-shell scattering amplitudes, the infinite set of interaction terms in the Lagrangian is not needed. They are only required to ensure off-shell diffeomorphism.

Let us now discuss formulas of graviton amplitudes for various helicity configurations, with designation N^kMHV by the MHV classification. Unlike gluon amplitudes, there is no color ordering, so the order of external particles is irrelevant. At tree level, graviton amplitudes vanish for all-plus and all-minus configurations,

$$M_n^{\text{tree}}(1^-2^-\dots n^-) = M_n^{\text{tree}}(1^+2^+\dots n^+) = 0, \tag{3.13}$$

as well as configurations with exactly one negative or positive helicity graviton for $n > 3$,

$$M_n^{\text{tree}}(1^+2^-3^-\dots n^-) = M_n^{\text{tree}}(1^-2^+3^+\dots n^+) = 0, \quad n > 3. \tag{3.14}$$

These relations are most easily proven using the supersymmetry Ward identities. For MHV graviton tree-level amplitudes, there are several formulas available in the literature, including the BGK (Berends, Giele, and Kuijf) formula [36],

$$M_n^{\text{tree}}(1^-2^-3^+\dots n^+) = \sum_{P(3,4,\dots,n-1)} \frac{\langle 12\rangle^8 \prod_{l=3}^{n-1}\langle n|2+3+\dots+(l-1)|l]}{\left(\prod_{i=1}^{n-2}\langle i,i+1\rangle\right)\langle 1,n-1\rangle\langle 1n\rangle^2\langle 2n\rangle^2\left(\prod_{l=3}^{n-1}\langle ln\rangle\right)}, \quad n > 4, \tag{3.15}$$

where the sum is over permutations of $(3, 4, \dots, n-1)$.

3.2. Supergravity

Supergravity is the beautiful union of gravity and supersymmetry, which can also be viewed as the result of making the supersymmetry transformations local in the sense that the supersymmetry parameter is dependent on spacetime. For a 4d supergravity theory with $\mathcal{N}$ supersymmetry generators, we can construct the spectrum by acting the generators Q^A on the positive helicity graviton h^+ as the highest-weight state to generate a supermultiplet. We also have to include the CPT conjugate states in the full spectrum.

For example, the spectrum of $\mathcal{N} = 1$ pure supergravity consists of two CPT conjugate pairs (h^+, ψ^+) and (h^-, ψ^-) each with a graviton and a gravitino,

$$\begin{array}{cccc} h=+2 & h=+\frac{3}{2} & h=-\frac{3}{2} & h=-2 \\ \text{graviton} & \text{gravitino} & \text{gravitino} & \text{graviton} \\ h^+ & \psi^+ & \psi^- & h^- \end{array} \tag{3.16}$$

The two pairs can be written as two superfields in the on-shell superspace formalism, here adopting the $\Phi - \Psi$ form, by introducing a Grassmann variable η such that $Q = \eta\lambda$,

$$\Phi = h^+ + \psi^+ \eta, \quad \Psi = \psi^- + h^- \eta. \tag{3.17}$$

For pure $\mathcal{N} = 2$ supergravity, the spectrum is

$$\begin{array}{cccccc} 1 & 2 & 1 & 1 & 2 & 1 \\ h=+2 & h=+\frac{3}{2} & h=+1 & h=-1 & h=-\frac{3}{2} & h=-2 \\ \text{graviton} & \text{gravitinos} & \text{gravi-photon} & \text{gravi-photon} & \text{gravitinos} & \text{graviton} \\ h^+ & \psi^+_A & v^+ & v^- & \psi^{-A} & h^- \end{array} \tag{3.18}$$

labeled by the SU(2) index $A = 1, 2$. On-shell superfields can be formed by introducing Grassmann variables η^A such that $Q^A = \eta^A \lambda$,

$$\Phi = h^+ + \psi^+_A \eta^A + v^+ (\eta)^2, \tag{3.19a}$$

$$\Psi = v^- + \psi^{-A} \eta_A + h^- (\eta)^2. \tag{3.19b}$$

where $(\eta)^2 = \eta^1 \eta^2$, and $\eta_A = \frac{1}{2!}\epsilon_{AB}\eta^B$ so $\eta_1 = \eta^2$.

For pure $\mathcal{N} = 4$ supergravity, the spectrum is

$$\begin{array}{ccccccccc} 1 & 4 & 6 & 4 & 1+1 & 4 & 6 & 4 & 1 \\ h=+2 & h=+\frac{3}{2} & h=+1 & h=+\frac{1}{2} & h=0 & h=-\frac{1}{2} & h=-1 & h=-\frac{3}{2} & h=-2 \\ \text{graviton} & \text{gravitinos} & \text{gravi-photons} & \text{gravi-photinos} & \text{scalars} & \text{gravi-photinos} & \text{gravi-photons} & \text{gravitinos} & \text{graviton} \\ h^+ & \psi^+_A & v^+_{AB} & \chi^{+A} & S, \tilde{S} & \chi^-_A & v_{-AB} & \psi^{-A} & h^- \end{array} \tag{3.20}$$

labeled by the SU(4) index $A = 1, 2, 3, 4$. On-shell superfields can be formed by introducing Grassmann variables η^A such that $Q^A = \eta^A \lambda$,

$$\Phi = h^+ + \psi^+_A \eta^A + \frac{1}{2!} v^+_{AB} \eta^A \eta^B + \chi^{+A} \eta^3_A + S\eta^4, \tag{3.21a}$$

$$\Psi = \tilde{S} + \chi^-_A \eta^A + \frac{1}{2!} v^-_{AB} \eta^A \eta^B + \psi^{-A} \eta^3_A + h^- \eta^4. \tag{3.21b}$$

where $\eta^3_A = \frac{1}{4!}\epsilon_{ABCD}\eta^B\eta^C\eta^D$, and $\eta^4 = \eta^1\eta^2\eta^3\eta^4$.

Since applying Q^A once lowers helicity by $\frac{1}{2}$, by acting Q^A on the graviton top state h^+, we see that if $\mathcal{N} > 8$ we cannot avoid states with spin grater than 2. Since there are no consistent interactions in flat space for particles with spin greater than 2, the theory with maximal supersymmetry in 4d is $\mathcal{N} = 8$ supergravity.

3.3. $\mathcal{N} = 8$ supergravity

The spectrum of $\mathcal{N} = 8$ supergravity consists of 128 bosons and 128 fermions,

1	8	28	56	70	56	28	8	1
$h = +2$	$h = +\frac{3}{2}$	$h = +1$	$h = +\frac{1}{2}$	$h = 0$	$h = -\frac{1}{2}$	$h = -1$	$h = -\frac{3}{2}$	$h = -2$
graviton	gravitinos	gravi-photons	gravi-photinos	scalars	gravi-photinos	gravi-photons	gravitinos	graviton
h^+	ψ_A	v_{AB}	χ_{ABC}	S_{ABCD}	χ^{ABC}	v^{AB}	ψ^A	h_-

(3.22)

We can formulate $\mathcal{N} = 8$ supergravity using an on-shell superspace by introducing eight Grassmann variables η^A, labeled by the SU(8) index $A = 1, ..., 8$. This allows us to associate the states of various helicities in the $\mathcal{N} = 8$ theory with components of different orders of η in an on-shell chiral superfield, which we write as

$$\begin{aligned}\Omega = h^+ + \psi_A\eta^A + \frac{1}{2!}v_{AB}\eta^A\eta^B + \frac{1}{3!}\chi_{ABC}\eta^A\eta^B\eta^C + \frac{1}{4!}S_{ABCD}\eta^A\eta^B\eta^C\eta^D \\ + \frac{1}{3!}\chi^{ABC}\eta^5_{ABC} + \frac{1}{2!}v^{AB}\eta^6_{AB} + \psi^A\eta^7_A + h_-\eta^8,\end{aligned} \tag{3.23}$$

where $\eta^5_{ABC} \equiv \frac{1}{5!}\epsilon_{ABCDEFGH}\eta^D\eta^E\eta^F\eta^G\eta^H$, and other η polynomials are similarly defined.

3.3.1. $\mathcal{N} = 7$ formalism of $\mathcal{N} = 8$ supergravity

The effectiveness of using the $\mathcal{N} = 7$ formalism instead of the standard $\mathcal{N} = 8$ was noted by Hodges [14]. There it was shown that the $\mathcal{N} = 7$ BCFW lead to new forms for 6- and 7-point graviton MHV amplitudes with greater symmetry. In particular, term by term $1/z^2$ fall off was noted for the 6-point amplitude. It is from these observations that we will extend to a more general amplitudes in this thesis. In the following we review the derivation of $\mathcal{N} = 7$ supergravity amplitudes from its $\mathcal{N} = 8$ counterpart, as well as its large z behavior. This discussion follows [37].

Following the $\Phi - \Psi$ formalism introduced by [37], we reduce the manifest supersymmetry from $\mathcal{N} = 8$ to $\mathcal{N} = 7$, such that the on-shell states separate into two superfields, which are obtained respectively from two different ways of reducing supersymmetry: setting η^8 to zero or integrating away η^8.

$$\Phi \equiv \Omega|_{\eta^8 \to 0} = \int d\eta^8 \; \eta^8 \, \Omega \,, \tag{3.24a}$$

$$\Psi \equiv \int d\eta^8 \; \Omega \,. \tag{3.24b}$$

The explicit forms of the superfields are:

$$\begin{aligned}\Phi &= h^+ + \psi_A\eta^A + \frac{1}{2!}v_{AB}\eta^A\eta^B + \frac{1}{3!}\chi_{ABC}\eta^A\eta^B\eta^C + \frac{1}{3!}S^{8ABC}\eta^4_{ABC} \\ &\quad + \frac{1}{2!}\chi^{8AB}\eta^5_{AB} + v^{8A}\eta^6_A + \psi^8\eta^7, \qquad (3.25a)\\ \Psi &= \psi_8 + v_{8A}\eta^A + \frac{1}{2!}\chi_{8AB}\eta^A\eta^B + \frac{1}{3!}S_{8ABC}\eta^A\eta^B\eta^C + \frac{1}{3!}\chi^{ABC}\eta^4_{ABC} \\ &\quad + \frac{1}{2!}v^{AB}\eta^5_{AB} + \psi^A\eta^6_A + h^-\eta^7. \qquad (3.25b)\end{aligned}$$

The indices are now summed from 1 to 7, and $\eta^4_{ABC} \equiv \frac{1}{4!}\epsilon_{ABCDEFG}\eta^D\eta^E\eta^F\eta^G$. Note that setting η^8 to zero can be represented by an integration over η^8 after multiplying by η^8. The Φ multiplet has helicity $+2$, and contains the positive helicity graviton h^+, while Ψ has helicity $+3/2$, and contains the negative helicity graviton h^-. Although the superfields contain more than the positive and negative helicity gravitons, we will continue to use the "+" and the "−" sign to mark quantities associated with the Φ and Ψ multiplet respectively, as in the non-supersymmetric case.

Using the same operations, $\mathcal{N} = 7$ amplitudes can be derived from the corresponding $\mathcal{N} = 8$ amplitudes. As an example, the $\mathcal{N} = 7$ MHV 3-point graviton scattering amplitude is obtained from the $\mathcal{N} = 8$ MHV 3-point amplitude as follows:

$$\mathcal{M}_3^{\mathcal{N}=7}[1^-2^-3^+] = \int \mathrm{d}\eta_1^8\mathrm{d}\eta_2^8\mathrm{d}\eta_3^8\ \eta_3^8\,\mathcal{M}_3^{\mathrm{MHV}}(123). \qquad (3.26)$$

Here the subscript of η refers to the associated particle number, while the superscript refers to the SU(8) index.

For a general $\mathrm{N^kMHV}$ amplitude, there will be $k+2$ external legs in the Ψ multiplet, which we denote by the set $\{x\}$, and $n-k-2$ external legs in the Φ multiplet, which we denote by the set $\{y\}$. Then we have the following map between $\mathcal{N} = 7$ and $\mathcal{N} = 8$ amplitudes:

$$\mathcal{M}^{\mathcal{N}=7}[\{x\},\{y\}] = \int \left[\prod_{a=1}^{n}\mathrm{d}\eta_a^8\right]\left[\prod_{b\in\{y\}}\eta_b^8\right]\mathcal{M}^{\mathcal{N}=8}. \qquad (3.27)$$

Or more explicitly,

$$\mathcal{M}_n^{\mathcal{N}=7}[1^-,\cdots,(k+2)^-,(k+3)^+,\cdots,n^+] = \int \mathrm{d}\eta_1^8\cdots\mathrm{d}\eta_n^8\ \eta_{k+3}^8\cdots\eta_n^8\,\mathcal{M}_n^{\mathcal{N}=8}(1,\cdots,n). \qquad (3.28)$$

In the following we will introduce super-BCFW recursion in the formalism of $\mathcal{N} = 7$ supergravity, and examine its validity under different scenarios, leading us to investigate the large z behavior of the $[+,-\rangle$ "bad shift" representation. We then present a proof for term-by-term $\mathcal{O}(z^{-2})$ scaling of the "bad shift" representation under a correspondingly chosen test shift. Furthermore, we discover the improved scaling in $\mathcal{N} = 7$ is related to bonus relations in $\mathcal{N} = 8$.

3.3.2. $\mathcal{N}=7$ BCFW recursion

Validity of a BCFW representation requires the amplitude vanish as the deformation parameter z goes to infinity:

$$|\hat{i}] = |i] + z|j], \quad |\hat{j}\rangle = |j\rangle + z|i\rangle, \quad \hat{\eta}_i = \eta_i + z\,\eta_j,$$
$$\hat{\mathcal{M}}_n(z) \longrightarrow 0 \quad \text{as} \quad z \longrightarrow \infty\,. \tag{3.29}$$

$\mathcal{N}=8$ amplitudes scale as $\mathcal{O}\left(z^{-2}\right)$ for large z. In the case of $\mathcal{N}=7$, we can deduce the large z behavior by relating the $\mathcal{N}=7$ amplitude to the parent $\mathcal{N}=8$ using (3.27). Unlike in the case of $\mathcal{N}=8$, amplitudes in $\mathcal{N}=7$ specialize into different supermultiplet configurations for lines i,j which may show different large z behavior.

Note that in order to deduce the large z behavior of $\mathcal{N}=7$ from $\mathcal{N}=8$ using (3.27), we need to take into the subtlety that for $\mathcal{N}=8$, we shift $\hat{\eta}_i^A$ for $A=1...8$, while for $\mathcal{N}=7$, we only shift for $A=1...7$. Thus we need to somehow unshift $\hat{\eta}_i^8$. This can easily be done by a change of variables. We define

$$\eta_i^8 = \overline{\eta}_i^8 - z\overline{\eta}_j^8, \quad \eta_a^8 = \overline{\eta}_a^8 \quad \text{for } a \neq i. \tag{3.30}$$

The Jacobian is simply 1. Now we can promote (3.27) into a relation for the shifted variables:

$$\mathcal{M}^{\mathcal{N}=7}(z) = \int \left[\prod_{a=1}^{n} \mathrm{d}\overline{\eta}_a^8\right] \left[\prod_{b\in\{y\}} \eta_b^8(\overline{\eta}_c^8)\right] \mathcal{M}^{\mathcal{N}=8}(z)\,, \tag{3.31}$$

where η_b^8 is a function of $\overline{\eta}_c^8$, as defined by (3.30).

We can now analyze different scenarios for which multiplet the lines i,j in our $[i,j\rangle$ shift sits in:

- For $[i^-,j^+\rangle$ and $[i^-,j^-\rangle$: Since i is not in the Φ multiplet, η_b^8 does not contain any z dependence, and hence the $\mathcal{N}=7$ amplitude behaves as $\mathcal{O}\left(z^{-2}\right)$ at large z exactly like $\mathcal{N}=8$.

- For $[i^+,j^+\rangle$: Now i belongs to the Φ multiplet, so naively applying a change of variable, one would pick up a z factor. However the z will be proportional to $\overline{\eta}_j$ which is already present in η_b^8 and thus this term drops out, i.e. $(\overline{\eta}_i - z\overline{\eta}_j)\overline{\eta}_j = \overline{\eta}_i\overline{\eta}_j$. Thus we see for this shift, the $\mathcal{N}=7$ amplitude again behaves as $\mathcal{O}\left(z^{-2}\right)$ at large z exactly like $\mathcal{N}=8$.

- For $[i^+,j^-\rangle$: Now i belongs to the Φ multiplet, while j does not, so η_b^8 obtains an overall factor of z. Thus the large z behavior for $\mathcal{N}=7$ amplitude behaves as $\mathcal{O}\left(z\right) \times \mathcal{O}\left(z^{-2}\right) = \mathcal{O}\left(z^{-1}\right)$.

From the above we conclude that for the "good" shifts $[i^-,j^+\rangle$, $[i^-,j^-\rangle$, $[i^+,j^+\rangle$, the $\mathcal{N}=7$ amplitude behaves as $1/z^2$ just as the $\mathcal{N}=8$ parent. The BCFW built for $\mathcal{N}=7$ from the good shifts will be using the same $1/z$ pole as the $\mathcal{N}=8$ parent. Thus the BCFW built from the $[+,-\rangle$ "bad" shift in $\mathcal{N}=7$ is secretly using information of the

$1/z^2$ behavior of the $\mathcal{N} = 8$ amplitude. In the following section, we will demonstrate that the $\mathcal{N} = 7$ BCFW expansion built from the $[j^+, i^-\rangle$ "bad shift" indeed has bonus behavior in the form of term-by-term $\mathcal{O}\left(z^{-2}\right)$ large-z scaling under the $[i^-, j^+\rangle$ test shift.

3.4. $\mathcal{N} = 7$ "bad shift" BCFW representation

3.4.1. A particular $[-, +\rangle$ test shift: $\mathrm{N^kMHV}$ amplitudes

We would like to prove that the $\mathcal{N} = 7$ $[j^+, i^-\rangle$ "bad shift" BCFW terms have $\mathcal{O}\left(z^{-2}\right)$ large z fall off under the secondary $[i^-, j^+\rangle$ test shift. Note our analysis can be easily applied to other helicity configurations as well, where the $\mathcal{O}\left(z^{-2}\right)$ fall off is no longer present. Therefore, we start without fixing which superfields particles i and j belong to and construct the $[j, i\rangle$ BCFW representation of the amplitude (see Fig. 3.1):

$$\mathcal{M}_n(1, \cdots, i, \cdots, j, \cdots, n) = \sum \int \mathrm{d}^7\eta_{\hat{P}}\ \mathcal{M}_L(-\hat{P}, \hat{j}, \cdots)\, \frac{1}{P^2}\, \mathcal{M}_R(\hat{P}, \hat{\imath}, \cdots)\ |_{\hat{P}^2=0}\,, \tag{3.32}$$

$$|\hat{j}] = |j] + w|i], \quad |\hat{i}\rangle = |i\rangle - w|j\rangle, \quad \hat{\eta}_j = \eta_j + w\,\eta_i\,. \tag{3.33}$$

We use w to denote the deformation parameter of the BCFW shift used to construct the representation, and z for the test shift used to probe high energy behavior.

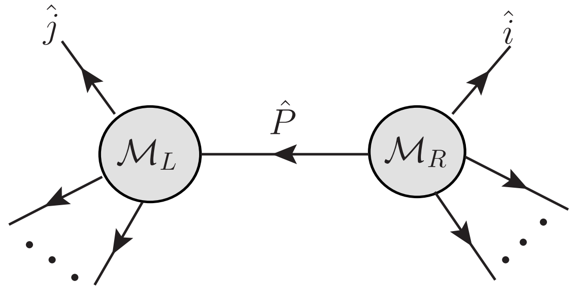


Figure 3.1.: Diagram of a BCFW term.

For the on-shell condition $\hat{P}^2 = (P + w\,|i]\langle j|)^2 = 0$, we can solve for w and $\hat{P}$ in terms of i, j and P.

Derivation of $\hat{P}$

Consider a $[j, i\rangle$ BCFW representation:

$$\mathcal{M}_n = \sum \int \mathrm{d}^7\eta_{\hat{P}}\ \mathcal{M}_L \frac{1}{P^2} \mathcal{M}_R\ |_{\hat{P}^2=m^2}\,, \tag{3.34}$$

$$|\hat{j}] = |j] + w|i], \quad |\hat{i}\rangle = |i\rangle - w|j\rangle, \quad \hat{\eta}_j = \eta_j + w\,\eta_i\,, \tag{3.35}$$

$$\hat{P} = P + w\,|i]\langle j|. \tag{3.36}$$

We can evaluate w using the on-shell condition $\hat{P}^2 = m^2$.

$$\begin{aligned}\hat{P}^2 &= (P + w\,|i]\langle j|)^2 \\ &= P^2 + 2P \cdot w\,|i]\langle j| \\ &= P^2 + w\,\langle j|P|i] = m^2.\end{aligned}$$

Therefore,

$$\boxed{w = \frac{-P^2 + m^2}{\langle j|P|i]}.} \tag{3.37}$$

Plugging the expression for w into $\hat{P}$,

$$\begin{aligned}\hat{P} &= P + \frac{(-P^2 + m^2)}{\langle j|P|i]}|i]\langle j| \\ &= \frac{[i|P|j\rangle P - P^2|i]\langle j| + m^2|i]\langle j|}{\langle j|P|i]}\,.\end{aligned}$$

This can be simplified by invoking the Schouten identity as follows:

$$\begin{aligned}\langle j|P|i]P_{a\dot{b}} &= j_{\dot{c}}P^{\dot{c}d}i_d P_{a\dot{b}} \\ &= -P_{\dot{c}}{}^{d} i_d P_a{}^{\dot{c}} j_{\dot{b}} - P_{a\dot{c}} j^{\dot{c}} P_{\dot{b}}{}^{d} i_d \\ &= P_{a\dot{c}}P^{\dot{c}d} i_d j_{\dot{b}} + P_{a\dot{c}} j^{\dot{c}} P_{\dot{b}d} i^d.\end{aligned}$$

Using $P^a_{\dot{c}}P^{\dot{c}d} = P^2\epsilon^{ad}$, we have

$$\begin{aligned}\langle j|P|i]P &= P^2\delta_a{}^d i_d j_{\dot{b}} + P_{a\dot{c}} j^{\dot{c}} P_{\dot{b}d} i^d \\ &= P^2 i_a j_{\dot{b}} + P_{a\dot{c}} j^{\dot{c}} P_{\dot{b}d} i^d \\ &= P^2|i]\langle j| + P|j\rangle[i|P.\end{aligned}$$

We obtain for $\hat{P}$:

$$\boxed{\hat{P} = \frac{P|j\rangle[i|P + m^2|i]\langle j|}{\langle j|P|i]}.} \tag{3.38}$$

The result for the massless case is

$$w = -\frac{P^2}{\langle j|P|i]}\,, \tag{3.39}$$

$$\hat{P} = \frac{P|j\rangle[i|P}{\langle j|P|i]}\,. \tag{3.40}$$

Let us now deform (3.32) by an $[i, j\rangle$ test shift:

$$|i](z) = |i] + z|j], \quad |j\rangle(z) = |j\rangle - z|i\rangle, \quad \eta_i(z) = \eta_i + z\,\eta_j. \tag{3.41}$$

Under the test shift, the amplitude is deformed into

$$\mathcal{M}_n(z) = \sum \int d^7\eta_{\hat{P}}\; \mathcal{M}_L(-\hat{P}(z), \hat{j}(z), \cdots)\,\frac{1}{P^2(z)}\,\mathcal{M}_R(\hat{P}(z), \hat{i}(z), \cdots) \tag{3.42}$$

Now $|i], |j\rangle, \eta_i, P^2, |\hat{j}], |\hat{i}\rangle, \hat{\eta}_j, |\hat{P}], |\hat{P}\rangle$ have become functions of z. Since the BCFW terms must have zero little group weight in $\hat{P}$, the z dependence of the BCFW terms only comes from $|i], |j\rangle, \eta_i, P^2, |\hat{j}], |\hat{i}\rangle, \hat{\eta}_j, \hat{P}$. By analyzing their large z behavior individually, we can deduce the large z behavior of the BCFW term as a whole. We thus proceed to do so.

From the $[i, j\rangle$ test shift (3.41), deriving the large-z behavior of $|i], |j\rangle, \eta_i, P^2$ is straightforward:

$$|i](z) \longrightarrow \mathcal{O}(z), \quad |j\rangle(z) \longrightarrow \mathcal{O}(z), \quad \eta_i(z) \longrightarrow \mathcal{O}(z). \tag{3.43}$$

$$P^2(z) = P^2 - z\langle i|P|j] \longrightarrow \mathcal{O}(z). \tag{3.44}$$

The primary deformed quantities $|\hat{j}], |\hat{i}\rangle, \hat{\eta}_j, \hat{P}$ transform under the test shift as

$$|\hat{j}](z) = |j] + w(z)|i](z), \quad |\hat{i}\rangle(z) = |i\rangle - w(z)|j\rangle(z), \quad \hat{\eta}_j(z) = \eta_j + w(z)\eta_i(z), \tag{3.45}$$

$$\hat{P}(z) = \frac{-(P - p_j)|i\rangle[j|(P + p_i)}{\langle i|P|j]} + \mathcal{O}\left(z^{-1}\right) \longrightarrow \mathcal{O}\left(z^{0}\right). \tag{3.46}$$

To determine the large-z behavior of $|\hat{j}], |\hat{i}\rangle, \hat{\eta}_j$, we solve for the z-deformed primary shift parameter $w(z)$, and expand it in powers of z:

$$w(z) = -\frac{1}{z} + \frac{-P^2 - \langle j|P|j] + \langle i|(P - p_j)|i]}{\langle i|P|j]}\frac{1}{z^2} + \mathcal{O}\left(z^{-3}\right). \tag{3.47}$$

We expand to $\mathcal{O}\left(z^{-2}\right)$ since the leading term gets canceled when we plug in expressions (3.41) and (3.47) into (3.45). We get:

$$\begin{aligned}
|\hat{j}](z) &= \left(-|i] + \frac{-P^2 - \langle j|P|j] + \langle i|(P - p_j)|i]}{\langle i|P|j]}|j]\right)\frac{1}{z} + \mathcal{O}\left(z^{-2}\right), \\
|\hat{i}\rangle(z) &= \left(|j\rangle + \frac{-P^2 - \langle j|P|j] + \langle i|(P - p_j)|i]}{\langle i|P|j]}|i\rangle\right)\frac{1}{z} + \mathcal{O}\left(z^{-2}\right), \\
\hat{\eta}_j(z) &= \left(-\eta_i + \frac{-P^2 - \langle j|P|j] + \langle i|(P - p_j)|i]}{\langle i|P|j]}\eta_j\right)\frac{1}{z} + \mathcal{O}\left(z^{-2}\right).
\end{aligned} \tag{3.48}$$

Now we can read off their large-z behavior. The results are organized below:

$$\boxed{\begin{gathered}
|i](z) \longrightarrow \mathcal{O}(z), \quad |j\rangle(z) \longrightarrow \mathcal{O}(z), \quad \eta_i(z) \longrightarrow \mathcal{O}(z), \\
|\hat{j}](z) \longrightarrow \mathcal{O}\left(z^{-1}\right), \quad |\hat{i}\rangle(z) \longrightarrow \mathcal{O}\left(z^{-1}\right), \quad \hat{\eta}_j(z) \longrightarrow \mathcal{O}\left(z^{-1}\right), \\
\hat{P}(z) \longrightarrow \mathcal{O}\left(z^{0}\right), \\
P^2(z) \longrightarrow \mathcal{O}(z).
\end{gathered}} \tag{3.49}$$

With the large-z scaling of $|i], |j\rangle, \eta_i, P^2, |\hat{j}], |\hat{i}\rangle, \hat{\eta}_j, \hat{P}$ in hand, we can know how the BCFW term behaves at large z by counting the orders of these contributing components. From (3.49) we see that $|i], \eta_i$, which have helicity 1/2, behave as $\mathcal{O}(z)$. On the other hand, $|\hat{i}\rangle$, which has helicity $-1/2$, scales oppositely as $\mathcal{O}(z^{-1})$. We can write a general Ansatz that if particle i contributes to the amplitude in the form of $|i]^a \eta_i^b |\hat{i}\rangle^c$, then it scales as $\mathcal{O}\left(z^{a+b-c}\right)$.

In general, determining the orders of the spinors and the Grassmann variable can be nontrivial. However, in this case little group scaling of external leg i trivializes the counting by fixing $a+b-c=2h_i$, where h_i and h_j are the helicities of the superfield corresponding to legs i and j. Therefore, particle i contributes $\mathcal{O}\left(z^{2h_i}\right)$ at large z. A similar analysis shows that particle j contributes $\mathcal{O}\left(z^{-2h_j}\right)$ at large z. Since $\hat{P}$ approaches a constant at $z \to \infty$, the large z scaling of each BCFW term is of:

$$\boxed{\mathcal{O}\left(z^{2(h_i-h_j)-1}\right).} \tag{3.50}$$

Crucial to this result is the choice of the $[j,i\rangle$ primary shift followed by $[i,j\rangle$ test shift, which enjoys the cancellation of order z^0 terms while obtaining (3.48) and thus ensures that the square spinors and the Grassmann variable scale oppositely to the angle spinors. Other choices would not have allowed us to determine the large z scaling from the helicities alone. For example, if we chose a $[j,i\rangle$ primary shift followed by a $[k,j\rangle$ test shift, where $i \neq k$, then $|k]$ and η_k would scale as $\mathcal{O}(z)$ while $|k\rangle$ scale as $\mathcal{O}(z^0)$. If particle k contributes to the amplitude in the form of $|k]^a \eta_k^b |k\rangle^c$, then it would scale as $\mathcal{O}\left(z^{a+b}\right)$, so $a+b-c=2h_k$ would not be sufficient to determine the large z scaling contributed by particle k.

Note that up until this point we have not designated the helicities of superfields i and j. If we choose a $[j^+, i^-\rangle$ "bad" $\mathcal{N}$ supershift for supergravity, h_j and h_i would be separated by $\frac{8-\mathcal{N}}{2}$, such that the large z scaling of each BCFW term be:

$$\boxed{\mathcal{O}\left(z^{\mathcal{N}-9}\right).} \tag{3.51}$$

We now specialize to the $\mathcal{N}=7$ $[j^+, i^-\rangle$ "bad shift" BCFW expansion under the secondary $[i^-, j^+\rangle$ test shift. From the expressions for the $\mathcal{N}=7$ superfields (3.25), superfield i has helicity $+3/2$ and therefore contributes $\mathcal{O}(z^3)$ at large z, while superfield j has helicity $+2$ and gives us $\mathcal{O}(z^{-4})$. $1/P^2$ gives $\mathcal{O}(z^{-1})$. Collectively, we find that the large z scaling for the BCFW term is of:

$$\boxed{\mathcal{O}\left(z^{-2}\right).} \tag{3.52}$$

We are lead to this result only if we specialize to the case where the $[j^+, i^-\rangle$ bad shift is the primary shift. Other choices can result in $\mathcal{O}(z^{-1})$ or worse fall off. However, note that our counting is only indicative of the worst behavior, so the terms can actually have better fall off than shown by the counting. For example, both $\mathcal{N}=7$ $[j^+, i^+\rangle$ and $[j^-, i^-\rangle$ count to $\mathcal{O}(z^{-1})$, but explicit calculations have shown that some but not all of their BCFW terms behave as $\mathcal{O}(z^{-2})$.

Finally, note that the place where $\mathcal{N}=7$ plays a crucial role is the fact that the bad shift BCFW recursion is not valid for $\mathcal{N}<7$, while $\mathcal{N}=8$ does not distinguish between different shifts.

3.4.2. Large z behavior under general test shifts

Our proof above for the $\mathcal{N}=7$ "bad shift" BCFW representation is restricted to particular chosen test shift legs that match the ones used to construct the BCFW representation. In the following we examine its large z behavior under more general test shifts.

$[-,+\rangle$ test shifts: the MHV case

The above result fails for general BCFW test shifts other than the $[i^-, j^+\rangle$ shift, and an alternative analysis is required. In general, there are many combinations of test shifts that we can choose from, however we are mainly concerned with the $[-,+\rangle$ test shift, since it is the most relevant in the high energy limit. In the following, we analyze the large z scaling under general $[-,+\rangle$ test shifts in the MHV case. (See Fig. 3.2)

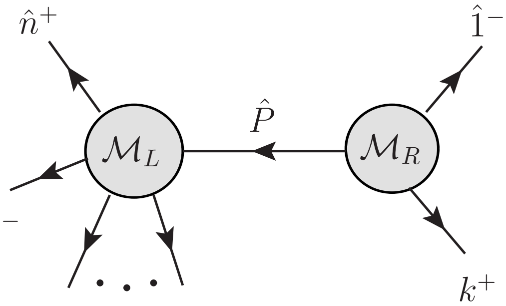


Figure 3.2.: Diagram of a MHV "bad shift" BCFW term.

Choosing the $[n^+, 1^-\rangle$ primary shift, the amplitude factorizes into an $n-1$ point MHV subamplitude and a 3-point $\overline{\text{MHV}}$ subamplitude. Similar to our previous analysis, first we solve for w and $\hat{P}$:

$$w = \frac{\langle 1k\rangle}{\langle nk\rangle}\,, \tag{3.53}$$

$$\hat{P} = -\left(|k] + \frac{\langle n1\rangle}{\langle nk\rangle}|1]\right)\langle k|. \tag{3.54}$$

We now analyze the large z scaling under different $[-,+\rangle$ test shifts:

- For the $[1^-, n^+\rangle$ shift: The proof in the previous section applies, and there is $\mathcal{O}\left(z^{-2}\right)$ term by term behavior.
- For the $[2^-, n^+\rangle$ shift: There is $\mathcal{O}\left(z^{-2}\right)$ term by term behavior. The large z behavior of the deformed quantities are:

$$\begin{aligned}
\hat{P} &\longrightarrow \mathcal{O}\left(z^0\right)\\
|2](z) &= |2] + z|n]\\
|n\rangle(z) &= |n\rangle - z|2\rangle\\
|\hat{n}](z) &= |n] + w|1] \longrightarrow |n]\\
|\hat{1}\rangle(z) &= |1\rangle - w(|n\rangle - z|2\rangle) \longrightarrow \mathcal{O}\left(z^0\right).
\end{aligned} \tag{3.55}$$

In the large z limit, dependence on z only comes from the $n-1$ point subamplitude $\mathcal{M}_L$, also we see that $|\hat{n}] \to |n]$. Therefore, the chosen test shift is precisely a BCFW shift on the subamplitude $\mathcal{M}_L$ at large z, so the BCFW term must scale as $\mathcal{O}\left(z^{-2}\right)$.

- For a $[2^-, m^+\rangle$ shift (where $m \neq n$): Individual terms scale as $\mathcal{O}\left(z^{-2}\right)$. The same argument as above applies if m is not on the 3 point amplitude, so terms scale as $\mathcal{O}\left(1/z^2\right)$. Moreover, the BCFW expansion is summed over all possible permutations, but there is only one diagram where m is on the 3 point amplitude, therefore this term must also scale as $\mathcal{O}\left(z^{-2}\right)$, since the existence of an $\mathcal{O}\left(z^{-1}\right)$ part cannot be canceled by other terms.

- For a $[1^-, m^+\rangle$ test shift: The above argument fails and there are terms which do not behave as $\mathcal{O}\left(z^{-2}\right)$.

Summarizing the results above, we have demonstrated that for the MHV case, the $\mathcal{N}=7$ $[n^+, 1^-\rangle$ bad shift BCFW representation has $\mathcal{O}\left(z^{-2}\right)$ term by term large z scaling under $[1^-, n^+\rangle$, $[2^-, n^+\rangle$ and $[2^-, m^+\rangle$ test shifts.

Explicit example: 5-point $\mathcal{N}=7$ "bad shift" representation

The $[5^+, 1^-\rangle$ BCFW representation of the $\mathcal{N}=7$ 5-point MHV amplitude $\mathcal{M}_5(12\tilde{3}\tilde{4}\tilde{5})$ has just two terms and is given by

$$\mathcal{M}_5(12\tilde{3}\tilde{4}\tilde{5}) = \frac{\delta^{(14)}(\widetilde{Q})[13][24]}{\langle 13\rangle\langle 15\rangle\langle 23\rangle\langle 25\rangle\langle 34\rangle\langle 35\rangle\langle 45\rangle} + (3 \leftrightarrow 4) \tag{3.56}$$

Note that the two terms can be related by either $1 \leftrightarrow 2$ or $3 \leftrightarrow 4$. The few number of terms and the symmetry between terms causes the two terms to be $\mathcal{O}(1/z^2)$ under many kinds of test shifts, which may not generalize to higher-point amplitudes.

- $[-,-\rangle$ test shifts

  Under a $[1^-, 2^-\rangle$ test shift, the two BCFW terms behave at large z respectively as (with some numerical factors ommited)

  $$-\frac{1}{z}[23][24]\langle 13\rangle\langle 14\rangle + \mathcal{O}\left(\frac{1}{z^2}\right)$$

  and

  $$\frac{1}{z}[23][24]\langle 13\rangle\langle 14\rangle + \mathcal{O}\left(\frac{1}{z^2}\right)$$

  Summing the two terms, the $\mathcal{O}(1/z)$ parts are cancelled, leaving only the $\mathcal{O}(1/z^2)$ part. This kind of cancellation is not desirable for our purposes. The $[2^-, 1^-\rangle$ test shift shows the same behavior due to $1 \leftrightarrow 2$ symmetry.

- $[+,+\rangle$ test shifts

  Under the $[+,+\rangle$ test shifts with one leg being the original shifted leg, i.e the $[3^+,5^+\rangle$, $[4^+,5^+\rangle$, $[5^+,3^+\rangle$, $[5^+,4^+\rangle$ test shifts, the individual BCFW terms fall off as $\mathcal{O}(1/z^2)$ for large z.

  Under all other $[+,+\rangle$ test shifts, i.e the $[3^+,4^+\rangle$, $[4^+,3^+\rangle$ test shifts, the individual BCFW terms $\mathcal{O}(1/z)$ for large z and the summation shows cancellation between the $\mathcal{O}(1/z)$ terms.

- $[-,+\rangle$ test shifts

  Under all the $[-,+\rangle$ test shifts, the individual BCFW terms behave as $\mathcal{O}(1/z^2)$ for large z.

Explicit example: 6-point $\mathcal{N}=7$ "bad shift" representation

The BCFW representation of the 6-point MHV amplitude $\mathcal{M}_6(12\tilde{3}45\tilde{6})$ constructed from the $[6^+,1^-\rangle$ shift is:

$$\mathcal{M}_6(12\tilde{3}45\tilde{6}) = \delta^{(14)}(\widetilde{Q})\left(\frac{[13][25](\langle 62\rangle[24]+\langle 65\rangle[54])}{\langle 62\rangle\langle 56\rangle\langle 13\rangle\langle 54\rangle\langle 16\rangle\langle 36\rangle\langle 43\rangle\langle 25\rangle\langle 46\rangle} + \frac{[13][24](\langle 62\rangle[25]+\langle 64\rangle[45])}{\langle 62\rangle\langle 56\rangle\langle 13\rangle\langle 53\rangle\langle 16\rangle\langle 36\rangle\langle 54\rangle\langle 24\rangle\langle 64\rangle}\right) \tag{3.57}$$

Under all $[-,+\rangle$ test shifts, the individual BCFW terms behave as $\mathcal{O}(1/z^2)$ for large z.

3.4.3. Comparison to other formulas for supergravity amplitudes

The large z scaling of the "bad shift" BCFW representation can be compared with the tree formula for MHV amplitudes by Nguyen, Spradlin, Volovich, and Wen [13], which also manifest $\mathcal{O}\left(z^{-2}\right)$ large z fall off term-by-term under certain test shifts. The formula chooses two legs as special, and involves a sum of terms each represented by a tree diagram. By directly counting the orders of z in the z deformed formula, we see that if at least one of test shift legs are special, then the term will scale as $\mathcal{O}\left(z^{-2}\right)$. Otherwise, for an $[i,j\rangle$ test shift where neither i or j is a special leg, the term scales as $\mathcal{O}\left(z^{\deg(i)+\deg(j)-4}\right)$. The degree of a leg refers to the number of propagators that connect to the leg in the tree diagram. The best fall off occurs when both leg i and j have only one connection, where the term scales as $\mathcal{O}\left(z^{-2}\right)$. The tree formula and the $\mathcal{N}=7$ BCFW is complementary in the sense that both manifest the $\mathcal{O}\left(z^{-2}\right)$ scaling term by term, but under different conditions of test shift legs.

3.4.4. $\mathcal{N}=8$ bonus relations and $\mathcal{N}=7$ bonus scaling: the MHV case

After demonstrating our proof, we would like to show that $\mathcal{N}=7$ BCFW terms manifest the improved scaling because they are using "bonus relations", which come from the $\mathcal{O}(z^{-2})$ fall off of $\mathcal{N}=8$ amplitudes. The bonus scaling of $\mathcal{N}=8$ amplitudes enables us to multiply a linear function of z on our amplitude and deform z as in BCFW recursion, except that we do not have to consider the boundary integral. These extra relations are called "bonus relations". Multiplying by the s channel, we have the sum over residues at $z=z_k$,

$$s(0)\,\mathcal{M}_n^{\mathcal{N}=8}=\sum_k s(z_k)\int \mathrm{d}^8\eta_{\hat{P}}\,\mathcal{M}_L\,\frac{1}{P^2}\,\mathcal{M}_R. \tag{3.58}$$

Our purpose is to use the bonus relations to recombine $\mathcal{N}=8$ terms and cancel out linear relations between terms, such that the remaining expression corresponds to the $\mathcal{N}=7$ representation. The following analysis focuses on the MHV case for simplicity and parallels Appendix C of [38]. Note that the BCFW representation for the $\mathcal{N}=8$ n-point MHV amplitude will always have one more diagram than $\mathcal{N}=7$. We will show that we can use the bonus relation to express the additional $\mathcal{N}=8$ term using terms appearing in $\mathcal{N}=7$. More explicitly, we write the $\mathcal{N}=8$ n-point MHV amplitude as $\mathcal{M}(123\cdots n)$ or $\mathcal{M}_n^{\mathcal{N}=8}$, the $\mathcal{N}=7$ amplitude as $\mathcal{M}(1^-2^-3^+\cdots n^+)$ or $\mathcal{M}_n^{\mathcal{N}=7}$, and construct the BCFW representation using the $[n^+,1^-\rangle$ shift:

$$|\hat{n}]=|n]+w|1],\quad |\hat{1}\rangle=|1\rangle-w|n\rangle,\quad \hat{\eta}_n=\eta_n+w\,\eta_1. \tag{3.59}$$

The $\mathcal{N}=8$ representation has $n-2$ diagrams while the $\mathcal{N}=7$ representation has $n-3$ diagrams. The additional term for $\mathcal{N}=8$ can be written as

$$\int \mathrm{d}^8\eta_{\hat{P}}\,\mathcal{M}_L\,\frac{1}{P^2}\,\mathcal{M}_R(\hat{1}\hat{P}2). \tag{3.60}$$

Intuitively, we want to expand this term into the other $n-3$ terms, so we separate the additional term and multiply s_{12} on each side

$$\mathcal{M}_n^{\mathcal{N}=8}=\int \mathrm{d}^8\eta_{\hat{P}}\,\mathcal{M}_L\,\frac{1}{P^2}\,\mathcal{M}_R(\hat{1}\hat{P}2)+\sum_{k=3}^{n-1}\int \mathrm{d}^8\eta_{\hat{P}}\,\mathcal{M}_L\,\frac{1}{P^2}\,\mathcal{M}_R(\hat{1}\hat{P}k), \tag{3.61}$$

$$s_{12}(0)\,\mathcal{M}_n^{\mathcal{N}=8}=\sum_{k=3}^{n-1}s_{12}(z_k)\int \mathrm{d}^8\eta_{\hat{P}}\,\mathcal{M}_L\,\frac{1}{P^2}\,\mathcal{M}_R(\hat{1}\hat{P}k). \tag{3.62}$$

After some manipulation, we successfully expand the additional term in $\mathcal{N}=8$ using others terms which have correspondence with $\mathcal{N}=7$.

$$\mathcal{M}_n^{\mathcal{N}=8}=\sum_{k=3}^{n-1}\frac{s_{12}(z_k)}{s_{12}(0)}\int \mathrm{d}^8\eta_{\hat{P}}\,\mathcal{M}_L\,\frac{1}{P^2}\,\mathcal{M}_R(\hat{1}\hat{P}k). \tag{3.63}$$

To compare with $\mathcal{N}=7$, we need to reduce the $\mathcal{N}=8$ terms to $\mathcal{N}=7$. In the MHV case, legs 1 and 2 are in multiplet Φ^-, which have helicity $+3/2$, while the other particles are in multiplet Φ^+, which has helicity $+2$, so we integrate out η_1^8, η_2^8 and $\hat{\eta_P}$ in the integral in (3.63) as follows:

$$
\begin{aligned}
&\int \mathrm{d}^8\eta_{\hat{P}}\, \mathcal{M}_L \, \frac{1}{P^2}\, \mathcal{M}_R(\hat{1}\hat{P}k) \\
&= \int \mathrm{d}\eta_1^8 \mathrm{d}\eta_2^8 \int \mathrm{d}\hat{\eta}_P^8 \,\delta(|n\rangle \hat{\eta}_n^8 + |\hat{P}\rangle \eta_{\hat{P}}^8 + \cdots)\, \delta([1k]\eta_{\hat{P}}^8 + [k\hat{P}]\eta_1^8 + [\hat{P}1]\eta_k^8) \int \mathrm{d}^7\eta_{\hat{P}}\, \widetilde{\mathcal{M}}_L \, \frac{1}{P^2}\, \widetilde{\mathcal{M}}_R \\
&= (w\langle 2n\rangle[1k] + [k\hat{P}]\langle 2\hat{P}\rangle) \int \mathrm{d}^7\eta_{\hat{P}}\, \widetilde{\mathcal{M}}_L \, \frac{1}{P^2}\, \widetilde{\mathcal{M}}_R \\
&= \langle 12\rangle[1k] \int \mathrm{d}^7\eta_{\hat{P}}\, \widetilde{\mathcal{M}}_L \, \frac{1}{P^2}\, \widetilde{\mathcal{M}}_R,
\end{aligned} \tag{3.64}
$$

where $\widetilde{\mathcal{M}}_L$ and $\widetilde{\mathcal{M}}_R$ are $\mathcal{M}_L$ and $\mathcal{M}_R$ with the supermomentum conservation delta function stripped off. Combining this result with (3.63), we obtain

$$
\sum_{k=3}^{n-1} \int \mathrm{d}^7\eta_{\hat{P}}\, \langle \hat{1}2\rangle[1k]\, \widetilde{\mathcal{M}}_L \, \frac{1}{P^2}\, \widetilde{\mathcal{M}}_R, \tag{3.65}
$$

which is exactly the explicit form for the corresponding $\mathcal{N}=7$ BCFW representation:

$$
\sum_{k=3}^{n-1} \int \mathrm{d}^7\eta_{\hat{P}}\, \mathcal{M}_L^{\mathcal{N}=7} \, \frac{1}{P^2}\, \mathcal{M}_R^{\mathcal{N}=7} = \sum_{k=3}^{n-1} \int \mathrm{d}^7\eta_{\hat{P}}\, \langle 2\hat{P}\rangle[\hat{P}k]\widetilde{\mathcal{M}}_L \frac{1}{P^2}\widetilde{\mathcal{M}}_R. \tag{3.66}
$$

What we have demonstrated is that we can use a bonus relation to relate $\mathcal{N}=8$ BCFW terms to $\mathcal{N}=7$ BCFW terms. In other words, the reason why $\mathcal{N}=7$ BCFW terms have nicer large z behavior in this example is precisely because they are implicitly using bonus relations to cancel out linear dependent terms which appear in the $\mathcal{N}=8$ representation.

The next question we can ask is whether the result applies to the general n-point $\mathrm{N^kMHV}$ case. To answer this question, we try the same analysis on the 6-point NMHV amplitude. Now we have 14 terms in $\mathcal{N}=8$ compared with 9 terms in $\mathcal{N}=7$, so we require 5 bonus relations to reduce the additional 5 terms to the other 9 terms. We cannot continue, since we only have one bonus relation and it is impossible to solve 5 parameters with one condition in general. This implies the $\mathcal{O}\left(z^{-2}\right)$ large z behavior of $\mathcal{N}=7$ individual terms include not only bonus relations which cancel out linear dependence but also some unknown property in $\mathcal{N}=7$.

3.5. Bonus scaling of "bad shift" BCFW for string amplitudes

Applications of BCFW recursion to string amplitudes have demonstrated improved large z scaling compared to field theory amplitudes in certain kinematic regimes [39] [40].

This not only validates the construction of a "bad shift" recursion formula without the requirement of $\mathcal{N}=7$ supersymmetry, but also enables the application of our previous argument to pursue even better term-by-term large z bonus scaling.

Since we encounter an infinite tower of massive states in string theory, we first demonstrate the validity of our argument in the case of a massive propagator. The previous derivation is modified such that the on-shell condition becomes $\hat{P}^2 = (P + w\,|i]\langle j|)^2 = m^2$. The primary shift parameter w and $|\hat{j}], |\hat{i}\rangle, \hat{\eta}_j, \hat{P}$ become:

$$w_m = \frac{-P^2 + m^2}{\langle j|P|i]} \tag{3.67}$$

$$|\hat{j}]_m = |j] + w_m|i], \quad |\hat{i}\rangle_m = |i\rangle - w_m|j\rangle, \quad \hat{\eta}_{j_m} = \eta_j + w_m\eta_i, \tag{3.68}$$

$$\hat{P}_m = \frac{P|j\rangle[i|P - m^2|i]\langle j|}{\langle j|P|i]}\,. \tag{3.69}$$

In the numerator of w_m, the additional m^2 term scales as z^0 while the original P^2 scales as z, so the large z scaling of w_m and hence $|\hat{j}], |\hat{i}\rangle, \hat{\eta}_j$ are not affected. The large z scaling of $\hat{P}_m$ is $\mathcal{O}(z^0)$, which is also unchanged compared to that of the massless $\hat{P}$. Hence making the propagator massive does not affect the large z behavior under the $[i, j\rangle$ test shift.

It was shown in [40] that the large z scaling under an $[i, j\rangle$ shift of superstring gluon amplitudes is improved by $z^{-\alpha' s_{ij}}$ compared to the corresponding field theory amplitude. For a $[j^+, i^-\rangle$ adjacent bad shift, the superstring amplitude scales as $z^{-\alpha' s_{ij}+3-\mathcal{N}}$ since the corresponding super-Yang-Mills amplitude scales as $z^{3-\mathcal{N}}$, thus by requiring the amplitude fall off faster than z^0, this leads to the kinematic condition $\mathrm{Re}\,[3 - \mathcal{N} - \alpha' s_{ij}] < 0$ for a valid representation. Following our previous result (3.50), under an $[i^-, j^+\rangle$ test shift the $\mathcal{N}$ bad shift representation has $z^{\mathcal{N}-5}$ term-by-term scaling, compared to the $z^{-\alpha' s_{ij}-1}$ large z fall off of the whole amplitude. Note the curious result that for $3 - \mathcal{N} < \mathrm{Re}\,[\alpha' s_{ij}] < 4 - \mathcal{N}$, the term-by-term scaling is actually better than the whole amplitude. We turn to a specific amplitude for further investigation.

As an example, we look at the superstring four-point gluon component amplitude, which is given by:

$$A_4(1^-, 2^-, 3^+, 4^+) = \frac{\langle 12\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle}\frac{\Gamma(1+\alpha' s)\Gamma(1+\alpha' t)}{\Gamma(1+\alpha'(s+t))} \tag{3.70}$$

Here the s and t are the usual Mandelstam variables, which in our convention read as $s = s_{12} = (p_1 + p_2)^2$, $t = s_{23} = (p_2 + p_3)^2$, and $u = s_{13} = (p_1 + p_3)^2$. The kinematic constraint for a valid recursion for this amplitude $\mathrm{Re}\,[3 - \alpha' t] < 0$ was first given in [39] by demonstrating the vanishing of the boundary term. We construct a bad shift representation by first deforming the amplitude with a $\mathcal{N} = 0$ $[3^+, 2^-\rangle$ shift,

$$A_4(w) = \frac{(\langle 12\rangle - w\langle 13\rangle)^3}{\langle 23\rangle\langle 34\rangle\langle 41\rangle}\frac{\Gamma(1+\alpha' s + w\alpha'[12]\langle 13\rangle)\Gamma(1+\alpha' t)}{\Gamma(1+\alpha'(s+t) + w\alpha'[12]\langle 13\rangle)}. \tag{3.71}$$

From the asymptotic expansion of the ratio of gamma functions, which can be obtained by using Stirling's series,

$$\frac{\Gamma(z+\alpha)}{\Gamma(z+\beta)} = z^{\alpha-\beta}\left[1+\frac{(\alpha-\beta)(\alpha+\beta-1)}{2z}\mathcal{O}\left(z^{-2}\right)\right], \tag{3.72}$$

we can readily see that $A_4(w)$ indeed scales as $w^{-\alpha' t+3}$.

Using the function $\frac{A_4(w)}{z}$, we can form the $[3^+,2^-\rangle$ representation of the amplitude as the sum of the residues at $w = -\frac{k+\alpha' s}{\alpha'[12]\langle 13\rangle}, k \in \mathbb{N}$. This representation can be simplified into

$$A_4(1^-,2^-,3^+,4^+) = \frac{\langle 12\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle}\frac{-1}{\alpha'^3 s^3}\sum_{k=1}^{\infty}\binom{\alpha' t}{k}\frac{(-1)^k k^4}{k+\alpha' s}. \tag{3.73}$$

Through direct summation using Mathematica, we can observe the convergence of the bad shift representation (3.73) to the closed form of the amplitude (3.70) within the kinematic regime $\mathrm{Re}\,[3-\alpha' t] < 0$. Another way to look at the convergence of the series is through the alternating series test. The ratio between terms of the series a_k expands at large k as

$$r = \left|\frac{a_{k+1}}{a_k}\right| = 1 + \frac{3-\alpha' t}{k} + \mathcal{O}\left(k^{-2}\right). \tag{3.74}$$

We obtain the condition $3-\alpha' t < 0$ by requiring $r < 1$ for sufficiently large k such that the series converges.

Under the $[2^-,3^+\rangle$ test shift, the $[3^+,2^-\rangle$ bad shift representation deforms into

$$A_4(z) = \frac{\langle 12\rangle^4}{\langle 12\rangle\langle 23\rangle(\langle 34\rangle + z\langle 24\rangle)\langle 41\rangle}\frac{-1}{\alpha'^3(s-z\langle 12\rangle[13])^3}\sum_{k=1}^{\infty}\binom{\alpha' t}{k}\frac{(-1)^k k^4}{k+\alpha'(s-z\langle 12\rangle[13])}. \tag{3.75}$$

From this form, we can observe directly that individual terms of the series fall off as z^{-5} as predicted. Also note that for $\alpha' t = n \in \mathbb{N}$, the series terminates after n terms and $A_4(z)$ has finite poles, in contrast to the case for $\alpha' t$ at generic values. This property can also be observed by shifting the closed form formula for A_4.

We now turn to the previously mentioned curiosity at $3 < \mathrm{Re}[\alpha' t] < 4$. Firstly, it is tested numerically by Mathematica that the series converges in this kinematic region and that under the $[2^-,3^+\rangle$ test shift, individual terms scale as z^{-5} at large z, better than the $z^{-\alpha' t-1}$ scaling of the amplitude in its closed form. We observe that the series converges slower at larger z, such that the number of terms required to sum to a certain fraction of the amplitude increases with z. From this, we expect that convergence issues may arise at the large z limit, allowing the large z fall off for individual terms to be better than the closed form in this kinematic region.

Similar analysis can be applied to the closed superstring. In our previous reasoning for supergravity, we noted that our argument for bonus scaling only applies to $\mathcal{N} = 7$ since the amplitude scales as $z^{6-\mathcal{N}}$ under the bad shift, and thus only offers a valid representation for $\mathcal{N} > 6$. For gravitons in the superstring, the condition for a valid $[j^+,i^-\rangle$ "bad shift" representation is:

$$\mathrm{Re}\,[6-\mathcal{N}-2\alpha' s_{ij}] < 0. \tag{3.76}$$

In this kinematic regime, the $[j^+, i^-\rangle$ bad shift representation has $z^{\mathcal{N}-9}$ term-by-term large z scaling under an $[i^-, j^+\rangle$ test shift according to (3.50), compared to the $z^{-2\alpha' s_{ij}^2 - 2}$ scaling of the whole amplitude. Similarly, note that the term-by-term large z fall off is better than the whole amplitude for $6 - \mathcal{N} < \mathrm{Re}\,[2\alpha' s_{ij}] < 7 - \mathcal{N}$.

4. Conclusion and Future directions

In this thesis, we proved that the "bad shift" BCFW representation of $\mathcal{N}=7$ supergravity gives building blocks that exhibit term by term bonus $\mathcal{O}(z^{-2})$ fall off. In particular, we prove that using the $[j^+, i^-\rangle$ BCFW representation of $\mathrm{N^kMHV}$ amplitudes, each term vanishes as $\mathcal{O}(z^{-2})$ under the $[i^-, j^+\rangle$ deformation. Focusing on the MHV case, we find that the $\mathcal{O}(z^{-2})$ behavior is also present for a large number of other $[-,+\rangle$ deformations. For example, in the $[n^+, 1^-\rangle$ representation, all $[2^-, m^+\rangle$ deformation exhibits term by term $\mathcal{O}(z^{-2})$ asymptotic behavior. The reason that the "bad shift" is a valid BCFW shift can be traced back to the $\mathcal{O}(z^{-2})$ fall off of $\mathcal{N}=8$ supergravity, which allows for the susy reduction to still have vanishing asymptotic, i.e. the shift behaves as $\mathcal{O}(z^{-2})$. Thus the "bad shift" BCFW representation of $\mathcal{N}=7$ supergravity is the only BCFW recursion that utilizes the $\mathcal{O}(z^{-2})$ fall off of the amplitude. We demonstrate this claim by showing that for the MHV case, we can use the bonus relation to recombine building blocks in $\mathcal{N}=8$ BCFW into building blocks of the $\mathcal{N}=7$ bad shift.

Our previous analysis only allows us to relate the BCFW representation of $\mathcal{N}=8$ supergravity to the $\mathcal{N}=7$ bad shift representation for the MHV amplitude. This relation is no longer straightforward for NMHV amplitude and beyond. For example the six-point NMHV contains 14 diagrams in $\mathcal{N}=8$ supergravity versus 9 diagrams for $\mathcal{N}=7$ bad-shift representation. Since there is only one bonus relation at each multiplicity, it is insufficient to convert one representation to the other, unless one incorporates the information of the bonus relations for the lower point amplitudes. This would require us to further expand the BCFW representation. Indeed it is known that using all bonus relation, one can express the supergravity amplitudes in terms of $(n-3)!$ building blocks [41]. It will be interesting to see if one can utilize these building blocks to form term by term $\mathcal{O}(z^{-2})$ fall off for all deformations.

Recent studies [12] have shown how BCFW terms of gravitational amplitudes can pair into combinations with improved permutation invariance, such that leading $\mathcal{O}(z^{-1})$ pieces cancel and $\mathcal{O}(z^{-2})$ fall off is exposed. However, it appears that to have $\mathcal{O}(z^{-2})$ fall off for all shifts, one eventually requires the combination of everything and end up with the full amplitude, which is similar to the $\mathcal{N}=7$ bad shift result. Thus it would appear that the improved fall off obtained by implementing partial permutation invariance can be similarly achieved without. It might be interesting to perform a general search of rational functions of spinor products that satisfies the correct helicity weight, mass dimension, at most simple poles and $\mathcal{O}(z^{-2})$ fall off for all shifts. These are very stringent constraints, and it is likely that the solution can serve as the true building blocks for the amplitude.

Finally, we note that the "bad shift" BCFW recursion is also valid for string amplitudes under certain kinematic conditions. Unlike the story for the $\mathcal{N}=7$ theory, whose validity of the "bad shift" BCFW is attributed to the bonus fall off of $\mathcal{N}=8$ gravity, here the validity of the string amplitude representation is tied to its improved high-energy behavior. Due to the enhanced large z scaling of string amplitudes, the restriction to the $\mathcal{N} = 7$ representation is lifted and we can further reduce supersymmetry to expose better term-by-term large z fall off compared to field theory. Furthermore, just as the bonus scaling of the $\mathcal{N} = 7$ bad shift representation may be considered as the incorporation of $\mathcal{N} = 8$ bonus relations, the improved behavior of BCFW terms of string amplitudes hint at possible relations inviting deeper investigation. It would be interesting to understand further, whether or not new symmetry or new amplitude relations emerge from this picture.

A. Amplitudes of Yang-Mills

In this appendix, we discuss topics specific to amplitudes of Yang-Mills and its supersymmetric extensions.

A.1. Yang-Mills and super-Yang-Mills

In the Lagrangian formulation, pure Yang-Mills theory is described by the Lagrangian

$$\mathcal{L} = -\frac{1}{4}\operatorname{tr} F_{\mu\nu}F^{\mu\nu}. \tag{A.1}$$

The field strength is given by

$$\begin{aligned} F &= (\tfrac{ig}{\sqrt{2}})^{-1}D^2 = [d, A] - \tfrac{ig}{\sqrt{2}}A^2 \\ &= F_{\mu\nu}dx^\mu dx^\nu = (\tfrac{ig}{\sqrt{2}})^{-1}[D_\mu, D_\nu]dx^\mu dx^\nu = 2(\partial_{[\mu}A_{\nu]} - \tfrac{ig}{\sqrt{2}}A_{[\mu}A_{\nu]})dx^\mu dx^\nu, \end{aligned} \tag{A.2}$$

which is related to the gauge covariant derivative,

$$\begin{aligned} D &= d + \tfrac{ig}{\sqrt{2}}A \\ &= D_\mu dx^\mu = (\partial_\mu + \tfrac{ig}{\sqrt{2}}A_\mu)dx^\mu \end{aligned} \tag{A.3}$$

of the non-Abelian gauge potential one-form

$$A = A_\mu dx^\mu = A^a_\mu T_a dx^\mu. \tag{A.4}$$

The gauge group for QCD is $G = \mathrm{SU}(3)$, however we will consider the more general case and take the gauge group to be $G = \mathrm{SU}(N)$.

$\mathfrak{su}(N)$ Lie algebra

Consider $U = e^{iA} = e^{iA^aT_a} \in \mathrm{SU}(N)$, where $T_a, a = 1, \ldots N^2 - 1$ are the generators of the associated Lie algebra $\mathfrak{su}(N)$, which can be represented by traceless Hermitian $N \times N$ matrices. We choose the normalization such that

$$T_aT_b = \frac{1}{N}\delta_{ab}I + \frac{1}{2}(if_{ab}{}^c + d_{ab}{}^c)T_c, \tag{A.5}$$

where the structure constants f are antisymmetric in all indices and the d-coefficients are symmetric in all indices. Then it follows that

$$\operatorname{tr} T_aT_b = \delta_{ab} \tag{A.6}$$

and

$$\frac{1}{2}\{T_a, T_b\} = \frac{1}{N}\delta_{ab} I + \frac{1}{2}d_{ab}{}^c T_c, \tag{A.7}$$

$$[T_a, T_b] = i f_{ab}{}^c T_c. \tag{A.8}$$

Completeness relations and Fierz identities

We can expand any $N \times N$ matrix A by the basis $\{I, T_a\}$:

$$A = A^0 I + A^a T_a. \tag{A.9}$$

Then the coefficients can be determined by the traces

$$A^0 = \frac{1}{N}\operatorname{tr} IA$$
$$A^a = \operatorname{tr} T^a A. \tag{A.10}$$

Therefore

$$A = (\frac{1}{N}\operatorname{tr} A)I + (\operatorname{tr} T^a A)T_a. \tag{A.11}$$

Writing in index form reads

$$A^i{}_j = \frac{1}{N}A^k{}_k \delta^i{}_j + (T^a)^k{}_l A^l{}_k (T_a)^i{}_j$$
$$\Rightarrow (T_a)^i{}_j (T^a)^k{}_l A^l{}_k = \delta^i{}_l \delta^k{}_j A^l{}_k - \frac{1}{N}\delta^i{}_j \delta^k{}_l A^l{}_k.$$

We obtain the completeness relation

$$(T_a)^i{}_j (T^a)^k{}_l = \delta^i{}_l \delta^k{}_j - \frac{1}{N}\delta^i{}_j \delta^k{}_l. \tag{A.12}$$

Gervais-Neveu gauge

In order to extract Feynman rules from the Lagrangian, we have to choose a gauge. A good choice for calculating amplitudes is the "Gervais-Neveu gauge", with gauge fixing term $\mathcal{L}_{gf} = -\frac{1}{2}\operatorname{tr}(H_\mu{}^\mu)^2$, $H_{\mu\nu} = \partial_\mu A_\nu - \frac{ig}{\sqrt{2}}A_\mu A_\nu$, which has the simplification that some of the terms in the Yang-Mills self-interaction have been canceled. In this gauge, the Lagrangian is

$$\mathcal{L} = \operatorname{tr}\left(\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu - i\sqrt{2}g\partial^\mu A^\nu A_\nu A_\mu + \frac{g^2}{4}A^\mu A^\nu A_\nu A_\mu.\right) \tag{A.13}$$

The Feynman rules then give a gluon propagator $\delta^{ab}\frac{\eta_{\mu\nu}}{p^2}$ and 3-point and 4-point gluon vertices dressed with color factors f^{abc} and $f^{abx}f^{xcd}$ + (permutations), respectively.

A.2. Color structure of Yang-Mills amplitudes

It follows from (A.10) and the definition of the structure constants (A.8) that

$$if_{abc} = [T_a, T_b]_c = \operatorname{tr} T_c[T_a, T_b]. \tag{A.14}$$

We can use the completeness relation in index form (A.12) to contract structure constant products into generator product traces. However, it is more straightforward to carry out the contraction by combining the use of (A.10) and (A.14). For example,

$$\begin{aligned} f_{a_1a_2b}f^{b}{}_{a_3a_4} &= -[T_{a_1}, T_{a_2}]_b\, \delta^{bc}\, [T_{a_3}, T_{a_4}]_c \\ &= -[T_{a_1}, T_{a_2}]_b\, (\operatorname{tr} T^b T^c)\, [T_{a_3}, T_{a_4}]_c \\ &= -\operatorname{tr}[T_{a_1}, T_{a_2}]_b T^b [T_{a_3}, T_{a_4}]_c T^c \\ &= -\operatorname{tr}[T_{a_1}, T_{a_2}][T_{a_3}, T_{a_4}] \\ &= -\operatorname{tr} T_{a_1}T_{a_2}T_{a_3}T_{a_4} + \operatorname{tr} T_{a_1}T_{a_2}T_{a_4}T_{a_3} + \operatorname{tr} T_{a_1}T_{a_3}T_{a_4}T_{a_2} - \operatorname{tr} T_{a_1}T_{a_4}T_{a_3}T_{a_2}. \end{aligned}$$

Similarly, generic color factors of tree amplitudes in terms of contractions of the structure constants can be grouped into single-trace color factors. We can therefore separate color and kinematic degrees of freedom by using trace color factors as a basis,

$$A_n^{tree} = \sum_{\sigma} A_n[1\sigma(2\ldots n)] \operatorname{tr}(T_{a_1}T_{(a_2}\ldots T_{a_n)}), \tag{A.15}$$

where the summation is over the permutations σ of $(2\ldots n)$ due to the cyclic symmetry of the trace. The partial amplitudes in the summation are called color-ordered amplitudes, which are gauge invariant quantities. More information on color-ordered amplitudes can be found in [42].

To calculate color-ordered amplitudes, we can use color ordered Feynman rules accordingly. The color ordered Feynman rules in the Gervais-Neveu gauge are

- 3-gluon vertex

$$V^{\mu_1\mu_2\mu_3}(p_1, p_2, p_3) = -\sqrt{2}(\eta^{\mu_1\mu_2}p_1^{\mu_3} + \eta^{\mu_2\mu_3}p_2^{\mu_1} + \eta^{\mu_3\mu_1}p_3^{\mu_2}). \tag{A.16}$$

- 4-gluon vertex

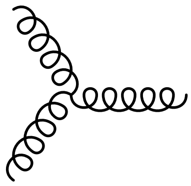
$$V^{\mu_1\mu_2\mu_3\mu_4}(p_1, p_2, p_3, p_4) = \eta^{\mu_1\mu_3}\eta^{\mu_2\mu_4}. \tag{A.17}$$

The color ordered amplitude is calculated in terms of diagrams with no lines crossing and the ordering of the external lines fixed as given $1, 2, 3, \ldots, n$. To illustrate the calculation explicitly, let us start from the simplest examples.

Example: YM color-ordered 3-point tree amplitude

The 3-gluon amplitude consists of a single Feynman diagram given by the 3-vertex rule,

$A_3[123] =$



$$= -\sqrt{2}\left[(\epsilon_1 \cdot \epsilon_2)(p_1 \cdot \epsilon_3) + (\epsilon_2 \cdot \epsilon_3)(p_2 \cdot \epsilon_1) + (\epsilon_3 \cdot \epsilon_1)(p_3 \cdot \epsilon_2)\right]. \tag{A.18}$$

Let us now choose the helicity configuration to be $1^-2^-3^+$, and translate to spinor helicity formalism to proceed,

$$A_3[1^-2^-3^+] = -\frac{\langle 12\rangle[q_1q_2]\langle q_3 1\rangle[13] + \langle q_3 2\rangle[q_2 3]\langle 12\rangle[q_1 2] + \langle q_3 1\rangle[q_1 3]\langle 23\rangle[q_2 3]}{[q_1 1][q_2 2]\langle q_3 3\rangle} \tag{A.19}$$

We must now consider 3-point special kinematics $|1\rangle \propto |2\rangle \propto |3\rangle$ or $|1] \propto |2] \propto |3]$. If $|1\rangle \propto |2\rangle \propto |3\rangle$, all three terms in the numerator of (A.19) vanishes, therefore $|1] \propto |2] \propto |3]$ and the first term vanishes,

$$A_3[1^-2^-3^+] = -\frac{\langle q_3 2\rangle[q_2 3]\langle 12\rangle[q_1 2] + \langle q_3 1\rangle[q_1 3]\langle 23\rangle[q_2 3]}{[q_1 1][q_2 2]\langle q_3 3\rangle} \tag{A.20}$$

By momentum conservation, $\langle 12\rangle[q_1 2] = -\langle 13\rangle[q_1 3]$. Then

$$\begin{aligned} A_3[1^-2^-3^+] &= -\frac{[q_1 3][q_2 3]\left(-\langle q_3 2\rangle\langle 31\rangle + \langle q_3 1\rangle\langle 23\rangle\right)}{[q_1 1][q_2 2]\langle q_3 3\rangle} \\ &= \frac{[q_1 3][q_2 3]\langle q_3 3\rangle\langle 12\rangle}{[q_1 1][q_2 2]\langle q_3 3\rangle} \\ &= \frac{[q_1 3][q_2 3]\langle 12\rangle}{[q_1 1][q_2 2]} \end{aligned} \tag{A.21}$$

by Schouten identity. Finally, we simplify the expression by momentum conservation,

$$\begin{aligned} A_3[1^-2^-3^+] &= \frac{[q_1 3]\langle 32\rangle[q_2 3]\langle 31\rangle\langle 12\rangle}{[q_1 1][q_2 2]\langle 32\rangle\langle 31\rangle} \\ &= \frac{[q_1 1]\langle 12\rangle[q_2 2]\langle 21\rangle\langle 12\rangle}{[q_1 1][q_2 2]\langle 32\rangle\langle 31\rangle} \end{aligned} \tag{A.22}$$

The result is remarkably simple,

$$A_3[1^-2^-3^+] = \frac{\langle 12\rangle^3}{\langle 23\rangle\langle 31\rangle}$$

This form can also be derived solely by considering the helicity structure with assumption of locality, as considered in Section 2.3.2.

A.3. $\mathcal{N} = 4$ super-Yang-Mills

From the helicity of the gluon, we can infer that the maximal supersymmetric extension to Yang-Mills theory is $\mathcal{N} = 4$ super-Yang-Mills (SYM). The theory is invariant under the a Yangian symmetry. The spectrum for the theory is

$$\begin{array}{ccccc} 1 & 4 & 6 & 4 & 1 \\ h=+1 & h=+\frac{1}{2} & h=0 & h=-\frac{1}{2} & h=-1 \\ \text{gluon} & \text{gluinos} & \text{scalars} & \text{gluinos} & \text{gluon} \\ g^{+} & \lambda_{A}^{+} & S_{AB} & \lambda^{-A} & g^{-} \end{array} \tag{A.23}$$

In the on-shell superspace formalism, the superfield for $\mathcal{N} = 4$ SYM is

$$\Phi = g^{+} + \lambda_{A}^{+}\eta^{A} + \frac{1}{2!}S_{AB}\eta^{A}\eta^{B} + \lambda^{-A}\eta_{A}^{3} + g^{-}\eta^{4}. \tag{A.24}$$

Bibliography

[1] R. Britto, F. Cachazo, B. Feng and E. Witten, Phys. Rev. Lett. **94**, 181602 (2005) [hep-th/0501052].

[2] N. Arkani-Hamed, F. Cachazo, C. Cheung and J. Kaplan, JHEP **1003**, 020 (2010) [arXiv:0907.5418 [hep-th]].

[3] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. B. Goncharov, A. Postnikov and J. Trnka, [arXiv:1212.5605 [hep-th]].

[4] H. Elvang and Y. t. Huang, [arXiv:1308.1697 [hep-th]].

[5] O. Aharony, O. Bergman, D. L. Jafferis and J. Maldacena, JHEP **0810**, 091 (2008) [arXiv:0806.1218 [hep-th]].

[6] K. Hosomichi, K. -M. Lee, S. Lee, S. Lee and J. Park, JHEP **0809**, 002 (2008) [arXiv: 0806.4977 [hep-th]].

[7] S. Lee, Phys. Rev. Lett. **105**, 151603 (2010) [arXiv:1007.4772 [hep-th]].

[8] Y. -T. Huang and C. Wen, "ABJM amplitudes and the positive orthogonal Grassmannian," JHEP **1402**, 104 (2014) [arXiv:1309.3252 [hep-th]].

[9] Y. -t. Huang, C. Wen and D. Xie, "The Positive orthogonal Grassmannian and loop amplitudes of ABJM," [arXiv:1402.1479 [hep-th]].

[10] J. M. Drummond, J. M. Henn and J. Plefka, JHEP **0905**, 046 (2009) [arXiv: 0902.2987 [hep-th]].

[11] N. Arkani-Hamed and J. Kaplan, JHEP **0804**, 076 (2008) [arXiv:0801.2385 [hep-th]].

[12] D. A. McGady and L. Rodina, [arXiv:1408.5125 [hep-th]].

[13] D. Nguyen, M. Spradlin, A. Volovich and C. Wen, JHEP **1007**, 045 (2010) [arXiv: 0907.2276 [hep-th]].

[14] A. Hodges, JHEP **1307** (2013) [arXiv:1108.2227 [hep-th]].

[15] N. Arkani-Hamed, F. Cachazo and J. Kaplan, JHEP **1009**, 016 (2010) [arXiv: 0808.1446 [hep-th]].

[16] B. Feng and M. Luo, Front. Phys. **7**, 533 (2012) [arXiv:1111.5759 [hep-th]].

[17] L. J. Dixon, In *Boulder 1995, QCD and beyond* 539-582 [hep-ph/9601359].

[18] L. J. Dixon, J. Phys. A **44**, 454001 (2011) [arXiv:1105.0771 [hep-th]].

[19] L. J. Dixon, [arXiv:1310.5353 [hep-ph]].

[20] R. Boels and F. Brümmer, "Introduction to Supersymmetry and Supergravity - or Supersymmetry in 12954 minutes "

[21] R. Boels "Introduction to Scattering Amplitudes Lecture 1: QCD and the Spinor-Helicity Formalism"

[22] R. Britto, J. Phys. A **44**, 454006 (2011) [arXiv:1012.4493 [hep-th]].

[23] R. Boels, JHEP **1001**, 010 (2010) [arXiv:0908.0738 [hep-th]].

[24] R. H. Boels and C. Schwinn, Phys. Rev. D **84**, 065006 (2011) [arXiv:1104.2280 [hep-th]].

[25] M. T. Grisaru and H. N. Pendleton, Some Properties Of Scattering Amplitudes In Supersymmetric Theories, Nucl. Phys. B 124, 81 (1977).
M. T. Grisaru, H. N. Pendleton and P. van Nieuwenhuizen, Supergravity And The S Matrix, Phys. Rev. D 15, 996 (1977).

[26] H. Elvang, D. Z. Freedman and M. Kiermaier, J. Phys. A **44**, 454009 (2011) [arXiv: 1012.3401 [hep-th]].

[27] J. Wess and J. Bagger, "Supersymmetry and supergravity," Princeton, USA: Univ. Pr. (1992) 259 p

[28] Y. Tachikawa, Lect. Notes Phys. **890**, 2014 [arXiv:1312.2684 [hep-th]].

[29] B. Feng, K. Zhou, C. Qiao and J. Rao, JHEP **1503**, 023 (2015) [arXiv:1411.0452 [hep-th]].

[30] B. Feng, J. Rao and K. Zhou, [arXiv:1504.06306 [hep-th]].

[31] P. Benincasa, C. Boucher-Veronneau and F. Cachazo, JHEP **0711**, 057 (2007) [hep-th/0702032 [HEP-TH]].

[32] P. Benincasa and F. Cachazo, [arXiv:0705.4305 [hep-th]].

[33] Z. Bern, J. J. Carrasco, D. Forde, H. Ita and H. Johansson, Phys. Rev. D **77**, 025010 (2008) [arXiv:0707.1035 [hep-th]].

[34] N. Arkani-Hamed and J. Kaplan, JHEP **0804**, 076 (2008) [arXiv:0801.2385 [hep-th]].

[35] S. Sannan, "Gravity As The Limit Of The Type II Superstring Theory," Phys. Rev. D **34**, 1749 (1986)

[36] F. A. Berends, W. T. Giele and H. Kuijf, "On relations between multi-gluon and multi-graviton scattering," Phys. Lett. B **211**, 91 (1988).

[37] H. Elvang, Y. t. Huang and C. Peng, JHEP **1109**, 031 (2011) [arXiv:1102.4843 [hep-th]].

[38] D. Nandan and C. Wen, JHEP **1208**, 040 (2012) [arXiv:1204.4841 [hep-th]].

[39] R. Boels, K. J. Larsen, N. A. Obers and M. Vonk, JHEP **0811**, 015 (2008) [arXiv: 0808.2598 [hep-th]].

[40] R. H. Boels, D. Marmiroli and N. A. Obers, JHEP **1010**, 034 (2010) [arXiv:1002.5029 [hep-th]].

[41] S. He, D. Nandan and C. Wen, JHEP **1102**, 005 (2011) [arXiv:1011.4287 [hep-th]].

[42] T. Schuster, Phys. Rev. D **89**, no. 10, 105022 (2014) [arXiv:1311.6296 [hep-ph]].